

Fundamental Quality Bound on Optical Quantum Communication

Tobias Rippchen^{*1}, Ludovico Lami², Gerardo Adesso³, and Mario Berta¹

¹*Institute for Quantum Information, RWTH Aachen University, Germany*

²*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy*

³*School of Mathematical Sciences and Centre for the Mathematical and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom*

Abstract

Sending quantum information reliably over long distances is a central challenge in quantum technology in general, and in quantum optics in particular, since most quantum communication relies on optical fibres or free-space links. Here, we address this problem by shifting the focus from the quantity of information sent to the quality of the transmission, i.e. the rate of decay of the transmission error with respect to the number of channel uses. For the general class of teleportation-simulable channels, which includes all channels arising in quantum optical communication, we prove that the single-letter reverse relative entropy of entanglement of the Choi state upper bounds the error exponent of two-way assisted quantum communication — paralleling the celebrated capacity bound of [Pirandola *et al.*, Nat. Comm. (2017)] in terms of the regularised relative entropy of entanglement. Remarkably, for Gaussian channels our bound can be computed efficiently through a convex program with simple constraints involving only finite-dimensional covariance matrices. As a prototypical application, we derive closed-form analytical expressions for several one-mode Gaussian channels that are fundamental to optical communication. Extending recent work [Lami *et al.*, arXiv:2408.07067 (2024)] to infinite-dimensional systems, we further endow the reverse relative entropy of entanglement with an exact operational interpretation in entanglement testing, and show that it characterises the rate of entanglement distillation under non-entangling operations. These findings offer a new perspective on entanglement as a resource and sharpen the theoretical benchmarks for future quantum optical networks.

The Challenge with Capacities. Quantum communication lies at the heart of emerging quantum technologies, from secure cryptography to distributed quantum computing [1]. It is thus of crucial importance to understand the fundamental limits to which the transmission of quantum information through noisy channels, such as optical fibres or free-space links, is subjected. Traditionally, this problem has been analysed in terms of *capacities*, which quantify the maximum amount of information that can be transmitted reliably. Of particular importance is the *quantum capacity*, which measures the ability of a channel to faithfully transmit qubits, enabling core primitives such as the distribution of entanglement.

Despite its importance, the quantum capacity of generic channels remains poorly understood. The main challenge is that optimal communication involves encoding the information across multiple uses of the channel. Consequently, the capacity is defined in an information-theoretic sense as an asymptotic rate, i.e. it quantifies how the performance scales with the number of channel uses. Notably, this is already true in the purely classical case of transmitting classical information via a classical channel. However, as Shannon showed in his foundational work [2], the capacity in this case

^{*}Contact information: tobias.rippchen@rwth-aachen.de

is mathematically given by a *single-letter* formula that involves optimising an entropic quantity—the mutual information—over a single use of the channel. In stark contrast, the (one-way assisted) quantum capacity $\mathcal{Q}$ of a quantum channel $\mathcal{N}$ is given by a *regularised* formula,

$$\mathcal{Q}(\mathcal{N}) := \lim_{n \rightarrow \infty} \frac{1}{n} \mathcal{Q}_n(\mathcal{N}), \quad \text{with} \quad \mathcal{Q}_n(\mathcal{N}) := \sup_{\rho_n} I_{\text{coh}}(\mathcal{N}^{\otimes n}, \rho_n), \quad (1)$$

i.e. by an optimisation of an entropic quantity—the coherent information I_{coh} —in the limit of arbitrarily many uses of the channel [3–5]. Here, $\mathcal{Q}_n(\mathcal{N})$ is the coherent information maximised for n uses of the channel $\mathcal{N}$ over a joint input state ρ_n .

These regularised formulas are notoriously difficult to compute, both analytically and numerically, because the underlying entropic measures are typically not additive [6, 7]. Moreover, even deciding whether a channel has non-zero quantum capacity generally requires an unbounded number of channel uses [8]. Because of these computational hurdles, we cannot calculate the quantum capacity even for very simple and physically relevant channels, e.g. thermal attenuators, which constitute the prototypical model for optical quantum communication.

Recently, a shift in perspective has gained traction in quantum information theory that may be summarised as *quality over quantity* [9–15]. Rather than asking “how much” information can be transmitted optimally, one may instead ask about the quality of the transmission. In fact, in many practical scenarios—such as distributing entanglement for cryptography—it is often more desirable to obtain less entanglement of higher quality than a lot of it of poorer quality (see Figure 1). The central question of this work may thus be phrased as:

What are the fundamental limitations on the quality of quantum communication and entanglement distribution over optical channels?

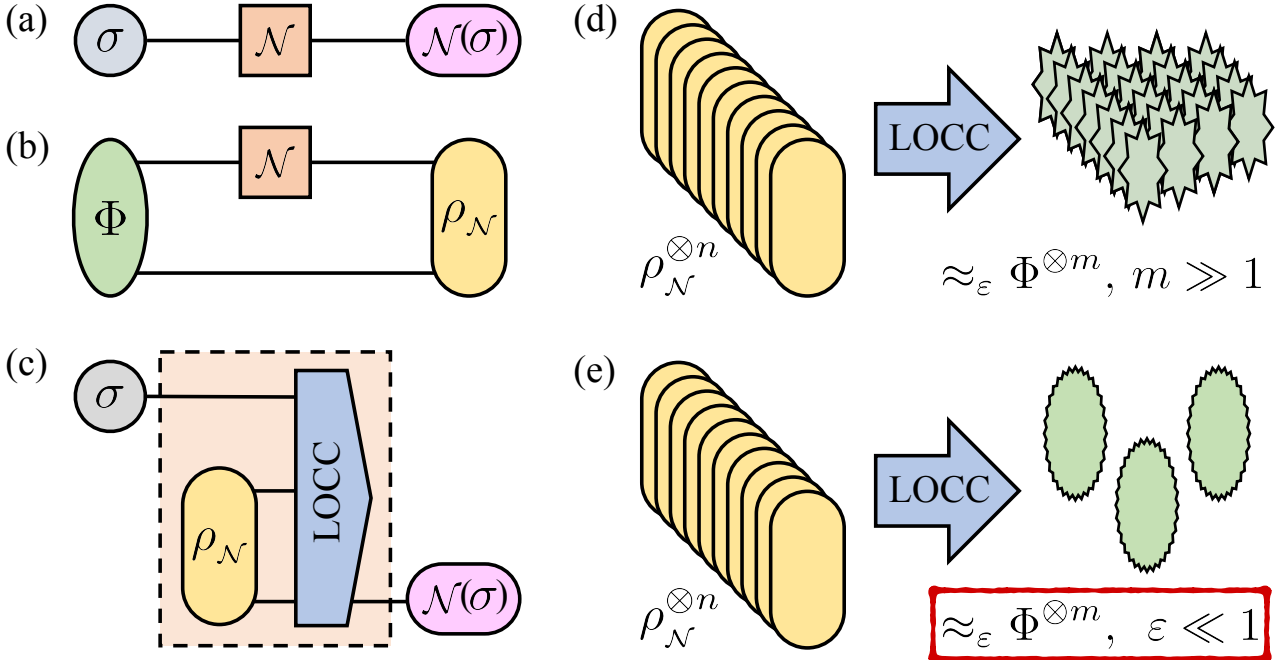


Figure 1: (a) A quantum channel $\mathcal{N}$ mapping an input state σ to an output state $\mathcal{N}(\sigma)$. (b) The Choi state $\rho_{\mathcal{N}} = (\mathcal{N} \otimes I)(\Phi)$ of the channel $\mathcal{N}$, where Φ is a maximally entangled state. (c) The channel $\mathcal{N}$ is said to be teleportation-simulable if its action on any input state σ can be reproduced by means of an LOCC protocol supplemented by the Choi state $\rho_{\mathcal{N}}$ distributed between sender and receiver. The quantum communication scenario studied in this work can be recast in terms of how well one can distill m copies of approximately maximally entangled states $\approx_{\varepsilon} \Phi^{\otimes m}$ from $n \gg 1$ copies $\rho_{\mathcal{N}}^{\otimes n}$ of the Choi state $\rho_{\mathcal{N}}$ of a channel $\mathcal{N}$, for which two paradigms can be considered: (d) aiming for *quantity*, i.e. maximising the number of output copies m , or (e) aiming for *quality*, i.e. maximising the exponent at which the error ε decays as $n \rightarrow \infty$. This work investigates the latter scenario.

A Link to Entanglement Theory. The answer to this question has important consequences beyond communication theory itself. In particular, the theory of entanglement measures is deeply linked with quantum communication. To make this connection explicit, it is instructive to consider the framework of (unlimited) two-way assistance, referred to in technical terms as *adaptive local operations and classical communication* (adaptive LOCC). This represents the most general class of protocols that sender and receiver may employ to aid quantum communication without relying on pre-shared entanglement. For example, when distributing entanglement this is a natural assistance model, since classical communication is considered inexpensive in this context. The associated *two-way assisted capacity* thus provides an important fundamental benchmark for quantum communication.

The two-way assisted capacity is substantially easier to analyse for the class of *teleportation-simulable channels*—those whose action can be simulated by running the quantum teleportation protocol using their own Choi state as a resource (see Figure 1 for a visual definition). Importantly, this encompasses all physically relevant channels in optical communication. For these channels, the two-way assisted quantum capacity coincides with the *distillable entanglement* of the Choi state [16, 17]. The latter, in turn, characterises the asymptotic yield of *entanglement distillation*, a fundamental primitive in entanglement theory, in which the goal is to convert noisy entangled states into pure, maximally entangled ones [16, 18, 19]. This establishes a rigorous and quantitative link between the theory of entanglement measures and quantum communication.

Naturally, the same non-additivity problems that plague quantum communication theory also arise in the theory of entanglement measures [20–22]. More precisely, the distillable entanglement is related to the *regularised relative entropy of entanglement* [23–25], defined via

$$D^\infty(\rho\|\text{SEP}) := \lim_{n \rightarrow \infty} \frac{1}{n} \inf_{\sigma_n \in \text{SEP}} D(\rho^{\otimes n} \|\sigma_n), \quad (2)$$

i.e. as the asymptotic relative-entropy distance between the bipartite product state $\rho^{\otimes n}$ and the set of states that is not entangled across the bipartite cut of the global system, denoted SEP. Crucially, the set SEP lacks the product structure of $\rho^{\otimes n}$, leading to a fundamental non-additivity that necessitates regularisation. The difficulty of evaluating this regularised formula remains one of the main bottlenecks to theoretical progress.

However, shifting the perspective from *quantity to quality* has already led to fundamentally new insights in entanglement theory. In a prior work [9], some of us discovered that the *reverse relative entropy of entanglement* [23, 26], defined as

$$D(\text{SEP}\|\rho) := \inf_{\sigma \in \text{SEP}} D(\sigma\|\rho), \quad (3)$$

plays a central role in the theory of entanglement measures when adopting the quality framework. Specifically, this measure characterises the *error exponent* of entanglement distillation under *non-entangling* operations. The latter class of operations is a useful relaxation of the traditional LOCC framework, whose study has already been fruitful in the past [25, 27–30]. Strikingly, the reverse relative entropy of entanglement enjoys a highly desirable property—it is additive on product inputs, which makes its regularisation superfluous. Thus, precisely for the previously poorly understood mixed quantum states, the shift of perspective from the *quantity* to the *quality* of the distilled entanglement enables a single-letter characterisation of the asymptotic performance. This gives us a strong intuition that this entanglement measure should also play a similar role in the theory of quantum communication.

Results

In following this intuition, we are faced with three immediate problems: (1) the previous work [9] dealt with states and not channels, (2) the proof is clearly limited to finite-dimensional systems and, most importantly, (3) even if you could reprove the generalised Sanov theorem that underpins the

main findings, the resulting expression would in general *not* be efficiently computable, as it would involve an optimisation over infinite-dimensional separable states. In this work, we solve all three of these problems, finding fundamental bounds on the *quality* of entanglement distribution and quantum communication over most optically relevant quantum channels.

Our finding complements the celebrated work by Pirandola et al. [17], which established fundamental bounds on the *quantity* of entanglement that can be distributed over the same class of optical channels (see also Bennett et al. [16]). The duality between our result and theirs is striking: while [17] showed that, for teleportation-simulable channels, the standard relative entropy of entanglement yields an upper bound on the two-way assisted quantum capacity, we prove that the *reverse* relative entropy of entanglement upper bounds the two-way assisted *error exponent*. Moreover, unlike in the case of [17], we can actually prove that the quantifiers we find have an exact operational interpretation in entanglement distillation, when relaxing the LOCC framework to non-entangling operations. The experience with entanglement measures suggests that our single-letter bound on the error exponent may already be among the tightest possible, in the sense that going substantially beyond it will unavoidably run into the fundamental problem of NPT bound entanglement, which is currently beyond our understanding [31] (see Remark 3.9 in the supplementary material for details).

The Error Exponent of Quantum Communication. Consider two parties, Alice and Bob, connected by a noisy bosonic channel $\mathcal{N}$, whose aim is to distribute entanglement between them. In order to achieve this task, they may employ the help of adaptive LOCC which we denote in the following by $\text{LOCC}_{\leftrightarrow}$ (see Figure 1 for the setup of this problem). Observe that in this setup once they can reliably establish entanglement, they can also transmit arbitrary quantum information using the quantum teleportation protocol.

More formally, the two parties will apply some protocol Λ_n , that uses the the channel n times, and produces at the output a state $\rho(\Lambda_n)$ that approximates m copies of the maximally entangled state $\Phi = |\Phi\rangle\langle\Phi|$ with $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Denoting the approximation error as ε_n , we can write this as $\rho(\Lambda_n) \approx_{\varepsilon_n} \Phi^{\otimes m}$, using a suitable measure of distance—either the fidelity or, equivalently, the trace distance. Although we allow for this finite error, we do require that it satisfies $\lim_{n \rightarrow \infty} \varepsilon_n = 0$; that is, as the channel can be used more often, the quality of the distributed entanglement should increase, becoming perfect in the asymptotic limit.

Traditionally, the focus was put on the *quantity* of distributed entanglement. In this setting, the rate $\frac{m}{n}$ of the protocol is the figure-of-merit and the goal is to find the largest asymptotic rate $\lim_{n \rightarrow \infty} \frac{m}{n}$ that can be achieved among the feasible protocols. This then leads to the notion of the two-way assisted quantum capacity. In this work, we shift the perspective to the *quality* of the obtained entanglement. That is, we require that $\varepsilon_n \sim 2^{-cn}$ and characterise the optimal achievable error exponent c . The *error exponent of two-way assisted quantum communication* is then defined as the largest such exponent that can be achieved in the asymptotic limit¹

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}) := \lim_{m \rightarrow \infty} \sup \left\{ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \varepsilon_n : \rho(\Lambda_n) \approx_{\varepsilon_n} \Phi^{\otimes m}, \Lambda_n \in \text{LOCC}_{\leftrightarrow}(\mathcal{N}^{\times n}) \right\}, \quad (4)$$

where we optimise over sequences of adaptive LOCC protocols that use the quantum channel n times. Observe that this definition no longer places any importance on the precise number of maximally entangled copies (provided that it can be made as large as desired), but only on the exponentially decreasing error. We then find the following *single-letter* upper bound.

¹To be technically precise, this defines the *zero-rate* error exponent in Shannon theory parlance.

Proposition 1. *For the class of teleportation-simulable channels, the reverse relative entropy of entanglement provides an upper bound on their error exponent of two-way assisted quantum communication. Specifically, we have for a channel $\mathcal{N}$ that acts on m bosonic modes that*

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}) \leq \liminf_{r \rightarrow \infty} D(\text{SEP} \| \rho_{\mathcal{N}}(r)), \quad (5)$$

where SEP denotes the set of separable states and the quasi-Choi state, $\rho_{\mathcal{N}}(r) := (\mathcal{N} \otimes \mathcal{I})(\Phi(r)^{\otimes m})$, is obtained by sending one half of a m -mode two-mode squeezed vacuum state $\Phi(r)^{\otimes m}$ through the channel.

In our proof, we start from the observation that for teleportation-simulable channels any adaptive LOCC protocol that employs n uses of the channel can be reduced into an equivalent protocol that acts on n copies of the associated Choi state (cf. [16] and [17]). Our bound then follows by exploiting the specific properties of the reverse relative entropy of entanglement, most notably its additivity, in combination with a suitable characterisation of the error exponent. The technical details can be found in Sec. 3.3 and 3.4 of the supplementary material.

Gaussian Reverse Relative Entropy of Entanglement. Although the reverse relative entropy of entanglement does not require regularisation, that does not mean it is efficiently computable in general; in fact, the optimisation over separable states is NP-hard (see e.g. [32] and references therein), and in infinite-dimensional systems one cannot even resort to SDP hierarchies (via extendibility) to soften this issue. Fortunately, we show that these problems evaporate in the case of N -mode bosonic Gaussian states. Specifically, we prove that the computation of the reverse relative entropy of entanglement reduces to a convex program over N -dimensional quantum covariance matrices with two simple positive semi-definite constraints.

Proposition 2. *The reverse relative entropy of entanglement of the Gaussian state ρ_G on the bipartite system $A \otimes B$ with quantum covariance matrix V_ρ can be computed via the **convex** program*

$$\begin{aligned} \min_{V_\sigma, \gamma_A > 0} \quad & \frac{\text{Tr}[V_\sigma(G[V_\rho] - G[V_\sigma])]}{2 \ln(2)} + \log_2 \sqrt{\frac{\det(V_\rho + i\Omega_{AB})}{\det(V_\sigma + i\Omega_{AB})}} \\ \text{s.t.} \quad & V_\sigma \geq \gamma_A \oplus i\Omega_B \quad \text{and} \quad \gamma_A \geq i\Omega_A \end{aligned} \quad (6)$$

where Ω is the canonical symplectic form of the system, and $G[V]$ is the Gibbs matrix associated with V .

Our result has fundamental consequences in the theory of entanglement measures in continuous-variable systems. While Gaussian states are fundamental and ubiquitous, prior to our work it was unknown whether any single Gaussian entanglement measure could be at the same time (A) operationally meaningful and (B) efficiently computable. The negativity [33] and Gaussian entanglement of formation [34] are both efficiently computable, but they do not enjoy a strong operational interpretation in themselves. It is the actual (regularised) entanglement of formation that has an operational interpretation, but we do not know in general if it coincides with its Gaussian version (except in special cases [35, 36]); and, crucially, we do not know how to regularise it. Therefore, the reverse relative entropy of entanglement is – to the best of the authors’ knowledge – the first *operational* entanglement measure that is *efficiently computable* for Gaussian states.

The key to this finding is that, for Gaussian inputs, the regular and Gaussian reverse relative entropy of entanglement coincide. The main technical hurdle to prove this is to show that the optimisation can be restricted to separable states with finite second moments. We achieve this with an argument based on the variational formula for the measured relative entropy from [37]. Once this has been

established, a Gaussification argument allows us to restrict the optimisation further to Gaussian separable states (see Sec. 3.1 of the supplementary material). This is especially noteworthy, as it is unknown whether a similar restriction holds for the standard relative entropy of entanglement. The convex program follows by combining the efficient description for Gaussian separability from [38] with the characterisation of the relative entropy in terms of statistical moments from [17]. Convexity of the resulting program is established by lifting the convexity of the relative entropy on states via our Gaussification argument. The technical details are presented in Sec. 3.2 of the supplementary material.

Thermal Attenuator Channel. As a paradigmatic example from optical communication, we consider the *thermal attenuator channel*, which serves as the predominant model for realistic optical quantum links. This is also the simplest non-trivial example because the error exponent of the (previously well understood) pure-loss channel diverges—as it should, because zero-error entanglement generation over a pure-loss channel is possible via dual-rail encoding. Moreover, it is also precisely this type of *noisier* channel, for which the capacity bound of Pirandola et al. [17] is no longer tight, making them harder to understand from the *quantity* perspective on quantum communication.

Mathematically, the single-mode thermal attenuator channel is modelled by mixing the input mode at a beamsplitter of transmissivity $\lambda \in [0, 1]$ with an ancillary mode prepared in a Gaussian thermal state with covariance matrix $n_{\text{th}}\sigma_0$ and vanishing first moments. The latter precisely corresponds to the Gibbs state of the free Hamiltonian associated with the thermal noise parameter $n_{\text{th}} \geq 1$.

We then find the following upper bound on its error exponent

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{\text{Att}}) \leq \lim_{r \rightarrow \infty} D(\text{SEP} \| \rho_{\text{Att}}(r)) = \frac{n_{\text{sep}} (\text{arccoth}(n_{\text{th}}) - \text{arccoth}(n_{\text{sep}}))}{\ln(2)} + \log_2 \left(\sqrt{\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}} \right) \quad (7)$$

for $1 \leq n_{\text{th}} \leq n_{\text{sep}}(\lambda)$ and zero otherwise (see Sec. 3.5 for details). Here, n_{sep} is defined as $n_{\text{sep}}(\lambda) := \frac{1+\lambda}{1-\lambda}$. As expected, this diverges for $n_{\text{th}} \rightarrow 1$, i.e. in the case of the pure-loss channel. Additionally, we find that, aesthetically pleasing, the asymptotically closest separable state coincides with the asymptotic Choi state obtained by sending the *other* half of the two-mode squeezed vacuum state through a thermal attenuator channel with the same transmissivity and thermal noise parameter given by $n_{\text{sep}}(\lambda)$.

Thermal Amplifier Channel. Another essential example for optical communication is the *thermal amplifier channel*, which models amplification in the presence of thermal noise. It can be described by the interaction with a thermal mode—with covariance matrix $n_{\text{th}}\sigma_0$, and null first moments—through a two-mode squeezing with gain $\eta \geq 1$.

We then find that its error exponent is bounded qualitatively by the same expression as the thermal attenuator

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{\text{Amp}}) \leq \frac{n_{\text{sep}} (\text{arccoth}(n_{\text{th}}) - \text{arccoth}(n_{\text{sep}}))}{\ln(2)} + \log_2 \left(\sqrt{\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}} \right) \quad (8)$$

for $1 \leq n_{\text{th}} \leq n_{\text{sep}}(\eta)$ and zero otherwise (see Sec. 3.5 for details). However, the definition of n_{sep} is now changed to $n_{\text{sep}}(\eta) := \frac{\eta+1}{\eta-1}$. As above, this diverges for $n_{\text{th}} \rightarrow 1$, which corresponds to the quantum limited amplifier.

Additive-Noise Channel. Perhaps the conceptually simplest model of decoherence in optical communication is given by the additive-noise Gaussian channel. This can be seen as the action of random Gaussian displacement on the input mode and is modelled by adding the covariance matrix $\mu\sigma_0$, with $\mu \geq 0$, to the input covariance matrix.

We find that its error exponent is bounded as

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{\text{Noise}}) \leq \frac{2 - \mu}{\mu \ln(2)} + \log_2 \left(\frac{\mu}{2} \right) \quad (9)$$

provided that $0 \leq \mu \leq 2$ and zero otherwise (see Sec. 3.5 for details).

Operational Interpretations. The reverse relative entropy of entanglement was endowed with an operational interpretation in [9], but this was limited to finite-dimensional systems only. In this work, we rigorously extend this interpretation to *general* (separable) infinite-dimensional quantum systems, establishing the reverse relative entropy of entanglement as an operationally meaningful entanglement measure.

We start by adopting the quality framework from the prior work [9] to analyse the task of entanglement distillation. Consider two parties Alice and Bob that share many copies of a bipartite (infinite-dimensional) state ρ_{AB} with the aim to convert them into pure, maximal entanglement. To that end, they apply some protocol Λ_n such that, when acting on n copies of ρ_{AB} , the final state approximates m copies of the maximally entangled state Φ . Regarding the set of feasible protocols, we follow [9] and relax the physically well-motivated but mathematically difficult LOCC framework to the class of non-entangling operations (NE) [25, 27–30] — by definition, $\Lambda_n(\sigma_n)$ must remain unentangled for all unentangled states σ_n .

Considering the quality of the distilled entanglement as the figure-of-merit, leads to the following definition for the *error exponent of entanglement distillation*:

$$E_{d, \text{err}}(\rho_{AB}) := \lim_{m \rightarrow \infty} \sup \left\{ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \varepsilon_n : \Lambda_n(\rho^{\otimes n}) \approx_{\varepsilon_n} \Phi^{\otimes m}, \Lambda_n \in \text{NE} \right\}. \quad (10)$$

We then show that, as in the finite-dimensional case, this exponent is fundamentally connected to a task from quantum state discrimination [39–41] known as entanglement testing (see Sec. 2.2 for the technical details). In entanglement testing, the goal is to distinguish the given entangled state $\rho_{AB}^{\otimes n}$ from the set of all separable states by performing a collective measurement on the global system. As usual, one can distinguish between two types of errors: the type-1 error occurs when mistaking $\rho_{AB}^{\otimes n}$ for a separable state; conversely, the type-2 error occurs when mistaking a separable state for $\rho_{AB}^{\otimes n}$. For a fixed type-2 error probability, the type-1 error decays exponentially and the asymptotically optimal error exponent is referred to as the *Sanov exponent*, denoted $\text{Sanov}(\rho_{AB} \| \mathcal{S}_{A:B})$.

As our main operational interpretation, we then prove that the latter exponent coincides exactly with the reverse relative entropy of entanglement, giving the solution to the *generalised quantum Sanov theorem* [9, 15, 42, 43] of entanglement testing (see Sec. 2.3 for the full argument). To summarise, we establish for (separable) infinite-dimensional systems the equalities:

$$E_{d, \text{err}}(\rho_{AB}) = \text{Sanov}(\rho_{AB} \| \mathcal{S}_{A:B}) = D(\mathcal{S}_{A:B} \| \rho_{AB}). \quad (11)$$

It is noteworthy that the proof does *not* follow from a standard truncation argument. Indeed, even if we truncate the space so that the tails of ρ_{AB} are eliminated, there is no a priori guarantee that the separable state we have to discriminate it from will lie in the same truncated space: the technical difficulty of our result lies in proving precisely this. Once that is done, our lifting procedure crucially relies on a semi-continuity property of the reverse relative entropy of entanglement, which we establish with a particular choice of operator topology.

Outlook. In this work, we have studied quantum communication from a quality perspective, establishing a fundamental bound on the error exponent of two-way assisted quantum communication. This bound parallels the bound of Pirandola et. al. [17] (see also Bennett et. al. [16, Section 5]) for the

two-way assisted capacity, which turns out to be very tight even for unassisted communication in certain regimes. Moreover, the derivation of the single-letter capacity bound follows by a generally non-tight relaxation of the regularised relative entropy of entanglement to its single-letter version. Crucially, this relaxation step does not occur in our analysis due to the additivity of the reversed entanglement measure. This makes a thorough investigation of the achievability of our bound a promising avenue for future work.

Additionally, in the Gaussian case, we showed that our bound can be computed via a finite-dimensional convex program with two positive semi-definite constraints. As such, it can straightforwardly be solved with off-the-shelf solvers based on interior-point methods, which are well-known to be efficient in praxis [44]. However, it might still be desirable to establish a precise complexity-theoretic result for the efficiency of this program. For this, one would (after rewriting it into a standard conic program using the epigraph formulation) need to construct a self-concordant logarithmic barrier function of the resulting convex cone. This would then give provable convergence guarantees using the barrier method, i.e. polynomial-time solvability in the number of modes. A possible path to achieve this would be to adapt the specialised literature on relative entropy optimisation such as [45] to this special case (see also Remark 3.5 in the supplementary material).

Another interesting direction for future work is to extend our analysis to general quantum resource theories [46]. Under certain assumptions, the reverse relative entropy of resource attains an operational interpretation in the task of resource testing (see the supplementary material and [9] for more details). Thus, another interesting open question is whether the resulting measure is also efficiently computable for other Gaussian resource theories, such as non-classicality or, more generally, λ -negativity.

Acknowledgments. The authors thank Bartosz Regula for discussions in the early stages of the project. TR thanks Julius Zeiss for insightful discussions at numerous stages of this work. MB and TR acknowledge support from the Excellence Cluster - Matter and Light for Quantum Computing (ML4Q) and funding by the European Research Council (ERC Grant Agreement No. 948139). G.A. acknowledges funding from the UK Research and Innovation (UKRI EPSRC Grant No. EP/X010929/1). LL acknowledges financial support from the European Union (ERC StG ETQO, Grant Agreement no. 101165230). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References

- [1] S. Pirandola et al. “Advances in Quantum Cryptography”. In: *Advances in Optics and Photonics* 12 (4 2020), pp. 1012–1236. eprint: <https://arxiv.org/abs/1906.01645>.
- [2] C. E. Shannon. “A Mathematical Theory of Communication”. In: *The Bell System Technical Journal* 27 (3 1948), pp. 379–423.
- [3] S. Lloyd. “Capacity of the Noisy Quantum Channel”. In: *Physical Review A* 55 (3 1997), pp. 1613–1622. eprint: <https://arxiv.org/abs/quant-ph/9604015>.
- [4] P. W. Shor. *The Quantum Channel Capacity and Coherent Information*. Lecture Notes, MSRI Workshop on Quantum Computation. 2002.
- [5] I. Devetak. “The Private Classical Capacity and Quantum Capacity of a Quantum Channel”. In: *IEEE Transactions on Information Theory* 51 (1 2005). eprint: <https://arxiv.org/abs/quant-ph/0304127>.
- [6] P. Shor. “Equivalence of Additivity Questions in Quantum Information Theory”. In: *Communications in Mathematical Physics* 246 (2004), p. 473. eprint: <https://arxiv.org/abs/quant-ph/0305035>.

- [7] M. B. Hastings. “Superadditivity of Communication Capacity Using Entangled Inputs”. In: *Nature Physics* 5 (2009), pp. 255–257. eprint: <https://arxiv.org/abs/0809.3972>.
- [8] T. Cubitt et al. “Unbounded Number of Channel Uses are Required to See Quantum Capacity”. In: *Nature Communications* 6.6739 (2015). eprint: <https://arxiv.org/abs/1408.5115>.
- [9] L. Lami, M. Berta, and B. Regula. *Asymptotic Quantification of Entanglement with a Single Copy*. 2024. eprint: <https://arxiv.org/abs/2408.07067>.
- [10] K. Ji et al. “Barycentric Bounds on the Error Exponents of Quantum Hypothesis Exclusion”. In: (2024). eprint: <https://arxiv.org/abs/2407.13728>.
- [11] T. Nuradha et al. “Multivariate Fidelities”. In: *Journal of Physics A: Mathematical and Theoretical* 58.16 (2025), p. 165304. eprint: <https://arxiv.org/abs/2404.16101>.
- [12] F. Girardi et al. “Umlaut information”. In: (2025). eprint: <https://arxiv.org/abs/2503.18910>.
- [13] F. Girardi et al. “Quantum Umlaut information”. In: (2025). eprint: <https://arxiv.org/abs/2503.21479>.
- [14] M. Hayashi. “General Detectability Measure”. In: (2025). eprint: <https://arxiv.org/abs/2501.09303v2>.
- [15] M. Hayashi. “Entanglement Measures for Detectability”. In: *IEEE Transactions on Information Theory* 71.6 (2025). eprint: <https://arxiv.org/abs/2311.11189>.
- [16] C. H. Bennett et al. “Mixed-State Entanglement and Quantum Error Correction”. In: *Physical Review A* 54 (5 1996), pp. 3824–3851. eprint: <https://arxiv.org/abs/quant-ph/9604024>.
- [17] S. Pirandola et al. “Fundamental Limits of Repeaterless Quantum Communications”. In: *Nature Communications* 8.15043 (2017). eprint: <https://arxiv.org/abs/1510.08863>.
- [18] C. H. Bennett et al. “Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels”. In: *Physical Review Letters* 76 (5 1996), pp. 722–725. eprint: <https://arxiv.org/abs/quant-ph/9511027>.
- [19] C. H. Bennett et al. “Concentrating Partial Entanglement by Local Operations”. In: *Physical Review A* 53 (4 1996), pp. 2046–2052. eprint: <https://arxiv.org/abs/quant-ph/9511030>.
- [20] K. G. H. Vollbrecht and R. F. Werner. “Entanglement Measures under Symmetry”. In: *Physical Review A* 64 (6 2001), p. 062307. eprint: <https://arxiv.org/abs/quant-ph/0010095>.
- [21] P. M. Hayden, M. Horodecki, and B. M. Terhal. “The Asymptotic Entanglement Cost of Preparing a Quantum State”. In: *Journal of Physics A: Mathematical and General* 34 (2001), p. 6891. eprint: <https://arxiv.org/abs/quant-ph/0008134>.
- [22] I. Devetak and A. Winter. “Distillation of Secret Key and Entanglement from Quantum States”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 461 (2005), pp. 207–235. eprint: <https://arxiv.org/abs/quant-ph/0306078>.
- [23] V. Vedral et al. “Quantifying Entanglement”. In: *Physical Review Letters* 78 (12 1997), pp. 2275–2279. eprint: <https://arxiv.org/abs/quant-ph/9702027>.
- [24] V. Vedral and M. B. Plenio. “Entanglement Measures and Purification Procedures”. In: *Physical Review A* 57 (3 1998), pp. 1619–1633. eprint: <https://arxiv.org/abs/quant-ph/9707035>.
- [25] F. G. S. L. Brandão and M. B. Plenio. “A Reversible Theory of Entanglement and its Relation to the Second Law”. In: *Communications in Mathematical Physics* 295 (2010), pp. 829–851. eprint: <https://arxiv.org/abs/0710.5827>.
- [26] J. Eisert, K. Audenaert, and M. B. Plenio. “Remarks on Entanglement Measures and Non-Local State Distinguishability”. In: *Journal of Physics A: Mathematical and General* 36 (2003), pp. 5605–5615. eprint: <https://arxiv.org/abs/quant-ph/0212007>.

- [27] F. G. S. L. Brandão and M. B. Plenio. “A Generalization of Quantum Stein’s Lemma”. In: *Communications in Mathematical Physics* 295 (2010), pp. 791–828. eprint: <https://arxiv.org/abs/0904.0281>.
- [28] F. G. S. L. Brandão and G. Gour. “Reversible Framework for Quantum Resource Theories”. In: *Physical Review Letters* 115 (7 2015), p. 070503. eprint: <https://arxiv.org/abs/1502.03149>.
- [29] M. Hayashi and H. Yamasaki. *Generalized Quantum Stein’s Lemma and Second Law of Quantum Resource Theories*. 2024. eprint: <https://arxiv.org/abs/2408.02722>.
- [30] L. Lami. *A Solution of the Generalised Quantum Stein’s Lemma*. 2024. eprint: <https://arxiv.org/abs/2408.06410>.
- [31] P. Horodecki, Ł. Rudnicki, and K. Życzkowski. “Five Open Problems in Quantum Information Theory”. In: *PRX Quantum* 3 (1 2022), p. 010101. doi: <https://doi.org/10.1103/PRXQuantum.3.010101>.
- [32] L. M. Ioannou. “Computational Complexity of the Quantum Separability Problem”. In: *Quantum Information and Computation* 7 (4 2007), pp. 335–370. eprint: <https://arxiv.org/abs/quant-ph/0603199>.
- [33] G. Vidal and R. F. Werner. “Computable Measure of Entanglement”. In: *Physical Review A* 65 (3 2002), p. 032314. eprint: <https://arxiv.org/abs/quant-ph/0102117>.
- [34] M. M. Wolf et al. “Gaussian Entanglement of Formation”. In: *Physical Review A* 69.5 (2004), p. 052320. eprint: <https://arxiv.org/abs/quant-ph/0306177>.
- [35] G. Giedke et al. “Entanglement of Formation for Symmetric Gaussian States”. In: *Physical Review Letters* 91.10 (2003), p. 107901. eprint: <https://arxiv.org/abs/quant-ph/0304042>.
- [36] Y. Akbari-Kourbolagh and H. Alijanzadeh-Boura. “On the Entanglement of Formation of Two-Mode Gaussian States: A Compact Form”. In: *Quantum Information Processing* 14 (11 2015), pp. 4179–4199. eprint: <https://arxiv.org/abs/1402.3321>.
- [37] G. Ferrari, L. Lami, and M. B. Plenio. “Asymptotic State Transformations of Continuous Variable Resources”. In: *Communications in Mathematical Physics* 398 (2023), pp. 291–351. eprint: <https://arxiv.org/abs/2010.00044>.
- [38] L. Lami, A. Serafini, and G. Adesso. “Gaussian Entanglement Revisited”. In: *New Journal of Physics* 20.2 (2018), p. 023030. eprint: <https://arxiv.org/abs/1612.05215>.
- [39] F. Hiai and D. Petz. “The Proper Formula for Relative Entropy and Its Asymptotics in Quantum Probability”. In: *Communications of Mathematical Physics* 143 (1991), pp. 99–114.
- [40] T. Ogawa and H. Nagaoka. “Strong Converse and Stein’s Lemma in Quantum Hypothesis Testing”. In: *IEEE Transactions on Information Theory* 46 (7 2000), pp. 2428–2433.
- [41] M. Hayashi. *Quantum Information Theory: Mathematical Foundation*. Springer Berlin, Heidelberg, 2017.
- [42] I. Bjelaković et al. “A Quantum Version of Sanov’s Theorem”. In: *Communications in Mathematical Physics* 269 (2005), pp. 659–671. eprint: <https://arxiv.org/abs/quant-ph/0412157>.
- [43] J. Nötzel. “Hypothesis Testing on Invariant Subspaces of the Symmetric Group: Part I. Quantum Sanov’s Theorem and Arbitrarily Varying Sources”. In: *Journal of Physics A: Mathematical and Theoretical* 47 (2014), p. 235303. eprint: <https://arxiv.org/abs/1310.5553>.
- [44] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [45] H. Fawzi and J. Saunderson. “Optimal Self-Concordant Barriers for Quantum Relative Entropies”. In: *SIAM Journal on Optimization* 33 (4 2023). eprint: <https://arxiv.org/abs/2205.04581>.
- [46] E. Chitambar and G. Gour. “Quantum Resource Theories”. In: *Reviews of Modern Physics* 91 (2 2019), p. 025001. eprint: <https://arxiv.org/abs/1806.06107>.

- [47] A. S. Holevo. *Quantum Systems, Channels, Information: A Mathematical Introduction*. De Gruyter, 2009.
- [48] A. Serafini. *Quantum Continuous Variables: A Primer of Theoretical Methods*. Taylor & Francis Group, 2017.
- [49] C. Weedbrook et al. “Gaussian Quantum Information”. In: *Reviews of Modern Physics* 84 (2 2012), pp. 621–669. eprint: <https://arxiv.org/abs/1110.3234>.
- [50] R. F. Werner. “Quantum States with Einstein-Podolsky-Rosen Correlations Admitting a Hidden-Variable Model”. In: *Physical Review A* 40 (8 1989), pp. 4277–4281.
- [51] J. Williamson. “On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems”. In: *American Journal of Mathematics* 58 (1 1936), pp. 141–163.
- [52] Arvind et al. “The Real Symplectic Groups in Quantum Mechanics and Optics”. In: *Pramana* 45 (1995), pp. 471–497. eprint: <https://arxiv.org/abs/quant-ph/9509002>.
- [53] L. Banchi, S. L. Braunstein, and S. Pirandola. “Quantum Fidelity for Arbitrary Gaussian States”. In: *Physical Review Letters* 115 (26 2015), p. 260501. eprint: <https://arxiv.org/abs/1507.01941>.
- [54] Xiao-yu Chen. “Gaussian Relative Entropy of Entanglement”. In: *Physical Review A* 71 (6 2005), p. 062320. eprint: <https://arxiv.org/abs/quant-ph/0402109>.
- [55] O. Krueger. “Quantum Information Theory with Gaussian Systems”. PhD thesis. Technische Universität Braunschweig, 2006. eprint: https://leopard.tu-braunschweig.de/receive/dbbs_mods_00020741.
- [56] A. S. Holevo. “The Entropy Gain of Infinite-Dimensional Quantum Evolutions”. In: *Doklady Mathematics* 82 (2 2010), pp. 730–731. eprint: <https://arxiv.org/abs/1003.5765>.
- [57] G. Giedke and J. I. Cirac. “Characterization of Gaussian Operations and Distillation of Gaussian States”. In: *Physical Review A* 66 (3 2002), p. 032316. eprint: <https://arxiv.org/abs/quant-ph/0204085>.
- [58] A. S. Holevo. “On the Choi–Jamiołkowski Correspondence in Infinite Dimensions”. In: *Journal of Mathematical Physics* 52 (2011), p. 042202. eprint: <https://arxiv.org/abs/1004.0196>.
- [59] M. Horodecki. “Entanglement Measures”. In: *Quantum Information and Computation* 1.1 (2001), pp. 3–26.
- [60] G. Vidal. “Entanglement Monotones”. In: *Journal of Modern Optics* 47 (2-3 2009), pp. 355–376. eprint: <https://arxiv.org/abs/quant-ph/9807077>.
- [61] H. Umegaki. “Conditional Expectation in an Operator Algebra. IV. Entropy and Information”. In: *Kodai Mathematical Seminar Reports* 14.2 (1962), pp. 59–85.
- [62] G. Lindblad. “Entropy, Information and Quantum Measurements”. In: *Communications in Mathematical Physics* 33 (1973), pp. 305–322.
- [63] L. Lami and M. E. Shirokov. “Attainability and Lower Semi-Continuity of the Relative Entropy of Entanglement and Variations on the Theme”. In: *Annales Henri Poincaré* 24 (2023), pp. 4069–4137. eprint: <https://arxiv.org/abs/2105.08091>.
- [64] L. Lami et al. “Framework for Resource Quantification in Infinite-Dimensional General Probabilistic Theories”. In: *Physical Review A* 103 (3 2021), p. 032424. eprint: <https://arxiv.org/abs/2009.11313>.
- [65] R. E. Megginson. *An Introduction to Banach Space Theory*. Springer New York, 2012.
- [66] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer Berlin, Heidelberg, 1993.
- [67] G. Lindblad. “Completely Positive Maps and Entropy Inequalities”. In: *Communications in Mathematical Physics* 40 (1975), pp. 147–151.

- [68] A. Uhlmann. "Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory". In: *Communications in Mathematical Physics* 54 (1977), pp. 21–32.
- [69] E. Chitambar et al. "Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)". In: *Communications in Mathematical Physics* 328 (2014), pp. 303–326. eprint: <https://arxiv.org/abs/1210.4583>.
- [70] F. G. S. L. Brandão and M. B. Plenio. "Entanglement Theory and The Second Law of Thermodynamics". In: *Nature Physics* 4 (2009), pp. 873–877. eprint: <https://arxiv.org/abs/0810.2319>.
- [71] E. M. Rains. "A Rigorous Treatment of Distillable Entanglement". In: *Physical Review A* 60 (1 1999), pp. 173–178. eprint: <https://arxiv.org/abs/quant-ph/9809078>.
- [72] A. Uhlmann. "The "Transition Probability" in the State Space of a $\star$ -algebra". In: *Reports on Mathematical Physics* 9 (2 1976), pp. 273–279.
- [73] R. Josza. "Fidelity for Mixed Quantum States". In: *Journal of Modern Optics* 41 (12 1993), pp. 2315–2323.
- [74] M. Hayashi. "Another Quantum Version of Sanov Theorem". In: *Annales Henri Poincaré* (2025). eprint: <https://arxiv.org/abs/2407.18566>.
- [75] F. Buscemi and N. Datta. "The Quantum Capacity of Channels with Arbitrarily Correlated Noise". In: *IEEE Transactions on Information Theory* 56 (3 2010), pp. 1447–1460. eprint: <https://arxiv.org/abs/0902.0158>.
- [76] L. Wang and R. Renner. "One-Shot Classical-Quantum Capacity and Hypothesis Testing". In: *Physical Review Letters* 108 (20 2012), p. 200501. eprint: <https://arxiv.org/abs/1007.5456>.
- [77] M. Sion. "On General Minimax Theorems". In: *Pacific Journal of Mathematics* 8.1 (1958), pp. 171–176.
- [78] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. "General Teleportation Channel, Singlet Fraction and Quasi-Distillation". In: *Physical Review A* 60 (3 1999), pp. 1888–1898. eprint: <https://arxiv.org/abs/quant-ph/9807091>.
- [79] Michal Horodecki and Pawel Horodecki. "Reduction Criterion of Separability and Limits for a Class of Distillation Protocols". In: *Physical Review A* 59 (6 1999), pp. 4206–4216. eprint: <https://arxiv.org/abs/quant-ph/9708015>.
- [80] M. Mosonyi. "The Strong Converse Exponent of Discriminating Infinite-Dimensional Quantum States". In: *Communications in Mathematical Physics* 400 (2023), pp. 83–132. eprint: <https://arxiv.org/abs/2107.08036>.
- [81] R. F. Werner and M. M. Wolf. "Bound Entangled Gaussian States". In: *Physical Review Letters* 86 (16 2001), pp. 3658–3661. eprint: <https://arxiv.org/abs/quant-ph/0009118>.
- [82] L. Lami et al. "From Log-Determinant Inequalities to Gaussian Entanglement via Recoverability Theory". In: *IEEE Transactions on Information Theory* 63.11 (2017), pp. 7553–7568. eprint: <https://arxiv.org/abs/1703.06149>.
- [83] L. Lami et al. "Gaussian Quantum Resource Theories". In: *Physical Review A* 98 (2 2018), p. 022335. eprint: <https://arxiv.org/abs/1801.05450>.
- [84] P. Marian and T. A. Marian. "Relative Entropy Is an Exact Measure of Non-Gaussianity". In: *Physical Review A* 88 (1 2013), p. 012322. eprint: <https://arxiv.org/abs/1308.2939>.
- [85] A. S. Holevo, M. Sohma, and O. Hirota. "Capacity of Quantum Gaussian Channels". In: *Physical Review A* 59 (3 1999), pp. 1820–1828.
- [86] M. M. Wolf, G. Giedke, and J. I. Cirac. "Extremality of Gaussian Quantum States". In: *Physical Review Letters* 96 (8 2006), p. 080502. eprint: <https://arxiv.org/abs/quant-ph/0509154>.
- [87] M. E. Shirokov. "Entropic Characteristics of Subsets of States". In: *Izvestiya: Mathematics* 70 (6 2006), pp. 1265–1292. eprint: <https://arxiv.org/abs/quant-ph/0510073>.

- [88] A. S. Holevo and R. F. Werner. “Evaluating Capacities of Bosonic Gaussian Channels”. In: *Physical Review A* 63 (3 2001), p. 032312. eprint: <https://arxiv.org/abs/quant-ph/9912067>.
- [89] L. Lami, K. K. Sabapathy, and A. Winter. “All Phase-Space Linear Bosonic Channels are Approximately Gaussian Distillable”. In: *New Journal of Physics* 20 (2018), p. 113012. eprint: <https://arxiv.org/abs/1806.11042>.
- [90] S. L. Braunstein and H. J. Kimble. “Teleportation of Continuous Quantum Variables”. In: *Physical Review Letters* 80 (4 1998), pp. 869–872.
- [91] M. Christandl and A. Winter. ““Squashed Entanglement”: An Additive Entanglement Measure”. In: *Journal of Mathematical Physics* 45 (3 2004), pp. 829–840. eprint: <https://arxiv.org/abs/quant-ph/0308088>.
- [92] M. M. Wilde. “Squashed Entanglement and Approximate Private States”. In: *Quantum Information Processing* 15 (2016), pp. 4563–4580. eprint: <https://arxiv.org/abs/1606.08028>.
- [93] E. M. Rains. “Bound on Distillable Entanglement”. In: *Physical Review A* 60 (1 1999), pp. 179–184. eprint: <https://arxiv.org/abs/quant-ph/9809082>.
- [94] L. Lami and B. Regula. “Distillable Entanglement under Dually Non-Entangling Operations”. In: *Nature Communications* 15.10120 (2024). eprint: <https://arxiv.org/abs/2307.11008>.
- [95] R. Simon. “Peres-Horodecki Separability Criterion for Continuous Variable Systems”. In: *Physical Review Letters* 84 (12 2000), pp. 2726–2729. eprint: <https://arxiv.org/abs/quant-ph/9909044>.
- [96] L.-M. Duan et al. “Inseparability Criterion for Continuous Variable Systems”. In: *Physical Review Letters* 84 (12 2000), pp. 2722–2725. eprint: <https://arxiv.org/abs/quant-ph/9908056>.
- [97] A. Serafini, F. Illuminati, and S. De Siena. “Symplectic Invariants, Entropic Measures and Correlations of Gaussian States”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 37.2 (2004), pp. L21–L28. eprint: <https://arxiv.org/abs/quant-ph/0307073>.
- [98] S. Pirandola, A. Serafini, and S. Lloyd. “Correlation Matrices of Two-Mode Bosonic Systems”. In: *Physical Review A* 79 (5 2009), p. 052327. eprint: <https://arxiv.org/abs/0902.1502>.
- [99] G. Spedalieri, C. Weedbrook, and S. Pirandola. “A Limit Formula for the Quantum Fidelity”. In: *Journal of Physics A: Mathematical and Theoretical* 46.2 (2013), p. 025304. eprint: <https://arxiv.org/abs/1204.2473>.

Supplementary Material

1 Notation and Preliminaries

Let us start with the relevant mathematical preliminaries. We refer the reader to Holevo’s textbook [47] for a detailed treatment of general quantum information theory and Serafini’s textbook [48] as well as the review article by Weedbrook et al. [49] for details on the Gaussian theory.

Throughout this work, we consider quantum systems that are represented mathematically by a separable (infinite-dimensional) Hilbert space $\mathcal{H}$. Two important sets of linear operators on $\mathcal{H}$ are the set of bounded operators $\mathcal{B}(\mathcal{H})$ and the set of trace-class operators $\mathcal{T}(\mathcal{H})$. The former consists of all operators with finite operator norm $\|\cdot\|_\infty$ and the latter of those with finite trace norm $\|\cdot\|_1$.²

²The trace norm of the operator X is defined via $\|X\|_1 := \text{Tr} \left[\sqrt{XX^\dagger} \right]$.

Equipping each set with the respective norm makes them Banach spaces, and we have the duality $\mathcal{T}(\mathcal{H})^* = \mathcal{B}(\mathcal{H})$ on the level of Banach spaces.

An operator X is called positive semi-definite (PSD), denoted $X \geq 0$, if $\langle \psi | X | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$. If the inequality is strict for all nonzero $|\psi\rangle$, the operator is referred to as positive definite, denoted as $X > 0$. The set of positive semi-definite bounded operators $\mathcal{B}_+(\mathcal{H})$ forms a convex cone that induces a partial order on $\mathcal{B}(\mathcal{H})$: for $X, Y \in \mathcal{B}(\mathcal{H})$, we write $X \leq Y$ if and only if $Y - X \in \mathcal{B}_+(\mathcal{H})$. Analogously, we will denote by $\mathcal{T}_+(\mathcal{H})$ the cone of positive semi-definite trace-class operators.

A *quantum state* is represented by a density operator on $\mathcal{H}$, i.e. a positive semi-definite operator that is normalised (trace-class with unit trace). The convex set of states on $\mathcal{H}$ will be denoted by

$$\mathcal{D}(\mathcal{H}) := \left\{ X \in \mathcal{T}(\mathcal{H}) : X \geq 0, \text{Tr}[X] = 1 \right\}. \quad (1.1)$$

If the density operator is a positive definite operator, i.e. $\rho > 0$, we will also refer to it as *faithful*.

Given N quantum systems with $N \in \mathbb{N}_+$ each associated with a separable Hilbert space $\mathcal{H}_i$, the composite system is represented by the tensor product Hilbert space $\mathcal{H} = \bigotimes_{i=1}^N \mathcal{H}_i$. A fundamental phenomenon that appears in these multipartite systems is entanglement. Considering a bipartite quantum system $\mathcal{H}_A \otimes \mathcal{H}_B$, we denote with $\mathcal{S} = \mathcal{S}_{A:B}$ the set of *separable* states defined as

$$\mathcal{S}(\mathcal{H}) := \text{cl}_{\text{tn}} \left(\text{conv} \left\{ |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_B : |\psi\rangle_A \in \mathcal{H}_A, |\phi\rangle_B \in \mathcal{H}_B, \langle\psi|\psi\rangle_A = 1 = \langle\phi|\phi\rangle_B \right\} \right), \quad (1.2)$$

i.e. the closed (w.r.t. the trace norm topology) convex hull of the set of pure product states [50]. Entanglement is then defined by negation, i.e. a state is called *entangled* if it is not separable.

A *quantum channel* is a bounded linear map $\Lambda : \mathcal{T}(\mathcal{H}_1) \mapsto \mathcal{T}(\mathcal{H}_2)$ that is completely positive and trace-preserving (CPTP). Here, complete positivity means that $\Lambda \otimes \mathcal{I}_{\mathcal{H}'}$ maps positive semi-definite operators on $\mathcal{H}_1 \otimes \mathcal{H}'$ to positive semi-definite operators on $\mathcal{H}_2 \otimes \mathcal{H}'$, for all auxiliary Hilbert spaces $\mathcal{H}'$. Trace preservation in turn means that $\text{Tr}[\Lambda(X)] = \text{Tr}[X]$ for all inputs X . The set of all such operations will be denoted $\text{CPTP}(\mathcal{H}_1 \rightarrow \mathcal{H}_2)$. Moreover, we define the adjoint $\Lambda^\dagger$ of the channel Λ as the linear map $\Lambda^\dagger : \mathcal{B}(\mathcal{H}_2) \mapsto \mathcal{B}(\mathcal{H}_1)$ that satisfies $\text{Tr}[X_1 \Lambda^\dagger(Y_2)] = \text{Tr}[\Lambda(X_1) Y_2]$ for all $X_1 \in \mathcal{T}(\mathcal{H}_1)$ and $Y_2 \in \mathcal{B}(\mathcal{H}_2)$.

An important subclass of channels are *quantum measurements*. Mathematically, they are represented by a Positive Operator-Valued Measures (POVM). For a measurement with finite number of outcomes, the associated POVM is a finite collection of positive semi-definite bounded operators $M_i \geq 0$ that form a resolution of the identity, i.e. $\sum_i M_i = \text{Id}_{\mathcal{H}}$ with $\text{Id}_{\mathcal{H}}$ the identity operator on $\mathcal{H}$.

Remark 1.1. We will usually denote Hilbert spaces by the capital letters A, B, C etc. and use subscripts to denote which space an operator acts on. However, to simplify the notation, we will drop the explicit reference to the Hilbert space whenever it is clear from context.

We now specialise to bosonic continuous-variable (CV) systems with a finite number of modes. A single bosonic mode is mathematically represented by a pair of self-adjoint quadrature operators $(\hat{x}, \hat{p})$ that satisfy the *canonical commutation relations* (CCR),

$$[\hat{x}, \hat{p}] = i, \quad (1.3)$$

where we used natural units with $\hbar = 1$. The Hilbert space of each bosonic mode is separable but necessarily infinite-dimensional. Its canonical basis is given by the Fock states $\{|n\rangle\}_{n=0}^\infty$, which are the eigenvectors of the associated number operator $\hat{n} = \hat{x}^2 + \hat{p}^2 - \frac{1}{2}$.

An N -mode CV quantum system with $N \in \mathbb{N}_+$ is associated with the tensor-product Hilbert space $\mathcal{H} = \bigotimes_{i=1}^N \mathcal{H}_i$, where $\mathcal{H}_i$ denotes the infinite-dimensional separable Hilbert space of the i -th mode. Using the vector notation $\hat{\mathbf{r}} := (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_N, \hat{p}_N)^T$, we can express the CCR compactly as $[\hat{\mathbf{r}}, \hat{\mathbf{r}}^T] =$

$i\Omega$. Here, the commutator on the left side denotes a $2N \times 2N$ matrix with entries $[\hat{r}_i, \hat{r}_j]$, and on the right side we have the symplectic form

$$\Omega := \bigoplus_{j=1}^N \Omega_1 \quad \text{with} \quad \Omega_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (1.4)$$

Remark 1.2. In the following, bold symbols will denote finite-dimensional vectors and matrices.

The most important class of CV quantum states are the so-called *Gaussian* states. These are the Gibbs states of quadratic Hamiltonians in the quadrature operators and completely characterised by their first and second statistical moments. A Gaussian state $\rho_G := \rho_G[\boldsymbol{\mu}, \mathbf{V}]$ is then uniquely specified by its displacement vector $\boldsymbol{\mu}$, defined via $\mu_j := \text{Tr}[\rho \hat{r}_j]$, and *covariance matrix* $\mathbf{V}$ with entries

$$V_{j,k} := \text{Tr}[\rho \{\hat{r}_j - \mu_j, \hat{r}_k - \mu_k\}] , \quad (1.5)$$

where $\{\cdot, \cdot\}$ denotes the anti-commutator.³

The covariance matrix is a $2N \times 2N$, real and symmetric matrix which must satisfy the Heisenberg uncertainty principle

$$\mathbf{V} + i\Omega \geq 0. \quad (1.6)$$

The latter ensures that $\mathbf{V}$ is a *bona fide* quantum covariance matrix. Note that Eq. (1.6) implies in particular that the covariance matrix must be positive definite. By Williamson's theorem [51], any positive definite $\mathbf{V}$ admits a *Williamson decomposition* $\mathbf{V} = \mathbf{S}\mathbf{D}\mathbf{S}^T$, where $\mathbf{S} \in \text{Sp}(2N)$ is a symplectic matrix and $\mathbf{D} = \text{diag}(v_1, v_1, \dots, v_N, v_N)$ contains the symplectic eigenvalues. Here, the symplectic group $\text{Sp}(2N)$ ⁴ is defined as the set of transformations that preserve Ω by congruence, i.e.

$$\mathbf{S} \in \text{Sp}(2N) \iff \mathbf{S}\Omega\mathbf{S}^T = \Omega, \quad (1.7)$$

and the symplectic spectrum $\{v_i\}_{i=1}^N$ is given by the standard eigenspectrum of $|i\Omega\mathbf{V}|$.

In the following, we will denote the set of Gaussian states on $\mathcal{H}$ by $\mathcal{G}(\mathcal{H})$. According to [53, Appendix A] (see also [54–56]), we can express the density operator of an arbitrary Gaussian state in Gibbs-type exponential form as

$$\rho_G[\boldsymbol{\mu}, \mathbf{V}] = \frac{1}{Z[\mathbf{V}]} \cdot \exp \left[-(\hat{\mathbf{r}} - \boldsymbol{\mu})^T \mathbf{G}[\mathbf{V}] (\hat{\mathbf{r}} - \boldsymbol{\mu}) \right] \quad (1.8)$$

with the Gibbs matrix $\mathbf{G}[\mathbf{V}] = i\Omega \text{arcoth}(\mathbf{V}i\Omega)$ and normalisation factor $Z[\mathbf{V}] = \sqrt{\det(\frac{1}{2}(\mathbf{V} + i\Omega))}$.

Remark 1.3. We use $\exp(\cdot)$ to denote the inverse of the natural logarithm $\ln(\cdot)$.

Due to the simple description of Gaussian states in terms of their statistical moments, a large and important class of quantum channels are similarly easy to describe. By definition, a *Gaussian* channel is a channel that maps Gaussian states to Gaussian states. Its action on Gaussian states is completely characterised by two $2N \times 2N$ real matrices $\mathbf{X}$ and $\mathbf{Y}$ with $\mathbf{Y} = \mathbf{Y}^T$, which act on the statistical moments of the state as

$$\boldsymbol{\mu} \mapsto \mathbf{X}\boldsymbol{\mu} \quad \text{and} \quad \mathbf{V} \mapsto \mathbf{X}\mathbf{V}\mathbf{X}^T + \mathbf{Y}. \quad (1.9)$$

The *Choi-Jamiołkowski isomorphism* is a well-known bijective mapping between quantum states and quantum channels. For finite-dimensional systems A , it is defined via the canonical maximally

³Note that different conventions for the covariance matrix are used in the literature. We follow here the convention used in Serafini's textbook [48].

⁴We refer the interested reader to [52] for a review of the symplectic group in physics.

entangled state $\Phi_d = \frac{1}{d} \sum_{i,j=1}^d |i\rangle\langle j|_A \otimes |i\rangle\langle j|_{A'}$, where $d = \dim(A)$ and $A' \simeq A$. The infinite-dimensional analogue of Φ_d is not normalisable and thus not a valid quantum state. However, one can approach it by a limit of normalisable Gaussian states [57] (see also [58] for an alternative definition). Following an operational definition, we consider the two-mode squeezed vacuum state (TMSV) given by

$$|\Phi(r)\rangle := \frac{1}{\cosh(r)} \sum_{n=0}^{\infty} \tanh^n(r) |n\rangle_A \otimes |n\rangle_{A'} \quad (1.10)$$

with the squeezing parameter $r \in [0, \infty)$. In the limit $r \rightarrow \infty$, the state $|\Psi(r)\rangle$ tends to an evenly weighted superposition of tensor products of Fock states and hence approaches the canonical maximally entangled state on $A \otimes A'$. With this, we then define the *quasi-Choi state* of the CPTP-map $\mathcal{N}$ that acts on N bosonic modes via

$$\rho_{\mathcal{N}}(r) := (\mathcal{N} \otimes \mathcal{I})(|\Phi(r)\rangle\langle\Phi(r)|^{\otimes N}), \quad (1.11)$$

where $\mathcal{I}$ denotes the identity channel.

Given a Gaussian channel $\mathcal{N} := \mathcal{N}(X, Y)$, its quasi-Choi state $\rho_{\mathcal{N}}(r)$ is a Gaussian state with covariance matrix

$$V_{\mathcal{N}}(r) = \begin{pmatrix} \cosh(2r)XX^T + Y & \sinh(2r)X\Sigma_3 \\ \sinh(2r)\Sigma_3X^T & \cosh(2r)\Sigma_0 \end{pmatrix} \quad \text{with} \quad \Sigma_i := \bigoplus_{j=1}^N \sigma_i. \quad (1.12)$$

Here, we introduced the Pauli sigma matrices, defined as

$$\sigma_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.13)$$

2 Infinite-Dimensional Quantum Systems

We start our technical analysis with general infinite-dimensional quantum systems (of type I). Our main goal is to formally establish the reverse relative entropy of entanglement as an entanglement measure with a precise operational interpretation in the context of entanglement distillation. This extends operational interpretation done in the finite-dimensional work [9].

2.1 Reverse Relative Entropy of Entanglement

In the literature, a plethora of entanglement measures have been studied (see e.g. [59, 60] for reviews). A prominent subclass is formed by measures derived from a distance function on state space; here, the state's distance to the set of separable states (w.r.t. the chosen distance function) is used as a quantifier of the state's entanglement. In this work, we investigate one such measure derived from the relative entropy function.

Given two positive semi-definite trace class operators $X, Y \in \mathcal{T}_+(\mathcal{H})$ – with spectral decomposition $X = \sum_i x_i |e_i\rangle\langle e_i|$ and $Y = \sum_j y_j |f_j\rangle\langle f_j|$ – their *relative entropy* [61, 62] is defined as

$$D(X\|Y) := \text{Tr}[X(\log X - \log Y) + Y - X] \quad (2.1)$$

$$:= \sum_{i,j} |\langle e_i | f_j \rangle|^2 (x_i \log x_i - x_i \log y_i + y_i - x_i). \quad (2.2)$$

Hereby, the expression in the first line is to be understood as specified by the second line. As explained in [62, Section 2], the convexity of $x \log x$ ensures that all summands in Eq. (2.2) are positive.⁵ Hence, the sum is well-defined (albeit possibly infinite) and the order of summation is

⁵Explicitly, one uses that a convex differentiable function satisfies $f(y) \geq f(x) + f'(x) \cdot (y - x)$ for all x, y in its domain.

irrelevant. By convention, we set $D(0\|0) = 0$ and $D(X\|0) = +\infty$ if $X \neq 0$. It is then evident from Eq. 2.2 that a necessary condition for $D(X\|Y) < \infty$ is that $X \ll Y$, i.e. the support of the first argument is contained in the support of the second.⁶

Let us also briefly mention when we can simplify Eq. (2.2) into a more familiar form. For this, we define the (von-Neumann) *entropy* of the positive semi-definite trace-class operator X via

$$H(X) := -\text{Tr}[X \log X] := \sum_i -x_i \log x_i. \quad (2.3)$$

Note that the function $f(x) := -x \log x$ is non-negative for all $x \in [0, 1]$. Since X is trace-class, only finitely many eigenvalues can lie outside this interval; thus, the sum is well-defined (albeit possibly infinite). Provided that $H(X) < \infty$ holds, Eq. (2.2) can be reduced to the well-known expression

$$D(X\|Y) := -H(X) - \text{Tr}[X \log Y] + \text{Tr}[Y] - \text{Tr}[X] \quad (2.4)$$

(see [63, Section II.B] for more details). Under the finite entropy assumption, the above expression is well-defined as the only term that can possibly diverge is the second, which is to be understood as the series

$$-\text{Tr}[X \log Y] := -\sum_{i,j} |\langle e_i | f_j \rangle|^2 x_i \log y_i. \quad (2.5)$$

Remark 2.1. We will use $\log(\cdot)$ to denote the logarithm to base two, corresponding to the canonical choice of measuring (quantum) information in (qu-)bits.

The quantum Stein lemma [39, 40] endows the relative entropy with an operational interpretation in asymmetric quantum hypothesis testing. Consequently, we can interpret it as a statistical distance measure on the set of quantum states. However, importantly, it is not a distance function in the strict mathematical sense, as it is not symmetric in general. This asymmetry gives rise to two possible entanglement measures, depending on whether the optimisation over the separable set is carried out in the first or second argument. The standard choice is the second argument, resulting in the well-known *relative entropy of entanglement* [23, 24]. Here, we consider instead the other case, which – using the terminology of [23, 26] – yields the *reverse relative entropy of entanglement*.

Definition 2.2. Let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ be a bipartite separable (possibly infinite-dimensional) Hilbert space and $\mathcal{S} = \mathcal{S}_{A:B}$ be the set of separable states on $\mathcal{H}_{AB}$. The reverse relative entropy of entanglement of the state $\rho \in \mathcal{D}$ is then defined as

$$D(\mathcal{S}\|\rho) := \inf_{\sigma \in \mathcal{S}} D(\sigma\|\rho). \quad (2.6)$$

Remark 2.3. Note that this definition can be extended straightforwardly to the framework of general quantum resource theories (see [46] for a review). Here, one considers a set of *free states* $\mathcal{F} \subseteq \mathcal{D}$ that are said to contain no resource in the context of the specific theory. The *reverse relative entropy of resource* $D(\mathcal{F}\|\rho)$ of the state ρ is then analogously defined as the relative entropy distance to the set of free states when optimising w.r.t. the first argument.

It is straightforward to show that the reverse relative entropy of entanglement is a measure of entanglement in the resource-theoretic sense of [46]. We collect its properties in the following lemma using the language of general quantum resource theories. Note that the main computational advantage of this measure is that it is additive, therefore eliminating the need for regularisation. Additionally, we obtain a semi-continuity result that is similar to [63, Theorem 5] and will be a key ingredient in the later proofs.

⁶Note that only in the case of a finite-dimensional Hilbert space, this condition is also sufficient.

Lemma 2.4. *Let $\mathcal{H}$ be a separable (infinite-dimensional) Hilbert space and $\mathcal{F} \subseteq \mathcal{D}$ a trace-norm closed set of free states. Then, the reverse relative entropy of resource $D(\mathcal{F}||\rho)$ is a valid resource monotone. That is, the functional is*

1. *Faithful, i.e. $D(\mathcal{F}||\rho) \geq 0$ and $D(\mathcal{F}||\rho) = 0$ if and only if $\rho \in \mathcal{F}$.*
2. *Monotone under free operations, i.e. $D(\mathcal{F}||\Lambda(\rho)) \leq D(\mathcal{F}||\rho)$ for any channel Λ such that $\Lambda(\rho) \in \mathcal{F}$ for all $\rho \in \mathcal{F}$.*

Moreover, the function $\rho \mapsto D(\mathcal{F}||\rho)$ is

3. *Convex if $\mathcal{F}$ is closed under convex combinations.*
4. *Additive on tensor products if $\mathcal{F}$ is closed under tensor products and partial traces.*
5. *Lower semicontinuous in the trace norm topology if $\text{cone}(\mathcal{F}) := \{\lambda\sigma : \lambda \in [0, \infty), \sigma \in \mathcal{F}\}$ is weak*-closed. Moreover, in this case there always exists $\sigma_* \in \mathcal{F}$ such that $D(\sigma_*||\rho) = D(\mathcal{F}||\rho)$.*

In particular, all the above properties hold when specialising to $\mathcal{F} = \mathcal{S}$ (see [64, Lemma 25] for the closure of $\text{cone}(\mathcal{S})$ in the weak-topology).*

Remark 2.5. The weak*-topology on $\mathcal{T}(\mathcal{H})$ is the topology induced on the space of trace-class operators by thinking of it as the dual of the space of compact operators on $\mathcal{H}$ (see e.g. [65, Chapter 2] for details). The closure of $\text{cone}(\mathcal{F})$ can in principle be established e.g. via the method from [63, Theorem 7].

Remark 2.6. Note that in finite-dimensional resource theories, one typically demands that a resource monotone should satisfy asymptotic continuity (see e.g. [46, Section VI.A]). However, as argued in [64, Section II.C], in the infinite-dimensional setting lower semicontinuity is the more natural continuity requirement.

Proof. Most of these properties – except the lower semi-continuity result – were already established for the finite-dimensional case in [26, Section III]. For the sake of completeness, we verify that they carry over into our infinite-dimensional setting. The proof is mostly standard using the properties of the relative entropy (cf. e.g. [66, Chapter 5]). However, for our proof of lower semi-continuity, we additionally need some results from general Banach space theory (cf. [65, Chapter 2]).

1. Non-negativity follows immediately from the respective property of the relative entropy. Then, assume that $D(\mathcal{F}||\rho) = 0$. By definition of the infimum, this implies the existence of a sequence of free states $\{\sigma_k\}_{k \in \mathbb{N}}$ such that $D(\sigma_k||\rho) \rightarrow 0$ as $k \rightarrow \infty$. By the Pinsker inequality (see [66, Theorem 5.5]), this implies that

$$\lim_{k \rightarrow \infty} \frac{1}{2} \|\sigma_k - \rho\|_1 \leq \lim_{k \rightarrow \infty} \sqrt{\frac{1}{2} D(\sigma_k||\rho)} = 0. \quad (2.7)$$

Since $\mathcal{F}$ is closed in the trace norm topology (by assumption), we conclude that $\rho \in \mathcal{F}$.

2. Monotonicity is a direct consequence of the monotonicity of the relative entropy under general CPTP maps [67, 68]. Consider any $\sigma \in \mathcal{F}$, we then have for any free operation $\Lambda \in \mathcal{O} \subseteq \text{CPTP}$ that

$$D(\sigma||\rho) \geq D(\Lambda(\sigma)||\Lambda(\rho)) \geq D(\mathcal{F}||\Lambda(\rho)), \quad (2.8)$$

where the second inequality holds since $\Lambda(\sigma) \in \mathcal{F}$ (by definition). As this holds for any $\sigma \in \mathcal{F}$, we can take the infimum on the left-hand side to obtain the claim.

3. This claim follows by the argument given in [44, Section 3.2.5] using that the relative entropy is a jointly convex function (see [66, Theorem 5.4]).

Let $\rho_1, \rho_2 \in \mathcal{D}$ and $\lambda \in [0, 1]$. Due to the approximation property of the infimum, we can find for any $\varepsilon > 0$ two states $\sigma_1, \sigma_2 \in \mathcal{F}$ such that

$$D(\sigma_1 \| \rho_1) \leq D(\mathcal{F} \| \rho_1) + \varepsilon \quad \text{and} \quad D(\sigma_2 \| \rho_2) \leq D(\mathcal{F} \| \rho_2) + \varepsilon. \quad (2.9)$$

Observe that $\lambda\sigma_1 + (1 - \lambda)\sigma_2 \in \mathcal{F}$ by assumption and thus

$$D(\mathcal{F} \| \lambda\rho_1 + (1 - \lambda)\rho_2) \leq D(\lambda\sigma_1 + (1 - \lambda)\sigma_2 \| \lambda\rho_1 + (1 - \lambda)\rho_2) \quad (2.10)$$

$$\leq \lambda D(\sigma_1 \| \rho_1) + (1 - \lambda) D(\sigma_2 \| \rho_2) \quad (2.11)$$

$$\leq \lambda D(\mathcal{F} \| \rho_1) + (1 - \lambda) D(\mathcal{F} \| \rho_2) + \varepsilon, \quad (2.12)$$

where the second inequality follows from joint convexity of the relative entropy. Taking the limit $\varepsilon \rightarrow 0$ proves the claim.

4. Let $\mathcal{F} \subseteq \mathcal{D}_{AB}$ be a set of free states on a bipartite Hilbert space $\mathcal{H}_{AB}$, that is closed under tensor products and partial traces.

The closure of $\mathcal{F}$ under partial traces implies super-additivity. W.l.o.g. we may assume $D(\mathcal{F} \| \rho_A \otimes \omega_B) < \infty$. For any $\varepsilon > 0$, there exists $\sigma_{AB} \in \mathcal{F}$ that satisfies

$$D(\sigma_{AB} \| \rho_A \otimes \omega_B) \leq D(\mathcal{F} \| \rho_A \otimes \omega_B) + \varepsilon. \quad (2.13)$$

It then holds that

$$\varepsilon + D(\mathcal{F} \| \rho_A \otimes \omega_B) \geq D(\sigma_{AB} \| \rho_A \otimes \omega_B) \geq D(\sigma_A \| \rho_A) + D(\sigma_B \| \omega_B) \quad (2.14)$$

$$\geq D(\mathcal{F} \| \rho_A) + D(\mathcal{F} \| \omega_B), \quad (2.15)$$

where the second inequality holds due the super-additivity of relative entropy (see [66, Corollary 5.2.1]) and the third since the set of free states is closed under taking the partial trace (by assumption). The claim follows upon taking $\varepsilon \rightarrow 0$.

Similarly, sub-additivity is a consequence of the closure of $\mathcal{F}$ under tensor products. We may assume w.l.o.g. that both $D(\mathcal{F} \| \rho_A)$ and $D(\mathcal{F} \| \omega_B)$ are finite (since otherwise the claim holds trivially). Then, for all $\varepsilon > 0$ there exists $\sigma, \tau \in \mathcal{F}$ such that

$$D(\sigma \| \rho_A) \leq D(\mathcal{F} \| \rho_A) + \frac{\varepsilon}{2} \quad \text{and} \quad D(\tau \| \omega_B) \leq D(\mathcal{F} \| \omega_B) + \frac{\varepsilon}{2}. \quad (2.16)$$

Since both relative entropies are finite (by assumption), the support condition implies that $\sigma \in \mathcal{D}_A$ and $\tau \in \mathcal{D}_B$. Consequently, $\sigma \otimes \tau \in \mathcal{F}$ by our assumption and the claim follows from

$$D(\mathcal{F} \| \rho_A \otimes \omega_B) \leq D(\sigma \otimes \tau \| \rho_A \otimes \omega_B) = D(\mathcal{F} \| \rho_A) + D(\mathcal{F} \| \omega_B) + \varepsilon, \quad (2.17)$$

where in the second step we used the additivity of the relative entropy on tensor product states (see [66, Corollary 5.2.1]).

5. The result is similar to [63, Theorem 5], but the proof works a bit differently.

Let $\{\rho_m\}_{m \in \mathbb{N}}$ be a sequence of bipartite states ρ_m that converges to some $\rho \in \mathcal{D}_{AB}$ in the trace-norm topology, i.e. $\lim_{m \rightarrow \infty} \frac{1}{2} \|\rho_m - \rho\|_1 = 0$. We need to prove that

$$D(\mathcal{F} \| \rho) \leq \liminf_{m \rightarrow \infty} D(\mathcal{F} \| \rho_m). \quad (2.18)$$

If the right-hand side is infinite, there is nothing to prove. Otherwise, we can assume w.l.o.g. (up to extracting sub-sequences) that

$$\lim_{m \rightarrow \infty} D(\mathcal{F} \| \rho_m) = R < \infty. \quad (2.19)$$

Using the approximation property of infimum, we can then construct a sequence of free states $\{\sigma_m\}_{m \in \mathbb{N}}$ such that

$$D(\mathcal{F} \parallel \rho_m) \leq D(\sigma_m \parallel \rho_m) \leq D(\mathcal{F} \parallel \rho_m) + \frac{1}{m} \leq R + 1 < \infty \quad (2.20)$$

holds for infinitely many m .

Observe that the set of sub-normalised free states $\mathcal{F}_\leq := \{\lambda\sigma : \lambda \in [0, 1], \sigma \in \mathcal{F}\}$ is weak*-compact and hence sequentially weak*-compact.⁷ To verify this, note that we can write $\mathcal{F}_\leq = \text{cone}(\mathcal{F}) \cap B_1$, where $B_1 := \{X \in \mathcal{T}(\mathcal{H}) : \|X\|_1 \leq 1\}$ is weak*-compact by the Banach-Alaoglu theorem [65, Theorem 2.6.18], and $\text{cone}(\mathcal{F})$ is per assumption weak*-closed. It is then an elementary fact from topology that the intersection of a compact set with a closed set is itself compact.

Since $\mathcal{F}_\leq$ is sequentially weak*-compact, we can extract from $\{\sigma_m\}_{m \in \mathbb{N}}$ a sub-sequence $\{\sigma_{m_k}\}_{k \in \mathbb{N}}$ such that $\sigma_{m_k} \rightarrow \lambda\sigma_*$ converges in the weak*-topology as $k \rightarrow \infty$, for some state $\sigma_* \in \mathcal{F}$ and some real coefficient $\lambda \in [0, 1]$.⁸ If we could show that $\lambda = 1$ we would be done, because then the claim follows from

$$D(\mathcal{F} \parallel \rho) \stackrel{(1)}{\leq} D(\sigma_* \parallel \rho) \stackrel{(2)}{\leq} \liminf_{k \rightarrow \infty} D(\sigma_{m_k} \parallel \rho_{m_k}) \stackrel{(3)}{\leq} \liminf_{k \rightarrow \infty} \left(D(\mathcal{F} \parallel \rho_{m_k}) + \frac{1}{m_k} \right) \stackrel{(4)}{=} R. \quad (2.21)$$

Here, (1) uses the assumption that $\lambda = 1$, (2) follows from the lower semi-continuity of the relative entropy in the product weak*-topology (cf. [63, Lemma 4]), (3) by construction of the sequence $\{\sigma_m\}_{m \in \mathbb{N}}$ and (4) since the limit exists and we can thus take it along any subsequence.

Now, the point of the proof is to show that indeed $\lambda = 1$. For this, let $\{P_m\}_{m \in \mathbb{N}}$ be a sequence of finite-rank projectors that converges strongly to the identity on $A \otimes B$, i.e. $P_m \xrightarrow{s} \text{Id}_{AB}$ as $m \rightarrow \infty$. Note that this directly implies $\lim_{m \rightarrow \infty} \text{Tr}[P_m \rho] = 1$ for any state $\rho \in \mathcal{D}$.

For all sufficiently large k and m , we then have that

$$R + 1 \stackrel{(1)}{\geq} D(\sigma_{m_k} \parallel \rho_{m_k}) \quad (2.22)$$

$$\stackrel{(2)}{\geq} D_{\text{bin}}(\text{Tr}[P_m \sigma_{m_k}] \parallel \text{Tr}[P_m \rho_{m_k}]) \quad (2.23)$$

$$\stackrel{(3)}{\geq} -1 - (1 - \text{Tr}[P_m \sigma_{m_k}]) \log(1 - \text{Tr}[P_m \rho_{m_k}]), \quad (2.24)$$

where (1) follows by Eq. (2.20), in (2) we used the monotonicity of the relative entropy and in (3) the elementary inequality

$$D_{\text{bin}}(q \parallel p) = -h_{\text{bin}}(q) - q \log p - (1 - q) \log(1 - p) \geq -1 - (1 - q) \log(1 - p). \quad (2.25)$$

Here, $D_{\text{bin}}(q \parallel p)$ denotes the relative entropy between the binary probability distributions $(q, 1 - q)$ and $(p, 1 - p)$. Re-arranging, we find that

$$1 - \text{Tr}[P_m \sigma_{m_k}] \leq \frac{R + 2}{-\log(1 - \text{Tr}[P_m \rho_{m_k}])}. \quad (2.26)$$

We can now take the limit $k \rightarrow \infty$ and – using that P_m is finite-rank and thus compact – the above inequality becomes

$$1 - \lambda \text{Tr}[P_m \sigma_*] \leq \frac{R + 2}{-\log(1 - \text{Tr}[P_m \rho])} \quad (2.27)$$

⁷It is a known fact that given a separable Hilbert space any weak*-compact subset of $\mathcal{T}(\mathcal{H})$ is also sequentially weak*-compact (cf. [63, Remark 1]).

⁸In the case $\lambda = 0$, it would be more accurate to say that we can choose the state to be free.

by the weak*-convergence of the sub-sequence $\{\sigma_{m_k}\}_{k \in \mathbb{N}}$. Finally, taking $m \rightarrow \infty$ and using the fact that $\lim_{m \rightarrow \infty} \text{Tr}[P_m \sigma_*] = 1 = \lim_{m \rightarrow \infty} \text{Tr}[P_m \rho]$ yields $1 - \lambda \leq 0$, which together with $\lambda \leq 1$ shows that indeed $\lambda = 1$.

Lastly, the attainability of the minimum follows by specialising the above argument to the particular sequence $\rho_m = \rho$ for all m . In this case, $R = D(\mathcal{F} \parallel \rho)$ and Eq. 2.21 immediately implies that $D(\sigma_* \parallel \rho) = D(\mathcal{F} \parallel \rho)$ proving attainability.

□

2.2 The Error Exponent of Entanglement Distillation and Entanglement Testing

A practically meaningful entanglement measure should admit an operational interpretation, i.e. it should be directly linked to a concrete task in quantum information processing. For the reverse relative entropy of entanglement, this was established for finite-dimensional systems by Lami et al. in [9] in the context of entanglement distillation. We will now extend this operational interpretation to the general infinite-dimensional setting.

Let us start with a formal description of entanglement distillation. Two parties, called Alice and Bob, are given n copies of a bipartite quantum state ρ_{AB} on a separable Hilbert space $A \otimes B$. Their task is to convert the given state into as much pure entanglement as possible. To be precise, the goal is to maximise the number of copies m of the maximally entangled state on the fixed two-qubit system $A_0 \otimes B_0$, i.e. $|\Phi\rangle_{A_0 B_0} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Moreover, we do not require that they achieve their goal perfectly, but allow for some small non-zero error ε_n . Thus, the goal is to find a quantum channel Λ that achieves

$$\Lambda(\rho_{AB}^{\otimes n}) \approx_{\varepsilon_n} \Phi_{A_0 B_0}^{\otimes m} \quad \text{with} \quad \Phi_{A_0 B_0} := |\Phi\rangle\langle\Phi|_{A_0 B_0}. \quad (2.28)$$

In order to get an interesting theory, not all possible quantum channels Λ are allowed to achieve this goal, but only a subset are deemed *free operations* $\mathcal{O}$. Historically, entanglement distillation was studied with $\mathcal{O}$ given by the *local operations and classical communication* (LOCC) paradigm [16, 18, 19]. Although operationally well-defined, the set of LOCC channels has a very complicated mathematical structure [69]. Therefore, we consider instead the largest physically consistent class of transformations. Specifically, we consider the set of *non-entangling* operations, denoted NE, which encompasses all CPTP maps that do not add entanglement to the state [25, 70]. Formally, we define

$$\text{NE}_{n \rightarrow m} := \left\{ \Lambda \in \text{CPTP}(A^n B^n \rightarrow A_0^m B_0^m) : \Lambda(\sigma) \in \mathcal{S}_{A_0^m B_0^m} \forall \sigma \in \mathcal{S}_{A^n B^n} \right\}. \quad (2.29)$$

The last piece of the theory is a figure-of-merit that allows us to compare the performance of two different distillation protocols. For this, we require that the associated error satisfies $\varepsilon_n \rightarrow 0$ as $n \rightarrow \infty$, i.e. as more and more copies are available the quality of distilled entanglement should increase, becoming perfect in the asymptotic limit. Previous works then focussed on the *quantity* of the obtained entanglement, i.e. the asymptotic yield of the protocol was used to quantify performance. This in turn leads to the notion of *distillable entanglement* as figure-of-merit (see e.g. [71] for a formal definition).

In this work, we instead follow [9] and consider the *quality* of the distilled entanglement as our performance quantifier. This means we study the distillation error exponent – that is, we ask the question how fast the quality of the distilled entanglement improves. Specifically, we require that the error behaves as $\varepsilon \sim 2^{-c \cdot n}$ and characterise the optimal achievable exponent c . Formally, we define the *distillable entanglement error exponent* for m copies via

$$E_{d, \text{err}}^{(m)}(\rho_{AB}) := \sup \left\{ \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \varepsilon_n : F(\Lambda_n(\rho_{AB}^{\otimes n}), \Phi_{A_0 B_0}^{\otimes m}) \geq 1 - \varepsilon_n, \Lambda_n \in \text{NE}_{n \rightarrow m} \right\} \quad (2.30)$$

with the Uhlmann fidelity function $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$ [72, 73].⁹ The asymptotic error exponent of entanglement distillation is then given by

$$E_{d,\text{err}}(\rho) = \lim_{m \rightarrow \infty} E_{d,\text{err}}^{(m)}(\rho). \quad (2.31)$$

The error exponent of entanglement distillation is closely connected to another primitive in quantum information theory: quantum state discrimination [39–41]. Specifically, Lami et. al. [9] showed that in finite-dimensions this exponent exactly coincides with the *Sanov exponent* of the composite testing problem known as *entanglement testing*. In entanglement testing, the goal is to determine if a given quantum state is entangled or not.

More formally, the null hypothesis is that the unknown state is given by $\rho_{AB}^{\otimes n}$, i.e. n i.i.d. copies of the bipartite entangled state ρ_{AB} . The alternative is that the state is separable w.r.t. the bipartite cut $A_n : B_n$, i.e. the whole set $\mathcal{S}_{A_n:B_n}$. Crucially, while the null hypothesis is simple the alternative is composite, significantly complicating the analysis. The task is then to discriminate between these two hypotheses using a measurement on the global system. Mathematically, this can be modelled by a binary POVM T_n on the whole Hilbert space, a so-called *quantum test*, where the measurement outcome corresponds to the acceptance or rejection of the null hypothesis.

As usual, we can associate two types of error with each test T_n . The type-1 error occurs when we mistake the entangled state $\rho_{AB}^{\otimes n}$ for a separable state; while the type-2 error occurs when we identify a separable state with the entangled $\rho_{AB}^{\otimes n}$. Requiring that the type-2 error is bounded by some nonzero threshold ε , we ask for the optimal *asymptotic* decay of the type-1 error as $\varepsilon \rightarrow 0$. This is termed the (zero-rate) *Sanov exponent* of entanglement testing.¹⁰

At this point, it is convenient to introduce the hypothesis-testing relative entropy [75, 76] given by

$$D_H^\varepsilon(\sigma \parallel \rho) := -\log \inf \left\{ \text{Tr}[M\rho] : 0 \leq M \leq \text{Id}, \text{Tr}[(\text{Id} - M)\sigma] \leq \varepsilon \right\}. \quad (2.32)$$

If we interpret M as an arbitrary POVM element, we can understand the pair $\{M, \text{Id} - M\}$ as the most general test we can use to discriminate between ρ and σ . Assigning the first outcome of this measurement to the state σ and the second to ρ , $D_H^\varepsilon(\sigma \parallel \rho)$ quantifies exactly the optimal type-1 error of testing ρ against σ given a threshold of ε on the type-2 error. Based on this, we can formally define the hypothesis-testing relative entropy of entanglement testing via

$$D_H^\varepsilon(\mathcal{S}_{A:B} \parallel \rho_{AB}) := \inf_{\sigma \in \mathcal{S}_{A:B}} D_H^\varepsilon(\sigma \parallel \rho_{AB}) \quad (2.33)$$

$$= -\log \sup_{\sigma \in \mathcal{S}_{A:B}} \inf \left\{ \text{Tr}[M\rho_{AB}] : 0 \leq M \leq \text{Id}, \text{Tr}[M\sigma] \geq 1 - \varepsilon \right\}. \quad (2.34)$$

The Sanov exponent of entanglement testing can then formally be defined as

$$\text{Sanov}(\rho_{AB} \parallel \mathcal{S}_{A:B}) := \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n} \parallel \rho_{AB}^{\otimes n}). \quad (2.35)$$

The importance of the Sanov exponent in our analysis stems from the following lemma, that generalises [9, Lemma 1] to the infinite-dimensional setting and establishes a direct connection between the Sanov exponent and the error exponent of entanglement distillation.

Lemma 2.7. *Let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ be a bipartite separable (infinite-dimensional) Hilbert space and let $\mathcal{D}_{AB}$ be the set of quantum states. Using the definitions introduced above, the asymptotic error exponent of entanglement distillation under non-entangling operations equals the Sanov exponent of entanglement testing. Formally, we have for any $\rho_{AB} \in \mathcal{D}_{AB}$ that*

$$E_{d,\text{err}}(\rho_{AB}) = \text{Sanov}(\rho_{AB} \parallel \mathcal{S}_{A:B}). \quad (2.36)$$

⁹Note that we could equivalently use the trace-norm distance to quantify the ε_n -closeness.

¹⁰Note that this is conceptually different from other versions of the quantum Sanov theorem such as [15, 42, 43, 74].

The proof generally follows along the same lines as the finite-dimensional one for [9, Lemma 1], because the output space is always finite-dimensional. We provide a full proof below for the sake of completeness, where we additionally simplify the second part of the original argument. However, in order to apply their proof, we first need to verify that a crucial rewriting of the hypothesis-testing divergence also goes through in the infinite-dimensional case.

Lemma 2.8. *Using the definitions introduced above, we have*

$$D_H^\varepsilon(\mathcal{S}_{A:B} \parallel \rho_{AB}) = -\log \min \left\{ \text{Tr}[M\rho_{AB}] : 0 \leq M \leq \text{Id}, \text{Tr}[(\text{Id} - M)\sigma] \leq \varepsilon \forall \sigma \in \mathcal{S}_{A:B} \right\}. \quad (2.37)$$

Proof. The general idea for the proof comes from [14, Lemma 1].

To begin, let us introduce some notation for the two relevant sets of quantum tests:

$$\mathfrak{T}_{\varepsilon, \sigma} := \left\{ T \in \mathcal{B}(\mathcal{H}) : 0 \leq T \leq \text{Id}, \text{Tr}[(\text{Id} - T)\sigma] \leq \varepsilon \right\} \quad (2.38)$$

and

$$\mathfrak{T}_{\varepsilon, \mathcal{S}} := \left\{ T \in \mathcal{B}(\mathcal{H}) : 0 \leq T \leq \text{Id}, \text{Tr}[(\text{Id} - T)\sigma] \leq \varepsilon \forall \sigma \in \mathcal{S} \right\}. \quad (2.39)$$

With this, the claim of the lemma can be rewritten as

$$\sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\varepsilon, \sigma}} \text{Tr}[T\rho] = \min_{T \in \mathfrak{T}_{\varepsilon, \mathcal{S}}} \text{Tr}[T\rho]. \quad (2.40)$$

We first argue that both minima are indeed attained. For this, we work with the weak*-topology on $\mathcal{B}(\mathcal{H})$ (induced on $\mathcal{B}(\mathcal{H})$ by its pre-dual $\mathcal{T}(\mathcal{H})$, cf. also Remark 2.5). By the Banach-Alaoglu theorem [65, Theorem 2.6.18], the operator interval $0 \leq T \leq \text{Id}$, i.e. the set of POVM elements, is compact in this topology. Moreover, by definition all functionals of the form $T \mapsto \text{Tr}[T\rho]$ with $\rho \in \mathcal{T}(\mathcal{H})$ are continuous. Both constraints thus define a weak*-closed half-space. Consequently, each set of quantum tests can be written as the intersection of a compact with a closed set, and is therefore compact as well. As we minimise a continuous function w.r.t. to a compact set, both minima are indeed attained.

Now, since $\mathfrak{T}_{\varepsilon, \mathcal{S}} \subseteq \mathfrak{T}_{\varepsilon, \sigma}$, it immediately follows by a subset argument that

$$\sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\varepsilon, \sigma}} \text{Tr}[T\rho] \leq \min_{T \in \mathfrak{T}_{\varepsilon, \mathcal{S}}} \text{Tr}[T\rho]. \quad (2.41)$$

For the reverse direction, consider any $\sigma \in \mathcal{S}$ and note that

$$\min_{T \in \mathfrak{T}_{\varepsilon, \sigma}} \text{Tr}[T\rho] \leq \sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\varepsilon, \sigma}} \text{Tr}[T\rho] =: \delta. \quad (2.42)$$

Note that the optimal test T^* on the LHS in Eq. (2.42) satisfies $\text{Tr}[T^*\rho] \leq \delta$ and $\text{Tr}[(\text{Id} - T^*)\sigma] \leq \varepsilon$. Therefore, $\text{Id} - T^* \in \mathfrak{T}_{\delta, \rho}$ and we have

$$\min_{T \in \mathfrak{T}_{\delta, \rho}} \text{Tr}[T\sigma] \leq \text{Tr}[(\text{Id} - T^*)\sigma] \leq \varepsilon \quad (2.43)$$

As this holds for arbitrary $\sigma \in \mathcal{S}$, we must have

$$\sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\delta, \rho}} \text{Tr}[T\sigma] \leq \varepsilon. \quad (2.44)$$

Now, note that $\mathcal{S}$ is a convex subset of the space of trace class operators (endowed with the trace norm topology) and $\mathfrak{T}_{\delta,\rho}$ is a convex and compact subset of the space of bounded linear operators (endowed with the weak*-topology). Moreover, $T \mapsto \text{Tr}[T\sigma]$ is a linear and weak*-continuous function for all $\sigma \in \mathcal{S}$ and $\sigma \mapsto \text{Tr}[T\sigma]$ is a linear and trace-norm-continuous function for all quantum tests T .¹¹ Consequently, we can apply Sion's minimax theorem [77] and interchange the optimisations to obtain

$$\min_{T \in \mathfrak{T}_{\delta,\rho}} \sup_{\sigma \in \mathcal{S}} \text{Tr}[T\sigma] = \sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\delta,\rho}} \text{Tr}[T\sigma] \leq \varepsilon. \quad (2.46)$$

This implies that there exists a test T^* such that

$$\sup_{\sigma \in \mathcal{S}} \text{Tr}[T^*\sigma] \leq \varepsilon \quad \text{and} \quad \text{Tr}[(\text{Id} - T^*)\rho] \leq \delta. \quad (2.47)$$

Note that $\text{Id} - T^* \in \mathfrak{T}_{\varepsilon,\mathcal{S}}$ and we can complete the proof with

$$\min_{T \in \mathfrak{T}_{\varepsilon,\mathcal{S}}} \text{Tr}[T\rho] \leq \text{Tr}[(\text{Id} - T^*)\rho] \leq \delta = \sup_{\sigma \in \mathcal{S}} \min_{T \in \mathfrak{T}_{\varepsilon,\sigma}} \text{Tr}[T\rho]. \quad (2.48)$$

Together with Eq. (2.41) this establishes the claimed equality. $\square$

Proof of Lemma 2.7. First, we show how to construct a feasible test for entanglement testing from any feasible distillation protocol. Consider an arbitrary non-entangling distillation protocol $\Lambda_n \in \text{NE}_{n \rightarrow m}$ that achieves an error ε_n . By definition, we have

$$1 - \varepsilon_n \leq F\left(\Lambda_n(\rho_{AB}^{\otimes n}), \Phi_{A_0 B_0}^{\otimes m}\right) = \text{Tr}\left[\Lambda_n(\rho_{AB}^{\otimes n}) \Phi_{A_0 B_0}^{\otimes m}\right] = \text{Tr}\left[\rho_{AB}^{\otimes n} \Lambda_n^\dagger(\Phi_{A_0 B_0}^{\otimes m})\right], \quad (2.49)$$

where in the first equality we used that the second argument is pure and in the second step we introduced the adjoint $\Lambda_n^\dagger$ of the protocol. Since $\Lambda_n^\dagger$ is completely positive and unital as the adjoint of a CPTP map [47], we can identify the operator $M_n := \text{Id} - \Lambda_n^\dagger(\Phi_{A_0 B_0}^{\otimes m})$ as a valid POVM element on $\mathcal{H}^{\otimes n}$. Moreover, for any $\sigma_n \in \mathcal{S}_{A_n:B_n}$, it holds that

$$\text{Tr}[M_n \sigma_n] = 1 - \text{Tr}\left[\Lambda_n(\sigma_n) \Phi_{A_0 B_0}^{\otimes m}\right] \geq 1 - 2^{-m}, \quad (2.50)$$

where we observed that $\Lambda_n(\sigma_n)$ is separable since the channel is non-entangling and then invoked a standard bound on the so-called singlet fraction in a separable state from [78]. Note that the latter directly applies as the output space of Λ_n is finite-dimensional.

Thus, the binary POVM $\{M_n, \text{Id} - M_n\}$ is a feasible test for the hypothesis-testing relative entropy of entanglement testing with a type-2 error threshold of 2^{-m} . Consequently, we obtain the general bound

$$\inf_{\sigma_n \in \mathcal{S}_{A_n:B_n}} D_H^{2^{-m}}(\sigma_n \| \rho_{AB}^{\otimes n}) \geq -\log \text{Tr}[M_n \rho_{AB}^{\otimes n}] \geq -\log \varepsilon_n. \quad (2.51)$$

Dividing by n , taking the limit $n \rightarrow \infty$ and finally the supremum over all non-entangling protocols $\{\Lambda_n\}_{n \in \mathbb{N}}$ yields

$$E_{d,\text{err}}^{(m)}(\rho_{AB}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \inf_{\sigma_n \in \mathcal{S}_{A:B}} D_H^{2^{-m}}(\sigma_n \| \rho_{AB}^{\otimes n}) = \liminf_{n \rightarrow \infty} \frac{1}{n} D_H^{2^{-m}}(\mathcal{S}_{A_n:B_n} \| \rho_{AB}^{\otimes n}). \quad (2.52)$$

¹¹Continuity holds in the first case by definition. In the second, this follows from Hölder's inequality stating that

$$\|\text{Tr}[T\sigma] - \text{Tr}[T\rho]\| \leq \|T\|_\infty \cdot \|\sigma - \rho\|_1 \quad (2.45)$$

For the other direction, we construct a distillation protocol from any feasible test for entanglement testing. Consider any test operator M_n satisfying $\text{Tr}[M_n \sigma_n] \geq 1 - 2^{-m}$ for all $\sigma_n \in \mathcal{S}_{A_n:B_n}$. With this, we construct the CPTP map $\Lambda_n \in \text{NE}_{n \rightarrow m}$ via the mapping

$$X_{A_n B_n} \mapsto \Lambda_n(X_{A_n B_n}) = \text{Tr}[(\text{Id} - M_n)X_{A_n B_n}] \Phi_{A_0 B_0}^{\otimes m} + \text{Tr}[M_n X_{A_n B_n}] \frac{\text{Id} - \Phi_{A_0 B_0}^{\otimes m}}{2^{2m} - 1}. \quad (2.53)$$

To verify that this map is non-entangling, let us first observe that the output of the map is by construction an isotropic state. It is known that for isotropic states separability is completely characterised in terms of the singlet fraction [79]. Now, observe that the singlet fraction at the output satisfies

$$F(\Lambda_n(\sigma_n), \Phi_{A_0 B_0}^{\otimes m}) = 1 - \text{Tr}[M_n \sigma_n] \leq 2^{-m} \quad (2.54)$$

by assumption for any separable state σ_n . This in turn is equivalent to the separability of $\Lambda_n(\sigma_n)$ [79]. As this holds for any separable input state, the map is indeed non-entangling. By construction, we then have $F(\Lambda_n(\rho_{AB}^{\otimes n}), \Phi_{A_0 B_0}^{\otimes m}) = 1 - \text{Tr}[M_n \rho_{AB}^{\otimes n}]$. Consequently, we obtain a feasible protocol for entanglement distillation that achieves an error $\varepsilon_n \leq \text{Tr}[M_n \rho_{AB}^{\otimes n}]$.

Now, picking the sequence of optimal tests M_n^* that achieves the hypothesis-testing relative entropy we obtain the general bound

$$E_{d,\text{err}}^{(m)}(\rho_{AB}) \geq \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr}[M_n^* \rho_{AB}^{\otimes n}] = \liminf_{n \rightarrow \infty} \frac{1}{n} D_H^{2^{-m}}(\mathcal{S}_{A_n:B_n} \parallel \rho_{AB}^{\otimes n}). \quad (2.55)$$

Finally, we observe that the function $\varepsilon \rightarrow D_H^\varepsilon(\mathcal{S}_{A_n:B_n} \parallel \rho_{AB}^{\otimes n})$ is monotone non-decreasing. Therefore, the left-sided limit $\varepsilon \rightarrow 0$ exists and we may take it along any sequence, in particular $\varepsilon_m = 2^{-m}$ with $m \rightarrow \infty$.

Combining all partial results, we finally conclude that

$$E_{d,\text{err}}(\rho_{AB}) = \lim_{m \rightarrow \infty} E_{d,\text{err}}^{(m)}(\rho) = \lim_{m \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} D_H^{2^{-m}}(\mathcal{S}_{A_n:B_n} \parallel \rho_{AB}^{\otimes n}) = \text{Sanov}(\rho_{AB} \parallel \mathcal{S}_{A:B}). \quad (2.56)$$

□

2.3 Generalised Sanov Theorem of Entanglement Testing

The main technical result of this first part of the manuscript is then the extension of [9, Corollary 15] to infinite-dimensional systems. Specifically, we establish that the Sanov exponent of entanglement testing is exactly given by the single-letter reverse relative entropy of entanglement. By the previous analysis, this endows the reverse relative entropy of entanglement with an operational interpretation in entanglement distillation as well, independent of the Hilbert space dimension.

Theorem 2.9. *Let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ be a bipartite separable (infinite-dimensional) Hilbert space and let $\mathcal{D}_{AB}$ be the set of quantum states on $\mathcal{H}_{AB}$. Using the definitions from above, for any state $\rho_{AB} \in \mathcal{D}_{AB}$, it holds that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n} \parallel \rho_{AB}^{\otimes n}) = D(\mathcal{S}_{A:B} \parallel \rho_{AB}) \quad \forall \varepsilon \in (0, 1). \quad (2.57)$$

Consequently, we have the following equalities:

$$E_{d,\text{err}}(\rho_{AB}) = \text{Sanov}(\rho_{AB} \parallel \mathcal{S}_{A:B}) = D(\mathcal{S}_{A:B} \parallel \rho_{AB}). \quad (2.58)$$

Let us begin with a small technical lemmata that we will need in our proof of the main theorem.

Lemma 2.10. Let $D_{\text{bin}}(q\|p)$ denote the relative entropy between the binary probability distributions $(q, 1 - q)$ and $(p, 1 - p)$. Then, for all $T \in \mathbb{R}$ we have

$$\lim_{p \rightarrow 1^-} \inf_{q \in [0,1]} \{D_{\text{bin}}(q\|p) + qT\} = T. \quad (2.59)$$

Proof. For all $p \in (0, 1)$, the function $q \rightarrow D_{\text{bin}}(q\|p) + qT$ is convex and differentiable on the domain $(0, 1)$. Its minimum can be found by setting the derivative to zero; it is achieved at

$$q = q(p, T) := \frac{p}{2^T(1 - p) + p}. \quad (2.60)$$

Since $\lim_{p \rightarrow 1^-} q(p, T) = 1$, by lower semi-continuity we see that

$$\liminf_{p \rightarrow 1^-} \inf_{q \in [0,1]} \{D_{\text{bin}}(q\|p) + qT\} = \liminf_{p \rightarrow 1^-} \{D_{\text{bin}}(q(p, T)\|p) + q(p, T)T\} \geq D_{\text{bin}}(1\|1) + T = T. \quad (2.61)$$

The converse bound can be obtained with the simple ansatz $q = p$. $\square$

Proof of Theorem 2.9. First, note that the finite-dimensional case is exactly [9, Corollary 15].

The converse is an immediate consequence of the standard quantum Stein's lemma in infinite dimensions. Note that here the strong converse exponents are known as well (see [80] for more details). For any given separable state $\sigma \in \mathcal{S}_{A:B}$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n} \|\rho^{\otimes n}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\sigma^{\otimes n} \|\rho^{\otimes n}) = D(\sigma\|\rho) \quad (2.62)$$

and the claim then follows by taking the infimum over $\sigma \in \mathcal{S}_{A:B}$ on the right-hand side.

The achievability of this exponent is obtained by lifting the finite-dimensional result. Let us start by considering a finite-rank tensor product projector $P = \Pi_{A \rightarrow A'} \otimes \Pi_{B \rightarrow B'}$ and apply to every copy the LOCC channel $\Lambda : AB \rightarrow A'X_A : B'X_B$ defined via

$$\rho \mapsto \Lambda(\rho) = P\rho P^\dagger \otimes |00\rangle\langle 00|_{X_A X_B} + \tau \cdot \text{Tr}[(\text{Id} - P^\dagger P)\rho] \otimes |11\rangle\langle 11|_{X_A X_B}, \quad (2.63)$$

where $\tau = \tau_{A'B'}$ is an arbitrary separable state on the finite-dimensional bipartite space $A' \otimes B'$ where P is supported and X_A and X_B are fixed classical single-bit systems.

Observe that this channel can be implemented by the following LOCC protocol: Alice and Bob perform the local test associated with the finite-rank projector and communicate their outcomes to each other. If both tests succeeded, they keep the state and set their classical registers to 0. Otherwise, they discard the state, prepare the fixed separable state τ and then set their classical registers to 1.

We can then write

$$\liminf_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n} \|\rho^{\otimes n}) \stackrel{(1)}{\geq} \liminf_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A'_n X_A : B'_n X_B} \|\Lambda(\rho)^{\otimes n}) \quad (2.64)$$

$$\stackrel{(2)}{=} D(\mathcal{S}_{A'X_A : B'X_B} \|\Lambda(\rho)) \quad (2.65)$$

$$\stackrel{(3)}{=} \inf_{\substack{\sigma_1, \sigma_2 \in \mathcal{S}_{A':B'} \\ q \in [0,1]}} D\left(q\sigma_1 \otimes |00\rangle\langle 00| + (1 - q)\sigma_2 \otimes |11\rangle\langle 11| \|\Lambda(\rho)\right) \quad (2.66)$$

$$\stackrel{(4)}{=} \inf_{\substack{\sigma_1, \sigma_2 \in \mathcal{S}_{A':B'} \\ q \in [0,1]}} D_{\text{bin}}\left(q\left\|\text{Tr}\left[P^\dagger P\rho\right]\right.\right) + qD(\sigma_1\|\rho_P) + (1 - q)D(\sigma_2\|\tau) \quad (2.67)$$

$$\stackrel{(5)}{=} \inf_{q \in [0,1]} D_{\text{bin}}\left(q\left\|\text{Tr}\left[P^\dagger P\rho\right]\right.\right) + qD(\mathcal{S}_{A':B'}\|\rho_P). \quad (2.68)$$

Here, (1) follows from monotonicity under LOCC maps and in (2) we applied the finite-dimensional generalised Sanov theorem from [9, Corollary 15]. To see that we used the most general ansatz in (3), recall that a necessary condition to keep the relative entropy finite is that the support of the first argument has to be contained in the support of the second. Any non-zero cross term in the classical register would violate this support condition. Furthermore, the state σ_1 is the post-measurement state when Alice and Bob both measure 0 in their classical register. As the global state is separable it therefore has to be separable as well, since it is not possible to create entanglement with local measurements and post-selection alone. The same reasoning then shows that σ_2 is also separable. In (4), we expanded the previous expression and introduced the simplifying notation

$$\rho_P := \frac{P\rho P^\dagger}{\text{Tr}[P^\dagger P\rho]}, \quad (2.69)$$

and in (5) we simply observed that the choice $\sigma_2 = \tau$ is optimal.

In the last line, we can now replace $D(\mathcal{S}_{A':B'}\|\rho_P)$ w.l.o.g. by $D(\mathcal{S}_{A:B}\|\rho_P)$. To see this, observe that the support of ρ_P is by construction contained in $A' \otimes B'$. Any $\sigma \in \mathcal{S}_{A:B}$ that has support outside of $A' \otimes B'$ leads to an infinite relative entropy, as it violates the support condition. By definition as an infimum, we can thus safely expand the feasible set without changing the optimal value.

We now take $P = P_m$ as the m -th element of a sequence $\{P_m\}_m$ of finite-rank projectors that converges strongly to the identity (note that this also implies the convergence $\rho_{P_m} \rightarrow \rho$ in trace-norm). Since $\lim_{m \rightarrow \infty} \frac{1}{2} \|\rho_{P_m} - \rho\|_1 = 0$, by the lower semi-continuity of the reverse relative entropy of entanglement (see Lemma 2.4), we have for all $\delta > 0$ that $D(\mathcal{S}_{A:B}\|\rho_{P_m}) \geq D(\mathcal{S}_{A:B}\|\rho) - \delta$ for all sufficiently large m .

Hence, the above argument shows that for all sufficiently large m we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n}\|\rho^{\otimes n}) \geq \inf_{q \in [0,1]} \left\{ D_{\text{bin}}\left(q\left\|\text{Tr}\left[P_m^\dagger P_m \rho\right]\right)\right) + q D(\mathcal{S}_{A:B}\|\rho_{P_m}) \right\} \quad (2.70)$$

$$\geq \inf_{q \in [0,1]} \left\{ D_{\text{bin}}\left(q\left\|\text{Tr}\left[P_m^\dagger P_m \rho\right]\right)\right) + q \left(D(\mathcal{S}_{A:B}\|\rho) - \delta \right) \right\}. \quad (2.71)$$

We can then take the limit $m \rightarrow \infty$ using Lemma 2.10 (note that $\text{Tr}[P_m^\dagger P_m \rho] \rightarrow 1$ from below) to deduce that for all $\delta > 0$ we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} D_H^\varepsilon(\mathcal{S}_{A_n:B_n}\|\rho^{\otimes n}) \geq D(\mathcal{S}_{A:B}\|\rho) - \delta. \quad (2.72)$$

We conclude the proof by letting $\delta \rightarrow 0$. □

Remark 2.11. Our argument extends to general quantum resource testing provided that the finite-dimensional Sanov theorem holds. We only require that the cone generated by the set of free states $\text{cone}(\mathcal{F})$ is weak*-closed (which is often the case) and the set of free operations includes our choice of measure-and-prepare channel.

3 Bosonic Continuous-Variable Systems

We now specialise our analysis to bosonic continuous-variable (CV) systems with a finite number of modes. We prove that, for Gaussian states, the reverse relative entropy of entanglement is efficiently computable by a convex program. We then derive an upper bound on the error exponent of quantum communication for the class of teleportation-simulable channels. Finally, we derive explicit analytical expressions of our bound for the most relevant one-mode Gaussian channels.

3.1 Gaussian Reverse Relative Entropy of Entanglement

Let us first revisit the definition of the reverse relative entropy of entanglement in Sec. 2.1. The issue with it is that even though the reverse relative entropy of entanglement is an operational entanglement measure, it is not efficiently computable in general (see also the discussion in the main text). However, suppose now that Alice and Bob share a bipartite quantum state of a bosonic CV system with $N = N_A + N_B$ modes, where Alice holds the first N_A modes and Bob the remaining N_B . If the quantum state they share is Gaussian, the characterisation of its separability considerably simplifies. This is because the entanglement properties of bipartite Gaussian states can be conveniently translated at the level of quantum covariance matrices in terms of simple positive-semi definite constraints [81] (see also [38] for an improvement).

In light of this simple characterisation of Gaussian separability, a common approach in the theory of entanglement measures of CV systems is to *Gaussify* the measure under consideration. This then leads to an efficiently computable entanglement measure with notable examples being the entanglement of formation [34], squashed entanglement [82] or the standard relative entropy of entanglement [54]. Naturally, we can also define a Gaussian version of the reverse relative entropy of entanglement via

$$D(\mathcal{S}_G \parallel \rho) := \min_{\sigma \in \mathcal{S}_G} D(\sigma \parallel \rho), \quad (3.1)$$

where $\mathcal{S}_G := \mathcal{S} \cap \mathcal{G}$ denotes the set states that are separable and Gaussian.

Remark 3.1. Note that both the set of Gaussian states $\mathcal{G}$ [83, Lemma 1 in Appendix A] and the separable set are closed w.r.t. the trace norm topology; hence, $\mathcal{S}_G$ is trace-norm closed as the intersection of two closed sets. Moreover, the cones generated by both sets are known to be weak*-closed (see [63, Lemma 34] for the Gaussian case). Since $\text{cone}(\mathcal{S}_G) = \text{cone}(\mathcal{S}) \cap \text{cone}(\mathcal{G})$, the cone generated by $\mathcal{S}_G$ is weak*-closed as well. By Lemma 2.4, the Gaussian relative entropy of entanglement is a lower-semicontinuous entanglement monotone and the infimum is always attained. However, note also that the set of Gaussian states is *not* convex in general.

However, a common issue with these Gaussian measures is that they typically lose their operational interpretation and are not known to coincide with their regular counterparts, except in special cases [35, 36]. However, the reverse relative entropy of entanglement is an exception to this rule. Specifically, we show below that the regular and Gaussian reverse relative entropy of entanglement coincide on a Gaussian input. This is particularly noteworthy, as it is believed that the standard relative entropy of entanglement does *not* have this property.

Lemma 3.2. *Let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ denote the Hilbert space of a bosonic $(N_A + N_B)$ -mode CV system, and let $\mathcal{S}$ and $\mathcal{G}$ denote the set of separable and Gaussian quantum states on $\mathcal{H}_{AB}$, respectively. Using the definitions introduced above, for any $\rho_G \in \mathcal{G}$, we have the equality*

$$D(\mathcal{S} \parallel \rho_G) = D(\mathcal{S}_G \parallel \rho_G), \quad (3.2)$$

where $\mathcal{S}_G := \mathcal{S} \cap \mathcal{G}$ denotes the set of states that are both separable and Gaussian.

The key technical tool in our proof is a method known as Gaussification of quantum states (see e.g. [84]). Let ρ be an arbitrary N -mode CV quantum state with finite first and second moments. We can then associate with ρ the unique Gaussian state that has the same first and second statistical moments. This is often termed the *Gaussification* of ρ and will be denoted in the following by ρ_G . In the following lemma, we collect the key properties of the Gaussification we will need.

Lemma 3.3. Let $\mathcal{H}_{AB}$ be the bipartite Hilbert space of a bosonic $(N_A + N_B)$ -mode CV system and consider an arbitrary quantum state ρ with finite first and second moment. Its Gaussification ρ_G , as defined above, satisfies the following properties:

1. $\text{Tr} [\rho \log \tau_G] = \text{Tr} [\rho_G \log \tau_G]$ for any Gaussian state $\tau_G \in \mathcal{G}(\mathcal{H})$.
2. $H(\rho) \leq H(\rho_G)$ provided that $H(\rho) < \infty$.
3. If ρ is separable w.r.t. to the cut $A : B$, then ρ_G is separable as well.

Proof. Note that all of these results are well-known in the literature and we only provide their proof here for the sake of completeness.

1. This observation was made e.g. in [85, Appendix] (see also [84, Theorem 1]).

Note that by definition the Gaussification satisfies for all $1 \leq j, k \leq 2N$ that

$$\text{Tr} [(\rho - \rho_G) \hat{r}_k] = 0 \quad \text{and} \quad \text{Tr} [(\rho - \rho_G) \{\hat{r}_j, \hat{r}_k\}] = 0. \quad (3.3)$$

Moreover, we also have $\text{Tr} [(\rho - \rho_G) [\hat{r}_j, \hat{r}_k]] = 0$ due to the CCR.

Consequently, $\text{Tr} [(\rho - \rho_G) f(\hat{q})] = 0$ holds for any second-order polynomial f in the canonical quadrature operators. Given an arbitrary Gaussian state τ_G , the operator $\log \tau_G$ is such a second-order polynomial. The latter follows immediately from the Gibbs exponential form given in Eq. 1.8.

2. This is referred to as the maximum entropy principle (see e.g. [85, Appendix]). This in turn can be seen as a special case of the general extremality principle from [86, Lemma 1].

We have that

$$H(\rho_G) - H(\rho) = \text{Tr} [\rho(\log \rho - \log \rho_G)] + \text{Tr} [(\rho - \rho_G) \log \rho_G] \geq 0, \quad (3.4)$$

where we identified the first term as the relative entropy $D(\rho \| \rho_G)$, which is known to be non-negative, and the second term vanishes by the previous property.

3. This was first established in [81, Proposition 1]. We follow the proof given in [48, Chapter 7.2].

We first show that for an arbitrary separable state with covariance matrix V_{AB} , there always exists covariance matrices V_A and V_B such that $V_{AB} \geq V_A \oplus V_B$. Note that by definition any separable state can be decomposed into a convex mixture of product states with convex weights p_k as

$$\rho_{AB} = \sum_k p_k (\rho_A^k \otimes \rho_B^k). \quad (3.5)$$

Let the component product states $\rho_A^k \otimes \rho_B^k$ have displacement vectors μ^k and block diagonal covariance matrices $V^k = V_A^k \oplus V_B^k$.

By linearity of the trace, the displacement vector of ρ_{AB} has components $\mu_i = \sum_k p_k \mu_i^k$. Similarly, the components of its covariance matrix satisfy

$$V_{i,j} + 2\mu_i \mu_j = \sum_k p_k (V_{i,j}^k + 2\mu_i^k \mu_j^k). \quad (3.6)$$

The difference between V and the block-diagonal matrix $\sum_k p_k V^k$ is thus given by

$$\Delta_{i,j} = 2 \left(\sum_k p_k \mu_i^k \mu_j^k - \sum_{k,l} p_k p_l \mu_i^k \mu_j^l \right). \quad (3.7)$$

This defines a positive semi-definite matrix since for any $\mathbf{x}$ we have

$$\mathbf{x}^T \Delta \mathbf{x} = \sum_{j,l} x_i \Delta_{i,j} x_j = \sum_{k,l} p_k p_l (s_k - s_l)^2 \geq 0, \quad (3.8)$$

where we introduced $s_k = \sum_i x_i \mu_i^k$.

The above observation implies that we can write $\mathbf{V}_{AB} = \mathbf{V}_A \oplus \mathbf{V}_B + \mathbf{Y}$, where $\mathbf{Y}$ is a positive semi-definite matrix. Hence, the Gaussian state with covariance matrix $\mathbf{V}_{AB}$ may be obtained from an uncorrelated Gaussian state with covariance matrix $\mathbf{V}_A \oplus \mathbf{V}_B$ by the action of an additive noise channel. Mathematically, this can be represented as the action of random local unitaries weighted by a Gaussian probability distribution (see e.g. [48, Chapter 5.3.2] for more details). Crucially, the action of this channel preserves the separability of the input state, proving the claim. □

The main technical hurdle in our proof is to show that we can restrict without loss of generality to separable states with finite second moments. Once this is established, the proof follows almost immediately by the properties of the Gaussification discussed above.

Proof of Lemma 3.2. First, observe that since $\mathcal{S}_G \subseteq \mathcal{S}$, we have by a subset argument that

$$D(\mathcal{S} \parallel \rho) \leq D(\mathcal{S}_G \parallel \rho) \quad (3.9)$$

for any quantum state $\rho \in \mathcal{D}(\mathcal{H})$. The point of the proof is to show that the reverse holds provided that the state is Gaussian, i.e. we will show for any $\rho_G \in \mathcal{G}$ that

$$D(\mathcal{S}_G \parallel \rho_G) \leq D(\mathcal{S} \parallel \rho_G). \quad (3.10)$$

We now show that we can restrict the optimisation to states with finite second moments. Note that this automatically implies the finiteness of its first moments. The idea for this part of the proof comes from [63, Corollary 35].

Without loss of generality, we assume that $D(\mathcal{S} \parallel \rho_G)$ is finite (since otherwise the claimed inequality holds trivially). Thus, consider an arbitrary state $\sigma \in \mathcal{S}$ with $D(\sigma \parallel \rho_G) < \infty$. Following [48, Chapter 3], we can write the Gaussian state as

$$\rho_G = U_G \left(|0\rangle\langle 0|_k \otimes \frac{1}{Z} \exp(-\hat{H}) \right) U_G^\dagger,$$

where $|0\rangle\langle 0|_k$ is the k -mode vacuum state, U_G a Gaussian unitary operator,

$$\hat{H} = \sum_{j=1}^{N-k} \frac{\omega_j}{2} (\hat{x}_j^2 + \hat{p}_j^2) \quad (3.11)$$

is the canonical Hamiltonian of the system with $\omega_j > 0$ and Z a normalisation constant. Due to the unitary invariance of the relative entropy, we have

$$D(\sigma \parallel \rho_G) = D \left(U_G^\dagger \sigma U_G \parallel |0\rangle\langle 0|_k \otimes Z^{-1} \exp(-\hat{H}) \right). \quad (3.12)$$

Given that $D(\sigma \parallel \rho_G) < \infty$ holds, we can write $U_G^\dagger \sigma U_G = |0\rangle\langle 0|_k \otimes \sigma'$ due to the support condition. Consequently, we have established the equality $D(\sigma \parallel \rho_G) = D(\sigma' \parallel Z^{-1} \exp(-\hat{H}))$.

As the state $Z^{-1} \exp(-\hat{H})$ is faithful, we can now invoke the variational expression for the measured relative entropy from [37, Lemma 20] to obtain

$$D(\sigma' \| Z^{-1} \exp(-\hat{H})) \geq \sup_{L>0} \left\{ \text{Tr}[\sigma' \log L] - \log \text{Tr}[Z^{-1} \exp(-\hat{H}) L] \right\} \quad (3.13)$$

$$\geq \frac{1}{2} \text{Tr}[\sigma' \hat{H}_n] - \log \text{Tr} \left[Z^{-1} \exp \left(-\frac{\hat{H}_n}{2} \right) \right], \quad (3.14)$$

where in the last step we simply picked the feasible operator $L = \exp(\hat{H}_n/2) > 0$ with $\hat{H}_n = P_n \hat{H}$, where P_n denotes the spectral projector of H corresponding to the interval $[0, n]$. As this holds for any n , we can take the limit $n \rightarrow \infty$, resulting (after rewriting) in

$$\frac{1}{2} \text{Tr}[\sigma' \hat{H}] \leq D(\sigma' \| Z^{-1} \exp(-\hat{H})) + \log \text{Tr} \left[Z^{-1} \exp \left(-\frac{\hat{H}}{2} \right) \right]. \quad (3.15)$$

Note that the RHS is finite as

$$\log \text{Tr} \left[Z^{-1} \exp \left(-\frac{\hat{H}}{2} \right) \right] = \sum_{j=1}^{N-k} \log \left(2 \cosh \left(\frac{\omega_j}{4} \right) \right) < \infty. \quad (3.16)$$

We conclude that $\text{Tr}[\sigma' \hat{H}] < \infty$, which in turn implies that $\text{Tr}[\sigma \hat{H}] < \infty$ and we must have $\text{Tr}[\sigma(\hat{x}_j^2 + \hat{p}_j^2)] < \infty$ for all modes of the system. This is equivalent to σ having finite first and second moments. Note that this also implies that the state has finite entropy (see e.g. [87]).

Having established the viability of this restriction, we can now apply the method of Gaussification. Fix an arbitrary $\sigma \in \mathcal{S}$ with finite second moments (which implies finite first moments) and its associated Gaussian state σ_G as defined above. Using the properties from Lemma 3.3, we have that

$$D(\sigma \| \rho_G) = -H(\sigma) - \text{Tr}[\sigma \log \rho_G] \geq -H(\sigma_G) - \text{Tr}[\sigma_G \log \rho_G] = D(\sigma_G \| \rho_G) \geq D(\mathcal{S}_G \| \rho_G), \quad (3.17)$$

where the first inequality follows by Property 1 and 2 of Lemma 3.3 and the second by Property 3 of Lemma 3.3. We then finish the proof by taking the infimum over $\sigma \in \mathcal{S}$ on the LHS. $\square$

3.2 Efficiently Computable Entanglement Measure

Lemma 3.2 is the key to the main insight in this part of the manuscript, which shows that the reverse relative entropy of entanglement is an operational measure that is efficiently computable for Gaussian inputs. Specifically, we can use it to derive a characterisation of the reverse relative entropy of entanglement for Gaussian states as a *convex* program on the level of quantum covariance matrices with two simple PSD constraints. Consequently, the reverse relative entropy of entanglement is – to the best of the authors' knowledge – the first *operational* entanglement measure that is also efficiently computable for Gaussian states.

Proposition 3.4 (Proposition 2 in main text). *Let $\mathcal{H}_{AB}$ be the Hilbert space of a bosonic $(N_A + N_B)$ -mode CV system and $\mathcal{G}$ be the set of Gaussian quantum states on $\mathcal{H}_{AB}$. The reverse relative entropy of entanglement of $\rho_G \in \mathcal{G}$ with covariance matrix $\mathbf{V}_\rho$ can be computed by the following convex program:*

$$\begin{aligned} \min_{\mathbf{V}_{AB}, \gamma_A} \quad & \frac{\text{Tr}[\mathbf{V}_{AB}(\mathbf{G}[\mathbf{V}_\rho] - \mathbf{G}[\mathbf{V}_{AB}])]}{2 \ln(2)} + \log \sqrt{\frac{\det(\mathbf{V}_\rho + i\mathbf{\Omega})}{\det(\mathbf{V}_{AB} + i\mathbf{\Omega})}} \\ \text{s.t.} \quad & \mathbf{V}_{AB} \geq \gamma_A \oplus i\mathbf{\Omega}_B \quad \text{and} \quad \gamma_A \geq i\mathbf{\Omega}_A \end{aligned} \quad (3.18)$$

Proof. Building on Lemma 3.2, the proof essentially follows by combining the separability criterion from [38, Theorem 5] with the characterisation of the relative entropy between two Gaussian states in terms of their statistical moments from [17, Theorem 7].

To begin with, observe that we can shift the vector of first moments of ρ_G , denoted $\boldsymbol{\mu}_\rho$, to zero using local Gaussian unitaries [48]. Moreover, note that the reverse relative entropy of entanglement is invariant under such local unitary operations. This follows directly from its monotonicity under local operations (cf. Lemma 2.4) combined with the reversibility of unitary channels. In what follows, we can therefore assume $\boldsymbol{\mu}_\rho = 0$ without loss of generality.

Thus, let $\rho_G := \rho_G[0, \mathbf{V}_\rho]$ and consider an arbitrary Gaussian state $\sigma_G := \sigma_G[\boldsymbol{\mu}_\sigma, \mathbf{V}_\sigma]$. According to [17, Theorem 7], we can write their relative entropy as

$$D(\sigma_G[\boldsymbol{\mu}_\sigma, \mathbf{V}_\sigma] \parallel \rho_G[0, \mathbf{V}_\rho]) = -\Sigma(\mathbf{V}_\sigma, \mathbf{V}_\sigma, 0) + \Sigma(\mathbf{V}_\sigma, \mathbf{V}_\rho, \boldsymbol{\mu}_\sigma), \quad (3.19)$$

where the functional Σ is given by

$$2 \ln(2) \cdot \Sigma(\mathbf{V}_1, \mathbf{V}_2, \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) := 2 \ln Z[\mathbf{V}_2] + \text{Tr}[\mathbf{V}_1 \mathbf{G}[\mathbf{V}_2]] + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \mathbf{G}[\mathbf{V}_2] (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \quad (3.20)$$

The Gibbs matrix is positive-definite and thus $(\boldsymbol{\mu}_\sigma - \boldsymbol{\mu}_\rho)^T \mathbf{G}(\mathbf{V}_2) (\boldsymbol{\mu}_\sigma - \boldsymbol{\mu}_\rho) > 0$ for all $\boldsymbol{\mu}_\sigma \neq \boldsymbol{\mu}_\rho = 0$. Since we are searching for a minimum, we can thus assume w.l.o.g. that $\boldsymbol{\mu}_\sigma = \boldsymbol{\mu}_\rho = 0$. Consequently, the objective function is given explicitly by

$$D(\sigma_G[0, \mathbf{V}_\sigma] \parallel \rho_G[0, \mathbf{V}_\rho]) = \frac{\text{Tr}[\mathbf{V}_\sigma (\mathbf{G}[\mathbf{V}_\rho] - \mathbf{G}[\mathbf{V}_\sigma])]}{2 \ln(2)} + \log \sqrt{\frac{\det(\mathbf{V}_\rho + i\boldsymbol{\Omega})}{\det(\mathbf{V}_\sigma + i\boldsymbol{\Omega})}}. \quad (3.21)$$

Moreover, the separability of Gaussian states can be expressed in terms of their covariance matrix only. Concretely, we take the criterion from [38, Theorem 5] which states that a $(N_A + N_B)$ -mode Gaussian state with covariance matrix $\mathbf{V}_{AB}$ is separable if and only if there exists a *bona-fide* N_A -mode quantum covariance matrix $\gamma_A \geq i\boldsymbol{\Omega}_A$ such that $\mathbf{V}_{AB} \geq \gamma_A \oplus i\boldsymbol{\Omega}_B$. Combining this with the above expression results in the program as stated in the proposition.

Lastly, we prove the convexity of this program. The convexity of the feasible set is clear. Regarding the objective function, since we can assume that $\boldsymbol{\mu}_\sigma = \boldsymbol{\mu}_\rho = 0$, we can lift the convexity of the relative entropy on the level of states can to the level of covariance matrices. That is, we will show convexity of the map

$$\mathbf{V}_\sigma \mapsto D(\sigma_G[0, \mathbf{V}_\sigma] \parallel \rho_G[0, \mathbf{V}_\rho]). \quad (3.22)$$

Consider an arbitrary collection $\{p_i, \mathbf{V}_i\}_i$ with $\sum_i p_i = 1$. It then follows that

$$\sum_i p_i D(\sigma_G[0, \mathbf{V}_i] \parallel \rho_G) \geq D\left(\sum_i p_i \sigma_G[0, \mathbf{V}_i] \parallel \rho_G\right) \quad (3.23)$$

$$\geq D\left(\left(\sum_i p_i \sigma_G[0, \mathbf{V}_i]\right)_G \parallel \rho_G\right) = D\left(\sigma_G\left[0, \sum_i p_i \mathbf{V}_i\right] \parallel \rho_G\right), \quad (3.24)$$

where in the first inequality step we used the convexity of the relative entropy and in the second we used our Gaussification argument from the proof of Lemma 3.2. In the final step, we then observed that $\sum_i p_i \sigma_G[0, \mathbf{V}_i]$ has zero first moments and its covariance matrix is given by the convex mixture of the individual covariance matrices (cf. Eq. 3.6 to see this). $\square$

Remark 3.5. The claim of efficiency for this program derives from the fact that it is a finite-dimensional program with two simple PSD constraints. As such, it can straightforwardly be solved with off-the-shelf solvers based on interior-point methods, which are well-known to be efficient in praxis (cf. e.g. [44]). For a rigorous complexity-theoretic efficiency statement, one would (after rewriting it into a standard conic program using the epigraph formulation) need to construct

a self-concordant logarithmic barrier function of the resulting convex cone. There exist universal constructions of such barrier functions that would lead to provable convergence guarantees using the barrier method, i.e. polynomial-time solvability. However, these construction themselves are not efficient in general. A possible solution would be to adapt the specialised literature on relative entropy optimisation such as [45] to this special case.

3.3 The Error Exponent of Two-Way Assisted Quantum Communication

We now turn our focus to the task from quantum communication theory described in detail in the main text. Consider two parties, Alice and Bob, connected by a noisy quantum channel $\mathcal{N} = \mathcal{N}_{A \rightarrow B}$. Their task is the distribution of entanglement, i.e. they are asked to establish copies of the maximally entangled state between them. In order to achieve this goal, they may use the channel n times assisted by unlimited two-way classical communication and adaptive local operations, referred to as adaptive LOCC. We denote the set of such protocols as $\text{LOCC}_{\leftrightarrow}(\mathcal{N}^{\times n})$.

Previous works in this setting characterised performance in terms of the *quantity* of the obtained entanglement, asking at what rate entanglement can be distributed asymptotically (see e.g. [16, 17]). This leads to the *two-way assisted quantum capacity* of the channel as the relevant figure-of-merit. In this work, we instead consider again the *quality* of the entanglement as our performance measure. Specifically, we study the *error exponent of two-way assisted quantum communication*, defined as

$$Q_{\leftrightarrow, \text{err}}^{(m)}(\mathcal{N}) := \sup \left\{ \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \varepsilon_n : \frac{1}{2} \left\| \rho(\Lambda_n) - \Phi_{A_0 B_0}^{\otimes m} \right\|_1 \leq \varepsilon_n, \Lambda_n \in \text{LOCC}_{\leftrightarrow}(\mathcal{N}^{\times n}) \right\}, \quad (3.25)$$

where $\rho(\Lambda_n)$ denotes the output of the protocol Λ_n . The asymptotic error exponent of two-way assisted quantum communication is then given by

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}) := \lim_{m \rightarrow \infty} Q_{\leftrightarrow, \text{err}}^{(m)}(\mathcal{N}) \quad (3.26)$$

Remark 3.6. Note that the distribution of entanglement is a special case of the transmission of arbitrary qubits. Moreover, once Alice and Bob can reliably establish maximally entangled qubits, they can also transmit an arbitrary qubit with the help of a teleportation protocol. Thus, we are justified to refer to this exponent in more general terms as the error exponent of *quantum communication*.

Lemma 3.7. We call a real number $s \geq 0$ an achievable exponent for the channel $\mathcal{N}_{A \rightarrow B}$ if for every $d \in \mathbb{N}$ there exists a $N_0 \in \mathbb{N}_+$ such that for all $n \geq N_0$ one can find an adaptive LOCC protocol that uses the channel n times and outputs a state of the form

$$F_n \Phi_d + (1 - F_n) \tau_d, \quad (3.27)$$

where Φ_d is the maximally entangled state of local dimension d , $\tau_d := (1 - \Phi_d)/(d^2 - 1)$ its orthogonal complement and $F_n \geq 1 - 2^{-ns}$. With this definition, we have

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{A \rightarrow B}) = \sup \left\{ s : s \text{ is an achievable exponent} \right\}. \quad (3.28)$$

Proof. Let s be an achievable exponent. Fix an arbitrary $m \in \mathbb{N}_+$ and set $d = 2^m$. By definition, there exists $N_0 \in \mathbb{N}_+$ such that for all $n \geq N_0$ some protocol $\Lambda_n \in \text{LOCC}_{\leftrightarrow}(\mathcal{N}^{\times n})$ exists that produces

$$\rho(\Lambda_n) = F_n \Phi_d + (1 - F_n) \tau_d. \quad (3.29)$$

The output of this protocol satisfies $\frac{1}{2} \left\| \rho(\Lambda_n) - \Phi_{2^m} \right\|_1 = (1 - F_n) \leq 2^{-ns}$; hence, we have

$$Q_{\leftrightarrow, \text{err}}^{(m)}(\mathcal{N}_{A \rightarrow B}) \geq s \quad (3.30)$$

for every m . Taking the limit $m \rightarrow \infty$, we get $Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{A \rightarrow B}) \geq s$ and finally taking the supremum over all such s completes the first part of the proof.

For the reverse direction, set $S := Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{A \rightarrow B})$. By definition of the limit, for every $\delta > 0$ there exists M_0 such that for all $m \geq M_0$ we have

$$Q_{\leftrightarrow, \text{err}}^{(m)}(\mathcal{N}_{A \rightarrow B}) \geq S - \delta. \quad (3.31)$$

Fix an arbitrary $d \in \mathbb{N}_+$ and consider $m = \max\{M_0, \lceil \log d \rceil\}$. By definition of $Q_{\leftrightarrow, \text{err}}^{(m)}(\mathcal{N}_{A \rightarrow B})$, there exists a sequence of protocols $\Lambda_n \in \text{LOCC}_{\leftrightarrow}(\mathcal{N}^{\times n})$ and errors ε_n that satisfy

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \varepsilon_n \geq S - \delta \quad \text{and} \quad \frac{1}{2} \|\rho(\Lambda_n) - \Phi_{2^m}\|_1 \leq \varepsilon_n. \quad (3.32)$$

We can straightforwardly convert these protocols into ones that produce an isotropic state by implementing an isotropic twirl at the output [16, 79]. Mathematically, the isotropic twirl of dimension d is given by

$$\mathcal{T}_d(\cdot) = \int_{\text{U}(d)} U \otimes \bar{U}(\cdot) (U \otimes \bar{U})^\dagger d\mu_H(U), \quad (3.33)$$

where $d\mu_H(U)$ denotes the Haar measure over the unitary group of dimension d . Crucially, it is known that this map can be implemented by two-way assisted LOCC using only a finite number of unitaries (by resorting to a unitary 2-design). The modified protocols Λ_n now produce

$$\mathcal{T}_d(\rho(\Lambda_n)) = F_n \Phi_d + (1 - F_n) \tau_d. \quad (3.34)$$

Due to the monotonicity of the trace-norm under CPTP-maps, we have

$$(1 - F_n) = \frac{1}{2} \|\mathcal{T}_d(\rho(\Lambda_n)) - \Phi_d\|_1 \leq \frac{1}{2} \|\rho(\Lambda_n) - \Phi_{2^m}\|_1 \leq \varepsilon_n. \quad (3.35)$$

Moreover, by definition of the limit inferior, there exists $N_0 \in \mathbb{N}_+$ such that for all $n \geq N_0$ we have $\varepsilon_n \leq 2^{-n(S-\delta)}$. Consequently, this implies by definition that $S - \delta$ is an achievable exponent; hence, we have for all $\delta > 0$ that

$$\sup \left\{ s : s \text{ is an achievable exponent} \right\} \geq S - \delta \quad (3.36)$$

and the proof is complete by letting $\delta \rightarrow 0$. $\square$

3.4 Fundamental Bound on the Quality of Quantum Communication

In the following, we focus on the special class of bosonic channels that can be simulated with a teleportation protocol using their Choi state. A CV quantum channel $\mathcal{N}_{A \rightarrow B}$ is said to be *teleportation-simulable* with its Choi state if for all $r \geq 0$ there is an LOCC protocol $\Lambda_r \in \text{LOCC}(AA' : B \rightarrow B)$ (with $A' \simeq A$) such that for all states on σ on $\mathcal{H}$, it holds that

$$\lim_{r \rightarrow \infty} \Lambda_r(\sigma_A \otimes \rho_{\mathcal{N}}(r)) = \mathcal{N}(\sigma)_B, \quad (3.37)$$

where convergence is understood in trace norm and $\rho_{\mathcal{N}}(r)$ is the quasi-Choi state of $\mathcal{N}$ defined in Eq. (1.11).

Notably, all Gaussian channels, and in fact all *linear bosonic channels* [88, 89], are known to be teleportation-simulable, with Λ_r being based on the Braunstein-Kimble CV teleportation protocol [90]. The error exponent of two-way assisted quantum communication of these channels can then be upper-bounded in terms of the reverse relative entropy of entanglement. Our bound thus nicely parallels the capacity bound of Pirandola et al. [17] (see also Bennett et al. [16]) in terms of the (regularised) relative entropy of entanglement.

Proposition 3.8 (Proposition 1 in main text). Let $\mathcal{H}_{AB}$ be the Hilbert space of a bosonic $(N_A + N_B)$ -mode CV system and $\mathcal{S}_{A:B}$ the set of separable states on $\mathcal{H}_{AB}$. Using the definitions introduced above, for any teleportation-simulable channel $\mathcal{N}_{A \rightarrow B}$, we have that

$$Q_{\leftrightarrow, \text{err}}(\mathcal{N}_{A \rightarrow B}) \leq \liminf_{r \rightarrow \infty} D(\mathcal{S}_{A:B} \| \rho_{\mathcal{N}}(r)), \quad (3.38)$$

where $\rho_{\mathcal{N}}(r)$ is the quasi-Choi state obtained by sending through $\mathcal{N}_{A \rightarrow B}$ one half of the two-mode squeezed vacuum state with squeezing parameter $r \in [0, \infty)$.

Proof. Let $s \geq 0$ be an achievable error exponent as defined in Lemma 3.7 and fix $d \in \mathbb{N}_+$. By definition, for all large enough n there exists an adaptive LOCC protocol Θ_n that uses $\mathcal{N} = \mathcal{N}_{A \rightarrow B}$ n times and outputs an isotropic state of local dimension d .

Since the channel is teleportation-simulable by assumption, we can replace any use of $\mathcal{N}$ with a copy of its quasi-Choi state $\rho_{\mathcal{N}}(r)$ (up to adding further LOCC processing). This approximation will become better and better as $r \rightarrow \infty$ (for fixed n). Consequently, as proven in [17, Lemma 3], there exists a family of LOCC protocols Λ_r such that for all sufficiently large enough n we have

$$\lim_{r \rightarrow \infty} \Lambda_r(\rho_{\mathcal{N}}(r)^{\otimes n}) = F\Phi_d + (1 - F)\tau_d =: \rho_{\text{iso}}(F), \quad (3.39)$$

where the convergence is understood in trace norm and $F \geq 1 - 2^{-ns}$.

Let us now consider the reverse relative entropy of entanglement of the output state. First, we get an upper bound by

$$D\left(\mathcal{S} \left\| \lim_{r \rightarrow \infty} \Lambda_r(\rho_{\mathcal{N}}(r)^{\otimes n})\right.\right) \leq \liminf_{r \rightarrow \infty} D(\mathcal{S} \| \Lambda_r(\rho_{\mathcal{N}}(r)^{\otimes n})) \quad (3.40)$$

$$\leq \liminf_{r \rightarrow \infty} D(\mathcal{S} \| \rho_{\mathcal{N}}(r)^{\otimes n}) \quad (3.41)$$

$$= n \liminf_{r \rightarrow \infty} D(\mathcal{S} \| \rho_{\mathcal{N}}(r)), \quad (3.42)$$

where we first used the lower-semicontinuity of the reverse relative entropy of entanglement, then its monotonicity under LOCC processing and in the last step its additivity on tensor products (cf. Lemma 2.4 for these properties).

For the lower bound, we exploit that the right-hand side of Eq. (3.39) is an isotropic state of local dimension d . Due to the inherent symmetries of isotropic states, its reverse relative entropy of entanglement can be evaluated analytically using the methods from [20]. Recall that, by definition, an isotropic state is invariant under all unitaries of the form $\mathcal{U} := U \otimes \bar{U}$ for any d -dimensional unitary $U \in \text{U}(d)$. Since such operation preserve separability, we can then invoke unitary invariance and the joint convexity of the relative entropy to deduce that for all $\sigma \in \mathcal{S}$:

$$D(\sigma \| \rho_{\text{iso}}(F)) = \int_{\text{U}(d)} D(\mathcal{U}\sigma\mathcal{U}^\dagger \| \rho_{\text{iso}}(F)) d\mu_H(U) \quad (3.43)$$

$$\geq D\left(\int_{\text{U}(d)} \mathcal{U}\sigma\mathcal{U}^\dagger d\mu_H(U) \left\| \rho_{\text{iso}}(F)\right.\right) \quad (3.44)$$

$$= D(x\Phi_d + (1 - x)\tau_d \| \rho_{\text{iso}}(F)) = x \log\left(\frac{x}{F}\right) + (1 - x) \log\left(\frac{1 - x}{1 - F}\right), \quad (3.45)$$

where $d\mu_H(U)$ denotes the Haar measure of the unitary group $\text{U}(d)$. Since we are searching a minimum, it follows that we can restrict the optimisation w.l.o.g. to isotropic states. Recall that the separability of an isotropic state is equivalent to $x \in [0, 1/d]$ [79]. A straightforward minimisation then yields for all sufficiently large n that

$$D(\mathcal{S} \| \rho_{\text{iso}}(F)) = D_{\text{bin}}(d^{-1} \| F). \quad (3.46)$$

Using the elementary inequality from Eq. (2.25) and $F \geq 1 - 2^{-ns}$, we get the lower bound

$$D_{\text{bin}} \left(d^{-1} \parallel F \right) \geq -1 - \left(1 - \frac{1}{d} \right) \log 2^{-ns} = -1 + ns \left(1 - \frac{1}{d} \right). \quad (3.47)$$

Combining both bounds, we therefore have

$$-1 + ns \left(1 - \frac{1}{d} \right) \leq n \liminf_{r \rightarrow \infty} D(\mathcal{S} \parallel \rho_{\mathcal{N}}(r)) \quad (3.48)$$

Dividing by n , then first taking the limit $n \rightarrow \infty$ followed by the limit $d \rightarrow \infty$ and finally taking the supremum over all achievable exponents s , we get the desired claim. $\square$

Remark 3.9. Recalling the tightest available bounds on the distillable entanglement, there are *essentially* three: (1) the regularised relative entropy of entanglement, which corresponds to relaxing the LOCC framework to non-entangling operations [24]; (2) the negativity, which corresponds to the relaxation to PPT operations [33]; and (3) the squashed entanglement, which is conceptually different and cannot be phrased in the language of LOCC relaxations [91, 92]. Moreover, combining (1) and (2) yields the Rains bound, which is tighter than both on their own [93]. To be precise, [94] investigated dually non-entangling operations, which give a strictly better bound than (1) that is, however, also regularised. It is an interesting open problem to investigate if there are versions of (2) and (3) in the error exponent setting that we study in this work. However, as is evident by the above, there are not many options that can compete with the relative entropy approach of (1). Moreover, note that none of the above approaches (1)–(3) can certify NPT bound entanglement.

3.5 One-Mode Gaussian Channels

Lastly, we compute our bound for the most important one-mode Gaussian channels in optical communication. In the following, we use the mathematical description provided in Serafini's textbook (see [48, Section 5.3] for more details).

1. The *thermal attenuator channel* is described mathematically by

$$X = \cos(\theta) \sigma_0 \quad \text{and} \quad Y = n_{\text{th}} \sin(\theta)^2 \sigma_0 \quad (3.49)$$

with $\theta \in [0, 2\pi)$ describing the transmissivity of the channel and $n_{\text{th}} \geq 1$ a thermal noise parameter. We find that the reverse relative entropy of entanglement of the asymptotic Choi state is given by

$$\lim_{r \rightarrow \infty} D(\mathcal{S} \parallel \rho_G[\mathbf{V}_{\text{Att}}(r)]) = \frac{n_{\text{sep}} (\text{arcoth}(n_{\text{th}}) - \text{arcoth}(n_{\text{sep}}))}{\ln(2)} + \log \left(\sqrt{\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}} \right) \quad (3.50)$$

for $1 \leq n_{\text{th}} \leq n_{\text{sep}}(\theta)$ with $n_{\text{sep}}(\theta) := \frac{1 + \cos(\theta)^2}{1 - \cos(\theta)^2}$ and zero otherwise. Note that this diverges for $n_{\text{th}} \rightarrow 1$, i.e. in the case of the pure loss channel.

The transmissivity $\lambda \in [0, 1]$ introduced in the main body is defined as $\lambda := \cos(\theta)^2$.

2. The *thermal amplifier channel* is described mathematically by

$$X = \cosh(s) \sigma_0 \quad \text{and} \quad Y = n_{\text{th}} \sinh(s)^2 \sigma_0 \quad (3.51)$$

with $s \in [0, \infty)$ describing the amplification of the channel and $n_{\text{th}} \geq 1$ a thermal noise parameter. We find that the reverse relative entropy of entanglement of the asymptotic Choi state is given by

$$\lim_{r \rightarrow \infty} D(\mathcal{S} \| \rho_G[\mathbf{V}_{\text{Amp}}(r)]) = \frac{n_{\text{sep}}(\text{arccoth}(n_{\text{th}}) - \text{arccoth}(n_{\text{sep}}))}{\ln(2)} + \log \left(\sqrt{\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}} \right) \quad (3.52)$$

for $1 \leq n_{\text{th}} \leq n_{\text{sep}}(\theta)$ with $n_{\text{sep}}(\theta) := \frac{\cosh(s)^2 + 1}{\cosh(s)^2 - 1}$ and zero otherwise. Note that this diverges for $n_{\text{th}} \rightarrow 1$, i.e. in the case of the quantum limited amplifier.

The gain $\eta \geq 1$ introduced in the main body is defined as $\eta := \cosh(s)^2$.

3. Finally, the most relevant example of an *additive noise channel* is described by

$$\mathbf{X} = \sigma_0 \quad \text{and} \quad \mathbf{Y} = \mu \sigma_0 \quad (3.53)$$

with $\mu \in [0, \infty)$ describing the induced noise in the channel. We find that the reverse relative entropy of entanglement of the asymptotic Choi state is given by

$$\lim_{r \rightarrow \infty} D(\mathcal{S} \| \rho_G[\mathbf{V}_{\text{Noise}}(r)]) = \frac{2 - \mu}{\mu \ln(2)} + \log \left(\frac{\mu}{2} \right) \quad (3.54)$$

for $0 \leq \mu \leq 2$ and zero otherwise. As expected this diverges in the limit $\mu \rightarrow 0$, i.e. the identity channel.

In the following, we provide the details how we obtain these results. We start with a general prescription that works for the complete class of phase-insensitive channels to reduce the problem to a standard multivariate optimisation problem and then delve into details for our specific channels.

3.5.1 Exploiting the Symplectic Symmetries

In order to tackle the optimisation problem analytically, we first take advantage of the inherent symmetries in order to simplify it as far as possible. Observe that the input covariance matrix in our examples is of the general form

$$\mathbf{V}_\rho = \begin{pmatrix} x\sigma_0 & z\sigma_3 \\ z\sigma_3 & y\sigma_0 \end{pmatrix}. \quad (3.55)$$

In the literature, this is known as the *normal form* of two-mode covariance matrices [95, 96].

Consider the generic symplectic transformation $\mathbf{S} = \mathbf{S}_1 \oplus \mathbf{S}_2$. This acts on the covariance matrix by congruence as

$$\begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & \mathbf{S}_2 \end{pmatrix} \begin{pmatrix} x\sigma_0 & z\sigma_3 \\ z\sigma_3 & y\sigma_0 \end{pmatrix} \begin{pmatrix} \mathbf{S}_1^T & 0 \\ 0 & \mathbf{S}_2^T \end{pmatrix} = \begin{pmatrix} x\mathbf{S}_1\mathbf{S}_1^T & z\mathbf{S}_1\sigma_3\mathbf{S}_2^T \\ y\mathbf{S}_2\sigma_3\mathbf{S}_1^T & z\mathbf{S}_2\mathbf{S}_2^T \end{pmatrix}. \quad (3.56)$$

The matrix is an invariant of this transformation provided the following three conditions hold:

$$\mathbf{S}_1\mathbf{S}_1^T = \sigma_0 \quad \mathbf{S}_2\mathbf{S}_2^T = \sigma_0 \quad \text{and} \quad \mathbf{S}_1\sigma_3\mathbf{S}_2^T = \sigma_3 \implies \mathbf{S}_2 = \sigma_3\mathbf{S}_1\sigma_3. \quad (3.57)$$

The first condition forces $\mathbf{S}_1$ to be orthogonal. Moreover, since $\sigma_3\Omega_1\sigma_3 = -\Omega_1$ holds, it follows that $\mathbf{S}_2 = \sigma_3\mathbf{S}_1\sigma_3$ inherits orthogonality and symplecticity from $\mathbf{S}_1$:

$$\mathbf{S}_2\Omega_1\mathbf{S}_2^T = \sigma_3\mathbf{S}_1\sigma_3\Omega_1\sigma_3\mathbf{S}_1^T\sigma_3 = -\sigma_3\mathbf{S}_1\Omega_1\mathbf{S}_1^T\sigma_3 = -\sigma_3\Omega_1\sigma_3 = \Omega_1. \quad (3.58)$$

Hence, a covariance matrix in normal form is invariant under all transformations of the form

$$\mathbf{S} = \mathbf{O} \oplus (\sigma_3\mathbf{O}\sigma_3), \quad (3.59)$$

where $\mathbf{O}$ is any 2×2 orthogonal symplectic matrix.

In the Hilbert space picture, such a transformation is represented by an infinite-dimensional local Gaussian unitary U_S . Using the unitary invariance of the relative entropy, we obtain

$$D(\sigma_G[\mathbf{V}] \parallel \rho_G[\mathbf{V}_\rho]) = D\left(U_S \sigma_G[\mathbf{V}] U_S^\dagger \parallel U_S \rho_G[\mathbf{V}_\rho] U_S^\dagger\right) \quad (3.60)$$

$$= D\left(\sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] \parallel \rho_G[\mathbf{S} \mathbf{V}_\rho \mathbf{S}^T]\right) \quad (3.61)$$

$$= D\left(\sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] \parallel \rho_G[\mathbf{V}_\rho]\right) = \int D\left(\sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] \parallel \rho_G[\mathbf{V}_\rho]\right) d\mu_H(\mathbf{O}_S), \quad (3.62)$$

where $d\mu_H(\mathbf{O}_S)$ denotes the Haar measure over the orthogonal subgroup of the symplectic group. Note that this exists as this is a compact subgroup of the symplectic group [52].

Then, using joint convexity of the relative entropy, it follows that

$$D(\sigma_G[\mathbf{V}] \parallel \rho_G[\mathbf{V}_\rho]) = \int D\left(\sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] \parallel \rho_G[\mathbf{V}_\rho]\right) d\mu_H(\mathbf{O}_S) \quad (3.63)$$

$$\geq D\left(\int \sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] d\mu_H(\mathbf{O}_S) \parallel \rho_G[\mathbf{V}_\rho]\right) \quad (3.64)$$

$$\geq D\left(\left(\int \sigma_G[\mathbf{S} \mathbf{V} \mathbf{S}^T] d\mu_H(\mathbf{O}_S)\right)_G \parallel \rho_G[\mathbf{V}_\rho]\right) \quad (3.65)$$

$$= D\left(\sigma_G\left[\int \mathbf{S} \mathbf{V} \mathbf{S}^T d\mu_H(\mathbf{O}_S)\right] \parallel \rho_G[\mathbf{V}_\rho]\right) \quad (3.66)$$

where we first used the convexity in the Hilbert space picture followed by our Gaussification argument from the proof of Lemma 3.2. For the latter, recall that both states have zero first moments and thus the convex mixture on the state level corresponds to a convex mixture of the covariance matrices.

As we are searching for a minimum, we can thus assume w.l.o.g. that the optimiser is of the form

$$\int \mathbf{S} \mathbf{V} \mathbf{S}^T d\mu_H(\mathbf{O}_S) = \begin{pmatrix} x\sigma_0 & z\sigma_3 \\ z\sigma_3 & y\sigma_0 \end{pmatrix} \quad (3.67)$$

with three real-valued parameters x, y and z , i.e. a two-mode covariance matrix in normal form.

3.5.2 Details on Two-Mode Normal Forms

The normal form of two-mode covariance matrices is well studied in the literature (see e.g. [97, 98] and [49, Section II.C] for detailed discussions). Using standard techniques from linear algebra, one finds that the *bona-fide* condition is equivalent to the following constraints on the normal form parameters:

$$x \geq 1, \quad y \geq 1, \quad |z| \leq z_{\max} := \sqrt{\min\left\{(x-1)(y+1), (x+1)(y-1)\right\}} = \sqrt{xy - 1 - |x - y|}. \quad (3.68)$$

Moreover, by [95, Theorem], the Peres-Horodecki criterion provides a necessary and sufficient criterion for the separability of two-mode Gaussian states. In terms of the normal form parameters, this can be shown to be equivalent to

$$z^2 \leq (x-1)(y-1). \quad (3.69)$$

We can then further simplify the optimisation by restricting to the boundary of the separable set.

To see that this entails no loss of generality, let $\mathbf{V}_{\text{opt}}$ be the optimal separable covariance matrix and define $f(t) := D((1-t)\mathbf{V}_{\text{opt}} + t\mathbf{V}_\rho \parallel \mathbf{V}_\rho)$ for $t \in [0, 1]$, where $D(\mathbf{V}_1 \parallel \mathbf{V}_2) := D(\sigma_G[\mathbf{V}_1] \parallel \rho_G[\mathbf{V}_2])$. The

function f is convex and non-negative, achieving its minimal value of zero at $t = 1$. Hence, it is monotone decreasing on the interval $[0, 1]$. Since V_{opt} is by assumption the minimiser, it follows that $(1 - t)V_{\text{opt}} + tV_\rho$ is not separable for all $t > 0$. Consequently, V_{opt} must lie on the boundary of the feasible set, i.e. the optimiser is so-called border-separable.

Furthermore, the symplectic eigenvalues of a two-mode covariance matrix are given by the handy formula:

$$\nu_\pm = \sqrt{\frac{\Delta(V) \pm \sqrt{\Delta(V)^2 - 4 \det(V)}}{2}} \quad (3.70)$$

using the two global symplectic invariants

$$\Delta(V) = x^2 + y^2 - 2z^2 \quad \text{and} \quad \det(V) = (xy - z^2)^2. \quad (3.71)$$

In terms of the normal form parameters, the two symplectic eigenvalues are given explicitly by

$$\nu_1 := \frac{\sqrt{(x+y)^2 - 4z^2} + (x-y)}{2} \quad \nu_2 := \frac{\sqrt{(x+y)^2 - 4z^2} + (y-x)}{2}, \quad (3.72)$$

where we used the identity

$$A \pm \sqrt{A^2 - B^2} = \left(\sqrt{\frac{A+B}{2}} \pm \sqrt{\frac{A-B}{2}} \right)^2. \quad (3.73)$$

Recall that Williamson's theorem ensures the existence of a symplectic matrix S that diagonalises the covariance matrix by congruence – $V := SWS^T$ – into its Williamson form $W := (\nu_1\sigma_0 \oplus \nu_2\sigma_0)$. For a two-mode covariance matrix in normal form, the symplectic matrix S that achieves this transformation is given explicitly by

$$S = \begin{pmatrix} \omega_+\sigma_0 & \text{sgn}(z)\omega_-\sigma_3 \\ \text{sgn}(z)\omega_-\sigma_3 & \omega_+\sigma_0 \end{pmatrix} \quad \text{with} \quad \omega_\pm := \sqrt{\frac{x+y \pm \sqrt{(x+y)^2 - 4z^2}}{2\sqrt{(x+y)^2 - 4z^2}}}. \quad (3.74)$$

Using this, we can compute the Gibbs matrix explicitly. We make use of its characterisation in terms of the symplectic action of the real function $f(x) := \text{arccoth}(x)$. Following [99, Section IV.B], the symplectic action of f is defined as $f_\star(V) := Sf(W)S^T$, where $f(W)$ is the standard matrix function of the Williamson form W . Note that this is in general different from the standard matrix function, since the symplectic and spectral decomposition need not coincide. It can be shown that the Gibbs matrix is then given by (cf. [53, Appendix A]):

$$G[V] = -\Omega \text{arccoth}_\star(V) \Omega. \quad (3.75)$$

Using the notation from above, the Gibbs matrix of a normal form covariance matrix is then also in normal form,

$$G[V] = \begin{pmatrix} \alpha\sigma_0 & \gamma\sigma_3 \\ \gamma\sigma_3 & \beta\sigma_0 \end{pmatrix}, \quad (3.76)$$

with the normal form parameters given by

$$\alpha = \omega_+^2 \text{arccoth}(\nu_1) + \omega_-^2 \text{arccoth}(\nu_2) \quad (3.77)$$

$$\beta = \omega_-^2 \text{arccoth}(\nu_1) + \omega_+^2 \text{arccoth}(\nu_2) \quad (3.78)$$

$$\gamma = -\text{sgn}(z)\omega_-\omega_+(\text{arccoth}(\nu_1) + \text{arccoth}(\nu_2)). \quad (3.79)$$

3.5.3 Setting Up the Optimisation Problem

We have now everything in place in order to express the objective function solely in terms of the three normal form parameters. We start from the following equivalent characterisations of the relative entropy

$$D(\sigma_G[\mathbf{V}] \parallel \rho_G[\mathbf{V}_\rho]) = \frac{\text{Tr} [\mathbf{V}(\mathbf{G}[\mathbf{V}_\rho] - \mathbf{G}[\mathbf{V}])]}{2 \ln 2} + \log \frac{Z[\mathbf{V}_\rho]}{Z[\mathbf{V}]} \quad (3.80)$$

$$= \frac{\text{Tr} [\mathbf{V}\mathbf{G}[\mathbf{V}_\rho]]}{2 \ln 2} - H(\sigma_G[\mathbf{V}]) + \log Z[\mathbf{V}_\rho] \quad (3.81)$$

The normalisation constant $Z[\mathbf{V}]$ can be expressed solely in terms of the symplectic spectrum. Namely, we can write it as

$$\log Z[\mathbf{V}] = \log \sqrt{\det \left(\frac{1}{2}(\mathbf{V} + i\mathbf{\Omega}_2) \right)} = \frac{1}{2} \left(\log(\nu_1^2 - 1) + \log(\nu_2^2 - 1) \right) - 2 \quad (3.82)$$

Moreover, a similar handy expression in terms of the symplectic spectrum is available for the von-Neumann entropy. Following [97, Proposition 1] (see also [85]), we can write it as

$$H(\sigma_G[\mathbf{V}]) = g(\nu_1) + g(\nu_2) \quad \text{with} \quad g(x) := \left(\frac{x+1}{2} \right) \log \left(\frac{x+1}{2} \right) - \left(\frac{x-1}{2} \right) \log \left(\frac{x-1}{2} \right). \quad (3.83)$$

Using Eq. (3.72), this is then given implicitly in terms of the normal form coefficients. Lastly, since both covariance matrices are given in normal form, we can write the overlap term $\text{Tr} [\mathbf{V}_1 \mathbf{G}[\mathbf{V}_2]]$ solely in terms of the respective normal form coefficients:

$$\text{Tr} [\mathbf{V}_1 \mathbf{G}[\mathbf{V}_2]] = 2x_1\alpha_2 + 2y_1\beta_2 + 4z_1\gamma_2, \quad (3.84)$$

where $(\alpha_2, \beta_2, \gamma_2)$ denote the normal form coefficients of the Gibbs matrix $\mathbf{G}[\mathbf{V}_2]$.

Hence, starting from Eq. (3.81) the objective function becomes

$$F(x, y, z) := \frac{x\alpha_\rho + y\beta_\rho + 2z\gamma_\rho}{\ln 2} - g(\nu_1(x, y, z)) - g(\nu_2(x, y, z)) + \log Z[\mathbf{V}_\rho], \quad (3.85)$$

where (x, y, z) have to satisfy the constraints discussed in the previous sections. The problem is now set up in principle as a multivariate optimisation problem that can be solved with the methods from standard calculus. However, in general this leads to intractable transcendental equations, thus we have to resort in the following to an asymptotic analysis in the large squeezing regime.

3.5.4 Analysis for Thermal Attenuator Channel

Let us begin with consider the thermal attenuator channel. The quasi-Choi state of this channel is a Gaussian state with covariance matrix

$$\mathbf{V}_{\text{Att}}(r) = \begin{pmatrix} (\cosh(2r) \cos(\theta)^2 + n_{\text{th}} \sin(\theta)^2) \sigma_0 & \sinh(2r) \cos(\theta) \sigma_3 \\ \sinh(2r) \cos(\theta) \sigma_3 & \cosh(2r) \sigma_0 \end{pmatrix}. \quad (3.86)$$

We find that it is entangled provided that

$$n_{\text{th}} \leq n_{\text{sep}}(\theta) := \frac{1 + \cos(\theta)^2}{1 - \cos(\theta)^2} \quad (3.87)$$

independent of the squeezing parameter r . Thus, we assume $1 \leq n_{\text{th}} \leq n_{\text{sep}}(\theta)$ in the following.

As we are ultimately interested in the limit $r \rightarrow \infty$, we focus on the large squeezing regime to make further analytical progress. For $\theta \neq \{0, \pi, 2\pi\}$, we find that the symplectic spectrum of the quasi-Choi states in Eq. (3.86) satisfies

$$\nu_1(\theta, n_{\text{th}}, r) = n_{\text{th}} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right) \quad (3.88)$$

$$\nu_2(\theta, n_{\text{th}}, r) = \cosh(2r) \sin(\theta)^2 + n_{\text{th}} \cos(\theta)^2 + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right) \quad (3.89)$$

With this, we can then derive the associated Gibbs matrix. We find for $\theta \neq \{0, \pi, 2\pi\}$ that

$$\mathbf{G}[\mathbf{V}_{\text{Att}}(\theta, n_{\text{th}}, r)] = \frac{\text{arcoth}(n_{\text{th}})}{\sin(\theta)^2} \begin{pmatrix} \cos(\theta)^2 \sigma_0 & -\cos(\theta) \sigma_3 \\ -\cos(\theta) \sigma_3 & \sigma_0 \end{pmatrix} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right). \quad (3.90)$$

Then, focusing on the first term in Eq. (3.85), we can rewrite it (ignoring the normalisation) as

$$\frac{x+y}{2}(\alpha_{\text{Att}} + \beta_{\text{Att}}) + \frac{x-y}{2}(\alpha_{\text{Att}} - \beta_{\text{Att}}) + 2z\gamma_{\text{Att}}. \quad (3.91)$$

From this, we are able to deduce that $x \geq y \geq 1$ may be assumed and we can replace $z \rightarrow |z|$ in the above w.l.o.g. as we are searching for a minimum. The former follows from the fact that $\alpha_{\text{Att}} \leq \beta_{\text{Att}}$ asymptotically and that the entropy term is invariant under swapping x and y . Similarly, the latter is due to the optimal z having the reversed sign of γ_{Att} (note that the symplectic eigenvalues are invariant under swapping the sign of z).

We can then explicitly re-parametrise the problem in terms of the two symplectic eigenvalues (ν_1, ν_2) of the optimisation variable, using that $\nu_1 \geq \nu_2 \geq 1$ must hold due to $x \geq y$. We explicitly have

$$x = \frac{1 + \nu_1 \nu_2 + \nu_1 - \nu_2}{2} \quad y = \frac{1 + \nu_1 \nu_2 + \nu_2 - \nu_1}{2} \quad (3.92)$$

and

$$|z| = \frac{\sqrt{(1 + \nu_1 \nu_2)^2 - (\nu_1 + \nu_2)^2}}{2}. \quad (3.93)$$

The first-order condition for the minimum then requires

$$\frac{\partial F}{\partial \nu_1} = \alpha_\rho \frac{\nu_2 + 1}{2} + \beta_\rho \frac{\nu_2 - 1}{2} + \gamma_\rho \nu_1 \sqrt{\frac{\nu_2^2 - 1}{\nu_1^2 - 1}} - \text{arcoth}(\nu_1) = 0 \quad (3.94)$$

and

$$\frac{\partial F}{\partial \nu_2} = \alpha_\rho \frac{\nu_1 - 1}{2} + \beta_\rho \frac{\nu_1 + 1}{2} + \gamma_\rho \nu_2 \sqrt{\frac{\nu_1^2 - 1}{\nu_2^2 - 1}} - \text{arcoth}(\nu_2) = 0. \quad (3.95)$$

However, this is a transcendental system of equations and cannot be solved analytically for all parameters. Observe that in the large squeezing regime, we have

$$\log Z_{\text{TA}} = \log(\sin(\theta)^2 \cosh(2r)) + \frac{1}{2} \log(n_{\text{th}}^2 - 1) - 2 + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right), \quad (3.96)$$

which grows without bounds as $r \rightarrow \infty$. As we are looking for a minimum, we need that $\nu_1 \nu_2 = \mathcal{O}(\cosh(2r))$ must hold to precisely cancel this growth and keep the relative entropy finite. Making the ansatz $\nu_1 = A \cosh(2r) + B$ and $\nu_2 = D$ with $r \gg 1$, we have

$$\frac{\partial F}{\partial \nu_1} = \frac{\alpha_\rho + \beta_\rho}{2} D + \frac{\alpha_\rho - \beta_\rho}{2} + \gamma_\rho \sqrt{D^2 - 1} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right), \quad (3.97)$$

which leads to the solution

$$v_2 = \frac{1 + \cos(\theta)^2}{1 - \cos(\theta)^2} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right) = n_{\text{sep}}(\theta) + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right). \quad (3.98)$$

Moreover, we have

$$\frac{\partial F}{\partial v_2} = A \left(\frac{\alpha_\rho + \beta_\rho}{2} + \gamma_\rho \frac{D}{\sqrt{D^2 - 1}} \right) \cosh(2r) + \mathcal{O}(1) \quad (3.99)$$

Plugging in the optimal D , the bracket evaluates to zero; hence, A is not fixed by this and a careful analysis of the next-order reveals the same happens there. Consequently, we have to resort to a numerical solution of the problem given in Eq. 3.85 to determine that $v_1 = \cosh(2r) \sin(\theta)^2 + \mathcal{O}(1)$ holds in the large squeezing regime. Using this, we can then determine that

$$F_{\text{opt}} = \frac{n_{\text{sep}}(\text{arccoth}(n_{\text{th}}) - \text{arccoth}(n_{\text{sep}}))}{\log(2)} + \frac{1}{2} \log\left(\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}\right) + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right). \quad (3.100)$$

3.5.5 Analysis of Thermal Amplifier Channel

The covariance matrix for the quasi-Choi of the thermal amplifier is given by

$$V_{\text{Amp}}(r) = \begin{pmatrix} (\cosh(2r) \cosh(s)^2 + n_{\text{th}} \sinh(s)^2) \sigma_0 & \sinh(2r) \cosh(s) \sigma_3 \\ \sinh(2r) \cosh(s) \sigma_3 & \cosh(2r) \sigma_0 \end{pmatrix}. \quad (3.101)$$

with squeezing parameter $s \in [0, \infty)$ and thermal noise parameter $n_{\text{th}} \geq 1$. We find that it is entangled if

$$n \leq \frac{\cosh(s)^2 + 1}{\cosh(s)^2 - 1} := n_{\text{sep}}(s) \quad (3.102)$$

independent of the squeezing parameter. For the symplectic eigenvalues, we find for $s > 0$ that

$$v_1(s, n_{\text{th}}, r) = \cosh(2r) \sinh(s)^2 + n_{\text{th}} \cosh(s)^2 + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right) \quad (3.103)$$

$$v_2(s, n_{\text{th}}, r) = n_{\text{th}} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right) \quad (3.104)$$

and the associated Gibbs matrix satisfies

$$G[V_{\text{Amp}}(s, n_{\text{th}}, r)] = \frac{\text{arccoth}(n_{\text{th}})}{\sinh(s)^2} \begin{pmatrix} \sigma_0 & -\cosh(s) \sigma_3 \\ -\cosh(s) \sigma_3 & \cosh(s)^2 \sigma_0 \end{pmatrix} + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right). \quad (3.105)$$

Thus, compared to the previous example of the attenuator, the roles of x and y are reversed here. Re-running the above analysis with this consideration, the optimal value is given by

$$F_{\text{opt}} = \frac{n_{\text{sep}}(\text{arccoth}(n_{\text{th}}) - \text{arccoth}(n_{\text{sep}}))}{\ln(2)} + \frac{1}{2} \log\left(\frac{n_{\text{th}}^2 - 1}{n_{\text{sep}}^2 - 1}\right) + \mathcal{O}\left(\frac{1}{\cosh(2r)}\right). \quad (3.106)$$

using the re-defined value $n_{\text{sep}}(s)$. Notably, this diverges in the case of $n_{\text{th}} \rightarrow 1$, which corresponds to the quantum limited amplifier.

3.5.6 Analysis of Additive Noise Channel

The quasi-Choi of the additive Gaussian noise channel has covariance matrix given by

$$V_{\text{Noise}}(r) = \begin{pmatrix} (\cosh(2r) + \eta) \sigma_0 & \sinh(2r) \sigma_3 \\ \sinh(2r) \sigma_3 & \cosh(2r) \sigma_0 \end{pmatrix}. \quad (3.107)$$

with noise parameter $\mu \in [0, \infty)$ and we find that it is entangled if $\mu \in [0, 2)$. The symplectic eigenvalues for $\mu > 0$ are given by

$$v_1(\mu, r) = \sqrt{\mu \cosh(2r)} + \frac{\mu}{2} + \mathcal{O}\left(\frac{1}{\sqrt{\cosh(2r)}}\right) \quad (3.108)$$

$$v_2(\mu, r) = \sqrt{\mu \cosh(2r)} - \frac{\mu}{2} + \mathcal{O}\left(\frac{1}{\sqrt{\cosh(2r)}}\right) \quad (3.109)$$

and the Gibbs matrix is given by

$$\mathbf{G}[\mathbf{V}_{\text{Noise}}(\mu, r)] = \frac{1}{\mu} \begin{pmatrix} \sigma_0 & -\sigma_3 \\ -\sigma_3 & \sigma_0 \end{pmatrix} + \mathcal{O}\left(\frac{1}{\sqrt{\cosh(2r)}}\right). \quad (3.110)$$

The solution can thus be assumed symmetric in x and y and we find with a similar analysis as before that

$$F_{\text{opt}} = \frac{2 - \mu}{\mu \ln(2)} + \log\left(\frac{\mu}{2}\right) + \mathcal{O}\left(\frac{1}{\sqrt{\cosh(2r)}}\right). \quad (3.111)$$