

One-Shot Coherence Distillation: Towards Completing the Picture

Qi Zhao^{ID}, Yunchao Liu, Xiao Yuan, Eric Chitambar^{ID}, and Andreas Winter^{ID}

Abstract—The resource framework of quantum coherence was introduced by Baumgratz, Cramer, and Plenio [*Phys. Rev. Lett.* 113, 140401 (2014)] and further developed by Winter and Yang [*Phys. Rev. Lett.* 116, 120404 (2016)]. We consider the one-shot problem of distilling pure coherence from a single instance of a given resource state. Specifically, we determine the distillable coherence with a given fidelity under incoherent operations (IO) through a generalization of the Winter-Yang protocol. This is compared to the distillable coherence under maximal incoherent operations (MIO) and dephasing-covariant incoherent operations (DIO), which can be cast as a semidefinite programme, that has been presented previously by Regula *et al.* [*Phys. Rev. Lett.* 121, 010401 (2018)]. Our results are given in terms of a smoothed min-relative entropy distance from the incoherent set of states, and a variant of the hypothesis-testing relative entropy distance, respectively. The one-shot distillable coherence is also related to one-shot randomness extraction. Moreover, from the one-shot formulas under IO, MIO, and DIO, we can recover the optimal distillable rate in the many-copy asymptotics, yielding the relative entropy of coherence. These results can be compared with previous work by some of the present authors [Zhao *et al.*, *Phys. Rev. Lett.* 120, 070403 (2018)] on one-shot coherence formation under IO, MIO, DIO and also SIO. This shows that the amount of distillable coherence is essentially the same for IO, DIO, and MIO, despite the fact that the three classes of operations are very different. We also relate the distillable coherence under strictly incoherent operations (SIO) to a constrained hypothesis testing

Manuscript received August 24, 2018; revised February 14, 2019; accepted March 31, 2019. Date of publication April 15, 2019; date of current version September 13, 2019. Q. Zhao was supported by the National Natural Science Foundation of China under Grant 11674193. X. Yuan was supported in part by BP plc and in part by the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology under Grant EP/M013243/1. E. Chitambar was supported by the National Science Foundation (NSF) Early CAREER Award under Grant 1352326. A. Winter was supported in part by the Spanish MINECO under Project FIS2016-80681-P, with the support of FEDER funds, and in part by the Generalitat de Catalunya (CIRIT) under Project 2017-SGR-1127.

Q. Zhao and Y. Liu are with the Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China (e-mail: zhaoqithu10@gmail.com).

X. Yuan is with the Department of Materials, University of Oxford, Oxford OX1 3PH, U.K.

E. Chitambar is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Champaign, IL 61801 USA (e-mail: echitamb@illinois.edu).

A. Winter is with ICREA—Institució Catalana de Recerca i Estudis Avançats, ES-08001 Barcelona, Spain, and also with Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain (e-mail: andreas.winter@uab.cat).

Communicated by M. M. Wilde, Associate Editor for Quantum Information Theory.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2019.2911102

problem and explicitly show the existence of bound coherence under SIO in the asymptotic regime.

Index Terms—Quantum coherence, coherence distillation, one-shot, quantum resource theory.

I. INTRODUCTION

COHERENCE, as the signature of non-classicality, has applications in many quantum information processing tasks, including cryptography [1], metrology [2], randomness generation [3], [4], biological systems [5], [6] and thermodynamics [7]–[11]. Recently, a resource theory framework of quantum coherence has been introduced by Baumgratz *et al.* [12] (after prior work of Åberg [13], as well as Braun and Georgeot [14]). A general quantum resource theory consists of two objects: a set of free states and a set of free operations, with the latter acting invariantly on the former. In the resource theory of coherence, the free states \mathcal{I} for a d -dimensional Hilbert space is the set of density matrices that are invariant under conjugation by the phase unitary

$$Z = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & e^{2\pi i/d} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & e^{2\pi i(d-1)/d} \end{pmatrix}, \quad (1)$$

with Z being expressed in an *a priori* fixed computational basis $\{|x\rangle\}_{x=1}^d$. Furthermore, different physical and mathematical motivations have led to the proposal and study of different classes of free operations, most notably the maximally incoherent operations (MIO) [13], the dephasing-covariant incoherent operations (DIO) [15], [16], the incoherent operations (IO) [12], and the strictly incoherent operations (SIO) [17], [18]. For instance, SIO emerges as a natural class of operations to consider when quantifying visibility in interferometer experiments [19]. Several coherence measures have been introduced to quantify the amount of coherence in a state, including the relative entropy and ℓ_1 -norm of coherence [12], the coherence of formation [3], [13], and the robustness of coherence [20], which have all been shown to possess different operational meanings [3], [17], [20], [21]. An overview on recent developments of the resource theory of coherence can be found in Ref. [22].

An important problem in the resource theory is convertibility of states via free operations, especially those between an arbitrary state ρ and a maximal resource state $|\Psi_M\rangle$. In an M -level system, the superposition state $|\Psi_M\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^M |i\rangle$

has the maximal coherence with respect to the aforementioned, and indeed all, coherence measures, and the qubit state $|\Psi_2\rangle$ serves as resource unit in the theory, the “cosbit” [23]. The process of converting a given state ρ to $|\Psi_M\rangle$ is referred to as *coherence distillation* and the reverse process as *coherence dilution*. In the asymptotic case where an unbounded number of independent and identically distributed (i.i.d.) copies of the initial state are provided, the optimal rates of asymptotic distillation and dilution are quantified by the relative entropy of coherence and the coherence of formation, respectively [17].

In practical scenarios, i.i.d. resources are not available and an analysis of non-asymptotic tasks becomes crucial. In recent work [24], the formation problem was addressed in the one-shot setting under all four of the above operational classes. Subsequently, the converse question of one-shot coherence distillation was solved for the classes MIO and DIO [25]. In the present paper, we close the remaining gap by providing an analysis of general one-shot coherence distillation under IO and SIO. Our results show that, while not being exactly identical, the three classes MIO, DIO, and IO all lead to essentially the same expression for the distillable coherence in terms of a min-relative entropy distance from the incoherent states; the differences are in the smoothing parameters and universal additive terms. When extended to the asymptotic case, our one-shot results imply that MIO, DIO, and IO all have distillable coherence given by the relative entropy, thereby recovering earlier work found in [17]. In contrast, we show that the coherence distillation by SIO can behave quite differently than its more general counterparts, and we characterize the one-shot distillable coherence as a constrained hypothesis testing problem. We further show that coherence can be a *bound* resource under SIO, meaning that states exist with zero distillable coherence but nonzero coherence cost. We also show the relationship among the extractable randomness, IO and DIO distillable rate, which are indeed closely related. The reason is that the distillation proceeds by a kind of decoupling process, and can thus be related to randomness extraction.

The structure of the remainder of the paper is as follows: In Section II we review the four mentioned classes of incoherent operations and introduce the dilution and distillation tasks for coherence in the one-shot setting. We present the prior results on dilution in all four cases, and in Section III review the prior results on distillation under MIO and DIO. In Section IV we come to the first main result of the present paper, a tight one-shot characterization of distillation with IO, showing in particular a distillation protocol achieving the lower bound with an operation that is at the same time IO and DIO. Then, in Section V we give a one-shot formula for distillation under SIO, and most importantly, show the existence of bound coherence in this model. In Section VI we show how the obtained one-shot formulas for IO, DIO and MIO imply the previously known asymptotic distillation rate $C_r(\rho)$, and finally Section VII contains the discussion on the relation between randomness extraction and coherence distillation, after which we conclude.

II. COHERENCE DILUTION AND DISTILLATION

A. Classes of Incoherent Operations

A general resource theory of coherence is constructed by imposing restrictions on the allowed completely positive trace-preserving (CPTP) maps $\Lambda : \mathcal{L}(A) \rightarrow \mathcal{L}(B)$, where $\mathcal{L}(A)$ denotes the space of linear operators acting on a finite-dimensional Hilbert space A and likewise for $\mathcal{L}(B)$. One necessary restriction is that Λ acts invariantly on the set of incoherent states \mathcal{I} ; i.e. $\Lambda(\delta) \in \mathcal{I}$ whenever $\delta \in \mathcal{I}$. The completely dephasing map $\Delta(\rho) = \sum_{x=1}^d |x\rangle\langle x|\rho|x\rangle\langle x|$ plays an important role in this theory as it destroys all coherence in a state, and it therefore maps any resource state to a free one. The following classes of operations have been proposed in [12], [13], [15]–[18], each motivated by different physical considerations.

MIO *Maximal Incoherent Operations* [13] are characterized by $\Lambda(\delta) \in \mathcal{I}$ for all $\delta \in \mathcal{I}$, which may be expressed as the identity $\Lambda \circ \Delta = \Delta \circ \Lambda \circ \Delta$;

DIO *Dephasing-Covariant Incoherent Operations* [15], [16] satisfy the stronger condition $\Lambda \circ \Delta = \Delta \circ \Lambda$;

IO *Incoherent Operations* [12] are those MIO with a Kraus decomposition $\Lambda(\rho) = \sum_a K_a \rho K_a^\dagger$, such that $K_a \delta K_a^\dagger / \text{Tr}(K_a \delta K_a^\dagger) \in \mathcal{I}$ for all a and $\delta \in \mathcal{I}$;

SIO *Strictly Incoherent Operations* [17] are those IO with a Kraus decomposition $\Lambda(\rho) = \sum_a K_a \rho K_a^\dagger$ such that $K_a = \sum_x c_{a,x} |f_a(x)\rangle\langle x|$, and $f_a : [d] = \{1, \dots, d\} \rightarrow [d]$ is a one-to-one function for all a .

The relationships among these classes are

$$\begin{aligned} \text{SIO} &\subsetneq \text{IO} \subsetneq \text{MIO}, \\ \text{SIO} &\subsetneq \text{DIO} \subsetneq \text{MIO}; \end{aligned} \quad (2)$$

note that IO is not included in DIO and vice versa.

For later use, and because it will turn out to be significant, we also give a name to the intersection of IO and DIO,

$$\text{DIO} = \text{IO} \cap \text{DIO},$$

which we dub *dephasing-covariant incoherent IO*.

B. Coherence Dilution

The one-shot coherence dilution problem characterizes the minimal resource required for the formation of a target state with an allowed error ε . The definition of one-shot coherence dilution is as follows.

Definition 1. Let $\mathcal{O} \in \{\text{MIO}, \text{DIO}, \text{IO}, \text{SIO}\}$ denote some class of incoherent operations. Then for a given state ρ and $\varepsilon \geq 0$, the one-shot coherence formation cost with error ε under \mathcal{O} is defined as

$$C_{c,\mathcal{O}}^\varepsilon(\rho) = \min_{\Lambda \in \mathcal{O}} \{ \log M : F(\Lambda(\Psi_M), \rho)^2 \geq 1 - \varepsilon \}. \quad (3)$$

where $F(\rho, \sigma) = \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}$ is the usual mixed state fidelity. By default, here and throughout the paper, $\log \equiv \log_2$ is the binary logarithm, in accordance with information theoretic use.

In Ref. [24], several coherence monotones were proposed to estimate the one-shot coherence cost. The first two are based

TABLE I

COHERENCE DISTILLATION AND FORMATION RATES, WITH ATTRIBUTIONS WHERE THE CORRESPONDING RESULT WAS FIRST OBTAINED. OUR NEW CONTRIBUTIONS (IN BOLD, AND MARKED BY ASTERISKS) ARE THE IO AND SIO DISTILLATION RATES IN THE ONE-SHOT SETTING; FURTHERMORE, WE PROVE THAT MIO, DIO AND IO ONE-SHOT DISTILLATION RATES DIFFER ONLY BY VARYING THE SMOOTHING PARAMETER AND BY UNIVERSAL ADDITIVE TERMS. THE ONE-SHOT DISTILLATION RATE UNDER SIO IS CONVERTED INTO AN OPTIMIZATION PROBLEM IN THEOREM 10, AND IN THE ASYMPTOTIC CASE, WE PROVE THE EXISTENCE OF BOUND COHERENCE UNDER SIO, SHOWING THAT THE RATE IS DIFFERENT FROM THE RELATIVE ENTROPY OF COHERENCE IN GENERAL

		MIO	DIO	IO	SIO
One-shot	distillation	\tilde{C}_H^ε [25]	\tilde{C}_H^ε [25]	$\mathbf{C}_{\min}^{\varepsilon'}$ [*]	Thm. 10 [*]
	formation	C_{\max}^ε [24]	$C_{\Delta,\max}^\varepsilon$ [24]	C_0^ε [24]	C_0^ε [24]
Asymptotic	distillation	C_r [25]	C_r [25]	C_r [17]	$\neq \mathbf{C}_r$ [*]
	formation	C_r [24], [26]	C_r [24], [26]	C_f [17]	C_f [17]

on the max relative entropy $D_{\max}(\rho\|\sigma) = \log \min\{\lambda : \rho \leq \lambda\sigma\}$. Alternatively, one can introduce this quantity using the sandwiched quantum α -Rényi divergence

$$\tilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha-1} \log \left(\text{Tr} \left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right] \right), \quad (4)$$

letting $\alpha \rightarrow \infty$ [27], [28]. One then defines

$$C_{\max}(\rho) = \min_{\delta \in \mathcal{I}} D_{\max}(\rho\|\delta) = \log \min_{\delta \in \mathcal{I}} \{\lambda : \rho \leq \lambda\delta\}, \quad (5)$$

$$C_{\Delta,\max}(\rho) = \min_{\sigma \in \overline{A}_\rho} D_{\max}(\rho\|\sigma) = \log \min\{\lambda : \rho \leq \lambda\Delta(\rho)\}, \quad (6)$$

where $A_\rho = \{\sigma \geq 0 : \sigma + t\rho = (1+t)\Delta(\rho), t > 0\}$ [29] and \overline{A}_ρ is its closure. The quantity $C_{\max}(\rho)$ is a monotone under MIO, while $C_{\Delta,\max}(\rho)$ is a monotone under DIO, but not IO in general. However, for a pure state $|\psi\rangle$, $C_{\Delta,\max}$ reduces to its *incoherent rank* [29], defined as

$$C_0(\psi) := S_0(\Delta(\psi)) = \log \text{rank}[\Delta(\psi)], \quad (7)$$

and the incoherent rank is a monotone under IO even for stochastic pure state transformations [12]. Thus, one can obtain a general mixed state monotone for IO by taking a convex roof extension:

$$C_0(\rho) := \min_{p_i, |\psi_i\rangle} \max_i C_0(\psi_i), \quad (8)$$

where the minimization is over all pure state ensembles satisfying $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Note, the minimization is followed by a maximization rather than an averaging over the $|\psi_i\rangle$ (as typically done in such measures) because $C_0(\psi)$ is a stochastic IO monotone.

Since we allow ε error in our one-shot tasks, we likewise apply an ε smoothing to our coherence measures. Depending on the quantity, this is done by either minimizing or maximizing the measure over all states ρ' lying in an ε -ball around ρ . For example, if C represents any one of the measures C_{\max} , $C_{\Delta,\max}$ or C_0 , its ε -smoothed variant is given by

$$C^\varepsilon(\rho) = \min_{\rho' \in B_\varepsilon(\rho)} C(\rho'). \quad (9)$$

Here $B_\varepsilon(\rho) = \{\rho' : P(\rho', \rho) \leq \varepsilon\}$ is defined with respect to the purified distance $P(\rho', \rho) = \sqrt{1 - F(\rho', \rho)^2}$. Note that the purified distance is related to the trace distance by [30]

$$1 - \sqrt{1 - P(\rho', \rho)^2} \leq \frac{1}{2} \|\rho - \rho'\|_1 \leq P(\rho', \rho).$$

As shown in [24], the measures C_{\max}^ε , $C_{\Delta,\max}^\varepsilon$ and C_0^ε precisely characterize the one-shot coherence formation for MIO, DIO, and IO/SIO, respectively (see Table I). Note that our definition of $B_\varepsilon(\rho)$ differs from [24] by replacing $\varepsilon \leftrightarrow \sqrt{\varepsilon}$.

The regularized (many-copy) coherence cost of formation can then be defined using the one-shot quantities. For $\mathcal{O} \in \{\text{MIO, DIO, IO, SIO}\}$, the asymptotic rate of coherence cost for a state ρ under operations \mathcal{O} is defined as

$$C_{c,\mathcal{O}}^\infty(\rho) := \limsup_{\varepsilon \rightarrow 0^+} \limsup_{n \rightarrow \infty} \frac{1}{n} C_{c,\mathcal{O}}^\varepsilon(\rho^{\otimes n}). \quad (10)$$

Remarkably, for all four operational classes, the coherence cost has a single-letter characterization. For MIO and DIO, the regularized coherence cost is given by the relative entropy of coherence, which we call the asymptotic coherence cost under MIO, DIO [24], [26],

$$C_r(\rho) = \min_{\delta \in \mathcal{I}} D(\rho\|\delta) = S(\Delta(\rho)) - S(\rho), \quad (11)$$

where $D(X\|Y) = -\text{Tr}[X(\log Y - \log X)]$ is the relative entropy, and we have

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{n} C_{c,\text{MIO/DIO}}^\varepsilon(\rho^{\otimes n}) = C_r(\rho). \quad (12)$$

In the above, we use \lim instead of \limsup or \liminf for simplicity. On the other hand, for IO and SIO, the regularized coherence cost is given by the so-called coherence of formation [17], which we call the asymptotic coherence cost under IO, SIO and is defined as

$$C_f(\rho) = \min_{p_i, |\psi_i\rangle} \sum_i p_i C_r(\psi_i), \quad (13)$$

where

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{n} C_{c,\text{IO/SIO}}^\varepsilon(\rho^{\otimes n}) = C_f(\rho). \quad (14)$$

In contrast to the entanglement of formation, the coherence of formation is an additive function [17].

C. Coherence Distillation

We next turn to the main focus of this paper, which is the task of transforming a given state ρ into a maximally coherent pure state. The one-shot distillation of the problem with ε error is stated as follows.

Definition 2. Let $\mathcal{O} \in \{\text{MIO, DIO, IO, SIO}\}$ denote some class of incoherent operations. Then for a given state ρ and $\varepsilon \geq 0$, the one-shot coherence distillation rate with error ε under \mathcal{O} is defined as

$$C_{d,\mathcal{O}}^\varepsilon(\rho) = \max_{\Lambda \in \mathcal{O}} \{ \log M : F(\Lambda(\rho), \Psi_M)^2 \geq 1 - \varepsilon \}. \quad (15)$$

In order to quantify the one-shot distillable rates, we first introduce two additional functions. The first one is based on the min quantum Rényi divergence $D_{\min}(\rho\|\sigma) = -\log F(\rho, \sigma)^2 = \tilde{D}_{1/2}(\rho\|\sigma)$ [27], $\alpha = \frac{1}{2}$ in Eq. (4). Using this quantity, the *min-entropy of coherence* $C_{\min}(\rho)$ can be defined as

$$C_{\min}(\rho) = \min_{\delta \in \mathcal{I}} D_{\min}(\rho\|\delta). \quad (16)$$

The properties of C_{\min} and its relationship between other coherence monotones are discussed in [31].

The second quantifier of coherence relevant to our study is based on the hypothesis testing relative entropy [32]–[34]. In the task of hypothesis testing, a positive operator valued measure (POVM) with two elements $\{W, \mathbb{1} - W\}$ is used to distinguish two possible states ρ and σ . The probability of obtaining a correct guess on ρ is $\text{Tr}(\rho W)$ and the probability of obtaining a wrong guess on σ is $\text{Tr}(\sigma W)$. The hypothesis testing relative entropy characterizes the minimal wrong guessing probability on σ , with the constraint that the probability of the correct guess on ρ is no less than $1 - \varepsilon$:

$$D_H^\varepsilon(\rho\|\sigma) = -\log \min \{ \text{Tr} \sigma W : 0 \leq W \leq \mathbb{1}, \text{Tr} \rho W \geq 1 - \varepsilon \}. \quad (17)$$

This serves as a parent quantity for two other coherence functions. Namely, we have

$$C_H^\varepsilon(\rho) = \min \{ D_H^\varepsilon(\rho\|\delta) : \delta \in \mathcal{I} \}, \quad (18)$$

$$\tilde{C}_H^\varepsilon(\rho) = \min \{ D_H^\varepsilon(\rho\|\sigma) : \sigma = \Delta(\sigma), \text{Tr} \sigma = 1 \}. \quad (19)$$

The function $\tilde{C}_H^\varepsilon(\rho)$ characterizes the one-shot distillable coherence under MIO and DIO. Note that the set of σ in $\tilde{C}_H^\varepsilon(\rho)$ is a larger set of operators than the set of incoherent states \mathcal{I} and does not need to be positive semi-definite. In the many-copy scenario, the regularized coherence distillation rate by operations $\mathcal{O} \in \{\text{MIO, DIO, IO, SIO}\}$ is defined as

$$C_{d,\mathcal{O}}^\infty(\rho) = \liminf_{\varepsilon \rightarrow 0^+} \liminf_{n \rightarrow \infty} \frac{1}{n} C_{d,\mathcal{O}}^\varepsilon(\rho^{\otimes n}). \quad (20)$$

For MIO, DIO, and IO, the distillable coherence of ρ is given by the relative entropy of coherence, which we call the asymptotic coherence distillation rate [17], where

$$\lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{1}{n} C_{d,\text{MIO/DIO/IO}}^\varepsilon(\rho^{\otimes n}) = C_r(\rho). \quad (21)$$

Combining Eq. (12), (21), we can conclude that coherence is asymptotically reversible under MIO/DIO. In contrast, as we show below, certain coherent states are undistillable under SIO. A summary of the distillation/formation rates for different operations is given in Table I.

III. ONE-SHOT DISTILLATION UNDER MIO AND DIO

In this section, we review the results of [25] on MIO and DIO distillation.

For a quantum channel $\Lambda : \mathcal{L}(A) \rightarrow \mathcal{L}(B)$, consider its Choi operator

$$\Gamma = (\text{id} \otimes \Lambda)\Phi = \sum_{i,j} |i\rangle\langle j|_R \otimes \Lambda(|i\rangle\langle j|_A), \quad (22)$$

where R is isomorphic to A . Let $\Gamma_{ij} = \Lambda(|i\rangle\langle j|_A)$. By Choi's theorem, Λ is completely positive if and only if $\Gamma \geq 0$. Also, Λ is trace preserving if and only if $\text{Tr} \Gamma_{ij} = \delta_{ij}$.

Now consider $\Lambda \in \text{MIO}$, meaning $\Lambda(\delta) \in \mathcal{I}$ for all $\delta \in \mathcal{I}$, which is equivalent to $\Lambda(|i\rangle\langle i|) \in \mathcal{I}$ for all i . For the Choi matrix this says $\Gamma_{ii} \in \mathcal{I}$ for all i . For $\Lambda \in \text{DIO}$, from $\Lambda(\Delta(|i\rangle\langle j|)) = \Delta(\Lambda(|i\rangle\langle j|))$ for all i and j , we get $\delta_{ij} \Gamma_{ij} = \Delta(\Gamma_{ij})$.

Since $\Lambda(\rho) = \sum_{i,j} \rho_{ij} \Gamma_{ij}$ (by linearity), we have now,

$$F(\Lambda(\rho), \Psi_M)^2 = \text{Tr}[\Lambda(\rho)\Psi_M] = \sum_{i,j} \rho_{ij} \text{Tr} \Gamma_{ij} \Psi_M. \quad (23)$$

The one-shot distillation under MIO/DIO can thus be expressed as

$$\begin{aligned} \max M \text{ s.t. } & \sum_{i,j} \rho_{ij} \text{Tr} \Gamma_{ij} \Psi_M \geq 1 - \varepsilon, \sum_{i,j} |i\rangle\langle j| \otimes \Gamma_{ij} \geq 0, \\ & \text{Tr} \Gamma_{ij} = \delta_{ij}, \begin{cases} \Gamma_{ii} \in \mathcal{I}[\text{MIO}], \\ \delta_{ij} \Gamma_{ij} = \Delta(\Gamma_{ij})[\text{DIO}]. \end{cases} \end{aligned} \quad (24)$$

Suppose that the Γ_{ij} satisfy the above constraints. Consider the twirling transformation

$$\tilde{\Gamma}_{ij} = \frac{1}{M!} \sum_{\pi \in S_M} U_\pi \Gamma_{ij} U_\pi^\dagger, \quad (25)$$

where S_M denotes the permutation group on M objects, and U_π is the unitary representation of π as a permutation matrix. It is easy to see that then the $\tilde{\Gamma}_{ij}$ also satisfy the constraints, and lead to the same objective function, due to the permutation invariance of Ψ_M . Therefore, without loss of generality, we can assume that each of the Γ_{ij} is permutation-invariant and has hence the form

$$\Gamma_{ij} = \alpha_{ij} \Psi_M + \beta_{ij} \frac{\mathbb{1} - \Psi_M}{M - 1}. \quad (26)$$

Furthermore, $\text{Tr} \Gamma_{ij} = \alpha_{ij} + \beta_{ij}$. We can see that both MIO and DIO lead to the following constraints: When $i \neq j$, we have $\beta_{ij} = -\alpha_{ij}$ and when $i = j$, we have $\alpha_{ii} = \frac{1}{M}$, $\beta_{ii} = 1 - \frac{1}{M}$. For the matrices $A = (\alpha_{ij})_{i,j}$ and $B = (\beta_{ij})_{i,j}$, this translates into the simple relation $B = \mathbb{1} - A$. Notice that

$$\Gamma = \sum_{i,j} \alpha_{ij} |i\rangle\langle j| \otimes \Psi_M + \sum_{i,j} \beta_{ij} |i\rangle\langle j| \otimes \frac{\mathbb{1} - \Psi_M}{M - 1}, \quad (27)$$

and we must satisfy $\langle \psi | \Gamma | \psi \rangle \geq 0$ for all $|\psi\rangle$. Consider the special case $|\psi\rangle = |u\rangle \otimes |v\rangle$; by choosing $|v\rangle = |\Psi_M\rangle$, we learn that $A \geq 0$. By choosing $|v\rangle$ orthogonal with $|\Psi_M\rangle$,

we see likewise that $B \geq 0$. Conversely, $A, B \geq 0$ implies that $\Gamma \geq 0$.

In conclusion, one-shot distillation under MIO/DIO can be solved by the following SDP: $C_{d,\text{MIO/DIO}}^{\varepsilon}(\rho) = \log M_{\text{opt}}$, with

$$\begin{aligned} M_{\text{opt}} &= \max M \text{ s.t. } \text{Tr } \rho^T A \geq 1 - \varepsilon, \\ 0 \leq A &\leq \mathbb{1}, \quad A_{ii} = \frac{1}{M} \forall i. \end{aligned} \quad (28)$$

To express this result in terms of a suitable relative-entropy distance, we recall the *hypothesis testing relative entropy* introduced above [32]–[34],

$$\begin{aligned} D_H^{\varepsilon}(\rho\|\sigma) &= -\log \min \{ \text{Tr } \sigma W : 0 \leq W \leq \mathbb{1}, \\ \text{Tr } \rho W &\geq 1 - \varepsilon \}, \end{aligned} \quad (29)$$

and the corresponding coherence measure

$$C_H^{\varepsilon}(\rho) = \min_{\delta \in \mathcal{I}} D_H^{\varepsilon}(\rho\|\delta). \quad (30)$$

We show that the coherence measure $C_H^{\varepsilon}(\rho)$ can be computed by SDP. Notice that

$$\begin{aligned} C_H^{\varepsilon}(\rho) &= -\log \max_{\delta \in \mathcal{I}} \min_{\substack{W: 0 \leq W \leq \mathbb{1} \\ \text{Tr } \rho W \geq 1 - \varepsilon}} \text{Tr } \delta W \\ &= -\log \min_{\substack{W: 0 \leq W \leq \mathbb{1} \\ \text{Tr } \rho W \geq 1 - \varepsilon}} \max_{\delta \in \mathcal{I}} \text{Tr } \delta W \\ &= -\log \min_{\substack{W: 0 \leq W \leq \mathbb{1} \\ \text{Tr } \rho W \geq 1 - \varepsilon}} \max_i W_{ii}, \end{aligned} \quad (31)$$

where in the second line we have appealed to the minimax theorem [35], noting that the objective function, $\text{Tr } \delta W$, is linear in each of the two arguments, and that the domains of optimization are closed convex sets.

The third line in Eq. (31) can now be expressed as an SDP as follows:

$$\begin{aligned} \mu_{\text{opt}} &= \max \mu \text{ s.t. } \text{Tr } \rho W \geq 1 - \varepsilon, \\ 0 \leq W &\leq \mathbb{1}, \quad W_{ii} \leq \frac{1}{\mu} \forall i, \end{aligned} \quad (32)$$

where $C_H^{\varepsilon}(\rho) = \log \mu_{\text{opt}}$. Notice that in Eq. (28) we have a transpose on ρ , while in Eq. (32) we don't. We can simply change A into A^T in the SDP (28) without changing its value, so that the two SDPs have similar form, except that one has an equality sign where the other other has a “ \leq ”. In [25], it was shown that the r.h.s. of Eq. (28) can be expressed in terms of D_H^{ε} , as well:

$$\begin{aligned} C_{d,\text{MIO/DIO}}^{\varepsilon}(\rho) &= \max \log M \\ \text{s.t. } \text{Tr } \rho^T A &\geq 1 - \varepsilon, \\ 0 \leq A &\leq \mathbb{1}, \quad A_{ii} = \frac{1}{M} \forall i \\ &= \min D_H^{\varepsilon}(\rho\|\delta) \\ \text{s.t. } \delta \text{ diagonal and } \text{Tr } \delta &= 1 \\ &=: \tilde{C}_H^{\varepsilon}(\rho), \end{aligned} \quad (33)$$

where the minimization is crucially over Hermitian, but not necessarily positive semidefinite matrices δ .

We record these findings in the following theorem.

Theorem 3 (Regula *et al.* [25]). *For any state ρ , the one-shot MIO- and DIO-distillable coherence is given by*

$$C_{d,\text{DIO}}^{\varepsilon}(\rho) = C_{d,\text{MIO}}^{\varepsilon}(\rho) = \tilde{C}_H^{\varepsilon}(\rho) \leq C_H^{\varepsilon}(\rho). \quad (34)$$

IV. ONE-SHOT DISTILLATION UNDER IO

Now we come to the main contribution of the present paper, the extension of the one-shot distillation protocols to the class IO.

To start, recall the definition of the min-relative entropy and its smoothed version,

$$\begin{aligned} D_{\min}(\rho\|\sigma) &= -\log F(\rho, \sigma)^2, \\ D_{\min}^{\varepsilon}(\rho\|\sigma) &= \max_{\rho' \in B^{\varepsilon}(\rho)} D_{\min}(\rho'\|\sigma). \end{aligned} \quad (35)$$

The *smooth min-relative entropy of coherence* is defined as

$$C_{\min}^{\varepsilon}(\rho) = \max_{\rho' \in B^{\varepsilon}(\rho)} \min_{\delta \in \mathcal{I}} D_{\min}(\rho\|\delta). \quad (36)$$

One might wonder why we do not interchange min and max here, but this is the version of the quantity that appears naturally both in the achievability bound we will derive on Subsection IV-A, and in the upper bound in Subsection IV-B.

A. An Achievable Lower Bound

We will generalize the protocol in [17, Thm. 6] to obtain a min-relative-entropic lower bound on the distillable coherence. In the process, the privacy amplification aspect will become even more apparent.

Theorem 4. *For an arbitrary state ρ and $0 < \varepsilon < 1$,*

$$C_{d,\text{IO}}^{\varepsilon}(\rho) \geq C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2 \log \frac{1}{\eta}, \quad (37)$$

for any $0 < \eta < \frac{\varepsilon}{2}$.

Proof. In order to accomplish our proof, we need to introduce the conditional min/max entropy and their smoothed versions. For a bipartite quantum state ρ^{AB} , the min-entropy of A conditioned on B is defined as

$$H_{\min}(A|B)_{\rho^{AB}} := -\min_{\sigma^B} D_{\max}(\rho^{AB}\|\mathbb{1}^A \otimes \sigma^B), \quad (38)$$

and the max-entropy of A conditioned on B is defined as

$$H_{\max}(A|B)_{\rho^{AB}} := -H_{\min}(A|C)_{\rho^{AC}}, \quad (39)$$

where ρ^{ABC} is a purification of ρ^{AB} and $\rho^{AC} = \text{Tr}_B \rho^{ABC}$. It has been proven that the max-entropy has an alternative form [36]

$$H_{\max}(A|B)_{\rho^{AB}} := -\min_{\sigma^B} D_{\min}(\rho^{AB}\|\mathbb{1}^A \otimes \sigma^B). \quad (40)$$

Furthermore, the smoothed conditional min- and max-entropy are defined by

$$\begin{aligned} H_{\min}^{\varepsilon}(A|B)_{\rho^{AB}} &:= \max_{\rho' \in B^{\varepsilon}(\rho^{AB})} H_{\min}^{\varepsilon}(A|B)_{\rho'}, \\ H_{\max}^{\varepsilon}(A|B)_{\rho^{AB}} &:= \min_{\rho' \in B^{\varepsilon}(\rho^{AB})} H_{\max}^{\varepsilon}(A|B)_{\rho'}. \end{aligned} \quad (41)$$

In the following proof, we denote the system of interest as A , i.e. $\rho^A := \rho$, and the incoherent basis as $\{|x\rangle\}$. Denote $\psi := |\psi\rangle\langle\psi|$. Choose a purification of ρ^A ,

$$|\psi\rangle^{AE} = \sum_x \sqrt{p_x} |x\rangle^A |\psi_x\rangle^E,$$

where $\text{Tr}_E \psi^{AE} = \rho^A$, and introduce the dephased cq-state

$$\omega^{AE} = (\Delta \otimes \text{id})\psi = \sum_x p_x |x\rangle\langle x|^A \otimes \psi_x^E.$$

According to [37, Cor. 5.6.1] (actually, a slight adaptation to get rid of a factor of 2), for every $\log M \leq H_{\min}^{\frac{\varepsilon}{2}-\eta}(A|E)_\omega - 2\log\frac{1}{\eta}$ there exists a function $G : \mathcal{X} \rightarrow [M] = \{1, 2, \dots, M\}$ such that the state

$$\Omega^{KE} := (G \otimes \text{id})\omega^{AE} = \sum_x p_x |G(x)\rangle\langle G(x)|^K \otimes \psi_x^E$$

satisfies

$$\frac{1}{2} \|\Omega^{KE} - \tau_K \otimes \sigma^E\|_1 \leq \frac{\varepsilon}{2},$$

for $\tau^K = \frac{1}{M} \mathbb{1}_K$ and a suitable state σ^E (which may be equal to ω^E , but it doesn't have to be). Hence, because of the well-known relation between trace distance and fidelity [30],

$$1 - F(\rho, \sigma) \leq \frac{1}{2} \|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2},$$

we have $F(\Omega^{KE}, \tau_K \otimes \sigma^E) \geq 1 - \frac{\varepsilon}{2}$.

To turn this property, which encapsulates the uniformity and independence of the “key” $K = G(X)$ from E , into a coherence distillation protocol, consider the following purification of Ω^{KE} :

$$|\Omega\rangle^{KAE} := \sum_x \sqrt{p_x} |G(x)\rangle^K |x\rangle^A |\psi_x\rangle^E,$$

which can be obtained from $|\psi\rangle^{AE}$ by applying the isometry

$$U : |x\rangle^A \mapsto |G(x)\rangle^K |x\rangle^A.$$

Crucially, U is incoherent (even SIO). For $\tau_K \otimes \sigma^E$, on the other hand, we choose a purification $|\Phi\rangle^{KL} \otimes |\zeta\rangle^{EF}$, with the standard maximally entangled state $|\Phi\rangle = \frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} |i\rangle|i\rangle$. By Uhlmann's theorem [38], there exists an isometry $V : A \hookrightarrow LF$ such that

$$|\langle\Phi|\langle\zeta| \cdot (\mathbb{1}_{KE} \otimes V)|\Omega\rangle| \geq 1 - \frac{\varepsilon}{2}. \quad (42)$$

If we had $|\Phi\rangle^{KL} |\zeta\rangle^{EF}$, we could clearly create a maximally coherent state Ψ_M on system K by tracing out F , and destructively measuring L in the Fourier conjugate basis $|\widehat{\alpha}\rangle := Z^\alpha |\Psi_M\rangle$. Here Z is the phase unitary defined in Eq. (1).

With this, the protocol is clear: starting from ρ^A , first apply U ; then apply V , followed by tracing out F , measuring L in the Fourier basis $\{|\widehat{\alpha}\rangle\}_{\alpha=0,\dots,M-1}$; finally, apply Z^α on K . The first and the third step clearly are incoherent (even SIO); the second seems suspicious, and indeed V may not be incoherent at all, but notice that we follow it by a destructive measurement, and these are all IO. We can write Kraus operators of the resulting map $\Lambda : A \rightarrow K$ as follows:

$$M_{\alpha\beta} := (Z_K^\alpha \otimes \langle\widehat{\alpha}|^L \langle\beta|F V)U,$$

where $\{|\beta\rangle\}$ is an arbitrary basis of F .

To analyze the fidelity of the protocol, we pass to the purifications and look at eq. (42); by the monotonicity of the fidelity under the CPTP maps Λ and Tr_E , we obtain $F(\Lambda(\rho), \Psi_M) \geq 1 - \frac{\varepsilon}{2}$ and hence $F(\Lambda(\rho), \Psi_M)^2 \geq 1 - \varepsilon$.

This shows that $\log M \approx H_{\min}^{\frac{\varepsilon}{2}-\eta}(A|E)_\omega - 2\log\frac{1}{\eta}$ is achievable. Now, introducing the purification $|\omega\rangle^{ABE} = \sum_x \sqrt{p_x} |x\rangle^A |x\rangle^B |\psi_x\rangle$, we have

$$\begin{aligned} H_{\min}^a(A|E)_\omega &= -H_{\max}^a(A|B)_\omega \\ &= -\log \min_{\omega' \in B^a(\omega)} \max_{\sigma} F(\omega', \mathbb{1} \otimes \sigma)^2 \\ &\geq -\log \min_{\substack{\omega' \in B^a(\omega) \\ \text{max. correlated}}} \max_{\sigma} F(\omega', \mathbb{1} \otimes \sigma)^2 \\ &= -\log \min_{\substack{\omega' \in B^a(\omega) \\ \text{max. correlated}}} \max_{\sigma \in \mathcal{I}} F(\omega', \mathbb{1} \otimes \sigma)^2 \\ &= -\log \min_{\rho' \in B^a(\rho)} \max_{\delta \in \mathcal{I}} F(\rho', \delta)^2 \\ &= \max_{\rho' \in B^a(\rho)} \min_{\delta \in \mathcal{I}} D_{\min}(\rho' \|\delta) \\ &= C_{\min}^a(\rho). \end{aligned} \quad (43)$$

Here, the second line follows from Eq. (40), and the third line is motivated by the observation that $\omega^{AB} = \sum_{xy} \sqrt{p_x p_y} \langle\psi_y|\psi_x\rangle |xx\rangle\langle yy|$ is a maximally correlated state, so it is natural to impose the same structure on ω' ; the fourth line follows from the $Z \otimes Z^\dagger$ -invariance of maximally correlated states, so by the concavity of the fidelity we can impose w.l.o.g. the same structure on $\mathbb{1} \otimes \sigma$, meaning that σ is diagonal. The rest is straightforward algebra. ■

Note that the equality in Eq. (43) can also be achieved by directly using the results by Coles [39], or [31]:

$$H_{\min}(X|E)_{\rho_{XE}} = \min_{\delta \in \mathcal{I}} D_{\min}(\rho_A \|\delta). \quad (44)$$

Remark 5. The CPTP map $\Lambda(\cdot) = \sum_{\alpha,\beta} M_{\alpha\beta}(\cdot) M_{\alpha\beta}^\dagger$ we constructed in the proof is not only IO but also DIO. To see this, first expand

$$\Lambda(|x\rangle\langle y|) = \sum_{\alpha,\beta} Z^\alpha |G(x)\rangle\langle G(y)| Z^{\alpha\dagger} \langle\widetilde{\alpha}\beta| V|x\rangle\langle y| V^\dagger |\widetilde{\alpha}\beta\rangle$$

for any incoherent basis states $|x\rangle$ and $|y\rangle$ of system A . The key observation is that fully dephased states are invariant under conjugation by Z^α , and moreover this conjugation commutes with Δ , i.e. $\Delta[Z^\alpha(\cdot)Z^{\alpha\dagger}] = \Delta(\cdot)$. Hence if we dephase system K after applying the map Λ , we find

$$\begin{aligned} \Delta[\Lambda(|x\rangle\langle y|)] &= \delta_{G(x)G(y)} |G(x)\rangle\langle G(y)| \\ &\quad \times \sum_{\alpha,\beta} \langle\widetilde{\alpha}\beta| V|x\rangle\langle y| V^\dagger |\widetilde{\alpha}\beta\rangle \\ &= \delta_{xy} |G(x)\rangle\langle G(y)| = \Lambda[\Delta(|x\rangle\langle y|)], \end{aligned} \quad (45)$$

where the second equality follows from the fact that $\{|\widetilde{\alpha}\beta\rangle\}_{\alpha,\beta}$ forms a complete basis. Thus, Λ is a DIO map.

Corollary 6. For an arbitrary state ρ and $0 < \varepsilon < 1$,

$$C_{d,\text{DIO}}^\varepsilon(\rho) \geq C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2\log\frac{1}{\eta}, \quad (46)$$

for any $0 < \eta < \frac{\varepsilon}{2}$, where DIO refers to the intersection of IO and DIO. ■

B. Upper Bound and Comparison With MIO Distillation

We have a partial converse theorem to Theorem 4 which can bound $C_{d,\text{IO}}^\varepsilon(\rho)$ from both sides, as follows.

Theorem 7. *For an arbitrary state ρ and $0 < \varepsilon < 1$,*

$$C_{d,\text{IO}}^\varepsilon(\rho) \leq C_{\min}^{\sqrt{\varepsilon(2-\varepsilon)}}(\rho). \quad (47)$$

Proof. Due to the inclusion of the classes of operations, and Theorems 3 and 4, we have the first four of the following (in)equalities:

$$\begin{aligned} C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2 \log \frac{1}{\eta} &\leq C_{d,\text{IO}}^\varepsilon(\rho) \\ &\leq C_{d,\text{M/DIO}}^\varepsilon(\rho) = \tilde{C}_H^\varepsilon(\rho) \\ &\leq C_H^\varepsilon(\rho) \\ &\leq C_{\min}^{\sqrt{\varepsilon(2-\varepsilon)}}(\rho). \end{aligned} \quad (48)$$

The last one also follows essentially from known facts, namely [34, Prop. 4.2]. Note only that our definition of D_H^ε differs from [34] by $\varepsilon \leftrightarrow 1 - \varepsilon$, and an additional term of $\log(1 - \varepsilon)$ added. We made this choice for easier comparison with the results from [25]. With this in mind, [34, Eq. (51)] reads

$$D_H^\varepsilon(\rho\|\sigma) \leq D_{\min}^{\sqrt{2\varepsilon}}(\rho\|\sigma),$$

and looking at the last step of the proof, one observes that $\sqrt{2\varepsilon}$ can be improved to $\sqrt{\varepsilon(2-\varepsilon)}$. We can adapt the proof in [34] to include the minimization over $\delta \in \mathcal{I}$, according to the following Lemma 9. By applying it with $\mathcal{S} = \mathcal{I}$, and maximizing over the $\sqrt{\varepsilon(2-\varepsilon)}$ -ball on the right hand side, we precisely obtain the last inequality in Eq. 48. ■

Remark 8. *Combining Theorem 4 and Theorem 7, we conclude that $C_{d,\text{IO}}^\varepsilon(\rho) \approx C_{\min}^{\varepsilon'}(\rho)$, with $\varepsilon' \in [\frac{\varepsilon}{2}, \sqrt{\varepsilon(2-\varepsilon)}]$.*

Lemma 9. *Let \mathcal{S} be a closed convex set of states on a Hilbert space A , and ρ a state. Then, for every $0 < \varepsilon < 1$ there exists a subnormalized density matrix ρ' with $P(\rho, \rho') \leq \sqrt{\varepsilon(2-\varepsilon)}$, such that*

$$\min_{\sigma \in \mathcal{S}} D_H^\varepsilon(\rho\|\sigma) \leq \min_{\sigma \in \mathcal{S}} D_{\min}^{\sqrt{\varepsilon(2-\varepsilon)}}(\rho'\|\sigma).$$

Proof. The crucial observation is that due to the convexity of the sets of operators, \mathcal{S} and $\{0 \leq W \leq \mathbb{1} : \text{Tr } \rho W \geq 1 - \varepsilon\}$, we can invoke the minimax theorem [35], to obtain

$$\max_{\sigma \in \mathcal{S}} \min_{\substack{0 \leq W \leq \mathbb{1} \\ \text{Tr } \rho W \geq 1 - \varepsilon}} \text{Tr } \sigma W = \min_{\substack{0 \leq W \leq \mathbb{1} \\ \text{Tr } \rho W \geq 1 - \varepsilon}} \max_{\sigma \in \mathcal{S}} \text{Tr } \sigma W.$$

Thus, there exists an optimizer W_0 of the second expression, $0 \leq W_0 \leq \mathbb{1}$, $\text{Tr } \rho W_0 \geq 1 - \varepsilon$, with

$$\min_{\sigma \in \mathcal{S}} D_H^\varepsilon(\rho\|\sigma) = \min_{\sigma \in \mathcal{S}} -\log \text{Tr } \sigma W_0. \quad (49)$$

Following the example of [34, Prop. 4.2], we define a subnormalized state $\rho' = \sqrt{W_0} \rho \sqrt{W_0}$, which we claim to be the sought-after object.

To start with, from optimality of W_0 , we have $\text{Tr } \rho' = \text{Tr } \rho W_0 = 1 - \varepsilon$, hence from [34, Lemma A.3] (see also [40, Lemma 7]), $P(\rho', \rho) \leq \sqrt{1 - (\text{Tr } \rho W_0)^2} = \sqrt{\varepsilon(2 - \varepsilon)}$.

At the same time, choosing a purification of $\rho^A = \text{Tr}_B |\varphi\rangle\langle\varphi|^{AB}$, we get a purification of ρ' by letting $|\varphi'\rangle = (\sqrt{W_0} \otimes \mathbb{1})|\varphi\rangle$. Conjugating the inequality $\varphi \leq \mathbb{1}$ by $\sqrt{W_0} \otimes \mathbb{1}$ this results in $\varphi' \leq W_0 \otimes \mathbb{1}$. Now, just as in the proof of [34, Prop. 4.2], we employ the dual variational characterization of the fidelity,

$$F(\rho', \sigma)^2 = \min \text{Tr } \sigma Z \text{ s.t. } \varphi' \leq Z \otimes \mathbb{1}.$$

This implies, that $\text{Tr } \sigma W_0 \geq F(\rho', \sigma)$ in (49), for all $\sigma \in \mathcal{S}$, and so $\min_{\sigma \in \mathcal{S}} D_H^\varepsilon(\rho\|\sigma) \leq \min_{\sigma \in \mathcal{S}} -\log F(\rho', \sigma)^2$, as claimed. ■

V. DISTILLATION UNDER SIO

A. Characterizing One-Shot SIO Distillation

We now turn to coherence distillation under strictly incoherent operations (SIO). Ever since [17], it has been an open question whether coherence distillation such as the protocol in Sec. IV-A, or in [17, Thm. 6], really requires IO, or can be performed within the smaller class of SIO (as all other protocols discussed in [17] can). The crucial object in this setting turns out to be the incoherent rank. Recall that the incoherent rank of a positive operator Ω is defined by

$$C_0(\Omega) = \min_{\{\lambda_j, |\phi_j\rangle\}} \max_j \log \text{rank}[\Delta(\phi_j)], \quad (50)$$

where the minimization is taken over all positive rank-one decompositions of Ω .

Theorem 10. *For any state ρ , the one-shot distillable coherence under SIO is given by*

$$\begin{aligned} C_{d,\text{SIO}}^\varepsilon(\rho) &= \max \log M \text{ s.t. } \text{Tr } \rho A \geq 1 - \varepsilon, \\ &0 \leq A \leq \mathbb{1}, \quad A_{ii} = \frac{1}{M} \forall i \\ &C_0(A) \leq \log M. \end{aligned} \quad (51)$$

Proof. Suppose that $\text{Tr}[\Lambda(\rho)\Psi_M] \geq 1 - \varepsilon$ for some SIO map Λ and $|\Psi_M\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^M |i\rangle$. Let $\Pi_{>M} = \frac{1}{M} \sum_{x=M+1}^d |x\rangle\langle x|$. Notice that

$$\begin{aligned} 1 - \varepsilon &\leq \text{Tr}[\Lambda(\rho)\Psi_M] \\ &\leq \text{Tr}[\Lambda(\rho)(\Psi_M + \Pi_{>M})] \\ &= \text{Tr}[\rho A], \end{aligned} \quad (52)$$

where $A := \Lambda^*(\Psi_M) + \Lambda^*(\Pi_{>M})$ and Λ^* is the adjoint channel of Λ . Using the form of SIO Kraus operators, we have

$$\begin{aligned} \Lambda^*(\Psi_M) &= \frac{1}{M} \sum_{\alpha} \sum_{\substack{x \text{ s.t.} \\ f_{\alpha}(x) \in [M]}} \sum_{\substack{x' \text{ s.t.} \\ f_{\alpha}(x') \in [M]}} c_{\alpha,x}^* |x\rangle\langle x'|c_{\alpha,x'} \\ &= \frac{1}{M} \sum_{\alpha} |\phi_{\alpha}\rangle\langle\phi_{\alpha}|, \end{aligned} \quad (53)$$

and likewise

$$\Lambda^*(\Pi_{>M}) = \frac{1}{M} \sum_{\alpha} \sum_{\substack{x \text{ s.t.} \\ f_{\alpha}(x) \notin [M]}} |c_{\alpha,x}|^2 |x\rangle\langle x|. \quad (54)$$

Thus A has a decomposition into rank-one vectors each having an incoherent rank no greater than M . Also, for any $y \in \{1, \dots, d_A\}$, we see that

$$\langle y|A|y\rangle = \frac{1}{M} \left(\sum_{\substack{\alpha \text{ s.t.} \\ f_{\alpha}(y) \in [M]}} + \sum_{\substack{\alpha \text{ s.t.} \\ f_{\alpha}(y) \notin [M]}} \right) |c_{\alpha,y}|^2 = \frac{1}{M}, \quad (55)$$

i.e. $A_{yy} = \frac{1}{M}$ for all y .

The converse involves essentially reversing these steps. Suppose that $\text{Tr } \rho A \geq 1 - \varepsilon$ for some operator $0 \leq A \leq \mathbb{1}$ with $C_0(A) \leq \log M$ and $A_{ii} = \frac{1}{M}$. Then there exists a decomposition

$$\begin{aligned} A &= \frac{1}{M} \sum_{\alpha} |\phi_{\alpha}\rangle\langle\phi_{\alpha}| \\ &= \frac{1}{M} \sum_{\alpha} \sum_{x,x'=1}^{d_A} c_{\alpha,x} c_{\alpha,x'}^* |x\rangle\langle x'|, \end{aligned} \quad (56)$$

where $(c_{\alpha,x})_x$ contains at most M nonzero elements for each α . Hence we can define permutations f_{α} on the set $\{1, \dots, |A|\}$ such that $f_{\alpha}(x) \in [M]$ for every x and α with $c_{\alpha,x} \neq 0$. The Kraus operators $K_{\alpha} = \sum_{x=1}^{d_A} c_{\alpha,x}^* |f_{\alpha}(x)\rangle\langle x|$ satisfy

$$\sum_{\alpha} K_{\alpha}^{\dagger} \Psi_M K_{\alpha} = \Omega. \quad (57)$$

Furthermore, $\sum_{\alpha} K_{\alpha}^{\dagger} K_{\alpha} = \sum_{\alpha} \sum_{x=1}^{d_A} |c_{\alpha,x}|^2 |x\rangle\langle x| = \mathbb{1}$, since by assumption

$$\frac{1}{M} = \langle x|A|x\rangle = \frac{1}{M} \sum_{\alpha} |c_{\alpha,x}|^2.$$

Therefore, the $\{K_{\alpha}\}_{\alpha}$ define a CPTP SIO map Λ satisfying $\text{Tr } \Lambda(\rho) \Psi_M \geq 1 - \varepsilon$. \blacksquare

Remark 11. Comparing with Eq. (33), we see that the one-shot distillable coherence under SIO takes the form of DIO distillation with the added constraint of $C_0(A) \leq \log M$.

Remark 12. An explicit calculation of the incoherent rank C_0 can be made through semi-definite programming techniques [41]. However the number of computational constraints scales as $\binom{d}{M}$ for certifying whether a d -dimensional state ρ has $C_0(\rho) \geq \log(M+1)$.

B. Bound Coherence Exists Under SIO

The constraint on the incoherent rank of A in Theorem 10 greatly diminishes the power of SIO to distill coherence. Here we illustrate this effect by a dramatic example. Consider the state

$$\rho = \frac{1}{2} (|+\rangle\langle +| \otimes |+\rangle\langle +| + |-\rangle\langle -| \otimes |\tilde{+}\rangle\langle \tilde{+}|), \quad (58)$$

where

$$\begin{aligned} |+\rangle|+\rangle &= \frac{1}{2} (|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\ |-\rangle|\tilde{+}\rangle &= \frac{1}{2} (|00\rangle + i|01\rangle - |10\rangle - i|11\rangle). \end{aligned}$$

The n -copy state $\rho^{\otimes n}$ is then an equal mixture of states belonging to the ensemble

$$\mathfrak{E}_n := \{|+\rangle|+\rangle, |-\rangle|\tilde{+}\rangle\}^{\otimes n}.$$

We will show that not a single cosbit of coherence can be distilled from $\rho^{\otimes n}$ by SIO with an error smaller than the minimal one-copy error. In comparison, n bits of coherence can be distilled by IO error-free: the first system in each copy of ρ is simply measured with the IO Kraus operations $\{|0\rangle\langle +|, |1\rangle\langle -|\}$ followed by a suitable controlled phase on the second qubit. Such a measurement is not possible by SIO.

Theorem 13. For the state ρ defined in Eq. (58), $C_{d, \text{SIO}}^{\infty}(\rho) = 0$.

The proof of this will follow by studying the structure of $\rho^{\otimes n}$ and showing that for a fixed value of ε (independent of n), $\text{Tr } \rho^{\otimes n} A < 1 - \varepsilon$ for any operator A having an incoherent rank of two and satisfying the conditions of Theorem 10. The key property we use is that the eigenvectors of $\rho^{\otimes n}$ will always be maximally coherent states with complex phases belonging to $\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. For the n -copy analysis to be tractable, we need to introduce some new notations. Let $\mathbf{b}_j = (b_{j,0}, b_{j,1}, \dots, b_{j,n-1}) \in \{0, 1\}^n$ denote the j^{th} binary sequence of length n . We then define an ensemble of 2^n equiprobable states $\{|\mathbf{b}_j\rangle\}_{j=1}^{2^n}$, where

$$|\mathbf{b}_j\rangle := \frac{1}{\sqrt{4^n}} \sum_{m_0, \dots, m_{n-1}=0}^3 \exp\left(i \frac{\pi}{2} \sum_{k=0}^{n-1} b_{j,k} m_k\right) \left| \sum_{k=0}^{n-1} 4^k m_k \right\rangle. \quad (59)$$

We claim that, up to relabelling, this ensemble is precisely \mathfrak{E}_n . For example, when $n = 1$ we have

$$\begin{aligned} |\mathbf{b}_1\rangle &= \frac{1}{2} (|0\rangle + |1\rangle + |2\rangle + |3\rangle) \\ |\mathbf{b}_2\rangle &= \frac{1}{2} (|0\rangle + i|1\rangle - |2\rangle - i|3\rangle), \end{aligned}$$

and for $n = 2$ we have

$$\begin{aligned} |\mathbf{b}_1\rangle &= \frac{1}{4} (|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle \\ &\quad + |5\rangle + |6\rangle + \dots + |14\rangle + |15\rangle) \\ |\mathbf{b}_2\rangle &= \frac{1}{4} (|0\rangle + i|1\rangle - |2\rangle - i|3\rangle + |4\rangle \\ &\quad + i|5\rangle - |6\rangle - \dots - |14\rangle - i|15\rangle) \\ |\mathbf{b}_3\rangle &= \frac{1}{4} (|0\rangle + |1\rangle + |2\rangle + |3\rangle + i|4\rangle \\ &\quad + i|5\rangle + i|6\rangle + \dots - i|14\rangle - i|15\rangle) \\ |\mathbf{b}_4\rangle &= \frac{1}{4} (|0\rangle + i|1\rangle - |2\rangle - i|3\rangle + i|4\rangle \\ &\quad - |5\rangle - i|6\rangle + \dots + i|14\rangle - |15\rangle). \end{aligned}$$

The case of general n can be checked by induction.

Now for any two distinct vectors $\vec{m} = (m_0, \dots, m_{n-1})$ and $\vec{m}' = (m'_0, \dots, m'_{n-1})$ belonging to $\{0, 1, 2, 3\}^n$, let us denote the kets

$$|\vec{m}\rangle := \left| \sum_{k=0}^{n-1} 4^k m_k \right\rangle, \quad |\vec{m}'\rangle := \left| \sum_{k=0}^{n-1} 4^k m'_k \right\rangle. \quad (60)$$

The relative phase between $|\vec{m}\rangle$ and $|\vec{m}'\rangle$ in any $|\mathbf{b}_j\rangle$ is defined to be

$$\frac{\pi}{2} \sum_{k=0}^{n-1} b_{j,k} (m'_k - m_k) \in \left\{ 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2} \right\}. \quad (61)$$

We now make a crucial observation about the distribution of relative phases among the $|\mathbf{b}_j\rangle$ in \mathfrak{E}_n .

Proposition 14. *For any fixed pair of distinct vectors \vec{m} and \vec{m}' , at most half of the $|\mathbf{b}_i\rangle$ in \mathfrak{E}_n have the same relative phase between $|\vec{m}\rangle$ and $|\vec{m}'\rangle$.*

Proof. Let $k \in \{0, \dots, n-1\}$ be chosen such that $m'_k - m_k \neq 0$. Let $\Delta m = m'_k - m_k$. Consider all the states $|\mathbf{b}_j\rangle \in \mathfrak{E}_n$ with binary sequence \mathbf{b}_j such that $b_{j,k} = 0$; this represents exactly half of all states in the ensemble \mathfrak{E}_n . We partition these states into four groups, A_0 , $A_{\pi/2}$, A_π , and $A_{3\pi/2}$, according to their respective relative phases between $|\vec{m}\rangle$ and $|\vec{m}'\rangle$. Now we consider the other half of the states in \mathfrak{E}_n , those having $b_{j,k} = 1$. We likewise partition these states into sets B_0 , $B_{\pi/2}$, B_π , and $B_{3\pi/2}$ of common relative phase between $|\vec{m}\rangle$ and $|\vec{m}'\rangle$. Notice that for any $|\mathbf{b}_j\rangle$ in, say, $A_{\pi/2}$, there will be a corresponding $|\mathbf{b}_{j'}\rangle$ in $B_{\pi/2(1+\Delta m)}$ and vice versa, the only difference between \mathbf{b}_j and $\mathbf{b}_{j'}$ being their k^{th} component. Hence $|A_0| = |B_{\pi/2(2\Delta m)}|$, $|A_{\pi/2}| = |B_{\pi/2(1+\Delta m)}|$, $|A_\pi| = |B_{\pi/2(2+\Delta m)}|$ and $|A_{3\pi/2}| = |B_{\pi/2(3+\Delta m)}|$, where all arithmetic is done modular 4. Therefore, the total number of states in the ensemble having a relative phase of, say, $\pi/2$ is

$$|A_{\pi/2}| + |B_{\pi/2}| = |A_{\pi/2}| + |A_{\pi/2(1-\Delta m)}| \leq 2^{n-1}.$$

The same bound likewise holds for the other three relative phases. \blacksquare

Proof of Theorem 13. Consider an arbitrary vector in the complex linear span of $|\vec{m}\rangle$, $|\vec{m}'\rangle$,

$$|\psi\rangle = \cos\theta|\vec{m}\rangle + \sin\theta e^{i\phi}|\vec{m}'\rangle. \quad (62)$$

Let N_0 , $N_{\pi/2}$, N_π , $N_{3\pi/2}$ denote the number of states in \mathfrak{E}_n having a relative phase between $|\vec{m}\rangle$ and $|\vec{m}'\rangle$ of 0 , $\pi/2$, π , and $3\pi/2$, respectively. We can then explicitly compute

$$\begin{aligned} \langle \psi | \rho^{\otimes n} | \psi \rangle &= \frac{1}{2^n 4^n} \sum_{i=1}^{2^n} |\langle \psi | \mathbf{b}_i \rangle|^2 \\ &= \frac{N_0}{2^n 4^n} |\cos\theta + e^{-i\phi} \sin\theta|^2 + \frac{N_{\pi/2}}{2^n 4^n} |\cos\theta + i e^{-i\phi} \sin\theta|^2 \\ &\quad + \frac{N_\pi}{2^n 4^n} |\cos\theta - e^{-i\phi} \sin\theta|^2 + \frac{N_{3\pi/2}}{2^n 4^n} |\cos\theta - i e^{-i\phi} \sin\theta|^2 \\ &= \frac{1}{4^n} + \frac{\sin 2\theta}{2^n 4^n} [(N_0 - N_\pi) \cos\phi + (N_{3\pi/2} - N_{\pi/2}) \sin\phi], \end{aligned} \quad (63)$$

where the last line follows by expanding out the squared amplitudes, the identity $2 \cos\theta \sin\theta = \sin 2\theta$, and using the fact that $N_0 + N_{\pi/2} + N_\pi + N_{3\pi/2} = 2^n$. Our goal is to maximize (63) under the constraint that $\max\{N_0, N_{\pi/2}, N_\pi, N_{3\pi/2}\} \leq 2^{n-1}$. This constraint implies that $|N_0 - N_\pi| \leq 2^{n-1}$ and $|N_{3\pi/2} - N_{\pi/2}| \leq 2^{n-1}$, and so

$$\begin{aligned} \langle \psi | \rho^{\otimes n} | \psi \rangle &\leq \frac{1}{4^n} (1 + \cos\theta \sin\theta (|\cos\phi| + |\sin\phi|)) \\ &\leq \frac{1}{4^n} \left(1 + \frac{\sqrt{2}}{2} \right). \end{aligned} \quad (64)$$

Suppose now that A has an incoherent rank of two and satisfies $\text{Tr}[A] = \frac{4^n}{2}$. Then by the previous calculation we have the fidelity bound

$$\begin{aligned} \text{Tr} \rho^{\otimes n} A &\leq (\text{Tr } A) \frac{1}{4^n} \left(1 + \frac{\sqrt{2}}{2} \right) = \frac{1}{2} \left(1 + \frac{\sqrt{2}}{2} \right) \\ &= 1 - \varepsilon, \end{aligned} \quad (65)$$

where $\varepsilon = \frac{1}{2} - \frac{\sqrt{2}}{4}$ is independent of n . This is precisely the single-copy error bound. As a consequence, it follows that $C_{d,SIO}^\infty(\rho) = 0$, proving the theorem. \blacksquare

Remark 15. *This result should be compared with the recent proof, by Marvian [42], that coherence distillation is generally impossible in the resource theory of energy conservation, which is characterized by the class of so-called time-translation-covariant operations (TIO). That class is difficult to compare with DIO, as at the single-system level, TIO is contained in DIO, but since the composition of systems works differently, it may result in TIO operations outside DIO on the multi-system level.*

The result of [42] shows that for generic mixed states, the rate of distilling cosbit states $|\Psi_2\rangle$ is zero, but that at the same time it is possible to obtain a single cosbit (or a sublinear number) with fidelity going to 1 as asymptotically many copies of the mixed resource become available. In contrast, here we showed that under SIO the fidelity remains bounded away from 1, irrespective of the number of resource states.

VI. RECOVERING THE INFORMATION THEORETIC LIMIT

In the asymptotic limit, the coherence distillation rate under operation class \mathcal{O} is defined as

$$C_{d,\mathcal{O}}^\infty(\rho) = \liminf_{\varepsilon \rightarrow 0^+} \liminf_{n \rightarrow \infty} \frac{1}{n} C_{d,\mathcal{O}}^\varepsilon(\rho^{\otimes n}). \quad (66)$$

From [17] and [25] (see also [26]) we know that $C_{d,\text{IO}}^\infty(\rho) = C_{d,\text{DIO}}^\infty(\rho) = C_{d,\text{MIO}}^\infty(\rho) = C_r(\rho)$. Below we show that our results on one-shot IO distillation can be used to recover the asymptotic limit, at the same time improving the result by showing that the limit exists and equals $C_r(\rho)$ for any fixed $0 < \varepsilon < 1$; such a statement is known as a strong converse in information theory.

Theorem 16. *For any state ρ and any $0 < \varepsilon < 1$,*

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} C_{d,\text{IO}}^\varepsilon(\rho^{\otimes n}) &= \lim_{n \rightarrow \infty} \frac{1}{n} C_{d,\text{DIO}}^\varepsilon(\rho^{\otimes n}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} C_{d,\text{MIO}}^\varepsilon(\rho^{\otimes n}) = C_r(\rho). \end{aligned}$$

Proof. Recall the results in Theorems 4 and 7, which state that for any $\eta < \frac{1}{2}\varepsilon$,

$$\begin{aligned} C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2 \log \frac{1}{\eta} &\leq C_{d,\text{IO}}^{\varepsilon}(\rho) \\ &\leq C_{d,\text{MIO}}^{\varepsilon}(\rho) \\ &= C_{d,\text{DIO}}^{\varepsilon}(\rho) \leq C_{\min}^{\sqrt{\varepsilon(2-\varepsilon)}}(\rho). \end{aligned} \quad (67)$$

Hence, to show the theorem, we only need to prove that for all $0 < \delta < 1$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} C_{\min}^{\delta}(\rho^{\otimes n}) = C_r(\rho), \quad (68)$$

which is equivalent to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \max_{\rho' \in B_{\delta}(\rho^{\otimes n})} H_{\min}(X^n | E^n)_{\omega'} = C_r(\rho). \quad (69)$$

Here, $\omega'_{X^n E^n} = (\Delta_{A^n} \otimes \text{id}_{E^n})|\psi'\rangle\langle\psi'|$ and $|\psi'\rangle$ is an arbitrary purification of ρ'_A . Recall the quantum asymptotic equipartition theorem [43], which states that for any $0 < \eta < 1$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \max_{\rho'_{AB} \in B_{\eta}(\rho_{AB}^{\otimes n})} H_{\min}(A|B)_{\rho'_{AB}} = H(A|B)_{\rho_{AB}}. \quad (70)$$

Now, in one direction, if we have a state $\rho' \in B_{\delta}(\rho^{\otimes n})$, then by Uhlmann's characterization of the fidelity there exists a purification $\psi' \in B_{\delta}(\rho^{\otimes n})$, hence $\omega' = (\Delta_{A^n} \otimes \text{id}_{E^n})\psi' \in B_{\delta}(\omega^{\otimes n})$, with $\omega^{X^n} = (\Delta_A \otimes \text{id}_E)\psi$. In the other direction, for $\omega'' \in B_{\delta}(\omega^{\otimes n})$, it is known that since ω is a cq-state, an optimal ω'' for $H_{\min}(X^n | E^n)$ may be assumed to be a cq-state as well [37], hence we can find a $\psi' \in B_{\delta}(\rho^{\otimes n})$ such that $\omega'' = (\Delta_{A^n} \otimes \text{id}_{E^n})\psi'$. Thus, we can conclude that

$$\max_{\rho'_A \in B_{\delta}(\rho^{\otimes n})} H_{\min}(X^n | E^n)_{\omega'} = \max_{\omega'' \in B_{\delta}(\omega^{\otimes n})} H_{\min}(X^n | E^n)_{\omega''},$$

where the left hand side corresponds to Eq. (69).

But this means that we can apply the quantum AEP directly, and get

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} C_{\min}^{\delta}(\rho^{\otimes n}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\omega' \in B_{\delta}(\omega^{\otimes n})} H_{\min}(X^n | E^n)_{\omega'} \\ &= H(X|E)_{\omega} = C_r(\rho), \end{aligned}$$

as claimed. \blacksquare

VII. COHERENCE DISTILLATION AND RANDOMNESS GENERATION

Suppose a purification of ρ^A is written as

$$|\psi\rangle^{AE} = \sum_x \sqrt{p_x} |x\rangle^A |\psi_x\rangle^E,$$

we use this state to generate randomness by first performing a computational basis measurement. The dephased cq-state after measuring A in the computational basis is

$$\omega^{AE} = (\Delta \otimes \text{id})\psi = \sum_x p_x |x\rangle\langle x|^A \otimes \psi_x^E.$$

Considering the measurement as a raw randomness generation process, a subsequent randomness extraction (via a deterministic function G) can further extract a random string that

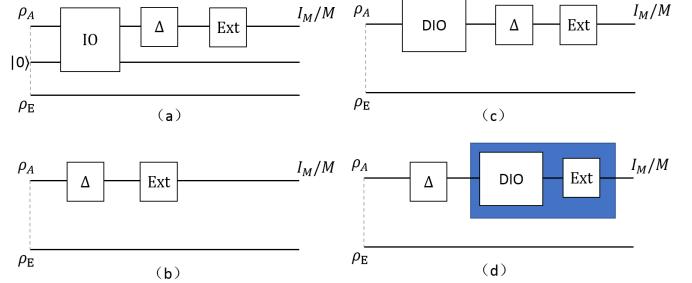


Fig. 1. The different schemes for extracting randomness. (a) Extracting randomness using incoherent strategy in [44]. Alice is allowed to perform the unitary IO operation on her system together with another ancilla system prepared in $|0\rangle\langle 0|$. (b) The randomness extraction process defined in our protocol. The extractor is implemented by a deterministic function G (c) and (d) Applying DIO distillation for extracting secure randomness. Due to the property of DIO, we can achieve a new operation (blue part) combining the DIO distillation operation and original extractor.

is almost uniform and independent of E . We think of the function as an incoherent operation, by letting $G(|x\rangle\langle y|) = \delta_{xy}|G(x)\rangle\langle G(x)|$. This identification is natural as every incoherent (MIO) operation Λ defines an associated classical channel via $\Lambda(|x\rangle\langle x|) = \sum_y \Lambda(y|x)\langle y|$.

Denote $\ell_{\text{ext}}^{\varepsilon}$ to be the maximum length of the extractable randomness that is ε -close to a string that is perfectly uniform and independent of E , i.e.

$$\ell_{\text{ext}}^{\varepsilon}(\rho_A) = \max_G \left\{ \log M : \frac{1}{2} \left\| (G \otimes \text{id})\omega^{AE} - \tau^K \otimes \omega^E \right\|_1 \leq \varepsilon \right\}, \quad (71)$$

where we recall the notation $\tau^K = \frac{1}{M} \mathbb{1}_K$ for the maximally mixed state of the M -dimensional key system.

Note that our definition of extractable randomness differs somewhat from the one in [44]; in that work, a model based on incoherent operations was proposed, which is shown in Fig. 1(a). The main process consists of three parts, incoherent operations Λ , dephasing operation Δ , a random hashing function as the extractor Ext . Here our definition is more straightforward and we do not need to perform the real incoherent operations. As shown in Fig. 1(b), after the dephasing operation Δ , we use a function G as an extractor to extract the secure randomness. This function has to be deterministic, as opposed to a noisy channels, since otherwise infinite randomness can be generated independent of E .

Moreover, in order to obtain the optimal G in our definition, we first consider the randomness extraction process via DIO which is shown in Fig. 1(c), where we apply the DIO distillation followed by dephasing operation Δ , a classical extractor (which may not be needed). Benefiting from the property of DIO, we can change the order of DIO and Δ , which implies that the DIO distillation may act as a good extractor (the blue part in Fig. 1(d)). The only remaining problem here is that we would have to show that this DIO operation gives rise to a deterministic classical channel, which is in general not true. For instance, the optimal coherence distillation process under DIO derived in Theorem 3 has the property that $\Lambda(|x\rangle\langle x|) = \tau_K$ for all x . Via the permutation

twirling Eq. (25), this can be imposed equally on any optimal IO distillation process, and hence even on DIO = IO \cap DIO.

Instead, inspired by Remark 5, we know that a suboptimal but achievable IO distillation operation is also a DIO operation and after a modification (another dephasing channel) we can construct a valid extractor G from it, which is shown in the proof of Theorem 17.

We first recall that every valid G can be used for IO distillation. Here, we consider G as a deterministic function, and $\omega^E = \psi^E$ is the reduced state of ω^{AE} on system E. For every function G satisfying $\frac{1}{2} \| (G \otimes \text{id})\omega^{AE} - \tau^K \otimes \omega^E \|_1 \leq \varepsilon$, we can substitute it in the proof of Theorem 4 and obtain an incoherent distillation channel. Thus as the maximal distillable rate $C_{d,\text{IO}}^\varepsilon(\rho)$, it satisfies

$$C_{d,\text{IO}}^{2\varepsilon}(\rho) \geq \ell_{\text{ext}}^\varepsilon(\rho). \quad (72)$$

The IO map achieving $C_{d,\text{IO}}^{2\varepsilon}(\rho)$ can also be applied to extract randomness by the following theorem.

Theorem 17. *For an arbitrary state ρ and $0 < \varepsilon < 1$,*

$$\ell_{\text{ext}}^{\sqrt{2\varepsilon}}(\rho) \geq C_{d,\text{IO}}^{(\frac{\varepsilon}{2}-\eta)^2/2}(\rho) - 2 \log \frac{1}{\eta}, \quad (73)$$

for any $0 < \eta < \frac{\varepsilon}{2}$.

Proof. Recall that the distillable coherence under IO is given by

$$C_{d,\text{IO}}^\varepsilon(\rho) = \max_{\Lambda \in \text{IO}} \{ \log M : F(\Lambda(\rho), \Psi_M)^2 \geq 1 - \varepsilon \}.$$

Suppose Λ is the IO that achieves the right hand side of Theorem 4, that is $C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2 \log \frac{1}{\eta} = \log M$ where $0 < \eta < \frac{\varepsilon}{2}$ and $F(\Lambda(\rho), \Psi_M)^2 \geq 1 - \varepsilon$. Note that Λ is not necessarily an optimal IO coherence distillation operation. For the purification state $|\psi\rangle^{AE}$, the resulting state by applying Λ on system A is given by

$$\Omega^{AE} = (\Lambda \otimes \text{id})(|\psi\rangle\langle\psi|^{AE}),$$

and it follows that

$$F(\Omega^{AE}, \Psi_M \otimes \Omega^E)^2 \geq (1 - \varepsilon)^2.$$

To prove that, suppose a purification of Ω^{AE} is $|\Omega\rangle^{AA'E}$. Then, considering an orthogonal basis $\{|\phi_x\rangle^A\}$ of system A such that $|\phi_0\rangle^A = |\Psi_M\rangle$, we can write

$$|\Omega\rangle^{AA'E} = \sum_x \alpha_x |\phi_x\rangle^A |\psi_x\rangle^{A'E}.$$

As $F(\Lambda(\rho), \Psi_M)^2 \geq 1 - \varepsilon$ and

$$\Lambda(\rho) = \sum_{xx'} \alpha_x \alpha_x' \langle\psi_{x'}|\psi_x\rangle^{A'E} |\phi_x\rangle\langle\phi_{x'}|^A,$$

we have $F(\Lambda(\rho), \Psi_M)^2 = |\alpha_0|^2 \geq 1 - \varepsilon$. The fidelity between $\Omega^{AA'E}$ and $\Psi_M^A \otimes \psi_0^{A'E}$ is

$$F(\Omega^{AA'E}, \Psi_M^A \otimes \psi_0^{A'E})^2 = |\alpha_0|^2 \geq 1 - \varepsilon.$$

Denoting $\Omega^{A'E} = \text{Tr}_A \Omega^{AA'E}$, then the fidelity between $\Omega^{AA'E}$ and $\Psi_M^A \otimes \Omega^{A'E}$ is

$$F(\Omega^{AA'E}, \Psi_M^A \otimes \Omega^{A'E})^2 = |\alpha_0|^4 \geq (1 - \varepsilon)^2.$$

Then $F(\Omega^{AE}, \Psi_M \otimes \Omega^E)^2 \geq (1 - \varepsilon)^2$ can be obtained by tracing out system A'.

Applying the dephasing operation Δ on system A,

$$F((\Delta \otimes \text{id})\Omega^{AE}, \tau^K \otimes \Omega^E)^2 \geq F(\Omega^{AE}, \Psi_M \otimes \Omega^E)^2.$$

From the Remark 5, we know that Λ is also a DIO which commutes with Δ , so we have equivalently

$$\begin{aligned} F((\Lambda \otimes \text{id})\omega^{AE}, \tau^K \otimes \Omega^E)^2 &\geq F(\Omega^{AE}, \Psi_M \otimes \Omega^E)^2 \\ &\geq (1 - \varepsilon)^2. \end{aligned} \quad (74)$$

In order to construct a deterministic G , we apply another dephasing operation after the incoherent channel Λ ,

$$F((\Delta \circ \Lambda \otimes \text{id})\omega^{AE}, \tau^K \otimes \Omega^E)^2 \geq (1 - \varepsilon)^2, \quad (75)$$

hence

$$\frac{1}{2} \|(\Delta \circ \Lambda \otimes \text{id})\omega^{AE} - \tau^K \otimes \Omega^E\|_1 \leq \sqrt{2\varepsilon}.$$

Note that from Remark 5, the map $\Delta \circ \Lambda$ acts on the incoherent basis states $|x\rangle\langle x|$ as

$$\Delta[\Lambda(|x\rangle\langle x|)] = |G(x)\rangle\langle G(x)|, \quad (76)$$

which is deterministic. From the achievable distillation IO map, we can construct an extractor and obtain

$$\ell_{\text{ext}}^{\sqrt{2\varepsilon}}(\rho) \geq C_{\min}^{\frac{\varepsilon}{2}-\eta}(\rho) - 2 \log \frac{1}{\eta}. \quad (77)$$

Recall the result in Theorem 7,

$$C_{d,\text{IO}}^\varepsilon(\rho) \leq C_{\min}^{\sqrt{\varepsilon}(2-\varepsilon)}(\rho) \leq C_{\min}^{\sqrt{2\varepsilon}}(\rho), \quad (78)$$

and we obtain

$$\ell_{\text{ext}}^{\sqrt{2\varepsilon}}(\rho) \geq C_{d,\text{IO}}^{(\frac{\varepsilon}{2}-\eta)^2/2}(\rho) - 2 \log \frac{1}{\eta}, \quad (79)$$

finishing the proof. \blacksquare

Combining with Eq. (72), we have

$$C_{d,\text{IO}}^{\sqrt{8\varepsilon}}(\rho) \geq \ell_{\text{ext}}^{\sqrt{2\varepsilon}}(\rho) \geq C_{d,\text{IO}}^{(\frac{\varepsilon}{2}-\eta)^2/2}(\rho) - 2 \log \frac{1}{\eta}. \quad (80)$$

In the regime of vanishingly small ε , the distillable coherence rate $C_{d,\text{IO}}(\rho)$ and $\ell_{\text{ext}}(\rho)$ are hence essentially the same. Whether $C_{d,\text{DIO}}(\rho)$ and $\ell_{\text{ext}}(\rho)$ are the same is still an open problem. Though DIO can commute with dephasing operation, the difficulty stems from that the combination of DIO and extractor (the blue part in Fig. 1(d)) may be not deterministic thus not a valid extractor.

VIII. DISCUSSION

We have considered the problem of one-shot coherence distillation under the classes MIO, DIO, IO, and SIO of incoherent operations. Our results indicate that the distillation rates under IO, MIO, DIO are roughly the same, up to different smoothing parameters and universal additive terms. The results allow us to recover the asymptotic (many-copy) limit, in which the distillation rates for all these three classes tend to be

relative entropy of coherence, which is consistent with the previous results in [17], [25], [26].

The smallest class for which we have been able to show a non-trivial distillation of coherence is that of dephasing-covariant incoherent IOs, $\text{DIO} = \text{IO} \cap \text{DIO}$. On the other hand, interestingly, there is a gap between distillation rates under SIO and DIO, both in the one-shot, and more importantly in the asymptotic regime. As a matter of fact, we showed that there is *bound coherence* under SIO; no pure coherence can be distilled under SIO from these states though they possess coherence, as shown by distillable coherence under IO.

An interesting future direction is then to study the case involving another system to help this distillation process, which referred as assisted coherence distillation for such bound coherence with SIO [45]–[48]. Furthermore, our work also connects the distillation of coherence to randomness extraction. The distillation process is also related to the decoupling in cryptography. Thus our results also shed light on other quantum information processing tasks like random number generation, extraction and cryptography.

Note added: After completion of this work and circulating a preliminary preprint, Ludovico Lami *et al.* [49] have shown that the bound coherence under SIO is in fact a generic phenomenon, showing that the fidelity of distilling even a single cosbit is bounded away from 1 for all but a measure-zero set of mixed states.

ACKNOWLEDGMENT

The authors thank María García Díaz, Xiongfeng Ma, Hongyi Zhou, Ludovico Lami and Bartosz Regula for various insightful discussions on coherence theory.

REFERENCES

- [1] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, “Numerical approach for unstructured quantum key distribution,” *Nature Commun.*, vol. 7, May 2016, Art. no. 11712.
- [2] V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology,” *Nature Photon.*, vol. 5, no. 4, pp. 222–229, Mar. 2011.
- [3] X. Yuan, H. Zhou, Z. Cao, and X. Ma, “Intrinsic randomness as a measure of quantum coherence,” *Phys. Rev. A, Gen. Phys.*, vol. 92, Aug. 2015, Art. no. 022124. doi: [10.1103/PhysRevA.92.022124](https://doi.org/10.1103/PhysRevA.92.022124).
- [4] J. Ma, X. Yuan, A. Hakande, and X. Ma, (2017). “Source-independent quantum random number generation via measuring coherence.” [Online]. Available: <https://arxiv.org/abs/1704.06915>
- [5] M. B. Plenio and S. F. Huelga, “Dephasing-assisted transport: Quantum networks and biomolecules,” *New J. Phys.*, vol. 10, no. 11, Nov. 2008, Art. no. 113019.
- [6] P. Rebentrost, M. Mohseni, and A. Aspuru-Guzik, “Role of quantum coherence and environmental fluctuations in chromophoric energy transport,” *J. Phys. Chem. B*, vol. 113, no. 29, pp. 9942–9947, Jun. 2009.
- [7] J. Åberg, “Catalytic Coherence,” *Phys. Rev. Lett.*, vol. 113, Oct. 2014, Art. no. 150402. doi: [10.1103/PhysRevLett.113.150402](https://doi.org/10.1103/PhysRevLett.113.150402).
- [8] P. Ćwikliński, M. Studziński, M. Horodecki, and J. Oppenheim, “Limitations on the evolution of quantum coherences: Towards fully quantum second laws of thermodynamics,” *Phys. Rev. Lett.*, vol. 115, Nov. 2015, Art. no. 210403. doi: [10.1103/PhysRevLett.115.210403](https://doi.org/10.1103/PhysRevLett.115.210403).
- [9] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, “Quantum coherence, time-translation symmetry, and thermodynamics,” *Phys. Rev. X*, vol. 5, Apr. 2015, Art. no. 021001. doi: [10.1103/PhysRevX.5.021001](https://doi.org/10.1103/PhysRevX.5.021001).
- [10] M. Lostaglio, D. Jennings, and T. Rudolph, “Description of quantum coherence in thermodynamic processes requires constraints beyond free energy,” *Nature Commun.*, vol. 6, Mar. 2015, Art. no. 6383.
- [11] V. Narasimhachar and G. Gour, “Low-temperature thermodynamics with quantum coherence,” *Nature Commun.*, vol. 6, Jul. 2015, Art. no. 7689.
- [12] T. Baumgratz, M. Cramer, and M. B. Plenio, “Quantifying coherence,” *Phys. Rev. Lett.*, vol. 113, Sep. 2014, Art. no. 140401.
- [13] J. Åberg, (2006). “Quantifying superposition.” [Online]. Available: <https://arxiv.org/abs/quant-ph/0612146>
- [14] D. Braun and B. Georgeot, “Quantitative measure of interference,” *Phys. Rev. A, Gen. Phys.*, vol. 73, Feb. 2006, Art. no. 022314.
- [15] E. Chitambar and G. Gour, “Critical examination of incoherent operations and a physically consistent resource theory of quantum coherence,” *Phys. Rev. Lett.*, vol. 117, Jul. 2016, Art. no. 030401.
- [16] I. Marvian and R. W. Spekkens, “How to quantify coherence: Distinguishing speakable and unspeakable notions,” *Phys. Rev. A, Gen. Phys.*, vol. 94, Nov. 2016, Art. no. 052324. doi: [10.1103/PhysRevA.94.052324](https://doi.org/10.1103/PhysRevA.94.052324).
- [17] A. Winter and D. Yang, “Operational resource theory of coherence,” *Phys. Rev. Lett.*, vol. 116, Mar. 2016, Art. no. 120404. doi: [10.1103/PhysRevLett.116.120404](https://doi.org/10.1103/PhysRevLett.116.120404).
- [18] B. Yadin, J. Ma, D. Girolami, M. Gu, and V. Vedral, “Quantum processes which do not use coherence,” *Phys. Rev. X*, vol. 6, Nov. 2016, Art. no. 041028. doi: [10.1103/PhysRevX.6.041028](https://doi.org/10.1103/PhysRevX.6.041028).
- [19] T. Biswas, M. G. Díaz, and A. Winter, “Interferometric visibility and coherence,” *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 473, no. 2203, Art. no. 20170170, Jul. 2017.
- [20] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, “Robustness of coherence: An operational and observable measure of quantum coherence,” *Phys. Rev. Lett.*, vol. 116, Apr. 2016, Art. no. 150502. doi: [10.1103/PhysRevLett.116.150502](https://doi.org/10.1103/PhysRevLett.116.150502).
- [21] S. Rana, P. Parashar, A. Winter, and M. Lewenstein, “Logarithmic coherence: Operational interpretation of l_1 -norm coherence,” *Phys. Rev. A, Gen. Phys.*, vol. 96, Nov. 2017, Art. no. 052336. doi: [10.1103/PhysRevA.96.052336](https://doi.org/10.1103/PhysRevA.96.052336).
- [22] A. Streltsov, G. Adesso, and M. B. Plenio, “Colloquium: Quantum coherence as a resource,” *Rev. Mod. Phys.*, vol. 89, Oct. 2017, Art. no. 041003. [Online]. Available: <https://arxiv.org/abs/1609.02439>
- [23] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, “Using and reusing coherence to realize quantum processes,” *Quantum*, vol. 2, p. 100, Oct. 2018. [Online]. Available: <https://arxiv.org/abs/1805.04045>
- [24] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and X. Ma, “One-shot coherence dilution,” *Phys. Rev. Lett.*, vol. 120, Feb. , Art. no. 070403. <https://arxiv.org/abs/1707.02522>
- [25] B. Regula, K. Fang, X. Wang, and G. Adesso, “One-shot coherence distillation,” *Phys. Rev. Lett.*, vol. 121, Jul. 2018, Art. no. 010401. doi: [10.1103/PhysRevLett.121.010401](https://doi.org/10.1103/PhysRevLett.121.010401).
- [26] E. Chitambar, “Dephasing-covariant operations enable asymptotic reversibility of quantum resources,” *Phys. Rev. A, Gen. Phys.*, vol. 97, May 2018, Art. no. 050301.
- [27] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Rényi entropies: A new generalization and some properties,” *J. Math. Phys.*, vol. 54, no. 12, Nov. 2013, Art. no. 122203. [Online]. Available: <https://doi.org/10.1063/1.4838856>.
- [28] M. M. Wilde, A. Winter, and D. Yang, “Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy,” *Commun. Math. Phys.*, vol. 331, no. 2, pp. 593–622, Oct. 2014.
- [29] E. Chitambar and G. Gour, “Comparison of incoherent operations and measures of coherence,” *Phys. Rev. A, Gen. Phys.*, vol. 94, Nov. 2016, Art. no. 052336.
- [30] C. A. Fuchs and J. V. D. Graaf, “Cryptographic distinguishability measures for quantum-mechanical states,” *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1216–1227, May 1999.
- [31] Y. Liu, Q. Zhao, and X. Yuan, “Quantum coherence via conditional entropy,” *J. Phys. A, Math. Theor.*, vol. 51, no. 41, 2018, Art. no. 414018. [Online]. Available: <http://stacks.iop.org/1751-8121/51/i=41/a=414018>
- [32] F. Buscemi and N. Datta, “The quantum capacity of channels with arbitrarily correlated noise,” *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1447–1460, Mar. 2010.
- [33] L. Wang and R. Renner, “One-shot classical-quantum capacity and hypothesis testing,” *Phys. Rev. Lett.*, vol. 108, May 2012, Art. no. 200501. doi: [10.1103/PhysRevLett.108.200501](https://doi.org/10.1103/PhysRevLett.108.200501).
- [34] F. Dupuis, L. Krämer, P. Faist, J. M. Renes, and R. Renner, “Generalized Entropies,” in *Proc. 17th Int. Congr. Math. Phys.*, Aalborg, Denmark, Aug. 2012, pp. 134–153, [Online]. Available: <https://arxiv.org/abs/1211.3141>
- [35] M. Sion, “On general minimax theorem,” *Pacific J. Math.*, vol. 8, pp. 171–176, Jun. 1957.

- [36] R. König, R. Renner, and C. Schaffner, “The operational meaning of min- and max-entropy,” *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4337–4347, Sep. 2009.
- [37] R. Renner, “Security of quantum key distribution,” Ph.D. dissertation, ETH Zurich, Zürich, Switzerland, 2005.
- [38] A. Uhlmann, “The ‘transition probability’ in the state space of a $*$ -algebra,” *Rep. Math. Phys.*, vol. 9, no. 2, pp. 273–279, Apr. 1976.
- [39] P. J. Coles, “Unification of different views of decoherence and discord,” *Phys. Rev. A, Gen. Phys.*, vol. 85, Apr. 2012, Art. no. 042103, [Online]. Available: <https://arxiv.org/abs/1110.1664>. doi: [10.1103/PhysRevA.85.042103](https://doi.org/10.1103/PhysRevA.85.042103).
- [40] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, “The uncertainty principle in the presence of quantum memory,” *Nature Phys.*, vol. 6, pp. 659–662, Jul. 2010.
- [41] M. Ringbauer *et al.* (2017). “Certification and quantification of multilevel quantum coherence.” [Online]. Available: <https://arxiv.org/abs/1707.05282>
- [42] I. Marvian. (2018). “Coherence distillation machines are impossible in quantum thermodynamics.” [Online]. Available: <https://arxiv.org/abs/1805.01989>
- [43] M. Tomamichel, R. Colbeck, and R. Renner, “A fully quantum asymptotic equipartition property,” *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5840–5847, Dec. 2009.
- [44] M. Hayashi and H. Zhu, “Secure uniform random-number extraction via incoherent strategies,” *Phys. Rev. A, Gen. Phys.*, vol. 97, Jan. 2018, Art. no. 012302. doi: [10.1103/PhysRevA.97.012302](https://doi.org/10.1103/PhysRevA.97.012302).
- [45] M. K. Vijayan, E. Chitambar, and M.-H. Hsieh. (2018). “One-shot assisted concentration of coherence.” [Online]. Available: <https://arxiv.org/abs/1804.06554>
- [46] B. Regula, L. Lami, and A. Streltsov. (2018). “Non-asymptotic assisted distillation of quantum coherence.” [Online]. Available: <https://arxiv.org/abs/1807.04705>
- [47] E. Chitambar, A. Streltsov, S. Rana, M. N. Bera, G. Adesso, and M. Lewenstein, “Assisted distillation of quantum coherence,” *Phys. Rev. Lett.*, vol. 116, Feb. 2016, Art. no. 070402. doi: [10.1103/PhysRevLett.116.070402](https://doi.org/10.1103/PhysRevLett.116.070402)
- [48] A. Streltsov, S. Rana, M. N. Bera, and M. Lewenstein, “Towards resource theory of coherence in distributed scenarios,” *Phys. Rev. X*, vol. 7, Mar. 2017, Art. no. 011024. doi: [10.1103/PhysRevX.7.011024](https://doi.org/10.1103/PhysRevX.7.011024).
- [49] L. Lami, B. Regula, and G. Adesso. (2018). “Generic bound coherence under strictly incoherent operations.” [Online]. Available: <https://arxiv.org/abs/1809.06880>

Qi Zhao received his B.S. degree from the Department of Mathematics, Tsinghua University in 2014 and his Ph.D. degree in physics from Tsinghua University in 2018 under the supervision of Professor Xiongfeng Ma. His research interests include quantum resource theory, self-testing quantum information and entanglement detection.

Yunchao Liu is an undergraduate student at Institute for Interdisciplinary Information Sciences, Tsinghua University. His research interests include quantum information theory, quantum computing and machine learning.

Xiao Yuan received his B.S. degree in physics from Peking University in 2012 and his Ph.D. degree in physics from Tsinghua University in 2016 under the supervision of Professor Xiongfeng Ma. He is currently a postdoctoral researcher at the Department of Materials, University of Oxford. His main research interests are in quantum resource theory, quantum cryptography, and quantum simulation.

Eric Chitambar received his B.S. degree in physics from the University of Notre Dame in 2005 and his Ph.D. degree in physics from the University of Michigan in 2010 under the direction of Prof. Yaoyun Shi. He served as a post-doctoral researcher at the University of Toronto and later at the Perimeter Institute for Theoretical Physics in Waterloo, Canada. In 2012, Dr. Chitambar became a member of the physics department at Southern Illinois University Carbondale, and in 2018 he joined the department of Electrical and Computer Engineering at the University of Illinois Urbana-Champaign. His research interests include the LOCC processing of quantum information, entanglement theory, and general quantum resource theories.

Andreas Winter received a Diploma degree in Mathematics from the Freie Universität Berlin, Berlin, Germany, in 1997, and a Ph.D. degree from the Fakultät für Mathematik, Universität Bielefeld, Bielefeld, Germany, in 1999. He was Research Associate at the University of Bielefeld until 2001, and then with the Department of Computer Science at the University of Bristol, Bristol, UK. In 2003, still with the University of Bristol, he was appointed Lecturer in Mathematics, and in 2006 Professor of Physics of Information. Since 2012 he has been ICREA Research Professor with the Universitat Autònoma de Barcelona, Barcelona, Spain. His research interests include quantum and classical Shannon theory, and discrete mathematics. He is recipient, along with Bennett, Devetak, Harrow and Shor, of the 2017 Information Theory Society Paper Award.