Generic bound coherence under strictly incoherent operations

Ludovico Lami, Bartosz Regula, and Gerardo Adesso

1School of Mathematical Sciences and Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

We fully characterize the distillability of quantum coherence under strictly incoherent operations (SIO), providing a complete description of the phenomenon of bound coherence. In particular, we establish a simple, analytically computable necessary and sufficient criterion for the asymptotic distillability under SIO. We use this result to show that almost every quantum state is undistillable — only pure states as well as states whose density matrix contains a rank-one submatrix allow for coherence distillation under SIO, while every other quantum state exhibits bound coherence. This demonstrates fundamental operational limitations of SIO in the resource theory of quantum coherence. We further show that the fidelity of distillation of a single bit of coherence can be efficiently computed as a semidefinite program, and investigate the generalization of this result to provide an understanding of asymptotically achievable distillation fidelity.

Introduction.—The resource theory of quantum coherence [1–4] has found extensive use in the characterization of a signature intrinsic feature of quantum mechanics — superposition — and our ability to manipulate it efficiently within a resource-theoretic framework [5–8]. Typically, the properties of a resource are investigated under a suitable set of allowed free operations, reflecting the constraints placed on the manipulation of the given resource [5–8]. In spite of the fact that the resource theory of coherence has found use in a variety of practical settings [4], no physically compelling set of assumptions has yet emerged which could single out a unique class of free operations under which the operational features of coherence should be investigated, mirroring the fundamental role of local operations and classical communication in the resource theory of entanglement [9]. This has motivated the definition and characterization of a multitude of possible sets of free operations [10–26]. However, many definitions of free operations stemming from meaningful considerations, such as the physically incoherent operations [13], the translationally-covariant incoherent operations [14], or the genuinely incoherent operations [16], were found to be too limited in their operational capabilities, suggesting that any useful resource theory of coherence would require a larger set of maps. On the other hand, strictly larger sets of maps such as the maximally incoherent operations (MIO) [11], the incoherent operations (IO) [2], or the dephasing-covariant incoherent operations (DIO) [13,14], while operationally powerful, might be considered as too permissive and lacking a physically implementable form.

The class of strictly incoherent operations (SIO) [3,12] appeared to be a suitable candidate for the smallest class of operations satisfying desirable resource-theoretic criteria, and has found widespread use in the resource theory of coherence [3,12,15,18,27]. SIO is a physically-motivated, easy to characterize, and seemingly powerful choice of free operations, allowing in particular for a “golden unit” of coherence represented by the maximally coherent state $|\Psi_m\rangle$, which can be transformed into any other state using SIO [2]. Although strictly smaller than the sets IO and DIO, the operational capabilities of SIO did not appear to be too limited — for instance, SIO have exactly the same power as IO as far as pure-to-pure state transformations are concerned [15,28], as well as in the context of coherence dilution [3,19]: they match the power of DIO in probabilistic distillation from pure states [23]; and even the largest class of free operations, MIO, cannot perform better than SIO in one-shot distillation from pure states [20] or in assisted coherence distillation [25]. On the other hand, there do exist tasks in which the limitations of SIO become apparent — in particular, unlike the larger sets IO, DIO, and MIO, the class SIO has recently been found to exhibit bound coherence [20] (see Fig. 1), i.e., the existence of coherent states from which no coherence can be distilled by such operations. It is, however, not known how common this property is among all quantum states, nor to what extent it limits the operational capabilities of SIO beyond specific examples.

In this work, we fully characterize distillability of arbitrary quantum states under SIO. By introducing a SIO coherence monotone which does not change when multiple copies of a state are considered, we establish a simple analytical criterion to decide whether a given quantum system is asymptotically distillable or not. In the former case, our method further yields a nontrivial lower bound on the distillable coherence under SIO. Instrumental in our approach is an efficiently computable semidefinite programming (SDP) expression characterizing the maximal achievable fidelity in the distillation of a single bit of coherence starting from a state of arbitrary dimension.

FIG. 1. To create $n$ copies of a coherent state $\rho$ (puzzle pieces) by strictly incoherent operations (SIO), one needs to consume a certain number of copies of the maximally coherent state $|\Psi_2\rangle$ (golden unit), in a process known as dilution (leftmost arrow). The state $\rho$ exhibits bound coherence when no maximally coherent state can be recovered back from it via a distillation process (rightmost arrow), even for arbitrarily large $n$. Here we show that this phenomenon is generic under SIO for mixed quantum states of arbitrary dimension.
where the density matrix contains a submatrix proportional to a pure state. This demonstrates fundamental restrictions on the power of SIO in general state transformations, and shows in a precise operational sense that SIO is a rather limited class of free maps in the resource theory of coherence.

A new SIO monotone. — Let us begin by recalling the basic formalism of the resource theory of quantum coherence. The set of free states, known as the incoherent states $I$, consists of all density matrices diagonal in a given $d$-dimensional orthonormal basis $\{|i\rangle\}$. We will denote by $\Delta$ the dephasing map, defined by $\Delta(\cdot) = \sum_i |i\rangle \langle i| \otimes |i\rangle \langle i|$, and by $|\Psi_m\rangle = \frac{1}{\sqrt{m}} \sum_i |i\rangle$ the maximally coherent state of dimension $m$. As for the choice of free operations, we will especially focus on the so-called strictly incoherent operations (SIO), defined as those functions $\Lambda$ which admit a Kraus decomposition $\Lambda(\cdot) = \sum_\alpha K_\alpha(\cdot)K_\alpha^\dagger$ such that $K_\alpha \Delta(\rho) K_\alpha^\dagger = \Delta(\rho) K_\alpha K_\alpha^\dagger$ for all $\alpha$ and $\rho$.

We use $A_{ij}$ to refer to the entries of a matrix in the basis $\{|i\rangle\}$. We will write $A \geq B$ if $A - B$ is positive semidefinite, and $A \geq B$ if $A_{ij} \geq B_{ij}$ for all $1 \leq i,j \leq d$. Throughout the manuscript, when applied to a matrix the modulus is to be intended entrywise, i.e., $|A|$ will be the matrix defined as $|A|_{ij} := |A_{ij}|$.

We now introduce a straightforwardly computable quantity that we name maximal coherence of $\rho$, defined by

$$\eta(\rho) := \max_{i \neq j} \frac{|\rho_{ij}|}{\sqrt{\rho_{ii}\rho_{jj}}},$$

(1)

where the optimization is over all choices of indices such that $\rho_{ii} \neq 0 \neq \rho_{jj}$, with $\eta(\rho) = 0$ if no such choice exists. Alternatively, this quantity can be understood as the largest modulus of an off-diagonal element of the matrix $\Delta(\rho)^{-1/2} \rho \Delta(\rho)^{-1/2}$.

We first notice that for all states $\rho$, one has $0 \leq \eta(\rho) \leq 1$. This follows from the positivity of the principal minor of $\rho$ of order 2 corresponding to the rows and columns identified by indices $i$ and $j$, which implies that $|\rho_{ij}|^2 / \rho_{ii}\rho_{jj} \leq 1$ for any choice of $i,j$. Moreover, we see by definition that $\eta(\rho) = 0$ iff $\rho$ is incoherent, and $\eta(\rho) = 1$ iff there are indices $1 \leq i \neq j \leq d$ such that $\Pi_{ij}\rho\Pi_{ij}$ is proportional to a pure state, where $\Pi_{ij}$ is the projector onto span$\{|i\rangle, |j\rangle\}$.

An important property of $\eta$ is that it is, in fact, monotonically non-increasing under SIO. Precisely, consider an SIO operation $\Lambda$ acting on a $d$-dimensional system, which can be written as

$$\Lambda(\cdot) = \sum_\alpha U^\dagger_{\alpha} D_{\alpha}(\cdot) D^\dagger_{\alpha} U_{\alpha},$$

(2)

where the $\pi_\alpha$ are permutations, $U_{\alpha} := \sum_{i=1}^d |\pi_\alpha(i)\rangle \langle i|$ are the unitaries that implement them, and the matrices $D_{\alpha} := \sum_{i=1}^d d_{\alpha}(i) |i\rangle \langle i|$ are all diagonal. This representation has some technical issues when input and output dimensions are different, but this is irrelevant for the present argument, as we argue in the Supplemental Material [29]. In our case, for two arbitrary indices $1 \leq i \neq j \leq d$, we can write

$$|\Lambda(\rho)_{ij}| \leq \sum_\alpha |\rho_{ij}| \left| \sum_i U^\dagger_{\alpha} d_{\alpha}(i) d_{\alpha}(j) \rho_{\pi_\alpha(i),\pi_\alpha(j)} U_{\alpha} \right| \leq \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) |d_{\alpha}(j)\rangle \langle \rho_{\pi_\alpha(i),\pi_\alpha(j)} | d_{\alpha}(j) \right| \leq \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) |d_{\alpha}(j)\rangle \langle \rho_{\pi_\alpha(i),\pi_\alpha(j)} | d_{\alpha}(j) \right| \leq \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) \rho_{\pi_\alpha(i),\pi_\alpha(j)} \right| \leq \eta(\rho) \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) \right| \rho_{\pi_\alpha(i),\pi_\alpha(j)} \leq \eta(\rho) \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) \right|^2 \rho_{\pi_\alpha(i),\pi_\alpha(j)} \leq \eta(\rho) \sum_\alpha |\rho_{ij}| \left| \sum_i d_{\alpha}(i) \right|^2 \rho_{\pi_\alpha(i),\pi_\alpha(j)} \leq \eta(\rho) \sqrt{\Lambda(\rho)_{ii} \Lambda(\rho)_{jj}}.$$
vanishing error; precisely, we have

$$C_{d,O}(\rho) := \sup \left\{ r \mid \lim_{n \to \infty} F_O(\rho^{\otimes n}, 2^{r n}) = 1 \right\}. \quad (4)$$

We will say that a state \( \rho \) is distillable under \( O \) if \( C_{d,O}(\rho) > 0 \).

We now make a crucial observation which lets us immediately relate the maximal coherence \( \eta \) to the problem of coherence distillation under SIO. It is the fact that \( \eta \) obeys the so-called tensorization property \([37]\), that is,

$$\eta(\rho \otimes \sigma) = \max \{ \eta(\rho), \eta(\sigma) \} \quad \forall \rho, \sigma. \quad (5)$$

To prove the above identity, observe that, according to Eq. (1), computing \( \eta(\rho \otimes \sigma) \) corresponds to maximizing the function \([\rho \otimes \sigma]_{ik,jl} \left[ \sqrt{\rho_{ii} \rho_{jj}} \right] \) over all pairs of indices \((i,k) \neq (j,l)\), and that this is equivalent to maximizing \([\rho_{ij} / \sqrt{\rho_{ii} \rho_{jj}}] \) over choices such that \( i \neq j \) or \( k \neq l \). The latter maximum is achieved on pairs either of the form \((i,k,ik)\) with \( i \neq j \) or of the form \((i,k,l)\) with \( k \neq l \), which corresponds precisely to the larger of \( \eta(\rho) \) and \( \eta(\sigma) \).

The tensorization and the monotonicity of \( \eta \), we readily obtain our main result: a necessary and sufficient criterion for the distillability of an arbitrary quantum state under SIO.

**Theorem 1.** For all density matrices \( \rho \) of any dimension, \( C_{d,SIO}(\rho) > 0 \) if and only if \( \eta(\rho) = 1 \).

**Proof.** Noting that \( \eta(\psi) = 1 \) and remembering that \( \eta \) is l.s.c., we see that for distillation to be possible, there needs to exist a sequence of SIO operations \( \Lambda_n \) such that \( \eta(\lim_{n \to \infty} \Lambda_n(\rho^{\otimes n})) = 1 \). However, it holds that

$$\eta(\lim_{n \to \infty} \Lambda_n(\rho^{\otimes n})) \leq \liminf_{n \to \infty} \eta(\Lambda_n(\rho^{\otimes n})) \leq \lim_{n \to \infty} \eta(\rho^{\otimes n}) = \eta(\rho),$$

where the first inequality is due to the l.s.c. of \( \eta \), the second inequality due to its monotonicity under SIO, and the final equality follows from the tensorization property. This shows in particular that any state with \( \eta(\rho) < 1 \) is undistillable. Conversely, if \( \eta(\rho) = 1 \), then we know that there exists a two-dimensional orthogonal projection \( \Pi_{ij} \), such that \( \Pi_{ij} \rho \Pi_{ij} =: p\psi \), where \( \psi \) is a pure coherent state and \( p > 0 \) a normalization coefficient. By performing the measurement defined by the positive operator-valued measure \( \{\Pi_{ij}, \mathbb{1} - \Pi_{ij}\} \) (straightforwardly verified to be SIO) on each of the \( n \) copies of \( \rho \), we obtain on average \( np \) copies of the state \( \psi \). Using the fact that all pure coherent states are distillable under SIO with a rate given by the entropy of their diagonal part \([3]\), we have that the rate of distillation associated with \( \rho \) is lower bounded by \( p S(\Delta(\psi)) > 0 \), and so we conclude that \( \rho \) is asymptotically distillable.

The above Theorem 1 establishes a complete characterization of distillability under SIO. In particular, it is not difficult to see that any generic quantum state exhibits bound coherence, as illustrated in Fig. 1 and that the condition for distillability of a state \( \rho \) — i.e., the existence of a submatrix of \( \rho \) in the basis \( \{|i\}\), proportional to a pure state — is an extremely restrictive property, satisfied only by a zero-measure class of mixed states. We further remark that the proof of the Theorem in fact establishes the stronger statement that any state \( \rho \) such that \( \eta(\rho) < 1 \) cannot be used to distill even a single coherence bit, no matter how large the number of available copies of \( \rho \) is.

**Fidelity of distillation.** — It follows from [28, Thm. 9] that the maximal fidelity of distillation of an \( m \)-dimensional maximally coherent state \( \Psi_m \) under SIO for any state \( \rho \) is

$$F_{SIO}(\rho, m) = \max_{\Lambda \in SIO} F(\Lambda(\rho), \Psi_m) \quad (6)$$

$$= \max \left\{ \text{Tr} \rho W \mid 0 \leq W \leq 1, \quad \Lambda(W) = \frac{1}{m}, \quad \text{CN}(W) \leq m \right\}$$

where \( \text{CN} \) denotes the coherence number \([40, 41]\).

For the case of distilling a coherence bit \( \Psi_2 \), we are able to obtain the following simplified characterization.

**Theorem 2.** The fidelity of distillation of a single bit of coherence under SIO is given by the SDP

$$F_{SIO}(\rho, 2) = \max_{-1 \leq X \leq 1, \Delta X = 0} \frac{1}{2} \left( \text{Tr} |\rho| X + 1 \right) \quad (7a)$$

$$= \min_{D \geq \Delta(D)} \frac{1}{2} \left( \| |\rho| + D + N \|_1 + 1 \right). \quad (7b)$$

**Proof.** We sketch the main idea of the argument, deferring the quite laborious details to the Supplemental Material \([29]\).

Our starting point is Eq. (6), which expresses the distillation fidelity as an optimization that involves a nontrivial constraint on the coherence number of the variable \( W \). For \( m = 2 \), this can be cast into an analytically manageable form thanks to Theorem 1 of [42], which states that a positive semidefinite matrix \( W \) satisfies \( \text{CN}(W) \leq 2 \iff \Delta(\Delta(W) - |W|) \geq 0 \). By leveraging this criterion and choosing carefully the optimization variables, one can turn Eq. (6) for \( m = 2 \) into Eq. (7a). Finally, Eq. (7b) is obtained by taking the SDP dual.

The results of Theorem 2 should be compared with the expressions for the fidelity of distillation associated with the larger sets DIO and MIO, given by \([20]\).

$$F_{MIO}(\rho, 2) = F_{DIO}(\rho, 2) = \frac{1}{2} \left( \min_{D \geq \Delta(D)} \| |\rho| + D \|_1 + 1 \right). \quad (8)$$

We recall in particular that \( F_{MIO}(\psi, m) = F_{SIO}(\psi, m) \) for all pure states \( \psi \) and all \( m \) \([20]\).

It is left to determine how closely one can approximate distillation of a perfect bit of coherence by means of SIO when one is given a large number of copies of an input state. This leads us to investigate the quantity \( F_{SIO}(\rho^{\otimes n}, 2) \) as a function of \( \rho \) and \( n \), and in particular its asymptotic properties in the limit of large \( n \). We formalize this in the following result, whose full proof we provide in the Supplemental Material \([29]\).
Theorem 3. For all states $\rho$ and all integers $n$, one has

$$\frac{1 + \eta(\rho)}{2} - \frac{\eta(\rho)^n}{2} \leq F_{\text{SIO}}(\rho^{\otimes n}, 2) \leq \frac{1 + \eta(\rho)}{2},$$

(9)

where $0 < \mu_\rho < 1$ is a number that depends only on $\rho$. Hence

$$\lim_{n \to \infty} F_{\text{SIO}}(\rho^{\otimes n}, 2) = \frac{1 + \eta(\rho)}{2},$$

(10)

and the convergence in the above identity is exponentially fast.

As a particularly strong example of undistillability under SIO, consider the class of qubit states $\rho_\lambda = \lambda |\psi\rangle\langle\psi| + (1 - \lambda) \frac{I}{2}$ with $\lambda \in [0,1]$. An explicit computation yields $\eta(\rho_\lambda) = \lambda$. By constructing a suitable choice of feasible solutions for the SDP (7) [29], it can be shown that the fidelity of distillation satisfies $F_{\text{SIO}}(\rho_\lambda^{\otimes n}, 2) = (1 + \lambda)/2$ for any number of copies $n$. Therefore, not only is the distillation of $\rho_\lambda$ impossible under SIO for $\lambda \neq 1$, it actually is impossible to increase the fidelity of distillation whatsoever by adding more copies of the state.

**Distillable coherence under SIO.**—Although we have proven that most states are bound coherent under SIO, it could be nevertheless interesting to compute the amount of coherence $C_{d,SIO}$ that can be extracted from the distillable states. This is a very different scenario from that considered in Theorem 3, while there we were interested in the distillation of a single coherence bit with good fidelity, here we look at the maximal rate of distillation of bits of coherence with vanishing errors.

As usual, the process of evaluating a maximal distillation rate is composed of two parts. First, one designs a protocol that achieves the conjectured rate in the limit of a large number of copies (direct part). Second, one shows that the performance of this protocol can not be beaten at least asymptotically (converse part). This is usually obtained by employing a monotone that is sensitive to how many coherence bits one has distilled, namely, an additive monotone. Unfortunately, we seem to lack easily computable, nontrivial, additive SIO monotones that are not at the same time also IO monotones, and it is for this very reason that we are not yet able to compute $C_{d,SIO}$ exactly on all states.

Nevertheless, we now present a direct construction, motivated by the properties of the monotone $\eta(\rho)$, that leads to a nontrivial lower bound to $C_{d,SIO}$. For a state $\rho$ such that $\Delta(\rho) > 0$, construct the set $E_\rho := \{(i,j) : |\rho_{ij}| = \sqrt{\rho_{ii}\rho_{jj}}\}$. As we show in the Supplemental Material [29], it turns out that there is a partition $\{\nu_i\}$, of $\{1, \ldots, d\}$ such that $(i,j) \in E_\rho$ iff $i, j$ belong to the same set $\nu_i$. Using this observation, it is possible to show that the operator $\tilde{\rho} := \sum_{(i,j) \in E_\rho} \rho_{ij} |i\rangle\langle j|$ is a legitimate density matrix, and that the quantifier

$$Q(\rho) := S(\Delta(\rho)) - S(\tilde{\rho})$$

(11)

is: (i) nonnegative; (ii) strictly positive iff $\eta(\rho) < 1$; and (iii) additive over tensor products. Our last result shows that $Q$ in fact lower bounds the SIO distillable coherence.

**Proposition 4.** For all states $\rho$ in any dimension, the SIO distillable coherence satisfies $C_{d,SIO}(\rho) \geq Q(\rho)$.

**Proof.** Given $n$ copies of the state $\rho$, we perform independently on each of them the measurement $\{\Pi_{\nu_i}\}$, where $\Pi_{\nu_i} := \sum_{\{i \in \nu_i\}} |i\rangle\langle i|$ and $\{\nu_i\}$ is the partition of $\{1, \ldots, d\}$ identified above. Setting $P(\nu) := Tr[\rho \Pi_{\nu}]$ and $\tilde{\rho}_{\nu} := P(\nu)^{-1} \Pi_{\nu} \rho \Pi_{\nu}$, we see that this protocol produces an average of $nP(\nu)$ copies of the states $\tilde{\rho}_{\nu}$, which can be shown to be all pure. It is known [11] that there exists an SIO protocol that extracts $S(\Delta(\rho))$ coherence bits per copy out of any pure state $\psi$. Applying this procedure to each $\tilde{\rho}_{\nu}$ leads to an expected number of coherence bits produced equal to $\sum_{\nu} n P(\nu) S(\Delta(\tilde{\rho}_{\nu})) = nQ(\rho)$, achieving a rate $Q(\rho)$. See [29] for further technical details.

We conjecture that the bound in Proposition 4 is in fact tight.

**Conjecture 5.** For all states $\rho$, the distillable coherence under SIO is given by $C_{d,SIO}(\rho) = Q(\rho)$.

**Conclusions.**—We fully characterized the problem of asymptotic distillability of quantum coherence under strictly incoherent operations (SIO), showing in particular that almost all states — with the sole exception of states whose density matrix contains a rank-one submatrix — are bound coherent. We introduced a new SIO monotone, the maximal coherence $\eta$, which plays a crucial role in forming a necessary and sufficient criterion for distillability. We established a computable SDP expression for the fidelity of one-shot distillation of a coherence bit under SIO and computed its exact value in the asymptotic many-copy limit in terms of the monotone $\eta$.

Our results reveal that, despite being as useful as the larger classes of free operations IO, DIO, and MIO in some tasks, the operational capabilities of SIO are ultimately very limited, and SIO are unfit for any application relying on coherence distillation. Since many resource manipulation protocols specifically require the use of coherence in pure, distilled form [11, 43–45], this suggests that the class SIO might not be a suitable choice of operations for general information-theoretic tasks. We remark that such a fairly drastic conclusion was a priori unexpected, and not suggested by any previous work.

We note the similarity of our main result to Ref. [46], where a generic phenomenon of bound coherence was also found in the related resource theory of unspeakable coherence (a.k.a. asymmetry) with respect to the set of translationally-covariant incoherent operations [14]; however, it does not appear possible to make this qualitative correspondence also quantitative, as the two settings are fundamentally different.

In light of the considerations of our work and the exposed weakness of SIO in performing coherence distillation, it remains an important open question to understand what the smallest physically-motivated set of free operations for manipulating coherence without such hindering operational limitations could be, and hence the ongoing quest for a satisfactory resource theory of coherence [4] becomes even more enthralling.

**Acknowledgments.**—We are grateful to Iman Marvian, Alexander Streltsov, and Andreas Winter for useful discussions, and to the authors of [26] for sharing a preliminary draft of their work with us. We thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the
programme ‘Beyond I.I.D. in Information Theory’ when part of the work on this paper was undertaken. We acknowledge financial support from the European Research Council (ERC) under the Starting Grant GQCOP (Grant No. 637352).

[1] J. Aberg, (2006), arXiv:quant-ph/0612146.
[2] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. 113, 140401 (2014).
[3] A. Winter and D. Yang, Phys. Rev. Lett. 116, 120404 (2016).
[4] A. Streltsov, G. Adesso, and M. B. Plenio, Rev. Mod. Phys. 89, 041003 (2017).
[5] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B 27, 1345019 (2012).
[6] L. del Rio, L. Kraemer, and R. Renner, (2015), arXiv:1511.08818.
[7] B. Coecke, T. Fritz, and R. W. Spekkens, Inf. Comput. 250, 59 (2016).
[8] E. Chitambar and G. Gour, (2018), arXiv:1806.06107.
[9] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. 81, 865 (2009).
[10] S. Du, Z. Bai, and X. Qi, Quantum Inf Comput 15, 1307 (2015).
[11] X. Yuan, H. Zhou, Z. Cao, and X. Ma, Phys. Rev. A 92, 022124 (2015).
[12] B. Yadin, J. Ma, D. Girolami, M. Gu, and V. Vedral, Phys. Rev. X 6, 041028 (2016).
[13] E. Chitambar and G. Gour, Phys. Rev. Lett. 117, 030401 (2016).
[14] I. Marvian and R. W. Spekkens, Phys. Rev. A 94, 052324 (2016).
[15] E. Chitambar and G. Gour, Phys. Rev. A 94, 052336 (2016).
[16] J. I. de Vicente and A. Streltsov, J. Phys. A: Math. Theor. 50, 045301 (2017).
[17] Z.-W. Liu, X. Hu, and S. Lloyd, Phys. Rev. Lett. 118, 060502 (2017).
[18] A. Streltsov, S. Rana, P. Boes, and J. Eisert, Phys. Rev. Lett. 119, 140402 (2017).
[19] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and X. Ma, Phys. Rev. Lett. 120, 070403 (2018).
[20] B. Regula, K. Fang, X. Wang, and G. Adesso, Phys. Rev. Lett. 121, 010401 (2018).
[21] E. Chitambar, Phys. Rev. A 97, 050301 (2018).
[22] D. Egloff, J. M. Matera, T. Theurer, and M. B. Plenio, Phys. Rev. X 8, 031005 (2018).
[23] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso, Phys. Rev. Lett. 121, 070404 (2018).
[24] T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, (2018), arXiv:1806.07332.
[25] B. Regula, L. Lami, and A. Streltsov, (2018), arXiv:1807.04705.
[26] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and A. Winter, (2018), arXiv:1808.01885.
[27] T. Biswas, M. García Díaz, and A. Winter, Proc. Roy. Soc. London A 473, 2203 (2017).
[28] S. Du, Z. Bai, and Y. Guo, Phys. Rev. A 91, 052120 (2015).
[29] See the Supplemental Material below, where we provide explicit proofs of some of the results discussed in the main text.
[30] H. O. Hirschfeld, Math. Proc. Camb. Philos. Soc. 31, 520 (1935).
[31] H. Gebelein, SIAM J. Appl. Math. 28, 100 (1975).
Supplemental Material

SIo AND THE CASE OF DIFFERENT INPUT AND OUTPUT DIMENSIONS

As reported in the main text, an quantum channel \( \Lambda \) acting on a \( d \)-dimensional system and outputting an \( m \)-dimensional system is defined to be an SIO if it admits a Kraus representation \( \Lambda(\cdot) = \sum_{\alpha} K_{\alpha}(\cdot) K_{\alpha}^\dagger \) such that

\[
K_{\alpha} \Delta(\rho) K_{\alpha}^\dagger = \Delta(K_{\alpha} \rho K_{\alpha}^\dagger) \quad \forall \alpha, \forall \rho,
\]

(S1)

where \( \Delta \) denotes the dephasing map acting on systems of the appropriate dimension. It is easy to verify that when \( m = d \) the validity of Eq. (S1) entails that

\[
K_{\alpha} = U_{\pi_{\alpha}}^\dagger D_{\alpha},
\]

(S2)

\[
U_{\pi_{\alpha}} = \sum_{i=1}^{d} |\pi_{\alpha}(i)\rangle \langle i|,
\]

(S3)

\[
D_{\alpha} = \sum_{i=1}^{d} d_{\alpha}(i) |i\rangle \langle i|,
\]

(S4)

where the \( \pi_{\alpha} \) are permutations. However, it is no longer possible to obtain this simple form when \( m \neq d \). To see why, start by observing that each Kraus operator should now become a \( m \times d \) rectangular matrix. In the case where \( m \geq d \), it is still possible to represent \( K_{\alpha} \) as in Eq. (S2), provided that one makes the sum in Eq. (S3) run all the way to \( m \). The opposite case \( m \leq d \) can be treated in a similar way by exchanging the order of the product in Eq. (S2), which corresponds to applying the above procedure to \( \Lambda^\dagger \) instead of \( \Lambda \).

It is very convenient to have a representation of the Kraus operators of an SIO operation, that is valid for all \( d \) and \( m \). This can be obtained in two different ways. On the other hand, we can write

\[
K_{\alpha} = U_{\pi_{\alpha}}^\dagger D_{\alpha},
\]

(S5)

\[
U_{\pi_{\alpha}} = \sum_{i \in J_{\alpha}} |\pi_{\alpha}(i)\rangle \langle i|,
\]

(S6)

\[
D_{\alpha} = \sum_{i \in J_{\alpha}} d_{\alpha}(i) |i\rangle \langle i|,
\]

(S7)

where \( J_{\alpha} \subseteq \{1, \ldots, m\} \) are subsets and \( \pi_{\alpha} : J_{\alpha} \to \{1, \ldots, d\} \) are injective functions. On the other hand, we can resort to the (non-unique) decomposition

\[
K_{\alpha} = U_{\pi_{\alpha}} U_{\varphi_{\alpha}}^\dagger D_{\alpha},
\]

(S8)

where \( \pi_{\alpha} \) and \( \sigma_{\alpha} \) are permutations acting on \( \{1, \ldots, d\} \) and \( \{1, \ldots, m\} \), respectively.

The careful reader will have noticed that in our proof of the monotonicity of the maximal coherence \( \eta \) under SIO, we restricted ourselves to the case where input and output dimensions coincide. This is possible without loss of generality, because of the following “lift and compress” argument. Given an SIO channel \( \Lambda \) that acts on a \( d \)-dimensional system and outputs an \( m \)-dimensional system, and taken some \( d' \geq \max\{d, m\} \), construct the modified SIO channel \( \Lambda' \) that acts on a \( d' \)-dimensional system and is defined by the formula

\[
\Lambda'(\rho) := \Pi_m^\dagger \Lambda \left( \Pi_d \rho \Pi_d^\dagger \right) \Pi_m,
\]

(S9)

where \( \rho \) is \( d' \times d' \), and \( \Pi_r : \mathbb{C}^d \to \mathbb{C}^r \) denotes the projector onto the subspace spanned by the first \( r \) basis vector. Observe that from Eq. (S9) we can deduce the identity

\[
\Lambda(\rho) = \Pi_m \Lambda' \left( \Pi_d \rho \Pi_d^\dagger \right) \Pi_m^\dagger,
\]

(S10)

valid for all \( d \times d \) matrices \( \rho \). Since in the main text we proved that \( \eta \) is monotonic at least under strictly incoherent operations that do not change the input dimension, we know that it is monotonic in particular under \( \Lambda' \). We now show how to deduce
from this that it is also monotonic under $\Lambda$. Evaluating $\eta$ on both sides of Eq. (S10), and using the elementary observation that
\[ \eta(\Pi r \rho \Pi d^\dagger) \leq \eta(\rho), \]
as an inspection of Eq. (1) immediately reveals, we deduce that
\[ \eta(\Lambda(\rho)) = \eta\left(\Pi_m \Lambda^\dagger \left(\Pi_d^\dagger \rho \Pi_d\right) \Pi_m^\dagger\right) \leq \eta\left(\Lambda^\dagger \left(\Pi_d^\dagger \rho \Pi_d\right)\right) \leq \eta(\Pi_d^\dagger \rho \Pi_d) = \eta(\rho), \]
for all $d \times d$ density matrices $\rho$. This proves that the maximal coherence is in fact monotonic under general SIO operations.

**FIDELITY OF DISTILLATION**

In what follows, we will denote by $A \circ B$ the Hadamard (or Schur, or entrywise) product of two matrices $A$ and $B$ of the same size. Explicitly, we have that $(A \circ B)_{ij} := A_{ij} B_{ij}$.

**Theorem 2** The fidelity of distillation of a single bit of coherence is given by the semidefinite program
\[
F_{\text{SIO}}(\rho, 2) = \max_{\Delta(X) = 0} \int_{-1 \leq X \leq 1} \frac{1}{2} \left(\text{Tr}[\rho |X + 1]\right)
\]
(7a)
\[
= \min_{D = \Delta(D), N \geq 0} \frac{1}{2} \left(\|\rho\| + D + N \|1 + 1\right).
\]
(7b)

**Proof.** Introducing the alternative parametrisation $W = \frac{1}{2} Y$, Eq. (6) can be expressed as the maximization of the function
\[
\frac{1}{2} (\text{Tr} (\rho Y) + 1)
\]
subject to the constraints $-1 \leq Y \leq 1$ and $\Delta(Y) = 0$. The remaining condition $\text{CN}(W) \leq 2$ can be imposed by means of Theorem 1 of [22], which states that a positive semidefinite matrix has $\text{CN}(W) \leq 2$ if and only if the matrix $2 \Delta(W) - |W|$ is positive semidefinite. We thus have
\[
0 \leq 2 \Delta(A) - |A| = 1 - \frac{1 + |Y|}{2} = \frac{1 - |Y|}{2},
\]
(S12)
i.e. $|Y| \leq 1$. Now, we want to argue that this latter condition automatically implies that $-1 \leq Y \leq 1$, or equivalently that $\|Y\|_\infty \leq 1$, which makes this constraint superfluous. To see why, write $\|Y\|_\infty \leq \|Y\|_\infty = \Lambda_{\text{max}}(|Y|) \leq 1$, where the steps are justified as follows: the first inequality is well-known, and can be explicitly seen to hold by writing
\[
\|Y\|_\infty = \max_{\|v\|_2 = 1} |v^\dagger Y v| = \max_{\|v\|_2 = 1} \left|\sum_{i,j} v_i^* v_j Y_{ij}\right|
\]
\[
\leq \max_{\|v\|_2 = 1} \sum_{i,j} \left|v_i^* v_j Y_{ij}\right|
\]
\[
= \max_{\|w\|_2 = 1, w \geq 0} w^\dagger |Y| w \leq \max_{\|w\|_2 = 1} w^\dagger |Y| w = \|Y\|_\infty;
\]
the middle equality follows from the Perron–Frobenius theorem, which implies that the spectral radius of every entrywise nonnegative matrix is itself an eigenvalue, which then by the hermiticity of $|Y|$ coincides with its operator norm; finally, the last inequality is a consequence of the assumption that $|Y| \leq 1$.

Putting everything together, we see that $Y$ is only subjected to the two constraints $\Delta(Y) = 0$ and $|Y| \leq 1$ (equivalently, $-1 \leq Y \leq 1$). We can thus parametrise $Y = |Y| \circ \omega$, where $|Y| := X$ satisfies $\Delta(X) = 0$, $X \geq 0$ and $X \leq 1$ (equivalently, $-1 \leq X \leq 1$), while $\omega = \omega^\dagger$ is any Hermitian matrix composed only of phases (complex numbers of unit modulus). Since the objective function takes the form
\[
\text{Tr}[\rho A] = \frac{1}{2} \left(1 + \sum_{i,j} \rho_{ij} Y_{ij}\right) = \frac{1}{2} \left(1 + \sum_{i,j} \omega_{ij} \rho_{ij} X_{ij}\right),
\]
(S13)
it is maximized by the choices $\omega_{ij} = \omega^\dagger_{ij} = \frac{\rho_{ij}}{|\rho_{ij}|}$ (and $\omega_{ij} = 1$ if $\rho_{ij} = 0$), which — importantly — identify a Hermitian matrix $\omega = \omega^\dagger$. The resulting value of the objective function is
\[
\max_{\omega} \frac{1}{2} \left(1 + \sum_{i,j} \omega_{ij} \rho_{ij} X_{ij}\right) = \frac{1}{2} \left(1 + \sum_{i,j} |\rho_{ij}| X_{ij}\right)
\]
(14)
The maximization over $X$ subjected to the aforementioned constraints yields the first line in the statement of the Theorem. The second line is then simply the corresponding dual SDP — the fact that strong duality holds, and thus the two problems have the same optimal value, can be straightforwardly seen by choosing any matrix $N$ with strictly positive entries as a feasible solution to \cite{[7]} and employing Slater’s theorem \cite{[27]}.

**Theorem 3.** For all states $\rho$ and all integers $n$, one has

\[
\frac{1 + \eta(\rho)}{2} - \frac{\eta(\rho)}{2} \mu^* \leq F_{\text{SIO}}(\rho^\otimes n, 2) \leq \frac{1 + \eta(\rho)}{2},
\]

where $0 < \mu < 1$ is a number that depends only on $\rho$. Hence

\[
\lim_{n \to \infty} F_{\text{SIO}}(\rho^\otimes n, 2) = \frac{1 + \eta(\rho)}{2},
\]

and the convergence in the above identity is exponentially fast.

**Proof.** We start by proving the upper bound in Eq. (9). Consider an arbitrary SIO operation $\Lambda$ that maps a system of dimension $d^n$ into a single qubit. Because of the monotonicity and tensorization properties of the $\eta$ function, we can write $\eta' := \eta(\Lambda(\rho^\otimes n)) \leq \eta(\rho^\otimes n) = \eta(\rho)$. Remembering that $\Lambda(\rho^\otimes n)$ is a qubit state, this means that there are $0 < p < 1$ and $\varphi \in \mathbb{R}$ such that

\[
\Lambda(\rho^\otimes n) := \left( \eta' \frac{p}{\sqrt{p(1-p)}} e^{-i\varphi} \eta' \sqrt{p(1-p)} e^{i\varphi} \right).
\]

The fidelity between the above state and a coherence bit reads

\[
F(\Lambda(\rho^\otimes n), \Psi_2) = \text{Tr} \left[ \Lambda(\rho^\otimes n) \Psi_2 \right] = \frac{1}{2} \left( 1 + 2 \eta' \sqrt{p(1-p)} \cos(\varphi) \right) \leq \frac{1}{2} (1 + \eta') \leq \frac{1}{2} (1 + \eta(\rho)),
\]

where for the first inequality we noted that $2 \sqrt{p(1-p)} \leq 1$. Taking the supremum over all SIO $\Lambda$ yields the upper bound in Eq. (9).

The lower bound in Eq. (9) can be proved by designing a suitable SIO protocol that achieves the prescribed fidelity on $n$ copies. To do this, without loss of generality we are going to assume that for the particular state $\rho$ we are considering: (a) the maximum in Eq. (1) is achieved on the pair $(i, j) = (1, 2)$; and (b) $\rho_{12}$ is real. These two assumptions imply that

\[
\eta(\rho) = \frac{\rho_{12}}{\sqrt{\rho_{11} \rho_{22}}}.
\]

Now, we construct a suitable “diagonal filtering” SIO instrument $\Lambda_{DF}$ that maps a $d$-dimensional system into a qubit and is defined by the Kraus operators

\[
K_0 := \sqrt{\min\{\rho_{11}, \rho_{22}\}} \left( \rho_{11}^{-1/2} |1\rangle\langle 1| + \rho_{22}^{-1/2} |2\rangle\langle 2| \right),
\]

\[
K_1 := \sqrt{1 - \min\left\{ \frac{\rho_{22}}{\rho_{11}}, |1\rangle\langle 1| \right\}},
\]

\[
K_2 := \sqrt{1 - \min\left\{ \frac{\rho_{11}}{\rho_{22}}, |1\rangle\langle 2| \right\}},
\]

\[
K_\alpha := |1\rangle\langle \alpha| \text{ for } \alpha = 3, \ldots, d.
\]

The probability of getting the outcome $\alpha = 0$ when applying the instrument $\Lambda_{DF}$ on $\rho$ is clearly

\[
P(0) = \text{Tr} \left[ K_0 \rho K_0^\dagger \right] = 2 \min\{\rho_{11}, \rho_{22}\}.
\]
The post-measurement state conditioned on the outcome 0 is then
\[ \tilde{\rho}_0 = \frac{K_0 \rho K_0^\dagger}{P(0)} = \frac{1}{2} \begin{pmatrix} 1 & \eta(\rho) \\ \eta(\rho) & 1 \end{pmatrix}. \]

Let us apply the instrument \( \Lambda_{DF} \) separately on each one of the \( n \) copies of \( \rho \) we have at our disposal. Since
\[ F(\tilde{\rho}_0, \Psi_2) = \frac{1 + \eta(\rho)}{2} \]
matches the upper bound in Eq. (9), we have achieved maximal distillation fidelity whenever at least one of the \( n \) outcomes we obtain is \( \alpha = 0 \). This happens with probability
\[ P_{\text{success}} = 1 - (1 - P(0))^n = 1 - (1 - 2 \min \{ \rho_{11}, \rho_{22} \})^n = 1 - \mu_\rho^n, \]
where \( \mu_\rho := 2 \min \{ \rho_{11}, \rho_{22} \} \). If none of the outcomes is \( \alpha = 0 \), then we can simply output the fixed state \( |1\rangle \). The average distillation fidelity of this protocol is
\[ \bar{F} = (1 - \mu_\rho^n) \cdot \frac{1 + \eta(\rho)}{2} + \mu_\rho^n \cdot \frac{1}{2} = \frac{1 + \eta(\rho)}{2} - \frac{\eta(\rho)}{2} \mu_\rho^n, \]
reproducing the lower bound in Eq. (9).

**Proposition 8.** For the state \( \rho_\lambda = \lambda \Psi_2 + (1 - \lambda) \frac{1}{2} \) with \( \lambda \in [0, 1] \), it holds that \( F_{\text{SIO}}(\rho_{\lambda \otimes n}^{\otimes n}, 2) = \frac{\lambda + 1}{2} \) for any \( n \in \mathbb{N} \).

**Proof.** We obtain the lower bound of \( \frac{\lambda + 1}{2} \) by considering \( n = 1 \) and simply noting that \( F(\rho_\lambda, \Psi_2) = \frac{\lambda + 1}{2} \). To show the upper bound, we consider the SDP (7b) for \( F_{\text{SIO}}(\rho_{\lambda \otimes n}^{\otimes n}, 2) \) and take as feasible solutions the following choices:
\[
D = -\frac{1 - \lambda}{2^n} 1, \\
N = \lambda \Psi_2^{\otimes n} + \frac{1 - \lambda}{2^n} 1 - \rho_{\lambda \otimes n}^{\otimes n},
\]
so that \( |\rho_{\lambda \otimes n}^{\otimes n}| + D + N = \lambda \Psi_2^{\otimes n} \). It remains to verify that \( N \) is a valid feasible solution, that is, that all of its coefficients are non-negative. This follows by noticing that the diagonal elements of \( N \) are given as
\[ N_{ii} = \lambda \frac{1}{2^n} + (1 - \lambda) \frac{1}{2^n} - [\rho_{\lambda \otimes n}^{\otimes n}]_{ii} = \frac{1}{2^n} - \frac{1}{2^n} = 0 \]
and similarly the off-diagonal elements are \( N_{ij} = \frac{1}{2^n} - [\rho_{\lambda \otimes n}^{\otimes n}]_{ij} \). Since the off-diagonal elements of \( \rho_{\lambda \otimes n}^{\otimes n} \) are always of the form \( \frac{d^m}{2^m} \) for some \( 1 \leq m \leq n \), and moreover \( \lambda \leq 1 \), we get \( N_{ij} \geq 0 \). Hence
\[
F_{\text{SIO}}(\rho_{\lambda \otimes n}^{\otimes n}, 2) \leq \frac{1}{2} \left( ||\rho_{\lambda \otimes n}^{\otimes n}| + D + N||_1 + 1 \right) = \frac{1}{2} \left( \|\lambda \Psi_2^{\otimes n}\|_1 + 1 \right) = \frac{1}{2} (\lambda + 1)
\]
as required.

**RATE OF DISTILLATION**

We start from a \( d \)-dimensional state \( \rho \), which is as usual assumed to satisfy \( \Delta(\rho) > 0 \) without loss of generality. Let us define the positive matrix
\[ A_\rho := \Delta(\rho)^{-1/2} \rho \Delta(\rho)^{-1/2}, \]
which satisfies \( (A_\rho)_{ii} = 1 \) for all \( i \). Consider the graph \( G_\rho = (V_\rho, E_\rho) \) with vertices \( V_\rho := \{1, \ldots, d\} \) and edges
\[ E_\rho := \{ (i, j) : ||(A_\rho)_{ij}||_1 = 1 \} = \{ (i, j) : |\rho_{ij}| = \sqrt{\rho_{ii} \rho_{jj}} \}. \]
For simplicity, we have included into \( E_\rho \) all “diagonal” pairs of the form \( (i, i) \). If \( E_\rho \) contains only these elements, we say that \( E_\rho \) is trivial.
Moreover, setting which concludes the proof.

The following result summarizes the main elementary properties of this object.

Corollary 10. For all states $\rho$ there exists a completely positive, trace-preserving and unital channel $\mathcal{P}_\rho$ such that the “trimmed” state $\bar{\rho}$ of Eq. (S19) satisfies $\bar{\rho} = \mathcal{P}_\rho(\rho)$. In particular, $\bar{\rho}$ is a legitimate density matrix, and moreover

$$0 \leq Q(\rho) \leq S(\Delta(\rho)) - S(\rho) = C_{d, IO}(\rho).$$

(S21)
The distillation rate associated with this protocol is then

\[ C \rho := \sum_{\nu} P(\nu) S(\Delta(\rho)), \]

where \( P(\nu) := \text{Tr}[\Pi_{\nu} / \rho] \). By virtue of Eq. (S20), it is not difficult to check that indeed \( \tilde{\rho} = \mathcal{P}_\rho(\rho) \), which shows that \( \tilde{\rho} \) is a density matrix. As \( S(\Delta(S)) - S(S) \geq 0 \) for all density matrices, and \( \Delta(\tilde{\rho}) = \Delta(\rho) \), this shows immediately that \( Q(\rho) \geq 0 \). Moreover, since \( \mathcal{P}_\rho \) is clearly completely positive, trace-preserving and unital, and unital channels never decrease the entropy, we have that

\[ Q(\rho) = S(\Delta(\rho)) - S(\mathcal{P}_\rho(\rho)) \leq S(\Delta(\rho)) - S(\rho). \]

The final inequality in Eq. (S21) follows from Theorem 6 in [3].

It is not difficult to show that a state \( \rho \) satisfies \( Q(\rho) > 0 \) if and only if the set of edges \( E_\rho \) is nontrivial, i.e. if and only if \( \eta(\rho) = 1 \). This observation, together with Theorem 1, provides some quantitative evidence in favour of the conjectured equality \( C_{d_\text{SIO}} = Q \) (Conjecture 5). In fact, it shows that \( C_{d_\text{SIO}} > 0 \) if and only if \( Q(\rho) > 0 \). We now strengthen this result by showing that \( Q \) always lower bounds \( C_{d_\text{SIO}} \).

**Proposition 4** For all states \( \rho \) in any dimension, the SIO distillable coherence satisfies

\[ C_{d_\text{SIO}}(\rho) \geq Q(\rho). \]  

**Proof.** As described in the main text, there is a simple SIO protocol that achieves a rate \( Q(\rho) \). This is composed of three steps, that we recall below referring to Lemma 9 for notation.

(i) One applies the instrument with Kraus operators \( \{ \Pi_{\nu} \}_{\nu} \) on each of the \( n \) copies of \( \rho \) that are initially available.

(ii) In the limit of large \( n \), each outcome \( \nu \) is obtained an average number of times equal to \( n P(\nu) \), where \( P(\nu) := \text{Tr}[\Pi_{\nu} / \rho] \).

(iii) The post-measurement state corresponding to the outcome \( \nu \), denoted by \( \tilde{\rho}_\nu := P(\nu)^{-1} \Pi_{\nu} \rho \Pi_{\nu}^\dagger \), is pure by Lemma 9; it is then known [10] that there is an SIO protocol that extracts coherence bits at a rate \( S(\Delta(\tilde{\rho}_\nu)) \); since we started with \( n P(\nu) \) states, we obtain \( n P(\nu) S(\Delta(\tilde{\rho}_\nu)) \) bits of coherence at the output.

The distillation rate associated with this protocol is then

\[ r = \sum_{\nu} P(\nu) S(\Delta(\tilde{\rho}_\nu)) = \sum_{\nu} P(\nu) \left( P(\nu)^{-1} \Pi_{\nu} \Delta(\rho) \Pi_{\nu}^\dagger \right) = S(\Delta(\tilde{\rho})) - S(\rho) = Q(\rho), \]

as claimed. This intuitive yet sketchy description of the protocol can be complemented with the following rigorous analysis. Fix an \( \epsilon, \delta > 0 \). By the weak law of large numbers, the number of times the outcome \( \nu \) is obtained in step (ii) will be \( N_\nu \geq n(P(\nu) - \delta) \) for all \( \nu \) with probability \( P_{n,\delta} \) converging to 1 as \( n \to \infty \). For a fixed \( \nu \), the SIO protocol described in [11] is able to extract from \( \tilde{\rho}_\nu^{\otimes N_\nu} \) a number \( |N_\nu (S(\Delta(\tilde{\rho}_\nu)) - \delta)| \) of coherence bits with vanishing error. Using the fact that \( S(\Delta(\tilde{\rho}_\nu)) \leq \log d \) and that \( P(\nu) \leq 1 \), observe that

\[ N_\nu (S(\Delta(\tilde{\rho}_\nu)) - \delta) \geq n(P(\nu) - \delta) (S(\Delta(\tilde{\rho}_\nu)) - \delta) \geq n(P(\nu) S(\Delta(\tilde{\rho}_\nu)) - (1 + \log d)\delta). \]

Up to discarding some of the produced coherence bits, we are thus able to convert \( \tilde{\rho}_\nu^{\otimes N_\nu} \) into a state \( \sigma_\nu \) such that \( \| \sigma_\nu - \sigma_\nu^{\text{ideal}} \|_1 \leq \epsilon \), where \( \sigma_\nu^{\text{ideal}} := P_{n,\delta} \left( P(\nu) S(\Delta(\tilde{\rho}_\nu)) - (1 + \log d)\delta \right) \). With probability \( P_{n,\delta} \) approaching one, the outcome of the protocol will then be \( \otimes \nu \sigma_\nu \). Using a standard telescopic technique together with the fact that the index \( \nu \) can take at most \( d \) values, it is not difficult to verify that

\[ \left\| \otimes \nu \sigma_\nu - \otimes \nu \sigma_\nu^{\text{ideal}} \right\|_1 \leq \sum \nu \| \sigma_\nu - \sigma_\nu^{\text{ideal}} \|_1 \leq d \epsilon. \]
Moreover,
\[
\bigotimes_v \sigma_v^\text{ideal} = \Psi_2^\otimes \sum_v \left( P(v) S(\Delta \rho_v) - (1 + \log d) \delta \right) = \Psi_2^\otimes Q(\rho) - (1 + \log d) \delta.
\]

Since \(d\) is fixed and \(\epsilon\) is arbitrary, we conclude that \(Q(\rho) - (1 + \log d) \delta\) is an achievable rate for all \(\delta > 0\). Taking the supremum in \(\delta\), this shows that \(C_{d,\text{STO}}(\rho) \geq Q(\rho)\).

One could speculate that a better lower bound on \(C_{d,\text{STO}}\) can be obtained by applying the above distillation protocol to many copies of the state \(\rho\) simultaneously, which leads to the bound
\[
C_{d,\text{STO}}(\rho) \geq \lim_{n \to \infty} \frac{1}{n} Q(\rho^\otimes n).
\]

However, it turns out that \(Q\) is additive over tensor products, and thus the r.h.s. of the above equation coincides with \(Q(\rho)\) itself. We conclude our discussion by proving this last property.

**Lemma 12.** For all states \(\rho\) and \(\sigma\) of any dimension,
\[
\rho \otimes \sigma = \bar{\rho} \otimes \bar{\sigma}
\]
and consequently
\[
Q(\rho \otimes \sigma) = Q(\rho) + Q(\sigma).
\]

**Proof.** Observe that our usual assumption that \(\Delta(\rho \otimes \sigma) > 0\) is equivalent to requiring that both \(\Delta(\rho) > 0\) and \(\Delta(\sigma) > 0\). Call \(d_1\) the dimension of the space on which \(\rho\) acts, and \(d_2\) that of the space on which \(\sigma\) acts. For all \(i, k \in \{1, \ldots, d_1\}\) and \(j, l \in \{1, \ldots, d_2\}\), one has
\[
|\langle \rho \otimes \sigma \rangle_{ij,kl}| = |\rho_{ik}| |\sigma_{jl}| \leq \sqrt{\rho_{ii}\rho_{kk}} \sqrt{\sigma_{jj}\sigma_{ll}} = \sqrt{\langle \rho \otimes \sigma \rangle_{ij,ij} (\rho \otimes \sigma)_{kl,kl}},
\]
Hence, the following facts are easily seen to be equivalent: (i) \((ij, kl) \in E_{\rho \otimes \sigma}\); (ii) the above inequality is saturated; (iii) \(|\rho_{ik}| = \sqrt{\rho_{ii}\rho_{kk}}\) and \(|\sigma_{jl}| = \sqrt{\sigma_{jj}\sigma_{ll}}\); (iv) \((i, k) \in E_\rho\) and \((j, l) \in E_\sigma\). Hence,
\[
\bar{\rho} \otimes \bar{\sigma} = \sum_{(ij, kl) \in E_{\rho \otimes \sigma}} (\rho \otimes \sigma)_{ij,kl} |ij\rangle \langle kl| = \sum_{(i, k) \in E_\rho, (j, l) \in E_\sigma} \rho_{ik} |i\rangle \langle k| \otimes |l\rangle \langle j|
\]
\[
= \left( \sum_{(i, k) \in E_\rho} \rho_{ik} |i\rangle \langle k| \right) \otimes \left( \sum_{(j, l) \in E_\sigma} \sigma_{jl} |j\rangle \langle l| \right)
\]
\[
= \bar{\rho} \otimes \bar{\sigma}.
\]