Bosonic quantum communication across arbitrarily high loss channels

Ludovico Lami,[1, *] Martin B. Plenio,[1, †] Vittorio Giovannetti,[2, ‡] and Alexander S. Holevo[3, §]

[1]*Institut für Theoretische Physik und IQST, Universität Ulm,*
Albert-Einstein-Allee 11, D-89069 Ulm, Germany
[2]*NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56127 Pisa, Italy*
[3]*Steklov Mathematical Institute, Gubkina 8, 119991 Moscow, Russia*

A general attenuator $\Phi_{\lambda,\sigma}$ is a bosonic quantum channel that acts by combining the input with a fixed environment state σ in a beam splitter of transmissivity λ. If σ is a thermal state the resulting channel is a thermal attenuator, whose quantum capacity vanishes for $\lambda \leq 1/2$. We study the quantum capacity of these objects for generic σ, proving a number of unexpected results. Most notably, we show that for any arbitrary value of $\lambda > 0$ there exists a suitable single-mode state $\sigma(\lambda)$ such that the quantum capacity of $\Phi_{\lambda,\sigma(\lambda)}$ is larger than a universal constant $c > 0$. Our result holds even when we fix an energy constraint at the input of the channel, and implies that quantum communication at a constant rate is possible even in the limit of arbitrarily low transmissivity, provided that the environment state is appropriately controlled. We also find examples of states σ such that the quantum capacity of $\Phi_{\lambda,\sigma}$ is not monotonic in λ. These findings may have implications for the study of communication lines running across integrated optical circuits, of which general attenuators provide natural models.

Introduction.— Quantum optics will likely play a major role in the future of quantum communication [1–4]. Indeed, practically all quantum communication in the foreseeable future will rely on optical platforms. For this reason, the study of quantum channels acting on continuous variable (CV) systems, that is, finite ensembles of electromagnetic modes, is a core area of the rapidly developing field of quantum information [5–7].

In the best studied models of optical communication, one represents an optical fibre as a memoryless thermal attenuator channel. Mathematically, its action can be thought of as that of a beam splitter with a certain transmissivity $0 \leq \lambda \leq 1$, where the input state is mixed with a fixed environment state σ that is assumed to be thermal. This approximation is well justified when the signal rate is sufficiently low that memory effects are negligible, and when the optical fibre is so long that the 'effective' environment state, resulting from averaging several elementary interactions that are effectively independent, due to the limited correlation length of the environment, is practically Gaussian and thermal, as follows from the quantum central limit theorem [8, 9]. And indeed, an impressive amount of literature has been devoted to finding bounds on the quantum capacity of the thermal attenuator. We now have exact formulae for the zero-temperature case [10–15], and tight upper [15–18] and lower [10, 19] bounds in all other cases.

However, the thermal noise approximation is challenged when memory effects become important [20], or when the communication channel is so short that the averaging process cannot possibly take place, as may happen e.g. in miniaturised quantum optical circuits [21–24]. In both cases, it is conceivable that the environment state may be manipulated and engineered to facilitate communication. Namely, one could exploit memory effects to send pulses that alter it and precede the actual transmission, or one could design the integrated optical circuit that surrounds the communication line in order to control the noise that comes from other elements of the same circuit. We are thus led to investigate general attenuator channels, hereafter denoted with $\Phi_{\lambda,\sigma}$, where the environment state σ is no longer thermal. Unsurprisingly, such models have received increasing attention recently [9, 25–29]. As discussed above, we will be interested in optimising over the environment state so as to increase the capacity [30, 31].

Other motivations for considering general attenuators stem on the one hand from the need to go beyond the Gaussian formalism to accomplish several tasks that are critical to quantum information, e.g. universal quantum computation [32, 33] entanglement distillation [34–36], entanglement swapping [37, 38], error correction [39], and state transformations in general resource theories [40, 41]. On the other hand, general attenuators are among the simplest examples of non-Gaussian channels that are nevertheless *Gaussian dilatable*, meaning that they can be Stinespring dilated [42] by means of a symplectic unitary [27, 28]. This makes them amenable to a quantitative analysis in many respects. For example, it has been shown that making the environment state non-Gaussian, e.g. by means of a photon addition, can be advantageous when transmitting quantum or private information [27]. In spite of their increased complexity compared to Gaussian channels, the entanglement-assisted capacity of a general attenuator can nevertheless be upper bounded thanks to the conditional entropy power inequality [25, 26]. Similar bounds can be obtained for the quantum [29] and private [43] capacity as well, by making use of the solution to the minimum output entropy conjecture [44–46] combined with known extremality properties of Gaussian states [47, 48]. Finally, we have mentioned that by concatenating a large number n of general

attenuators with a fixed total transmissivity one typically obtains an effective channel that resembles a thermal attenuator. In this regime of large but finite n, the associated quantum capacity can be bounded thanks to the quantum Berry–Esseen inequality [9, Corollary 13].

Here we investigate the quantum capacity of general attenuators $\Phi_{\lambda,\sigma}$, uncovering some unexpected phenomena. It has been observed [9, Lemma 16] that output states of general attenuators with transmissivity $\lambda = 1/2$ have non-negative Wigner functions [49, 50]. At first sight, this may suggest that such channels are somewhat 'classical' [51–53]. Indeed, we show that for all convex combinations of symmetric states – and in particular for all Gaussian states – $\Phi_{1/2,\sigma}$ is anti-degradable and therefore its quantum capacity satisfies $Q\left(\Phi_{1/2,\sigma}\right) = 0$ [54]. Here we call a state symmetric if it remains invariant under phase space inversion up to displacements. However, we also find an example of a state σ that does not belong to this class and that makes $Q\left(\Phi_{1/2,\sigma}\right) > 0$.

Next, we tackle the question of whether transmission of quantum information is possible even for very low values of the transmissivity $0 < \lambda \ll 1$. Intuitively, a beam splitter of transmissivity $\lambda \leq 1/2$ should give away to the environment more than it transmits. By the no-cloning theorem, we could be led to conjecture that the quantum capacity $Q\left(\Phi_{\lambda,\sigma}\right)$ vanishes for all σ as soon as $\lambda \leq 1/2$. Indeed, this is exactly what happens for thermal attenuators. This intuition is further supported by the analysis of general finite-dimensional depolarising channels $\Delta_{\lambda,\sigma}(\rho)$, defined by $\Delta_{\lambda,\sigma}(\rho) := \lambda\rho + (1 - \lambda)\sigma$, whose quantum capacity also vanishes for $\lambda \leq 1/2$.

However, we establish the following surprising result: *for all values of $\lambda > 0$ one can find suitable states $\sigma(\lambda)$ that make $Q\left(\Phi_{\lambda,\sigma(\lambda)}\right) \geq c$, where the constant $c > 0$ is universal* (Theorem 2). This implies, but is stronger than, the fact that $\Phi_{\lambda,\sigma(\lambda)}$ can be used to distribute entanglement [55]. As a corollary, we also see that $Q\left(\Phi_{\lambda,\sigma}\right)$ is in general not monotonic in λ for fixed σ. All this marks a striking difference with the aforementioned behaviour of thermal attenuators and depolarising channels, and reveals that the phenomenology of general attenuators is richer than perhaps expected. Our proof is fully analytical, and goes by analysing the single-copy coherent information associated with a specific transmission scheme. By a tour-de-force of inequalities we show that the output state of the channel is *majorised* by that of the associated complementary channel. In turn, this makes it possible to lower bound the coherent information by applying a beautiful inequality recently proved by Ho and Verdú [56].

Notation.— The Hilbert space corresponding to an m-mode CV comprises all square-integrable functions $\mathbb{R}^m \to \mathbb{C}$, and is denoted by $\mathcal{H}_m := L^2(\mathbb{R}^m)$. Quantum states are represented by *density operators* on $\mathcal{H}_m$, i.e. positive semi-definite trace class operators with unit trace. We will denote with $a_j, a_j^\dagger$, respectively, the *anni-*
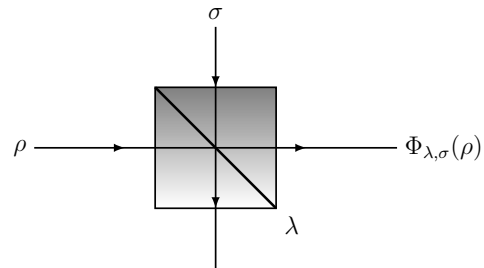

FIG. 1. A general attenuator acts by mixing the input state ρ in a beam splitter of transmissivity λ with an environment in a fixed state σ.

hilation and *creation* operators corresponding to the j-th mode, and with $|0\rangle$ the *vacuum state*. The *canonical commutation relations* read $[a_j, a_k^\dagger] = \delta_{jk}I$, $[a_j, a_k] = 0$. The unitary *displacement operators* on $\mathcal{H}_m$ are constructed as $D(\alpha) := e^{\sum_j (\alpha_j a_j^\dagger - \alpha_j^* a_j)}$, where $\alpha \in \mathbb{C}^m$; they satisfy $D(\alpha)D(\beta) = e^{\frac{1}{2}(\alpha^\intercal \beta^* - \alpha^\dagger \beta)}D(\alpha + \beta)$ for all $\alpha, \beta \in \mathbb{C}^m$.

For every trace class operator T on $\mathcal{H}_m$, its *characteristic function* $\chi_T : \mathbb{C}^m \to \mathbb{C}$ is defined by [6, 57]

$$\chi_T(\alpha) := \operatorname{Tr}[TD(\alpha)]. \tag{1}$$

The *Wigner function* W_T of T is the Fourier transform of χ_T [6, 49, 50, 57]. Note that W_ρ is typically not pointwise positive for a generic quantum state ρ [51–53].

A *beam splitter* of transmissivity $0 \leq \lambda \leq 1$ acting on two systems of m modes each is represented by the unitary operator

$$U_\lambda := e^{\arccos \sqrt{\lambda} \sum_j (a_j^\dagger b_j - a_j b_j^\dagger)}, \tag{2}$$

where a_j, b_j are the annihilation operators on the j-th modes of the first and second system, respectively. Our main object of study is the *general attenuator* channel $\Phi_{\lambda,\sigma}$, which acts on an m-mode system B as

$$\Phi_{\lambda,\sigma}^B(\rho_B) := \operatorname{Tr}_E\left[U_\lambda^{BE}(\rho_B \otimes \sigma_E)\left(U_\lambda^{BE}\right)^\dagger\right]. \tag{3}$$

Dropping the system labels for simplicity, this can be cast in the language of characteristic functions as

$$\chi_{\Phi_{\lambda,\sigma}(\rho)}(\alpha) = \chi_\rho\left(\sqrt{\lambda}\,\alpha\right)\chi_\sigma\left(\sqrt{1-\lambda}\,\alpha\right). \tag{4}$$

A pictorial representation of the action of a general attenuator is provided in Figure 1.

The *thermal attenuators* $\mathscr{E}_{\lambda,\nu} := \Phi_{\lambda,\tau_\nu}$ as well as the *pure loss channels* $\mathscr{E}_\lambda := \mathscr{E}_{\lambda,0} = \Phi_{\lambda,|0\rangle\langle 0|}$ are standard examples of single-mode attenuators, obtained by taking the environment to be in a thermal state $\tau_\nu := \frac{1}{\nu+1}\sum_{n=0}^\infty \left(\frac{\nu}{\nu+1}\right)^n |n\rangle\langle n|$, where $|n\rangle$ is the n-th *Fock state*.

Quantum channels are useful because they can transmit quantum information. The maximum rate at which independent copies of a channel Φ acting on a system

B can simulate instances of the noiseless qubit channel I_2 is called the *quantum capacity* of Φ, and denoted with $Q(\Phi)$. For CV systems, physical transmission of quantum data must be subjected to an energy constraint. We shall assume that the relevant Hamiltonian is the total photon number: for an m-mode system, $H_m := \sum_{j=1}^m a_j^\dagger a_j$. The *energy-constrained quantum capacity* can be obtained thanks to the following modified version [14, Theorem 5] of the Lloyd–Shor–Devetak theorem [58–61]:

$$Q\left(\Phi, N\right) = \sup_k \frac{1}{k} Q_1\big(\Phi^{\otimes k}, kN\big), \qquad (5)$$

$$Q_1(\Phi, N) := \sup_{\mathrm{Tr}[\Psi_B H_B] \leq N} I_{\mathrm{coh}}(A\rangle B)_{(I_A \otimes \Phi_B)(\Psi_{AB})}. \qquad (6)$$

where $\Psi_{AB} := |\Psi\rangle\langle\Psi|_{AB}$ is pure, and $I_{\mathrm{coh}}(A\rangle B)_\rho := \mathrm{Tr}\left[\rho_{AB}\left(\log_2 \rho_{AB} - \log_2 \rho_B\right)\right]$ is the coherent information. The unconstrained quantum capacity is obtained as $Q(\Phi) := \lim_{N\to\infty} Q(\Phi, N)$. In general, the expression in (5) is intractable. However, for the pure loss channel the regularisation is not needed, and the quantum capacity can be expressed in closed form as [10, 12–15]

$$Q\left(\mathscr{E}_\lambda, N\right) = \max\left\{g(\lambda N) - g((1-\lambda)N), 0\right\}, \qquad (7)$$

where $g(x) := (x+1)\log_2(x+1) - x\log_2 x$ is the bosonic entropy. No such formula is known for the thermal attenuators, although sharp bounds are available [10, 15–19].

Results.— Before expounding our findings, let us forge our intuition by looking at other channels that present some analogies with general attenuators. An obvious starting point is the thermal attenuator $\mathscr{E}_{\lambda,\nu} = \Phi_{\lambda,\tau_\nu}$. When $\lambda \leq 1/2$, $\mathscr{E}_{\lambda,\nu}$ is anti-degradable, meaning that tracing out B instead of E in (3) results in a channel that can simulate $\mathscr{E}_{\lambda,\nu}$ via post-processing [11, 54, 62]. This implies that $Q(\mathscr{E}_{\lambda,\nu}) = 0$ for $\lambda \leq 1/2$ [11, p. 3]. On a different note, we can also consider a generalised depolarising channel in finite dimension d, acting as $\rho \mapsto \Delta_{\lambda,\sigma}(\rho) = \lambda\rho + (1-\lambda)\sigma$. As it turns out, its quantum capacity is again zero for $\lambda \leq 1/2$. In fact, $\Delta_{\lambda,\sigma}$ can be obtained from an erasure channel [63] via post-processing. Since the quantum capacity of this latter object is known [64], by data processing we obtain that $Q(\Delta_{\lambda,\sigma}) \leq \max\left\{(1-2\lambda)\log_2 d, 0\right\}$ for all σ. In particular, $Q(\Delta_{\lambda,\sigma}) = 0$ for $\lambda \leq 1/2$.

Our results show that the phenomenology of general attenuators is way richer than these considerations may have suggested. We start by looking at the role of the special point $\lambda = 1/2$.

Theorem 1. *Let σ be an m-mode state of the form $\sigma = \int d\mu(\alpha)\, D(\alpha)\sigma_0(\alpha)D(\alpha)^\dagger$, where $\alpha \in \mathbb{C}^m$, μ is a probability measure on $\mathbb{C}^m$, and the states $\sigma_0(\alpha) = V\sigma_0(\alpha)V^\dagger$ are symmetric under the phase space inversion operation $V := (-1)^{H_m}$, with H_m being the total photon number. Then the channel $\Phi_{1/2,\sigma}$ is anti-degradable [54], and in particular $Q\left(\Phi_{1/2,\sigma}\right) = 0$.*

Proof. Under our assumptions it holds that $\Phi_{1/2,\sigma} = \int d\mu(\alpha)\, \Phi_{1/2, D(\alpha)\sigma_0(\alpha)D(\alpha)^\dagger}$. Now, since the set of anti-degradable channels is convex [65, Appendix A.2], we can directly assume that μ is a Dirac measure, i.e. $\sigma = D(\alpha)\sigma_0 D(\alpha)^\dagger$ with σ_0 symmetric under phase space inversion. Acting on $\rho \otimes \sigma$ with the beam splitter unitary U_λ yields a global state with characteristic function

$$\chi_\rho\left(\sqrt{\lambda}\alpha - \sqrt{1-\lambda}\beta\right)\chi_\sigma\left(\sqrt{1-\lambda}\alpha + \sqrt{\lambda}\beta\right).$$

While the reduced state on the first system is given by (4), that on the second system has characteristic function $\chi_\rho\left(-\sqrt{1-\lambda}\beta\right)\chi_\sigma\left(\sqrt{\lambda}\beta\right)$, which coincides with that of $V\Phi_{1-\lambda, V\sigma V^\dagger}(\rho)V^\dagger$. Therefore, the weak complementary channel associated to $\Phi_{\lambda,\sigma}$ via the representation (3) can be expressed as

$$\Phi_{\lambda,\sigma}^{\mathrm{wc}} = \mathcal{V} \circ \Phi_{1-\lambda, \mathcal{V}(\sigma)},$$

where $\mathcal{V}(\cdot) := V(\cdot)V^\dagger$.

Using the identity $VD(\alpha)V^\dagger = D(-\alpha)$, we see that when $\sigma = D(\alpha)\sigma_0 D(\alpha)^\dagger$ we also have that $\mathcal{V}(\sigma) = \mathcal{D}_{-2\alpha}(\sigma)$, where $\mathcal{D}_z(\cdot) := D(z)(\cdot)D(z)^\dagger$. Noting that $\Phi_{1-\lambda, \mathcal{D}_z(\sigma)} = \mathcal{D}_{\sqrt{\lambda}z} \circ \Phi_{1-\lambda,\sigma}$, we finally obtain that

$$\Phi_{\lambda,\sigma}^{\mathrm{wc}} = \mathcal{V} \circ \mathcal{D}_{-2\sqrt{\lambda}\alpha} \circ \Phi_{1-\lambda,\sigma}.$$

Thus, if $\lambda = 1/2$ the channel is equivalent to its weak complementary up to a unitary post-processing. $\square$

The class of states σ to which Theorem 1 applies is invariant under symplectic unitaries and displacement operators, and it includes many states that are relevant for applications, for instance all convex combinations of Gaussian states (e.g. *classical states* [66, 67]) and all Fock-diagonal states. Remarkably, the above result no longer holds if we weaken the assumption on σ. To see this, for $0 \leq \eta \leq 1$ consider the family of single-mode states $\xi(\eta) = |\xi(\eta)\rangle\langle\xi(\eta)|$, with $|\xi(\eta)\rangle := \sqrt{\eta}\,|0\rangle - \sqrt{1-\eta}\,|1\rangle$. A lower bound on the energy-constrained quantum capacity of the channels $\Phi_{1/2, \xi(\eta)}$ can be obtained by setting $|\Psi(\eta)\rangle_{AB} := \sqrt{\eta(1-\eta)}\,|00\rangle + (1-\eta)\,|01\rangle + \sqrt{\eta}\,|10\rangle$ and by considering that [55]

$$Q\left(\Phi_{1/2, \xi(\eta)}, (1-\eta)^2\right) \geq I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(\eta)}, \qquad (8)$$

where $\zeta_{AB}(1/2,\eta) := \left(I^A \otimes \Phi_{1/2, \xi(\eta)}^B\right)(\Psi_{AB}(\eta))$, and $\Psi(\eta) := |\Psi(\eta)\rangle\langle\Psi(\eta)|$. The function on the r.h.s. of (8) is strictly positive for all $0 < \eta < 1$ [55].

The above example shows that quantum communication can be possible on a general attenuator even for transmissivity $\lambda = 1/2$. At this point, we may wonder whether at least for a fixed energy constraint at the input there exists a threshold value for λ below which quantum communication becomes impossible. Our main result states that this is not the case; on the contrary, the quantum capacity can be bounded away from 0 even

when λ approaches 0, if the environment state σ is chosen appropriately. Note that the bounds by Lim et al. [29] cannot possibly be used to draw such a conclusion [55].

Theorem 2. *For all $0 < \lambda \leq 1$ there exists a single-mode (pure) state $\sigma(\lambda)$ such that*

$$Q\big(\Phi_{\lambda,\,\sigma(\lambda)}\big) \geq Q\big(\Phi_{\lambda,\,\sigma(\lambda)}, 1/2\big) \geq c \qquad (9)$$

for some universal constant $c > 0$. Depending on λ, we can take $\sigma(\lambda)$ to be either the vacuum $|0\rangle$, or a superposition $\alpha|0\rangle + \beta|1\rangle$, or a Fock state $|n\rangle$ with $n \geq 2$.

Sketch of the proof. When $1/2 < \lambda \leq 1$, it suffices to set $\sigma(\lambda) = |0\rangle\langle 0|$ and leverage (S26). Around $\lambda = 1/2$, positive quantum capacity follows by perturbing the lower bound in (8) thanks to the Alicki–Fannes–Winter inequality [68, 69]. It remains to establish the result for $0 < \lambda \leq 1/2 - \epsilon$, where $\epsilon > 0$ is fixed. We start by making an ansatz for a state $|\Psi\rangle_{AB}$ to be plugged into (6). Let us set $|\Psi\rangle_{AB} := \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big)$ and $\sigma(n) := |n\rangle\langle n|$. The output state $\omega_{AB}(n, \lambda) := \big(I^A \otimes \Phi^B_{\lambda,\,\sigma(n)}\big)(\Psi_{AB})$ can be computed e.g. thanks to the formulae derived by Sabapathy and Winter [27, Section III.B]. One obtains that

$$Q\big(\Phi_{\lambda,\,\sigma(n)}, 1/2\big) \geq \mathscr{I}(n, \lambda) := I_{\text{coh}}(A\rangle B)_{\omega_{AB}(n,\lambda)}$$
$$= H\big(p(n, \lambda)\big) - H\big(q(n, \lambda)\big),$$

where the two probability distributions $p(n, \lambda)$ and $q(n, \lambda)$ over the alphabet $\{0, \ldots, n+1\}$ are defined by

$$p_\ell(n, \lambda) := \frac{1}{2(n+1)(1-\lambda)}\binom{n+1}{\ell}(1-\lambda)^\ell \lambda^{n-\ell}$$
$$\times \Big((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2\Big),$$

$$q_\ell(n, \lambda) := \frac{1}{2(n+1)(1-\lambda)}\binom{n+1}{\ell}(1-\lambda)^\ell \lambda^{n-\ell}$$
$$\times \Big(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2\Big).$$

In Figure 2 we plotted $\mathscr{I}(n, \lambda)$ as a function of λ for increasing values of n. The lower endpoint of the range for which $\mathscr{I}(n, \lambda) \geq c$ for some fixed $c > 0$ seems to move closer and closer to 0 as n grows. However, an analytical proof of this fact is technically challenging. The crux of our argument is to show that $p(n, \lambda)$ and $q(n, \lambda)$ are in a majorisation relation, that is, $p(n, \lambda) \prec q(n, \lambda)$ for all $n \geq 2$ and all $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. Given two probability distributions r and s over the same alphabet $\{0, \ldots, N\}$, we say that r is *majorised* by s, and we write $r \prec s$, if $\sum_{\ell=0}^{k} r_\ell^\uparrow \geq \sum_{\ell=0}^{k} s_\ell^\uparrow$ holds for all $k = 0, \ldots, N$, where $r^\uparrow$ and $s^\uparrow$ are obtained by sorting r and s in ascending order [70]. This definition captures the intuitive notion of r being 'more disordered' than s. An immediate consequence is that the entropy of r is never smaller than that of s. But more is true: a beautiful inequality recently established by Ho and Verdú [56, Theorem 3] allows us to lower bound the entropy difference as

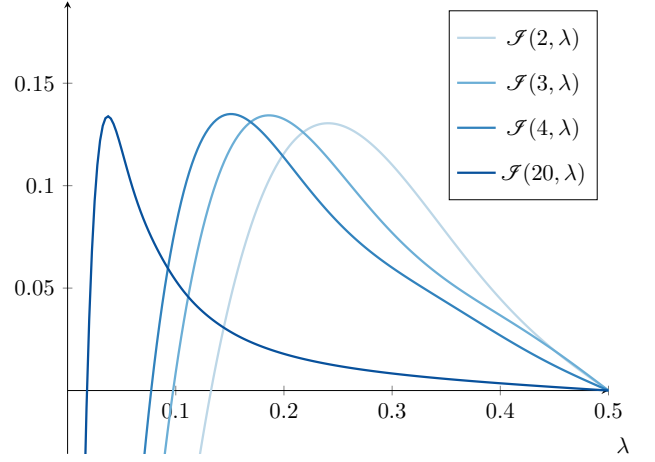$$H(s) - H(r) \geq D\big(s^\uparrow \| r^\uparrow\big), \qquad (10)$$

FIG. 2. The functions $\mathscr{I}(n, \lambda)$ plotted with respect to the variable λ for several values of n.

where $D(u\|v) := \sum_\ell u_\ell \log_2 \frac{u_\ell}{v_\ell}$ is the Kullback–Leibler divergence. This latter quantity can be in turn lower bounded as $D(u\|v) \geq \frac{1}{2\ln 2}\|u - v\|_1^2$ in term of the *total variation distance* $\|u - v\|_1 := \sum_\ell |u_\ell - v_\ell|$ thanks to Pinsker's inequality [71]. We find that

$$\mathscr{I}(n, \lambda) = H(p(n, \lambda)) - H(q(n, \lambda))$$
$$\geq D\big(q^\uparrow(n, \lambda)\big\|p^\uparrow(n, \lambda)\big)$$
$$\geq \frac{1}{2\ln 2}\big\|q^\uparrow(n, \lambda) - p^\uparrow(n, \lambda)\big\|_1^2$$
$$\geq \frac{2}{\ln 2}\big|q_{n+1}^\uparrow(n, \lambda) - p_{n+1}^\uparrow(n, \lambda)\big|^2$$
$$= \frac{2}{\ln 2}\big|p_{n-1}(n, \lambda) - q_{n+1}(n, \lambda)\big|^2,$$

where in the last line we used the fact, proven in the SM [55], that $p_{n-1}(n, \lambda) = \max_\ell p_\ell(n, \lambda)$ and $q_{n+1}(n, \lambda) = \max_\ell q_\ell(n, \lambda)$ for all $n \geq 2$ and $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. It remains to lower bound $k(n, \lambda) := |p_{n-1}(n, \lambda) - q_{n+1}(n, \lambda)|$, which is done by inspection. We find that (a) $k(2, \lambda) \geq \epsilon/4$ for all $1/3 \leq \lambda \leq 1/2 - \epsilon$; and (b) $k(n, \lambda) \geq c$ for some universal constant $c > 0$ for all $n \geq 3$ and $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$, concluding the proof. $\square$

Note that $Q\big(\Phi_{1/2,\,|n\rangle\langle n|}\big) \equiv 0$ for all n by Theorem 1, while we have just shown that $Q\big(\Phi_{\lambda,\,|n\rangle\langle n|}\big) > 0$ when $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. This illustrates the rather surprising fact that $Q\big(\Phi_{\lambda,\sigma}\big)$ can happen not to be monotonic in λ for a fixed σ. In the SM [55] we prove that monotonicity still holds under certain circumstances, e.g. when $\sigma = \sigma_{\text{G}}$ is Gaussian. Combining this with Theorem 1 also shows that $Q\big(\Phi_{\lambda,\sigma_{\text{G}}}\big) \equiv 0$ for all $\lambda \leq 1/2$ and all Gaussian σ_{G}.

From the proof we see that while the energy of the input of the channel in Theorem 2 is fixed, that of the environment state diverges as λ approaches 0. Intuitively, this may be due to the need for the receiver to distinguish the faint low-energy signals, which requires environmental states with highly oscillatory phase space structures

and thus high energy. Whether this reasoning can be made rigorous is left as an open problem.

We now look at the optimal value of the constant c in (9). Our argument yields $c \geq 5.133 \times 10^{-6}$, while numerical investigations suggest that $c \gtrsim 0.066$. If only sufficiently small values of λ are taken into account, we can prove that $c \geq 0.0244$. To put this into perspective, elementary considerations show that $c \leq 1.377$ [55].

Conclusions.— We have studied the transmission of quantum information on general attenuator channels, which are among the simplest examples of non-Gaussian channels and may be relevant for applications. We have shown that their quantum capacity vanishes for transmissivity 1/2 and for a wide class of environment states. At the same time, we have uncovered an unexpected phenomenon: namely, for any non-zero value of the transmissivity there exists an environment state that makes the quantum capacity of the corresponding general attenuator larger than a universal constant. This also implies that said quantum capacity is not necessarily monotonically increasing in the transmissivity for a fixed environment state.

---

* ludovico.lami@gmail.com
† martin.plenio@uni-ulm.de
‡ vittorio.giovannetti@sns.it
§ holevo@mi-ras.ru

[1] E. Knill, R. Laflamme, and G. J. Milburn, Nature **409**, 46 (2001).
[2] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. **77**, 513 (2005).
[3] N. J. Cerf, G. Leuchs, and E. S. Polzik, *Quantum information with continuous variables of atoms and light* (Imperial College Press, 2007).
[4] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).
[5] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (CRC Press, Taylor & Francis Group, 2017).
[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Publications of the Scuola Normale Superiore (Scuola Normale Superiore, 2011).
[7] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction*, 2nd ed., Texts and Monographs in Theoretical Physics (De Gruyter, 2019).
[8] C. D. Cushen and R. L. Hudson, J. Appl. Probab. **8**, 454 (1971).
[9] S. Becker, N. Datta, L. Lami, and C. Rouzé, Preprint arXiv:1912.06129 (2019).
[10] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001).
[11] F. Caruso, V. Giovannetti, and A. S. Holevo, New J. Phys. **8**, 310 (2006).
[12] M. M. Wolf, D. Pérez-García, and G. Giedke, Phys. Rev. Lett. **98**, 130501 (2007).
[13] M. M. Wilde, P. Hayden, and S. Guha, Phys. Rev. A **86**, 062306 (2012).
[14] M. M. Wilde and H. Qi, IEEE Trans. Inf. Theory **64**, 7802 (2018).
[15] K. Noh, V. V. Albert, and L. Jiang, IEEE Trans. Inf. Theory **65**, 2563 (2019).
[16] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Nat. Commun. **8**, 15043 (2017).
[17] M. Rosati, A. Mari, and V. Giovannetti, Nat. Commun. **9**, 4339 (2018).
[18] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, New J. Phys. **20**, 063025 (2018).
[19] K. Noh, S. Pirandola, and L. Jiang, Nat. Commun. **11**, 457 (2020).
[20] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Rev. Mod. Phys. **86**, 1203 (2014).
[21] J. L. O'Brien, A. Furusawa, and J. Vučković, Nat. Photonics **3**, 687 (2009).
[22] A. Politi, J. C. F. Matthews, M. G. Thompson, and J. L. O'Brien, IEEE J. Sel. Top. Quantum Electron. **15**, 1673 (2009).
[23] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Science **349**, 711 (2015).
[24] P. P. Rohde and J. P. Dowling, Science **349**, 696 (2015).
[25] R. Koenig, J. Math. Phys. **56**, 022201 (2015).
[26] G. De Palma and D. Trevisan, Commun. Math. Phys. **360**, 639 (2018).
[27] K. K. Sabapathy and A. Winter, Phys. Rev. A **95**, 062309 (2017).
[28] L. Lami, K. K. Sabapathy, and A. Winter, New J. Phys. **20**, 113012 (2018).
[29] Y. Lim, S. Lee, J. Kim, and K. Jeong, Phys. Rev. A **99**, 052326 (2019).
[30] S. Karumanchi, S. Mancini, A. Winter, and D. Yang, IEEE Trans. Inf. Theory **62**, 1733 (2016).
[31] S. Karumanchi, S. Mancini, A. Winter, and D. Yang, Probl. Inf. Transm. **52**, 214 (2016).
[32] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, Phys. Rev. Lett. **97**, 110501 (2006).
[33] M. Ohliger, K. Kieling, and J. Eisert, Phys. Rev. A **82**, 042336 (2010).
[34] J. Eisert, S. Scheel, and M. B. Plenio, Phys. Rev. Lett. **89**, 137903 (2002).
[35] J. Fiurášek, Phys. Rev. Lett. **89**, 137904 (2002).
[36] G. Giedke and I. J. Cirac, Phys. Rev. A **66**, 032316 (2002).
[37] J. Hoelscher-Obermaier and P. van Loock, Phys. Rev. A **83**, 012319 (2011).
[38] R. Namiki, O. Gittsovich, S. Guha, and N. Lütkenhaus, Phys. Rev. A **90**, 062316 (2014).
[39] J. Niset, J. Fiurášek, and N. J. Cerf, Phys. Rev. Lett. **102**, 120501 (2009).
[40] L. Lami, B. Regula, X. Wang, R. Nichols, A. Winter, and G. Adesso, Phys. Rev. A **98**, 022335 (2018).
[41] L. Lami, R. Takagi, and G. Adesso, Phys. Rev. A **101**, 052305 (2020).

[42] W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).

[43] K. Jeong, Preprint arXiv:2001.01179 (2020).

[44] V. Giovannetti, R. García-Patrón, N. J. Cerf, and A. S. Holevo, Nat. Photonics **8**, 796 (2014).

[45] V. Giovannetti, A. S. Holevo, and R. García-Patrón, Commun. Math. Phys. **334**, 1553 (2015).

[46] G. De Palma, D. Trevisan, and V. Giovannetti, IEEE Trans. Inf. Theory **63**, 728 (2017).

[47] J. Eisert and M. M. Wolf, "Gaussian quantum channels," in *Quantum Information with Continuous Variables of Atoms and Light*, edited by N. J. Cerf, G. Leuchs, and E. S. Polzik (Imperial College Press, 2007) pp. 23–42.

[48] M. M. Wolf, G. Giedke, and J. I. Cirac, Phys. Rev. Lett. **96**, 080502 (2006).

[49] E. Wigner, Phys. Rev. **40**, 749 (1932).

[50] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, Phys. Rep. **106**, 121 (1984).

[51] R. L. Hudson, Rep. Math. Phys. **6**, 249 (1974).

[52] F. Soto-Eguibar and P. Claverie, J. Math. Phys. **24**, 1104 (1983).

[53] T. Bröcker and R. F. Werner, J. Math. Phys. **36**, 62 (1995).

[54] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).

[55] See the Supplemental Material, which contains the references [72–82], for complete proofs of some of the results discussed in the main text.

[56] S. Ho and S. Verdú, IEEE Trans. Inf. Theory **56**, 5930 (2010).

[57] R. Werner, J. Math. Phys. **25**, 1404 (1984).

[58] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).

[59] P. Shor, "Lecture notes," (2002), MSRI Workshop on Quantum Computation.

[60] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005).

[61] A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction*, De Gruyter Studies in Mathematical Physics (De Gruyter, 2012).

[62] L. Lami, S. Khatri, G. Adesso, and M. M. Wilde, Phys. Rev. Lett. **123**, 050501 (2019).

[63] An erasure channel acts as $\rho \mapsto \mathcal{N}_\lambda(\rho) := \lambda\rho + (1 - \lambda)|e\rangle\langle e|$, where $|e\rangle$ is an error flag that is orthogonal to every input state. Constructing the post-processing channel $\rho \mapsto \mathcal{M}_\sigma(\rho) := (\mathbb{1} - |e\rangle\langle e|)\rho(\mathbb{1} - |e\rangle\langle e|) + \langle e|\rho|e\rangle\sigma$, we see that $\Delta_{\lambda,\sigma} = \mathcal{M}_\sigma \circ \mathcal{N}_\lambda$.

[64] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Phys. Rev. Lett. **78**, 3217 (1997).

[65] T. S. Cubitt, M. B. Ruskai, and G. Smith, J. Math. Phys. **49**, 102104 (2008).

[66] A. Bach and U. Lüxmann-Ellinghaus, Commun. Math. Phys. **107**, 553 (1986).

[67] B. Yadin, F. C. Binder, J. Thompson, V. Narasimhachar, M. Gu, and M. S. Kim, Phys. Rev. X **8**, 041038 (2018).

[68] R. Alicki and M. Fannes, J. Phys. A **37**, L55 (2004).

[69] A. Winter, Commun. Math. Phys. **347**, 291 (2016).

[70] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of majorization and its applications*, 2nd ed., Vol. 143 (Springer, 2011).

[71] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*, Probability and Mathematical Statistics (Academic Press, Inc., New York-London, 1981).

[72] P. Jordan, Z. Phys. **94**, 531 (1935).

[73] L. Lami, S. Das, and M. M. Wilde, J. Phys. A **51**, 125301 (2018).

[74] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Graduate Texts in Mathematics (Springer International Publishing, 2015).

[75] S. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics*, Oxford Series in Optical and Imaging Sciences (Clarendon Press, 2002).

[76] A. S. Holevo, Probl. Pered. Inform. **44**, 3 (2008), (English translation: Probl. Inf. Transm. 44(3):171–184, 2008).

[77] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[78] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).

[79] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[80] L. Masanes, Phys. Rev. Lett. **97**, 050503 (2006).

[81] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).

[82] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[83] A small note on the meaning of this word for us. The estimate reported here for $c(1/2)$ is found through a numerical search of the zero of a sum of elementary functions, carried out with Wolfram Mathematica. Apart from this numerical step, whose accuracy is guaranteed by the programme's algorithms, it does not involve any other non-analytical insight, such as – for instance – 'verifying' that a certain function is positive in a certain interval by looking at its graph. A claim of this latter kind would be very far away from a rigorous proof, as it involves keeping under control the infinite number of points that form an interval.

Supplemental material:
Bosonic quantum communication across arbitrarily high loss channels

I. GENERALITIES

A. Quantum entropy

The (von Neumann) *entropy* of a quantum state ρ is defined as

$$S(\rho) := - \operatorname{tr} \left[\rho \log_2 \rho \right], \tag{S1}$$

which is well defined although possibly infinite. Indeed, one way to understand it is via the infinite sum $S(\rho) = \sum_i (-p_i \log_2 p_i)$, where $\rho = \sum_i p_i \, |e_i\rangle\langle e_i|$ is the spectral decomposition of ρ. Since all terms in the above sum are non-negative, the sum itself is well defined but possibly infinite.

Consider an m-mode system with Hilbert space $\mathcal{H}_m = L^2 \left(\mathbb{R}^m \right) \simeq \mathcal{H}_1^{\otimes m}$. The *total photon number* is a densely defined operator on $\mathcal{H}_m$ that takes the form

$$H_m := \sum_{j=1}^{m} a_j^\dagger a_j \tag{S2}$$

when written in terms of the creation and annihilation operators. It is well known to have a discrete spectrum of the form $\left\{ \sum_{j=1}^{m} n_j : n_j \in \mathbb{N} \right\}$, with the eigenvector corresponding to $\sum_{j=1}^{m} n_j$ being given by the tensor product of Fock states $|n_1\rangle \dots |n_m\rangle$. The single-mode thermal state with mean photon number $\nu \geq 0$ is given by

$$\tau_\nu := \frac{1}{\nu+1} \sum_{n=0}^{\infty} \left(\frac{\nu}{\nu+1} \right)^n |n\rangle\langle n| \, . \tag{S3}$$

The thermal state over m modes with *total* mean photon number ν can be easily obtained as the m-fold tensor product $\tau_{\nu/m}^{\otimes m}$. Thermal states are important because they are the maximisers of the entropy among all states with a fixed mean photon number. That is,

$$\max \left\{ S(\rho) : \operatorname{Tr} \left[\rho \, H_m \right] \leq \nu \right\} = S \left(\tau_{\nu/m}^{\otimes m} \right) = m \, g \left(\frac{\nu}{m} \right) \tag{S4}$$

holds for all $\nu \geq 0$, where

$$g(x) := (x+1) \log_2(x+1) - x \log_2 x \, , \tag{S5}$$

sometimes called the *bosonic entropy*, expresses the entropy of a single-mode thermal state in terms of its mean photon number. The function g has many notable properties: (a) it is monotonically increasing; (b) it is subadditive, meaning that

$$g(x+y) \leq g(x) + g(y) \qquad \forall \, x, y \geq 0 \, ; \tag{S6}$$

(c) it is concave; and (d) it has the asymptotic behaviour

$$g(x) = \log_2(ex) + o(1) \qquad (x \to \infty) \, . \tag{S7}$$

B. Beam splitters

A beam splitter is perhaps the simplest example of a passive unitary acting on an $(m+m)$-mode bipartite CV system. As reported in the main text (2), it is defined by $U_\lambda := e^{\arccos \sqrt{\lambda} \sum_j (a_j^\dagger b_j - a_j b_j^\dagger)}$, where a_j, b_j are the annihilation operators on the j-th modes belonging to the first and second system, respectively. This exponential can be

decomposed thanks to a well-known trick. Consider the annihilation operators $a_1, \ldots, a_m$ of m independent modes. The *Jordan map* [72]

$$J: \ X \longmapsto \sum_{j,k=1}^{m} X_{jk} a_j^\dagger a_k, \tag{S8}$$

is a Lie algebra isomorphism between the set of $m \times m$ matrices and that of the operators on the Hilbert space $\mathcal{H}_m$ of m modes that are bilinear in the $a_j^\dagger$ and a_k. Let us note in passing that the Jordan map (S8) can be extended so as to include all operators that can be expressed as polynomials of degree up to 2 in the creation and annihilation operators [73, Appendix A]. Now, since the Baker–Campbell–Hausdorff formula shows that $\ln(e^X e^Y)$ just depends on (nested) commutators between elements of the Lie algebra generated by X and Y, one has that [74, Corollary 3.4]

$$J\left(\ln(e^X e^Y)\right) = \ln\left(e^{J(X)} e^{J(Y)}\right). \tag{S9}$$

In other words, if $e^X e^Y = e^Z$ as matrices, then $e^{J(X)} e^{J(Y)} = e^{J(Z)}$ as operators. This is extremely useful in practical computations. In our case, the exponential that defines a two-mode beam splitter involves only two independent operators a and b. Therefore, the matrix Lie algebra that corresponds to it via the Jordan map is composed of 2×2 matrices. We obtain the explicit correspondence

$$a^\dagger b \longleftrightarrow \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \tag{S10}$$

$$a b^\dagger \longleftrightarrow \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \tag{S11}$$

$$\frac{1}{2}(a^\dagger a - b^\dagger b) \longleftrightarrow \left(\begin{smallmatrix} 1/2 & 0 \\ 0 & -1/2 \end{smallmatrix}\right), \tag{S12}$$

$$\frac{1}{2}(a^\dagger b + a b^\dagger) \longleftrightarrow \left(\begin{smallmatrix} 0 & 1/2 \\ 1/2 & 0 \end{smallmatrix}\right), \tag{S13}$$

$$\frac{1}{2i}(a^\dagger b - a b^\dagger) \longleftrightarrow \left(\begin{smallmatrix} 0 & -i/2 \\ i/2 & 0 \end{smallmatrix}\right). \tag{S14}$$

By exploiting these formulae, performing the computations for 2×2 matrices, and bringing back the result with the Jordan map, it is possible to prove that [75, Appendix 5]

$$U_\lambda = e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b} \, e^{-\frac{1}{2} \ln \lambda \, (a^\dagger a - b^\dagger b)} \, e^{-\sqrt{\frac{1-\lambda}{\lambda}}\, a b^\dagger}. \tag{S15}$$

This decomposition can be employed to derive an expression for the output state obtained by mixing in a beam splitter of arbitrary transmissivity the vacuum $|0\rangle$ or the first Fock state $|1\rangle$ with another Fock states $|n\rangle$. Namely,

$$U_\lambda |0\rangle |n\rangle = \sum_{\ell=0}^{n} \sqrt{\binom{n}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} |\ell\rangle |n-\ell\rangle, \tag{S16}$$

$$U_\lambda |1\rangle |n\rangle = -\frac{1}{\sqrt{(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} ((n+1)(1-\lambda) - \ell) |\ell\rangle |n+1-\ell\rangle. \tag{S17}$$

To prove (S16), write

$$U_\lambda |0\rangle |n\rangle = e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b} \, e^{-\frac{1}{2} \ln \lambda \, (a^\dagger a - b^\dagger b)} \, e^{-\sqrt{\frac{1-\lambda}{\lambda}}\, a b^\dagger} |0\rangle |n\rangle$$

$$= e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b} \, e^{-\frac{1}{2} \ln \lambda \, (a^\dagger a - b^\dagger b)} |0\rangle |n\rangle$$

$$= e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b} \, \lambda^{\frac{n}{2}} |0\rangle |n\rangle$$

$$= \lambda^{\frac{n}{2}} \sum_{\ell=0}^{n} \frac{1}{\ell!} \left(\frac{1-\lambda}{\lambda}\right)^{\frac{\ell}{2}} \left(\sqrt{\ell!} \, |\ell\rangle\right) \left(\sqrt{\frac{n!}{(n-\ell)!}} \, |n-\ell\rangle\right)$$

$$= \sum_{\ell=0}^{n} \sqrt{\binom{n}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} |\ell\rangle |n-\ell\rangle.$$

In the same spirit, one can compute

$$
\begin{aligned}
U_\lambda \left|1\right\rangle \left|n\right\rangle &= e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b}\, e^{-\frac{1}{2}\ln\lambda\,(a^\dagger a - b^\dagger b)}\, e^{-\sqrt{\frac{1-\lambda}{\lambda}}\, ab^\dagger} \left|1\right\rangle \left|n\right\rangle \\
&= e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b}\, e^{-\frac{1}{2}\ln\lambda\,(a^\dagger a - b^\dagger b)} \left(\left|1\right\rangle \left|n\right\rangle - \sqrt{\frac{1-\lambda}{\lambda}}\sqrt{n+1}\left|0\right\rangle \left|n+1\right\rangle \right) \\
&= e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b} \left(\lambda^{\frac{n-1}{2}} \left|1\right\rangle \left|n\right\rangle - \sqrt{\frac{1-\lambda}{\lambda}}\sqrt{n+1}\,\lambda^{\frac{n+1}{2}}\left|0\right\rangle \left|n+1\right\rangle \right) \\
&= \lambda^{\frac{n-1}{2}} e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b}\left|1\right\rangle \left|n\right\rangle - \lambda^{\frac{n}{2}}\sqrt{(1-\lambda)(n+1)}\, e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b}\left|0\right\rangle \left|n+1\right\rangle \\
&= \lambda^{\frac{n-1}{2}} \sum_{\ell=0}^{n} \frac{1}{\ell!}\left(\frac{1-\lambda}{\lambda}\right)^{\frac{\ell}{2}}\left(\sqrt{(\ell+1)!}\left|\ell+1\right\rangle\right)\left(\sqrt{\frac{n!}{(n-\ell)!}}\left|n-\ell\right\rangle\right) \\
&\quad - \lambda^{\frac{n}{2}}\sqrt{(1-\lambda)(n+1)}\sum_{\ell=0}^{n+1}\frac{1}{\ell!}\left(\frac{1-\lambda}{\lambda}\right)^{\frac{\ell}{2}}\left(\sqrt{\ell!}\left|\ell\right\rangle\right)\left(\sqrt{\frac{(n+1)!}{(n+1-\ell)!}}\left|n+1-\ell\right\rangle\right) \\
&= \frac{1}{\sqrt{n+1}}\sum_{\ell=1}^{n+1}\sqrt{\binom{n+1}{\ell}}(1-\lambda)^{\frac{\ell}{2}}\lambda^{\frac{n-\ell}{2}}\,\ell\,\left|\ell\right\rangle\left|n+1-\ell\right\rangle \\
&\quad - \sqrt{(n+1)(1-\lambda)}\sum_{\ell=0}^{n+1}\sqrt{\binom{n+1}{\ell}}(1-\lambda)^{\frac{\ell}{2}}\lambda^{\frac{n-\ell}{2}}\left|\ell\right\rangle\left|n+1-\ell\right\rangle \\
&= \frac{1}{\sqrt{(n+1)(1-\lambda)}}\sum_{\ell=0}^{n+1}\sqrt{\binom{n+1}{\ell}}(1-\lambda)^{\frac{\ell}{2}}\lambda^{\frac{n-\ell}{2}}(\ell - (n+1)(1-\lambda))\left|\ell\right\rangle\left|n+1-\ell\right\rangle\,,
\end{aligned}
$$

which proves (S17).

If instead of (S15) one employs the alternative decomposition

$$
U_\lambda = e^{-\sqrt{\frac{1-\lambda}{\lambda}}\, ab^\dagger}\, e^{\frac{1}{2}\ln\lambda\,(a^\dagger a - b^\dagger b)}\, e^{\sqrt{\frac{1-\lambda}{\lambda}}\, a^\dagger b}\,, \tag{S18}
$$

one finds that

$$
U_\lambda \left|n\right\rangle \left|0\right\rangle = \sum_{\ell=0}^{n}(-1)^\ell \sqrt{\binom{n}{\ell}}(1-\lambda)^{\frac{\ell}{2}}\lambda^{\frac{n-\ell}{2}}\left|n-\ell\right\rangle\left|\ell\right\rangle\,, \tag{S19}
$$

$$
U_\lambda \left|n\right\rangle \left|1\right\rangle = -\frac{1}{\sqrt{(n+1)(1-\lambda)}}\sum_{\ell=0}^{n+1}(-1)^\ell \sqrt{\binom{n+1}{\ell}}(1-\lambda)^{\frac{\ell}{2}}\lambda^{\frac{n-\ell}{2}}((n+1)(1-\lambda)-\ell)\left|n+1-\ell\right\rangle\left|\ell\right\rangle\,. \tag{S20}
$$

Note that (S19) and (S20) can also be derived from (S16) and (S17) by applying the swap operator to both sides of the equations.

Finally, for future convenience we report the expressions of the matrices that represent U_λ on subspaces with low total photon number. By applying (S15) or (S18) one can verify that

$$
U_\lambda\big|_{\mathrm{span}\{\left|0\right\rangle\left|1\right\rangle,\,\left|1\right\rangle\left|0\right\rangle\}} = \begin{pmatrix} \sqrt{\lambda} & -\sqrt{1-\lambda} \\ \sqrt{1-\lambda} & \sqrt{\lambda} \end{pmatrix}\,, \tag{S21}
$$

$$
U_\lambda\big|_{\mathrm{span}\{\left|0\right\rangle\left|2\right\rangle,\,\left|1\right\rangle\left|1\right\rangle,\,\left|2\right\rangle\left|0\right\rangle\}} = \begin{pmatrix} \lambda & -\sqrt{2\lambda(1-\lambda)} & 1-\lambda \\ \sqrt{2\lambda(1-\lambda)} & 2\lambda - 1 & -\sqrt{2\lambda(1-\lambda)} \\ 1-\lambda & \sqrt{2\lambda(1-\lambda)} & \lambda \end{pmatrix}\,. \tag{S22}
$$

C. General attenuators

The family of channels that we consider here is that of *general attenuators* [25–27, 29], sometimes called additive noise channels [25]. They are parametrised by a generic m-mode quantum state σ and by a value of the associated

transmissivity $0 \leq \lambda \leq 1$. As reported in the main text (3)–(4), the action of a general attenuator $\Phi_{\lambda,\sigma}$ on a system B is defined by $\Phi_{\lambda,\sigma}^B(\rho_B) := \mathrm{Tr}_E\left[U_\lambda^{BE}(\rho_B \otimes \sigma_E)\left(U_\lambda^{BE}\right)^\dagger\right]$, which – dropping the system labels for simplicity – translates to $\chi_{\Phi_{\lambda,\sigma}(\rho)}(\alpha) = \chi_\rho\left(\sqrt{\lambda}\,\alpha\right)\chi_\sigma\left(\sqrt{1-\lambda}\,\alpha\right)$ at the level of characteristic functions. This particularly simple expression can be used in conjunction with the composition rule for displacement operators to prove the covariance formulae

$$\Phi_{\lambda,\sigma} \circ \mathcal{D}_z = \mathcal{D}_{\sqrt{\lambda}\,z} \circ \Phi_{\lambda,\sigma}\,, \tag{S23}$$

$$\Phi_{\lambda,\,\mathcal{D}_z(\sigma)} = \mathcal{D}_{\sqrt{1-\lambda}\,z} \circ \Phi_{\lambda,\sigma}\,, \tag{S24}$$

where the displacement channel is defined as

$$\mathcal{D}_z(\rho) := D(z)\,\rho\,D(z)^\dagger\,. \tag{S25}$$

Note that the identity (S24) has been used in the proof of Theorem 1.

The canonical example of a general attenuator channel $\Phi_{\lambda,\sigma}$ – say, in the single-mode case – is obtained by setting $\sigma = \tau_\nu$, where the thermal state with mean photon number ν is defined by (S3). The resulting map $\mathscr{E}_{\lambda,\nu} := \Phi_{\lambda,\,\tau_\nu}$ is usually referred to as a *thermal attenuator*. An even simpler yet extremely important channel, called the *quantum-limited attenuator* (or the *pure loss channel*) and usually denoted with $\mathscr{E}_\lambda := \mathscr{E}_{\lambda,0} = \Phi_{\lambda,\,|0\rangle\langle 0|}$, is obtained by setting the temperature of the environment equal to zero.

The energy-constrained quantum capacity of the pure loss channel has been determined exactly. It reads

$$Q\left(\mathscr{E}_\lambda, N\right) = \max\left\{g(\lambda N) - g\left((1-\lambda)N\right),\, 0\right\}. \tag{S26}$$

The decisive step towards establishing (S26) has been done by Wolf et al. [12, Eq. (12)], who proved that for this particular channel the regularisation in (5) is not needed. This implies that the quantum capacity is simply given by the coherent information (6), which had been previously computed by Holevo and Werner [10, Eq. (5.9)]. A more complete discussion of these latter calculations, and in particular of why it suffices to consider thermal states at the input, can be found in Holevo's monograph [61, Propositions 12.38 and 12.47] (see also the more recent version [7, Propositions 12.40 and 12.62]). The problem of completeness of the original argument was recently raised by Wilde and Qi [14, Remark 4], and further elaborated on by Noh et al. [15, Theorem 9]. An alternative derivation of the formula (S26) has been put forward by Wilde et al. [13].

We do not yet have an exact expression for the energy-constrained capacity of all thermal attenuators. However, many upper [15–18] as well as lower [10, 19] bounds have been discovered so far. We do not report the corresponding formulae here, as we do not need them. What we will need, instead, is a much simpler observation due to Caruso and Giovannetti [11].

Lemma S1 [11, p. 3]. *For all $0 \leq \lambda \leq \frac{1}{2}$ and all $\nu \geq 0$, the thermal attenuator $\mathscr{E}_{\lambda,\nu}$ is anti-degradable, and thus $Q\left(\mathscr{E}_{\lambda,\nu}\right) = 0$.*

The above result can be further generalised thanks to the concept of channel k-extendibility. Here, anti-degradable channels are precisely those that are 2-extendible. The complete characterisation of the k-extendibility regions of all thermal attenuators has been put forward recently [62].

We now turn to the problem of estimating the quantum capacity of general attenuators. We start by recalling the following elementary fact, that is part of the folklore.

Lemma S2. *Let Φ be a quantum channel acting on a system of m modes. For all $N \geq 0$, its energy-constrained quantum capacity satisfies*

$$Q\left(\Phi,\, N\right) \leq m\,g\left(\frac{N}{m}\right),$$

where the bosonic entropy is given by (S5).

Proof. For any bipartite quantum system AB we have that $H(AB) \geq |H(A) - H(B)|$. From this we deduce that the coherent information in (5) satisfies

$$I_{\mathrm{coh}}(A\rangle B)_{(I_A \otimes \Phi_B)(\Psi_{AB})} = \left(H(B) - H(AB)\right)_{(I_A \otimes \Phi_B)(\Psi_{AB})} \leq H(A)_{\Psi_A} = S(\Psi_A) = S(\Psi_B) \leq m\,g\left(\frac{N}{m}\right),$$

where we used (i) the fact that the initial state $\Psi_{AB} = |\Psi\rangle\langle\Psi|_{AB}$ is pure; and (ii) the fact that the thermal state maximises the entropy for a given mean photon number, as stated in (S4). Since the above upper bound is additive, applying the LSD theorem (6) yields the claim. $\square$

Exploiting known extremality properties of Gaussian states [47, 48], the recent solution of the minimum output entropy conjecture [46] (see also [44, 45]), and the even more recently established conditional entropy power inequality [25, 26], Lim et al. [29] were able to prove the following more sophisticated bounds.

Lemma S3 [29, Sections III and IV]. *Let σ be a single-mode state with mean photon number ν_σ and entropy $S(\sigma)$. Then, for all $0 \leq \lambda \leq 1$ and $N \geq 0$ the energy-constrained quantum capacity of the corresponding general attenuator satisfies*

$$g\left((1-\lambda)g^{-1}(S(\sigma)) + \lambda N\right) - S(\sigma) - g\left(\lambda\nu_\sigma + (1-\lambda)N\right) \leq Q\left(\Phi_{\lambda,\sigma}, N\right) \leq g\left(\lambda N + (1-\lambda)\nu_\sigma\right) - \ln\left(\lambda + (1-\lambda)e^{S(\sigma)}\right), \tag{S27}$$

where g^{-1} is the inverse function of the bosonic entropy g defined by (S5).

Remark S4. Note that the lower bound in (S27) always vanishes when $\lambda \leq 1/2$. Indeed, using the subadditivity (S6) and monotonicity of the bosonic entropy yields

$$\begin{aligned}
g\left((1-\lambda)g^{-1}(S(\sigma)) + \lambda N\right) &- S(\sigma) - g\left(\lambda\nu_\sigma + (1-\lambda)N\right) \\
&\leq g\left((1-\lambda)g^{-1}(S(\sigma))\right) + g\left(\lambda N\right) - S(\sigma) - g\left(\lambda\nu_\sigma + (1-\lambda)N\right) \\
&\leq g\left(g^{-1}(S(\sigma))\right) + g\left(\lambda N\right) - S(\sigma) - g\left((1-\lambda)N\right) \\
&= g\left(\lambda N\right) - g\left((1-\lambda)N\right) \\
&\leq 0\,,
\end{aligned}$$

where the last inequality holds provided that $\lambda \leq 1/2$, again using the monotonicity of g. It follows that the recent results by Lim et al. [29] cannot be possibly used to detect a positive quantum capacity below the threshold value $\lambda = 1/2$.

Remark S5. The upper bound in (S27) diverges for every fixed N and $\lambda > 0$ when $\nu_\sigma \to \infty$. However, we have already seen in Lemma S2 that the maximum capacity $Q\left(\Phi_{\lambda,\sigma}, N\right)$ stays finite in the same limit.

D. Quantum capacity and entanglement distribution

Before we move on, let us briefly comment on the problem of entanglement distribution with general attenuators. It is well known that the channel $\mathscr{E}_{\lambda,\nu}$ is entanglement-breaking if and only if $0 \leq \lambda \leq \frac{\nu}{\nu+1}$ [76, Eq. (38)] (see also [62] for the generalisation to k-extendibility). This in particular implies that the pure loss channel, corresponding to the case $\nu = 0$, is not entanglement-breaking for any $\lambda > 0$. In other words, it is possible to distribute entanglement using the pure loss channel for arbitrary small non-zero values of the transmissivity. This can be easily verified with a one-line computation by sending one half of a Bell state $|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right)_{AB}$ through the channel $\mathscr{E}_\lambda^B$. The result is

$$\left(I^A \otimes \mathscr{E}_\lambda^B\right)(\Psi_{AB}) = \frac{1-\lambda}{2}|00\rangle\langle 00|_{AB} + \frac{\lambda}{2}|01\rangle\langle 01|_{AB} + \frac{1}{2}|10\rangle\langle 10|_{AB} + \frac{\sqrt{\lambda}}{2}\left(|01\rangle\langle 10| + |10\rangle\langle 01|\right)_{AB}. \tag{S28}$$

This is effectively a two-qubit state, and the fact that it is indeed entangled for all values of $\lambda > 0$ can be straightforwardly verified by an application of the partial transposition criterion [77]. Since all entangled two-qubit states are distillable [78, 79], we conclude that the state in (S28) is also distillable. A result by Masanes [80] then guarantees that a large number of copies of (S28) can be processed locally – without communication – so as to yield a *non-local state*, that is, a state that violates a Bell inequality [81]. Incidentally, the Bell inequality in question can be taken to be the Clauser–Horne–Shimony–Holt inequality [82]. Therefore, the pure loss channel $\mathscr{E}_\lambda$ can be used (in conjunction with local processing) to construct states that violate a Bell inequality for all $\lambda > 0$.

At first glance, this may appear to contradict the above Lemma S1 [11]. The contradiction is resolved once one observes that sending quantum messages through a channel is a different – and indeed harder – task than that of using it to distribute a non-local state. While any channel with positive quantum capacity can be used to create maximally entangled (and hence non-local) states between sender and receiver, the above example shows that there exist channels with vanishing quantum capacity that can anyway distribute non-local states. What makes our main result (Theorem 2) non-trivial is that while it is relatively easy to show that general attenuators with arbitrary low non-zero values of the transmissivity are fit to the latter task, this does not imply anything about the former (and harder) one.

II. CONVEX COMBINATIONS OF GAUSSIAN STATES

Throughout this section, we look at general attenuators whose environment state is a convex combination of Gaussian states. Note that this family of states encompasses the so-called *classical states* [66, 67], which by definition can be written as convex combinations of coherent states, i.e.

$$\sigma = \int d\mu(\alpha) \, |\alpha\rangle\langle\alpha| \, , \tag{S29}$$

where μ is a probability measure on $\mathbb{C}^m$.

We start by showing how to apply the data processing bound to constrain the quantum capacity of general attenuators.

Lemma S6. *Let $0 \leq \lambda, \mu \leq 1$, and let σ, ω be m-mode states. Then we have the composition rule*

$$\Phi_{\lambda,\sigma} \circ \Phi_{\mu,\omega} = \Phi_{\lambda\mu,\tau} \, , \tag{S30}$$

where

$$\tau := \Phi_{\frac{\lambda(1-\mu)}{1-\lambda\mu},\sigma}(\omega) = \Phi_{\frac{1-\lambda}{1-\lambda\mu},\omega}(\sigma) \, . \tag{S31}$$

Proof. The easiest way to verify (S30) is by looking at the transformation rules for characteristic functions. For an arbitrary input state ρ, using (4) multiple times we obtain that

$$
\begin{aligned}
\chi_{(\Phi_{\lambda,\sigma} \circ \Phi_{\mu,\omega})(\rho)}(\alpha) &= \chi_{\Phi_{\mu,\omega}(\rho)}\left(\sqrt{\lambda}\,\alpha\right) \chi_\sigma\left(\sqrt{1-\lambda}\,\alpha\right) \\
&= \chi_\rho\left(\sqrt{\lambda\mu}\,\alpha\right) \chi_\omega\left(\sqrt{\lambda(1-\mu)}\,\alpha\right) \chi_\sigma\left(\sqrt{1-\lambda}\,\alpha\right) \\
&= \chi_\rho\left(\sqrt{\lambda\mu}\,\alpha\right) \chi_\eta\left(\sqrt{1-\lambda\mu}\,\alpha\right) \\
&= \chi_{\Phi_{\lambda\mu,\eta}(\rho)}(\alpha) \, .
\end{aligned}
$$

Since quantum states are in one-to-one correspondence with characteristic functions, this implies that $\Phi_{\lambda\mu,\eta}(\rho) = (\Phi_{\lambda,\sigma} \circ \Phi_{\mu,\omega})(\rho)$. Given that ρ was arbitrary, the proof is complete. $\square$

A first immediate corollary is as follows.

Corollary S7. *Let $0 \leq \lambda, \mu \leq 1$, and let σ, ω be m-mode states. Define τ as in (S31). Then: (a) if $\Phi_{\mu,\omega}$ is anti-degradable, then so is $\Phi_{\lambda\mu,\tau}$; (b) it holds that*

$$Q\left(\Phi_{\lambda\mu,\tau}\right) \leq \min\left\{\Phi_{\lambda,\sigma}, \, \Phi_{\mu,\omega}\right\}. \tag{S32}$$

Proof. Claim (a) is a consequence of the fact that set of anti-degradable channels is invariant by post-processing. Claim (b), instead, follows from the observation that in the definition of quantum capacity any pre- or post-processing can be included into the encoding or decoding transformations. $\square$

We now show that the phenomenon illustrated in Theorem 2 does not occur for general attenuators whose environment state is a convex combination of Gaussian states.

Corollary S8. *Let σ be a state in the convex hull of all Gaussian states. Then $\Phi_{\lambda,\sigma}$ is anti-degradable for all $0 \leq \lambda \leq 1/2$, and in particular*

$$Q\left(\Phi_{\lambda,\sigma}\right) \equiv 0 \qquad \forall \, 0 \leq \lambda \leq \frac{1}{2}\,. \tag{S33}$$

Proof. For σ satisfying the hypothesis, we have that $\Phi_{\lambda,\sigma}$ is a convex combination of channels of the form $\Phi_{\lambda,\sigma_{\mathrm{G}}}$, where σ_{G} is Gaussian. Since the set of anti-degradable channels is convex [65, Appendix A.2], it suffices to prove that $\Phi_{\lambda,\sigma_{\mathrm{G}}}$ is anti-degradable for all $0 \leq \lambda \leq 1/2$.

Moreover, we can assume without loss of generality that σ_{G} is centred, i.e. that $\mathrm{Tr}[\sigma_{\mathrm{G}}\,a_j] \equiv 0$ for all $j = 1, \ldots, m$, where a_j are the annihilation operators. In fact, σ_{G} can always be displaced by an arbitrary amount by means of a

unitary post-processing as in (S24). Note that unitary post-processing does not affect anti-degradability, and that $D(z)a_j D(z)^\dagger = a_j - z_j$. Thus, we can make sure that $\mathcal{D}_z(\sigma_G)$ is centred by choosing z appropriately.

In light of the above reasoning, from now on we shall assume that σ_G is centred. The characteristic function of σ_G then is a centred Gaussian, entailing that

$$\chi_{\sigma_G}(\sqrt{\eta}\,\alpha)\,\chi_{\sigma_G}\left(\sqrt{1-\eta}\,\alpha\right) \equiv \chi_{\sigma_G}(\alpha) \qquad \forall\,\alpha \in \mathbb{C}^m, \quad \forall\, 0 \le \eta \le 1\,.$$

Using (4), this translates to

$$\Phi_{\eta,\,\sigma_G}(\sigma_G) \equiv \sigma_G \qquad \forall\, 0 \le \eta \le 1\,. \tag{S34}$$

Leveraging (S30)–(S31), we see that

$$\Phi_{\lambda,\,\sigma_G} = \Phi_{2\lambda,\,\sigma_G} \circ \Phi_{1/2,\,\sigma_G}$$

for all $0 \le \lambda \le 1/2$. Note that $\Phi_{1/2,\,\sigma_G}$ is anti-degradable by Theorem 1. Since anti-degradable channels remain such upon post-processing [65, Lemma 17], we conclude that also $\Phi_{\lambda,\,\sigma_G}$ is anti-degradable, completing the proof. $\qquad\square$

Another consequence of Lemma S6 is that the quantum capacity of a general attenuator is monotonically increasing as a function of the transmissivity for a fixed Gaussian environment state. By comparison, remember that in the main text we have instead shown that monotonicity fails to hold when the environment state is a Fock state.

Corollary S9. *Let $\sigma = \sigma_G$ be an arbitrary m-mode Gaussian state. Then the function*

$$\lambda \longmapsto Q\left(\Phi_{\lambda,\,\sigma_G}\right) \tag{S35}$$

is monotonically increasing for all $0 \le \lambda \le 1$, and strictly zero for $0 \le \lambda \le 1/2$.

Proof. The proof is along the same lines as that of Corollary S8. We can assume without loss of generality that σ_G is centred, which in turn implies that (S34) holds. Picking $0 \le \lambda' \le \lambda \le 1$ and setting $\mu := \frac{\lambda'}{\lambda}$ and $\eta := \frac{1-\lambda}{1-\lambda\mu} = \frac{1-\lambda}{1-\lambda'}$ in (S31)–(S32), we deduce that

$$\Phi_{\lambda',\,\sigma_G} = \Phi_{\lambda\mu,\,\sigma_G} = \Phi_{\lambda\mu,\,\Phi_{\eta,\,\sigma_G}(\sigma_G)} = \Phi_{\lambda,\sigma_G} \circ \Phi_{\mu,\,\sigma_G}\,.$$

Then, applying (S32) we conclude that $Q\left(\Phi_{\lambda',\,\sigma_G}\right) \le Q\left(\Phi_{\lambda,\sigma_G}\right)$, completing the proof. $\qquad\square$

III.   POSITIVE CAPACITY AT $\lambda = 1/2$

Given the fact that general attenuator channels of the form $\Phi_{1/2,\,\sigma}$ always output states with positive Wigner functions [9, Lemma 16], one may be tempted to conjecture that their quantum capacities vanish. Interestingly, this is not the case, as the next example shows.

Example S10. For $0 \le \eta \le 1$, set $\xi(\eta) = |\xi(\eta)\rangle\langle\xi(\eta)|$, with

$$|\xi(\eta)\rangle := \sqrt{\eta}\,|0\rangle - \sqrt{1-\eta}\,|1\rangle\,. \tag{S36}$$

We will see (numerically) that for all $\eta \in (0,1)$ the channel $\Phi_{1/2,\,\xi(\eta)}$ has nonzero quantum capacity. A lower bound on $Q\left(\Phi_{1/2,\,\xi(\eta)}\right)$ is plotted in Figure 2.

To estimate the quantum capacity of $\Phi_{1/2,\,\xi(\eta)}$ from below, we apply the achievability part of the LSD theorem. This is done by finding a suitable ansatz for the state $|\Psi\rangle_{AB}$ to be plugged into (6). Let us define the family of two-mode states

$$|\Psi(\eta)\rangle_{AB} := \sqrt{\eta(1-\eta)}\,|0\rangle_A\,|0\rangle_B + (1-\eta)\,|0\rangle_A\,|1\rangle_B + \sqrt{\eta}\,|1\rangle_A\,|0\rangle_B\,. \tag{S37}$$

Upon re-ordering the terms, the joint state reads

$$\begin{aligned}|\Psi(\eta)\rangle_{AB}\,|\xi(\eta)\rangle_E = {}& \eta\left(|1\rangle_A + \sqrt{1-\eta}\,|0\rangle_A\right)|0\rangle_B\,|0\rangle_E - \sqrt{\eta(1-\eta)}\left(\sqrt{1-\eta}\,|0\rangle_A + |1\rangle_A\right)|0\rangle_B\,|1\rangle_E \\ & + \sqrt{\eta}(1-\eta)\,|0\rangle_A\,|1\rangle_B\,|0\rangle_E - (1-\eta)^{3/2}\,|0\rangle_A\,|1\rangle_B\,|1\rangle_E\,.\end{aligned} \tag{S38}$$

Using the explicit representations (S21)–(S22) of the action of the beam splitter unitary on the low photon number subspaces, it is not difficult to see that the tripartite output state, which we denote as

$$|\zeta(\lambda,\eta)\rangle_{ABE} := U_\lambda^{BE} |\Psi(\eta)\rangle_{AB} |\xi(\eta)\rangle_E , \tag{S39}$$

reads

$$
\begin{aligned}
|\zeta(\lambda,\eta)\rangle_{ABE} &= \eta \left(|1\rangle_A + \sqrt{1-\eta}\,|0\rangle_A\right) |0\rangle_B |0\rangle_E - \sqrt{\eta(1-\eta)} \left(\sqrt{1-\eta}\,|0\rangle_A + |1\rangle_A\right)\left(\sqrt{\lambda}\,|0\rangle_B |1\rangle_E + \sqrt{1-\lambda}\,|1\rangle_B |0\rangle_E\right) \\
&\quad + \sqrt{\eta}(1-\eta)\,|0\rangle_A \left(-\sqrt{1-\lambda}\,|0\rangle_B |1\rangle_E + \sqrt{\lambda}\,|1\rangle_B |0\rangle_E\right) \\
&\quad - (1-\eta)^{3/2}\,|0\rangle_A \left(-\sqrt{2\lambda(1-\lambda)}\,|0\rangle_B |2\rangle_E + (2\lambda-1)\,|1\rangle_B |1\rangle_E + \sqrt{2\lambda(1-\lambda)}\,|2\rangle_B |0\rangle_E\right) \\
&= \left(\eta\sqrt{1-\eta}\,|0\rangle_A |0\rangle_B + \sqrt{\eta}(1-\eta)\left(\sqrt{\lambda}-\sqrt{1-\lambda}\right)|0\rangle_A |1\rangle_B \right. \\
&\qquad \left. + \eta\,|1\rangle_A |0\rangle_B - \sqrt{\eta(1-\eta)(1-\lambda)}\,|1\rangle_A |1\rangle_B - (1-\eta)^{3/2}\sqrt{1-\lambda}\,|0\rangle_A |2\rangle_B \right) |0\rangle_E \\
&\quad - \left(\sqrt{\eta}(1-\eta)\left(\sqrt{\lambda}+\sqrt{1-\lambda}\right)|0\rangle_A |0\rangle_B + (1-\eta)^{3/2}(2\lambda-1)\,|0\rangle_A |1\rangle_B + \sqrt{\eta(1-\eta)\lambda}\,|1\rangle_A |0\rangle_B\right) |1\rangle_E \\
&\quad + (1-\eta)^{3/2}\sqrt{2\lambda(1-\lambda)}\,|0\rangle_A |0\rangle_B |2\rangle_E \,.
\end{aligned}
$$

From now on, we focus only on the case $\lambda = 1/2$. Upon tedious yet straightforward calculations, we find that with respect to the lexicographically ordered product basis $\{|0\rangle_A, |1\rangle_A\} \otimes \{|0\rangle_B, |1\rangle_B, |2\rangle_B\}$ we have that

$$
\zeta_{AB}(1/2,\,\eta) = \left(I^A \otimes \Phi_{1/2,\,\xi(\eta)}^B\right)\left(\Psi_{AB}(\eta)\right) =
\begin{pmatrix}
\frac{1}{2}\left(1+\eta-3\eta^2+\eta^3\right) & 0 & -\frac{(1-\eta)^2\eta}{\sqrt{2}} & \eta\sqrt{1-\eta} & -\frac{(1-\eta)\eta^{3/2}}{\sqrt{2}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{(1-\eta)^2\eta}{\sqrt{2}} & 0 & \frac{1}{2}(1-\eta)^3 & -\frac{(1-\eta)^{3/2}\eta}{\sqrt{2}} & \frac{1}{2}(1-\eta)^2\sqrt{\eta} & 0 \\
\eta\sqrt{1-\eta} & 0 & -\frac{(1-\eta)^{3/2}\eta}{\sqrt{2}} & \frac{1}{2}\eta(1+\eta) & -\frac{\eta^{3/2}\sqrt{1-\eta}}{\sqrt{2}} & 0 \\
-\frac{(1-\eta)\eta^{3/2}}{\sqrt{2}} & 0 & \frac{1}{2}(1-\eta)^2\sqrt{\eta} & -\frac{\eta^{3/2}\sqrt{1-\eta}}{\sqrt{2}} & \frac{1}{2}(1-\eta)\eta & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\tag{S40}
$$

and that

$$
\zeta_B(1/2,\,\eta) = \Phi_{1/2,\,\xi(\eta)}^B\left(\Psi_B(\eta)\right) =
\begin{pmatrix}
\frac{1}{2}\left(1+2\eta-2\eta^2+\eta^3\right) & -\frac{\eta^{3/2}\sqrt{1-\eta}}{\sqrt{2}} & -\frac{(1-\eta)^2\eta}{\sqrt{2}} \\
-\frac{\eta^{3/2}\sqrt{1-\eta}}{\sqrt{2}} & \frac{1}{2}(1-\eta)\eta & 0 \\
-\frac{(1-\eta)^2\eta}{\sqrt{2}} & 0 & \frac{1}{2}(1-\eta)^3
\end{pmatrix}.
\tag{S41}
$$

We are now ready to apply the LSD theorem to our case. Note that the mean photon number of $\phi_B(\eta)$ is precisely $(1-\eta)^2$. Then, employing (5)–(6) we find that

$$Q\left(\Phi_{1/2,\,\xi(\eta)}\right) \geq Q\left(\Phi_{1/2,\,\xi(\eta)},\,(1-\eta)^2\right) \geq I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(1/2,\,\eta)}. \tag{S42}$$

The coherent information $I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(1/2,\,\eta)}$ is plotted in Figure 2. The numerics shows clearly that this is strictly positive for all $\eta \in (0,1)$. We do not provide an analytical proof of this claim, because it is not necessary for what follows. In our proof of Theorem 2 we will only use the easily verified fact that $I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(1/2,\,\eta)} > 0$ for *some* values of η.

Incidentally, general attenuators can have a substantially larger transmission capacity if one allows for a higher input power to be deployed.

Example S11. For $n \geq 3$ to be fixed, consider the environment state $\xi'(n) = |\xi'(n)\rangle\langle\xi'(n)|$, with

$$|\xi'(n)\rangle := \frac{|n-1\rangle + |n\rangle}{\sqrt{2}}\,. \tag{S43}$$

We look at the transmission scheme identified by an initial state

$$|\Psi'(n)\rangle_{AB} := \frac{1}{2}\left(|0\rangle_A \left(|n-1\rangle_B + |n\rangle_B\right) + |1\rangle_A \left(|n-3\rangle_B + |n-2\rangle_B\right)\right). \tag{S44}$$
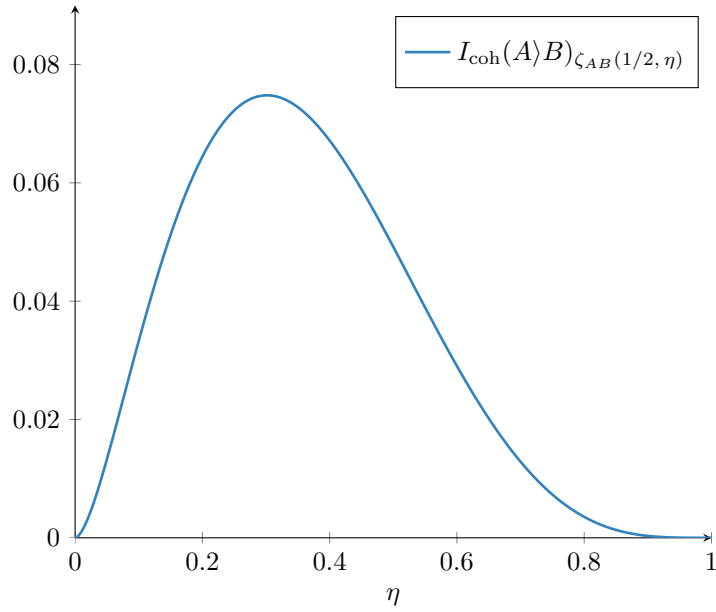
FIG. 1. The coherent information lower bound (8) on the quantum capacity of the channel $\Phi_{1/2,\,\xi(\eta)}$ defined by the environment state (S36). The maximum can be numerically evaluated, yielding $\max_{0 \leq \eta \leq 1} I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(1/2,\,\eta)} \approx 0.0748$.

Note that the mean photon number of $\Psi'_B(n)$ is $n - \frac{3}{2}$. Applying the LSD theorem in the form of (5)–(6) then yields

$$Q\left(\Phi_{1/2,\,\xi'(n)}\right) \geq Q\left(\Phi_{1/2,\,\xi'(n)},\, n - \frac{3}{2}\right) \geq I_{\mathrm{coh}}(A\rangle B)_{\zeta'_{AB}(1/2,\,n)}, \tag{S45}$$

with $\zeta'_{AB}(1/2,\, n) := \left(I^A \otimes \Phi_{1/2,\,\xi'(n)}\right)(\Psi'_{AB}(n))$. The values of the right-hand side of (S45) for $n = 3, \ldots, 35$ are reported in Figure 2. For $n = 54$ the lower bound evaluates to around 0.3530.
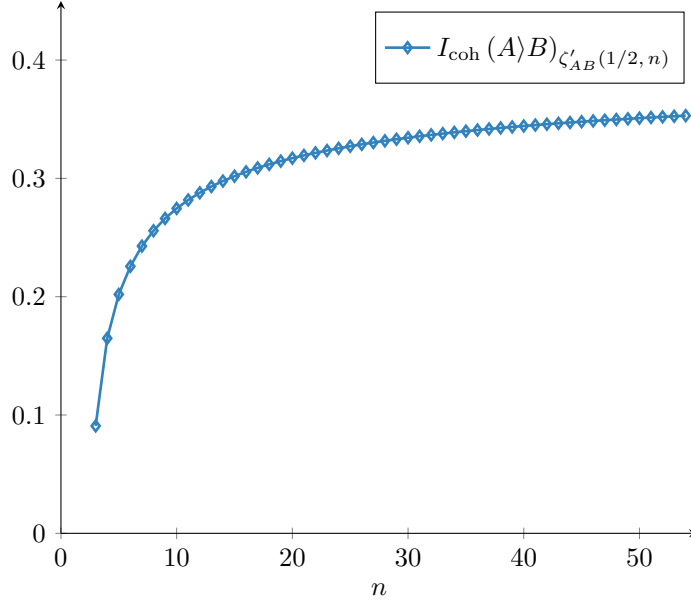

FIG. 2. The coherent information lower bound (S45) on the quantum capacity of the channel $\Phi_{1/2,\,\xi'(n)}$ defined by the environment state in (S43). For $n = 54$ we obtain that $I_{\mathrm{coh}}(A\rangle B)_{\zeta'_{AB}(1/2,\,54)} \approx 0.3530$.

IV. POSITIVE CAPACITY AT ARBITRARY TRANSMISSIVITY

Theorem 2. For all $0 < \lambda \leq 1$ there exists a single-mode pure state $\sigma(\lambda)$ such that

$$Q\big(\Phi_{\lambda,\,\sigma(\lambda)}\big) \geq Q\big(\Phi_{\lambda,\,\sigma(\lambda)},\,1/2\big) \geq c \tag{S46}$$

for some universal constant $c > 0$. For an arbitrary $\epsilon > 0$, one can take $\sigma(\lambda)$ to be: (a) the vacuum for $1/2 + \epsilon \leq \lambda \leq 1$; (b) the state $\xi(1/3)$ defined by (S36) for $1/2 - \epsilon \leq \lambda \leq 1/2 + \epsilon$; (c) the Fock state $|2\rangle\langle 2|$ for $1/3 \leq \lambda \leq 1/2 - \epsilon$; and (d) the Fock state $|n\rangle\langle n|$ for $1/(n+1) \leq \lambda \leq 1/n$, for all integers $n \geq 3$.

Remark S12. The value of the constant c in Theorem 2 will depend on ϵ. For the numerical determination of the optimal value yielded by our argument, see Remark S22. We also examine a closely related question in the subsequent Remark S23.

Throughout this section we will provide a complete proof of the above result. In light of its complexity, we will break it down into several elementary steps. Further considerations and some extensions of Theorem 2 are postponed to the end of this section. Here is a brief account of the content of the various subsections:

IV A. Here we fix a transmission scheme, that is, a family of environment states $\sigma(n)$ (for $n = 1, 2, \ldots$) and an ansatz $|\Psi\rangle_{AB}$ to be plugged into the coherent information (6). The result will be a lower bound of the form $Q\left(\Phi_{\lambda,\,\sigma(\lambda)}\right) \geq H(p(n,\lambda)) - H(q(n,\lambda))$, where $p(n,\lambda)$ and $q(n,\lambda)$ are appropriate probability distributions over some index $\ell \in \{0, \ldots, n+1\}$ (Proposition S13).

IV B. We will then proceed to identify a range of values of λ (depending on n) for which $q(n,\lambda)$ can be sorted in ascending order by a fixed permutation (luckily enough, this turns out to be the identity). The result is contained in Proposition S14.

IV C. The same is then done for $p(n,\lambda)$, with considerably more effort and by keeping three distinct possibilities on the table (Proposition S18).

IV D. The crux of the argument is to verify that for a sufficiently large range of values of λ (for varying n) the probability distribution $q(n,\lambda)$ actually *majorises* $p(n,\lambda)$.

IV E. The existence of a majorisation relation between $q(n,\lambda)$ and $p(n,\lambda)$ allows us to exploit a beautiful inequality due to Ho and Verdú [56, Theorem 3] to lower bound their entropy difference by means of the relative entropy distance. In turn, this can be lower bounded in terms of their total variation distance thanks to Pinsker's inequality (see also [71, p. 58] and references therein). We look at the resulting bounds and draw our conclusions.

IV F As a small addition to the actual proof of Theorem 2 (that concludes itself in Section IV E), we show that, if one excludes the two base cases $n = 2, 3$, the majorisation between $q(n,\lambda)$ and $p(n,\lambda)$ becomes of a very special type (Proposition S29).

A. A transmission scheme

Proposition S13. Set $|\Psi\rangle_{AB} := \frac{1}{\sqrt{2}}\big(|0\rangle_A |1\rangle_B + |1\rangle_A |0\rangle_B\big)$ and $\sigma(n) := |n\rangle\langle n|$. Then for all $\lambda \in (0,1)$ it holds that

$$Q\left(\Phi_{\lambda,\,\sigma(n)}, 1/2\right) \geq \mathscr{I}(n,\lambda) := I_{\mathrm{coh}}(A\rangle B)_{\left(I^A \otimes \Phi^B_{\lambda,\,\sigma(n)}\right)(\Psi_{AB})} = H\left(p(n,\lambda)\right) - H\left(q(n,\lambda)\right), \tag{S47}$$

where H denotes the Shannon entropy, and the two probability distributions $p(n,\lambda) = (p_0(n,\lambda), \ldots, p_{n+1}(n,\lambda))$ and $q(n,\lambda) = (q_0(n,\lambda), \ldots, q_{n+1}(n,\lambda))$ are defined by

$$p_\ell(n,\lambda) := \frac{1}{2(n+1)(1-\lambda)} \binom{n+1}{\ell} (1-\lambda)^\ell \lambda^{n-\ell} \left((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2\right) \tag{S48}$$

$$q_\ell(n,\lambda) := \frac{1}{2(n+1)(1-\lambda)} \binom{n+1}{\ell} (1-\lambda)^\ell \lambda^{n-\ell} \left(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2\right) \tag{S49}$$

Proof. Thanks to (S16) and (S17), the action of the beam splitter on the BE system can be expressed as

$$U_\lambda^{BE} |\Psi\rangle_{AB} |n\rangle_E$$

$$= -\frac{1}{\sqrt{2}} |0\rangle_A \left(\frac{1}{\sqrt{(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} ((n+1)(1-\lambda) - \ell) |\ell\rangle_B |n+1-\ell\rangle_E \right)$$

$$+ \frac{1}{\sqrt{2}} |1\rangle_A \left(\sum_{\ell=0}^{n} \sqrt{\binom{n}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} |\ell\rangle_B |n-\ell\rangle_E \right)$$

$$= \frac{1}{\sqrt{2(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}}$$

$$\times \left(-((n+1)(1-\lambda) - \ell) |0\rangle_A |\ell\rangle_B |n+1-\ell\rangle_E + \sqrt{(1-\lambda)(n-\ell+1)} |1\rangle_A |\ell\rangle_B |n-\ell\rangle_E \right)$$

$$= \frac{1}{\sqrt{2(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} \left(-((n+1)(1-\lambda) - \ell) |0\rangle_A |\ell\rangle_B + \sqrt{\lambda\ell} |1\rangle_A |\ell-1\rangle_B \right) |n+1-\ell\rangle_E,$$

with the convention that $|-1\rangle \equiv 0$. Introducing the normalised vectors

$$|\zeta_\ell(n,\lambda)\rangle := \frac{1}{\sqrt{(1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2}} \left(-((n+1)(1-\lambda) - \ell) |0\rangle|n+1-\ell\rangle + \sqrt{(1-\lambda)(n-\ell+1)} |1\rangle|n-\ell\rangle \right)$$

$$\text{(S50)}$$

$$|\eta_\ell(n,\lambda)\rangle := \frac{1}{\sqrt{\lambda\ell + ((n+1)(1-\lambda) - \ell)^2}} \left(-((n+1)(1-\lambda) - \ell) |0\rangle|\ell\rangle + \sqrt{\lambda\ell} |1\rangle|\ell-1\rangle \right), \qquad \text{(S51)}$$

for $\ell = 0, \ldots, n+1$, we finally arrive at

$$U_\lambda^{BE} |\Psi\rangle_{AB} |n\rangle_E$$

$$= \frac{1}{\sqrt{2(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} \sqrt{(1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2} \, |\zeta_\ell(n,\lambda)\rangle_{AE} |\ell\rangle_B. \quad \text{(S52)}$$

$$= \frac{1}{\sqrt{2(n+1)(1-\lambda)}} \sum_{\ell=0}^{n+1} \sqrt{\binom{n+1}{\ell}} (1-\lambda)^{\frac{\ell}{2}} \lambda^{\frac{n-\ell}{2}} \sqrt{\lambda\ell + ((n+1)(1-\lambda) - \ell)^2} \, |\eta_\ell(n,\lambda)\rangle_{AB} |n+1-\ell\rangle_E. \quad \text{(S53)}$$

Tracing away the subsystem E from (S53) yields the output state of the channel as

$$\omega_{AB}(n,\lambda) := \left(I^A \otimes \Phi_{\lambda,\sigma(n)}^B \right) (\Psi_{AB})$$

$$= \frac{1}{2(n+1)(1-\lambda)} \sum_{\ell=0}^{n+1} \binom{n+1}{\ell} (1-\lambda)^\ell \lambda^{n-\ell} \left(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2 \right) |\eta_\ell(n,\lambda)\rangle\langle\eta_\ell(n,\lambda)|_{AB}. \quad \text{(S54)}$$

Note that the total photon number of the state $|\eta_\ell\rangle$ is exactly ℓ, for all $\ell = 0, \ldots, n+1$:

$$(a^\dagger a + b^\dagger b) |\eta_\ell\rangle_{AB} = \ell |\eta_\ell\rangle_{AB}. \qquad \text{(S55)}$$

Hence, the vectors $|\eta_\ell\rangle$ are all orthogonal to each other. This allows us to immediately deduce the spectrum of $\omega_{AB}(n,\lambda)$. We obtain that

$$\mathrm{sp}\,(\omega_{AB}(n,\lambda)) = \{q_0(n,\lambda), \ldots, q_{n+1}(n,\lambda)\}, \qquad \text{(S56)}$$

where the probability distribution $q(n,\lambda)$ is given by (S49).

To derive an expression for $\omega_B(n,\lambda) = \Phi_{\lambda,\sigma(n)}^B (\Psi_B)$ we could trace away A from (S54). However, it is slightly more convenient to read off the result directly from (S52). We obtain that

$$\omega_B(n,\lambda) = \Phi_{\lambda,\sigma(n)}^B (\Psi_B) = \frac{1}{2(n+1)(1-\lambda)} \sum_{\ell=0}^{n+1} \binom{n+1}{\ell} (1-\lambda)^\ell \lambda^{n-\ell} \left((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2 \right) |\ell\rangle\langle\ell|_B$$

$$\text{(S57)}$$

The above decomposition allows us to write down the spectrum of the reduced output state on the B system immediately. We obtain that

$$\mathrm{sp}\,(\omega_B(n,\lambda)) = \{p_0(n,\lambda),\ldots,p_{n+1}(n,\lambda)\}\,, \tag{S58}$$

where the probability distribution $p(n,\lambda)$ is given by (S48).

Since the reduced input state Ψ_B on the B system has mean photon number $1/2$, the (energy-constrained) LSD theorem (5)–(6) yields the estimate in (S47), thus concluding the proof. $\qquad\square$

B. Sorting $q(n,\lambda)$

In the following, for a given probability distribution $r = (r_0,\ldots,r_N)$, we denote with $r^\uparrow = \left(r_0^\uparrow,\ldots,r_N^\uparrow\right)$ the distribution obtained by sorting it in ascending order, so that e.g. $r_0^\uparrow = \min_{\ell=0,\ldots,N} r_\ell$. Our first result tells us that for a wide range of values of λ the distribution $q(n,\lambda)$ is actually already sorted. It is useful to define the two functions

$$\lambda_+(n) := \frac{3}{n+2}\left(1 - \sqrt{\frac{n-1}{3(n+1)}}\right), \tag{S59}$$

$$\lambda_-(n) := \frac{2}{n+2}\left(1 - \sqrt{\frac{n}{2(n+1)}}\right). \tag{S60}$$

We are now ready to state and prove our first result.

Proposition S14. *For all $n \geq 2$,*

$$q^\uparrow(n,\lambda) = q(n,\lambda) \qquad \forall\ \ \frac{1}{n+1} \leq \lambda \leq \lambda_+(n)\,. \tag{S61}$$

Proof. For $\ell = 0,\ldots,n$, leveraging the fact that

$$\frac{\binom{n+1}{\ell+1}}{\binom{n+1}{\ell}} = \frac{n-\ell+1}{\ell+1} \tag{S62}$$

the formula (S49) yields

$$\lambda(\ell+1)\left(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2\right)\left(\frac{q_{\ell+1}(n,\lambda)}{q_\ell(n,\lambda)} - 1\right)$$

$$= (n-\ell+1)(1-\lambda)\left(\lambda(\ell+1) + ((n+1)(1-\lambda) - \ell - 1)^2\right) - \lambda(\ell+1)\left(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2\right)$$

$$= \lambda(\ell+1)\left((n-\ell+1)(1-\lambda) - \lambda\ell - ((n+1)(1-\lambda) - \ell)^2\right) + (n-\ell+1)(1-\lambda)((n+1)(1-\lambda) - \ell - 1)^2$$

$$= -\lambda(\ell+1)((n+1)(1-\lambda) - \ell - 1)((n+1)(1-\lambda) - \ell) + (n-\ell+1)(1-\lambda)((n+1)(1-\lambda) - \ell - 1)^2$$

$$= ((n+1)(1-\lambda) - \ell - 1)\left(-\lambda(\ell+1)((n+1)(1-\lambda) - \ell) + (n-\ell+1)(1-\lambda)((n+1)(1-\lambda) - \ell - 1)\right)$$

$$= ((n+1)(1-\lambda) - \ell - 1)\left(\ell^2 - 2\left((n+1)(1-\lambda) - \frac{1}{2}\right)\ell + (n+1)(1-\lambda)(n - (n+2)\lambda)\right).$$

Setting

$$f_{n,\lambda}(\ell) := \ell^2 - 2\left((n+1)(1-\lambda) - \frac{1}{2}\right)\ell + (n+1)(1-\lambda)(n - (n+2)\lambda)\,, \tag{S63}$$

we arrive at the identity

$$\lambda(\ell+1)\left(\lambda\ell + ((n+1)(1-\lambda) - \ell)^2\right)\left(\frac{q_{\ell+1}(n,\lambda)}{q_\ell(n,\lambda)} - 1\right) = ((n+1)(1-\lambda) - \ell - 1)\,f_{n,\lambda}(\ell)\,. \tag{S64}$$

Now, the function $f_{n,\lambda}(\ell)$ is a second-degree polynomial in the variable ℓ. By finding its roots we can determine its sign on the whole real line. We see that

$$f_{n,\lambda}(\ell) \leq 0 \quad \text{if } \ell_-(n,\lambda) \leq \ell \leq \ell_+(n,\lambda),$$
$$f_{n,\lambda}(\ell) \geq 0 \quad \text{otherwise,}$$

where

$$\ell_\pm(n,\lambda) := n + \frac{1}{2} - (n+1)(1-\lambda) \pm \sqrt{\frac{1}{4} + (n+1)\lambda(1-\lambda)} \,.$$

One can show that

$$\ell_-(n,\lambda) \geq n - 2 \quad \forall \quad 0 \leq \lambda \leq \min\left\{\frac{5}{2(n+1)}, \lambda_+(n)\right\} = \lambda_+(n) \,.$$

Moreover,

$$\ell_-(n,\lambda) \leq n - 1 \quad \forall \quad \lambda \geq \lambda_-(n) \,.$$

Putting all together, we find that

$$n - 2 \leq \ell_-(n,\lambda) \leq n - 1 \quad \forall \quad \lambda_-(n) \leq \lambda \leq \lambda_+(n) \,.$$

It is also easy to verify that

$$\ell_+(n,\lambda) \geq n \quad \forall \quad 0 \leq \lambda \leq \frac{2}{n+2} \,.$$

Since $\frac{2}{n+2} \geq \lambda_+(n)$ for all $n \geq 2$, we deduce that

$$\ell_+(n,\lambda) \geq n \quad \forall \quad 0 \leq \lambda \leq \lambda_+(n) \,.$$

Going back to the function $f_{n,\lambda}(\ell)$, the above discussion implies that

$$\begin{aligned} f_{n,\lambda}(\ell) &\leq 0 \quad \text{if } \ell = n-1, n, \\ f_{n,\lambda}(\ell) &\geq 0 \quad \text{if } \ell = 0, \ldots, n-2, \end{aligned} \quad \forall \quad \lambda_-(n) \leq \lambda \leq \lambda_+(n) \,. \tag{S65}$$

Also, it is not difficult to verify that

$$n - 2 \leq (n+1)(1-\lambda) - 1 \leq n - 1 \quad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+1} \,;$$

we infer that

$$\begin{aligned} (n+1)(1-\lambda) - \ell - 1 &\leq 0 \quad \text{if } \ell = n-1, n, \\ (n+1)(1-\lambda) - \ell - 1 &\geq 0 \quad \text{if } \ell = 0, \ldots, n-2, \end{aligned} \quad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+1} \,. \tag{S66}$$

Using the fact that $\frac{1}{n+1} \geq \lambda_-(n)$ and $\lambda_+(n) \leq \frac{2}{n+1}$ for all n, and combining (S64) on the one hand with (S65)–(S66) on the other, we finally see that

$$q_{\ell+1}(n,\lambda) \geq q_\ell(n,\lambda) \quad \forall \quad \frac{1}{n+1} \leq \lambda \leq \lambda_+(n) \,,$$

which proves the claim. $\qquad \square$

C. Sorting $p(n,\lambda)$

As it turns out, for an analogous range of values of λ the probability distribution $p(n,\lambda)$, unlike $q(n,\lambda)$, is *not* automatically sorted in ascending order. The next lemma represents a first step in the direction of ascertaining how $p(n,\lambda)$ can be sorted.

Lemma S15. *For all $n \geq 2$,*

$$p_0(n, \lambda) \leq p_1(n, \lambda) \leq \ldots \leq p_{n-1}(n, \lambda) \geq p_n(n, \lambda) \leq p_{n+1}(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \lambda_+(n). \tag{S67}$$

Proof. For all $\ell = 0, \ldots, n$, employing (S48) and (S62) we compute

$$\lambda(\ell+1)\left((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2\right)\left(\frac{p_{\ell+1}(n, \lambda)}{p_\ell(n, \lambda)} - 1\right)$$

$$= (n-\ell+1)(1-\lambda)\left((1-\lambda)(n-\ell) + ((n+1)(1-\lambda) - \ell - 1)^2\right)$$

$$\quad - \lambda(\ell+1)\left((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2\right)$$

$$= (n-\ell+1)(1-\lambda)\left((1-\lambda)(n-\ell) + ((n+1)(1-\lambda) - \ell - 1)^2 - \lambda(\ell+1)\right) - \lambda(\ell+1)\left((n+1)(1-\lambda) - \ell\right)^2$$

$$= (n-\ell+1)(1-\lambda)\left((n+1)(1-\lambda) - \ell\right)\left((n+1)(1-\lambda) - \ell - 1\right) - \lambda(\ell+1)\left((n+1)(1-\lambda) - \ell\right)^2$$

$$= \left((n+1)(1-\lambda) - \ell\right)\left((n-\ell+1)(1-\lambda)\left((n+1)(1-\lambda) - \ell - 1\right) - \lambda(\ell+1)\left((n+1)(1-\lambda) - \ell\right)\right)$$

$$= \left((n+1)(1-\lambda) - \ell\right)\left(\ell^2 - 2\left((n+1)(1-\lambda) - \frac{1}{2}\right)\ell + (n+1)(1-\lambda)\left(n - (n+2)\lambda\right)\right).$$

Thus,

$$\lambda(\ell+1)\left((1-\lambda)(n-\ell+1) + ((n+1)(1-\lambda) - \ell)^2\right)\left(\frac{p_{\ell+1}(n, \lambda)}{p_\ell(n, \lambda)} - 1\right) = \left((n+1)(1-\lambda) - \ell\right) f_{n,\lambda}(\ell), \tag{S68}$$

where $f_{n,\lambda}(\ell)$, defined by (S63), is – luckily enough – the same function that we already encountered in the proof of Proposition S14, which makes (S65) available. Since

$$n - 1 \leq (n+1)(1-\lambda) \leq n \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+1},$$

we obtain that

$$\begin{array}{ll} (n+1)(1-\lambda) - \ell \leq 0 & \text{if } \ell = n, \\ (n+1)(1-\lambda) - \ell - 1 \geq 0 & \text{if } \ell = 0, \ldots, n-1, \end{array} \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+1}. \tag{S69}$$

Combining (S68) with (S65) and (S69) shows that for all $\frac{1}{n+1} \leq \lambda \leq \lambda_+(n)$ the inequalities $p_\ell(n, \lambda) \leq p_{\ell+1}(n, \lambda)$ hold true for $\ell = 0, \ldots, n-2$ or $\ell = n$, while for $\ell = n-1$ we have the opposite relation $p_{n-1}(n, \lambda) \geq p_n(n, \lambda)$. This completes the proof. $\square$

Lemma S16. *For all $n \geq 3$,*

$$p_{n-3}(n, \lambda) \leq p_n(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \widetilde{\lambda}_+(n), \tag{S70}$$

where

$$\widetilde{\lambda}_+(n) := \frac{3^{1/3}}{2^{1/3}n + 3^{1/3} - 2^{1/3}}. \tag{S71}$$

Proof. Using the explicit formulae (S48) and (S49), we compute

$$\frac{p_{n-3}(n, \lambda)}{p_n(n, \lambda)} = \frac{1}{24} n(n-1)(n-2)\left(\frac{\lambda}{1-\lambda}\right)^3 \frac{4(1-\lambda) + (4 - (n+1)\lambda)^2}{1-\lambda + (1 - (n+1)\lambda)^2}$$

$$= \left(\frac{2}{3} n(n-1)(n-2)\left(\frac{\lambda}{1-\lambda}\right)^3\right)\left(\frac{1}{16} \frac{4(1-\lambda) + (4 - (n+1)\lambda)^2}{1-\lambda + (1 - (n+1)\lambda)^2}\right)$$

We now evaluate separately the above two factors, and show that they are both upper bounded by 1. The first one can be estimated by resorting to the elementary inequality $n(n-2) \le (n-1)^2$; one obtains that

$$\frac{2}{3} \, n(n-1)(n-2) \left(\frac{\lambda}{1-\lambda} \right)^3 \le \frac{2}{3} \, (n-1)^3 \left(\frac{\lambda}{1-\lambda} \right)^3 \le 1 \qquad \forall \quad 0 \le \lambda \le \widetilde{\lambda}_+(n) \,,$$

where the last inequality can be easily proved by taking the cubic root of both sides. Upon simple algebraic manipulations, the inequality

$$\frac{1}{16} \, \frac{4(1-\lambda) + (4 - (n+1)\lambda)^2}{1 - \lambda + (1 - (n+1)\lambda)^2} \le 1 \,,$$

which is to be proved, becomes

$$\frac{5}{4}(1+n)^2 \lambda^2 - (3+2n)\lambda - 1 \ge 0 \,.$$

The discriminant of the second-degree polynomial on the left-hand side is $-n^2 + 2n + 4$. This is negative for all $n \ge 4$, and hence in this case the above inequality is satisfied for all $0 \le \lambda \le 1$ and a fortiori in the prescribed range. If $n = 3$, an explicit calculation shows that the inequality holds true for $\lambda \le \frac{1}{5}$ or $\lambda \ge \frac{1}{4}$, i.e. in particular for all $\lambda \ge \frac{1}{n+1} = \frac{1}{4}$. This completes the proof. □

Lemma S17. *For all $n \ge 2$,*

$$p_{n-1}(n,\lambda) = \max_{\ell=0,\ldots,n+1} p_\ell(n,\lambda) \qquad \forall \quad \frac{1}{n+1} \le \lambda \le \lambda_+(n) \,.$$

Proof. Since Lemma S15 holds in the prescribed interval in λ, we need only to prove that $p_{n-1}(n,\lambda) \ge p_{n+1}(n,\lambda)$. Indeed, one verifies that

$$(n+1)^2 (1-\lambda)^2 \left(\frac{p_{n-1}(n,\lambda)}{p_{n+1}(n,\lambda)} - 1 \right) = \frac{1}{2}(n-1)(n+1)(n+2)\lambda^2 - \left(2n^2 + n - 2\right)\lambda + 2n - 1 \ge 0 \,,$$

where the last inequality holds because the above second-degree polynomial in λ has discriminant $n(-2n^2 + n + 2) < 0$ as soon as $n \ge 2$. □

Proposition S18. *For all $n \ge 2$ and all $\frac{1}{n+1} \le \lambda \le \frac{1}{n}$,*

$$\text{either} \qquad p^\uparrow(n,\lambda) = \Big(p_0(n,\lambda), \, p_1(n,\lambda), \ldots, \, p_{n-3}(n,\lambda), \, p_n(n,\lambda), \, p_{n+1}(n,\lambda), \, p_{n-2}(n,\lambda), \, p_{n-1}(n,\lambda) \Big) \,,$$

$$\text{or} \qquad p^\uparrow(n,\lambda) = \Big(p_0(n,\lambda), \, p_1(n,\lambda), \ldots, \, p_{n-3}(n,\lambda), \, p_n(n,\lambda), \, p_{n-2}(n,\lambda), \, p_{n+1}(n,\lambda), \, p_{n-1}(n,\lambda) \Big) \,, \qquad \text{(S72)}$$

$$\text{or} \qquad p^\uparrow(n,\lambda) = \Big(p_0(n,\lambda), \, p_1(n,\lambda), \ldots, \, p_{n-3}(n,\lambda), \, p_{n-2}(n,\lambda), \, p_n(n,\lambda), \, p_{n+1}(n,\lambda), \, p_{n-1}(n,\lambda) \Big) \,.$$

When $n = 2$, it is understood that only the last 4 entries are to be taken into account.

Proof. It suffices to combine Lemmata S15, S16 and S17. Note that $\frac{1}{n} \le \min\left\{ \lambda_+(n), \widetilde{\lambda}_+(n) \right\}$ for all $n \ge 2$. □

D. Majorisation

Let $r = (r_0, \ldots, r_N)$ and $s = (s_0, \ldots, s_N)$ be two probability distributions. We remind the reader that r is said to be *majorised* by s, and we write $r \prec s$, if

$$\sum_{\ell=0}^{k} r_\ell^\uparrow \ge \sum_{\ell=0}^{k} s_\ell^\uparrow \qquad \forall \, k = 0, \ldots, N \,. \qquad \text{(S73)}$$

Of course, the above inequality becomes an equality for $k = N$, since the elements of both distributions add up to 1. For a complete introduction to the theory of majorisation, we refer the reader to the excellent monograph by Marshall and Olkin [70].

The goal of this subsection is to show that the two probability distributions $p(n,\lambda)$ and $q(n,\lambda)$ obey precisely a majorisation relation $p(n,\lambda) \prec q(n,\lambda)$. Our first step in this direction is a simple lemma.

Lemma S19. *For all $n \geq 2$,*

$$q_\ell(n, \lambda) \leq p_\ell(n, \lambda) \qquad \forall \quad \ell = 0, \ldots, n-1, \qquad \forall \quad 0 \leq \lambda \leq \frac{2}{n+1} \,. \qquad \text{(S74)}$$

Proof. Using the expressions (S48) and (S49), one verifies that

$$\left(\lambda \ell + ((n+1)(1-\lambda) - \ell)^2 \right) \left(\frac{p_\ell(n, \lambda)}{q_\ell(n, \lambda)} - 1 \right)$$
$$= (1-\lambda)(n - \ell + 1) + ((n+1)(1-\lambda) - \ell)^2 - \lambda \ell - ((n+1)(1-\lambda) - \ell)^2$$
$$= n - \ell + 1 - (n+1)\lambda$$
$$\geq 0 \,,$$

where the last inequality holds provided that $\lambda \leq \frac{2}{n+1}$ and $\ell \leq n-1$. $\qquad \square$

Lemma S20. *For all $n \geq 3$,*

$$q_{n+1}(n, \lambda) - p_{n-1}(n, \lambda) \geq \frac{(n+1)(n-2)}{4n(n-1)} \left(1 - \frac{1}{n} \right)^n \qquad \forall \quad 0 \leq \lambda \leq \frac{1}{n} \,. \qquad \text{(S75)}$$

When $n = 2$, we have instead that

$$q_3(2, \lambda) - p_1(2, \lambda) \geq \frac{\epsilon}{4} \qquad \forall \quad 0 \leq \lambda \leq \frac{1}{2} - \epsilon \qquad \text{(S76)}$$

for any fixed $\epsilon > 0$.

Proof. For all $n \geq 2$, one verifies that

$$\frac{\partial}{\partial \lambda} \left(q_{n+1}(n, \lambda) - p_{n-1}(n, \lambda) \right)$$
$$= \frac{1}{4}(1-\lambda)^{n-3} \left(2 - 6n + 2\left(6n^2 + n - 3 \right)\lambda - (n+1)^2(7n-6)\lambda^2 + (n+1)^2 \left(n^2 + n - 2 \right)\lambda^3 \right)$$
$$=: \frac{1}{4}(1-\lambda)^{n-3} g_n(\lambda) \,.$$

Now, since

$$\frac{d^2 g_n(\lambda)}{d\lambda} = 2(n+1)^2 \left(6 - 7n + 3\left(n^2 + n - 2 \right)\lambda \right)$$
$$\leq 2(n+1)^2 \left(6 - 7n + 3\left(n^2 + n - 2 \right)\frac{1}{n} \right)$$
$$= -\frac{2}{n}(n+1)^2 \left(4n^2 - 9n + 6 \right) \leq 0$$

for $0 \leq \lambda \leq \frac{1}{n}$, the first derivative $\frac{dg_n(\lambda)}{d\lambda}$ of $g_n(\lambda)$ is a decreasing function of λ in the same interval $\left[0, \frac{1}{n} \right]$. Hence,

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{dg_n(\lambda)}{d\lambda} = \frac{dg_n(\lambda)}{d\lambda} \bigg|_{\lambda = \frac{1}{n}} = \frac{n+1}{n^2} \left(6 + n(n-1)(n-3) \right) \geq 0 \,.$$

In turn, this implies that

$$\max_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} g_n(\lambda) = g_n \left(\frac{1}{n} \right) = -\frac{(n-1)^2(n+2)}{n^3} \leq 0 \,.$$

Thus, $q_{n+1}(n, \lambda) - p_{n-1}(n, \lambda)$ is decreasing in λ. Finally, we conclude from this that

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \left\{ q_{n+1}(n, \lambda) - p_{n-1}(n, \lambda) \right\} = q_{n+1} \left(n, \frac{1}{n} \right) - p_{n-1} \left(n, \frac{1}{n} \right) = \frac{(n+1)(n-2)}{4n(n-1)} \left(1 - \frac{1}{n} \right)^n \,.$$

When $n = 2$, we have instead that

$$\min_{\frac{1}{3} \leq \lambda \leq \frac{1}{2} - \epsilon} \left\{ q_3(2, \lambda) - p_1(2, \lambda) \right\} = q_3 \left(2, \frac{1}{2} - \epsilon \right) - p_1 \left(2, \frac{1}{2} - \epsilon \right) = \frac{\epsilon}{4} + 3\epsilon^3 \geq \frac{\epsilon}{4} \,.$$

This concludes the proof. $\qquad \square$

Proposition S21. *For all $n \geq 2$,*

$$p(n, \lambda) \prec q(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n}. \tag{S77}$$

Proof. According to (S73), we need to verify that

$$\sum_{\ell=0}^{k} p_\ell^\uparrow(n, \lambda) \geq \sum_{\ell=0}^{k} q_\ell^\uparrow(n, \lambda) \qquad \forall \quad k = 0, \ldots, n, \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n}, \tag{S78}$$

where we used the fact that the inequality corresponding to $k = n + 1$ is in fact an equality, by normalisation. Using Proposition S18 and Lemma S19, and observing that $\frac{1}{n} \leq \frac{2}{n+1}$ for all $n \geq 1$, we see that (when $n \geq 3$) the cases $k = 0, \ldots, n - 3$ of (S78) are automatically satisfied. Exploiting again normalisation, we recast the difference of the two sides of (S78) (for arbitrary k) as

$$\sum_{\ell=0}^{k} p_\ell^\uparrow(n, \lambda) - \sum_{\ell=0}^{k} q_\ell^\uparrow(n, \lambda) = \sum_{\ell=k+1}^{n+1} q_\ell^\uparrow(n, \lambda) - \sum_{\ell=k+1}^{n+1} p_\ell^\uparrow(n, \lambda) = \sum_{\ell=k+1}^{n+1} q_\ell(n, \lambda) - \sum_{\ell=k+1}^{n+1} p_\ell^\uparrow(n, \lambda), \tag{S79}$$

where the last identity follows from Proposition S14, once one observes that $\frac{1}{n} \leq \lambda_+(n)$ as long as $n \geq 2$. It remains to check the positivity of (S79) for $k = n, n-1, n-2$ and for $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. When $k = n - 1$ and $k = 2$ we have to reckon the (two) distinct possibilities offered by Proposition S18. This makes a total of 5 different cases to vet. We break down the proof into the separate analysis of each of these cases.

- $k = n$. Thanks to Lemma S17 (or Proposition S18) and Lemma S20,

$$q_{n+1}(n, \lambda) - p_{n+1}^\uparrow(n, \lambda) = q_{n+1}(n, \lambda) - p_{n-1}(n, \lambda) \geq 0.$$

- $k = n - 1$ and $p_n^\uparrow(n, \lambda) = p_{n-2}(n, \lambda)$. Let us write

$$q_{n+1}(n, \lambda) + q_n(n, \lambda) - p_{n-1}(n, \lambda) - p_{n-2}(n, \lambda)$$

$$= \frac{1}{2}(1 - \lambda)^{n-3} \left(2 - 3(n+2)\lambda + \left(6 + 9n + n^2\right)\lambda^2 + \frac{1}{2}(n+1)\left(n^2 - 10n - 4\right)\lambda^3 - \frac{1}{6}n(n+1)(n+2)(n-5)\lambda^4 \right)$$

$$=: \frac{1}{2}(1 - \lambda)^{n-3} h_n(\lambda).$$

Now, since

$$\frac{d^3 h_n(\lambda)}{d\lambda^3} = (n+1)\left(3\left(n^2 - 10n - 4\right) - 4n(n+2)(n-5)\lambda\right)$$

is a linear function of λ, we have that

$$\max_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{d^3 h_n(\lambda)}{d\lambda^3} = \max \left\{ \frac{d^3 h_n(\lambda)}{d\lambda^3}\bigg|_{\lambda=\frac{1}{n+1}}, \frac{d^3 h_n(\lambda)}{d\lambda^3}\bigg|_{\lambda=\frac{1}{n}} \right\}$$

$$= \max \left\{ -n^3 - 15n^2 - 2n - 12, \; -(n+1)\left(n^2 + 18n - 28\right) \right\}$$

$$\leq 0.$$

That is to say, the function $\frac{d^2 h_n(\lambda)}{d\lambda^2}$ is non-increasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. Therefore,

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{d^2 h_n(\lambda)}{d\lambda^2} = \frac{d^2 h_n(\lambda)}{d\lambda^2}\bigg|_{\lambda=\frac{1}{n}} = 3n^2 - 5n - 4 + \frac{8}{n} \geq n - 4 + \frac{8}{n} \geq 4\left(\sqrt{2} - 1\right) \geq 0,$$

where we exploited the fact that $n \geq 2$. This shows that the function $\frac{dh_n(\lambda)}{d\lambda}$ is non-decreasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. We infer that

$$\max_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{dh_n(\lambda)}{d\lambda} = \frac{dh_n(\lambda)}{d\lambda}\bigg|_{\lambda=\frac{1}{n+1}} = -\frac{n+1}{6n^2}\left(n^2 + 2n + 4\right) \leq 0.$$

Finally, given that $h_n(\lambda)$ has been shown to be non-increasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, we have that

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} h_n(\lambda) = h_n\left(\frac{1}{n}\right) = \frac{1}{6n^3}(n-1)(n-2)(2n-1) \geq 0,$$

which shows that $q_{n+1}(n, \lambda) + q_n(n, \lambda) - p_{n-1}(n, \lambda) - p_{n-2}(n, \lambda)$ for all $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$ and concludes the analysis of this case.

- $k = n - 1$ and $p_n^\uparrow(n, \lambda) = p_{n+1}(n, \lambda)$. We compute

$$q_{n+1}(n, \lambda) + q_n(n, \lambda) - p_{n-1}(n, \lambda) - p_{n+1}(n, \lambda)$$
$$= (1-\lambda)^{n-2}\left(1 - \frac{1}{2}(4n+5)\lambda + \frac{1}{2}(3n^2+6n+4)\lambda^2 - \frac{1}{4}(n+1)^2(n+2)\lambda^3\right)$$
$$=: (1-\lambda)^{n-2} s_n(\lambda).$$

Let us first deal with the case $n = 2$; note that $s_2(\lambda) = 1 - \frac{13}{2}\lambda + 14\lambda^2 - 9\lambda^3$. Now, $\frac{ds_2(\lambda)}{d\lambda} = -\frac{13}{2} + 28\lambda - 27\lambda^2 \geq 0$ for $0.351 \approx \frac{28-\sqrt{82}}{54} \leq \lambda \leq \frac{28+\sqrt{82}}{54} \approx 0.686$, and $\frac{ds_2(\lambda)}{d\lambda} \leq 0$ outside of that interval. Hence,

$$\min_{\frac{1}{3} \leq \lambda \leq \frac{1}{2}} s_2(\lambda) = s_2\left(\frac{28-\sqrt{82}}{54}\right) \approx 0.054 \geq 0.$$

We now consider the case where $n \geq 3$. Since

$$\frac{d^2 s_n(\lambda)}{d\lambda^2} = 3n^2 + 6n + 4 - \frac{3}{2}(n+1)^2(n+2)\lambda$$

is decreasing in λ, we obtain that

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{d^2 s_n(\lambda)}{d\lambda^2} = \frac{d^2 s_n(\lambda)}{d\lambda^2}\bigg|_{\lambda=1/n} = \frac{3}{2}n^2 - \frac{3}{n} - \frac{7}{2} \geq \frac{27}{2} - 1 - \frac{7}{2} = 9 \geq 0,$$

where we used the fact that $n \geq 3$. This proves that $\frac{ds_n(\lambda)}{d\lambda}$ is non-decreasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. Hence,

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{ds_n(\lambda)}{d\lambda} = \frac{ds_n(\lambda)}{d\lambda}\bigg|_{\lambda=\frac{1}{n+1}} = \frac{n(n-3)}{4(n+1)} \geq 0,$$

where the last estimate holds because $n \geq 3$. We have just shown that $s_n(\lambda)$ is non-decreasing in the interval $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. We infer that

$$\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} s_n(\lambda) = s_n\left(\frac{1}{n+1}\right) = \frac{n(n-1)}{4(n+1)^2} \geq 0,$$

which shows that $q_{n+1}(n, \lambda) + q_n(n, \lambda) - p_{n-1}(n, \lambda) - p_{n+1}(n, \lambda) \geq 0$ for all $n \geq 2$ and all $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$, thus completing the argument for this case.

- $k = n - 2$ and $\left\{p_{n-1}^\uparrow(n, \lambda), p_n^\uparrow(n, \lambda)\right\} = \{p_{n+1}(n, \lambda), p_{n-2}(n, \lambda)\}$. The relevant quantity is now

$$q_{n+1}(n, \lambda) + q_n(n, \lambda) + q_{n-1}(n, \lambda) - p_{n-1}(n, \lambda) - p_{n+1}(n, \lambda) - p_{n-2}(n, \lambda)$$
$$= (1-\lambda)^{n-3}\left(1 - \left(\frac{7}{2} + n\right)\lambda - \frac{1}{4}(n^2 - 15n - 18)\lambda^2 + \frac{1}{2}(n^3 - 2n^2 - 7n - 5)\lambda^3 - \frac{1}{12}(n+1)^2(n+2)(n-3)\lambda^4\right)$$
$$=: (1-\lambda)^{n-3} t_n(\lambda).$$

To study the polynomial $t_n(\lambda)$, let us treat separately the cases $n = 2$ and $n \geq 3$. Note that

$$t_2(\lambda) = (1-\lambda)^2\left(1 - \frac{1}{2}\lambda(7 - 6\lambda)\right) \geq 0 \qquad \forall \quad \frac{1}{3} \leq \lambda \leq \frac{1}{2},$$

where the last inequality is a consequence of the fact that the function $\lambda \mapsto 1 - \frac{1}{2}\lambda(7 - 6\lambda)$ is decreasing on $\left(-\infty, \frac{7}{12}\right] \supset \left[\frac{1}{3}, \frac{1}{2}\right]$ and vanishes for $\lambda = \frac{1}{2}$.

We now look at the case where $n \geq 3$. Since $\frac{d^4 t_n(\lambda)}{d\lambda^4} = -2(n+1)^2(n+2)(n-3) \leq 0$, the function $\frac{d^2 t_n(\lambda)}{d\lambda^2}$ is concave. Hence,

$$
\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{d^2 t_n(\lambda)}{d\lambda^2} = \min\left\{ \frac{d^2 t_n(\lambda)}{d\lambda^2}\Big|_{\lambda = \frac{1}{n+1}}, \ \frac{d^2 t_n(\lambda)}{d\lambda^2}\Big|_{\lambda = \frac{1}{n}} \right\}
$$
$$
= \min\left\{ \frac{n}{2(n+1)}\left(3n^2 + 2n + 5\right), \ \frac{1}{2n^2}\left(3n^4 + n^3 - 10n^2 - 4n + 12\right) \right\}
$$
$$
\geq 0,
$$

where in the last step we used the fact that $n \geq 3$. We deduce that $\frac{dt_n(\lambda)}{d\lambda}$ is non-decreasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, in turn implying that

$$
\max_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} \frac{dt_n(\lambda)}{d\lambda} = \frac{dt_n(\lambda)}{d\lambda}\Big|_{\lambda = 1/n} = -\frac{(n-1)}{n^4}\left(2n^3 - 2n^2 - 7n + 12\right) \leq 0,
$$

where the last inequality holds because $2n^3 - 2n^2 - 7n + 12 \geq 4n^2 - 7n + 12 \geq 9 \geq 0$ for $n \geq 3$. Since we have just shown that $t_n(\lambda)$ is non-increasing on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, we conclude that

$$
\min_{\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}} t_n(\lambda) = t_n\left(\frac{1}{n}\right) = \frac{(n-2)(n-1)(2n^2 - 2n + 3)}{12n^4} \geq 0,
$$

concluding the argument.

- $k = n - 2$ and $\left(p_{n-1}^{\uparrow}(n, \lambda), \ p_n^{\uparrow}(n, \lambda)\right) = (p_n(n, \lambda), \ p_{n+1}(n, \lambda))$. The analysis of this last case is much simpler. It suffices to verify that

$$
q_{n+1}(n, \lambda) + q_n(n, \lambda) + q_{n-1}(n, \lambda) - p_{n-1}(n, \lambda) - p_{n+1}(n, \lambda) - p_n(n, \lambda) = \frac{1}{4}n(n-1)\lambda^2(1-\lambda)^{n-2} \geq 0.
$$

This completes the proof.

$\square$

E. Concluding the proof

Proof of Theorem 2. Let us partition the $(0, 1]$ into the three regions

$$
(0, 1] = \left(0, \frac{1}{2} - \epsilon\right] \cup \left[\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon\right] \cup \left[\frac{1}{2} + \epsilon, 1\right], \tag{S80}
$$

where $\epsilon > 0$ is a small constant to be determined later. In the third region, i.e. for $\frac{1}{2} + \epsilon \leq \lambda \leq 1$, the claim follows elementarily from the ansatz $\sigma = |0\rangle\langle 0|$, which brings us back to the case of the pure loss channel. Thanks to (S26), we know that

$$
Q\left(\Phi_{\lambda, |0\rangle\langle 0|}, \frac{1}{2}\right) = Q\left(\mathscr{E}_\lambda, \frac{1}{2}\right) = \max\left\{ g\left(\frac{\lambda}{2}\right) - g\left(\frac{1-\lambda}{2}\right), \ 0 \right\} \geq g\left(\frac{1}{2}\left(\frac{1}{2} + \epsilon\right)\right) - g\left(\frac{1}{2}\left(\frac{1}{2} - \epsilon\right)\right) > 0
$$

as long as $\epsilon > 0$. In the second region, that is, for $\frac{1}{2} - \epsilon \leq \lambda \leq \frac{1}{2} + \epsilon$, one can use Example S10 and some standard continuity arguments. Namely, consider the state $\sigma = \xi(1/3)$ as defined by (S36); specialising (S42) we find that

$$
Q\left(\Phi_{1/2, \xi(1/3)}, \frac{1}{2}\right) \geq I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(1/2, 1/3)} \approx 0.07392 > 0,
$$

where $\zeta_{AB}(\lambda, \eta)$ is the reduced state on AB corresponding to (S39). The density matrices $\zeta_{AB}(\lambda, 1/3)$ clearly depend continuously on λ; moreover, they live in a qubit–qutrit system for all values of λ. Hence, the Alicki–Fannes–Winter

inequality [68, 69] implies that $I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(\lambda,\,1/3)}$ is a continuous function of λ. By choosing $\epsilon > 0$ small enough, we can therefore insure that

$$Q\left(\Phi_{\lambda,\,\xi(1/3)},\,\frac{1}{2}\right) \geq I_{\mathrm{coh}}(A\rangle B)_{\zeta_{AB}(\lambda,\,1/3)} \geq c_1 \qquad \forall \quad \frac{1}{2} - \epsilon \leq \lambda \leq \frac{1}{2} + \epsilon, \tag{S81}$$

where $c_1 > 0$ is a universal constant.

We are thus left with the first region, corresponding to $0 < \lambda \leq \frac{1}{2} - \epsilon$. We further split it according to

$$\left(0,\,\frac{1}{2} - \epsilon\right] = \left[\frac{1}{3},\,\frac{1}{2} - \epsilon\right] \cup \bigcup_{n=3}^{\infty} \left[\frac{1}{n+1},\,\frac{1}{n}\right].$$

Thanks to Proposition S13, we need only to show that

$$\mathscr{I}(n,\lambda) = H(p(n,\lambda)) - H(q(n,\lambda)) \geq c \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n} \tag{S82}$$

for all $n \geq 3$ and for some universal constant $c_2 > 0$, and also that

$$\mathscr{I}(2,\lambda) = H(p(2,\lambda)) - H(q(2,\lambda)) \geq c_3 \qquad \forall \quad \frac{1}{3} \leq \lambda \leq \frac{1}{2} - \epsilon \tag{S83}$$

for some other constant $c_3 > 0$.

Our main tool here will be a beautiful inequality proved by Ho and Verdú [56, Theorem 3]. This states that whenever r and s are two probability distributions such that $r \prec s$, it holds that

$$H(s) - H(r) \geq D\left(s^\uparrow \,\middle\|\, r^\uparrow\right), \tag{S84}$$

where $D(u\|v) := \sum_\ell u_\ell \log_2 \frac{u_\ell}{v_\ell}$ is the Kullback–Leibler divergence, i.e. the relative entropy. Let us first deal with the case $n \geq 3$. We obtain that

$$\mathscr{I}(n,\lambda) = H(p(n,\lambda)) - H(q(n,\lambda))$$
$$\overset{1}{\geq} D\left(q^\uparrow(n,\lambda)\,\middle\|\,p^\uparrow(n,\lambda)\right)$$
$$\overset{2}{\geq} \frac{1}{2\ln 2}\left\|q^\uparrow(n,\lambda) - p^\uparrow(n,\lambda)\right\|_1^2$$
$$\overset{3}{\geq} \frac{2}{\ln 2}\left|q^\uparrow_{n+1}(n,\lambda) - p^\uparrow_{n+1}(n,\lambda)\right|^2$$
$$\overset{4}{=} \frac{2}{\ln 2}\left|p_{n-1}(n,\lambda) - q_{n+1}(n,\lambda)\right|^2$$
$$\overset{5}{\geq} \frac{2}{\ln 2}\left(\frac{(n+1)(n-2)}{4n(n-1)}\right)^2\left(1 - \frac{1}{n}\right)^{2n}$$
$$\overset{6}{\geq} \frac{32}{6561\ln 2} > 0.$$

Here, 1 comes from applying the Ho–Verdú inequality (S84) to the case of $r = p(n,\lambda)$ and $s = q(n,\lambda)$, which is possible by Proposition S21. The estimate in 2 is just Pinsker's inequality (see [71, p.58] and references therein). In 3 we estimated the total variation or L_1 distance between $p^\uparrow(n,\lambda)$ and $q^\uparrow(n,\lambda)$ from below as twice their L_∞ distance, namely

$$\left\|q^\uparrow(n,\lambda) - p^\uparrow(n)\right\|_1 \geq 2\max_{\ell=0,\dots,n+1}\left|q^\uparrow_\ell(n,\lambda) - p^\uparrow_\ell(n)\right| \geq \left|q^\uparrow_{n+1}(n,\lambda) - p^\uparrow_{n+1}(n)\right|.$$

Then, in 4 we used Proposition S14 and Lemma S17, together with the observation that $\lambda_+(n) \geq \frac{1}{n}$ for all $n \geq 2$. The estimate in 5 follows from (S75), while in 6 we noted that both $n \mapsto \frac{(n+1)(n-2)}{4n(n-1)}$ and $n \mapsto \left(1 - \frac{1}{n}\right)^n$ are increasing function of n for $n \geq 3$, and therefore their product can be lower bounded by evaluating it for $n = 3$. Thus, (S82) holds with $c_3 = \frac{32}{6561\ln 2}$.

It remains to deal with the $n = 2$ case. We can repeat the same reasoning as above all the way until step 5, where we have to use instead the estimate in (S76), thus obtaining

$$\mathscr{I}(2,\lambda) \geq \frac{\epsilon^2}{8\ln 2} =: c_2 > 0,$$

which proves (S83). Setting $c := \min\{c_1, c_2, c_3\}$ completes the argument. $\qquad\square$

Remark S22. The optimal constant in Theorem 2 can be expressed as a function of the energy constraint N as

$$c(N) := \inf_{0 < \lambda \leq 1} \sup_{\sigma} Q\left(\Phi_{\lambda, \sigma}, N\right), \tag{S85}$$

where the supremum is over all single-mode states σ. Using the explicit form of the Alicki–Fannes–Winter inequality [68, 69] could yield the explicit, rigorous,[83] yet very small lower bound $c(1/2) \geq 5.133 \times 10^{-6}$. Numerical investigations suggest that this is very far away from the truth, and that one could take at least $c(1/2) \gtrsim 0.066$, which is four orders of magnitude larger than the former estimate. This must be confronted with the 'trivial' upper bound descending from Lemma S2, which reads $c(1/2) \leq g(1/2) \approx 1.377$.

Remark S23. It is perhaps more interesting to look at the slightly different quantities

$$c_0(N) := \lim_{\lambda \to 0^+} \sup_{\substack{\sigma, \\ 0 < \lambda' \leq \lambda}} Q\left(\Phi_{\lambda', \sigma}, N\right), \tag{S86}$$

which represent the best-case-scenario quantum communication rates when the transmissivity approaches 0 but the single-mode environment state σ is chosen optimally. Since

$$\lim_{n \to \infty} \frac{(n+1)(n-2)}{4n(n-1)} \left(1 - \frac{1}{n}\right)^n = \frac{1}{4e},$$

it can be seen that our argument yields the rigorous estimate

$$c_0(1/2) \geq \frac{1}{8e^2 \ln 2} \approx 0.0244.$$

Numerical investigations produce a substantially higher bound $c_0(1/2) \gtrsim 0.133$, which again must be confronted with the upper bound $c_0(1/2) \leq g(1/2) \approx 1.377$.

F. Further considerations

It turns out that one can get rid of the multiple options in Proposition S18 if one is willing to exclude the special cases $n = 2$ and $n = 3$. When this is done something more happens. Namely, the majorisation $p(n, \lambda) \prec q(n, \lambda)$ of Proposition S21 is of a very special type. It actually holds that $p_\ell^\uparrow(n, \lambda) \geq q_\ell^\uparrow(n, \lambda)$ for all $n \geq 4$ and $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. Throughout this section we prove these claims.

Lemma S24. *For all $n \geq 4$,*

$$p_{n+1}(n, \lambda) \leq p_{n-2}(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq 1. \tag{S87}$$

Proof. Employing the expressions (S48), we see that

$$\frac{p_{n-2}(n, \lambda)}{p_{n+1}(n, \lambda)} = \frac{n(n-1)}{6(n+1)} \frac{\lambda}{(1-\lambda)^2} \left(3 + \frac{1}{1-\lambda} \left(3 - (n+1)\lambda\right)^2\right) =: \frac{n(n-1)}{6(n+1)} g_n(\lambda).$$

It is not difficult to see that

$$\frac{dg_n(\lambda)}{d\lambda} = \frac{6}{(1-\lambda)^4} \left(\frac{1}{2}(n^2 - 2)\lambda^2 - (2n-1)\lambda + 2\right) \geq 0 \qquad \forall \quad 0 \leq \lambda \leq 1,$$

because the discriminant of the second-degree polynomial on the right-hand side equals $9 - 4n$ and is therefore negative as long as $n \geq 3$. Thus,

$$\min_{\frac{1}{n+1} \leq \lambda \leq 1} \left\{\frac{p_{n-2}(n, \lambda)}{p_{n+1}(n, \lambda)} - 1\right\} = \frac{n(n-1)}{6(n+1)} g_n\left(\frac{1}{n+1}\right) - 1 = \frac{1}{6n^2}(n+1)(n-4) \geq 0$$

for all $n \geq 4$. $\qquad \square$

Proposition S25. *For all $n \geq 4$,*

$$p^{\uparrow}(n, \lambda) = \left(p_0(n, \lambda), p_1(n, \lambda), \dots, p_{n-3}(n, \lambda), p_n(n, \lambda), p_{n+1}(n, \lambda), p_{n-2}(n, \lambda), p_{n-1}(n, \lambda)\right) \qquad \forall \; \frac{1}{n+1} \leq \lambda \leq \widetilde{\lambda}_+(n),$$
(S88)

where $\widetilde{\lambda}_+(n)$ is defined by (S71). In other words, for the stated range of values of λ the probability vector $p(n, \lambda)$ can be sorted in ascending order by exchanging the last two pairs of entries.

Proof. It suffices to combine Lemmata S15, S16 and S24. Note that $\widetilde{\lambda}_+(n) \leq \lambda_+(n)$ for all $n \geq 4$. This can be shown e.g. by noting that

$$\widetilde{\lambda}_+(n) \leq \frac{3 - \sqrt{3}}{n + 2} \leq \lambda_+(n) \qquad \forall \, n \geq 18,$$

where the first relation is equivalent to a linear inequality upon elementary algebraic manipulations, while the second is easily seen to hold for all $n \geq 1$ by direct inspection of (S59). In the remaining cases $n = 4, \dots, 17$, the fact that $\widetilde{\lambda}_+(n) \leq \lambda_+(n)$ can be checked numerically. $\qquad \square$

Now that the probability distribution $p(n, \lambda)$ has been sorted in ascending order by a fixed permutation, we proceed to check that indeed $p_{\ell}^{\uparrow}(n, \lambda) \geq q_{\ell}^{\uparrow}(n, \lambda)$ for all $n \geq 4$ and $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$.

Lemma S26. *For all $n \geq 2$,*

$$q_{n-2}(n, \lambda) \leq p_n(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n}.$$
(S89)

Proof. One verifies that

$$\frac{\partial}{\partial \lambda} \left(p_n(n, \lambda) - q_{n-2}(n, \lambda)\right)$$

$$= \frac{1}{12}(1 - \lambda)^{n-3} \left(2 - (n + 1)\lambda\right) \left(6 - 3(n + 6)\lambda - 3 \left(n^2 - 4n - 6\right) \lambda^2 + (n + 1)(n + 2)(n - 3)\lambda^3\right)$$

$$=: \frac{1}{12}(1 - \lambda)^{n-3} \left(2 - (n + 1)\lambda\right) h_n(\lambda).$$

We will now show that $h_n(\lambda) \geq 0$ for all $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$. To this end, compute

$$\frac{1}{6} \frac{d^2 h_n(\lambda)}{d\lambda} = -\left(n^2 - 4n - 6\right) + (n + 1)(n + 2)(n - 3)\lambda \geq 0 \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n},$$

where the last inequality holds because: (i) it can be verified explicitly for $n = 2$ and $n = 3$; (ii) for $n \geq 4$, one has that

$$\frac{n^2 - 4n - 6}{(n + 1)(n + 2)(n - 3)} \leq \frac{1}{n + 1} \qquad \forall \, n \geq 4,$$

with equality for $n = 4$. Since we have shown that $\frac{dh_n(\lambda)}{d\lambda}$ is increasing in λ on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, there it holds that

$$\frac{1}{3} \frac{dh_n(\lambda)}{d\lambda} \leq \frac{1}{3} \frac{dh_n(\lambda)}{d\lambda}\Big|_{\lambda=1/n}$$

$$= \left(-(n + 6) - 2\left(n^2 - 4n - 6\right)\lambda + (n + 1)(n + 2)(n - 3)\lambda^2\right)\Big|_{\lambda=1/n}$$

$$= -2(n - 1) + \frac{5}{n} - \frac{6}{n^2} \leq -(n - 1)\left(2 - \frac{5}{n^2}\right) \leq 0.$$

Thus, $h_n(\lambda)$ is decreasing in λ on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. From this we deduce that

$$h_n(\lambda) \geq h_n\left(\frac{1}{n}\right) = \frac{(n - 1)(n - 2)(n - 3)}{n^3} \geq 0$$

for all $n = 2, 3, 4, \dots$. $\qquad \square$

Lemma S27. *For all $n \geq 2$,*

$$q_{n-1}(n, \lambda) \leq p_{n+1}(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+2}. \tag{S90}$$

Proof. A simple calculation shows that

$$2(n+1)(1-\lambda)^2 \left(1 - \frac{q_{n-1}(n,\lambda)}{p_{n+1}(n,\lambda)}\right) = (n-1)\left((n+1)\lambda - 1\right)\left(2 - (n+2)\lambda\right) \geq 0 \qquad \frac{1}{n+1} \leq \lambda \leq \frac{2}{n+2},$$

completing the proof. $\qquad\square$

Lemma S28. *For all $n \geq 4$,*

$$q_n(n, \lambda) \leq p_{n-2}(n, \lambda) \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n}. \tag{S91}$$

Proof. One finds that

$$\frac{n(n-1)}{6} \lambda^2 \left(3(1-\lambda) + (3 - (n+1)\lambda)^2\right) \left(1 - \frac{q_n(n,\lambda)}{p_{n-2}(n,\lambda)}\right)$$

$$= -1 + (n+4)\lambda + \left(n^2 - 6n - 6\right)\lambda^2 - (n+1)\left(n^2 - \frac{5}{2}n - 4\right)\lambda^3 + \frac{1}{6}(n+1)^2(n+2)(n-3)\lambda^4$$

$$=: r_n(\lambda).$$

We look at the polynomial $r_n(\lambda)$ and its derivatives in the interval $\left[\frac{1}{n+1}, \frac{1}{n}\right]$. Since $\frac{d^4 r_n(\lambda)}{d\lambda} = 4(n+1)^2(n+2)(n-3) \geq 0$, the function $\frac{d^2 r_n(\lambda)}{d\lambda}$ is convex. Therefore, on the larger interval $\left[\frac{1}{n+1}, \frac{2}{n+1}\right] \supset \left[\frac{1}{n+1}, \frac{1}{n}\right]$ it holds that

$$\frac{d^2 r_n(\lambda)}{d\lambda} \leq \max\left\{\frac{d^2 r_n(\lambda)}{d\lambda}\Big|_{\lambda=1/(n+1)}, \frac{d^2 r_n(\lambda)}{d\lambda}\Big|_{\lambda=2/(n+1)}\right\}$$

$$= \max\left\{-n(2n-1), -2(n-2)(n-3)\right\}$$

$$\leq 0.$$

In turn, this tells us that $r_n(\lambda)$ is concave. Thus, on $\left[\frac{1}{n+1}, \frac{1}{n}\right]$ it holds that

$$r_n(\lambda) \leq \max\left\{r_n\left(\frac{1}{n+1}\right), r_n\left(\frac{1}{n}\right)\right\}$$

$$= \min\left\{\frac{n(n-4)}{6(n+1)^2}, \frac{(n-1)(n-2)(n^2+n-3)}{6n^4}\right\}$$

$$\geq 0.$$

This proves the claim. $\qquad\square$

We are finally ready to prove our last claim.

Proposition S29. *Let $n \geq 4$ be an integer. Then*

$$p_\ell^\uparrow(n, \lambda) \geq q_\ell^\uparrow(n, \lambda) \qquad \forall \quad \ell = 0, \ldots, n, \qquad \forall \quad \frac{1}{n+1} \leq \lambda \leq \frac{1}{n}, \tag{S92}$$

with the reverse inequality holding instead for $\ell = n+1$. In particular, $p(n, \lambda) \prec q(n, \lambda)$ for all $\frac{1}{n+1} \leq \lambda \leq \frac{1}{n}$.

Proof. Since $\widetilde{\lambda}_+(n) \geq \frac{1}{n}$ for all $n \geq 1$, Proposition S25 applies and tell us that the ordering of $p(n, \lambda)$ is as in (S88). Now, the cases $\ell = 0, \ldots, n-3$ of (S92) follow from Lemma S19, as usual. When $\ell = n-2$, we have instead to verify that $p_n(n, \lambda) \geq q_{n-2}(n, \lambda)$, which is a consequence of Lemma S26. For $\ell = n-1$, the claim amounts to the inequality $p_{n+1}(n, \lambda) \geq q_{n-1}(n, \lambda)$, which holds by Lemma S27, because $\frac{2}{n+2} \geq \frac{1}{n}$ whenever $n \geq 2$. The last case is $\ell = n$, for which we have to show that $p_{n-2}(n, \lambda) \geq q_n(n, \lambda)$; this is guaranteed to hold by Lemma S28. The reverse inequality holds for $\ell = n+1$ by normalisation:

$$p_{n+1}^\uparrow(n, \lambda) = 1 - \sum_{\ell=0}^{n} p^\uparrow(n, \lambda) \geq 1 - \sum_{\ell=0}^{n} q^\uparrow(n, \lambda) = q_{n+1}^\uparrow(n, \lambda).$$

Finally, majorisation follows by direct inspection. $\qquad\square$

V.   SOME EXTENSIONS OF THEOREM 2

Throughout this section, we discuss some possible extensions of Theorem 2. In particular, we look into the case where the transmission channel results from a concatenation of multiple beam splitters instead of a single one. The scenario we consider is depicted in Figure 3.
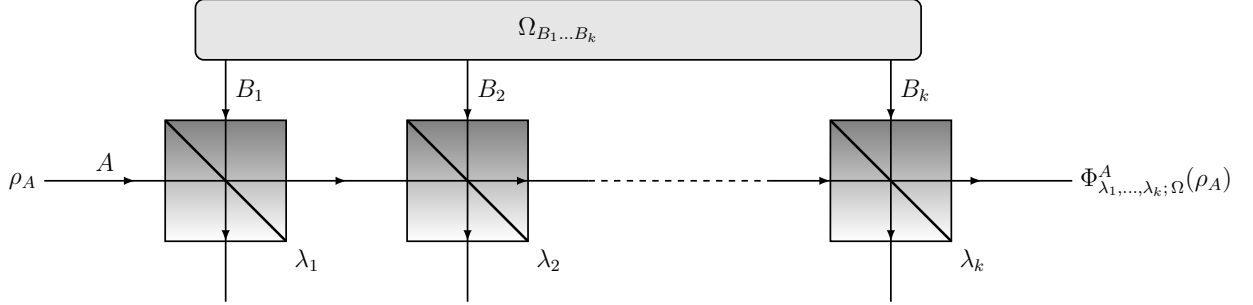

FIG. 3. The input state ρ_A is sent through a sequence of beam splitters with transmissivities $\lambda_1, \ldots, \lambda_k$. The vertical output arms coming down from each beam splitter are simply traced away. The action of the channel $\Phi_{\lambda_1, \ldots, \lambda_k; \Omega}$ is given by (S93).

As usual, we will fix the values of the transmissivities and optimise the environment state so as to maximise the quantum capacity of the resulting channel. There are at least three different scenarios we could look into:

(a) $\Omega_{B_1 \ldots B_k}$ is a generic multipartite entangled state;

(b) $\Omega_{B_1 \ldots B_k}$ is constrained to be either bi-separable or fully separable;

(c) $\Omega_{B_1 \ldots B_k}$ is constrained to be a product state.

We will now show that in setting (a) our Theorem 2 still holds. We need a preliminary lemma.

Lemma S30. *Let $0 \le \lambda_1, \ldots, \lambda_k \le 1$ be transmissivities, and let σ be an arbitrary m-mode state. Set $\lambda := \lambda_1 \ldots \lambda_k$. Then there exists a k-partite state $\Omega_{B_1 \ldots B_k}$ of k systems of m modes each such that the channel $\Phi_{\lambda_1, \ldots, \lambda_k; \Omega}$ defined by*

$$\Phi_{\lambda_1, \ldots, \lambda_k; \Omega}^A(\rho_A) := \mathrm{Tr}_{B_1 \ldots B_k} \left[U_{\lambda_k}^{AB_k} \ldots U_{\lambda_1}^{AB_1} (\rho_A \otimes \Omega_{B_1 \ldots B_k}) \left(U_{\lambda_1}^{AB_1} \right)^\dagger \ldots \left(U_{\lambda_k}^{AB_k} \right)^\dagger \right] \tag{S93}$$

satisfies that

$$\Phi_{\lambda, \sigma} = \Phi_{\lambda_1, \ldots, \lambda_k; \Omega}. \tag{S94}$$

Proof. Iterating the same argument as in the proof of Lemma S6 shows that

$$\chi_{\Phi_{\lambda_1, \ldots, \lambda_k; \Omega}(\rho)}(\alpha) = \chi_\rho \left(\sqrt{\lambda_1 \lambda_2 \ldots \lambda_k}\, \alpha \right) \chi_\Omega \left(\sqrt{(1 - \lambda_1)\lambda_2 \ldots \lambda_k}\, \alpha, \; \sqrt{(1 - \lambda_2)\lambda_3 \ldots \lambda_k}\, \alpha, \; \ldots, \; \sqrt{1 - \lambda_k}\, \alpha \right),$$

where $\alpha \in \mathbb{C}^m$ is generic. This coincides with $\chi_{\Phi_{\lambda, \sigma}(\rho)}(\alpha) = \chi_\rho \left(\sqrt{\lambda}\, \alpha \right) \chi_\sigma \left(\sqrt{1 - \lambda}\, \alpha \right)$ for all ρ if

$$\chi_\Omega \left(\sqrt{\frac{(1 - \lambda_1)\lambda_2 \ldots \lambda_k}{1 - \lambda_1 \ldots \lambda_k}}\, \alpha, \; \sqrt{\frac{(1 - \lambda_2)\lambda_3 \ldots \lambda_k}{1 - \lambda_1 \ldots \lambda_k}}\, \alpha, \ldots, \; \sqrt{\frac{1 - \lambda_k}{1 - \lambda_1 \ldots \lambda_k}}\, \alpha \right) = \chi_\sigma (\alpha). \tag{S95}$$

To construct a state $\Omega_{B_1 \ldots B_k}$ such that (S95) holds, we start by defining the numbers

$$\eta_i := \frac{\lambda_i (1 - \lambda_1 \ldots \lambda_{i-1})}{1 - \lambda_1 \ldots \lambda_i}, \qquad \forall \, i = 2, \ldots, k. \tag{S96}$$

We can then rephrase (S95) as

$$\chi_\Omega \left(\sqrt{\eta_2 \ldots \eta_k}\, \alpha, \; \sqrt{(1 - \eta_2)\eta_3 \ldots \eta_k}\, \alpha, \; \sqrt{(1 - \eta_3)\eta_4 \ldots \eta_k}\, \alpha, \ldots, \; \sqrt{1 - \eta_k}\, \alpha \right) = \chi_\sigma(\alpha). \tag{S97}$$

Now we observe that (S97) is satisfied if and only if

$$\sigma_{B_1} = \text{Tr}_{B_2 \dots B_k} \left[U_{\eta_k}^{B_1 B_k} \dots U_{\eta_2}^{B_1 B_2} \, \Omega_{B_1 \dots B_k} \left(U_{\eta_2}^{B_1 B_2} \right)^\dagger \dots \left(U_{\eta_k}^{B_1 B_k} \right)^\dagger \right]. \tag{S98}$$

In order to meet (S98), it suffices to set e.g.

$$\Omega_{B_1 \dots B_k} := \left(U_{\eta_2}^{B_1 B_2} \right)^\dagger \dots \left(U_{\eta_k}^{B_1 B_k} \right)^\dagger \left(\sigma_{B_1} \otimes |0\rangle\langle 0|_{B_2 \dots B_k} \right) U_{\eta_k}^{B_1 B_k} \dots U_{\eta_2}^{B_1 B_2}. \tag{S99}$$

This construction concludes the proof. Incidentally, note that we could replace $\sigma_{B_1} \otimes |0\rangle\langle 0|_{B_2 \dots B_k}$ in (S99) with any extension of σ_{B_1}. □

Remark S31. The argument in the above proof amounts to an equivalence between the channel depicted in Figure 3 and that in Figure 4 below.
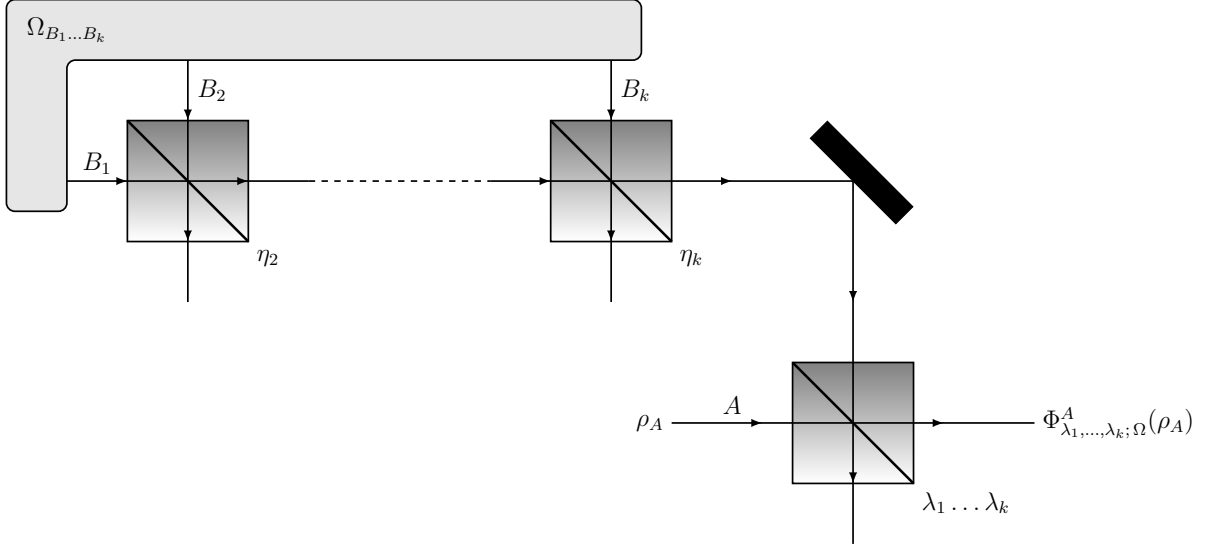

FIG. 4. An equivalent representation of the channel in Figure 3, as constructed in the above proof of Lemma S30. The transmissivities η_i $(i = 2, \dots, k)$ are given by (S96).

We are now ready to formulate the sought extension of Theorem 2.

Theorem S32. *Let $0 < \lambda_1, \dots, \lambda_k \leq 1$ be positive transmissivities. There exists a state $\Omega_{B_1 \dots B_k}$ of k single-mode systems such that the channel $\Phi_{\lambda_1, \dots, \lambda_k; \Omega}$ defined by (S93) (see Figure 3) satisfies that*

$$Q\big(\Phi_{\lambda_1, \dots, \lambda_k; \Omega}\big) \geq Q\big(\Phi_{\lambda_1, \dots, \lambda_k; \Omega}, 1/2\big) \geq c, \tag{S100}$$

where $c > 0$ is the same universal constant as in Theorem 2.

Proof. Thanks to Lemma S30, by varying the state Ω we can reproduce all general attenuator channels $\Phi_{\lambda_1 \dots \lambda_k, \sigma}$. Noting that $\lambda_1 \dots \lambda_k > 0$, it suffices to apply Theorem 2 to conclude. □

As a final remark, let us point out that, while Theorem S32 clarifies the situation in scenario (a) above, i.e. when $\Omega_{B_1 \dots B_k}$ is a generic multipartite entangled state, it would also be of interest to explore scenarios (b) and (c). We leave this for future investigations.