PAPER • OPEN ACCESS

One-shot entanglement distillation beyond local operations and classical communication

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
5 July 2019

REVISED
2 September 2019

ACCEPTED FOR PUBLICATION
24 September 2019

PUBLISHED
9 October 2019

PAPER

One-shot entanglement distillation beyond local operations and classical communication

Bartosz Regula[1,2], **Kun Fang**[3], **Xin Wang**[4] and **Mile Gu**[1,2,5]

[1] School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore
[2] Complexity Institute, Nanyang Technological University, 637335, Singapore
[3] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, CB3 0WA, United Kingdom
[4] Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD 20742, United States of America
[5] Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore

E-mail: bartosz.regula@gmail.com, kf383@cam.ac.uk, wangxinfelix@gmail.com and mgu@quantumcomplexity.org

Keywords: entanglement distillation, entanglement measures, convex optimisation

Abstract

We study the task of entanglement distillation in the one-shot setting under different classes of quantum operations which extend the set of local operations and classical communication (LOCC). Establishing a general formalism which allows for a straightforward comparison of their exact achievable performance, we relate the fidelity of distillation under these classes of operations with a family of entanglement monotones and the rates of distillation with a class of smoothed entropic quantities based on the hypothesis testing relative entropy. We then characterise exactly the one-shot distillable entanglement of several classes of quantum states and reveal many simplifications in their manipulation. We show in particular that the ε-error one-shot distillable entanglement of any pure state is the same under all sets of operations ranging from one-way LOCC to separability-preserving operations or operations preserving the set of states with positive partial transpose, and can be computed exactly as a quadratically constrained linear program. We establish similar operational equivalences in the distillation of isotropic and maximally correlated states, reducing the computation of the relevant quantities to linear or semidefinite programs. We also show that all considered sets of operations achieve the same performance in environment-assisted entanglement distillation from any state.

1. Introduction

Quantum entanglement plays a fundamental role in quantum information processing by serving as a resource which underlies many important protocols such as quantum teleportation [1] or superdense coding [2] as well as quantum technological applications such as quantum repeaters and networks [3, 4]. Many such schemes require the use of entanglement in the pure, maximal form of singlets—the efficient conversion of entanglement into such form, dubbed *entanglement distillation* [5, 6], is thus of vital importance, and the development of effective theoretical and practical methods to characterise entanglement distillation remains at the forefront of quantum information research [7]. First studied in the asymptotic regime under the assumption of being able to manipulate an unbounded number of independent and identically distributed copies of a quantum system [5, 6, 8–11], distillation later attracted a significant amount of research using the tools of non-asymptotic quantum information theory [12–21]. The latter setting is of particular importance due to the physical limitations of near-term quantum technologies, preventing us from being able to manipulate large numbers of quantum systems effectively. In particular, to efficiently exploit entanglement in practical settings it is necessary to obtain a thorough understanding of *one-shot* distillation of entanglement, which takes into account the realistic, non-asymptotic restrictions on state transformations and aims to understand how finite accuracy limits our ability to manipulate entanglement.

The characterisation of entanglement as a resource in practical settings is rooted in the so-called distant labs paradigm [7], in which experimenters are free to perform any local operation within their own labs and

communicate with each other classically, but any use of quantum communication has an associated resource cost since it requires the use of entanglement. This formalism led to the definition of local operations and classical communication (LOCC) as the set of allowed ('free') operations, and the operational characterisation of entanglement distillation is concerned precisely with delimiting the capabilities of LOCC in manipulating entanglement. However, the mathematical description of LOCC is known to have a highly complicated structure [22], making many important questions in the resource theory of entanglement either very challenging or downright unanswerable. This motivated the investigation of several relaxations of the class LOCC [23–27], whose simplified description can provide accessible upper bounds on the capabilities of LOCC as well as establish the ultimate limitations on entanglement transformations. Understanding the properties of such relaxations and characterising their precise operational power can therefore shed light on the fundamental structure of entanglement as a resource.

In this work, we develop a comprehensive framework for the study of one-shot entanglement distillation under several different classes of operations—separable maps (SEP), separability-preserving maps (SEPP), positive partial transpose (PPT) maps, two types of PPT-preserving maps, as well as two types of maps based on the so-called Rains set—many of which have been considered in the literature as a relaxation of LOCC in various contexts, but whose one-shot distillation capabilities in relation to other operations remained unknown. Such extensions are still bound by operationally motivated constraints (e.g. SEPP can never generate entanglement from an unentangled state, just as LOCC), but they can often be understood as allowing for additional resources to be used in entanglement manipulation (e.g. any PPT operation can be stochastically implemented by LOCC if one is additionally given access to a bound entangled state [28]). We compare the performance of these sets of maps in distiling entanglement in the one-shot setting, establishing in particular a general formalism which allows us to describe the distillation under the different operations together in a unified framework. We make use of tools from convex analysis and convex optimisation to relate the rates of distillation with a family of entanglement monotones. By evaluating these monotones for all pure states, isotropic states, and maximally correlated states, we simplify the description of distillation in these cases and show that many of the relaxations coincide in their distillation power, facilitating an efficient quantification of fundamental entanglement properties and revealing many operational similarities in entanglement manipulation under different classes of channels.

Our work improves many earlier results in the characterisation of one-shot entanglement distillation [16, 17, 19, 20, 29], which relied on approximate bounds and were only exact asymptotically; crucially, our formalism allows for a precise description of distillation already at the one-shot level, providing an exact characterisation of the operational power of several classes of operations which extend LOCC and shedding light on the capabilities of LOCC themselves.

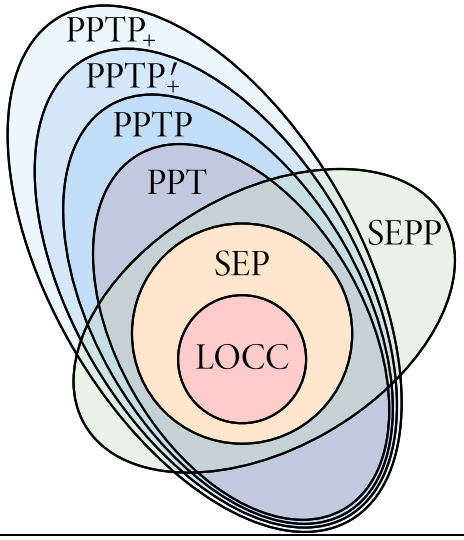
1.1. Summary of results

We begin our work in section 2 with a brief introduction to a family of entanglement monotones $T_S^{(m)}$ which will play an important role in the later investigation of entanglement distillation. We characterise their properties and in particular show that the class of monotones generalises two known measures of entanglement—the robustness of entanglement and a distance-based quantifier based on trace distance—which will allow us to endow the measures with a direct operational meaning.

Our characterisation of entanglement distillation begins in section 3 where we establish explicit general connections between the quantifiers $T_S^{(m)}$, quantum hypothesis testing, and one-shot entanglement distillation through convex duality. The methods will form the foundations of the framework developed in this work.

We commence the explicit applications of our framework in section 3.1 by quantifying the distillation capabilities of several classes of operations based on the set of PPT states, recovering previous results of [25, 26] as well as describing new classes of operations in this context. The results additionally allow for an understanding of important asymptotic quantities, such as the regularised PPT relative entropy of entanglement or the Rains bound [24, 25], not just as bounds for distillable entanglement but as quantities with a precise operational meaning of their own. This section serves also as an introduction to the formalism considered in the manuscript and showcases the generality of our methods.

In section 3.2, we consider the class of separability-preserving operations [18, 27]. By relating the achievable fidelity of distillation with the monotones $T_S^{(m)}$ again, we establish an operational interpretation of the generalised robustness of entanglement in the context of distillation. Furthermore, we demonstrate a general operational equivalence in the distillation from pure states: all sets of operations, ranging from one-way LOCC to SEPP and PPT-preserving operations, achieve exactly the same performance in one-shot pure-state distillation. Although such an equivalence in the asymptotic regime was already known [10, 30], the correspondence already in the one-shot setting is remarkable, considering that the one-shot manipulation power of the larger sets of operations is generally much greater than that of LOCC. The results allow us to explicitly relate the fidelity of distillation of any pure state with an analytically computable norm of its Schmidt



Class of operations	Acronym
Local operations and classical communication	LOCC
Separable operations	SEP
Separability-preserving operations (see Sec. 3.2)	SEPP
PPT operations (see Def. 6)	PPT
PPT-preserving operations (see Def. 8)	PPTP
Rains-preserving operations (see Sec. 3.1.1)	PPTP'_+
PPT_+-preserving operations (see Def. 7)	PPTP_+

Figure 1. Schematic hierarchy of operations considered in this work. The pictured inclusions between PPTP_+, PPTP'_+, PPTP, PPT, SEP, and LOCC are all strict; there is no inclusion between SEPP and any of the sets PPTP_+, PPTP, PPT in general.

coefficients and express the computation of the ε-error one-shot distillable entanglement of a pure state as a convex quadratically-constrained linear program.

We continue in section 3.3 by establishing a similar operational equivalence in the distillation of isotropic states, showing that any class of operations ranging from separable operations to PPT- and separability-preserving operations achieve the same one-shot rates of distillation. Analogously, in section 3.4 we show that separability-preserving operations provide no advantage over PPT operations in the distillation from maximally correlated states, and furthermore, by relating the entanglement monotones with measures of quantum coherence, the achievable rates and fidelities of distillation can be computed efficiently as semidefinite programs.

In section 3.5 we show how our results immediately imply that in the setting of environment-assisted entanglement distillation [17, 31, 32], all considered operations—from one-way LOCC to PPT- and separability-preserving—achieve exactly the same performance. We furthermore recover the one-shot characterisation of [17] in a simplified fashion by employing the formalism introduced herein.

We conclude in section 3.6 with a discussion of zero-error distillation under the different sets of operations, obtaining in particular a single-letter formula for the asymptotic zero-error distillable entanglement under Rains-preserving operations which recovers a bound of [33] and endows it with an operational interpretation as a zero-error Rains bound.

An overview of the sets of operations considered in this work as well as their relative power in one-shot entanglement distillation is provided in figure 1 and table 1.

2. Preliminaries

We will work in the real vector space of Hermitian matrices $\mathbb{H}$ with the Hilbert–Schmidt inner product $\langle X, Y\rangle = \mathrm{Tr}(X^\dagger Y)$. We will denote by $\mathbb{H}_+$ the cone of positive semidefinite matrices and by $\succcurlyeq$ the inequality with respect to this cone, that is, $X \in \mathbb{H}_+ \iff X \succcurlyeq 0$. We will denote by $\mathbb{H}_1$ the set of unit trace Hermitian matrices, and by $\mathbb{D} = \mathbb{H}_1 \cap \mathbb{H}_+$ the set of density matrices. The notation $|x\rangle$ will be used to refer to general vectors in $\mathbb{C}^d$, with Greek letters such as $|\psi\rangle$ reserved for normalised vectors corresponding to quantum states; in the latter case we will often refer to the projector $|\psi\rangle\langle\psi|$ as ψ. We will use $\||x\rangle\|_{\ell_p} = (\sum_i |x_i|^p)^{1/p}$ for the p-norms in $\mathbb{C}^d$ and $\|X\|_p = \mathrm{Tr}(|X|^p)^{1/p}$ for the Schatten p-norms in $\mathbb{H}$.

Table 1. Comparison between the one-shot ε-error entanglement distillation rates achievable under the different sets of operations considered in this work. We use $>$ when the given inequality can be strict for some state, and $\geqslant$ if—to the best of our knowledge—the strictness of the inequality remains an open question.

General states	Pure states	Isotropic states	Max. corr. states
SEPP	SEPP	SEPP	SEPP
$\vee$	$\|\|$	$\|\|$	$\|\|$
PPTP_+	PPTP_+	PPTP_+	PPTP_+
$\vee$	$\|\|$	$\|\|$	$\|\|$
PPTP'_+	PPTP'_+	PPTP'_+	PPTP'_+
$\vee$	$\|\|$	$\|\|$	$\|\|$
PPTP	PPTP	PPTP	PPTP
$\|\|$	$\|\|$	$\|\|$	$\|\|$
PPT	PPT	PPT	PPT
$\vee$	$\|\|$	$\|\|$	$\|\vee$
SEP	SEP	SEP	SEP
$\|\vee$	$\|\|$	$\|\vee$	$\|\vee$
(1-)LOCC	(1-)LOCC	(1-)LOCC	(1-)LOCC

For any set $\mathcal{Q}$, we define the dual cone $\mathcal{Q}^*=\{X\,|\,\langle X,\,Q\rangle \geqslant 0\ \forall\, Q\in\mathcal{Q}\}$ and the polar set $\mathcal{Q}^\circ=\{X\,|\,\langle X,\,Q\rangle\leqslant 1\ \forall\ Q\in\mathcal{Q}\}$. We have in particular $\mathcal{Q}^{**}:=(\mathcal{Q}^*)^*=\mathrm{cl}\,\mathrm{conv}\{\lambda Q|\ \lambda\geqslant 0,\ \ Q\in\mathcal{Q}\}$ (the closure of the conic hull of $\mathcal{Q}$) and $\mathcal{Q}^{\circ\circ}:=(\mathcal{Q}^\circ)^\circ=\mathrm{cl}\,\mathrm{conv}(\mathcal{Q}\cup\{0\})$ where cl denotes closure and conv the convex hull of a set.

All logarithms in this work are base 2. We will use the shorthand

$$\lfloor x\rfloor_{\log}:=\log\lfloor 2^x\rfloor, \tag{1}$$

and analogously for $\lceil x\rceil_{\log}$.

2.1. A family of entanglement monotones

The analysis of this work will focus on understanding the achievable fidelity of distillation under different sets of operations, and establishing methods allowing us to relate it with convex optimisation problems which admit an efficient characterisation. To this end, we will introduce a family of entanglement monotones, which we will later explicitly endow with an operational interpretation and show to play a fundamental role in characterising entanglement distillation.

Consider a bipartite system shared between parties A and B, with d_A and d_B denoting the dimensions of the corresponding spaces. Let $d=\min\{d_A,\,d_B\}$. We will consider the following sets of Hermitian matrices:

$$\begin{aligned}\mathsf{PPT}&=\{X|\ \mathrm{Tr}(X)=1,\ X^{T_B}\succcurlyeq 0\}\\ \mathsf{PPT}_+&=\{X|\ \mathrm{Tr}(X)=1,\ X\succcurlyeq 0,\ X^{T_B}\succcurlyeq 0\}\\ \mathsf{SEP}&=\mathrm{conv}\{|\psi\rangle\langle\psi|\,||\psi\rangle=|\phi\rangle_A\otimes|\eta\rangle_B\},\end{aligned} \tag{2}$$

where X^{T_B} is the partial transpose of X. Letting $\mathcal{S}$ denote one of the above sets, a quantifier which found use in measuring the entanglement of quantum states in several contexts is the *generalised robustness*, defined as [34]

$$\begin{aligned}R_S^{\mathbb{D}}(\rho)&=\min\{\lambda\geqslant 0|\ \rho\preceq(\lambda+1)X,\ X\in S\}\\&=\max\{\langle\rho,\,W\rangle\,|-\mathbb{1}\preceq W,\ \ W\in -S^*\}.\end{aligned} \tag{3}$$

We then extend this definition to a class of measures

$$T_S^{(m)}(\rho):=\max\{\langle\rho,\,W\rangle\,|-\mathbb{1}\preceq W\preceq m\mathbb{1},\ \ W\in -S^*\} \tag{4}$$

for some parameter $m\in\mathbb{R}_+$. We note this class of measures has been considered in [35], but we have found that some of the results concerning the quantification and characterisation of $T_S^{(m)}$ stated there are in fact incorrect, so we present a self-contained investigation of their basic properties below and in the appendix.

A useful characterisation of the quantifiers is obtained by considering their dual form, which can be obtained as follows.

Proposition 1. *The measures $T_S^{(m)}$ can be equivalently expressed as*

$$T_S^{(m)}(\rho)=\min\{m\,\mathrm{Tr}\,(\rho-X)_+ + \mathrm{Tr}\,(\rho-X)_-|\ X\in\mathcal{S}^{**}\}, \tag{5}$$

where $(\rho-X)_+$ (respectively, $(\rho-X)_-$) denotes the positive (negative) part of the Hermitian operator $\rho-X$.

The proof follows well-known methods in matrix analysis and we include it in the appendix for completeness.

In particular, for $m = 1$ the measures take the remarkably simple form $T_{\mathcal{S}}^{(1)}(\rho) = \min_{X\in\mathcal{S}^{**}} \|\rho - X\|_1$. This quantity, considered first in the resource theory of coherence as the modified trace distance [36], generalises the commonly employed trace distance measure $\min_{\sigma\in\mathcal{S}} \|\rho - \sigma\|_1$. The reason why $T_{\mathcal{S}}^{(1)}$ is a more suitable measure of entanglement than the trace distance itself is the fact that, contrary to $T_{\mathcal{S}}^{(1)}$, the trace distance does not satisfy strong monotonicity under LOCC [36, 37] (i.e. the requirement that a measure M obeys $M(\rho) \geqslant \sum_i p_i M(\Lambda_i(\rho))$ for any probabilistic protocol which applies an LOCC transformation Λ_i to ρ with probability p_i), which is often considered as one of the basic requirements that a measure of entanglement should satisfy [38]. This demonstrates a case where it becomes necessary to consider the distance with respect to the unnormalised cone $\mathcal{S}^{**}$ rather than the set $\mathcal{S}$ in order to ensure strong monotonicity.

Another interesting case is $m = d - 1$, for which we obtain the following.

Proposition 2. *For any* $\mathcal{S} \in \{\mathsf{PPT}, \mathsf{PPT}_+, \mathsf{SEP}\}$ *it holds that*

$$T_{\mathcal{S}}^{(d-1)}(\rho) = R_{\mathcal{S}}^{\mathbb{D}}(\rho). \tag{6}$$

Proof. Let $-W \in \mathsf{SEP}^*$, and notice that $\mathsf{SEP} \subseteq \mathsf{PPT}_+ \subseteq \mathsf{PPT} \Rightarrow \mathsf{PPT}^* \subseteq \mathsf{PPT}_+^* \subseteq \mathsf{SEP}^*$. Letting $\lambda_{\min}$ denote the smallest and $\lambda_{\max}$ the largest eigenvalue of a given matrix, we now use the property that $-W \in \mathsf{SEP}^* \Rightarrow \lambda_{\min}(-W) \geqslant (1 - d)\lambda_{\max}(-W)$ [39, Cor. 5.5] together with the constraint $W \succcurlyeq -\mathbb{1}$ to obtain

$$\lambda_{\max}(W) \leqslant d - 1. \tag{7}$$

It follows that the feasible sets for $R_{\mathcal{S}}^{\mathbb{D}}$ and $T_{\mathcal{S}}^{(d-1)}$ are equal and so the problems are equivalent. ■

From the above two Propositions, we have that the family of measures $T_{\mathcal{S}}^{(m)}$ can be understood as interpolating between the robustness of $\mathcal{S}$ for $m = d - 1$, the modified trace distance for $m = 1$, and the trivial value of 0 for $m = 0$. It is furthermore easy to see that $T_{\mathcal{S}}^{(m)}(\rho) = T_{\mathcal{S}}^{(d-1)}(\rho)$ for any $m > d - 1$.

2.2. Generalising $T_{\mathcal{S}}^{(m)}$ to arbitrary sets

In the operational characterisation of entanglement distillation, it will be necessary to consider also generalisations of the above measures beyond sets of normalised (unit trace) Hermitian operators. To allow for this, we will now consider arbitrary compact sets of Hermitian operators $\mathcal{Q}$ and define the quantity

$$G_{\mathcal{Q}}^{(m)}(\rho) := \sup\left\{\langle\rho,\, W\rangle \;\middle|\; 0 \preceq W \preceq \mathbb{1},\;\; W \in \frac{1}{m}\mathcal{Q}^\circ\right\}. \tag{8}$$

It is straightforward to see that if $\mathcal{Q}$ is a set of unit trace operators, then $G_{\mathcal{Q}}^{(m)}(\rho)$ is equal to $\frac{1}{m}(T_{\mathcal{Q}}^{(m-1)}(\rho) + 1)$, although this relation does not hold in general.

To obtain a general dual formulation of $G_{\mathcal{Q}}^{(m)}$, we will employ the formalism of gauge functions [40, 41]. The convex gauge function of a set $\mathcal{Q}$ is defined as [40]

$$\Gamma_{\mathcal{Q}}(\rho) = \inf\{\lambda \geqslant 0 |\; \rho \in \lambda\,\mathrm{conv}(\mathcal{Q})\} = \sup\{\langle\rho,\, W\rangle \mid W \in \mathcal{Q}^\circ\}, \tag{9}$$

and can be thought of as an extension of the concept of a norm associated with a set—indeed, all norms are gauge functions, but the latter are more general. One can further notice that

$$\Gamma_{\mathcal{Q}^\circ}(\rho) = \inf\{\lambda \geqslant 0 |\; \rho \in \lambda\mathcal{Q}^\circ\} = \sup\{\langle\rho,\, W\rangle \mid W \in \mathcal{Q}^{\circ\circ}\}, \tag{10}$$

where the first equality follows because the set $\mathcal{Q}^\circ$ is always convex. We will take $\inf\varnothing = -\sup\varnothing = \infty$ and note that when the set $\mathcal{Q}$ is compact, the infima and suprema in the definitions of the gauge functions are attained as long as they are finite. We then have the following.

Proposition 3. *For any compact set* $\mathcal{Q} \subseteq \mathbb{H}$*, we have that*

$$G_{\mathcal{Q}}^{(m)}(\rho) = \inf_{Z\in\mathbb{H}} \mathrm{Tr}(\rho - Z)_+ + \frac{1}{m}\Gamma_{\mathcal{Q}}(Z). \tag{11}$$

Proof. The definition of $G_{\mathcal{Q}}^{(m)}$ imposes that $W \in (-\mathbb{H}_+)^\circ \cap \mathbb{D}^\circ \cap (m\mathcal{Q})^\circ = [(-\mathbb{H}_+) \cup \mathbb{D} \cup m\,\mathrm{conv}(\mathcal{Q})]^\circ$, from which we get $G_{\mathcal{Q}}^{(m)} = \Gamma_{(-\mathbb{H}_+)\cup\mathbb{D}\cup m\,\mathrm{conv}(\mathcal{Q})}$. Since $\mathbb{D}$ and $\mathcal{Q}$ are compact and $\mathbb{H}_+$ is closed, we get [40, Theorem 16.4]

$$\begin{aligned} G_{\mathcal{Q}}^{(m)}(\rho) &= \inf\{\Gamma_{-\mathbb{H}_+}(X) + \Gamma_{\mathbb{D}}(Y) + \Gamma_{m\mathcal{Q}}(Z) |\; \rho = X + Y + Z\} \\ &= \inf\{\Gamma_{\mathbb{D}}(Y) + \Gamma_{m\mathcal{Q}}(Z) |\; \rho = X + Y + Z,\;\; X \in -\mathbb{H}_+\} \\ &= \inf\left\{\mathrm{Tr}\,Y + \frac{1}{m}\Gamma_{\mathcal{Q}}(Z) \;\middle|\; \rho - Z \preceq Y,\;\; Y \in \mathbb{H}_+\right\} = \inf \mathrm{Tr}(\rho - Z)_+ + \frac{1}{m}\Gamma_{\mathcal{Q}}(Z), \end{aligned} \tag{12}$$

where we have used that $\Gamma_{-\mathbb{H}_+}(X) = 0$ if $X \in -\mathbb{H}_+$ and ∞ otherwise, and $\Gamma_{\mathbb{D}}(Y) = \mathrm{Tr}\,Y$ for any positive semidefinite Y and ∞ if Y is not positive semidefinite. ■

Remark. The above formula effectively constraints the optimisation to be over $Z \in \mathcal{Q}^{**}$, since for any $Z \notin \mathcal{Q}^{**}$, we have $\Gamma_{\mathcal{Q}}(Z) = \infty$. In particular, if $\mathcal{Q}$ consists only of trace-one matrices, we can equivalently write

$$G_{\mathcal{Q}}^{(m)}(\rho) = \inf_{X \in \mathcal{Q}^{**}} \mathrm{Tr}\,(\rho - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \tag{13}$$

which reduces to the form in Proposition 1.

3. One-shot entanglement distillation

Denoting by Ψ_m the maximally entangled state $|\Psi_m\rangle = \sum_{i=1}^{m} \frac{1}{\sqrt{m}}|ii\rangle$, we consider the task of distiling the state Ψ_m under a given class of completely positive trace-preserving (CPTP) maps $\mathcal{O}$. The fidelity of distillation under $\mathcal{O}$ is defined by

$$F_{\mathcal{O}}(\rho,\, m) \coloneqq \sup_{\Lambda \in \mathcal{O}} \langle \Lambda(\rho),\, \Psi_m \rangle. \tag{14}$$

Here, without loss of generality we constrain ourselves to operations $\Lambda \in \mathcal{O}$ whose output dimension matches the dimension of Ψ_m in order to make the inner product well-defined; more general cases can be considered by suitably embedding Ψ_m or $\Lambda(\rho)$ in a larger space. The one-shot ε-error distillable entanglement is then defined as the maximum size of Ψ_m which can be obtained with the given class of operations within an error tolerance of ε, that is,

$$E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) \coloneqq \log \max\{m \in \mathbb{N} |\; F_{\mathcal{O}}(\rho,\, m) \geqslant 1 - \varepsilon\}. \tag{15}$$

In the asymptotic i.i.d. limit, distillable entanglement can then be expressed as

$$E_{d,\mathcal{O}}^{\infty}(\rho) = \limsup_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho^{\otimes n}). \tag{16}$$

To begin the general description of one-shot distillation, we will make explicit the connection between the quantifiers discussed earlier and distillation rates. The precise link will be established through the *hypothesis testing relative entropy* [42–44], defined as

$$D_H^{\varepsilon}(\rho||\, X) \coloneqq -\log\, \min\{\langle M,\, X\rangle |\; 0 \leqslant M \leqslant \mathbb{1},\;\; 1 - \langle M,\, \rho\rangle \leqslant \varepsilon\}, \tag{17}$$

where we have extended the standard definition (limited to positive semidefinite X) by taking $\log(x) = -\infty\; \forall\, x \leqslant 0$. This quantity characterises the fundamental task of quantum hypothesis testing [45, 46], where one is interested in distinguishing between two quantum states—ρ and σ—by performing a test measurement $\{M,\, \mathbb{1} - M\}$ where $0 \preceq M \preceq \mathbb{1}$. The probability of incorrectly accepting state σ as true (type-I error) is given by $\langle \mathbb{1} - M,\, \rho\rangle$, and the probability of incorrectly accepting state ρ as true (type-II error) is given by $\langle M,\, \sigma\rangle$. The entropy $D_H^{\varepsilon}(\rho||\sigma)$ then quantifies the minimum type-II error while constraining the type-I error to be no greater than ε. We note that it is not clear if such an operational understanding of D_H^{ε} can be obtained when X is not a positive semidefinite operator, but we will find it useful to consider the quantity D_H^{ε} regardless. Furthermore, we remark that for any operator X, $D_H^{\varepsilon}(\rho||X)$ is efficiently computable as a semidefinite program.

Let us first note a general correspondence between the hypothesis testing relative entropy and gauge functions, showing that D_H^{ε} minimised over a set of operators gives a suitably 'smoothed' gauge function.

Proposition 4. *Let $\mathcal{Q}$ be a closed set of Hermitian operators. Then*

$$\inf_{X \in\, \mathrm{conv}(\mathcal{Q})} D_H^{\varepsilon}(\rho\|X) = -\log \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \Gamma_{\mathcal{Q}^{\circ}}(W). \tag{18}$$

Proof. We have

$$\begin{aligned} -\log \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \Gamma_{\mathcal{Q}^{\circ}}(W) &= -\log \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \sup_{X \in \mathcal{Q}^{\circ\circ}} \langle X,\, W\rangle = -\log \sup_{X \in \mathcal{Q}^{\circ\circ}} \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \langle X,\, W\rangle \\ &= -\log \sup_{X \in\, \mathrm{conv}(\mathcal{Q})} \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \langle X,\, W\rangle = \inf_{X \in \mathrm{conv}(\mathcal{Q})} -\log \inf_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \langle X,\, W\rangle \\ &= \inf_{X \in\, \mathrm{conv}(\mathcal{Q})} D_H^{\varepsilon}(\rho\|X), \end{aligned} \tag{19}$$

where the second equality follows by Sion's minimax Theorem, since the sets $\{W |\; \langle \rho,\, W\rangle \geqslant 1 - \varepsilon,\;\; 0 \preceq W \preceq \mathbb{1}\}$ and $\mathcal{Q}^{\circ\circ}$ are both convex and the former is compact. We have replaced the optimisation over $\mathcal{Q}^{\circ\circ}$ with an optimisation over $\mathrm{conv}(\mathcal{Q})$ without loss of generality, since the problem has the same optimal value in both

cases—either there exists an $X \in \mathrm{conv}(\mathcal{Q})$ such that $D_H^\varepsilon(\rho\|X) < \infty$, or we have $D_H^\varepsilon(\rho\|X) = D_H^\varepsilon(\rho\|0) = \infty$ for all $X \in \mathcal{Q}^{\circ\circ}$. ■

The above can be directly applied in the context of entanglement distillation. Specifically, if one can show that the fidelity of distillation under a given class of operations is given by $G_{\mathcal{Q}}^{(m)}$ for some set $\mathcal{Q}$, then the optimal rate of distillation can be computed exactly as the hypothesis testing entropy minimised over $\mathrm{conv}(\mathcal{Q})$. Although we leave open the question of when exactly a given class of operations leads to a fidelity of distillation of the form given by $G_{\mathcal{Q}}^{(m)}$, we will see below that this is a very common phenomenon among different classes of operations relevant to the resource theory of entanglement.

Formally, we have the following.

Theorem 5. *Let $\mathcal{O}$ be a class of CPTP operations, and $\mathcal{Q}$ a compact set of Hermitian operators. If a given state ρ satisfies*

$$F_{\mathcal{O}}(\rho, m) = G_{\mathcal{Q}}^{(m)}(\rho) \quad \forall\ m \in \mathbb{N}, \tag{20}$$

then

$$E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) = \left\lfloor \min_{X \in\, \mathrm{conv}(\mathcal{Q})} D_H^\varepsilon(\rho\|X) \right\rfloor_{\log}. \tag{21}$$

Remark. The Theorem includes in particular the case when $F_{\mathcal{O}}(\rho, m) = \frac{1}{m}(T_{\mathcal{S}}^{(m-1)}(\rho) + 1)$ for $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+\}$. However, it is more general than that—for example, $\mathcal{Q}$ can be the set $\mathsf{PPT}' = \{X \,|\, \|X^{T_B}\|_1 \leqslant 1\}$, in which case we recover a result of [20].

Proof. By assumption, we have

$$\begin{aligned} E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho) &= \log \max\left\{ m \in \mathbb{N} \;\middle|\; \langle \rho, W\rangle \geqslant 1-\varepsilon, \quad 0 \preceq W \preceq \mathbb{1}, \;\; W \in \frac{1}{m}\mathcal{Q}^\circ \right\} \\ &= \lfloor -\log \min\{k \in \mathbb{R} |\, \langle \rho, W\rangle \geqslant 1-\varepsilon, \;\; 0 \preceq W \preceq \mathbb{1}, \;\; W \in k\mathcal{Q}^\circ\} \rfloor_{\log} \\ &= \left\lfloor -\log \min_{\substack{\langle \rho, W\rangle \geqslant 1-\varepsilon \\ 0 \preceq W \preceq \mathbb{1}}} \Gamma_{\mathcal{Q}^\circ}(W) \right\rfloor_{\log} = \left\lfloor \min_{X \in\, \mathrm{conv}(\mathcal{Q})} D_H^\varepsilon(\rho\|X) \right\rfloor_{\log}, \end{aligned} \tag{22}$$

where the last equality follows from Proposition 4. ■

The application of the above result will allow us to employ the powerful framework of convex optimisation in the description of entanglement distillation.

3.1. PPT and PPT-preserving operations

One of the first relaxations of LOCC in the literature was the class of separable operations (SEP) [23, 47], corresponding to all quantum channels $\Lambda\colon AB \to A'B'$ whose Choi matrix is separable across the bipartition $AA'|BB'$. This set of maps has been shown to be strictly larger than LOCC [48], thus providing an upper bound on the capabilities of LOCC in distillation. However, the fact that the definition of SEP relies on the separability of the Choi matrix means that the set is not amenable to an efficient analytical characterisation, which then motivated the definitions of larger sets of operations. We begin with the investigation of several classes of such operations based on the set PPT.

The class of PPT operations, due to Rains [24, 25], is defined to consist of all CPTP maps $\Lambda\colon AB \to A'B'$ whose Choi matrix J_Λ satisfies $J_\Lambda^{T_{BB'}} \succcurlyeq 0$. In some works, a closely related class of 'PPT-preserving operations' has been considered [26, 49], motivated by the fact that $X \in \mathsf{PPT} \Rightarrow \Lambda(X) \in \mathsf{PPT}$ for any PPT operation Λ. Although the two classes have sometimes been claimed to be equal, it is not difficult to see that only imposing the PPT-preserving constraint leads to a strictly larger class of quantum channels—consider, for instance, the channel which swaps subsystems A and B—so it is in fact incorrect to use the names 'PPT' and 'PPT-preserving' interchangeably when referring to operations. Interestingly, however, the two sets of channels lead to exactly the same rates of one-shot entanglement distillation (as well as dilution), as we will shortly see explicitly.

More recently, the name 'PPT-preserving operations' was also used to denote operations which map any PPT state to a PPT state, in the sense that $\sigma \in \mathsf{PPT}_+ \Rightarrow \Lambda(\sigma) \in \mathsf{PPT}_+$ [50]. It is well-known that this leads to a strictly larger class of operations than Rains' PPT operations [51], although an accurate way of referring to the class of PPT operations could be *completely* PPT-preserving [52], since the condition $J_\Lambda^{T_{BB'}} \succcurlyeq 0$ ensures the preservation of positivity when the map acts on a part of a larger system, akin to completely positive maps.

For clarity, let us begin with the precise definitions.

Definition 6. A CPTP map $\Lambda\colon AB \to A'B'$ is called PPT if any one of the following equivalent conditions is satisfied [24].

(i) The Choi matrix J_Λ is PPT with respect to the partition $AA'|BB'$, i.e. $J_\Lambda^{T_{BB'}} \succcurlyeq 0$.

(ii) The map $T_{B'} \circ \Lambda \circ T_B$ is completely positive.

(iii) For any spaces C, D such that $d_C = d_A$ and $d_D = d_B$ it holds that
$$\sigma \in \mathsf{PPT}_+(AC|BD) \Rightarrow \Lambda \otimes \mathrm{id}(\sigma) \in \mathsf{PPT}_+(A'C|B'D)$$
where id is the identity channel.

We will use PPT to denote the set of all such maps.

Definition 7. A CPTP map $\Lambda\colon AB \to A'B'$ is called PPT_+-preserving if
$$\sigma \in \mathsf{PPT}_+ \Rightarrow \Lambda(\sigma) \in \mathsf{PPT}_+. \tag{23}$$
We will use PPTP_+ to denote the set of all such maps.

Definition 8. A CPTP map $\Lambda\colon AB \to A'B'$ is called PPT-preserving if either of the following equivalent conditions is satisfied:

(i) $X \in \mathsf{PPT} \Rightarrow \Lambda(X) \in \mathsf{PPT}$.

(ii) The map $T_{B'} \circ \Lambda \circ T_B$ is positive, i.e. $X \in \mathbb{H}_+ \Rightarrow T_{B'} \circ \Lambda \circ T_B(X) \in \mathbb{H}_+$.

We will use PPTP to denote the set of all such maps.

We now characterise the operational capabilities of the different sets of operations. Note that the fidelity of distillation under the class PPT has previously been obtained by Rains [25], and an explicit expression for the rate of distillation in terms of D_H^ε appeared more recently in [20]. Reference [26] considered the class PPTP in this context, but the capabilities of PPTP_+ have not been explicitly investigated before.

Theorem 9. *The fidelity of distillation under the classes of operations* PPT, PPTP, *and* PPTP_+ *is given by*
$$\begin{aligned} F_{\mathrm{PPTP}_+}(\rho, m) &= G^{(m)}_{\mathsf{PPT}_+}(\rho) \\ F_{\mathrm{PPT}}(\rho, m) = F_{\mathrm{PPTP}}(\rho, m) &= G^{(m)}_{\mathsf{PPT}'}(\rho), \end{aligned} \tag{24}$$
where $\mathsf{PPT}' = \mathrm{conv}(\mathsf{PPT} \cup -\mathsf{PPT}) = \{X \,|\, \|X^{T_B}\|_1 \leqslant 1\}$, *and hence the one-shot distillable entanglement can be expressed as*
$$E^{(1),\varepsilon}_{d,\mathrm{PPTP}_+}(\rho) = \left\lfloor \min_{\sigma \in \mathsf{PPT}_+} D_H^\varepsilon(\rho\|\sigma) \right\rfloor_{\log} \tag{25}$$
$$E^{(1),\varepsilon}_{d,\mathrm{PPT}}(\rho) = E^{(1),\varepsilon}_{d,\mathrm{PPTP}}(\rho) = \left\lfloor \min_{X \in \mathsf{PPT}'} D_H^\varepsilon(\rho\|X) \right\rfloor_{\log}. \tag{26}$$

Proof. Since Ψ_m is invariant under any unitary of the form $U \otimes U^*$, it is in particular invariant under the twirling $\mathcal{T}(\cdot) \coloneqq \int (U \otimes U^*) \cdot (U \otimes U^*)^\dagger \, \mathrm{d}U$, where the integration is performed with respect to the Haar measure of the unitary group. We can then without loss of generality consider only trace-preserving operations of the form $\Lambda = \mathcal{T} \circ \Lambda$, giving
$$\Lambda(Z) = \langle Z, X\rangle \Psi_m + \frac{\langle Z, \mathbb{1} - X\rangle}{m^2 - 1}(\mathbb{1} - \Psi_m) \tag{27}$$
as this is the most general form of an operator invariant under twirling [53]. Since $\Psi_m^{T_B} = \frac{1}{m}(P_m^+ - P_m^-)$, where P_m^+ (respectively, P_m^-) denote the projector onto the symmetric (antisymmetric) subspace, we have
$$\Lambda(Z)^{T_B} = \frac{P_m^+}{m+1}\left(\frac{\mathrm{Tr}\, Z}{m} + \langle Z, X\rangle\right) + \frac{P_m^-}{m-1}\left(\frac{\mathrm{Tr}\, Z}{m} - \langle Z, X\rangle\right). \tag{28}$$
Using the mutual orthogonality of $P_m^\pm$, we obtain the general conditions
$$\begin{aligned} \Lambda(Z) \succcurlyeq 0 &\iff \langle Z, X\rangle \geqslant 0 \\ \Lambda(Z)^{T_B} \succcurlyeq 0 &\iff -\frac{\mathrm{Tr}\, Z}{m} \leqslant \langle Z, X\rangle \leqslant \frac{\mathrm{Tr}\, Z}{m}. \end{aligned} \tag{29}$$
Noting in addition that the complete positivity of Λ imposes $0 \preceq X \preceq \mathbb{1}$, we can constrain the map Λ such that $\Lambda(\sigma) \in \mathsf{PPT}_+$ for any $\sigma \in \mathsf{PPT}_+$ to get

$$F_{\mathrm{PPTP}_+}(\rho,\, m) = \max\left\{\langle \rho,\, X\rangle \;\middle|\; 0 \preceq X \preceq \mathbb{1},\;\; \langle \sigma,\, X\rangle \leqslant \frac{1}{m}\;\forall\, \sigma \in \mathsf{PPT}_+\right\}, \tag{30}$$

and similarly, by imposing that $S \in \mathsf{PPT} \Rightarrow \Lambda(S) \in \mathsf{PPT}$ we have

$$\begin{aligned} F_{\mathrm{PPTP}}(\rho,\, m) &= \max\left\{\langle \rho,\, X\rangle \;\middle|\; 0 \preceq X \preceq \mathbb{1},\;\; |\langle S,\, X\rangle| \leqslant \frac{1}{m}\;\forall\, S \in \mathsf{PPT}\right\} \\ &= \max\left\{\langle \rho,\, X\rangle \;\middle|\; 0 \preceq X \preceq \mathbb{1},\;\; \langle S,\, X\rangle \leqslant \frac{1}{m}\;\forall\, S \in \mathsf{PPT}'\right\}. \end{aligned} \tag{31}$$

Noting that the Choi matrix of the map (27) is given by $J_\Lambda = X_{AB} \otimes \Psi_{m_{A'B'}} + \frac{1}{m^2-1}(\mathbb{1} - X)_{AB} \otimes (\mathbb{1} - \Psi_m)_{A'B'}$, an explicit computation yields

$$J_\Lambda^{T_{BB'}} \succcurlyeq 0 \iff -\frac{1}{m}\mathbb{1} \preceq X^{T_B} \preceq \frac{1}{m}\mathbb{1}, \tag{32}$$

which is precisely the condition $\langle S,\, X\rangle \leqslant \frac{1}{m}\;\forall\, S \in \mathsf{PPT}'$, yielding the equality between $F_{d,\mathrm{PPTP}}(\rho,\, m)$ and $F_{d,\mathrm{PPT}}(\rho,\, m)$.

Since the distillation fidelities F_{PPTP_+} and F_{PPTP} are precisely of the form G_{PPT_+} and $G_{\mathsf{PPT}'}$, respectively, the result follows by Theorem 5. ■

The Theorem establishes an operational equivalence between the sets of operations PPT and PPTP, although we stress again that in fact PPT $\subsetneq$ PPTP: in particular, the swap operation, defined as $\Lambda(|ij\rangle\langle kl|) = |ji\rangle\langle lk|$ in a basis and extended by linearity, trivially preserves the positivity of the partial transpose of any operator, while the partial transpose of the Choi matrix J_Λ can be verified to be non-positive. This can be understood by noting that the swap operation does not preserve PPT states when acting only on a part of a larger system—indeed, if Alice and Bob each possess a singlet and exchange only half of it, they will have generated (maximal) entanglement. Notice also that we have explicitly shown a difference between the distillation rates of PPTP and PPTP_+, thus immediately implying that PPTP $\subsetneq$ PPTP_+.

In addition, we recall an argument in [50] which investigated a gap between PPT and PPTP_+ operations by showing that the negativity (a known monotone under PPT [54]) can increase under PPTP_+. This argument no longer applies to PPTP—the negativity can be expressed as a robustness-type quantifier with respect to the set PPT [54] and it follows straightforwardly that this is a strong monotone under PPTP [41]. The gap between PPTP and PPT is therefore much more subtle.

Although it is not easy to characterise the asymptotic rates of distillation under PPT and PPTP maps, we have the following characterisation of distillable entanglement under PPTP_+, thus establishing a limit on the asymptotic performance of PPT and PPTP (see also [55]).

Corollary 10. *The asymptotic distillable entanglement under* PPTP_+ *is given by the regularised relative entropy of entanglement with respect to the set* PPT_+,

$$E_{d,\mathrm{PPTP}_+}^{\infty}(\rho) = E_{R,\mathsf{PPT}_+}^{\infty}(\rho) \coloneqq \lim_{n\to\infty} \min_{\sigma\in\mathsf{PPT}_+} \frac{1}{n} D(\rho^{\otimes n}\|\sigma) \tag{33}$$

with D denoting the quantum relative entropy.

Proof. Follows directly from the generalised quantum Stein's Lemma [56], which shows precisely that the regularisation of the hypothesis testing relative entropy $\min_{\sigma\in\mathsf{PPT}_+} D_H^\varepsilon(\rho\|\sigma)$ in the asymptotic limit with ε going to 0 is given by the regularised relative entropy. ■

We remark that, although $E_{d,\mathrm{PPTP}_+}^{\infty}(\rho)$ is not known in general, it has been computed exactly for classes of all orthogonally invariant states (including isotropic and Werner states) [57], and it has been shown that there exist states such that $E_{d,\mathrm{PPT}}^{\infty}(\rho) < E_{d,\mathrm{PPTP}_+}^{\infty}(\rho)$ [58].

3.1.1. Rains set and distillation

The above Corollary in particular gives an operational interpretation to the regularised relative entropy $E_{R,\mathsf{PPT}_+}^{\infty}$, introduced first as a bound for distillable entanglement in [24]. One can then wonder whether similar operational interpretation can be given to other asymptotic quantities in entanglement distillation theory. We will show that it is indeed the case for one of the most fundamental of such bounds, the regularised Rains bound [25, 33, 57], constituting the tightest known bound for the asymptotically distillable entanglement. It is defined as

$$E^{\infty}_{\mathrm{Rains}}(\rho) \coloneqq \lim_{n\to\infty} \min_{X\in \mathsf{PPT}'_+} \frac{1}{n} D(\rho^{\otimes n}\|X), \tag{34}$$

where $\mathsf{PPT}'_+ = \{X \succcurlyeq 0 | \|X^{T_B}\|_1 \leqslant 1\}$ is the so-called Rains set. To relate this quantity with the distillation of entanglement, we will define the class of Rains-preserving operations PPTP'_+ as all maps such that $X \in \mathsf{PPT}'_+ \Rightarrow \Lambda(X) \in \mathsf{PPT}'_+$. We then have the following.

Theorem 11. *The fidelity of distillation and one-shot distillable entanglement under Rains-preserving operations* PPTP'_+ *are given by*

$$\begin{aligned} F_{\mathrm{PPTP}'_+}(\rho, m) &= G^{(m)}_{\mathsf{PPT}'_+}(\rho) \\ E^{(1),\varepsilon}_{d,\mathrm{PPTP}'_+}(\rho) &= \left\lfloor \min_{X\in\mathsf{PPT}'_+} D^{\varepsilon}_H(\rho\|X) \right\rfloor_{\log}. \end{aligned} \tag{35}$$

Proof. The proof proceeds analogously to Theorem 9. The crucial step is to notice that for the isotropic operator $\Lambda(Z) = \langle Z, X\rangle \Psi_m + \frac{\langle Z, \mathbb{1}-X\rangle}{m^2-1}(\mathbb{1} - \Psi_m)$ we have

$$\|\Lambda(Z)^{T_B}\|_1 = \frac{m}{2}\left|\frac{\mathrm{Tr}\, Z}{m} + \langle Z, X\rangle\right| + \frac{m}{2}\left|\frac{\mathrm{Tr}\, Z}{m} - \langle Z, X\rangle\right|, \tag{36}$$

where we have used that $\frac{1}{m+1}\|P^+_m\|_1 = \frac{1}{m-1}\|P^-_m\|_1 = \frac{m}{2}$ and that $P^{\pm}_m$ are mutually orthogonal projections. For any $Z \in \mathsf{PPT}'_+$, it is then easy to verify that we have $\|\Lambda(Z)^{T_B}\|_1 \leqslant 1 \iff \langle Z, X\rangle \leqslant \frac{1}{m}$. This gives

$$F_{\mathrm{PPTP}'_+}(\rho, m) = \max\left\{\langle \rho, X\rangle \;\middle|\; 0 \preceq X \preceq \mathbb{1},\ \langle Z, X\rangle \leqslant \frac{1}{m}\ \forall Z \in \mathsf{PPT}'_+\right\} = G^{(m)}_{\mathsf{PPT}'_+}(\rho). \tag{37}$$

The statement about $E^{(1),\varepsilon}_{d,\mathrm{PPTP}'_+}$ then follows directly from Theorem 5. ∎

Once again, an application of the generalised quantum Stein's Lemma [56] then gives

$$E^{\infty}_{d,\mathrm{PPTP}'_+}(\rho) = E^{\infty}_{\mathrm{Rains}}(\rho), \tag{38}$$

which establishes an explicit operational interpretation of the regularised Rains bound as the asymptotic rate of entanglement distillation under the class of Rains-preserving operations. Noting that $\mathrm{PPTP}'_+ \supseteq \mathrm{PPTP}$ by definition, we recover the result that $E^{\infty}_{\mathrm{Rains}}$ upper bounds the asymptotic distillable entanglement under PPT operations [25]. It is interesting to conjecture that we have equality between $E^{\infty}_{d,\mathrm{PPT}}$ and $E^{\infty}_{d,\mathrm{PPTP}'_+}$ (see [20]), but we were not able to establish this.

To obtain a tighter bound on distillable entanglement, one could then ask about distillation under operations which *completely* preserve the Rains set, in the sense that

$$X \in \mathsf{PPT}'_+(AC|BD) \Rightarrow \Lambda \otimes \mathrm{id}(X) \in \mathsf{PPT}'_+(A'C|B'D) \tag{39}$$

for some spaces C, D such that $d_C = d_A$ and $d_D = d_B$. We will call any such channel completely Rains-preserving. In other words, a map $\Lambda\colon AB \to A'B'$ is completely Rains-preserving iff it is CPTP and

$$\|\Lambda \otimes \mathrm{id}(X)^{T_{B'D'}}\|_1 \leqslant 1 \quad \forall\ X\colon X \succcurlyeq 0,\ \|X^{T_{BD}}\|_1 \leqslant 1. \tag{40}$$

We will now show that these maps are precisely the set of PPT channels.

Theorem 12. *A quantum channel is PPT iff it is completely Rains-preserving.*

Proof. One direction is straightforward: if Λ is completely positive and completely Rains-preserving, then for any $\sigma \in \mathsf{PPT}_+(AC|BD)$ we necessarily have $\Lambda \otimes \mathrm{id}(\sigma) \in \mathsf{PPT}_+(A'C'|B'D')$ due to the fact that $\mathsf{PPT}_+ = \mathsf{PPT}'_+ \cap \mathbb{H}_+$. This means that Λ is completely PPT_+-preserving, i.e. PPT.

To see the opposite inclusion, define a 'PPT-diamond norm' of any map Γ as

$$\begin{aligned} \|\Gamma\|_{\blacklozenge} &\coloneqq \max\{\|\Gamma \otimes \mathrm{id}(X^{T_{BD}})\|_1 | \|X^{T_{BD}}\|_1 \leqslant 1,\ X \succcurlyeq 0\} \\ &= \max\{\|\Gamma \otimes \mathrm{id}(X)\|_1 | \|X\|_1 \leqslant 1,\ X^{T_{BD}} \succcurlyeq 0\}. \end{aligned}$$

Rewriting equation (40) one can see that Λ is completely Rains-preserving iff

$$\|T_{B'} \circ \Lambda \circ T_B\|_{\blacklozenge} \leqslant 1.$$

Notice then that, for any Hermiticity-preserving map Γ it holds that

$$\|\Gamma\|_{\blacklozenge} \leqslant \max\{\|\Gamma \otimes \mathrm{id}(X)\|_1 | \|X\|_1 \leqslant 1\} = \|\Gamma\|_{\diamond},$$

where $\|\cdot\|_{\diamond}$ is the diamond norm (completely bounded trace norm) [59, 60]. Since for any PPT channel Λ the map $T_{B'} \circ \Lambda \circ T_B$ is CPTP, we have that any PPT channel satisfies $\|T_{B'} \circ \Lambda \circ T_B\|_{\diamond} = 1$ [60] and therefore is completely Rains-preserving. ■

The above result establishes an operational connection between the sets PPTP_+ and PPTP'_+, showing that their 'completely preserving' variants reduce to the same set of operations (PPT).

3.2. Pure-state distillation and separability-preserving operations

The class of separability-preserving operations SEPP is defined as all CPTP maps Λ such that $\sigma \in \mathsf{SEP} \Rightarrow \Lambda(\sigma) \in \mathsf{SEP}$, that is, as the maximal class of free (non-entangling) operations in the resource theory of entanglement. Notice that this class does not *completely* preserve separability, in the sense that it could generate entanglement if applied to a part of a larger system; if such complete preservation is imposed, we instead recover the class of separable operations. The inclusions between the different classes of operations are shown in figure 1.

The fidelity of distillation under SEPP was first derived in [27], and can be used to characterise the distillable entanglement as follows.

Lemma 13 ([18]). *It holds that* $F_{\mathrm{SEPP}}(\rho, m) = G^{(m)}_{\mathsf{SEP}}(\rho) = \frac{1}{m}(T^{(m-1)}_{\mathsf{SEP}}(\rho) + 1)$, *and hence*

$$E^{(1),\varepsilon}_{d,\mathrm{SEPP}}(\rho) = \left\lfloor \min_{\sigma \in \mathsf{SEP}} D^{\varepsilon}_H(\rho\|\sigma) \right\rfloor_{\log}. \tag{41}$$

Proof. Follows in exactly the same way as the proof of Theorem 9, since isotropic states of the form

$$\langle Z, X\rangle \Psi_m + \frac{\langle Z, \mathbb{1} - X\rangle}{m^2 - 1}(\mathbb{1} - \Psi_m) \tag{42}$$

are separable if and only if they are PPT [53]. ■

By the inclusion $\mathsf{SEP} \subseteq \mathsf{PPT}_+ \subset \mathsf{PPT}'$, we immediately have that

$$F_{\mathrm{SEPP}}(\rho, m) \geqslant F_{\mathrm{PPTP}_+}(\rho, m) \geqslant F_{\mathrm{PPT}}(\rho, m) \tag{43}$$

thus establishing a hierarchy of rates of distillation between the operations SEPP, PPTP_+, and PPT. Notice that this does not follow from their definition, as there is no inclusion between the sets of maps SEPP and PPT, nor between SEPP and PPTP_+.

Crucially, for any pure state, the fidelity of distillation can be computed exactly. To establish this result, we will employ the so-called *m-distillation norm*, introduced in [61] as

$$\||x\rangle\|_{[m]} := \min_{|x\rangle = |y\rangle + |z\rangle} \||y\rangle\|_{\ell_1} + \sqrt{m}\ \||z\rangle\|_{\ell_2} = \max\{|\langle x|w\rangle|\ |\ \||w\rangle\|_{\ell_\infty} \leqslant 1,\ \ \||w\rangle\|_{\ell_2} \leqslant \sqrt{m}\} \tag{44}$$

for any vector $|x\rangle \in \mathbb{C}^d$. One can immediately notice from the inequality $\|\cdot\|_{\ell_2} \leqslant \|\cdot\|_{\ell_1} \leqslant \sqrt{d}\ \|\cdot\|_{\ell_2}$ that we have $\||x\rangle\|_{[1]} = \||x\rangle\|_{\ell_2}$ and $\||x\rangle\|_{[d]} = \||x\rangle\|_{\ell_1}$. Notably, for any normalised vector $|x\rangle \in \mathbb{C}^d$ and any integer $m \in \{1,\ldots,d\}$, the norm admits an exact expression as [61]

$$\||x\rangle\|_{[m]} = \||x^{\downarrow}_{1:m-k^\star}\rangle\|_{\ell_1} + \sqrt{k^\star}\ \||x^{\downarrow}_{m-k^\star+1:d}\rangle\|_{\ell_2},$$

where $|x^{\downarrow}_{1:k}\rangle$ denotes the vector consisting of the k largest (by magnitude) coefficients of $|x\rangle$, analogously $|x^{\downarrow}_{k+1:d}\rangle$ denotes the vector of the $d - k$ smallest coefficients of $|x\rangle$ with $|x^{\downarrow}_{1:0}\rangle$ being the zero vector, and we define

$$k^\star := \arg\min_{1\leqslant k\leqslant m} \frac{1}{k}\ \||x^{\downarrow}_{m-k+1:d}\rangle\|^2_{\ell_2}. \tag{45}$$

We stress that the computation of $\||x\rangle\|_{[m]}$ is thus reduced to evaluating $m - 1$ inequalities.

We will now use $|\xi_\psi\rangle \in \mathbb{R}^d$ to denote the vector of Schmidt coefficients of a pure state $|\psi\rangle$, in the sense that $|\xi_\psi\rangle = (\alpha_1,\ldots,\alpha_d)^T$ where $|\psi\rangle = \sum_i \alpha_i |i\rangle_A |i\rangle_B$ for some orthonormal bases $\{|i_A\rangle\}$, $\{|i\rangle_B\}$. Employing the m-distillation norm, we then have the following.

Theorem 14. *For any $m \geqslant 1$, it holds that*

$$T_{\mathsf{SEP}}^{(m-1)}(\psi) = \| |\xi_\psi\rangle \|_{[m]}^2 - 1, \tag{46}$$

and in particular $F_{\mathrm{SEPP}}(\psi, m) = \frac{1}{m} \| |\xi_\psi\rangle \|_{[m]}^2$.

Proof. To begin, notice that with a simple rearrangement of terms $T_{\mathsf{SEP}}^{(m-1)}$ can be written as

$$\max\{\langle \rho, W\rangle \mid 0 \preceq W \preceq m\mathbb{1},\ W \in \mathsf{SEP}^\circ\} = \max\left\{\langle \rho, W\rangle \,\middle|\, W \succcurlyeq 0,\ W \in \left(\mathsf{SEP} \cup \frac{1}{m}\mathbb{D}\right)^\circ\right\}, \tag{47}$$

where we used that $(\mathcal{C} \cup \mathcal{D})^\circ = \mathcal{C}^\circ \cap \mathcal{D}^\circ$ for convex and closed sets. The set $\mathrm{conv}(\mathsf{SEP} \cup \frac{1}{m}\mathbb{D}) =: \mathcal{Q}_m$ can be noticed to be the convex hull of rank-one terms as $\mathcal{Q}_m = \mathrm{conv}\{|x\rangle\langle x| \,|\, |x\rangle \in \mathcal{V} \cup \mathcal{N}_m\}$ where $\mathcal{V} = \{|\phi\rangle_A \otimes |\eta\rangle_B\}$ is the set of all normalised product state vectors and $\mathcal{N}_m := \{|x\rangle \,|\, \| |x\rangle \|_{\ell_2} = 1/\sqrt{m}\}$.

By Theorem 10 in [41], for any pure state $|\psi\rangle$ we then have[6]

$$\begin{aligned} T_{\mathsf{SEP}}^{(m-1)}(\psi) + 1 &= \max\{\langle\psi|W|\psi\rangle \mid W \succcurlyeq 0,\ W \in \mathcal{Q}_m^\circ\} \\ &= \max\{|\langle\psi|w\rangle|^2 \mid |w\rangle \in (V \cup N_m)^\circ\} = \Gamma_{V\cup N_m}(|\psi\rangle)^2 \end{aligned} \tag{48}$$

which means that the value of $T_{\mathsf{SEP}}^{(m-1)}$ will be given by a corresponding gauge function $\Gamma_{\mathcal{V}\cup\mathcal{N}_m}(|\psi\rangle)^2$ defined at the level of the underlying Hilbert space, instead of the whole space of Hermitian operators. Since $\mathcal{V}$ and $\mathcal{N}_m$ are both compact sets, by standard results in convex analysis (see e.g. [40], 16.4.1 and 15.1.2), this gauge can be obtained as

$$\Gamma_{\mathcal{V}\cup\mathcal{N}_m}(|\psi\rangle) = \min_{|\psi\rangle = |x\rangle + |y\rangle} \Gamma_{\mathcal{V}}(|x\rangle) + \Gamma_{\mathcal{N}_m}(|y\rangle). \tag{49}$$

Now, for any vector $|x\rangle$ we have $\Gamma_{\mathcal{N}_m}(|x\rangle) = \sqrt{m}\ \|x\|_{\ell_2} = \sqrt{m}\ \| |\xi_x\rangle \|_{\ell_2}$ and it is known that $\Gamma_{\mathcal{V}}(|x\rangle)$ can be computed as $\| |\xi_x\rangle \|_{\ell_1}$ (see e.g. [62, 63]). By optimising over vectors $|x\rangle$, $|y\rangle$ in the Schmidt basis of $|\psi\rangle$ only, the problem reduces to the m-distillation norm of the Schmidt vector $|\xi_\psi\rangle$, and we thus have $T_{\mathsf{SEP}}^{(m-1)}(\psi) \leqslant \| |\xi_\psi\rangle \|_{[m]}^2 - 1$.

To show the opposite inequality, we use the fact that $(\mathcal{V} \cup \mathcal{N}_m)^\circ = \mathcal{V}^\circ \cap \mathcal{N}_m^\circ$ to write the gauge $\Gamma_{\mathcal{V}\cup\mathcal{N}_m}$ in its dual form as

$$T_{\mathsf{SEP}}^{(m-1)}(\psi) + 1 = \max\{|\langle\psi|x\rangle|^2 \mid \Gamma_{N_m}^\circ(|x\rangle) \leqslant 1,\ \Gamma_V^\circ(|x\rangle) \leqslant 1\}, \tag{50}$$

where we have, for any $|x\rangle$, $\Gamma_{\mathcal{V}}^\circ(|x\rangle) = \| |\xi_x\rangle \|_{\ell_\infty}$ [64]. By optimising over all vectors $|x\rangle$ in the Schmidt basis of $|\psi\rangle$, we recover again the m-distillation norm of $|\xi_\psi\rangle$ and the result follows. ∎

From the above, we then have the fidelity of distillation under SEPP of a pure state $|\psi\rangle$ as $\frac{1}{m} \| |\xi_\psi\rangle \|_{[m]}^2$. Crucially, the m-distillation norm can be closely connected with the concept of majorisation, allowing us to relate it to the optimal fidelity of pure-state distillation under LOCC and one-way LOCC (1–LOCC), which was previously considered in [65]. We will now rederive the exact expression for the fidelity of pure-state distillation under LOCC in terms of the m-distillation norm, and in particular establish an operational equivalence between all relevant sets of operations in the distillation of entanglement from pure states.

Theorem 15. *For any pure state $|\psi\rangle$, any integer $m \geqslant 1$, and any set of operations $\mathcal{O} \in \{1\text{–LOCC}, \mathrm{LOCC}, \mathrm{PPT}, \mathrm{PPTP}_+', \mathrm{PPTP}_+, \mathrm{SEPP}\}$, the fidelity of distillation is given by*

$$F_{\mathcal{O}}(\psi, m) = \frac{1}{m} \| |\xi_\psi\rangle \|_{[m]}^2. \tag{51}$$

Proof. We begin by recalling that the m-distillation norm of $|\xi_\psi\rangle := (\alpha_1, \ldots, \alpha_d)^T$ can be computed as

$$\| |\xi_\psi\rangle \|_{[m]} = \| |\alpha_{1:m-k^\star}^\downarrow\rangle \|_{\ell_1} + \sqrt{k^\star}\ \| |\alpha_{m-k^\star+1:d}^\downarrow\rangle \|_{\ell_2}, \tag{52}$$

with $k^\star := \arg\min_{1\leqslant k\leqslant m} \frac{\| |\alpha_{m-k+1:d}^\downarrow\rangle \|_{\ell_2}^2}{k}$.

Now, it is well-known that the deterministic transformation from $|\psi\rangle$ to another pure state $|\eta\rangle$ is possible with (either one-way and two-way) LOCC if and only if the Schmidt vector $|\xi_\psi\rangle = (\alpha_1, \ldots, \alpha_d)^T$ is majorised by the Schmidt vector $|\xi_\eta\rangle := (\beta_1, \ldots, \beta_d)^T$ [66], that is

[6] Strictly speaking, [41, Theorem 10] is obtained for sets of *normalised* operators; it is easy to notice, however, that the proof does not rely on normalisation and the Theorem applies in full generality also for unnormalised sets of operators, such as $\mathcal{Q}_m$ in our proof.

$$\sum_{i=1}^{k} \alpha_i^2 \leqslant \sum_{i=1}^{k} \beta_i^2 \quad \forall\, k \in \{1,\ldots,d\}, \tag{53}$$

where we have assumed without loss of generality that the Schmidt coefficients are given in non-increasing order. Let us then define the ansatz

$$|\eta\rangle = \sum_{i=1}^{m-k^\star} \alpha_i |ii\rangle + \sum_{i=m-k^\star+1}^{m} \frac{\| |\alpha^{\downarrow}_{m-k^\star+1:d}\rangle \|_{\ell_2}}{\sqrt{k^\star}} |ii\rangle \tag{54}$$

expressed in the Schmidt basis of $|\Psi_m\rangle$, where $k^\star$ is defined as above. To see that the Schmidt coefficients of $|\eta\rangle$ majorise the ones of $|\psi\rangle$, let us assume that $k^\star > 1$ (as otherwise the desired relation is trivial) and consider the following chain of equivalent inequalities:

$$\begin{aligned}
&\frac{\| |\, \alpha^{\downarrow}_{m-k^\star+1:d}\rangle \|^2_{\ell_2}}{k^\star} \leqslant \frac{\| |\, \alpha^{\downarrow}_{m-k^\star+2:d}\rangle \|^2_{\ell_2}}{k^\star - 1} \\
&\alpha^2_{m-k^\star+1} + \| |\, \alpha^{\downarrow}_{m-k^\star+2:d}\rangle \|^2_{\ell_2} \leqslant \frac{k^\star}{k^\star - 1} \| |\, \alpha^{\downarrow}_{m-k^\star+2:d}\rangle \|^2_{\ell_2} \\
&\alpha^2_{m-k^\star+1} \leqslant \frac{1}{k^\star - 1}(1 - \| |\, \alpha^{\downarrow}_{1:m-k^\star+1}\rangle \|^2_{\ell_2}) \\
&k^\star \alpha^2_{m-k^\star+1} \leqslant (1 - \| |\, \alpha^{\downarrow}_{1:m-k^\star}\rangle \|^2_{\ell_2}) \\
&\alpha^2_{m-k^\star+1} \leqslant \frac{\| |\, \alpha^{\downarrow}_{m-k^\star+1:d}\rangle \|^2_{\ell_2}}{k^\star},
\end{aligned} \tag{55}$$

where the first line follows by definition of $k^\star$, and in the third and fifth lines we have used the fact that $|\psi\rangle$ is a normalised pure state. Thus, we have

$$F_{1\text{–LOCC}}(\psi, m) \geqslant |\langle \Psi_m | \eta \rangle|^2 = \frac{1}{m}(\| \alpha^{\downarrow}_{1:m-k^\star} \|_{\ell_1} + \sqrt{k^\star}\, \| \alpha^{\downarrow}_{m-k^\star+1:d} \|_{\ell_2})^2 = \frac{1}{m}\, \| |\xi_\psi\rangle \|^2_{[m]}\,. \tag{56}$$

On the other hand, since 1–LOCC $\subseteq$ SEPP, we have

$$F_{1\text{–LOCC}}(\psi, m) \leqslant F_{\text{SEPP}}(\psi, m) = \frac{1}{m}(T^{(m-1)}_{\text{SEP}}(\psi) + 1) = \frac{1}{m}\, \| |\xi_\psi\rangle \|^2_{[m]} \tag{57}$$

by Theorem 14, which concludes the proof. ■

The exact correspondence is rather surprising. The operations PPT and SEPP are known to be much more powerful than LOCC in general—indeed, SEPP exhibit no bound entanglement whatsoever [27], and even in the manipulation of pure states PPT operations can, for instance, arbitrarily increase the Schmidt rank (number of Schmidt coefficients) of a pure state [52, 67], which cannot increase whatsoever under LOCC or SEP [12]. The result then shows that even such large sets cannot outperform one-way LOCC in entanglement distillation from pure states, even in the one-shot setting.

Using the above expression, we can furthermore show that the computation of $E^{(1),\varepsilon}_{d,\mathcal{O}}$ belongs to a class of efficiently solvable optimisation problems known as quadratically-constrained linear programs [68].

Corollary 16. *For any set of operations* $\mathcal{O} \in \{1\text{–LOCC}, \text{LOCC}, \text{PPT}, \text{PPTP}'_+, \text{PPTP}_+, \text{SEPP}\}$, *the one-shot distillable entanglement of a pure state can be expressed exactly as the optimal value of the convex quadratically-constrained linear program*

$$E^{(1),\varepsilon}_{d,\mathcal{O}}(\psi) = \lfloor -\log \min\{ \| |\omega\rangle \|^2_{\ell_\infty} \mid \langle \xi_\psi | \omega \rangle \geqslant \sqrt{1-\varepsilon},\ \| |\omega\rangle \|_{\ell_2} \leqslant 1,\ |\omega\rangle \in \mathbb{R}^d_+ \} \rfloor_{\log}\,. \tag{58}$$

Proof. The dual form of the m-distillation norm, which we recall here as

$$\| |x\rangle \|_{[m]} = \max\{ |\langle x | w\rangle| \mid \| |w\rangle \|_{\ell_\infty} \leqslant 1,\ \| |w\rangle \|_{\ell_2} \leqslant \sqrt{m}\, \}, \tag{59}$$

gives

$$\begin{aligned}
E^{(1),\varepsilon}_{d,\mathcal{O}}(\psi) &= \log \max \left\{ m \in \mathbb{N} \;\middle|\; \frac{1}{m}\, \| |\xi_\psi\rangle \|^2_{[m]} \geqslant 1 - \varepsilon \right\} \\
&= \lfloor -\log \min\{ \| |\omega\rangle \|^2_{\ell_\infty} \mid |\langle \xi_\psi | \omega \rangle|^2 \geqslant 1 - \varepsilon,\ \| |\omega\rangle \|_{\ell_2} \leqslant 1 \} \rfloor_{\log}
\end{aligned} \tag{60}$$

and we conclude by noting that it suffices to optimise over vectors with non-negative coefficients since $|\xi_\psi\rangle$ is also non-negative. ■

The above result can be compared with the bounds obtained for LOCC distillable entanglement in [16, 17], and in fact we have tightened the bounds to an exact expression for the one-shot distillable entanglement:

$$E_{d,\mathcal{O}}^{(1),\varepsilon}(\psi) = \lfloor -\log \min\{\|\rho_B\|_\infty \mid F(\rho_B, \mathrm{Tr}_A(\psi)) \geqslant 1 - \varepsilon\} \rfloor_{\log}. \tag{61}$$

The Theorem also leads to an interesting characterisation of the m-distillation norm in two different ways. Notice that the proof of Theorem 14 in fact shows that the m-distillation norm of the Schmidt vector $|\xi_x\rangle$ of a vector $|x\rangle$ can be equivalently written as a norm at the level of the vector $|x\rangle$ itself:

$$\begin{aligned} \||\xi_x\rangle\|_{[m]} &= \min_{|x\rangle=|y\rangle+|z\rangle} \||\xi_y\rangle\|_{\ell_1} + \sqrt{m}\ \||\xi_z\rangle\|_{\ell_2} \\ &= \max\{|\langle x|w\rangle|\,|\,\||\xi_w\rangle\|_{\ell_\infty} \leqslant 1,\ \ \||\xi_w\rangle\|_{\ell_2} \leqslant \sqrt{m}\}. \end{aligned} \tag{62}$$

Writing $|\psi\rangle \xrightarrow{\text{LOCC}} |\eta\rangle$ to denote that the pure state transformation is possible with LOCC, we then have

$$\begin{aligned} \||\xi_\psi\rangle\|_{[m]} &= \sqrt{m}\ \max\{|\langle\Psi_m|\ \eta\rangle|\,|\,|\psi\rangle \xrightarrow{\text{LOCC}} |\eta\rangle\} \\ &= \sqrt{m}\ \max\{|\langle\psi|\ \eta\rangle|\,|\,|\eta\rangle \xrightarrow{\text{LOCC}} |\Psi_m\rangle\}, \end{aligned} \tag{63}$$

where the maximisation is over normalised state vectors $|\eta\rangle$, and the second line is precisely equation (62).

As a straightforward Corollary of the results above, we can establish the value of the quantifiers $G_{\mathcal{S}}^{(m)}$ for several sets other than SEP.

Corollary 17. *For any* $1 \leqslant m \leqslant d$, *any pure state* $|\psi\rangle$, *and any* $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+, \mathsf{PPT}'\}$ *we have*

$$G_{\mathcal{S}}^{(m)}(|\psi\rangle\langle\psi|) = \frac{1}{m}\||\xi_\psi\rangle\|_{[m]}^2. \tag{64}$$

Going beyond pure states, combined with Proposition 2 the expression for $F_{\mathrm{SEPP}}(\rho, d)$ gives a direct operational interpretation to the generalised robustness of entanglement $R_{\mathsf{SEP}}^{\mathbb{D}}$ by showing that

$$R_{\mathsf{SEP}}^{\mathbb{D}}(\rho) = dF_{\mathrm{SEPP}}(\rho, d) - 1 \tag{65}$$

for any state ρ, and in fact by Theorem 9 also for the class PPTP_+ we have the relation $R_{\mathsf{PPT}_+}^{\mathbb{D}}(\rho) = dF_{\mathrm{PPTP}_+}(\rho, d) - 1$. This complements the known operational applications of this quantity [18, 69, 70]. Note also that all of the other measures in the family $T_{\mathsf{SEP}}^{(m)}$ and $T_{\mathsf{PPT}_+}^{(m)}$, including the modified trace distance quantifiers, are given similar interpretations.

An important use of the fidelity of distillation $F_{\mathrm{LOCC}}(\rho, m)$ in the particular case $m = d$ is as the fidelity of teleportation, that is, the best average fidelity one can achieve in the task of quantum teleportation by employing an LOCC protocol on the state ρ [71]. Notably, in [72] it was then shown that for $d_A = d_B = 2$, we have $F_{\mathrm{LOCC}}(\rho, d) = F_{\mathrm{PPT}}(\rho, d)$ showing that even PPT protocols (or SEPP protocols, by Lemma 13) cannot enhance the fidelity of teleportation of the given state. By Theorem 15, we know that this relation extends to all pure states in all dimensions; that is, $F_{\mathrm{LOCC}}(\psi, d) = F_{\mathrm{PPT}}(\psi, d) = F_{\mathrm{SEPP}}(\psi, d)$.

Remark. In [35] (Proposition 9) it was claimed that the asymptotic distillable entanglement $E_{d,\mathrm{LOCC}}^{\infty}(\rho)$ of any state is upper bounded by $\log(T_{\mathsf{SEP}}^{(m)}(\rho) + 1)$ for any $m \geqslant 1$. This is clearly not true, as $T_{\mathsf{SEP}}^{(m)}(\rho) \leqslant m\ \forall\ \rho$ and the distillable entanglement obeys no such restriction. We have seen, however, that the quantifiers $T_{\mathsf{SEP}}^{(m)}(\rho)$ characterise exactly the *fidelity* of distillation.

As a side note, noticing the similarity between the distillation under PPT_+-preserving, Rains-preserving, and SEP-preserving operations, it might appear that the hypothesis testing relative entropy D_H^ε in general quantifies the rate of distillation under a set of operations which is defined to preserve a given set of operators. This claim is supported by recent independent results concerned with distillation in a class of general quantum resources [73], but it does not hold in full generality as the distillation under PPTP operations shows (see Theorem 9), and indeed also distillation in the resource theory of coherence [61] is a counterexample.

3.3. Isotropic states

Consider $d = d_A = d_B$ and define the isotropic states as

$$\rho_f = f\Psi_d + \frac{1-f}{d^2-1}(\mathbb{1} - \Psi_d) \tag{66}$$

with $0 \leqslant f \leqslant 1$. This class of states is particularly useful due to its strong symmetry, allowing for a much easier evaluation of their entanglement properties [53]. We then have the following result, showing the operational equivalence of all sets of channels from SEP to PPTP_+ and SEPP in distiling entanglement from isotropic states, and extending the known characterisation of isotropic state distillation under PPT operations considered in [25].

Theorem 18. *For any isotropic state ρ_f and any $\mathcal{O} \in \{\mathrm{SEP},\ \mathrm{PPT},\ \mathrm{PPTP}'_+,\ \mathrm{PPTP}_+,\ \mathrm{SEPP}\}$, it holds that*

$$F_{\mathcal{O}}(\rho_f,\, m) = \begin{cases} \dfrac{1}{m} & f \leqslant \dfrac{1}{d} \\ \dfrac{df-1}{d-1} + \dfrac{d(1-f)}{m(d-1)} & f \geqslant \dfrac{1}{d}. \end{cases} \tag{67}$$

Proof. Take $1 \leqslant m \leqslant d$. If $f \leqslant \frac{1}{d}$, then $\rho_f \in \mathsf{SEP}$, and we have $F_{\mathrm{SEP}}(\rho_f,\, m) = F_{\mathrm{PPT}}(\rho_f,\, m) = F_{\mathrm{SEPP}}(\rho_f,\, m) = \frac{1}{m}$; we will therefore assume that $f \geqslant \frac{1}{d}$ in the sequel.

Recall by our previous arguments that, due to twirling, we can limit ourselves to considering operations of the form

$$\Lambda_W(Z) = \langle Z,\, W\rangle \Psi_m + \frac{\langle Z,\, \mathbb{1} - W\rangle}{m^2-1}(\mathbb{1} - \Psi_m), \tag{68}$$

and the fidelity of distillation under a set $\mathcal{O}$ is then given by

$$F_{\mathcal{O}}(\rho_f,\, m) = \max\{\langle \rho_f,\, W\rangle | \Lambda_W \in \mathcal{O}\}. \tag{69}$$

Here, noting the invariance of ρ_f under twirling, we can twirl once more; in particular, $\langle \rho_f,\, W\rangle = \langle \mathcal{T}(\rho_f),\, \mathcal{T}(W)\rangle = \langle \rho_f,\, \mathcal{T}(W)\rangle\ \forall\, W$, so we can again limit the considered W to be of the form

$$W = \alpha \Psi_d + \beta \mathbb{1}, \tag{70}$$

where $0 \leqslant \beta \leqslant 1$, $0 \leqslant \alpha + \beta \leqslant 1$. The Choi operator of the corresponding map Λ_W is then of the form

$$J_{\Lambda_W} = \frac{\alpha m^2}{m^2-1}\Psi_{dm} + \frac{\beta m^2 - 1}{m^2-1}\mathbb{1}_{AB} \otimes \Psi_m - \frac{\alpha}{m^2-1}\Psi_d \otimes \mathbb{1}_{A'B'} + \frac{1-\beta}{m^2-1}\mathbb{1}_{ABA'B'}, \tag{71}$$

where $d_{A'} = d_{B'} = m$. By [74], Theorem 6, $J_{\Lambda_W} \in \mathsf{SEP}(AA' \colon BB')$ if and only if the following conditions are all satisfied:

$$d - m\alpha + dm\beta \geqslant 0,\ \ d + m\alpha - dm\beta \geqslant 0,\ \ d - m\alpha - dm\beta \geqslant 0,\ \ m\alpha + d^2 m\beta - d(\alpha+\beta) \geqslant 0. \tag{72}$$

Let us choose

$$W = \frac{d(m-1)}{(d-1)m}\Psi_d + \frac{d-m}{(d-1)m}\mathbb{1} \tag{73}$$

for which the inequalities (72) can be readily verified to hold. This gives

$$F_{\mathrm{SEP}}(\rho_f,\, m) \geqslant \langle \rho_f,\, W\rangle = \frac{df-1}{d-1} + \frac{d(1-f)}{m(d-1)}. \tag{74}$$

On the other hand, take

$$X = \frac{d(1-f)}{d^2-1}\left(\Psi_d + \frac{1}{d}\mathbb{1}\right). \tag{75}$$

It is known that $\Psi_d + \frac{1}{d}\mathbb{1} \in \mathsf{SEP}^{**}$ [34], which gives

$$F_{\mathrm{SEPP}}(\rho_f,\, m) \leqslant \mathrm{Tr}\,(\rho_f - X)_+ + \frac{1}{m}\mathrm{Tr}\,X = \frac{df-1}{d-1} + \frac{d(1-f)}{m(d-1)}, \tag{76}$$

where we used Lemma 13 together with the dual form of $G^{(m)}_{\mathsf{SEP}}$ from Proposition 3. ■

We remark that the above also gives a general way of lower-bounding the fidelity of distillation under separable operations of any state with a simple linear program, tight for all isotropic states:

$$F_{\mathrm{SEP}}(\rho,\, m) \geqslant \max\{\alpha\langle \rho,\, \Psi_d\rangle + \beta |\ 0 \leqslant \beta \leqslant 1,\ \ 0 \leqslant \alpha + \beta \leqslant 1,\ \ \text{equation(72)}\}. \tag{77}$$

Unsurprisingly, however, a numerical investigation reveals this bound to be rather ineffective beyond the set of isotropic states.

Note also that a general investigation of one- and multi-shot entanglement distillation from isotropic states under PPT operations as a linear program has been explored in [20, 25].

3.4. Maximally correlated states

Let us consider a bipartite system with $d = d_A = d_B$. A *maximally correlated state* is any state of the form $\rho_{\mathrm{mc}} = \sum_{i,j}\rho_{ij}|ii\rangle\langle jj|$ for some local orthonormal bases $\{|i\rangle\}$ [24]. The name for this class of states comes from the fact is that the two parties are guaranteed to obtain the same measurement results for any measurement in their local basis $\{|i\rangle\}$.

Notice that any maximally correlated state has a corresponding single-party state $\widetilde{\rho_{\mathrm{mc}}} \coloneqq \sum_{i,j}\rho_{ij}|i\rangle\langle j|$ with the same coefficients in an orthonormal basis $\{|i\rangle\}$. This led to comparisons between the manipulation of maximally correlated states and the resource theory of coherence, which studies the properties of superposition as a quantum resource [75, 76]. In particular, it has been conjectured in several works that the resource theory of coherence is equivalent to the resource theory of entanglement restricted to maximally correlated states [77, 78]. Although this conjecture is still unsolved in full generality, we know that many operational quantifiers such as the entanglement of formation, relative entropy of entanglement (and other Rényi entropy–based entanglement measures), and asymptotic distillable entanglement can be evaluated on maximally correlated states by quantifying the corresponding coherence quantifiers, typically significantly simpler to evaluate and satisfying useful properties such as additivity [77, 79]. Furthermore, an operational equivalence between transformations acting on $\widetilde{\rho_{\mathrm{mc}}}$ and LOCC operations acting on ρ_{mc} has been suggested, although so far this conjecture has been shown only in specific cases [77, 78].

To obtain a result allowing us to quantify the one-shot distillable entanglement of maximally correlated states, we will fix a choice of basis $\{|i\rangle\}_{i=1}^{d}$ for the single-party state $\widetilde{\rho_{\mathrm{mc}}}$, and use $\mathcal{I} \coloneqq \mathrm{conv}\{|i\rangle\langle i|\}_i$ to denote the set of all incoherent (diagonal) states in this basis. Furthermore, we define the subset of separable states $\mathsf{SEP}_{\mathrm{mc}} \coloneqq \mathrm{conv}\{|ii\rangle\langle ii|\}_i$ where $\{|ii\rangle\}_{i=1}^{d}$ is the maximally correlated basis of the state ρ_{mc}. We then get the following.

Theorem 19. *For any maximally correlated state, any $m \geqslant 1$, and any choice of operations $\mathcal{O} \in \{\mathrm{PPT}, \mathrm{PPTP}_+', \mathrm{PPTP}_+, \mathrm{SEPP}\}$ it holds that*

$$F_{\mathcal{O}}(\rho_{\mathrm{mc}}, m) = G_{\mathcal{I}}^{(m)}(\widetilde{\rho_{\mathrm{mc}}}). \tag{78}$$

Proof. Using Lemma 13 together with Proposition 3, we have

$$\begin{aligned} F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m) \leqslant F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m) &= \min_{X\in\mathsf{SEP}^{**}} \mathrm{Tr}(\rho_{\mathrm{mc}} - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \\ &\leqslant \min_{X\in\mathrm{SEP}_{\mathrm{mc}}^{**}} \mathrm{Tr}(\rho_{\mathrm{mc}} - X)_+ + \frac{1}{m}\mathrm{Tr}(X) \\ &= \min_{X\in\mathcal{I}^{**}} \mathrm{Tr}(\widetilde{\rho_{\mathrm{mc}}} - X)_+ + \frac{1}{m}\mathrm{Tr}(X) = G_{\mathcal{I}}^{(m)}(\widetilde{\rho_{\mathrm{mc}}}). \end{aligned} \tag{79}$$

On the other hand, let $W^\star = \sum_{i,j}W_{ij}|i\rangle\langle j|$ be the optimal solution to the dual problem of

$$G_I^{(m)}(\widetilde{\rho_{\mathrm{mc}}}) = \max\left\{\langle\widetilde{\rho_{\mathrm{mc}}}, W\rangle \,\middle|\, 0 \preceq W \preceq \mathbb{1},\ W \in \frac{1}{m}I^\circ\right\}. \tag{80}$$

Notice that $W^\star \in \frac{1}{m}\mathcal{I}^\circ$ is equivalent to $\max_i W_{ii} \leqslant \frac{1}{m}$. Consider then the matrix $W_{\mathrm{mc}} \coloneqq \sum_{i,j}W_{ij}|ii\rangle\langle jj|$, defined in the basis of ρ_{mc}. The eigenvalues of $W_{\mathrm{mc}}^{T_B}$ can be straightforwardly verified to be $\{W_{ii},\ \pm|W_{ij}|\}_{i,j=1}^{d}$. Positivity of $W^\star$ imposes that

$$\max_{i,j}|W_{ij}| \leqslant \max_i W_{ii} \tag{81}$$

from which it follows that

$$\Gamma_{\mathsf{PPT}'^\circ}(W_{\mathrm{mc}}) = \|W_{\mathrm{mc}}^{T_B}\|_\infty = \max_i W_{ii} \leqslant \frac{1}{m}. \tag{82}$$

This means that $W_{\mathrm{mc}} \in \frac{1}{m}\mathsf{PPT}'^\circ$, and so

$$\begin{aligned} F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m) = G_{\mathsf{PPT}'}^{(m)}(\rho_{\mathrm{mc}}) &= \max\left\{\langle\rho_{\mathrm{mc}}, W\rangle \,\middle|\, 0 \preceq W \preceq \mathbb{1},\ W \in \frac{1}{m}\mathsf{PPT}'^\circ\right\} \\ &\geqslant \langle\rho_{\mathrm{mc}}, W_{\mathrm{mc}}\rangle = G_{\mathcal{I}}^{(m)}(\widetilde{\rho_{\mathrm{mc}}}). \end{aligned} \tag{83}$$

■

Notice that $G_{\mathcal{I}}^{(m)}(\widetilde{\rho_{\mathrm{mc}}}) = \frac{1}{m}(T_{\mathcal{I}}^{(m-1)}(\widetilde{\rho_{\mathrm{mc}}}) + 1)$, where $T_{\mathcal{I}}^{(m-1)}$ have been considered as coherence measures in [61]. Further, using Theorem 5 we have that

$$E_{d,\mathcal{O}}^{(1),\varepsilon}(\rho_{\mathrm{mc}}) = \left\lfloor \min_{\sigma\in\mathcal{I}} D_H^\varepsilon(\widetilde{\rho_{\mathrm{mc}}}\|\sigma)\right\rfloor_{\log} = \left\lfloor -\log \min_{\substack{\langle\widetilde{\rho_{\mathrm{mc}}}, W\rangle \geqslant 1-\varepsilon \\ 0\leq W\leq\mathbb{1}}} \|\Delta(W)\|_\infty \right\rfloor_{\log}, \tag{84}$$

where $\Delta(\cdot) = \sum_i|i\rangle\langle i|\cdot|i\rangle\langle i|$ is the completely dephasing map. We stress that these optimisation problems are all efficiently computable as simple semidefinite programs [80], facilitating an efficient quantification of the fidelity as well as rates of one-shot distillation of all maximally correlated states.

Interestingly, in contrast to many other results which show an exact equality between operational quantities in the resource theory of coherence and the resource theory of entanglement of maximally correlated states, our

result above shows a slight discrepancy between the two resources: in particular, it is not difficult to find numerical examples of states such that

$$F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m) = G_{\mathcal{I}}^{(m)}(\rho) > G_{\mathcal{J}}^{(m)}(\rho) = F_{\mathrm{MIO}}(\rho, m) \tag{85}$$

$\forall\, m < d$, where MIO denotes the class of maximally incoherent operations in the resource theory of coherence, defined to be channels Λ such that $\sigma \in \mathcal{I} \Rightarrow \Lambda(\sigma) \in \mathcal{I}$, and $\mathcal{J}$ is the set of all unit-trace diagonal Hermitian operators (see [61] for the rightmost equality). Therefore, the one-shot distillable entanglement of a maximally correlated state under the largest set of free operations in the resource theory of entanglement (SEPP) can be strictly larger than the distillable coherence of the corresponding single-partite state under the largest set of free operations in the resource theory of coherence (MIO). This shows in particular that, in the distillation of entanglement from maximally correlated states under SEPP, it is not sufficient to consider operations whose output remains in the maximally correlated subspace —indeed, if this were the case, any such operation could always be mapped to a corresponding MIO operation, and the fidelities $F_{\mathrm{MIO}}(\rho, m)$ and $F_{\mathrm{SEPP}}(\rho_{\mathrm{mc}}, m)$ would be equal. This also motivates a rather curious conjecture that, should there exist a smaller class of operations for which it suffices to consider only maximally correlated output states, then it is plausible that $F_{\mathrm{LOCC}}(\rho_{\mathrm{mc}}, m) \leqslant F_{\mathrm{MIO}}(\rho, m) < F_{\mathrm{PPT}}(\rho_{\mathrm{mc}}, m)$ for general maximally correlated states. This could be surprising, as it is known that the gap between LOCC and PPT distillation of ρ_{mc} disappears at the asymptotic level [11, 81] or even when considering the second-order non-asymptotic expansion of the rate of distillation [20].

3.5. Assisted distillation

The setting of (environment-)assisted distillation of entanglement, considered first in [31, 32], has been studied in the non-asymptotic regime in [17]. It is based on a scenario in which the two parties A and B are assisted by a third party C who holds a purifying state of the system ρ_{AB}, i.e. such that the joint state is ψ_{ABC}, and aims to increase the entanglement distillable from ρ_{AB} by performing a measurement on their local system C and communicating its result classically to A and B. A particular property of this setting is that the optimal protocol always involves a rank-1 measurement on subsystem C [17], giving parties A and B access to arbitrary pure-state decompositions of the system ρ_{AB}. Specifically, the best achievable rate of distillation is given by

$$E_{A,\mathcal{O}}^{(1),\varepsilon}(\rho) \coloneqq \log \max\{m \in \mathbb{N} |\ F_{A,\mathcal{O}}(\rho, m) \geqslant 1 - \varepsilon\}, \tag{86}$$

where the fidelity of assisted distillation is the best average fidelity optimised over all decompositions, i.e.

$$F_{A,\mathcal{O}}(\rho, m) \coloneqq = \max\left\{ \left\langle \sum_i p_i \Lambda_i(|\psi_i\rangle\langle\psi_i|), \Psi_m \right\rangle \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|,\ \Lambda_i \in \mathcal{O}\ \forall\, i \right\}. \tag{87}$$

This in particular means that, having obtained the measurement result from party C, the distillation is performed from a *pure* state—therefore, employing our results in Theorem 15, we immediately obtain the result that the rate of assisted entanglement disitillation is the same under all sets of operations from 1–LOCC up to PPTP_+ and SEPP.

Additionally, the proofs of the main results of [17] can be significantly simplified by employing the formalism introduced in this work, in fact strengthening the one-shot characterisation of Theorems 1 and 2 of [17] and tightening the bounds derived therein. In particular, our pure state-results in Theorem 15 allow us to straightforwardly obtain the following.

Theorem 20. *For any* $\mathcal{O} \in \{1\text{–LOCC}, \mathrm{LOCC}, \mathrm{PPT}, \mathrm{PPTP}, \mathrm{PPTP}_+, \mathrm{SEPP}\}$, *the fidelity and one-shot rate of assisted distillation of any state are given by*

$$F_{A,\mathcal{O}}(\rho, m) = \max\{F(\rho, \omega) | \omega \in \mathcal{M}_m\}, \tag{88}$$

$$E_{A,\mathcal{O}}^{(1),\varepsilon}(\rho) = \lfloor -\log \min\{\vartheta(\omega) |\ \omega \in \mathbb{D},\ F(\rho, \omega) \geqslant 1 - \varepsilon\} \rfloor_{\log}, \tag{89}$$

where

$$\mathcal{M}_m \coloneqq \mathrm{conv}\left\{ |\phi\rangle\langle\phi| \;\middle|\; \||\phi\rangle\|_{\ell_2} = 1,\ \||\xi_\phi\rangle\|_{\ell_\infty} \leqslant \frac{1}{\sqrt{m}} \right\} \tag{90}$$

and

$$\vartheta(\omega) \coloneqq \min\left\{ \max_i \||\xi_{\psi_i}\rangle\|_{\ell_\infty}^2 \;\middle|\; \omega = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}. \tag{91}$$

Proof. The derivation follows the approach taken for quantum coherence in [82]. We begin by writing the fidelity of assisted distillation as

$$F_{A,\mathcal{O}}(\rho, m) = \max\left\{ \sum_i p_i F_{\mathcal{O}}(\psi_i, m) \;\middle|\; \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\} \tag{92}$$

with $F_{\mathcal{O}}$ denoting the fidelity of distillation as before. Notice now that $F_{\mathcal{O}}(\psi_i, m) = \frac{1}{m}\||\xi_{\psi_i}\rangle\|^2_{[m]}$ can be written as $F_{\mathcal{O}}(\psi_i, m) = \max_{\omega\in\mathcal{M}_m} F(\psi_i, \omega)$, where we employed the dual characterisation of the m-distillation norm of the Schmidt vector. Since $\mathcal{M}_m$ is defined as the convex hull of rank-one projectors, we can now use the result of Streltsov *et al* [83] (see also [82]) to obtain

$$F_{A,\mathcal{O}}(\rho, m) = \max\left\{\sum_i p_i \max_{\omega_i\in\mathcal{M}_m} F(\psi_i, \omega_i) \;\middle|\; \rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|\right\} = \max_{\omega\in\mathcal{M}_m} F(\rho, \omega) \tag{93}$$

as required. The quantity ϑ is simply a function defined so that any state ω satisfies $\omega \in \mathcal{M}_m \iff \vartheta(\omega) \leqslant \frac{1}{m}$, allowing us to obtain

$$\begin{aligned} E^{(1),\varepsilon}_{A,\mathcal{O}}(\rho) &\coloneqq \log\max\left\{m\in\mathbb{N} \;\middle|\; \max_{\omega\in\mathcal{M}_m} F(\rho, \omega) \geqslant 1-\varepsilon\right\} \\ &= \log\max\left\{m\in\mathbb{N} \;\middle|\; \omega\in\mathbb{D},\ \vartheta(\omega)\leqslant\frac{1}{m}, F(\rho,\omega)\geqslant 1-\varepsilon\right\} \end{aligned} \tag{94}$$

and thus completing the proof. ■

3.6. Zero-error distillation

Taking $\varepsilon = 0$ in the task of one-shot distillation corresponds to the problem of characterising the exact transformation $\rho \to \Psi_m$ with a given class of free operations. One is then interested in understanding not only the one-shot capabilities in such a task, but also the asymptotically achievable rate

$$E^{\infty,0}_{d,\mathcal{O}}(\rho) \coloneqq \limsup_{n\to\infty} \frac{1}{n} E^{(1),0}_{d,\mathcal{O}}(\rho^{\otimes n}). \tag{95}$$

To apply our methods in this setting, let us focus on the classes of operations for which we have shown that $F_{\mathcal{O}}(\rho, m) = G^{(m)}_{\mathcal{S}}(\rho)$ for some set $\mathcal{S}$; recall from our previous results that $F_{\mathrm{SEPP}}(\cdot,m) = G^{(m)}_{\mathsf{SEP}}$, $F_{\mathrm{PPTP}_+}(\cdot,m) = G^{(m)}_{\mathsf{PPT}_+}$, $F_{\mathrm{PPTP}'_+} = G^{(m)}_{\mathsf{PPT}'_+}$, and $F_{\mathrm{PPTP}}(\cdot,m) = F_{\mathrm{PPT}}(\cdot,m) = G^{(m)}_{\mathsf{PPT}'}$.

Lemma 21. *Take $\mathcal{O} \in \{\mathrm{PPT}, \mathrm{PPTP}, \mathrm{PPTP}'_+, \mathrm{PPTP}_+, \mathrm{SEPP}\}$ and let $\mathcal{S}$ be the set such that $F_{\mathcal{O}}(\rho, m) = G^{(m)}_{\mathcal{S}}(\rho)$ for the given class. Then, for any ρ, it holds that*

$$E^{(1),0}_{d,\mathcal{O}}(\rho) = \lfloor -\log\min\{\Gamma_{\mathcal{S}^\circ}(W) |\ \Pi_\rho \preceq W \preceq \mathbb{1}\}\rfloor_{\log}, \tag{96}$$

where Π_ρ is the projector onto the support of ρ.

Proof. Using the characterisation in Proposition 4 and Theorem 5, we can write

$$E^{(1),0}_{d,\mathcal{O}}(\rho) = \lfloor -\log\min\{\Gamma_{\mathcal{S}^\circ}(W) | \langle\rho, W\rangle = 1,\ 0\preceq W\preceq\mathbb{1}\}\rfloor_{\log}. \tag{97}$$

Write ρ in its spectral decomposition as $\rho = \sum_i \lambda_i|\psi_i\rangle\langle\psi_i|$. We then have

$$\sum_i \lambda_i = \mathrm{Tr}\,\rho = 1 = \langle W, \rho\rangle = \sum_i \lambda_i \langle W, |\psi_i\rangle\langle\psi_i|\rangle \tag{98}$$

and so $\langle\psi_i|W|\psi_i\rangle = 1$ for each $i \in \{1,\ldots,r\}$ since $0 \preceq W \preceq \mathbb{1}$. The constraints then imply that every feasible solution will have the form $W = \Pi_\rho + P$ with $0 \preceq P \preceq \mathbb{1}$ and $\mathrm{supp}(P) \subseteq \ker(\rho)$, and in particular $\Pi_\rho \preceq W \preceq \mathbb{1}$. Conversely, every W such that $\Pi_\rho \preceq W \preceq \mathbb{1}$ satisfies $1 \geqslant \langle\rho, W\rangle \geqslant 1$ and $0 \preceq W \preceq \mathbb{1}$, so the feasible sets of the two problems are equal. ■

Note that for any positive semidefinite W, from equation (10) we have $\Gamma_{\mathcal{S}^\circ}(W) = \max_{X\in\mathcal{S}}\langle X, W\rangle$. For the case of PPT operations, where $\Gamma_{\mathrm{PPT}'^\circ}(W) = \|W^{T_B}\|_\infty$, the above recovers a result of [84].

Notice that the above implies that one-shot zero-error distillation is impossible from any full-rank state under any class of free operations, as for $\Pi_\rho = \mathbb{1}$ the only feasible W is $\mathbb{1}$ itself and so we have $E^{(1),0}_{d,\mathcal{O}}(\rho) = \log\lfloor\Gamma_{\mathcal{S}^\circ}(\mathbb{1})^{-1}\rfloor = \log 1 = 0$. We will shortly improve this characterisation of zero-error undistillability.

Interestingly, in the case of PPTP_+, PPTP'_+, and SEPP, the set $\mathcal{S}$ consists of positive semidefinite operators, which means that for any $P \in \mathbb{H}_+$ it holds that

$$\Gamma_{\mathcal{S}^\circ}(\Pi_\rho + P) = \max_{\sigma\in\mathcal{S}}\langle\sigma, \Pi_\rho + P\rangle \geqslant \max_{\sigma\in\mathcal{S}}\langle\sigma, \Pi_\rho\rangle = \Gamma_{\mathcal{S}^\circ}(\Pi_\rho) \tag{99}$$

and so Π_ρ itself will be the optimal solution to the minimisation in equation (96). This gives the following.

Corollary 22. *For the classes of operations $\mathcal{O} \in \{\mathrm{PPTP}_+, \mathrm{PPTP}'_+, \mathrm{SEPP}\}$, the one-shot zero-error distillable entanglement is given exactly by*

$$E_{d,\mathcal{O}}^{(1),0}(\rho) = \log\lfloor \Gamma_{\mathcal{S}^\circ}(\Pi_\rho)^{-1}\rfloor. \tag{100}$$

Noticing further that $\Pi_{\rho^{\otimes n}} = \Pi_\rho^{\otimes n}$, we can easily see that $\Gamma_{\mathcal{S}^\circ}(\Pi_{\rho^{\otimes n}}) \geqslant \Gamma_{\mathcal{S}^\circ}(\Pi_\rho)^n$ due to the fact that $\sigma \in \mathcal{S} \Rightarrow \sigma^{\otimes n} \in \mathcal{S}$. This gives the relation $E_{d,\mathcal{O}}^{(1),0}(\rho^{\otimes n}) \leqslant \lfloor -n\log\Gamma_{\mathcal{S}^\circ}(\Pi_\rho)\rfloor_{\log}$, and in particular we see that $-\log\Gamma_{\mathcal{S}^\circ}(\Pi_\rho)$ upper bounds the asymptotically achievable zero-error distillable entanglement $E_{d,\mathcal{O}}^{\infty,0}(\rho)$. Equality does not generally hold since the quantities $\Gamma_{\mathsf{SEP}^\circ}$, $\Gamma_{\mathsf{PPT}_+^\circ}$ are not multiplicative (a counterexample being any state supported on the antisymmetric subspace [85, 86]). Interestingly, multiplicativity is indeed satisfied for $\Gamma_{\mathsf{PPT}'^\circ_+}$—this can be seen explicitly by expressing the computation of $\Gamma_{\mathsf{PPT}'^\circ_+}$ in its dual form as

$$\Gamma_{\mathsf{PPT}'^\circ_+}(\Pi_\rho) = \min_{Q\succcurlyeq\Pi_\rho} \|Q^{T_B}\|_\infty, \tag{101}$$

from which it straightforwardly follows that $\Gamma_{\mathsf{PPT}'^\circ_+}(\Pi_\rho^{\otimes n}) \leqslant \Gamma_{\mathsf{PPT}'^\circ_+}(\Pi_\rho)^n$. This gives in particular the following.

Corollary 23. *The asymptotic zero-error distillable entanglement under Rains-preserving operations is given by*

$$E_{d,\mathcal{O}}^{\infty,0}(\rho) = -\log\Gamma_{\mathsf{PPT}'^\circ_+}(\Pi_\rho). \tag{102}$$

The result therefore ensures the computability of both one-shot and asymptotic zero-error distillable entanglement under PPTP'_+, showing that it constitutes an efficiently computable upper bound for zero-error LOCC distillation. Note that $\Gamma_{\mathsf{PPT}'^\circ_+}(\Pi_\rho)$ appeared previously in the works [33, 58] as a bound on entanglement cost and zero-error distillable entanglement. Our result gives this quantity a precise operational meaning, establishing it as a zero-error equivalent of the Rains bound (see section 3.1.1; see also discussion in [33]).

Evaluating $\Gamma_{\mathsf{SEP}^\circ}$ is significantly more difficult [87, 88]. One can write this quantity more explicitly as [89]

$$\Gamma_{\mathsf{SEP}^\circ}(\Pi_\rho) = \max\{\||\xi_\psi\rangle\|^2_{\ell_\infty} \,|\, |\psi\rangle \in \mathrm{supp}(\rho)\} \tag{103}$$

which makes it easy to see that if the support of ρ contains a product state, then no class of free operations can distil any entanglement without error (even asymptotically). By a result of Parthasarathy [90], if $\mathrm{rank}(\rho) > (d_A - 1)(d_B - 1)$, then $\Gamma_{\mathsf{SEP}^\circ}(\Pi_\rho) = 1$ and so $E_{d,\mathrm{SEPP}}^{(1),0}(\rho) = 0$. In a very similar manner, if $\mathrm{supp}(\rho)$ contains a PPT state, then $E_{d,\mathrm{PPTP}_+}^{(1),0}(\rho) = 0$; this, however, does not give a better universal bound for the rank of ρ which ensures undistillability [91].

Our results in previous sections can further simplify the characterisation of zero-error distillable entanglement for several classes of states. In particular, any pure state has

$$E_{d,\mathcal{O}}^{(1),0}(\psi) = \log\lfloor \||\xi_\psi\rangle\|^{-2}_{\ell_\infty}\rfloor \tag{104}$$

for any class of operations $\mathcal{O}$ considered in this work (which was already known in the case of LOCC [12] and PPT operations [52]), and a maximally correlated state satisfies

$$E_{d,\mathcal{O}}^{(1),0}(\rho_{\mathrm{mc}}) = \log\lfloor \|\Delta(\Pi_{\rho_{\mathrm{mc}}})\|^{-1}_\infty\rfloor \tag{105}$$

for any $\mathcal{O} \in \{\mathrm{PPT},\ \mathrm{PPTP}'_+,\ \mathrm{PPTP}_+,\ \mathrm{SEPP}\}$, where Δ is the completely dephasing channel (diagonal map) in the maximally correlated basis. One can furthermore notice that in both of the above cases the quantity $\Gamma_{\mathsf{SEP}^\circ}(\Pi_\rho)$ is multiplicative, which means that in the asymptotic limit we have $E_{d,\mathcal{O}}^{\infty,0}(\psi) = -\log\||\xi_\psi\rangle\|^2_{\ell_\infty}$ and $E_{d,\mathcal{O}}^{\infty,0}(\rho_{\mathrm{mc}}) = -\log\|\Delta(\Pi_{\rho_{\mathrm{mc}}})\|_\infty$.

Finally, we remark that the quantity $\Gamma_{\mathsf{SEP}^\circ}$, often encountered under the name h_{SEP}, has found a plethora of uses beyond the resource theory of entanglement—in particular, in the theory of quantum Merlin–Arthur games [92] as well as in characterising the maximum output norms of quantum channels [85, 92]. Indeed, the non-multiplicativity of $\Gamma_{\mathsf{SEP}^\circ}$ is equivalent to the non-multiplicativity of the norm $\|\Lambda\|_{1\to\infty} := \max\{\|\Lambda(\rho)\|_\infty \,|\, \rho \in \mathbb{D}\}$ of a channel Λ; specifically, if Λ takes operators on a Hilbert space $\mathcal{H}_{\mathrm{in}}$ to operators on Hilbert space $\mathcal{H}_{\mathrm{out}}$ and $V\colon \mathcal{H}_{\mathrm{in}} \to \mathcal{H}_{\mathrm{out}} \otimes \mathcal{H}_R$ is an isometry such that $\Lambda(\cdot) = \mathrm{Tr}_R\, V\cdot V^\dagger$ for some auxiliary Hilbert space $\mathcal{H}_R$, then $\|\Lambda\|_{1\to\infty} = \Gamma_{\mathsf{SEP}^\circ}(VV^\dagger)$ [92]. This interpretation provides an understanding of the cases in which $\Gamma_{\mathsf{SEP}^\circ}(VV^\dagger)$ is multiplicative: these are the cases in which the protocol Λ obeys so-called perfect parallel repetition [92]. It is furthermore known that, although not multiplicative, the quantity $\Gamma_{\mathsf{SEP}^\circ}$ obeys a form of weaker multiplicativity relations [93, 94]. In the context of entanglement distillation we can see that additivity, in the sense that $E_{d,\mathrm{SEPP}}^{\infty,0}(\rho) = -\log\Gamma_{\mathsf{SEP}^\circ}(\Pi_\rho)$, holds when the optimal operation $\Lambda \in \mathrm{SEPP}$ which distils entanglement from ρ satisfies $\Lambda^{\otimes n} \in \mathrm{SEPP}$ for any n. We stress that an additive lower bound on $E_{d,\mathrm{SEPP}}^{\infty,0}(\rho)$, and therefore also an upper bound on the regularisation of $\|\cdot\|_{1\to\infty}$, is given by Corollary 23.

4. Discussion

The contribution of our work is twofold.

First, we established a comprehensive set of theoretical tools for the study of entanglement distillation. Employing a general framework based on convex analysis, we were able to relate many operational quantities to convex optimisation problems which can be efficiently characterised, in particular allowing for a significant simplification of the optimisation in many relevant cases. Our results revealed general connections between entanglement monotones $G_{\mathcal{O}}^{(m)}$ and the hypothesis testing relative entropy D_H^{ε}, uncovering the fundamental role that both of the quantities play in the task of one-shot entanglement distillation.

Second, the methods found immediate operational applications in characterising the capabilities of several sets of quantum channels which extend the set LOCC. We not only established a precise and accessible one-shot description of entanglement distillation under a wide variety of relevant operations, we revealed several operational equivalences in distillation in the one-shot regime—showing in particular that all sets of free operations achieve exactly the same performance in pure-state distillation, with similar simplifications occurring also in the distillation from isotropic and maximally correlated states. The theoretical framework allowed us to establish computable expressions for the distillation fidelities and rates in such cases, thus providing an exact characterisation of entanglement distillation for these classes of states. The insight from the one-shot characterisation allowed for an operational interpretation of quantities which did not enjoy a direct interpretation of this kind, including asymptotic bounds such as the Rains bound and its zero-error equivalent as well as entanglement monotones such as the generalised robustness or the modified trace distance of entanglement.

Our work thus sheds light on fundamental problems in the study of manipulating entanglement as a resource. By providing a powerful theoretical framework, establishing a precise description of entanglement distillation in the practically relevant one-shot setting, as well as uncovering several novel relations in the operational description of LOCC and beyond, our results will contribute to the ongoing effort to efficiently utilise entanglement in technological applications and optimise the performance of quantum technologies.

Due to the high generality of our framework, we expect it to find use in a variety of contexts not explicitly considered in this work, facilitating the precise description of other classes of states and operations. We hope the results can aid not only the further study of entanglement, but also other quantum resources whose distillation enjoys a similar structure [55, 73], including for example coherence [61, 95, 96] or thermodynamics [97–99].

Acknowledgments

We thank Ryuji Takagi, Hayata Yamasaki, and in particular Ludovico Lami for useful discussions. BR and MG acknowledge the financial support of the National Research Foundation of Singapore Fellowship No. NRF-NRFF2016-02 and the National Research Foundation and L'Agence Nationale de la Recherche joint Project No. NRF2017-NRFANR004 VanQuTe. KF was supported by the University of Cambridge Isaac Newton Trust Early Career grant RG74916. X W was supported by the Department of Defense.

Appendix. Properties of the monotones $T_{\mathcal{S}}^{(m)}$

First of all, we establish that the considered quantities are valid measures of entanglement. A common set of requirements that an entanglement monotone M should obey is [38]: faithfulness (i.e. $M(\rho) = 0$ iff $\rho \in \mathsf{SEP}$), convexity, and strong monotonicity (i.e. the requirement that $M(\rho) \geqslant \sum_i p_i M(\Lambda_i(\rho))$ for any probabilistic protocol which applies an LOCC transformation Λ_i to ρ with probability p_i). By a direct application of Theorem 20 in [41], we have the following.

Proposition 24. *Let $\mathcal{S} \in \{\mathsf{PPT}, \mathsf{PPT}_+, \mathsf{SEP}\}$, and consider the class of CPTP operations $\mathcal{O}$ such that $X \in \mathcal{S} \Rightarrow \Lambda(X) \in \mathcal{S}$. Then, for each $m \geqslant 1$, $T_{\mathcal{S}}^{(m)}$ is faithful with respect to the set $\mathcal{S}$, convex, and strongly monotonic under the operations $\mathcal{O}$.*

The above establishes in particular that all of the measures are strong monotones under LOCC. Note, however, that $T_{\mathsf{PPT}}^{(m)}$ and $T_{\mathsf{PPT}_+}^{(m)}$ are not faithful as entanglement measures, since they are zero for all PPT states.

The following result establishes a dual form for the measures.

Proposition 1. *The measures $T_{\mathcal{S}}^{(m)}$ can be equivalently expressed as*

$$T_{\mathcal{S}}^{(m)}(\rho) = \min\{m\,\mathrm{Tr}(\rho - X)_+ + \mathrm{Tr}(\rho - X)_- \,|\, X \in \mathcal{S}^{**}\}, \tag{A1}$$

where $(\rho - X)_+$ (respectively, $(\rho - X)_-$) denotes the positive (negative) part of the Hermitian operator $\rho - X$.

Proof. For any self-adjoint operator X, let $\{X \succcurlyeq 0\}$ (respectively, $\{X \preccurlyeq 0\}$) denote the orthogonal projection operator onto the span of the eigenvectors corresponding to non-negative (non-positive) eigenvalues of X. The positive and negative parts of X are then given by $X_+ = \{X \succcurlyeq 0\} X \{X \succcurlyeq 0\}$ and $X_- = -\{X \preccurlyeq 0\} X \{X \preccurlyeq 0\}$, such that $X = X_+ - X_-$.

By strong Lagrange duality we have

$$T_{\mathcal{S}}^{(m)}(\rho) = \min \{m \operatorname{Tr} A + \operatorname{Tr} B |\ \rho - X = A - B,\ A, B \succcurlyeq 0,\ X \in \mathcal{S}^{**}\}. \tag{A2}$$

We will now show that for each feasible X, the optimal value of the optimisation problem

$$\min \{m \operatorname{Tr} A + \operatorname{Tr} B |\ \rho - X = A - B,\ A, B \succcurlyeq 0\} \tag{A3}$$

is given by $m \operatorname{Tr}(\rho - X)_+ + \operatorname{Tr}(\rho - X)_-$. To see this, note that on the one hand we can take $A = (\rho - X)_+$ and $B = (\rho - X)_-$, and on the other hand by strong Lagrange duality we have

$$\min \{m \operatorname{Tr} A + \operatorname{Tr} B |\ \rho - X = A - B,\ A, B \succcurlyeq 0\} = \max \{\langle \rho - X, W \rangle | -\mathbb{1} \preccurlyeq W \preccurlyeq m\mathbb{1}\} \tag{A4}$$

for which $W = m\{\rho \succcurlyeq X\} - \{\rho \preccurlyeq X\}$ is a feasible solution. ■

We additionally establish an equality between the two PPT-based monotones in the case $m = d - 1$.

Proposition 26. *For any state we have*

$$R^{\mathbb{D}}_{\mathsf{PPT}}(\rho) = R^{\mathbb{D}}_{\mathsf{PPT}_+}(\rho). \tag{A5}$$

Proof. Let $W \in -\mathsf{PPT}_+^*$ be the optimal dual solution for $R^{\mathbb{D}}_{\mathsf{PPT}_+}$, which means it satisfies $-\mathbb{1} \preccurlyeq W$ and $W = N + Q^{T_B}$ for $N, Q \preccurlyeq 0$. But then notice that $W' \coloneqq W - N \in -\mathsf{PPT}^*$ is also feasible, and we have $\langle \rho, W' \rangle = \langle \rho, W \rangle - \langle \rho, N \rangle \geqslant \langle \rho, W \rangle$, so in fact it suffices to optimise over $W \in -\mathsf{PPT}^*$. ■

Remark. In [35], it was claimed that $T^{(m-1)}_{\mathsf{PPT}_+}(\rho_f) = mf - 1$ for all $1 \leqslant m \leqslant d - 1$ and $f \geqslant \frac{1}{d}$. One can see that this is incorrect by comparing it with the result of Theorem 18 which shows that $T^{(m-1)}_{\mathsf{PPT}_+}(\rho_f) = \frac{(m-1)(df-1)}{d-1}$.

Remark. A formula was given for $R^{\mathbb{D}}_{\mathsf{PPT}_+}$ in [35] as

$$R^{\mathbb{D}}_{\mathsf{PPT}_+}(\rho) = N(\rho) / \lambda_{\max}(\{\rho^{T_B} \preccurlyeq 0\}^{T_B}), \tag{A6}$$

with $N(\rho)$ denoting the negativity. However, numerical counterexamples to this result can be readily constructed, and we find that this only provides a lower bound for the value of $R^{\mathbb{D}}_{\mathsf{PPT}_+}$ in general. For completeness, we give an explicit counterexample.

Consider the symmetric and antisymmetric subspaces in $d = d_A = d_B = 3$, spanned by the sets of mutually orthogonal vectors

$$\begin{cases} |\psi_1\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |\psi_2\rangle = \frac{1}{\sqrt{2}}(|02\rangle + |20\rangle) \\ |\psi_3\rangle = \frac{1}{\sqrt{2}}(|12\rangle + |21\rangle) \end{cases} \begin{cases} |\alpha_1\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \\ |\alpha_2\rangle = \frac{1}{\sqrt{2}}(|02\rangle - |20\rangle) \\ |\alpha_3\rangle = \frac{1}{\sqrt{2}}(|12\rangle - |21\rangle) \end{cases} \tag{A7}$$

respectively.

Take the ansatz $\rho = \frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2|$. An explicit calculation gives $N(\rho) = \frac{1}{2\sqrt{2}}$ and $\lambda_{\max}(\{\rho^{T_B} \preccurlyeq 0\}^{T_B}) = \frac{1}{2}$. Now, take the operator given by

$$W = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| - |\psi_3\rangle\langle\psi_3| - |\alpha_1\rangle\langle\alpha_1| - |\alpha_2\rangle\langle\alpha_2| + |\alpha_3\rangle\langle\alpha_3| - |00\rangle\langle 00| - |11\rangle\langle 11| - |22\rangle\langle 22|, \tag{A8}$$

whose partial transpose can be computed as $W^{T_B} = -3|w\rangle\langle w|$ with $|w\rangle = \frac{1}{\sqrt{3}}(|00\rangle - |11\rangle - |22\rangle)$. We then clearly have $W \succcurlyeq -\mathbb{1}$ and $W^{T_B} \preccurlyeq 0$, which gives

$$R_{\mathsf{PPT}_+}(\rho) \geqslant \langle \rho, W \rangle = 1 > \frac{1}{\sqrt{2}} = \frac{N(\rho)}{\lambda_{\max}(\{\rho^{T_B} \preccurlyeq 0\}^{T_B})}. \tag{A9}$$

We remark that, despite the lack of an exact analytical expression, the quantity $R^{\mathbb{D}}_{\mathsf{PPT}_+}$ can nevertheless be evaluated efficiently as a semidefinite program.

Remark. For completeness, we collect the results obtained in the manuscript which simplify the computation of $T_{\mathcal{S}}^{(m)}$ in several cases.

- For any pure state, any $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+\}$, and any integer $1 \leqslant m \leqslant d-1$, it holds that
$$T_{\mathcal{S}}^{(m-1)}(\psi) = \| |\xi_\psi\rangle \|_{[m]}^2 - 1. \tag{A10}$$
(See Theorems 14 and 15.)
- For any isotropic state, any $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+\}$, and any $1 \leqslant m \leqslant d-1$, it holds that
$$T_{\mathcal{S}}^{(m)}(\rho_f) = \begin{cases} 0 & f \leqslant \dfrac{1}{d}, \\ \dfrac{m(df-1)}{d-1} & f \geqslant \dfrac{1}{d}. \end{cases} \tag{A11}$$
(See Theorem 18.)
- For any maximally correlated state, any $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+\}$, and any $1 \leqslant m \leqslant d-1$, it holds that
$$T_{\mathcal{S}}^{(m)}(\rho_{\mathrm{mc}}) = T_{\mathcal{I}}^{(m)}(\widetilde{\rho_{\mathrm{mc}}}), \tag{A12}$$
where $\rho_{\mathrm{mc}} = \sum_{i,j}\rho_{ij}|ii\rangle\langle jj|$, $\widetilde{\rho_{\mathrm{mc}}} = \sum_{i,j}\rho_{ij}|i\rangle\langle j|$, and $\mathcal{I}$ is the set of states diagonal in the given basis $\{|i\rangle\}$. (See Theorem 19.)
- For any $\mathcal{S} \in \{\mathsf{SEP}, \mathsf{PPT}, \mathsf{PPT}_+\}$, it holds that
$$T_{\mathcal{S}}^{(d-1)}(\rho) = R_{\mathcal{S}}^{\mathbb{D}}(\rho). \tag{A13}$$
(See Proposition 2.)

ORCID iDs

Bartosz Regula https://orcid.org/0000-0001-7225-071X
Kun Fang https://orcid.org/0000-0002-9232-6846
Xin Wang https://orcid.org/0000-0002-0641-3186

References

[1] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 *Phys. Rev. Lett.* **70** 1895
[2] Bennett C H and Wiesner S J 1992 *Phys. Rev. Lett.* **69** 2881
[3] Briegel H-J, Dür W, Cirac J I and Zoller P 1998 *Phys. Rev. Lett.* **81** 5932
[4] Kimble H J 2008 *Nature* **453** 1023
[5] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 *Phys. Rev.* A **54** 3824
[6] Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J A and Wootters W K 1996 *Phys. Rev. Lett.* **76** 722
[7] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 *Rev. Mod. Phys.* **81** 865
[8] Rains E M 1999 *Phys. Rev.* A **60** 173
[9] Horodecki M, Horodecki P and Horodecki R 1998 *Phys. Rev. Lett.* **80** 5239
[10] Vedral V and Plenio M B 1998 *Phys. Rev.* A **57** 1619
[11] Devetak I and Winter A 2005 *Proc. R. Soc.* A **461** 207
[12] Lo H-K and Popescu S 2001 *Phys. Rev.* A **63** 022301
[13] Morikoshi F and Koashi M 2001 *Phys. Rev.* A **64** 022316
[14] Martín-Delgado M A and Navascués M 2003 *Phys. Rev.* A **68** 012322
[15] Hayashi M 2006 *IEEE Trans. Inf. Theory* **52** 1904
[16] Buscemi F and Datta N 2010 *J. Math. Phys.* **51** 102201
[17] Buscemi F and Datta N 2013 *IEEE Trans. Inf. Theory* **59** 1940
[18] Brandão F G S L and Datta N 2011 *IEEE Trans. Inf. Theory* **57** 1754
[19] Datta N and Leditzky F 2015 *IEEE Trans. Inf. Theory* **61** 582
[20] Fang K, Wang X, Tomamichel M and Duan R 2019 *IEEE Trans. Inf. Theory* **65** 6454
[21] Rozpedek F, Schiet T, Thinh L P, Elkouss D, Doherty A C and Wehner S 2018 *Phys. Rev.* A **97** 062333
[22] Chitambar E, Leung D, Mančinska L, Ozols M and Winter A 2014 *Commun. Math. Phys.* **328** 303
[23] Rains E M 1997 arXiv:quant-ph/9707002
[24] Rains E M 1999 *Phys. Rev.* A **60** 179
[25] Rains E M 2001 *IEEE Trans. Inf. Theory* **47** 2921
[26] Eggeling T, Vollbrecht K G H, Werner R F and Wolf M M 2001 *Phys. Rev. Lett.* **87** 257902
[27] Brandão F G S L and Plenio M B 2010 *Commun. Math. Phys.* **295** 829
[28] Cirac J I, Dür W, Kraus B and Lewenstein M 2001 *Phys. Rev. Lett.* **86** 544
[29] Wilde M M, Tomamichel M and Berta M 2017 *IEEE Trans. Inf. Theory* **63** 1792
[30] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 *Phys. Rev.* A **53** 2046

[31] DiVincenzo D P, Fuchs C A, Mabuchi H, Smolin J A, Thapliyal A and Uhlmann A 1999 *Quantum Computing and Quantum Communications (Lecture Notes in Computer Science)* (Berlin: Springer) pp 247–57
[32] Smolin J A, Verstraete F and Winter A 2005 *Phys. Rev.* A **72** 052317
[33] Wang X and Duan R 2017 *Phys. Rev.* A **95** 062322
[34] Vidal G and Tarrach R 1999 *Phys. Rev.* A **59** 141
[35] Brandão F G S L 2005 *Phys. Rev.* A **72** 022310
[36] Yu X-D, Zhang D-J, Xu G F and Tong D M 2016 *Phys. Rev.* A **94** 060302
[37] Qiao L-F *et al* 2018 *Phys. Rev.* A **98** 052351
[38] Vidal G 2000 *J. Mod. Opt.* **47** 355
[39] Johnston N and Kribs D W 2010 *J. Math. Phys.* **51** 082202
[40] Rockafellar R T 1970 *Convex Analysis* (Princeton, NJ: Princeton University Press)
[41] Regula B 2018 *A: Math. Theor.* **51** 045303
[42] Buscemi F and Datta N 2010 *IEEE Trans. Inf. Theory* **56** 1447
[43] Wang L and Renner R 2012 *Phys. Rev. Lett.* **108** 200501
[44] Tomamichel M and Hayashi M 2013 *IEEE Trans. Inf. Theory* **59** 7693
[45] Hayashi M 2016 *Quantum Information Theory: Mathematical Foundation* (Berlin: Springer)
[46] Hayashi M 2017 arXiv:1709.07701 [quant-ph]
[47] Vedral V, Plenio M B, Rippin M A and Knight P L 1997 *Phys. Rev. Lett.* **78** 2275
[48] Bennett C H, DiVincenzo D P, Fuchs C A, Mor T, Rains E, Shor P W, Smolin J A and Wootters W K 1999 *Phys. Rev.* A **59** 1070
[49] Audenaert K, Plenio M B and Eisert J 2003 *Phys. Rev. Lett.* **90** 027901
[50] Chitambar E, de Vicente J I, Girard M W and Gour G 2017 arXiv:1711.03835
[51] Horodecki M 2001 *Quant. Inf. Comput.* **1** 3
[52] Matthews W and Winter A 2008 *Phys. Rev.* A **78** 012317
[53] Horodecki M and Horodecki P 1999 *Phys. Rev.* A **59** 4206
[54] Vidal G and Werner R F 2002 *Phys. Rev.* A **65** 032314
[55] Brandão F G S L and Gour G 2015 *Phys. Rev. Lett.* **115** 070503
[56] Brandão F G S L and Plenio M B 2010 *Commun. Math. Phys.* **295** 791
[57] Audenaert K, De Moor B, Vollbrecht K G H and Werner R F 2002 *Phys. Rev.* A **66** 032310
[58] Wang X and Duan R 2017 *Phys. Rev. Lett.* **119** 180506
[59] Paulsen V 2002 *Completely Bounded Maps and Operator Algebras* (Cambridge: Cambridge University Press)
[60] Watrous J 2004 arXiv:quant-ph/0411077
[61] Regula B, Fang K, Wang X and Adesso G 2018 *Phys. Rev. Lett.* **121** 010401
[62] Rudolph O 2005 *Quant. Inf. Proc.* **4** 219
[63] Johnston N and Kribs D W 2015 *Houst. J. Math.* **41** 831
[64] Shimony A 1995 *Ann. New York Acad. Sci.* **755** 675
[65] Vidal G, Jonathan D and Nielsen M A 2000 *Phys. Rev.* A **62** 012304
[66] Nielsen M A 1999 *Phys. Rev. Lett.* **83** 436
[67] Ishizaka S and Plenio M B 2005 *Phys. Rev.* A **71** 052303
[68] Lobo M S, Vandenberghe L, Boyd S and Lebret H 1998 *Linear Algebr. Appl.* **284** 193
[69] Takagi R, Regula B, Bu K, Liu Z-W and Adesso G 2019 *Phys. Rev. Lett.* **122** 140402
[70] Bae J, Chruściński D and Piani M 2019 *Phys. Rev. Lett.* **122** 140404
[71] Horodecki M, Horodecki P and Horodecki R 1999 *Phys. Rev.* A **60** 1888
[72] Verstraete F and Verschelde H 2003 *Phys. Rev. Lett.* **90** 097901
[73] Liu Z-W, Bu K and Takagi R 2019 *Phys. Rev. Lett.* **123** 020401
[74] Lami L and Huber M 2016 *J. Math. Phys.* **57** 092201
[75] Baumgratz T, Cramer M and Plenio M B 2014 *Phys. Rev. Lett.* **113** 140401
[76] Streltsov A, Adesso G and Plenio M B 2017 *Rev. Mod. Phys.* **89** 041003
[77] Winter A and Yang D 2016 *Phys. Rev. Lett.* **116** 120404
[78] Chitambar E and Gour G 2016 *Phys. Rev. Lett.* **117** 030401
[79] Zhu H, Hayashi M and Chen L 2017 *J. Phys. A: Math. Theor.* **50** 475303
[80] Vandenberghe L and Boyd S 1996 *SIAM Rev.* **38** 49
[81] Hiroshima T and Hayashi M 2004 *Phys. Rev.* A **70** 030302
[82] Regula B, Lami L and Streltsov A 2018 *Phys. Rev.* A **98** 052329
[83] Streltsov A, Kampermann H and Bruß D 2010 *New J. Phys.* **12** 123004
[84] Wang X and Duan R 2016 *Phys. Rev.* A **94** 050301
[85] Werner R F and Holevo A S 2002 *J. Math. Phys.* **43** 4353
[86] Grudka A, Horodecki M and Pankowski Ł 2010 *J. Phys. A: Math. Theor.* **43** 425304
[87] Gurvits L 2003 *Proc. 35th Annual ACM Symp. on Theory of Computing, STOC '03* (New York: ACM) pp 10–19
[88] Harrow A W and Montanaro A 2010 *2010 IEEE 51st Annual Symp. on Foundations of Computer Science* pp 633–42
[89] Acin A, Vidal G and Cirac J I 2003 *Quantum. Inf. Comput.* **3** 55
[90] Parthasarathy K R 2004 *Proc. Math. Sci.* **114** 365
[91] Johnston N 2013 *Phys. Rev.* A **87** 064302
[92] Harrow A W and Montanaro A 2013 *J ACM* **60** 1
[93] Montanaro A 2013 *Commun. Math. Phys.* **319** 535
[94] Lancien C and Winter A 2017 *J. Math. Phys.* **58** 092203
[95] Zhao Q, Liu Y, Yuan X, Chitambar E and Winter A 2019 *IEEE Trans. Inf. Theory* **65** 6441
[96] Lami L, Regula B and Adesso G 2019 *Phys. Rev. Lett.* **122** 150402
[97] Åerg J 2013 *Nat. Commun.* **4** 1925
[98] Horodecki M and Oppenheim J 2013 *Nat. Commun.* **4** 2059
[99] Yunger Halpern N and Renes J M 2016 *Phys. Rev.* E **93** 022126