

Moderate deviation analysis of majorisation-based resource interconversion

Christopher T. Chubb,^{1,*} Marco Tomamichel,² and Kamil Korzekwa¹

¹*Centre for Engineered Quantum Systems, School of Physics,
University of Sydney, Sydney NSW 2006, Australia.*

²*Centre for Quantum Software and Information, School of Software,
University of Technology Sydney, Sydney NSW 2007, Australia.*

We consider the problem of interconverting a finite amount of resources within all theories whose single-shot transformation rules are based on a majorisation relation, e.g. the resource theories of entanglement and coherence (for pure state transformations), as well as thermodynamics (for energy-incoherent transformations). When only finite resources are available we expect to see a non-trivial trade-off between the rate r_n at which n copies of a resource state ρ can be transformed into nr_n copies of another resource state σ , and the error level ϵ_n of the interconversion process, as a function of n . In this work we derive the optimal trade-off in the so-called moderate deviation regime, where the rate of interconversion r_n approaches its optimum in the asymptotic limit of unbounded resources ($n \rightarrow \infty$), while the error ϵ_n vanishes in the same limit. We find that the moderate deviation analysis exhibits a resonance behaviour which implies that certain pairs of resource states can be interconverted at the asymptotically optimal rate with negligible error, even in the finite n regime.

I. INTRODUCTION

In principle, while processing quantum information, any initial state can be transformed into any final state. One could thus conclude that all quantum states are equally valuable or resourceful. In reality, however, some transformations are harder to implement than others, which results in a partial ordering of the set of quantum states, with the hardest to prepare at the top, and easiest at the bottom. Such a resource hierarchy arises naturally when we face any kind of restrictions: from the locality constraint, through experimental difficulties in preparing particular superpositions, to fundamental constraints induced by physical laws like energy conservation. The mathematical framework developed to study possible state transformations under such restrictions is known under the collective name of resource theories [1, 2].

Inspired by classical information theory, the early resource-theoretic works considered optimal conversion rates between different resource states in the asymptotic regime, i.e., the limit of processing infinitely many copies of a given state. This led to the discovery of asymptotic resource measures, which provided operational meaning to quantities such as entropy of entanglement [3] and non-equilibrium free energy [4]. Namely, a given transformation becomes asymptotically possible if and only if the corresponding asymptotic resource measure is non-increasing, which allows one to reversibly interconvert between all resource states.

On the other hand, almost simultaneously to the asymptotic studies, the single-shot regime was investigated, where one aims at deciding whether it is possible to convert a single copy of an initial state into the final state. Here, probably the most famous contributions

are the Nielsen's theorem [5] within the resource theory of entanglement, and more recently the family of second laws for the resource theory of thermodynamics [6]. In general, in the single-shot regime simple asymptotic transformation rules get replaced by more complex sets of conditions, which also give rise to irreversible transformation.

In this paper we focus on the interconversion process in the intermediate regime, when the number of processed resource states is large, but finite. This way we aim at keeping the simplicity of the asymptotic analysis, but also at preserving the irreversible nature of single-shot regime. The first steps in this direction were recently made in Refs. [7] and [8] for the resource theories of entanglement and thermodynamics, where the corrections to asymptotic conversion rates were found in the scenario with a constant transformation error (i.e., in the small deviation regime [9]). Here, we present a moderate deviation analysis [10] (see also [11, 12] for applications in the quantum domain) of the interconversion problem within a unified framework that includes all resource theories for which the single-shot transformation rules can be expressed via majorisation or thermo-majorisation. This way we find finite-size corrections to conversion rates in resource theories of entanglement [13], coherence [14] and thermodynamics [15], in the regime where the transformation error, measured by either infidelity or total variation distance, asymptotically vanishes.

Our results can be directly applied to the study of important problems such as entanglement distillation [16] or coherence dilution [17], but also allow one for a rigorous analysis of the irreversibility arising when finite-size resources are interconverted. Most intriguingly, we find that if a pair of states satisfies a particular resonance condition, one can achieve lossless interconversion, i.e., transformation that is arbitrarily close to reversible even for finite n . In the accompanying paper [18] we discuss how this effect can be employed to avoid irreversibility,

* christopher.chubb@sydney.edu.au

which directly affects, e.g., the performance of heat engines working with finite-size working bodies [19].

This paper is structured in the following way. First, in Sec. II, we set the scene by introducing necessary tools and concepts. Next, in Sec. III we state our main result concerning moderate deviation corrections to the asymptotic interconversion rates for majorisation-based resource theories. We then proceed to Sec. IV that contains auxiliary technical results concerning tail bounds, which are used in the formal proof that can be found in Sec. V. Finally, we provide conclusions and outlook in Sec. VI.

II. SETTING THE SCENE

A. Resource theories in different regimes

Every quantum resource theory [1, 2] is defined by a set of quantum operations that are considered free, and a set of restrictions that make other operations impossible without an additional cost. Such restrictions may arise from practical difficulties, e.g., when preparing a system in a superposition of particular states is experimentally challenging, but may also be of fundamental nature, as with the laws of thermodynamics constraining possible transformations to preserve energy and increase entropy. A resource is then defined as a quantum system that allows one to lift a given restriction. Typical examples of resources include an excited pure state that acts as a work storage, and thus can be used to reduce the entropy of another system (overcoming thermodynamic constraints [15]); an entangled Bell pair, which can be used to teleport a quantum state (overcoming locality constraints [13]); or a system in the superposition of energy eigenstates, which can be used as a reference frame for time (overcoming symmetry constraints [20]).

Once the restrictions and the corresponding resources are defined, the central question concerns resource interconversion, i.e., what final states can be obtained from the initial one. This problem so far was mainly approached in either the *single-shot regime* [21], or in an idealised *asymptotic limit* [22]. The first approach, due to its generality and the corresponding complexity of the answer, provides only a limited insight into the nature of different resource states. The second one provides an elegant and simple answer in the form of optimal *conversion rate*, which tells us how many copies of the final state can be obtained per one copy of the initial state, if one assumes access to an infinite source of initial states. From a practical point of view, however, such an assumption is unjustified, as most quantum resources will be available only in small amounts in the foreseeable future. More fundamentally, finite-size effects may be of interest themselves, as it is the case within quantum thermodynamics [23], where one aims at accurate description of heat and work processes involving small number of particles.

	Ent.	Coh.	Thermo.
$n = 1, \epsilon = 0$	[5]	[25]	[15]
$n \rightarrow \infty, \epsilon \rightarrow 0$		[3]	[4]
$n < \infty, \epsilon > 0$		[7]	[8]
$n < \infty, \epsilon \rightarrow 0$	This work		

TABLE I. *Interconversion in various regimes.* Exposition of works on state interconversion within resource theories of ent(anglement), coh(erence) and thermo(dynamics) in single-shot regime ($n = 1, \epsilon = 0$), asymptotic limit ($n \rightarrow \infty, \epsilon \rightarrow 0$), small deviation regime ($n < \infty, \epsilon > 0$) and moderate deviation regime ($n < \infty, \epsilon \rightarrow 0$).

Very recently the first steps have been made to study the intermediate regime, where one focuses on the interconversion of large but finite number n of resource states. First, in Ref. [7] the authors focused on transformations within the resource theory of entanglement. Their results were then generalised and adapted to the studies of the interconversion process in the resource theory of thermodynamics by the present authors [8]. In both these works the second-order correction to the asymptotic rate was found in the so-called *small deviation* regime [9], where the conversion rate approaches the asymptotic one for $n \rightarrow \infty$, but the transformation is realised with a constant error. In the current work we solve the issue of constant error by deriving corrections to the asymptotic rate in the *moderate deviation* regime [10], where the correction term still vanishes as $n \rightarrow \infty$, but also the transformation is asymptotically error-free. For the completeness of discussion, we also note that the interconversion problem may be studied in the *large deviation* regime [24], where the error is exponentially vanishing for the price of the constant gap between the realised conversion rate and the asymptotic one. In Table I we collect references to central results concerning state interconversion within the investigated resource theories in various regimes.

B. Exact single-shot interconversion

Irrespective of the investigated regime, the first step is to find single-shot interconversion rules, which form the basis of further analysis. In this work we study the interconversion problem within all majorisation-based resource theories, i.e., when conditions for single-shot transformations can be expressed as majorisation partial order [26], or a variant known as thermo-majorisation [15, 27]. Within such theories, each resource state can be represented by a probability distribution, and the conversion process is possible when the distribution representing the initial state majorises (or is majorised) by the distribution representing the final state, with majorisation \succ defined by

$$\mathbf{a} \succ \mathbf{b} \iff \forall j : \sum_{i=1}^j a_i^\downarrow \geq \sum_{i=1}^j b_i^\downarrow, \quad (1)$$

where \mathbf{a}^\downarrow denotes the vector \mathbf{a} in a decreasing order.

Three prominent examples of majorisation-based resource theories include the resource theories of entanglement, coherence and thermodynamics. These are defined via the relevant sets of free operations and free states: Local Operations and Classical Communication (LOCC) and separable states in entanglement theory [13]; Incoherent Operations and incoherent states in coherence theory [14]; Thermal Operations and the thermal equilibrium state γ in the resource theory of thermodynamics (with respect to a fixed background temperature $T = 1/\beta$) [28]. As mentioned above, within each of these theories there exists a representation of initial and target quantum states, ρ and σ , as probability distributions \mathbf{p} and \mathbf{q} . For entanglement theory, given initial and target pure bipartite states, $\rho = |\Psi\rangle\langle\Psi|$ and $\sigma = |\Phi\rangle\langle\Phi|$, with the Schmidt decomposition given by

$$|\Psi\rangle = \sum_i a_i |\psi_i \psi_i\rangle, \quad |\Phi\rangle = \sum_i b_i |\phi_i \phi_i\rangle, \quad (2)$$

we can represent them via probability distributions

$$p_i = |a_i|^2, \quad q_i = |b_i|^2. \quad (3)$$

For coherence theory, with respect to a fixed basis $\{|i\rangle\}$, one can represent pure initial and target states, $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$, using

$$p_i = \langle i|\psi\rangle^2, \quad q_i = \langle i|\phi\rangle^2. \quad (4)$$

Finally, in the resource theory of thermodynamics, the initial and target energy-incoherent mixed states ρ and σ can be represented by

$$p_i = \langle E_i|\rho|E_i\rangle, \quad q_i = \langle E_i|\sigma|E_i\rangle, \quad (5)$$

where $\{|E_i\rangle\}$ denotes the energy eigenbasis of the system. We will denote distributions representing free states by \mathbf{f} . In entanglement and coherence theories these are represented by sharp probability distributions \mathbf{s} with a single non-zero entry; whereas in the thermodynamic case \mathbf{f} is given by a thermal Gibbs distribution γ with $\gamma_i \propto \exp(-\beta E_i)$.

The celebrated Nielsen's theorem [5] (for entanglement) and the recent result of Ref. [25] (for coherence) state that the initial state represented by \mathbf{p} can be transformed into the target state represented by \mathbf{q} if and only if $\mathbf{p} \prec \mathbf{q}$. Similarly, in Ref. [15], it was found that a thermodynamic transformations between states represented by \mathbf{p} and \mathbf{q} is possible if and only if $\hat{\mathbf{p}} \succ \hat{\mathbf{q}}$, where $\hat{\mathbf{a}}$ can be obtained from \mathbf{a} via a straightforward application of an *embedding map* Γ^β [6, 29]. For the sake of our analysis, it is only crucial to note that Γ^β maps d -dimensional distributions to \hat{d} -dimensional ones with $\hat{d} \geq d$; and that an embedded version of the free thermal distribution is given by a maximally mixed distribution on a larger subspace, i.e., $\hat{\gamma} = \boldsymbol{\eta}$ with $\boldsymbol{\eta} = [1/\hat{d}, \dots, 1/\hat{d}]$.

C. Approximate multi-copy interconversion

When considering transformations between many copies of initial and target states, represented by $\mathbf{p}^{\otimes n}$ and $\mathbf{q}^{\otimes m}$, we need to make sure that the dimensionality of the input and output spaces match. Since one can always append any number of free states \mathbf{f} to both the initial and target states, we introduce total initial and target distributions,

$$\mathbf{P}^{n,m} := \mathbf{p}^{\otimes n} \otimes \mathbf{f}^{\otimes m}, \quad \mathbf{Q}^{n,m} := \mathbf{q}^{\otimes m} \otimes \mathbf{f}^{\otimes n}. \quad (6)$$

Our main object of interest will be the *conversion rate* $r_n := m/n$, i.e., the number of target states one can obtain per one copy of the initial state. For notational clarity we will denote total initial and target distributions by \mathbf{P}^n and \mathbf{Q}^n , with the dependence on m (so, in fact, on r_n) kept implicit. The single-shot interconversion conditions can now be expressed as $\mathbf{P}^n \prec \mathbf{Q}^n$ for the entanglement and coherence transformations, and $\hat{\mathbf{P}}^n \succ \hat{\mathbf{Q}}^n$ for the thermodynamic transformations.

We also need to introduce the concept of approximate interconversion. Assume that for given \mathbf{P}^n and \mathbf{Q}^n the relevant majorisation relation does not hold, so that the interconversion is impossible. However, there may exist $\tilde{\mathbf{Q}}^n$ that is ϵ -close to \mathbf{Q}^n and such that the interconversion is possible. We then say that an approximate transformation is possible with the error level ϵ quantified by either the infidelity, $1 - F$, or total variation distance (TVD), δ , between target and final states, with

$$F(\mathbf{Q}^n, \tilde{\mathbf{Q}}^n) := \left(\sum_i \sqrt{Q_i^n \tilde{Q}_i^n} \right)^2, \quad (7a)$$

$$\delta(\mathbf{Q}^n, \tilde{\mathbf{Q}}^n) := \frac{1}{2} \sum_i |Q_i^n - \tilde{Q}_i^n|. \quad (7b)$$

The concept of approximate interconversion gives rise to two notions of approximate majorisation introduced in Ref. [8], ϵ -post-majorisation \succ_ϵ and ϵ -pre-majorisation \prec_ϵ , defined by

$$\mathbf{a} \succ_\epsilon \mathbf{b} \iff \exists \tilde{\mathbf{b}} : \mathbf{a} \succ \tilde{\mathbf{b}} \text{ and } \delta(\mathbf{b}, \tilde{\mathbf{b}}) \leq \epsilon, \quad (8a)$$

$$\mathbf{a} \prec_\epsilon \mathbf{b} \iff \exists \tilde{\mathbf{a}} : \tilde{\mathbf{a}} \succ \mathbf{b} \text{ and } \delta(\mathbf{a}, \tilde{\mathbf{a}}) \leq \epsilon, \quad (8b)$$

where, depending on the context, δ can be replaced by $1 - F$. Crucially, in Ref. [8] the present authors showed that these two notions are equivalent and, moreover, that ϵ -post-majorisation between embedded vectors, $\hat{\mathbf{a}} \succ_\epsilon \hat{\mathbf{b}}$, is a necessary and sufficient condition for the existence of an approximate thermodynamic transformation between \mathbf{a} and \mathbf{b} with error level ϵ .

We conclude that an approximate transformation between initial and target states, represented by $\mathbf{p}^{\otimes n}$ and $\mathbf{q}^{\otimes nr_n}$, is possible within resource theories of entanglement and coherence if and only if

$$\mathbf{P}^n \prec_\epsilon \mathbf{Q}^n, \quad (9)$$

with the free state $\mathbf{f} = \mathbf{s}$. We will refer to the above relation as the approximate majorisation relation for the *entanglement direction*. Similarly, such a transformation is possible within resource theory of thermodynamics if and only if

$$\hat{\mathbf{P}}^n \succ_{\epsilon} \hat{\mathbf{Q}}^n, \quad (10)$$

with the free state $\mathbf{f} = \gamma$. We will refer to this relation as the approximate majorisation relation for the *thermodynamic direction*.

D. Information-theoretic notions

The main role in the quantitative analysis of the interconversion process for the entanglement direction will be played by the Shannon entropy H and entropy variance V . For a given probability distribution \mathbf{a} these are defined by

$$H(\mathbf{a}) = -\sum_i a_i \ln a_i, \quad (11a)$$

$$V(\mathbf{a}) = \sum_i a_i [\ln a_i + H(\mathbf{a})]^2. \quad (11b)$$

The analogous role for the thermodynamic direction will be played by the relative entropy D and relative entropy variance V . Given two probability distributions, \mathbf{a} and \mathbf{b} , these are defined by

$$D(\mathbf{a}||\mathbf{b}) = \sum_i a_i \ln \frac{a_i}{b_i}, \quad (12a)$$

$$V(\mathbf{a}||\mathbf{b}) = \sum_i a_i \left[\ln \frac{a_i}{b_i} - D(\mathbf{a}||\mathbf{b}) \right]^2. \quad (12b)$$

An important fact, that can be verified by direct calculation, is that the relative quantities are invariant under embedding, i.e., $D(\mathbf{a}||\mathbf{b}) = D(\hat{\mathbf{a}}||\hat{\mathbf{b}})$ and $V(\mathbf{a}||\mathbf{b}) = V(\hat{\mathbf{a}}||\hat{\mathbf{b}})$ [8].

In order to formally state our main result we also need to introduce the notion of a *moderate sequence*:

Definition 1 (Moderate sequence). A sequence of real numbers $\{t_n\}_n$ is a *moderate sequence* if its scaling is strictly between $1/\sqrt{n}$ and 1, meaning that $t_n \rightarrow 0$ and $\sqrt{n}t_n \rightarrow +\infty$ as $n \rightarrow \infty$.

Note that an important family of moderate sequences is given by $t_n \sim n^{-\alpha}$ for $\alpha \in (0, 1/2)$, which can be used to obtain a particularly simple version of our main results.

Finally, as we will be interested in asymptotic expansions in n , we will employ the standard asymptotic notation: $o(f(n))$, $O(f(n))$ and $\Theta(f(n))$. We will also use $\overset{\text{ev.}}{>}$ and $\overset{\text{ev.}}{<}$ to denote eventual inequalities, specifically we write $a_n \overset{\text{ev.}}{>} b_n$ if and only if there exists N such that $a_n > b_n$ for all $n \geq N$. Moreover, we will denote equalities and inequalities up to terms of order $o(t_n)$ by \simeq , \lesssim and \gtrsim .

III. INTERCONVERSION RATES BEYOND THE ASYMPTOTIC REGIME

We are now ready to state our central technical result, which may be of interest outside the resource-theoretic studies due to ubiquity of majorisation partial order in the broad field of applied mathematics [26]. We split it into three theorems. The first two concern state interconversion below the asymptotic rate and with asymptotically vanishing error (one for each majorisation direction). The third one concerns practically less relevant scenario of state interconversion above the asymptotic rate and with error asymptotically approaching 1.

For the entanglement direction we introduce the *optimal conversion rate* $R_n^{\text{ent}}(\epsilon)$ as the largest conversion rate r_n for which the approximate majorisation relation for the entanglement direction, $\hat{\mathbf{P}}^n \prec_{\epsilon} \hat{\mathbf{Q}}^n$, holds. Due to the discussion presented in Sec. II, $R_n^{\text{ent}}(\epsilon)$ is the maximal rate for which the approximate interconversion, with error ϵ , is possible between states represented by \mathbf{p} and \mathbf{q} within resource theories of entanglement and coherence. We also define the *asymptotic rate*,

$$R_{\infty}^{\text{ent}} = \frac{H(\mathbf{p})}{H(\mathbf{q})}, \quad (13)$$

and the *irreversibility parameter*,

$$\nu^{\text{ent}} = \frac{V(\mathbf{p})/H(\mathbf{p})}{V(\mathbf{q})/H(\mathbf{q})}. \quad (14)$$

We then have:

Theorem 1 (Entanglement direction). *For any moderate sequence t_n and the accepted error level of*

$$\epsilon_n = e^{-nt_n^2}, \quad (15)$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{ent}}(\epsilon_n)$ is

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_{\infty}^{\text{ent}} - \sqrt{\frac{2V(\mathbf{p})}{H(\mathbf{q})^2}} \left| 1 - 1/\sqrt{\nu^{\text{ent}}} \right| t_n. \quad (16)$$

Analogously, for the thermodynamic direction we introduce the *optimal conversion rate* $R_n^{\text{th}}(\epsilon)$ as the largest conversion rate r_n for which the approximate majorisation relation for the thermodynamic direction, $\hat{\mathbf{P}}^n \succ_{\epsilon} \hat{\mathbf{Q}}^n$, holds. As before, $R_n^{\text{th}}(\epsilon)$ is the maximal rate for which the approximate interconversion, with error ϵ , is possible between states represented by \mathbf{p} and \mathbf{q} within the resource theory of thermodynamics. We also define the *asymptotic rate*,

$$R_{\infty}^{\text{th}} = \frac{D(\mathbf{p}||\gamma)}{D(\mathbf{q}||\gamma)}, \quad (17)$$

and the *irreversibility parameter*,

$$\nu^{\text{th}} = \frac{V(\mathbf{p}||\gamma)/D(\mathbf{p}||\gamma)}{V(\mathbf{q}||\gamma)/D(\mathbf{q}||\gamma)}. \quad (18)$$

We then have:

Theorem 2 (Thermodynamic direction). *For any moderate sequence t_n and the accepted error level of*

$$\epsilon_n = e^{-nt_n^2}, \quad (19)$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{th}}(\epsilon_n)$ is

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} - \sqrt{\frac{2V(\mathbf{p}|\gamma)}{D(\mathbf{q}|\gamma)^2}} \left| 1 - 1/\sqrt{\nu^{\text{th}}} \right| t_n. \quad (20)$$

Finally, one expects that conversion above the asymptotic rate leads to transformation error approaching 1. This is formalised in the following theorem which, unlike the previous two theorems (that hold for the error level measured by both infidelity and total variation distance), applies only to TVD. In Appendix A, where we relate our current results to the small deviation analysis of Refs. [7, 8], we also conjecture the analogue of Theorem 3 with the error measured by infidelity.

Theorem 3 (Converse regime). *For any moderate sequence t_n and the accepted TVD error of*

$$\epsilon_n = 1 - e^{-nt_n^2}, \quad (21)$$

the asymptotic expansion of the optimal conversion rate $R_n^{\text{ent}}(\epsilon_n)$ is

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} + \sqrt{\frac{2V(\mathbf{p})}{H(\mathbf{q})^2}} \left(1 + 1/\sqrt{\nu^{\text{ent}}} \right) t_n, \quad (22a)$$

and similarly for $R_n^{\text{th}}(\epsilon_n)$ we have

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} + \sqrt{\frac{2V(\mathbf{p}|\gamma)}{D(\mathbf{q}|\gamma)^2}} \left(1 + 1/\sqrt{\nu^{\text{th}}} \right) t_n. \quad (22b)$$

We present the proofs in Sec. V, after we introduce the necessary tools in Sec. IV. Before that let us make two important remarks.

Remark 4. For initial and target states satisfying $\nu^{\text{ent}} = 1$, the optimal conversion rate R_n^{ent} in the regime of vanishing error is given by the asymptotic rate R_∞^{ent} . This means that, up to terms of order $o(t_n)$, such a transformation is reversible even for finite n . Analogous observation holds for the thermodynamic direction. We discuss the implications of this particularly interesting scenario in an accompanying paper [18].

Remark 5. When $V(\mathbf{p}) = 0$, resulting in $1/\sqrt{\nu^{\text{ent}}}$ diverging to infinity and the apparent multiplication of zero times infinity, one can simply use the definition of ν^{ent} to replace Eq. (16) with

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} \pm \sqrt{\frac{2V(\mathbf{q})H(\mathbf{p})}{H(\mathbf{q})^3}} t_n. \quad (23)$$

Analogous observation holds for the thermodynamic direction.

IV. MODERATE DEVIATION TOOLKIT

A. Preliminaries

The central result of the moderate deviation analysis can be stated as follows.

Lemma 6 (Moderate deviation bound). *Let $\{X_i\}_{1 \leq i \leq n}$ be independent and identically distributed (i.i.d.) random variables with zero-mean and variance v . For any moderate sequence $\{t_n\}_n$ the following hold:*

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\Pr \left(\frac{1}{n} \sum_{i=1}^n X_i \geq t_n \right) \right] = -\frac{1}{2v}, \quad (24a)$$

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\Pr \left(\frac{1}{n} \sum_{i=1}^n X_i \leq -t_n \right) \right] = -\frac{1}{2v}. \quad (24b)$$

The proof of the above lemma can be found, e.g., in Appendix A of Ref. [11]. For the remainder of the paper, consider $\{t_n\}_n$ to be a fixed moderate sequence. For clarity we will henceforth omit the dependence of all implicit constants on this sequence. It should be noted that the above lemma also holds when $v = 0$, where we henceforth adopt the convention that $1/v = +\infty$ in this case.

B. Two variations on tail bounds

We now want to adapt Lemma 6 to our purposes of majorisation-based analysis. For a probability vector \mathbf{a} we thus introduce the following quantity

$$k_n(\mathbf{a}, x) := \exp \left(H(\mathbf{a}^{\otimes n}) + xnt_n \right), \quad (25)$$

which allows us to formulate the magnitude-based version of the moderate deviation bound for products of distributions.

Lemma 7 (Magnitude-based tail bound). *Consider an arbitrary probability distribution \mathbf{a} . For $x \leq 0$ we have*

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \sum_i \left\{ (\mathbf{a}^{\otimes n})_i \left| (\mathbf{a}^{\otimes n})_i \geq \frac{1}{k_n(\mathbf{a}, x)} \right. \right\} = \frac{-x^2}{2V(\mathbf{a})}, \quad (26a)$$

and similarly for $x \geq 0$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \sum_i \left\{ (\mathbf{a}^{\otimes n})_i \left| (\mathbf{a}^{\otimes n})_i \leq \frac{1}{k_n(\mathbf{a}, x)} \right. \right\} = \frac{-x^2}{2V(\mathbf{a})}. \quad (26b)$$

Proof. Consider the random variable $L := -\log a$, distributed according to \mathbf{a} , such that the expectation value $\langle L \rangle$ and the variance $\text{Var}(L)$ are equal to $H(\mathbf{a})$ and $V(\mathbf{a})$ respectively. We can express k_n in terms of L as

$$\log k_n(\mathbf{a}, x) = n\langle L \rangle_{\mathbf{a}} + xnt_n. \quad (27)$$

If we let $\{L_j\}_{1 \leq j \leq n}$ be i.i.d. copies of L , then we can write the tail bound of $\mathbf{a}^{\otimes n}$ in terms of tail bounds on the average of these variables,

$$\begin{aligned} & \sum_i \left\{ (\mathbf{a}^{\otimes n})_i \mid (\mathbf{a}^{\otimes n})_i \geq \frac{1}{k_n(\mathbf{a}, x)} \right\} \\ &= \sum_{i_1, \dots, i_n} \left\{ \prod_{j=1}^n a_{i_j} \mid \prod_{j=1}^n a_{i_j} \geq \frac{1}{k_n(\mathbf{a}, x)} \right\} \\ &= \sum_{i_1, \dots, i_n} \left\{ \prod_{j=1}^n a_{i_j} \mid \sum_{j=1}^n \log a_{i_j} \geq \log \frac{1}{k_n(\mathbf{a}, x)} \right\} \\ &= \Pr \left[\sum_{j=1}^n L_j \leq n\langle L \rangle + xnt_n \right]. \end{aligned} \quad (28)$$

For $x < 0$, we can now apply Lemma 7 to the variables $X_j := (L_j - \langle L \rangle) / x$ to obtain Eq. (26a). An analogous argument can be employed for $x > 0$, with all of the above inequalities reversed, yielding Eq. (26b). Finally, for $x = 0$ case, we can appeal to the Central Limit Theorem, which gives

$$\begin{aligned} & \sum_i \left\{ (\mathbf{a}^{\otimes n})_i \mid (\mathbf{a}^{\otimes n})_i \geq \frac{1}{k_n(\mathbf{a}, 0)} \right\} \\ &= \Pr \left[\frac{1}{n} \sum_{j=1}^n L_j \leq \langle L \rangle \right] \xrightarrow{n \rightarrow \infty} \frac{1}{2}, \end{aligned} \quad (29)$$

implying Eqs. (26a)-(26b). \square

Using the above result we can now prove the majorisation-based version of the moderate deviation bound.

Lemma 8 (Majorisation-based tail bound). *Consider an arbitrary probability distribution \mathbf{a} satisfying $V(\mathbf{a}) > 0$. For $x \leq 0$ we have*

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq k_n(\mathbf{a}, x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] = -\frac{x^2}{2V(\mathbf{a})}, \quad (30a)$$

and similarly for $x \geq 0$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq k_n(\mathbf{a}, x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] = -\frac{x^2}{2V(\mathbf{a})}. \quad (30b)$$

Proof. Here we follow the proof of the small-deviation analogue of this result, Lemmas 15 and 16 of Ref. [7]. Consider first the $x \leq 0$ case, and define two sets of indices

$$S_n(x) := \{1, \dots, \lfloor k_n(\mathbf{a}, x) \rfloor\}, \quad (31a)$$

$$\tilde{S}_n(x) := \left\{ i \mid (\mathbf{a}^{\otimes n})_i^\downarrow \geq 1/k_n(\mathbf{a}, x) \right\}. \quad (31b)$$

We note that Lemma 7 gives that

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \in \tilde{S}_n(x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] = -\frac{x^2}{2V(\mathbf{a})} \quad (32)$$

for any $x \leq 0$, and we wish to show an analogous result for $S_n(x)$. We will achieve this by showing, for any $\delta > 0$, that $\tilde{S}_n(x) \subseteq S_n(x) \subseteq \tilde{S}_n(x + \delta)$ holds eventually, i.e., for large enough n . The first inclusion follows trivially from the normalisation of our distribution, and so it is left only to show that $\tilde{S}_n(x) \subseteq S_n(x + \delta)$.

Noting that $(\mathbf{a}^{\otimes n})_i^\downarrow - 1/k_n(\mathbf{a}, x + \delta/2) \geq 0$ if and only if $i \in \tilde{S}_n(x + \delta/2)$, we see that

$$\begin{aligned} & \sum_{i \in \tilde{S}_n(x + \delta/2)} \left[(\mathbf{a}^{\otimes n})_i^\downarrow - \frac{1}{k_n(\mathbf{a}, x + \delta/2)} \right] \\ & \geq \sum_{i \in T} \left[(\mathbf{a}^{\otimes n})_i^\downarrow - \frac{1}{k_n(\mathbf{a}, x + \delta/2)} \right], \end{aligned} \quad (33)$$

for any set of indices T . Taking $T = \tilde{S}_n(x + \delta)$, this gives

$$\frac{|\tilde{S}_n(x + \delta) \setminus \tilde{S}_n(x + \delta/2)|}{k_n(\mathbf{a}, x + \delta/2)} \geq \sum_{i \in \tilde{S}_n(x + \delta) \setminus \tilde{S}_n(x + \delta/2)} (\mathbf{a}^{\otimes n})_i^\downarrow. \quad (34)$$

Lemma 7 tells us that the summation on the RHS scales as $e^{-\Theta(nt_n^2)}$, specifically that there is a lower bound of the form $e^{-Cnt_n^2}$ for some constant C . Since $t_n \rightarrow 0$, we eventually have that $Ct_n < \delta/2$, and so this sum can be lower bounded as follows

$$\sum_{i \in \tilde{S}_n(x + \delta) \setminus \tilde{S}_n(x + \delta/2)} (\mathbf{a}^{\otimes n})_i^\downarrow \stackrel{\text{ev.}}{>} e^{-\delta nt_n/2}. \quad (35)$$

Applying this bound to Eq. (34) allows us to conclude

$$\begin{aligned} \left| \tilde{S}_n(x + \delta) \right| & \geq \left| \tilde{S}_n(x + \delta) \setminus \tilde{S}_n(x + \delta/2) \right| \\ & \stackrel{\text{ev.}}{>} e^{-\delta nt_n} k_n(\mathbf{a}, x + \delta/2), \\ & = k_n(\mathbf{a}, x), \end{aligned} \quad (36)$$

and therefore that $S_n(x) \subseteq \tilde{S}_n(x + \delta)$ as required.

The inclusions $\tilde{S}_n(x) \subseteq S_n(x) \subseteq \tilde{S}_n(x + \delta)$, together with Lemma 7, give us the following inequalities

$$\liminf_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \in S_n(x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] \geq -\frac{(x + \delta)^2}{2V(\mathbf{a})}, \quad (37a)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \in S_n(x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] \leq -\frac{x^2}{2V(\mathbf{a})}. \quad (37b)$$

As this holds for any $\delta > 0$, we conclude that

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \in S_n(x)} (\mathbf{a}^{\otimes n})_i^\downarrow \right] = -\frac{x^2}{2V(\mathbf{a})}, \quad (38)$$

which is equivalent to Equation (30a). An analogous proof can be performed for $x \geq 0$, resulting in Eq. (30b). \square

Remark 9. One can extend Lemma 8 to probability distributions \mathbf{a} with $V(\mathbf{a}) = 0$ by a direct calculation, since $V(\mathbf{a}) = 0$ means all non-zero entries of \mathbf{a} are equal. One then obtains that Eq. (30a) holds for $x < 0$ and Eq. (30b) for $x > 0$, i.e., both expressions diverge to $-\infty$.

C. Tail bounds for total distributions

Recall that in Sec. II C we defined total initial and target states for a given rate r_n as

$$\mathbf{P}^n := \mathbf{p}^{\otimes n} \otimes \mathbf{f}^{\otimes nr_n}, \quad \mathbf{Q}^n := \mathbf{q}^{\otimes nr_n} \otimes \mathbf{f}^{\otimes n}, \quad (39)$$

where \mathbf{f} stands for the free state of a given resource theory, i.e., \mathbf{f} is a sharp state \mathbf{s} for entanglement and coherence transformations, and \mathbf{f} is the maximally mixed state $\boldsymbol{\eta}$ in the case of thermodynamic transformations (corresponding to the embedded thermal state $\boldsymbol{\gamma}$). For notational clarity we will henceforth omit the \downarrow superscripts on these total states, assuming them to be ordered (i.e. we denote $\mathbf{P}^{n\downarrow}$ and $\mathbf{Q}^{n\downarrow}$ simply by \mathbf{P}^n and \mathbf{Q}^n).

Analogous to the quantity which appears in our moderate deviation bounds, consider the quantity

$$K_n(x) := \exp(H(\mathbf{Q}^n) + xnt_n). \quad (40)$$

Using Lemma 8 we can prove the following tail bounds for the total distributions.

Lemma 10 (Tail bound for \mathbf{P}^n and \mathbf{Q}^n). *For any $\mu \in \mathbb{R}$, consider the conversion rate*

$$r_n(\mu) = \frac{H(\mathbf{f}) - H(\mathbf{p}) + \mu t_n}{H(\mathbf{f}) - H(\mathbf{q})}, \quad (41)$$

and the irreversibility parameter

$$\nu := \frac{V(\mathbf{p})}{V(\mathbf{q})} \cdot \frac{H(\mathbf{f}) - H(\mathbf{q})}{H(\mathbf{f}) - H(\mathbf{p})}. \quad (42)$$

The total output state \mathbf{Q}^n has the tail bounds

$$x \leq 0 : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(x)} Q_i^n \right] = -\frac{\nu x^2}{2V(\mathbf{p})}, \quad (43a)$$

$$x \geq 0 : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(x)} Q_i^n \right] = -\frac{\nu x^2}{2V(\mathbf{p})}. \quad (43b)$$

Similarly, the total input state \mathbf{P}^n has the tail bounds

$$x \leq \mu : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(x)} P_i^n \right] = -\frac{(x - \mu)^2}{2V(\mathbf{p})}, \quad (44a)$$

$$x \geq \mu : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(x)} P_i^n \right] = -\frac{(x - \mu)^2}{2V(\mathbf{p})}. \quad (44b)$$

Proof. For $\mathbf{f} = \mathbf{s}$ we have

$$\begin{aligned} \sum_{i \leq K_n(x)} Q_i^n &= \sum_{i \leq k_{nr_n}(\mathbf{q}, x/r_n)} (\mathbf{q}^{\otimes nr_n} \otimes \mathbf{s}^{\otimes n})_i^\downarrow \\ &= \sum_{i \leq k_{nr_n}(\mathbf{q}, x/r_n)} (\mathbf{q}^{\otimes nr_n})_i^\downarrow. \end{aligned} \quad (45)$$

Similarly for $\mathbf{f} = \boldsymbol{\eta}$ we have

$$\begin{aligned} \sum_{i \leq K_n(x)} Q_i^n &= \sum_{i \leq d^n k_{nr_n}(\mathbf{q}, x/r_n)} (\mathbf{q}^{\otimes nr_n} \otimes \boldsymbol{\eta}^{\otimes n})_i^\downarrow \\ &= \sum_{i \leq k_{nr_n}(\mathbf{q}, x/r_n)} (\mathbf{q}^{\otimes nr_n})_i^\downarrow. \end{aligned} \quad (46)$$

Applying Lemma 8 to both of the above equations yields the desired bounds.

Next, define $K_n^P(y) := \exp(H(\mathbf{P}^n) + ynt_n)$. By analogy to \mathbf{Q}^n , we have the following tail bounds on \mathbf{P}^n

$$y \leq 0 : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n^P(y)} P_i^n \right] = -\frac{y^2}{2V(\mathbf{p})}, \quad (47a)$$

$$y \geq 0 : \quad \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n^P(y)} P_i^n \right] = -\frac{y^2}{2V(\mathbf{p})}. \quad (47b)$$

Using the rate $r_n(\mu)$, and expanding out both K_n and K_n^P , we find that $K_n^P(x - \mu) = K_n(x)$. Substituting this into the above expressions, we get the desired tail bounds purely in terms of $K_n(x)$. \square

We now want to consider the regions in which our two total distributions majorise each other. To do this, we first define the values of x for which the tail bounds for \mathbf{P}^n and \mathbf{Q}^n coincide. Let us introduce

$$z_C := \frac{\mu}{1 - \sqrt{\nu}}, \quad \text{and} \quad z_T := \frac{\mu}{1 + \sqrt{\nu}}. \quad (48)$$

These correspond to the values of x for which the moderate deviation tail bounds of the total distributions meet on the same side (cis) or on opposite sides (trans), respectively. More precisely, as a consequence of Lemma 10, z_C and z_T are the solutions to the following equations

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_C)} P_i^n \right] = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_C)} Q_i^n \right], \quad (49a)$$

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_C)} P_i^n \right] = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_C)} Q_i^n \right], \quad (49b)$$

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_T)} P_i^n \right] = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_T)} Q_i^n \right], \quad (49c)$$

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_T)} P_i^n \right] = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_T)} Q_i^n \right]. \quad (49d)$$

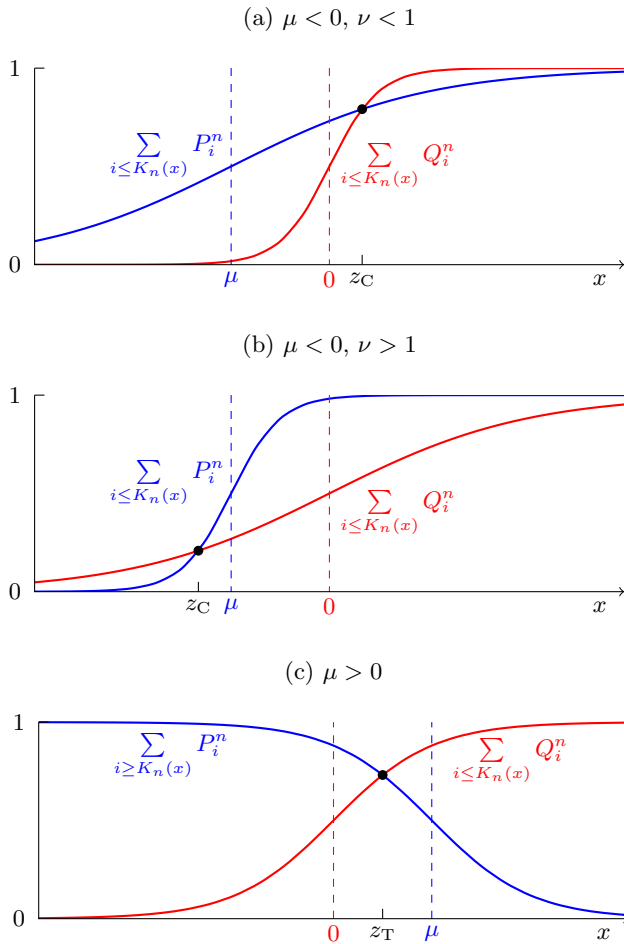


FIG. 1. Schematic representation of cumulative distribution functions for P^n and Q^n , and the positions of z_C and z_T in different regimes.

We schematically present the positions of z_C and z_T in Fig. 1, which also serves to illustrate the proof of the following lemma.

Lemma 11 (Dominance of total states). *For a bounded interval $[a, b]$, such that $a > z_C$ (for $\nu < 1$) or $b < z_C$ (for $\nu > 1$), the inequalities*

$$\sum_{i \leq K_n(x)} P_i^n > \sum_{i \leq K_n(x)} Q_i^n \quad (50)$$

hold for all $x \in [a, b]$, for sufficiently large n . Similarly, for any bounded interval $[a, b]$ with $b < z_T$, the inequalities

$$\sum_{i \geq K_n(x)} P_i^n > \sum_{i \geq K_n(x)} Q_i^n \quad (51)$$

hold for all $x \in [a, b]$, for sufficiently large n .

Proof. We will prove Eq. (50) and explain how Eq. (51) can be proven in an analogous way. Consider the function

$L(y) = \log \frac{y}{1-y}$, which is strictly increasing for $y \in (0, 1)$. Next, define two sequences of functions

$$f_n(x) := \frac{2V(\mathbf{p})}{nt_n^2} L \left(\sum_{i \leq K_n(x)} P_i^n \right), \quad (52a)$$

$$g_n(x) := \frac{2V(\mathbf{p})}{nt_n^2} L \left(\sum_{i \leq K_n(x)} Q_i^n \right). \quad (52b)$$

We can combine the direct and converse parts of Lemma 8, i.e., Eqs. (30a) and (30b), to obtain the limits $f_n \rightarrow f$ and $g_n \rightarrow g$ as $n \rightarrow \infty$, where

$$f(x) = \frac{(x - \mu)^3}{|x - \mu|} \quad \text{and} \quad g(x) = \frac{\nu x^3}{|x|}. \quad (53a)$$

We now find that

$$f(x) > g(x) \iff \begin{cases} x > z_C & \text{for } \nu < 1, \\ x < z_C & \text{for } \nu > 1. \end{cases} \quad (54)$$

Therefore, for any x in the above regions we have that $f_n(x) \stackrel{\text{ev.}}{>} g_n(x)$. As L is strictly monotone, this in turn implies that Eq. (50) holds eventually in the same region.

The above argument only ensures that $f_n \rightarrow f$ and $g_n \rightarrow g$ point-wise. This does not yet allow us to conclude that there exists an N such that Eq. (50) will hold for all $x \in [a, b]$ and all $n \geq N$. In Appendix B we close this gap by proving that this convergence is compact.

Finally, Eq. (51) can be proven in an analogous way by substituting: $g_n \rightarrow -g_n$ and $g \rightarrow -g$. \square

V. PROOF OF THE MAIN THEOREM

We are now ready to prove Theorems 1, 2 and 3. We will achieve this in a series of steps. First, we will prove the following result.

Proposition 12 (Thermodynamic direction, $\gamma = \eta$, TVD). *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \quad (55)$$

the approximation majorisation condition

$$P^n \succ_{\epsilon_n^\pm} Q^n \quad (56)$$

with $\mathbf{f} = \eta$ holds with an optimal interconversion rate of

$$R_n^{\text{th}}(\epsilon_n^\pm) \simeq \frac{D(\mathbf{p} \parallel \eta) \pm \sqrt{2V(\mathbf{p} \parallel \eta)} \left| 1 \pm 1/\sqrt{\nu^{\text{th}}} \right| t_n}{D(\mathbf{q} \parallel \eta)}. \quad (57)$$

The proof of Proposition 12 will consist of two parts: first, in Sec. VA, we will show that the claimed rate is achievable for the given error; and then, in Sec. VB, that it is also optimal. This way we will prove a special case of Theorem 2 for the case of infinite temperature

(when $\gamma = \eta$) and error level measured only by TVD; and Theorem 3 for the thermodynamic direction and infinite temperature.

The next step is to generalise Proposition 12 to arbitrary finite temperatures (arbitrary thermal state γ). It is enough to note that the approximate interconversion condition for the thermodynamic direction, $\hat{\mathbf{P}}^n \succ_\epsilon \hat{\mathbf{Q}}^n$, is exactly captured by Eq. (56) if one only replaces \mathbf{p} and \mathbf{q} with $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$, respectively. Moreover, since the relative entropy and relative entropy variance are invariant under the embedding map [8], one can obtain the optimal rate by ‘‘unembedding’’ Eq. (57), i.e., replacing $\hat{\mathbf{p}}$, $\hat{\mathbf{q}}$ and $\eta = \hat{\gamma}$ with \mathbf{p} , \mathbf{q} and γ respectively. Thus, by proving Proposition 12, we in fact prove Theorem 2 for any temperature and the error level measured by TVD; and Theorem 3 for the thermodynamic direction with arbitrary temperature.

Then, in Sec. VC, we will prove the following result

Proposition 13 (Entanglement direction, TVD). *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \quad (58)$$

the approximation majorisation condition

$$\mathbf{P}^n \prec_{\epsilon_n^\pm} \mathbf{Q}^n \quad (59)$$

with $\mathbf{f} = \mathbf{s}$ holds with an optimal interconversion rate of

$$R_n^{\text{ent}}(\epsilon_n^\pm) \simeq \frac{H(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu^{\text{ent}}}| t_n}{H(\mathbf{q})}. \quad (60)$$

To prove Proposition 13 we will leverage the proof of Proposition 12. More precisely, we will explain how to adapt that proof, so that the thermodynamic direction gets replaced by the entanglement direction. This way we will prove Theorem 1 for the error level measured by TVD; and Theorem 3 for the entanglement direction.

The final missing piece is to show that Theorems 1 and 2 also hold for the error level measured by infidelity. We will prove this in Sec. VD, again by explaining the necessary modifications of the reasoning that will result in replacing TVD with infidelity distance.

A. Proof of Proposition 12 (Achievability)

We will start by considering the *achievability* of Proposition 12, i.e., a lower bound on the optimal conversion rate for the thermodynamic direction. For notational convenience, we will drop the superscripts on both R_n^{th} and ν^{th} , adopt the convention $D(\cdot) := D(\cdot \| \eta)$, and note that $V(\cdot \| \eta) = V(\cdot)$. Specifically, we will prove the following:

Lemma 14 (Proposition 12: Achievability). *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \quad (61)$$

the optimal rate is lower bounded,

$$R_n(\epsilon_n^\pm) \gtrsim \frac{D(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}. \quad (62)$$

We will prove this lemma by constructing a family of distributions $\tilde{\mathbf{P}}^{(\mu)}$, which eventually obey the required majorisation condition $\tilde{\mathbf{P}}^{(\mu)} \succ \mathbf{Q}^n$. We will then show that by picking appropriate values of $\mu = \mu_+$ and $\mu = \mu_-$, one can obtain distributions $\tilde{\mathbf{P}}^{(\mu_\pm)}$ such that $\delta(\mathbf{P}^n, \tilde{\mathbf{P}}^{(\mu_\pm)}) \stackrel{\text{ev.}}{\leq} \epsilon_n^\pm$.

1. Constructing the approximate distribution $\tilde{\mathbf{P}}^{(\mu)}$

To prove achievability, we will construct a family of distributions $\tilde{\mathbf{P}}^{(\mu)}$ which, for any fixed μ , eventually majorise \mathbf{Q}^n . As in Lemma 10, consider the rate

$$r_n(\mu) = \frac{D(\mathbf{p}) + \mu t_n}{D(\mathbf{q})}, \quad (63)$$

where $\mu \in \mathbb{R}$ is a parameter of our construction. We will construct $\tilde{\mathbf{P}}^{(\mu)}$ by the *cut-and-pile* method. Specifically we will consider starting with \mathbf{P}^n , removing mass from its tail, and adding it to the largest element. This construction allows us to construct a nearby state which is higher in the majorisation order. We start by defining the cutting point,

$$z_{\mu,\nu} := \begin{cases} 2\mu - z_C & : \mu < 0, \quad \nu < 1, \\ z_C & : \mu < 0, \quad \nu > 1, \\ z_T & : \mu > 0. \end{cases} \quad (64)$$

If we let $\zeta > 0$ be a small slack parameter, then $\tilde{\mathbf{P}}^{(\mu)}$ is defined as

$$\tilde{P}_i^{(\mu)} := \begin{cases} P_1^n + \sum_{i \geq K_n(y)} P_i^n & : i = 1, \\ P_i^n & : 1 < i < K_n(y), \\ 0 & : i \geq K_n(y), \end{cases} \quad (65)$$

with $y = z_{\mu,\nu} - \zeta$.

2. Showing majorisation $\tilde{\mathbf{P}}^{(\mu)} \succ \tilde{\mathbf{Q}}^n$

Given the above construction, we now want to prove the majorisation condition $\tilde{\mathbf{P}}^{(\mu)} \succ \mathbf{Q}^n$ eventually holds. The idea here is to leverage Lemma 11, and show that a cut-and-pile construction with a cut at $K_n(x)$ for any $x < z_{\mu,\nu}$ will always eventually majorise \mathbf{Q}^n .

Lemma 15. *For any fixed μ , $\tilde{\mathbf{P}}^{(\mu)} \stackrel{\text{ev.}}{\succ} \mathbf{Q}^n$.*

Proof. To prove majorisation we need to show that, eventually, the inequalities

$$\sum_{i=1}^k \tilde{P}_i^{(\mu)} \geq \sum_{i=1}^k Q_i^n \quad (66)$$

hold for all k . The ‘cut’ of the cut-and-pile construction implies Eq. (66) for large k . Specifically, the restricted support of $\tilde{\mathbf{P}}^{(\mu)}$,

$$\sum_{i < K_n(z_{\mu, \nu} - \zeta)} \tilde{P}_i^{(\mu)} = 1, \quad (67)$$

implies that Eq. (66) holds trivially for any $k \geq K_n(z_{\mu, \nu} - \zeta)$.

The idea now is to show that the ‘pile’ similarly gives us majorisation for small k , and then to leverage Lemma 11 to argue that \mathbf{P}^n already majorises \mathbf{Q}^n for intermediate k . We will split this argument into the three cases given in the definition of $z_{\mu, \nu}$, and illustrated by panels (a)-(c) of Fig. 1.

Case 1: $\mu < 0, \nu < 1$. Noticing that \mathbf{P}^n tail bounds in Lemma 10 are symmetric under $x \rightarrow 2\mu - x$, we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_{\mu, \nu} - \zeta)} P_i^n \right] \\ = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(z_C + \zeta)} P_i^n \right]. \end{aligned} \quad (68)$$

Applying Lemma 11, we therefore have that

$$\sum_{i \geq K_n(z_{\mu, \nu} - \zeta)} P_i^n \stackrel{\text{ev.}}{>} \sum_{i \leq K_n(z_C + \zeta)} Q_i^n. \quad (69)$$

Using this, for any $k \leq K_n(z_C + \zeta)$ we can leverage the ‘pile’ in the $\tilde{\mathbf{P}}^{(\mu)}$ construction to yield

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} \geq \tilde{P}_1^{(\mu)} > \sum_{i \geq K_n(z_{\mu, \nu} - \zeta)} P_i^n \stackrel{\text{ev.}}{>} \sum_{i \leq K_n(z_C + \zeta)} Q_i^n \geq \sum_{i \leq k} Q_i^n, \quad (70)$$

which implies Eq. (66). Applying Lemma 11, we have that Eq. (66) must also eventually hold on the remaining intermediate indices $k \in [K_n(z_C + \zeta), K_n(z_{\mu, \nu} - \zeta)]$.

Case 2: $\mu < 0, \nu > 1$. Noticing that \mathbf{Q}^n tail bounds in Lemma 10 are symmetric under $x \rightarrow -x$, we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \geq K_n(z_{\mu, \nu})} Q_i^n \right] \\ = \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\sum_{i \leq K_n(-z_C)} Q_i^n \right]. \end{aligned} \quad (71)$$

We now use Lemma 11 again, giving for any $k \leq K_n(-z_C)$ that the ‘pile’ of $\tilde{\mathbf{P}}^{(\mu)}$ implies Eq. (66),

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} > \sum_{i \geq K_n(z_{\mu, \nu} - \zeta)} P_i^n \stackrel{\text{ev.}}{>} \sum_{i \leq K_n(-z_C)} Q_i^n \geq \sum_{i \leq k} Q_i^n. \quad (72)$$

A direct application of Eq. (11) implies Equation (66) also holds on the remaining intermediate indices $k \in [K_n(-z_C), K_n(z_{\mu, \nu} - \zeta)]$.

Case 3: $\mu > 0$. For $k \leq K_n(z_T - \zeta)$, we use Lemma 11 to eventually give

$$\sum_{i \leq k} \tilde{P}_i^{(\mu)} > \sum_{i \geq K_n(z_T - \zeta)} P_i^n \stackrel{\text{ev.}}{>} \sum_{i \leq K_n(z_T - \zeta)} Q_i^n \geq \sum_{i \leq k} Q_i^n. \quad (73)$$

In this case there is no intermediate region, which implies Eq. (66) holds for all k . \square

3. Showing that $\tilde{\mathbf{P}}^{(\mu)}$ is close to \mathbf{P}^n

Now that we have shown that $\tilde{\mathbf{P}}^{(\mu)}$ eventually majorises \mathbf{Q}^n , we want to ask how close $\tilde{\mathbf{P}}^{(\mu)}$ is to \mathbf{P}^n in total variation distance. The answer is provided by the following lemma.

Lemma 16. For a fixed $\mu < 0$,

$$\delta(\tilde{\mathbf{P}}^{(\mu)}, \mathbf{P}^n) \stackrel{\text{ev.}}{<} \exp \left(-\frac{(z_{\mu, \nu} - \mu - 2\zeta)^2}{2V(\mathbf{p})} nt_n^2 \right), \quad (74)$$

and similarly for $\mu > 0$,

$$\delta(\tilde{\mathbf{P}}^{(\mu)}, \mathbf{P}^n) \stackrel{\text{ev.}}{<} 1 - \exp \left(-\frac{(z_{\mu, \nu} - \mu - 2\zeta)^2}{2V(\mathbf{p})} nt_n^2 \right). \quad (75)$$

Proof. A convenient feature of the cut-and-pile construction is that the total variation distance to the original distribution takes a particularly simple form. Specifically,

$$\delta(\tilde{\mathbf{P}}^{(\mu)}, \mathbf{P}^n) = \sum_{i \geq K_n(z_{\mu, \nu} - \zeta)} P_i^n. \quad (76)$$

We can now apply Lemma 10 to give

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\delta(\tilde{\mathbf{P}}^{(\mu < 0)}, \mathbf{P}^n) \right] = -\frac{(z_{\mu, \nu} - \mu - \zeta)^2}{2V(\mathbf{p})}, \quad (77a)$$

$$\lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[1 - \delta(\tilde{\mathbf{P}}^{(\mu > 0)}, \mathbf{P}^n) \right] = -\frac{(z_{\mu, \nu} - \mu - \zeta)^2}{2V(\mathbf{p})}. \quad (77b)$$

This, in turn, implies the eventual inequalities

$$\ln \left[\delta(\tilde{\mathbf{P}}^{(\mu < 0)}, \mathbf{P}^n) \right] \stackrel{\text{ev.}}{<} -\frac{(z_{\mu, \nu} - \mu - 2\zeta)^2}{2V(\mathbf{p})} nt_n^2, \quad (78a)$$

$$\ln \left[1 - \delta(\tilde{\mathbf{P}}^{(\mu > 0)}, \mathbf{P}^n) \right] \stackrel{\text{ev.}}{>} -\frac{(z_{\mu, \nu} - \mu - 2\zeta)^2}{2V(\mathbf{p})} nt_n^2, \quad (78b)$$

which are equivalent to our desired bounds. \square

4. Proof of Achievability

We now put together the above to prove the achievability of Proposition 12.

Proof of Lemma 14. Consider our construction of $\tilde{\mathbf{P}}^{(\mu)}$ for a specific choice of μ . Specifically, let

$$\mu_{\pm} = |1 \pm 1/\sqrt{\nu}| \left(\pm \sqrt{2V(\mathbf{p})} - 2\zeta \right), \quad (79)$$

where μ_- will give the direct result (with vanishing error ϵ_n^-), and μ_+ will give the converse result (with error ϵ_n^+ approaching 1).

From Lemma 15 we have that $\tilde{\mathbf{P}}^{(\mu_{\pm})} \stackrel{\text{ev.}}{\succ} \mathbf{Q}^n$ as required. Substituting μ_{\pm} in Lemma 16, we have that the TVD error is bounded,

$$\delta \left(\tilde{\mathbf{P}}^{(\mu_{\pm})}, \mathbf{P}^n \right) \stackrel{\text{ev.}}{<} \epsilon_n^{\pm}, \quad (80)$$

as required. The rate in these cases takes the form

$$\begin{aligned} r_n(\mu_{\pm}) &= \frac{D(\mathbf{p}) + \mu_{\pm} t_n}{D(\mathbf{q})} \\ &= \frac{D(\mathbf{p}) + |1 \pm 1/\sqrt{\nu}| \left(\pm \sqrt{2V(\mathbf{p})} - 2\zeta \right) t_n}{D(\mathbf{q})}. \end{aligned} \quad (81)$$

As the above analysis holds for any $\zeta > 0$, we can therefore conclude that

$$R_n(\epsilon_n^{\pm}) \gtrsim \frac{D(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}, \quad (82)$$

as desired. \square

B. Proof of Proposition 12 (Optimality)

We now move on to showing the *optimality* of Proposition 12, i.e. an upper bound on the optimal conversion rate for the thermodynamic direction. For convenience we will reuse the notation used in the proof of achievability.

Lemma 17 (Proposition 12: Optimality). *For a TVD error level of*

$$\epsilon_n^- = e^{-nt_n^2} \quad \text{or} \quad \epsilon_n^+ = 1 - e^{-nt_n^2}, \quad (83)$$

the optimal rate is upper bounded,

$$R_n(\epsilon_n^{\pm}) \lesssim \frac{D(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}. \quad (84)$$

We will prove the above by showing that, for any distribution $\tilde{\mathbf{P}}^n$ obeying the majorisation condition $\tilde{\mathbf{P}}^n \succ \mathbf{Q}^n$ with a rate $r_n(\mu)$, the TVD distance between $\tilde{\mathbf{P}}^n$ and \mathbf{P}^n is eventually lower bounded by ϵ_n^{\pm} . To achieve this, we will use the fact that the total variation distance is monotonically decreasing under coarse-graining. Specifically, for any distributions \mathbf{a} and \mathbf{b} , and index k , the triangle inequality gives

$$\delta(\mathbf{a}, \mathbf{b}) \geq \left| \sum_{i \leq k} a_i - \sum_{j \leq k} b_j \right|, \quad (85)$$

Applying this to distributions $\tilde{\mathbf{P}}^n$ and \mathbf{P}^n , and index $k = K_n(x)$, gives

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \geq \left| \sum_{i \leq K_n(x)} \tilde{P}_i^n - \sum_{j \leq K_n(x)} P_j^n \right|. \quad (86)$$

The idea is that, with a careful choice of x , we will be able to use the majorisation $\tilde{\mathbf{P}}^n \succ \mathbf{Q}^n$ to replace the summations over $\tilde{\mathbf{P}}^n$ with those over \mathbf{Q}^n . This will then allow us to apply Lemma 10 to arrive at our final bound on the error.

We will first present this argument in detail for the case where $\mu < 0$ and $\nu < 1$, and then present the modifications required for the remaining cases of $\mu < 0$ and $\nu > 1$, and $\mu > 0$.

1. Case 1: $\mu < 0, \nu < 1$

Here we will perform our coarse-grained binning at $x = z_C - \zeta$. Recalling that $\tilde{\mathbf{P}}^n \succ \mathbf{Q}^n$, we have that

$$\sum_{i \leq K_n(z_C - \zeta)} \tilde{P}_i^n \geq \sum_{i \leq K_n(z_C - \zeta)} Q_i^n. \quad (87)$$

Using the positivity of ζ , Lemma 11 allows us to also conclude that

$$\sum_{i \leq K_n(z_C - \zeta)} \tilde{P}_i^n \stackrel{\text{ev.}}{>} \sum_{i \leq K_n(z_C - \zeta)} P_i^n. \quad (88)$$

We now use the fact that $|\alpha - \beta|$ is monotonically increasing in α for $\alpha \geq \beta$. Using this, we have an eventual lower bound on TVD purely in terms of the total states \mathbf{P}^n and \mathbf{Q}^n ,

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \geq \left| \sum_{i \leq K_n(z_C - \zeta)} Q_i^n - \sum_{j \leq K_n(z_C - \zeta)} P_j^n \right|. \quad (89)$$

Applying Lemma 10, we see that the tail sum of \mathbf{Q}^n asymptotically dominates. Specifically, we have

$$\liminf_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \right] \geq -\frac{\nu(z_C - \zeta)^2}{2V(\mathbf{p})}, \quad (90)$$

and therefore

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \stackrel{\text{ev.}}{>} \exp \left(-\frac{\nu(z_C - \zeta)^2}{2V(\mathbf{p})} nt_n^2 \right). \quad (91)$$

We now choose

$$\mu = |1 - 1/\sqrt{\nu}| \left(-\sqrt{2V(\mathbf{p})} + 2\zeta\sqrt{\nu} \right). \quad (92)$$

This gives us that $\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \stackrel{\text{ev.}}{>} \epsilon_n^-$, with a rate of

$$r_n(\mu) = \frac{D(\mathbf{p}) + |1 - 1/\sqrt{\nu}| \left(-\sqrt{2V(\mathbf{p})} + 2\zeta\sqrt{\nu} \right) t_n}{D(\mathbf{q})}. \quad (93)$$

As this is true for all $\zeta > 0$, we can conclude that

$$R_n(\epsilon_n^-) \lesssim \frac{D(\mathbf{p}) - \sqrt{2V(\mathbf{p})} |1 - 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}. \quad (94)$$

2. *Case 2: $\mu < 0, \nu > 1$*

Now we consider the case of $\mu < 0$ and $\nu > 1$. The proof here is similar, starting with a cut at $x = z_C + \zeta$. Here the tail sum of \mathbf{P}^n asymptotically dominates, with Lemma 10 giving

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \stackrel{\text{ev.}}{>} \exp\left(-\frac{(z_C - \mu + 2\zeta)^2 n t_n^2}{2V(\mathbf{p})}\right). \quad (95)$$

We now make the choice

$$\mu = |1 - 1/\sqrt{\nu}| \left(-\sqrt{2V(\mathbf{p})} + 2\zeta\right), \quad (96)$$

which gives that $\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \stackrel{\text{ev.}}{>} \epsilon_n^-$, and thus

$$R_n(\epsilon_n^-) \lesssim \frac{D(\mathbf{p}) - \sqrt{2V(\mathbf{p})} |1 - 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}. \quad (97)$$

3. *Case 3: $\mu > 0$*

Lastly, we consider the case of $\mu > 0$. Here, we perform our cut at $x = z_T$ and, instead of using an argument based on majorisation *from above*, we use majorisation *from below*. Specifically, we note that $\tilde{\mathbf{P}}^n \succ \mathbf{Q}^n$ implies

$$\sum_{i > K_n(z_T)} \tilde{P}_i^n \leq \sum_{i > K_n(z_T)} Q_i^n. \quad (98)$$

We now use that $|\alpha - \beta|$ is monotonically *decreasing* in α for $\alpha \leq \beta$. This gives us the analogous upper bound

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \geq \left| \sum_{i \leq K_n(z_T)} Q_i^n - \sum_{j \leq K_n(z_T)} P_j^n \right|. \quad (99)$$

Applying Lemma 10, we find that the tail sums of \mathbf{P}^n and \mathbf{Q}^n both dominate¹, yielding

$$\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) \stackrel{\text{ev.}}{>} 1 - \exp\left(-\frac{(z_T - \mu + \zeta)^2 n t_n^2}{2V(\mathbf{p})}\right). \quad (100)$$

for sufficiently large n . Substituting

$$\mu = |1 + 1/\sqrt{\nu}| \left(\sqrt{2V(\mathbf{p})} + \zeta\right), \quad (101)$$

we eventually have that $\delta(\tilde{\mathbf{P}}^n, \mathbf{P}^n) > \epsilon_n^+$, and thus

$$R_n(\epsilon_n^+) \lesssim \frac{D(\mathbf{p}) + \sqrt{2V(\mathbf{p})} |1 + 1/\sqrt{\nu}| t_n}{D(\mathbf{q})}. \quad (102)$$

¹ The lack of a slack parameter in the cut means that these two sums compete. In the direct case (vanishing error ϵ_n^-) this leads to a catastrophic cancellation, but in the converse case, this causes no problem.

C. **Proof of Proposition 13**

Proposition 12 states that the largest rate R_n such that the majorisation condition

$$\mathbf{p}^{\otimes n} \otimes \boldsymbol{\eta}^{\otimes n R_n} \succ_{\epsilon_n^\pm} \mathbf{q}^{\otimes n R_n} \otimes \boldsymbol{\eta}^{\otimes n} \quad (103)$$

holds is of the form

$$R_n(\epsilon_n^\pm) \simeq \frac{H(\boldsymbol{\eta}) - H(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}| t_n}{H(\boldsymbol{\eta}) - H(\mathbf{q})}, \quad (104)$$

where

$$\nu = \frac{V(\mathbf{p})}{V(\mathbf{q})} \cdot \frac{H(\boldsymbol{\eta}) - H(\mathbf{q})}{H(\boldsymbol{\eta}) - H(\mathbf{p})}. \quad (105)$$

The proof of the above relied entirely on Lemma 10, which holds for $\mathbf{f} = \boldsymbol{\eta}$ as well as $\mathbf{f} = \mathbf{s}$. Thus, in order to rigorously prove Proposition 13, one could perform steps analogous to the ones presented in Sec. V A-V B, with $\boldsymbol{\eta}$ replaced by \mathbf{s} . Instead, below we present a shorter proof that directly employs the result stated by Proposition 12.

Proof of Proposition 13. Since Lemma 10 holds both for $\mathbf{f} = \boldsymbol{\eta}$ and for $\mathbf{f} = \mathbf{s}$, the statement of Proposition 12 can be extended to show that the smallest R_n such that the majorisation condition

$$\mathbf{p}^{\otimes n} \otimes \mathbf{s}^{\otimes n R_n} \succ_{\epsilon_n^\pm} \mathbf{q}^{\otimes n R_n} \otimes \mathbf{s}^{\otimes n} \quad (106)$$

holds is of the form

$$R_n(\epsilon_n^\pm) \simeq \frac{H(\mathbf{s}) - H(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}| t_n}{H(\mathbf{s}) - H(\mathbf{q})}, \quad (107)$$

where

$$\nu = \frac{V(\mathbf{p})}{V(\mathbf{q})} \cdot \frac{H(\mathbf{s}) - H(\mathbf{q})}{H(\mathbf{s}) - H(\mathbf{p})}. \quad (108)$$

We now want to reverse the direction of majorisation. The first step is simply to swap $\mathbf{p} \leftrightarrow \mathbf{q}$, and use the fact ϵ -post-majorisation is equivalent to ϵ -pre-majorisation [8]. This transforms the considered majorisation relation into the following form

$$\mathbf{p}^{\otimes n R_n} \otimes \mathbf{s}^{\otimes n} \prec_{\epsilon_n^\pm} \mathbf{q}^{\otimes n} \otimes \mathbf{s}^{\otimes n R_n}. \quad (109)$$

We see that R_n now forms the ‘inverse rate’ of the transformation between states represented by \mathbf{p} and the ones represented by \mathbf{q} . To find the true rate of this transformation, we make the following substitutions

$$n \leftarrow n R_n, \quad t_n \leftarrow t_n / \sqrt{R_n}, \quad R_n \leftarrow 1/R_n. \quad (110)$$

As a result, the desired majorisation condition,

$$\mathbf{p}^{\otimes n} \otimes \mathbf{s}^{\otimes n R_n} \prec_{\epsilon_n^\pm} \mathbf{q}^{\otimes n R_n} \otimes \mathbf{s}^{\otimes n} \quad (111)$$

holds with the optimal rate of the form

$$\begin{aligned} R_n(\epsilon_n^\pm) &\simeq \frac{H(\mathbf{p})}{H(\mathbf{q}) \mp \sqrt{2V(\mathbf{q}) \frac{H(\mathbf{q})}{H(\mathbf{p})}} |1 \pm \sqrt{\nu}| t_n} \\ &\simeq \frac{H(\mathbf{p}) \pm \sqrt{2V(\mathbf{p})} |1 \pm 1/\sqrt{\nu}|}{H(\mathbf{q})}, \end{aligned} \quad (112)$$

and

$$\nu = \frac{V(\mathbf{p})}{V(\mathbf{q})} \cdot \frac{H(\mathbf{q})}{H(\mathbf{p})}. \quad (113)$$

□

D. Extension to infidelity

We now want to argue that, in the direct regime (for vanishing error ϵ_n^-), our results extend to the case where we consider error in terms of infidelity instead of TVD. To show that the achievability argument, presented in Sec. VA, extends to infidelity, we leverage the Fuchs-van de Graaf inequality [30], specifically

$$1 - \sqrt{F(\mathbf{P}^n, \tilde{\mathbf{P}}^n)} \leq \delta(\mathbf{P}^n, \tilde{\mathbf{P}}^n). \quad (114)$$

Using this, in the direct regime where $\delta(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \rightarrow 0$, we have that the corresponding moderate exponential of infidelity must be bounded by that of the TVD,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[1 - F(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \right] \\ \leq \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\delta(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \right]. \end{aligned} \quad (115)$$

Therefore, the TVD upper bounds employed in the proof of achievability naturally extend to the infidelity.

In the proof of optimality, presented in Sec. VB, we used coarse-graining and a monotonicity argument to bound the TVD as

$$\delta(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \geq \left| \sum_{i \leq k} P_i^n - \sum_{j \leq k} Q_j^n \right|. \quad (116)$$

By using the data-processing inequality and an analogous monotonicity argument, we can similarly bound the fidelity as

$$\begin{aligned} \sqrt{F(\mathbf{P}^n, \tilde{\mathbf{P}}^n)} &\leq \sqrt{\sum_{i \leq k} P_i^n \cdot \sum_{j \leq k} Q_j^n} \\ &\quad + \sqrt{\sum_{i > k} P_i^n \cdot \sum_{j > k} Q_j^n}. \end{aligned} \quad (117)$$

In the case where $\nu < 1$, we chose k such that

$$\sum_{i \leq k} P_i^n \ll \sum_{i \leq k} Q_i^n \ll 1, \quad (118)$$

and similarly for $\nu > 1$

$$\sum_{i > k} Q_i^n \ll \sum_{i > k} P_i^n \ll 1. \quad (119)$$

In either case, a single tail sum dominates in the bound upon both fidelity and TVD, similarly allowing us to lower bound the moderate exponent of the infidelity,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[1 - F(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \right] \\ \geq \lim_{n \rightarrow \infty} \frac{1}{nt_n^2} \ln \left[\delta(\mathbf{P}^n, \tilde{\mathbf{P}}^n) \right]. \end{aligned} \quad (120)$$

VI. OUTLOOK

We have performed the moderate deviation analysis of resource interconversion problems for which single-shot transformation rules are based on majorisation and thermo-majorisation. As a result, in the regime of asymptotically vanishing error, we have found unified expressions for second-order corrections to asymptotic conversion rates within resource theories of entanglement, coherence and thermodynamics. More precisely, we obtained a family of results that specifies the optimal trade-off between the speed at which the conversion rate approaches the asymptotic rate, and the speed at which the error vanishes, when the number of transformed states n grows. Crucially, we have found that the correction term can vanish independently of n when a certain resonance condition between the initial and final states is satisfied. This opens the path to transformation reversibility beyond the asymptotic limit, the phenomenon that we discuss in detail in the accompanying paper [18].

There are quite a few research directions that one may want to take in order to generalise and extend the results presented in this paper. Since the small deviation analysis of the majorisation-based resource interconversion has been performed in Refs. [7, 8], and the moderate deviation analysis was the focus of the current work, the straightforward extension would be to investigate the interconversion problem in the large deviation regime. This may be of particular interest in the context of fluctuation-free work extraction, where one may want to sacrifice a constant fraction of possible work output for its quality [31]. On the other hand, one could also look for the exact expression for the third-order term of the asymptotic expansion of R_n , which one can conjecture to scale as $O(t_n^2 + \log n)$ [11]. Another obvious generalisation is to go beyond the restrictions of pure states (for entanglement direction) and energy-incoherent states (for thermodynamic direction). This, however, is a much harder task, as the single-shot transformation rules are still not known for these unrestricted cases. Finally, since the second-order analysis of resource interconversion led us not only to quantitative results, but also to qualitatively new predictions concerning finite-size reversibility, a sim-

ilar analysis for other resource theories is now very well justified.

Acknowledgements

CC and KK acknowledge support from the ARC via the Centre of Excellence in Engineered Quantum Sys-

tems (EQuS), project number CE110001013. CC also acknowledges support from the Australian Institute for Nanoscale Science and Technology Postgraduate Scholarship (John Makepeace Bennett Gift). MT is funded by an ARC Discovery Early Career Researcher Award, project number DE160100821.

-
- [1] Bob Coecke, Tobias Fritz, and Robert W Spekkens, “A mathematical theory of resources,” *Inform. Comput.* **250**, 59–86 (2016).
- [2] E. Chitambar and G. Gour, “Quantum Resource Theories,” [arXiv:1806.06107](https://arxiv.org/abs/1806.06107) (2018).
- [3] Charles H Bennett, Herbert J Bernstein, Sandu Popescu, and Benjamin Schumacher, “Concentrating partial entanglement by local operations,” *Phys. Rev. A* **53**, 2046 (1996).
- [4] Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens, “Resource theory of quantum states out of thermal equilibrium,” *Phys. Rev. Lett.* **111**, 250404 (2013).
- [5] Michael A Nielsen, “Conditions for a class of entanglement transformations,” *Phys. Rev. Lett.* **83**, 436 (1999).
- [6] F. G. S. L. Brandão, M. Horodecki, N. H. Y. Ng, J. Oppenheim, and S. Wehner, “The second laws of quantum thermodynamics,” *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3275 (2015).
- [7] Wataru Kumagai and Masahito Hayashi, “Second-order asymptotics of conversions of distributions and entangled states based on Rayleigh-normal probability distributions,” *IEEE Trans. Inf. Theory* **63**, 1829–1857 (2017).
- [8] Christopher T Chubb, Marco Tomamichel, and Kamil Korzekwa, “Beyond the thermodynamic limit: finite-size corrections to state interconversion rates,” [arXiv:1711.01193](https://arxiv.org/abs/1711.01193) (2017).
- [9] V Strassen, “Asymptotische Abschätzungen in Shannons Informationstheorie,” in *Trans. Third Prague Conf. Inf. Theory* (Prague, 1962) pp. 689–723.
- [10] Yucel Altug and Aaron B. Wagner, “Moderate Deviations in Channel Coding,” *IEEE Trans. Inf. Theory* **60**, 4417–4426 (2014).
- [11] C.T. Chubb, V.Y.F. Tan, and M. Tomamichel, “Moderate Deviation Analysis for Classical Communication over Quantum Channels,” *Commun. Math. Phys.* **355** (2017).
- [12] Hao-Chung Cheng and Min-Hsiu Hsieh, “Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing,” (2017), [arXiv:1701.03195](https://arxiv.org/abs/1701.03195).
- [13] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.* **81**, 865–942 (2009).
- [14] T Baumgratz, M Cramer, and MB Plenio, “Quantifying coherence,” *Phys. Rev. Lett.* **113**, 140401 (2014).
- [15] M. Horodecki and J. Oppenheim, “Fundamental limitations for quantum and nanoscale thermodynamics,” *Nat. Commun.* **4**, 2059 (2013).
- [16] Paul G Kwiat, Salvador Barraza-Lopez, Andre Stefanov, and Nicolas Gisin, “Experimental entanglement distillation and hidden non-locality,” *Nature* **409**, 1014 (2001).
- [17] Andreas Winter and Dong Yang, “Operational resource theory of coherence,” *Phys. Rev. Lett.* **116**, 120404 (2016).
- [18] Kamil Korzekwa, Christopher T Chubb, and Marco Tomamichel, “Avoiding irreversibility: engineering resonant conversions of quantum resources,” (In preparation).
- [19] Hiroyasu Tajima and Masahito Hayashi, “Finite-size effect on optimal efficiency of heat engines,” *Phys. Rev. E* **96**, 012128 (2017).
- [20] Stephen D Bartlett, Terry Rudolph, and Robert W Spekkens, “Reference frames, superselection rules, and quantum information,” *Rev. Mod. Phys.* **79**, 555 (2007).
- [21] Gilad Gour, “Quantum resource theories in the single-shot regime,” *Phys. Rev. A* **95**, 062314 (2017).
- [22] Fernando GSL Brandão and Gilad Gour, “Reversible framework for quantum resource theories,” *Phys. Rev. Lett.* **115**, 070503 (2015).
- [23] John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk, “The role of quantum information in thermodynamics – a topical review,” *J. Phys. A* **49**, 143001 (2016).
- [24] Amir Dembo and Ofer Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Stochastic Modelling and Applied Probability (Springer, 1998).
- [25] Shuanping Du, Zhaofang Bai, and Yu Guo, “Conditions for coherence transformations under incoherent operations,” *Phys. Rev. A* **91**, 052120 (2015).
- [26] Albert W Marshall, Ingram Olkin, and Barry C Arnold, *Inequalities: theory of majorization and its applications*, Vol. 143 (Springer, 1979).
- [27] Ernst Ruch, Rudolf Schraner, and Thomas H. Seligman, “The mixing distance,” *J. Chem. Phys.* **69** (1978).
- [28] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, “Thermodynamic cost of reliability and low temperatures: tightening Landauer’s principle and the second law,” *Int. J. Theor. Phys.* **39**, 2717–2753 (2000).
- [29] Matteo Lostaglio, “Thermodynamic laws for populations and quantum coherence: A self-contained introduction to the resource theory approach to thermodynamics,” [arXiv:1807.11549](https://arxiv.org/abs/1807.11549) (2018).
- [30] Christopher A Fuchs and Jeroen Van De Graaf, “Cryptographic distinguishability measures for quantum-mechanical states,” *IEEE Trans. Inf. Theory* **45**, 1216–1227 (1999).
- [31] Johan Åberg, “Truly work-like work extraction via a single-shot analysis,” *Nat. Commun.* **4**, 1925 (2013).
- [32] Sidney I Resnick, *Extreme values, regular variation and point processes* (Springer, 2013).

Appendix A: Relation to small deviation bound

In this section we explore the relationship between our moderate deviation results of Theorems 1–3, and the small deviation results of Refs. [7, 8]. In these papers, a second-order expansion of the rate is given for a *constant* infidelity error, in terms of the Rayleigh-normal distribution. Below we will consider the expansions of the Rayleigh-normal cumulative distribution function around $-\infty$, and show consistency with Theorems 1 and 2. Moreover, we will explain how the expansion around $+\infty$ leads to a conjecture analogous to Theorem 3 with error level measured by infidelity.

The Rayleigh-normal distributions are a parameterised family of distributions, depending on a parameter $\nu \geq 0$. For a formal definition and properties, see Ref. [7]. We will denote the associated cumulative distribution functions by $Z_\nu(\mu)$, for $\nu \geq 0$ and $\mu \in \mathbb{R}$. We will also adopt the notation for Gaussian probability and cumulative distribution functions of

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad (\text{A1})$$

$$\phi_{\mu,\nu}(x) = \frac{1}{\sqrt{2\pi\nu}} e^{-(x-\mu)^2/2\nu}, \quad (\text{A2})$$

$$\Phi(x) = \int_{-\infty}^x \phi(t) dt, \quad (\text{A3})$$

$$\Phi_{\mu,\nu}(x) = \int_{-\infty}^x \phi_{\mu,\nu}(t) dt. \quad (\text{A4})$$

1. Expansion around $\mu = -\infty$

The crossing point $\alpha_{\mu,\nu}$ is defined by the equation [7]

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} = \frac{\Phi(\alpha_{\mu,\nu})}{\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \quad (\text{A5})$$

As $\mu \rightarrow -\infty$, $\alpha_{\mu,\nu} \rightarrow +\infty$. We will use the $x \rightarrow \infty$ approximation $\Phi(x) \approx 1$, which leads to $\phi(\alpha_{\mu,\nu})/\phi_{\mu,\nu}(\alpha_{\mu,\nu}) \approx 1$, resulting in

$$\alpha_{\mu,\nu} \approx \frac{\mu}{1 - \sqrt{\nu}}. \quad (\text{A6})$$

We now look at the Rayleigh-normal distribution, which takes the form

$$\begin{aligned} \sqrt{1 - Z_\nu(\mu)} &= \sqrt{\frac{2\sqrt{\nu}}{1+\nu}} e^{-\mu^2/4(1+\nu)} \Phi_{\frac{-\mu}{1+\nu}, \frac{2\nu}{1+\nu}}(-\alpha_{\mu,\nu}) \\ &+ \sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \end{aligned} \quad (\text{A7})$$

We now note that, as $\alpha_{\mu,\nu} \rightarrow +\infty$, the first term on the RHS is exponentially vanishing, and second expo-

nentially approaching 1. Specifically we have

$$\ln \left[1 - \sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})} \right] \approx -\frac{1}{2} \frac{\mu^2}{(1 - \sqrt{\nu})^2}, \quad (\text{A8a})$$

$$\ln \left[e^{-\mu^2/4(1+\nu)} \Phi_{\frac{-\mu}{1+\nu}, \frac{2\nu}{1+\nu}}(\alpha_{\mu,\nu}) \right] \approx -\frac{1}{2} \frac{\mu^2}{(1 - \sqrt{\nu})^2}, \quad (\text{A8b})$$

and thus for $\mu \rightarrow -\infty$ we have

$$\ln [Z_\nu(\mu)] \approx -\frac{1}{2} \frac{\mu^2}{(1 - \sqrt{\nu})^2}. \quad (\text{A9})$$

2. Consistency with Theorems 1 and 2

The small deviation analyses of Refs. [7, 8] consider transformations with a fixed infidelity, bounded away from zero. Despite this, using the above expansion one can naively substitute $1 - F = \epsilon_n^-$ into the expressions for the optimal rates obtained within the small deviation regime. Whilst this analysis is no longer rigorous, we will see that it gives a rate which is consistent with our rigorous results, Theorems 1 and 2.

Inverting the expansion around $\mu = -\infty$, we have

$$Z_\nu^{-1}(\epsilon) \approx |1 - \sqrt{\nu}| \sqrt{-2 \ln \epsilon}, \quad (\text{A10})$$

for small positive ϵ . In particular, for $\epsilon_n^- := e^{-nt_n^2}$ we have

$$Z_\nu^{-1}(\epsilon_n^-) \simeq |1 - \sqrt{\nu}| \sqrt{2nt_n}. \quad (\text{A11})$$

Substituting the above into the results of Refs. [7, 8] yields the expressions for optimal rates from Theorems 1 and 2.

3. Expansion around $\mu = +\infty$.

The crossing point $\alpha_{\mu,\nu}$ is defined by the equation [7]

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} = \frac{\Phi(\alpha_{\mu,\nu})}{\Phi_{\mu,\nu}(\alpha_{\mu,\nu})}. \quad (\text{A12})$$

As $\mu \rightarrow +\infty$, $\alpha_{\mu,\nu} \rightarrow -\infty$. We can now use the approximation $\Phi(x) \approx \phi(x)/x$ for $x \rightarrow -\infty$. Applying this to the above, we have

$$\frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} \approx \frac{\phi(\alpha_{\mu,\nu})}{\phi_{\mu,\nu}(\alpha_{\mu,\nu})} \cdot \frac{\alpha_{\mu,\nu} - \mu}{\nu \alpha_{\mu,\nu}}, \quad (\text{A13})$$

which in turn implies that

$$\alpha_{\mu,\nu} \approx \frac{\mu}{1 - \nu}. \quad (\text{A14})$$

Returning to the Rayleigh-normal distribution, we note that $\sqrt{1 - Z_\nu(\mu)}$ exponentially vanishes as

$\alpha_{\mu,\nu} \rightarrow -\infty$, specifically

$$\ln \left[\sqrt{\Phi(\alpha_{\mu,\nu})\Phi_{\mu,\nu}(\alpha_{\mu,\nu})} \right] \approx -\frac{1+\nu}{(1-\nu)^2} \frac{\mu^2}{4}, \quad (\text{A15a})$$

$$\ln \left[e^{-\mu^2/4(1+\nu)} \Phi_{\frac{-\mu}{1+\nu}, \frac{2\nu}{1+\nu}}(\alpha_{\mu,\nu}) \right] \approx -\frac{1}{1+\nu} \frac{\mu^2}{4}. \quad (\text{A15b})$$

As such, we can see that the second expression dominates, leading us to conclude that, for $\mu \rightarrow +\infty$, the Rayleigh-normal distribution may be approximated as

$$\ln[1 - Z_\nu(\mu)] \approx -\frac{\mu^2}{4(1+\nu)}. \quad (\text{A16})$$

4. Conjectured infidelity analogue of Theorem 3

Similar to the case of the expansion around $\mu = -\infty$, we can use the expansion around $\mu = +\infty$ to (non-rigorously) obtain a conjecture for the analogue of Theorem 3 for error measured in infidelity. Inverting the expansion around $\mu = +\infty$, we find

$$Z_\nu^{-1}(\epsilon) \approx \sqrt{-4(1+\nu) \ln(1-\epsilon)}, \quad (\text{A17})$$

for ϵ close to 1. Specifically, for $\epsilon_n^+ = 1 - e^{-nt_n^2}$ we have

$$Z_\nu^{-1}(\epsilon_n^+) \approx \sqrt{4(1+\nu)nt_n}. \quad (\text{A18})$$

Substituting this approximation into the results of Refs. [7, 8], we have that the optimal rates for an infidelity error level of ϵ_n^+ are of the following form

$$R_n^{\text{ent}}(\epsilon_n) \simeq R_\infty^{\text{ent}} + \sqrt{\frac{2V(\mathbf{p})}{H(\mathbf{q})^2}} \sqrt{1 + 1/\nu^{\text{th}}} t_n, \quad (\text{A19a})$$

$$R_n^{\text{th}}(\epsilon_n) \simeq R_\infty^{\text{th}} + \sqrt{\frac{2V(\mathbf{p}||\gamma)}{D(\mathbf{q}||\gamma)^2}} \sqrt{1 + 1/\nu^{\text{th}}} t_n. \quad (\text{A19b})$$

Appendix B: Compact convergence lemma

Lemma 18. *Let $\{f_n\}_n$ and $\{g_n\}_n$ be sequences of non-decreasing real functions, both of which point-wise converge to continuous functions f and g , respectively. If $f > g$, then $f_n|_X > g_n|_X$ eventually on all compact X .*

Proof. First we note (see, e.g., page 1 of Ref. [32]) that sequences of non-decreasing real functions which point-wise converge to continuous functions do so compactly. Let $\Delta_n : X \rightarrow \mathbb{R}$ be defined by $\Delta_n(x) = f_n(x) - g_n(x)$ for all $x \in X$. As $\Delta = f - g$, we know that $\Delta > 0$. Indeed, because Δ is a continuous function on a compact domain, the Extreme Value Theorem tells us that Δ is bounded away from zero, i.e., there exists an $\epsilon > 0$ such that $\Delta \geq \epsilon$. As $\Delta_n \rightarrow \Delta$ uniformly, we must eventually have that $\Delta_n \geq \Delta - \epsilon/2 \geq \epsilon/2$, and so $f_n|_X > g_n|_X$. \square