

No-Go Theorems for Quantum Resource Purification: New Approach and Channel Theory

Kun Fang^{1,2,*} and Zi-Wen Liu^{3,†}

¹*Institute for Quantum Computing, Baidu Research, Beijing 100193, China*

²*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

³*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

 (Received 3 November 2020; revised 19 January 2022; accepted 9 February 2022; published 8 March 2022)

It has been recently shown that there exist universal fundamental limits to the accuracy and efficiency of the transformation from noisy resource states to pure ones (e.g., distillation) in any well-behaved quantum resource theory [Fang and Liu, Phys. Rev. Lett. **125**, 060405 (2020)]. Here, we develop a novel and powerful method for analyzing the limitations on quantum resource purification, which not only leads to improved bounds that rule out exact purification for a broader range of noisy states and are tight in certain cases, but also enable us to establish a robust no-purification theory for quantum channel (dynamical) resources. More specifically, we employ the new method to derive universal bounds on the error and cost of transforming generic noisy channels (where multiple instances can be used adaptively, in contrast to the state theory) to some unitary resource channel under any free channel-to-channel map. We address several cases of practical interest in more concrete terms, and discuss the connections and applications of our general results to distillation, quantum error correction, quantum Shannon theory, and quantum circuit synthesis.

DOI: [10.1103/PRXQuantum.3.010337](https://doi.org/10.1103/PRXQuantum.3.010337)

I. INTRODUCTION

Quantum technologies, such as quantum computing, quantum communication, and quantum cryptography, are an exciting frontier of science, due to their promising potential of achieving substantial advantages over conventional methods that may spark an important technological revolution. However, quantum systems are inherently highly susceptible to noise and errors in real-world scenarios, which often make them unreliable or difficult to scale up. This poses a serious challenge to realizing the potential power of quantum technologies in practice. The noise problem is particularly pressing at the moment, as we are now at a critical juncture where we are starting to make real effort to put the theoretically blueprinted quantum technologies into practice [1,2]. In order to ease the effects of noise, we would generally need techniques that can “purify” the noisy systems. To this end, methods

such as quantum error correction [3] and distillation [4–7] are developed and have become central research topics in quantum information.

Behind the power of quantum technologies is the manipulation and utilization of various forms of quantum “resources” such as entanglement [8], coherence [9], and “magic” [7,10]. These different kinds a quantum resources can be commonly understood and characterized using the universal framework of “quantum resource theory” (see, e.g., Ref. [11] for an introduction), which have been under active development in recent years. Recently, Ref. [12] revealed a fundamental principle of quantum mechanics that there exists universal limitations on the accuracy and efficiency of purifying noisy states in general quantum resource theories, by employing one-shot resource theory ideas [13]. However, Ref. [12] is only part of the story and there are two gaps that we would like to fill to make the picture more complete. First, the results there assume the input states to be full-rank and it is not fully understood whether there are no-purification rules when the input state is noisy but not of full rank. Second, the approach developed there is primarily designed for state or static resources, but given that the manipulation of channel or dynamical resources plays intrinsic roles in many scenarios including quantum computation, communication, and error correction, it is also important to understand whether the no-purification principles extend to quantum channels.

*fangkun02@baidu.com

†zliu1@perimeterinstitute.ca

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

In this work, we develop a novel approach to establishing fundamental limits of general quantum resource purification tasks, which addresses the above problems. This approach is built upon decompositions of the input that separate out the free parts. As we demonstrate, such decompositions link the weight of the free parts, a key quantity that we call *free component*, to the optimal error of purification. We apply this approach to both quantum states and channel resource theories. For state theories, we use the new method to derive new bounds on the error and efficiency of deterministic purification or distillation tasks, which significantly improve those in Ref. [12]. More specifically, the new results lift the full-rank assumption and imply no-purification principles for a broader range of mixed states. Furthermore, they are quantitatively better and are shown to be tight in certain simple cases. We use several concrete examples to demonstrate the improvements and show that the new bounds are tight in certain cases. Next, as a major contribution of this work, we develop a comprehensive no-purification theory for quantum channels (Ref. [12] presents only a zero-error result). Most importantly, there are two key complications of the channel theory that does not come up in the state theory: (i) there are several different ways to define channel fidelity measures; (ii) multiple instances of channels can be used or consumed in various presumably inequivalent ways, such as in parallel, sequentially, or adaptively. Using the free-component method, we derive bounds on the purification errors and costs for all cases. To provide a more concrete understanding, we discuss the roles and features of common noise channels in different types of channel resource theories, as well as providing guidelines for applying the no-purification bounds to a broad range of fields of great theoretical and practical interest, including distillation, quantum error correction, Shannon theory, and circuit (gate) synthesis.

We emphasize a particularly remarkable and counter-intuitive feature of the no-purification principles, which is that they rule out any noisy-to-pure transformation for noisy input states or channels with free component, where the noisy inputs can be much more “resourceful” in terms of common resource measures or operational tasks than the pure targets. This is in sharp contrast with generic (such as pure-to-pure) transformation tasks where the transformability is naturally determined by the resource content in general. Also notably, our theory is applicable to virtually all well-defined resource theories (not even requiring the standard convexity assumption), highlighting the fundamental nature of the no-purification principles.

The paper is organized as follows. In Sec. II, we apply the free-component method to state theories, and in particular discuss the improvements over previous results in Ref. [12]. In Sec. III, we establish the no-purification theory for quantum channels using the free-component method. We first present general-form results in Sec. III A,

and then elaborate on specific scenarios and applications in Sec. III B. Finally in Sec. IV we summarize the work and discuss future directions.

II. STATE THEORY

We first consider state resource theories, which are built upon the notions of *free states* and *free operations* that represent the allowed transformation among states. Here, we consider the most general resource theory framework with the “minimalist” requirement—the *golden rule* that any free operation must map a free state to another free state, or in other words, cannot create resource (see, e.g., Refs. [11,14,15]). This golden rule defines the largest possible set of operations that encompasses any legitimate set of free operations, and thus the fundamental limits induced by it apply universally to any nontrivial resource theory. Also, for mathematical rigor, we assume that the set of free states \mathcal{F} has the following two reasonable, commonly held properties: (i) the composition of free states should be free, namely if $\rho_1, \rho_2 \in \mathcal{F}$ then $\rho_1 \otimes \rho_2 \in \mathcal{F}$; (ii) \mathcal{F} is closed.

The following quantity that we call *free component* will play a central role in our theory.

Definition 1 (Free component). *The free component of quantum state ρ is defined as*

$$\Gamma_\rho := \max \{ \gamma : \rho - \gamma\sigma \geq 0, \sigma \in \mathcal{F} \}. \quad (1)$$

Equivalently,

$$\Gamma_\rho = \max \{ \gamma : \rho = \gamma\sigma + (1 - \gamma)\tau, \sigma \in \mathcal{F}, \tau \in \mathcal{D} \}, \quad (2)$$

where \mathcal{D} is the set of all density matrices. That is, the free component is directly related to the “weight of resource” W , which is recently studied in general resource theory contexts [16,17], by $\Gamma_\rho = 1 - W_\rho$. Another equivalent form is $\Gamma(\rho) = \min_{\sigma \in \mathcal{F}} 2^{D_{\max}(\sigma \parallel \rho)}$ where the max-relative entropy is defined by $D_{\max}(\sigma \parallel \rho) := \log \min \{ t : \sigma \leq t\rho \}$ if $\text{supp}(\sigma) \subseteq \text{supp}(\rho)$ and $+\infty$ otherwise [18]. Note that, if \mathcal{F} can be characterized by semidefinite conditions (which is quite common, e.g., in coherence theory $\mathcal{F} = \{ \sigma : \sigma \geq 0, \text{Tr} \sigma = 1, \sigma = \Delta(\sigma) \}$, where Δ is the dephasing channel erasing the off-diagonal entries), then Γ_ρ can be efficiently computed by semidefinite programming (SDP) for given ρ . In the resource theory of thermodynamics, for Hamiltonian H and inverse temperature β the Gibbs (thermal) state $\sigma := e^{-\beta H} / \text{Tr} e^{-\beta H}$ is the only free state and we thus have a closed-form formula for free component as $\Gamma_\rho = 1 / \lambda^{\max}(\rho^{-1}\sigma)$ (where λ^{\max} denotes the largest eigenvalue) if $\text{supp}(\rho) \supseteq \text{supp}(\sigma)$ and zero otherwise [19, Theorem 2].

It can be easily seen that the free component obeys the desirable monotonicity property that it cannot be reduced by free operations.

Proposition 2 (Monotonicity). *For any state ρ and any free operation \mathcal{N} , it holds that $\Gamma_\rho \leq \Gamma_{\mathcal{N}(\rho)}$.*

Proof. Suppose Γ_ρ is achieved by $\sigma \in \mathcal{F}$. Then by definition $\rho - \Gamma_\rho \sigma \geq 0$, and thus $\mathcal{N}(\rho) - \Gamma_\rho \mathcal{N}(\sigma) \geq 0$. Since $\mathcal{N}(\sigma) \in \mathcal{F}$ by the golden rule, we have that $\Gamma_{\mathcal{N}(\rho)} \geq \Gamma_\rho$ by definition. ■

Moreover, it is supermultiplicative under tensor product of states:

Proposition 3 (Supermultiplicity). *For any quantum states ρ_1, ρ_2 , it holds that $\Gamma_{\rho_1 \otimes \rho_2} \geq \Gamma_{\rho_1} \Gamma_{\rho_2}$.*

Proof. Suppose that the maximization in $\Gamma_{\rho_1}, \Gamma_{\rho_2}$ are, respectively, achieved by $\sigma_1, \sigma_2 \in \mathcal{F}$, that is, $\rho_i \geq \Gamma_{\rho_i} \sigma_i, i = 1, 2$. It holds that $\rho_1 \otimes \rho_2 \geq (\Gamma_{\rho_1} \sigma_1) \otimes (\Gamma_{\rho_2} \sigma_2) = \Gamma_{\rho_1} \Gamma_{\rho_2} \sigma_1 \otimes \sigma_2$. Also note that $\sigma_1 \otimes \sigma_2 \in \mathcal{F}$ axiomatically. Therefore, $\Gamma_{\rho_1 \otimes \rho_2} \geq \Gamma_{\rho_1} \Gamma_{\rho_2}$ by definition. ■

Consider the task of purification, namely transforming a noisy state ρ to a certain target pure state $\psi = |\psi\rangle\langle\psi|$ up to some error. Formally, the error of purification is defined by the infidelity with the target state: for input state ρ , transformation operation \mathcal{N} and target pure state ψ , the error

$$\varepsilon(\rho \xrightarrow{\mathcal{N}} \psi) := 1 - \text{Tr} \psi \mathcal{N}(\rho) = 1 - \langle \psi | \mathcal{N}(\rho) | \psi \rangle. \quad (3)$$

Also, let f_ψ denote the maximum overlap of pure state $\psi = |\psi\rangle\langle\psi|$ with free states, namely,

$$f_\psi := \max_{\sigma \in \mathcal{F}} \text{Tr} \psi \sigma = \max_{\sigma \in \mathcal{F}} \langle \psi | \sigma | \psi \rangle. \quad (4)$$

We now prove an improved deterministic no-purification theorem using a method different from Ref. [12], which directly connects the accuracy of purifying a noisy state with its free component.

Theorem 4. *Given any state ρ and any pure state ψ , there is no free operation that transforms ρ to ψ with error smaller than $\Gamma_\rho(1 - f_\psi)$. That is, it holds for any free operation \mathcal{N} that*

$$\varepsilon(\rho \xrightarrow{\mathcal{N}} \psi) \geq \Gamma_\rho(1 - f_\psi). \quad (5)$$

Proof. By the definition of Γ_ρ , there exists free state $\sigma \in \mathcal{F}$ and state τ such that ρ can be decomposed as

follows:

$$\rho = \Gamma_\rho \sigma + (1 - \Gamma_\rho) \tau. \quad (6)$$

Let \mathcal{N} be any free operation. By linearity,

$$\mathcal{N}(\rho) = \Gamma_\rho \mathcal{N}(\sigma) + (1 - \Gamma_\rho) \mathcal{N}(\tau). \quad (7)$$

Then it holds that

$$\varepsilon(\rho \xrightarrow{\mathcal{N}} \psi) = 1 - \text{Tr} \mathcal{N}(\rho) \psi \quad (8)$$

$$= 1 - \Gamma_\rho \text{Tr} \mathcal{N}(\sigma) \psi - (1 - \Gamma_\rho) \text{Tr} \mathcal{N}(\tau) \psi \quad (9)$$

$$\geq 1 - \Gamma_\rho f_\psi - (1 - \Gamma_\rho) \quad (10)$$

$$= \Gamma_\rho(1 - f_\psi), \quad (11)$$

where the inequality follows from $\text{Tr} \mathcal{N}(\sigma) \psi \leq f_\psi$ since $\mathcal{N}(\sigma) \in \mathcal{F}$ by the golden rule, and $\text{Tr} \mathcal{N}(\tau) \psi \leq 1$. ■

As first noted in Ref. [12], we can translate the upper bounds on transformation accuracy into lower bounds on the ‘‘amount’’ of input resources required to achieve a certain target, in particular, the cost of many-copy distillation procedures, which are widely considered for various purposes in quantum computation and information [4–7]. The above Theorem 4 induces the following general lower bound on distillation overhead.

Corollary 5. *Consider distillation procedures represented by a free operation that transform n copies of noisy states ρ to a target pure state ψ within error ϵ . Then n must satisfy*

$$n \geq \left\lceil \log \frac{1 - f_\psi}{\epsilon} \right\rceil \left\lceil \log \frac{1}{\Gamma_\rho} \right\rceil^{-1}. \quad (12)$$

Proof. Suppose the transformation is given by the free operation \mathcal{N} . Then it holds from Theorem 4 that

$$\epsilon \geq \varepsilon(\rho^{\otimes n} \xrightarrow{\mathcal{N}} \psi) \geq \Gamma_{\rho^{\otimes n}}(1 - f_\psi). \quad (13)$$

Note that, due to the supermultiplicity property from Proposition 3, we get $\Gamma_{\rho^{\otimes n}} \geq (\Gamma_\rho)^n$. This gives

$$\epsilon \geq (\Gamma_\rho)^n(1 - f_\psi), \quad (14)$$

which is equivalent to the above assertion. ■

Our new method essentially replaces the minimum eigenvalue of ρ in the corresponding bounds in Ref. [12] (which we refer to as the min-eigenvalue bounds) by its free component, which represents a significant improvement from both qualitative and quantitative perspectives, as detailed in the following.

First, the range of applicability of the no-purification theorem is significantly extended. The proof using the quantum hypothesis testing relative entropy presented in Ref. [12] applies only to full-rank input states. However, Theorem 4 implies that the no-purification rule actually holds more broadly (see also Ref. [20, Proposition 2]):

Corollary 6. *There is no free operation that exactly transforms a state ρ to any pure state $\psi \notin \mathcal{F}$ if $\Gamma_\rho > 0$.*

Proof. Since \mathcal{F} is closed by assumption and $\psi \notin \mathcal{F}$, we have $f_\psi < 1$. Then due to Theorem 4, the transformation error $\varepsilon > 0$, indicating that exact transformation is impossible. ■

Below we give some alternative useful characterizations of the $\Gamma > 0$ condition.

Proposition 7. *For any quantum state ρ , the following conditions are equivalent:*

- (a) *free component: $\Gamma_\rho > 0$;*
- (b) *support: there exists a free state $\sigma \in \mathcal{F}$ such that the support condition $\text{supp}(\rho) \supseteq \text{supp}(\sigma)$ holds;*
- (c) *resource measure: the min-relative entropy of resource $\mathfrak{D}_{\min}(\rho) := \min_{\sigma \in \mathcal{F}} D_{\min}(\rho \parallel \sigma) = 0$, where $D_{\min}(\rho \parallel \sigma) := -\log \text{Tr} \Pi_\rho \sigma$ is the min-relative entropy, and Π_ρ is the projector onto $\text{supp}(\rho)$.*

Proof. It is easy to see that (a) implies (b) and (b) implies (c). It remains to show that (c) implies (a). Suppose (c) holds, then there exists $\sigma \in \mathcal{F}$ such that $D_{\min}(\rho \parallel \sigma) = 0$. By definition, we have $\text{Tr}(I - \Pi_\rho)\sigma = 0$. That is, $\ker(\rho) \perp \text{supp}(\sigma)$, and equivalently, $\text{supp}(\rho) \supseteq \text{supp}(\sigma)$. If $\rho = \sigma$, then (a) holds. Otherwise, we have $D_{\max}(\sigma \parallel \rho) > 0$. [Note that $D_{\max}(\sigma \parallel \rho) = 0$ if and only if $\rho = \sigma$.] This implies that there exists $t > 1$ such that $\sigma \leq t\rho$. Thus $\Gamma_\rho \geq 1/t > 0$, implying (a). ■

It is clear that for any pure resource state $\psi \notin \mathcal{F}$ we have $\Gamma_\psi = 0$, so the no-purification bounds can only be nontrivial for mixed states. Meanwhile, it can be immediately seen [e.g., from (b)] that the $\Gamma > 0$ condition is weaker than the full-rank condition. In fact, it holds as long as the support of ρ contains some free state in its support, which is generically the case for mixed states in common resource theories. Also note that the $\Gamma > 0$ condition does not necessarily hold for all mixed states. For a concrete example, consider the coherence theory defined by an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$. Consider the state $\rho = (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)/2$ where $|\psi_1\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|\psi_2\rangle = (|2\rangle + |3\rangle)/\sqrt{2}$. It can be verified that ρ is mixed but $\Gamma_\rho = 0$ because any decreasing of the diagonal entries will render the matrix negative. It would

be interesting to further understand and characterize the $\Gamma > 0$ condition in specific theories.

Furthermore, note that the derivation and results (also the channel versions below) apply to continuous variable or infinite-dimensional quantum systems: the relevant quantities, the free component Γ and the maximum overlap f , can be defined likewise (supremum instead of maximum over \mathcal{F}), and the proof steps follow. In particular, $\Gamma > 0, f < 1$ still indicate no purification. An elementary continuous variable example will be given later.

We remark that if we only require the purification transformation to succeed with some probability (the probabilistic setting), the $\Gamma > 0$ condition is not sufficient to rule out purification and it seems that the full-rank condition cannot be alleviated. For example, consider the following state with a flag register F :

$$\rho = p|0\rangle\langle 0|_F \otimes \psi_A + (1-p)|1\rangle\langle 1|_F \otimes \tau_A, \quad (15)$$

where ψ is the target pure state and τ is a state such that $\Gamma_\tau > 0$. Then we have $\Gamma_\rho > 0$ (ρ is not full rank), but we can obtain ψ with probability p simply by measuring F (which is conventionally free) and postselect on 0.

Second, the new results are quantitatively better than the corresponding ones in Ref. [12] for full-rank input states. It is first straightforward to see that $\Gamma_\rho \geq \lambda_\rho^{\min}$, where λ_ρ^{\min} denotes the minimum nonzero eigenvalue of ρ , because $\rho \geq \lambda_\rho^{\min} \cdot I \geq \lambda_\rho^{\min} \cdot \sigma$ for any state σ where I denotes the identity matrix on $\text{supp}(\rho)$. So by definition, $\Gamma_\rho \geq \lambda_\rho^{\min}$. In sum, the new free-component bounds cover the min-eigenvalue bounds. In particular, when the noisy state ρ is close to the set of free states \mathcal{F} , the minimum eigenvalue λ_ρ^{\min} could still be small but Γ_ρ approaches 1. This indicates that the free-component bounds potentially exhibit much tighter behaviors in the large error regime like when ρ is close to \mathcal{F} . Importantly, the distillation overhead bound Corollary 5 indicates the key behavior that as ρ approaches \mathcal{F} , it holds that $n \rightarrow \infty$, i.e., the number of copies needed diverges, because $\Gamma_\rho \rightarrow 1$. This cannot be deduced from the min-eigenvalue bounds in Ref. [12].

Now we discuss the application of our general bounds in a few important specific scenarios that are of practical interest in diverse manners, showcasing the versatility of our theory. In particular, it is concretely demonstrated that the free-component bounds can strictly outperform the corresponding min-eigenvalue bounds in Ref. [12] and notably, could be tight, in key scenarios.

Example 1 (Magic state distillation): Consider T states $|T\rangle = T|+\rangle = (|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$ contaminated by depolarizing or dephasing noise, given by

$$\tau = (1 - \zeta)|T\rangle\langle T| + \zeta \frac{I}{2}, \quad (16)$$

where ζ is the noise rate, as the input. Note that we are interested in $\zeta \in (0, 1 - 1/\sqrt{2})$ so that τ is a mixed state

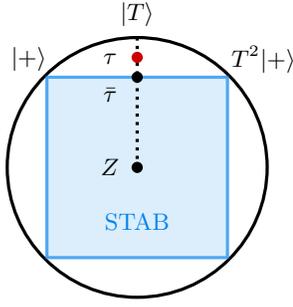


FIG. 1. The cross section of the Bloch sphere through the center perpendicular to the Z axis. The blue square represents the corresponding cross section of the stabilizer hull (STAB). $\bar{\tau}$ is actually $|T\rangle$ subject to $p = 1 - 1/\sqrt{2}$ depolarizing noise and lies on the edge of STAB. τ is the noisy input state that lies between $|T\rangle$ and $\bar{\tau}$.

outside of the stabilizer hull. On the one hand, it can be directly checked that $\lambda_{\bar{\tau}}^{\min} = \zeta/2$. On the other hand, Γ_{τ} is bounded as follows. Consider the free state

$$\bar{\tau} = \frac{1}{2}(|+\rangle\langle +| + T^2|+\rangle\langle +|T^{\dagger 2}) = \frac{1}{4} \begin{pmatrix} 2 & 1-i \\ 1+i & 2 \end{pmatrix} \in \mathcal{F},$$

which sits at the edge of the stabilizer hull closest to $|T\rangle$ (as depicted in Fig. 1). Then by definition we have

$$\Gamma_{\tau} \geq \max\{\gamma : \tau - \gamma\bar{\tau} \geq 0\} = \max\left\{\gamma : \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \alpha \end{pmatrix} \geq 0\right\},$$

with $\alpha = (1/2)(1 - \gamma)$ and $\beta = (1 - \zeta/2)e^{-i\pi/4} - \gamma(1 - i/4)$. By solving the determinant we obtain that

$$\Gamma_{\tau} \geq (2 + \sqrt{2})\zeta, \quad (17)$$

when $\zeta \in (0, 1 - 1/\sqrt{2})$. This implies $\Gamma_{\tau} > \lambda_{\bar{\tau}}^{\min}$, and thus the previous error bound is outperformed for any pure target state by a constant factor. As a sanity check, the bound indeed approaches 1 as $\zeta \rightarrow 1 - 1/\sqrt{2}$, in contrast to the λ^{\min} bound. This indeed implies the expected phenomenon that the total distillation overhead blows up as τ approaches the stabilizer hull. In particular, for the standard task of distilling T states, we thus obtain an improved bound on the average overhead following the proof of Theorem 3 in Ref. [12].

Corollary 8. Consider the following common formulation of magic state distillation task: given n copies of noisy states τ [defined in Eq. (16)], output an m -qubit state σ such that $\text{Tr} \sigma_i T = \langle T | \sigma_i | T \rangle \geq 1 - \epsilon, \forall i = 1, \dots, m$ where $\sigma_i = \text{Tr}_i \sigma$ is the i th qubit, by some free (stabilizer-preserving) operation. Then n must satisfy

$$n \geq \left\lceil \log \frac{(4 - 2\sqrt{2})^m - 1}{(4 - 2\sqrt{2})^m m \epsilon} \right\rceil \left\lceil \log \frac{2 - \sqrt{2}}{2\zeta} \right\rceil^{-1}. \quad (18)$$

Proof. By applying the union bound, we have $\langle T^{\otimes m} | \sigma | T^{\otimes m} \rangle \geq 1 - m\epsilon$. Recall that $\Gamma_{\tau} \geq (2 + \sqrt{2})\zeta$ in Eq. (17) and notice that $f_{T^{\otimes m}} = (4 - 2\sqrt{2})^{-m}$ [13,21–23]. By plugging everything into Eq. (12) we obtain the claimed bound. ■

Example 2 (Coherence): Consider the maximally coherent qubit state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ contaminated by typical noise channels, including depolarizing, dephasing, and amplitude damping, as the input.

For the depolarizing noise (here the dephasing noise has an equivalent effect), the noisy state is given by $\rho = [|0\rangle\langle 0| + |1\rangle\langle 1| + (1 - \mu)(|0\rangle\langle 1| + |1\rangle\langle 0|)]/2$ where $\mu \in (0, 1)$ is the noise rate. Then $\lambda_{\rho}^{\min} = \mu/2$, and it can be easily calculated that $\Gamma_{\rho} = \mu$. That is, the new error bound is twice the min-eigenvalue bound for any pure target state.

For the amplitude damping noise, the free-component bounds have a more remarkable advantage. Here the noisy state is given by $\rho = [(1 + \nu)|0\rangle\langle 0| + (1 - \nu)|1\rangle\langle 1| + \sqrt{1 - \nu}(|0\rangle\langle 1| + |1\rangle\langle 0|)]/2$ where $\nu \in (0, 1)$ is the noise rate. Then $\lambda_{\rho}^{\min} = (1/2)(1 - \sqrt{1 - \nu + \nu^2})$. We numerically solve Γ_{ρ} , and compare it with λ_{ρ}^{\min} in Fig. 2(b) (note that the values plotted are all multiplied by a factor 1/2; see below). Note that, as ν increases, i.e., ρ is more heavily damped towards the free state $|0\rangle$, the error of purification is expected to grow. As can be seen from Fig. 2(b), as $\nu \rightarrow 1$, λ_{ρ}^{\min} vanishes and so do corresponding bounds, but Γ_{ρ} indeed keeps growing, showcasing an important scenario where only the free-component bounds are nontrivial.

Let us explicitly consider $|+\rangle$ as the target state. It is known that the optimal fidelity of transforming ρ to $|+\rangle$

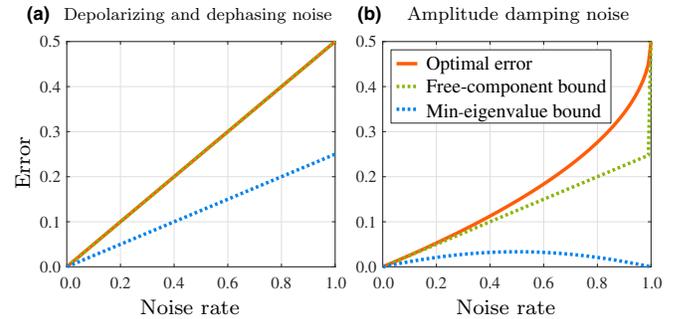


FIG. 2. Comparisons between the optimal achievable error of a standard purification task and the lower bounds induced by Γ (this work) and λ^{\min} (Ref. [12]) in coherence theory. The task is to recover the maximally coherent qubit state $|+\rangle$ under typical noise channels: (a) depolarizing and dephasing; (b) amplitude damping. The green and blue dashed lines are, respectively, the free component and min-eigenvalue lower bounds on the error, and the red line is the minimum error achieved by MIO computed by SDP Eq. (19). In (a) the green dashed line actually overlaps with the red line, indicating that the free-component error bound is tight.

by free operations (MIO) can be solved by the following SDP [24, Theorem 3]:

$$\max \left\{ \text{Tr } G\rho : 0 \leq G \leq I, \Delta(G) = \frac{I}{2} \right\}, \quad (19)$$

where Δ takes the diagonal part of a given matrix. In Fig. 2, we plot the optimal error obtained by the above SDP as well as the free component and min-eigenvalue lower bounds for comparison. In particular, for depolarizing or dephasing noise, the free-component error bound turns out to be tight, i.e., is achievable, for any noise rate.

Example 3 (Constrained quantum error correction): Here we demonstrate how the state no-purification bounds can be used to find limits on quantum error correction (QEC). In particular, we consider the broadly important situations where the QEC procedures are subject to certain constraints (such as stabilizer or Clifford constraints, symmetries) so that resource theory becomes useful. Notice that the decoding procedures are aimed at recovering all logical states from noisy physical states, indicating connections between the no-purification bounds and the overall recovery accuracy. More specifically, we have the general result (L, S denote the logical and physical systems, respectively).

Corollary 9. *Suppose the decoding operation is free. Then given encoding operation $\mathcal{E}_{L \rightarrow S}$ and noise channel \mathcal{N}_S acting on the physical system S , the error of the recovery of pure logical state ψ_L obeys $\varepsilon \geq \Gamma_{\mathcal{N}_S \circ \mathcal{E}_{L \rightarrow S}(\psi_L)}(1 - f_{\psi_L})$, based on which we directly obtain bounds on measures of the overall accuracy of the code, such as the worst-case error given by maximization over ψ_L , and the average-case error given by a certain (e.g., Haar) average over ψ_L .*

We further remark on the case of covariant (symmetry-constrained) codes, which play fundamental roles in quantum computing and physics and has drawn considerable recent interest [25–31]. Suppose we consider some compact continuous symmetry group G . Based on Lemma 2 in Ref. [29,32], it can be seen that when the noise channel \mathcal{N}_S is covariant (which is usually the case), then we can construct a covariant decoding operation that achieves the optimal error. That is, we can actually remove the freeness assumption of the decoder to apply the no-purification bounds, leading to the following adapted version.

Corollary 10 (Covariant code). *Let G be a compact continuous symmetry group. Let $\mathcal{E}_{L \rightarrow S}$ be a G -covariant encoding operation. Suppose the noise channel \mathcal{N}_S is G -covariant. Then Corollary 9 (where the parameters are defined in terms of the G -asymmetry theory) holds for any decoder.*

See Sec. III B 2 for related discussions and results in the channel setting.

Example 4 (Continuous variable): Lastly, we provide an elementary example of the application to continuous-variable theories. Consider continuous-variable nonclassicality, a characteristic resource feature in quantum optics that is closely relevant to, e.g., linear optical quantum computation [33] and metrology [34–36]. Here the coherent states of light and their probabilistic mixtures are considered free (classical). The coherent state corresponding to complex amplitude $\alpha \in \mathbb{C}$ takes the form

$$|\alpha\rangle = \exp(-|\alpha|^2/2) \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (20)$$

in the number state (Fock) basis $\{|n\rangle\}$. A prototypical type of nonclassical resource states is the (single-mode) squeezed states [37,38]

$$|s_r\rangle := \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} (\tanh r)^n |2n\rangle \quad (21)$$

generated by the squeezing operator $S(r) := \exp\{r[\hat{a}^2 - (\hat{a}^\dagger)^2]/2\}$ (\hat{a} and \hat{a}^\dagger are, respectively, the annihilation and creation operators) acting on the vacuum state $|0\rangle$, where $r \geq 0$ is the squeezing parameter. It can be calculated that

$$\langle s_r | \alpha \rangle = \sqrt{\frac{e^{-|\alpha|^2}}{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} (\tanh r)^n \times \frac{\alpha^{2n}}{\sqrt{(2n)!}} \quad (22)$$

$$= \sqrt{\frac{e^{-|\alpha|^2}}{\cosh r}} \sum_{n=0}^{\infty} \frac{[(\alpha^2 \tanh r)/2]^n}{n!} \quad (23)$$

$$= \sqrt{\frac{e^{-|\alpha|^2}}{\cosh r}} e^{(\alpha^2 \tanh r)/2}, \quad (24)$$

using which we obtain

$$f_{s_r} := \sup_{\sigma \in \mathcal{F}} \langle s_r | \sigma | s_r \rangle = \sup_{\alpha \in \mathbb{C}} |\langle s_r | \alpha \rangle|^2 \quad (25)$$

$$= \sup_{\alpha \in \mathbb{C}} \frac{e^{-|\alpha|^2}}{\cosh r} e^{\text{Re}(\alpha^2) \tanh r} \quad (26)$$

$$= (\cosh r)^{-1}, \quad (27)$$

where we used $\tanh r < 1$. Then, to showcase an example of a no-purification bound, consider the task of distilling some squeezed state $|s_r\rangle$ from noisy state ρ using free, namely classicality-preserving operations (which, in particular, include passive linear optical operations) [35,39]. Then Theorem 4 directly implies that the transformation error $\varepsilon \geq \Gamma_\rho[1 - (\cosh r)^{-1}]$, from which it can be

observed that the task indeed becomes more demanding as the squeezing parameter increases. Like the discrete-variable setting, for specific noise models, it is often easy to calculate or bound Γ_ρ so that the error bound can be further specified.

III. CHANNEL THEORY

We now extend the no-purification theory to quantum channels or dynamical settings. The channel analog of purification is to transform a noisy channel (or noisy channels, as will be discussed) to a unitary (noiseless) channel, or equivalently, to simulate the unitary channel by the noisy ones. The free-component approach directly enables us to study these problems in the channel resource theory setting where the resource objects are quantum channels instead of states (note that it is not clear how to fully extend the hypothesis testing approach in Ref. [12] to channels). It is worth noting again that the structure of channel theories is much richer than the state theories since multiple instances of channels can be used in different ways, such as in parallel, sequentially, or adaptively. Here, we first present error bounds in the most general forms, and then specifically investigate the adaptive or sequential simulation setting, which represents a fundamental difference from state theories. To demonstrate the practical relevance of the general no-go rules and bounds, we discuss them in more specific contexts, and, in particular, outline the applications to quantum error correction, gate and circuit synthesis, and channel capacities.

Note that we often specify the input and output systems of channels in the subscripts (a channel \mathcal{N} from system A to system B is denoted as $\mathcal{N}_{A \rightarrow B}$, and if the input and output systems are the same one A it is simply denoted as \mathcal{N}_A), but when there is no ambiguity we omit the labels. Given linear maps \mathcal{N}, \mathcal{M} , the order $\mathcal{N} - \mathcal{M} \geq 0$ means $\mathcal{N} - \mathcal{M}$ is a completely positive map. To simplify the notation, given some input state ρ on A and reference system R , we also denote the output state of the channel $\mathcal{N}_{A \rightarrow B}$ acting on A by

$$\rho_{\mathcal{N}} := \mathcal{N}_{A \rightarrow B} \otimes \text{id}_R(\rho_{AR}). \quad (28)$$

In particular, the Choi state of \mathcal{N} is given by

$$\Phi_{\mathcal{N}} := \mathcal{N}_{A \rightarrow B} \otimes \text{id}_R(\Phi_{AR}), \quad (29)$$

where $\Phi_{AR} = \sum_j |j\rangle_A |j\rangle_R / \sqrt{d}$ is the maximally entangled state between A and reference system R of the same dimension d .

A. General theory and results

1. Setups and basic error bounds

For channel resource theories, the building blocks analogous to free states and free operations are *free channels* and

free superchannels, where superchannels map channels to channels. Like the state case, we consider the most general framework where the free superchannels are required only to obey the *golden rule* that any free superchannel must map a free channel to another free channel. Note again that this golden rule gives rise to the largest possible set of superchannels that encompasses any legitimate set of free superchannels, so that the fundamental limits induced by it apply universally. We also assume the following two commonly held properties of the set of free channels (which we still denote by \mathcal{F}): (i) the composition of free channels [for channels there are two fundamental types of composition—parallel composition (represented by tensor product \otimes), and sequential composition (represented by \circ)] should be free, that is, if $\mathcal{N}_1, \mathcal{N}_2 \in \mathcal{F}$, then both $\mathcal{N}_1 \otimes \mathcal{N}_2 \in \mathcal{F}$ and $\mathcal{N}_2 \circ \mathcal{N}_1 \in \mathcal{F}$ hold; (ii) \mathcal{F} is closed. We refer readers to, e.g., Refs. [40,41] for more comprehensive discussions of the general framework of channel resource theories.

We now define the channel version of free component as follows.

Definition 11 (Channel free component). *The free component of quantum channel \mathcal{N} is defined as*

$$\Gamma_{\mathcal{N}} := \max \{ \gamma : \mathcal{N} - \gamma \mathcal{M} \geq 0, \mathcal{M} \in \mathcal{F} \}. \quad (30)$$

Equivalently,

$$\Gamma_{\mathcal{N}} = \max \{ \gamma : \mathcal{N} = \gamma \mathcal{M} + (1 - \gamma) \mathcal{R}, \mathcal{M} \in \mathcal{F}, \mathcal{R} \in \mathcal{C} \}, \quad (31)$$

where \mathcal{C} is the set of all completely positive and trace-preserving maps (quantum channels). Since $\mathcal{N} \geq \gamma \mathcal{M}$ is equivalent to $\Phi_{\mathcal{N}} \geq \gamma \Phi_{\mathcal{M}}$, we also have the relation

$$\Gamma_{\mathcal{N}} = \Gamma_{\Phi_{\mathcal{N}}}, \quad (32)$$

where on the RHS, $\Phi_{\mathcal{N}}$ is the Choi state of \mathcal{N} and the free component Γ is defined with respect to the set of free states consisting of the Choi states of all free channels. Similar to the state case, as long as \mathcal{F} can be characterized by semidefinite conditions, the channel free component $\Gamma_{\mathcal{N}}$ can be efficiently computed by SDP.

The channel free component also exhibits monotonicity and super-multiplicity properties.

Proposition 12 (Monotonicity). *For any quantum channel \mathcal{N} and free superchannel Π , it holds that*

$$\Gamma_{\mathcal{N}} \leq \Gamma_{\Pi(\mathcal{N})}. \quad (33)$$

Proof. Suppose $\Gamma_{\mathcal{N}}$ is achieved by $\mathcal{M} \in \mathcal{F}$. Then by definition $\mathcal{N} - \Gamma_{\mathcal{N}} \mathcal{M} \geq 0$, and thus $\Pi(\mathcal{N}) - \Gamma_{\mathcal{N}} \Pi(\mathcal{M}) \geq 0$. Since $\Pi(\mathcal{M}) \in \mathcal{F}$ by the golden rule, we have that $\Gamma_{\Pi(\mathcal{N})} \geq \Gamma_{\mathcal{N}}$ by definition. ■

For channels, we need to consider sequential composition in addition to parallel composition represented by tensor product. The channel free component is supermultiplicative under both types of composition.

Proposition 13 (Supermultiplicity). *For any quantum channels $\mathcal{N}_1, \mathcal{N}_2$, it holds that*

$$\Gamma_{\mathcal{N}_1 \otimes \mathcal{N}_2} \geq \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2}, \quad (34)$$

$$\Gamma_{\mathcal{N}_2 \circ \mathcal{N}_1} \geq \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2}. \quad (35)$$

Proof. Suppose that the maximization in $\Gamma_{\mathcal{N}_1}, \Gamma_{\mathcal{N}_2}$ are, respectively, achieved by $\mathcal{M}_1, \mathcal{M}_2 \in \mathcal{F}$, that is, $\mathcal{N}_i \geq \Gamma_{\mathcal{N}_i} \mathcal{M}_i, i = 1, 2$. It holds that $\mathcal{N}_1 \otimes \mathcal{N}_2 \geq (\Gamma_{\mathcal{N}_1} \mathcal{M}_1) \otimes (\Gamma_{\mathcal{N}_2} \mathcal{M}_2) = \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2} \mathcal{M}_1 \otimes \mathcal{M}_2$, and similarly, $\mathcal{N}_2 \circ \mathcal{N}_1 \geq (\Gamma_{\mathcal{N}_2} \mathcal{M}_2) \circ (\Gamma_{\mathcal{N}_1} \mathcal{M}_1) = \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2} \mathcal{M}_2 \circ \mathcal{M}_1$. Also note that $\mathcal{M}_1 \otimes \mathcal{M}_2 \in \mathcal{F}$ and $\mathcal{M}_1 \circ \mathcal{M}_2 \in \mathcal{F}$ axiomatically. Therefore, $\Gamma_{\mathcal{N}_1 \otimes \mathcal{N}_2} \geq \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2}$ and $\Gamma_{\mathcal{N}_2 \circ \mathcal{N}_1} \geq \Gamma_{\mathcal{N}_1} \Gamma_{\mathcal{N}_2}$ by definition. ■

Here we are interested in the channel simulation task of transforming a given quantum channel \mathcal{N} to a target unitary channel \mathcal{U} via some superchannel up to some error that is measured by certain choices of channel distances. Let $F(\rho, \sigma) = \|\sqrt{\rho} \sqrt{\sigma}\|_1^2$ be the Uhlmann fidelity between general states ρ and σ . Consider the following three typical versions of channel fidelity that are commonly used.

(a) Worst-case (entanglement) fidelity:

$$F_W(\mathcal{N}, \mathcal{M}) := \inf_{\rho_{AR}} F(\rho_{\mathcal{N}}, \rho_{\mathcal{M}}), \quad (36)$$

where $\rho_{\mathcal{N}}, \rho_{\mathcal{M}}$ are, respectively, the channel output states of \mathcal{N}, \mathcal{M} defined in Eq. (28), and the optimization includes system R . Note that it is equivalent to optimize over pure input states due to the joint concavity of fidelity F [42].

(b) Choi fidelity:

$$F_C(\mathcal{N}, \mathcal{M}) := F(\Phi_{\mathcal{N}}, \Phi_{\mathcal{M}}), \quad (37)$$

where $\Phi_{\mathcal{N}}, \Phi_{\mathcal{M}}$ are, respectively, the Choi states of \mathcal{N}, \mathcal{M} .

(c) Average-case fidelity [43]:

$$F_A(\mathcal{N}, \mathcal{M}) := \int d\psi F(\mathcal{N}(\psi), \mathcal{M}(\psi)), \quad (38)$$

where the integral is over the Haar measure on the input state space.

The corresponding versions of infidelity are then

$$\varepsilon_x(\mathcal{N}, \mathcal{M}) := 1 - F_x(\mathcal{N}, \mathcal{M}), \quad x \in \{W, C, A\}. \quad (39)$$

Also, a standard measure of distance between channels is given by the diamond norm distance:

$$\varepsilon_{\diamond}(\mathcal{N}, \mathcal{M}) := \frac{1}{2} \|\mathcal{N} - \mathcal{M}\|_{\diamond}, \quad (40)$$

where $\|\mathcal{N}\|_{\diamond} := \sup_{\rho_{AR}} \|\mathcal{N}_{A \rightarrow B} \otimes \text{id}_R(\rho_{AR})\|_1$. Again, it is equivalent to optimize over pure input states due to the convexity of trace norm $\|\cdot\|_1$. All the above channel distance measures are symmetric in its arguments.

Note that these channel distance measures are commonly used in different scenarios [43]. For example, the worst-case entanglement fidelity and the diamond norm error are commonly used in quantum computation scenarios like circuit synthesis (see, e.g., Refs. [44,45], Sec. III B 4) and approximate quantum error correction (see, e.g., Ref. [46], Sec. III B 2); the Choi fidelity is used in quantum Shannon theory to evaluate the performance of quantum communication (see, e.g., Refs. [47,48], Sec. III B 3); the average-case fidelity is easier to estimate in experiments (see, e.g., Refs. [49–52]).

In this work, we are mostly interested in the case where an argument is a unitary channel \mathcal{U} . Note that for pure state ψ , we have the inequality [3]

$$1 - F(\rho, \psi) \leq \frac{1}{2} \|\rho - \psi\|_1 \leq \sqrt{1 - F(\rho, \psi)}. \quad (41)$$

Applying the above result to channels, we can conclude

$$\varepsilon_W(\mathcal{N}, \mathcal{U}) \leq \varepsilon_{\diamond}(\mathcal{N}, \mathcal{U}) \leq \sqrt{\varepsilon_W(\mathcal{N}, \mathcal{U})}. \quad (42)$$

Also, it is known [43,53] that the average-case fidelity and the Choi fidelity have the following direct relation:

$$F_A(\mathcal{N}, \mathcal{U}) = \frac{F_C(\mathcal{N}, \mathcal{U})d + 1}{d + 1}, \quad (43)$$

and thus

$$\varepsilon_A(\mathcal{N}, \mathcal{U}) = \frac{d}{d + 1} \varepsilon_C(\mathcal{N}, \mathcal{U}), \quad (44)$$

where d is the dimension of the input system. Furthermore, it is clear from definition that

$$F_C(\mathcal{N}, \mathcal{M}) \geq F_W(\mathcal{N}, \mathcal{M}), \quad (45)$$

$$\varepsilon_C(\mathcal{N}, \mathcal{M}) \leq \varepsilon_W(\mathcal{N}, \mathcal{M}), \quad (46)$$

for any channels \mathcal{N}, \mathcal{M} . To summarize, for the case of comparing with unitary channel \mathcal{U} , which is of interest in

this work, the four channel distance measures are ordered as follows:

$$\varepsilon_{\diamond}(\mathcal{N}, \mathcal{U}) \geq \varepsilon_W(\mathcal{N}, \mathcal{U}) \geq \varepsilon_C(\mathcal{N}, \mathcal{U}) = \frac{d+1}{d} \varepsilon_A(\mathcal{N}, \mathcal{U}). \quad (47)$$

We are interested in the task of using channel \mathcal{N} to simulate unitary target channel \mathcal{U} via transformation superchannel Π . The (different versions of) simulation error is simply given by

$$\varepsilon_x(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) := \varepsilon_x(\Pi(\mathcal{N}), \mathcal{U}), \quad x \in \{W, C, A, \diamond\}. \quad (48)$$

Also define corresponding versions of the maximum overlap of channel \mathcal{N} with free channels as

$$f_{\mathcal{N}}^x := \max_{\mathcal{M} \in \mathcal{F}} F_x(\mathcal{N}, \mathcal{M}), \quad x \in \{W, C, A\}. \quad (49)$$

Note the following simple fact.

Proposition 14 (Faithfulness). *For any quantum channels \mathcal{N} and \mathcal{M} , for $x \in \{W, C, A\}$,*

$$F_x(\mathcal{N}, \mathcal{M}) = 1 \iff \mathcal{N} = \mathcal{M}, \quad (50)$$

and as a consequence,

$$f_{\mathcal{N}}^x = 1 \iff \mathcal{N} \in \mathcal{F}. \quad (51)$$

Proof. The first equivalence follows from the fact of state fidelity that $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. The second equivalence follows since \mathcal{F} is closed by assumption. ■

We now present error bounds for these channel error measures. For the Choi and average-case fidelities, note the following linearity property.

Lemma 15 (Linearity). *Let $x \in \{C, A\}$ and \mathcal{U} be a unitary channel. Then $F_x(\mathcal{N}, \mathcal{U})$ is linear in \mathcal{N} . That is, given $\mathcal{N} = p\mathcal{N}_1 + (1-p)\mathcal{N}_2$ for $p \in [0, 1]$ and quantum channels $\mathcal{N}_1, \mathcal{N}_2$, it holds that*

$$F_x(\mathcal{N}, \mathcal{U}) = pF_x(\mathcal{N}_1, \mathcal{U}) + (1-p)F_x(\mathcal{N}_2, \mathcal{U}). \quad (52)$$

Proof. Consider the Choi fidelity first. We have

$$\begin{aligned} F_C(p\mathcal{N}_1 + (1-p)\mathcal{N}_2, \mathcal{U}) \\ = F(\Phi_{p\mathcal{N}_1 + (1-p)\mathcal{N}_2}, \Phi_{\mathcal{U}}) \end{aligned} \quad (53)$$

$$= \text{Tr}(\Phi_{p\mathcal{N}_1 + (1-p)\mathcal{N}_2} \Phi_{\mathcal{U}}) \quad (54)$$

$$= p \text{Tr}(\Phi_{\mathcal{N}_1} \Phi_{\mathcal{U}}) + (1-p) \text{Tr}(\Phi_{\mathcal{N}_2} \Phi_{\mathcal{U}}) \quad (55)$$

$$= pF_C(\mathcal{N}_1, \mathcal{U}) + (1-p)F_C(\mathcal{N}_2, \mathcal{U}), \quad (56)$$

where the second equality follows since the Choi state $\Phi_{\mathcal{U}}$ is a pure state, and the third equality follows from the

linearity of the trace function. Then due to Eq. (43), we conclude that F_A has the same linearity property. ■

Collectively, our best bounds are the following.

Theorem 16. *Given any quantum channel \mathcal{N} and any unitary target channel \mathcal{U} , it holds for any free superchannel Π that*

$$\begin{aligned} \varepsilon_{\diamond}(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) &\geq \varepsilon_W(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) \\ &\geq \varepsilon_C(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) \geq \Gamma_{\mathcal{N}}(1 - f_{\mathcal{U}}^C), \end{aligned} \quad (57)$$

and

$$\varepsilon_A(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) \geq \Gamma_{\mathcal{N}}(1 - f_{\mathcal{U}}^A) = \frac{d}{d+1} \Gamma_{\mathcal{N}}(1 - f_{\mathcal{U}}^C), \quad (58)$$

where d is the dimension of the input system of \mathcal{U} .

Proof. The proof is analogous to that of Theorem 4. By the definition of $\Gamma_{\mathcal{N}}$, there exists free channel $\mathcal{M} \in \mathcal{F}$ and channel \mathcal{R} such that \mathcal{N} can be decomposed as follows:

$$\mathcal{N} = \Gamma_{\mathcal{N}}\mathcal{M} + (1 - \Gamma_{\mathcal{N}})\mathcal{R}. \quad (59)$$

Let Π be any free superchannel. By the linearity of superchannels,

$$\Pi(\mathcal{N}) = \Gamma_{\mathcal{N}}\Pi(\mathcal{M}) + (1 - \Gamma_{\mathcal{N}})\Pi(\mathcal{R}). \quad (60)$$

Then for $x \in \{C, A\}$, it holds that

$$\begin{aligned} \varepsilon_x(\mathcal{N} \xrightarrow{\Pi} \mathcal{U}) \\ = 1 - F_x(\Pi(\mathcal{N}), \mathcal{U}) \end{aligned} \quad (61)$$

$$= 1 - F_x(\Gamma_{\mathcal{N}}\Pi(\mathcal{M}) + (1 - \Gamma_{\mathcal{N}})\Pi(\mathcal{R}), \mathcal{U}) \quad (62)$$

$$= 1 - \Gamma_{\mathcal{N}}F_x(\Pi(\mathcal{M}), \mathcal{U}) - (1 - \Gamma_{\mathcal{N}})F_x(\Pi(\mathcal{R}), \mathcal{U}) \quad (63)$$

$$\geq 1 - \Gamma_{\mathcal{N}}f_{\mathcal{U}}^x - (1 - \Gamma_{\mathcal{N}}) \quad (64)$$

$$= \Gamma_{\mathcal{N}}(1 - f_{\mathcal{U}}^x), \quad (65)$$

where the third line follows from the linearity property Lemma 15, and the inequality follows from the fact that $F_C(\Pi(\mathcal{M}), \mathcal{U}) \leq f_{\mathcal{U}}^C$ since $\Pi(\mathcal{M}) \in \mathcal{F}$ by the golden rule, and $F_C(\Pi(\mathcal{R}), \mathcal{U}) \leq 1$. Then by Eq. (47) we obtain Eq. (57), and Eq. (58) follows from the relation Eq. (43) and is essentially the same bound as the last one of Eq. (57) upto a dimension factor. ■

Note that the best bounds we can get for all error measures are in terms of the Choi overlap $f_{\mathcal{U}}^C$. A natural question is whether one can directly use $f_{\mathcal{U}}^W$ in the bound for ε_W , which would improve the bound. The problem is we do not have a linearity property analogous to Lemma 15

for the worst-case fidelity F_W , so the third line does not go through.

As long as the target channel $\mathcal{U} \notin \mathcal{F}$, it is clear by definition that $f_{\mathcal{U}}^C < 1$. That is, for any channel \mathcal{N} satisfying the $\Gamma_{\mathcal{N}} > 0$ condition and any resource unitary channel, all the above error bounds are nontrivial and thus imply a nonzero error.

2. Multiple channel uses and adaptive channel simulation

Now we discuss the scenario where one takes multiple noisy channels as inputs and intends to simulate some unitary channel, which is analogous to the standard task of distilling high-quality resources from many noisy resources in the state setting. However, the multiple instance setting represents a very important difference between channels and states. The composition of multiple states has a simple parallel structure represented by tensor products. In contrast, multiple channels can be used sequentially and adaptively, which is not simply described by tensor products and may be more powerful than the parallel scheme. Whether the adaptive scheme can outperform the parallel one is a crucial problem in many research areas about quantum channels, such as channel simulation, discrimination, and estimation (see, e.g., Refs. [54–68]).

First, note that the parallel use of multiple channels $\mathcal{N}_1, \dots, \mathcal{N}_n$ is again represented by tensor product and thus can be simply regarded as a single channel $\mathcal{N} = \bigotimes_{i=1}^n \mathcal{N}_i$. Therefore, the results above can be directly applied. In addition to error bounds, using the supermultiplicity property (Proposition 13), we directly bound the cost or overhead of unitary channel simulation, defined by the number of instances of a certain channel needed to simulate some unitary channel, using parallel strategies.

Corollary 17 (Parallel simulation cost). *Suppose some free superchannel Π transforms n instances of noisy channels \mathcal{N} to target unitary channel \mathcal{U} with a certain type of error $\epsilon_x(\mathcal{N}^{\otimes n} \xrightarrow{\Pi} \mathcal{U}) \leq \epsilon_x, x \in \{\diamond, W, C, A\}$. Then n must satisfy*

$$n \geq \left\lceil \log \frac{1 - f_{\mathcal{U}}^C}{\epsilon_x} \right\rceil \left\lceil \log \frac{1}{\Gamma_{\mathcal{N}}} \right\rceil^{-1}, \quad (66)$$

for any $x \in \{\diamond, W, C\}$. The bound on n in terms of the average-case error ϵ_A is equivalent to that in terms of the Choi error ϵ_C .

Now we consider the adaptive scheme, which represents a more general way to use multiple input channels to simulate an output channel. Here, the action on input channels $\mathcal{N}_1, \dots, \mathcal{N}_n$ is represented by a “quantum comb” [69,70] Π with appropriate dimensions realized by channels $\mathcal{P}_1, \dots, \mathcal{P}_{n+1}$, and the input channels are inserted into

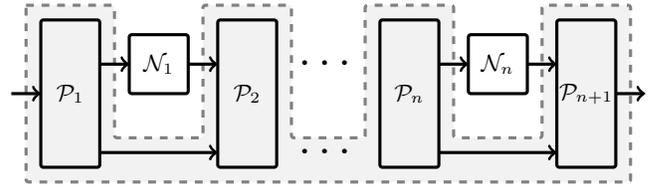


FIG. 3. Quantum comb. Given input channels $\mathcal{N}_1, \dots, \mathcal{N}_n$, the general map that outputs a channel can be represented by a quantum comb (gray area) realized by channels $\mathcal{P}_1, \dots, \mathcal{P}_{n+1}$, and the input channels are used by inserting them into the slots.

the slots (as depicted in Fig. 3). In resource theory contexts, there is again a golden rule on the combs that a free comb must map free channels to a free channel, that is, if one inserts free channels $\mathcal{N}_1, \dots, \mathcal{N}_n \in \mathcal{F}$ in the slots of comb Π_n then the overall channel $\Pi_n(\mathcal{N}_{[n]}) \in \mathcal{F}$ (where $\mathcal{N}_{[n]}$ is short for the channel collection $[\mathcal{N}_1, \dots, \mathcal{N}_n]$). Note that, in the case where the comb is realized by free channels $\mathcal{P}_1, \dots, \mathcal{P}_{n+1} \in \mathcal{F}$ (and the identities on the ancilla systems are considered free), it obviously obey the golden rule, because axiomatically the composition of free channels is free. However, the converse is not necessarily true, that is, the notion of free combs is more general than free realization.

Note that the channel free component obeys the following monotonicity property under free combs.

Proposition 18 (Monotonicity). *Given any channels $\mathcal{N}_1, \dots, \mathcal{N}_n$ (collectively denoted by $\mathcal{N}_{[n]}$), it holds that, for any free comb Π_n acting on $\mathcal{N}_{[n]}$,*

$$\Gamma_{\Pi_n(\mathcal{N}_{[n]})} \geq \prod_{i=1}^n \Gamma_{\mathcal{N}_i}. \quad (67)$$

Proof. Suppose the quantum comb Π_n is realized by channels \mathcal{P}_i with $i = 1, \dots, n+1$, as depicted in Fig. 3. We emphasize that \mathcal{P}_i are not necessarily free channels themselves; the only requirement here is that the whole comb obeys the golden rule, i.e., $\Pi_n(\mathcal{M}_{[n]}) \in \mathcal{F}$ as long as $\mathcal{M}_{[n]} = [\mathcal{M}_1, \dots, \mathcal{M}_n]$ and $\mathcal{M}_i \in \mathcal{F}$ for $i \in [n]$. Suppose that the maximization in $\Gamma_{\mathcal{N}_i}$ is achieved by $\mathcal{M}_i \in \mathcal{F}$, that is, $\mathcal{N}_i \geq \Gamma_{\mathcal{N}_i} \mathcal{M}_i, \forall i$. Then we have

$$\Pi_n(\mathcal{N}_{[n]}) = \mathcal{P}_{n+1} \prod_{i=1}^n (\mathcal{N}_i \otimes \text{id}_i) \circ \mathcal{P}_i \quad (68)$$

$$\geq \mathcal{P}_{n+1} \prod_{i=1}^n \Gamma_{\mathcal{N}_i} (\mathcal{M}_i \otimes \text{id}_i) \circ \mathcal{P}_i \quad (69)$$

$$= \prod_{i=1}^n \Gamma_{\mathcal{N}_i} \Pi_n(\mathcal{M}_{[n]}), \quad (70)$$

where the inequality follows from the fact that channel tensorizations and compositions preserve the channel order $\mathcal{N}_i \geq \Gamma_{\mathcal{N}_i} \mathcal{M}_i$. Since $\Pi_n(\mathcal{M}_{[n]}) \in \mathcal{F}$ by the golden rule, we get $\Gamma_{\Pi_n(\mathcal{N}_{[n]})} \geq \prod_{i=1}^n \Gamma_{\mathcal{N}_i}$ by definition. ■

Now for input channels $\mathcal{N}_{[n]} = [\mathcal{N}_1, \dots, \mathcal{N}_n]$, comb Π_n and unitary target channel \mathcal{U} , the simulation error is defined as

$$\varepsilon_x(\mathcal{N}_{[n]} \xrightarrow{\Pi_n} \mathcal{U}) := \varepsilon_x(\Pi_n(\mathcal{N}_{[n]}), \mathcal{U}), \quad (71)$$

for $x \in \{\diamond, W, C, A\}$.

By a little tweak of the proofs above, we establish bounds on the error and cost for adaptive simulation, which match those for the parallel case.

Corollary 19 (Adaptive simulation error). *Given any channels $\mathcal{N}_1, \dots, \mathcal{N}_n$ (collectively denoted by $\mathcal{N}_{[n]}$) and any unitary target channel \mathcal{U} , it holds that, for any free comb Π_n acting on $\mathcal{N}_{[n]}$,*

$$\begin{aligned} \varepsilon_\diamond(\mathcal{N}_{[n]} \xrightarrow{\Pi_n} \mathcal{U}) &\geq \varepsilon_W(\mathcal{N}_{[n]} \xrightarrow{\Pi_n} \mathcal{U}) \\ &\geq \varepsilon_C(\mathcal{N}_{[n]} \xrightarrow{\Pi_n} \mathcal{U}) \geq \prod_{i=1}^n \Gamma_{\mathcal{N}_i} (1 - f_{\mathcal{U}}^C), \end{aligned} \quad (72)$$

and

$$\begin{aligned} \varepsilon_A(\mathcal{N}_{[n]} \xrightarrow{\Pi_n} \mathcal{U}) &\geq \prod_{i=1}^n \Gamma_{\mathcal{N}_i} (1 - f_{\mathcal{U}}^A) \\ &= \frac{d}{d+1} \prod_{i=1}^n \Gamma_{\mathcal{N}_i} (1 - f_{\mathcal{U}}^C), \end{aligned} \quad (73)$$

where d is the dimension of the input system of \mathcal{U} .

Proof. Simply note that, according to Eq. (70), we have the following decomposition:

$$\Pi_n(\mathcal{N}_{[n]}) = \prod_{i=1}^n \Gamma_{\mathcal{N}_i} \Pi_n(\mathcal{M}_{[n]}) + \left(1 - \prod_{i=1}^n \Gamma_{\mathcal{N}_i}\right) \mathcal{R}, \quad (74)$$

for some channel \mathcal{R} . By following the arguments in the proof of Theorem 16, one can establish similar error bounds where $\Gamma_{\mathcal{N}}$ is replaced by $\prod_{i=1}^n \Gamma_{\mathcal{N}_i}$. ■

Therefore, we can establish the same bound on the simulation cost for the adaptive scheme.

Corollary 20 (Adaptive simulation cost). *Suppose some free comb Π_n transforms n instances of noisy channels \mathcal{N} to target unitary channel \mathcal{U} with a certain type of error*

$\varepsilon_x([\mathcal{N}, \dots, \mathcal{N}] \xrightarrow{\Pi_n} \mathcal{U}) \leq \epsilon_x, x \in \{\diamond, W, C, A\}$. Then n must satisfy

$$n \geq \left\lceil \log \frac{1 - f_{\mathcal{U}}^C}{\epsilon_x} \right\rceil \left\lceil \log \frac{1}{\Gamma_{\mathcal{N}}} \right\rceil^{-1}, \quad (75)$$

for any $x \in \{\diamond, W, C\}$. The bound on n in terms of the average-case error ε_A is equivalent to that in terms of the Choi error ε_C .

Note that the adaptive strategies may potentially reduce the error or cost of simulation compared to parallel ones, so the adaptive simulation bounds can be regarded stronger.

A general observation is that the simulation cost asymptotically scales at least as $\Omega(\log(1/\epsilon_x))$ as target error $\epsilon_x \rightarrow 0$ even if we allow adaptive usages of the input channels, no matter which kind of error measure x is chosen.

3. No-purification conditions

Here, we discuss the situations where no-go rules are in place for channel resource purification, i.e., no unitary resource channels can be exactly simulated. For both the cases of single and multiple input channels, the basic statement goes as follows.

Corollary 21. *There is no free superchannel (or comb) that exactly transforms channel \mathcal{N} (or a collection of channels $\{\mathcal{N}_i\}$) to any unitary resource channel $\mathcal{U} \notin \mathcal{F}$ if $\Gamma_{\mathcal{N}} > 0$ (or $\Gamma_{\mathcal{N}_i} > 0, \forall i$).*

Proof. Since \mathcal{F} is closed by assumption and $\mathcal{U} \notin \mathcal{F}$, we have $f_{\mathcal{U}} < 1$. Then due to Theorem 16 the transformation error (in whichever measure) is strictly positive, indicating that the exact transformation is impossible. ■

Now similar to Proposition 7, we give a series of alternative characterizations of the $\Gamma > 0$ condition for channels, which could be illustrative or useful in certain scenarios.

Proposition 22. *For any quantum channel \mathcal{N} , the following conditions are equivalent:*

- (a) channel free component: $\Gamma_{\mathcal{N}} > 0$;
- (b) state free component:
 - (b1) worst case: for any input state ρ , $\Gamma_{\rho_{\mathcal{N}}} > 0$ (defined with respect to the set of free states $\{\rho_{\mathcal{M}} : \mathcal{M} \in \mathcal{F}\}$);
 - (b2) Choi state: $\Gamma_{\Phi_{\mathcal{N}}} > 0$ (defined with respect to the set of free states $\{\Phi_{\mathcal{M}} : \mathcal{M} \in \mathcal{F}\}$);
- (c) support:
 - (c1) worst case: there exists $\mathcal{M} \in \mathcal{F}$ such that, for any ρ , $\text{supp}(\rho_{\mathcal{N}}) \supseteq \text{supp}(\rho_{\mathcal{M}})$;

(c2) *Choi state: there exists $\mathcal{M} \in \mathcal{F}$ such that $\text{supp}(\Phi_{\mathcal{N}}) \supseteq \text{supp}(\Phi_{\mathcal{M}})$;*

(d) *resource measure:*

(d1) *worst case: $\min_{\mathcal{M} \in \mathcal{F}} D_{\min}(\mathcal{N} \parallel \mathcal{M}) = 0$, where $D_{\min}(\mathcal{N} \parallel \mathcal{M}) := \sup_{\rho} D_{\min}(\rho_{\mathcal{N}} \parallel \rho_{\mathcal{M}})$ is the channel min-entropy between \mathcal{N} and \mathcal{M} [59, 71];*

(d2) *Choi state: $\min_{\mathcal{M} \in \mathcal{F}} D_{\min}(\Phi_{\mathcal{N}} \parallel \Phi_{\mathcal{M}}) = 0$.*

Proof. First it is clear that (b1), (c1), (d1) are equivalent, and (b2), (c2), (d2) are equivalent, due to the state theory result, Proposition 7. The equivalence between (b1), (b2), and (a) follows from the fact that $\mathcal{N} \geq \gamma \mathcal{M}$ is equivalent to $\Phi_{\mathcal{N}} \geq \gamma \Phi_{\mathcal{M}}$ as well as $\rho_{\mathcal{N}} \geq \gamma \rho_{\mathcal{M}}$ for all ρ . ■

Worth noting, in the channel theory, the counterparts of min-relative entropy monotones also nicely contrast noisy entities with pure ones.

B. Practical scenarios and applications

The above no-purification rules and bounds are given in general forms so that their range of applicability is as wide as possible. To provide some concrete understanding and guideline of their practical relevance, we now discuss some specific scenarios and applications of interest. We shall start with a general discussion on typical noise models and the corresponding no-purification bounds in the contexts of different kinds of channel resource theories, and then specifically consider the roles of no-purification bounds in the contexts of quantum error correction, quantum communication, and circuit synthesis. Note that the main objective of our discussion here is to establish the frameworks for linking the no-purification principles to these practical problems. We shall mostly present general-form bounds, which are expected to be crude for certain specific resource features, noise models, system features etc., leaving refined analyses elsewhere.

1. Channel resource theories and practical noises

At a high level, we have the following two major different types of channel resource theories, signified by the role of the identity channel.

- (a) **Information preservation theories.** In such theories, one is primarily interested in the noise channels and their abilities to simulate noiseless channels so as to preserve or transmit information. Typical scenarios include quantum error correction and quantum communication. A signature of such theories is that the identity channel (between certain systems) is an ideal resource channel, representing no error or loss of quantum information occurring. The set of free channels commonly involve e.g., certain constant (replacer) channels, which represent complete loss of information. Here the free channels are in

general directly induced by physical restrictions on the implementable operations that, e.g., perform the tasks of encoding and decoding.

- (b) **Resource generation theories.** Such theories are commonly based on some resource theory defined at the level of states (such as entanglement, coherence, magic states). The features of channels and simulation tasks of interest are related to their ability of generating the state resource. Here the set of free channels are derived from state theories and thus obey the resource nongenerating property (for example, the identity channel is axiomatically free). A typical scenario of this kind is synthesis, where a common task is to simulate, or “synthesize” some complicated target channel by elementary resource channels. See further discussions in the next part.

In some sense, theories of the first kind are intrinsically based on channels, and those of the second kind are induced by state theories. Such classification may help elucidate the interplay between channel and state resource theories.

Now we discuss typical noisy channels of interest in these two different kinds of channel resource theories.

First, consider the first kind, i.e., information preservation theories, where the identity channel id is a resource. Here, the simulation capabilities (capacities) of noise channels themselves are of interest. A general observation is that, for stochastic noise

$$\mathcal{N}_{\mu} = (1 - \mu)\text{id} + \mu\mathcal{N}, \quad (76)$$

where $\mu \in (0, 1)$ is the noise rate, if the noise channel \mathcal{N} is considered free in the theory in consideration, then $\Gamma_{\mathcal{N}_{\mu}} \geq \mu$, which can be directly used to establish bounds on simulation error and cost. We list a few important noise models that are special cases: (i) depolarizing noise: $\mathcal{N}(\rho) = I/d$ is just a constant channel that outputs the maximally mixed state; (ii) erasure noise: $\mathcal{N}(\rho) = |\perp\rangle\langle\perp|$ is also a constant channel that outputs an orthogonal garbage state; (for these two cases \mathcal{N} is normally free as it essentially erases information completely.) (iii) dephasing noise: $\mathcal{N} = \Delta$, which erases the off-diagonal entries and thus is typically free in quantum scenarios since all coherence-related information is lost; (iv) Pauli noise: $\mathcal{N}(\rho) = \sum_i \mu_i P_i \rho P_i$ where $\forall i \mu_i \geq 0$, $\sum_i \mu_i = \mu$ and P_i 's are nonidentity Pauli operators (note that this model encompasses the depolarizing and dephasing noises); here \mathcal{N} is a stabilizer operation, and thus the global Pauli noise has free component in the stabilizer theory, leading to limitations on stabilizer codes. We shall demonstrate the connections to quantum error correction in more detail in Sec. III B 2. Quantum communication is another important scenario of this kind, which we shall discuss more specifically in Sec. III B 3.

For the second kind, i.e., resource generation theories, the input channels of practical interest are usually not the noise channels themselves but the resource-generating channels contaminated by noises. For example, consider $\mathcal{N}_\mu \circ \mathcal{G} = (1 - \mu)\mathcal{G} + \mu\mathcal{N} \circ \mathcal{G}$ where \mathcal{N}_μ is a stochastic noise and \mathcal{G} is a noiseless resource-generating channel. Also note that, in contrast to the first kind, the theory is commonly built upon a clear notion of free states. Then a general observation for this case is that if \mathcal{N} always output a free state, then $\Gamma_{\mathcal{N}_\mu \circ \mathcal{G}} \geq \mu$. Again, this holds for the depolarizing and erasure noises in normal theories where the maximally mixed state and the garbage state are free (note that the bound can be loose in, e.g., magic theory; see Sec. III B 4). Then by definition, it also applies to dephasing noise in theories where the diagonal states are free (such as coherence and certain asymmetry theories). As mentioned, a particularly important problem in such theories is gate synthesis. In Sec. III B 4, we shall discuss the implications of our general results to practical synthesis problems in more detail.

Notably, certain communication problems and gate synthesis correspond to adaptive channel simulation, which cannot be understood in the single-channel or parallel simulation schemes.

2. Quantum error correction

As a cornerstone of quantum computing and information [3], quantum error correction (QEC) serves to reduce noise effects and errors in physical systems by the idea of encoding the quantum information in a suitable way so that after noise and errors occur the original logical information can be restored (decoded). It is clearly important to understand various kinds of limits on QEC. Our results here are relevant to the broadly important scenario where the QEC procedures and codes obey certain rules or constraints. Typical examples include the well-studied stabilizer codes [3,72], and covariant codes [25–31], which has recently drawn considerable interest in quantum computing and physics. In Sec. II we presented general limits on the QEC accuracy based on understanding the decoding as a purification task. Here the channel framework provides an alternate formulation: notice that the QEC task is essentially to simulate an identity channel on the logical system; then the channel no-purification bounds induce fundamental limits on this channel simulation task. As a result, we have the following general bounds on the accuracy and cost of constrained QEC when the system is subject to generic nonunitary noises (L, S denote the logical and physical systems, respectively):

Corollary 23 (Constrained quantum error correction). *Suppose that the encoder and decoder are free channels (subject to certain resource theory constraints) Π . Then given noise channel \mathcal{N}_S acting on the physical system*

S , the commonly considered overall error measures for approximate QEC $\varepsilon_{x,x} \in \{\diamond, W, C\}$ obey

$$\varepsilon_x(\mathcal{N}_S \xrightarrow{\Pi} \text{id}_L) \geq \Gamma_{\mathcal{N}_S}(1 - f_{\text{id}_L}^C). \quad (77)$$

For example, consider the natural independent noise model where the noise channel \mathcal{N} acts independently and uniformly on each subsystem (e.g., qubit), i.e., the overall noise channel has the form $\mathcal{N}_S = \mathcal{N}^{\otimes n}$. Then

$$\varepsilon_x(\mathcal{N}_S \xrightarrow{\Pi} \text{id}_L) \geq (\Gamma_{\mathcal{N}})^n(1 - f_{\text{id}_L}^C), \quad (78)$$

and therefore, to achieve target error $\varepsilon_x(\mathcal{N}_S \rightarrow \text{id}_L) \leq \epsilon_x$, the number of physical subsystems n obeys

$$n \geq \left\lceil \log \frac{1 - f_{\text{id}_L}^C}{\epsilon_x} \right\rceil \left\lceil \log \frac{1}{\Gamma_{\mathcal{N}}} \right\rceil^{-1}. \quad (79)$$

In the case of stochastic noise $\mathcal{N} = (1 - \mu)\text{id} + \mu\mathcal{M}$ where $\mathcal{M} \in \mathcal{F}$, $\Gamma_{\mathcal{N}}$ in the above bounds can be replaced by μ .

As previously noted, this general result applies to the important cases of stabilizer and covariant QEC, which, respectively, correspond to Clifford [72] and symmetry [29] constraints. Note again that, in covariant QEC, under the commonly held assumption that the noise channel \mathcal{N}_S is covariant, we have the stronger conclusion that the error bound Eq. (77) holds for any decoder, meaning that covariant codes are no better than $[\Gamma_{\mathcal{N}_S}(1 - f_{\text{id}_L}^C)]$ -correctable for any decoder [29, Lemma 2]. For independent Pauli and erasure noises in the stabilizer case, and depolarizing, dephasing, and erasure noises in the covariant case, $\Gamma_{\mathcal{N}}$ can be replaced by μ in the bounds.

The bounds here are given in the most general forms, indicating universal limitations on the accuracy and cost of constrained QEC schemes for any noise channel with free component like typical global noise channels, which are naturally important but underinvestigated in the context of QEC. It would be interesting to perform more refined analysis of the bounds for specific constraints and noise models, which we leave for future work.

3. Quantum communication and Shannon theory

The central problem in quantum Shannon theory is to determine the capability of quantum channels to reliably transmit information. Depending on the purpose of transmission (e.g., transmitting classical or quantum information) and the resources that can be used at hand, there are many different variants of channel capacities, each of which corresponds to a channel simulation task in the language of resource theory (see, e.g., Refs. [40,56,73–77]).

Here we discuss quantum capacities, which correspond to the task of transforming a given channel to an identity

channel between two distinct, distant parties (labs). Note that we need to distinguish the identity channel shared between distant labs from the local identity channel whose input and output systems belong to the same lab. The former is regarded as the ideal resource while the latter is completely free. In resource theory language, channel capacities are determined by the choice of free superchannels or combs Π , which correspond to specific coding strategies. Some important cases include the following [75–77]:

- (a) unassisted code: superchannel Π can be decomposed into an encoder $\mathcal{K}_{A \rightarrow A'}$ by Alice composed with a decoder $\mathcal{D}_{B \rightarrow B'}$ by Bob, i.e., $\Pi = \mathcal{D}_{B \rightarrow B'} \mathcal{K}_{A \rightarrow A'}$;
- (b) entanglement-assisted code: superchannel Π acts as $\Pi(\mathcal{N})(\rho_A) = \mathcal{D}_{B\bar{B} \rightarrow B'} \mathcal{N}_{A' \rightarrow B} \mathcal{K}_{A\bar{A} \rightarrow A'}(\rho_A \otimes \omega_{\bar{A}\bar{B}})$ with encoder $\mathcal{K}_{A\bar{A} \rightarrow A'}$, decoder $\mathcal{D}_{B\bar{B} \rightarrow B'}$ and shared quantum state $\omega_{\bar{A}\bar{B}}$;
- (c) nonsignalling assisted code: superchannel Π is nonsignalling from Alice and Bob and vice versa;
- (d) two-way classical communication assisted code: quantum comb Π can be realized by local operations and classical communication (LOCC) operations $\mathcal{P}_1, \dots, \mathcal{P}_{n+1}$ between Alice and Bob (see Fig. 3).

Once the free superchannels or combs are set, the set of free channels is then implicitly defined as the channels that can be generated via these superchannels or combs. Note that the first three coding strategies correspond to parallel channel simulation while the last one corresponds to adaptive channel simulation.

The performance of quantum communication can be characterized by an achievable triplet (n, k, ϵ) , meaning that there exists a Π -assisted coding strategy that uses n instances of the resource channel to transmit k qubits, or simulate id_{2^k} (identity channel on the system of dimension 2^k), within ϵ error (here we consider Choi error, which is the standard choice of error measure for quantum communication). Then by Corollary 19, we can obtain the following bounds on these parameters for general quantum communication in the nonasymptotic regime.

Corollary 24 (Quantum communication). *Suppose (n, k, ϵ_C) is an achievable quantum communication triplet by noise channel \mathcal{N} with an Π -assisted code. Then the Choi error ϵ_C obeys*

$$\epsilon_C \geq (\Gamma_{\mathcal{N}})^n (1 - f_{\text{id}_{2^k}}^C). \quad (80)$$

In other words, the minimum number of channel uses required to enable reliable transmission of k qubits within

Choi error ϵ_C must satisfy

$$n \geq \left\lceil \log \frac{1 - f_{\text{id}_{2^k}}^C}{\epsilon_C} \right\rceil \left\lceil \log \frac{1}{\Gamma_{\mathcal{N}}} \right\rceil^{-1}. \quad (81)$$

We now discuss in more detail the two-way assisted quantum capacity, which is of particular importance due to its close relation to the practical scenario of distributed quantum computing and quantum key distribution. Due to the notorious difficulty of adaptive communication strategies and the involved structure of LOCC operations, this quantum communication scenario is not well understood in spite of its practical importance. The corresponding asymptotic setting that assumes infinite access to the resource channels was recently investigated by a relaxation of LOCC operations to the mathematically more tractable PPT operations (see e.g., Refs. [76–78]). In this case, we have the maximum overlap $f_{\text{id}_{2^k}}^C \leq 1/2^k$ [79]. As the quantum capacity concerns the maximum number of qubits that can be reliably transmitted per use of the channel, we can equivalently obtain from Corollary 24 a nontrivial trade-off (which can be interpreted as a bound on the nonasymptotic two-way assisted quantum capacity):

$$\frac{k}{n} \leq -\frac{1}{n} \log \left(1 - \frac{\epsilon_C}{(\Gamma_{\mathcal{N}})^n} \right) \quad \text{if } \epsilon_C \leq (\Gamma_{\mathcal{N}})^n. \quad (82)$$

Also note that, since PPT operations are semidefinite representable, $\Gamma_{\mathcal{N}}$ here can be efficiently computed by

$$\Gamma_{\mathcal{N}} = \max \left\{ \text{Tr } W : \Phi_{\mathcal{N}} \geq W, W \geq 0, W^{T_B} \geq 0, \text{Tr}_B W_{AB} = (\text{Tr } W_{AB}) I_A / |A| \right\}, \quad (83)$$

where $\Phi_{\mathcal{N}}$ is the Choi state of $\mathcal{N}_{A \rightarrow B}$. Fitting this into Eq. (82) can help us do analysis beyond the asymptotic treatment and understand the intricate trade-off between different operational parameters of concern.

4. Noisy circuit synthesis

The problem of approximating some desired transformation by quantum circuits consisted of certain elementary gates, commonly studied under the name of quantum circuit or gate or unitary synthesis (or sometimes known as “compiling”), is crucial to the practical implementation of quantum computation. Depending on the practical setting, it is often the case that some gates are considered particularly costly as compared to other gates, and thus we are mostly interested in the amount of costly gates needed for the desired synthesis task. A key observation here is that such synthesis tasks can be formalized as adaptive channel simulation problems, where free gates form a comb and the costly gates are input channels that are inserted into the slots of the comb. A particularly important

case is “Clifford + T ,” where we would like to decompose the target transformation into Clifford gates, which are assumed to be free since they can be rather easily implemented fault tolerantly, and the “expensive” T gates $T = |0\rangle\langle 0| + e^{i\pi/4}|1\rangle\langle 1|$. Note that the T gates are often implemented by “state-injection” gadgets [80] that make use of T states produced by magic state distillation (studied in Sec. II), which is a resource-intensive procedure. Therefore, the key figure of merit we would like to optimize is the number of T gates used (namely the “ T count”); see, e.g., Refs. [81–87] for a host of previous studies related to this problem. Notably, resource theory is helpful for finding good bounds on T counts in certain cases [23,88].

Existing literature on the synthesis problem mostly focuses on the noiseless scenario, where the elementary gates are unitary. The noisy nature of practical (especially near-term) devices motivates us to consider the scenario where certain gates are intrinsically associated with noise and such noisy gates are the elementary components of the circuit for synthesis. For example, a key incentive for the Clifford+ T model is that the non-Clifford gates are much harder to protect compared to Clifford gates, so that one may want to consider intrinsically noisy non-Clifford gates (see below). We note that there are fundamental differences between this noisy synthesis setting and the noiseless one, as seen later. Now, the central question is how many noisy resource gates are needed to approximate a target unitary. Based on the observation mentioned above which links the synthesis problem to adaptive channel simulation, we establish the following universal lower bounds on such “noisy gate count” from Corollary 20 (note that for synthesis problems we often use the diamond norm error).

Corollary 25 (Noisy gate count). *Consider the synthesis task of simulating unitary channel \mathcal{U} by channel (noisy gate) \mathcal{G} and arbitrary use of a set of free channels, which compose a free comb, within diamond norm error ϵ_\diamond . Then the number of instances of \mathcal{G} needed must satisfy*

$$n \geq \left\lceil \log \frac{1 - f_{\mathcal{U}}^C}{\epsilon_\diamond} \right\rceil \left\lceil \log \frac{1}{\Gamma_{\mathcal{G}}} \right\rceil^{-1}. \quad (84)$$

We now investigate the Clifford+ T case specifically, where the T gate is associated with noise, and we are interested in the number of such noisy T gates, or the “noisy T count”. Let $\mathcal{C}_n = \{\mathcal{U}_j\}_{j=1}^N$ be the n -qubit Clifford group consisting of N discrete elements. Let the set of free channels be the convex hull of \mathcal{C}_n , i.e., $\mathcal{F} = \text{conv}(\mathcal{C}_n)$, meaning that we allow mixtures of Clifford gates. Any free channel $\mathcal{M} \in \mathcal{F}$ can be written as a convex combination $\mathcal{M} = \sum_{j=1}^N p_j \mathcal{U}_j$ with $p_j \geq 0$ and $\sum_{j=1}^N p_j = 1$. For condition $\mathcal{N} \geq \gamma \mathcal{M}$, we can replace $q_j = \gamma p_j$ and obtain an equivalent condition $\mathcal{M} \geq \sum_{j=1}^N q_j \mathcal{U}_j$ with $q_j \geq 0$ and $\gamma = \sum_{j=1}^N q_j$. Therefore, the free component can be

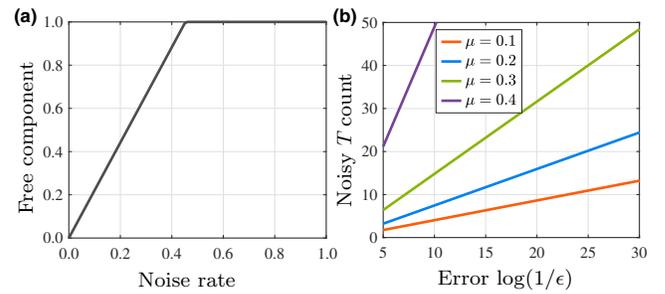


FIG. 4. Noisy Clifford+ T synthesis. (a) Free component of the noisy T gate $\mathcal{N}_\mu \circ T$, where T is followed by depolarizing noise of strength μ . (b) Lower bounds on the noisy T count (the number of $\mathcal{N}_\mu \circ T$) needed to simulate a CCZ gate within diamond norm error ϵ , depicted for noise rates $\mu = 0.1, 0.2, 0.3, 0.4$.

computed by a semidefinite program

$$\Gamma_{\mathcal{N}} = \max \left\{ \sum_{j=1}^N q_j : \Phi_{\mathcal{N}} \geq \sum_{j=1}^N q_j \Phi_{\mathcal{U}_j}, q_j \geq 0 \right\}, \quad (85)$$

where $\Phi_{\mathcal{N}}$ and $\Phi_{\mathcal{U}_j}$ are the Choi states of \mathcal{N} and \mathcal{U}_j , respectively. As a concrete example, consider T gate followed by depolarizing noise $\mathcal{N}_\mu(\rho) = (1 - \mu)\rho + \mu I/2$ as the elementary channel. Its free component $\Gamma_{\mathcal{N}_\mu \circ T}$ is computed by the SDP Eq. (85) where $N = 24$ (see, e.g., Ref. [89] for an explicit enumeration of \mathcal{C}_1), and depicted in Fig. 4(a). When $\mu \geq 1 - \sqrt{3}/3 \approx 0.42$ we see that $\Gamma_{\mathcal{N}_\mu \circ T} = 1$, as \mathcal{N}_μ compresses the entire Bloch sphere into the stabilizer octahedron so any output is a stabilizer state. Note that this explicit calculation improves the general bound μ as discussed in Sec. III B 1. In Fig. 4(b), as an example, we plot the lower bounds on the noisy T count in order to approximate a CCZ gate, obtained from Corollary 25 [where we used $f_{\text{CCZ}}^W \leq 9/16$ [23, Eq.(33)]].

Recently, Ref. [90, Proposition 26] also gave an expression for the noisy gate counts in magic theory of odd dimensions using the mana monotone. Note that our result applies to any dimension, and is expected to outperform the mana bound especially in the small target error regime. In particular, our bound implies diverging cost as the target error $\epsilon \rightarrow 0$, which is in line with intuitions, but the mana bound cannot.

Finally, we remark that the noisy synthesis results here are fundamentally different from the existing ones on noiseless synthesis, in spite of some apparent relations. Most notably, it is known that for any universal gate set, the number of gates needed to approximate all unitaries up to error ϵ (which can essentially be measured by any channel error measure discussed earlier) scales at least as $\Omega(\log(1/\epsilon))$ [91] (note that the well-known Solovay-Kitaev theorem [3,44,45] concerns the upper bound). Although the $\Omega(\log(1/\epsilon))$ scaling is similar to our lower bound on noisy gate counts, there are two

key differences: (i) our noisy synthesis result bounds the number of resource gates needed and says nothing about the number of free gates, while the previous noiseless-case result counts the total number of gates; (ii) our noisy synthesis result is universal for *any* target resource unitary, while the previous noiseless-case result examines the worst case and there could well be target unitaries with lower or even trivial cost (some target unitaries can be exactly simulated, such as T in Clifford+ T). Relatedly, the geometric covering argument used in Ref. [91] is not useful for the noisy case. In general, the noiseless and noisy synthesis and gate counts are fundamentally disparate problems contingent on different factors. This can again be seen from Clifford+ T , where intricate number theory properties and techniques play decisive roles in the noiseless case [81–85] while being irrelevant in the noisy case.

IV. CONCLUDING REMARKS

We introduced a simple, universal framework for understanding and analyzing the limitations on quantum resource purification tasks that applies to virtually any resource theory, based on the notion of “free component” of noisy resources. We developed the theory in detail for both quantum states and channels. For the state theory, our new results significantly improve over corresponding ones discovered in Ref. [12] in terms of both the regime of the no-purification rules and the quantitative limits. This framework also enabled us to quantitatively understand the no-purification principles for quantum channels or dynamical resources. Specifically, the channel theory involves complications concerning error measures and the possibility of adaptively using multiple resource instances, as compared to the state theory. We demonstrated broad theoretical and practical relevance of our techniques and results by discussing their applications to several key areas of quantum information science and physics. The simplicity and generality of our theory highlight the fundamental nature of the no-purification principles.

Several technical problems are worth further study. First, we considered channel simulation with a single target channel here, but more generally the output can also be a comb [69]; it would be interesting to further study the no-purification bounds for such cases and explore their relevance. Second, we formulated the results in terms of deterministic one-shot transformation and only left preliminary remarks on the probabilistic case; a comprehensive understanding of the probabilistic case is left for future work. Third, it is worth further study purification tasks for continuous variables, especially resource (e.g., non-Gaussianity) distillation tasks and their applications in optical quantum information processing, given that there are some sharp distinctions known concerning the feasibility and behaviors of distillation procedures [39,92] between continuous

and discrete variables, but the understanding of the full correspondence is still preliminary.

Furthermore, it would be interesting to further analyze our bounds and associated parameters in specific theories and problems. The discussion on the applications we gave here mainly serve to establish the general, conceptual connections and are thus preliminary. Further developments of these connections, taking specific features of the system, resource, and noise etc. into account, may be fruitful. In particular, for the extensively studied topics of quantum error correction and quantum Shannon theory, it would be interesting to further optimize the bounds and compare them with existing results in specific scenarios. We eventually hope that our demonstrations here will spark explorations of further applications or consequences of the no-purification principles in quantum information and physics.

ACKNOWLEDGMENTS

We thank Newton Cheng, Gilad Gour, Liang Jiang, Seth Lloyd, Milad Marvian, Kyungjoo Noh, and Sisi Zhou for discussions and feedback, Bartosz Regula and Ryuji Takagi for discussions and correspondence about their work [93], and especially Andreas Winter for inspiring discussions about the initial ideas. Z.W.L. is supported by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

Note added.—Recently, we became aware that Regula and Takagi independently considered the resource weight and obtained results related to ours, which later developed into Ref. [93]. The two papers were arranged to be released concurrently on arXiv.

-
- [1] John Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 - [2] Frank Arute *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
 - [3] Michael A Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
 - [4] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046 (1996).
 - [5] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, *Phys. Rev. Lett.* **76**, 722 (1996).

- [6] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [7] Sergey Bravyi and Alexei Kitaev, Universal quantum computation with ideal clifford gates and noisy ancillas, *Phys. Rev. A* **71**, 022316 (2005).
- [8] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, Quantum entanglement, *Rev. Mod. Phys.* **81**, 865 (2009).
- [9] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio, Colloquium: Quantum coherence as a resource, *Rev. Mod. Phys.* **89**, 041003 (2017).
- [10] Victor Veitch, S. A. Hamed Mousavian, Daniel Gottesman, and Joseph Emerson, The resource theory of stabilizer quantum computation, *New J. Phys.* **16**, 013009 (2014).
- [11] Eric Chitambar and Gilad Gour, Quantum resource theories, *Rev. Mod. Phys.* **91**, 025001 (2019).
- [12] Kun Fang and Zi-Wen Liu, No-Go Theorems for Quantum Resource Purification, *Phys. Rev. Lett.* **125**, 060405 (2020).
- [13] Zi-Wen Liu Kaifeng Bu, and Ryuji Takagi, One-Shot Operational Quantum Resource Theory, *Phys. Rev. Lett.* **123**, 020401 (2019).
- [14] Fernando G. S. L. Brandão and Gilad Gour, Reversible Framework for Quantum Resource Theories, *Phys. Rev. Lett.* **115**, 070503 (2015).
- [15] Zi-Wen Liu, Xueyuan Hu, and Seth Lloyd, Resource Destroying Maps, *Phys. Rev. Lett.* **118**, 060502 (2017).
- [16] Andrés F. Ducuara and Paul Skrzypczyk, Operational interpretation of weight-based resource quantifiers in convex quantum resource theories of states, *Phys. Rev. Lett.* **125**, 110401 (2020).
- [17] Roope Uola, Tom Bullock, Tristan Kraft, Juha-Pekka Pellonpää, and Nicolas Brunner, All quantum resources provide an advantage in exclusion tasks, *Phys. Rev. Lett.* **125**, 110402 (2020).
- [18] Nilanjana Datta, Min-and max-relative entropies and a new entanglement monotone, *IEEE Trans. Inf. Theory* **55**, 2816 (2009).
- [19] Terry Rudolph and Robert W. Spekkens, Quantum state targeting, *Phys. Rev. A* **70**, 052306 (2004).
- [20] Bartosz Regula, Kaifeng Bu, Ryuji Takagi, and Zi-Wen Liu, Benchmarking one-shot distillation in general quantum resource theories, *Phys. Rev. A* **101**, 062315 (2020).
- [21] Earl T. Campbell, Catalysis and activation of magic states in fault-tolerant architectures, *Phys. Rev. A* **83**, 032317 (2011).
- [22] Sergey Bravyi and David Gosset, Improved Classical Simulation of Quantum Circuits Dominated by Clifford Gates, *Phys. Rev. Lett.* **116**, 250501 (2016).
- [23] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, *Quantum* **3**, 181 (2019).
- [24] Bartosz Regula, Kun Fang, Xin Wang, and Gerardo Adesso, One-Shot Coherence Distillation, *Phys. Rev. Lett.* **121**, 010401 (2018).
- [25] Patrick Hayden, Sepehr Nezami, Sandu Popescu, and Grant Salton, Error correction of quantum reference frame information, *PRX Quantum* **2**, 010326 (2021).
- [26] Philippe Faist, Sepehr Nezami, Victor V. Albert, Grant Salton, Fernando Pastawski, Patrick Hayden, and John Preskill, Continuous Symmetries and Approximate Quantum Error Correction, *Phys. Rev. X* **10**, 041018 (2020).
- [27] Mischa P. Woods and Álvaro M. Alhambra, Continuous groups of transversal gates for quantum error correcting codes from finite clock reference frames, *Quantum* **4**, 245 (2020).
- [28] Aleksander Kubica and Rafał Demkowicz-Dobrzański, Using Quantum Metrological Bounds in Quantum Error Correction: A Simple Proof of the Approximate Eastin-Knill Theorem, *Phys. Rev. Lett.* **126**, 150503 (2021).
- [29] Sisi Zhou, Zi-Wen Liu, and Liang Jiang, New perspectives on covariant quantum error correction, *Quantum* **5**, 521 (2021).
- [30] Yuxiang Yang, Yin Mo, Joseph M. Renes, Giulio Chiribella, and Mischa P. Woods, Covariant quantum error correcting codes via reference frames, arXiv e-prints, arXiv:2007.09154 (2020).
- [31] Linghang Kong and Zi-Wen Liu, Near-optimal covariant quantum error-correcting codes from random unitaries with symmetries, arXiv e-prints, arXiv:2112.01498 (2021).
- [32] Replace $\frac{1}{\tau} \int_0^\tau d\theta$ by the integration over the Haar measure on G .
- [33] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, *Nature* **409**, 46 (2001).
- [34] Roman Schnabel, Squeezed states of light and their applications in laser interferometers, *Phys. Rep.* **684**, 1 (2017).
- [35] Benjamin Yadin, Felix C. Binder, Jayne Thompson, Varun Narasimhachar, Mile Gu, and M. S. Kim, Operational Resource Theory of Continuous-Variable Nonclassicality, *Phys. Rev. X* **8**, 041038 (2018).
- [36] Hyukjoon Kwon, Kok Chuan Tan, Tyler Volkoff, and Hyunseok Jeong, Nonclassicality as a Quantifiable Resource for Quantum Metrology, *Phys. Rev. Lett.* **122**, 040503 (2019).
- [37] Horace P. Yuen, Two-photon coherent states of the radiation field, *Phys. Rev. A* **13**, 2226 (1976).
- [38] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd, Gaussian quantum information, *Rev. Mod. Phys.* **84**, 621 (2012).
- [39] Ludovico Lami, Bartosz Regula, Xin Wang, Rosanna Nichols, Andreas Winter, and Gerardo Adesso, Gaussian quantum resource theories, *Phys. Rev. A* **98**, 022335 (2018).
- [40] Zi-Wen Liu and Andreas Winter, Resource theories of quantum channels and the universal role of resource erasure, arXiv e-prints, arXiv:1904.04201 (2019).
- [41] Gilad Gour and Carlo Maria Scandolo, The entanglement of a bipartite channel, arXiv preprint arXiv:1907.02552 (2019).
- [42] Mark M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
- [43] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen, Distance measures to compare real and ideal quantum processes, *Phys. Rev. A* **71**, 062310 (2005).
- [44] A. Yu. Kitaev, Quantum computations: Algorithms and error correction, *Russ. Math. Surveys* **52**, 1191 (1997).

- [45] Christopher M. Dawson and Michael A. Nielsen, The solovay-kitaev algorithm, *Quantum Info. Comput.* **6**, 81 (2006).
- [46] Debbie W. Leung, Michael A. Nielsen, Isaac L. Chuang, and Yoshihisa Yamamoto, Approximate quantum error correction can lead to better codes, *Phys. Rev. A* **56**, 2567 (1997).
- [47] Marco Tomamichel, Mario Berta, and Joseph M. Renes, Quantum coding with finite resources, *Nat. Commun.* **7**, 1 (2016).
- [48] Xin Wang, Kun Fang, and Runyao Duan, Semidefinite programming converse bounds for quantum communication, *IEEE Trans. Inf. Theory* **65**, 2583 (2018).
- [49] Mark D. Bowdrey, Daniel K. L. Oi, Anthony J. Short, Konrad Banaszek, and Jonathan A. Jones, Fidelity of single qubit maps, *Phys. Lett. A* **294**, 258 (2002).
- [50] J. M. Chow, J. M. Gambetta, L. Tornberg, Jens Koch, Lev S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized Benchmarking and Process Tomography for Gate Errors in a Solid-State Qubit, *Phys. Rev. Lett.* **102**, 090502 (2009).
- [51] Steven T. Flammia and Yi-Kai Liu, Direct Fidelity Estimation from few Pauli Measurements, *Phys. Rev. Lett.* **106**, 230501 (2011).
- [52] Dawei Lu, Hang Li, Denis-Alexandre Trottier, Jun Li, Aharon Brodutch, Anthony P. Krismanich, Ahmad Ghavami, Gary I. Dmitrienko, Guilu Long, Jonathan Baugh, and Raymond Laflamme, Experimental Estimation of Average Fidelity of a Clifford Gate on a 7-Qubit Quantum Processor, *Phys. Rev. Lett.* **114**, 140505 (2015).
- [53] Pawel Horodecki, Michal Horodecki, and Ryszard Horodecki, General teleportation channel, singlet fraction and quasi-distillation, arXiv e-prints, [arXiv:quant-ph/9807091](https://arxiv.org/abs/quant-ph/9807091) (1998).
- [54] M. Hayashi, Discrimination of two channels by adaptive methods and its application to quantum system, *IEEE Trans. Inf. Theory* **55**, 3807 (2009).
- [55] Aram W. Harrow, Avinatan Hassidim, Debbie W. Leung, and John Watrous, Adaptive versus nonadaptive strategies for quantum channel discrimination, *Phys. Rev. A* **81**, 032339 (2010).
- [56] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi, Fundamental limits of repeaterless quantum communications, *Nat. Commun.* **8**, 1 (2017).
- [57] Stefano Pirandola, Riccardo Laurenza, Cosmo Lupo, and Jason L. Pereira, Fundamental limits to quantum channel discrimination, *Npj Quantum Inf.* **5**, 50 (2019).
- [58] Leonardo Banchi, Jason Pereira, Seth Lloyd, and Stefano Pirandola, Convex optimization of programmable quantum computers, *Npj Quantum Inf.* **6**, 42 (2020).
- [59] Xin Wang and Mark M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, *Phys. Rev. Res.* **1**, 033169 (2019).
- [60] Kun Fang, Omar Fawzi, Renato Renner, and David Sutter, Chain Rule for the Quantum Relative Entropy, *Phys. Rev. Lett.* **124**, 100501 (2020).
- [61] Mark M. Wilde, Mario Berta, Christoph Hirche, and Eneet Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, *Lett. Math. Phys.* **110**, 2277 (2020).
- [62] Rafal Demkowicz-Dobrzański and Lorenzo Maccone, Using Entanglement against Noise in Quantum Metrology, *Phys. Rev. Lett.* **113**, 250801 (2014).
- [63] Stefano Pirandola and Cosmo Lupo, Ultimate Precision of Adaptive Noise Estimation, *Phys. Rev. Lett.* **118**, 100502 (2017).
- [64] Haidong Yuan and Chi-Hang Fred Fung, Fidelity and fisher information on quantum channels, *New J. Phys.* **19**, 113039 (2017).
- [65] Riccardo Laurenza, Cosmo Lupo, Gaetana Spedalieri, Samuel L. Braunstein, and Stefano Pirandola, Channel simulation in quantum metrology, *Quantum Measurements Quantum Metrology* **5**, 1 (2018).
- [66] Thomas P. W. Cope and Stefano Pirandola, Adaptive estimation and discrimination of holevo-werner channels, *Quantum Measurements Quantum Metrology* **4**, 44 (2017).
- [67] Sisi Zhou and Liang Jiang, Asymptotic theory of quantum channel estimation, arXiv e-prints, [arXiv:2003.10559](https://arxiv.org/abs/2003.10559) (2020).
- [68] Vishal Katariya and Mark M. Wilde, Geometric distinguishability measures limit quantum channel estimation and discrimination, arXiv e-prints, [arXiv:2004.10708](https://arxiv.org/abs/2004.10708) (2020).
- [69] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [70] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, Theoretical framework for quantum networks, *Phys. Rev. A* **80**, 022339 (2009).
- [71] Tom Cooney, Milán Mosonyi, and Mark M. Wilde, Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication, *Commun. Math. Phys.* **344**, 797 (2016).
- [72] Daniel Gottesman, Ph.D. thesis, California Institute of Technology, 1997.
- [73] Kun Fang, Xin Wang, Marco Tomamichel, and Mario Berta, Quantum channel simulation and the channel's smooth max-information, *IEEE Trans. Inf. Theory* **66**, 2129 (2019).
- [74] Ryuji Takagi, Kun Wang, and Masahito Hayashi, Application of the Resource Theory of Channels to Communication Scenarios, *Phys. Rev. Lett.* **124**, 120502 (2020).
- [75] Xin Wang, Kun Fang, and Marco Tomamichel, On converse bounds for classical communication over quantum channels, *IEEE Trans. Inf. Theory* **65**, 4609 (2019).
- [76] Mario Berta and Mark M. Wilde, Amortization does not enhance the max-rains information of a quantum channel, *New J. Phys.* **20**, 053044 (2018).
- [77] Kun Fang and Hamza Fawzi, Geometric Rényi divergence and its applications in quantum channel capacities, arXiv preprint [arXiv:1909.05758](https://arxiv.org/abs/1909.05758) (2019).
- [78] Hamza Fawzi and Omar Fawzi, Defining quantum divergences via convex optimization, arXiv preprint [arXiv:2007.12576](https://arxiv.org/abs/2007.12576) (2020).
- [79] Eric M. Rains, A semidefinite program for distillable entanglement, *IEEE Trans. Inf. Theory* **47**, 2921 (2001).
- [80] Daniel Gottesman and Isaac L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390 (1999).
- [81] Brett Giles and Peter Selinger, Exact synthesis of multiqubit clifford+T circuits, *Phys. Rev. A* **87**, 032332 (2013).

- [82] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca, Asymptotically Optimal Approximation of Single Qubit Unitaries by Clifford and T Circuits Using a Constant Number of Ancillary Qubits, *Phys. Rev. Lett.* **110**, 190502 (2013).
- [83] David Gosset, Vadym Kliuchnikov, Michele Mosca, and Vincent Russo, An algorithm for the T-count, arXiv e-prints, [arXiv:1308.4134](https://arxiv.org/abs/1308.4134) (2013).
- [84] Peter Selinger, Efficient clifford+T approximation of single-qubit operators, *Quantum Info. Comput.* **15**, 159 (2015).
- [85] Vadym Kliuchnikov and Jon Yard, A framework for exact synthesis, arXiv e-prints, [arXiv:1504.04350](https://arxiv.org/abs/1504.04350) (2015).
- [86] Ryan Babbush, Craig Gidney, Dominic W Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven, Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity, *Phys. Rev. X* **8**, 041015 (2018).
- [87] Guang Hao Low, Vadym Kliuchnikov, and Luke Schaeffer, Trading T-gates for dirty qubits in state preparation and unitary synthesis, arXiv preprint [arXiv:1812.00954](https://arxiv.org/abs/1812.00954) (2018).
- [88] Mark Howard and Earl Campbell, Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing, *Phys. Rev. Lett.* **118**, 090501 (2017).
- [89] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman, Randomized Benchmarking of Single-Qubit Gates in a 2d Array of Neutral-Atom Qubits, *Phys. Rev. Lett.* **114**, 100503 (2015).
- [90] Xin Wang, Mark M. Wilde, and Yuan Su, Quantifying the magic of quantum channels, *New J. Phys.* **21**, 103002 (2019).
- [91] Aram W. Harrow, Benjamin Recht, and Isaac L. Chuang, Efficient discrete approximations of quantum gates, *J. Math. Phys.* **43**, 4445 (2002).
- [92] Ryuji Takagi and Quntao Zhuang, Convex resource theory of non-gaussianity, *Phys. Rev. A* **97**, 062337 (2018).
- [93] Bartosz Regula and Ryuji Takagi, Fundamental limitations on quantum channel manipulation, arXiv e-prints, [arXiv:2010.11942](https://arxiv.org/abs/2010.11942) (2020).