PAPER • OPEN ACCESS

Symmetric distinguishability as a quantum resource

To cite this article: Robert Salzmann *et al* 2021 *New J. Phys.* **23** 083016

View the article online for updates and enhancements.

New Journal of Physics

The open access journal at the forefront of physics

PAPER

Symmetric distinguishability as a quantum resource

Robert Salzmann[1,*] , Nilanjana Datta[1], Gilad Gour[2,3], Xin Wang[4] and Mark M Wilde[5,6]

[1] University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge CB3 0WA, United Kingdom
[2] Department of Mathematics and Statistics, University of Calgary, Alberta, T2N 1N4 Canada
[3] Institute for Quantum Science and Technology, University of Calgary, Alberta, T2N 1N4 Canada
[4] Institute for Quantum Computing, Baidu Research, Beijing 100193, People's Republic of China
[5] Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, United States of America
[6] Stanford Institute for Theoretical Physics, Stanford University, Stanford, CA 94305, United States of America
* Author to whom any correspondence should be addressed.

E-mail: rs2059@cam.ac.uk

Keywords: quantum information theory, quantum resource theory, quantum hypothesis testing, quantum Chernoff divergence, symmetric distinguishability

Abstract

We develop a resource theory of symmetric distinguishability, the fundamental objects of which are elementary quantum information sources, i.e. sources that emit one of two possible quantum states with given prior probabilities. Such a source can be represented by a classical-quantum state of a composite system *XA*, corresponding to an ensemble of two quantum states, with *X* being classical and *A* being quantum. We study the resource theory for two different classes of free operations: (i)CPTP$_A$, which consists of quantum channels acting only on *A*, and (ii) conditional doubly stochastic maps acting on *XA*. We introduce the notion of symmetric distinguishability of an elementary source and prove that it is a monotone under both these classes of free operations. We study the tasks of distillation and dilution of symmetric distinguishability, both in the one-shot and asymptotic regimes. We prove that in the asymptotic regime, the optimal rate of converting one elementary source to another is equal to the ratio of their quantum Chernoff divergences, under both these classes of free operations. This imparts a new operational interpretation to the quantum Chernoff divergence. We also obtain interesting operational interpretations of the Thompson metric, in the context of the dilution of symmetric distinguishability.

1. Introduction

Distinguishability plays a central role in all of modern science. The ability to distinguish one possibility from another allows for making inferences from experimental data and making decisions based on these inferences or developing new theories. Thus, it is essential to understand distinguishability from a fundamental perspective. Furthermore, distinguishability is a resource, in the sense that fewer trials of an experiment are needed to arrive at conclusions when two different possibilities are more distinguishable from one another.

In this paper, we adopt a resource-theoretic approach to distinguishability in quantum mechanics that ultimately is helpful in and enriches our fundamental understanding of distinguishability. We note here that, more generally, the resource-theoretic approach to quantum information processing [CG19] has illuminated not only quantum information science but also other areas of research in physics and mathematical statistics. Our work differs from prior developments with a related motivation [Mat10, WW19a, WW19b], in that here we focus instead on what we call symmetric distinguishability (SD), or alternatively, the Bayesian approach to distinguishability. The outcome of our efforts is a resource theory with a plethora of appealing features, including *asymptotic reversibility* with the optimal conversion rate being given by a ratio of Chernoff divergences. We explain these concepts in more detail in what follows.

Our theory is similar in spirit to that proposed previously in [Mor09], but there are some notable differences and our conclusions are arguably stronger than those presented in [Mor09]. We refer to the resource theory that we propose here as the resource theory of symmetric distinguishability (RTSD).

1.1. Overview of the resource theory of symmetric distinguishability

The basic objects of this resource theory are elementary quantum information sources, which emit one of two quantum states with certain prior probabilities (i.e. the source emits a state ρ_0 with probability p or a state ρ_1 with probability $1 - p$). Such a source can be represented by the following classical–quantum (c–q) state:

$$\rho_{XA} := p|0\rangle\langle 0| \otimes \rho_0 + \left(1 - p\right)|1\rangle\langle 1| \otimes \rho_1, \tag{1.1}$$

where $p \in [0, 1]$ is a prior probability and ρ_0 and ρ_1 are quantum states[7]. Note that the classical system is equivalently specified by a random variable that we also denote as X. In analogy with the notation used in the resource theory of asymmetric distinguishability [WW19a, WW19b], it can be equivalently represented by a *quantum box* given by the triple (p, ρ_0, ρ_1). The nomenclature 'box' is used here to indicate that a quantum system is prepared in the state ρ_0 with probability p and ρ_1 with probability $1 - p$ and it is not known which is the case (thus, the system is analogous to an unopened box).

An important goal of the resource theory is to transform a state of the above form to the following state

$$\sigma_{XB} := q|0\rangle\langle 0| \otimes \sigma_0 + \left(1 - q\right)|1\rangle\langle 1| \otimes \sigma_1, \tag{1.2}$$

via a chosen set of free operations, where $q \in [0, 1]$ and σ_0 and σ_1 are quantum states. Note that the target system B need not be isomorphic to the initial system A. This corresponds to the following transformation between boxes: $(p, \rho_0, \rho_1) \mapsto (q, \sigma_0, \sigma_1)$. Note that in what follows, we often suppress the subscripts denoting the quantum systems, for notational simplicity.

Given such an elementary quantum source $\rho_{XA} \equiv (p, \rho_0, \rho_1)$, a natural way to study distinguishability is to consider the binary hypothesis testing task of discriminating between the states ρ_0 and ρ_1. There are two possible errors that can be incurred in the process, namely, the *type I error* (mistaking ρ_0 to be ρ_1) and the *type II error* (mistaking ρ_1 to be ρ_0). In the setting of asymmetric hypothesis testing, one minimizes the type II error probability under the constraint that the type I error probability is below a given threshold. In contrast, in symmetric hypothesis testing, the two error probabilities are considered on the same footing and weighted by the prior distribution. The latter (also known as Bayesian discrimination) is arguably the first problem ever considered in the field of quantum information theory and solved in the single-copy case by Helstrom [Hel67, Hel69] and Holevo [Hol72]. The operational quantity in this task is the minimum (average) error probability, which we formally define in (3.2) and denote as $p_{\mathrm{err}}(\rho_{XA})$.

Let $\rho_{XA}^{(n)}$ denote the c–q state corresponding to a source that emits the state $\rho_0^{\otimes n}$ (resp. $\rho_1^{\otimes n}$) with probability p (resp. $(1 - p)$). It is known that $p_{\mathrm{err}}(\rho_{XA}^{(n)})$ decays exponentially in n, with exponent given by the Chernoff divergence (also known as the quantum Chernoff bound) of the states ρ_0 and ρ_1 [NS09, ACMnT⁺07]. Just as the resource theory of asymmetric distinguishability (RTAD) provides a resource-theoretic perspective to asymmetric hypothesis testing [Mat10, Mat11, WW19a], our resource theory (RTSD) provides a resource-theoretic framework for symmetric hypothesis testing. By the quantum Stein's lemma, the relevant operational quantity in asymmetric hypothesis testing is known to be characterized by the quantum relative entropy [HP91, ON00], and in the RTAD, the optimal rate of transformation between quantum boxes was proved to be given by a ratio of quantum relative entropies [WW19a] (see also [BST19] in this context). In analogy and given the result of [ACMnT⁺07], it is natural to expect that, in the RTSD, the corresponding optimal asymptotic rate of transformation between quantum boxes is given by a ratio of Chernoff divergences. It is pleasing to see that this is indeed the case. In contrast, in [Mor09], only one-shot transformations are considered and hence, in contrast to our work, asymptotic transformations are not studied.

We consider the RTSD for two different choices of free operations: (i) local quantum channels (i.e. linear, completely positive, trace-preserving maps) acting on the system A alone, and (ii) the more general class of conditional doubly stochastic (CDS) maps. We denote the former class of free operations by CPTP_A. A CDS map acting on a c–q state ρ_{XA} defined through (1.1) consists of quantum operations acting on the system A and associated permutations of the letters $x \in \mathcal{X}$. A detailed justification behind the choice of CDS maps as free operations is given in section 2, where the RTSD is introduced via an axiomatic approach.

In any quantum resource theory, there are several pertinent questions to address. What are the conditions for the feasibility of transforming a source state to a target state? If one cannot perform a

---

[7] Throughout this paper we restrict attention for the most part to c–q states of the form in (1.1).

transformation exactly, how well can one do so approximately? What is an appropriate measure for approximation when converting a source state to a target state? Is there a 'golden unit' resource that one can go through as an intermediate step when converting a source state to a target state? At what rate can one convert repetitions (i.e. multiple copies) of a source state to repetitions of a target state, either exactly or approximately? More specifically, at what rate can one distill repetitions of a source state to the golden unit resource, either exactly or approximately? Conversely, at what rate can one dilute the golden unit resource to repetitions of a target state? Is the resource theory asymptotically reversible? In this paper, we address all of these questions within the context of the RTSD.

Given an elementary quantum source ρ_{XA}, a natural measure of SD is given by the minimum error probability $p_{\mathrm{err}}(\rho_{XA})$ in the context of Bayesian state discrimination (mentioned above). Then

$$\mathrm{SD}(\rho_{XA}) := -\log\left(2p_{\mathrm{err}}(\rho_{XA})\right) \tag{1.3}$$

is a natural measure of the SD contained in ρ_{XA}.[8] A justification of this choice arises from the consideration of *free states* and *infinite-resource states*. For a detailed discussion of the notion of SD in a more general setting, see section 2. A natural choice of a *free state* for this resource theory is a c–q state of the form (1.1), for which $p = 1/2$ and ρ_0 and ρ_1 are identical and hence indistinguishable. For such a state, $p_{\mathrm{err}}(\rho_{XA}) = 1/2$, which is achieved by random guessing, and hence $\mathrm{SD}(\rho_{XA}) = 0$. Note that the converse is true also (i.e. $\mathrm{SD}(\rho_{XA}) = 0$ implies that ρ_0 and ρ_1 are identical and $p = 1/2$). Hence our choice of SD respects the requirement that a state has zero SD if and only if it is free. On the other hand, a natural choice of an *infinite-resource state* is a c–q state of the form (1.1), for which ρ_0 and ρ_1 have mutually orthogonal supports. This corresponds to an elementary quantum information source that emits perfectly distinguishable states. For such a state, $p_{\mathrm{err}}(\rho_{XA}) = 0$ and hence $\mathrm{SD}(\rho_{XA}) = +\infty$. Thus our choice of SD validates the identification of such states as infinite-resource states.

An infinite-resource state has the desirable property that it can be converted to any other c–q state via CDS maps. Under CPTP_A, it can be transformed to any other c–q state with the same prior. Moreover, any given c–q state cannot be transformed to an infinite-resource state unless it is itself an infinite-resource state.

We coin the word *SD-bit* to refer to the unit of symmetric distinguishability, the basic currency of this resource theory. In fact, for every positive real number $m \geqslant 0$, it is useful to identify a family of c–q states that have m SD-bits. In definition 3.4 of section 3.1, we consider a natural choice for such a family of c–q states. They are parametrized by $M \equiv 2^m$ and denoted as $\gamma_{XQ}^{(M)}$. Such a state, which we call an M-golden unit, has the following key property: $p_{\mathrm{err}}(\gamma_{XQ}^{(M)}) = \frac{1}{2M}$ and hence $\mathrm{SD}(\gamma_{XQ}^{(M)}) = \log M = m$.

Consideration of an M-golden unit also leads naturally to a clear definition of the fundamental tasks of distillation and dilution in the RTSD as the conversions of a given c–q state to and from, respectively, an M-golden unit under free operations[9]. The previously mentioned task of transformations between two arbitrary c–q states under free operations can be achieved by first distilling an M-golden unit (for the maximal possible M) from the initial state and then diluting the distilled M-golden unit to the desired target state. This is discussed in detail in sections 4 and 5.

One fundamental setting of interest is the one-shot setting, with the question being to determine the minimum error in converting an initial source to a target source. We prove that this minimum error can be calculated by means of a semi-definite program (SDP). Thus, we can efficiently calculate this error in time polynomial in the dimensions of the A and A' systems.

Moving on to the case of asymptotic transformations, we consider the following fundamental conversion task via free operations:

$$\rho_{XA}^{(n)} \equiv (p, \rho_0^{\otimes n}, \rho_1^{\otimes n}) \mapsto \sigma_{XB}^{(m)} \equiv (q, \sigma_0^{\otimes m}, \sigma_1^{\otimes m}), \tag{1.4}$$

where $p, q \in [0, 1]$ and, for $i \in \{0, 1\}$, ρ_i and σ_i are states of quantum systems A and B, respectively. The goal here is, for a fixed n, to make m as large as possible, and to evaluate the optimal asymptotic rate $\frac{m}{n}$ of the transformation in the limit as n becomes arbitrarily large. We allow for approximations in the transformation and require the approximation error to vanish in the asymptotic limit ($n \to \infty$). This approximation error will be measured with respect to an error measure (defined for c–q states of the form considered in this paper) that we denote by the symbol D'. The precise definition of D' and its mathematical properties are given in section 3.2.

---

[8] Note, however, that this measure is not unique and it is possible to define other measures. In this paper, logarithms are taken to base 2.
[9] [Mor09] also considers transformations to and from a certain unit of resource. However, instead of the M-golden unit that we consider, there the unit of resource is a *dbit*, which is a pair of orthogonal states and therefore corresponds to the case $M = \infty$ of our golden unit. The advantage of our choice (M-golden unit) is that it yields a more refined analysis also in the case of finite resources.

1.2. Main results

In this paper, we develop a consistent and systematic RTSD that answers the most important questions associated with a resource theory. In brief, the main contributions of this paper can be summarized as follows. In the following, all c–q states are assumed to be of the form (1.1), and they hence represent elementary quantum information sources.

- We define two new examples of *generalized divergences*, each of which satisfies the data-processing inequality (DPI). These are denoted as $\xi_{\min}$ and $\xi_{\max}$. Moreover, we define a quantity, denoted as $\xi_{\max}^\star$, on c–q states of the form (1.1), and we prove that it satisfies monotonicity under CDS maps.
- All of the above quantities are of operational significance in the RTSD:
  * The one-shot exact distillable-SD of ρ_{XA} under CPTP_A maps is given by $\xi_{\min}(\rho_{XA})$ (theorem 4.3);
  * The one-shot exact SD-cost of ρ_{XA} under CPTP_A maps is given by $\xi_{\max}(\rho_{XA})$ (theorem 5.3);
  * The one-shot exact distillable-SD of ρ_{XA} under CDS maps is given by its symmetric distinguishability, $\mathrm{SD}(\rho_{XA})$ (theorem 4.5);
  * In addition, the one-shot exact SD-cost of ρ_{XA} under CDS maps is given by $\xi_{\max}^\star(\rho_{XA})$ (theorem 5.5).
- $\xi_{\max}(\rho_{XA})$ and $\xi_{\max}^\star(\rho_{XA})$ are both defined in terms of the *Thompson metric* of the state ρ_{XA} (see theorems 5.3 and 5.5), thus providing operational interpretations of the latter in the context of the RTSD[10].
- The optimal asymptotic rate of exact and approximate SD-distillation for a state $\rho_{XA} \equiv (p, \rho_0, \rho_1)$, under both CPTP_A and CDS maps, is equal to its *quantum Chernoff divergence*, $\xi(\rho_0, \rho_1)$ (theorem 4.7 and 4.17), where

$$\xi(\rho_0, \rho_1) := \sup_{s \in [0,1]} \left(-\log \mathrm{Tr}[\rho_0^s \rho_1^{1-s}] \right). \tag{1.5}$$

- The optimal asymptotic rate of exact SD-dilution for a state ρ_{XA} is equal to its Thompson metric (see theorem 5.12). This provides another clear operational interpretation for the latter.
- The optimal asymptotic rate of approximate SD-dilution for a state ρ_{XA} is equal to its quantum Chernoff divergence (see theorem 5.13).
- The optimal asymptotic rate of transforming one c–q state to another, under both CPTP_A and CDS maps, is equal to the ratio of their quantum Chernoff divergences (see theorem 7.4). This result constitutes a novel operational interpretation of the Chernoff divergence beyond that reported in earlier work on symmetric quantum hypothesis testing [NS09, ACMnT+07]. It also demonstrates that the RTSD is *asymptotically reversible.*

In the following sections, we develop all of the above claims in more detail. In particular, in section 2, we introduce a general RTSD, for arbitrary quantum information sources (given by an ensemble $\{p_x, \rho_x\}_{x \in \mathcal{X}}$ of quantum states), via an axiomatic approach. The resource theory studied in the rest of the paper is a special case of the above, namely, the one for elementary quantum information sources (i.e. corresponding to the choice $|\mathcal{X}| = 2$). Certain necessary ingredients of the RTSD are introduced in this section and in section 3. These include the notion of golden units, which facilitates a study of distillation and dilution of SD. SD-distillation and SD-dilution are studied in sections 4 and 5, respectively, both in the one-shot and asymptotic regimes. In section 6, we elucidate the salient features of the RTSD for certain examples of elementary quantum information sources. The interesting task of converting one elementary quantum information source to another via free operations is studied in section 7. We conclude the main part of the paper with a summary and some open questions for future research. Various relevant quantities of the RTSD can be formulated as SDPs. These are stated in sections 3 and 4, but some of their proofs appear in the appendices.

2. An axiomatic approach to the resource theory of symmetric distinguishability

In this section, we introduce an axiomatic approach to a RTSD, from which the particular resource theory that we study in this paper arises as a natural special case, corresponding to the choice $|X| \equiv |\mathcal{X}| = 2$ in what follows.

---

[10] Even though the Thompson metric has been widely studied in the mathematics literature, to the best of our knowledge, this is the first time an operational meaning has been given to it.

Consider an ensemble $\{p_x, \rho_x\}_{x \in \mathcal{X}}$ of quantum states. Such an ensemble can be described by a c−q state

$$\rho_{XA} = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x| \otimes \rho_x. \tag{2.1}$$

There are many functions that can be used to quantify the distinguishability of the ensemble of states above. Perhaps the function that is most operationally motivated is the *guessing probability*:

$$p_{\text{guess}}(X|A) := \max_{\{\Lambda_x\}_{x \in \mathcal{X}}} \sum_{x \in \mathcal{X}} p_x \operatorname{Tr}(\Lambda_x \rho_x), \tag{2.2}$$

where the maximum is over every possible POVM $\{\Lambda_x\}_{x \in \mathcal{X}}$.

We are interested in a notion of ensemble distinguishability that takes into account the prior distribution $\{p_x\}_{x \in \mathcal{X}}$, while distinguishing between the states in the ensemble, as opposed to state distinguishability, which is concerned only with the distinguishability of the states in the set $\{\rho_x\}_{x \in \mathcal{X}}$.

Motivated by the guessing probability and symmetric hypothesis testing, we identify an ensemble as free in the RTSD if the guessing probability takes on its minimum value; i.e. if the guessing probability is equal to $\frac{1}{|\mathcal{X}|}$, which is the same value attained by a random guessing strategy. Note that the guessing probability is equal to $\frac{1}{|\mathcal{X}|}$ if and only if all the states of the ensemble $\{p_x, \rho_x\}_{x \in \mathcal{X}}$ are identical and the prior distribution is uniform. For the sake of completeness, we include a proof of this fact at the end of this section (see lemma 2.6 below).

Thus, for all such ensembles, the SD is equal to zero. The corresponding c−q state $\rho_{XA} = \pi_X \otimes \omega_A$, where $\pi_X := I_X/|\mathcal{X}|$ is the completely mixed state, is then a 'free' state of the RTSD because it has zero SD. This leads us to identify the set of free states in the RTSD as follows:

$$\mathfrak{F}(XA) := \{\pi_X \otimes \omega_A : \omega_A \in \mathcal{D}(A)\}. \tag{2.3}$$

Note that, in the above, we use the notation $\mathcal{D}(A)$ to denote the set of quantum states (i.e. density matrices) of the quantum system A.

To make the notion of SD precise, we introduce an axiomatic approach. Here, we define a preorder relation $\prec$ on the set of all c−q states, which satisfies axioms I–V below (see definition 2.1). We say a c−q state ρ_{XA} has less SD than $\sigma_{X'A'}$ if $\rho_{XA} \prec \sigma_{X'A'}$. In this approach, SD is a property of a composite physical system shared between two parties, say, Xiao and Alice, with Xiao possessing classical systems (denoted by X, X', etc) and Alice possessing quantum systems denoted by A, A', etc. The word 'symmetric' refers to the fact that the distinguishability is symmetric with respect to the ordering in the ensemble; i.e. for every permutation π, the ensemble $\{p_x, \rho_x\}_x$ has the same SD as the ensemble $\{p_{\pi(x)}, \rho_{\pi(x)}\}_x$ because the latter is just a relabeling of the former. We write this equivalence as the following relation:

Axiom I $\quad \rho_{XA} \sim \mathcal{P}_{X \to X}(\rho_{XA}) \quad \forall\, \mathcal{P} - \text{permutation channel}. \tag{2.4}$

Similarly, every isometric channel $\mathcal{V} \in \text{CPTP}(A \to A')$ that acts on the states $\{\rho_x\}_x$ does not change their 'overlap' (i.e. their Hilbert–Schmidt inner product), and this leads to the next axiom:

Axiom II $\quad \rho_{XA} \sim \mathcal{V}_{A \to A'}(\rho_{XA}) \quad \forall\, \mathcal{V} - \text{isometric channel}. \tag{2.5}$

The fact that states with zero SD cannot add SD leads to the next axiom: for every classical system X' on Xiao's side and every state $\omega_{A'}$ on Alice's side,

Axiom III $\quad \rho_{XA} \sim \rho_{XA} \otimes (\pi_{X'} \otimes \omega_{A'}). \tag{2.6}$

The next axiom concerns two c−q states $\rho_{XAA'}, \sigma_{XAA'} \in \mathcal{D}(XAA')$ that have the form

$$\rho_{XAA'} = \sum_{y \in \mathcal{Y}} q_y \rho_{XA}^y \otimes |y\rangle\langle y|_{A'} \quad \text{and} \quad \sigma_{XAA'} = \sum_{y \in \mathcal{Y}} q_y \sigma_{XA}^y \otimes |y\rangle\langle y|_{A'}, \tag{2.7}$$

with $\mathcal{Y}$ a finite alphabet, $q_y \in [0, 1]$ the elements of a probability distribution, and $\{|y\rangle\}_{y \in \mathcal{Y}}$ an orthonormal basis. Since the label y is distinguishable in such states, we assume that

Axiom IV $\quad \rho_{XA}^y \sim \sigma_{XA}^y \quad \forall\, y \in \mathcal{Y} \;\Rightarrow\; \rho_{XAA'} \sim \sigma_{XAA'}. \tag{2.8}$

The motivation behind axiom IV is the following: since the label y can be perfectly inferred by a measurement of system A', the SD of the states $\rho_{XAA'}$ and $\sigma_{XAA'}$ should be fully determined by the SD of the individual states ρ_{XA}^y and σ_{XA}^y. Hence, if $\rho_{XA}^y \sim \sigma_{XA}^y$ for all y, then their mixtures $\rho_{XAA'}$ and $\sigma_{XAA'}$ should also be equivalent.

The last axiom is the following natural assumption: the SD of a c–q state does not increase by discarding subsystems; i.e.

$$\textbf{Axiom V} \quad \mathrm{Tr}_{X'A'}(\rho_{XX'AA'}) \prec \rho_{XX'AA'}. \tag{2.9}$$

The above axioms now lead to the formal definition of the preorder of SD:

Definition 2.1 (Preorder of SD). Let $\mathfrak{D}(XA)$ denote the set of c–q states on a classical system X and a quantum system A, and let

$$\mathfrak{D}_{cq} := \cup_{X,A} \mathfrak{D}(XA)$$

be the union over all finite-dimensional systems X and A. The preorder of SD is the smallest preorder relation on $\mathfrak{D}_{cq}$ that satisfies axioms I–V.[11]

In appendix A, we discuss a number of consequences of axioms I–V and the preorder of SD. There, we also define general resource measures to quantify SD, and we provide several examples.

We now use this preorder to define the set of free operations.

Definition 2.2 (Free operations of the RTSD). A map $\mathcal{N} \in \mathrm{CPTP}(XA \to X'A')$ is said to be a free operation if

$$\mathcal{N}(\rho_{XA}) \prec \rho_{XA} \tag{2.10}$$

for every c–q state ρ_{XA}. We denote the set of such free operations by $\mathfrak{F}(XA \to X'A')$.

The above definition of the free operations defines the RTSD. Specifically, we identify SD as a property of a composite physical system, shared between two parties, Xiao and Alice (with Xiao's systems being classical and Alice's being quantum), that can neither be generated nor increased by the set $\mathfrak{F}$ of free operations.

There is an important class of operations that play a central role in the RTSD. These are referred to as *CDS maps* and were first introduced in [GGH+18]. In fact, when the classical input and output systems of the resource theory are the same (i.e. $X = X'$), then the set of free operations, defined above, reduces to the set of CDS maps. This is stated in lemma 2.3 below. Before proceeding to the lemma, we introduce the notion of CDS maps in the next section.

2.1. Conditional doubly stochastic (CDS) maps

Consider the following problem: Xiao picks a state $\rho_x \in \{\rho_1, \dots, \rho_{|\mathcal{X}|}\}$ at random, with prior probability p_x, and then sends the state ρ_x to Alice through a noiseless quantum channel. Alice knows the probability distribution $\{p_x\}_x$ from which Xiao sampled the state ρ_x, but she does not know the value of x. Therefore, the overall state can be represented as a c–q state of the form:

$$\rho_{XA} = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x| \otimes \rho_x, \tag{2.11}$$

where X is a classical register to which Alice does not have access and A is a quantum register to which she does have access. What are the ways in which Alice can manipulate the state ρ_{XA}? There are two basic operations that she can perform:

(a) Alice can perform a generalized measurement on her system A.

(b) Alice can partially lose her knowledge of the distribution p_x, by performing random relabeling on the alphabet of the classical system $\mathcal{X} := \{1, 2, \dots, |\mathcal{X}|\}$.

That is, Alice can perform a generalized measurement on the quantum system A, and based on the outcome, say j, apply a random relabeling map $\mathcal{D}^{(j)}$ on the classical system X. Hence, the most general operation that Alice can perform is a CPTP map $\mathcal{N} \in \mathrm{CPTP}(XA \to XA')$ of the form

$$\mathcal{N} = \sum_j \mathcal{D}^j_{X \to X} \otimes \mathcal{E}^j_{A \to A'}, \tag{2.12}$$

where each $\mathcal{D}^j$ is a classical doubly stochastic channel and each $\mathcal{E}^j$ is a completely positive (CP) map such that $\sum_j \mathcal{E}^j$ is CPTP. Alternatively, since every doubly stochastic matrix can be expressed as a convex

---

[11] Note that an arbitrary preorder relation $\prec$ on $\mathfrak{D}_{cq}$ can be interpreted as a set $R \subset \mathfrak{D}_{cq} \times \mathfrak{D}_{cq}$, such that $\rho_{XA} \prec \sigma_{X'A'}$ if and only if $(\rho_{XA}, \sigma_{X'A'}) \in R$. The smallest preorder relation satisfying axioms I–V is then given by the intersection $\cap_{R \in \mathcal{I}} R$, with the index set $\mathcal{I}$ defined as $\mathcal{I} = \{R \mid R \text{ preorder relation on } \mathfrak{D}_{cq} \text{ satisfying axioms I–V}\}$. As the intersection of preorder relations is again a preorder relation, we see that $\cap_{R \in \mathcal{I}} R$ gives a well-defined preorder on $\mathfrak{D}_{cq}$.

combination of permutation matrices, there exist conditional probabilities $t_{z|j}$ such that

$$\mathcal{D}^j = \sum_{z=1}^{|\mathcal{X}|!} t_{z|j} \mathcal{P}^z, \quad t_{z|j} \geqslant 0 \ \forall \ z,j, \quad \sum_{z=1}^{|\mathcal{X}|!} t_{z|j} = 1 \ \forall \ j, \tag{2.13}$$

where $\mathcal{P}^z$ is the permutation channel defined by $\mathcal{P}^z(|x\rangle\langle x|) = |\pi_z(x)\rangle\langle\pi_z(x)|$, with π_z being one of the $|\mathcal{X}|!$ permutations. We therefore conclude that

$$\mathcal{N} = \sum_{z=1}^{|\mathcal{X}|!} \mathcal{P}^z_{X \to X} \otimes \tilde{\mathcal{E}}^z_{A \to A'}, \tag{2.14}$$

where $\tilde{\mathcal{E}}^z \equiv \sum_j t_{z|j} \mathcal{E}^j$. Hence, we can assume, without loss of generality, that in (2.12) the map $\mathcal{D}^z = \mathcal{P}^z$, with $z = 1, \ldots, |\mathcal{X}|!$, so that the maps that Alice can perform are given by

$$\mathcal{N} = \sum_{z=1}^{|\mathcal{X}|!} \mathcal{P}^z_{X \to X} \otimes \mathcal{E}^z_{A \to A'}. \tag{2.15}$$

We call such CPTP maps CDS maps.

The following lemma states that, when the input and output classical systems are the same (i.e. $X = X'$), the free operations for the RTSD, introduced in definition 2.2, are given by CDS maps, denoted by $CDS(XA \to XA')$.

Lemma 2.3. *For a classical system X and quantum systems A and A′, the following set equivalence holds*

$$CDS(XA \to XA') = \mathfrak{F}(XA \to XA'). \tag{2.16}$$

Proof. We first prove that

$$CDS(XA \to XA') \subseteq \mathfrak{F}(XA \to XA'). \tag{2.17}$$

Since partial trace and isometries acting on Alice's systems are free, it follows that every quantum instrument on Alice's side is free. Let

$$\mathcal{E}_{A \to A'A''}(\omega_A) := \sum_{y \in \mathcal{Y}} \mathcal{E}^y_{A \to A'}(\omega_A) \otimes |y\rangle\langle y|_{A''} \tag{2.18}$$

be a quantum instrument on Alice's side, and let $\rho_{XA} \in \mathfrak{D}(XA)$. Then, the action of the quantum instrument on ρ_{XA} yields the state

$$\sum_{y \in \mathcal{Y}} \mathcal{E}^y_{A \to A'}(\rho_{XA}) \otimes |y\rangle\langle y|_{A''}, \tag{2.19}$$

and from axiom I, we have, for all $y \in \mathcal{Y}$, the equivalence

$$\frac{\mathcal{E}^y_{A \to A'}(\rho_{XA})}{\text{Tr}\left[\mathcal{E}^y_{A \to A'}(\rho_{XA})\right]} \sim \frac{\mathcal{P}^y_{X \to X} \otimes \mathcal{E}^y_{A \to A'}(\rho_{XA})}{\text{Tr}\left[\mathcal{E}^y_{A \to A'}(\rho_{XA})\right]}, \tag{2.20}$$

where each $\mathcal{P}^y_{X \to X}$ is a permutation channel. Combining this with axiom IV, we conclude that the c−q state in (2.19) can further be transformed to

$$\sum_{y \in \mathcal{Y}} (\mathcal{P}^y_{X \to X} \otimes \mathcal{E}^y_{A \to A'})(\rho_{XA}) \otimes |y\rangle\langle y|_{A''}. \tag{2.21}$$

Finally, tracing system A'' yields the overall transformation

$$\rho_{XA} \mapsto \sum_{y \in \mathcal{Y}} (\mathcal{P}^y_{X \to X} \otimes \mathcal{E}^y_{A \to A'})(\rho_{XA}), \tag{2.22}$$

which is the general form of a CDS map. This completes the proof of (2.17).

Conversely, note that only axiom III is not covered by CDS maps. Therefore, the most general transformation $\mathcal{N}_{XA \to XA'} \in \mathfrak{F}(XA \to XA')$ has the form

$$\mathcal{N}_{XA \to XA'}(\rho_{XA}) = \sum_{y \in \mathcal{Y}} (\text{Tr}_{X'} \circ \mathcal{P}^y_{XX'} \otimes \mathcal{E}^y_{A \to A'})(\rho_{XA} \otimes \pi_{X'}), \tag{2.23}$$

where $\pi_{X'}$ is the maximally mixed state, $\{\mathcal{E}_y\}_y$ is a quantum instrument, and $\mathcal{P}^y_{XX'} \in \mathrm{CPTP}(XX' \to XX')$ are joint permutation channels. However, observe that

$$\mathcal{D}^y_{X \to X}(\omega_X) := (\mathrm{Tr}_{X'} \circ \mathcal{P}^y_{XX'})\,(\omega_X \otimes \pi_{X'}) \quad \forall\, \omega_X \in \mathfrak{D}(X), \tag{2.24}$$

is a classical doubly stochastic channel and therefore can be expressed as a convex combination of permutation channels. We therefore conclude that $\mathcal{N}_{XA \to XA'} \in \mathrm{CDS}(XA \to XA')$. $\qquad\square$

Lemma 2.3 above demonstrates that if the dimension of the classical system is fixed, then the free operations in the resource theory of SD are CDS maps. However, we point out that the lemma above can also be used to characterize $\mathfrak{F}(XA \to X'A')$ where $|X| \neq |X'|$. In particular, observe that

$$\rho_{XA} \xrightarrow{\mathfrak{F}} \sigma_{X'A'} \quad \Leftrightarrow \quad \pi_{X'} \otimes \rho_{XA} \xrightarrow{\mathrm{CDS}} \pi_X \otimes \sigma_{X'A'} \tag{2.25}$$

for all $\rho_{XA}, \sigma_{X'A'} \in \mathfrak{D}_{\mathrm{cq}}$ because the maximally mixed states π_X and $\pi_{X'}$ are free. In the general case, when $|X| \neq |X'|$, the set $\mathfrak{F}(AX \to A'X')$ can be viewed as a special subset of conditional thermal operations [NG17], corresponding to a thermodynamical system with a completely degenerate Hamiltonian. We therefore call it the set of *conditional noisy operations*.

In the rest of the paper, we focus on the case in which the input classical system X has the same dimension as the output classical system X', and furthermore, we constrain both of them to have dimension equal to two. Thus, according to lemma 2.3, the set of free operations reduces to CDS maps in all of our discussions that follow.

Moreover, in addition to CDS maps, we will also consider the set

$$\mathrm{CPTP}_A := \{\mathrm{id} \otimes \mathcal{E} \mid \mathcal{E}\ \mathrm{CPTP\ on\ system}\ A\} \tag{2.26}$$

as a possible set of transformations. Note that CPTP_A has the clear physical interpretation of applying a fixed quantum channel onto the quantum part of the c–q state without changing its classical probability distribution. Furthermore, it is clear from the definition of CDS maps that we have the inclusion

$$\mathrm{CPTP}_A \subset \mathrm{CDS}. \tag{2.27}$$

The following lemma shows that the minimum error probability $p_{\mathrm{err}}(\rho_{XA})$ can only be increased by application of a CDS map.

Lemma 2.4 (Monotonicity of minimum error probability under CDS maps). *Let* $\mathcal{N} \in CDS(XA \to XA')$ *and* ρ_{XA} *be a c–q state. Then*

$$p_{\mathrm{err}}(\mathcal{N}(\rho_{XA})) \geqslant p_{\mathrm{err}}(\rho_{XA}), \tag{2.28}$$

where $p_{\mathrm{err}}(\rho_{XA}) := 1 - p_{\mathrm{guess}}(X|A)$, *with* $p_{\mathrm{guess}}(X|A)$ *defined in* (2.2).

Proof. The guessing probability $p_{\mathrm{guess}}(X|A)$ can be written as follows [KRS09]:

$$p_{\mathrm{guess}}(X|A) = 2^{\inf_{\omega_A} D_{\max}(\rho_{XA} \| I_X \otimes \omega_A)},$$

where $D_{\max}$ denotes the max-relative entropy, which for a state ρ and a positive semi-definite operator σ is defined as follows [Dat09]:

$$D_{\max}(\rho \| \sigma) := \inf\{\lambda : \rho \leqslant 2^\lambda \sigma\}. \tag{2.29}$$

Thus, we conclude that

$$p_{\mathrm{err}}(\mathcal{N}(\rho_{XA})) = 1 - 2^{\inf_{\sigma_{A'}} D_{\max}(\mathcal{N}(\rho_{XA}) \| I_X \otimes \sigma_{A'})} \geqslant 1 - 2^{\inf_{\omega_A} D_{\max}(\mathcal{N}(\rho_{XA}) \| \mathcal{N}(I_X \otimes \omega_A))}$$

$$\geqslant 1 - 2^{\inf_{\omega_A} D_{\max}(\rho_{XA} \| I_X \otimes \omega_A)} = p_{\mathrm{err}}(\rho_{XA}).$$

The first inequality follows from the fact that for each quantum state ω_A on system A, the image under the CDS map $\mathcal{N}$ can be written as $\mathcal{N}(I_X \otimes \omega_A) = I_X \otimes \sigma_{A'}$ with $\sigma_{A'}$ a state on system A' (which is evident from the definition of CDS maps). The second inequality follows from the DPI for the max-relative entropy [Dat09]. $\qquad\square$

The above lemma immediately leads to a natural choice of a measure of SD for the particular case of the RTSD that we study in this paper, namely, one in which the dimension of the classical system X is fixed to $|X| = 2$. This measure was mentioned in (1.3), and we recall its definition here:

Definition 2.5. For a fixed dimension $|X| = 2$, we define

$$\mathrm{SD}(\rho_{XA}) := -\log\left(2p_{\mathrm{err}}(\rho_{XA})\right). \tag{2.30}$$

This function is equal to zero on free states and behaves monotonically under CDS maps, as is evident from lemma 2.4 above.

Lemma 2.6. *For an ensemble $\{p_x, \rho_A^x\}_{x \in \mathcal{X}}$, the following lower bound holds for the guessing probability:*

$$p_{guess}(X|A) \geqslant \frac{1}{|\mathcal{X}|}, \tag{2.31}$$

with $p_{guess}(X|A)$ defined in (2.2). This lower bound is saturated (i.e. $p_{guess}(X|A) = \frac{1}{|\mathcal{X}|}$) if and only if

$$\exists \, \sigma_A \in \mathcal{D}(A) \quad \forall \, x \in \mathcal{X}, \qquad \rho_x = \sigma_A, \quad and \quad p_x = \frac{1}{|\mathcal{X}|}. \tag{2.32}$$

Proof. It is known that [KRS09]

$$p_{guess}(X|A) = 2^{\inf_{\omega_A} D_{\max}(\rho_{XA} \| I_X \otimes \omega_A)} = \frac{1}{|\mathcal{X}|} 2^{\inf_{\omega_A} D_{\max}(\rho_{XA} \| \pi_X \otimes \omega_A)}, \tag{2.33}$$

where the infimum is over every state ω_A. In the above, ρ_{XA} is the c–q state corresponding to the ensemble and is defined in (2.1), and $\pi_X = I_X / |\mathcal{X}|$ is the completely mixed state. The lower bound in (2.31) is then a direct consequence of (2.33) and the fact that $D_{\max}(\rho \| \sigma) \geqslant 0$ for states ρ and σ.

As a consequence of (2.33), we also conclude that

$$p_{guess}(X|A) = \frac{1}{|\mathcal{X}|} \iff \inf_{\omega_A} D_{\max}(\rho_{XA} \| \pi_X \otimes \omega_A) = 0. \tag{2.34}$$

Now note that the infimum on the right-hand side of (2.34) is actually a minimum (due to the finiteness of $\inf_{\omega_A} D_{\max}(\rho_{XA} \| \pi_X \otimes \omega_A)$ and continuity of $2^\lambda \sigma$ in λ). Therefore,

$$p_{guess}(X|A) = \frac{1}{|\mathcal{X}|} \iff \exists \, \sigma_A \text{ s.t. } \rho_{XA} = \pi_X \otimes \sigma_A, \tag{2.35}$$

where we also employed the property of max-relative entropy that, for states ρ and σ, $D_{\max}(\rho \| \sigma) = 0$ if and only if $\rho = \sigma$. □

Remark 2.7. By rearranging (2.33), we find that

$$\log(|\mathcal{X}| p_{guess}(X|A)) = \inf_{\omega_A} D_{\max}(\rho_{XA} \| \pi_X \otimes \omega_A). \tag{2.36}$$

This quantity is a measure of symmetric distinguishability alternative to $SD(\rho_{XA})$, as defined in (1.3). Indeed, this measure of SD has the appealing feature that it is the max-relative entropy from the state ρ_{XA} of interest to the set of free states. As is common in quantum resource theories [CG19], one could then define an infinite number of SD measures based on the generalized divergence of the state of interest with the set of free states.

3. Some key ingredients of the resource theory of symmetric distinguishability

3.1. Infinite-resource states and golden units

In the following sections, we study the information-theoretic tasks of distillation and dilution within the framework of the RTSD, with respect to the golden unit of definition 3.4 below. We consider two different choices of free operations: (i) CPTP maps on the quantum system A and the identity channel on the classical system, which we denote as CPTP$_A$, and (ii) CDS maps, which were introduced in the previous section. The reason for considering both of these is that they lead to novel and interesting results. Also, in some cases, a proof for one choice of free operations follows as a simple corollary from that for the other choice. We refer to these tasks as SD-distillation and SD-dilution, respectively. We study the exact and approximate one-shot cases, as well as the asymptotic case for both of these tasks.

In this section, we prove certain key results that serve as prerequisites for the above study; they involve *infinite-resource states* and *golden units*—notions that were introduced in section 1. We recall their definitions before stating the relevant results. As mentioned in the introduction, the basic objects of the RTSD are elementary quantum information sources represented by

$$\rho_{XA} := p|0\rangle\langle0| \otimes \rho_0 + (1 - p)|1\rangle\langle1| \otimes \rho_1, \tag{3.1}$$

where $p \in [0, 1]$ is a prior probability and ρ_0 and ρ_1 are quantum states.

The main operational quantity associated with such a state, in the context of symmetric hypothesis testing, is the minimum error probability of Bayesian state discrimination of the states ρ_0 and ρ_1:

$$p_{\mathrm{err}}(\rho_{XA}) = \min_{0 \leqslant \Lambda \leqslant \mathbb{1}} \left(p\,\mathrm{Tr}(\Lambda\rho_0) + (1-p)\mathrm{Tr}((\mathbb{1}-\Lambda)\rho_1) \right). \tag{3.2}$$

Remark 3.1. Since the c–q state ρ_{XA} is also represented by the quantum box (p, ρ_0, ρ_1) (as mentioned in the introduction), we sometimes use the notation $p_{\mathrm{err}}(p, \rho_0, \rho_1)$ instead of $p_{\mathrm{err}}(\rho_{XA})$ in the following.

The following well-known theorem gives an explicit expression for this minimum error probability [Hel67, Hel69, Hol72].

Theorem 3.2 (Helstrom–Holevo theorem). *For a c–q state ρ_{XA} of the form in* (3.1), *the following equality holds*

$$p_{\mathrm{err}}(\rho_{XA}) = \frac{1}{2} \left(1 - \| p\rho_0 - (1-p)\rho_1 \|_1 \right). \tag{3.3}$$

Recall from section 1 that a c–q state ρ_{XA} is said to be an *infinite-resource state* if $p_{\mathrm{err}}(\rho_{XA}) = 0$, which is equivalent to the quantum states ρ_0 and ρ_1 having mutually orthogonal support. Hence, the symmetric distinguishability given by $\mathrm{SD}(\rho_{XA}) = -\log\left(2p_{\mathrm{err}}(\rho_{XA})\right)$ is infinite in this case.

The following lemma shows that we can transform any infinite-resource state to any other c–q state of the form in (3.1) via CDS maps.

Lemma 3.3. *Let ω_{XA} be an infinite-resource state, and let σ_{XB} be a general c–q state. Then there exists a CDS map $\mathcal{N} : XA \to XB$ such that*

$$\mathcal{N}(\omega_{XA}) = \sigma_{XB}. \tag{3.4}$$

Proof. We write the c–q states explicitly as

$$\omega_{XA} = p\,|0\rangle\langle 0| \otimes \omega_0 + (1-p)\,|1\rangle\langle 1| \otimes \omega_1,$$

$$\sigma_{XB} = q\,|0\rangle\langle 0| \otimes \sigma_0 + (1-q)\,|1\rangle\langle 1| \otimes \sigma_1,$$

for some $p, q \in [0,1]$ and quantum states $\omega_0, \omega_1, \sigma_0$, and σ_1. As ω_{XA} is an infinite-resource state, and hence ω_0 and ω_1 have mutually orthogonal supports, we can pick a POVM $\{\Lambda, \mathbb{1} - \Lambda\}$ such that $\mathrm{Tr}(\Lambda\omega_0) = \mathrm{Tr}((\mathbb{1}-\Lambda)\omega_1) = 1$ and consequently $\mathrm{Tr}(\Lambda\omega_1) = \mathrm{Tr}((\mathbb{1}-\Lambda)\omega_0) = 0$. Consider a pair $(\mathcal{E}_0, \mathcal{E}_1)$ of quantum operations, i.e. CP, trace non-increasing linear maps that sum to a CPTP map, defined as follows:

$$\mathcal{E}_0(\cdot) = q\,\mathrm{Tr}(\Lambda\cdot)\sigma_0 + (1-q)\mathrm{Tr}((\mathbb{1}-\Lambda)\cdot)\sigma_1, \qquad \mathcal{E}_1(\cdot) = q\,\mathrm{Tr}((\mathbb{1}-\Lambda)\cdot)\sigma_0 + (1-q)\mathrm{Tr}(\Lambda\cdot)\sigma_1,$$

and consider the corresponding CDS map

$$\mathcal{N} = \mathrm{id}_X \otimes \mathcal{E}_0 + \mathcal{F}_X \otimes \mathcal{E}_1,$$

where $\mathcal{F}_X$ denotes the flip channel on the classical system X. We then immediately get

$$\mathcal{N}(\omega_{XA}) = \sigma_{XB},$$

which concludes the proof. $\qquad\square$

As mentioned in the introduction, it is useful to consider a particular class of c–q states that lead naturally to a clear definition of the fundamental tasks of distillation and dilution in the RTSD. These states are parametrized by $M \in [1, \infty]$ and $q \in (0, 1)$, and for M large enough have SD equal to $\log M$. We refer to such a state as an (M, q)-golden unit. It is defined as follows:

Definition 3.4 (Golden unit). We choose the following class of c–q states of a composite system XQ, where Q is a qubit. Each state is labeled by a parameter $M \in [1, \infty]$ and a probability $q \in (0, 1)$ and is defined as follows:

$$\gamma_{XQ}^{(M,q)} := q\,|0\rangle\langle 0|_X \otimes \pi_M + (1-q)\,|1\rangle\langle 1| \otimes \sigma^{(1)}\pi_M\sigma^{(1)}, \tag{3.5}$$

where

$$\pi_M := \left(1 - \frac{1}{2M}\right)|0\rangle\langle 0| + \frac{1}{2M}|1\rangle\langle 1|$$

is a state of a qubit Q and $\sigma^{(1)}$ denotes the Pauli-x matrix. We call the state $\gamma_{XQ}^{(M,q)}$ an (M, q)-golden unit. Note that for $M = \infty$, we have $\pi_\infty = |0\rangle\langle 0|$ and hence the golden unit reduces to an infinite-resource state.

The goodness of this choice of the golden unit lies in the fact that its SD has a useful scaling property, as stated in the following lemma.

Lemma 3.5. *For all $M \in [1, \infty)$ such that $2M \geqslant \max\{1/q, 1/(1-q)\}$ and $q \in (0, 1)$,*

$$p_{\text{err}}(\gamma_{XQ}^{(M,q)}) = \frac{1}{2M}, \tag{3.6}$$

and hence its symmetric distinguishability is given by $\text{SD}(\gamma_{XQ}^{(M,q)}) = \log M$. *For $M = \infty$, and hence $p_{\text{err}}(\gamma_{XQ}^{(M,q)}) = 0$, the SD of $\gamma_{XQ}^{(M,q)}$ is infinite.*

Proof. In the case $M = \infty$, we trivially have $p_{\text{err}}(\gamma_{XQ}^{(M,q)}) = 0$ as $\pi_\infty = |0\rangle\langle 0|$ and $\sigma^{(1)}\pi_\infty\sigma^{(1)} = |1\rangle\langle 1|$ are orthogonal. For the case $M \in [1, \infty)$, we use the Helstrom–Holevo theorem to conclude that

$$p_{\text{err}}(\gamma_{XQ}^{(M,q)}) = \frac{1}{2}\left(1 - \left\| q\,\pi_M - (1-q)\sigma^{(1)}\pi_M\sigma^{(1)} \right\|_1\right)$$

$$= \frac{1}{2}\left(1 - \left\| \left(q - \frac{1}{2M}\right)|0\rangle\langle 0| - \left(1 - q - \frac{1}{2M}\right)|1\rangle\langle 1| \right\|_1\right)$$

$$= \frac{1}{2}\left(1 - \left| q - \frac{1}{2M}\right| - \left| 1 - q - \frac{1}{2M}\right|\right). \tag{3.7}$$

Hence, for $2M \geqslant \max\{1/q, 1/(1-q)\}$, we get

$$p_{\text{err}}(\gamma_{XQ}^{(M,q)}) = \frac{1}{2M}, \tag{3.8}$$

and hence $\text{SD}(\gamma_{XQ}^{(M,q)}) = -\log\left(2p_{\text{err}}(\gamma_{XQ}^{(M,q)})\right) = \log M$. $\qquad\square$

Furthermore, we note that for $q_1, q_2 \in [0, 1]$ such that the corresponding distribution $\vec{q}_1 := (q_1, 1 - q_1)$ is majorised by $\vec{q}_2 := (q_2, 1 - q_2)$ (i.e. $\vec{q}_1 \prec \vec{q}_2$), the golden unit $\gamma_{XQ}^{(M,q_1)}$ is dominated by $\gamma_{XQ}^{(M,q_2)}$ in the preorder of SD (see definition 2.1). This is the statement of the following lemma.

Lemma 3.6. *Let $M \in [1, \infty]$ and $q_1, q_2 \in [0, 1]$ be such that the corresponding distribution vectors satisfy $\vec{q}_1 \prec \vec{q}_2$ in majorisation order. Then*

$$\gamma_{XQ}^{(M,q_1)} \prec \gamma_{XQ}^{(M,q_2)}, \tag{3.9}$$

and hence $\gamma_{XQ}^{(M,q_2)}$ can be transformed to $\gamma_{XQ}^{(M,q_1)}$ via a CDS map.

Proof. As $\vec{q}_1 \prec \vec{q}_2$, there exists a $\lambda \in [0, 1]$ such that $\lambda q_2 + (1 - \lambda)(1 - q_2) = q_1$. Now consider the CDS map

$$\mathcal{N} = \lambda\, \text{id}_X \otimes \text{id}_Q + (1 - \lambda)\mathcal{F}_X \otimes \mathcal{F}_Q,$$

where $\mathcal{F}_X$ and $\mathcal{F}_Q$ denotes the flip channel on systems X and Q, respectively. This directly gives

$$\mathcal{N}\left(\gamma_{XQ}^{(M,q_2)}\right) = |0\rangle\langle 0| \otimes \left(\lambda q_2\, \pi_M + (1 - \lambda)(1 - q_2)\,\mathcal{F}_Q(\sigma^{(1)}\pi_M\sigma^{(1)})\right)$$

$$+ |1\rangle\langle 1| \otimes \left((1 - \lambda)q_2\,\mathcal{F}_Q(\pi_M) + \lambda(1 - q_2)\,\sigma^{(1)}\pi_M\sigma^{(1)}\right)$$

$$= q_1 |0\rangle\langle 0| \otimes \pi_M + (1 - q_1)|1\rangle\langle 1| \otimes \sigma^{(1)}\pi_M\sigma^{(1)} = \gamma_{XQ}^{(M,q_1)},$$

which finishes the proof. $\qquad\square$

When we consider CDS maps as free operations, it suffices to focus on the case $q = 1/2$. We denote the corresponding golden unit simply as $\gamma_{XQ}^{(M)}$, and call it the *M-golden unit*. For future reference, we write it out explicitly:

$$\gamma_{XQ}^{(M)} \equiv \gamma_{XQ}^{(M,1/2)} = \frac{1}{2}|0\rangle\langle 0| \otimes \pi_M + \frac{1}{2}|1\rangle\langle 1| \otimes \sigma^{(1)}\pi_M\sigma^{(1)}, \tag{3.10}$$

where

$$\pi_M = \left(1 - \frac{1}{2M}\right)|0\rangle\langle 0| + \frac{1}{2M}|1\rangle\langle 1|.$$

Remark 3.7. Note that the golden unit $\gamma_{XQ}^{(M)}$ is equivalent under CDS maps to $\pi_M \otimes \omega$, for an arbitrary quantum state ω. To see that, consider first the CDS map

$$\mathcal{N} = \text{id}_X \otimes \langle 0| \cdot |0\rangle\, \omega + \mathcal{F}_X \otimes \langle 1| \cdot |1\rangle\, \omega. \tag{3.11}$$

and note that $\mathcal{N}(\gamma_{XQ}^{(M)}) = \pi_M \otimes \omega$. To see the other direction, consider the CDS map

$$\mathcal{M} = \frac{1}{2}\left(\mathrm{id}_X \otimes \mathrm{Tr}(\cdot)\,|0\rangle\langle 0| + \mathcal{F}_X \otimes \mathrm{Tr}(\cdot)\,|1\rangle\langle 1|\right), \tag{3.12}$$

which gives $\mathcal{M}(\pi_M \otimes \omega) = \frac{1}{2}\left(\pi_M \otimes |0\rangle\langle 0| + \sigma^{(1)}\pi_M\sigma^{(1)} \otimes |1\rangle\langle 1|\right) = \gamma_{XQ}^{(M)}$.

3.2. A suitable error measure for approximate transformation tasks

In this section, we introduce the notion of a *minimum conversion error* for the transformation of one c–q state to another (both of the form defined in (3.1)) via free operations. For the transformation $\rho_{XA} \mapsto \sigma_{XB}$, we denote this quantity as $d'_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB})$. The latter is defined in terms of a *scaled trace distance*, which we denote as $D'(\rho_{XA}, \sigma_{XB})$. We also discuss some of the properties of the above quantities. These quantities are then used to define the *one-shot approximate distillable-SD* and the *one-shot approximate SD-cost* in the following sections.

Definition 3.8. For general c–q states ρ_{XA} and σ_{XA}, define the *scaled trace distance*

$$D'(\rho_{XA}, \sigma_{XA}) = \begin{cases} \dfrac{\frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1}{p_{\mathrm{err}}(\sigma_{XA})}, & \text{if } p_{\mathrm{err}}(\sigma_{XA}) > 0, \\ 0, & \text{if } p_{\mathrm{err}}(\sigma_{XA}) = 0 \text{ and } \rho_{XA} = \sigma_{XA}, \\ \infty, & \text{if } p_{\mathrm{err}}(\sigma_{XA}) = 0 \text{ and } \rho_{XA} \neq \sigma_{XA}. \end{cases} \tag{3.13}$$

Remark 3.9. Note that the scaling factor in the definition of the scaled trace distance $D'(\rho_{XA}, \sigma_{XA})$ depends on the state σ_{XA} in the second slot.

Definition 3.10. For a set of free operations denoted by FO, we define the minimum conversion error corresponding to the scaled trace distance $D'(\cdot, \cdot)$ as follows:

$$d'_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}) = \min_{\mathcal{A} \in \mathrm{FO}} D'(\mathcal{A}(\rho_{XA}), \sigma_{XB}), \tag{3.14}$$

with ρ_{XA} and σ_{XB} being general c–q states on the classical system X and the quantum systems A and B, respectively.

Proposition 3.11. *Let $\rho_{XA} \equiv (p, \rho_0, \rho_1)$ and $\sigma_{XB} \equiv (q, \sigma_0, \sigma_1)$ be c–q states. The minimum conversion error $d'_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB})$ for $\mathrm{FO} \in \{\mathrm{CDS}, \mathrm{CPTP}_A\}$ can be written as*

$$d'_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}) = \frac{\min\limits_{\mathcal{M} \in \mathrm{FO}} \|\mathcal{M}(\rho_{XA}) - \sigma_{XB}\|_1}{\min\limits_{\mathcal{N} \in \mathrm{FO}} \left\|\mathcal{N}(\sigma_{XB}) - \gamma_{XQ}^{(\infty,q)}\right\|_1}, \tag{3.15}$$

where $\gamma_{XQ}^{(\infty,q)} = q\,|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q + (1-q)\,|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q$ is an infinite resource state.

Proof. By inspecting the definition of $d'_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB})$, we see that it remains to prove the following:

$$p_{\mathrm{err}}(\sigma_{XA}) = \min_{\mathcal{N} \in \mathrm{FO}} \left\|\mathcal{N}(\sigma_{XB}) - \gamma_{XQ}^{(\infty,q)}\right\|_1, \tag{3.16}$$

where FO is either CPTP_A or CDS. We leave the proof of the equality above to appendix B (see lemmas B.3 and B.4 therein). $\qquad\square$

Remark 3.12. The reason for considering the scaled trace distance $D'(\cdot, \cdot)$ instead of the usual trace distance as an error measure for transformations in the RTSD is that using the latter would allow for the unreasonable possibility of a finite-resource state being arbitrarily close to an infinite-resource state. In particular, as any infinite-resource state can be transformed to any other c–q state via CDS maps (see lemma 3.3), this would imply that, for any finite allowed error measured in trace norm, the transformation

$$\rho_{XA}^{(n)} \equiv \left(p, \rho_0^{\otimes n}, \rho_1^{\otimes n}\right) \mapsto \sigma_{XB}^{(m)} \equiv \left(q, \sigma_0^{\otimes m}, \sigma_1^{\otimes m}\right) \tag{3.17}$$

would be possible at an infinite rate (as long as $\rho_0 \neq \rho_1$). To see this, for all $n \in \mathbb{N}$, pick a POVM $\{\Lambda_n, \mathbb{1} - \Lambda_n\}$ on the composite system of n copies of system A such that both type I and II error probabilities corresponding to the source $(p, \rho_0^{\otimes n}, \rho_1^{\otimes n})$ vanish asymptotically, i.e.

$$\lim_{n \to \infty} \mathrm{Tr}\left((\mathbb{1} - \Lambda_n)\rho_0^{\otimes n}\right) = \lim_{n \to \infty} \mathrm{Tr}\left(\Lambda_n \rho_1^{\otimes n}\right) = 0. \tag{3.18}$$

This is possible because $\rho_0 \neq \rho_1$. Hence, considering the infinite-resource state

$$\omega_{XQ} = p \, |0\rangle\langle 0| \otimes |0\rangle\langle 0| + (1-p) \, |1\rangle\langle 1| \otimes |1\rangle\langle 1|,$$

with the quantum system being a qubit Q, and the measure-prepare channel $\mathcal{E}_n(\cdot) = \mathrm{Tr}(\Lambda_n \cdot) \, |0\rangle\langle 0| + \mathrm{Tr}((\mathbb{1} - \Lambda_n)\cdot) \, |1\rangle\langle 1|$, we see that

$$\lim_{n\to\infty} \frac{1}{2} \left\| (\mathrm{id} \otimes \mathcal{E}_n)(\rho_{XA}^{(n)}) - \omega_{XQ} \right\|_1 = \lim_{n\to\infty} \left(p \, \mathrm{Tr}\big((\mathbb{1} - \Lambda_n)\rho_0^{\otimes n}\big) + (1-p)\mathrm{Tr}\big(\Lambda_n \rho_1^{\otimes n}\big) \big) \right) = 0.$$

Therefore, as the infinite-resource state ω_{XQ} can be transformed to any other c–q state without error, we see that for every $\varepsilon > 0$, we can pick $n \in \mathbb{N}$ large enough such that for every c–q state $\sigma_{XB} \equiv (q, \sigma_0, \sigma_1)$ and $m \in \mathbb{N}$ there exists a CDS map $\mathcal{N}$ such that

$$\frac{1}{2} \left\| \mathcal{N}(\rho_{XA}^{(n)}) - \sigma_{XB}^{(m)} \right\|_1 \leqslant \varepsilon. \tag{3.19}$$

Hence, the transformation $\rho_{XA}^{(n)} \equiv \big(p, \rho_0^{\otimes n}, \rho_1^{\otimes n}\big) \mapsto \sigma_{XB}^{(m)} \equiv \big(q, \sigma_0^{\otimes m}, \sigma_1^{\otimes m}\big)$ is possible at an infinite rate and with arbitrarily small error measured in trace distance. For $q = p$, this transformation can also be performed at an infinite rate by just using CPTP$_A$ operations.

Note that $D'(\cdot, \cdot)$ is exactly defined in such a way that for every infinite-resource state ω_{XA} and $\rho_{XA} \neq \omega_{XA}$, we have $D'(\rho_{XA}, \omega_{XA}) = \infty$. Therefore, the problem discussed above, in which we obtain unreasonable infinite rates in the transformations in the RTSD, does not occur when we choose $D'(\cdot, \cdot)$ as the error measure. Moreover, in the following, we see that $D'(\cdot, \cdot)$ has many desirable properties that lead to reasonable asymptotic rates in SD distillation, SD dilution, and the transformation of general elementary quantum sources.

The scaled trace distance D' satisfies the DPI under CDS maps:

Lemma 3.13 (DPI for D' under CDS maps). *Let ρ_{XA}, σ_{XA} be two c–q states and $\mathcal{N}$ a CDS map. Then*

$$D'(\mathcal{N}(\rho_{XA}), \mathcal{N}(\sigma_{XA})) \leqslant D'(\rho_{XA}, \sigma_{XA}). \tag{3.20}$$

Proof. The statement directly follows from the DPI for the trace distance under general CPTP maps, i.e.

$$\frac{1}{2}\|\mathcal{N}(\rho_{XA}) - \mathcal{N}(\sigma_{XA})\|_1 \leqslant \frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1 \tag{3.21}$$

and the monotonicity of the minimum error probability under CDS maps proven in lemma 2.4, i.e.

$$p_{\mathrm{err}}(\mathcal{N}(\sigma_{XA})) \geqslant p_{\mathrm{err}}(\sigma_{XA}). \tag{3.22}$$

This concludes the proof. □

The following lemma now establishes a bound relating the minimum error probabilities of two c–q states ρ_{XA} and σ_{XA}, involving a multiplicative term related to their scaled trace distance D'. This lemma is the key ingredient for proving all converses in approximate asymptotic SD-distillation, SD-dilution, and the transformation of general elementary quantum sources.

Lemma 3.14. *Let ρ_{XA} and σ_{XA} be c–q states such that $D'(\rho_{XA}, \sigma_{XA})$ is finite. Then*

$$p_{\mathrm{err}}(\rho_{XA}) \leqslant \big(D'(\rho_{XA}, \sigma_{XA}) + 1\big) p_{\mathrm{err}}(\sigma_{XA}). \tag{3.23}$$

Proof. First, it is helpful to note that by writing the c–q states explicitly as

$$\rho_{XA} = p \, |0\rangle\langle 0| \otimes \rho_0 + (1-p) \, |1\rangle\langle 1| \otimes \rho_1, \tag{3.24}$$

$$\sigma_{XA} = q \, |0\rangle\langle 0| \otimes \sigma_0 + (1-q) \, |1\rangle\langle 1| \otimes \sigma_1, \tag{3.25}$$

they can be block-diagonalised in the same basis and hence we can write

$$\frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1 = \frac{1}{2}\big(\|p\rho_0 - q\sigma_0\|_1 + \|(1-p)\rho_1 - (1-q)\sigma_1\|_1\big). \tag{3.26}$$

We can hence bound the minimum error probability of ρ_{XA} as

$$p_{\mathrm{err}}(\rho_{XA}) = \min_{0\leqslant\Lambda\leqslant\mathbb{1}} \big(p \, \mathrm{Tr}(\Lambda\rho_0) + (1-p)\mathrm{Tr}((\mathbb{1} - \Lambda)\rho_1)\big)$$

$$= \min_{0\leqslant\Lambda\leqslant\mathbb{1}} \big(q \, \mathrm{Tr}(\Lambda\sigma_0) + (1-q)\mathrm{Tr}((\mathbb{1} - \Lambda)\sigma_1)\big)$$

$$+ \operatorname{Tr}\left(\Lambda(p\rho_0 - q\sigma_0)\right) + \operatorname{Tr}\left((\mathbb{1} - \Lambda)((1-p)\rho_1 - (1-q)\sigma_1)\right)\right)$$

$$= \min_{0 \leqslant \Lambda \leqslant \mathbb{1}} \left(q \operatorname{Tr}(\Lambda\sigma_0) + (1-q)\operatorname{Tr}((\mathbb{1} - \Lambda)\sigma_1)\right.$$

$$+ \operatorname{Tr}\left((|0\rangle\langle0| \otimes \Lambda + |1\rangle\langle1| \otimes (\mathbb{1} - \Lambda))(\rho_{XA} - \sigma_{XA})\right)\right)$$

$$\leqslant \min_{0 \leqslant \Lambda \leqslant \mathbb{1}} \left(q \operatorname{Tr}(\Lambda\sigma_0) + (1-q)\operatorname{Tr}((\mathbb{1} - \Lambda)\sigma_1)\right) + \frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1$$

$$= p_{\mathrm{err}}(\sigma_{XA}) + \frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1$$

$$= \left(D'(\rho_{XA}, \sigma_{XA}) + 1\right) p_{\mathrm{err}}(\sigma_{XA}). \tag{3.27}$$

This concludes the proof. $\qquad\square$

3.2.1. Semi-definite program for the scaled trace distance D'

We now prove that the scaled trace distance $D'(\cdot, \cdot)$ can be calculated by means of a SDP. SDPs can be computed efficiently by numerical solvers [VB96]. As semi-definite programming is a powerful theoretical and numerical tool for quantum information theory, with a plethora of applications, we expect that the following SDP characterizations of $D'(\cdot, \cdot)$ may be useful for a further understanding of this quantity.

Proposition 3.15. *For general c–q states ρ_{XA} and σ_{XA} with $p_{\mathrm{err}}(\sigma_{XA}) > 0$, the scaled trace distance, $D'(\rho_{XA}, \sigma_{XA})$, in definition 3.8 is given by the following SDP:*

$$\max t$$

$$s.t. \quad -I_{XA} \leqslant L_{XA} \leqslant I_{XA},$$

$$-tI_A \leqslant P_A \leqslant tI_A, \tag{3.28}$$

$$t - \operatorname{Tr} P_A(q\sigma_0 - (1-q)\sigma_1) = \operatorname{Tr} L_{XA}(\rho_{XA} - \sigma_{XA}).$$

The dual SDP is as follows:

$$\min \operatorname{Tr}(B_{XA} + C_{XA})$$

$$s.t. \ B_{XA}, C_{XA}, D_A, E_A \geqslant 0, s \in \mathbb{R},$$

$$B_{XA} - C_{XA} = s(\rho_{XA} - \sigma_{XA}), \tag{3.29}$$

$$D_A - E_A = s(q\sigma_0 - (1-q)\sigma_1),$$

$$\operatorname{Tr}(D_A + E_A) \leqslant s - 1.$$

If $p_{\mathrm{err}}(\sigma_{XA}) > 0$, then strong duality holds, so that the optimal value in (3.28) is equal to the optimal value in (3.29).

The proof of proposition 3.15 can be found in appendix C.

3.2.2. Semi-definite program for the minimum conversion error

The one-shot transformation task from a source $\rho_{XA} := p|0\rangle\langle0| \otimes \rho_0 + (1-p)|1\rangle\langle1| \otimes \rho_1$ to a target $\sigma_{XA'} := q|0\rangle\langle0| \otimes \sigma_0 + (1-q)|1\rangle\langle1| \otimes \sigma_1$ using CDS as the set of free operations can be phrased as the following optimization task:

$$d'_{\mathrm{CDS}}(\rho_{XA} \mapsto \sigma_{XA'}) = \min_{\mathcal{N}_{XA \to XA'} \in \mathrm{CDS}} D'(\mathcal{N}_{XA \to XA'}(\rho_{XA}), \sigma_{XA'}). \tag{3.30}$$

We now prove that the minimum conversion error in (3.30) can be calculated by means of a SDP.

Proposition 3.16. *The minimum conversion error in (3.30) can be evaluated by the following SDP:*

$$\min_{\substack{B_{XA'}, C_{XA'}, D_{A'}, E_{A'} \geqslant 0, \\ \Omega^0_{AA'}, \Omega^1_{AA'} \geqslant 0, s \geqslant 1}} \left\{ \begin{array}{c} \operatorname{Tr}[B_{XA'} + C_{XA'}]: \\ B_{XA'} - C_{XA'} = \tau_{XA'} - s\sigma_{XA'}, \\ D_{A'} - E_{A'} = s(q\sigma_0 - (1-q)\,\sigma_1), \\ \operatorname{Tr}[D_{A'} + E_{A'}] \leqslant s - 1, \\ \operatorname{Tr}_{A'}[\Omega^0_{AA'} + \Omega^1_{AA'}] = sI_A, \\ \langle0|_X \tau_{XA'}|0\rangle_X = \operatorname{Tr}_A[(p\rho_A^0)^{\mathrm{T}}\Omega^0_{AA'}] + \operatorname{Tr}_A[((1-p)\,\rho_A^1)^{\mathrm{T}}\Omega^1_{AA'}], \\ \langle1|_X \tau_{XA'}|1\rangle_X = \operatorname{Tr}_A[(p\rho_A^0)^{\mathrm{T}}\Omega^1_{AA'}] + \operatorname{Tr}_A[((1-p)\,\rho_A^1)^{\mathrm{T}}\Omega^0_{AA'}], \\ \langle1|_X \tau_{XA'}|0\rangle_X = \langle0|_X \tau_{XA'}|1\rangle_X = 0 \end{array} \right\}. \tag{3.31}$$

The proof of proposition 3.16 can be found in appendix C.

4. SD-distillation

In this section, we study the fundamental task of distillation of symmetric distinguishability (SD-distillation), both in the one-shot and asymptotic settings.

4.1. One-shot exact SD-distillation

One-shot exact SD-distillation of a given c−q state

$$\rho_{XA} := p|0\rangle\langle0| \otimes \rho_0 + (1 - p)\,|1\rangle\langle1| \otimes \rho_1, \tag{4.1}$$

with $p \in [0, 1]$ and ρ_0, ρ_1 states of a quantum system A, is the task of converting a single copy of it to an M-golden unit via free operations. The maximal value of $\log M$ for which this conversion is possible is equal to the *one-shot exact distillable-SD* for the chosen set of free operations. This is defined formally as follows:We state the following theorems now.

Definition 4.1. For a set of free operations denoted by FO and $q \in [0, 1]$, the one-shot exact distillable-SD of the c−q state ρ_{XA} defined in (4.1) is given by

$$\xi_d^{FO,q}(\rho_{XA}) := \log\left(\sup\left\{M \mid \mathcal{A}(\rho_{XA}) = \gamma_{XQ}^{(M,q)}, \; \mathcal{A} \in \text{free operations}\,(\text{FO})\right\}\right). \tag{4.2}$$

For the choice

$$\text{FO} \equiv \{\text{id} \otimes \mathcal{E} \mid \mathcal{E}\;\text{CPTP on system A}\} \equiv \text{CPTP}_A, \tag{4.3}$$

the only sensible choice in (4.2) is $q = p$, as free operations of the form $\text{id} \otimes \mathcal{E}$ cannot change the prior in the c−q state. In that case, the above quantity is called the *one-shot exact distillable-SD under* CPTP_A *maps* and we simply write

$$\xi_d(\rho_{XA}) \equiv \xi_d^{\text{CPTP}_A, p}(\rho_{XA}). \tag{4.4}$$

Whereas for the choice $\text{FO} \equiv \text{CDS}$ and $q = 1/2$, the above quantity is called the *one-shot exact distillable-SD under CDS maps* and we use the notation

$$\xi_d^{\star}(\rho_{XA}) \equiv \xi_d^{\text{CDS}, 1/2}(\rho_{XA}). \tag{4.5}$$

Explicitly, for a c−q state ρ_{XA} given by (4.1) we then have

$$\xi_d(\rho_{XA}) = \log\left(\sup\left\{M \mid (\text{id} \otimes \mathcal{E})\,(\rho_{XA}) = \gamma_{XQ}^{(M,p)}, \mathcal{E} \in \text{CPTP}\right\}\right) \tag{4.6}$$

and

$$\xi_d^{\star}(\rho_{XA}) = \log\left(\sup\left\{M \mid \mathcal{N}(\rho_{XA}) = \gamma_{XQ}^{(M)}, \mathcal{N} \in \text{CDS}\right\}\right), \tag{4.7}$$

where $\gamma_{XQ}^{(M)} \equiv \gamma_{XQ}^{(M,1/2)}$ as stated previously.

Remark 4.2. Note that in the case of prior $p \in (0, 1)$ and the free operations being CPTP_A, the distillable-SD is, by definition, independent of p. In fact, in that case it can be equivalently written as

$$\xi_d(\rho_{XA}) \equiv \xi_d(\rho_0, \rho_1) = \log\left(\sup\left\{M \mid \mathcal{E}(\rho_0) = \pi_M \text{ and } \mathcal{E}(\rho_1) = \sigma^{(1)}\pi_M\sigma^{(1)}, \mathcal{E} \in \text{CPTP}\right\}\right). \tag{4.8}$$

For $p \in \{0, 1\}$ one easily sees that

$$\xi_d(\rho_{XA}) = \infty. \tag{4.9}$$

Therefore, we restrict to the non-singular case $p \in (0, 1)$ in the following theorem 4.3. We state the following theorems now.

Theorem 4.3. *The one-shot exact distillable-SD under* CPTP_A *maps of a c−q state* ρ_{XA}, *defined through* (4.1), *with* $p \in (0, 1)$ *is given by*

$$\xi_d(\rho_{XA}) = \xi_{\min}(\rho_{XA}), \tag{4.10}$$

where

$$\xi_{\min}(\rho_{XA}) \equiv \xi_{\min}(\rho_0, \rho_1) = -\log Q_{\min}(\rho_0, \rho_1), \tag{4.11}$$

and $Q_{\min}(\rho_0, \rho_1)$ *is given by the following SDP:*

$$Q_{\min}(\rho_0, \rho_1) := 2 \min\left\{\text{Tr}(\Lambda\rho_0) \mid \text{Tr}(\Lambda(\rho_0 + \rho_1)) = 1, \; 0 \leqslant \Lambda \leqslant \mathbb{1}\right\}. \tag{4.12}$$

Remark 4.4. An alternative way of writing $Q_{min}(\rho_0, \rho_1)$ is as follows:

$$Q_{min}(\rho_0, \rho_1) = 2 \min \left\{ Tr(\Lambda \rho_0) \,|\, Tr(\Lambda \rho_0) = Tr((\mathbb{1} - \Lambda)\rho_1), \; 0 \leqslant \Lambda \leqslant \mathbb{1} \right\}$$

$$= 2 \min \left\{ Tr(\Lambda \rho_0) \,|\, Tr(\Lambda \rho_0) \geqslant Tr((\mathbb{1} - \Lambda)\rho_1), \; 0 \leqslant \Lambda \leqslant \mathbb{1} \right\}. \tag{4.13}$$

This clarifies that the optimization is over all POVMs $\{\Lambda, \mathbb{1} - \Lambda\}$ such that the type I error probability $Tr(\Lambda \rho_0)$ is (greater than or) equal to the type II error probability $Tr((\mathbb{1} - \Lambda)\rho_1)$.

To see the second equality in (4.13), note first that the last line is trivially smaller than or equal to the right-hand side of the first line, since we are minimising over a larger set. To arrive at the other inequality, let $\tilde{\Lambda}_{min}$ be a minimiser of the last line of (4.13) and $c := Tr(\tilde{\Lambda}_{min}(\rho_0 + \rho_1)) \geqslant 1$. Let $\Lambda_{min} = \tilde{\Lambda}_{min}/c$. Clearly $0 \leqslant \Lambda_{min} \leqslant \tilde{\Lambda}_{min} \leqslant \mathbb{1}$ and $Tr(\Lambda_{min}(\rho_0 + \rho_1)) = 1$. Moreover,

$$Q_{min}(\rho_0, \rho_1) \leqslant 2 Tr(\Lambda_{min}\rho_0) = \frac{2 Tr\left(\tilde{\Lambda}_{min}\rho_0\right)}{c} \leqslant 2 Tr\left(\tilde{\Lambda}_{min}\rho_0\right),$$

from which we conclude the equality in (4.13).

Theorem 4.5. *The one-shot exact distillable-SD under CDS maps of a c−q state ρ_{XA}, defined through (4.1), is given by*

$$\xi_d^\star(\rho_{XA}) = -\log(2 p_{err}(\rho_{XA})) = SD(\rho_{XA}). \tag{4.14}$$

Remark 4.6. As a consequence of theorem 4.3, it follows that the one-shot exact distillable-SD under CPTP_A maps can be calculated by means of a SDP, due to the form of $Q_{min}(\rho_0, \rho_1)$ in (4.12). As a consequence of theorem 4.5, it follows that the one-shot exact distillable-SD under CDS maps can be calculated by means of a SDP, due to the expression for $p_{err}(\rho_{XA})$ in (3.2).

Proofs of the above theorems are given in sections 4.4 and 4.5, respectively. The quantities ξ_{min} and Q_{min} appearing in theorem 4.3 have several useful and interesting properties, which are given in section 4.3.

4.2. Optimal asymptotic rate of exact SD-distillation
Consider the c−q state

$$\rho_{XA}^{(n)} := p \,|0\rangle\langle 0| \otimes \rho_0^{\otimes n} + (1 - p) \,|1\rangle\langle 1| \otimes \rho_1^{\otimes n},$$

and let $\xi_d(\rho_{XB}^{(n)})$ and $\xi_d^\star(\rho_{XB}^{(n)})$ denote its one-shot exact distillable-SD under CPTP_A maps and CDS maps, respectively. Then the *optimal asymptotic rates of exact SD-distillation* under CPTP_A maps and CDS maps are defined by the following two quantities, respectively:

$$\liminf_{n \to \infty} \frac{\xi_d(\rho_{XA}^{(n)})}{n}; \qquad \liminf_{n \to \infty} \frac{\xi_d^\star(\rho_{XA}^{(n)})}{n}. \tag{4.15}$$

The next theorem asserts that both limits in (4.15) actually exist and are equal to the well-known quantum Chernoff divergence [ACMnT$^+$07, NS09]:

Theorem 4.7 (Optimal asymptotic rate of exact SD-distillation). *For $p \in (0, 1)$, the optimal asymptotic rates of exact SD-distillation under CPTP_A and CDS maps are given by the following expression:*

$$\lim_{n \to \infty} \frac{\xi_d(\rho_{XA}^{(n)})}{n} = \lim_{n \to \infty} \frac{\xi_d^\star(\rho_{XA}^{(n)})}{n} = \lim_{n \to \infty} \frac{\xi_{min}(\rho_0^{\otimes n}, \rho_1^{\otimes n})}{n} = \xi(\rho_0, \rho_1), \tag{4.16}$$

where $\xi(\rho_0, \rho_1) := -\log \min_{0 \leqslant s \leqslant 1} Tr(\rho_0^s \rho_1^{1-s})$ denotes the quantum Chernoff divergence.

Here, the restriction to $p \in (0, 1)$ is sensible as for $p \in \{0, 1\}$ we directly get $\xi_d(\rho_{XA}^{(n)}) = \xi_d^\star(\rho_{XA}^{(n)}) = \infty$ for all $n \in \mathbb{N}$.

A proof of the above theorem is given in section 4.6.

4.3. Properties of Q_{min} and ξ_{min}
In this section, we establish some basic properties of the distinguishability measures Q_{min} and ξ_{min}.

Lemma 4.8. *The distinguishability measures $Q_{min}(\rho_0, \rho_1)$ and $\xi_{min}(\rho_0, \rho_1)$ are symmetric in their arguments.*

Proof. To see this, note that if Λ_{min} is a minimiser for $Q_{min}(\rho_0, \rho_1)$, i.e. satisfying $2 Tr(\Lambda_{min}\rho_0) = Q_{min}(\rho_0, \rho_1), 0 \leqslant \Lambda_{min} \leqslant \mathbb{1}$, and $Tr(\Lambda_{min}(\rho_0 + \rho_1)) = 1$, then $\tilde{\Lambda}_{min} = \mathbb{1} - \Lambda_{min}$ also satisfies $0 \leqslant \tilde{\Lambda}_{min} \leqslant \mathbb{1}$ and $Tr(\tilde{\Lambda}_{min}(\rho_0 + \rho_1)) = 1$. Moreover, $2 Tr(\tilde{\Lambda}_{min}\rho_1) = 2 Tr(\Lambda_{min}\rho_0) = Q_{min}(\rho_0, \rho_1)$. From that we see

$$Q_{min}(\rho_1, \rho_0) \leqslant Q_{min}(\rho_0, \rho_1).$$

The reversed inequality can be obtained by symmetry, which yields

$$Q_{\min}(\rho_0, \rho_1) = Q_{\min}(\rho_1, \rho_0). \tag{4.17}$$

By a straightforward consequence of the definition of $\xi_{\min}(\rho_0, \rho_1)$ in terms of $Q_{\min}(\rho_0, \rho_1)$, it follows that $\xi_{\min}(\rho_0, \rho_1)$ is symmetric in its arguments. □

The quantity $\xi_{\min}$ also satisfies a DPI under CPTP maps, which is the statement of the following lemma:

Lemma 4.9 (DPI for $\xi_{\min}$ under CPTP maps). *Let $\mathcal{E}$ be a quantum channel. Then*

$$\xi_{\min}(\mathcal{E}(\rho_0), \mathcal{E}(\rho_1)) \leqslant \xi_{\min}(\rho_0, \rho_1). \tag{4.18}$$

Proof. Consider

$$
\begin{aligned}
Q_{\min}(\mathcal{E}(\rho_0), \mathcal{E}(\rho_1)) &= 2 \min \left\{ \mathrm{Tr}(\Lambda \mathcal{E}(\rho_0)) \,\middle|\, \mathrm{Tr}(\Lambda \mathcal{E}(\rho_0 + \rho_1)) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1} \right\} \\
&= 2 \min \left\{ \mathrm{Tr}(\mathcal{E}^*(\Lambda) \rho_0) \,\middle|\, \mathrm{Tr}(\mathcal{E}^*(\Lambda)(\rho_0 + \rho_1)) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1} \right\} \\
&\geqslant 2 \min \left\{ \mathrm{Tr}(\Lambda \rho_0) \,\middle|\, \mathrm{Tr}(\Lambda(\rho_0 + \rho_1)) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1} \right\} \\
&= Q_{\min}(\rho_0, \rho_1).
\end{aligned}
\tag{4.19}
$$

In the last inequality, we have used that $0 \leqslant \mathcal{E}^*(\Lambda) \leqslant \mathbb{1}$ for each $0 \leqslant \Lambda \leqslant \mathbb{1}$, implying that we are effectively minimising over a smaller set. Hence, directly from the definition of $\xi_{\min}$ in (4.11), we conclude the DPI (4.18). □

The next lemma gives upper and lower bounds on $Q_{\min}$, which turn out to be the key ingredients for proving the asymptotic result in (4.16).

Lemma 4.10. *For all $q \in [0, 1]$, we have*

$$p_{\mathrm{err}}(q, \rho_0, \rho_1) \leqslant \frac{1}{2} Q_{\min}(\rho_0, \rho_1) \leqslant 2 p_{\mathrm{err}}(1/2, \rho_0, \rho_1). \tag{4.20}$$

Proof. Let $\Lambda_{\min}$ be a minimiser of $Q_{\min}(\rho_0, \rho_1)$. As $\mathrm{Tr}(\Lambda_{\min}(\rho_0 + \rho_1)) = 1$, we have $\mathrm{Tr}(\Lambda_{\min} \rho_0) = \mathrm{Tr}((\mathbb{1} - \Lambda_{\min})\rho_1)$. Hence, for each $q \in [0, 1]$

$$
\begin{aligned}
Q_{\min}(\rho_0, \rho_1)/2 = \mathrm{Tr}(\Lambda_{\min} \rho_0) &= q\, \mathrm{Tr}(\Lambda_{\min} \rho_0) + (1 - q)\mathrm{Tr}((\mathbb{1} - \Lambda_{\min})\rho_1) \\
&\geqslant \min_{0 \leqslant \Lambda \leqslant \mathbb{1}} \left(q\, \mathrm{Tr}(\Lambda \rho_0) + (1 - q)\mathrm{Tr}((\mathbb{1} - \Lambda)\rho_1) \right) = p_{\mathrm{err}}(q, \rho_0, \rho_1).
\end{aligned}
\tag{4.21}
$$

For the remaining inequality in (4.20), we use the specific choice of

$$\Lambda = (\rho_0 + \rho_1)^{-1/2} \rho_1 (\rho_0 + \rho_1)^{-1/2},$$

with $(\rho_0 + \rho_1)^{-1/2}$ being defined via the pseudo inverse. Note that $\{\Lambda, \mathbb{1} - \Lambda\}$ forms the so-called *pretty good measurement* [Bel75a, Bel75b, Hol79, Hau93, HW94]. As

$$(\rho_0 + \rho_1)^{-1}(\rho_0 + \rho_1) = \Pi_{\rho_0 + \rho_1},$$

with $\Pi_{\rho_0 + \rho_1}$ being the projector onto the support of the operator $\rho_0 + \rho_1$, we see that

$$\mathrm{Tr}(\Lambda(\rho_0 + \rho_1)) = \mathrm{Tr}(\rho_1 \Pi_{\rho_0 + \rho_1}) = 1. \tag{4.22}$$

In the above, we have used that the null spaces satisfy

$$N(\rho_0 + \rho_1) = N(\rho_0) \cap N(\rho_1),$$

which follows by the positive semi-definiteness of ρ_0 and ρ_1. Moreover, by (4.22) we already get $\mathrm{Tr}((\mathbb{1} - \Lambda)\rho_1) = \mathrm{Tr}(\Lambda \rho_0)$. Hence, we see

$$Q_{\min}/2 \leqslant \mathrm{Tr}(\Lambda \rho_0) = \frac{1}{2}\left(\mathrm{Tr}(\Lambda \rho_0) + \mathrm{Tr}((\mathbb{1} - \Lambda)\rho_1)\right). \tag{4.23}$$

Using the well known fact that the error probability of the pretty good measurement is upper bounded by twice the minimum error probability [BK02] (also compare [HW12, CMTM$^+$08]), i.e.

$$p_{\mathrm{err}}^{\mathrm{pgm}}(1/2, \rho_0, \rho_1) = \frac{1}{2}\left(\mathrm{Tr}(\Lambda \rho_0) + \mathrm{Tr}(\mathbb{1} - \Lambda)\rho_1\right) \leqslant 2 p_{\mathrm{err}}(1/2, \rho_0, \rho_1), \tag{4.24}$$

this completes the proof.

Note that in order to see (4.24), [BK02, equation (13)] gives the following lower bound on the guessing probability of the pretty good measurement:

$$p_{\text{guess}}(1/2, \rho_0, \rho_1)^2 \leqslant p_{\text{guess}}^{\text{pgm}}(1/2, \rho_0, \rho_1), \tag{4.25}$$

where $p_{\text{guess}}(1/2, \rho_0, \rho_1)$ is the optimal guessing probability in the discrimination task. This bound immediately implies (4.24) since $(1 - x)^2 \geqslant 1 - 2x$ for every real number x. □

Lemma 4.11. *For states ρ_0 and ρ_1, the following bound holds*

$$\xi_{\min}(\rho_0, \rho_1) \geqslant \xi(\rho_0, \rho_1) - 1, \tag{4.26}$$

where $\xi_{\min}(\rho_0, \rho_1)$ is defined in (4.11) and $\xi(\rho_0, \rho_1)$ in (1.5).

Proof. From the Helstrom–Holevo theorem and lemma 4.10, we conclude the upper bound

$$Q_{\min}(\rho_0, \rho_1) \leqslant 2 - \|\rho_0 - \rho_1\|_1.$$

Using theorem 1 in [ACMnT$^+$07] (also compare equation (6) therein) we see

$$Q_{\min}(\rho_0, \rho_1) \leqslant 2 - \|\rho_0 - \rho_1\|_1 \leqslant 2Q(\rho_0, \rho_1),$$

with $Q(\rho_0, \rho_1) = \min_{0 \leqslant s \leqslant 1} \text{Tr}(\rho_0^s \rho_1^{1-s})$. This directly yields the desired lower bound on $\xi_{\min}$. □

Remark 4.12. Using lemma 4.8, $\xi_{\min}$ can be written as

$$\xi_{\min}(\rho_0, \rho_1) := -\log\big(\min\left\{\text{Tr}(\Lambda\rho_1) \,|\, \text{Tr}(\Lambda(\rho_0 + \rho_1)) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1}\right\}\big) - 1. \tag{4.27}$$

The use of the subscript in $\xi_{\min}$ is motivated by the similarity of the above expression (modulo the additive constant) with $D_{\min}$:

$$D_{\min}(\rho_0 \| \rho_1) = -\log \min\left\{\text{Tr}(\Lambda\rho_1) \,|\, \text{Tr}(\Lambda\rho_0) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1}\right\}. \tag{4.28}$$

The notation $\xi_{\min}$ is further motivated by analogy with the RTAD [WW19a], where the quantity analogous to $\xi_{\min}$ is the min-relative entropy $D_{\min}$ [Dat09].

4.4. Proof of theorem 4.3—one-shot exact distillable-SD under CPTP$_A$ maps

Proof. We first prove the achievability, i.e. the lower bound $\xi_d(\rho_{XA}) \geqslant \xi_{\min}(\rho_{XA}) \equiv \xi_{\min}(\rho_0, \rho_1)$.

Let $\Lambda_{\min}$ be a minimiser of the optimisation problem corresponding to $Q_{\min}$ given in (4.12), i.e. $2\,\text{Tr}(\Lambda_{\min}\rho_0) = Q_{\min}(\rho_0, \rho_1)$. Let $\mathcal{E}$ be the following measure-and-prepare channel:

$$\mathcal{E}(\rho) = \text{Tr}((\mathbb{1} - \Lambda_{\min})\rho) \, |0\rangle\langle 0| + \text{Tr}(\Lambda_{\min}\rho) \, |1\rangle\langle 1|. \tag{4.29}$$

Hence, using $\text{Tr}(\Lambda_{\min}\rho_0) = \text{Tr}((\mathbb{1} - \Lambda_{\min})\rho_1)$ we get

$$(\mathbb{1} \otimes \mathcal{E})(\rho_{XA}) = p \, |0\rangle\langle 0| \otimes \pi_M + (1 - p) \, |1\rangle\langle 1| \otimes \sigma^{(1)}\pi_M\sigma^{(1)}, \tag{4.30}$$

with $M = 1/(2\,\text{Tr}(\Lambda_{\min}\rho_0))$. This implies that

$$\xi_{\min}(\rho_0, \rho_1) := -\log Q_{\min}(\rho_0, \rho_1) = -\log(2\,\text{Tr}(\Lambda_{\min}\rho_0)) \leqslant \xi_d(\rho_{XA}). \tag{4.31}$$

To obtain the converse, i.e. the reverse inequality $\xi_d(\rho_0, \rho_1) \leqslant \xi_{\min}(\rho_0, \rho_1)$, we first note that for all $M \geqslant 0$ we have $\pi_M + \sigma^{(1)}\pi_M\sigma^{(1)} = \mathbb{1}$ and hence by picking $\Lambda = |1\rangle\langle 1|$

$$Q_{\min}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)}) = 2 \min\left\{\text{Tr}(\Lambda\pi_M) \,|\, \text{Tr}(\Lambda) = 1, \ 0 \leqslant \Lambda \leqslant \mathbb{1}\right\}$$

$$\leqslant 2\langle 1|\pi_M|1\rangle = \frac{1}{M},$$

and therefore

$$\xi_{\min}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)}) \geqslant \log M.$$

Now considering $\mathcal{E}$ to be an arbitrary CPTP map such that

$$(\mathbb{1} \otimes \mathcal{E})(\rho_{XA}) = p \, |0\rangle\langle 0| \otimes \pi_M + (1 - p) \, |1\rangle\langle 1| \otimes \sigma^{(1)}\pi_M\sigma^{(1)},$$

for some $M \geqslant 1$, we see by the data processing inequality in lemma 4.9 that

$$\xi_{\min}(\rho_0, \rho_1) \geqslant \xi_{\min}(\mathcal{E}(\rho_0), \mathcal{E}(\rho_1)) = \xi_{\min}(\pi_M, \sigma^{(1)} \pi_M \sigma^{(1)}) \geqslant \log M.$$

As $\mathcal{E}$ is an arbitrary CPTP map satisfying the constraint in (4.2), we get

$$\xi_{\min}(\rho_0, \rho_1) \geqslant \xi_d(\rho_{XA}),$$

and hence $\xi_{\min}(\rho_{XA}) \equiv \xi_{\min}(\rho_0, \rho_1) = \xi_d(\rho_{XA})$. □

4.5. Proof of theorem 4.5—one-shot exact distillable-SD under CDS maps

Proof. We start with the achievability part, i.e. the lower bound

$$\xi_d^\star(\rho_{XA}) \geqslant -\log(2p_{\mathrm{err}}(\rho_{XA})).$$

Let $\Lambda_{\min}$ be the minimiser of

$$p_{\mathrm{err}}(\rho_{XA}) = \min_{0 \leqslant \Lambda \leqslant \mathbb{1}} \left(p \operatorname{Tr}(\Lambda \rho_0) + (1-p)\operatorname{Tr}((\mathbb{1} - \Lambda)\rho_1) \right). \tag{4.32}$$

Define the quantum operations, i.e. CP, trace non-increasing maps

$$\mathcal{E}_0(\rho) = \frac{1}{2} \left(\operatorname{Tr}((\mathbb{1} - \Lambda_{\min})\rho) \, |0\rangle\langle 0| + \operatorname{Tr}(\Lambda_{\min}\rho) \, |1\rangle\langle 1| \right),$$

$$\mathcal{E}_1(\rho) = \frac{1}{2} \left(\operatorname{Tr}((\mathbb{1} - \Lambda_{\min})\rho) \, |1\rangle\langle 1| + \operatorname{Tr}(\Lambda_{\min}\rho) \, |0\rangle\langle 0| \right), \tag{4.33}$$

and note that $\mathcal{E}_0$ and $\mathcal{E}_1$ sum to a CPTP map. Hence, we can define the corresponding CDS map as

$$\mathcal{N} = \mathrm{id}_X \otimes \mathcal{E}_0 + \mathcal{F}_X \otimes \mathcal{E}_1, \tag{4.34}$$

where id_X and $\mathcal{F}_X$ denote the identity and flip channel on the classical system X, respectively. Noting that

$$\begin{aligned}
p\,\mathcal{E}_0(\rho_0) + (1-p)\mathcal{E}_1(\rho_1) &= \frac{1}{2} \left(p \operatorname{Tr}((\mathbb{1} - \Lambda_{\min})\rho_0) + (1-p)\operatorname{Tr}(\Lambda_{\min}\rho_1) \right) |0\rangle\langle 0| \\
&\quad + \frac{1}{2} \left(p \operatorname{Tr}(\Lambda_{\min}\rho_0) + (1-p)\operatorname{Tr}((\mathbb{1} - \Lambda_{\min})\rho_1) \right) |1\rangle\langle 1| \\
&= \frac{1}{2} \left((1 - p_{\mathrm{err}}(\rho_{XA})) \, |0\rangle\langle 0| + p_{\mathrm{err}}(\rho_{XA}) \, |1\rangle\langle 1| \right) \\
&= \frac{1}{2} \pi_{1/2p_{\mathrm{err}}(\rho_{XA})},
\end{aligned} \tag{4.35}$$

and by symmetry

$$\begin{aligned}
p\,\mathcal{E}_1(\rho_0) + (1-p)\mathcal{E}_0(\rho_1) &= \frac{1}{2} \left((1 - p_{\mathrm{err}}) \, |1\rangle\langle 1| + p_{\mathrm{err}} \, |0\rangle\langle 0| \right) \\
&= \frac{1}{2} \sigma^{(1)} \pi_{1/2p_{\mathrm{err}}(\rho_{XA})} \sigma^{(1)}.
\end{aligned} \tag{4.36}$$

Hence,

$$\begin{aligned}
\mathcal{N}(\rho_{XA}) &= |0\rangle\langle 0| \otimes \left(p\mathcal{E}_0(\rho_0) + (1-p)\mathcal{E}_1(\rho_1) \right) + |1\rangle\langle 1| \otimes \left(p\mathcal{E}_1(\rho_0) + (1-p)\mathcal{E}_0(\rho_1) \right) \\
&= \frac{1}{2} \left(|0\rangle\langle 0| \otimes \pi_{1/2p_{\mathrm{err}}(\rho_{XA})} + |1\rangle\langle 1| \otimes \sigma^{(1)} \pi_{1/2p_{\mathrm{err}}(\rho_{XA})} \sigma^{(1)} \right) \\
&= \gamma_{XQ}^{(1/2p_{\mathrm{err}}(\rho_{XA}))},
\end{aligned} \tag{4.37}$$

which implies

$$\xi_d^\star(\rho_{XA}) \geqslant -\log(2p_{\mathrm{err}}(\rho_{XA})). \tag{4.38}$$

To obtain the upper bound in (4.14), we use monotonicity of the minimum error probability under CDS maps. More precisely, let $M \geqslant 1$ satisfy the constraint in (4.7); i.e. there exists a CDS map $\mathcal{N}$ such that

$$\mathcal{N}(\rho_{XA}) = \gamma_{XQ}^{(M)}. \tag{4.39}$$

Using the monotonicity of the minimum error probability p_{err} under CDS maps, we obtain

$$-\log\bigl(2p_{\mathrm{err}}(\rho_{XA})\bigr) \geqslant -\log\bigl(2p_{\mathrm{err}}(\mathcal{N}(\rho_{XA}))\bigr) = -\log\Bigl(2p_{\mathrm{err}}(\gamma_{XQ}^{(M)})\Bigr)$$

$$= \log M. \tag{4.40}$$

As M is arbitrary under the constraints in (4.7), we have shown that

$$-\log(2p_{\mathrm{err}}(\rho_{XA})) \geqslant \xi_d^{\star}(\rho_{XA}), \tag{4.41}$$

which finishes the proof. $\square$

4.6. Proof of theorem 4.7—optimal asymptotic rate of exact SD-distillation

Proof. We first prove the result in the case of free operations being CPTP_A maps. As a consequence of theorem 4.3, we have the equality

$$\frac{\xi_d(\rho_{XA}^{(n)})}{n} = \frac{\xi_{\min}(\rho_0^{\otimes n}, \rho_1^{\otimes n})}{n}, \tag{4.42}$$

and so it suffices to prove the asymptotic result for $\xi_{\min}$. Using (4.26) and the fact that the quantum Chernoff divergence ξ is additive, we get

$$\liminf_{n\to\infty} \frac{\xi_{\min}(\rho_0^{\otimes n}, \rho_1^{\otimes n})}{n} \geqslant \xi(\rho_0, \rho_1).$$

To get the other inequality, we use the lower bound in (4.20) to see that for every $q \in (0, 1)$

$$\frac{\xi_{\min}(\rho_0^{\otimes n}, \rho_1^{\otimes n})}{n} \leqslant \frac{-\log(2p_{\mathrm{err}}(q, \rho_0^{\otimes n}, \rho_1^{\otimes n}))}{n}. \tag{4.43}$$

Hence, by using the main result in [ACMnT+07], we conclude that

$$\limsup_{n\to\infty} \frac{\xi_{\min}(\rho_0^{\otimes n}, \rho_1^{\otimes n})}{n} \leqslant \lim_{n\to\infty} \frac{-\log(p_{\mathrm{err}}(q, \rho_0^{\otimes n}, \rho_1^{\otimes n})) - 1}{n} = \xi(\rho_0, \rho_1). \tag{4.44}$$

In the case of free operations being CDS maps, the equality

$$\lim_{n\to\infty} \frac{\xi_d^{\star}(\rho_{XA}^{(n)})}{n} = \lim_{n\to\infty} \frac{-\log(p_{\mathrm{err}}(\rho_{XA}^{(n)})) - 1}{n} = \xi(\rho_0, \rho_1) \tag{4.45}$$

directly follows from theorem 4.5, together with the main result of [ACMnT+07], which finishes the proof. $\square$

4.7. Approximate SD-distillation

We now define the *one-shot approximate distillable-SD* for a general c−q state

$$\rho_{XA} = p\,|0\rangle\langle 0| \otimes \rho_0 + (1-p)\,|1\rangle\langle 1| \otimes \rho_1. \tag{4.46}$$

Definition 4.13. For $\varepsilon \geqslant 0$ and golden unit $\gamma_{XQ}^{(M,q)}$, the *one-shot approximate distillable-SD* of the c−q state ρ_{XA} is given by

$$\xi_d^{\mathrm{FO},q,\varepsilon}(\rho_{XA}) := \log\left(\sup\left\{M \,\middle|\, d_{\mathrm{FO}}'(\rho_{XA} \mapsto \gamma_{XQ}^{(M,q)}) \leqslant \varepsilon\right\}\right)$$

$$= \log\left(\sup\left\{M \,\middle|\, D'\Bigl(\mathcal{A}(\rho_{XA}), \gamma_{XQ}^{(M,q)}\Bigr) \leqslant \varepsilon,\ \mathcal{A} \in \mathrm{FO}\right\}\right), \tag{4.47}$$

where the minimum conversion error d_{FO}' is defined in definition 3.10. For the choice

$$\mathrm{FO} \equiv \{\mathrm{id} \otimes \mathcal{E} \,|\, \mathcal{E}\ \mathrm{CPTP\ on\ system\ A}\} \equiv \mathrm{CPTP}_A, \tag{4.48}$$

the only sensible choice in (4.47) is $q = p$, as free operations of the form $\mathrm{id} \otimes \mathcal{E}$ cannot change the prior in the c−q state. In that case, we simply write

$$\xi_d^{\varepsilon}(\rho_{XA}) \equiv \xi_d^{\mathrm{CPTP}_A, p, \varepsilon}(\rho_{XA}). \tag{4.49}$$

Whereas for the choice $\mathrm{FO} \equiv \mathrm{CDS}$ and $q = 1/2$, we use the notation

$$\xi_d^{\star,\varepsilon}(\rho_{XA}) \equiv \xi_d^{\mathrm{CDS}, 1/2, \varepsilon}(\rho_{XA}). \tag{4.50}$$

4.7.1. One-shot approximate distillable-SD as a semi-definite program

In this section, we prove that the one-shot approximate distillable-SD under CPTP_A can be evaluated by means of a SDP and comment on SDP formulations of the one-shot approximate distillable-SD under CDS. Under CPTP_A this quantity is defined as follows:

$$\xi_d^\varepsilon(\rho_{XA}) := \log \sup_{\mathcal{A} \in \text{CPTP}_A} \left\{ M \mid D'(\mathcal{A}(\rho_{XA}), \gamma_{XQ}^{(M,p)}) \leqslant \varepsilon \right\}, \tag{4.51}$$

and for CDS maps as

$$\xi_d^{\star,\varepsilon}(\rho_{XA}) := \log \sup_{\mathcal{A} \in \text{CDS}} \left\{ M \mid D'(\mathcal{A}(\rho_{XA}), \gamma_{XQ}^{(M,1/2)}) \leqslant \varepsilon \right\}. \tag{4.52}$$

We begin with the following:

Proposition 4.14. *For all $\varepsilon \geqslant 0$ and $p \in (0,1)$ and for every elementary source described by the c−q state ρ_{XA}, the one-shot approximate distillable-SD under CPTP_A maps can be evaluated by the following SDP:*

$$\xi_d^\varepsilon(\rho_{XA}) = -\log \left[2 \inf_{\substack{0 \leqslant \Lambda \leqslant \mathbb{1}, \\ r \in [0,1/2]}} \left\{ r \mid p \, |1 - r - \text{Tr}[\Lambda \rho_0]| + (1-p)|r - \text{Tr}[\Lambda \rho_1]| \leqslant \varepsilon \, \min\{r, p, 1-p\} \right\} \right]. \tag{4.53}$$

Proof. By definition the quantity $\xi_d^\varepsilon(\rho_{XA})$ is equal to the negative logarithm of

$$\inf_{\mathcal{A} \in \text{CPTP}_A, r \in [0,1]} \left\{ r \mid D'\left(\mathcal{A}(\rho_{XA}), \gamma_{XQ}^{(1/r,p)} \right) \leqslant \varepsilon \right\}. \tag{4.54}$$

Let $\mathcal{A} \in \text{CPTP}_A$. Then by applying a completely dephasing channel to the Q system, the state $\gamma_{XQ}^{(1/r,p)}$ does not change, whereas the local channel becomes a measurement channel of the following form:

$$\mathcal{M}(\omega) := \text{Tr}[\Lambda \omega]|0\rangle\langle 0|_Q + \text{Tr}[(I - \Lambda)\,\omega]|1\rangle\langle 1|_Q. \tag{4.55}$$

So we find that the optimal value is given by

$$\inf_{\mathcal{M}} \left\{ r \mid D'\left(\mathcal{M}(\rho_{XA}), \gamma_{XQ}^{(1/r,p)} \right) \leqslant \varepsilon \right\}, \tag{4.56}$$

as a consequence of the DPI for D' given in lemma 3.13 and with the optimization over every measurement channel $\mathcal{M}$. Now consider that

$$\mathcal{M}(\rho_{XA}) = p|0\rangle\langle 0|_X \otimes \text{Tr}[\Lambda \rho_0]|0\rangle\langle 0|_Q + p|0\rangle\langle 0|_X \otimes \text{Tr}[(I - \Lambda)\rho_0]|1\rangle\langle 1|_Q$$
$$+ (1-p)|1\rangle\langle 1|_X \otimes \text{Tr}[\Lambda \rho_1]|0\rangle\langle 0|_Q + (1-p)|1\rangle\langle 1|_X \otimes \text{Tr}[(I - \Lambda)\rho_1]|1\rangle\langle 1|_Q \tag{4.57}$$

and

$$D'\left(\mathcal{M}(\rho_{XA}), \gamma_{XQ}^{(1/r,p)} \right) = \frac{\frac{1}{2}\left\| \mathcal{M}(\rho_{XA}) - \gamma_{XQ}^{(1/r,p)} \right\|_1}{p_{\text{err}}(\gamma_{XQ}^{(1/r,p)})}. \tag{4.58}$$

By applying (3.7), we know that

$$p_{\text{err}}(\gamma_{XQ}^{(1/r,p)}) = \frac{1}{2}\left(1 - \left| p - \frac{r}{2} \right| - \left| 1 - p - \frac{r}{2} \right| \right) \tag{4.59}$$

$$= \min\left\{ \frac{r}{2}, p, 1 - p \right\}, \tag{4.60}$$

where the last equality follows from a simplification that holds for $p \in [0,1]$ and $r \in [0,1]$. Now consider that

$$\gamma_{XQ}^{(1/r,p)} = p|0\rangle\langle 0|_X \otimes \left(1 - \frac{r}{2} \right) |0\rangle\langle 0|_Q + p|0\rangle\langle 0|_X \otimes \frac{r}{2}|1\rangle\langle 1|_Q$$
$$+ (1-p)|1\rangle\langle 1|_X \otimes \frac{r}{2}|0\rangle\langle 0|_Q + (1-p)|1\rangle\langle 1|_X \otimes \left(1 - \frac{r}{2} \right) |1\rangle\langle 1|_Q. \tag{4.61}$$

So then we find that

$$\left\|\mathcal{M}(\rho_{XA}) - \gamma_{XQ}^{(1/r,p)}\right\|_1 = p\left|1 - \frac{r}{2} - \mathrm{Tr}[\Lambda\rho_0]\right| + p\left|\frac{r}{2} - \mathrm{Tr}[(I - \Lambda)\rho_0]\right|$$

$$+ (1 - p)\left|\frac{r}{2} - \mathrm{Tr}[\Lambda\rho_1]\right| + (1 - p)\left|\left(1 - \frac{r}{2}\right) - \mathrm{Tr}[(I - \Lambda)\rho_1]\right| \tag{4.62}$$

$$= 2p\left|1 - \frac{r}{2} - \mathrm{Tr}[\Lambda\rho_0]\right| + 2(1 - p)\left|\frac{r}{2} - \mathrm{Tr}[\Lambda\rho_1]\right|. \tag{4.63}$$

So the optimization problem is equivalent to the following:

$$\inf_{\Lambda \geqslant 0, r \in [0,1]} \left\{ r : p\left|1 - \frac{r}{2} - \mathrm{Tr}[\Lambda\rho_0]\right| + (1 - p)\left|\frac{r}{2} - \mathrm{Tr}[\Lambda\rho_1]\right| \atop \leqslant \varepsilon \min\left\{\frac{r}{2}, p, 1 - p\right\}, \ \Lambda \leqslant I \right\}. \tag{4.64}$$

Let us make the substitution $r \to 2r$, and the above becomes the following:

$$2 \inf_{\Lambda \geqslant 0, r \in [0,1/2]} \left\{ r : p\left|1 - r - \mathrm{Tr}[\Lambda\rho_0]\right| + (1 - p)\left|r - \mathrm{Tr}[\Lambda\rho_1]\right| \atop \leqslant \varepsilon \min\{r, p, 1 - p\}, \ \Lambda \leqslant I \right\}. \tag{4.65}$$

This concludes the proof.   □

Remark 4.15. We can see that if $\varepsilon = 0$ on the right-hand side of (4.53), this expression reduces to the $\xi_{\min}$ quantity defined in (4.11).

Remark 4.16. An alternate SDP formulation for the one-shot approximate distillable-SD under CPTP_A, stemming from an SDP formulation for the minimum conversion error, is also possible. Similarly, the one-shot approximate distillable-SD under CDS maps can also be formulated as an SDP. We do not include details of these SDPs here, but note that they are available in the arXiv posting of our paper [SDG$^+$21].

4.7.2. Optimal asymptotic rate of approximate SD-distillation
We now consider the asymptotic case of approximate SD-distillation. Theorem 4.7 already established that in the exact case the optimal rates for free operations being CDS or CPTP_A are given by the quantum Chernoff divergence. The following theorem shows that also when an error is allowed in the transformation, the corresponding asymptotic rates are still given by the quantum Chernoff divergence.

Theorem 4.17 (Optimal asymptotic rate of approximate SD-distillation). *For all $\varepsilon \geqslant 0$ and $p \in (0, 1)$, the optimal asymptotic rates of approximate SD-distillation under CPTP_A and CDS maps are given by*

$$\lim_{n \to \infty} \frac{\xi_d^\varepsilon(\rho_{XA}^{(n)})}{n} = \lim_{n \to \infty} \frac{\xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} = \xi(\rho_0, \rho_1), \tag{4.66}$$

where $\xi(\rho_0, \rho_1) = -\log\left(\min_{0 \leqslant s \leqslant 1} \mathrm{Tr}\left(\rho_0^s \rho_1^{1-s}\right)\right)$ denotes the quantum Chernoff divergence.

Here, the restriction to $p \in (0, 1)$ is sensible, as for $p \in \{0, 1\}$, we directly get $\xi_d^\varepsilon(\rho_{XA}^{(n)}) = \xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)}) = \infty$ for all $n \in \mathbb{N}$.

Remark 4.18. Note that theorem 4.17 establishes the strong converse property for the task of asymptotic distillation in the RTSD. Another way of interpreting this statement is as follows: for a sequence of SD distillation protocols with rate above the asymptotic distillable-SD, the error necessarily converges to ∞ in the limit as n becomes large.

Proof of theorem 4.17. As we have, for every c–q state ρ_{XA},

$$\xi_d^\varepsilon(\rho_{XA}) \geqslant \xi_d(\rho_{XA}), \tag{4.67}$$

$$\xi_d^{\star,\varepsilon}(\rho_{XA}) \geqslant \xi_d^\star(\rho_{XA}), \tag{4.68}$$

we immediately get the lower bounds

$$\liminf_{n \to \infty} \frac{\xi_d^\varepsilon(\rho_{XA}^{(n)})}{n} \geqslant \xi(\rho_0, \rho_1), \tag{4.69}$$

$$\liminf_{n \to \infty} \frac{\xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} \geqslant \xi(\rho_0, \rho_1), \tag{4.70}$$

from theorem 4.7.

To establish the upper bounds, we use lemma 3.14. We only prove the upper bound for $\xi_d^{\star,\varepsilon}$, as the one for ξ_d^ε exactly follows the same lines. Let M be such that it satisfies the constraint in the following optimization:

$$\xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)}) = \log\left(\sup\left\{M \,\middle|\, d'_{\text{CDS}}(\rho_{XA}^{(n)} \mapsto \gamma_{XQ}^{(M)}) \leqslant \varepsilon\right\}\right). \tag{4.71}$$

Hence, there exists a CDS map $\mathcal{N}$ such that

$$D'\left(\mathcal{N}(\rho_{XA}^{(n)}), \gamma_{XQ}^{(M)}\right) \leqslant \varepsilon.$$

Then, by the monotonicity of the minimum error probability under CDS maps and the bound in lemma 3.14, we see

$$p_{\text{err}}\left(\rho_{XA}^{(n)}\right) \leqslant p_{\text{err}}\left(\mathcal{N}(\rho_{XA}^{(n)})\right) \leqslant (\varepsilon + 1)\, p_{\text{err}}\left(\gamma_{XQ}^{(M)}\right) = \frac{\varepsilon + 1}{2M}. \tag{4.72}$$

As M is arbitrary under the constraint in (4.71), we get

$$\xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)}) \leqslant \left(-\log(p_{\text{err}}(\rho_{XA}^{(n)})) - 1\right) + \log(\varepsilon + 1), \tag{4.73}$$

and hence

$$\limsup_{n\to\infty} \frac{\xi_d^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} \leqslant \lim_{n\to\infty} \frac{\left(-\log(p_{\text{err}}(\rho_{XA}^{(n)})) - 1\right) + \log(\varepsilon + 1)}{n} = \xi(\rho_0, \rho_1), \tag{4.74}$$

which finishes the proof of theorem 4.17. $\qquad\square$

5. SD-dilution

We now turn to the case of dilution of SD. We begin with the exact one-shot case in section 5.1, establish some properties of relevant divergences in section 5.2, provide proofs in sections 5.3 and 5.4, consider the one-shot approximate case in section 5.5, and evaluate asymptotic quantities in sections 5.6 and 5.7.

5.1. One-shot exact SD-dilution
One-shot exact SD-dilution of a given c–q state

$$\rho_{XA} := p|0\rangle\langle0| \otimes \rho_0 + (1 - p)\,|1\rangle\langle1| \otimes \rho_1, \tag{5.1}$$

with $p \in [0, 1]$, is the task of converting an M-golden unit to the target state ρ_{XA} via free operations. The minimal value of $\log M$ for which this conversion is possible is equal to the *one-shot exact SD-cost* for the chosen set of free operations. This is formally defined as follows:

Definition 5.1. For a set of free operations denoted by FO and $q \in [0, 1]$, the one-shot exact SD-cost of the c–q state ρ_{XA} defined in (5.1) is given by

$$\xi_c^{\text{FO},q}(\rho_{XA}) := \log\left(\inf\left\{M \,\middle|\, \mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right) = \rho_{XA}, \;\mathcal{A} \in \text{free operations (FO)}\right\}\right). \tag{5.2}$$

For the choice

$$\text{FO} \equiv \{\text{id} \otimes \mathcal{E} \mid \mathcal{E} \text{ CPTP on system } A\} \equiv \text{CPTP}_A, \tag{5.3}$$

the only sensible choice in (5.2) is $q = p$, as free operations of the form $\text{id} \otimes \mathcal{E}$ cannot change the prior in the c–q state. In that case, the above quantity is called the *one-shot exact SD-cost under* CPTP_A *maps* and we simply write

$$\xi_c(\rho_{XA}) \equiv \xi_c^{\text{CPTP}_A, p}(\rho_{XA}). \tag{5.4}$$

whereas for the choice $\text{FO} \equiv \text{CDS}$ and $q = 1/2$, the above quantity is called the *one-shot exact SD-cost under CDS maps* and we use the notation

$$\xi_c^\star(\rho_{XA}) \equiv \xi_c^{\text{CDS},1/2}(\rho_{XA}). \tag{5.5}$$

Explicitly, for a c–q state ρ_{XA} given by (5.1), we then have

$$\xi_c(\rho_{XA}) = \log\left(\inf\left\{M \,\middle|\, (\text{id} \otimes \mathcal{E})\left(\gamma_{XQ}^{(M,p)}\right) = \rho_{XA}, \;\mathcal{E} \in \text{CPTP}\right\}\right) \tag{5.6}$$

and

$$\xi_c^\star(\rho_{XA}) = \log\left(\inf\left\{M \,\middle|\, \mathcal{N}\left(\gamma_{XQ}^{(M)}\right) = \rho_{XA}, \;\mathcal{N} \in \text{CDS}\right\}\right), \tag{5.7}$$

with $\gamma_{XQ}^{(M)} \equiv \gamma_{XQ}^{(M,1/2)}$ as defined previously.

Remark 5.2. Note that in the case of the free operations being CPTP$_A$ and prior $p \in (0,1)$ the SD-cost, just as the distillable-SD (see remark 4.2), is independent of p by definition. In fact, in that case, it can be equivalently written as

$$\xi_c(\rho_{XA}) \equiv \xi_c(\rho_0, \rho_1) = \log\left(\inf\left\{M \,\middle|\, \mathcal{E}(\pi_M) = \rho_0 \text{ and } \mathcal{E}(\sigma^{(1)} \pi_M \sigma^{(1)}) = \rho_1, \, \mathcal{E} \in \text{CPTP}\right\}\right). \tag{5.8}$$

For $p \in \{0,1\}$, one easily sees that

$$\xi_c(\rho_{XA}) = 0. \tag{5.9}$$

Therefore, we restrict to the non-singular case $p \in (0,1)$ in the following theorem 5.3.

Theorem 5.3. *The one-shot exact SD-cost under* CPTP$_A$ *maps of a c–q state ρ_{XA} with $p \in (0,1)$, as defined through (5.1), is given by*

$$\xi_c(\rho_{XA}) \equiv \xi_c(\rho_0, \rho_1) = \xi_{\max}(\rho_0, \rho_1), \tag{5.10}$$

where

$$\xi_{\max}(\rho_0, \rho_1) := \log Q_{\max}(\rho_0, \rho_1), \tag{5.11}$$

and

$$Q_{\max}(\rho_0, \rho_1) := \inf\left\{M \,\middle|\, \rho_0 \leqslant (2M-1)\rho_1, \; \rho_1 \leqslant (2M-1)\rho_0\right\}$$

$$= \frac{1}{2}\left(2^{d_{\mathrm{T}}(\rho_0, \rho_1)} + 1\right) \geqslant 1. \tag{5.12}$$

Here, for two positive semi-definite operators ω_1, ω_2,

$$d_{\mathrm{T}}(\omega_1, \omega_2) := \max\left\{D_{\max}(\omega_1 \| \omega_2), D_{\max}(\omega_2 \| \omega_1)\right\} \tag{5.13}$$

denotes the Thompson metric [Tho63].

Remark 5.4. Note that $\xi_{\max}$ can be written as

$$\xi_{\max}(\rho_0, \rho_1) = \max\left\{\log\left(2^{D_{\max}(\rho_0, \rho_1)} + 1\right), \log\left(2^{D_{\max}(\rho_1, \rho_0)} + 1\right)\right\} - 1. \tag{5.14}$$

The use of the subscript in $\xi_{\max}$ is motivated by the fact that it is a divergence that is essentially a symmetrized version of $D_{\max}$.

The notation $\xi_{\max}$ is further motivated by analogy with the RTAD [WW19a], in which the quantity analogous to $\xi_{\max}$, arising as the cost of exact dilution of asymmetric distinguishability, is the max-relative entropy $D_{\max}$ [Dat09].

We also note that theorem 5.3 can be inferred from [BST19, lemma 3.1], but we include a self-contained proof below for completeness.

Theorem 5.5. *The one-shot exact SD-cost under CDS maps of a c–q state ρ_{XA}, defined through (5.1), is given by*

$$\xi_c^\star(\rho_{XA}) = \xi_{\max}^\star(\rho_{XA}), \tag{5.15}$$

where

$$\xi_{\max}^\star(\rho_{XA}) = \log Q_{\max}^\star(\rho_{XA}), \tag{5.16}$$

and

$$Q_{\max}^\star(\rho_{XA}) := \inf\left\{M \,\middle|\, p\rho_0 \leqslant (2M-1)(1-p)\rho_1, \; (1-p)\rho_1 \leqslant (2M-1)p\rho_0\right\}$$

$$= \frac{1}{2}\left(2^{d_{\mathrm{T}}(p\rho_0, (1-p)\rho_1)} + 1\right) \geqslant \frac{1}{2}\max\left\{1/p, 1/(1-p)\right\}. \tag{5.17}$$

Remark 5.6. Note that theorems 5.3 and 5.5 give an operational interpretation to the Thompson metric in the context of exact one-shot dilution in the RTSD. To the best of our knowledge, this is the first time an operational meaning in quantum information has been given to the Thompson metric.

Remark 5.7. As a consequence of theorem 5.3, it follows that the one-shot exact SD-cost under CPTP$_A$ maps can be calculated by means of a SDP, due to the expression in (5.12). As a consequence of theorem 5.5, it follows that the one-shot exact SD-cost under CDS maps can be calculated by means of a SDP, due to the expression in (5.17).

Properties of $\xi_{\max}$, $Q_{\max}$, $\xi_{\max}^\star$, and $Q_{\max}^\star$ are discussed in section 5.2. Proofs of the above theorems are given in sections 5.3 and 5.4.

5.2. Properties of $\xi_{\max}$, $Q_{\max}$, $\xi_{\max}^\star$, and $Q_{\max}^\star$

The quantity $\xi_{\max}$ satisfies the DPI under CPTP maps:

Lemma 5.8 (**DPI for** $\xi_{\max}$). *Let $\mathcal{E}$ be a CPTP map. Then*

$$\xi_{\max}(\mathcal{E}(\rho_0), \mathcal{E}(\rho_1)) \leqslant \xi_{\max}(\rho_0, \rho_1). \tag{5.18}$$

Proof. Follows immediately from the DPI for $D_{\max}$ under CPTP maps [Dat09]. □

We now prove that $\xi_{\max}^\star$ is decreasing under CDS maps. For that, we find that $\xi_{\max}^\star$ can also be expressed in the following ways, as a consequence of the definition (5.17) of $Q_{\max}^\star$:

$$
\begin{aligned}
Q_{\max}^\star(\rho_{XA}) &= \frac{1}{2} \left(\max \left\{ 2^{D_{\max}(p\rho_0 \| (1-p)\rho_1)}, 2^{D_{\max}((1-p)\rho_1 \| p\rho_0)} \right\} + 1 \right) \\
&= \frac{1}{2} \left(\max \left\{ 2^{D_{\max}(\rho_0 \| \rho_1) + \log(p/(1-p))}, 2^{D_{\max}(\rho_1 \| \rho_0) + \log((1-p)/p)} \right\} + 1 \right) \\
&= \frac{1}{2} \left(2^{D_{\max}(\rho_{XA} \| (\mathcal{F}_X \otimes \mathrm{id}_A)(\rho_{XA}))} + 1 \right),
\end{aligned}
\tag{5.19}
$$

where $\mathcal{F}_X(\cdot) = \sigma^{(1)} \cdot \sigma^{(1)}$ denotes the flip channel on the system X.

Lemma 5.9. *Let $\mathcal{N} \in$ CDS. Then*

$$\xi_{\max}^\star(\mathcal{N}(\rho_{XA})) \leqslant \xi_{\max}^\star(\rho_{XA}). \tag{5.20}$$

Proof. Noting that every CDS channel $\mathcal{N}$ commutes with the channel $\mathcal{F}_X \otimes \mathrm{id}_A$, (5.20) follows directly from the DPI for $D_{\max}$. □

5.3. Proof of theorem 5.3—one-shot exact SD-cost under CPTP$_A$ maps

Proof. We prove that $\xi_c(\rho_{XA}) = \xi_{\max}(\rho_0, \rho_1)$ and start with the achievability part, i.e. the upper bound $\xi_c(\rho_{XA}) \leqslant \xi_{\max}(\rho_0, \rho_1)$. Without loss of generality, suppose that $\xi_{\max}(\rho_0, \rho_1)$ is finite because, otherwise, the upper bound is trivially satisfied.

Let us first consider the case $Q_{\max}(\rho_0, \rho_1) = 1$ (and hence $\xi_{\max} = 0$) for which necessarily $\rho_0 = \rho_1 \equiv \rho$. Let us choose the measure-and-prepare channel

$$\mathcal{E}(\cdot) = \mathrm{Tr}(\cdot)\rho.$$

Note that

$$\mathcal{E}(\pi_M) = \mathcal{E}(\sigma^{(1)} \pi_M \sigma^{(1)}) = \rho \quad \forall \, M \geqslant 1.$$

Hence,

$$(\mathrm{id} \otimes \mathcal{E})(\gamma_{XQ}^{(1,p)}) = \rho_{XA},$$

which in turn implies that $\xi_c(\rho_{XA}) \leqslant 0 = \xi_{\max}(\rho_0, \rho_1)$. As $\xi_c(\rho_{XA}) \geqslant 0$ by definition, this also gives $\xi_c(\rho_{XA}) = \xi_{\max}(\rho_0, \rho_1)$.

Next consider the case in which ρ_{XA} is a state for which $M = Q_{\max}(\rho_0, \rho_1) > 1$ and hence $\xi_{\max}(\rho_0, \rho_1) = \log M > 0$. We first prove the achievability bound $\xi_c(\rho_{XA}) \leqslant \xi_{\max}(\rho_0, \rho_1)$. To do this, let us define

$$\tilde{\rho}_0 = \frac{1}{2M-2} \left((2M-1)\rho_0 - \rho_1 \right), \tag{5.21}$$

$$\tilde{\rho}_1 = \frac{1}{2M-2} \left((2M-1)\rho_1 - \rho_0 \right). \tag{5.22}$$

By assumption on M, we have that $\tilde{\rho}_0$ and $\tilde{\rho}_1$ are quantum states. Consider now the measure-and-prepare channel $\mathcal{E}$ given by

$$\mathcal{E}(\rho) = \langle 0 | \rho | 0 \rangle \, \tilde{\rho}_0 + \langle 1 | \rho | 1 \rangle \, \tilde{\rho}_1. \tag{5.23}$$

Then

$$
\begin{aligned}
\mathcal{E}(\pi_M) &= \left(1 - \frac{1}{2M} \right) \tilde{\rho}_0 + \frac{1}{2M} \tilde{\rho}_1 \\
&= \frac{2M-1}{4M^2 - 4M} \left((2M-1)\rho_0 - \rho_1 \right) + \frac{1}{4M^2 - 4M} \left((2M-1)\rho_1 - \rho_0 \right) \\
&= \frac{1}{4M^2 - 4M} \left((2M-1)^2 - 1 \right) \rho_0 = \rho_0.
\end{aligned}
\tag{5.24}
$$

By symmetry, we also get

$$\mathcal{E}(\sigma^{(1)}\pi_M\sigma^{(1)}) = \rho_1. \tag{5.25}$$

This implies that $(\mathrm{id} \otimes \mathcal{E})(\gamma_{XQ}^{(M,p)}) = \rho_{XA}$, which in turn implies that

$$\xi_c(\rho_0, \rho_1) \leqslant \log M = \xi_{\max}(\rho_0, \rho_1). \tag{5.26}$$

To show the reverse inequality, we first note that for all $M \geqslant 1$,

$$\xi_{\max}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)}) = \log M. \tag{5.27}$$

To see this, note that

$$\pi_M \leqslant (2\tilde{M} - 1)\sigma^{(1)}\pi_M\sigma^{(1)} \tag{5.28}$$

if and only if $M \leqslant \tilde{M}$. The reversed constraint in (5.12) is then also satisfied, which establishes that

$$Q_{\max}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)}) = M,$$

and hence (5.27). Let $M \geqslant 1$ satisfy the constraint in the definition of $\xi_c(\rho_{XA})$ in (5.6). By definition there exists a CPTP map $\mathcal{E}$ such that

$$(\mathrm{id} \otimes \mathcal{E})\left(\gamma_{XQ}^{(M,p)}\right) = \rho_{XA}. \tag{5.29}$$

By the DPI for $\xi_{\max}$, we get

$$\xi_{\max}(\rho_0, \rho_1) \leqslant \xi_{\max}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)}) = \log M. \tag{5.30}$$

And hence, as M can be chosen arbitrarily under the constraint in (5.2), we see that

$$\xi_{\max}(\rho_0, \rho_1) \leqslant \xi_c(\rho_{XA}), \tag{5.31}$$

and therefore in total

$$\xi_c(\rho_{XA}) = \xi_{\max}(\rho_0, \rho_1), \tag{5.32}$$

concluding the proof. $\qquad\square$

5.4. Proof of theorem 5.5—one-shot exact SD-cost under CDS maps

Proof. We first prove the achievability part, i.e. $\xi_c^\star(\rho_{XA}) \leqslant \xi_{\max}^\star(\rho_{XA})$. Without loss of generality, suppose that $\xi_{\max}^\star(\rho_{XA})$ is finite because, otherwise, the upper bound is trivially satisfied. Moreover, note that

$$
\begin{aligned}
Q_{\max}^\star(\rho_{XA}) &= \frac{1}{2}\left(\max\left\{2^{D_{\max}(\rho_0\|\rho_1)+\log(p/(1-p))}, 2^{D_{\max}(\rho_1\|\rho_0)+\log((1-p)/p)}\right\} + 1\right)\\
&\geqslant \frac{1}{2}\left(2^{\max\{\log(p/(1-p)),\log((1-p)/p)\}} + 1\right) = \frac{1}{2}\left(\max\left\{\frac{p}{1-p}, \frac{1-p}{p}\right\} + 1\right)\\
&= \frac{1}{2}\max\{1/p, 1/(1-p)\}.
\end{aligned} \tag{5.33}
$$

We first treat the case $Q_{\max}^\star(\rho_{XA}) = \frac{1}{2}\max\{1/p, 1/(1-p)\}$ for which necessarily $\rho_0 = \rho_1 \equiv \rho$. Assume without loss of generality $p \leqslant 1 - p$ (otherwise just flip the classical system X of ρ_{XA}). Let now

$$\mathcal{E}_0(\cdot) = \langle 1|\cdot|1\rangle\,\rho, \qquad \mathcal{E}_1(\cdot) = \langle 0|\cdot|0\rangle\,\rho, \tag{5.34}$$

which are quantum operations, i.e. CP and trace non-increasing maps that sum to a trace-preserving map. Define the corresponding CDS map

$$\mathcal{N} = \mathrm{id}_X \otimes \mathcal{E}_0 + \mathcal{F}_X \otimes \mathcal{E}_1. \tag{5.35}$$

Writing now $M = Q_{\max}^\star(\rho_{XA}) = \frac{1}{2p}$, we see

$$
\begin{aligned}
\mathcal{N}(\gamma_{XQ}^{(M)}) &= |0\rangle\langle 0| \otimes \frac{1}{2}\left(\mathcal{E}_0(\pi_M) + \mathcal{E}_1(\sigma^{(1)}\pi_M\sigma^{(1)})\right) + |1\rangle\langle 1| \otimes \frac{1}{2}\left(\mathcal{E}_1(\pi_M) + \mathcal{E}_0(\sigma^{(1)}\pi_M\sigma^{(1)})\right)\\
&= p\,|0\rangle\langle 0| \otimes \rho + (1-p)\,|1\rangle\langle 1| \otimes \rho = \rho_{XA}.
\end{aligned} \tag{5.36}
$$

Hence, this shows $\xi_c^\star(\rho_{XA}) \leqslant \log M = \xi_{\max}^\star(\rho_{XA})$.

Let now $Q^\star_{\max}(\rho_{XA}) > \frac{1}{2}\max\{1/p, 1/(1-p)\} \geqslant 1$ and define for $M = Q^\star_{\max}(\rho_{XA})$ the operators

$$\tilde{\rho}_0 = \frac{1}{2Mp-1}\left(p(2M-1)\rho_0 - (1-p)\rho_1\right), \tag{5.37}$$

$$\tilde{\rho}_1 = \frac{1}{2M(1-p)-1}\left((1-p)(2M-1)\rho_1 - p\rho_0\right). \tag{5.38}$$

By assumption on M we have that $\tilde{\rho}_0$ and $\tilde{\rho}_1$ are states. Moreover, let

$$q = \frac{2Mp-1}{2M-2}, \tag{5.39}$$

and hence

$$1-q = \frac{2M(1-p)-1}{2M-2}. \tag{5.40}$$

Consider now the quantum operations

$$\mathcal{E}_0(\rho) = q\,\langle 0|\,\rho\,|0\rangle\,\tilde{\rho}_0 + (1-q)\,\langle 1|\,\rho\,|1\rangle\,\tilde{\rho}_1, \tag{5.41}$$

$$\mathcal{E}_1(\rho) = (1-q)\,\langle 0|\,\rho\,|0\rangle\,\tilde{\rho}_1 + q\,\langle 1|\,\rho\,|1\rangle\,\tilde{\rho}_0. \tag{5.42}$$

Note that $\mathcal{E}_0 + \mathcal{E}_1$ is trace preserving. The corresponding CDS map is given by

$$\mathcal{N} = \mathrm{id}_X \otimes \mathcal{E}_0 + \mathcal{F}_X \otimes \mathcal{E}_1. \tag{5.43}$$

Then

$$\mathcal{N}(\gamma_{XQ}^{(M)}) = |0\rangle\langle 0| \otimes \frac{1}{2}\left(\mathcal{E}_0(\pi_M) + \mathcal{E}_1(\sigma^{(1)}\pi_M\sigma^{(1)})\right) + |1\rangle\langle 1| \otimes \frac{1}{2}\left(\mathcal{E}_1(\pi_M) + \mathcal{E}_0(\sigma^{(1)}\pi_M\sigma^{(1)})\right). \tag{5.44}$$

Moreover, we see

$$\begin{aligned}
\frac{1}{2}\left(\mathcal{E}_0(\pi_M) + \mathcal{E}_1(\sigma^{(1)}\pi_M\sigma^{(1)})\right) &= q\left(1 - \frac{1}{2M}\right)\tilde{\rho}_0 + \frac{1-q}{2M}\tilde{\rho}_1 \\
&= \frac{2M-1}{4M^2-4M}\left(p(2M-1)\rho_0 - (1-p)\rho_1\right) \\
&\quad + \frac{1}{4M^2-4M}\left((1-p)(2M-1)\rho_1 - p\rho_0\right) \\
&= \frac{p}{4M^2-4M}\left((2M-1)^2 - 1\right)\rho_0 = p\rho_0.
\end{aligned} \tag{5.45}$$

By symmetry we also get

$$\frac{1}{2}\left(\mathcal{E}_1(\pi_M) + \mathcal{E}_0(\sigma^{(1)}\pi_M\sigma^{(1)})\right) = (1-p)\rho_1. \tag{5.46}$$

This proves the achievability

$$\xi_c^\star(\rho_{XA}) \leqslant \xi_{\max}^\star(\rho_{XA}). \tag{5.47}$$

For the other inequality, we first note that

$$\xi_{\max}^\star(\gamma_{XQ}^{(M)}) = \log M, \tag{5.48}$$

which follows from the observation $\xi_{\max}^\star(\gamma_{XQ}^{(M)}) = \xi_{\max}(\pi_M, \sigma^{(1)}\pi_M\sigma^{(1)})$ and the arguments in the case of free operations being CPTP$_A$ maps (in particular consider the discussion of (5.27)). Let $M \geqslant 1$ satisfy the constraint in the definition of $\xi_c^\star(\rho_{XA})$ in (5.7). By definition there exists a CDS map $\mathcal{N}$ such that

$$\mathcal{N}(\gamma_{XQ}^{(M)}) = \rho_{XA}. \tag{5.49}$$

We use monotonicity of $\xi_{\max}^\star$ under CDS maps (compare lemma 5.9) to get that

$$\xi_{\max}^\star(\rho_{XA}) \leqslant \xi_{\max}^\star(\gamma_{XQ}^{(M)}) = \log M. \tag{5.50}$$

As M is arbitrary, we see that $\xi_{\max}^\star(\rho_{XA}) \leqslant \xi_c^\star(\rho_{XA})$ and hence $\xi_{\max}^\star(\rho_{XA}) = \xi_c^\star(\rho_{XA})$. □

5.5. One-shot approximate SD-dilution

We can now define the *one-shot approximate SD-cost* for a general c−q state

$$\rho_{XA} = p\,|0\rangle\langle 0| \otimes \rho_0 + (1-p)\,|1\rangle\langle 1| \otimes \rho_1. \tag{5.51}$$

Definition 5.10. For $\varepsilon \geqslant 0$ and golden unit $\gamma_{XQ}^{(M,q)}$, the *one-shot approximate SD-cost* of the c−q state ρ_{XA} is given by

$$
\begin{aligned}
\xi_c^{FO,q,\varepsilon}(\rho_{XA}) &:= \log\left(\inf\left\{ M \,\middle|\, d_{FO}'\left(\gamma_{XQ}^{(M,q)} \to \rho_{XA}\right) \leqslant \varepsilon \right\}\right) \\
&= \log\left(\inf\left\{ M \,\middle|\, D'\left(\mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right), \rho_{XA}\right) \leqslant \varepsilon,\ \mathcal{A} \in FO \right\}\right),
\end{aligned} \tag{5.52}
$$

where the minimum conversion error d_{FO}' is defined in definition 3.10. For the choice

$$FO \equiv \{\mathrm{id} \otimes \mathcal{E}\,|\,\mathcal{E}\ \text{CPTP on system}\ A\} \equiv \mathrm{CPTP}_A, \tag{5.53}$$

the only sensible choice in (4.47) is $q = p$, as free operations of the form $\mathrm{id} \otimes \mathcal{E}$ cannot change the prior in the c−q state. In that case we simply write

$$\xi_c^{\varepsilon}(\rho_{XA}) \equiv \xi_c^{\mathrm{CPTP}_A,p,\varepsilon}(\rho_{XA}). \tag{5.54}$$

Whereas for the choice $FO \equiv CDS$ and $q = 1/2$, we use the notation

$$\xi_c^{\star,\varepsilon}(\rho_{XA}) \equiv \xi_c^{\mathrm{CDS},1/2,\varepsilon}(\rho_{XA}). \tag{5.55}$$

In the case of the dilution task, the *one-shot approximate SD-cost* can be directly obtained from the corresponding exact quantity.

Lemma 5.11. *For* FO *being c−q state preserving and* $\varepsilon \geqslant 0$, *we have*

$$\xi_c^{FO,q,\varepsilon}(\rho_{XA}) = \inf_{\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})} \xi_c^{FO,q}(\tilde{\rho}_{XA}), \tag{5.56}$$

where we have defined the ball of c−q states[12] around ρ_{XA} *with radius* ε *with respect to the scaled trace distance* $D'(\cdot, \cdot)$

$$B_{\varepsilon}'(\rho_{XA}) = \left\{\tilde{\rho}_{XA}\ c\text{−}q\ state\,\middle|\,D'(\tilde{\rho}_{XA}, \rho_{XA}) \leqslant \varepsilon\right\}. \tag{5.57}$$

Hence, in particular we get for free operations being CPTP_A *or CDS maps*

$$\xi_c^{\varepsilon}(\rho_{XA}) = \inf_{\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})} \xi_c(\tilde{\rho}_{XA}), \tag{5.58}$$

$$\xi_c^{\star,\varepsilon}(\rho_{XA}) = \inf_{\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})} \xi_c^{\star}(\tilde{\rho}_{XA}). \tag{5.59}$$

Proof. The proof simply follows by

$$
\begin{aligned}
\xi_c^{FO,q,\varepsilon}(\rho_{XA}) &= \log\left(\inf\left\{ M \,\middle|\, d_{FO}'(\gamma_{XQ}^{(M,q)} \to \rho_{XA}) \leqslant \varepsilon \right\}\right) \\
&= \log\left(\inf\left\{ M \,\middle|\, D'\left(\mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right), \rho_{XA}\right) \leqslant \varepsilon,\ \mathcal{A} \in FO \right\}\right) \\
&= \log\left(\inf\left\{ M \,\middle|\, \mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right) = \tilde{\rho}_{XA},\ \tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA}),\ \mathcal{A} \in FO \right\}\right) \\
&= \inf_{\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})} \log\left(\inf\left\{ M \,\middle|\, \mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right) = \tilde{\rho}_{XA},\ \mathcal{A} \in FO \right\}\right) \\
&= \inf_{\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})} \xi_c^{FO,q}(\tilde{\rho}_{XA}).
\end{aligned} \tag{5.60}
$$

Here, for the third equality we have used that $\tilde{\rho}_{XA} = \mathcal{A}\left(\gamma_{XQ}^{(M,q)}\right)$ is a c−q state because FO is c−q state preserving. Hence $D'(\tilde{\rho}_{XA}, \rho_{XA}) \leqslant \varepsilon$ already implies $\tilde{\rho}_{XA} \in B_{\varepsilon}'(\rho_{XA})$. $\qquad\square$

---

[12] Note that (5.56) is also true if we replace $B_{\varepsilon}'(\rho_{XA})$ with the full D'-ball with radius ε of all linear operators and not just c−q states. The reason for that is that $\xi_c^{FO,q}(\tilde{\rho}_{XA})$ is infinite for $\tilde{\rho}_{XA}$ not a c−q state because the set of free operations FO assumed to be c−q state preserving.

5.6. Optimal asymptotic rates of exact and approximate SD-dilution

Consider the c–q state

$$\rho_{XA}^{(n)} := p \, |0\rangle\langle 0| \otimes \rho_0^{\otimes n} + (1-p) \, |1\rangle\langle 1| \otimes \rho_1^{\otimes n}.$$

Similarly, as in the case of distillation, we are interested in the asymptotic quantities

$$\limsup_{n\to\infty} \frac{\xi_c(\rho_{XA}^{(n)})}{n}; \qquad \limsup_{n\to\infty} \frac{\xi_c^\star(\rho_{XA}^{(n)})}{n}. \tag{5.61}$$

Using theorems 5.3 and 5.5 and the additivity of $D_{\max}$ and hence also of the Thompson metric d_{T}, we can directly read off that both limits in (5.61) exist and are given by the Thompson metric:

Theorem 5.12 (Exact asymptotic SD-cost).　*For all $p \in (0,1)$, the optimal asymptotic rates of exact SD-dilution under* CPTP_A *and CDS maps is given by*

$$\lim_{n\to\infty} \frac{\xi_c(\rho_{XA}^{(n)})}{n} = \lim_{n\to\infty} \frac{\xi_c^\star(\rho_{XA}^{(n)})}{n} = d_{\mathrm{T}}(\rho_0, \rho_1). \tag{5.62}$$

Hence, unlike the case of distillation, the optimal asymptotic rates in the case of exact dilution do not match the quantum Chernoff bound, as they are too large. We also note here that (5.62) gives an operational interpretation of the Thompson metric in quantum information theory.

However, allowing errors with respect to the scaled trace distance D' in the conversion, the corresponding approximate quantities converge to the Chernoff divergence.

Theorem 5.13 (Approximate asymptotic SD-cost).　*For all $\varepsilon > 0$ and $p \in (0,1)$, the optimal asymptotic rates of approximate SD-dilution under* CPTP_A *and CDS maps is given by*

$$\lim_{n\to\infty} \frac{\xi_c^\varepsilon(\rho_{XA}^{(n)})}{n} = \lim_{n\to\infty} \frac{\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} = \xi(\rho_0, \rho_1), \tag{5.63}$$

where $\xi(\rho_0, \rho_1) = -\log\!\big(\min_{0\leqslant s\leqslant 1} \mathrm{Tr}\big(\rho_0^s \rho_1^{1-s}\big)\big)$ denotes the quantum Chernoff divergence. In particular this gives

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{\xi_c^\varepsilon(\rho_{XA}^{(n)})}{n} = \lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} = \xi(\rho_0, \rho_1). \tag{5.64}$$

Here, the restriction to $p \in (0,1)$ is sensible, as for $p \in \{0,1\}$, we directly get $\xi_c^\varepsilon(\rho_{XA}^{(n)}) = 0$ and $\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)}) = \infty$ for all $n \in \mathbb{N}$ and $\varepsilon \geqslant 0$.

Remark 5.14.　The fact that the limits in (5.63) hold without any restriction on the value of $\varepsilon > 0$ implies that the strong converse holds for the asymptotic SD-cost. Another way of interpreting this statement is as follows: for a sequence of SD dilution protocols with rate below the asymptotic SD-cost, the error necessarily converges to infinity as $n \to \infty$.

Remark 5.15.　Given that the asymptotic distillable-SD and SD-cost are equal to the quantum Chernoff divergence, it follows that the RTSD is asymptotically reversible. This means that, in the asymptotic limit of large n, one can convert the source state $\rho_{XA}^{(n)}$ to a target state $\sigma_{XA'}^{(m)}$ at a rate $\frac{m}{n}$ given by the ratio of the Chernoff divergences. Then one can go back to $\rho_{XA}^{(n)}$ at the inverse rate with no loss (in the asymptotic limit). The procedure to do so, for the first aforementioned state conversion, is to distill golden-unit states from $\rho_{XA}^{(n)}$ at a rate equal to the Chernoff divergence. Then we dilute these golden-unit states to $\sigma_{XA'}^{(m)}$, and the overall conversion rate is equal to the ratio of Chernoff divergences. Then we go back from $\sigma_{XA'}^{(m)}$ to $\rho_{XA}^{(n)}$ in a similar manner, and there is no loss in the asymptotic limit. We discuss these points in much more detail in section 7.

We note here that other resource theories, such as pure-state bipartite entanglement [BBPS96], coherence [WY16], thermodynamics [BaHO$^+$13], and asymmetric distinguishability [WW19a] are all asymptotically reversible in a similar sense.

5.7. Proof of theorem 5.13

5.7.1. Lower bound for the asymptotic SD-cost in equation (5.63)

Using an argument similar to that given in the proof of theorem 4.17, we establish the following asymptotic lower bound on the approximate SD-cost:

Lemma 5.16.　*For all $\varepsilon \geqslant 0$,*

$$\liminf_{n\to\infty} \frac{\xi_c^\varepsilon(\rho_{XA}^{(n)})}{n} \geqslant \xi(\rho_0, \rho_1), \tag{5.65}$$

$$\liminf_{n\to\infty} \frac{\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} \geqslant \xi(\rho_0, \rho_1). \tag{5.66}$$

Proof. We only consider the lower bound for $\xi_c^{\star,\varepsilon}$, as the one for ξ_c^{ε} exactly follows the same line of reasoning. Let M be such that it satisfies the constraint in the following optimization:

$$\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)}) = \log\left(\inf\left\{ M \,\middle|\, d'_{\mathrm{CDS}}(\gamma_{XQ}^{(M)} \to \rho_{XA}^{(n)}) \leqslant \varepsilon \right\}\right). \tag{5.67}$$

Hence, there exists a CDS map $\mathcal{N}$ satisfying

$$D'\left(\mathcal{N}(\gamma_{XQ}^{(M)}), \rho_{XA}^{(n)}\right) \leqslant \varepsilon.$$

Then, by the monotonicity of the minimum error probability under CDS maps and lemma 3.14, we see

$$\frac{1}{2M} = p_{\mathrm{err}}\left(\gamma_{XQ}^{(M)}\right) \leqslant p_{\mathrm{err}}\left(\mathcal{N}(\gamma_{XQ}^{(M)})\right) \leqslant (\varepsilon + 1)p_{\mathrm{err}}(\rho_{XA}^{(n)}). \tag{5.68}$$

As M is arbitrary under the constraint in (5.67), we get

$$\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)}) \geqslant -\log(p_{\mathrm{err}}(\rho_{XA}^{(n)})) - \log(2(\varepsilon + 1)), \tag{5.69}$$

and hence

$$\liminf_{n\to\infty} \frac{\xi_c^{\star,\varepsilon}(\rho_{XA}^{(n)})}{n} \geqslant \lim_{n\to\infty} \frac{-\log(p_{\mathrm{err}}(\rho_{XA}^{(n)})) - \log(2(\varepsilon + 1))}{n} = \xi(\rho_0, \rho_1), \tag{5.70}$$

concluding the proof. $\square$

5.7.2. Upper bound for the asymptotic SD-cost in equation (5.63) and smoothed Thompson metric
Denote the set of sub-normalised states on a Hilbert space $\mathcal{H}$ by

$$\mathcal{S}_\leqslant(\mathcal{H}) := \{\omega \in \mathcal{B}(\mathcal{H}) \mid \omega \geqslant 0, \ \mathrm{Tr}(\omega) \leqslant 1\}, \tag{5.71}$$

with $\mathcal{B}(\mathcal{H})$ the set of bounded operators on $\mathcal{H}$. Moreover, for a sub-normalised state ω on $\mathcal{H}$, we define the trace ball of sub-normalised states with radius $\varepsilon \geqslant 0$ around ω as

$$\mathcal{B}_\varepsilon(\omega) := \left\{ \tilde{\omega} \in \mathcal{S}_\leqslant(\mathcal{H}) \,\middle|\, \frac{1}{2}\|\omega - \tilde{\omega}\|_1 \leqslant \varepsilon \right\}. \tag{5.72}$$

In order to prove the desired upper bound on the asymptotic dilution cost (5.63), we consider the following smoothed quantity:

$$d_{\mathrm{T}}^{\varepsilon}(\omega_0, \omega_1) := \inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0), \\ \tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1)}} d_{\mathrm{T}}(\tilde{\omega}_0, \tilde{\omega}_1) \tag{5.73}$$

$$= \inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0), \\ \tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1)}} \max\left\{ D_{\max}(\tilde{\omega}_0 \| \tilde{\omega}_1), D_{\max}(\tilde{\omega}_1 \| \tilde{\omega}_0) \right\}. \tag{5.74}$$

We get the following upper bound on the smoothed Thompson metric:

Lemma 5.17. *For all $\omega_0, \omega_1 \in \mathcal{S}_\leqslant(\mathcal{H})$ and $\varepsilon \in (0, 1]$, we have the bound*

$$d_{\mathrm{T}}^{\varepsilon}(\omega_0, \omega_1) \leqslant \log\left(\frac{4}{\varepsilon}\right). \tag{5.75}$$

For the special case of $\mathrm{Tr}(\omega_0) + \mathrm{Tr}(\omega_1) = 1$, *we get the slightly stronger bound*

$$\inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0), \, \tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1), \\ \mathrm{Tr}(\tilde{\omega}_0) + \mathrm{Tr}(\tilde{\omega}_1) = 1}} d_{\mathrm{T}}(\tilde{\omega}_0, \tilde{\omega}_1) \leqslant \log\left(\frac{4}{\varepsilon}\right), \tag{5.76}$$

and for the case $\mathrm{Tr}(\omega_0) = 1$ *and* $\mathrm{Tr}(\omega_1) = 1$, *we get*

$$\inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0), \, \tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1), \\ \mathrm{Tr}(\tilde{\omega}_0) = 1, \, \mathrm{Tr}(\tilde{\omega}_1) = 1}} d_{\mathrm{T}}(\tilde{\omega}_0, \tilde{\omega}_1) \leqslant \log\left(\frac{2}{\varepsilon}\right). \tag{5.77}$$

Proof. For $\varepsilon \in (0, 1]$, fix

$$\lambda_0 = \max\left\{\frac{2\,\mathrm{Tr}((\omega_1 - \omega_0)_+)}{\varepsilon}, 1\right\},$$

$$\lambda_1 = \max\left\{\frac{2\,\mathrm{Tr}((\omega_0 - \omega_1)_+)}{\varepsilon}, 1\right\},$$

(5.78)

and write

$$0 \leqslant \varepsilon_0 := \mathrm{Tr}((\omega_1 - \omega_0)_+)/\lambda_0 \leqslant \varepsilon/2,$$

(5.79)

$$0 \leqslant \varepsilon_1 := \mathrm{Tr}((\omega_0 - \omega_1)_+)/\lambda_1 \leqslant \varepsilon/2.$$

(5.80)

Define now the positive semi-definite operators

$$\tilde{\omega}_0 := \frac{\omega_0 + \frac{1}{\lambda_0}(\omega_1 - \omega_0)_+}{1 + \varepsilon_0 + \varepsilon_1},$$

(5.81)

$$\tilde{\omega}_1 := \frac{\omega_1 + \frac{1}{\lambda_1}(\omega_0 - \omega_1)_+}{1 + \varepsilon_0 + \varepsilon_1},$$

(5.82)

where A_+ denotes the positive part of a self-adjoint operator A. Firstly, note that by definition

$$\mathrm{Tr}(\tilde{\omega}_0) = \frac{\mathrm{Tr}(\omega_0) + \varepsilon_0}{1 + \varepsilon_0 + \varepsilon_1} \leqslant 1,$$

(5.83)

and analogously $\mathrm{Tr}(\tilde{\omega}_1) \leqslant 1$, which gives $\tilde{\omega}_0, \tilde{\omega}_1 \in \mathcal{S}_{\leqslant}(\mathcal{H})$. Moreover, note that $\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0)$ and $\tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1)$, which follows by

$$
\begin{aligned}
\frac{1}{2}\|\omega_0 - \tilde{\omega}_0\|_1 &= \frac{1}{2(1 + \varepsilon_0 + \varepsilon_1)}\left\|(\varepsilon_0 + \varepsilon_1)\omega_0 - \frac{1}{\lambda_0}(\omega_1 - \omega_0)_+\right\|_1 \\
&\leqslant \frac{1}{2(1 + \varepsilon_0 + \varepsilon_1)}\left((\varepsilon_0 + \varepsilon_1)\|\omega_0\|_1 + \frac{\|(\omega_1 - \omega_0)_+\|_1}{\lambda_0}\right) \\
&\leqslant \frac{2\varepsilon_0 + \varepsilon_1}{2(1 + \varepsilon_0 + \varepsilon_1)} \leqslant \varepsilon,
\end{aligned}
$$

(5.84)

and analogously for $\tilde{\omega}_1$.

We now note that

$$\lambda_0\tilde{\omega}_0 = \frac{1}{1 + \varepsilon_0 + \varepsilon_1}(\lambda_0\omega_0 + (\omega_1 - \omega_0)_+) \geqslant \frac{1}{1 + \varepsilon_0 + \varepsilon_1}((\lambda_0 - 1)\omega_0 + \omega_1) \geqslant \frac{\omega_1}{1 + \varepsilon_0 + \varepsilon_1},$$

(5.85)

where we have used $\lambda_0 \geqslant 1$. Furthermore, we see

$$
\begin{aligned}
\frac{\lambda_0}{\lambda_1}\tilde{\omega}_0 &- \frac{1}{\lambda_1(1 + \varepsilon_0 + \varepsilon_1)}(\omega_0 - \omega_1) \\
&= \frac{1}{1 + \varepsilon_0 + \varepsilon_1}\left(\frac{\lambda_0}{\lambda_1}\omega_0 + \frac{1}{\lambda_1}((\omega_1 - \omega_0)_+ - (\omega_0 - \omega_1)_+)\right) \\
&= \frac{1}{1 + \varepsilon_0 + \varepsilon_1}\left(\frac{\lambda_0}{\lambda_1}\omega_0 + \frac{1}{\lambda_1}((\omega_1 - \omega_0)_+ - (\omega_1 - \omega_0)_-)\right) \\
&= \frac{1}{1 + \varepsilon_0 + \varepsilon_1}\left(\frac{\lambda_0}{\lambda_1}\omega_0 + \frac{1}{\lambda_1}(\omega_1 - \omega_0)\right) \\
&= \frac{1}{1 + \varepsilon_0 + \varepsilon_1}\left(\frac{\lambda_0 - 1}{\lambda_1}\omega_0 + \frac{1}{\lambda_1}\omega_1\right) \geqslant 0.
\end{aligned}
$$

(5.86)

Here, in the third line, we have denoted the negative part of an self-adjoint operator A by A_- and used $(-A)_+ = A_-$ and in the fourth line we used $A = A_+ - A_-$. Hence, combining (5.85) and (5.86) together with the definition of $\tilde{\omega}_1$ (5.82), this in total gives

$$\left(\lambda_0 + \frac{\lambda_0}{\lambda_1}\right)\tilde{\omega}_0 \geqslant \tilde{\omega}_1.$$

(5.87)

Analogously we get

$$\left(\lambda_1 + \frac{\lambda_1}{\lambda_0}\right)\tilde{\omega}_1 \geqslant \tilde{\omega}_0. \tag{5.88}$$

Therefore, using $\mathrm{Tr}((\omega_0 - \omega_1)_+) \leqslant \mathrm{Tr}(\omega_0) \leqslant 1$ and $\mathrm{Tr}((\omega_1 - \omega_0)_+) \leqslant \mathrm{Tr}(\omega_1) \leqslant 1$ and hence $1 \leqslant \lambda_0, \lambda_1 \leqslant \max\{2/\varepsilon, 1\} = 2/\varepsilon$, (5.87) and (5.88) give

$$D_{\max}\left(\tilde{\omega}_1 \| \tilde{\omega}_0\right) \leqslant \log\left(\lambda_0 + \frac{\lambda_0}{\lambda_1}\right) \leqslant \log\left(\frac{4}{\varepsilon}\right) \tag{5.89}$$

and

$$D_{\max}(\tilde{\omega}_0 \| \tilde{\omega}_1) \leqslant \log\left(\lambda_1 + \frac{\lambda_1}{\lambda_0}\right) \leqslant \log\left(\frac{4}{\varepsilon}\right). \tag{5.90}$$

Therefore, we obtain for the Thompson metric of $\tilde{\omega}_0$ and $\tilde{\omega}_1$

$$d_{\mathrm{T}}(\tilde{\omega}_0, \tilde{\omega}_1) = \max\left\{D_{\max}(\tilde{\omega}_0 \| \tilde{\omega}_1), D_{\max}(\tilde{\omega}_1 \| \tilde{\omega}_0)\right\} \leqslant \log\left(\frac{4}{\varepsilon}\right). \tag{5.91}$$

Hence, by definition of the smoothed Thompson metric this shows (5.75). Moreover, noting that in case $\mathrm{Tr}(\omega_0) + \mathrm{Tr}(\omega_1) = 1$, we have by construction

$$\mathrm{Tr}(\tilde{\omega}_0) + \mathrm{Tr}(\tilde{\omega}_1) = \frac{\mathrm{Tr}(\omega_0) + \mathrm{Tr}(\omega_1) + \varepsilon_0 + \varepsilon_1}{1 + \varepsilon_0 + \varepsilon_1} = 1,$$

which immediately also gives (5.76).

In order to also conclude (5.77), we slightly change the above construction in the following way: first note that in the case $\mathrm{Tr}(\omega_0) = \mathrm{Tr}(\omega_1) = 1$ we have $\mathrm{Tr}((\omega_0 - \omega_1)_+) = \mathrm{Tr}((\omega_1 - \omega_0)_+)$. Let now

$$\lambda = \max\left\{\frac{\mathrm{Tr}((\omega_0 - \omega_1)_+)}{\varepsilon}, 1\right\}, \tag{5.92}$$

and write $0 \leqslant \varepsilon' = \mathrm{Tr}((\omega_0 - \omega_1))_+)/\lambda = \mathrm{Tr}((\omega_1 - \omega_0)_+)/\lambda \leqslant \varepsilon$. Define now the positive semi-definite operators

$$\tilde{\omega}_0 = \frac{\omega_0 + \frac{1}{\lambda}(\omega_1 - \omega_0)_+}{1 + \varepsilon'}, \tag{5.93}$$

and

$$\tilde{\omega}_1 = \frac{\omega_1 + \frac{1}{\lambda}(\omega_0 - \omega_1)_+}{1 + \varepsilon'}. \tag{5.94}$$

Note that by $\mathrm{Tr}(\omega_0) = \mathrm{Tr}(\omega_1) = 1$ we get $\mathrm{Tr}(\tilde{\omega}_0) = \mathrm{Tr}(\tilde{\omega}_1) = 1$. Moreover, similarly to the above, we see $\tilde{\omega}_0 \in \mathcal{B}_\varepsilon(\omega_0)$ and $\tilde{\omega}_1 \in \mathcal{B}_\varepsilon(\omega_1)$. Lastly, by using the same arguments as in (5.85) and (5.86) we see that

$$(\lambda + 1)\tilde{\omega}_0 \geqslant \tilde{\omega}_1 \tag{5.95}$$

and

$$(\lambda + 1)\tilde{\omega}_1 \geqslant \tilde{\omega}_0, \tag{5.96}$$

which gives by using $\lambda \leqslant 1/\varepsilon$

$$D_{\max}\left(\tilde{\omega}_1 \| \tilde{\omega}_0\right) \leqslant \log(\lambda + 1) \leqslant \log\left(\frac{1 + \varepsilon}{\varepsilon}\right) \leqslant \log\left(\frac{2}{\varepsilon}\right) \tag{5.97}$$

and

$$D_{\max}(\tilde{\omega}_0 \| \tilde{\omega}_1) \leqslant \log(c + 1) \leqslant \log\left(\frac{1 + \varepsilon}{\varepsilon}\right) \leqslant \log\left(\frac{2}{\varepsilon}\right). \tag{5.98}$$

Then we conclude (5.77). $\qquad\square$

The desired upper bound (5.63) in theorem 5.13 can now be deduced from the following lemma.

Lemma 5.18. *For all $\varepsilon > 0$, we have*

$$\limsup_{n \to \infty} \frac{\xi_c^\varepsilon(\rho_{XA}^{(n)})}{n} \leqslant \xi(\rho_0, \rho_1), \tag{5.99}$$

$$\limsup_{n \to \infty} \frac{\xi_c^{\star, \varepsilon}(\rho_{XA}^{(n)})}{n} \leqslant \xi(\rho_0, \rho_1). \tag{5.100}$$

Proof. We prove the statement only for $\xi_c^{\star,\varepsilon}$, as the proof for ξ_c^{ε} follows exactly along the same lines. Let A_n denote the quantum system consisting of n copies of the quantum system A. Moreover, for a generic c–q state $\tilde{\rho}_{XA_n}$, we use the notation

$$\tilde{\rho}_{XA_n} = \tilde{p} \, |0\rangle\langle 0| \otimes \tilde{\rho}_0 + (1-\tilde{p}) \, |1\rangle\langle 1| \otimes \tilde{\rho}_1. \tag{5.101}$$

By using theorem 5.5 and lemma 5.11, we see that

$$\xi_c^{\varepsilon,\star}(\rho_{XA}^{(n)}) = \inf_{\substack{\tilde{\rho}_{XA_n} \text{ c–q state,} \\ D'(\tilde{\rho}_{XA_n},\rho_{XA}^{(n)}) \leqslant \varepsilon}} \xi_c^{\star}(\tilde{\rho}_{XA_n}) = \inf_{\substack{\tilde{\rho}_{XA_n} \text{ c–q state,} \\ D'(\tilde{\rho}_{XA_n},\rho_{XA}^{(n)}) \leqslant \varepsilon}} \xi_{\max}^{\star}(\tilde{\rho}_{XA_n})$$

$$= \inf_{\substack{\tilde{\rho}_{XA_n} \text{ c–q state,} \\ D'(\tilde{\rho}_{XA_n},\rho_{XA}^{(n)}) \leqslant \varepsilon}} \log\left(2^{d_{\mathrm{T}}(\tilde{p}\tilde{\rho}_0,(1-\tilde{p})\tilde{\rho}_1)} + 1 \right) - 1. \tag{5.102}$$

Define now

$$\tilde{\varepsilon}(\varepsilon,n) := \frac{\varepsilon p_{\mathrm{err}}(\rho_{XA}^{(n)})}{2}. \tag{5.103}$$

Note that if a given $\tilde{\rho}_{XA_n} \equiv (\tilde{p}, \tilde{\rho}_0, \tilde{\rho}_1)$ is such that $\tilde{p}\tilde{\rho}_0 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}(p\rho_0^{\otimes n})$ and $(1-\tilde{p})\tilde{\rho}_1 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}((1-p)\rho_1^{\otimes n})$, this implies that $\frac{1}{2}\|\tilde{\rho}_{XA_n} - \rho_{XA}^{(n)}\|_1 \leqslant 2\tilde{\varepsilon}(\varepsilon,n)$ and hence

$$D'(\tilde{\rho}_{XA_n}, \rho_{XA}^{(n)}) \leqslant \varepsilon. \tag{5.104}$$

Using that, by writing $\tilde{\omega}_0 = \tilde{p}\tilde{\rho}_0$ and $\tilde{\omega}_1 = (1-\tilde{p})\tilde{\rho}_1$, we see that

$$\xi_c^{\varepsilon,\star}(\rho_{XA}^{(n)}) + 1 = \inf_{\substack{\tilde{\rho}_{XA_n} \text{ c–q state,} \\ D'(\tilde{\rho}_{XA_n},\rho_{XA}^{(n)}) \leqslant \varepsilon}} \log\left(2^{d_{\mathrm{T}}(\tilde{p}\tilde{\rho}_0,(1-\tilde{p})\tilde{\rho}_1)} + 1 \right)$$

$$\leqslant \inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}(p\rho_0^{\otimes n}),\, \tilde{\omega}_1 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}((1-p)\rho_1^{\otimes n}) \\ \mathrm{Tr}(\tilde{\omega}_0) + \mathrm{Tr}(\tilde{\omega}_1) = 1}} \log\left(2^{d_{\mathrm{T}}(\tilde{\omega}_0,\tilde{\omega}_1)} + 1 \right)$$

$$= \log\left(\inf_{\substack{\tilde{\omega}_0 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}(p\rho_0^{\otimes n}),\, \tilde{\omega}_1 \in \mathcal{B}_{\tilde{\varepsilon}(\varepsilon,n)}((1-p)\rho_1^{\otimes n}) \\ \mathrm{Tr}(\tilde{\omega}_0) + \mathrm{Tr}(\tilde{\omega}_1) = 1}} 2^{d_{\mathrm{T}}(\tilde{\omega}_0,\tilde{\omega}_1)} + 1 \right)$$

$$\leqslant \log\left(\frac{4}{\tilde{\varepsilon}(\varepsilon,n)} + 1 \right)$$

$$= \log\left(\frac{1}{\tilde{\varepsilon}(\varepsilon,n)} \right) + \log(4 + \tilde{\varepsilon}(\varepsilon,n))$$

$$\leqslant \log\left(\frac{1}{\tilde{\varepsilon}(\varepsilon,n)} \right) + \log(4 + \varepsilon), \tag{5.105}$$

where we have used equation (5.76) in lemma 5.17 for the fourth line and $\tilde{\varepsilon}(\varepsilon,n) \leqslant \varepsilon$ for the last inequality. Hence, we finally see by definition of $\tilde{\varepsilon}(\varepsilon,n)$

$$\frac{\xi_c^{\varepsilon,\star}(\rho_{XA}^{(n)})}{n} \leqslant \frac{\log\left(\frac{1}{\tilde{\varepsilon}(\varepsilon,n)} \right)}{n} + \frac{\log(4+\varepsilon) - 1}{n}$$

$$= \frac{-\log(p_{\mathrm{err}}(\rho_{XA}^{(n)}))}{n} - \frac{\log(\varepsilon/2)}{n} + \frac{\log(4+\varepsilon) - 1}{n}. \tag{5.106}$$

Noting once again that

$$\lim_{n\to\infty} \frac{-\log(p_{\mathrm{err}}(\rho_{XA}^{(n)}))}{n} = \xi(\rho_0,\rho_1)$$

and the other terms on the right-hand side of (5.106) are $\mathcal{O}(1)$ in n finishes the proof. $\square$

6. Examples

In this section, we detail a few examples of the RTSD to illustrate some of the key theoretical concepts developed in the previous sections. We begin with a first example. Let ρ_0 and ρ_1 be the following states:

$$\rho_0 := \mathcal{A}^{\gamma,N}(|0\rangle\langle 0|), \tag{6.1}$$

$$\rho_1 := \mathcal{A}^{\gamma,N}(|1\rangle\langle 1|), \tag{6.2}$$

where $\mathcal{A}^{\gamma,N}$ is the generalized amplitude damping channel, defined as

$$\mathcal{A}^{\gamma,N}(\omega) := \sum_{i=0}^{3} A_i \omega A_i^\dagger, \tag{6.3}$$

with $\gamma, N \in [0,1]$ and

$$A_0 := \sqrt{1-N}\left(|0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|\right), \tag{6.4}$$

$$A_1 := \sqrt{\gamma(1-N)}|0\rangle\langle 1|, \tag{6.5}$$

$$A_2 := \sqrt{N}\left(\sqrt{1-\gamma}|0\rangle\langle 0| + |1\rangle\langle 1|\right), \tag{6.6}$$

$$A_3 := \sqrt{\gamma N}|1\rangle\langle 0|. \tag{6.7}$$

The parameter $\gamma \in [0,1]$ is a damping parameter and $N \in [0,1]$ is a thermal noise parameter. The generalized amplitude damping channel models the dynamics of a two-level system in contact with a thermal bath at non-zero temperature [NC10] and can be used as a phenomenological model for relaxation noise in superconducting qubits [CB08]. See [KSW20] for an in-depth study of the information-theoretic properties of this channel and for a discussion of how this channel can be interpreted as a qubit thermal attenuator channel. We choose the prior probabilities for the states ρ_0 and ρ_1 to be q and $1-q$ respectively, with $q = 1/3$, so that the c–q state describing the elementary quantum source is

$$\rho_{XA} := q|0\rangle\langle 0| \otimes \rho_0 + (1-q)|1\rangle\langle 1| \otimes \rho_1. \tag{6.8}$$

In figure 1, we set the thermal noise parameter $N = 0.1$ and plot the exact one-shot distillable-SD of ρ_{XA} under CPTP_A maps, and under CDS maps, and the exact one-shot SD-cost of ρ_{XA} under CPTP_A maps, and under CDS maps. As expected, when the damping parameter γ increases, each measure of SD decreases. The exact distillable-SD and SD-cost under CDS maps do not decrease to zero due to the non-uniform prior ($q = 1/3$). However, they do decrease to zero under CPTP_A maps because the prior q does not play a role in this case. Additionally, the SD-cost under CPTP_A maps is strictly larger than the distillable-SD under CPTP_A maps for all $\gamma \in (0,1)$, demonstrating that the RTSD is not reversible in this one-shot scenario. The same holds for distillable-SD and SD-cost under CDS maps.

The SD-cost under CPTP_A maps is smaller than that under CDS maps because we use different golden units in these two cases. In this context, recall definition 5.1. It is not possible for CPTP_A maps to change the prior q. So we are forced to use the golden unit with prior q, i.e. $\gamma^{(M,q)}$, which in this case we chose to be $q = 1/3$. CDS maps, however, can change the prior, and as mentioned in definition 5.1, we pick the prior of the golden unit to be the canonical choice of $1/2$. Also, note that $\gamma^{(M,1/3)}$ has more SD than $\gamma^{(M,1/2)}$; i.e. it dominates $\gamma^{(M,1/2)}$ in the preorder of SD and hence can be transformed into the prior $1/2$ golden unit via CDS (see lemma 3.6). So the SD costs under CPTP_A and CDS maps are different, as we are paying with a less valuable currency in the case of CDS maps.

We next consider the following example:

$$\rho_0 := \mathcal{A}^{\gamma,N}(|0\rangle\langle 0|), \tag{6.9}$$

$$\rho_1 := e^{i\phi\sigma^{(1)}}\mathcal{A}^{\gamma,N}(|1\rangle\langle 1|)e^{-i\phi\sigma^{(1)}}, \tag{6.10}$$

where the angle $\phi \in [0, \pi/2]$, we choose the prior q to be the same (i.e. $q = 1/3$). All of the quantities mentioned above are plotted in figure 2. As the angle ϕ increases from zero to $\pi/2$, the states $\mathcal{A}^{\gamma,N}(|0\rangle\langle 0|)$ and $e^{i\phi\sigma^{(1)}}\mathcal{A}^{\gamma,N}(|1\rangle\langle 1|)e^{-i\phi\sigma^{(1)}}$ become less distinguishable and become the same state when $\phi = \pi/2$. Thus, we expect for the various measures of SD to decrease as ϕ increases from zero to $\pi/2$. Similar statements as given above apply regarding the difference between the SD quantities under CPTP_A and CDS maps.

As another example, we plot the logarithm of the one-shot approximate distillable-SD of the states in (6.1) and (6.2) under both CPTP_A and CDS maps, as a function of the damping parameter γ. We set the
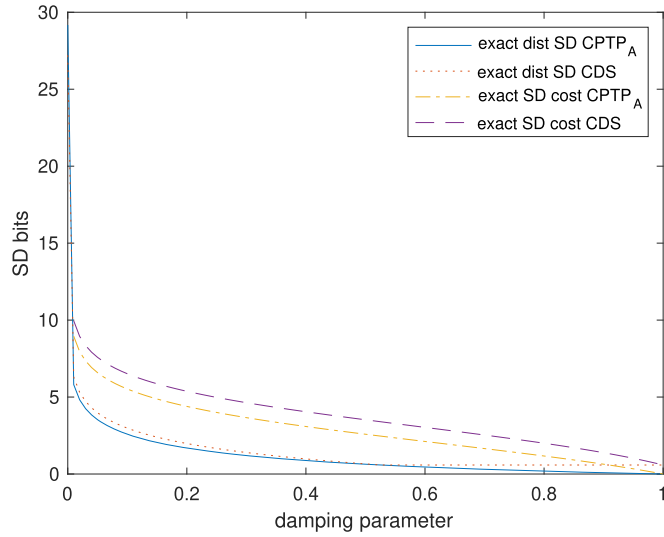
Figure 1. Various operational measures of SD for the states in (6.1) and (6.2) with prior $q = 1/3$, as a function of the damping parameter γ. The noise parameter $N = 0.1$.
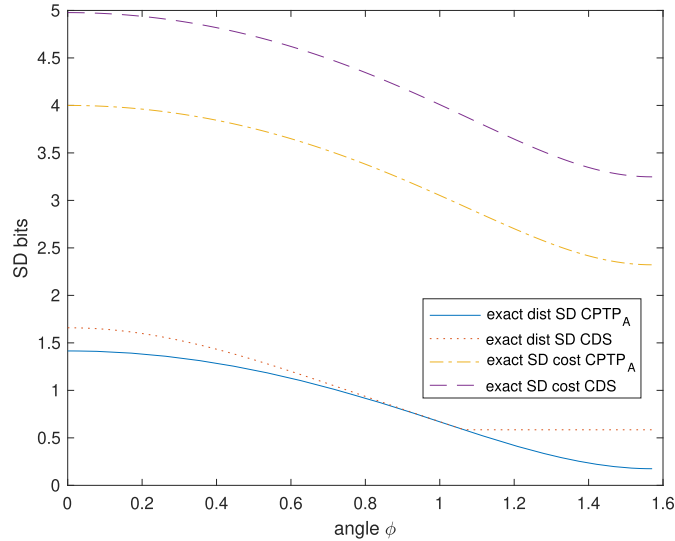

Figure 2. Various operational measures of SD for the states in (6.9)–(6.10) with prior $q = 1/3$, as a function of the angle parameter ϕ. The damping parameter $\gamma = 1/4$ and the noise parameter $N = 0.1$.

approximate error $\varepsilon = 0.1$, the prior probability $q = 1/3$, and the noise parameter $N = 0.1$. For reference, we also plot the logarithm of the exact distillable-SD under both CPTP$_A$ and CDS maps. See figure 3. We have plotted the logarithm of the number of SD bits in order to distinguish the curves more clearly. The difference in the behavior of the curves has to do with the fact that CDS maps can change the prior probability while CPTP$_A$ maps cannot. Here, the distillable-SD under CDS maps (both in the approximate and exact cases) flattens out for values of the damping parameter greater than $\gamma \approx 0.5$ as in this case SD of the considered box is exclusively due to the non-uniform prior ($q = 1/3$) and does not decrease further even if the quantum states themselves become less distinguishable.

As a final example, we plot the minimum conversion error in (3.30) when transforming the box

$$(1/3, \mathcal{A}^{\gamma_1, N_1}(|0\rangle\langle 0|), \mathcal{A}^{\gamma_1, N_1}(|1\rangle\langle 1|))$$

to the box

$$(1/4, \mathcal{A}^{\gamma_2, N_2}(|0\rangle\langle 0|), e^{i\phi\sigma^{(1)}} \mathcal{A}^{\gamma_2, N_2}(|1\rangle\langle 1|) e^{-i\phi\sigma^{(1)}})$$

as a function of the angle $\phi \in [0, \pi/2]$, with $\gamma_1 = 0.5$, $N_1 = 0.3$, $\gamma_2 = 0.25$, and $N_2 = 0.1$. To do so, we make use of the SDP from proposition 3.16. The minimum conversion error is plotted in figure 4 as a function of the angle ϕ. Intuitively, for small values of the angle ϕ, it should be more difficult to perform
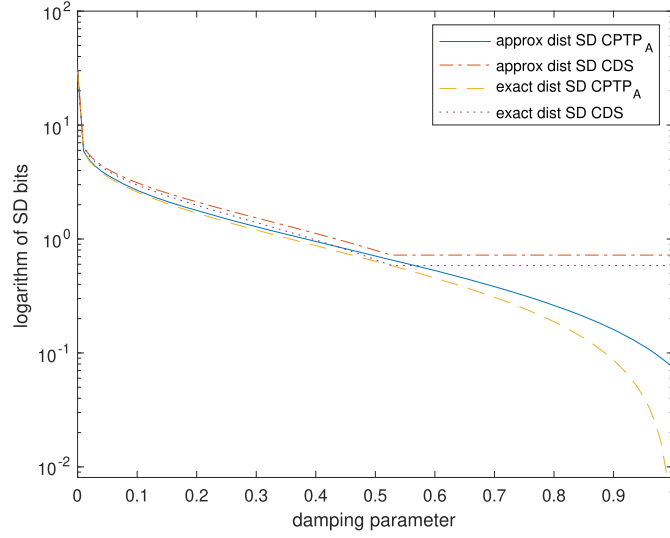
Figure 3. Logarithm of various operational measures of SD for the states in (6.1) and (6.2) with prior $q = 1/3$, as a function of the damping parameter γ. The noise parameter $N = 0.1$, and the approximation error $\varepsilon = 0.1$. We have plotted the logarithm of the number of SD bits in order to distinguish the curves more clearly.
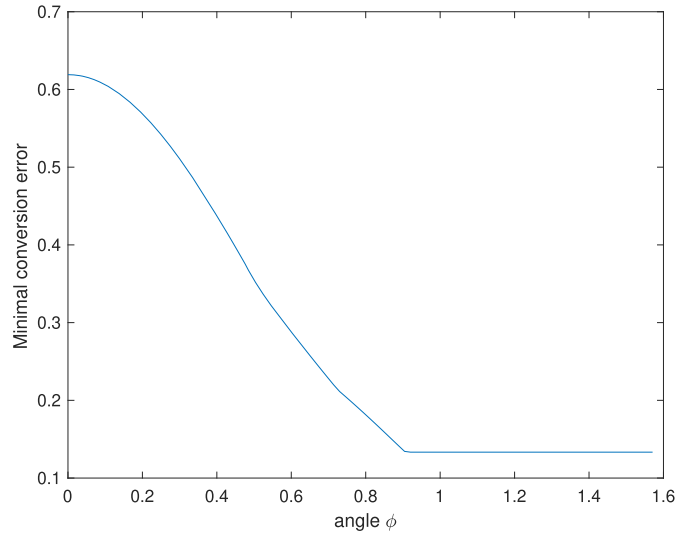

Figure 4. Minimum conversion error when transforming $(1/3, \mathcal{A}^{\gamma_1,N_1}(|0\rangle\langle0|), \mathcal{A}^{\gamma_1,N_1}(|1\rangle\langle1|))$ to $(1/4, \mathcal{A}^{\gamma_2,N_2}(|0\rangle\langle0|), e^{i\phi\sigma^{(1)}} \mathcal{A}^{\gamma_2,N_2}(|1\rangle\langle1|)e^{-i\phi\sigma^{(1)}})$ as a function of the angle $\phi \in [0, \pi/2]$, with $\gamma_1 = 0.5$, $N_1 = 0.3$, $\gamma_2 = 0.25$, and $N_2 = 0.1$.

the conversion because the states in the first box are less distinguishable than those in the second, and so we expect the error to be higher. However, as the angle ϕ increases, the states in the second box become less distinguishable and so the transformation becomes easier. The difference in the prior probabilities of the boxes is a fundamental limitation that cannot be overcome, even as ϕ becomes closer to $\pi/2$, so that the minimum conversion error plateaus for angle values greater than ≈ 0.9.

All Matlab programs that generate the above plots (along with the SDPs) are available with the arXiv ancillary files of this paper.

7. Asymptotic transformation task

Let ρ_{XA} and σ_{XB} be c–q states explicitly given by

$$\rho_{XA} = p \, |0\rangle\langle0| \otimes \rho_0 + (1-p) \, |1\rangle\langle1| \otimes \rho_1, \tag{7.1}$$

$$\sigma_{XB} = q \, |0\rangle\langle0| \otimes \sigma_0 + (1-q) \, |1\rangle\langle1| \otimes \sigma_1, \tag{7.2}$$

with ρ_0, ρ_1 states of a quantum system A, and σ_0, σ_1 states of a quantum system B. Moreover, we assume that $p, q \in [0, 1]$. We use the short-hand notation $\vec{p} := (p, 1 - p)$ and $\vec{q} := (q, 1 - q)$ for the prior (distribution) of ρ_{XA} and σ_{XB} respectively and write $\vec{p} \succ \vec{q}$ if $\vec{p}$ majorizes $\vec{q}$. Let $\rho_{XA}^{(n)}$ and $\sigma_{XB}^{(m)}$ be as follows:

$$\rho_{XA}^{(n)} = p \, |0\rangle\langle 0| \otimes \rho_0^{\otimes n} + (1 - p) \, |1\rangle\langle 1| \otimes \rho_1^{\otimes n}, \tag{7.3}$$

$$\sigma_{XB}^{(m)} = q \, |0\rangle\langle 0| \otimes \sigma_0^{\otimes m} + (1 - q) \, |1\rangle\langle 1| \otimes \sigma_1^{\otimes m}. \tag{7.4}$$

Definition 7.1. Let ρ_{XA}, σ_{XB} be c−q states, and let FO denote the set of free operations. For $n, m \in \mathbb{N}$ and $\varepsilon > 0$, we say that there exists a (n, m, ε) FO-*transformation protocol* for the states ρ_{XA} and σ_{XB} if

$$d'_{\mathrm{FO}}(\rho_{XA}^{(n)} \mapsto \sigma_{XB}^{(m)}) \leqslant \varepsilon. \tag{7.5}$$

That is, there exists an $\mathcal{A} \in \mathrm{FO}$ such that

$$D'\left(\mathcal{A}(\rho_{XA}^{(n)}), \sigma_{XB}^{(m)}\right) \leqslant \varepsilon. \tag{7.6}$$

We denote such a transformation protocol in short by the notation $\rho_{XA} \mapsto \sigma_{XB}$.

Definition 7.2. The rate $R \geqslant 0$ is an achievable rate for the transformation $\rho_{XA} \mapsto \sigma_{XB}$ under free operations FO, if for all $\varepsilon, \delta > 0$ and $n \in \mathbb{N}$ large enough there exists an $(n, \lfloor n(R - \delta) \rfloor, \varepsilon)$ FO-transformation protocol. The optimal rate is given by the supremum over all achievable rates, and is denoted by

$$R_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}) = \sup \{ R \geqslant 0 \, | R \, \text{achievable rate under FO} \} . \tag{7.7}$$

In particular, in the case of free operations being CPTP_A we write

$$R(\rho_{XA} \mapsto \sigma_{XB}) \equiv R_{\mathrm{CPTP}_A}(\rho_{XA} \mapsto \sigma_{XB}), \tag{7.8}$$

and in the case of free operations being CDS we write

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) \equiv R_{\mathrm{CDS}}(\rho_{XA} \mapsto \sigma_{XB}). \tag{7.9}$$

Note that by the inclusion $\mathrm{CPTP}_A \subset \mathrm{CDS}$ we immediately get the inequality

$$R(\rho_{XA} \mapsto \sigma_{XB}) \leqslant R^\star(\rho_{XA} \mapsto \sigma_{XB}). \tag{7.10}$$

Definition 7.3 (Strong converse rate). The rate $R \geqslant 0$ is a strong converse rate for the transformation $\rho_{XA} \mapsto \sigma_{XB}$ under free operations FO, if for all $\varepsilon, \delta > 0$ and $n \in \mathbb{N}$ large enough there does not exist an $(n, \lceil n(R + \delta) \rceil, \varepsilon)$ FO-transformation protocol. The optimal strong converse rate is given by the infimum over all strong converse rates, and is denoted by

$$\tilde{R}_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}) = \inf \{ R \geqslant 0 \, | \, R \, \text{strong converse rate under FO} \} . \tag{7.11}$$

In particular, in the case of free operations being CPTP_A we write

$$\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) \equiv \tilde{R}_{\mathrm{CPTP}_A}(\rho_{XA} \mapsto \sigma_{XB}), \tag{7.12}$$

and in the case of free operations being CDS we write

$$\tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) \equiv \tilde{R}_{\mathrm{CDS}}(\rho_{XA} \mapsto \sigma_{XB}). \tag{7.13}$$

By definition, we have

$$R_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}) \leqslant \tilde{R}_{\mathrm{FO}}(\rho_{XA} \mapsto \sigma_{XB}), \tag{7.14}$$

and, moreover, by again using the fact that $\mathrm{CPTP}_A \subset \mathrm{CDS}$, we get the inequality

$$\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) \leqslant \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}). \tag{7.15}$$

The following theorem gives expressions for the optimal achievable and strong converse rates for the transformation $\rho_{XA} \mapsto \sigma_{XB}$ under both CDS and CPTP_A. Note that in the following we interpret $\frac{\infty}{\infty}$ as ∞.

Theorem 7.4. *Let ρ_{XA} and σ_{XB} be the c−q states defined through (7.1) and (7.2) with $p, q \in (0, 1)$.*
For free operations being CDS *we have: for $\xi(\sigma_0, \sigma_1) > 0$*

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}. \tag{7.16}$$

For $\xi(\sigma_0, \sigma_1) = 0$ and $\xi(\rho_0, \rho_1) > 0$ we have

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty. \tag{7.17}$$

For $\xi(\sigma_0, \sigma_1) = \xi(\rho_0, \rho_1) = 0$, we have

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, \vec{p} \succ \vec{q}, \tag{7.18}$$

$$whereas \quad R^\star(\rho_{XA} \mapsto \sigma_{XB}) = 0 \quad and \quad \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad else. \tag{7.19}$$

For free operations being CPTP_A we have: in the case of ρ_{XA} and σ_{XB} having equal priors

$$R(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}. \tag{7.20}$$

Here, we interpreted $\frac{0}{0}$ as ∞.
In the case of the priors being different we get

$$R(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = 0, \quad if\, \xi(\sigma_0, \sigma_1) > 0, \tag{7.21}$$

$$R(\rho_{XA} \mapsto \sigma_{XB}) = 0, \qquad \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, \xi(\sigma_0, \sigma_1) = 0. \tag{7.22}$$

Remark 7.5. For simplicity we excluded in theorem 7.4 the case of singular priors, i.e. $p \in \{0, 1\}$ or $q \in \{0, 1\}$. For completeness we now state the corresponding results on optimal and strong converse rates in these cases:

For free operations being CDS we have

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, p \in \{0, 1\}, \tag{7.23}$$

and for $p \in (0, 1)$ and $q \in \{0, 1\}$ we have

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, \xi(\rho_0, \rho_1) = \infty, \tag{7.24}$$

$$whereas \quad R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) = 0, \quad if\, \xi(\rho_0, \rho_1) < \infty. \tag{7.25}$$

For free operations being CPTP_A and $q \in \{0, 1\}$ we have

$$R(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, p = q, \tag{7.26}$$

$$R(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = 0, \quad if\, p \neq q. \tag{7.27}$$

And for $p \in \{0, 1\}$ and $q \in (0, 1)$ we have similarly to (7.21) and (7.22)

$$R(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = 0, \quad if\, \xi(\sigma_0, \sigma_1) > 0, \tag{7.28}$$

$$R(\rho_{XA} \mapsto \sigma_{XB}) = 0, \qquad \tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \quad if\, \xi(\sigma_0, \sigma_1) = 0. \tag{7.29}$$

Remark 7.6. Further to what was already stated in remark 5.15, theorem 7.4 expresses the fact that the RTSD is asymptotically reversible. Indeed, the optimal asymptotic rate at which one can convert ρ_{XA} to σ_{XB} is equal to the ratio of quantum Chernoff divergences. The rate at which one can convert back is thus equal to the reciprocal of the forward rate. Since the product of these two rates is equal to one, we conclude that the RTSD is asymptotically reversible.

7.1. Proof of theorem 7.4

7.1.1. Achievability
We start the proof of theorem 7.4 by proving the achievability part. In particular, we show the following lemma.

Lemma 7.7. *Let ρ_{XA} and σ_{XB} be the c–q states defined through (7.1) and (7.2) with $p, q \in (0, 1)$. For $\xi(\sigma_0, \sigma_1) > 0$, we have*

$$R^\star(\rho_{XA} \mapsto \sigma_{XB}) \geqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}. \tag{7.30}$$

Moreover, for $\xi(\sigma_0, \sigma_1) = 0$ and $\xi(\rho_0, \rho_1) > 0$, we get $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty$ and in the case $\xi(\sigma_0, \sigma_1) = \xi(\rho_0, \rho_1) = 0$ we get $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty$ if $\vec{p} \succ \vec{q}$, and $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = 0$ otherwise.
Moreover, in the case of ρ_{XA} and σ_{XB} having the same priors we have

$$R(\rho_{XA} \mapsto \sigma_{XB}) \geqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}. \tag{7.31}$$

Here, we interpreted $\frac{\infty}{\infty}$ and $\frac{0}{0}$ as ∞. In the case of priors being different we have $R(\rho_{XA} \mapsto \sigma_{XB}) = 0$.

Proof. We prove the result for free operations being CDS, since for free operations being $CPTP_A$ the proof follows the same lines. Let us first consider the case $\xi(\sigma_0, \sigma_1) = 0$, in which case necessarily $\sigma_0 = \sigma_1 \equiv \sigma$. Moreover, first assume that $\xi(\rho_0, \rho_1) = 0$ (which implies that $\rho_0 = \rho_1 \equiv \rho$) and $\vec{p} \succ \vec{q}$. In that case, there exists a $\lambda \in [0, 1]$ such that

$$\lambda p + (1 - \lambda)(1 - p) = q,$$

and consequently

$$\lambda(1 - p) + (1 - \lambda)p = 1 - q.$$

Therefore, considering for every $m \in \mathbb{N}$

$$\mathcal{E}_0^{(m)}(\cdot) = \lambda \operatorname{Tr}(\cdot)\sigma^{\otimes m}, \qquad \mathcal{E}_1^{(m)}(\cdot) = (1 - \lambda)\operatorname{Tr}(\cdot)\sigma^{\otimes m},$$

which are quantum operations summing to a CPTP map, and the corresponding CDS map

$$\mathcal{N}^{(m)} = \operatorname{id} \otimes \mathcal{E}_0^{(m)} + \mathcal{F} \otimes \mathcal{E}_1^{(m)}, \tag{7.32}$$

we get

$$\mathcal{N}^{(m)}(\rho_{XA}) = \left(\lambda p + (1 - \lambda)(1 - p)\right) |0\rangle\langle 0| \otimes \sigma^{\otimes m} + \left(\lambda(1 - p) + (1 - \lambda)p\right) |1\rangle\langle 1| \otimes \sigma^{\otimes m}$$

$$= \sigma_{XB}^{(m)}.$$

Consequently, $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty$, since ρ_{XA} can be transformed to $\sigma_{BX}^{(m)}$ without error via a CDS map for an arbitrary $m \in \mathbb{N}$.

Consider now the case $\xi(\rho_0, \rho_1) = \xi(\sigma_0, \sigma_1) = 0$ and $\vec{p} \not\succ \vec{q}$. In this case, using lemma 3.14, we see that for all $n, m \in \mathbb{N}$

$$\min\{p, (1 - p)\} = p_{\operatorname{err}}(\rho_{XA}) = p_{\operatorname{err}}(\rho_{XA}^{(n)}) \leqslant \left(d'_{\operatorname{CDS}}(\rho_{XA}^{(n)} \mapsto \sigma_{XB}^{(m)}) + 1\right) p_{\operatorname{err}}(\sigma_{XB}^{(m)})$$

$$= \left(d'_{\operatorname{CDS}}(\rho_{XA}^{(n)} \mapsto \sigma_{XB}^{(m)}) + 1\right) p_{\operatorname{err}}(\sigma_{XB})$$

$$= \left(d'_{\operatorname{CDS}}(\rho_{XA}^{(n)} \mapsto \sigma_{XB}^{(m)}) + 1\right) \min\{q, 1 - q\},$$

and since $\min\{p, (1 - p)\} = (c + 1)\min\{q, (1 - q)\}$ for some $c > 0$, we get

$$d'_{\operatorname{CDS}}(\rho_{XA}^{(n)} \mapsto \sigma_{XB}^{(m)}) \geqslant c > 0$$

for all $n, m \in \mathbb{N}$, and hence $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = 0$.

Now consider the case in which $\xi(\sigma_0, \sigma_1) = 0$ but $\xi(\rho_0, \rho_1) > 0$. Assume without loss of generality that $q \leqslant 1 - q$, and for all $n \in \mathbb{N}$, let $\{\Lambda_0^{(n)}, \Lambda_1^{(n)}\}$ denote the optimal POVM for discriminating the quantum states $\rho_0^{\otimes n}$ and $\rho_1^{\otimes n}$ of the c–q state $\rho_{XA}^{(n)}$. Moreover, for all $m \in \mathbb{N}$, let

$$\mathcal{E}_0^{(n,m)}(\cdot) = \operatorname{Tr}(\Lambda_1^{(n)}\cdot)\sigma^{\otimes m}, \qquad \mathcal{E}_1^{(n,m)}(\cdot) = \operatorname{Tr}(\Lambda_0^{(n)}\cdot)\sigma^{\otimes m}, \tag{7.33}$$

which are quantum operations summing to a CPTP map, and define the corresponding CDS map

$$\mathcal{N}^{(n,m)} = \operatorname{id} \otimes \mathcal{E}_0^{(n,m)} + \mathcal{F} \otimes \mathcal{E}_1^{(n,m)}.$$

This gives

$$\mathcal{N}^{(n,m)}(\rho_{XA}^{(n)}) = p_{\operatorname{err}}(\rho_{XA}^{(n)}) |0\rangle\langle 0| \otimes \sigma^{\otimes m} + (1 - p_{\operatorname{err}}(\rho_{XA}^{(n)})) |1\rangle\langle 1| \otimes \sigma^{\otimes m}.$$

Now choose n large enough such that $p_{\operatorname{err}}(\rho_{XA}^{(n)}) \leqslant q$, which gives that the prior of $\mathcal{N}^{(n,m)}(\rho_{XA}^{(n)})$ majorises the prior of $\sigma_{XB}^{(m)}$. Using the above, i.e. the construction around (7.32), we can find a CDS map transforming $\mathcal{N}^{(n,m)}(\rho_{XA}^{(n)})$ to $\sigma_{XB}^{(m)}$ for all m without error. This implies that $R^\star(\rho_{XA} \mapsto \sigma_{XB}) = \infty$.

Next consider the case in which $\xi(\sigma_0, \sigma_1) > 0$. We can assume without loss of generality that $\xi(\rho_0, \rho_1) > 0$, since otherwise (7.31) is trivially satisfied. The case $\xi(\rho_0, \rho_1) = \infty$ follows by lemma 3.3, i.e. by the fact that we can transform any infinite-resource state to any other c–q state via CDS maps without error. Note here that $\xi(\rho_0, \rho_1) = \infty$ if and only if ρ_0 and ρ_1 have orthogonal supports. Furthermore, the case in which $\xi(\rho_0, \rho_1) < \infty$ and $\xi(\sigma_0, \sigma_1) = \infty$ follows from the fact that any transformation from a finite resource to an infinite-resource has infinite error with respect to the scaled trace distance D'.

Let us therefore now finally consider the case in which $0 < \xi(\rho_0, \rho_1), \xi(\sigma_0, \sigma_1) < \infty$, and fix $\varepsilon, \delta > 0$. By theorem 4.7 we can find for all $\delta_1 > 0$, an $N_1 \in \mathbb{N}$ such that for all $n \geqslant N_1$ there exists a CDS map $\mathcal{N}_1$ such that

$$\mathcal{N}_1(\rho_{XA}^{(n)}) = \gamma_{XQ}^{(M_n)},$$

where $\gamma_{XQ}^{(M_n)}$ is the M_n-golden unit defined in (3.10) with

$$\frac{\log M_n}{n} \geqslant \xi(\rho_0, \rho_1) - \delta_1. \tag{7.34}$$

Since $\xi(\rho_0, \rho_1) > 0$, without loss of generality even $\xi(\rho_0, \rho_1) > \delta_1$ by picking δ_1 small enough, we also see that the sequence $(M_n)_{n \geqslant N}$ goes to infinity for $n \to \infty$. For each M_n consider $m_n \in \mathbb{N}$ to be the unique number such that

$$\xi_c^{\star,\varepsilon}(\sigma_{XB}^{(m_n)}) \leqslant \log M_n \leqslant \xi_c^{\star,\varepsilon}(\sigma_{XB}^{(m_n+1)}). \tag{7.35}$$

First, note that we also get $m_n \xrightarrow[n\to\infty]{} \infty$ because $M_n \xrightarrow[n\to\infty]{} \infty$. Also note that as each $\xi_c^{\star,\varepsilon}(\sigma_{XB}^{(m_n)})$ is finite there exists a CDS map $\mathcal{N}_2$ such that

$$D'\left(\mathcal{N}_2(\gamma_{XQ}^{(M_n)}), \sigma_{XB}^{(m_n)}\right) \leqslant \varepsilon$$

and hence in total

$$D'\left(\mathcal{N}_2 \circ \mathcal{N}_1(\rho_{XA'}^{(n)}), \sigma_{XB}^{(m_n)}\right) \leqslant \varepsilon.$$

Moreover, we have by (7.34) and (7.35)

$$\frac{\xi_c^{\star,\varepsilon}(\sigma_{XB}^{(m_n+1)})}{n} \geqslant \frac{\log M_n}{n} \geqslant \xi(\rho_0, \rho_1) - \delta_1. \tag{7.36}$$

By theorem 5.13 we can pick for all $\delta_2 > 0$, an $N_2 \in \mathbb{N}$ such that for all $n \geqslant \max\{N_1, N_2\}$ we have

$$\frac{\xi_c^{\star,\varepsilon}(\sigma_{XB}^{(m_n+1)})}{m_n + 1} \leqslant \xi(\sigma_0, \sigma_1) + \delta_2$$

and therefore

$$\frac{m_n + 1}{n}(\xi(\sigma_0, \sigma_1) + \delta_2) \geqslant \frac{\log M_n}{n} \geqslant \xi(\rho_0, \rho_1) - \delta_1. \tag{7.37}$$

As $\delta_1, \delta_2 > 0$ are arbitrary, we can choose them small enough and moreover possibly n even larger such that

$$\frac{\xi(\rho_0, \rho_1) - \delta_1}{\xi(\sigma_0, \sigma_1) + \delta_2} - \frac{1}{n} \geqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)} - \delta.$$

In summary, we have proven that for all $\varepsilon, \delta > 0$ and n large enough, there exists an (n, m_n, ε)CDS-transformation protocol with

$$\frac{m_n}{n} \geqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)} - \delta, \tag{7.38}$$

which shows that $R := \frac{\xi(\rho_0,\rho_1)}{\xi(\sigma_0,\sigma_1)}$ is an achievable rate and hence proves (7.30). $\qquad\square$

7.1.2. Strong converse

We now prove the strong converse part of theorem 7.4. The proof follows a similar idea as the one for proving the strong converse in the RTAD [WW19a]. There, the key insight was the pseudo continuity bound for either the sandwiched Rényi relative entropy [WW19a, lemma 1] or Petz–Rényi relative entropy [WW19a, lemma 3]. However, here we instead use a pseudo continuity bound for the operational quantity $-\log(p_{\mathrm{err}}(\cdot))$, involving the scaled trace distance D', and holding for two arbitrary cq-states ρ_{XA} and σ_{XA}:

$$-\log(p_{\mathrm{err}}(\rho_{XA})) + \log(p_{\mathrm{err}}(\sigma_{XA})) \geqslant -\log\left(1 + D'(\rho_{XA}, \sigma_{XA})\right). \tag{7.39}$$

Note that (7.39) directly follows from lemma 3.14. With that we prove the following lemma:

Lemma 7.8. *Let ρ_{XA} and σ_{XB} be the c–q states defined through (7.1) and (7.2) with $p, q \in (0,1)$. Then we have*

$$\tilde{R}^\star(\rho_{XA} \mapsto \sigma_{XB}) \leqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}, \tag{7.40}$$

and

$$\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) \leqslant \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}. \tag{7.41}$$

Here, we interpreted $\frac{\infty}{\infty}$ and $\frac{0}{0}$ as ∞.
Furthermore, in the case in which $\xi(\sigma_0, \sigma_1) = 0$ we get

$$\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^{\star}(\rho_{XA} \mapsto \sigma_{XB}) = \infty. \tag{7.42}$$

Proof. As $\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) \leqslant \tilde{R}^{\star}(\rho_{XA} \mapsto \sigma_{XB})$ we only need to show (7.40) as (7.41) directly follows. Moreover, consider $\xi(\sigma_0, \sigma_1) > 0$ as otherwise (7.40) is trivially true. Let $n, m \in \mathbb{N}$ and $\varepsilon > 0$ be such that there exists an (n, m, ε) CDS-transformation protocol. Hence, there exists a CDS map $\mathcal{N}$ such that

$$D'\left(\mathcal{N}(\rho_{XA}^{(n)}), \sigma_{XB}^{(m)}\right) \leqslant \varepsilon. \tag{7.43}$$

Using first monotonicity of the minimum error probability under CDS maps and then (7.39) we get

$$-\log\left(p_{\mathrm{err}}(\rho_{XA}^{(n)})\right) \geqslant -\log\left(p_{\mathrm{err}}(\mathcal{N}(\rho_{XA}^{(n)}))\right) \geqslant -\log\left(p_{\mathrm{err}}(\sigma_{XB}^{(m)})\right) - \log(1 + \varepsilon)$$
$$\geqslant m\xi(\sigma_0, \sigma_1) - \log(1 + \varepsilon), \tag{7.44}$$

where we have used [ACMnT$^+$07, theorem 1] for the last inequality. This gives

$$\frac{-\log\left(p_{\mathrm{err}}(\rho_{XA}^{(n)})\right)}{\xi(\sigma_0, \sigma_1)} \geqslant m - \frac{\log(1 + \varepsilon)}{\xi(\sigma_0, \sigma_1)}. \tag{7.45}$$

Let now $\delta_1 > 0$ and consider $n \in \mathbb{N}$ large enough such that

$$\frac{-\log\left(p_{\mathrm{err}}(\rho_{XA}^{(n)})\right)}{n} \leqslant \xi(\rho_0, \rho_1) + \delta_1, \tag{7.46}$$

which is possible by the results of [NS09]. Substituting this into (7.45) we get

$$\frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)} \geqslant \frac{m}{n} - \frac{\log(1 + \varepsilon)}{n\xi(\sigma_0, \sigma_1)} - \frac{\delta_1}{\xi(\sigma_0, \sigma_1)}. \tag{7.47}$$

Hence, by picking $\delta_1 > 0$ small enough, we see that for all $\varepsilon, \delta > 0$ and n large enough a possible (n, m, ε)CDS-transformation protocol satisfies

$$\frac{m}{n} < \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)} + \delta, \tag{7.48}$$

which shows that $R := \frac{\xi(\rho_0, \rho_1)}{\xi(\sigma_0, \sigma_1)}$ is a strong converse rate and finishes the proof.

It remains to discuss the case $\xi(\sigma_0, \sigma_1) = 0$, in which case $\sigma_0 = \sigma_1 \equiv \sigma$, and to show that then $\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^{\star}(\rho_{XA} \mapsto \sigma_{XB}) = \infty$. Consider for all $n, m \in \mathbb{N}$ the CPTP map $\mathcal{E}^{(m)} = \mathrm{Tr}(\cdot)\sigma^{\otimes m}$ and note that

$$\left(\mathrm{id} \otimes \mathcal{E}^{(m)}\right)(\rho_{XA}^{(n)}) = p\,|0\rangle\langle 0| \otimes \sigma^{\otimes m} + (1 - p)\,|1\rangle\langle 1| \otimes \sigma^{\otimes m}.$$

Hence,

$$D'\left(\left(\mathrm{id} \otimes \mathcal{E}^{(m)}\right)(\rho_{XA}^{(n)}), \sigma_{XB}^{(m)}\right) = \frac{\left\|\left(\mathrm{id} \otimes \mathcal{E}^{(m)}\right)(\rho_{XA}^{(n)}) - \sigma_{XB}^{(m)}\right\|_1}{2p_{\mathrm{err}}(\sigma_{XB}^{(m)})} = \frac{|p - q|}{\min\{q, 1 - q\}} < \infty.$$

Therefore, for $\varepsilon := \frac{|p - q|}{\min\{q, 1 - q\}}$ fixed and all $n, m \in \mathbb{N}$ we have found a (n, m, ε) transformation protocol (under both CPTP$_A$ and CDS) which implies that

$$\tilde{R}(\rho_{XA} \mapsto \sigma_{XB}) = \tilde{R}^{\star}(\rho_{XA} \mapsto \sigma_{XB}) = \infty, \tag{7.49}$$

concluding the proof. $\qquad \square$

8. Summary and open questions

In summary, we have introduced the RTSD and have answered many of the fundamental questions associated with it. In particular, we have developed an axiomatic approach to the RTSD, which led to the conclusion that CDS maps are the natural choice for free operations, with $CPTP_A$ maps being a special case. We then introduced the golden units of the RTSD and argued why a particular scaled trace distance is a more appropriate figure of merit for approximate transformations, instead of the standard trace distance. We finally defined and studied the tasks of dilution, distillation, and transformation, in the exact and approximate cases, both in the one-shot and asymptotic scenarios. We proved that the rate at which asymptotic transformations are possible is equal to the ratio of quantum Chernoff divergences of the elementary information sources, and we thus concluded that the RTSD is asymptotically reversible.

Going forward from here, it would be interesting to generalize the RTSD that we developed in this paper for elementary information sources to more general information sources, i.e. to c–q states for which the classical alphabet has a size greater than two. We note here that many of the concepts considered in our paper, such as the basic axioms for the RTSD, CDS maps, and the scaled trace distance $D'(\cdot, \cdot)$ already apply to this more general setting. In light of the seminal result in [Li16], it is a tantalizing possibility that the optimal conversion rate between quantum information sources would be equal to a ratio of multiple-state Chernoff divergences, as a generalization of theorem 7.4, but it remains open to determine if it is the case. It is also interesting to determine expressions for the one-shot distillable-SD and SD-cost, as generalizations of $\xi_{\min}$ and $\xi_{\max}$. As an additional open direction, it is worth exploring whether there is an operational interpretation of the scaled trace distance $D'(\cdot, \cdot)$ that we introduced in section 3.2. Finally, it is an open question to determine if the one-shot approximate SD-cost can be evaluated by a SDP. We prove in appendix D that the two variants of approximate SD-cost (based on $CPTP_A$ and CDS maps) can be evaluated by means of bilinear programs, so that the methods of [HKT20] can be used to evaluate these quantities. However, it is not clear to us if these bilinear programs can be simplified further to SDPs.

Acknowledgments

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://arxiv.org/abs/2102.12512.

Appendix A. General properties of the axiomatic framework of the RTSD

In this appendix, we discuss a number of consequences of axioms I–V, and the preorder of SD, introduced in section 2. We also define general resource measures to quantify SD, and we provide several examples of such measures.

Lemma A.1. *The preorder of SD satisfies the following properties:*

(a) **Minimal elements.** *For all $\rho_{XA} \in \mathfrak{D}(XA)$ and $\pi_{X'} \otimes \omega_{A'} \in \mathfrak{D}(X'A')$ we have*

$$\pi_{X'} \otimes \omega_{A'} \prec \rho_{XA}. \tag{A.1}$$

(b) **Maximal elements.** *For all $\rho_{XA} \in \mathfrak{D}(XA)$, $\{p_x\}_{x=1}^{|X|}$ a probability distribution, and A' a quantum system with $|A'| \geqslant |X|$, we have*

$$\rho_{XA} \prec |1\rangle\langle 1|_X \sim \sum_{x=1}^{|X|} p_x |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{A'}. \tag{A.2}$$

(c) **Reduction to majorisation.** *For $\rho_X, \sigma_X \in \mathcal{D}(X)$ classical states of the same dimension, the preorder of SD is equivalent to the majorisation preorder.*

Proof. The first property follows immediately from axiom III and axiom V. For the second property, first observe that axiom III implies that $|1\rangle\langle 1|_X \sim |1\rangle\langle 1|_X \otimes \omega_A$ for some ω_A. From the form of CDS maps we get

that $|1\rangle\langle 1|_X \otimes \omega_A$ can be converted to any cq-state in $\mathfrak{D}(XA')$. Hence, since $\mathfrak{F}(XA \to XA') = \mathrm{CDS}$ $(XA \to XA')$ (see lemma 2.3) the assertion follows. For the third property, note that by axioms I, III, and V,

$$\rho_X \prec \sigma_X \tag{A.3}$$

with respect to the preorder of SD if and only if

$$\rho_X = \mathrm{Tr}_{X'} \circ \mathcal{P}_{XX'} (\sigma_X \otimes \pi_{X'}), \tag{A.4}$$

for some classical system X' and permutation channel $\mathcal{P}_{XX'}$ on the joint classical system XX'. The above channels are known as noisy operations [HHO03]. First, observe that noisy operations are doubly stochastic, so that if such a permutation channel $\mathcal{P}_{XX'}$ exists, then σ_X majorizes ρ_X.

Conversely, suppose $\rho = \sum_{y=1}^{m} t_y U_y \sigma U_y$ where $\{t_y\}_{y=1}^{m}$ are the components of a probability distribution, U_y are permutation matrices on system X, and for convenience of the exposition here we removed the subscript X. It is well known that such $\{t_y\}$ and $\{U_y\}$ exist iff $\rho \prec \sigma$. Suppose that $t_y = \frac{n_y}{n}$ are rational components where $n = \sum_{y=1}^{m} n_y$ is the common denominator and each $n_y \in \mathbb{N}$. Let X' be a classical system of dimension $|X'| = n$. Define a permutation matrix $P_{XX'}$ by its action on the basis elements

$$P_{XX'} \left(|x\rangle_X \otimes |x'\rangle_{X'}\right) = \left(U_{y_{x'}} |x\rangle_X\right) \otimes |x'\rangle_{X'} \quad \forall \, x \in [|X|] \,, \; x' \in [n],$$

where $y_{x'} \in [m]$ is the index satisfying

$$\sum_{z=1}^{y_{x'}-1} n_z \leqslant x' \leqslant \sum_{z=1}^{y_{x'}} n_z, \tag{A.5}$$

and we used the convention that the left-hand side of the above inequality is zero for $y_{x'} = 1$. With this definition we have

$$\mathrm{Tr}_{X'} \circ \mathcal{P}_{XX'} (\sigma_X \otimes \pi_{X'}) = \frac{1}{m} \sum_{x'=1}^{n} U_{y_{x'}} \sigma_X U_{y_{x'}}^{\dagger} = \frac{1}{m} \sum_{y=1}^{m} n_y U_y \sigma U_y^{\dagger} = \sum_{y=1}^{m} t_y U_y \sigma_X U_y = \rho_X.$$

Therefore, noisy operations can approximate any mixture of unitaries arbitrarily well. □

Note that the lemma above indicates that $|1\rangle\langle 1|_X$ is the maximal resource in the fixed dimension of X. If, for example, X' is another system with a higher dimension $|X'| > |X|$ then

$$|1\rangle\langle 1|_X \prec |1\rangle\langle 1|_{X'}. \tag{A.6}$$

That is, the 'embedding' of X into X' by adding zero components to matrices/vectors is not allowed in this resource theory since it can increase the value of the resource. To get the intuition behind it, consider Xiao possessing either one of the two classical states $|1\rangle\langle 1|_{X\tilde{X}}$ or $|1\rangle\langle 1|_X \otimes \pi_{\tilde{X}}$. In the first case, Xiao has complete information of the state in her possession since she knows the values of *both* the random variables X and $\tilde{X}$. On the other hand, in the second case Xiao has no information about $\tilde{X}$, since it is in a uniform state. Therefore, the first state is more distinguishable than the second one and we get

$$|1\rangle\langle 1|_X \sim |1\rangle\langle 1|_X \otimes \pi_{\tilde{X}} \prec |1\rangle\langle 1|_{X\tilde{X}}. \tag{A.7}$$

Hence, for the case $X' = X\tilde{X}$, (A.7) reduces (A.6). More generally, the SD of a cq-state ρ_{XA} represents the ability of Alice to distinguish the elements in Xiao's system. Therefore, the greater $|X|$ is, the more elements there are to distinguish, and consequently, the maximal resource has greater SD. This, in particular, applies to the minimum error probability $p_{\mathrm{err}}(\rho_{XA})$. That is, suppose for example that two states $\rho_{XA}, \sigma_{X'A'} \in \mathfrak{D}_{\mathrm{cq}}$, with $|X| < |X'|$ satisfy

$$p_{\mathrm{err}}(\rho_{XA}) = p_{\mathrm{err}}(\sigma_{X'A'}). \tag{A.8}$$

Then, we can expect that $\sigma_{X'A'}$ has more SD since Alice is able to distinguish among $|X'| > |X|$ elements with the same error as she would have if she held ρ_{XA}. This means in particular that if we consider inter-conversions among c–q states with different classical dimensions then the minimum error probability is not a good measure of SD. In the following subsection we show how the minimum error probability needs to be re-scaled with the classical dimension so that it becomes a proper measure of SD.

A.1. Quantification of SD
SD is quantified with functions that preserve the preorder of SD.

Definition A.2. A function $f : \mathfrak{D}_{\mathrm{cq}} \to \mathbb{R}$ is called a *measure of SD* if:

(a) For any $\rho_{XA}, \sigma_{X'A'} \in \mathfrak{D}_{cq}$ we have

$$\rho_{XA} \prec \sigma_{X'A'} \Rightarrow f(\rho_{XA}) \leqslant f(\sigma_{X'A'}). \tag{A.9}$$

(b) For the trivial state 1 (i.e. $|X| = |A| = 1$) we have

$$f(1) = 0. \tag{A.10}$$

Note that from axiom III and the second condition above, all measures of SD vanish on free states. Combining this with the property in (A.1) we conclude that measures of SD are non-negative.

Example A.3. Let

$$\mathbb{D} : \bigcup_A \mathcal{D}(A) \times \mathcal{D}(A) \to \mathbb{R} : (\rho, \sigma) \mapsto \mathbb{D}(\rho\|\sigma)$$

be a relative entropy; i.e. it satisfies the DPI, additivity (under tensor products), and the normalization $\mathbb{D}(1\|1) = 0$. Then, the function

$$\min_{\omega \in \mathcal{D}(A)} \mathbb{D}\left(\rho_{XA}\|\pi_X \otimes \omega_A\right) := \log|X| - \mathbb{H}(X|A)_\rho \tag{A.11}$$

is a measure of SD, since in any quantum resource theory, a function of the form $\min_{\omega \in \mathfrak{F}(A)} \mathbb{D}(\rho_A\|\omega_A)$ is a measure of a resource [CG19].

Example A.4 (Normalized guessing probability). A special example of the above family of measures of SD is obtained when setting $\mathbb{D}$ to be the max-relative entropy $D_{\max}$. Specifically, in [KRS09] it was shown that the guessing probability can be written as

$$p_{\text{guess}}(X|A)_\rho = 2^{\min_{\omega \in \mathcal{D}(A)} D_{\max}(\rho_{XA}\|I_X \otimes \omega_A)}.$$

Therefore, replacing I_X with the maximally mixed state π_X we get that the function

$$\log\left(|X| p_{\text{guess}}(X|A)_\rho\right) = \min_{\omega \in \mathcal{D}(A)} D_{\max}(\rho_{XA}\|\pi_X \otimes \omega_A) \tag{A.12}$$

is a measure of SD. Note that the dimension of the classical system is included on the left-side so that the expression remains invariant under replacement of ρ_{XA} with $\rho_{XA} \otimes \pi_{X'}$.

In this paper we have focused on the RTSD for the particular case in which the dimension of the classical system X is fixed to $|X| = 2$. In this case, as mentioned in the main text, it suffices to consider measures of SD that behave monotonically under CDS but not necessarily under conditional noisy operations (i.e. under free operations that change the dimension of X). The measure of SD that we have chosen in the paper is given by definition 2.5. As shown in theorem 4.5 it has the particularly pleasing feature of having an operational meaning in the context of SD distillation.

Appendix B. Proof of equation (3.16)—minimum trace distance to infinite-resource states under free operations

Let $\left(p\rho^0, (1-p)\rho^1\right)$ be a pair of subnormalized states, with $p \in (0,1)$, and ρ^0 and ρ^1 states. Then this pair is in one-to-one correspondence with the following c–q state:

$$\rho_{XB} := p|0\rangle\langle 0|_X \otimes \rho_B^0 + (1-p)|1\rangle\langle 1|_X \otimes \rho_B^1. \tag{B.1}$$

Let $\left(q\sigma^0, (1-q)\sigma^1\right)$ be another pair of subnormalized states, with $q \in (0,1)$, and σ^0 and σ^1 states. Then this pair is in one-to-one correspondence with the following c–q state:

$$\sigma_{XB'} := q|0\rangle\langle 0|_X \otimes \sigma_{B'}^0 + (1-q)|1\rangle\langle 1|_X \otimes \sigma_{B'}^1. \tag{B.2}$$

The trace-distance conversion error of ρ_{XB} to $\sigma_{XB'}$ is defined as follows:

$$\min_{\mathcal{N}_{XB \to XB'} \in \text{CDS}} \frac{1}{2}\|\mathcal{N}_{XB \to XB'}(\rho_{XB}) - \sigma_{XB'}\|_1, \tag{B.3}$$

where CDS is the set of conditional doubly stochastic (CDS) maps.

We first show that the trace-distance conversion error can be computed by means of a SDP.

Proposition B.1. *The trace-distance conversion error*

$$\min_{\mathcal{N}_{XB\rightarrow XB'}\in\text{CDS}} \frac{1}{2}\|\mathcal{N}_{XB\rightarrow X'B'}(\rho_{XB}) - \sigma_{XB'}\|_1 \tag{B.4}$$

between the initial pair $(p\rho_B^0, (1-p)\rho_B^1)$ and the target pair $(q\sigma_{B'}^0, (1-q)\sigma_{B'}^1)$ can be calculated by means of the following SDP:

$$\min_{Y_{B'}^0, Y_{B'}^1, \Gamma_{BB'}^{\mathcal{N}^0}, \Gamma_{BB'}^{\mathcal{N}^1} \geqslant 0} \text{Tr}[Y_{B'}^0 + Y_{B'}^1] \tag{B.5}$$

subject to

$$\text{Tr}_{B'}[\Gamma_{BB'}^{\mathcal{N}^0} + \Gamma_{BB'}^{\mathcal{N}^1}] = I_B, \tag{B.6}$$

$$Y_{B'}^0 \geqslant \text{Tr}_B[(p\rho_B^0)^{\text{T}}\Gamma_{BB'}^{\mathcal{N}^0} + ((1-p)\rho_B^1)^{\text{T}}\Gamma_{BB'}^{\mathcal{N}^1}] - q\sigma_{B'}^0, \tag{B.7}$$

$$Y_{B'}^1 \geqslant \text{Tr}_B[(p\rho_B^0)^{\text{T}}\Gamma_{BB'}^{\mathcal{N}^1} + ((1-p)\rho_B^1)^{\text{T}}\Gamma_{BB'}^{\mathcal{N}^0}] - (1-q)\sigma_{B'}^1. \tag{B.8}$$

The dual program is given by

$$\max_{Y_B \in \text{Herm}, W_{B'}, Z_{B'} \geqslant 0} \text{Tr}[Y_B] - q\,\text{Tr}[W_{B'}\sigma_{B'}^0] - (1-q)\,\text{Tr}[Z_{B'}\sigma_{B'}^1], \tag{B.9}$$

subject to

$$W_{B'}, Z_{B'} \leqslant I_{B'}, \tag{B.10}$$

$$Y_B \otimes I_{B'} \leqslant p\rho_B^0 \otimes W_{B'} + (1-p)\rho_B^1 \otimes Z_{B'}, \tag{B.11}$$

$$Y_B \otimes I_{B'} \leqslant (1-p)\rho_B^1 \otimes W_{B'} + p\rho_B^0 \otimes Z_{B'}. \tag{B.12}$$

Proof. Recall that an arbitrary CDS channel has the following form:

$$id_X \otimes \mathcal{N}_{B\rightarrow B'}^0 + \mathcal{P}_X \otimes \mathcal{N}_{B\rightarrow B'}^1, \tag{B.13}$$

where $\mathcal{N}_{B\rightarrow B'}^0$ and $\mathcal{N}_{B\rightarrow B'}^1$ are CP maps such that $\mathcal{N}_{B\rightarrow B'}^0 + \mathcal{N}_{B\rightarrow B'}^1$ is trace preserving, and $\mathcal{P}_X$ is a unitary channel that flips $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. This means that its action on an input

$$p|0\rangle\langle 0|_X \otimes \rho_B^0 + (1-p)|1\rangle\langle 1|_X \otimes \rho_B^1 \tag{B.14}$$

is as follows:

$$\begin{aligned}
&\left(id_X \otimes \mathcal{N}_{B\rightarrow B'}^0 + \mathcal{P}_X \otimes \mathcal{N}_{B\rightarrow B'}^1\right)\left(p|0\rangle\langle 0|_X \otimes \rho_B^0 + (1-p)|1\rangle\langle 1|_X \otimes \rho_B^1\right) \\
&= \left(id_X \otimes \mathcal{N}_{B\rightarrow B'}^0\right)\left(p|0\rangle\langle 0|_X \otimes \rho_B^0 + (1-p)|1\rangle\langle 1|_X \otimes \rho_B^1\right) \\
&\quad + \left(\mathcal{P}_X \otimes \mathcal{N}_{B\rightarrow B'}^1\right)\left(p|0\rangle\langle 0|_X \otimes \rho_B^0 + (1-p)|1\rangle\langle 1|_X \otimes \rho_B^1\right) \tag{B.15} \\
&= p|0\rangle\langle 0|_X \otimes \mathcal{N}_{B\rightarrow B'}^0(\rho_B^0) + (1-p)|1\rangle\langle 1|_X \otimes \mathcal{N}_{B\rightarrow B'}^0(\rho_B^1) \\
&\quad + p\mathcal{P}_X(|0\rangle\langle 0|_X) \otimes \mathcal{N}_{B\rightarrow B'}^1(\rho_B^0) + (1-p)\mathcal{P}_X(|1\rangle\langle 1|_X) \otimes \mathcal{N}_{B\rightarrow B'}^1(\rho_B^1) \tag{B.16} \\
&= p|0\rangle\langle 0|_X \otimes \mathcal{N}_{B\rightarrow B'}^0(\rho_B^0) + (1-p)|1\rangle\langle 1|_X \otimes \mathcal{N}_{B\rightarrow B'}^0(\rho_B^1) \\
&\quad + p|1\rangle\langle 1|_X \otimes \mathcal{N}_{B\rightarrow B'}^1(\rho_B^0) + (1-p)|0\rangle\langle 0|_X \otimes \mathcal{N}_{B\rightarrow B'}^1(\rho_B^1) \tag{B.17} \\
&= |0\rangle\langle 0|_X \otimes \left[\mathcal{N}_{B\rightarrow B'}^0(p\rho_B^0) + \mathcal{N}_{B\rightarrow B'}^1((1-p)\rho_B^1)\right] \\
&\quad + |1\rangle\langle 1|_X \otimes \left[\mathcal{N}_{B\rightarrow B'}^1(p\rho_B^0) + \mathcal{N}_{B\rightarrow B'}^0((1-p)\rho_B^1)\right]. \tag{B.18}
\end{aligned}$$

The semi-definite specifications for the CP maps $\mathcal{N}_{B\rightarrow B'}^0$ and $\mathcal{N}_{B\rightarrow B'}^1$ are as follows:

$$\Gamma_{BB'}^{\mathcal{N}^0}, \Gamma_{BB'}^{\mathcal{N}^1} \geqslant 0, \tag{B.19}$$

$$\text{Tr}_{B'}[\Gamma_{BB'}^{\mathcal{N}^0} + \Gamma_{BB'}^{\mathcal{N}^1}] = I_B. \tag{B.20}$$

Furthermore, the output state is given by

$$|0\rangle\langle 0|_X \otimes \left[\mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] \right]$$

$$+ |1\rangle\langle 1|_X \otimes \left[\mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] \right]. \tag{B.21}$$

Recall that the dual SDP for computing the normalized trace distance of two quantum states ρ and σ is as follows (see, e.g. [WW19a]):

$$\frac{1}{2}\|\rho - \sigma\|_1 = \min_{Y \geqslant 0} \left\{ \mathrm{Tr}[Y] : Y \geqslant \rho - \sigma \right\}. \tag{B.22}$$

So in this case, it follows that

$$\min_{\mathcal{N}_{XB\to XB'}\in\mathrm{CDS}} \frac{1}{2}\|\mathcal{N}_{XB\to XB'}(\rho_{XB}) - \sigma_{XB'}\|_1$$

$$= \min_{Y_{XB'} \geqslant 0} \left\{ \mathrm{Tr}[Y_{XB'}] : Y_{XB'} \geqslant \mathcal{N}_{XB\to XB'}(\rho_{XB}) - \sigma_{XB'} \right\}, \tag{B.23}$$

where we have now called the output system X for clarity. Then the SDP for the trace-distance conversion error is given by

$$\min_{Y_{XB'} \geqslant 0} \mathrm{Tr}[Y_{XB'}] \tag{B.24}$$

subject to

$$\Gamma_{BB'}^{\mathcal{N}^0}, \Gamma_{BB'}^{\mathcal{N}^1} \geqslant 0, \tag{B.25}$$

$$\mathrm{Tr}_{B'}[\Gamma_{BB'}^{\mathcal{N}^0} + \Gamma_{BB'}^{\mathcal{N}^1}] = I_B, \tag{B.26}$$

$$Y_{XB'} \geqslant |0\rangle\langle 0|_X \otimes \left[\mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] - q\sigma_{B'}^0 \right]$$

$$+ |1\rangle\langle 1|_X \otimes \left[\mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] - (1-q)\,\sigma_{B'}^1 \right]. \tag{B.27}$$

It is clear that the optimal $Y_{XB'}$ respects the c–q structure. So this means that the final SDP can be written as follows:

$$\min_{Y_{B'}^0, Y_{B'}^1, \Gamma_{BB'}^{\mathcal{N}^0}, \Gamma_{BB'}^{\mathcal{N}^1} \geqslant 0} \mathrm{Tr}[Y_{B'}^0 + Y_{B'}^1] \tag{B.28}$$

subject to

$$\mathrm{Tr}_{B'}[\Gamma_{BB'}^{\mathcal{N}^0} + \Gamma_{BB'}^{\mathcal{N}^1}] = I_B, \tag{B.29}$$

$$Y_{B'}^0 \geqslant \mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] - q\sigma_{B'}^0, \tag{B.30}$$

$$Y_{B'}^1 \geqslant \mathrm{Tr}_B[(p\rho_B^0)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^1}] + \mathrm{Tr}_B[((1-p)\,\rho_B^1)^\mathrm{T}\Gamma_{BB'}^{\mathcal{N}^0}] - (1-q)\,\sigma_{B'}^1. \tag{B.31}$$

Now we compute the dual of the SDP above. Recall the standard form of primal and dual SDPs:

$$\max_{X \geqslant 0} \left\{ \mathrm{Tr}[AX] : \Phi(X) \leqslant B \right\}, \tag{B.32}$$

$$\min_{Y \geqslant 0} \left\{ \mathrm{Tr}[BY] : \Phi^\dagger(Y) \geqslant A \right\}. \tag{B.33}$$

From inspecting the above, we see that

$$Y = \begin{bmatrix} Y_{B'}^0 & 0 & 0 & 0 \\ 0 & Y_{B'}^1 & 0 & 0 \\ 0 & 0 & \Gamma_{BB'}^{\mathcal{N}^0} & 0 \\ 0 & 0 & 0 & \Gamma_{BB'}^{\mathcal{N}^1} \end{bmatrix}, \qquad B = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{B.34}$$

$$\Phi^\dagger(Y) = \begin{bmatrix} \mathrm{Tr}_{B'}[\Gamma^{\mathcal{N}^0}_{BB'} + \Gamma^{\mathcal{N}^1}_{BB'}] & 0 & 0 & 0 \\ 0 & -\mathrm{Tr}_{B'}[\Gamma^{\mathcal{N}^0}_{BB'} + \Gamma^{\mathcal{N}^1}_{BB'}] & 0 & 0 \\ 0 & 0 & Y^0_{B'} - Z^0_{B'} & 0 \\ 0 & 0 & 0 & Y^1_{B'} - Z^1_{B'} \end{bmatrix}, \tag{B.35}$$

$$A = \begin{bmatrix} I_B & 0 & 0 & 0 \\ 0 & -I_B & 0 & 0 \\ 0 & 0 & -q\sigma^0_{B'} & 0 \\ 0 & 0 & 0 & -(1-q)\,\sigma^1_{B'} \end{bmatrix}. \tag{B.36}$$

where

$$Z^0_{B'} := \mathrm{Tr}_B[(p\rho^0_B)^\mathrm{T}\Gamma^{\mathcal{N}^0}_{BB'} + ((1-p)\,\rho^1_B)^\mathrm{T}\Gamma^{\mathcal{N}^1}_{BB'}], \tag{B.37}$$

$$Z^1_{B'} := \mathrm{Tr}_B[(p\rho^0_B)^\mathrm{T}\Gamma^{\mathcal{N}^1}_{BB'} + ((1-p)\,\rho^1_B)^\mathrm{T}\Gamma^{\mathcal{N}^0}_{BB'}]. \tag{B.38}$$

Now we need to find the adjoint map of $\Phi^\dagger$. Consider that

$$\mathrm{Tr}[\Phi^\dagger(Y)X] = \mathrm{Tr}[\mathrm{Tr}_{B'}[\Gamma^{\mathcal{N}^0}_{BB'} + \Gamma^{\mathcal{N}^1}_{BB'}](X^1_B - X^2_B)]$$

$$+ \mathrm{Tr}[(Y^0_{B'} - \mathrm{Tr}_B[(p\rho^0_B)^\mathrm{T}\Gamma^{\mathcal{N}^0}_{BB'} + ((1-p)\,\rho^1_B)^\mathrm{T}\Gamma^{\mathcal{N}^1}_{BB'}])X^3_{B'}]$$

$$+ \mathrm{Tr}[(Y^1_{B'} - \mathrm{Tr}_B[(p\rho^0_B)^\mathrm{T}\Gamma^{\mathcal{N}^1}_{BB'} + ((1-p)\,\rho^1_B)^\mathrm{T}\Gamma^{\mathcal{N}^0}_{BB'}])X^4_{B'}] \tag{B.39}$$

$$= \mathrm{Tr}[(\Gamma^{\mathcal{N}^0}_{BB'} + \Gamma^{\mathcal{N}^1}_{BB'})((X^1_B - X^2_B) \otimes I_{B'})] + \mathrm{Tr}[Y^0_{B'}X^3_{B'}] - \mathrm{Tr}[\Gamma^{\mathcal{N}^0}_{BB'}((p\rho^0_B)^\mathrm{T} \otimes X^3_{B'})]$$

$$- \mathrm{Tr}[(\Gamma^{\mathcal{N}^1}_{BB'})((1-p)\,\rho^1_B)^\mathrm{T} \otimes X^3_{B'})] + \mathrm{Tr}[Y^1_{B'}X^4_{B'}] - \mathrm{Tr}[\Gamma^{\mathcal{N}^1}_{BB'}((p\rho^0_B)^\mathrm{T} \otimes X^4_{B'})]$$

$$- \mathrm{Tr}[\Gamma^{\mathcal{N}^0}_{BB'}(((1-p)\,\rho^1_B)^\mathrm{T} \otimes X^4_{B'})] \tag{B.40}$$

$$= \mathrm{Tr}[Y^0_{B'}X^3_{B'}] + \mathrm{Tr}[Y^1_{B'}X^4_{B'}]$$

$$+ \mathrm{Tr}[\Gamma^{\mathcal{N}^0}_{BB'}((X^1_B - X^2_B) \otimes I_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^3_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^4_{B'})]$$

$$+ \mathrm{Tr}[\Gamma^{\mathcal{N}^1}_{BB'}((X^1_B - X^2_B) \otimes I_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^3_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^4_{B'})]. \tag{B.41}$$

So this means that

$$\Phi(X) = \begin{bmatrix} X^3_{B'} & 0 & 0 & 0 \\ 0 & X^4_{B'} & 0 & 0 \\ 0 & 0 & W^0_B & 0 \\ 0 & 0 & 0 & W^1_B \end{bmatrix}, \tag{B.42}$$

$$W^0_B := (X^1_B - X^2_B) \otimes I_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^3_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^4_{B'}, \tag{B.43}$$

$$W^1_B := (X^1_B - X^2_B) \otimes I_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^3_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^4_{B'}. \tag{B.44}$$

Then the dual program is given by

$$\max_{X^1_B, X^2_B, X^3_{B'}, X^4_{B'} \geqslant 0} \mathrm{Tr}\left[\begin{bmatrix} I_B & 0 & 0 & 0 \\ 0 & -I_B & 0 & 0 \\ 0 & 0 & -q\sigma^0_{B'} & 0 \\ 0 & 0 & 0 & -(1-q)\,\sigma^1_{B'} \end{bmatrix}\begin{bmatrix} X^1_B & 0 & 0 & 0 \\ 0 & X^2_B & 0 & 0 \\ 0 & 0 & X^3_{B'} & 0 \\ 0 & 0 & 0 & X^4_{B'} \end{bmatrix}\right] \tag{B.45}$$

subject to

$$X^3_{B'} \leqslant I_{B'}, \qquad X^4_{B'} \leqslant I_{B'}, \tag{B.46}$$

$$(X^1_B - X^2_B) \otimes I_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^3_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^4_{B'} \leqslant 0, \tag{B.47}$$

$$(X^1_B - X^2_B) \otimes I_{B'} - (1-p)\,(\rho^1_B)^\mathrm{T} \otimes X^3_{B'} - p(\rho^0_B)^\mathrm{T} \otimes X^4_{B'} \leqslant 0. \tag{B.48}$$

This can be simplified to the following:

$$\max_{Y_B \in \mathrm{Herm}, X^3_{B'}, X^4_{B'} \geqslant 0} \mathrm{Tr}[Y_B] - q\,\mathrm{Tr}[X^3_{B'}\sigma^0_{B'}] - (1-q)\,\mathrm{Tr}[X^4_{B'}\sigma^1_{B'}], \tag{B.49}$$

subject to

$$X_{B'}^3 \leqslant I_{B'}, \qquad X_{B'}^4 \leqslant I_{B'}, \tag{B.50}$$

$$Y_B \otimes I_{B'} \leqslant p(\rho_B^0)^{\mathrm{T}} \otimes X_{B'}^3 + (1-p)(\rho_B^1)^{\mathrm{T}} \otimes X_{B'}^4, \tag{B.51}$$

$$Y_B \otimes I_{B'} \leqslant (1-p)(\rho_B^1)^{\mathrm{T}} \otimes X_{B'}^3 + p(\rho_B^0)^{\mathrm{T}} \otimes X_{B'}^4. \tag{B.52}$$

Then we can set $X_{B'}^3 = W_{B'}$ and $X_{B'}^4 = Z_{B'}$ to get

$$\max_{Y_B \in \mathrm{Herm}, W_{B'}, Z_{B'} \geqslant 0} \mathrm{Tr}[Y_B] - q\,\mathrm{Tr}[W_{B'}\sigma_{B'}^0] - (1-q)\,\mathrm{Tr}[Z_{B'}\sigma_{B'}^1], \tag{B.53}$$

subject to

$$W_{B'}, Z_{B'} \leqslant I_{B'}, \tag{B.54}$$

$$Y_B \otimes I_{B'} \leqslant p(\rho_B^0)^{\mathrm{T}} \otimes W_{B'} + (1-p)(\rho_B^1)^{\mathrm{T}} \otimes Z_{B'}, \tag{B.55}$$

$$Y_B \otimes I_{B'} \leqslant (1-p)(\rho_B^1)^{\mathrm{T}} \otimes W_{B'} + p(\rho_B^0)^{\mathrm{T}} \otimes Z_{B'}. \tag{B.56}$$

We can finally make the substitution $Y_B \to Y_B^{\mathrm{T}}$ and the optimal value is unchanged. Since the operators on the right-hand side of the inequalities just above are separable, the partial transpose has no effect and can be removed. This concludes the proof. $\qquad\square$

B.1. Minimum error probability and minimum conversion error (in terms of trace distance) to infinite-resource states

As mentioned in the main text, the minimum error probability is given by

$$p_{\mathrm{err}}(\rho_{XB}) = \frac{1}{2}\left(1 - \left\|p\rho_B^0 - (1-p)\rho_B^1\right\|_1\right). \tag{B.57}$$

An alternative expression for it is given by

$$p_{\mathrm{err}}(\rho_{XB}) = \max_{Y_B \in \mathrm{Herm}} \left\{ \mathrm{Tr}[Y_B] : Y_B \leqslant p\rho_B^0, Y_B \leqslant (1-p)\rho_B^1 \right\}, \tag{B.58}$$

where *Herm* denotes the set of Hermitian operators acting on the system *B*. Note that the maximising operator Y_B on the right-hand side of the above equation is called the 'greatest lower bound (GLB) operator' of the operators $p\rho_B^0$ and $(1-p)\rho_B^1$. The GLB operator is defined in equation (84) of [AM14], and the above result was established as lemma A.7 of the same paper.

Consider the infinite-resource state

$$\gamma_{XQ}^{(\infty,q)} := q|0\rangle\langle0|_X \otimes |0\rangle\langle0|_Y + (1-q)|1\rangle\langle1|_X \otimes |1\rangle\langle1|_Y, \tag{B.59}$$

which is the (M,q)-golden unit (definition 3.4) with $M = \infty$. It follows from lemma 3.3 that for all $q, q' \in [0,1]$, it is possible to perform the transformation $\mathcal{N}_{XQ}(\gamma_{XQ}^{(\infty,q)}) = \gamma_{XQ}^{(\infty,q')}$, where $\mathcal{N}_{XQ}$ is a CDS map.

Lemma B.2. *Let $\rho_{XB} \equiv (p, \rho^0, \rho^1)$ be a c–q state. Then the following equality holds for all $q, q' \in [0,1]$:*

$$\min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|\mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q)}\right\|_1 = \min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|\mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q')}\right\|_1. \tag{B.60}$$

Proof. We first establish the inequality $\geqslant$. Let $\mathcal{N}'_{XQ}$ be the CDS channel that converts $\gamma_{XQ}^{(\infty,q)}$ to $\gamma_{XQ}^{(\infty,q')}$. Then consider that

$$\min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|\mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q')}\right\|_1$$

$$= \min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|\mathcal{N}_{XB \to XQ}(\rho_{XB}) - \mathcal{N}'_{XQ}(\gamma_{XQ}^{(\infty,q)})\right\| \tag{B.61}$$

$$\leqslant \min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|(\mathcal{N}'_{XQ} \circ \mathcal{N}_{XB \to XQ})(\rho_{XB}) - \mathcal{N}'_{XQ}(\gamma_{XQ}^{(\infty,q)})\right\|_1 \tag{B.62}$$

$$\leqslant \min_{\mathcal{N}_{XB \to XQ} \in \mathrm{CDS}} \frac{1}{2}\left\|\mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q)}\right\|_1. \tag{B.63}$$

The first inequality follows because $\mathcal{N}'_{XQ} \circ \mathcal{N}_{XB \to XQ}$ is a member of the set of CDS channels. The second inequality follows from the DPI under the trace distance. We can then apply the same argument to arrive at the opposite inequality. □

Lemma B.3. *Let $\rho_{XB} \equiv (p, \rho^0, \rho^1)$ be a c–q state. Then the following equality holds*

$$\min_{\substack{q \in [0,1], \\ \mathcal{N}_{XB \to XQ} \in \text{CDS}}} \frac{1}{2} \left\| \mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q)} \right\|_1 = p_{\text{err}}(\rho_{XB}). \tag{B.64}$$

Proof. We first establish the inequality

$$\min_{\substack{q \in [0,1], \\ \mathcal{N}_{XB \to XQ} \in \text{CDS}}} \frac{1}{2} \left\| \mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q)} \right\|_1 \leqslant p_{\text{err}}(\rho_{XB}), \tag{B.65}$$

by demonstrating the existence of a value of $q \in [0,1]$ and a CDS channel for which the left-hand side is equal to $p_{\text{err}}(\rho_{XB})$. From lemma B.2, it follows that the left-hand side of (B.64) is independent of $q \in [0,1]$. So we can pick $q = p$, and the value is unchanged. Now consider that the channel used in state discrimination is a simple local channel of the following form:

$$\mathcal{M}_{B \to Q}(\sigma_B) := \text{Tr}[\Lambda_B \sigma_B]|0\rangle\langle 0|_Q + \text{Tr}[(I_B - \Lambda_B)\sigma_B]|1\rangle\langle 1|_Q, \tag{B.66}$$

and so $id_X \otimes \mathcal{M}_{B \to Q}$ is a CDS channel. Acting with it on ρ_{XB} leads to the following state:

$$\begin{aligned}
\mathcal{M}_{B \to Q}(\rho_{XB}) &= p|0\rangle\langle 0|_X \otimes \left(\text{Tr}[\Lambda_B \rho_B^0]|0\rangle\langle 0|_Q + \text{Tr}[(I_B - \Lambda_B)\rho_B^0]|1\rangle\langle 1|_Q \right) \\
&\quad + (1-p)\,|1\rangle\langle 1|_X \otimes \left(\text{Tr}[\Lambda_B \rho_B^1]|0\rangle\langle 0|_Q + \text{Tr}[(I_B - \Lambda_B)\rho_B^1]|1\rangle\langle 1|_Q \right) \tag{B.67} \\
&= p\,\text{Tr}[\Lambda_B \rho_B^0]|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q + p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0]|0\rangle\langle 0|_X \otimes |1\rangle\langle 1|_Q \\
&\quad + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1]|1\rangle\langle 1|_X \otimes |0\rangle\langle 0|_Q + (1-p)\,\text{Tr}[(I_B - \Lambda_B)\rho_B^1]|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q. \tag{B.68}
\end{aligned}$$

Since this is a particular choice, it follows that

$$\min_{\substack{q \in [0,1], \\ \mathcal{N}_{XB \to XQ} \in \text{CDS}}} \frac{1}{2} \left\| \mathcal{N}_{XB \to XQ}(\rho_{XB}) - \gamma_{XQ}^{(\infty,q)} \right\|_1 \leqslant \frac{1}{2} \left\| \mathcal{M}_{B \to Q}(\rho_{XB}) - \gamma_{XQ}^{(\infty,p)} \right\|_1. \tag{B.69}$$

Now let us compute the trace distance between $\mathcal{M}_{B \to Q}(\rho_{XB})$ and the simple state $\gamma_{XQ}^{(\infty,p)}$:

$$\begin{aligned}
& \left\| \mathcal{M}_{B \to Q}(\rho_{XB}) - \gamma_{XQ}^{(\infty,p)} \right\|_1 \\
&= \left\| \begin{matrix} p\,\text{Tr}[\Lambda_B \rho_B^0]|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q + p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0]|0\rangle\langle 0|_X \otimes |1\rangle\langle 1|_Q \\ + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1]|1\rangle\langle 1|_X \otimes |0\rangle\langle 0|_Q + (1-p)\,\text{Tr}[(I_B - \Lambda_B)\rho_B^1]|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q \\ -p|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q - (1-p)\,|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q \end{matrix} \right\|_1 \tag{B.70} \\
&= \left\| \begin{matrix} (p\,\text{Tr}[\Lambda_B \rho_B^0] - p)\,|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q + p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0]|0\rangle\langle 0|_X \otimes |1\rangle\langle 1|_Q \\ + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1]|1\rangle\langle 1|_X \otimes |0\rangle\langle 0|_Q \\ + \left[(1-p)\,\text{Tr}[(I_B - \Lambda_B)\rho_B^1] - (1-p) \right] |1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q \end{matrix} \right\|_1 \tag{B.71} \\
&= \left| p\,\text{Tr}[\Lambda_B \rho_B^0] - p \right| + p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0] \\
&\quad + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1] + \left| (1-p)\,\text{Tr}[(I_B - \Lambda_B)\rho_B^1] - (1-p) \right| \tag{B.72} \\
&= \left| p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0] \right| + p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0] \\
&\quad + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1] + \left| (1-p)\,\text{Tr}[\Lambda_B \rho_B^1] \right| \tag{B.73} \\
&= 2\left(p\,\text{Tr}[(I_B - \Lambda_B)\rho_B^0] + (1-p)\,\text{Tr}[\Lambda_B \rho_B^1] \right) \tag{B.74} \\
&= 2p_{\text{err}}(\rho_{XB}). \tag{B.75}
\end{aligned}$$

We now establish the opposite inequality. Since the value of q does not matter, let us set it to $1/2$, so that

$$\omega_{XY}(q) = \frac{1}{2}|0\rangle\langle 0|_X \otimes |0\rangle\langle 0|_Q + \frac{1}{2}|1\rangle\langle 1|_X \otimes |1\rangle\langle 1|_Q. \tag{B.76}$$

By applying proposition B.1 and weak duality of semi-definite programming, we conclude that the trace-distance conversion error

$$\min_{\substack{q\in[0,1],\\ \mathcal{N}_{XB\to XQ}\in\mathrm{CDS}}} \frac{1}{2}\|\mathcal{N}_{XB\to XQ}(\rho_{XB}) - \omega_{XQ}(q)\|_1 \tag{B.77}$$

is not smaller than the optimal value of the following SDP:

$$\max_{Y_B\in\mathrm{Herm},W_{B'},Z_{B'}\geqslant 0} \mathrm{Tr}[Y_B] - \frac{1}{2}\left(\mathrm{Tr}[W_{B'}|0\rangle\langle 0|] + \mathrm{Tr}[Z_{B'}|1\rangle\langle 1|]\right), \tag{B.78}$$

subject to

$$W_{B'}, Z_{B'} \leqslant I_{B'}, \tag{B.79}$$

$$Y_B \otimes I_{B'} \leqslant p\rho_B^0 \otimes W_{B'} + (1-p)\,\rho_B^1 \otimes Z_{B'}, \tag{B.80}$$

$$Y_B \otimes I_{B'} \leqslant (1-p)\,\rho_B^1 \otimes W_{B'} + p\rho_B^0 \otimes Z_{B'}. \tag{B.81}$$

Let us then pick $W_{B'} = |1\rangle\langle 1|_{B'}$ and $Z_{B'} = |0\rangle\langle 0|_{B'}$. Then the SDP simplifies as follows:

$$\max_{Y_B\in\mathrm{Herm}} \mathrm{Tr}[Y_B], \tag{B.82}$$

subject to

$$Y_B \otimes I_{B'} \leqslant p\rho_B^0 \otimes |1\rangle\langle 1|_{B'} + (1-p)\,\rho_B^1 \otimes |0\rangle\langle 0|_{B'}, \tag{B.83}$$

$$Y_B \otimes I_{B'} \leqslant (1-p)\,\rho_B^1 \otimes |1\rangle\langle 1|_{B'} + p\rho_B^0 \otimes |0\rangle\langle 0|_{B'}. \tag{B.84}$$

Since we can write

$$Y_B \otimes I_{B'} = Y_B \otimes |0\rangle\langle 0|_{B'} + Y_B \otimes |1\rangle\langle 1|_{B'}, \tag{B.85}$$

the above constraints are equivalent to the following:

$$Y_B \leqslant p\rho_B^0, \qquad Y_B \leqslant (1-p)\,\rho_B^1. \tag{B.86}$$

This SDP is thus equal to the following one:

$$\max_{Y_B\in\mathrm{Herm}} \left\{\mathrm{Tr}[Y_B] : Y_B \leqslant p\rho_B^0, Y_B \leqslant (1-p)\,\rho_B^1\right\}. \tag{B.87}$$

This quantity is precisely the trace of the GLB operator, and so we conclude by applying (B.58).   □

Lemma B.4. *Let* $\rho_{XB} \equiv \left(p, \rho_B^0, \rho_B^1\right)$ *be a c–q state. Then the following equality holds*

$$\min_{\mathcal{N}_{B\to Q}\in\mathrm{CPTP}_A} \frac{1}{2}\left\|\mathcal{N}_{B\to Q}(\rho_{XB}) - \gamma_{XQ}^{(\infty,p)}\right\|_1 = p_{\mathrm{err}}(\rho_{XB}). \tag{B.88}$$

Proof. The inequality $\leqslant$ follows by the same reasoning given at the beginning of the proof of the previous theorem. The opposite inequality follows because we can apply a completely dephasing channel to the Q system and the state $\gamma_{XQ}^{(\infty,p)}$ remains invariant, while the channel $\mathcal{N}$ is transformed to a measurement channel. The trace distance does not increase under such a channel and evaluating it leads to an expression for the error probability under a particular measurement.   □

Appendix C. Derivation of the SDPs for scaled trace distance D' and minimum conversion error in propositions 3.15 and 3.16

Proof of proposition 3.15. We begin by rewriting the scaled trace distance $D'(\rho_{XA}, \sigma_{XA})$ as follows:

$$D'(\rho_{XA}, \sigma_{XA}) = \frac{\frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1}{\frac{1}{2}\left(1 - \|q\sigma_0 - (1-q)\sigma_1\|_1\right)} \tag{C.1}$$

$$= \frac{\|\rho_{XA} - \sigma_{XA}\|_1}{1 - \|q\sigma_0 - (1-q)\sigma_1\|_1} \tag{C.2}$$

$$= \frac{\max_{-I_{XA} \leqslant L_{XA} \leqslant I_{XA}} \operatorname{Tr} L_{XA}(\rho_{XA} - \sigma_{XA})}{1 - \max_{-I_A \leqslant P_A \leqslant I_A} \operatorname{Tr} P_A(q\sigma_0 - (1-q)\sigma_1)} \tag{C.3}$$

$$= \max_{-I_{XA} \leqslant L_{XA} \leqslant I_{XA}, -I_A \leqslant P_A \leqslant I_A} \frac{\operatorname{Tr} L_{XA}(\rho_{XA} - \sigma_{XA})}{1 - \operatorname{Tr} P_A(q\sigma_0 - (1-q)\sigma_1)}. \tag{C.4}$$

Furthermore, let us introduce $t = \frac{\operatorname{Tr} L_{XA}(\rho_{XA} - \sigma_{XA})}{1 - \operatorname{Tr} P_A(q\sigma_0 - (1-q)\sigma_1)}$ and obtain

$$D'(\rho_{XA}, \sigma_{XA}) = \{\max t : t - t \operatorname{Tr} P_A(q\sigma_0 - (1-q)\sigma_1) = \operatorname{Tr} L_{XA}(\rho_{XA} - \sigma_{XA}),$$
$$-I_{XA} \leqslant L_{XA} \leqslant I_{XA}, -I_A \leqslant P_A \leqslant I_A\}. \tag{C.5}$$

The constraints in the optimization above still have bilinear conditions. However, we can absorb tP_A into a single variable and obtain the simplified SDP in (3.28).

We now continue with the derivation of the dual SDP stated in (3.29). The standard form of primal and dual SDPs is as follows [Wat18]:

$$\sup_{X \geqslant 0} \{\operatorname{Tr}[AX] : \Phi(X) \leqslant B\}, \tag{C.6}$$

$$\inf_{Y \geqslant 0} \{\operatorname{Tr}[BY] : \Phi^\dagger(Y) \geqslant A\}. \tag{C.7}$$

The SDP for the scaled trace distance can be written as

$$\sup_{t, L_{XA}^0, L_{XA}^1, P_A^0, P_A^1 \geqslant 0} \left\{ \begin{array}{c} t : -I_{XA} \leqslant L_{XA}^1 - L_{XA}^0 \leqslant I_{XA}, \\ -tI_A \leqslant P_A^1 - P_A^0 \leqslant tI_A, \\ t - \operatorname{Tr}[(P_A^1 - P_A^0)(q\sigma_0 - (1-q)\sigma_1)] = \operatorname{Tr}[(L_{XA}^1 - L_{XA}^0)(\rho_{XA} - \sigma_{XA})] \end{array} \right\}. \tag{C.8}$$

In standard form, this SDP is as follows:

$$X = \operatorname{diag}(t, L_{XA}^0, L_{XA}^1, P_A^0, P_A^1), \tag{C.9}$$

$$A = \operatorname{diag}(1, 0, 0, 0, 0), \tag{C.10}$$

$$\Phi(X) = \operatorname{diag}\left(L_{XA}^1 - L_{XA}^0, L_{XA}^0 - L_{XA}^1, P_A^1 - P_A^0 - tI_A, P_A^0 - P_A^1 - tI_A,\right.$$
$$\operatorname{Tr}[(L_{XA}^1 - L_{XA}^0)(\rho_{XA} - \sigma_{XA})] - (t - \operatorname{Tr}[(P_A^1 - P_A^0)(q\sigma_0 - (1-q)\sigma_1)]),$$
$$\left. t - \operatorname{Tr}[(P_A^1 - P_A^0)(q\sigma_0 - (1-q)\sigma_1)] - \operatorname{Tr}[(L_{XA}^1 - L_{XA}^0)(\rho_{XA} - \sigma_{XA})]\right), \tag{C.11}$$

$$B = \operatorname{diag}(I_{XA}, I_{XA}, 0, 0, 0, 0). \tag{C.12}$$

So we need to derive the adjoint map $\Phi^\dagger$, satisfying $\operatorname{Tr}[Y\Phi(X)] = \operatorname{Tr}[\Phi^\dagger(Y)X]$. Consider that the dual variables are given as

$$Y = \operatorname{diag}(B_{XA}, C_{XA}, D_A, E_A, s_1, s_2). \tag{C.13}$$

Then we find that

$$\operatorname{Tr}[Y\Phi(X)] = \operatorname{Tr}[B_{XA}(L_{XA}^1 - L_{XA}^0)] + \operatorname{Tr}[C_{XA}(L_{XA}^0 - L_{XA}^1)]$$
$$+ \operatorname{Tr}[D_A(P_A^1 - P_A^0 - tI_A)] + \operatorname{Tr}[E_A(P_A^0 - P_A^1 - tI_A)]$$
$$+ (s_1 - s_2)\left(\operatorname{Tr}[(L_{XA}^1 - L_{XA}^0)(\rho_{XA} - \sigma_{XA})] - (t - \operatorname{Tr}[(P_A^1 - P_A^0)(q\sigma_0 - (1-q)\sigma_1)])\right) \tag{C.14}$$

$$= \operatorname{Tr}[(B_{XA} - C_{XA} + (s_1 - s_2)(\rho_{XA} - \sigma_{XA}))L_{XA}^1]$$
$$- \operatorname{Tr}[(B_{XA} - C_{XA} + (s_1 - s_2)(\rho_{XA} - \sigma_{XA}))L_{XA}^0]$$
$$+ \operatorname{Tr}[(D_A - E_A + (s_1 - s_2)(q\sigma_0 - (1-q)\sigma_1))P_A^1]$$
$$- \operatorname{Tr}[(D_A - E_A + (s_1 - s_2)(q\sigma_0 - (1-q)\sigma_1))P_A^0]$$
$$- ((s_1 - s_2) + \operatorname{Tr}[D_A + E_A])t. \tag{C.15}$$

So we conclude that

$$
\begin{aligned}
\Phi^\dagger(Y) = \mathrm{diag}\,(&- ((s_1 - s_2) + \mathrm{Tr}[D_A + E_A]),\, -(B_{XA} - C_{XA} + (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA})),\\
&(B_{XA} - C_{XA} + (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA})),\\
&-\left(D_A - E_A + (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right)\right),\\
&\left(D_A - E_A + (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right)\right)).
\end{aligned}
\tag{C.16}
$$

So then $\Phi^\dagger(Y) \geqslant A$ is equivalent to the following constraints:

$$- ((s_1 - s_2) + \mathrm{Tr}[D_A + E_A]) \geqslant 1, \tag{C.17}$$

$$- (B_{XA} - C_{XA} + (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA})) \geqslant 0, \tag{C.18}$$

$$(B_{XA} - C_{XA} + (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA})) \geqslant 0, \tag{C.19}$$

$$- \left(D_A - E_A + (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right)\right) \geqslant 0, \tag{C.20}$$

$$\left(D_A - E_A + (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right)\right) \geqslant 0, \tag{C.21}$$

which is the same as

$$- (s_1 - s_2) - 1 \geqslant \mathrm{Tr}[D_A + E_A], \tag{C.22}$$

$$B_{XA} - C_{XA} = - (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA}), \tag{C.23}$$

$$D_A - E_A = - (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right). \tag{C.24}$$

Then the dual SDP is as follows:

$$
\inf_{B_{XA}, C_{XA}, D_A, E_A, s_1, s_2 \geqslant 0}
\left\{
\begin{aligned}
&\mathrm{Tr}[B_{XA} + C_{XA}]:\\
&- (s_1 - s_2) - 1 \geqslant \mathrm{Tr}[D_A + E_A],\\
&B_{XA} - C_{XA} = - (s_1 - s_2)\,(\rho_{XA} - \sigma_{XA}),\\
&D_A - E_A = - (s_1 - s_2)\left(q\sigma_0 - (1-q)\,\sigma_1\right)
\end{aligned}
\right\}.
\tag{C.25}
$$

Now setting $s = -(s_1 - s_2)$, this can be rewritten as

$$
\inf_{\substack{B_{XA}, C_{XA}, D_A, E_A \geqslant 0,\\ s \in \mathbb{R}}}
\left\{
\begin{aligned}
&\mathrm{Tr}[B_{XA} + C_{XA}]:\\
&s - 1 \geqslant \mathrm{Tr}[D_A + E_A],\\
&B_{XA} - C_{XA} = s\,(\rho_{XA} - \sigma_{XA}),\\
&D_A - E_A = s\left(q\sigma_0 - (1-q)\,\sigma_1\right)
\end{aligned}
\right\}.
\tag{C.26}
$$

We now prove that strong duality holds under the following conditions:

$$\|\rho_{XA} - \sigma_{XA}\|_1 > 0, \tag{C.27}$$

$$\left\|q\sigma_0 - (1-q)\,\sigma_1\right\|_1 < 1. \tag{C.28}$$

Then strong duality holds by picking the primal variables as

$$t = \frac{1}{2}\|\rho_{XA} - \sigma_{XA}\|_1, \tag{C.29}$$

$$L_{XA}^1 = \frac{1}{2}\Pi_{\rho_{XA} \geqslant \sigma_{XA}} + \frac{1}{2}I_{XA}, \tag{C.30}$$

$$L_{XA}^0 = \frac{1}{2}\Pi_{\rho_{XA} < \sigma_{XA}} + \frac{1}{2}I_{XA}, \tag{C.31}$$

$$P_A^1 = P_A^0 = I_A, \tag{C.32}$$

where $\Pi_{\rho_{XA} \geqslant \sigma_{XA}}$ is the projection onto the non-negative eigenspace of $\rho_{XA} \geqslant \sigma_{XA}$ and $\Pi_{\rho_{XA} < \sigma_{XA}}$ is the projection onto the strictly negative eigenspace. Then all of the following constraints are satisfied with strict inequality (except for the final equality):

$$t, L_{XA}^0, L_{XA}^1, P_A^0, P_A^1 \geqslant 0, \tag{C.33}$$

$$-I_{XA} \leqslant L_{XA}^1 - L_{XA}^0 \leqslant I_{XA}, \tag{C.34}$$

$$-tI_A \leqslant P_A^1 - P_A^0 \leqslant tI_A, \tag{C.35}$$

$$t - \text{Tr}[\left(P_A^1 - P_A^0\right)\left(q\sigma_0 - \left(1 - q\right)\sigma_1\right)] = \text{Tr}[\left(L_{XA}^1 - L_{XA}^0\right)\left(\rho_{XA} - \sigma_{XA}\right)]. \tag{C.36}$$

For the dual program, we pick

$$s = \frac{1}{1 - \left\|q\sigma_0 - \left(1 - q\right)\sigma_1\right\|_1}, \tag{C.37}$$

$$B_{XA} = I_{XA} + s\rho_{XA}, \tag{C.38}$$

$$C_{XA} = I_{XA} + s\sigma_{XA}, \tag{C.39}$$

$$D_A = sP, \tag{C.40}$$

$$E_A = sN, \tag{C.41}$$

where P is the positive part of $q\sigma_0 - \left(1 - q\right)\sigma_1$ and N is the negative part of $q\sigma_0 - \left(1 - q\right)\sigma_1$. Under these choices, we find that the constraints from the dual program are met, i.e. as follows:

$$B_{XA}, C_{XA}, D_A, E_A \geqslant 0, s \in \mathbb{R}, \tag{C.42}$$

$$s - 1 = \text{Tr}[D_A + E_A], \tag{C.43}$$

$$B_{XA} - C_{XA} = s\left(\rho_{XA} - \sigma_{XA}\right), \tag{C.44}$$

$$D_A - E_A = s\left(q\sigma_0 - \left(1 - q\right)\sigma_1\right). \tag{C.45}$$

Thus, strong duality holds under the conditions given in (C.27) and (C.28).

Finally, suppose now that $\left\|\rho_{XA} - \sigma_{XA}\right\|_1 = 0$ and $\left\|q\sigma_0 - \left(1 - q\right)\sigma_1\right\|_1 < 1$. Then the choices $t = 0$, $L_{XA} = 0$, and $P_A = 0$ are feasible for the primal and lead to a value of zero for the objective function. Also, setting B_{XA} to be the positive part of $s(\rho_{XA} - \sigma_{XA})$ and C_{XA} to be the negative part of $s(\rho_{XA} - \sigma_{XA})$, with the same choices for s, D_A, and E_A as given above, leads to feasible choices for the dual, for which the objective function also evaluates to zero. So strong duality holds in this case also. $\qquad\square$

Proof of proposition 3.16. Using (3.29), the scaled trace distance $D'(\cdot, \cdot)$ for states $\tau_{XA} := t|0\rangle\langle0| \otimes \tau_0 + (1 - t)|1\rangle\langle1| \otimes \tau_1$ and $\omega_{XA} := w|0\rangle\langle0| \otimes \omega_0 + (1 - w)|1\rangle\langle1| \otimes \omega_1$ can be written as the following SDP

$$D'(\tau_{XA}, \omega_{XA}) = \min_{B_{XA}, C_{XA}, D_A, E_A \geqslant 0, s \in \mathbb{R}} \left\{ \begin{array}{c} \text{Tr}[B_{XA} + C_{XA}]: \\ B_{XA} - C_{XA} = s\left(\tau_{XA} - \omega_{XA}\right), \\ D_A - E_A = s(w\omega_0 - (1 - w)\omega_1), \\ \text{Tr}[D_A + E_A] \leqslant s - 1 \end{array} \right\}. \tag{C.46}$$

Following the development in (B.13)–(B.21) and combining with (C.46), we conclude the following form for the optimization task

$$\min_{\substack{B_{XA'}, C_{XA'}, D_{A'}, E_{A'} \geqslant 0, \\ \Gamma_{AA'}^{\mathcal{N}^0}, \Gamma_{AA'}^{\mathcal{N}^1} \geqslant 0, s \in \mathbb{R}}} \left\{ \begin{array}{c} \text{Tr}[B_{XA'} + C_{XA'}]: \\ B_{XA'} - C_{XA'} = s\left(\tau_{XA'} - \omega_{XA'}\right), \\ D_{A'} - E_{A'} = s(q\sigma_0 - \left(1 - q\right)\sigma_1), \\ \text{Tr}[D_{A'} + E_{A'}] \leqslant s - 1, \\ \text{Tr}_{A'}[\Gamma_{AA'}^{\mathcal{N}^0} + \Gamma_{AA'}^{\mathcal{N}^1}] = I_A, \\ \langle0|_X\tau_{XA'}|0\rangle_X = \text{Tr}_A[(p\rho_A^0)^{\text{T}}\Gamma_{AA'}^{\mathcal{N}^0}] + \text{Tr}_A[((1 - p)\rho_A^1)^{\text{T}}\Gamma_{AA'}^{\mathcal{N}^1}], \\ \langle1|_X\tau_{XA'}|1\rangle_X = \text{Tr}_A[(p\rho_A^0)^{\text{T}}\Gamma_{AA'}^{\mathcal{N}^1}] + \text{Tr}_A[((1 - p)\rho_A^1)^{\text{T}}\Gamma_{AA'}^{\mathcal{N}^0}], \\ \langle1|_X\tau_{XA'}|0\rangle_X = \langle0|_X\tau_{XA'}|1\rangle_X = 0 \end{array} \right\}. \tag{C.47}$$

As written, this is not a SDP, due to the bilinear term $s\tau_{XA'}$ in the second line above, given that s is an optimization variable and $\tau_{XA'}$ includes the optimization variables $\Gamma_{AA'}^{\mathcal{N}^0}$ and $\Gamma_{AA'}^{\mathcal{N}^1}$. However, we observe that $s \geqslant 1$, due to the constraints $\text{Tr}[D_{A'} + E_{A'}] \leqslant s - 1$ and $D_{A'}, E_{A'} \geqslant 0$. We can then make the reassignments

$s\Gamma_{AA'}^{\mathcal{N}^0} \to \Omega_{AA'}^0$ and $s\Gamma_{AA'}^{\mathcal{N}^1} \to \Omega_{AA'}^1$ to rewrite the above optimization as follows:

$$
\min_{\substack{B_{XA'},C_{XA'},D_{A'},E_{A'} \geqslant 0, \\ \Omega_{AA'}^0, \Omega_{AA'}^1 \geqslant 0, s \geqslant 1}}
\left\{
\begin{array}{c}
\mathrm{Tr}[B_{XA'} + C_{XA'}]: \\
B_{XA'} - C_{XA'} = \tau_{XA'} - s\sigma_{XA'}, \\
D_{A'} - E_{A'} = s(q\sigma_0 - (1-q)\sigma_1), \\
\mathrm{Tr}[D_{A'} + E_{A'}] \leqslant s - 1, \\
\mathrm{Tr}_{A'}[\Omega_{AA'}^0 + \Omega_{AA'}^1] = sI_A, \\
\langle 0|_X \tau_{XA'}|0\rangle_X = \mathrm{Tr}_A[(p\rho_A^0)^{\mathrm{T}}\Omega_{AA'}^0] + \mathrm{Tr}_A[((1-p)\rho_A^1)^{\mathrm{T}}\Omega_{AA'}^1], \\
\langle 1|_X \tau_{XA'}|1\rangle_X = \mathrm{Tr}_A[(p\rho_A^0)^{\mathrm{T}}\Omega_{AA'}^1] + \mathrm{Tr}_A[((1-p)\rho_A^1)^{\mathrm{T}}\Omega_{AA'}^0], \\
\langle 1|_X \tau_{XA'}|0\rangle_X = \langle 0|_X \tau_{XA'}|1\rangle_X = 0
\end{array}
\right\}. \quad \text{(C.48)}
$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Appendix D. Optimizations for approximate one-shot SD-cost

In this appendix, we detail the optimization problem for approximate one-shot SD-cost under CPTP_A and CDS maps. We prove that both of these operational quantities can be calculated by means of bilinear programs.

Consider from lemma 5.11 that the approximate one-shot SD-cost is equal to

$$
\xi_c^{\mathrm{FO},q,\varepsilon}(\rho_{XA}) = \inf_{\tilde{\rho}_{XA} \in B'_\varepsilon(\rho_{XA})} \xi_c^{\mathrm{FO},q}(\tilde{\rho}_{XA}), \quad \text{(D.1)}
$$

where $\xi_c^{\mathrm{FO},q}(\tilde{\rho}_{XA})$ is the exact cost. Under CPTP_A maps, we know from theorem 5.3 that

$$
\xi_c^{\mathrm{CPTP}_A,p}(\tilde{\rho}_{XA}) = \log \inf_{M \geqslant 1} \{M : \tilde{\rho}_0 \leqslant (2M-1)\tilde{\rho}_1,\ \tilde{\rho}_1 \leqslant (2M-1)\tilde{\rho}_0\}. \quad \text{(D.2)}
$$

Combining with the SDP for the minimum conversion error (from proposition 3.16), we conclude that

$$
\xi_c^\varepsilon(\rho_{XA}) = \xi_c^{\mathrm{CPTP}_A,p,\varepsilon}(\rho_{XA}) \quad \text{(D.3)}
$$

$$
= \log \inf_{\substack{M \geqslant 1, B_{XA}, C_{XA}, \\ D_A, E_A \geqslant 0, s \in \mathbb{R}, \\ \tilde{p} \in [0,1], \tilde{\rho}_0, \tilde{\rho}_1 \geqslant 0}}
\left\{
\begin{array}{c}
M: \\
\tilde{\rho}_0 \leqslant (2M-1)\tilde{\rho}_1, \\
\tilde{\rho}_1 \leqslant (2M-1)\tilde{\rho}_0, \\
\mathrm{Tr}[B_{XA} + C_{XA}] \leqslant \varepsilon, \\
s - 1 \geqslant \mathrm{Tr}[D_A + E_A], \\
B_{XA} - C_{XA} = s(\rho_{XA} - \tilde{\rho}_{XA}), \\
D_A - E_A = s(\tilde{p}\tilde{\rho}_0 - (1-\tilde{p})\tilde{\rho}_1), \\
\tilde{\rho}_{XA} = |0\rangle\langle 0| \otimes \tilde{p}\tilde{\rho}_0 + |1\rangle\langle 1| \otimes (1-\tilde{p})\tilde{\rho}_1, \\
\mathrm{Tr}[\tilde{\rho}_0] = \mathrm{Tr}[\tilde{\rho}_1] = 1
\end{array}
\right\} \quad \text{(D.4)}
$$

$$
= \log \inf_{\substack{M \geqslant 1, B_{XA}, C_{XA}, \\ D_A, E_A \geqslant 0, s \geqslant 1, \\ \tilde{p} \in [0,1], \tilde{\rho}_0, \tilde{\rho}_1 \geqslant 0}}
\left\{
\begin{array}{c}
M: \\
\tilde{\rho}_0 \leqslant (2M-1)\tilde{\rho}_1, \\
\tilde{\rho}_1 \leqslant (2M-1)\tilde{\rho}_0, \\
\mathrm{Tr}[B_{XA} + C_{XA}] \leqslant \varepsilon, \\
s - 1 \geqslant \mathrm{Tr}[D_A + E_A], \\
B_{XA} - C_{XA} = s\rho_{XA} - \tilde{\rho}_{XA}, \\
D_A - E_A = \tilde{p}\tilde{\rho}_0 - (1-\tilde{p})\tilde{\rho}_1, \\
\tilde{\rho}_{XA} = |0\rangle\langle 0| \otimes \tilde{p}\tilde{\rho}_0 + |1\rangle\langle 1| \otimes (1-\tilde{p})\tilde{\rho}_1, \\
\mathrm{Tr}[\tilde{\rho}_0] = \mathrm{Tr}[\tilde{\rho}_1] = s
\end{array}
\right\}. \quad \text{(D.5)}
$$

This is a bilinear program, due to terms like $(2M-1)\tilde{\rho}_1$, $(2M-1)\tilde{\rho}_0$, $\tilde{p}\tilde{\rho}_0$, and $(1-\tilde{p})\tilde{\rho}_1$ appearing in the optimization.

Under CDS maps, we know from theorem 5.5 that

$$
\xi_c^{\mathrm{CDS},\frac{1}{2}}(\tilde{\rho}_{XA}) = \log \inf_{M \geqslant 1} \left\{ M : \begin{array}{c} \tilde{p}\tilde{\rho}_0 \leqslant (2M-1)(1-\tilde{p})\tilde{\rho}_1, \\ (1-\tilde{p})\tilde{\rho}_1 \leqslant (2M-1)\tilde{p}\tilde{\rho}_0 \end{array} \right\}. \quad \text{(D.6)}
$$

Combining with the SDP for the minimum conversion error (from proposition 3.16), we conclude that

$$
\xi_c^{\star,\varepsilon}(\rho_{XA}) = \xi_c^{\mathrm{CDS},\frac{1}{2},\varepsilon}(\rho_{XA}) \quad \text{(D.7)}
$$

$$= \log \inf_{\substack{M \geqslant 1, B_{XA}, C_{XA}, \\ D_A, E_A \geqslant 0, s \in \mathbb{R}, \\ \tilde{p} \in [0,1], \tilde{\rho}_0, \tilde{\rho}_1 \geqslant 0}} \left\{ \begin{array}{c} M: \\ \tilde{p}\tilde{\rho}_0 \leqslant (2M-1)(1-\tilde{p})\,\tilde{\rho}_1, \\ (1-\tilde{p})\,\tilde{\rho}_1 \leqslant (2M-1)\tilde{p}\tilde{\rho}_0, \\ \mathrm{Tr}[B_{XA} + C_{XA}] \leqslant \varepsilon, \\ s - 1 \geqslant \mathrm{Tr}[D_A + E_A], \\ B_{XA} - C_{XA} = s\,(\rho_{XA} - \tilde{\rho}_{XA})\,, \\ D_A - E_A = s\,(\tilde{p}\tilde{\rho}_0 - (1-\tilde{p})\,\tilde{\rho}_1)\,, \\ \tilde{\rho}_{XA} = |0\rangle\langle 0| \otimes \tilde{p}\tilde{\rho}_0 + |1\rangle\langle 1| \otimes (1-\tilde{p})\,\tilde{\rho}_1, \\ \mathrm{Tr}[\tilde{\rho}_0] = \mathrm{Tr}[\tilde{\rho}_1] = 1 \end{array} \right\} \tag{D.8}$$

$$= \log \inf_{\substack{M \geqslant 1, B_{XA}, C_{XA}, \\ D_A, E_A \geqslant 0, s \in \mathbb{R}, \\ \widehat{\rho}_0, \widehat{\rho}_1 \geqslant 0}} \left\{ \begin{array}{c} M: \\ \widehat{\rho}_0 \leqslant (2M-1)\,\widehat{\rho}_1, \\ \widehat{\rho}_1 \leqslant (2M-1)\,\widehat{\rho}_0, \\ \mathrm{Tr}[B_{XA} + C_{XA}] \leqslant \varepsilon, \\ s - 1 \geqslant \mathrm{Tr}[D_A + E_A], \\ B_{XA} - C_{XA} = s\,(\rho_{XA} - \tilde{\rho}_{XA})\,, \\ D_A - E_A = s\,(\widehat{\rho}_0 - \widehat{\rho}_1)\,, \\ \tilde{\rho}_{XA} = |0\rangle\langle 0| \otimes \widehat{\rho}_0 + |1\rangle\langle 1| \otimes \widehat{\rho}_1, \\ \mathrm{Tr}[\widehat{\rho}_0] + \mathrm{Tr}[\widehat{\rho}_1] = 1 \end{array} \right\} \tag{D.9}$$

$$= \log \inf_{\substack{M \geqslant 1, B_{XA}, C_{XA}, \\ D_A, E_A \geqslant 0, s \geqslant 0, \\ \widehat{\rho}_0, \widehat{\rho}_1 \geqslant 0}} \left\{ \begin{array}{c} M: \\ \widehat{\rho}_0 \leqslant (2M-1)\,\widehat{\rho}_1, \\ \widehat{\rho}_1 \leqslant (2M-1)\,\widehat{\rho}_0, \\ \mathrm{Tr}[B_{XA} + C_{XA}] \leqslant \varepsilon, \\ s - 1 \geqslant \mathrm{Tr}[D_A + E_A], \\ B_{XA} - C_{XA} = s\rho_{XA} - \tilde{\rho}_{XA}, \\ D_A - E_A = \widehat{\rho}_0 - \widehat{\rho}_1, \\ \tilde{\rho}_{XA} = |0\rangle\langle 0| \otimes \widehat{\rho}_0 + |1\rangle\langle 1| \otimes \widehat{\rho}_1, \\ \mathrm{Tr}[\widehat{\rho}_0] + \mathrm{Tr}[\widehat{\rho}_1] = s \end{array} \right\}. \tag{D.10}$$

This is also a bilinear program, due to terms like $(2M-1)\,\tilde{\rho}_1$ and $(2M-1)\,\widehat{\rho}_1$ appearing in the optimization.

ORCID iDs

Robert Salzmann https://orcid.org/0000-0003-4002-240X
Xin Wang https://orcid.org/0000-0002-0641-3186
Mark M Wilde https://orcid.org/0000-0002-3916-4462

References

[ACMnT+07]   Audenaert K M R, Calsamiglia J, Muñoz Tapia R, Bagan E, Masanes L, Acin A and Verstraete F 2007 Discriminating states: the quantum Chernoff bound *Phys. Rev. Lett.* **98** 160501

[AM14]   Audenaert K M R and Mosonyi M 2014 Upper bounds on the error probabilities and asymptotic error exponents in quantum multiple state discrimination *J. Math. Phys.* **55** 102201

[BaHO+13]   Brandão F G S L, Horodecki M, Oppenheim J, Renes J M and Spekkens R W 2013 Resource theory of quantum states out of thermal equilibrium *Phys. Rev. Lett.* **111** 250404

[BBPS96]   Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 Concentrating partial entanglement by local operations *Phys. Rev.* A **53** 2046–52

[Bel75a]   Belavkin V 1975 Optimal distinction of non-orthogonal quantum signals *Radio Eng. Electron. Phys.* **20** 39–47

[Bel75b]   Belavkin V P 1975 Optimal multiple quantum statistical hypothesis testing *Stochastics* **1** 315–45

[BK02]   Barnum H and Knill E 2002 Reversing quantum dynamics with near-optimal quantum and classical fidelity *J. Math. Phys.* **43** 2097–106

[BST19]   Buscemi F, Sutter D and Tomamichel M 2019 An information-theoretic treatment of quantum dichotomies *Quantum* **3** 209

[CB08]   Chirolli L and Burkard G 2008 Decoherence in solid-state qubits *Adv. Phys.* **57** 225–85

[CG19]   Chitambar E and Gour G 2019 Quantum resource theories *Rev. Mod. Phys.* **91** 025001

[CMTM+08]   Calsamiglia J, Munoz-Tapia R, Masanes L, Acin A and Bagan E 2008 Quantum Chernoff bound as a measure of distinguishability between density matrices: application to qubit and Gaussian states *Phys. Rev.* A **77** 032311

[Dat09]   Datta N 2009 Min- and max-relative entropies and a new entanglement monotone *IEEE Trans. Inf. Theory* **55** 2816–26

[GGH+18]   Gour G, Grudka A, Horodecki M, Kłobus W, Łodyga J and Narasimhachar V 2018 Conditional uncertainty principle *Phys. Rev.* A **97** 042130

[Hau93]   Hausladen P 1993 On the quantum mechanical channel capacity as a function of the density matrix *Bachelor's Thesis* Williams College, Williamstown, Massachusetts

[Hel67]   Helstrom C W 1967 Detection theory and quantum mechanics *Inf. Control* **10** 254–91

[Hel69]   Helstrom C W 1969 Quantum detection and estimation theory *J. Stat. Phys.* **1** 231–52

[HHO03]   Horodecki M, Horodecki P and Oppenheim J 2003 Reversible transformations from pure to mixed states and the unique measure of information *Phys. Rev.* A **67** 062104

[HKT20]   Huber S, König R and Tomamichel M 2020 Jointly constrained semidefinite bilinear programming with an application to Dobrushin curves *IEEE Trans. Inf. Theory* **66** 2934–50

[Hol72]   Holevo A S 1972 An analogue of statistical decision theory and noncommutative probability theory *Tr. Mosk. Mat. Obshchestva* **26** 133–49

[Hol79]   Holevo A S 1979 On asymptotically optimal hypothesis testing in quantum statistics *Theory Probab. Appl.* **23** 411–5

[HP91]   Hiai F and Petz D 1991 The proper formula for relative entropy and its asymptotics in quantum probability *Commun. Math. Phys.* **143** 99–114

[HW94]   Hausladen P and Wootters W K 1994 A 'pretty good' measurement for distinguishing quantum states *J. Mod. Opt.* **41** 2385–90

[HW12]   Harrow A W and Winter A 2012 How many copies are needed for state discrimination? *IEEE Trans. Inf. Theory* **58** 1–2

[KRS09]   Koenig R, Renner R and Schaffner C 2009 The operational meaning of min- and max-entropy *IEEE Trans. Inf. Theory* **55** 4337–47

[KSW20]   Khatri S, Sharma K and Wilde M M 2020 Information-theoretic aspects of the generalized amplitude-damping channel *Phys. Rev.* A **102** 012401

[Li16]   Li K 2016 Discriminating quantum states: the multiple Chernoff distance *Ann. Stat.* **44** 1661–79

[Mat10]   Matsumoto K 2010 Reverse test and characterization of quantum relative entropy (arXiv:1010.1030)

[Mat11]   Matsumoto K 2011 Reverse test and characterization of quantum relative entropy Slides available at https://sites.google.com/site/nww2011/home/talks-slides

[Mor09]   Morris R 2009 Topics in quantum foundations: ontological models, and distinguishability as a resource *Master's Thesis* University of Waterloo, Applied Mathematics

[NC10]   Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge: Cambridge University Press)

[NG17]   Narasimhachar V and Gour G 2017 Resource theory under conditioned thermal operations *Phys. Rev.* A **95** 012313

[NS09]   Nussbaum M and Szkoła A 2009 The Chernoff lower bound for symmetric quantum hypothesis testing *Ann. Stat.* **37** 1040–57

[ON00]   Ogawa T and Nagaoka H 2000 Strong converse and Stein's lemma in quantum hypothesis testing *IEEE Trans. Inf. Theory* **46** 2428–33

[SDG+21]   Salzmann R, Datta N, Gour G, Wang X and Wilde M M 2021 Symmetric distinguishability as a quantum resource (arXiv:2102.12512)

[Tho63]   Thompson A C 1963 On certain contraction mappings in a partially ordered vector space *Proc. Am. Math. Soc.* **14** 438–43

[VB96]   Vandenberghe L and Boyd S 1996 Semidefinite programming *SIAM Rev.* **38** 49–95

[Wat18]   Watrous J 2018 *The Theory of Quantum Information* (Cambridge: Cambridge University Press)

[WW19a]   Wang X and Wilde M M 2019 Resource theory of asymmetric distinguishability *Phys. Rev. Res.* **1** 033170

[WW19b]   Wang X and Wilde M M 2019 Resource theory of asymmetric distinguishability for quantum channels *Phys. Rev. Res.* **1** 033169

[WY16]   Winter A and Yang D 2016 Operational resource theory of coherence *Phys. Rev. Lett.* **116** 120404