arXiv:2511.16299v3 [quant-ph] 5 Oct 2026

Emulation Capacity between Idempotent Channels*

Idris Delsol[1], Omar Fawzi[1], Li Gao[2,3], and Mizanur Rahaman[1,4,5]

[1] *Inria, ENS de Lyon, UCBL, LIP, Lyon, France*
[2] *School of Mathematics and Statistics, Wuhan University, Wuhan, China*
[3] *Wuhan Institute of Quantum Technology, Wuhan, China*
[4] *Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden*
[5] *Wallenberg Centre for Quantum Technology, Chalmers University of Technology, Gothenburg, Sweden*

Abstract

We study the optimal rates of emulation (also called interconversion) between quantum channels, in three regimes: when the encoder and the decoder are unassisted, when they share classical randomness, and when they share an unlimited amount of entanglement. When the source and the target channels are idempotent, we completely characterise these three regimes, in terms of the shape vectors of the two channels, which describe the structure of their ranges. In the unassisted and shared randomness regimes, we give a single-letter expression for the emulation capacity, show that it coincides with the zero-error emulation capacity, and prove a finite-blocklength converse which implies a strong converse. In these two regimes, shared randomness does not increase the emulation capacity, and the emulation of idempotent channels is not reversible. Moreover, we provide an efficient algorithm to approximate this capacity to arbitrary precision. When an unlimited amount of shared entanglement is allowed, we determine the minimal emulation error exactly at every finite blocklength. It is governed by a single quantity, the ℓ_2-norm of the shape vector, so that the emulation becomes reversible, and exact emulation is achievable with finitely many maximally entangled states. Altogether, these results draw a comprehensive picture of the task of emulating idempotent channels with each other.

Contents

1 Introduction **2**

2 Preliminaries and notation **6**
2.1 Emulation of quantum channels 7

3 Decomposition properties of idempotent channels **8**

4 Unassisted emulation of idempotent channels with zero error **16**
4.1 Achievability in item (i) of Theorem A 18

5 Minimal error for unassisted and shared-randomness-assisted emulation for finite blocklength **22**

*A short version of this work was presented at ISIT 2026 [1]. The present version additionally treats the shared randomness and entanglement-assisted regimes, and establishes a strong converse that matches the unassisted capacity $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F})$ for the full range $p \in [1, +\infty]$, whereas the converse of the short version was restricted to $p \in \{1, +\infty\}$.

6 An efficient algorithm for the emulation capacity **28**
6.1 Main result and model of computation . . . 29
6.2 Reduction to a bounded range of p . . . 30
6.3 Polynomial approximation on a dyadic interval . . . 32
6.4 Certified minimisation of a rational function . . . 35
6.5 The algorithm and the proof of Theorem 24 . . . 36

7 Entanglement-assisted emulation of idempotent channels **38**
7.1 Entanglement-assisted protocols and exact emulations . . . 39
7.2 The operational meaning of $\|\lambda(\mathcal{F})\|_2^2$ . . . 40
7.3 Entanglement-assisted emulation of noiseless classical channels . . . 43
7.4 Proof of Theorem B . . . 45
7.5 Consequences and examples . . . 46

A An algebraic proof of the converse of item (i) of Theorem A in the unassisted setting **51**

B An application of approximate C^*-algebras **54**
B.1 The Error Reduction Theorem . . . 55
B.2 Approximate Multiplicative Domains . . . 56
B.3 Proof of Theorem 49 . . . 57

1 Introduction

Given target and source quantum channels $\mathcal{F}$ and $\mathcal{G}$ (respectively), a central problem in quantum information theory is to understand the optimal interconversion rates k/n such that one can use n copies of $\mathcal{G}$, combined with an encoder and a decoder channel, to emulate the action of k copies of $\mathcal{F}$. When $\mathcal{F}$ is a noiseless channel, this is the problem of channel coding for $\mathcal{G}$, which is at the heart of Shannon theory [2] and of its quantum generalisations [3, 4]. When $\mathcal{G}$ is a noiseless channel, this is the problem of channel simulation for $\mathcal{F}$, which is also well-studied in both classical and quantum Shannon theory, where it gives rise to reverse Shannon theorems [5, 6, 7, 8]. More generally, the interconversion of channels is a natural question in resource theories of quantum channels [9, 10, 11].

Determining such optimal rates is notoriously difficult. Already for channel coding, the capacities of a quantum channel are in general given by regularised formulas, which involve an optimisation over an unbounded number of uses of the channel, and superadditivity phenomena show that this regularisation cannot be dispensed with in general, see e.g. [12] and the references therein. Even for classical channels, no single-letter formula is known for the zero-error capacity, and its computability is a delicate question [13]. The interconversion of two general channels is a fortiori harder: it contains both channel coding and channel simulation as special cases, and neither the source nor the target channel is noiseless. It is therefore natural to look for families of channels for which the interconversion problem can be solved completely.

In this article, we solve this problem when both $\mathcal{F}$ and $\mathcal{G}$ are *idempotent*, i.e. when they satisfy $\mathcal{F}\circ\mathcal{F} = \mathcal{F}$ and $\mathcal{G}\circ\mathcal{G} = \mathcal{G}$. Examples of idempotent channels include the identity channel, the completely depolarising channel, the completely dephasing channel and channels on bipartite systems $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ of the form $x \mapsto \mathrm{tr}_B(x) \otimes \rho_B$ with ρ_B a state in $\mathcal{L}(\mathcal{H}_B)$. Idempotent channels naturally appear as the limits of iterated noise: for an arbitrary channel Φ, a subsequence of the sequential compositions $\Phi^n = \Phi\circ\cdots\circ\Phi$ converges to an idempotent channel [14, 15], which describes the information that survives an arbitrarily long Markovian evolution. The capacities of quantum Markovian noise for large times were studied from this perspective in [16, 17, 18]. As we will see, all the properties of an idempotent channel $\mathcal{F}$ relevant for interconversion are captured by an integer vector $\lambda(\mathcal{F})$, called its *shape vector*. The range of $\mathcal{F}$ decomposes as a direct sum of blocks of the form $\mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k$, with ρ_k a fixed state, and $\lambda(\mathcal{F})$ is the vector of the dimensions of the spaces $\mathcal{H}_{k,1}$, sorted in non-increasing order (see Proposition 1 for the precise statement). For instance, $\lambda(\mathrm{Id}_d) = (d)$ and $\lambda(\Delta_d) = (1, \ldots, 1) \in \mathbb{N}^d$, with Id_d and Δ_d the identity and the completely dephasing channels on $\mathcal{L}(\mathbb{C}^d)$. Shape vectors were introduced by Kuperberg [14] to compare hybrid quantum memories, and, given an idempotent channel $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, $\lambda(\mathcal{F})$ can be computed in time polynomial in the dimension of $\mathcal{H}$ [16]. We write $\|u\|_p$ for the ℓ_p-norm of a vector u (see eq. (5)).

We consider three regimes, according to the resources that the encoder and the decoder share: in the *unassisted* regime (UA), they share nothing; in the *shared randomness* regime (SR), they can use a common random variable to choose the pair of encoding and decoding channels that they apply; in the *entanglement-assisted* regime (EA), they share an arbitrary entangled state, of arbitrary finite dimension. In each regime $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$ and for $k, n \in \mathbb{N}_{>0}$, we denote by $\delta^{\mathrm{R}}_{k,n}(\mathcal{G} \mapsto \mathcal{F})$ the minimal error with which k copies of $\mathcal{F}$ can be approximated by a channel built from n uses of $\mathcal{G}$, preceded by an encoding and followed by a decoding channel, possibly with the help of the shared resource allowed in the regime, the error being measured in normalised diamond distance. In each regime R, the *emulation capacity* $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})$ is the largest rate r such that, along some sequence of values of n going to infinity, $\lceil rn \rceil$ uses of $\mathcal{F}$ can be emulated by n uses of $\mathcal{G}$ with an error $\delta^{\mathrm{R}}_{\lceil rn \rceil, n}(\mathcal{G} \mapsto \mathcal{F})$ tending to zero. We say that an emulation is *exact*, or has *zero error*, when the channel it implements is exactly $\mathcal{F}^{\otimes k}$. The *zero-error emulation capacity* is the supremum of the ratios k/n over all pairs (k, n) of positive integers for which $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$. The zero-error emulation capacity is at most the emulation capacity, and the two quantities can differ for general channels. The precise definitions are given in Section 2.1.

Our first main result gives a complete description of the unassisted and shared randomness regimes for idempotent channels.

Theorem A (Unassisted and shared randomness regimes)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be two idempotent channels with $\lambda(\mathcal{F}) \neq (1)$.*

(i) Capacity. *The unassisted and the shared randomness emulation capacities of $\mathcal{F}$ by $\mathcal{G}$ coincide and are given by the single-letter formula*

$$C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{p \in [1, +\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}, \tag{1}$$

where the term $p = +\infty$ is omitted when $\|\lambda(\mathcal{F})\|_\infty = 1$. Moreover, in both regimes, this capacity coincides with the zero-error emulation capacity.

(ii) Finite blocklength. *For all $k, n \in \mathbb{N}_{>0}$,*

$$\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - \inf_{p \in [1, +\infty]} \frac{\|\lambda(\mathcal{G})\|_p^n}{\|\lambda(\mathcal{F})\|_p^k}. \tag{2}$$

(iii) Strong converse. *At rates above the capacity, the minimal error does not merely fail to vanish, but instead tends to its maximal value 1, exponentially fast in n. More precisely, for every $\varepsilon > 0$ there is a constant $c > 0$, depending only on ε, $\mathcal{F}$ and $\mathcal{G}$, such that*

$$\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - 2^{-cn} \quad \text{for all } k, n \in \mathbb{N}_{>0} \text{ with } k/n \geq C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) + \varepsilon.$$

In particular, for any sequences (k_ν), (n_ν) with $n_\nu \to \infty$ and $\liminf_\nu k_\nu/n_\nu > C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F})$, both $\delta^{\mathrm{UA}}_{k_\nu, n_\nu}(\mathcal{G} \mapsto \mathcal{F})$ and $\delta^{\mathrm{SR}}_{k_\nu, n_\nu}(\mathcal{G} \mapsto \mathcal{F})$ tend to 1 exponentially fast in n_ν.

(iv) Computability. *The capacity $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F})$ can be approximated up to an additive error 2^{-b} in time polynomial in b and in the dimensions of the input spaces of $\mathcal{F}$ and $\mathcal{G}$ (see Corollary 25 for the computational model).*

Item (i) of Theorem A is proved in Section 4 (achievability) and Section 5 (converse), items (ii) and (iii) in Section 5, and item (iv) in Section 6.

Our second main result gives a complete description of the entanglement-assisted regime for idempotent channels, at every finite blocklength.

Theorem B (Entanglement-assisted regime)**.** *Let $\mathcal{F}$ and $\mathcal{G}$ be two idempotent channels.*

(i) Capacity. *If $\lambda(\mathcal{F}) \neq (1)$, the entanglement-assisted emulation capacity of $\mathcal{F}$ by $\mathcal{G}$ is*

$$C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) = \frac{\log(\|\lambda(\mathcal{G})\|_2)}{\log(\|\lambda(\mathcal{F})\|_2)}, \tag{3}$$

and it coincides with the zero-error entanglement-assisted emulation capacity.

(ii) Exact finite blocklength error. *For all* $k, n \in \mathbb{N}_{>0}$,

$$\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \max\left\{0, 1 - \frac{\|\lambda(\mathcal{G})\|_2^{2n}}{\|\lambda(\mathcal{F})\|_2^{2k}}\right\}. \tag{4}$$

Furthermore, when this quantity vanishes, $\mathcal{F}^{\otimes k}$ *can be emulated exactly by* $\mathcal{G}^{\otimes n}$ *using finitely many maximally entangled states.*

(iii) Strong converse. *If* $\lambda(\mathcal{F}) \neq (1)$*, for every* $\varepsilon > 0$*, there is a constant* $c > 0$*, depending only on* ε *and* $\mathcal{F}$*, such that* $\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - 2^{-cn}$ *for all* $k, n \in \mathbb{N}_{>0}$ *with* $k/n \geq C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) + \varepsilon$.

All items of Theorem B are proved in Section 7.

Ideas of the proofs. Both theorems rest on a reduction to *block pinchings*. An idempotent channel $\mathcal{F}$ is, up to encoding and decoding, a very simple object: it transmits a classical index $k \in [K]$ together with a perfect quantum system of dimension d_k, and the list $\lambda(\mathcal{F}) = (d_1, \ldots, d_K)$ of these dimensions is its *shape vector*. More precisely, $\mathcal{F}$ can be converted exactly into the *block pinching* $\Pi_{\lambda(\mathcal{F})}$, which acts on $\bigoplus_k \mathbb{C}^{d_k}$ by erasing the coherences between different blocks and leaving each block untouched (Definition 3), and conversely $\Pi_{\lambda(\mathcal{F})}$ can be converted exactly into $\mathcal{F}$ (Proposition 3). Consequently, every emulation question only depends on the shape vectors of the two channels (Corollary 4). Since shape vectors are multiplicative under tensor products (Proposition 7), and so are their ℓ_p-norms, all the expressions above are single-letter.

The achievability part of Theorem A (Section 4) is proved in the unassisted regime, which suffices for the shared randomness regime as well, since an unassisted strategy is a shared randomness strategy. It replaces both channels by the tracial conditional expectations onto the $*$-algebras $\bigoplus_k \mathcal{L}(\mathbb{C}^{d_k})$, and relies on a theorem of Kuperberg [14]: when $\|\lambda(\mathcal{F})\|_p < \|\lambda(\mathcal{G})\|_p$ for all $p \in [1, +\infty]$, the n-th tensor power of the first algebra embeds into the n-th tensor power of the second one for n large enough. We lift this embedding to an encoding and a decoding channel. The converse (Section 5) is a testing argument: we check whether the emulated channel preserves a suitable mixture of maximally entangled states on the blocks of $\Pi_{\lambda(\mathcal{F})}$, and the Cauchy–Schwarz and Hölder inequalities bound the probability of passing this test by $\|\lambda(\mathcal{G})\|_p / \|\lambda(\mathcal{F})\|_p$, uniformly over the shared randomness. Taking tensor powers yields the finite blocklength bound of item (ii) in Theorem A, from which all the other converse statements follow. This argument is close to the proof of the Hölder-type inequality for tracial conditional expectations in [14, Theorem 4.1]. It nevertheless differs from it, as it bounds the diamond-norm error of arbitrary idempotent channels at every finite blocklength, instead of the fidelity of encoding some states, which is what yields a non-asymptotic strong converse. The algorithm of item (iv) (Section 6) first restricts p to a bounded range thanks to an explicit tail bound, then minimises certified polynomial approximations of the ratio $\log(\|\lambda(\mathcal{G})\|_p) / \log(\|\lambda(\mathcal{F})\|_p)$ on logarithmically many dyadic intervals, by real root isolation.

In the entanglement-assisted regime (Section 7), teleportation and superdense coding, applied block by block, show that $\Pi_{\lambda(\mathcal{F})}$, and hence $\mathcal{F}$, can be converted exactly with the help of shared entanglement, with a single use in each direction, into the noiseless classical channel on an alphabet of size $\|\lambda(\mathcal{F})\|_2^2$, and back. The problem therefore reduces to the entanglement-assisted emulation of noiseless classical channels with one another. A strategy based on a random permutation of the alphabet, which only uses shared randomness, is optimal by a no-signalling argument.

Relation to previous work in the unassisted and shared randomness regimes. Several special cases of Theorem A were known. When both $\mathcal{F}$ and $\mathcal{G}$ are identity channels, the capacity and a finite blocklength converse were obtained by Kretschmann and Werner [19, Section 3.2]. For these channels, our finite blocklength bound strengthens theirs by a factor of 2. When $\mathcal{F}$ is a noiseless quantum or classical channel, the infimum in Theorem A is attained at an endpoint and the formula simplifies. Writing $\lambda(\mathcal{G}) = (d_1, \ldots, d_K)$, the identity channel Id_d has shape vector (d) and the completely dephasing channel Δ_d has shape vector $(1, \ldots, 1)$, so that

$$C_{\mathrm{R}}(\mathcal{G} \mapsto \mathrm{Id}_d) = \frac{\log \max_k d_k}{\log d}, \qquad C_{\mathrm{R}}(\mathcal{G} \mapsto \Delta_d) = \frac{\log \sum_k d_k}{\log d},$$

for $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}\}$, with the infimum attained at $p = +\infty$ and $p = 1$ respectively. Taking $d = 2$, these are the quantum and classical capacities of the idempotent channel $\mathcal{G}$, equal to $\log_2 \max_k d_k$ qubits and $\log_2 \sum_k d_k$ bits per use. The quantum capacity and its strong converse were first obtained for self-adjoint

idempotent channels (also known as tracial conditional expectations) in [20], and then for the idempotent channels arising as large-time limits of Markovian noise in [16, Theorem 2.3 and Proposition 2.5], together with the classical capacity, and independently, in the zero-error case, in [17]; see also [18]. Theorem A complements these results with an exponential strong converse in diamond norm, which also holds when the encoder and decoder share randomness.

Relation to previous work in the entanglement-assisted regime. In the asymptotic regime with vanishing error, the value of $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F})$ in Theorem B already follows from known results. Indeed, the entanglement-assisted classical capacity of an idempotent channel is $C_E(\mathcal{F}) = \log \|\lambda(\mathcal{F})\|_2^2$ [20, Proposition 3.4]. Moreover, by the entanglement-assisted coding theorem [5] combined with the quantum reverse Shannon theorem [5, 6, 7], any channel $\mathcal{G}$ can emulate any channel $\mathcal{F}$ with shared entanglement and vanishing error at rate $C_E(\mathcal{G})/C_E(\mathcal{F})$, and at no higher rate, which gives the expression of item (ii) in Theorem B for idempotent channels. However, for general input states, the quantum reverse Shannon theorem requires entangled resources beyond maximally entangled states, such as embezzling states [6, 7]. Its statements are asymptotic, and its one-shot refinements [21] express the simulation cost through smooth entropic quantities or semidefinite programs rather than in closed form. Theorem B goes beyond this asymptotic statement in three respects. First, it is a single-shot result: for all $k, n \in \mathbb{N}_{>0}$, the minimal error is given by a closed formula, which involves no vanishing error, no smoothing and no optimisation. In particular, the zero-error capacity coincides with $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F})$, and the error tends to 1 exponentially fast above the capacity, with an explicit exponent. Second, the optimal protocols are explicit and use only finitely many maximally entangled states: teleportation and superdense coding, applied block by block, reduce both channels to noiseless classical channels, which are then emulated with a random permutation of the alphabet. Third, the converse holds for arbitrary shared entangled states, and even for arbitrary no-signalling assistance, for which only linear and semidefinite programming characterisations are available in general [22, 23, 21]. Hence, for idempotent channels, maximally entangled states are as powerful as any no-signalling correlation, at every finite blocklength. Zero-error capacities and exact simulation costs of classical channels assisted by non-local correlations were determined along these lines in [24].

Consequences. Theorems A and B have several consequences, which we discuss in Sections 4 and 7. First, in the unassisted and shared randomness regimes, the emulation of idempotent channels is in general not reversible: as shown in Example 9, there are idempotent channels for which $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F})C_{\mathrm{UA}}(\mathcal{F} \mapsto \mathcal{G}) \neq 1$, because the infimum in item (i) of Theorem A can be attained at different values of p for the two directions. Second, in these regimes, the emulation capacity can be strictly superadditive under tensor products of source channels, i.e. there are channels $\mathcal{G}_1, \mathcal{G}_2$ and $\mathcal{F}$ for which $C_{\mathrm{UA}}(\mathcal{G}_1 \otimes \mathcal{G}_2 \mapsto \mathcal{F}) > C_{\mathrm{UA}}(\mathcal{G}_1 \mapsto \mathcal{F}) + C_{\mathrm{UA}}(\mathcal{G}_2 \mapsto \mathcal{F})$ (Example 8). In contrast, with shared entanglement, only the value $p = 2$ matters: if $\lambda(\mathcal{F}) \neq (1)$ and $\lambda(\mathcal{G}) \neq (1)$, the emulation is reversible, $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F})C_{\mathrm{EA}}(\mathcal{F} \mapsto \mathcal{G}) = 1$, and the capacity is additive under tensor products of source channels (Corollary 43). At finite blocklength, the additive quantity $\log \|\lambda(\mathcal{F})\|_2$ completely characterises exact entanglement-assisted conversions of idempotent channels (Corollary 44). Finally, taking $\mathcal{F} = \mathrm{Id}_2$, we obtain the quantum capacity of arbitrary idempotent channels, together with a strong converse, in the unassisted, shared randomness and entanglement-assisted regimes.

Organisation of the article. This article is structured as follows. Section 2 introduces our notation, basic preliminaries on quantum information and operator algebras, and the precise definitions of the emulation tasks that we consider. Section 3 studies the structure of idempotent channels and their reduction to block pinchings. Section 4 proves the capacity formula of Theorem A in the unassisted and shared randomness regimes (item (i)), starting with the achievability part. Section 5 proves the finite blocklength bound (item (ii)), from which we deduce the converse part of item (i) and the strong converse (item (iii)). Section 6 gives an efficient algorithm to compute $C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F})$ (item (iv)). Section 7 proves Theorem B, together with its consequences for entanglement-assisted emulation. Finally, Appendix A gives a purely algebraic proof of the converse of item (i) for exact unassisted emulation, and Appendix B applies the theory of approximate C^*-algebras recently developed by Kitaev [25] to strengthen this converse for reduced idempotent channels.

2 Preliminaries and notation

All Hilbert spaces in this work are assumed to be finite-dimensional. Given a Hilbert space $\mathcal{H}$, we write $\mathcal{L}(\mathcal{H})$ for the set of linear operators on $\mathcal{H}$. This set is endowed with the Hilbert–Schmidt inner product, denoted $\langle x, y\rangle = \mathrm{tr}(x^* y)$ for $x, y \in \mathcal{L}(\mathcal{H})$, with x^* the conjugate transpose of x and $\mathrm{tr}(\cdot)$ the trace on $\mathcal{L}(\mathcal{H})$. Note that x^* is the adjoint of x for the Hermitian product of $\mathcal{H}$. We denote by $\mathbb{1} \in \mathcal{L}(\mathcal{H})$ the identity operator on $\mathcal{H}$. Given $x \in \mathcal{L}(\mathcal{H})$, we call the *(right) support* of x, denoted $\mathrm{supp}(x) \subseteq \mathcal{H}$, the orthogonal complement of its (right) kernel. We write x^0 for the orthogonal projector onto the support of x, parallel to its kernel. Furthermore, for $x \in \mathcal{L}(\mathcal{H})$, we write $\|x\|_1$ for its Schatten 1-norm and $\|x\|$ for its Schatten ∞-norm, which coincides with its operator norm as a linear operator on $\mathcal{H}$. Recall that the set of Hermitian operators in $\mathcal{L}(\mathcal{H})$ is a partially ordered set: for $x = x^*, y = y^* \in \mathcal{L}(\mathcal{H})$, we write $x \leq y$ whenever $y - x$ is positive semidefinite.

A state is represented by a positive semidefinite operator $\rho \in \mathcal{L}(\mathcal{H})$ with trace one, which is called a *density operator*. We will take $\mathcal{D}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ to be the set of density operators belonging to $\mathcal{L}(\mathcal{H})$. A measurement on a physical system $\mathcal{L}(\mathcal{H})$ is represented by a positive operator-valued measure, POVM for short, which is a set of positive operators $\mu_1, \ldots, \mu_K \in \mathcal{L}(\mathcal{H})$, called *effects*, which sum to the identity, i.e. $\sum_{k=1}^K \mu_k = \mathbb{1}$.

A quantum channel $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ is a completely positive trace preserving linear map. Each linear map $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ admits an adjoint $\mathcal{F}^* : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$ for the Hilbert–Schmidt inner product such that for all $x \in \mathcal{L}(\mathcal{H})$, $y \in \mathcal{L}(\mathcal{H}')$, $\langle \mathcal{F}(x), y\rangle = \langle x, \mathcal{F}^*(y)\rangle$. When $\mathcal{F}$ is a channel, its adjoint $\mathcal{F}^*$ is a unital completely positive map. We write $\mathrm{Id}_d : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d)$ for the identity channel on $\mathcal{L}(\mathbb{C}^d)$ and $\Delta_d : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d)$ for the completely dephasing one. The latter maps every $x \in \mathcal{L}(\mathbb{C}^d)$ onto its diagonal $\mathrm{Diag}(x_{11}, \ldots, x_{dd})$ (with respect to the standard orthonormal basis). It is a 'classical identity', as it perfectly preserves classical information while mapping its input onto a diagonal, hence classical, state.

We use the diamond norm as a figure of merit for the quality of an emulation. We recall that for $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, the diamond norm is defined as

$$\|\mathcal{F}\|_\diamond = \sup_{n \in \mathbb{N}_{>0}} \|\mathcal{F} \otimes \mathrm{Id}_n\|_{1\to 1},$$

with $\|\cdot\|_{1\to 1}$ the 'one-to-one' norm, i.e. the map norm when the input and output spaces are endowed with the Schatten 1-norm, denoted $\|\cdot\|_1$. That is, for all linear maps $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$,

$$\|\mathcal{F}\|_{1\to 1} = \sup_{\|x\|_1 \leq 1} \|\mathcal{F}(x)\|_1.$$

Then, we recall basic properties of the ℓ_p-norms. For $u \in \mathbb{C}^n$ and $p \in [1, +\infty)$, the ℓ_p-norm of u is defined as

$$\|u\|_p = \Big(\sum_{k=1}^n |u_k|^p\Big)^{\frac{1}{p}}, \tag{5}$$

and $\|u\|_\infty = \sup_k |u_k|$. In the discussions of the next sections, we will frequently make use of the fact that the ℓ_p-norms are *multiplicative* under the tensor product of vectors. Given $u \in \mathbb{C}^n$, $v \in \mathbb{C}^m$, the tensor product $u \otimes v \in \mathbb{C}^{nm}$ is the vector whose coordinates are the products of the coordinates of u and v, i.e. $(u_k v_l)_{k,l}$ for $k \in [n]$, $l \in [m]$. Then, for all $p \in [1, +\infty]$, $\|u \otimes v\|_p = \|u\|_p \|v\|_p$.

Finally, we end this section by introducing the notions of finite-dimensional $*$-algebras and their morphisms.

In this article, we actually consider only $*$-subalgebras of $\mathcal{L}(\mathcal{H})$. Such a $*$-subalgebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H})$ is a linear subspace of $\mathcal{L}(\mathcal{H})$ which is stable under the multiplication and the $*$ operation of $\mathcal{L}(\mathcal{H})$, i.e. taking the conjugate transpose. That is, for all $x, y \in \mathcal{A}$, $xy \in \mathcal{A}$ and for all $x \in \mathcal{A}$, $x^* \in \mathcal{A}$. Furthermore, all $*$-algebras in this work are assumed to be *unital*, that is there is an identity $\mathbb{1} \in \mathcal{A}$ satisfying $\mathbb{1}^* = \mathbb{1}$ and, for all $x \in \mathcal{A}$, $\mathbb{1}x = x\mathbb{1} = x$. Such an identity operator is necessarily unique in each $*$-algebra.

Given two $*$-algebras $\mathcal{A}$ and $\mathcal{B}$, a $*$-homomorphism $\iota : \mathcal{A} \to \mathcal{B}$ is a linear map that preserves the product and the $*$ operation in $\mathcal{A}$. That is, for all $x, y \in \mathcal{A}$, $\iota(xy) = \iota(x)\iota(y)$ and $\iota(x^*) = \iota(x)^*$. In many texts, $*$-homomorphisms between unital $*$-algebras are assumed to be unital as well, that is, they are assumed to map the identity of $\mathcal{A}$ onto the identity of $\mathcal{B}$. Nevertheless, in this article, we will work with more general **subunital** $*$-homomorphisms. Such a map $\iota : \mathcal{A} \to \mathcal{B}$ maps the identity of $\mathcal{A}$ onto $\iota(\mathbb{1}_\mathcal{A})$, which satisfies $\iota(\mathbb{1}_\mathcal{A}) \leq \mathbb{1}_\mathcal{B}$. Note that $*$-homomorphisms are completely positive [26].

2.1 Emulation of quantum channels

We now define precisely the emulation tasks that we consider. Throughout this subsection, $\mathcal{F}$ and $\mathcal{G}$ are two quantum channels, called the *target* and the *source* channel respectively, and $k, n \in \mathbb{N}_{>0}$.

Unassisted regime. An unassisted emulation of $\mathcal{F}^{\otimes k}$ by $\mathcal{G}^{\otimes n}$ is given by a pair of channels $(\mathcal{E}, \mathcal{D})$, where the encoding channel $\mathcal{E}$ maps the input space of $\mathcal{F}^{\otimes k}$ to the input space of $\mathcal{G}^{\otimes n}$ and the decoding channel $\mathcal{D}$ maps the output space of $\mathcal{G}^{\otimes n}$ to the output space of $\mathcal{F}^{\otimes k}$. Its error is $\frac{1}{2}\|\mathcal{F}^{\otimes k} - \mathcal{D} \circ \mathcal{G}^{\otimes n} \circ \mathcal{E}\|_\diamond$, and we write

$$\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{\mathcal{D},\mathcal{E}} \frac{1}{2}\|\mathcal{F}^{\otimes k} - \mathcal{D} \circ \mathcal{G}^{\otimes n} \circ \mathcal{E}\|_\diamond \tag{6}$$

for the minimal error, where the infimum is taken over all encoding and decoding channels $\mathcal{E}, \mathcal{D}$.

Shared randomness regime. We can naturally extend the power of the encoder and decoder by giving them access to *shared randomness*. In this regime, the pair of encoding and decoding channels is chosen at random from a set $\{(\mathcal{E}_\lambda, \mathcal{D}_\lambda) \mid \lambda \in [\Lambda]\}$, with λ drawn according to a probability distribution $(p_\lambda)_{\lambda\in[\Lambda]}$, and the error of the emulation is $\frac{1}{2}\|\mathcal{F}^{\otimes k} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \circ \mathcal{G}^{\otimes n} \circ \mathcal{E}_\lambda\|_\diamond$. We write

$$\delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{(p_\lambda, \mathcal{D}_\lambda, \mathcal{E}_\lambda)_{\lambda\in[\Lambda]}} \frac{1}{2}\|\mathcal{F}^{\otimes k} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \circ \mathcal{G}^{\otimes n} \circ \mathcal{E}_\lambda\|_\diamond, \tag{7}$$

where the infimum is taken over all $\Lambda \in \mathbb{N}_{>0}$, all probability distributions $(p_\lambda)_{\lambda\in[\Lambda]}$ and all sets of pairs of encoding and decoding channels $\{(\mathcal{E}_\lambda, \mathcal{D}_\lambda) \mid \lambda \in [\Lambda]\}$.

Entanglement-assisted regime. In the *entanglement-assisted* regime, the encoder and the decoder share a bipartite state $\omega_{A_0B_0}$ on systems A_0, B_0 of arbitrary finite dimension. The encoding channel $\mathcal{E}$ acts on the input of $\mathcal{F}^{\otimes k}$ and on A_0, and the decoding channel $\mathcal{D}$ acts on the output of $\mathcal{G}^{\otimes n}$ and on B_0. With $\mathcal{T}_{\omega_{A_0B_0}} : x \mapsto x \otimes \omega_{A_0B_0}$ for $x \in \mathcal{L}(\mathcal{H}_{\mathcal{F}}^{\otimes k})$ the input of $\mathcal{F}^{\otimes k}$, the error of the emulation is

$$\frac{1}{2}\|\mathcal{F}^{\otimes k} - \mathcal{D} \circ (\mathcal{G}^{\otimes n} \otimes \mathrm{Id}_{B_0}) \circ (\mathcal{E} \otimes \mathrm{Id}_{B_0}) \circ \mathcal{T}_{\omega_{A_0B_0}}\|_\diamond, \tag{8}$$

and we write

$$\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{\omega_{A_0B_0}, \mathcal{E}, \mathcal{D}} \frac{1}{2}\|\mathcal{F}^{\otimes k} - \mathcal{D} \circ (\mathcal{G}^{\otimes n} \otimes \mathrm{Id}_{B_0}) \circ (\mathcal{E} \otimes \mathrm{Id}_{B_0}) \circ \mathcal{T}_{\omega_{A_0B_0}}\|_\diamond, \tag{9}$$

where the infimum is taken over all bipartite states $\omega_{A_0B_0}$, with systems A_0, B_0 of arbitrary finite dimension, and over all encoding and decoding channels $\mathcal{E}, \mathcal{D}$.

Capacities. Let $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$ be one of the three regimes. A rate $r \geq 0$ is *achievable* in the regime R if there exist two sequences $(k_\nu)_{\nu\in\mathbb{N}}$ and $(n_\nu)_{\nu\in\mathbb{N}}$ of positive integers such that

$$\lim_{\nu\to\infty} n_\nu = +\infty, \qquad \liminf_{\nu\to\infty} \frac{k_\nu}{n_\nu} \geq r, \qquad \lim_{\nu\to\infty} \delta^{\mathrm{R}}_{k_\nu,n_\nu}(\mathcal{G} \mapsto \mathcal{F}) = 0. \tag{10}$$

The *emulation capacity* of $\mathcal{F}$ by $\mathcal{G}$ in the regime R, denoted $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})$, is the supremum of the achievable rates, and it is understood to be zero when no rate is achievable. We may also use the notation $C(\mathcal{G} \mapsto \mathcal{F})$ and $\delta_{k,n}(\mathcal{G} \mapsto \mathcal{F})$ when the regime is unambiguous.

We say that $\mathcal{F}^{\otimes k}$ can be emulated *exactly*, or with *zero error*, by $\mathcal{G}^{\otimes n}$ in the regime R if there is an emulation in this regime whose error is equal to zero. The *zero-error emulation capacity* in the regime R is the supremum of the ratios k/n over the pairs (k, n) for which $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$, and is zero when there is no such pair. If $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$, then, for every $m \in \mathbb{N}_{>0}$, $\mathcal{F}^{\otimes mk}$ can be emulated exactly by $\mathcal{G}^{\otimes mn}$, by taking the m-th tensor powers of the encoding and decoding channels, and of the shared probability distribution or shared state. Hence, every ratio k/n of this form is an achievable rate, and the zero-error emulation capacity is at most the emulation capacity $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})$. For idempotent channels, these two capacities coincide in all three regimes, by item (i) of Theorems A and B.

These three regimes form a hierarchy: an unassisted strategy is a shared randomness strategy with a single pair of encoding and decoding channels, and a shared randomness strategy is an entanglement-assisted protocol in which the shared state is a classical correlated state, a standard fact about shared randomness as a particular case of shared quantum states, see e.g. [27]. Hence, we have the inequalities

$$C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) \leq C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \leq C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}), \tag{11}$$

$$\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}). \tag{12}$$

Shared randomness allows one to average several strategies, which a single deterministic pair cannot do, and shared entanglement provides correlations that no classical randomness can reproduce. For idempotent channels, we will see that the first inclusion does not change the emulation capacity (item (i) of Theorem A), whereas the second one can increase it (Example 45).

We summarise the notation we use throughout this article in Table 1.

$\mathcal{H}$	A finite-dimensional Hilbert space	
$\mathcal{L}(\mathcal{H})$	Set of linear operators on $\mathcal{H}$	
$\mathcal{D}(\mathcal{H})$	Set of density operators in $\mathcal{L}(\mathcal{H})$	
$\mathcal{A}, \mathcal{B}$	$*$-subalgebras of $\mathcal{L}(\mathcal{H})$	
ρ, σ	Density operators	
$\mathrm{supp}(x)$	Support of x	
x^0	Orthogonal projector on the support of x	
x^*	Hermitian adjoint of x	
$\lVert \cdot \rVert_p$	ℓ_p-norm	Eq. (5)
$\mathcal{F}^*$	Adjoint of $\mathcal{F}$ for the Hilbert–Schmidt product	
$\widehat{\mathcal{F}}$	Reduced channel associated to $\mathcal{F}$	Prop. 1, eq. (16)
$\mathrm{Rg}(\mathcal{F})$	Range of $\mathcal{F}$	
$\lambda(\mathcal{F})$	Shape vector of $\mathcal{F}$	Prop. 1
$\mathcal{T}_{\omega_{A_0B_0}}$	Map $x \mapsto x \otimes \omega_{A_0B_0}$	Eq. (8)
$[n]$	The set of integers $\{1, \dots, n\}$	
$\mathbb{R}^K_{\geq 0}$, $\mathbb{R}^K_{>0}$	Vectors in $\mathbb{R}^K$ with non-negative (resp. positive) entries	Lem. 22
$\mathbb{N}^K_{>0}$	Elements of $\mathbb{N}^K$ with non-zero entries	Lem. 23
$H(w)$	Shannon entropy of a probability vector w	Lem. 29
$\mathbb{Q}[X]$	Set of univariate polynomials with rational coefficients	Lem. 27, 34
$\lfloor y \rfloor_k$	A dyadic lower approximation of y with error at most 2^{-k}	Eq. (86)

Table 1: Notation

3 Decomposition properties of idempotent channels

In this section, we discuss the decomposition properties of the range of idempotent channels, which are summarised in the following propositions. The range of a map $f : \Omega \to \Omega'$, denoted $\mathrm{Rg}(f)$, is the set of all the images of the elements of its domain Ω, i.e. $\mathrm{Rg}(f) = \{f(x) \mid x \in \Omega\}$. The range of an idempotent channel coincides with its set of fixed points and the structure of this space is well understood, see for example the lecture notes [15]. The following proposition summarises the properties that we use.

Proposition 1 (Structure of the range of idempotent channels, see [15])**.** *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be an idempotent channel, i.e. satisfying $\mathcal{F} = \mathcal{F} \circ \mathcal{F}$. We can write $\mathcal{H}$ as the direct sum*

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1, \tag{13}$$

with $\mathcal{H}_0 = \mathrm{supp}(\mathcal{F}(\mathbb{1}))$ and $\mathcal{H}_1 = \ker(\mathcal{F}(\mathbb{1}))$, and $\mathcal{H}_0$ can further be decomposed as the orthogonal direct sum

$$\mathcal{H}_0 = \bigoplus_{k=1}^{K} \mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}, \tag{14}$$

so that there is a set of density operators $\{\rho_k \in \mathcal{D}(\mathcal{H}_{k,2}) \mid k \in [K]\}$*, such that*

$$\mathrm{Rg}(\mathcal{F}) = \Big(\bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k\Big) \oplus 0. \tag{15}$$

Let $e = \mathcal{F}(\mathbb{1})^0 \in \mathcal{L}(\mathcal{H})$ *be the orthogonal projector onto* $\mathcal{H}_0$*, and* $V : \mathcal{H}_0 \to \mathcal{H}$ *be the isometry embedding* $\mathcal{H}_0$ *into* $\mathcal{H}$ *such that* $VV^* = e$ *and* $V^*V = \mathbb{1}_0$ *the identity operator in* $\mathcal{L}(\mathcal{H}_0)$*. The* reduced quantum channel *associated to* $\mathcal{F}$*, denoted* $\widehat{\mathcal{F}} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_0)$ *is then defined as*

$$\widehat{\mathcal{F}} : x \mapsto V^*\mathcal{F}(VxV^*)V, \tag{16}$$

which is idempotent and whose range can be written as

$$\mathrm{Rg}(\widehat{\mathcal{F}}) = \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k. \tag{17}$$

Furthermore, we can express $\widehat{\mathcal{F}}$ *and its adjoint* $\widehat{\mathcal{F}}^*$ *in the following way:*

$$\widehat{\mathcal{F}}(x) = \sum_{k=1}^{K} \mathrm{tr}_{k,2}(P_k x P_k) \otimes \rho_k, \tag{18}$$

and

$$\widehat{\mathcal{F}}^*(x) = \sum_{k=1}^{K} \mathrm{tr}_{k,2}(P_k x P_k(\mathbb{1}_{d_k} \otimes \rho_k)) \otimes \mathbb{1}_{m_k}, \tag{19}$$

where, for all $k \in [K]$*,* P_k *is the orthogonal projector onto* $\mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}$*,* $d_k = \dim(\mathcal{H}_{k,1}) \geq 1$*,* $m_k = \dim(\mathcal{H}_{k,2})$ *and* $\mathrm{tr}_{k,2} : \mathcal{L}(\mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}) \to \mathcal{L}(\mathcal{H}_{k,1})$ *is the partial trace over* $\mathcal{H}_{k,2}$*.*

Therefore, the range of $\widehat{\mathcal{F}}^*$ *is a unital* $*$*-subalgebra of* $\mathcal{L}(\mathcal{H}_0)$ *and the following decomposition holds:*

$$\mathrm{Rg}(\widehat{\mathcal{F}}^*) = \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \mathbb{1}_{m_k}. \tag{20}$$

We call the shape *of* $\mathcal{F}$ *the* K*-dimensional vector whose coordinates are the dimensions* $(d_k)_{k\in[K]}$ *sorted in non-increasing order, which we denote by* $\lambda(\mathcal{F})$*.*

Note that $\mathcal{F}$ and its reduced channel $\widehat{\mathcal{F}}$ both have the same shape vector, i.e. $\lambda(\mathcal{F}) = \lambda(\widehat{\mathcal{F}})$.

As an example of a shape vector, consider the qubit channel $\mathcal{F} : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^2)$ defined as

$$\mathcal{F} : x \mapsto \mathrm{tr}(x)|0\rangle\langle 0| = \begin{pmatrix} \mathrm{tr}(x) & 0 \\ 0 & 0 \end{pmatrix},$$

where the matrix expression is written in the basis $(|0\rangle, |1\rangle)$. We have

$$\mathcal{F}(\mathbb{1}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then, in the notation of eq. (14), we get $\mathcal{H}_0 = \mathrm{supp}(\mathcal{F}(\mathbb{1})) = \mathbb{C}\,|0\rangle$, $\mathcal{H}_1 = \ker(\mathcal{F}(\mathbb{1})) = \mathbb{C}\,|1\rangle$, we can write $\mathrm{Rg}(\mathcal{F}) = \mathbb{C} \otimes |0\rangle\langle 0| \oplus 0$ and thus $\lambda(\mathcal{F}) = (1)$. To construct the reduced version of $\mathcal{F}$ we can take the isometry $V = \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \mathcal{H}_0 \to \mathcal{H}$, where the expression of V is written in the basis $(|0\rangle, |1\rangle)$. Note that we clearly have

$$VV^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \mathcal{F}(\mathbb{1})^0.$$

Then, $\widehat{\mathcal{F}} : \mathcal{L}(\mathbb{C}) \to \mathcal{L}(\mathbb{C})$ is defined by the equation

$$\widehat{\mathcal{F}}(x) = V^*\mathcal{F}(VxV^*)V = x,$$

so that $\widehat{\mathcal{F}}$ is the identity on $\mathcal{L}(\mathbb{C})$.

Furthermore, we remark that $\lambda(\mathcal{F}) = (1)$ if and only if the reduced channel of $\mathcal{F}$ is a replacer channel, i.e. $\widehat{\mathcal{F}} : x \mapsto \mathrm{tr}(x)\rho$ with ρ a fixed state. This fact follows easily from eq. (18) taking $K = 1$.

Before we present the proof of Proposition 1, we need to state the following two lemmas, which will be used in the proof.

Lemma 1. *Let $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H})$ be a $*$-subalgebra containing the identity operator $\mathbb{1}_{\mathcal{H}}$, and let $\Phi : \mathcal{A} \to \mathcal{L}(\mathcal{H}')$ be a positive linear map. Let $\mathcal{H}_0 = \mathrm{supp}(\Phi(\mathbb{1}_{\mathcal{H}})) \subseteq \mathcal{H}'$ and let $V : \mathcal{H}_0 \to \mathcal{H}'$ be the inclusion isometry, so that VV^* is the orthogonal projector onto $\mathcal{H}_0$. Then, for every $x \in \mathcal{A}$,*

$$VV^*\Phi(x)VV^* = \Phi(x).$$

Proof. Let $x \in \mathcal{A}$ be positive semidefinite, then $0 \leq x \leq \|x\|\mathbb{1}$, so that by positivity of Φ, we have

$$0 \leq \Phi(x) \leq \|x\|\Phi(\mathbb{1}).$$

Hence, $\Phi(x)$ is supported on the support of $\Phi(\mathbb{1})$ which is, by definition, $\mathcal{H}_0$. Furthermore, as x and $\mathbb{1}$ are positive semidefinite, $\Phi(x)$ and $\Phi(\mathbb{1})$ are positive semidefinite as well so that their respective ranges (as linear maps on $\mathcal{H}'$) coincide with their supports. Hence, we have

$$VV^*\Phi(x)VV^* = \Phi(x).$$

Now, if $x \in \mathcal{A}$ is arbitrary, we can use the polarisation identity (see [15, eq. (1.45)]) to write

$$x = \frac{1}{4}\sum_{k=0}^{3} i^k(x + i^k\mathbb{1})^*(x + i^k\mathbb{1}),$$

so that x is a linear combination of four positive semidefinite elements of $\mathcal{A}$. Now, by linearity, we get

$$VV^*\Phi(x)VV^* = \Phi(x).$$

□

Furthermore, we remark that the set of fixed points and the range of an idempotent function coincide.

Lemma 2. *Let $f : \Omega \to \Omega$ be an arbitrary idempotent function, with Ω an arbitrary set, then*

$$\mathrm{Rg}(f) = \{x \in \Omega \mid f(x) = x\}. \tag{21}$$

Proof. If x is a fixed point, $x = f(x) \in \mathrm{Rg}(f)$. For the other inclusion, if $y = f(x)$, we also have $f(y) = f(f(x)) = f(x) = y$, so y is a fixed point. □

Proposition 1 then basically follows from statements in [15, Section 6.4] and we detail it here only for the convenience of the reader. A similar derivation can be found in [28] in the Heisenberg picture.

Proof of Proposition 1. We will first prove that the reduced channel $\widehat{\mathcal{F}}$, defined in eq. (16), is actually a channel. It is clearly completely positive as it is a composition of completely positive maps and it is trace preserving because for all $x \in \mathcal{L}(\mathcal{H}_0)$,

$$\mathrm{tr}(\widehat{\mathcal{F}}(x)) = \mathrm{tr}(V^*\mathcal{F}(VxV^*)V) = \mathrm{tr}(VV^*\mathcal{F}(VxV^*)VV^*) = \mathrm{tr}(\mathcal{F}(VxV^*)) = \mathrm{tr}(VxV^*) = \mathrm{tr}(x),$$

where the third equality follows from Lemma 1. Furthermore, $\widehat{\mathcal{F}}$ is idempotent as, for $x \in \mathcal{L}(\mathcal{H}_0)$,

$$\begin{aligned}\widehat{\mathcal{F}}(\widehat{\mathcal{F}}(x)) &= V^*\mathcal{F}(VV^*\mathcal{F}(VxV^*)VV^*)V \\ &= V^*\mathcal{F}(\mathcal{F}(VxV^*))V \\ &= V^*\mathcal{F}(VxV^*)V \\ &= \widehat{\mathcal{F}}(x).\end{aligned}$$

The second equation follows from Lemma 1, and the third from the idempotence of $\mathcal{F}$.

Furthermore, $\widehat{\mathcal{F}}$ admits a full-rank fixed point. Indeed, $\widehat{\mathcal{F}}(V^*\mathcal{F}(\mathbb{1})V)$ is a fixed point of $\widehat{\mathcal{F}}$, since $\widehat{\mathcal{F}}$ is idempotent, and we have

$$\widehat{\mathcal{F}}(V^*\mathcal{F}(\mathbb{1})V) = V^*\mathcal{F}(VV^*\mathcal{F}(\mathbb{1})VV^*)V = V^*\mathcal{F}\mathcal{F}(\mathbb{1})V = V^*\mathcal{F}(\mathbb{1})V,$$

which is full-rank in $\mathcal{H}_0 = \mathrm{supp}(\mathcal{F}(\mathbb{1}))$, again using Lemma 1 for the second equality.

Therefore, we can apply [15, Theorem 6.12] to the adjoint $\widehat{\mathcal{F}}^*$ of $\widehat{\mathcal{F}}$. Hence, the set of fixed points of $\widehat{\mathcal{F}}^*$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_0)$. By Lemma 2, since $\widehat{\mathcal{F}}^*$ is idempotent, its set of fixed points coincides with its range, so that its range is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_0)$. It is well known (see e.g. [29]) that if $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_0)$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_0)$, we can find a basis of $\mathcal{H}_0$ in which it can be expressed as an orthogonal direct sum as in eq. (20), which proves this equation. Furthermore, we get by [15, Proposition 1.5] that $\widehat{\mathcal{F}}^*$ takes the form of eq. (19). Then, taking the adjoint of this equation, we obtain that $\widehat{\mathcal{F}} = \widehat{\mathcal{F}}^{**}$ can be written as in eq. (18). Then, for $x \in \mathcal{L}(\mathcal{H}_0)$, we have that

$$\widehat{\mathcal{F}}(x) = \sum_{k=1}^{K} \mathrm{tr}_{k,2}(P_k x P_k) \otimes \rho_k \in \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k.$$

Conversely, if

$$x = \sum_{k=1}^{K} x_k \otimes \rho_k \in \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k,$$

where, for each $k \in [K]$, $x_k \in \mathcal{L}(\mathcal{H}_{k,1})$, we have that $P_l x_k \otimes \rho_k P_l = \delta_{k,l} x_k \otimes \rho_k$, since the spaces $\mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}$ are pairwise orthogonal, so that $x = \widehat{\mathcal{F}}(x)$, and finally $x \in \mathrm{Rg}(\widehat{\mathcal{F}})$. This proves eq. (17).

Finally, to prove eq. (15), let $x \in \mathrm{Rg}(\mathcal{F})$. We then have

$$x = \mathcal{F}(x) = VV^*\mathcal{F}(x)VV^* = VV^*xVV^*, \tag{22}$$

so that $x \in \mathcal{L}(\mathcal{H}_0) \oplus 0$ where the 0 operator acts on $\mathcal{H}_1$. Then, $V^*xV \in \mathcal{L}(\mathcal{H}_0)$ and

$$\widehat{\mathcal{F}}(V^*xV) = V^*\mathcal{F}(VV^*xVV^*)V = V^*\mathcal{F}(x)V = V^*xV,$$

by Lemma 1 and by eq. (22). Hence,

$$V^*xV \in \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k,$$

by eq. (17) and

$$x = VV^*xVV^* \in \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k \oplus 0.$$

Conversely, if

$$x = \begin{pmatrix} x_0 & 0 \\ 0 & 0 \end{pmatrix},$$

with

$$x_0 \in \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \rho_k,$$

then

$$x = Vx_0V^* = V\widehat{\mathcal{F}}(x_0)V^* = VV^*\mathcal{F}(Vx_0V^*)VV^* = \mathcal{F}(Vx_0V^*) = \mathcal{F}(x),$$

so that $x \in \mathrm{Rg}(\mathcal{F})$. □

We also show in this section that the shape vector introduced in Proposition 1 is the object that governs the interconversion of idempotent channels, that is, for idempotent $\mathcal{F}$ and $\mathcal{G}$, it suffices to know their shape vectors to specify the rates that will achieve any given error for the emulation of $\mathcal{F}$ by $\mathcal{G}$.

We first show that if $\mathcal{F}$ is idempotent, it can be interconverted exactly with its reduced channel $\widehat{\mathcal{F}}$, which can then be interconverted with the block pinching $\Pi_{\mathbf{f}}$, introduced in Definition 3, with $\mathbf{f} = \lambda(\mathcal{F})$.

Proposition 2. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be an idempotent channel, let $\mathcal{H}_0 = \mathrm{supp}(\mathcal{F}(\mathbb{1}))$ and let $\widehat{\mathcal{F}} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_0)$ be its reduced quantum channel. Then, there are channels $\widehat{\mathcal{E}} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_0)$, $\widehat{\mathcal{D}} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H})$, $\mathcal{E} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H})$ and $\mathcal{D} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_0)$ such that*

$$\mathcal{F} = \widehat{\mathcal{D}}\widehat{\mathcal{F}}\widehat{\mathcal{E}}, \text{ and } \widehat{\mathcal{F}} = \mathcal{D}\mathcal{F}\mathcal{E}. \tag{23}$$

Proof. Let $V : \mathcal{H}_0 \to \mathcal{H}$ be the isometry embedding $\mathcal{H}_0$ into $\mathcal{H}$ which defines $\widehat{\mathcal{F}}$ as in Proposition 1. For the first equality, let $\widehat{\mathcal{E}} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_0)$ be defined by

$$\widehat{\mathcal{E}} : x \mapsto V^*\mathcal{F}(x)V, \tag{24}$$

and $\widehat{\mathcal{D}} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H})$ be defined by

$$\widehat{\mathcal{D}} : x \mapsto VxV^*. \tag{25}$$

Both $\widehat{\mathcal{E}}$ and $\widehat{\mathcal{D}}$ are completely positive trace preserving maps. The complete positivity is clear. $\widehat{\mathcal{D}}$ is trace preserving because $\mathrm{tr}(VxV^*) = \mathrm{tr}(V^*Vx) = \mathrm{tr}(x)$ and $\widehat{\mathcal{E}}$ is trace preserving as well because for all $x \in \mathcal{L}(\mathcal{H})$,

$$\mathrm{tr}(V^*\mathcal{F}(x)V) = \mathrm{tr}(VV^*\mathcal{F}(x)VV^*) = \mathrm{tr}(\mathcal{F}(x)) = \mathrm{tr}(x), \tag{26}$$

where the second equality follows from Lemma 1. Then we have the factorisation

$$\mathcal{F} = \widehat{\mathcal{D}}\widehat{\mathcal{F}}\widehat{\mathcal{E}}. \tag{27}$$

Indeed, let $x \in \mathcal{L}(\mathcal{H})$, we have

$$\widehat{\mathcal{D}}\widehat{\mathcal{F}}\widehat{\mathcal{E}}(x) = VV^*\mathcal{F}(VV^*\mathcal{F}(x)VV^*)VV^* = \mathcal{F}(\mathcal{F}(x)) = \mathcal{F}(x),$$

where the second equality follows from Lemma 1. Now, for the factorisation of $\widehat{\mathcal{F}}$, let $\mathcal{E} = \widehat{\mathcal{D}}$ and $\mathcal{D} = \widehat{\mathcal{E}}$, for $x \in \mathcal{L}(\mathcal{H}_0)$, we have:

$$\mathcal{D}\mathcal{F}\mathcal{E}(x) = \widehat{\mathcal{E}}\mathcal{F}\widehat{\mathcal{D}}(x) = V^*\mathcal{F}(\mathcal{F}(VxV^*))V = V^*\mathcal{F}(VxV^*)V = \widehat{\mathcal{F}}(x).$$

□

We now introduce the block pinching $\Pi_{\mathbf{f}}$, which serves as the canonical idempotent channel with shape vector $\lambda(\Pi_{\mathbf{f}}) = \mathbf{f}$.

Definition 3 (Block pinching)**.** Let $\mathbf{d} = (d_1, \ldots, d_K) \in \mathbb{N}_{>0}^K$ and let

$$\mathcal{H}_{\mathbf{d}} = \bigoplus_{k=1}^{K} \mathbb{C}^{d_k}. \tag{28}$$

For $k \in [K]$, let $J_k : \mathbb{C}^{d_k} \to \mathcal{H}_{\mathbf{d}}$ be the isometry embedding $\mathbb{C}^{d_k}$ as the k-th summand. That is, for $\nu \in [d_k]$, $J_k |\nu\rangle = |\nu\rangle_k$, with $\cup_{k\in[K]}\{|\nu\rangle_k \mid \nu \in [d_k]\}$ an orthonormal basis of $\mathcal{H}_{\mathbf{d}}$. Let $\widetilde{P}_k = J_k J_k^* \in \mathcal{L}(\mathcal{H}_{\mathbf{d}})$ be the orthogonal projector onto this summand. The *block pinching* associated to $\mathbf{d}$ is the channel

$$\Pi_{\mathbf{d}} : \mathcal{L}(\mathcal{H}_{\mathbf{d}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{d}}), \qquad \Pi_{\mathbf{d}}(z) = \sum_{k=1}^{K} \widetilde{P}_k z \widetilde{P}_k. \tag{29}$$

Given an idempotent channel $\mathcal{F}$, we write $\Pi_{\mathbf{f}} = \Pi_{\lambda(\mathcal{F})}$ for the block pinching associated to its shape vector.

Operationally, $\Pi_{\mathbf{d}}$ perfectly transmits a classical label $k \in [K]$ together with a d_k-dimensional quantum system. It is a channel since $\sum_{k\in[K]} \widetilde{P}_k = \mathbb{1}_{\mathbf{d}}$ the identity of $\mathcal{L}(\mathcal{H}_{\mathbf{d}})$. Furthermore, it is idempotent, and its shape vector is $\mathbf{d}$, up to the order of the entries, since $\mathrm{Rg}(\Pi_{\mathbf{d}}) = \bigoplus_{k=1}^{K} \mathcal{L}(\mathbb{C}^{d_k})$. In particular $\lambda(\Pi_{\mathbf{f}}) = \lambda(\mathcal{F})$. Given $\mathcal{F}$ an idempotent channel, we can interconvert $\mathcal{F}$ with $\Pi_{\mathbf{f}}$ and vice versa, as shown in the following proposition.

Proposition 3. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ be an idempotent channel and let $\Pi_\mathbf{f}$ be the block pinching associated to its shape vector. Then there exist encoding and decoding channels $\mathcal{E}_1, \mathcal{D}_1, \mathcal{E}_2, \mathcal{D}_2$ such that*

$$\mathcal{F} = \mathcal{D}_1 \circ \Pi_\mathbf{f} \circ \mathcal{E}_1, \qquad \Pi_\mathbf{f} = \mathcal{D}_2 \circ \mathcal{F} \circ \mathcal{E}_2. \tag{30}$$

Proof. By Proposition 2, $\mathcal{F}$ and its reduced channel $\widehat{\mathcal{F}}$ can be converted one into the other by encoding and decoding channels, and $\lambda(\widehat{\mathcal{F}}) = \lambda(\mathcal{F})$ by Proposition 1. Since compositions of channels are channels, it therefore suffices to prove eq. (30) for $\widehat{\mathcal{F}}$, and we assume from now on that $\mathcal{F} = \widehat{\mathcal{F}}$.

We use the notation of Proposition 1: we have $\mathcal{H}_\mathcal{F} = \bigoplus_{k=1}^K \mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}$ with $d_k = \dim(\mathcal{H}_{k,1})$ and $\mathbf{f} = \lambda(\mathcal{F}) = (d_1, \dots, d_K)$ up to the order of the entries, P_k is the orthogonal projector onto $\mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}$, and

$$\mathcal{F}(x) = \sum_{k=1}^K \mathrm{tr}_{k,2}(P_k x P_k) \otimes \rho_k.$$

We identify $\mathcal{H}_\mathbf{f} = \bigoplus_{k=1}^K \mathbb{C}^{d_k}$ with $\bigoplus_{k=1}^K \mathcal{H}_{k,1}$, so that $\widetilde{P}_k$ is the orthogonal projector onto $\mathcal{H}_{k,1}$ and $\Pi_\mathbf{f}(z) = \sum_{k=1}^K \widetilde{P}_k z \widetilde{P}_k$. Define

$$\mathcal{F}_1(x) = \sum_{k=1}^K \mathrm{tr}_{k,2}(P_k x P_k), \qquad \mathcal{F}_2(z) = \sum_{k=1}^K \widetilde{P}_k z \widetilde{P}_k \otimes \rho_k, \tag{31}$$

so that $\mathcal{F}_1 : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathbf{f})$ and $\mathcal{F}_2 : \mathcal{L}(\mathcal{H}_\mathbf{f}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$. Both are channels: they are completely positive as compositions of conjugations by the P_k or $\widetilde{P}_k$, of partial traces and of tensorisations by the fixed states ρ_k, and they are trace preserving since $\sum_k P_k = \mathbb{1}_{\mathcal{H}_\mathcal{F}}$ and $\sum_k \widetilde{P}_k = \mathbb{1}_{\mathcal{H}_\mathbf{f}}$. Furthermore, we have

$$\mathcal{F} = \mathcal{F}_2 \circ \mathcal{F}_1, \quad \Pi_\mathbf{f} = \mathcal{F}_1 \circ \mathcal{F}_2. \tag{32}$$

$\Pi_\mathbf{f} \mapsto \mathcal{F}$: We have, by eq. (32) and by idempotence of $\mathcal{F}$

$$\mathcal{F} = \mathcal{F} \circ \mathcal{F} = \mathcal{F}_2 \circ \mathcal{F}_1 \circ \mathcal{F}_2 \circ \mathcal{F}_1 = \mathcal{F}_2 \circ \Pi_\mathbf{f} \circ \mathcal{F}_1.$$

$\mathcal{F} \mapsto \Pi_\mathbf{f}$: Again by eq. (32) and by idempotence of $\Pi_\mathbf{f}$

$$\Pi_\mathbf{f} = \Pi_\mathbf{f} \circ \Pi_\mathbf{f} = \mathcal{F}_1 \circ \mathcal{F}_2 \circ \mathcal{F}_1 \circ \mathcal{F}_2 = \mathcal{F}_1 \circ \mathcal{F} \circ \mathcal{F}_2.$$

□

Therefore, we can interconvert exactly any two idempotent channels having the same shape vector.

Proposition 4 (Exact emulation)**.** *Let $\mathcal{F}_1, \mathcal{F}_2$ be idempotent channels such that $\lambda(\mathcal{F}_1) = \lambda(\mathcal{F}_2)$, then there are encoding and decoding channels such that*

$$\mathcal{F}_1 = \mathcal{D}_1 \mathcal{F}_2 \mathcal{E}_1, \quad \mathcal{F}_2 = \mathcal{D}_2 \mathcal{F}_1 \mathcal{E}_2. \tag{33}$$

Proof. By Proposition 3, since $\lambda(\mathcal{F}_1) = \lambda(\mathcal{F}_2) = \mathbf{f}$ we can perform the emulation of $\mathcal{F}_2$ by $\mathcal{F}_1$ through $\Pi_\mathbf{f}$ and vice versa. □

We now show that if we can interconvert exactly general channels $\mathcal{F}_1$ with $\mathcal{F}_2$ as well as $\mathcal{G}_1$ with $\mathcal{G}_2$, then it is equivalent to study the emulation of $\mathcal{F}_1$ by $\mathcal{G}_1$ or that of $\mathcal{F}_2$ by $\mathcal{G}_2$.

Proposition 5. *Let $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{G}_1, \mathcal{G}_2$ be general quantum channels such that*

$$\mathcal{F}_1 = \mathcal{D}_{1,f} \mathcal{F}_2 \mathcal{E}_{1,f}, \qquad \mathcal{F}_2 = \mathcal{D}_{2,f} \mathcal{F}_1 \mathcal{E}_{2,f} \tag{34}$$

$$\mathcal{G}_1 = \mathcal{D}_{1,g} \mathcal{G}_2 \mathcal{E}_{1,g}, \qquad \mathcal{G}_2 = \mathcal{D}_{2,g} \mathcal{G}_1 \mathcal{E}_{2,g}, \tag{35}$$

with $\mathcal{E}_{i,a}, \mathcal{D}_{i,a}$ for $(i,a) \in \{1,2\} \times \{f,g\}$ pairs of encoding and decoding channels. Then, in every assistance regime R*, i.e.* $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$*, and for all integers $k, n \in \mathbb{N}_{>0}$*

$$\delta^\mathrm{R}_{k,n}(\mathcal{G}_1 \mapsto \mathcal{F}_1) = \delta^\mathrm{R}_{k,n}(\mathcal{G}_2 \mapsto \mathcal{F}_2). \tag{36}$$

In particular,

$$C_\mathrm{R}(\mathcal{G}_1 \mapsto \mathcal{F}_1) = C_\mathrm{R}(\mathcal{G}_2 \mapsto \mathcal{F}_2). \tag{37}$$

Proof. We prove the statement for $k = n = 1$. The general case follows by applying the same argument to the tensor powers $\mathcal{F}_i^{\otimes k}$, $\mathcal{G}_i^{\otimes n}$, $i \in [2]$, using the tensor powers of the conversion maps $\mathcal{E}_{i,a}, \mathcal{D}_{i,a}$ for $(i,a) \in \{1,2\} \times \{f,g\}$.

Let us first show $\delta^{\mathrm{SR}}_{1,1}(\mathcal{G}_1 \mapsto \mathcal{F}_1) = \delta^{\mathrm{SR}}_{1,1}(\mathcal{G}_2 \mapsto \mathcal{F}_2)$, since the same construction proves the unassisted case by taking $\Lambda = 1$. Let $\Lambda \in \mathbb{N}_{>0}$, $(p_\lambda)_{\lambda\in[\Lambda]}$ be a probability distribution and $\{(\mathcal{E}_\lambda, \mathcal{D}_\lambda) \mid \lambda \in [\Lambda]\}$ be a set of pairs of encoding and decoding channels. Then,

$$\begin{aligned}
\|\mathcal{F}_1 - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \mathcal{G}_1 \mathcal{E}_\lambda\|_\diamond &= \|\mathcal{F}_1 - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \mathcal{D}_{1,g}\mathcal{G}_2\mathcal{E}_{1,g}\mathcal{E}_\lambda\|_\diamond \\
&= \|\mathcal{D}_{2,f}\|_\diamond \|\mathcal{F}_1 - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \mathcal{D}_{1,g}\mathcal{G}_2\mathcal{E}_{1,g}\mathcal{E}_\lambda\|_\diamond \|\mathcal{E}_{2,f}\|_\diamond \\
&\geq \|\mathcal{D}_{2,f}\mathcal{F}_1\mathcal{E}_{2,f} - \mathcal{D}_{2,f} \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \mathcal{D}_{1,g}\mathcal{G}_2\mathcal{E}_{1,g}\mathcal{E}_\lambda\mathcal{E}_{2,f}\|_\diamond \\
&= \|\mathcal{F}_2 - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_{2,f}\mathcal{D}_\lambda \mathcal{D}_{1,g}\mathcal{G}_2\mathcal{E}_{1,g}\mathcal{E}_\lambda\mathcal{E}_{2,f}\|_\diamond,
\end{aligned}$$

where we used in the second equality that, as $\mathcal{E}_{2,f}$, $\mathcal{D}_{2,f}$ are channels, $\|\mathcal{E}_{2,f}\|_\diamond = \|\mathcal{D}_{2,f}\|_\diamond = 1$.

Therefore,

$$\delta^{\mathrm{SR}}_{1,1}(\mathcal{G}_2 \mapsto \mathcal{F}_2) \leq \delta^{\mathrm{SR}}_{1,1}(\mathcal{G}_1 \mapsto \mathcal{F}_1).$$

We can apply the same line of argument but converting $\mathcal{F}_2$ to $\mathcal{F}_1$ and $\mathcal{G}_1$ to $\mathcal{G}_2$ to prove the converse inequality.

The proof in the entanglement-assisted case is very similar. Recall that given a bipartite state $\omega_{A_0B_0}$ on systems A_0, B_0, we write $\mathcal{T}_{\omega_{A_0B_0}} : x \mapsto x \otimes \omega_{A_0B_0}$ for $x \in \mathcal{L}(\mathcal{H}_{\mathcal{F}_1})$. Hence given $(\mathcal{E}, \mathcal{D})$ a pair of encoding and decoding channels,

$$\|\mathcal{F}_1 - \mathcal{D}(\mathcal{G}_1\mathcal{E} \otimes \mathrm{Id}_{B_0})\mathcal{T}_{\omega_{A_0B_0}}\|_\diamond \geq \|\mathcal{F}_2 - \mathcal{D}_{2,f}\mathcal{D}(\mathcal{D}_{1,g}\mathcal{G}_2\mathcal{E}_{1,g}\mathcal{E} \otimes \mathrm{Id}_{B_0})\mathcal{T}_{\omega_{A_0B_0}}\mathcal{E}_{2,f}\|_\diamond. \tag{38}$$

Now, in the right-hand side of eq. (38), $\mathcal{T}_{\omega_{A_0B_0}}$ and $\mathcal{E}_{2,f}$ commute, in the sense that for every $x \in \mathcal{L}(\mathcal{H}_{\mathcal{F}_2})$,

$$\mathcal{T}_{\omega_{A_0B_0}}\mathcal{E}_{2,f}(x) = \mathcal{E}_{2,f}(x) \otimes \omega_{A_0B_0} = (\mathcal{E}_{2,f} \otimes \mathrm{Id}_{A_0B_0})\mathcal{T}_{\omega_{A_0B_0}}(x).$$

Hence,

$$\|\mathcal{F}_1 - \mathcal{D}(\mathcal{G}_1\mathcal{E} \otimes \mathrm{Id}_{B_0})\mathcal{T}_{\omega_{A_0B_0}}\|_\diamond \geq \|\mathcal{F}_2 - \mathcal{D}^\sharp(\mathcal{G}_2\mathcal{E}^\sharp \otimes \mathrm{Id}_{B_0})\mathcal{T}_{\omega_{A_0B_0}}\|_\diamond, \tag{39}$$

where $\mathcal{D}^\sharp = \mathcal{D}_{2,f}\mathcal{D}(\mathcal{D}_{1,g} \otimes \mathrm{Id}_{B_0})$ and $\mathcal{E}^\sharp = \mathcal{E}_{1,g}\mathcal{E}(\mathcal{E}_{2,f} \otimes \mathrm{Id}_{A_0})$. The converse inequality is proven by swapping the roles of $\mathcal{F}_1$ and $\mathcal{G}_1$ with those of $\mathcal{F}_2$ and $\mathcal{G}_2$. □

Therefore, the shape vector is the only relevant object governing the interconversion of idempotent channels.

Corollary 4 (Interconversion is governed by shape vectors)**.** *Let $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{G}_1, \mathcal{G}_2$ be idempotent channels such that $\lambda(\mathcal{F}_1) = \lambda(\mathcal{F}_2)$ and $\lambda(\mathcal{G}_1) = \lambda(\mathcal{G}_2)$. Then, in all regimes $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$ and for all $k, n \in \mathbb{N}_{>0}$*

$$\delta^{\mathrm{R}}_{k,n}(\mathcal{G}_1 \mapsto \mathcal{F}_1) = \delta^{\mathrm{R}}_{k,n}(\mathcal{G}_2 \mapsto \mathcal{F}_2). \tag{40}$$

In particular

$$C_{\mathrm{R}}(\mathcal{G}_1 \mapsto \mathcal{F}_1) = C_{\mathrm{R}}(\mathcal{G}_2 \mapsto \mathcal{F}_2). \tag{41}$$

Proof. Follows from Propositions 4 and 5. □

Throughout this work, we will therefore use the following corollary to swap the idempotent channels for their reduced versions or their associated block pinching maps to simplify our proofs.

Corollary 5. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels and $\widehat{\mathcal{F}}$, $\widehat{\mathcal{G}}$ be their associated reduced channels, as defined in Proposition 1, then*

$$\delta^{\mathrm{R}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \delta^{\mathrm{R}}_{k,n}(\widehat{\mathcal{G}} \mapsto \widehat{\mathcal{F}}) = \delta^{\mathrm{R}}_{k,n}(\Pi_{\mathbf{g}} \mapsto \Pi_{\mathbf{f}}), \tag{42}$$

so that

$$C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{R}}(\widehat{\mathcal{G}} \mapsto \widehat{\mathcal{F}}) = C_{\mathrm{R}}(\Pi_{\mathbf{g}} \mapsto \Pi_{\mathbf{f}}), \tag{43}$$

where R *denotes the assistance regime, i.e.* $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$ *and* $\mathbf{f} = \lambda(\mathcal{F})$, $\mathbf{g} = \lambda(\mathcal{G})$.

Proof. This follows directly from Propositions 2, 3 and 5. □

As mentioned in the introduction, the shape vector $\lambda(\mathcal{F})$ of an idempotent channel $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ can be computed efficiently, in time polynomial in the dimension of $\mathcal{H}$.

Proposition 6 ([16, Theorem 2.6 and Algorithm 1])**.** *Given $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ an idempotent channel, there is an algorithm that computes $\lambda(\mathcal{F})$ in time $\mathcal{O}(d^6 \log(d))$ where d is the dimension of $\mathcal{H}$, using the algorithm and computational cost convention of [16].*

We now show that the shape vector of idempotent channels preserves the tensor product, a property that will be used in the proofs of our main results, Theorems A and B. Indeed, it will allow us to reduce the study of the problem of emulating k copies of $\mathcal{F}$ using n copies of $\mathcal{G}$ for general pairs of integers $k, n \in \mathbb{N}_{>0}$ to the case where $k = n$.

Proposition 7. *Let $\mathcal{F}_1 : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_1)$ and $\mathcal{F}_2 : \mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_2)$ be two idempotent channels. Then $\mathcal{F}_1 \otimes \mathcal{F}_2$ is idempotent and*

$$\mathrm{Rg}(\mathcal{F}_1^* \otimes \mathcal{F}_2^*) = \mathrm{Rg}(\mathcal{F}_1^*) \otimes \mathrm{Rg}(\mathcal{F}_2^*), \tag{44}$$

$$\lambda(\mathcal{F}_1 \otimes \mathcal{F}_2) = \lambda(\mathcal{F}_1) \otimes \lambda(\mathcal{F}_2), \tag{45}$$

where the last equality is understood up to a permutation of the entries of the shape vectors.

Proof. That $\mathcal{F}_1 \otimes \mathcal{F}_2$ is idempotent follows from $(\mathcal{F}_1 \otimes \mathcal{F}_2) \circ (\mathcal{F}_1 \otimes \mathcal{F}_2) = (\mathcal{F}_1 \circ \mathcal{F}_1) \otimes (\mathcal{F}_2 \circ \mathcal{F}_2) = \mathcal{F}_1 \otimes \mathcal{F}_2$.

We first show that, for all linear maps $\Phi : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_1')$ and $\Psi : \mathcal{L}(\mathcal{H}_2) \to \mathcal{L}(\mathcal{H}_2')$,

$$\mathrm{Rg}(\Phi \otimes \Psi) = \mathrm{Rg}(\Phi) \otimes \mathrm{Rg}(\Psi), \tag{46}$$

which gives eq. (44) by taking $\Phi = \mathcal{F}_1^*$ and $\Psi = \mathcal{F}_2^*$. For the inclusion from right to left, $\mathrm{Rg}(\Phi) \otimes \mathrm{Rg}(\Psi)$ is spanned by the product elements $y_1 \otimes y_2$ with $y_1 = \Phi(x_1)$ and $y_2 = \Psi(x_2)$. Then $y_1 \otimes y_2 = (\Phi \otimes \Psi)(x_1 \otimes x_2) \in \mathrm{Rg}(\Phi \otimes \Psi)$. Conversely, every $x \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) = \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2)$ can be written $x = \sum_{j \in [J]} x_1^{(j)} \otimes x_2^{(j)}$ with $x_i^{(j)} \in \mathcal{L}(\mathcal{H}_i)$, $i \in [2]$, so that

$$(\Phi \otimes \Psi)(x) = \sum_{j \in [J]} \Phi(x_1^{(j)}) \otimes \Psi(x_2^{(j)}) \in \mathrm{Rg}(\Phi) \otimes \mathrm{Rg}(\Psi).$$

It remains to prove eq. (45). By Proposition 1, for $i \in [2]$ there are integers K_i, dimensions $(d_k^{(i)})_{k \in [K_i]}$ and states $(\rho_k^{(i)})_{k \in [K_i]}$ such that $\lambda(\mathcal{F}_i) = (d_k^{(i)})_{k \in [K_i]}$ up to the order of the entries and

$$\mathrm{Rg}(\mathcal{F}_i) = \Big(\bigoplus_{k \in [K_i]} \mathcal{L}(\mathcal{H}_{k,1}^{(i)}) \otimes \rho_k^{(i)} \Big) \oplus 0, \qquad d_k^{(i)} = \dim(\mathcal{H}_{k,1}^{(i)}).$$

Applying eq. (46) to $\Phi = \mathcal{F}_1$ and $\Psi = \mathcal{F}_2$ and distributing the tensor product over the direct sums, we get

$$\mathrm{Rg}(\mathcal{F}_1 \otimes \mathcal{F}_2) = \Big(\bigoplus_{(k,l) \in [K_1] \times [K_2]} \mathcal{L}(\mathcal{H}_{k,1}^{(1)} \otimes \mathcal{H}_{l,1}^{(2)}) \otimes (\rho_k^{(1)} \otimes \rho_l^{(2)}) \Big) \oplus 0,$$

where we used $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}') = \mathcal{L}(\mathcal{H} \otimes \mathcal{H}')$ and the fact that $\rho_k^{(1)} \otimes \rho_l^{(2)}$ is a state. This is a decomposition of $\mathrm{Rg}(\mathcal{F}_1 \otimes \mathcal{F}_2)$ of the form of eq. (15), indexed by the pairs (k, l), with blocks of dimensions $\dim(\mathcal{H}_{k,1}^{(1)} \otimes \mathcal{H}_{l,1}^{(2)}) = d_k^{(1)} d_l^{(2)}$. By the definition of the shape vector, $\lambda(\mathcal{F}_1 \otimes \mathcal{F}_2)$ is the family $(d_k^{(1)} d_l^{(2)})_{(k,l)}$ sorted in non-increasing order, that is $\lambda(\mathcal{F}_1) \otimes \lambda(\mathcal{F}_2)$. This identification does not depend on the chosen decomposition: by eq. (20), the shape vector of an idempotent channel is that of the $*$-algebra $\mathrm{Rg}(\widehat{\mathcal{F}}^*)$, which is unique up to the order of its entries, and, since $\widehat{\mathcal{F}_1 \otimes \mathcal{F}_2} = \widehat{\mathcal{F}}_1 \otimes \widehat{\mathcal{F}}_2$, eq. (44) gives $\mathrm{Rg}((\widehat{\mathcal{F}_1 \otimes \mathcal{F}_2})^*) = \mathrm{Rg}(\widehat{\mathcal{F}}_1^*) \otimes \mathrm{Rg}(\widehat{\mathcal{F}}_2^*)$, whose shape vector is $\lambda(\mathcal{F}_1) \otimes \lambda(\mathcal{F}_2)$. □

Hence, for $k \in \mathbb{N}_{>0}$, $\lambda(\mathcal{F}^{\otimes k}) = \lambda(\mathcal{F})^{\otimes k}$. Moreover, for block pinchings, we directly have the stability under tensor products: identifying $\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_{\mathbf{d}'}$ with $\mathcal{H}_{\mathbf{d} \otimes \mathbf{d}'}$, we have $\Pi_{\mathbf{d}} \otimes \Pi_{\mathbf{d}'} = \Pi_{\mathbf{d} \otimes \mathbf{d}'}$, so that $\Pi_{\mathbf{f}}^{\otimes k} = \Pi_{\lambda(\mathcal{F}^{\otimes k})}$.

We end this section with the following remark on the origin of the concept of shape vectors.

Remark 6. Note that the shape of an idempotent channel $\mathcal{F}$ coincides with the shape of the $*$-algebra $\mathrm{Rg}(\widehat{\mathcal{F}}^*)$, with $\widehat{\mathcal{F}}$ the reduced channel associated to $\mathcal{F}$. The notion of shape vector for $*$-algebras was introduced by Kuperberg in [14]. If $\mathcal{A}$ is a $*$-algebra, we can decompose

$$\mathcal{A} \cong \bigoplus_{k=1}^{K} \mathcal{L}(\mathbb{C}^{d_k}),$$

for a set of integers $\{d_k \mid k \in [K]\}$ and we write $\lambda(\mathcal{A})$ for the shape vector of $\mathcal{A}$, whose coordinates are the integers $(d_k)_{k\in[K]}$ sorted in non-increasing order[1].

4 Unassisted emulation of idempotent channels with zero error

In this section and the next one, we prove item (i) of Theorem A, that is, eq. (1), together with the fact that, in both regimes, the emulation capacity coincides with the zero-error emulation capacity, i.e. with the supremum of the ratios k/n for which $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$. Throughout, we assume that $\lambda(\mathcal{F}) \neq (1)$: otherwise, $\mathcal{F}$ is essentially a replacer channel (see Proposition 1), and it is simple to see that it can be emulated exactly by any other idempotent channel $\mathcal{G}$ with an arbitrarily large rate. When $\lambda(\mathcal{F}) = (1, \dots, 1)$, i.e. when $\|\lambda(\mathcal{F})\|_\infty = 1$, the denominator in eq. (1) vanishes at $p = +\infty$, and the term $p = +\infty$ is omitted from the infimum, as in the last row of Table 2. When $\mathcal{F}$ or $\mathcal{G}$ is either the identity channel or the completely dephasing channel, it suffices to restrict the infimum to $p \in \{1, +\infty\}$ (see Table 2 below):

$$C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) = \min_{p\in\{1,+\infty\}} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}. \tag{47}$$

On the right-hand side of eq. (1), the two terms $\lambda(\mathcal{F})$ and $\lambda(\mathcal{G})$ are the shape vectors of Proposition 1. A remarkable aspect of this result is that it is a single-letter formula, i.e. the expression does not involve a limit. This comes from the fact that the norm of the shape vector is multiplicative under tensor product.

Now that we have presented the decomposition properties of idempotent channels, we move on to proving the achievability part of item (i) of Theorem A in this section. Formally, we prove that every rate below

$$\inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} \tag{48}$$

is achievable in the unassisted regime, and even with zero error. This yields the bounds

$$\inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} \leq C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) \leq C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}). \tag{49}$$

We show in Section 5 that any rate strictly greater than eq. (48) is not achievable, as a consequence of item (ii) of Theorem A.

Before presenting these proofs, we provide some examples of the expression in eq. (48). In particular, we gather in Table 2 the expression of the emulation capacities $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F})$ when $\mathcal{F}$ or $\mathcal{G}$ is an identity quantum channel or a completely dephasing channel. In fact, when $\mathcal{F} = \mathrm{Id}_d$, then $\lambda(\mathcal{F}) = (d)$ and thus $\|\lambda(\mathcal{F})\|_p = d$ for all $p \geq 1$. As $\|\lambda(\mathcal{G})\|_p$ is non-increasing in p, this shows that $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = \frac{\log(\|\lambda(\mathcal{G})\|_\infty)}{\log(d)}$. The other cases are similar. Note that, when $\mathcal{G} = \Delta_d$, $\log(\|\lambda(\Delta_d)\|_\infty) = \log(1) = 0$. Therefore, in Table 2, we split the case $\mathcal{G} = \Delta_d$ into two subcases: the one where $\mathcal{F}$ is any idempotent channel such that at least one coordinate of $\lambda(\mathcal{F})$ is greater than or equal to 2, and the one where $\lambda(\mathcal{F}) = (1, \dots, 1)$, in which case, for all $p \in [1, +\infty)$,

$$\frac{\log(\|\lambda(\Delta_d)\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} = \frac{p\log(\|\lambda(\Delta_d)\|_p)}{p\log(\|\lambda(\mathcal{F})\|_p)} = \frac{\log(d)}{\log(\|\lambda(\mathcal{F})\|_1)}.$$

Note that in the first case, i.e. when $\lambda(\mathcal{F})$ has a coefficient greater than or equal to two, the emulation capacity of $\mathcal{F}$ by the completely dephasing channel is zero, that is it is impossible to emulate a channel preserving some quantum information with a classical one, without error.

[1] Throughout this article and in Kuperberg's work [14], the order with respect to which the coordinates of shape vectors are sorted is actually irrelevant.

$\mathcal{F}$	$\mathcal{G}$	$C_{\mathrm{UA}}(\mathcal{G}\mapsto\mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G}\mapsto\mathcal{F})$
Id_d	$\mathcal{G}$	$\frac{\log(\|\lambda(\mathcal{G})\|_\infty)}{\log(d)}$
$\mathcal{F}$	Id_d	$\frac{\log(d)}{\log(\|\lambda(\mathcal{F})\|_1)}$
Δ_d	$\mathcal{G}$	$\frac{\log(\|\lambda(\mathcal{G})\|_1)}{\log(d)}$
$\mathcal{F}$ s.t. $\lambda(\mathcal{F})\neq(1,\dots,1)$	Δ_d	0
$\mathcal{F}$ s.t. $\lambda(\mathcal{F})=(1,\dots,1)$	Δ_d	$\frac{\log(d)}{\log(\|\lambda(\mathcal{F})\|_1)}$

Table 2: Emulation capacities when either $\mathcal{F}$ or $\mathcal{G}$ is the identity or the completely dephasing channel (with $d\geq 2$ and $\lambda(\mathcal{F})\neq(1)$).

Then, we also note that, as a corollary of item (i) of Theorem A, the emulation capacity of idempotent channels is, in general, superadditive under tensor products of source channels. However, since the infimum in eq. (1) collapses to a single value of p when the target channel is either a quantum or a classical identity, the quantum and classical capacities of idempotent channels are additive.

Corollary 7. *Let $\mathcal{F}$, $\mathcal{G}_1$ and $\mathcal{G}_2$ be three idempotent channels, and* $\mathrm{R}\in\{\mathrm{UA},\mathrm{SR}\}$*, then*

$$C_{\mathrm{R}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathcal{F})\geq C_{\mathrm{R}}(\mathcal{G}_1\mapsto\mathcal{F})+C_{\mathrm{R}}(\mathcal{G}_2\mapsto\mathcal{F}). \tag{50}$$

However, when $\mathcal{F}$ is either the quantum or the classical identity, i.e. $\mathcal{F}\in\{\mathrm{Id}_d,\Delta_d\}$, then the emulation capacity becomes additive:

$$C_{\mathrm{R}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathcal{F})=C_{\mathrm{R}}(\mathcal{G}_1\mapsto\mathcal{F})+C_{\mathrm{R}}(\mathcal{G}_2\mapsto\mathcal{F}). \tag{51}$$

Proof. If $\lambda(\mathcal{F})=(1)$, all the capacities in the statement are infinite, so the claims are immediate. We therefore assume that $\lambda(\mathcal{F})\neq(1)$. By item (i) of Theorem A, we have, for $\mathrm{R}\in\{\mathrm{UA},\mathrm{SR}\}$,

$$\begin{aligned}
C_{\mathrm{R}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathcal{F}) &= \inf_{p\in[1,+\infty]}\frac{\log(\|\lambda(\mathcal{G}_1\otimes\mathcal{G}_2)\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}\\
&\geq \inf_{p\in[1,+\infty]}\frac{\log(\|\lambda(\mathcal{G}_1)\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}+\inf_{p\in[1,+\infty]}\frac{\log(\|\lambda(\mathcal{G}_2)\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}\\
&= C_{\mathrm{R}}(\mathcal{G}_1\mapsto\mathcal{F})+C_{\mathrm{R}}(\mathcal{G}_2\mapsto\mathcal{F}),
\end{aligned}$$

where we used Proposition 7 and the multiplicativity of the ℓ_p-norms over the tensor product in the inequality.

However, when $\mathcal{F}=\mathrm{Id}_d$, as per Table 2,

$$\begin{aligned}
C_{\mathrm{R}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathrm{Id}_d) &= \frac{\log(\|\lambda(\mathcal{G}_1\otimes\mathcal{G}_2)\|_\infty)}{\log(d)}\\
&= \frac{\log(\|\lambda(\mathcal{G}_1)\|_\infty)}{\log(d)}+\frac{\log(\|\lambda(\mathcal{G}_2)\|_\infty)}{\log(d)}\\
&= C_{\mathrm{R}}(\mathcal{G}_1\mapsto\mathrm{Id}_d)+C_{\mathrm{R}}(\mathcal{G}_2\mapsto\mathrm{Id}_d).
\end{aligned}$$

When $\mathcal{F}=\Delta_d$, the same line of argument holds replacing the ℓ_∞ norm by the ℓ_1 norm. □

The following example shows that the superadditivity of Corollary 7 can be strict in general.

Example 8 (Complementary source channels)**.** Let $\mathcal{G}_1=\mathrm{Id}_2$, $\mathcal{G}_2=\Delta_2$ and $\mathcal{F}=\mathcal{G}_1\otimes\mathcal{G}_2$. By Proposition 7,

$$\lambda(\mathcal{G}_1)=(2),\qquad \lambda(\mathcal{G}_2)=(1,1),\qquad \lambda(\mathcal{F})=\lambda(\mathcal{G}_1)\otimes\lambda(\mathcal{G}_2)=(2,2).$$

By the second row of Table 2, $C_{\mathrm{UA}}(\mathcal{G}_1\mapsto\mathcal{F})=\log(2)/\log(\|(2,2)\|_1)=1/2$, and by its fourth row, $C_{\mathrm{UA}}(\mathcal{G}_2\mapsto\mathcal{F})=0$, since $\lambda(\mathcal{F})$ has an entry equal to 2. However, $\lambda(\mathcal{G}_1\otimes\mathcal{G}_2)=\lambda(\mathcal{F})$, so that $C_{\mathrm{UA}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathcal{F})=1$. Hence the inequality of Corollary 7 is strict here:

$$C_{\mathrm{UA}}(\mathcal{G}_1\otimes\mathcal{G}_2\mapsto\mathcal{F})=1>\frac{1}{2}=C_{\mathrm{UA}}(\mathcal{G}_1\mapsto\mathcal{F})+C_{\mathrm{UA}}(\mathcal{G}_2\mapsto\mathcal{F}).$$

The two source channels are complementary: $\mathcal{G}_1$ carries the quantum part and $\mathcal{G}_2$ the classical part of the hybrid target channel $\mathcal{F}$, and neither of them can emulate the target efficiently without entanglement. Indeed, without entanglement, $\mathcal{G}_1$ can only reach the rate $1/2$, while $\mathcal{G}_2$ is even useless, as it cannot emulate $\mathcal{F}$ at any positive rate; it is only in combination with $\mathcal{G}_1$ that it contributes. We will see in Example 46 that shared entanglement restores additivity for these channels.

The following example shows that the emulation capacity is not *reversible* in the unassisted and shared-randomness regimes, i.e. that $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})C_{\mathrm{R}}(\mathcal{F} \mapsto \mathcal{G}) \neq 1$ in general, for $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}\}$, and that the failure of reversibility can be as extreme as possible.

Example 9 (Emulation of a quantum channel by a classical one)**.** Let $\mathcal{F} = \mathrm{Id}_2$ be the identity channel on $\mathcal{L}(\mathbb{C}^2)$ and $\mathcal{G} = \Delta_2$ be the completely dephasing channel on $\mathcal{L}(\mathbb{C}^2)$, so that $\lambda(\mathcal{F}) = (2)$ and $\lambda(\mathcal{G}) = (1, 1)$. Since $\lambda(\mathcal{F})$ has a coordinate equal to 2, the fourth row of Table 2 gives

$$C_{\mathrm{UA}}(\Delta_2 \mapsto \mathrm{Id}_2) = C_{\mathrm{SR}}(\Delta_2 \mapsto \mathrm{Id}_2) = 0,$$

that is, a noiseless qubit channel cannot be emulated with zero error by a classical channel, even with shared randomness and at an arbitrarily small rate. In the other direction, by the second row of Table 2,

$$C_{\mathrm{UA}}(\mathrm{Id}_2 \mapsto \Delta_2) = \frac{\log(2)}{\log(\|\lambda(\Delta_2)\|_1)} = \frac{\log(2)}{\log(2)} = 1,$$

since one use of a noiseless qubit channel transmits one bit. Therefore

$$C_{\mathrm{UA}}(\Delta_2 \mapsto \mathrm{Id}_2)\, C_{\mathrm{UA}}(\mathrm{Id}_2 \mapsto \Delta_2) = 0,$$

which is the most extreme failure of reversibility possible: not only is $C_{\mathrm{R}}(\mathcal{F} \mapsto \mathcal{G})$ different from $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})^{-1}$, but $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F})^{-1}$ is not even defined, as $C_{\mathrm{R}}(\mathcal{G} \mapsto \mathcal{F}) = 0$. We will see in Example 45 that shared entanglement restores reversibility here, the two capacities becoming $1/2$ and 2.

Although the infimum in eq. (1) is reached for $p \in \{1, +\infty\}$ when either $\mathcal{F}$ or $\mathcal{G}$ is an identity channel or a completely dephasing one, we show in the following example that this is not the case in general.

In particular, the fact that the infimum in eq. (1) can be reached for $p \notin \{1, +\infty\}$ implies that, for general idempotent channels $\mathcal{F}$ and $\mathcal{G}$, the chain of emulation $\mathcal{G} \mapsto \mathrm{Id}_d \mapsto \mathcal{F}$ (where each arrow represents an emulation) is less efficient than directly emulating $\mathcal{F}$ by $\mathcal{G}$, even when shared randomness is allowed.

Example 10. Let $\mathcal{F} : \mathcal{L}(\mathbb{C}^8) \to \mathcal{L}(\mathbb{C}^8)$ and $\mathcal{G} : \mathcal{L}(\mathbb{C}^{15}) \to \mathcal{L}(\mathbb{C}^{15})$ be two idempotent channels such that $\lambda(\mathcal{F}) = (5, 3)$ and $\lambda(\mathcal{G}) = (10, 3, 1, 1)$. Numerically, we find that the infimum

$$C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{p \in [1, +\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} \approx 1.29916,$$

is reached for $p \approx 1.15401$. In contrast, the infimum

$$C_{\mathrm{UA}}(\mathcal{F} \mapsto \mathcal{G}) = \inf_{p \in [1, +\infty]} \frac{\log(\|\lambda(\mathcal{F})\|_p)}{\log(\|\lambda(\mathcal{G})\|_p)} = \frac{\log(5)}{\log(10)},$$

is reached for $p = +\infty$.

4.1 Achievability in item (i) of Theorem A

To prove the achievability direction of item (i) of Theorem A, we first show the following proposition in the case where only one copy of $\mathcal{F}$ is emulated for each copy of $\mathcal{G}$ used; we will then use Proposition 7 on the compatibility of the shape vector with the tensor product to conclude.

Proposition 8. *Let $\mathcal{F}$ and $\mathcal{G}$ be two idempotent channels such that, for all $p \in [1, +\infty]$,*

$$\|\lambda(\mathcal{F})\|_p < \|\lambda(\mathcal{G})\|_p.$$

Then, there exists $n \in \mathbb{N}_{>0}$, $\mathcal{E}$ and $\mathcal{D}$ two channels such that

$$\mathcal{F}^{\otimes n} = \mathcal{D}\mathcal{G}^{\otimes n}\mathcal{E}.$$

To prove this proposition, we use the following theorem on embeddability of $*$-algebras.

Theorem 11 (Proven in [14])**.** *If $\mathcal{A}$ and $\mathcal{B}$ are two finite-dimensional $*$-algebras such that, for all $p \in [1, +\infty]$,*

$$\|\lambda(\mathcal{A})\|_p < \|\lambda(\mathcal{B})\|_p, \tag{52}$$

then, there is $n \in \mathbb{N}_{>0}$ such that there exists a subunital injective $*$-homomorphism $\iota : \mathcal{A}^{\otimes n} \to \mathcal{B}^{\otimes n}$.

Proof. This theorem, despite not being stated explicitly in [14], is proven in the proof of [14, Theorem 1.1]. □

To connect the problem of embedding $*$-algebras with that of interconverting idempotent channels, we use tracial conditional expectations. Given $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_\mathcal{A})$ a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{A})$ sharing its identity operator with $\mathcal{L}(\mathcal{H}_\mathcal{A})$, we can write the orthogonal direct sum decompositions

$$\mathcal{H}_\mathcal{A} = \bigoplus_{k\in[K]} \mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}, \quad \mathcal{A} = \bigoplus_{k\in[K]} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \mathbb{1}_{m_k},$$

with $m_k = \dim(\mathcal{H}_{k,2})$ for $k \in [K]$, and P_k the orthogonal projector onto $\mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2}$.

Definition 12 (Tracial conditional expectation, see e.g. [15, Proposition 1.5])**.** With the notation above, the *tracial conditional expectation onto* $\mathcal{A}$ is the map

$$E_\mathcal{A} : \mathcal{L}(\mathcal{H}_\mathcal{A}) \to \mathcal{A}, \qquad E_\mathcal{A}(x) = \sum_{k\in[K]} \mathrm{tr}_{k,2}(P_k x P_k) \otimes \frac{1}{m_k} \mathbb{1}_{m_k}. \tag{53}$$

It is unital, completely positive, and satisfies $E_\mathcal{A}(x) = x$ for every $x \in \mathcal{A}$.

Given an idempotent channel $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$, we write $E_\mathcal{F}$ for the tracial conditional expectation onto the $*$-algebra $\mathrm{Rg}(\widehat{\mathcal{F}}^*)$, where $\widehat{\mathcal{F}} : \mathcal{L}(\mathcal{H}_0) \to \mathcal{L}(\mathcal{H}_0)$ is the reduced channel associated to $\mathcal{F}$. This is well-defined: by Proposition 1, $\mathrm{Rg}(\widehat{\mathcal{F}}^*)$ is a unital $*$-subalgebra of $\mathcal{L}(\mathcal{H}_0)$ which shares its identity operator with $\mathcal{L}(\mathcal{H}_0)$.

The tracial conditional expectation onto $\mathcal{A}$ is moreover an idempotent channel whose shape vector is that of $\mathcal{A}$.

Proposition 9. *Let $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_\mathcal{A})$ be a unital $*$-subalgebra whose identity operator coincides with that of $\mathcal{L}(\mathcal{H}_\mathcal{A})$. Then $E_\mathcal{A}$ is an idempotent channel and $\lambda(E_\mathcal{A}) = \lambda(\mathcal{A})$. In particular, for every idempotent channel $\mathcal{F}$, the map $E_\mathcal{F}$ is an idempotent channel with $\lambda(E_\mathcal{F}) = \lambda(\mathcal{F})$.*

Proof. $E_\mathcal{A}$ is completely positive and trace preserving by eq. (53), and $E_\mathcal{A}(x) = x$ for every $x \in \mathcal{A}$, so that $E_\mathcal{A} \circ E_\mathcal{A} = E_\mathcal{A}$ since $\mathrm{Rg}(E_\mathcal{A}) \subseteq \mathcal{A}$. Moreover, with the notation above,

$$\mathrm{Rg}(E_\mathcal{A}) = \bigoplus_{k\in[K]} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \frac{\mathbb{1}_{m_k}}{m_k},$$

which is a decomposition of the form of eq. (15) with the states $\rho_k = \mathbb{1}_{m_k}/m_k$ and blocks of dimensions $\dim(\mathcal{H}_{k,1})$. Hence $\lambda(E_\mathcal{A}) = (\dim(\mathcal{H}_{k,1}))_{k\in[K]} = \lambda(\mathcal{A})$. The last statement follows by applying this to $\mathcal{A} = \mathrm{Rg}(\widehat{\mathcal{F}}^*)$, whose shape vector is $\lambda(\mathcal{F})$ by eq. (20) and by the definition of the shape vector of an idempotent channel. □

Since $\lambda(E_\mathcal{F}) = \lambda(\mathcal{F})$ and $\lambda(E_\mathcal{G}) = \lambda(\mathcal{G})$ for all idempotent channels $\mathcal{F}$ and $\mathcal{G}$, by Proposition 4 we can interconvert exactly $\mathcal{F}$ with $E_\mathcal{F}$, $\mathcal{G}$ with $E_\mathcal{G}$ and vice versa, so that $\mathcal{F} = \mathcal{D}\mathcal{G}\mathcal{E}$ for channels $\mathcal{D}, \mathcal{E}$ if and only if $E_\mathcal{F} = \mathcal{D}' E_\mathcal{G} \mathcal{E}'$. We may therefore replace $\mathcal{F}$ and $\mathcal{G}$ by the tracial conditional expectations $E_\mathcal{F}$ and $E_\mathcal{G}$, whose ranges are the $*$-algebras on which Theorem 11 operates directly.

We now lift the $*$-homomorphism ι of Theorem 11 into a unital completely positive map.

Lemma 13. *Let $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_\mathcal{A})$, $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_\mathcal{B})$ be two finite-dimensional unital $*$-subalgebras of $\mathcal{L}(\mathcal{H}_\mathcal{A})$ and $\mathcal{L}(\mathcal{H}_\mathcal{B})$ respectively, where we assume that the identity operator in $\mathcal{A}$ coincides with the one in $\mathcal{L}(\mathcal{H}_\mathcal{A})$, denoted $\mathbb{1}_\mathcal{A}$ and the one in $\mathcal{B}$ coincides with the one in $\mathcal{L}(\mathcal{H}_\mathcal{B})$, denoted $\mathbb{1}_\mathcal{B}$. Let $\iota : \mathcal{A} \to \mathcal{B}$ be an injective subunital $*$-homomorphism.*

Then, writing $\mathcal{H}_{\mathcal{B}}^{(0)} = \mathrm{supp}(\iota(\mathbb{1}_{\mathcal{A}})) \subseteq \mathcal{H}_{\mathcal{B}}$, *we have that* $P_\iota = \iota(\mathbb{1}_{\mathcal{A}}) \in \mathcal{L}(\mathcal{H}_{\mathcal{B}})$ *is the orthogonal projector onto* $\mathcal{H}_{\mathcal{B}}^{(0)}$. *We denote by* $V : \mathcal{H}_{\mathcal{B}}^{(0)} \to \mathcal{H}_{\mathcal{B}}$ *the isometry embedding* $\mathcal{H}_{\mathcal{B}}^{(0)}$ *into* $\mathcal{H}_{\mathcal{B}}$ *such that* $VV^* = P_\iota$ *and we define the* reduced $*$-homomorphism $\widehat{\iota} : \mathcal{A} \to \mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$ *by*

$$\widehat{\iota} : x \mapsto V^* \iota(x) V. \tag{54}$$

Then $\widehat{\iota} : \mathcal{A} \to \mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$ *is a* unital *injective* $*$*-homomorphism and* $\widehat{\iota}(\mathcal{A})$ *is a unital* $*$*-subalgebra of* $\mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$ *where both* $\widehat{\iota}(\mathcal{A})$ *and* $\mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$ *share the same identity operator. We will write* $\widehat{\iota}^{-1} : \widehat{\iota}(\mathcal{A}) \to \mathcal{A}$ *for the inverse of* $\widehat{\iota}$, *which is itself a unital injective* $*$*-homomorphism.*

We write $E_{\mathcal{A}} : \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \to \mathcal{A}$ *and* $E_{\widehat{\iota}(\mathcal{A})} : \mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)}) \to \widehat{\iota}(\mathcal{A})$ *for the tracial conditional expectations on* $\mathcal{A}$ *and* $\widehat{\iota}(\mathcal{A})$ *respectively. Then, the maps*

$$\begin{aligned} \widetilde{\iota} : \mathcal{L}(\mathcal{H}_{\mathcal{A}}) &\to \mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_{\mathcal{B}}) \\ x &\mapsto \iota(E_{\mathcal{A}}(x)) + \frac{\mathrm{tr}(x)}{\mathrm{tr}(\mathbb{1}_{\mathcal{A}})}(\mathbb{1}_{\mathcal{B}} - P_\iota), \end{aligned} \tag{55}$$

and

$$\begin{aligned} \widetilde{\iota}^{-1} : \mathcal{L}(\mathcal{H}_{\mathcal{B}}) &\to \mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_{\mathcal{A}}) \\ x &\mapsto \widehat{\iota}^{-1}(E_{\widehat{\iota}(\mathcal{A})}(V^* x V)) \end{aligned} \tag{56}$$

are unital completely positive and satisfy for all $x \in \mathcal{A}$,

$$\widetilde{\iota}^{-1} \circ \widetilde{\iota}(x) = x. \tag{57}$$

Proof. First, we prove that $P_\iota = \iota(\mathbb{1}_{\mathcal{A}})$ is an orthogonal projector on $\mathrm{supp}(\iota(\mathbb{1}_{\mathcal{A}}))$. As ι is a $*$-homomorphism, it preserves both the operation of taking the adjoint and the multiplication, so that $P_\iota^* = \iota(\mathbb{1}_{\mathcal{A}})^* = \iota(\mathbb{1}_{\mathcal{A}}^*) = P_\iota$ and $P_\iota^2 = \iota(\mathbb{1}_{\mathcal{A}})^2 = \iota(\mathbb{1}_{\mathcal{A}}^2) = P_\iota$.

We move on to prove that $\widehat{\iota} : \mathcal{A} \to \mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$, with $\mathcal{H}_{\mathcal{B}}^{(0)} = \mathrm{supp}(\iota(\mathbb{1}_{\mathcal{A}}))$ is a $*$-homomorphism, which is also unital and injective. First $\widehat{\iota}$ is clearly linear. Let $x \in \mathcal{A}$,

$$\widehat{\iota}(x)^* = (V^* \iota(x) V)^* = V^* \iota(x)^* V = V^* \iota(x^*) V = \widehat{\iota}(x^*),$$

and for $y \in \mathcal{A}$,

$$\widehat{\iota}(xy) = V^* \iota(xy) V = V^* \iota(x) \iota(y) V \overset{(a)}{=} V^* V V^* \iota(x) V V^* \iota(y) V = \widehat{\iota}(x)\widehat{\iota}(y).$$

The equality (a) holds by applying Lemma 1 on the map $\Phi = \iota$ which is positive since it is a $*$-homomorphism[2]. To show the unitality, we have

$$\widehat{\iota}(\mathbb{1}_{\mathcal{A}}) = V^* \iota(\mathbb{1}_{\mathcal{A}}) V = V^* P_\iota V = V^* V V^* V = \mathbb{1}_{\mathcal{H}_{\mathcal{B}}^{(0)}},$$

with $\mathbb{1}_{\mathcal{H}_{\mathcal{B}}^{(0)}}$ the identity operator in $\mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$. We finish with the injectivity property. Let $x \in \ker(\widehat{\iota})$,

$$0 = \widehat{\iota}(x) = V^* \iota(x) V.$$

Hence, by Lemma 1

$$\iota(x) = V V^* \iota(x) V V^* = 0,$$

so that $x \in \ker(\iota)$, thus $x = 0$ by injectivity of ι.

Since $\widehat{\iota}$ is a unital $*$-homomorphism, $\widehat{\iota}(\mathcal{A})$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$ and has the same identity operator as $\mathcal{L}(\mathcal{H}_{\mathcal{B}}^{(0)})$, so that the tracial conditional expectation $E_{\widehat{\iota}(\mathcal{A})}$ is well-defined by eq. (53). Furthermore, as $\widehat{\iota}$ is injective, $\widehat{\iota}^{-1}$ is well-defined and is a unital $*$-homomorphism as well.

Now, we show that $\widetilde{\iota}$ and $\widetilde{\iota}^{-1}$ are unital completely positive maps. By unitality of $E_{\mathcal{A}}$, $E_{\widehat{\iota}(\mathcal{A})}$ and $\widehat{\iota}^{-1}$, it is clear that $\widetilde{\iota}$ and $\widetilde{\iota}^{-1}$ are unital. Furthermore, $\widetilde{\iota}^{-1}$ is the composition of completely positive maps and is thus completely positive. Moreover, as ι was assumed to be subunital, $\iota(\mathbb{1}_{\mathcal{A}}) = P_\iota \leq \mathbb{1}_{\mathcal{B}}$ so that $\widetilde{\iota}$ is a sum of two completely positive maps and is thus completely positive itself.

[2] Actually, $*$-homomorphisms are even completely positive [30].

We also have that $\widetilde{\iota}(\mathcal{L}(\mathcal{H}_{\mathcal{A}})) \subseteq \mathcal{B}$. Indeed, as $\mathcal{B}$ is unital of unit $\mathbb{1}_{\mathcal{B}}$ and as the range of ι is in $\mathcal{B}$, $\mathbb{1}_{\mathcal{B}} - P_\iota \in \mathcal{B}$ and $\iota(E_{\mathcal{A}}(x)) \in \mathcal{B}$ for all $x \in \mathcal{L}(\mathcal{H}_{\mathcal{A}})$.

Finally, note that $\mathbb{1}_{\mathcal{B}} - P_\iota$ is the orthogonal projector on $\ker(P_\iota)$ which is orthogonal to $\mathrm{supp}(P_\iota) = \mathrm{Rg}(V)$, so that $(\mathbb{1}_{\mathcal{B}} - P_\iota)V = 0$. Hence, for $x \in \mathcal{A}$,

$$\widetilde{\iota}^{-1}(\widetilde{\iota}(x)) = \widehat{\iota}^{-1}(E_{\widehat{\iota}(\mathcal{A})}(V^*\iota(E_{\mathcal{A}}(x))V)) = \widehat{\iota}^{-1}(E_{\widehat{\iota}(\mathcal{A})}(\widehat{\iota}(x))) = \widehat{\iota}^{-1}(\widehat{\iota}(x)) = x.$$

□

We now prove Proposition 8.

Proof of Proposition 8. By Proposition 4, since $\lambda(\mathcal{F}) = \lambda(E_{\mathcal{F}})$ and $\lambda(\mathcal{G}) = \lambda(E_{\mathcal{G}})$, it suffices to show that there are channels $\mathcal{E}', \mathcal{D}'$ and an integer $n \in \mathbb{N}_{>0}$ such that $E_{\mathcal{F}}^{\otimes n} = \mathcal{D}' E_{\mathcal{G}}^{\otimes n} \mathcal{E}'$.

Let $\mathcal{A} = \mathrm{Rg}(E_{\mathcal{F}})$ and $\mathcal{B} = \mathrm{Rg}(E_{\mathcal{G}})$ respectively seen as unital $*$-subalgebras of $\mathcal{L}(\mathcal{H}_{\widehat{\mathcal{F}}})$ and $\mathcal{L}(\mathcal{H}_{\widehat{\mathcal{G}}})$, where $\mathcal{H}_{\widehat{\mathcal{F}}} = \mathrm{supp}(\mathcal{F}(\mathbb{1}))$ and $\mathcal{H}_{\widehat{\mathcal{G}}} = \mathrm{supp}(\mathcal{G}(\mathbb{1}))$ are the spaces on which the reduced versions of $\mathcal{F}$ and $\mathcal{G}$, respectively $\widehat{\mathcal{F}}$ and $\widehat{\mathcal{G}}$, act. By assumption, $\|\lambda(\mathcal{A})\|_p < \|\lambda(\mathcal{B})\|_p$ for all $p \in [1, +\infty]$, so that Theorem 11 provides an $n \in \mathbb{N}_{>0}$ and an injective subunital $*$-homomorphism $\iota : \mathcal{A}^{\otimes n} \to \mathcal{B}^{\otimes n}$. Let $\widetilde{\iota}$ and $\widetilde{\iota}^{-1}$ be the unital completely positive maps associated to ι by Lemma 13, and set

$$\mathcal{E} = (\widetilde{\iota}^{-1})^* \circ E_{\mathcal{A}}^{\otimes n}, \qquad \mathcal{D} = E_{\mathcal{A}}^{\otimes n} \circ \widetilde{\iota}^*,$$

which are channels, as compositions of channels: the adjoint of a unital completely positive map is completely positive and trace preserving, and $E_{\mathcal{A}}^{\otimes n}$ is a channel.

It remains to check that $\mathcal{D} \circ E_{\mathcal{B}}^{\otimes n} \circ \mathcal{E} = E_{\mathcal{A}}^{\otimes n}$. Taking adjoints, this is equivalent to

$$E_{\mathcal{A}}^{\otimes n} = \mathcal{E}^* \circ E_{\mathcal{B}}^{\otimes n} \circ \mathcal{D}^* = E_{\mathcal{A}}^{\otimes n} \circ \widetilde{\iota}^{-1} \circ E_{\mathcal{B}}^{\otimes n} \circ \widetilde{\iota} \circ E_{\mathcal{A}}^{\otimes n}, \tag{58}$$

where we used that $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ are self-adjoint, being tracial conditional expectations. Let $x \in \mathcal{L}(\mathcal{H}_{\mathcal{A}}^{\otimes n})$. Then $E_{\mathcal{A}}^{\otimes n}(x) \in \mathcal{A}^{\otimes n}$, so that $\widetilde{\iota}(E_{\mathcal{A}}^{\otimes n}(x)) \in \mathcal{B}^{\otimes n}$ by Lemma 13. Since $E_{\mathcal{B}}^{\otimes n}$ acts as the identity on $\mathcal{B}^{\otimes n}$, we get $E_{\mathcal{B}}^{\otimes n}(\widetilde{\iota}(E_{\mathcal{A}}^{\otimes n}(x))) = \widetilde{\iota}(E_{\mathcal{A}}^{\otimes n}(x))$. Applying eq. (57) to $E_{\mathcal{A}}^{\otimes n}(x) \in \mathcal{A}^{\otimes n}$ gives $\widetilde{\iota}^{-1}(\widetilde{\iota}(E_{\mathcal{A}}^{\otimes n}(x))) = E_{\mathcal{A}}^{\otimes n}(x)$, and a final application of $E_{\mathcal{A}}^{\otimes n}$, which fixes $\mathcal{A}^{\otimes n}$, yields eq. (58). □

Then we can prove the achievability in item (i) of Theorem A.

Proof of achievability in item (i) of Theorem A. Let $\mathcal{F}$, $\mathcal{G}$ be idempotent channels. If the infimum in eq. (1) is zero, the desired lower bound follows from the non-negativity of the capacity. We therefore assume that this infimum is positive. Let a rate $r \geq 0$ be such that

$$r < \inf_{p \in [1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}.$$

By density of the rational numbers in the reals, there exist $k, n \in \mathbb{N}_{>0}$ such that

$$r \leq \frac{k}{n} < \inf_{p \in [1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}. \tag{59}$$

If both shape vectors have only unit entries, that is both $\mathcal{F}$ and $\mathcal{G}$ are classical, then $\|\lambda(\mathcal{F})\|_1^k < \|\lambda(\mathcal{G})\|_1^n$, so embedding the corresponding classical alphabets gives $\mathcal{F}^{\otimes k} = \mathcal{D}\mathcal{G}^{\otimes n}\mathcal{E}$ for suitable channels $\mathcal{E}$ and $\mathcal{D}$. Suppose that $\mathcal{F}$ is not a classical channel, so that $\log(\|\lambda(\mathcal{F})\|_p) > 0$ for all $p \in [1, +\infty]$. Then, for every $p \in [1, +\infty]$, we have

$$\|\lambda(\mathcal{F}^{\otimes k})\|_p < \|\lambda(\mathcal{G}^{\otimes n})\|_p,$$

where we used Proposition 7 on the shape vectors of the tensor product of channels. If $\mathcal{F}$ is classical and $\mathcal{G}$ is not, we still have for all $p \in [1, +\infty)$, by eq. (59), $\|\lambda(\mathcal{F}^{\otimes k})\|_p < \|\lambda(\mathcal{G}^{\otimes n})\|_p$. Then, for $p = +\infty$, $\|\lambda(\mathcal{F}^{\otimes k})\|_\infty = 1 < 2 \leq \|\lambda(\mathcal{G}^{\otimes n})\|_\infty$. Hence, we still have in this case $\|\lambda(\mathcal{F}^{\otimes k})\|_p < \|\lambda(\mathcal{G}^{\otimes n})\|_p$ for all $p \in [1, +\infty]$.

Then, by Proposition 8, there is $m \in \mathbb{N}_{>0}$, $\mathcal{E}$ and $\mathcal{D}$ channels such that

$$\mathcal{F}^{\otimes km} = \mathcal{D}\mathcal{G}^{\otimes nm}\mathcal{E},$$

so that $\mathcal{F}^{\otimes km}$ can be emulated exactly by $\mathcal{G}^{\otimes nm}$ without assistance. By Section 2.1, $km/(nm) = k/n \geq r$ is thus an achievable rate, and the zero-error unassisted emulation capacity is at least $k/n \geq r$. □

5 Minimal error for unassisted and shared-randomness-assisted emulation for finite blocklength

In this section, we prove item (ii) of Theorem A, that is, the lower bound of eq. (2) on the minimal error for the emulation of idempotent channels at finite blocklength, both in the unassisted and in the shared randomness regimes, which holds for any two idempotent channels $\mathcal{F}$ and $\mathcal{G}$ and all $k, n \in \mathbb{N}_{>0}$. We then deduce from it the converse part of item (i) and the strong converse of item (iii).

Since, in general, $\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F})$, it actually suffices to show the right-most inequality in eq. (2), that is, for all $\mathcal{F}$ and $\mathcal{G}$ idempotent channels and every pair $k, n \in \mathbb{N}_{>0}$,

$$\delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - \inf_{p \in [1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p^n}{\|\lambda(\mathcal{F})\|_p^k}. \tag{60}$$

We first reduce the problem to the emulation of block pinchings. Let $\mathbf{f} = \lambda(\mathcal{F})$ and $\mathbf{g} = \lambda(\mathcal{G})$ be the shape vectors of $\mathcal{F}$ and $\mathcal{G}$. We write $\Pi_{\mathbf{f}}$ and $\Pi_{\mathbf{g}}$ for the pinchings of shape vectors $\lambda(\Pi_{\mathbf{f}}) = \lambda(\mathcal{F})$ and $\lambda(\Pi_{\mathbf{g}}) = \lambda(\mathcal{G})$, as per Definition 3, so that Corollary 4 gives, in every regime $\mathrm{R} \in \{\mathrm{UA}, \mathrm{SR}, \mathrm{EA}\}$ and for all $k, n \in \mathbb{N}_{>0}$,

$$\delta^{\mathrm{R}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \delta^{\mathrm{R}}_{k,n}(\Pi_{\mathbf{g}} \mapsto \Pi_{\mathbf{f}}). \tag{61}$$

It is therefore enough to prove eq. (60) for the block pinchings $\Pi_{\mathbf{f}}$ and $\Pi_{\mathbf{g}}$, which have the same shape vectors as $\mathcal{F}$ and $\mathcal{G}$. The gain is that $\Pi_{\mathbf{f}}$ and $\Pi_{\mathbf{g}}$ act on the spaces $\mathcal{H}_{\mathbf{f}}$ and $\mathcal{H}_{\mathbf{g}}$ of eq. (28), which carry no multiplicity spaces: every quantity below is expressed directly in terms of the entries of $\lambda(\mathcal{F}) = \mathbf{f}$ and $\lambda(\mathcal{G}) = \mathbf{g}$.

To prove this bound, we use the Holevo–Helstrom bound, applied to the problem of channel discrimination.

Lemma 14 (see e.g. [27, Theorem 3.52])**.** *Let $\Phi_1,\ \Phi_2 : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be two quantum channels. For any choice of auxiliary system $\mathcal{L}(\mathcal{H}'')$ of dimension $d = \dim(\mathcal{H}'')$, for any choice of positive semidefinite operator $\mu \in \mathcal{L}(\mathcal{H}' \otimes \mathcal{H}'')$ satisfying $\mu \leq \mathbb{1}$ and of density operator $\sigma \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}'')$,*

$$\frac{1}{2}\|\Phi_1 - \Phi_2\|_\diamond \geq \operatorname{tr}(\mu \Phi_1 \otimes \mathrm{Id}_d(\sigma)) - \operatorname{tr}(\mu \Phi_2 \otimes \mathrm{Id}_d(\sigma)). \tag{62}$$

Let $\Lambda \in \mathbb{N}_{>0}$, let $(p_\lambda)_{\lambda \in [\Lambda]}$ be a probability distribution and let $\{(\mathcal{D}_\lambda, \mathcal{E}_\lambda) \mid \lambda \in [\Lambda]\}$ be a set of pairs of encoding and decoding channels for the emulation of $\Pi_{\mathbf{f}}$ by $\Pi_{\mathbf{g}}$, with $\mathcal{E}_\lambda : \mathcal{L}(\mathcal{H}_{\mathbf{f}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{g}})$ and $\mathcal{D}_\lambda : \mathcal{L}(\mathcal{H}_{\mathbf{g}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{f}})$. For $\lambda \in [\Lambda]$, we write

$$\mathcal{N}_\lambda = \mathcal{D}_\lambda \circ \Pi_{\mathbf{g}} \circ \mathcal{E}_\lambda \ : \ \mathcal{L}(\mathcal{H}_{\mathbf{f}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{f}}), \tag{63}$$

which is a channel, as a composition of channels, so that $\sum_{\lambda \in [\Lambda]} p_\lambda \mathcal{N}_\lambda$ is a channel as well. Applying Lemma 14 to $\Pi_{\mathbf{f}}$ and to $\sum_{\lambda \in [\Lambda]} p_\lambda \mathcal{N}_\lambda$, the proof of item (ii) of Theorem A then proceeds by constructing ansätze $\mu, \sigma^{(p)}$, for $p \in [1, +\infty)$, such that

$$\frac{1}{2}\Big\|\Pi_{\mathbf{f}} - \sum_{\lambda \in [\Lambda]} p_\lambda \mathcal{D}_\lambda \Pi_{\mathbf{g}} \mathcal{E}_\lambda\Big\|_\diamond \geq \operatorname{tr}(\mu\, \Pi_{\mathbf{f}} \otimes \mathrm{Id}_d(\sigma^{(p)})) - \sum_{\lambda \in [\Lambda]} p_\lambda \operatorname{tr}(\mu\, \mathcal{N}_\lambda \otimes \mathrm{Id}_d(\sigma^{(p)})), \tag{64}$$

with

$$\operatorname{tr}(\mu\, \Pi_{\mathbf{f}} \otimes \mathrm{Id}_d(\sigma^{(p)})) = 1, \qquad \operatorname{tr}(\mu\, \mathcal{N}_\lambda \otimes \mathrm{Id}_d(\sigma^{(p)})) \leq \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}, \tag{65}$$

for every $\lambda \in [\Lambda]$ and $p \in [1, +\infty)$.

The only structural property of the maps $\mathcal{N}_\lambda$ that we shall use is the following Kraus representation, which exhibits the block projectors $\widetilde{P}'_b$ of $\mathcal{H}_{\mathbf{g}}$.

Lemma 15. *For $\lambda \in [\Lambda]$, the channel $\mathcal{N}_\lambda$ of eq. (63) admits the Kraus representation*

$$\mathcal{N}_\lambda(z) = \sum_{a,b,c} \left(D_a^{(\lambda)} \widetilde{P}'_b E_c^{(\lambda)}\right) z \left(D_a^{(\lambda)} \widetilde{P}'_b E_c^{(\lambda)}\right)^*, \tag{66}$$

where $\{\widetilde{P}'_b\}_{b \in [K']}$ are the block projectors of $\mathcal{H}_{\mathbf{g}}$ of Definition 3, $\{D_a^{(\lambda)}\}_a$ is any Kraus representation of $\mathcal{D}_\lambda$ and $\{E_c^{(\lambda)}\}_c$ is any Kraus representation of $\mathcal{E}_\lambda$.

Proof. By eq. (29), the operators $\{\widetilde{P}'_b\}_{b\in[K']}$ form a Kraus representation of $\Pi_{\mathbf{g}}$. Eq. (66) then follows from the factorisation $\mathcal{N}_\lambda = \mathcal{D}_\lambda \circ \Pi_{\mathbf{g}} \circ \mathcal{E}_\lambda$ of eq. (63). □

Now, we introduce the ansätze μ, σ which, as we will show, satisfy eq. (65). In the following, let $\mathcal{H}_R$ denote a copy of $\mathcal{H}_{\mathbf{f}}$, with blocks $\mathcal{H}_k^R \cong \mathbb{C}^{d_k}$ and orthonormal bases $\{|\nu\rangle_k^R\}_{\nu\in[d_k]}$, for $k \in [K]$, where $\mathbf{f} = (d_1, \ldots, d_K)$.

Definition 16. For $k \in [K]$ let

$$|\phi_k\rangle = \frac{1}{\sqrt{d_k}} \sum_{\nu=1}^{d_k} |\nu\rangle_k \otimes |\nu\rangle_k^R \;\in\; \mathbb{C}^{d_k} \otimes \mathcal{H}_k^R \subseteq \mathcal{H}_{\mathbf{f}} \otimes \mathcal{H}_R.$$

Given a probability vector $q = (q_1, \ldots, q_K)$, set

$$\sigma_q = \sum_{k=1}^{K} q_k \, |\phi_k\rangle\langle\phi_k| \;\in\; \mathcal{D}(\mathcal{H}_{\mathbf{f}} \otimes \mathcal{H}_R), \qquad \mu = \sum_{k=1}^{K} |\phi_k\rangle\langle\phi_k| \;\in\; \mathcal{L}(\mathcal{H}_{\mathbf{f}} \otimes \mathcal{H}_R).$$

For $p \in [1, +\infty)$ let $\sigma^{(p)} = \sigma_{q^{(p)}}$ with

$$q_k^{(p)} = \frac{d_k^p}{\sum_{j=1}^K d_j^p} = \frac{d_k^p}{\|\lambda(\mathcal{F})\|_p^p}. \tag{67}$$

We then directly have the following two lemmas.

Lemma 17 (μ is a valid test)**.** $0 \le \mu \le \mathbb{1}_{\mathcal{H}_{\mathbf{f}}\otimes\mathcal{H}_R}$*, since* μ *is an orthogonal projector.*

Proof. The subspaces $\mathbb{C}^{d_k} \otimes \mathcal{H}_k^R$, $k \in [K]$, are mutually orthogonal inside $\mathcal{H}_{\mathbf{f}} \otimes \mathcal{H}_R$. Hence the rank-one projectors $|\phi_k\rangle\langle\phi_k|$ have mutually orthogonal ranges and their sum is an orthogonal projector. □

Lemma 18 (The ideal statistic)**.** *For every probability vector* q, $(\Pi_{\mathbf{f}} \otimes \mathrm{Id})(\sigma_q) = \sigma_q$ *and*

$$\mathrm{tr}\big(\mu\,(\Pi_{\mathbf{f}} \otimes \mathrm{Id})(\sigma_q)\big) = 1. \tag{68}$$

Proof. We can write $|\phi_k\rangle\langle\phi_k| = \frac{1}{d_k}\sum_{\nu,\eta} |\nu\rangle\langle\eta|_k \otimes |\nu\rangle\langle\eta|_k^R$ and $\Pi_{\mathbf{f}}(|\nu\rangle\langle\eta|_k) = |\nu\rangle\langle\eta|_k$ since both $|\nu\rangle_k$ and $|\eta\rangle_k$ lie in the same block $\mathbb{C}^{d_k}$ of $\mathcal{H}_{\mathbf{f}}$. Hence $(\Pi_{\mathbf{f}} \otimes \mathrm{Id})(|\phi_k\rangle\langle\phi_k|) = |\phi_k\rangle\langle\phi_k|$ and $(\Pi_{\mathbf{f}} \otimes \mathrm{Id})(\sigma_q) = \sigma_q$. Since the $|\phi_k\rangle\langle\phi_k|$ are mutually orthogonal rank-one projectors, $\mathrm{tr}(\mu\sigma_q) = \sum_{k,l} q_l \,\mathrm{tr}(|\phi_k\rangle\langle\phi_k||\phi_l\rangle\langle\phi_l|) = \sum_k q_k = 1$. □

To finish the proof of item (ii) of Theorem A, it suffices now to show that, for every $\lambda \in [\Lambda]$ and every $p \in [1, +\infty)$

$$\mathrm{tr}\big(\mu\,(\mathcal{N}_\lambda \otimes \mathrm{Id}_d)(\sigma^{(p)})\big) \le \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}. \tag{69}$$

Bounding the trace $\mathrm{tr}\big(\mu\,(\mathcal{N}_\lambda \otimes \mathrm{Id}_d)(\sigma^{(p)})\big)$ requires more involved mathematical machinery, which we develop in the following lemmas.

Lemma 19 (The simulated statistic as an entanglement fidelity)**.** *Let* $\mathcal{M} : \mathcal{L}(\mathcal{H}_{\mathbf{f}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{f}})$ *be a channel with Kraus representation* $\{K_i\}_i$*. Then, for every probability vector* $q = (q_1, \ldots, q_K)$*,*

$$\mathrm{tr}\big(\mu\,(\mathcal{M} \otimes \mathrm{Id})(\sigma_q)\big) = \sum_{k=1}^{K} \frac{q_k}{d_k^2} \sum_i \big|\mathrm{tr}(\widetilde{P}_k K_i)\big|^2, \tag{70}$$

with $\mathbf{f} = (d_1, \ldots, d_K)$*.*

Proof. Write $|\phi_k\rangle\langle\phi_k| = \frac{1}{d_k}\sum_{\nu,\eta\in[d_k]} |\nu\rangle\langle\eta|_k \otimes |\nu\rangle\langle\eta|_k^R$, so

$$(\mathcal{M} \otimes \mathrm{Id})(\sigma_q) = \sum_k \frac{q_k}{d_k} \sum_{\nu,\eta\in[d_k]} \sum_i K_i|\nu\rangle\langle\eta|_k K_i^* \otimes |\nu\rangle\langle\eta|_k^R, \qquad \mu = \sum_l \frac{1}{d_l} \sum_{\nu',\eta'\in[d_l]} |\nu'\rangle\langle\eta'|_l \otimes |\nu'\rangle\langle\eta'|_l^R.$$

Using $\operatorname{tr}[(X\otimes Y)(X'\otimes Y')]=\operatorname{tr}(XX')\operatorname{tr}(YY')$, the reference-system factor is

$$\operatorname{tr}\big(|\nu'\rangle\!\langle\eta'|_l^R\,|\nu\rangle\!\langle\eta|_k^R\big)=\delta_{kl}\,\delta_{\eta'\nu}\,\delta_{\eta\nu'}.$$

Hence only the terms $l=k$, $\nu'=\eta$, $\eta'=\nu$ survive and

$$\begin{aligned}\operatorname{tr}\big(\mu(\mathcal{M}\otimes\mathrm{Id})(\sigma_q)\big)&=\sum_k\frac{q_k}{d_k^2}\sum_i\sum_{\nu,\eta\in[d_k]}\operatorname{tr}\big(|\eta\rangle\!\langle\nu|_k\,K_i|\nu\rangle\!\langle\eta|_kK_i^*\big)\\&=\sum_k\frac{q_k}{d_k^2}\sum_i\sum_{\nu,\eta\in[d_k]}\langle\nu|_k\,K_i\,|\nu\rangle_k\;\langle\eta|_k\,K_i^*\,|\eta\rangle_k\\&=\sum_k\frac{q_k}{d_k^2}\sum_i\Big(\sum_\nu\langle\nu|_k\,K_i\,|\nu\rangle_k\Big)\overline{\Big(\sum_\eta\langle\eta|_k\,K_i\,|\eta\rangle_k\Big)}=\sum_k\frac{q_k}{d_k^2}\sum_i\big|\operatorname{tr}(\widetilde{P}_kK_i)\big|^2.\end{aligned}$$

□

We can now use the Cauchy–Schwarz inequality to reduce the whole problem to a finite optimisation over two matrices. This method draws inspiration from the proof techniques of [16, Theorem 2.3].

Lemma 20. *Let $\lambda\in[\Lambda]$ and let $\mathcal{N}_\lambda$ be as in eq. (63), with the Kraus representation eq. (66). Define*

$$A_{k,b}^{(\lambda)}=\operatorname{tr}\big(\widetilde{P}_k\,\mathcal{D}_\lambda(\widetilde{P}'_b)\big),\qquad B_{k,b}^{(\lambda)}=\operatorname{tr}\big(\widetilde{P}_k\,\mathcal{E}_\lambda^*(\widetilde{P}'_b)\big),\qquad k\in[K],\ b\in[K'],\tag{71}$$

where $\{\widetilde{P}_k\}_{k\in[K]}$ and $\{\widetilde{P}'_b\}_{b\in[K']}$ are the block projectors of $\mathcal{H}_{\mathbf{f}}$ and $\mathcal{H}_{\mathbf{g}}$ of Definition 3. Then $A_{k,b}^{(\lambda)}\ge0$, $B_{k,b}^{(\lambda)}\ge0$,

$$\sum_{k=1}^{K}A_{k,b}^{(\lambda)}=d'_b\quad\text{for every } b\in[K'],\qquad\sum_{b=1}^{K'}B_{k,b}^{(\lambda)}=d_k\quad\text{for every } k\in[K],\tag{72}$$

and for every probability vector q,

$$\operatorname{tr}\big(\mu\,(\mathcal{N}_\lambda\otimes\mathrm{Id})(\sigma_q)\big)\;\le\;\sum_{k=1}^{K}\sum_{b=1}^{K'}\frac{q_k}{d_k^2}\,A_{k,b}^{(\lambda)}B_{k,b}^{(\lambda)},\tag{73}$$

with $\mathbf{f}=(d_1,\dots,d_K)$ and $\mathbf{g}=(d'_1,\dots,d'_{K'})$.

Proof. Since it is irrelevant in this proof, we omit the parameter λ.

Positivity. $\mathcal{D}$ is CP and $\widetilde{P}'_b\ge0$, so $\mathcal{D}(\widetilde{P}'_b)\ge0$ and $A_{k,b}=\operatorname{tr}(\widetilde{P}_k\mathcal{D}(\widetilde{P}'_b))\ge0$. Likewise $\mathcal{E}^*$ is CP and $B_{k,b}\ge0$.

Constraints. Since $\sum_k\widetilde{P}_k=\mathbb{1}_{\mathcal{H}_{\mathbf{f}}}$ and $\mathcal{D}$ is trace preserving, $\sum_k A_{k,b}=\operatorname{tr}\big(\mathcal{D}(\widetilde{P}'_b)\big)=\operatorname{tr}(\widetilde{P}'_b)=d'_b$. Since $\sum_b\widetilde{P}'_b=\mathbb{1}_{\mathcal{H}_{\mathbf{g}}}$ and $\mathcal{E}^*$ is unital (as $\mathcal{E}$ is a channel), $\sum_b B_{k,b}=\operatorname{tr}\big(\widetilde{P}_k\mathcal{E}^*(\mathbb{1}_{\mathcal{H}_{\mathbf{g}}})\big)=\operatorname{tr}(\widetilde{P}_k)=d_k$.

The bound. By Lemma 19 with $K_i=D_a\widetilde{P}'_bE_c$, $i=(a,b,c)$,

$$\operatorname{tr}\big(\mu(\mathcal{N}\otimes\mathrm{Id})(\sigma_q)\big)=\sum_k\frac{q_k}{d_k^2}\sum_{a,b,c}\big|\operatorname{tr}(\widetilde{P}_kD_a\widetilde{P}'_bE_c)\big|^2.$$

Fix k,a,b,c. Using $\widetilde{P}_k^2=\widetilde{P}_k$, $(\widetilde{P}'_b)^2=\widetilde{P}'_b$ and cyclicity of the trace,

$$\operatorname{tr}(\widetilde{P}_kD_a\widetilde{P}'_bE_c)=\operatorname{tr}\Big(\big(\widetilde{P}'_bD_a^*\widetilde{P}_k\big)^*\,\big(\widetilde{P}'_bE_c\widetilde{P}_k\big)\Big).$$

The Cauchy–Schwarz inequality for the Hilbert–Schmidt inner product, $|\operatorname{tr}(x^*y)|^2\le\operatorname{tr}(x^*x)\operatorname{tr}(y^*y)$, applied with $x=\widetilde{P}'_bD_a^*\widetilde{P}_k$ and $y=\widetilde{P}'_bE_c\widetilde{P}_k$, gives

$$\big|\operatorname{tr}(\widetilde{P}_kD_a\widetilde{P}'_bE_c)\big|^2\le\operatorname{tr}\big(\widetilde{P}_kD_a\widetilde{P}'_bD_a^*\big)\operatorname{tr}\big(\widetilde{P}_kE_c^*\widetilde{P}'_bE_c\big).$$

The right-hand side factorises over a and c once b and k are fixed, so

$$\sum_{a,b,c}\big|\operatorname{tr}(\widetilde{P}_kD_a\widetilde{P}'_bE_c)\big|^2\le\sum_b\Big(\sum_a\operatorname{tr}\big(\widetilde{P}_kD_a\widetilde{P}'_bD_a^*\big)\Big)\Big(\sum_c\operatorname{tr}\big(\widetilde{P}_kE_c^*\widetilde{P}'_bE_c\big)\Big)=\sum_bA_{k,b}B_{k,b},$$

where we used $\sum_aD_a\widetilde{P}'_bD_a^*=\mathcal{D}(\widetilde{P}'_b)$ and $\sum_cE_c^*\widetilde{P}'_bE_c=\mathcal{E}^*(\widetilde{P}'_b)$. Multiplying by q_k/d_k^2 and summing over k gives eq. (73). □

Remark 21. Note that eq. (72) is a pair of *complementary* normalisations: A is normalised along columns (indexed by the blocks of $\mathcal{G}$) and B along rows (indexed by the blocks of $\mathcal{F}$). This asymmetry is not an artefact: A carries the decoder's trace preservation and B the unitality of the adjoint of the encoder. It is precisely this asymmetry that forces the extremal configurations to be *injective partial matchings* between the two block sets.

Everything now reduces to a finite-dimensional optimisation which we solve exactly.

Lemma 22 (Extremal configurations)**.** *Let $c = (c_k)_{k\in[K]} \in \mathbb{R}^K_{\geq 0}$, $d = (d_k)_{k\in[K]} \in \mathbb{R}^K_{>0}$ and $d' = (d'_b)_{b\in[K']} \in \mathbb{R}^{K'}_{>0}$. Put*

$$\mathcal{A} = \Big\{ A \in \mathbb{R}^{K\times K'}_{\geq 0} \;:\; \sum_{k=1}^{K} A_{k,b} = d'_b, \quad \forall b \in [K'] \Big\}, \qquad \mathcal{B} = \Big\{ B \in \mathbb{R}^{K\times K'}_{\geq 0} \;:\; \sum_{b=1}^{K'} B_{k,b} = d_k, \quad \forall k \in [K] \Big\}.$$

Then

$$\max_{A\in\mathcal{A},\, B\in\mathcal{B}} \sum_{k,b} c_k\, A_{k,b} B_{k,b} \;=\; \max\Big\{ \sum_{b\in S} c_{\phi(b)}\, d'_b\, d_{\phi(b)} \;:\; S \subseteq [K'],\ \phi : S \to [K] \text{ injective} \Big\}. \tag{74}$$

Proof. Write $f(A,B) = \sum_{k,b} c_k A_{k,b} B_{k,b}$. The sets $\mathcal{A}$ and $\mathcal{B}$ are non-empty, both contain the matrices with entries d'_b/K and d_k/K' respectively, and are compact convex polytopes, so the maximum of the continuous function f on $\mathcal{A}\times\mathcal{B}$ is attained at a pair $(A^\star, B^\star)$, with value M.

(i) The maximum is attained at a pair of vertices. The map $A \mapsto f(A, B^\star)$ is linear, hence attains its maximum over the polytope $\mathcal{A}$ at an extreme point $A^\sharp$; then $f(A^\sharp, B^\star) \geq f(A^\star, B^\star) = M$, so $f(A^\sharp, B^\star) = M$. Likewise $B \mapsto f(A^\sharp, B)$ is linear and attains its maximum at an extreme point $B^\sharp$ with $f(A^\sharp, B^\sharp) = M$.

(ii) Description of the vertices. $\mathcal{A} = \prod_{b\in[K']} (d'_b \Delta_{K-1})$ is a product of scaled simplices (Δ_{K-1} the probability simplex on $[K]$), and the extreme points of a product of compact convex sets are exactly the products of extreme points. Hence the extreme points of $\mathcal{A}$ are the matrices $A_{k,b} = d'_b\, \delta_{k,\phi(b)}$ for functions $\phi : [K'] \to [K]$. Symmetrically, the extreme points of $\mathcal{B}$ are $B_{k,b} = d_k\, \delta_{b,\psi(k)}$ for functions $\psi : [K] \to [K']$.

(iii) Evaluation. For such a pair,

$$f(A,B) = \sum_{k,b} c_k\, d'_b\, d_k\, \delta_{k,\phi(b)}\delta_{b,\psi(k)} = \sum_{b\in S} c_{\phi(b)}\, d'_b\, d_{\phi(b)}, \qquad S = \{ b \in [K'] \;:\; \psi(\phi(b)) = b \}.$$

The restriction $\phi|_S$ is injective: if $b_1, b_2 \in S$ and $\phi(b_1) = \phi(b_2)$, then $b_1 = \psi(\phi(b_1)) = \psi(\phi(b_2)) = b_2$. This proves "$\leq$" in eq. (74).

(iv) Converse. Given $S \subseteq [K']$ and an injective $\phi : S \to [K]$, extend ϕ to $[K']$ arbitrarily and define $\psi : [K] \to [K']$ by $\psi(\phi(b)) = b$ for $b \in S$ (well defined by injectivity) and arbitrarily elsewhere. The corresponding vertices satisfy $\{b : \psi(\phi(b)) = b\} \supseteq S$, so $f(A,B) \geq \sum_{b\in S} c_{\phi(b)} d'_b d_{\phi(b)}$ (all terms are non-negative). This gives "$\geq$". □

Lemma 23. *Let $d = (d_1, \ldots, d_K) \in \mathbb{N}^K_{>0}$ and $d' = (d'_1, \ldots, d'_{K'}) \in \mathbb{N}^{K'}_{>0}$. Fix $p \in [1, +\infty)$ and let $q^{(p)}$ be as in eq. (67). Then for all $A \in \mathcal{A}$, $B \in \mathcal{B}$ (with the notation of Lemma 22),*

$$\sum_{k,b} \frac{q^{(p)}_k}{d_k^2} A_{k,b} B_{k,b} \;\leq\; \frac{\|d'\|_p}{\|d\|_p}. \tag{75}$$

Proof. Apply Lemma 22 with $c_k = q^{(p)}_k / d_k^2 = d_k^{p-2} / \|d\|_p^p$. For $S \subseteq [K']$ and $\phi : S \to [K]$ injective, the corresponding value is

$$\sum_{b\in S} \frac{d_{\phi(b)}^{p-2}}{\|d\|_p^p}\, d'_b\, d_{\phi(b)} = \frac{1}{\|d\|_p^p} \sum_{b\in S} d'_b\, d_{\phi(b)}^{p-1}.$$

Case $p = 1$. The exponent $p-1$ vanishes, so the value equals $\frac{1}{\|d\|_1} \sum_{b\in S} d'_b \leq \frac{\|d'\|_1}{\|d\|_1}$, which is eq. (75).

Case $p \in (1, +\infty)$. Let $p' = \frac{p}{p-1}$ be the Hölder conjugate, so that $(p-1)p' = p$. Hölder's inequality on the index set S gives

$$\sum_{b\in S} d'_b\, d_{\phi(b)}^{p-1} \;\le\; \Big(\sum_{b\in S}(d'_b)^p\Big)^{1/p}\Big(\sum_{b\in S} d_{\phi(b)}^{(p-1)p'}\Big)^{1/p'} = \Big(\sum_{b\in S}(d'_b)^p\Big)^{1/p}\Big(\sum_{b\in S} d_{\phi(b)}^{p}\Big)^{\frac{p-1}{p}}.$$

Now $\sum_{b\in S}(d'_b)^p \le \|d'\|_p^p$, and *because* ϕ *is injective* the indices $\phi(b)$, $b \in S$, are pairwise distinct, whence $\sum_{b\in S} d_{\phi(b)}^p \le \sum_{k\in[K]} d_k^p = \|d\|_p^p$. Therefore

$$\sum_{b\in S} d'_b\, d_{\phi(b)}^{p-1} \;\le\; \|d'\|_p\, \|d\|_p^{p-1},$$

and the value is at most $\|d'\|_p\|d\|_p^{p-1}/\|d\|_p^p = \|d'\|_p/\|d\|_p$. Since by Lemma 22 every admissible (A, B) gives a value at most the maximum over such (S, ϕ), eq. (75) follows. □

We can now prove item (ii) of Theorem A.

Proof of item (ii) of Theorem A. By eq. (61), it suffices to prove eq. (2) for the block pinchings $\Pi_{\mathbf{f}}$ and $\Pi_{\mathbf{g}}$, since $\lambda(\Pi_{\mathbf{f}}) = \lambda(\mathcal{F}) = \mathbf{f}$ and $\lambda(\Pi_{\mathbf{g}}) = \lambda(\mathcal{G}) = \mathbf{g}$.

Fix $p \in [1, +\infty)$, $(p_\lambda)_{\lambda\in[\Lambda]}$ a probability distribution and channels $\{(\mathcal{D}_\lambda, \mathcal{E}_\lambda) \mid \lambda \in [\Lambda]\}$, with $\mathcal{E}_\lambda : \mathcal{L}(\mathcal{H}_{\mathbf{f}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{g}})$ and $\mathcal{D}_\lambda : \mathcal{L}(\mathcal{H}_{\mathbf{g}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{f}})$. We first show that

$$\frac{1}{2}\Big\|\Pi_{\mathbf{f}} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \Pi_{\mathbf{g}} \mathcal{E}_\lambda\Big\|_\diamond \ge 1 - \inf_{p\in[1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}, \tag{76}$$

that is, we show the lower bound

$$\delta_{1,1}^{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \ge 1 - \inf_{p\in[1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}. \tag{77}$$

Let $\mathcal{N}_\lambda$ be as in eq. (63), and let $\sigma^{(p)}$, μ be as in Definition 16. By Lemma 17, $0 \le \mu \le \mathbb{1}$, and $\sigma^{(p)}$ is a state. Hence, Lemma 14 applies to the channels $\Pi_{\mathbf{f}}$ and $\sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{N}_\lambda$ on $\mathcal{L}(\mathcal{H}_{\mathbf{f}})$ with the auxiliary system $\mathcal{H}_R$:

$$\frac{1}{2}\Big\|\Pi_{\mathbf{f}} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{N}_\lambda\Big\|_\diamond \;\ge\; \operatorname{tr}\big(\mu(\Pi_{\mathbf{f}} \otimes \mathrm{Id})(\sigma^{(p)})\big) - \sum_{\lambda\in[\Lambda]} p_\lambda \operatorname{tr}\big(\mu(\mathcal{N}_\lambda \otimes \mathrm{Id})(\sigma^{(p)})\big).$$

By Lemma 18 the first term equals 1. For the second term, let $\lambda \in [\Lambda]$. By Lemma 20, $\operatorname{tr}\big(\mu(\mathcal{N}_\lambda \otimes \mathrm{Id})(\sigma^{(p)})\big)$ is at most $\sum_{k,b} \frac{q_k^{(p)}}{d_k^2} A_{k,b}^{(\lambda)} B_{k,b}^{(\lambda)}$ for matrices $A^{(\lambda)}, B^{(\lambda)}$ satisfying eq. (72), i.e. $A^{(\lambda)} \in \mathcal{A}$ and $B^{(\lambda)} \in \mathcal{B}$ in the notation of Lemma 22. By Lemma 23 this is at most $\|\lambda(\mathcal{G})\|_p/\|\lambda(\mathcal{F})\|_p$. Hence, for every $\lambda \in [\Lambda]$,

$$\operatorname{tr}\big(\mu(\mathcal{N}_\lambda \otimes \mathrm{Id})(\sigma^{(p)})\big) \le \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p},$$

which is eq. (69). Therefore,

$$\frac{1}{2}\Big\|\Pi_{\mathbf{f}} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{N}_\lambda\Big\|_\diamond \;\ge\; 1 - \sum_{\lambda\in[\Lambda]} p_\lambda \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p} = 1 - \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}.$$

This proves the pointwise bound for finite p, since $\mathcal{N}_\lambda = \mathcal{D}_\lambda \circ \Pi_{\mathbf{g}} \circ \mathcal{E}_\lambda$ by eq. (63).

For any fixed finite vector u, we have $\|u\|_p \to \|u\|_\infty$ as $p \to +\infty$. Taking this limit in the preceding pointwise bound therefore yields $\frac{1}{2}\|\Pi_{\mathbf{f}} - \sum_{\lambda\in[\Lambda]} p_\lambda \mathcal{D}_\lambda \Pi_{\mathbf{g}} \mathcal{E}_\lambda\|_\diamond \ge 1 - \|\lambda(\mathcal{G})\|_\infty/\|\lambda(\mathcal{F})\|_\infty$.

Since the preceding bound holds for every $p \in [1, +\infty]$ separately, it holds with the right-hand side maximised over p, i.e. with $\|\lambda(\mathcal{G})\|_p/\|\lambda(\mathcal{F})\|_p$ replaced by its infimum over $p \in [1, +\infty]$. Taking the infimum over Λ, over the probability distributions $(p_\lambda)_{\lambda\in[\Lambda]}$ and over the pairs of channels $\{(\mathcal{D}_\lambda, \mathcal{E}_\lambda)\}_{\lambda\in[\Lambda]}$, and using eq. (61), we obtain

$$\delta_{1,1}^{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \ge 1 - \inf_{p\in[1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p}{\|\lambda(\mathcal{F})\|_p}. \tag{78}$$

We can now show eq. (2) in full generality by noticing that $\delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = \delta^{\mathrm{SR}}_{1,1}(\mathcal{G}^{\otimes n} \mapsto \mathcal{F}^{\otimes k})$. Hence

$$\begin{aligned}\delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) &= \delta^{\mathrm{SR}}_{1,1}(\mathcal{G}^{\otimes n} \mapsto \mathcal{F}^{\otimes k}) \\ &\geq 1 - \inf_{p\in[1,+\infty]} \frac{\|\lambda(\mathcal{G}^{\otimes n})\|_p}{\|\lambda(\mathcal{F}^{\otimes k})\|_p} \\ &= 1 - \inf_{p\in[1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p^n}{\|\lambda(\mathcal{F})\|_p^k},\end{aligned}$$

where the last equality follows by Proposition 7. □

Now that item (ii) of Theorem A is proven, we can prove the converse part of item (i).

Proof of the converse part of item (i) of Theorem A. Let $\mathcal{F}$ and $\mathcal{G}$ be idempotent channels, with $\lambda(\mathcal{F}) \neq (1)$, and let $P_{\mathcal{F}} = [1, +\infty]$ if $\mathcal{F}$ is not classical, i.e. if $\lambda(\mathcal{F}) \neq (1, \dots, 1)$, and $P_{\mathcal{F}} = [1, +\infty)$ otherwise. In both cases, $\log(\|\lambda(\mathcal{F})\|_p) > 0$ for all $p \in P_{\mathcal{F}}$, and, following the convention of Section 4, the infimum in eq. (1) is taken over $p \in P_{\mathcal{F}}$.

Let r be an achievable rate in the shared randomness regime, and let $(k_\nu)_\nu$ and $(n_\nu)_\nu$ be two sequences of positive integers satisfying eq. (10) with $\delta_\nu = \delta^{\mathrm{SR}}_{k_\nu,n_\nu}(\mathcal{G} \mapsto \mathcal{F}) \to 0$. Let $p \in P_{\mathcal{F}}$. By item (ii) of Theorem A, for every ν,

$$\frac{\|\lambda(\mathcal{G})\|_p^{n_\nu}}{\|\lambda(\mathcal{F})\|_p^{k_\nu}} \geq 1 - \delta_\nu.$$

For ν large enough, $\delta_\nu < 1$, and, taking the logarithm and dividing by $n_\nu \log(\|\lambda(\mathcal{F})\|_p) > 0$, we get

$$\frac{k_\nu}{n_\nu} \leq \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} - \frac{\log(1-\delta_\nu)}{n_\nu \log(\|\lambda(\mathcal{F})\|_p)}.$$

Since $\delta_\nu \to 0$ and $n_\nu \to +\infty$, the last term vanishes as $\nu \to \infty$, so that

$$r \leq \liminf_{\nu\to\infty} \frac{k_\nu}{n_\nu} \leq \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}.$$

As this holds for every $p \in P_{\mathcal{F}}$ and every achievable rate r, and as every rate achievable in the unassisted regime is also achievable with shared randomness, we obtain

$$C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) \leq C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \leq \inf_{p\in P_{\mathcal{F}}} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}. \tag{79}$$

Finally, by Section 2.1, the zero-error emulation capacity in each regime is at most the corresponding emulation capacity, and the zero-error unassisted emulation capacity is at most the zero-error shared randomness emulation capacity. In Section 4, we showed that, for every rate r strictly below the right-hand side of eq. (79), there is a pair (k, n) with $r \leq k/n$ such that $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$ without assistance. Hence, the zero-error unassisted emulation capacity is at least the right-hand side of eq. (79), so that the four capacities coincide with it. This ends the proof of item (i) of Theorem A. □

Finally, item (ii) yields the strong converse of item (iii) of Theorem A. Since the normalised diamond distance between two channels is at most 1, item (iii) states that if one tries to emulate $\mathcal{F}$ with $\mathcal{G}$, with or without shared randomness, at a rate greater than the emulation capacity, the error tends to its maximal value, exponentially fast in the blocklength. As the lower bound of eq. (2) holds at every finite blocklength, it suffices to evaluate it at a value of p which almost achieves the infimum in eq. (1).

Proof of item (iii) of Theorem A. Let $\mathcal{F}$ and $\mathcal{G}$ be idempotent channels with $\lambda(\mathcal{F}) \neq (1)$. By item (i), $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F})$, and we write

$$C = C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) = \inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}.$$

Since $\delta^{\mathrm{UA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq \delta^{\mathrm{SR}}_{k,n}(\mathcal{G} \mapsto \mathcal{F})$ for all $k, n \in \mathbb{N}_{>0}$, it suffices to prove the first inequality of item (iii) in the shared randomness regime. Throughout this proof, we take the logarithm in base 2.

Let $\varepsilon > 0$. We first fix a value of p which almost achieves the infimum defining C. Since $\lambda(\mathcal{F}) \neq (1)$, we have $\|\lambda(\mathcal{F})\|_p > 1$ for all $p \in [1, +\infty)$, and the ratio in the infimum is well-defined for these values of p (the value $p = +\infty$ being also allowed when $\|\lambda(\mathcal{F})\|_\infty \geq 2$). Hence, by definition of the infimum, there is $p^\star \in [1, +\infty]$ such that $\|\lambda(\mathcal{F})\|_{p^\star} > 1$ and

$$\frac{\log(\|\lambda(\mathcal{G})\|_{p^\star})}{\log(\|\lambda(\mathcal{F})\|_{p^\star})} \leq C + \frac{\varepsilon}{2}. \tag{80}$$

Note that $p^\star$ depends only on $\mathcal{F}$, $\mathcal{G}$ and ε. We write $a = \log(\|\lambda(\mathcal{F})\|_{p^\star}) > 0$ and $g = \log(\|\lambda(\mathcal{G})\|_{p^\star}) \geq 0$, so that eq. (80) reads $g \leq (C + \varepsilon/2)a$.

Let $k, n \in \mathbb{N}_{>0}$ with $k/n \geq C + \varepsilon$. Multiplying this inequality by $na > 0$, we get

$$ka \geq n(C + \varepsilon)a \geq ng + \frac{\varepsilon}{2}na,$$

where we used eq. (80) in the second inequality. Therefore,

$$\frac{\|\lambda(\mathcal{G})\|_{p^\star}^n}{\|\lambda(\mathcal{F})\|_{p^\star}^k} = 2^{ng - ka} \leq 2^{-\frac{\varepsilon}{2}a\,n}.$$

Now, by item (ii), since the infimum over p in eq. (2) is upper-bounded by the value of the ratio at $p = p^\star$,

$$\delta_{k,n}^{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - \inf_{p \in [1,+\infty]} \frac{\|\lambda(\mathcal{G})\|_p^n}{\|\lambda(\mathcal{F})\|_p^k} \geq 1 - \frac{\|\lambda(\mathcal{G})\|_{p^\star}^n}{\|\lambda(\mathcal{F})\|_{p^\star}^k} \geq 1 - 2^{-cn},$$

with $c = \frac{\varepsilon}{2}\log(\|\lambda(\mathcal{F})\|_{p^\star}) > 0$, which depends only on ε, $\mathcal{F}$ and $\mathcal{G}$. This proves item (iii) in the shared randomness regime, and thus in the unassisted regime.

Finally, let $(k_\nu)_\nu$ and $(n_\nu)_\nu$ be two sequences of positive integers with $n_\nu \to +\infty$ and $\liminf_\nu k_\nu/n_\nu > C$. Let $\varepsilon > 0$ be such that $\liminf_\nu k_\nu/n_\nu > C + \varepsilon$, so that $k_\nu/n_\nu \geq C + \varepsilon$ for all ν large enough. For such ν, since the normalised diamond distance between two channels is at most 1,

$$1 - 2^{-cn_\nu} \leq \delta_{k_\nu,n_\nu}^{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) \leq \delta_{k_\nu,n_\nu}^{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) \leq 1,$$

so that both errors tend to 1 exponentially fast in n_ν. □

6 An efficient algorithm for the emulation capacity

Since the single-letter expression giving the emulation capacity, both in the unassisted and in the shared-randomness-assisted settings, involves computing an infimum over an unbounded interval, it is not a priori clear whether this expression can actually be computed given any two idempotent channels. It turns out that the infimum in the right-hand side of eq. (1) can be *efficiently* computed. We provide in this section an algorithm that, given two channels $\mathcal{F}$ and $\mathcal{G}$ and an error parameter $\varepsilon > 0$, outputs an approximation of $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F})$ up to the error ε. Note that since an algorithm computing $\lambda(\mathcal{F})$ given $\mathcal{F}$ in time $O(d^6 \log d)$, where $d = \dim \mathcal{H}_\mathcal{F}$, is provided in [16], it suffices to construct an algorithm that, given two integer vectors u, v and an error parameter $\varepsilon > 0$, computes an approximation of

$$\inf_{p \in [1,+\infty]} \frac{\log(\|u\|_p)}{\log(\|v\|_p)}. \tag{81}$$

The relevance of such an algorithm goes therefore beyond the problem of computing the emulation capacities for idempotent channels.

Throughout this section, we write, for $u \in \mathbb{N}_{>0}^{m_u}$, $v \in \mathbb{N}_{>0}^{m_v}$ and $p \in [1, +\infty]$,

$$R_{u,v}(p) = \frac{\log(\|u\|_p)}{\log(\|v\|_p)}, \qquad R_*(u, v) = \inf_{p \in [1,+\infty]} R_{u,v}(p), \tag{82}$$

so that, by item (i) of Theorem A, $C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}) = C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F}) = R_*(\lambda(\mathcal{G}), \lambda(\mathcal{F}))$. Here, m_u and m_v denote the number of entries of u and v, and we write $M_u = \|u\|_\infty$, $M_v = \|v\|_\infty$ their largest entries. Since the ratio in eq. (82) does not depend on the base of the logarithm, we take, in this section only,

log to be the natural logarithm, and we write $\log_2$ for the binary logarithm. We always assume that $v \neq (1)$, which corresponds to the assumption $\lambda(\mathcal{F}) \neq (1)$ of Theorem A. Then, $\|v\|_p > 1$ for every $p \in [1, +\infty)$, since v has either an entry greater than or equal to 2 or at least two entries equal to 1. However, $\|v\|_\infty = 1$ when all the entries of v are equal to 1, i.e. when $\mathcal{F}$ is a classical channel. In that case, the denominator of $R_{u,v}(\infty)$ vanishes and we use the convention that the term $p = +\infty$ is omitted in the infimum of eq. (82), which is consistent with the last row of Table 2.

6.1 Main result and model of computation

We work in the standard bit model of computable analysis, see e.g. [31]: the vectors u and v are given as explicit lists of positive integers written in binary, and the precision is given as an integer $b \in \mathbb{N}$. Hence, the size of the input is measured by b and by

$$S = m_u + m_v + \sum_{i=1}^{m_u} \lceil \log_2(u_i + 1) \rceil + \sum_{j=1}^{m_v} \lceil \log_2(v_j + 1) \rceil. \tag{83}$$

Note that we will frequently use the following elementary bounds: for $x \in \mathbb{N}_{>0}^{m_x}$,

$$\log(m_x) \leq m_x \leq S, \qquad \log\Big(\sum_{i=1}^{m_x} x_i\Big) \leq \log\Big(\prod_{i=1}^{m_x} (x_i + 1)\Big) \leq \sum_{i=1}^{m_x} \lceil \log_2(x_i + 1) \rceil \leq S, \tag{84}$$

where we used that the natural logarithm is upper-bounded by the binary one on $[1, +\infty)$. The main result of this section is the following theorem.

Theorem 24. *There is a deterministic algorithm which, given as input $u \in \mathbb{N}_{>0}^{m_u}$, $v \in \mathbb{N}_{>0}^{m_v}$ with $v \neq (1)$ and $b \in \mathbb{N}$, outputs a rational number $\widehat{R}$ such that*

$$|\widehat{R} - R_*(u, v)| \leq 2^{-b}. \tag{85}$$

The bit complexity of this algorithm is polynomial in S and b, where S is defined in eq. (83).

Combining this theorem with Proposition 6, we obtain that the emulation capacities of item (i) of Theorem A can be efficiently computed from the description of the channels.

Corollary 25. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels with $\lambda(\mathcal{F}) \neq (1)$, and let $b \in \mathbb{N}$. One can compute a rational number $\widehat{R}$ such that $|\widehat{R} - C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F})| = |\widehat{R} - C_{\mathrm{SR}}(\mathcal{G} \mapsto \mathcal{F})| \leq 2^{-b}$ by first computing the shape vectors $\lambda(\mathcal{F})$ and $\lambda(\mathcal{G})$ with the algorithm of Proposition 6, and then running the algorithm of Theorem 24 on $(\lambda(\mathcal{G}), \lambda(\mathcal{F}), b)$. The first step has complexity polynomial in $\dim(\mathcal{H}_\mathcal{F})$ and $\dim(\mathcal{H}_\mathcal{G})$ in the computational model of [16]. The second has bit complexity polynomial in b, $\dim(\mathcal{H}_\mathcal{F})$ and $\dim(\mathcal{H}_\mathcal{G})$. With these computational conventions, the overall complexity is polynomial in b, $\dim(\mathcal{H}_\mathcal{F})$ and $\dim(\mathcal{H}_\mathcal{G})$.*

Proof. The correctness follows directly from item (i) of Theorem A and Theorem 24. For the complexity, recall from Proposition 1 that $\lambda(\mathcal{F})$ has $K \leq \dim(\mathcal{H}_\mathcal{F})$ entries, and that each of them is the dimension d_k of a subspace of $\mathcal{H}_\mathcal{F}$, so that $d_k \leq \dim(\mathcal{H}_\mathcal{F})$. Hence, the contribution of $\lambda(\mathcal{F})$ to the input size S of eq. (83) is at most $\dim(\mathcal{H}_\mathcal{F})(1 + \lceil \log_2(\dim(\mathcal{H}_\mathcal{F}) + 1) \rceil)$, and similarly for $\lambda(\mathcal{G})$. Therefore, $S = O(\dim(\mathcal{H}_\mathcal{F}) \log(\dim(\mathcal{H}_\mathcal{F})) + \dim(\mathcal{H}_\mathcal{G}) \log(\dim(\mathcal{H}_\mathcal{G})))$ and the conclusion follows from Theorem 24. □

Before we move on to the proof of Theorem 24, we make the following remark on its scope.

Remark 26. Theorem 24 provides a certified approximation of the *value* of $R_*(u, v)$. It does not, by itself, decide symbolic questions such as whether the infimum in eq. (82) is attained at a given value of p, e.g. at $p = 2$ or at $p = +\infty$, nor whether $R_*(\lambda(\mathcal{G}), \lambda(\mathcal{F}))$ is larger than a given rational number k/n. Indeed, two distinct transcendental expressions can be arbitrarily close to each other, and deciding such questions would require additional structural information on the two shape vectors.

The algorithm relies on standard subroutines of computer algebra, which we now recall. First, given a real number $y \geq 0$ and an integer $k \in \mathbb{N}_{>0}$, we write $\lfloor y \rfloor_k$ for a rational number approximating y from below with k binary digits of accuracy, that is, such that

$$0 \leq y - \lfloor y \rfloor_k \leq 2^{-k}. \tag{86}$$

This is a slight abuse of notation, as such a number is not unique: whenever $\widetilde{y}$ is a rational number with $|\widetilde{y} - y| \leq 2^{-k-1}$, the number $\max\{0, \widetilde{y} - 2^{-k-1}\}$ satisfies eq. (86), and we fix one such choice for each quantity we truncate. We use this one-sided approximation instead of the exact truncation $2^{-k}\lfloor 2^k y \rfloor$ so that computing it requires no separation estimate between y and the dyadic boundaries. Also, we choose $\widetilde{y}$ to be a multiple of 2^{-k-1}, so that $\lfloor y \rfloor_k$ also has at most $k+1$ binary digits after the binary point. The only transcendental evaluations required by our algorithm are logarithms of positive rational numbers. Certified approximations of these logarithms can be computed in time polynomial in the requested precision and in the binary description length of their arguments, see e.g. [32]. The quotients of logarithms appearing below are evaluated as exact quotients of rational approximations. Their denominators approximate $\log(a)$ with $a \geq 2$ from below to error at most $1/4$, and are therefore at least $\log(2) - 1/4 > 1/4$.

We will also need these two technical lemmas.

Lemma 27 (Real root isolation, see e.g. [33])**.** *Let $Q \in \mathbb{Q}[X]$ be a non-zero polynomial of degree d whose coefficients have binary length at most τ.*

(i) The square-free part $Q_{\mathrm{sf}} = Q/\gcd(Q, Q')$ of Q can be computed in time polynomial in d and τ, and its degree and the binary length of its coefficients are polynomial in d and τ. Moreover, Q_{sf} is square-free and has the same real roots as Q, see e.g. [34].

(ii) If Q is square-free, then, given $\kappa \in \mathbb{N}$, one can compute in time polynomial in d, τ and κ a list of pairwise disjoint intervals with rational endpoints, each of width at most $2^{-\kappa}$, such that every interval contains exactly one real root of Q and every real root of Q belongs to one of these intervals.

Lemma 28. *Let $n, \delta, \tilde{n}, \tilde{\delta}$ be real numbers with $0 \leq n \leq S$ and $\delta \geq 1/2$, and let $\tau \in (0, 1/4]$ be such that $|\tilde{n} - n| \leq \tau$ and $|\tilde{\delta} - \delta| \leq \tau$. Then $\tilde{\delta} \geq 1/4$ and*

$$\Big|\frac{\tilde{n}}{\tilde{\delta}} - \frac{n}{\delta}\Big| \leq (4 + 8S)\tau. \tag{87}$$

Proof. Since $\tau \leq 1/4$ and $\delta \geq 1/2$, we have $\tilde{\delta} \geq \delta - \tau \geq 1/4 > 0$. Hence

$$\Big|\frac{\tilde{n}}{\tilde{\delta}} - \frac{n}{\delta}\Big| = \Big|\frac{\tilde{n} - n}{\tilde{\delta}} + n\,\frac{\delta - \tilde{\delta}}{\delta\tilde{\delta}}\Big| \leq \frac{\tau}{\tilde{\delta}} + \frac{n\tau}{\delta\tilde{\delta}} \leq 4\tau + \frac{S\tau}{(1/2)(1/4)} = (4 + 8S)\tau. \qquad \square$$

6.2 Reduction to a bounded range of p

The first step of the algorithm consists in reducing the optimisation over the unbounded interval $[1, +\infty]$ to an optimisation over a bounded interval $[1, P_{\max}]$. To that end, we change the variable $p \in [1, +\infty]$ into $t = 1/p \in [0, 1]$. Throughout this subsection and the next one, given $x \in \mathbb{N}_{>0}^{m_x}$, we write

$$Z_x(p) = \sum_{i=1}^{m_x} \Big(\frac{x_i}{M_x}\Big)^p, \qquad p \in [1, +\infty), \tag{88}$$

for the sum of the p-th powers of the entries of x, normalised by its largest entry. Every term of this sum is at most 1, and at least one of them equals 1, so that

$$1 \leq Z_x(p) \leq m_x, \qquad p \in [1, +\infty), \tag{89}$$

and, factoring out M_x,

$$\log(\|x\|_p) = \log(M_x) + \frac{1}{p}\log(Z_x(p)). \tag{90}$$

Furthermore, given a vector $x \in \mathbb{N}_{>0}^{m_x}$, we define the function $A_x : [0, 1] \to \mathbb{R}$ by

$$A_x(t) = t\log\Big(\sum_{i=1}^{m_x} x_i^{1/t}\Big) \quad \text{for } t \in (0, 1], \qquad A_x(0) = \log(M_x), \tag{91}$$

so that for all $p \in [1, +\infty]$, $A_x(1/p) = \log(\|x\|_p)$, with the convention $1/(+\infty) = 0$. Equivalently, by eq. (90),

$$A_x(t) = \log(M_x) + t\log(Z_x(1/t)) \tag{92}$$

for $t \in (0,1]$, which makes the value $A_x(0) = \log(M_x)$ the natural extension of A_x at $t = 0$. We also write, for $t \in (0,1]$, $w_x(t)$ for the probability vector with coordinates

$$w_{x,i}(t) = \frac{x_i^{1/t}}{\sum_{j=1}^{m_x} x_j^{1/t}} = \frac{(\frac{x_i}{M_x})^{1/t}}{Z_x(1/t)}, \quad i \in [m_x], \tag{93}$$

and $H(w) = -\sum_i w_i \log(w_i)$ for the Shannon entropy of a probability vector w.

Lemma 29. *Let $x \in \mathbb{N}_{>0}^{m_x}$. The function A_x is continuous and non-decreasing on $[0,1]$, differentiable on $(0,1]$ and, for all $t \in (0,1]$,*

$$A_x'(t) = H(w_x(t)) \in [0, \log(m_x)]. \tag{94}$$

In particular, for all $t \in [0,1]$, $\log(M_x) \le A_x(t) \le \log\big(\sum_{i=1}^{m_x} x_i\big)$.

Proof. Let $a_i = \log(x_i/M_x) \le 0$ for $i \in [m_x]$, so that, by eqs. (88) and (90), for $t \in (0,1]$,

$$Z_x(1/t) = \sum_{i=1}^{m_x} e^{a_i/t}, \qquad w_{x,i}(t) = \frac{e^{a_i/t}}{Z_x(1/t)}.$$

The derivative of $t \mapsto Z_x(1/t)$ is given by

$$\tfrac{d}{dt} Z_x(1/t) = -\sum_i \frac{a_i}{t^2} e^{a_i/t}, \tag{95}$$

for $t \in (0,1]$. Hence, we obtain the derivative of A_x:

$$\begin{aligned} A_x'(t) &= \log(Z_x(1/t)) - \frac{1}{t}\sum_{i=1}^{m_x} w_{x,i}(t) a_i \\ &= -\sum_{i=1}^{m_x} w_{x,i}(t)\Big(\frac{a_i}{t} - \log(Z_x(1/t))\Big) \\ &= -\sum_{i=1}^{m_x} w_{x,i}(t)\log(w_{x,i}(t)) \\ &= H(w_x(t)), \end{aligned}$$

where we used $\sum_i w_{x,i}(t) = 1$ in the second equality. This proves eq. (94), since the entropy of a probability vector with m_x coordinates lies in $[0, \log(m_x)]$. For the continuity at $t = 0$, eq. (89) gives

$$|A_x(t) - \log(M_x)| = t\log(Z_x(1/t)) \le t\log(m_x) \underset{t\to 0}{\to} 0. \tag{96}$$

Finally, A_x is non-decreasing as its derivative is non-negative, so that $A_x(0) \le A_x(t) \le A_x(1) = \log(\sum_{i=1}^{m_x} x_i)$. □

We can now control the behaviour of $R_{u,v}(p)$ for large p.

Lemma 30 (Tail bound)**.** *Let $u \in \mathbb{N}_{>0}^{m_u}$ and $v \in \mathbb{N}_{>0}^{m_v}$ with $M_v \ge 2$. Then, for all $p \in [1, +\infty)$,*

$$|R_{u,v}(p) - R_{u,v}(\infty)| \le \frac{L_{u,v}}{p}, \quad \text{with} \quad L_{u,v} = \frac{\log(m_u)}{\log(M_v)} + \frac{\log(\sum_{i=1}^{m_u} u_i)\log(m_v)}{\log(M_v)^2}. \tag{97}$$

Furthermore, $L_{u,v} \le \overline{L} = 2S + 4S^2$, with S defined in eq. (83).

Proof. Let $Q(t) = A_u(t)/A_v(t)$ for $t \in [0,1]$, so that $Q(1/p) = R_{u,v}(p)$ for $p \in [1, +\infty]$. By Lemma 29, $A_v(t) \ge \log(M_v) > 0$ for all $t \in [0,1]$, so that Q is well-defined and continuous on $[0,1]$ and differentiable on $(0,1]$ with

$$|Q'(t)| = \Big|\frac{A_u'(t)}{A_v(t)} - \frac{A_u(t)A_v'(t)}{A_v(t)^2}\Big| \le \frac{\log(m_u)}{\log(M_v)} + \frac{\log(\sum_i u_i)\log(m_v)}{\log(M_v)^2} = L_{u,v},$$

where we used $0 \le \log(M_u) \le A_u(t) \le \log(\sum_{i=1}^{m_u} u_i)$, $0 \le A_u'(t) \le \log(m_u)$, $\log(M_v) \le A_v(t)$ and $0 \le A_v'(t) \le \log(m_v)$ from Lemma 29. By the mean value theorem, $|R_{u,v}(p) - R_{u,v}(\infty)| = |Q(1/p) - Q(0)| \le L_{u,v}/p$. Finally, since $\log(M_v) \ge \log(2) \ge 1/2$, the bounds of eq. (84) give $L_{u,v} \le 2S + 4S^2$. □

The case $M_v = 1$, which is excluded in Lemma 30, is simple.

Lemma 31 (Classical target)**.** *Let $u \in \mathbb{N}_{>0}^{m_u}$ and $v = (1, \dots, 1) \in \mathbb{N}_{>0}^{m_v}$ with $m_v \geq 2$. Then $p \mapsto R_{u,v}(p)$ is non-decreasing on $[1, +\infty)$, so that*

$$R_*(u, v) = R_{u,v}(1) = \frac{\log(\sum_{i=1}^{m_u} u_i)}{\log(m_v)}. \tag{98}$$

Proof. For $p \in [1, +\infty)$, $\log(\|v\|_p) = \frac{1}{p}\log(m_v)$, so that

$$R_{u,v}(p) = \frac{p \log(\|u\|_p)}{\log(m_v)} = \frac{\log(\sum_i u_i^p)}{\log(m_v)},$$

which is non-decreasing in p since $u_i \geq 1$ for every i. □

Therefore, when $M_v \geq 2$, Lemma 30 allows us to discard the large values of p, as follows. Let $\varepsilon = 2^{-b} \in (0, 1]$ be the precision that we target for $\widehat{R}$ in eq. (85) and let $J \in \mathbb{N}$ be the smallest integer such that $2^J \geq 8\overline{L}/\varepsilon$, with $\overline{L}$ as in Lemma 30. Then, for every $p \geq 2^J$, eq. (97) gives $|R_{u,v}(p) - R_{u,v}(\infty)| \leq \overline{L}/2^J \leq \varepsilon/8$, so that

$$\inf_{p \in [2^J, +\infty]} R_{u,v}(p) \in \big[R_{u,v}(\infty) - \tfrac{\varepsilon}{8},\ R_{u,v}(\infty)\big]. \tag{99}$$

It is therefore enough to approximate $R_{u,v}(p)$ for $p \geq 2^J$ by $R_{u,v}(\infty) = \log(M_u)/\log(M_v)$, which is a ratio of logarithms of two positive integers, and to minimise $R_{u,v}$ only over the bounded interval $[1, 2^J]$. The latter is the union of the J intervals of the form $[2^j, 2^{j+1}]$, for $j \in \{0, \dots, J-1\}$, which we call *dyadic intervals*. Although the threshold 2^J itself is exponentially large in b, the number J of these dyadic intervals is only $O(\log_2(S) + b)$, since $\overline{L} = 2S + 4S^2$ is polynomial in S. It then suffices to approximate the minimum of $R_{u,v}$ on each of these dyadic intervals, which is the object of the next subsection, and then to extract the least of these approximants. The latter step is polynomial since there are only $O(\log_2(S) + b)$ such values to compare.

6.3 Polynomial approximation on a dyadic interval

Let us fix a dyadic interval $[P, 2P]$, with $P = 2^j$, $j \in \mathbb{N}$, and let us write $p = P\theta$ with $\theta \in [1, 2]$. By eq. (90), we have, for every $x \in \mathbb{N}_{>0}^{m_x}$,

$$\log(\|x\|_{P\theta}) = \log(M_x) + \frac{1}{P\theta}\log(Z_{x,P}(\theta)), \tag{100}$$

where

$$Z_{x,P}(\theta) = Z_x(P\theta) = \sum_{i=1}^{m_x} e^{-z_i\theta}, \quad z_i = P\log\Big(\frac{M_x}{x_i}\Big) \geq 0,$$

and $Z_{x,P}(\theta) \in [1, m_x]$ for all $\theta \in [1, 2]$, by eq. (89). Hence,

$$R_{u,v}(P\theta) = \frac{N_P(\theta)}{D_P(\theta)}, \tag{101}$$

with

$$N_P(\theta) = \theta\log(M_u) + \frac{1}{P}\log(Z_{u,P}(\theta)),$$
$$D_P(\theta) = \theta\log(M_v) + \frac{1}{P}\log(Z_{v,P}(\theta)),$$

where we multiplied both the numerator and the denominator of $R_{u,v}(P\theta)$ by θ. The functions $\theta \mapsto \log(Z_{x,P}(\theta))$ are analytic, but they are not polynomials. We now show that they can be uniformly approximated on $[1, 2]$ by polynomials with rational coefficients, whose degree and coefficient binary lengths are polynomial in S and $\log_2(1/\eta)$, where η is the required precision. The construction below works for an arbitrary $\eta \in (0, 1]$, and η will be fixed in step 1 of Algorithm 3 as $\eta = \varepsilon/[256(S+1)]$. It has to be smaller than the target precision ε by a factor polynomial in S, because η controls the error

on $\log(\|u\|_p)$ and on $\log(\|v\|_p)$ separately, whereas an error η on each of them induces an error of order $S\eta$ on their ratio, by Lemma 33.

The idea of the following algorithm is to discard the exponentially small terms in $Z_{x,P}$, to truncate the Taylor series of the remaining exponentials and to truncate the series of the logarithm. This is formalised in Algorithm 1, whose correctness and complexity are proven in Lemma 32.

Algorithm 1 `LogSumPolynomial`(x, P, η)

Input: a vector $x \in \mathbb{N}_{>0}^{m_x}$ written in binary, an integer $P \geq 1$ written in binary and promised to be a power of two, for which we write $P = 2^j$, and a rational number $\eta \in (0, 1]$.
Output: a polynomial $B \in \mathbb{Q}[X]$ such that $\sup_{\theta\in[1,2]} |B(\theta) - \log(Z_{x,P}(\theta))| \leq \eta$.

1. Set $h \leftarrow \eta/(64m_x)$ and $T \leftarrow \lceil \log_2(1/h) \rceil$, so that $2^{-T} \leq h$.
2. For each $i \in [m_x]$, set $r_i \leftarrow 0$ if $x_i = M_x$, and $r_i \leftarrow \big\lfloor P \log(M_x/x_i) \big\rfloor_T$ otherwise, so that $0 \leq r_i$ and $|r_i - P\log(M_x/x_i)| \leq 2^{-T} \leq h$.
3. Set $\mathcal{J} \leftarrow \{i \in [m_x] \mid r_i \leq T + 1\}$.
4. Set $d \leftarrow 12(T + 1)$ and
$$Z_{\mathrm{pol}} \leftarrow \sum_{i\in\mathcal{J}} \sum_{\ell=0}^{d} \frac{(-r_i X)^\ell}{\ell!} \in \mathbb{Q}[X].$$
5. Set $q \leftarrow 4m_x(T - 1)$ and $c \leftarrow \big\lfloor \log(2m_x) \big\rfloor_K$ with $K = \lceil \log_2(8/\eta) \rceil$, so that $|c - \log(2m_x)| \leq 2^{-K} \leq \eta/8$.
6. Compute, using exact arithmetic in $\mathbb{Q}$, the list of coefficients in the monomial basis of the polynomial
$$B = c - \sum_{\ell=1}^{q} \frac{1}{\ell} \Big(1 - \frac{Z_{\mathrm{pol}}}{2m_x}\Big)^\ell \in \mathbb{Q}[X],$$
and **return** it.

In step 1, since h is rational, T is obtained by comparing $1/h$ with the successive powers of two, that is, by integer comparisons on the binary expansion of $1/h$; no comparison between transcendental numbers is involved.

Lemma 32. *Let $x \in \mathbb{N}_{>0}^{m_x}$, $P = 2^j$ with $j \in \mathbb{N}$ and $\eta \in (0, 1]$ be rational. Then, the polynomial B returned by Algorithm 1 satisfies*

$$\sup_{\theta\in[1,2]} |B(\theta) - \log(Z_{x,P}(\theta))| \leq \eta. \tag{102}$$

Furthermore, $\deg(B) = O(m_x T^2)$ with $T = O(\log(m_x) + \log(1/\eta))$, and Algorithm 1 has bit complexity polynomial in S, j, and the binary description length of η.

Proof. We write $z_i = P\log(M_x/x_i)$ for $i \in [m_x]$, so that $|r_i - z_i| \leq h$. We proceed in three steps. First, we bound the sum of discarded terms $\sum_{i\notin\mathcal{J}} e^{-z_i\theta}$. Second, we bound the distance between $Z_{x,P}(\theta)$ and $Z_{\mathrm{pol}}(\theta)$, which yields an upper bound on $|\log(Z_{x,P}(\theta)) - \log(Z_{\mathrm{pol}}(\theta))|$. Third, truncating the Taylor series of the logarithm yields an upper bound on $|B(\theta) - \log(Z_{\mathrm{pol}}(\theta))|$, which finally yields eq. (102) by triangle inequality.

Discarded terms. If $i \notin \mathcal{J}$, then

$$z_i \geq r_i - h > T + 1 - h \geq T,$$

so that, for $\theta \in [1, 2]$,

$$0 \leq e^{-z_i\theta} \leq e^{-T} \leq 2^{-T} \leq h = \eta/(64m_x).$$

Hence the discarded terms contribute at most $m_x\eta/(64m_x) = \eta/64$ to $Z_{x,P}(\theta)$, that is

$$0 \leq \sum_{i\notin\mathcal{J}} e^{-z_i\theta} \leq \frac{\eta}{64}. \tag{103}$$

Approximating $Z_{x,P}$ by Z_{pol}: Let $\theta \in [1,2]$,

$$|Z_{x,P}(\theta) - Z_{\mathrm{pol}}(\theta)| = \left| \sum_{i=1}^{m_x} e^{-z_i\theta} - \sum_{i\in\mathcal{J}} \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!} \right| \tag{104}$$

$$\leq \sum_{i\notin\mathcal{J}} e^{-z_i\theta} + \left| \sum_{i\in\mathcal{J}} \big(e^{-z_i\theta} - \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!}\big) \right| \tag{105}$$

$$\leq \frac{\eta}{64} + \sum_{i\in\mathcal{J}} \left| e^{-z_i\theta} - \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!} \right|, \tag{106}$$

where the last inequality follows from eq. (103) and the triangle inequality. Let $i \in \mathcal{J}$ and $\theta \in [1,2]$. By the mean value theorem, we have $|e^{-z_i\theta} - e^{-r_i\theta}| \leq \theta|z_i - r_i| \leq 2h = \eta/(32m_x)$. Hence,

$$\left| e^{-z_i\theta} - \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!} \right| \leq \left| e^{-z_i\theta} - e^{-r_i\theta} \right| + \left| e^{-r_i\theta} - \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!} \right| \leq \frac{\eta}{32m_x} + \left| e^{-r_i\theta} - \sum_{\ell=0}^{d} \frac{(-r_i\theta)^\ell}{\ell!} \right|. \tag{107}$$

It suffices now to upper bound the right-most term in eq. (107). We have $r_i\theta \in [0, 2(T+1)]$, since $r_i \leq T+1$ for $i \in \mathcal{J}$ and $\theta \leq 2$. We write $V = 2(T+1)$. Let $y \in [0, V]$. By the Taylor formula with Lagrange remainder applied to $y \mapsto e^{-y}$ at the order d, there exists $\xi \in [0, y]$ such that

$$e^{-y} - \sum_{\ell=0}^{d} \frac{(-y)^\ell}{\ell!} = (-1)^{d+1} \frac{e^{-\xi}}{(d+1)!} y^{d+1}.$$

Since $0 \leq e^{-\xi} \leq 1$ and $0 \leq y \leq V$, and since $(d+1)! \geq ((d+1)/e)^{d+1}$, this gives

$$\left| e^{-y} - \sum_{\ell=0}^{d} \frac{(-y)^\ell}{\ell!} \right| \leq \frac{V^{d+1}}{(d+1)!} \leq \Big(\frac{eV}{d+1}\Big)^{d+1}.$$

By the choice $d = 12(T+1)$ made in step 4, we have $d+1 > 12(T+1) = 6V \geq 2eV$, so that $eV/(d+1) \leq 1/2$ and

$$\left| e^{-y} - \sum_{\ell=0}^{d} \frac{(-y)^\ell}{\ell!} \right| \leq 2^{-(d+1)} \leq 2^{-T} \leq h = \frac{\eta}{64m_x},$$

where the last two inequalities follow from $d+1 \geq T$ and from the choice $T = \lceil \log_2(1/h) \rceil$ made in step 1. Summing the two contributions over $i \in \mathcal{J}$ and adding the discarded terms, we obtain

$$\sup_{\theta\in[1,2]} |Z_{\mathrm{pol}}(\theta) - Z_{x,P}(\theta)| \leq m_x \Big(\frac{\eta}{32m_x} + \frac{\eta}{64m_x}\Big) + \frac{\eta}{64} = \frac{\eta}{16}.$$

In particular, since $Z_{x,P}(\theta) \in [1, m_x]$ and $\eta \leq 1$, we have $Z_{\mathrm{pol}}(\theta) \in [1/2, 2m_x]$ for all $\theta \in [1,2]$. As log is 2-Lipschitz on $[1/2, +\infty)$, i.e. for all $x, y \in [1/2, +\infty)$, $|\log(x) - \log(y)| \leq 2|x-y|$, which follows from the mean value theorem, we obtain $|\log(Z_{\mathrm{pol}}(\theta)) - \log(Z_{x,P}(\theta))| \leq \eta/8$.

Series of the logarithm. For $z \in [1/2, 2m_x]$, we let $y = 1 - z/(2m_x) \in [0, 1 - 1/(4m_x)]$ and

$$\log(z) = \log(2m_x) + \log(1-y) = \log(2m_x) - \sum_{\ell=1}^{\infty} \frac{y^\ell}{\ell}.$$

The tail of this series satisfies

$$\sum_{\ell>q} \frac{y^\ell}{\ell} \leq \frac{y^{q+1}}{1-y} \leq 4m_x \Big(1 - \frac{1}{4m_x}\Big)^q \leq 4m_x e^{-q/(4m_x)} \leq 4m_x 2^{-\lceil \log_2(32m_x/\eta) \rceil} \leq \frac{\eta}{8}.$$

Applying this bound with $z = Z_{\mathrm{pol}}(\theta)$, and taking into account the error $\eta/8$ on the constant c, we get that for all $\theta \in [1,2]$, $|B(\theta) - \log(Z_{\mathrm{pol}}(\theta))| \leq \eta/4$. Together with the previous step, this proves eq. (102), with the error $3\eta/8 \leq \eta$.

Complexity. We have $\deg(Z_{\mathrm{pol}}) \le d = O(T)$ and $\deg(B) \le dq = O(m_x T^2)$. Then, let us bound the cost and the size of the numbers r_i computed at step 2. Multiplying by $P = 2^j$ is an exact shift of the binary point, but it multiplies absolute errors by 2^j; hence r_i is obtained by computing $\log(M_x/x_i)$ with an absolute error at most $2^{-(T+j+1)}$, which takes time polynomial in $T+j$ and in the binary lengths of M_x and x_i [32], followed by a shift and a truncation. Since $T = \lceil \log_2(64 m_x/\eta) \rceil = O(\log_2(m_x) + \log_2(1/\eta))$ and $m_x \le S$ by eq. (84), this is polynomial in S, j and $\log_2(1/\eta)$. Moreover, $0 \le r_i \le P \log(M_x/x_i) \le 2^j S$, again by eq. (84), and r_i has at most $T+1$ binary digits after the binary point, so that its binary length is $O(j + \log_2(S) + T)$, which is polynomial in S, j and $\log_2(1/\eta)$ as well. The computation of these approximations is polynomial in S, j and $\log_2(1/\eta)$ [32]. The coefficients of Z_{pol} are sums of at most m_x terms of the form $(-r_i)^\ell/\ell!$ with $\ell \le d$, so that their binary length is polynomial in S, j and $\log_2(1/\eta)$. Finally, the expansion of the q powers of $1 - Z_{\mathrm{pol}}/(2m_x)$ involves polynomials of degree at most dq. Since the coefficients of Z_{pol} admit a common denominator of polynomial binary length, these powers and their weighted sum can be computed in exact rational arithmetic using numbers of polynomial binary length. The degrees, coefficient lengths and arithmetic operation counts are all polynomial in S, j and $\log_2(1/\eta)$. □

Applying Algorithm 1 to u and to v provides polynomial approximations $B_{u,P}$ and $B_{v,P}$ of $\log(Z_{u,P})$ and $\log(Z_{v,P})$. Together with rational approximations c_u, c_v of $\log(M_u)$ and $\log(M_v)$, we obtain a rational function

$$\widetilde{R}_P(\theta) = \frac{\widetilde{N}_P(\theta)}{\widetilde{D}_P(\theta)}, \quad \text{with} \quad \widetilde{N}_P(\theta) = \theta c_u + \frac{1}{P} B_{u,P}(\theta), \quad \widetilde{D}_P(\theta) = \theta c_v + \frac{1}{P} B_{v,P}(\theta), \tag{108}$$

which approximates $R_{u,v}(P\theta)$ uniformly for $\theta \in [1,2]$, as we now show.

Lemma 33. *Let $u \in \mathbb{N}_{>0}^{m_u}$ and $v \in \mathbb{N}_{>0}^{m_v}$ with $M_v \ge 2$, let $\eta \in (0, 1/8]$ and let $P = 2^j$ with $j \in \mathbb{N}$. Suppose that $|c_u - \log(M_u)| \le \eta$, $|c_v - \log(M_v)| \le \eta$ and that $B_{u,P}$ and $B_{v,P}$ satisfy eq. (102). Then, for all $\theta \in [1,2]$, $\widetilde{D}_P(\theta)/\theta \ge 1/4 > 0$ and*

$$|\widetilde{R}_P(\theta) - R_{u,v}(P\theta)| \le (8 + 16S)\eta. \tag{109}$$

Proof. Let $\theta \in [1,2]$ and $p = P\theta$. By eqs. (100) and (101), $N_P(\theta)/\theta = \log(\|u\|_p)$ and $D_P(\theta)/\theta = \log(\|v\|_p)$. Since $P\theta \ge 1$, we have

$$\left| \frac{\widetilde{N}_P(\theta)}{\theta} - \log(\|u\|_p) \right| \le |c_u - \log(M_u)| + \frac{1}{P\theta} |B_{u,P}(\theta) - \log(Z_{u,P}(\theta))| \le 2\eta,$$

and similarly $|\widetilde{D}_P(\theta)/\theta - \log(\|v\|_p)| \le 2\eta$. By Lemma 29 and eq. (84), $\log(\|v\|_p) \ge \log(M_v) \ge \log(2) \ge 1/2$ and $0 \le \log(\|u\|_p) \le S$. Hence, $\widetilde{D}_P(\theta)/\theta \ge 1/2 - 2\eta \ge 1/4$. Lemma 28, applied with $n = \log(\|u\|_p)$, $\delta = \log(\|v\|_p)$, $\tilde{n} = \widetilde{N}_P(\theta)/\theta$, $\tilde{\delta} = \widetilde{D}_P(\theta)/\theta$ and $\tau = 2\eta \le 1/4$, gives $\widetilde{D}_P(\theta)/\theta \ge 1/4$ and eq. (109), since $(4 + 8S) \cdot 2\eta = (8 + 16S)\eta$. □

6.4 Certified minimisation of a rational function

It remains to minimise the rational function $\widetilde{R}_P$ over $[1,2]$. Polynomial and rational optimisation can be approached using methods of real algebraic geometry [35] or semidefinite programming [36]. For the univariate problem considered here, we give an explicit subroutine based on real root isolation. It returns a rational interval of prescribed width ζ containing the minimum, with an explicit Lipschitz bound controlling the required root approximation accuracy.

Since the numerator and denominator of $\widetilde{R}_P$ are polynomials with rational coefficients and its denominator is positive on $[1,2]$, its minimum is attained either at one of the endpoints $\theta \in \{1,2\}$ or at a real root in $(1,2)$ of the numerator of the derivative of $\widetilde{R}_P$, which we write

$$K_P = \widetilde{N}_P' \widetilde{D}_P - \widetilde{N}_P \widetilde{D}_P'. \tag{110}$$

This leads to Algorithm 2.

Algorithm 2 `RationalMinimum`(N, D, ζ)

Input: two polynomials $N, D \in \mathbb{Q}[X]$ such that $D(\theta) \geq \theta/4$ for all $\theta \in [1,2]$, and a rational number $\zeta > 0$.

Output: an interval $[\ell, h]$ with rational endpoints, of width at most ζ, containing $\min_{\theta\in[1,2]} N(\theta)/D(\theta)$.

1. Form $K \leftarrow N'D - ND'$. **If** $K = 0$, **return** $\big[N(1)/D(1),\ N(1)/D(1)\big]$.
2. Set $\Lambda \leftarrow 16 \cdot 2^{\deg(K)} \|K\|_{\ell_1}$, where $\|K\|_{\ell_1}$ denotes the sum of the absolute values of the coefficients of K, and $\kappa \leftarrow \max\big\{0, \lceil \log_2(2\Lambda/\zeta) \rceil\big\}$, so that $2^{-\kappa} \leq \zeta/(2\Lambda)$.
3. Compute the square-free part $K_{\mathrm{sf}} \leftarrow K/\gcd(K, K')$ and isolate its real roots with Lemma 27, with isolating intervals of width at most $2^{-\kappa}$.
4. Form the finite set of candidates
$$\Theta \leftarrow \{1,2\} \cup \big\{\max\{\alpha, 1\} \ \big|\ [\alpha,\beta] \text{ an isolating interval of step 3 meeting } [1,2]\big\} \subseteq [1,2].$$
5. Compute $m \leftarrow \min_{\theta\in\Theta} N(\theta)/D(\theta)$ in exact rational arithmetic and **return** $[m - \zeta/2,\ m]$.

Lemma 34. *Let $N, D \in \mathbb{Q}[X]$ be such that $D(\theta) \geq \theta/4$ for all $\theta \in [1,2]$, and let $\zeta > 0$ be rational. Then Algorithm 2 returns an interval of width at most ζ containing $\min_{\theta\in[1,2]} N(\theta)/D(\theta)$. Its bit complexity is polynomial in the degrees of N and D, in the binary length of their coefficients and in the binary description length of ζ.*

Proof. Let $\mu = \min_{\theta\in[1,2]} N(\theta)/D(\theta)$, which exists by continuity, since D does not vanish on $[1,2]$. The derivative of N/D is K/D^2, so that, if μ is attained at $\theta^\star \in (1,2)$, then $K(\theta^\star) = 0$. Hence μ is the minimum of the values of N/D at the endpoints $1, 2$ and at the real roots of K in $(1,2)$, which are the real roots of K_{sf} by Lemma 27(i). If $K = 0$, then N/D is constant on $[1,2]$ and $\mu = N(1)/D(1)$, which is returned at step 1.

Suppose now that $K \neq 0$. For $\theta \in [1,2]$ and a polynomial Q, we have $|Q(\theta)| \leq 2^{\deg(Q)}\|Q\|_{\ell_1}$, so that, since $D(\theta) \geq \theta/4 \geq 1/4$,
$$\left|\left(\frac{N}{D}\right)'(\theta)\right| = \frac{|K(\theta)|}{D(\theta)^2} \leq 16 \cdot 2^{\deg(K)}\|K\|_{\ell_1} = \Lambda \tag{111}$$
for all $\theta \in [1,2]$. In particular N/D is Λ-Lipschitz on $[1,2]$, and the binary length of Λ is polynomial in the degrees and in the coefficient lengths of N and D.

Since $\Theta \subseteq [1,2]$, we have $m \geq \mu$. Conversely, let $\theta^\star \in [1,2]$ be such that $N(\theta^\star)/D(\theta^\star) = \mu$. If $\theta^\star \in \{1,2\}$, then $m \leq \mu$. Otherwise $K(\theta^\star) = 0$, hence $K_{\mathrm{sf}}(\theta^\star) = 0$, so that $\theta^\star$ belongs to one of the isolating intervals, say $[\alpha_0, \beta_0]$, of step 3, which therefore meets $[1,2]$. The corresponding candidate $\theta_0 = \max\{\alpha_0, 1\}$ is the projection of α_0 onto $[1,2]$, and $\theta^\star \in [1,2]$, so that $|\theta_0 - \theta^\star| \leq |\alpha_0 - \theta^\star| \leq \beta_0 - \alpha_0 \leq 2^{-\kappa} \leq \zeta/(2\Lambda)$. Since N/D is Λ-Lipschitz on $[1,2]$, we have
$$\begin{aligned} 0 \leq m - \mu &\leq \frac{N(\theta_0)}{D(\theta_0)} - \frac{N(\theta^\star)}{D(\theta^\star)} \\ &\leq \Lambda|\theta_0 - \theta^\star| \leq \frac{\zeta}{2}. \end{aligned}$$
Hence $\mu \in [m - \zeta/2, m]$, which is the returned interval, of width $\zeta/2 \leq \zeta$.

For the complexity, the polynomial K, its derivative, the greatest common divisor $\gcd(K, K')$, the quotient K_{sf} and the rational number Λ can be computed in time polynomial in the degrees and coefficient lengths of N and D, and so can the binary lengths of their coefficients. By Lemma 27(ii), the root isolation takes time polynomial in the degree and coefficient length of K_{sf} and in $\kappa = O(1 + \log_2(\max\{1,\Lambda\}) + \log_2(\max\{1, 1/\zeta\}))$, where the maxima in the definition of κ ensure the logarithms are non-negative. Finally, there are at most $\deg(K) + 2$ candidates, and each evaluation of $N(\theta)/D(\theta)$ is an exact rational computation of polynomial size. □

6.5 The algorithm and the proof of Theorem 24

We are now ready to present the full algorithm, which combines the reduction of Section 6.2, the polynomial approximations of Algorithm 1 on each dyadic interval and Algorithm 2.

Algorithm 3 `ApproximateRate`(u, v, b)

Input: two vectors $u \in \mathbb{N}_{>0}^{m_u}$, $v \in \mathbb{N}_{>0}^{m_v}$ written in binary, with $v \neq (1)$, and an integer $b \in \mathbb{N}$.
Output: an interval $I = [a, z]$ with rational endpoints containing $R_*(u, v)$, of width at most 2^{-b}, and its midpoint $\widehat{R} = (a + z)/2$.

1. Set $\varepsilon \leftarrow 2^{-b}$, compute S as in eq. (83), and set
$$\eta \leftarrow \frac{\varepsilon}{256(S+1)}, \qquad k \leftarrow b + 8 + \lceil \log_2(S+1) \rceil, \qquad \text{so that } 2^{-k} \le \eta.$$
2. Set $\overline{L} \leftarrow 2S + 4S^2$ and $J \leftarrow b + 3 + \lceil \log_2(\overline{L}) \rceil$, so that $2^J \ge 8\overline{L}/\varepsilon$.
3. **If** $u = (1)$, **return** $I = [0, 0]$ and $\widehat{R} = 0$.
4. **If** $M_v = 1$, set $w \leftarrow \big\lfloor \log(\sum_i u_i) \big\rfloor_k / \big\lfloor \log(m_v) \big\rfloor_k$ and **return** $I = [w - \varepsilon/8, w + \varepsilon/8]$ and $\widehat{R} = w$.
5. Set $c_u \leftarrow \lfloor \log(M_u) \rfloor_k$, $c_v \leftarrow \lfloor \log(M_v) \rfloor_k$ and $w_\infty \leftarrow c_u / c_v$.
6. Initialise $a \leftarrow w_\infty - \varepsilon/4$ and $z \leftarrow w_\infty + \varepsilon/4$.
7. **For** $j = 0, \dots, J - 1$:
   (a) Set $B_u \leftarrow$ `LogSumPolynomial`$(u, 2^j, \eta)$ and $B_v \leftarrow$ `LogSumPolynomial`$(v, 2^j, \eta)$.
   (b) Form $\widetilde{N}(\theta) \leftarrow \theta c_u + 2^{-j} B_u(\theta)$ and $\widetilde{D}(\theta) \leftarrow \theta c_v + 2^{-j} B_v(\theta)$.
   (c) Compute $[\ell_j, h_j] \leftarrow$ `RationalMinimum`$(\widetilde{N}, \widetilde{D}, \varepsilon/4)$ and update
   $$a \leftarrow \min\{a,\ \ell_j - \varepsilon/8\}, \qquad z \leftarrow \min\{z,\ h_j + \varepsilon/8\}.$$
8. **Return** $I = [a, z]$ and $\widehat{R} = (a + z)/2$.

Proof of Theorem 24. We show that Algorithm 3 satisfies the requirements of Theorem 24. Actually, we show the stronger statement that it outputs an interval of width at most $\varepsilon/2$ containing $R_*(u, v)$, so that its midpoint $\widehat{R}$ satisfies $|\widehat{R} - R_*(u, v)| \le \varepsilon/4 \le 2^{-b}$.

Trivial cases. If $u = (1)$, then $\|u\|_p = 1$ for all p and $R_{u,v}(p) = 0$ for all p, so that $R_*(u, v) = 0$. If $M_v = 1$, then $v = (1, \dots, 1)$ with $m_v \ge 2$ since $v \neq (1)$, and $R_*(u, v) = \log(\sum_i u_i)/\log(m_v)$ by Lemma 31. By Lemma 28, applied with $\tau = 2^{-k} \le \eta \le 1/4$ and with the bounds $\log(\sum_i u_i) \le S$ of eq. (84) and $\log(m_v) \ge \log(2) \ge 1/2$, we get $|w - R_*(u, v)| \le (4 + 8S)\eta \le 8(S + 1)\eta = \varepsilon/32$. In both cases the returned interval contains $R_*(u, v)$ and has width at most $\varepsilon/4$.

General case. Suppose now that $M_v \ge 2$. Since $[1, +\infty] = [2^J, +\infty] \cup \bigcup_{j=0}^{J-1} [2^j, 2^{j+1}]$, we have

$$R_*(u, v) = \min\Big\{\mu_\infty,\ \min_{0 \le j \le J-1} \mu_j\Big\}, \quad \text{with} \quad \mu_\infty = \inf_{p \in [2^J, +\infty]} R_{u,v}(p) \quad \text{and} \quad \mu_j = \min_{\theta \in [1,2]} R_{u,v}(2^j \theta). \tag{112}$$

For $c \in \mathcal{C} = \{\infty\} \cup \{0, \dots, J - 1\}$, let us introduce the intervals $[\alpha_c, \beta_c]$ implicitly considered by the algorithm, namely

$$[\alpha_\infty, \beta_\infty] = \Big[w_\infty - \frac{\varepsilon}{4},\ w_\infty + \frac{\varepsilon}{4}\Big], \qquad [\alpha_j, \beta_j] = \Big[\ell_j - \frac{\varepsilon}{8},\ h_j + \frac{\varepsilon}{8}\Big] \quad (0 \le j \le J - 1),$$

where $[\ell_j, h_j]$ is the interval returned by `RationalMinimum` at step 7(c). By construction, the values returned at step 8 are $a = \min_{c \in \mathcal{C}} \alpha_c$ and $z = \min_{c \in \mathcal{C}} \beta_c$. We first show that, for every $c \in \mathcal{C}$, the interval $[\alpha_c, \beta_c]$ contains the corresponding term μ_c of eq. (112) and has width at most $\varepsilon/2$.

Let $c = \infty$. By Lemma 30 and by the choice of J, we have $|R_{u,v}(p) - R_{u,v}(\infty)| \le \overline{L}/p \le \overline{L}/2^J \le \varepsilon/8$ for all $p \ge 2^J$, so that $\mu_\infty \in [R_{u,v}(\infty) - \varepsilon/8,\ R_{u,v}(\infty)]$. Moreover, $R_{u,v}(\infty) = \log(M_u)/\log(M_v)$, and $w_\infty = c_u/c_v$ with $|c_u - \log(M_u)| \le 2^{-k} \le \eta$ and $|c_v - \log(M_v)| \le \eta$ by step 5. Since $0 \le \log(M_u) \le S$ by eq. (84), $\log(M_v) \ge \log(2) \ge 1/2$ and $\eta \le 1/256 \le 1/4$, Lemma 28 gives

$$|w_\infty - R_{u,v}(\infty)| \le (4 + 8S)\eta \le 8(S + 1)\eta = \frac{\varepsilon}{32}.$$

Hence $\mu_\infty \in [w_\infty - \varepsilon/4, w_\infty + \varepsilon/4] = [\alpha_\infty, \beta_\infty]$, which has width $\varepsilon/2$.

Let now $c = j \in \{0, \dots, J-1\}$ and $P = 2^j$, and let $\widetilde{N}, \widetilde{D}$ be the polynomials formed at step 7(b), which are those of eq. (108) for this value of P. By Lemma 32, B_u and B_v satisfy eq. (102), and $|c_u - \log(M_u)| \leq \eta$, $|c_v - \log(M_v)| \leq \eta$ as above, with $\eta \leq 1/8$. Hence Lemma 33 applies and gives $\widetilde{D}(\theta) \geq \theta/4$ on $[1, 2]$, as required by `RationalMinimum`, together with

$$\sup_{\theta \in [1,2]} \left| \frac{\widetilde{N}(\theta)}{\widetilde{D}(\theta)} - R_{u,v}(P\theta) \right| \leq (8 + 16S)\eta \leq \frac{16(S+1)\varepsilon}{256(S+1)} \leq \frac{\varepsilon}{8}.$$

Taking the minimum over $\theta \in [1, 2]$ on both sides, the minimum of $\widetilde{N}/\widetilde{D}$ on $[1, 2]$ is at distance at most $\varepsilon/8$ from μ_j. By Lemma 34, $[\ell_j, h_j]$ contains that minimum and has width at most $\varepsilon/4$. Hence $\mu_j \in [\ell_j - \varepsilon/8, h_j + \varepsilon/8] = [\alpha_j, \beta_j]$, which has width at most $\varepsilon/4 + \varepsilon/4 = \varepsilon/2$.

It remains to combine these enclosures. Since min is non-decreasing in each of its arguments, $\mu_c \in [\alpha_c, \beta_c]$ for every $c \in \mathcal{C}$ implies

$$R_*(u, v) = \min_{c \in \mathcal{C}} \mu_c \in \left[\min_{c \in \mathcal{C}} \alpha_c,\ \min_{c \in \mathcal{C}} \beta_c \right] = [a, z],$$

which is why both endpoints are updated by a minimum in step 7(c). Furthermore, if $a = \alpha_{c_0}$, then $z \leq \beta_{c_0} \leq \alpha_{c_0} + \varepsilon/2 = a + \varepsilon/2$. Therefore, the returned interval $[a, z]$ contains $R_*(u, v)$ and has width at most $\varepsilon/2$, independently of the number $J + 1$ of terms in eq. (112).

Complexity. The number of iterations of the loop is $J = O(\log_2(S) + b)$ by the choice of $\overline{L} = 2S + 4S^2$. For each $j \leq J - 1$, the integer j and $\log_2(1/\eta) = b + O(\log_2(S))$ are polynomial in S and b. By Lemma 32, the polynomials B_u and B_v are computed in time polynomial in S and b, and so are their degrees and the binary lengths of their coefficients. The same holds for $\widetilde{N}$ and $\widetilde{D}$, as c_u and c_v are computed in polynomial time [32]. Then, by Lemma 34, each call to Algorithm 2 runs in time polynomial in S and b. Finally, steps 1 to 6 and step 8 are clearly polynomial in S and b. □

Remark 35. As specified before Theorem 24, the shape vectors are given as explicit lists of entries, so that S counts their numbers of entries m_u and m_v, and the complexity bound is polynomial in these numbers rather than in their binary lengths. This is intrinsic to our algorithm rather than an artefact of the encoding: the number of terms $q = 4m_x(T - 1)$ retained from the series of the logarithm in Algorithm 1 is linear in m_x, because the argument $1 - Z_{\mathrm{pol}}(\theta)/(2m_x)$ of that series is only bounded away from 1 by $1/(4m_x)$; consequently $\deg(B) = O(m_x T^2)$, and grouping equal entries and writing their multiplicities in binary would not help. Theorem 24 therefore says nothing about succinct descriptions of shape vectors, such as $\lambda(\mathcal{F}^{\otimes k}) = \lambda(\mathcal{F})^{\otimes k}$ described by $\lambda(\mathcal{F})$ and k, which has m^k entries, with m the number of entries of $\lambda(\mathcal{F})$. This is not a limitation for the computation of emulation capacities: by Proposition 7, $R_*(\lambda(\mathcal{G}^{\otimes n}), \lambda(\mathcal{F}^{\otimes k})) = \frac{n}{k} R_*(\lambda(\mathcal{G}), \lambda(\mathcal{F}))$, so that only the shape vectors of $\mathcal{F}$ and $\mathcal{G}$ themselves are ever needed, and these have at most $\dim(\mathcal{H}_\mathcal{F})$ and $\dim(\mathcal{H}_\mathcal{G})$ entries by Proposition 1.

7 Entanglement-assisted emulation of idempotent channels

In this section, we prove Theorem B. The proof relies on the following observation: in the presence of unlimited shared entanglement, every idempotent channel $\mathcal{F}$ is *exactly* interconvertible, with a single use in each direction, with the noiseless classical channel Δ_s on an alphabet of size $s = \|\lambda(\mathcal{F})\|_2^2$ (Corollary 39). The two directions of this interconversion are given by the teleportation [37] and the superdense coding [38] protocols, applied block by block to the decomposition of Proposition 1. This reduces the entanglement-assisted emulation of idempotent channels to the entanglement-assisted emulation of noiseless classical channels with one another, which we solve exactly at finite blocklength in Lemma 41. In contrast with the unassisted and shared randomness regimes, where the whole family of ℓ_p-norms of the shape vectors is needed, the emulation is thus governed by the single quantity $\|\lambda(\mathcal{F})\|_2$. This is reminiscent of the quantum reverse Shannon theorem [5, 6], according to which the asymptotic simulation of a channel, in the presence of free entanglement, is governed by a single quantity, as well as of its one-shot refinement in terms of the smooth max-information [21]. Theorem B is an exact, zero-error counterpart of this phenomenon for idempotent channels.

We prove item (ii) of Theorem B first, and deduce items (i) and (iii) from it. Note that eq. (4) involves the squares of the ℓ_2-norms of the shape vectors $\lambda(\mathcal{F})$ and $\lambda(\mathcal{G})$.

Throughout this section, we use the following notation. Given an idempotent channel $\mathcal{F}$, we write

$$s(\mathcal{F}) = \|\lambda(\mathcal{F})\|_2^2 = \sum_{k=1}^{K} d_k^2, \tag{113}$$

where $\lambda(\mathcal{F}) = (d_1, \ldots, d_K)$. Note that $s(\mathcal{F})$ is a positive integer, and that $s(\mathcal{F}) = 1$ if and only if $\lambda(\mathcal{F}) = (1)$, since the entries of $\lambda(\mathcal{F})$ are positive integers. Furthermore, for $M \in \mathbb{N}_{>0}$, we regard the completely dephasing channel $\Delta_M : \mathcal{L}(\mathbb{C}^M) \to \mathcal{L}(\mathbb{C}^M)$ as a noiseless classical channel on the alphabet $[M]$. Recall that $\lambda(\Delta_M) = (1, \ldots, 1) \in \mathbb{N}^M$, so that $s(\Delta_M) = M$, and that $\lambda(\mathrm{Id}_d) = (d)$, so that $s(\mathrm{Id}_d) = d^2$.

7.1 Entanglement-assisted protocols and exact emulations

We start by fixing some terminology on entanglement-assisted protocols, which allows us to compose them easily.

An entanglement-assisted protocol is a triple $\mathsf{P} = (\omega_{A_0B_0}, \mathcal{E}, \mathcal{D})$, where $\omega_{A_0B_0} \in \mathcal{D}(\mathcal{H}_{A_0} \otimes \mathcal{H}_{B_0})$ is a bipartite state on systems A_0, B_0 of arbitrary finite dimension, $\mathcal{E} : \mathcal{L}(\mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{A_0}) \to \mathcal{L}(\mathcal{H}'_{\mathrm{in}})$ is an encoding channel and $\mathcal{D} : \mathcal{L}(\mathcal{H}'_{\mathrm{out}} \otimes \mathcal{H}_{B_0}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}})$ is a decoding channel, and where $\mathcal{H}_{\mathrm{in}}$, $\mathcal{H}_{\mathrm{out}}$, $\mathcal{H}'_{\mathrm{in}}$ and $\mathcal{H}'_{\mathrm{out}}$ are Hilbert spaces. $\mathsf{P} = (\omega_{A_0B_0}, \mathcal{E}, \mathcal{D})$ maps $\mathcal{L}(\mathcal{H}'_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}'_{\mathrm{out}})$ to $\mathcal{L}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}})$. Given $\Psi : \mathcal{L}(\mathcal{H}'_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}'_{\mathrm{out}})$, we write

$$\mathsf{P}[\Psi] = \mathcal{D} \circ (\Psi \otimes \mathrm{Id}_{B_0}) \circ (\mathcal{E} \otimes \mathrm{Id}_{B_0}) \circ \mathcal{T}_{\omega_{A_0B_0}} : \mathcal{L}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}}), \tag{114}$$

where $\mathcal{T}_{\omega_{A_0B_0}} : x \mapsto x \otimes \omega_{A_0B_0}$ for $x \in \mathcal{L}(\mathcal{H}_{\mathrm{in}})$. If Ψ is a channel, $\mathsf{P}[\Psi]$ is a channel as well, which we call the channel *induced* by the protocol P from Ψ. Throughout, whenever we write $\mathsf{P}[\Psi]$, it is understood that the input and output spaces of Ψ are those that the protocol P expects, namely $\mathcal{L}(\mathcal{H}'_{\mathrm{in}})$ and $\mathcal{L}(\mathcal{H}'_{\mathrm{out}})$. The same convention applies to iterated expressions such as $\mathsf{P}_2[\mathsf{P}_1[\Theta]]$. We say that a channel Φ can be *exactly emulated* by a channel Ψ with shared entanglement when $\Phi = \mathsf{P}[\Psi]$ for some entanglement-assisted protocol P.

Entanglement-assisted protocols compose: given $\mathsf{P}_1 = (\omega_{A_1B_1}, \mathcal{E}_1, \mathcal{D}_1)$ and $\mathsf{P}_2 = (\omega_{A_2B_2}, \mathcal{E}_2, \mathcal{D}_2)$, and a channel Θ, we have $\mathsf{P}_2[\mathsf{P}_1[\Theta]] = \mathsf{P}_3[\Theta]$ for the entanglement-assisted protocol $\mathsf{P}_3 = (\omega_{A_1B_1} \otimes \omega_{A_2B_2}, \mathcal{E}_3, \mathcal{D}_3)$ given by

$$\mathcal{E}_3 = \mathcal{E}_1 \circ (\mathcal{E}_2 \otimes \mathrm{Id}_{A_1}), \qquad \mathcal{D}_3 = \mathcal{D}_2 \circ (\mathcal{D}_1 \otimes \mathrm{Id}_{B_2}), \tag{115}$$

where the shared state is on the systems $A_0 = A_2A_1$ and $B_0 = B_1B_2$. This follows by expanding both sides with eq. (114), identifying as is customary the tensor products of Hilbert spaces that only differ by the order of their factors. In particular, the shared state of the composed protocol is the tensor product of the shared states of P_1 and P_2.

With this notation, for any channels $\Phi : \mathcal{L}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}})$ and $\Psi : \mathcal{L}(\mathcal{H}'_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}'_{\mathrm{out}})$, we write

$$\delta^{\mathrm{EA}}_{1,1}(\Psi \mapsto \Phi) = \inf_{\mathsf{P}} \frac{1}{2} \|\Phi - \mathsf{P}[\Psi]\|_\diamond, \tag{116}$$

where the infimum is taken over all the entanglement-assisted protocols P from maps $\mathcal{L}(\mathcal{H}'_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}'_{\mathrm{out}})$ to $\mathcal{L}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}})$. Then, for $k, n \in \mathbb{N}_{>0}$, by definition

$$\delta^{\mathrm{EA}}_{k,n}(\Psi \mapsto \Phi) = \delta^{\mathrm{EA}}_{1,1}(\Psi^{\otimes n} \mapsto \Phi^{\otimes k}). \tag{117}$$

The maps induced by an entanglement-assisted protocol contract the diamond norm.

Lemma 36. *Let $\mathsf{P} = (\omega_{A_0B_0}, \mathcal{E}, \mathcal{D})$ be an entanglement-assisted protocol. Then, for every pair of linear maps Θ_1, Θ_2 from $\mathcal{L}(\mathcal{H}'_{in})$ to $\mathcal{L}(\mathcal{H}'_{out})$,*

$$\|\mathsf{P}[\Theta_1] - \mathsf{P}[\Theta_2]\|_\diamond \leq \|\Theta_1 - \Theta_2\|_\diamond. \tag{118}$$

Proof. By linearity, $\mathsf{P}[\Theta_1] - \mathsf{P}[\Theta_2] = \mathcal{D} \circ ((\Theta_1 - \Theta_2) \otimes \mathrm{Id}_{B_0}) \circ (\mathcal{E} \otimes \mathrm{Id}_{B_0}) \circ \mathcal{T}_{\omega_{A_0B_0}}$. The maps $\mathcal{D}$, $\mathcal{E} \otimes \mathrm{Id}_{B_0}$ and $\mathcal{T}_{\omega_{A_0B_0}}$ are channels, so that their diamond norm is equal to 1, and the diamond norm is submultiplicative under composition and stable under tensor product with identity maps, i.e. $\|(\Theta_1 - \Theta_2) \otimes \mathrm{Id}_{B_0}\|_\diamond = \|\Theta_1 - \Theta_2\|_\diamond$ (see e.g. [27, Chapter 3]). Eq. (118) follows. □

As a consequence, the minimal error of entanglement-assisted emulation is monotone under exact entanglement-assisted emulation.

Lemma 37. *Let Φ, Φ', Ψ and Ψ' be channels, and let P_1, P_2 be entanglement-assisted protocols such that $\Phi' = \mathsf{P}_1[\Phi]$ and $\Psi = \mathsf{P}_2[\Psi']$. Then*

$$\delta^{\mathrm{EA}}_{1,1}(\Psi' \mapsto \Phi') \leq \delta^{\mathrm{EA}}_{1,1}(\Psi \mapsto \Phi). \tag{119}$$

Proof. Let P be an arbitrary entanglement-assisted protocol. By eq. (115), applied twice, there is an entanglement-assisted protocol P' such that $\mathsf{P}_1[\mathsf{P}[\Psi]] = \mathsf{P}_1[\mathsf{P}[\mathsf{P}_2[\Psi']]] = \mathsf{P}'[\Psi']$. Then, by Lemma 36,

$$\delta^{\mathrm{EA}}_{1,1}(\Psi' \mapsto \Phi') \leq \frac{1}{2}\|\Phi' - \mathsf{P}'[\Psi']\|_\diamond = \frac{1}{2}\|\mathsf{P}_1[\Phi] - \mathsf{P}_1[\mathsf{P}[\Psi]]\|_\diamond \leq \frac{1}{2}\|\Phi - \mathsf{P}[\Psi]\|_\diamond.$$

Taking the infimum over P yields eq. (119). □

7.2 The operational meaning of $\|\lambda(\mathcal{F})\|_2^2$

By Proposition 3, an idempotent channel $\mathcal{F}$ can be converted into the block pinching $\Pi_{\mathbf{f}}$ associated to its shape vector, and back. As far as emulation is concerned, we may therefore work with block pinchings throughout this section, and it suffices to understand their entanglement-assisted emulation. We now show that the block pinching $\Pi_{\mathbf{d}}$ of Definition 3 is exactly interconvertible with the noiseless classical channel Δ_S, with $S = \|\mathbf{d}\|_2^2$, when shared entanglement is allowed.

We first recall the teleportation and superdense coding protocols, in a form adapted to our needs. Let $d \geq 1$ be an integer and let $\{|\nu\rangle\}_{\nu\in\{0,\dots,d-1\}}$ be the standard basis of $\mathbb{C}^d$, whose indices are taken modulo d. We consider the *Heisenberg–Weyl* unitaries

$$U_{a,b} = X^a Z^b, \quad (a,b) \in \{0,\dots,d-1\}^2, \quad \text{with} \quad X|\nu\rangle = |\nu+1\rangle, \quad Z|\nu\rangle = e^{2i\pi\nu/d}|\nu\rangle, \tag{120}$$

which satisfy the orthogonality relations $\mathrm{tr}(U^*_{a,b}U_{a',b'}) = d\,\delta_{a,a'}\delta_{b,b'}$. Here X and Z are the d-dimensional generalisations of the Pauli matrices σ_x and σ_z, to which they reduce for $d = 2$; note that they are unitary but not Hermitian for $d \geq 3$. As in Definition 16, we write $|\phi_d\rangle = \frac{1}{\sqrt{d}}\sum_{\nu=0}^{d-1}|\nu\rangle\otimes|\nu\rangle \in \mathbb{C}^d\otimes\mathbb{C}^d$ for the maximally entangled state in dimension d and, for $(a,b)\in\{0,\dots,d-1\}^2$, we let

$$|\phi_{a,b}\rangle = (U_{a,b}\otimes\mathbb{1})|\phi_d\rangle. \tag{121}$$

Since, for every $x, y \in \mathcal{L}(\mathbb{C}^d)$, $\langle\phi_d|(x\otimes\mathbb{1})^*(y\otimes\mathbb{1})|\phi_d\rangle = \frac{1}{d}\mathrm{tr}(x^*y)$, the orthogonality relations of the unitaries $U_{a,b}$ imply that $\{|\phi_{a,b}\rangle\}_{a,b}$ is an orthonormal basis of $\mathbb{C}^d\otimes\mathbb{C}^d$, called the *Bell basis*. In particular, $\sum_{a,b}|\phi_{a,b}\rangle\langle\phi_{a,b}| = \mathbb{1}_{\mathbb{C}^d\otimes\mathbb{C}^d}$.

Lemma 38 (Teleportation identity)**.** *Let X, A, B be three systems of dimension d, identified with $\mathbb{C}^d$, and let R be an arbitrary reference system. Then, for all $x \in \mathcal{L}(\mathcal{H}_X\otimes\mathcal{H}_R)$ and all $(a,b)\in\{0,\dots,d-1\}^2$,*

$$(U_{a,b})_B\,(\langle\phi_{a,b}|_{XA}\otimes\mathbb{1}_{BR})\big(x_{XR}\otimes|\phi_d\rangle\langle\phi_d|_{AB}\big)(|\phi_{a,b}\rangle_{XA}\otimes\mathbb{1}_{BR})\,(U_{a,b})^*_B = \frac{1}{d^2}\,x_{BR}, \tag{122}$$

where x_{BR} denotes the operator x in which the system X is relabelled B.

Proof. By linearity, it suffices to prove eq. (122) for $x = |\psi\rangle\langle\varphi|$ with $|\psi\rangle, |\varphi\rangle \in \mathcal{H}_X\otimes\mathcal{H}_R$, and even to prove that for every $|\psi\rangle\in\mathcal{H}_X\otimes\mathcal{H}_R$,

$$(U_{a,b})_B(\langle\phi_{a,b}|_{XA}\otimes\mathbb{1}_{BR})\,|\psi\rangle_{XR}\,|\phi_d\rangle_{AB} = \frac{1}{d}\,|\psi\rangle_{BR}.$$

Let us first consider the case $(a,b) = (0,0)$. Writing $|\psi\rangle = \sum_\nu|\nu\rangle_X\otimes w_\nu$, where the $w_\nu\in\mathcal{H}_R$ are vectors which need not be normalised, we get

$$(\langle\phi_d|_{XA}\otimes\mathbb{1}_{BR})\,|\psi\rangle_{XR}\,|\phi_d\rangle_{AB} = \frac{1}{d}\sum_{\nu,\mu,\kappa}\langle\mu|\nu\rangle_X\langle\mu|\kappa\rangle_A\,|\kappa\rangle_B\otimes w_\nu = \frac{1}{d}\sum_\nu|\nu\rangle_B\otimes w_\nu = \frac{1}{d}\,|\psi\rangle_{BR}.$$

In the general case, $\langle\phi_{a,b}|_{XA} = \langle\phi_d|_{XA}\,((U_{a,b})^*_X \otimes \mathbb{1}_A)$, so that, applying the previous case to the vector $(U_{a,b})^*_X\,|\psi\rangle$, we obtain

$$(\langle\phi_{a,b}|_{XA} \otimes \mathbb{1}_{BR})\,|\psi\rangle_{XR}\,|\phi_d\rangle_{AB} = \frac{1}{d}(U_{a,b})^*_B\,|\psi\rangle_{BR}\,,$$

and the unitary correction $(U_{a,b})_B$ yields the claim. □

Summing eq. (122) over the d^2 outcomes (a,b) shows that the teleportation protocol, i.e. a measurement of XA in the Bell basis followed by the correction $U_{a,b}$ on B, implements the identity channel from X to B, even on inputs which are entangled with an arbitrary reference system R. We now use teleportation and superdense coding block by block.

Proposition 10. *Let $\mathbf{d} = (d_1, \ldots, d_K) \in \mathbb{N}^K_{>0}$ and $S = \|\mathbf{d}\|_2^2 = \sum_{k=1}^K d_k^2$. Then there are entanglement-assisted protocols P and P', both with the shared state $\bigotimes_{k=1}^K |\phi_{d_k}\rangle\langle\phi_{d_k}|_{A_kB_k}$, such that*

$$\Pi_{\mathbf{d}} = \mathsf{P}[\Delta_S] \qquad \text{and} \qquad \Delta_S = \mathsf{P}'[\Pi_{\mathbf{d}}], \tag{123}$$

the protocol P being the teleportation protocol and P' the superdense coding protocol, applied block by block.

Proof. We identify the classical alphabet $[S]$ with the disjoint union

$$\Sigma = \bigsqcup_{k=1}^K \{k\} \times \{0, \ldots, d_k - 1\}^2 = \{(k,a,b) \mid k \in [K],\ (a,b) \in \{0,\ldots,d_k-1\}^2\},$$

which has cardinality $\sum_k d_k^2 = S$, so that Δ_S is the completely dephasing channel on $\mathcal{L}(\mathbb{C}^\Sigma)$ in the basis $\{|k,a,b\rangle\}_{(k,a,b)\in\Sigma}$. In both directions, Alice and Bob share the state $\omega_{A_0B_0} = \bigotimes_{k=1}^K |\phi_{d_k}\rangle\langle\phi_{d_k}|_{A_kB_k}$ with $A_0 = A_1 \cdots A_K$, $B_0 = B_1 \cdots B_K$ and $\mathcal{H}_{A_k} = \mathcal{H}_{B_k} = \mathbb{C}^{d_k}$. For $l \in [K]$, we write $A_{\bar{l}}$ for the systems A_m with $m \neq l$, and likewise $B_{\bar{l}}$. For $k \in [K]$ and $(a,b) \in \{0,\ldots,d_k-1\}^2$, we write $U^{(k)}_{a,b}$ and $|\phi^{(k)}_{a,b}\rangle$ for the Heisenberg–Weyl unitaries and the Bell basis in dimension d_k, and, with $J_k : \mathbb{C}^{d_k} \to \mathcal{H}_{\mathbf{d}}$ the isometry embedding $\mathbb{C}^{d_k}$ as the k-th summand of $\mathcal{H}_{\mathbf{d}}$ as in Definition 3, we define the orthogonal projector on $\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_{A_0}$

$$M_{k,a,b} = (J_k \otimes \mathbb{1}_{A_k})|\phi^{(k)}_{a,b}\rangle\langle\phi^{(k)}_{a,b}|(J_k \otimes \mathbb{1}_{A_k})^* \otimes \mathbb{1}_{A_{\bar{k}}}. \tag{124}$$

Then $\sum_{a,b} M_{k,a,b} = (J_kJ_k^* \otimes \mathbb{1}_{A_k}) \otimes \mathbb{1}_{A_{\bar{k}}} = \widetilde{P}_k \otimes \mathbb{1}_{A_0}$, since the Bell basis is an orthonormal basis, and thus $\sum_{(k,a,b)\in\Sigma} M_{k,a,b} = \mathbb{1}_{\mathcal{H}_{\mathbf{d}}\otimes\mathcal{H}_{A_0}}$, i.e. $\{M_{k,a,b}\}_{(k,a,b)\in\Sigma}$ is a projective measurement. We define in the same way the projective measurement $\{N_{k,a,b}\}_{(k,a,b)\in\Sigma}$ on $\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_{B_0}$ by replacing the systems A_l by B_l.

Emulation of $\Pi_{\mathbf{d}}$ by Δ_S (teleportation). Alice measures her input together with A_0 with the measurement $\{M_{k,a,b}\}$, which first reveals the block k of the input and then teleports its content with the pair A_kB_k, and she sends the outcome through Δ_S. Bob applies the teleportation correction on B_k and embeds B_k in the k-th block. Formally, let $\mathcal{E} : \mathcal{L}(\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_{A_0}) \to \mathcal{L}(\mathbb{C}^\Sigma)$ and $\mathcal{D} : \mathcal{L}(\mathbb{C}^\Sigma \otimes \mathcal{H}_{B_0}) \to \mathcal{L}(\mathcal{H}_{\mathbf{d}})$ be defined by

$$\begin{aligned}
\mathcal{E}(y) &= \sum_{(k,a,b)\in\Sigma} \operatorname{tr}(M_{k,a,b}\,y)|k,a,b\rangle\langle k,a,b|,\\
\mathcal{D}(y) &= \sum_{(k,a,b)\in\Sigma} J_kU^{(k)}_{a,b}\operatorname{tr}_{B_{\bar{k}}}\big((\langle k,a,b| \otimes \mathbb{1}_{B_0})y(|k,a,b\rangle \otimes \mathbb{1}_{B_0})\big)U^{(k)*}_{a,b}J_k^*.
\end{aligned}$$

The map $\mathcal{E}$ is a channel, since $\{M_{k,a,b}\}$ is a POVM, and $\mathcal{D}$ is a channel as it first measures the classical register in the basis $\{|k,a,b\rangle\}$ and then applies, conditionally on the outcome, a channel from $\mathcal{L}(\mathcal{H}_{B_0})$ to $\mathcal{L}(\mathcal{H}_{\mathbf{d}})$. Let R be a reference system and $x \in \mathcal{L}(\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_R)$. Since the output of $\mathcal{E}$ is diagonal, $\Delta_S \circ \mathcal{E} = \mathcal{E}$, and we get

$$\big(\mathsf{P}[\Delta_S] \otimes \mathrm{Id}_R\big)(x) = \sum_{(k,a,b)\in\Sigma} (J_kU^{(k)}_{a,b} \otimes \mathbb{1}_R)\,T_{k,a,b}\,(J_kU^{(k)}_{a,b} \otimes \mathbb{1}_R)^*,$$

with $\mathsf{P} = (\omega_{A_0B_0}, \mathcal{E}, \mathcal{D})$ and $T_{k,a,b} = \mathrm{tr}_{\mathcal{H}_{\mathbf{d}} A_0 B_{\bar{k}}}\big((M_{k,a,b} \otimes \mathbb{1}_{B_0 R})(x \otimes \omega_{A_0B_0})\big) \in \mathcal{L}(\mathcal{H}_{B_k} \otimes \mathcal{H}_R)$. Now, $M_{k,a,b}$ acts as the identity on the systems $A_{\bar{k}}$, and $\mathrm{tr}_{A_{\bar{k}}B_{\bar{k}}}(\bigotimes_{l\neq k} |\phi_{d_l}\rangle\langle\phi_{d_l}|) = 1$. Hence, writing $x_{kk} = (J_k^* \otimes \mathbb{1}_R)x(J_k \otimes \mathbb{1}_R) \in \mathcal{L}(\mathbb{C}^{d_k} \otimes \mathcal{H}_R)$ and abbreviating

$$V_{k,a,b} = J_k U^{(k)}_{a,b} \otimes \mathbb{1}_R, \qquad \Lambda_{k,a,b} = \langle \phi^{(k)}_{a,b}|_{XA_k} \otimes \mathbb{1}_{B_k R},$$

where $X \cong \mathbb{C}^{d_k}$ denotes the k-th block of the input, the (k,a,b) term of the sum reads

$$V_{k,a,b}\,\Lambda_{k,a,b}\big(x_{kk} \otimes |\phi_{d_k}\rangle\langle\phi_{d_k}|_{A_kB_k}\big)\Lambda^*_{k,a,b}\,V^*_{k,a,b} = \frac{1}{d_k^2}(J_k \otimes \mathbb{1}_R)x_{kk}(J_k \otimes \mathbb{1}_R)^*,$$

where the equality follows from Lemma 38. Summing over the d_k^2 pairs (a,b) and over $k \in [K]$, we obtain

$$\big(\mathsf{P}[\Delta_S] \otimes \mathrm{Id}_R\big)(x) = \sum_{k=1}^{K} (\widetilde{P}_k \otimes \mathbb{1}_R)x(\widetilde{P}_k \otimes \mathbb{1}_R) = (\Pi_{\mathbf{d}} \otimes \mathrm{Id}_R)(x),$$

that is, $\Pi_{\mathbf{d}} = \mathsf{P}[\Delta_S]$.

Emulation of Δ_S by $\Pi_{\mathbf{d}}$ (superdense coding). To send the symbol (k,a,b), Alice applies $U^{(k)}_{a,b}$ on her half A_k of the k-th maximally entangled state and sends A_k in the k-th block of $\Pi_{\mathbf{d}}$. Bob measures the received system together with B_0 with the measurement $\{N_{k,a,b}\}$. Formally, let $\mathcal{E}' : \mathcal{L}(\mathbb{C}^\Sigma \otimes \mathcal{H}_{A_0}) \to \mathcal{L}(\mathcal{H}_{\mathbf{d}})$ and $\mathcal{D}' : \mathcal{L}(\mathcal{H}_{\mathbf{d}} \otimes \mathcal{H}_{B_0}) \to \mathcal{L}(\mathbb{C}^\Sigma)$ be defined by

$$\mathcal{E}'(y) = \sum_{(k,a,b)\in\Sigma} J_k U^{(k)}_{a,b}\,\mathrm{tr}_{A_{\bar{k}}}\big((\langle k,a,b| \otimes \mathbb{1}_{A_0})y(|k,a,b\rangle \otimes \mathbb{1}_{A_0})\big)U^{(k)*}_{a,b}J_k^*,$$
$$\mathcal{D}'(y) = \sum_{(k,a,b)\in\Sigma} \mathrm{tr}(N_{k,a,b}\,y)|k,a,b\rangle\langle k,a,b|,$$

which are channels for the same reasons as $\mathcal{E}$ and $\mathcal{D}$, and let $\mathsf{P}' = (\omega_{A_0B_0}, \mathcal{E}', \mathcal{D}')$. Since $\mathcal{E}'$ starts by measuring its classical input in the basis $\{|k,a,b\rangle\}$ and $\mathcal{D}'$ outputs diagonal operators, $\mathsf{P}'[\Pi_{\mathbf{d}}]$ satisfies $\mathsf{P}'[\Pi_{\mathbf{d}}] = \Delta_S \circ \mathsf{P}'[\Pi_{\mathbf{d}}] \circ \Delta_S$. Hence, it suffices to show that $\mathsf{P}'[\Pi_{\mathbf{d}}](|k,a,b\rangle\langle k,a,b|) = |k,a,b\rangle\langle k,a,b|$ for every $(k,a,b) \in \Sigma$. Indeed, in that case, for every reference system R and every $x = \sum_{\sigma,\sigma'\in\Sigma} |\sigma\rangle\langle\sigma'| \otimes x_{\sigma\sigma'} \in \mathcal{L}(\mathbb{C}^\Sigma \otimes \mathcal{H}_R)$,

$$(\mathsf{P}'[\Pi_{\mathbf{d}}] \otimes \mathrm{Id}_R)(x) = \sum_{\sigma\in\Sigma} \mathsf{P}'[\Pi_{\mathbf{d}}](|\sigma\rangle\langle\sigma|) \otimes x_{\sigma\sigma} = \sum_{\sigma\in\Sigma} |\sigma\rangle\langle\sigma| \otimes x_{\sigma\sigma} = (\Delta_S \otimes \mathrm{Id}_R)(x).$$

Let us fix $(k,a,b) \in \Sigma$. We have

$$(\mathcal{E}' \otimes \mathrm{Id}_{B_0})(|k,a,b\rangle\langle k,a,b| \otimes \omega_{A_0B_0}) = (J_k \otimes \mathbb{1}_{B_k})|\phi^{(k)}_{a,b}\rangle\langle\phi^{(k)}_{a,b}|_{A_kB_k}(J_k \otimes \mathbb{1}_{B_k})^* \otimes \bigotimes_{l\neq k} \frac{\mathbb{1}_{B_l}}{d_l},$$

where we used $(U^{(k)}_{a,b} \otimes \mathbb{1})\,|\phi_{d_k}\rangle = |\phi^{(k)}_{a,b}\rangle$ and $\mathrm{tr}_{A_l}(|\phi_{d_l}\rangle\langle\phi_{d_l}|) = \mathbb{1}_{B_l}/d_l$. This operator is invariant under $\Pi_{\mathbf{d}} \otimes \mathrm{Id}_{B_0}$, as it is supported on the k-th block of $\mathcal{H}_{\mathbf{d}}$. Finally, for $(k',a',b') \in \Sigma$,

$$\mathrm{tr}\Big(N_{k',a',b'}\,(J_k \otimes \mathbb{1}_{B_k})|\phi^{(k)}_{a,b}\rangle\langle\phi^{(k)}_{a,b}|(J_k \otimes \mathbb{1}_{B_k})^* \otimes \bigotimes_{l\neq k} \frac{\mathbb{1}_{B_l}}{d_l}\Big) = \delta_{k,k'}|\langle\phi^{(k)}_{a',b'}|\phi^{(k)}_{a,b}\rangle|^2 = \delta_{k,k'}\delta_{a,a'}\delta_{b,b'},$$

where we used $J_{k'}^* J_k = \delta_{k,k'}\mathbb{1}$ and the orthonormality of the Bell basis. Hence $\mathsf{P}'[\Pi_{\mathbf{d}}](|k,a,b\rangle\langle k,a,b|) = |k,a,b\rangle\langle k,a,b|$, which shows that $\Delta_S = \mathsf{P}'[\Pi_{\mathbf{d}}]$. □

Combining Proposition 3 and Proposition 10 with the composability of entanglement-assisted protocols in eq. (115), we obtain the following corollary.

Corollary 39. *Let $\mathcal{F}$ be an idempotent channel and let $s(\mathcal{F}) = \|\lambda(\mathcal{F})\|_2^2$. Then there are entanglement-assisted protocols Q and Q' such that*

$$\mathcal{F} = \mathsf{Q}[\Delta_{s(\mathcal{F})}] \qquad \text{and} \qquad \Delta_{s(\mathcal{F})} = \mathsf{Q}'[\mathcal{F}]. \tag{125}$$

Remark 40. Corollary 39 explains why the *squares* d_k^2 of the entries of the shape vector, rather than the entries d_k themselves or other powers of them, govern the entanglement-assisted regime. A block of dimension d_k can carry d_k^2 classical symbols with superdense coding, and it can be simulated by the transmission of d_k^2 classical symbols with teleportation. In contrast, without entanglement, a d_k-dimensional block can only carry d_k classical symbols, and it cannot be simulated at all by a classical channel when $d_k \geq 2$, as shown by the fourth row of Table 2.

7.3 Entanglement-assisted emulation of noiseless classical channels

By Corollary 39, it remains to understand the entanglement-assisted emulation of a noiseless classical channel by another one. The following lemma provides the exact minimal error for this task. Its converse part is a no-signalling argument: it only uses the fact that the local operations of the encoder do not change the marginal state of the decoder, so that it also holds for the larger class of protocols assisted by arbitrary no-signalling correlations, for which linear and semidefinite programming characterisations are available at finite blocklength [22, 23]. Zero-error capacities and exact simulation costs of classical channels assisted by non-local correlations were determined along these lines in [24].

Lemma 41. *Let $M, N \in \mathbb{N}_{>0}$. Then*

$$\delta_{1,1}^{\mathrm{EA}}(\Delta_N \mapsto \Delta_M) = \max\Big\{0, 1 - \frac{N}{M}\Big\}. \tag{126}$$

Furthermore, if $N \geq M$, then $\Delta_M = \mathsf{P}[\Delta_N]$ for an unassisted protocol P.

Proof. Achievability. Suppose first that $N \geq M$ and let $\mathcal{E} : \mathcal{L}(\mathbb{C}^M) \to \mathcal{L}(\mathbb{C}^N)$ and $\mathcal{D} : \mathcal{L}(\mathbb{C}^N) \to \mathcal{L}(\mathbb{C}^M)$ be the channels

$$\mathcal{E}(x) = \sum_{m\in[M]} \langle m|\, x\, |m\rangle\, |m\rangle\langle m|, \qquad \mathcal{D}(y) = \sum_{z\in[M]} \langle z|\, y\, |z\rangle\, |z\rangle\langle z| + \sum_{z=M+1}^{N} \langle z|\, y\, |z\rangle\, |1\rangle\langle 1|.$$

The encoder $\mathcal{E}$ embeds $[M]$ into $[N]$, so that the symbols $z \in \{M+1, \ldots, N\}$ are never transmitted and the second sum defining $\mathcal{D}$, which sends them to the default symbol 1, never contributes. Hence $\mathcal{D} \circ \Delta_N \circ \mathcal{E} = \Delta_M$, so that $\Delta_M = \mathsf{P}[\Delta_N]$ for the unassisted protocol P with encoder $\mathcal{E}$ and decoder $\mathcal{D}$, and $\delta_{1,1}^{\mathrm{EA}}(\Delta_N \mapsto \Delta_M) = 0$.

Suppose now that $N < M$. We use shared randomness, which is a special case of shared entanglement (see Section 2.1). Let $\mathfrak{S}_M$ be the set of permutations of $[M]$ and, for $\pi \in \mathfrak{S}_M$, let $f_\pi : [M] \to [N]$ be defined by $f_\pi(m) = \pi(m)$ if $\pi(m) \leq N$ and $f_\pi(m) = 1$ otherwise. We define the channels $\mathcal{E}_\pi : \mathcal{L}(\mathbb{C}^M) \to \mathcal{L}(\mathbb{C}^N)$ and $\mathcal{D}_\pi : \mathcal{L}(\mathbb{C}^N) \to \mathcal{L}(\mathbb{C}^M)$ by

$$\mathcal{E}_\pi(x) = \sum_{m\in[M]} \langle m|\, x\, |m\rangle\, |f_\pi(m)\rangle\langle f_\pi(m)|, \qquad \mathcal{D}_\pi(y) = \sum_{z\in[N]} \langle z|\, y\, |z\rangle\, |\pi^{-1}(z)\rangle\langle \pi^{-1}(z)|.$$

In words, Alice and Bob agree on a uniformly random permutation π; Alice sends $\pi(m)$ when it belongs to $[N]$ and the symbol 1 otherwise, and Bob outputs π^{-1} of the received symbol. Let $\widetilde{\Delta} = \frac{1}{M!}\sum_{\pi\in\mathfrak{S}_M} \mathcal{D}_\pi \Delta_N \mathcal{E}_\pi$ be the channel induced by this shared randomness strategy. For $m \in [M]$, we have $\mathcal{D}_\pi \Delta_N \mathcal{E}_\pi(|m\rangle\langle m|) = |\pi^{-1}(f_\pi(m))\rangle\langle \pi^{-1}(f_\pi(m))|$, which is equal to $|m\rangle\langle m|$ if $\pi(m) \leq N$ and to $|\pi^{-1}(1)\rangle\langle\pi^{-1}(1)|$ with $\pi^{-1}(1) \neq m$ otherwise (since then $\pi(m) > N \geq 1$). Splitting the average over $\mathfrak{S}_M$ according to whether $\pi(m) \leq N$ or $\pi(m) > N$, we obtain

$$\widetilde{\Delta}(|m\rangle\langle m|) = \frac{1}{M!} \sum_{\substack{\pi\in\mathfrak{S}_M \\ \pi(m)\leq N}} |m\rangle\langle m| \;+\; \frac{1}{M!} \sum_{\substack{\pi\in\mathfrak{S}_M \\ \pi(m)> N}} |\pi^{-1}(1)\rangle\langle\pi^{-1}(1)|.$$

As $\pi(m)$ is uniformly distributed on $[M]$ when π is uniformly distributed on $\mathfrak{S}_M$, the first sum has weight N/M and the second has weight $1 - N/M$, so that

$$\widetilde{\Delta}(|m\rangle\langle m|) = \frac{N}{M}|m\rangle\langle m| + \Big(1 - \frac{N}{M}\Big)\tau_m, \qquad \tau_m = \frac{1}{|\{\pi \mid \pi(m) > N\}|} \sum_{\substack{\pi\in\mathfrak{S}_M \\ \pi(m)> N}} |\pi^{-1}(1)\rangle\langle\pi^{-1}(1)|.$$

Then, τ_m is a diagonal state satisfying $\langle m|\,\tau_m\,|m\rangle = 0$, since $\pi^{-1}(1) \neq m$ for every π occurring in the sum. Now, since each $\mathcal{E}_\pi$ first dephases its input, $\widetilde{\Delta} = \widetilde{\Delta} \circ \Delta_M$. Let R be a reference system and $\rho \in \mathcal{D}(\mathbb{C}^M \otimes \mathcal{H}_R)$. Then $(\Delta_M \otimes \mathrm{Id}_R)(\rho) = \sum_{m\in[M]} |m\rangle\langle m| \otimes \rho_m$ with $\rho_m = (\langle m| \otimes \mathbb{1}_R)\rho(|m\rangle \otimes \mathbb{1}_R) \geq 0$ and $\sum_{m\in[M]} \mathrm{tr}(\rho_m) = 1$, so that

$$\begin{aligned}
\frac{1}{2}\|((\Delta_M - \widetilde{\Delta}) \otimes \mathrm{Id}_R)(\rho)\|_1 &= \frac{1}{2}\|((\Delta_M - \widetilde{\Delta}) \circ \Delta_M \otimes \mathrm{Id}_R)(\rho)\|_1 \\
&= \frac{1}{2}\Big\| \sum_{m\in[M]} (\Delta_M - \widetilde{\Delta})(|m\rangle\langle m|) \otimes \rho_m \Big\|_1 \\
&\leq \sum_{m\in[M]} \mathrm{tr}(\rho_m)\, \frac{1}{2}\Big(1 - \frac{N}{M}\Big) \||m\rangle\langle m| - \tau_m\|_1 \\
&= 1 - \frac{N}{M},
\end{aligned}$$

where the first equality uses $\widetilde{\Delta} = \widetilde{\Delta} \circ \Delta_M$ together with the idempotence of Δ_M, and where we then used the triangle inequality, $\|y \otimes \rho_m\|_1 = \|y\|_1 \,\mathrm{tr}(\rho_m)$ and $\||m\rangle\langle m| - \tau_m\|_1 = 2$ since $|m\rangle\langle m|$ and τ_m are states with orthogonal supports. Since the diamond norm of the Hermitian-preserving map $\Delta_M - \widetilde{\Delta}$ is attained on states (see e.g. [27, Chapter 3]), we obtain $\frac{1}{2}\|\Delta_M - \widetilde{\Delta}\|_\diamond \leq 1 - N/M$, so that $\delta^{\mathrm{EA}}_{1,1}(\Delta_N \mapsto \Delta_M) \leq \delta^{\mathrm{SR}}_{1,1}(\Delta_N \mapsto \Delta_M) \leq 1 - N/M$.

Converse. Let $\mathsf{P} = (\omega_{A_0B_0}, \mathcal{E}, \mathcal{D})$ be an arbitrary entanglement-assisted protocol, with $\mathcal{E} : \mathcal{L}(\mathbb{C}^M \otimes \mathcal{H}_{A_0}) \to \mathcal{L}(\mathbb{C}^N)$ and $\mathcal{D} : \mathcal{L}(\mathbb{C}^N \otimes \mathcal{H}_{B_0}) \to \mathcal{L}(\mathbb{C}^M)$, and let $\widetilde{\Delta} = \mathsf{P}[\Delta_N]$. We first show that the average probability that $\widetilde{\Delta}$ transmits correctly a uniformly distributed symbol is at most N/M, that is

$$\frac{1}{M} \sum_{m\in[M]} \langle m|\,\widetilde{\Delta}(|m\rangle\langle m|)\,|m\rangle \leq \frac{N}{M}. \tag{127}$$

For $m \in [M]$, let $Y_m = (\mathcal{E} \otimes \mathrm{Id}_{B_0})(|m\rangle\langle m| \otimes \omega_{A_0B_0}) \in \mathcal{L}(\mathbb{C}^N \otimes \mathcal{H}_{B_0})$ and, for $z \in [N]$, let $\sigma_{m,z} = (\langle z| \otimes \mathbb{1}_{B_0})Y_m(|z\rangle \otimes \mathbb{1}_{B_0}) \in \mathcal{L}(\mathcal{H}_{B_0})$. Since Y_m is positive semidefinite, so are the operators $\sigma_{m,z}$, and

$$(\Delta_N \otimes \mathrm{Id}_{B_0})(Y_m) = \sum_{z\in[N]} |z\rangle\langle z| \otimes \sigma_{m,z}.$$

Furthermore, for every $w \in \mathcal{L}(\mathbb{C}^M \otimes \mathcal{H}_{A_0} \otimes \mathcal{H}_{B_0})$, we have $\mathrm{tr}_{\mathbb{C}^N}\big((\mathcal{E} \otimes \mathrm{Id}_{B_0})(w)\big) = \mathrm{tr}_{\mathbb{C}^M A_0}(w)$: indeed, for every $v \in \mathcal{L}(\mathcal{H}_{B_0})$, $\mathrm{tr}\big(\mathrm{tr}_{\mathbb{C}^N}((\mathcal{E} \otimes \mathrm{Id}_{B_0})(w))\, v\big) = \mathrm{tr}\big((\mathcal{E} \otimes \mathrm{Id}_{B_0})(w)\,(\mathbb{1}_{\mathbb{C}^N} \otimes v)\big) = \mathrm{tr}\big(w\,(\mathcal{E}^*(\mathbb{1}_{\mathbb{C}^N}) \otimes v)\big) = \mathrm{tr}\big(\mathrm{tr}_{\mathbb{C}^M A_0}(w)\, v\big)$, where the last equality uses that $\mathcal{E}^*$ is unital, which is equivalent to $\mathcal{E}$ being trace preserving. Applying this with $w = |m\rangle\langle m| \otimes \omega_{A_0B_0}$, we get

$$\sum_{z\in[N]} \sigma_{m,z} = \mathrm{tr}_{\mathbb{C}^N}(Y_m) = \mathrm{tr}_{\mathbb{C}^M A_0}(|m\rangle\langle m| \otimes \omega_{A_0B_0}) = \omega_{B_0}, \tag{128}$$

where $\omega_{B_0} = \mathrm{tr}_{A_0}(\omega_{A_0B_0})$. This equation expresses the no-signalling principle: the local operation of Alice does not change the marginal state of Bob. In particular, $0 \leq \sigma_{m,z} \leq \omega_{B_0}$ for all $m \in [M]$ and $z \in [N]$. Now, for $z \in [N]$ and $m \in [M]$, let

$$D^{(z)}_m = (\langle z| \otimes \mathbb{1}_{B_0})\mathcal{D}^*(|m\rangle\langle m|)(|z\rangle \otimes \mathbb{1}_{B_0}) \in \mathcal{L}(\mathcal{H}_{B_0}).$$

Since $\mathcal{D}^*$ is unital and completely positive, $D^{(z)}_m \geq 0$ and $\sum_{m\in[M]} D^{(z)}_m = (\langle z| \otimes \mathbb{1}_{B_0})\mathcal{D}^*(\mathbb{1})(|z\rangle \otimes \mathbb{1}_{B_0}) = \mathbb{1}_{B_0}$, i.e. $\{D^{(z)}_m\}_{m\in[M]}$ is the POVM that Bob applies on B_0 when he receives the symbol z. Then, for $m \in [M]$,

$$\langle m|\,\widetilde{\Delta}(|m\rangle\langle m|)\,|m\rangle = \mathrm{tr}\Big(\mathcal{D}^*(|m\rangle\langle m|) \sum_{z\in[N]} |z\rangle\langle z| \otimes \sigma_{m,z}\Big) = \sum_{z\in[N]} \mathrm{tr}(D^{(z)}_m \sigma_{m,z}) \leq \sum_{z\in[N]} \mathrm{tr}(D^{(z)}_m \omega_{B_0}),$$

where the inequality follows from $\sigma_{m,z} \leq \omega_{B_0}$ and $D^{(z)}_m \geq 0$. Summing over m, we get

$$\sum_{m\in[M]} \langle m|\,\widetilde{\Delta}(|m\rangle\langle m|)\,|m\rangle \leq \sum_{z\in[N]} \mathrm{tr}\Big(\sum_{m\in[M]} D^{(z)}_m\, \omega_{B_0}\Big) = \sum_{z\in[N]} \mathrm{tr}(\omega_{B_0}) = N,$$

which proves eq. (127).

We then apply Lemma 14 to the channels Δ_M and $\widetilde{\Delta}$, with the auxiliary system $\mathcal{H}_R = \mathbb{C}^M$ and

$$\sigma = \frac{1}{M} \sum_{m\in[M]} |m\rangle\langle m| \otimes |m\rangle\langle m|_R, \qquad \mu = \sum_{m\in[M]} |m\rangle\langle m| \otimes |m\rangle\langle m|_R.$$

Since μ is an orthogonal projector, $0 \leq \mu \leq \mathbb{1}$. Furthermore, $(\Delta_M \otimes \mathrm{Id}_R)(\sigma) = \sigma$, so that $\mathrm{tr}(\mu(\Delta_M \otimes \mathrm{Id}_R)(\sigma)) = \mathrm{tr}(\mu\sigma) = 1$, and

$$\mathrm{tr}(\mu(\widetilde{\Delta} \otimes \mathrm{Id}_R)(\sigma)) = \frac{1}{M} \sum_{m\in[M]} \langle m| \widetilde{\Delta}(|m\rangle\langle m|) |m\rangle \leq \frac{N}{M},$$

by eq. (127). Hence, by Lemma 14, $\frac{1}{2}\|\Delta_M - \mathsf{P}[\Delta_N]\|_\diamond \geq 1 - N/M$. As this holds for every entanglement-assisted protocol P, and as the diamond norm is non-negative, we get $\delta^{\mathrm{EA}}_{1,1}(\Delta_N \mapsto \Delta_M) \geq \max\{0, 1 - N/M\}$, which ends the proof. □

Note that the achievability part of Lemma 41 only uses shared randomness, and that the converse holds with arbitrary shared entanglement, and even with arbitrary no-signalling correlations [22]. Hence, for noiseless classical channels, shared entanglement does not help compared to shared randomness, and the minimal error in eq. (126) is also the minimal error with shared randomness.

7.4 Proof of Theorem B

We are now ready to prove the two main results of this section.

Proof of item (ii) of Theorem B. Let $\mathcal{F}$ and $\mathcal{G}$ be idempotent channels and $k, n \in \mathbb{N}_{>0}$. By Proposition 7, $\lambda(\mathcal{F}^{\otimes k}) = \lambda(\mathcal{F})^{\otimes k}$, so that, by the multiplicativity of the ℓ_2-norm under tensor products,

$$s(\mathcal{F}^{\otimes k}) = \|\lambda(\mathcal{F})^{\otimes k}\|_2^2 = \|\lambda(\mathcal{F})\|_2^{2k} = s(\mathcal{F})^k,$$

and similarly $s(\mathcal{G}^{\otimes n}) = s(\mathcal{G})^n$. By Corollary 39 applied to the idempotent channels $\mathcal{F}^{\otimes k}$ and $\mathcal{G}^{\otimes n}$, there are entanglement-assisted protocols $\mathsf{Q}_\mathcal{F}$, $\mathsf{Q}'_\mathcal{F}$, $\mathsf{Q}_\mathcal{G}$ and $\mathsf{Q}'_\mathcal{G}$ such that

$$\mathcal{F}^{\otimes k} = \mathsf{Q}_\mathcal{F}[\Delta_{s(\mathcal{F})^k}], \quad \Delta_{s(\mathcal{F})^k} = \mathsf{Q}'_\mathcal{F}[\mathcal{F}^{\otimes k}], \quad \mathcal{G}^{\otimes n} = \mathsf{Q}_\mathcal{G}[\Delta_{s(\mathcal{G})^n}], \quad \Delta_{s(\mathcal{G})^n} = \mathsf{Q}'_\mathcal{G}[\mathcal{G}^{\otimes n}]. \tag{129}$$

Therefore, by eq. (117), Lemma 37 and Lemma 41,

$$\begin{aligned}\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) &= \delta^{\mathrm{EA}}_{1,1}(\mathcal{G}^{\otimes n} \mapsto \mathcal{F}^{\otimes k}) = \delta^{\mathrm{EA}}_{1,1}(\Delta_{s(\mathcal{G})^n} \mapsto \Delta_{s(\mathcal{F})^k}) \\ &= \max\left\{0, 1 - \frac{s(\mathcal{G})^n}{s(\mathcal{F})^k}\right\} = \max\left\{0, 1 - \frac{\|\lambda(\mathcal{G})\|_2^{2n}}{\|\lambda(\mathcal{F})\|_2^{2k}}\right\}.\end{aligned}$$

□

Note that the proof of item (ii) of Theorem B also shows that, when $\|\lambda(\mathcal{G})\|_2^{2n} \geq \|\lambda(\mathcal{F})\|_2^{2k}$, the minimal error $\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) = 0$ is *attained*, i.e. there is an entanglement-assisted protocol P such that $\mathcal{F}^{\otimes k} = \mathsf{P}[\mathcal{G}^{\otimes n}]$. Indeed, in that case, $s(\mathcal{G})^n \geq s(\mathcal{F})^k$, so that $\Delta_{s(\mathcal{F})^k} = \mathsf{R}[\Delta_{s(\mathcal{G})^n}]$ for some entanglement-assisted protocol R by Lemma 41. Composing with eq. (129) gives $\mathcal{F}^{\otimes k} = \mathsf{Q}_\mathcal{F}[\mathsf{R}[\mathsf{Q}'_\mathcal{G}[\mathcal{G}^{\otimes n}]]]$, and the right-hand side is of the form $\mathsf{P}[\mathcal{G}^{\otimes n}]$ by eq. (115).

Remark 42. The protocol P above consumes only maximally entangled states: by eq. (115), its shared state is the tensor product of the shared states of $\mathsf{Q}_\mathcal{F}$, R and $\mathsf{Q}'_\mathcal{G}$, which are tensor products of finitely many maximally entangled states of finite dimension by Proposition 10, the protocol R being unassisted. Hence the exact interconversion of idempotent channels never requires entanglement beyond a finite supply of maximally entangled states, a feature of interest from the point of view of resource theories, where such states are the standard unit of entanglement.

Proof of item (i) of Theorem B. Let $\mathcal{F}$ and $\mathcal{G}$ be idempotent channels with $\lambda(\mathcal{F}) \neq (1)$, so that $s(\mathcal{F}) \geq 2$, and let

$$R = \frac{\log(s(\mathcal{G}))}{\log(s(\mathcal{F}))} = \frac{\log(\|\lambda(\mathcal{G})\|_2)}{\log(\|\lambda(\mathcal{F})\|_2)},$$

where we used $\log(s(\mathcal{G})) = 2\log(\|\lambda(\mathcal{G})\|_2)$ and $\log(s(\mathcal{F})) = 2\log(\|\lambda(\mathcal{F})\|_2) > 0$.

Converse. Let r be an achievable rate in the entanglement-assisted regime, and let $(k_\nu)_\nu$ and $(n_\nu)_\nu$ be two sequences of positive integers satisfying eq. (10) with $\delta_\nu = \delta^{\mathrm{EA}}_{k_\nu,n_\nu}(\mathcal{G} \mapsto \mathcal{F}) \to 0$. By item (ii) of Theorem B, $s(\mathcal{G})^{n_\nu}/s(\mathcal{F})^{k_\nu} \geq 1 - \delta_\nu$ for every ν. For ν large enough, $\delta_\nu < 1$, and, taking the logarithm and dividing by $n_\nu \log(s(\mathcal{F})) > 0$,

$$\frac{k_\nu}{n_\nu} \leq R - \frac{\log(1-\delta_\nu)}{n_\nu \log(s(\mathcal{F}))}.$$

The last term vanishes as $\nu \to \infty$, so that $r \leq \liminf_\nu k_\nu/n_\nu \leq R$. Hence $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) \leq R$, and the zero-error entanglement-assisted emulation capacity is at most R as well, by Section 2.1.

Achievability with zero error. By the discussion following the proof of item (ii) of Theorem B, for $k, n \in \mathbb{N}_{>0}$, $\mathcal{F}^{\otimes k}$ can be emulated exactly by $\mathcal{G}^{\otimes n}$ with shared entanglement as soon as $s(\mathcal{G})^n \geq s(\mathcal{F})^k$, that is as soon as $k/n \leq R$. If $\lambda(\mathcal{G}) = (1)$, then $R = 0$ and $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) = 0$, so that both capacities vanish, in accordance with eq. (3). Otherwise, $R > 0$ and, by density of the rational numbers in the reals, the supremum of the ratios $k/n \leq R$ is R, so that the zero-error entanglement-assisted emulation capacity is at least R. Since it is at most $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) \leq R$, both capacities are equal to R, which proves eq. (3) and the last statement of item (i). □

As in the unassisted and shared randomness regimes (see the end of Section 5), the finite blocklength result of item (ii) directly yields the strong converse of item (iii), here with an explicit exponent. No optimisation over p is needed, since the only relevant value is $p = 2$.

Proof of item (iii) of Theorem B. Let $\mathcal{F}$ and $\mathcal{G}$ be two idempotent channels such that $\lambda(\mathcal{F}) \neq (1)$, so that $\|\lambda(\mathcal{F})\|_2 > 1$, and let $\varepsilon > 0$. Throughout this proof, we take the logarithm in base 2. Let $k, n \in \mathbb{N}_{>0}$ with $k/n \geq C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) + \varepsilon$. Multiplying this inequality by $n\log(\|\lambda(\mathcal{F})\|_2) > 0$ and using eq. (3), we get

$$k\log(\|\lambda(\mathcal{F})\|_2) \geq n\log(\|\lambda(\mathcal{G})\|_2) + \varepsilon n \log(\|\lambda(\mathcal{F})\|_2),$$

and thus, by eq. (4),

$$\delta^{\mathrm{EA}}_{k,n}(\mathcal{G} \mapsto \mathcal{F}) \geq 1 - \frac{\|\lambda(\mathcal{G})\|_2^{2n}}{\|\lambda(\mathcal{F})\|_2^{2k}} = 1 - 2^{2(n\log(\|\lambda(\mathcal{G})\|_2) - k\log(\|\lambda(\mathcal{F})\|_2))} \geq 1 - 2^{-cn},$$

with $c = 2\varepsilon\log(\|\lambda(\mathcal{F})\|_2) > 0$, which proves item (iii). □

Since eq. (4) is an equality, the exponent $c = 2\varepsilon\log(\|\lambda(\mathcal{F})\|_2)$ is optimal: for $k/n = C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F}) + \varepsilon$, the error is exactly $1 - 2^{-cn}$.

7.5 Consequences and examples

We now draw some consequences of Theorem B, and compare them with the unassisted and shared randomness regimes. First, the entanglement-assisted emulation capacity is reversible and additive.

Corollary 43. *Let $\mathcal{F}$, $\mathcal{G}_1$ and $\mathcal{G}_2$ be idempotent channels with $\lambda(\mathcal{F}) \neq (1)$.*

(i) If $\lambda(\mathcal{G}_1) \neq (1)$, then $C_{\mathrm{EA}}(\mathcal{G}_1 \mapsto \mathcal{F})\, C_{\mathrm{EA}}(\mathcal{F} \mapsto \mathcal{G}_1) = 1$.

(ii) $C_{\mathrm{EA}}(\mathcal{G}_1 \otimes \mathcal{G}_2 \mapsto \mathcal{F}) = C_{\mathrm{EA}}(\mathcal{G}_1 \mapsto \mathcal{F}) + C_{\mathrm{EA}}(\mathcal{G}_2 \mapsto \mathcal{F})$.

Proof. (i) follows directly from eq. (3), since $\log(\|\lambda(\mathcal{F})\|_2)$ and $\log(\|\lambda(\mathcal{G}_1)\|_2)$ are both positive. For (ii), $\mathcal{G}_1 \otimes \mathcal{G}_2$ is idempotent and, by Proposition 7, $\lambda(\mathcal{G}_1 \otimes \mathcal{G}_2) = \lambda(\mathcal{G}_1) \otimes \lambda(\mathcal{G}_2)$, so that $\log(\|\lambda(\mathcal{G}_1 \otimes \mathcal{G}_2)\|_2) = \log(\|\lambda(\mathcal{G}_1)\|_2) + \log(\|\lambda(\mathcal{G}_2)\|_2)$, and (ii) follows from eq. (3). □

Second, exact emulation with a single use of the source channel is governed by a single integer.

Corollary 44. *Let $\mathcal{F}$ and $\mathcal{G}$ be idempotent channels. Then there is an entanglement-assisted protocol P such that $\mathcal{F} = \mathsf{P}[\mathcal{G}]$ if and only if $\|\lambda(\mathcal{F})\|_2 \leq \|\lambda(\mathcal{G})\|_2$. In particular, there are entanglement-assisted protocols P, P' with $\mathcal{F} = \mathsf{P}[\mathcal{G}]$ and $\mathcal{G} = \mathsf{P}'[\mathcal{F}]$ if and only if $\|\lambda(\mathcal{F})\|_2 = \|\lambda(\mathcal{G})\|_2$.*

Proof. If $\mathcal{F} = \mathsf{P}[\mathcal{G}]$ for some entanglement-assisted protocol P, then $\delta^{\mathrm{EA}}_{1,1}(\mathcal{G} \mapsto \mathcal{F}) = 0$, so that $\|\lambda(\mathcal{G})\|_2^2 \geq \|\lambda(\mathcal{F})\|_2^2$ by item (ii) of Theorem B with $k = n = 1$. Conversely, if $\|\lambda(\mathcal{F})\|_2 \leq \|\lambda(\mathcal{G})\|_2$, then such a protocol exists by the discussion following the proof of item (ii) of Theorem B. The second statement follows by applying the first one in both directions. □

Corollary 44 shows that, in the presence of shared entanglement, idempotent channels are totally preordered by exact emulation with a single use, and that their equivalence classes are labelled by the single integer $\|\lambda(\mathcal{F})\|_2^2$. This contrasts with the unassisted regime, where inequalities $\|\lambda(\mathcal{F})\|_p \leq \|\lambda(\mathcal{G})\|_p$ for every $p \in [1, +\infty]$ are necessary for exact unassisted emulation after taking a common tensor power, by item (ii) of Theorem A. For instance, let $\mathcal{G}$ be an idempotent channel with $\lambda(\mathcal{G}) = (2, 2, 1)$. Then $\|\lambda(\mathrm{Id}_3)\|_\infty = 3 > 2 = \|\lambda(\mathcal{G})\|_\infty$ and $\|\lambda(\mathcal{G})\|_1 = 5 > 3 = \|\lambda(\mathrm{Id}_3)\|_1$, so that neither of Id_3 and $\mathcal{G}$ can be emulated with zero error by one use of the other, even with shared randomness. However, $\|\lambda(\mathrm{Id}_3)\|_2^2 = 9 = \|\lambda(\mathcal{G})\|_2^2$, so that Id_3 and $\mathcal{G}$ can be exactly emulated by each other with shared entanglement, by Corollary 44.

As in the unassisted case, we gather in Table 3 the expressions of $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F})$ when either $\mathcal{F}$ or $\mathcal{G}$ is an identity or a completely dephasing channel, which follow directly from item (i) of Theorem B and from $\|\lambda(\mathrm{Id}_d)\|_2 = d$, $\|\lambda(\Delta_d)\|_2 = \sqrt{d}$. In particular, for every idempotent channel $\mathcal{G}$, the entanglement-assisted quantum and classical capacities of $\mathcal{G}$ (in qubits and bits per channel use) are respectively $C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathrm{Id}_2) = \log_2(\|\lambda(\mathcal{G})\|_2)$ and $C_{\mathrm{EA}}(\mathcal{G} \mapsto \Delta_2) = 2\log_2(\|\lambda(\mathcal{G})\|_2)$, as expected from teleportation and superdense coding.

$\mathcal{F}$	$\mathcal{G}$	$C_{\mathrm{EA}}(\mathcal{G} \mapsto \mathcal{F})$
Id_d	$\mathcal{G}$	$\frac{\log(\|\lambda(\mathcal{G})\|_2)}{\log(d)}$
$\mathcal{F}$	Id_d	$\frac{\log(d)}{\log(\|\lambda(\mathcal{F})\|_2)}$
Δ_d	$\mathcal{G}$	$\frac{2\log(\|\lambda(\mathcal{G})\|_2)}{\log(d)}$
$\mathcal{F}$	Δ_d	$\frac{\log(d)}{2\log(\|\lambda(\mathcal{F})\|_2)}$

Table 3: Entanglement-assisted emulation capacities when either $\mathcal{F}$ or $\mathcal{G}$ is the identity or the completely dephasing channel (with $d \geq 2$ and $\lambda(\mathcal{F}) \neq (1)$).

The following examples illustrate how shared entanglement modifies the emulation capacities of idempotent channels.

Example 45 (Emulation of a quantum channel by a classical one)**.** Let $\mathcal{F} = \mathrm{Id}_2$ and $\mathcal{G} = \Delta_2$, so that $\lambda(\mathcal{F}) = (2)$ and $\lambda(\mathcal{G}) = (1, 1)$. As computed in Example 9, $C_{\mathrm{UA}}(\Delta_2 \mapsto \mathrm{Id}_2) = C_{\mathrm{SR}}(\Delta_2 \mapsto \mathrm{Id}_2) = 0$ and $C_{\mathrm{UA}}(\mathrm{Id}_2 \mapsto \Delta_2) = 1$: a noiseless qubit channel cannot be emulated with zero error by a classical channel, even with shared randomness. In contrast, by Table 3,

$$C_{\mathrm{EA}}(\Delta_2 \mapsto \mathrm{Id}_2) = \frac{\log(\sqrt{2})}{\log(2)} = \frac{1}{2}, \qquad C_{\mathrm{EA}}(\mathrm{Id}_2 \mapsto \Delta_2) = \frac{2\log(2)}{\log(2)} = 2.$$

Operationally, two uses of Δ_2 transmit the two bits needed to teleport a qubit, and one use of Id_2 transmits two bits with superdense coding. Shared entanglement is known to strictly increase zero-error rates already for classical channels, both in the one-shot [39] and in the asymptotic [40] regimes; this example is an extreme instance of this phenomenon, since the emulation capacity jumps from 0 to 1/2. Note that the product of these two capacities is 1, in accordance with Corollary 43 (i), whereas the corresponding product is 0 in the unassisted regime.

Example 46 (Complementary source channels)**.** Let $\mathcal{G}_1 = \mathrm{Id}_2$, $\mathcal{G}_2 = \Delta_2$ and $\mathcal{F} = \mathcal{G}_1 \otimes \mathcal{G}_2$, so that

$$\lambda(\mathcal{G}_1) = (2), \qquad \lambda(\mathcal{G}_2) = (1, 1), \qquad \lambda(\mathcal{F}) = \lambda(\mathcal{G}_1) \otimes \lambda(\mathcal{G}_2) = (2, 2).$$

As computed in Example 8, the superadditivity of Corollary 7 is strict for these channels. With shared entanglement, since $\|\lambda(\mathcal{G}_1)\|_2^2 = 4$, $\|\lambda(\mathcal{G}_2)\|_2^2 = 2$ and $\|\lambda(\mathcal{F})\|_2^2 = 8$, we get instead

$$C_{\mathrm{EA}}(\mathcal{G}_1 \mapsto \mathcal{F}) = \frac{\log(4)}{\log(8)} = \frac{2}{3}, \qquad C_{\mathrm{EA}}(\mathcal{G}_2 \mapsto \mathcal{F}) = \frac{\log(2)}{\log(8)} = \frac{1}{3}, \qquad C_{\mathrm{EA}}(\mathcal{G}_1 \otimes \mathcal{G}_2 \mapsto \mathcal{F}) = 1,$$

which is additive, in accordance with Corollary 43 (ii). Entanglement therefore restores additivity here: the classical source channel $\mathcal{G}_2$, which was useless on its own without entanglement, now contributes the missing third of the rate.

Acknowledgements

We thank Robert Salzmann for helpful discussions about [25] and [19], as well as about sets of approximate fixed points of channels. We thank Satvik Singh and Arturo Konderak for pointing out a mistake in Proposition 1 in the first version of the present article. We also thank Arturo Konderak for drawing our attention to [28].

ID and OF acknowledge funding by the European Research Council (ERC Grant AlgoQIP, Agreement No. 851716). LG acknowledges funding by National Natural Science Foundation of China (NSFC) via grant No. 12401163. This work started when MR was a Marie Skłodowska-Curie Fellow at ENS de Lyon, and he acknowledges funding from the European Union's Horizon Europe research and innovation programme, call HORIZON-MSCA-2022-PF-01 (Project and agreement number: 101108117). MR is partially supported by the Wallenberg Centre for Quantum Technology (WACQT) funded by the Knut and Alice Wallenberg Foundation (KAW).

Use of AI tools

This article supersedes two earlier preprint versions on arXiv, as well as a short version presented at ISIT 2026 [1], which were prepared without the assistance of generative AI, except for minor polishing. In those versions, we established a weaker form of item (ii) of Theorem A where the infimum over $p \in [1, +\infty]$ was restricted to the set $\{1, +\infty\}$, leaving the full interval as an open problem. Considering recent developments in the use of AI to assist, if not automate, mathematical research, we tasked Anthropic's Claude Opus 5 with solving it. The model produced the statement and proofs of Lemmas 20, 22, and 23 (without shared randomness), as well as the ansatz in Definition 16.

Furthermore, we asked OpenAI's Astra model to find the algorithms of Section 6, as well as to prove their correctness and to analyse their complexity, which it achieved. Interestingly, Astra itself defined the relevant computational model. Astra also solved the entanglement-assisted regime. Finally, most of the writing of Sections 6 and 7 was left to Anthropic's Claude Opus 5, based on the material produced by OpenAI's Astra model.

The authors independently verified all AI-generated steps and assume full responsibility for the mathematical correctness of the results.

Marginally, Google's Gemini Flash was used to assist with the linguistic revision and polishing of the remaining, human-written parts of the manuscript.

References

[1] Idris Delsol, Omar Fawzi, Li Gao, and Mizanur Rahaman. "Emulation capacity between idempotent channels". In 2026 IEEE International Symposium on Information Theory (ISIT). Pages 1–6. IEEE (2026).

[2] Claude E. Shannon. "A mathematical theory of communication". The Bell System Technical Journal **27**, 379–423; 623–656 (1948).

[3] A. S. Holevo. "The capacity of the quantum channel with general signal states". IEEE Transactions on Information Theory **44**, 269–273 (1998). arXiv:quant-ph/9611023.

[4] Benjamin Schumacher and Michael D. Westmoreland. "Sending classical information via noisy quantum channels". Physical Review A **56**, 131–138 (1997).

[5] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem". IEEE Transactions on Information Theory **48**, 2637–2655 (2002). arXiv:quant-ph/0106052.

[6] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels". IEEE Transactions on Information Theory **60**, 2926–2959 (2014). arXiv:0912.5537.

[7] Mario Berta, Matthias Christandl, and Renato Renner. "The quantum reverse Shannon theorem based on one-shot information theory". Communications in Mathematical Physics **306**, 579–615 (2011). arXiv:0912.3805.

[8] Michael X. Cao, Navneeth Ramakrishnan, Mario Berta, and Marco Tomamichel. "Channel simulation: Finite blocklengths and broadcast channels". IEEE Transactions on Information Theory **70**, 6780–6808 (2024). arXiv:2212.11666.

[9] Yunchao Liu and Xiao Yuan. "Operational resource theory of quantum channels". Physical Review Research **2**, 012035 (2020). arXiv:1904.02680.

[10] Khaled Ben Dana, María García Díaz, Mohamed Mejatty, and Andreas Winter. "Resource theory of coherence: beyond states". Physical Review A **95**, 062327 (2017). arXiv:1704.03710.

[11] María García Díaz, Kun Fang, Xin Wang, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Andreas Winter. "Using and reusing coherence to realize quantum processes". Quantum **2**, 100 (2018). arXiv:1805.04045.

[12] Felix Leditzky, Debbie Leung, Vikesh Siddhu, Graeme Smith, and John A. Smolin. "The platypus of the quantum channel zoo". IEEE Transactions on Information Theory **69**, 3825–3849 (2023). arXiv:2202.08380.

[13] Holger Boche and Christian Deppe. "Computability of the zero-error capacity of noisy channels". In 2021 IEEE Information Theory Workshop (ITW). Pages 1–6. IEEE (2021). arXiv:2010.06873.

[14] Greg Kuperberg. "The capacity of hybrid quantum memory". IEEE Transactions on Information Theory **49**, 1465–1473 (2003). arXiv:quant-ph/0203105.

[15] Michael M. Wolf. "Quantum Channels & Operations. Guided Tour". Lecture notes, TUM. `https://mediatum.ub.tum.de/node?id=1701036` (2012).

[16] Omar Fawzi, Mizanur Rahaman, and Mostafa Taheri. "Capacities of quantum Markovian noise for large times". IEEE Transactions on Information Theory **72**, 6715–6727 (2026). arXiv:2408.00116.

[17] Satvik Singh, Mizanur Rahaman, and Nilanjana Datta. "Zero-error communication under discrete-time Markovian dynamics". Quantum **9**, 1910 (2025). arXiv:2402.18703.

[18] Satvik Singh and Nilanjana Datta. "Information storage and transmission under Markovian noise". PRX Quantum **7**, 020312 (2026). arXiv:2504.10436.

[19] Dennis Kretschmann and Reinhard F. Werner. "Tema con variazioni: quantum channel capacity". New Journal of Physics **6**, 26 (2004). arXiv:quant-ph/0311037.

[20] Li Gao, Marius Junge, and Nicholas LaRacuente. "Capacity estimates via comparison with TRO channels". Communications in Mathematical Physics **364**, 83–121 (2018). arXiv:1609.08594.

[21] Kun Fang, Xin Wang, Marco Tomamichel, and Mario Berta. "Quantum channel simulation and the channel's smooth max-information". IEEE Transactions on Information Theory **66**, 2129–2140 (2020). arXiv:1807.05354.

[22] William Matthews. "A linear program for the finite block length converse of Polyanskiy-Poor-Verdú via nonsignaling codes". IEEE Transactions on Information Theory **58**, 7036–7044 (2012). arXiv:1109.5417.

[23] Runyao Duan and Andreas Winter. "No-signalling-assisted zero-error capacity of quantum channels and an information theoretic interpretation of the Lovász number". IEEE Transactions on Information Theory **62**, 891–914 (2016). arXiv:1409.3426.

[24] Toby S. Cubitt, Debbie Leung, William Matthews, and Andreas Winter. "Zero-error channel capacity and simulation assisted by non-local correlations". IEEE Transactions on Information Theory **57**, 5509–5523 (2011). arXiv:1003.3195.

[25] Alexei Kitaev. "Almost-idempotent quantum channels and approximate C^*-algebras" (2025). arXiv:2405.02434.

[26] Kenneth R. Davidson. "C^*-algebras by example". Volume 6 of Fields Institute Monographs. American Mathematical Society. (1996).

[27] John Watrous. "The theory of quantum information". Cambridge University Press. (2018).

[28] Daniele Amato, Paolo Facchi, and Arturo Konderak. "Asymptotic dynamics in the Heisenberg picture: attractor subspace and Choi-Effros product". Open Systems & Information Dynamics **32**, 2550014 (2025). arXiv:2511.17770.

[29] Masamichi Takesaki. "Theory of Operator Algebras I". Springer Verlag. (1979).

[30] Jacques Dixmier. "Les algèbres d'opérateurs dans l'espace hilbertien (algèbres de von Neumann)". Gauthier-Villars. (1969). Second edition. url: https://www.gabay-editeur.com/DIXMIER-Les-algebres-doperateurs-dans-lespace-hilbertien-Algebres-de-von-Neumann-2e-ed-1969.

[31] Ker-I Ko. "Complexity theory of real functions". Progress in Theoretical Computer Science. Birkhäuser Boston. (1991).

[32] Richard P. Brent. "Fast multiple-precision evaluation of elementary functions". Journal of the ACM **23**, 242–251 (1976).

[33] Michael Sagraloff and Kurt Mehlhorn. "Computing real roots of real polynomials". Journal of Symbolic Computation **73**, 46–86 (2016). arXiv:1308.4088.

[34] Joachim von zur Gathen and Jürgen Gerhard. "Modern computer algebra". Cambridge University Press. (2013). Third edition.

[35] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. "Algorithms in real algebraic geometry". Volume 10 of Algorithms and Computation in Mathematics. Springer. (2006). Second edition.

[36] Dorina Jibetean and Etienne de Klerk. "Global optimization of rational functions: a semidefinite programming approach". Mathematical Programming **106**, 93–109 (2006).

[37] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels". Physical Review Letters **70**, 1895–1899 (1993).

[38] Charles H. Bennett and Stephen J. Wiesner. "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states". Physical Review Letters **69**, 2881–2884 (1992).

[39] Toby S. Cubitt, Debbie Leung, William Matthews, and Andreas Winter. "Improving zero-error classical communication with entanglement". Physical Review Letters **104**, 230503 (2010). arXiv:0911.5300.

[40] Debbie Leung, Laura Mančinska, William Matthews, Maris Ozols, and Aidan Roy. "Entanglement can increase asymptotic rates of zero-error classical communication over classical channels". Communications in Mathematical Physics **311**, 97–111 (2012). arXiv:1009.1195.

A An algebraic proof of the converse of item (i) of Theorem A in the unassisted setting

Since a reduced channel $\widehat{\mathcal{F}}$ has an adjoint $\widehat{\mathcal{F}}^*$ which projects its input onto a finite-dimensional $*$-algebra, it is natural to ask whether it is possible to approach the problem of the unassisted emulation of idempotent channels through purely algebraic arguments. We provide such a proof of the converse of item (i) of Theorem A for exact unassisted emulation in this section, that is, we show that for any two idempotent channels $\mathcal{F}$ and $\mathcal{G}$ with $\lambda(\mathcal{F}) \neq (1)$, the zero-error unassisted emulation capacity of $\mathcal{F}$ by $\mathcal{G}$, i.e. the supremum of the ratios k/n for which $\mathcal{F}^{\otimes k} = \mathcal{D}\mathcal{G}^{\otimes n}\mathcal{E}$ for some channels $\mathcal{D}$ and $\mathcal{E}$, is at most

$$\inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}. \tag{130}$$

This algebraic approach yields a slightly stronger version of the converse of item (i) of Theorem A for the interconversion of reduced idempotent channels when allowing for dimension-dependent errors. We provide the latter theorem and its proof in Appendix B, which is an application of approximate C^*-algebras. These algebraic structures were recently introduced by Kitaev in [25].

To prove this direction, we use a result of [14] which is almost a converse to Theorem 11.

Theorem 47 (Forward direction of [14, Theorem 1.1])**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be two finite-dimensional $*$-algebras. If there is an injective $*$-homomorphism of $\mathcal{A}$ into $\mathcal{B}$, then for all $p \in [1, +\infty]$,*

$$\|\lambda(\mathcal{A})\|_p \leq \|\lambda(\mathcal{B})\|_p. \tag{131}$$

Note that Theorems 11 and 47 are not converse to one another, because the inequality between the ℓ_p-norms of the shape vectors is strict in one case but not in the other.

The converse of item (i) of Theorem A then follows from Theorem 47 combined with the following proposition.

Proposition 11. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels equal to their reduced channels, i.e. $\mathcal{F} = \widehat{\mathcal{F}}$ and $\mathcal{G} = \widehat{\mathcal{G}}$. Suppose that there are two channels $\mathcal{E}$, $\mathcal{D}$ satisfying*

$$\mathcal{F} = \mathcal{D}\mathcal{G}\mathcal{E}. \tag{132}$$

We write $\mathbb{1}_\mathcal{F}$ and $\mathbb{1}_\mathcal{G}$ for the identity operators in $\mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{L}(\mathcal{H}_\mathcal{G})$ respectively. To simplify notation, we also write $e_{\mathcal{G}\mathcal{E}} = \mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F})^0$. Define

$$\widetilde{\mathcal{G}}^* : x \mapsto e_{\mathcal{G}\mathcal{E}}\mathcal{G}^*(x)e_{\mathcal{G}\mathcal{E}}. \tag{133}$$

Then, $e_{\mathcal{G}\mathcal{E}} \operatorname{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is a unital $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{G})$ whose identity operator is $e_{\mathcal{G}\mathcal{E}}$ and $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ restricted to the $*$-algebra $\operatorname{Rg}(\mathcal{F}^*)$ is an injective unital $*$-homomorphism which embeds $\operatorname{Rg}(\mathcal{F}^*)$ into $e_{\mathcal{G}\mathcal{E}} \operatorname{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ as a subalgebra.

Furthermore, for all $p \in [1, +\infty]$

$$\|\lambda(e_{\mathcal{G}\mathcal{E}} \operatorname{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}})\|_p \leq \|\lambda(\operatorname{Rg}(\mathcal{G}^*))\|_p.$$

The proof of this proposition uses the following lemma.

Lemma 48. *Let $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a quantum channel and write $e = \Phi(\mathbb{1})^0$. Then,*

i) for all $x \in \mathcal{L}(\mathcal{H}')$, $\Phi^*(x) = \Phi^*(exe)$,

ii) for all positive semidefinite $x \in \mathcal{L}(\mathcal{H}')$, $\Phi^*(x) = 0$ implies $exe = 0$.

Proof. Let $\mathcal{H}'_0 = \operatorname{supp}(\Phi(\mathbb{1}))$ and $V : \mathcal{H}'_0 \to \mathcal{H}'$ be the isometry such that $e = VV^*$. Let $y \in \mathcal{L}(\mathcal{H})$ and $x \in \mathcal{L}(\mathcal{H}')$, then

$$\operatorname{tr}(y\Phi^*(x)) = \operatorname{tr}(\Phi(y)x) = \operatorname{tr}(VV^*\Phi(y)VV^*x) = \operatorname{tr}(e\Phi(y)ex) = \operatorname{tr}(y\Phi^*(exe)),$$

where the second equality follows by Lemma 1. Therefore $\Phi^*(x) = \Phi^*(exe)$.

For the second statement, let $x \in \mathcal{L}(\mathcal{H}')$ be positive semidefinite such that $\Phi^*(x) = 0$. Then

$$0 = \operatorname{tr}(\Phi^*(x)) = \operatorname{tr}(\Phi(\mathbb{1})x) = \operatorname{tr}(e\Phi(\mathbb{1})ex) \geq \lambda_{\min}(\Phi(\mathbb{1}))\operatorname{tr}(exe),$$

with $\lambda_{\min}(\Phi(\mathbb{1}))$ the minimal non-zero eigenvalue of $\Phi(\mathbb{1})$. Therefore, $\operatorname{tr}(exe) = 0$ and, by positivity of x, $exe = 0$. □

Then, we move on to the proof of Proposition 11. This proof uses the Kadison–Schwarz inequality [15, eq. (5.2)], which is a generalisation of the well-known Cauchy–Schwarz inequality to maps acting on operators. For every channel $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, for all $x \in \mathcal{L}(\mathcal{H}')$,

$$\mathcal{F}^*(x^*)\mathcal{F}^*(x) \leq \mathcal{F}^*(x^*x). \tag{134}$$

Furthermore, in this proof, we will use the notion of *multiplicative domain* of linear maps. Given $\mathcal{F}^* : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$ a linear map, the multiplicative domain of $\mathcal{F}^*$ is the set

$$\{x \in \mathcal{L}(\mathcal{H}') \mid \forall y \in \mathcal{L}(\mathcal{H}'),\ \mathcal{F}^*(xy) = \mathcal{F}^*(x)\mathcal{F}^*(y) \text{ and } \mathcal{F}^*(yx) = \mathcal{F}^*(y)\mathcal{F}^*(x)\}.$$

If $\mathcal{F}^*$ is furthermore Hermitian preserving, then its multiplicative domain is closed under taking the adjoint and is therefore a $*$-algebra. Moreover, in this case, $\mathcal{F}^*$ acts as a (not necessarily unital) $*$-homomorphism on its multiplicative domain. Furthermore, when $\mathcal{F}^*$ satisfies the Kadison–Schwarz inequality, by [15, Theorem 5.7], we can write its multiplicative domain as

$$\{x \in \mathcal{L}(\mathcal{H}') \mid \mathcal{F}^*(x^*x) = \mathcal{F}^*(x^*)\mathcal{F}^*(x) \text{ and } \mathcal{F}^*(xx^*) = \mathcal{F}^*(x)\mathcal{F}^*(x^*)\}.$$

In particular, if $\mathcal{F}^*$ is unital completely positive[3], it satisfies the Kadison–Schwarz inequality (eq. (134)) and we can use this second characterisation of its multiplicative domain.

Proof of Proposition 11. Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels equal to their reduced channels, i.e. $\mathcal{F} = \widehat{\mathcal{F}}$ and $\mathcal{G} = \widehat{\mathcal{G}}$, and let $\mathcal{D}$, $\mathcal{E}$ be two channels such that

$$\mathcal{F} = \mathcal{D}\mathcal{G}\mathcal{E}.$$

Taking the adjoint of this equation, we get

$$\mathcal{F}^* = \mathcal{E}^*\mathcal{G}^*\mathcal{D}^* = \mathcal{E}^*\mathcal{G}^*\mathcal{G}^*\mathcal{D}^*,$$

where the right-hand side follows from the idempotence of $\mathcal{G}^*$. Let $x \in \mathcal{L}(\mathcal{H}_\mathcal{F})$, applying Lemma 48 to the channel $\Phi = \mathcal{G}\mathcal{E}$, we get that

$$\mathcal{F}^*(x) = \mathcal{E}^*\mathcal{G}^*(e_{\mathcal{G}\mathcal{E}}\mathcal{G}^*\mathcal{D}^*(x)e_{\mathcal{G}\mathcal{E}}) = \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x),$$

with $\widetilde{\mathcal{G}}^*$ defined in the statement of Proposition 11. Let $x \in \operatorname{Rg}(\mathcal{F}^*)$. By the assumption that $\mathcal{F} = \widehat{\mathcal{F}}$, $\operatorname{Rg}(\mathcal{F}^*)$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{F})$, thus $x^*x \in \operatorname{Rg}(\mathcal{F}^*)$, and as $\mathcal{F}^*$ is idempotent, x^*x is a fixed point of $\mathcal{F}^*$, so that we have:

$$\begin{aligned}
x^*x &= \mathcal{F}^*(x^*x) \\
&= \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) \\
&\overset{(a)}{\geq} \mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)) \\
&\overset{(b)}{\geq} \mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*))\mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)) \\
&= x^*x,
\end{aligned}$$

where we used the Kadison–Schwarz inequality to get both (a) and (b). Since the left-hand side and the right-hand side of this chain of inequalities coincide, we conclude that (a) and (b) are in fact equalities. Therefore,

$$\mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)) = 0,$$

[3] In fact, it suffices for $\mathcal{F}^*$ to be subunital 2-positive for it to satisfy the Kadison–Schwarz inequality, see [15, Section 5.2] for more details.

where, by the Kadison–Schwarz inequality,

$$\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x) \geq 0.$$

Therefore, by the second statement in Lemma 48 applied to the channel $\Phi = \mathcal{G}\mathcal{E}$, we obtain

$$\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x) = 0.$$

We can swap the roles of x and x^* in (a) and (b) to find that

$$\widetilde{\mathcal{G}}^*\mathcal{D}^*(xx^*) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*) = 0.$$

Therefore, x is in the multiplicative domain of $\widetilde{\mathcal{G}}^*\mathcal{D}^*$. Thus, $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ restricted to $\mathrm{Rg}(\mathcal{F}^*)$ is an injective unital $*$-homomorphism into the $*$-algebra $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$, which is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_{\mathcal{G}})$ and whose identity operator is $e_{\mathcal{G}\mathcal{E}}$. It is injective as, for all $x \in \mathrm{Rg}(\mathcal{F}^*)$,

$$x = \mathcal{F}^*(x) = \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x),$$

hence $\widetilde{\mathcal{G}}^*\mathcal{D}^*(x) = 0$ implies $x = 0$. It is unital as it maps the identity of $\mathcal{L}(\mathcal{H}_{\mathcal{F}})$ to $e_{\mathcal{G}\mathcal{E}}$ which is the identity operator of the $*$-algebra $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$. Indeed, as both $\mathcal{G}$ and $\mathcal{D}$ are channels, $\mathcal{G}^*$ and $\mathcal{D}^*$ are unital so that $\mathcal{G}^*\mathcal{D}^*(\mathbb{1}_{\mathcal{F}}) = \mathbb{1}_{\mathcal{G}}$ and $\widetilde{\mathcal{G}}^*\mathcal{D}^*(\mathbb{1}_{\mathcal{F}}) = e_{\mathcal{G}\mathcal{E}}$.

Now, to show that $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_{\mathcal{G}})$ which satisfies for all $p \in [1, +\infty]$, $\|\lambda(e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}})\|_p \leq \|\lambda(\mathrm{Rg}(\mathcal{G}^*))\|_p$, it suffices to remark that, by Proposition 1, we can decompose the Hilbert space $\mathcal{H}_{\mathcal{G}}$ as

$$\mathcal{H}_{\mathcal{G}} = \bigoplus_{k=1}^{K} \mathcal{H}_{k,1} \otimes \mathcal{H}_{k,2},$$

and the $*$-algebra $\mathrm{Rg}(\mathcal{G}^*)$ as

$$\mathrm{Rg}(\mathcal{G}^*) = \bigoplus_{k=1}^{K} \mathcal{L}(\mathcal{H}_{k,1}) \otimes \mathbb{1}_{m_k},$$

with $m_k = \dim(\mathcal{H}_{k,2})$ for all $k \in [K]$. Then $\mathcal{G}\mathcal{E}(\mathbb{1}_{\mathcal{F}}) \in \mathrm{Rg}(\mathcal{G})$ can itself be decomposed as

$$\mathcal{G}\mathcal{E}(\mathbb{1}_{\mathcal{F}}) = \sum_{k=1}^{K} x_k \otimes \rho_k,$$

with $\{x_k \in \mathcal{L}(\mathcal{H}_{k,1}) \mid k \in [K]\}$ and $\{\rho_k \in \mathcal{D}(\mathcal{H}_{k,2}) \mid k \in [K]\}$. Thus,

$$e_{\mathcal{G}\mathcal{E}} = \mathcal{G}\mathcal{E}(\mathbb{1}_{\mathcal{F}})^0 = \sum_{k=1}^{K} x_k^0 \otimes \rho_k^0.$$

Since $\mathcal{G}$ is reduced, $\mathcal{G}(\mathbb{1}_{\mathcal{G}})$ is full-rank, so that for all $k \in [K]$, $\rho_k^0 = \mathbb{1}_{m_k}$. Then,

$$e_{\mathcal{G}\mathcal{E}} = \sum_{k=1}^{K} x_k^0 \otimes \mathbb{1}_{m_k} \in \mathrm{Rg}(\mathcal{G}^*),$$

and

$$e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}} = \bigoplus_{k=1}^{K} x_k^0\mathcal{L}(\mathcal{H}_{k,1})x_k^0 \otimes \rho_k^0.$$

Hence, $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is a $*$-subalgebra whose shape vector consists of the non-zero ranks of the projections x_k^0, arranged in non-increasing order. Since $\mathrm{rank}(x_k^0) \leq \dim(\mathcal{H}_{k,1})$ for every $k \in [K]$, we obtain, for every $p \in [1, +\infty)$,

$$\|\lambda(e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}})\|_p^p = \sum_{k=1}^{K} \mathrm{rank}(x_k^0)^p \leq \sum_{k=1}^{K} \dim(\mathcal{H}_{k,1})^p = \|\lambda(\mathrm{Rg}(\mathcal{G}^*))\|_p^p.$$

For $p = +\infty$, the same conclusion follows by taking the maximum instead of the sum. This proves the claimed inequality. □

We can now prove the converse of item (i) of Theorem A for exact unassisted emulation.

Proof of the converse for exact unassisted emulation. Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels, let $k, n \in \mathbb{N}_{>0}$ such that there are two channels $\mathcal{E}$ and $\mathcal{D}$ satisfying

$$\mathcal{F}^{\otimes k} = \mathcal{D}\mathcal{G}^{\otimes n}\mathcal{E}.$$

By Corollary 5, then, if we denote $\widehat{\mathcal{F}^{\otimes k}}$ and $\widehat{\mathcal{G}^{\otimes n}}$ the reduced channels associated to $\mathcal{F}^{\otimes k}$ and $\mathcal{G}^{\otimes n}$, this condition is equivalent to

$$\widehat{\mathcal{F}^{\otimes k}} = \mathcal{D}'\widehat{\mathcal{G}^{\otimes n}}\mathcal{E}',$$

for some channels $\mathcal{D}'$, $\mathcal{E}'$. Thus, by Proposition 11, there is an injective unital $*$-homomorphism from $\mathrm{Rg}(\widehat{\mathcal{F}^{\otimes k}}^*)$ into the $*$-algebra $e_{\widehat{\mathcal{G}^{\otimes n}}\mathcal{E}'}\,\mathrm{Rg}(\widehat{\mathcal{G}^{\otimes n}}^*)e_{\widehat{\mathcal{G}^{\otimes n}}\mathcal{E}'}$, which is itself a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{G}^{\otimes n})$. Then, by Theorem 47 and the second statement of Proposition 11, we have for all $p \in [1, +\infty]$:

$$\|\lambda(\widehat{\mathcal{F}^{\otimes k}})\|_p = \|\lambda(\mathrm{Rg}(\widehat{\mathcal{F}^{\otimes k}}^*))\|_p \leq \|\lambda(e_{\widehat{\mathcal{G}^{\otimes n}}\mathcal{E}'}\,\mathrm{Rg}(\widehat{\mathcal{G}^{\otimes n}}^*)e_{\widehat{\mathcal{G}^{\otimes n}}\mathcal{E}'})\|_p \leq \|\lambda(\mathrm{Rg}(\widehat{\mathcal{G}^{\otimes n}}^*))\|_p = \|\lambda(\widehat{\mathcal{G}^{\otimes n}})\|_p.$$

By the multiplicativity property of the shape vectors and the ℓ_p-norms with the tensor product (Proposition 7), we get for all $p \in [1, +\infty]$:

$$\|\lambda(\mathcal{F})\|_p^k = \|\lambda(\widehat{\mathcal{F}})\|_p^k \leq \|\lambda(\widehat{\mathcal{G}})\|_p^n = \|\lambda(\mathcal{G})\|_p^n.$$

Since $\lambda(\mathcal{F}) \neq (1)$, we have $\log(\|\lambda(\mathcal{F})\|_p) > 0$ for every finite $p \geq 1$. The same holds at $p = +\infty$ unless $\mathcal{F}$ is classical. In that case, we omit $p = +\infty$, following the convention of Section 4. Taking logarithms and dividing therefore gives,

$$\frac{k}{n} \leq \inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)},$$

which ends the proof. □

B An application of approximate C^*-algebras

In a recent article, Kitaev introduced and studied approximate C^*-algebras and almost idempotent quantum channels, in order to find an algebraic framework which allows one to naturally express the effects of noise on quantum information [25]. As most of our proofs in Section 4 exploit the decomposition properties of finite-dimensional $*$-algebras, we use approximate C^*-algebras to strengthen the converse of item (i) of Theorem A, for the interconversion of reduced idempotent channels, and show that it holds even if a small error is allowed.

Theorem 49. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels, which we assume to be reduced, that is $\mathcal{F} = \widehat{\mathcal{F}}$ and $\mathcal{G} = \widehat{\mathcal{G}}$. Suppose $\lambda(\mathcal{F}) \neq (1)$. Let $k, n \in \mathbb{N}_{>0}$ and let $\mathcal{E} : \mathcal{L}(\mathcal{H}_\mathcal{F}^{\otimes k}) \to \mathcal{L}(\mathcal{H}_\mathcal{G}^{\otimes n})$, $\mathcal{D} : \mathcal{L}(\mathcal{H}_\mathcal{G}^{\otimes n}) \to \mathcal{L}(\mathcal{H}_\mathcal{F}^{\otimes k})$ be two channels. Let $d = \dim(\mathcal{H}_\mathcal{F})$, $\delta_{\max}$ be the absolute constant[4] of Theorem 52 and let $\lambda_{\min,n}$ be the minimal non-zero eigenvalue of $\mathcal{G}^{\otimes n}\mathcal{E}(\mathbb{1}_{\mathcal{F}^{\otimes k}})$. If*

$$\|\mathcal{F}^{\otimes k} - \mathcal{D}\mathcal{G}^{\otimes n}\mathcal{E}\|_\diamond \leq \frac{\delta_{\max}\lambda_{\min,n}}{6d^k},$$

then

$$\frac{k}{n} \leq \inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)}.$$

Note that in Theorem 49, the error allowed depends on the dimension of the input space of $\mathcal{F}$, on the number of copies of $\mathcal{F}$ obtained, on the number of copies of $\mathcal{G}$ used, as well as on the encoder used for the emulation. This theorem is therefore less operationally motivated than items (i) and (iii) of Theorem A, but it still strengthens the converse of item (i), and its proof leads to interesting mathematical developments, especially on multiplicative domains (see Theorem 53).

An interesting direction would be to extend item (i) of Theorem A to the problem of emulating almost idempotent channels with one another, but we leave this for future work.

[4] Here, 'absolute' means that $\delta_{\max}$ depends neither on the dimension of $\mathcal{H}_\mathcal{F}$ nor on that of $\mathcal{H}_\mathcal{G}$.

This section is organised as follows. In Subsection B.1, we recall, for completeness, the *Error Reduction Theorem* for approximate C^*-algebras. We then move on to proving Theorem 53 on the approximate multiplicative domain of maps satisfying the Kadison–Schwarz inequality in Subsection B.2, which will be used in the proof of Theorem 49 in Subsection B.3.

Note that this section assumes familiarity with basic concepts from operator algebras. Note also that, in this section, we measure the quality of an emulation with the diamond norm itself, and not with the normalised diamond norm $\frac{1}{2}\|\cdot\|_\diamond$ used in the rest of the article. Also, throughout this section, the idempotent channels $\mathcal{F}$ and $\mathcal{G}$ are assumed equal to their reduced channels.

B.1 The Error Reduction Theorem

We begin by recalling the definitions of ε-C^*-algebras and of δ-homomorphisms.

Definition 50 (ε-C^*-algebra, [25, Definition 2.1])**.** An *ε-Banach* algebra is a Banach space $\mathcal{A}$, endowed with a bilinear multiplication map $\mathcal{A}\times\mathcal{A}\to\mathcal{A}$ such that

$$\begin{aligned}&\forall x,y\in\mathcal{A},\quad \|xy\|\leq(1+\varepsilon)\|x\|\|y\|,\\&\forall x,y,z\in\mathcal{A},\quad \|(xy)z-x(yz)\|\leq\varepsilon\|x\|\|y\|\|z\|.\end{aligned}$$

A *$*\varepsilon$-Banach algebra* is a complex ε-Banach algebra with a conjugate linear involution $x\mapsto x^*$ satisfying the equations:

$$\forall x,y\in\mathcal{A},\ \|x^*\|=\|x\|,\quad (xy)^*=y^*x^*. \tag{135}$$

An ε-C^*-algebra is one satisfying the following property:

$$\forall x\in\mathcal{A},\ \|x^*x\|\geq(1-\varepsilon)\|x\|^2. \tag{136}$$

The unit element $\mathbb{1}\in\mathcal{A}$ should satisfy the following conditions:

$$\|x\mathbb{1}-x\|\leq\varepsilon\|x\|,\ \|\mathbb{1}x-x\|\leq\varepsilon\|x\|,\ |\|\mathbb{1}\|-1|\leq\varepsilon. \tag{137}$$

If $\mathcal{A}$ is involutive, we also have

$$\mathbb{1}^*=\mathbb{1}. \tag{138}$$

A δ-homomorphism is then defined in the following way.

Definition 51 (δ-homomorphism, [25, Definition 2.2])**.** A *δ-homomorphism* from an ε'-Banach algebra $\mathcal{A}'$ to an ε''-Banach algebra $\mathcal{A}''$ is a bounded linear map $v:\mathcal{A}'\to\mathcal{A}''$ that almost preserves the unit and the multiplication:

$$\begin{aligned}&\|v(\mathbb{1})-\mathbb{1}\|\leq\delta,\\&\forall x,y\in\mathcal{A}',\ \|v(xy)-v(x)v(y)\|\leq\delta\|x\|\|y\|.\end{aligned} \tag{139}$$

A *non-unital* δ-homomorphism is defined by imposing only the second condition. In the $*$-algebra setting, it is also required that $v(x^*)=v(x)^*$.

A *δ-inclusion* is a δ-homomorphism such that

$$\forall x\in\mathcal{A}',\ (1-\delta)\|x\|\leq\|v(x)\|\leq(1+\delta)\|x\|.$$

The Error Reduction Theorem then states that if there is a δ-inclusion of an exact C^*-algebra $\mathcal{A}$ into an ε-C^*-algebra $\mathcal{B}$ for $\delta\leq\delta_{\max}$ and $\varepsilon<\varepsilon_{\max}$, which are constants that do not depend on the dimensions of $\mathcal{A}$ and $\mathcal{B}$, one can 'correct' this δ-inclusion into a less noisy approximate inclusion.

Theorem 52 (Error Reduction Theorem, [25, Corollary 8.3])**.** *There exist positive constants $\varepsilon_{\max}$, $\delta_{\max}$ and c_0 such that for all $\varepsilon<\varepsilon_{max}$, if a finite-dimensional C^*-algebra $\mathcal{A}$ is δ_{max}-included into an ε-C^*-algebra $\mathcal{B}$, then there is also a $(c_0\varepsilon)$-inclusion. If the original inclusion is bijective, then so is the new inclusion.*

Note that as every C^*-algebra is a 0-C^*-algebra, Theorem 52 states that when we have a $\delta_{\max}$-inclusion between two C^*-algebras, it is possible to lift it to an *exact* inclusion. This can be interpreted as a case of *complete correction*, in the sense that the reduction of the error parameter δ is total.

B.2 Approximate Multiplicative Domains

We used in the proof of the converse of item (i) of Theorem A the fact that, for a map $\mathcal{F}^* : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$, which satisfies the Kadison–Schwarz inequality, its multiplicative domain can be expressed as

$$\{x \in \mathcal{L}(\mathcal{H}') \mid \mathcal{F}^*(x^*x) = \mathcal{F}^*(x^*)\mathcal{F}^*(x) \text{ and } \mathcal{F}^*(xx^*) = \mathcal{F}^*(x)\mathcal{F}^*(x^*)\}.$$

In order to prove Theorem 49, we have to extend this result to the approximate case. For $\mathcal{F}^* : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$ a linear map satisfying the Kadison–Schwarz inequality, we define, for $\delta \geq 0$ the following sets:

$$\begin{aligned}
\mathcal{C}_\delta^R &= \{x \in \mathcal{L}(\mathcal{H}') \mid \|\mathcal{F}^*(x^*x) - \mathcal{F}^*(x^*)\mathcal{F}^*(x)\| \leq \delta\|x\|^2\},\\
\mathcal{C}_\delta^L &= \{x \in \mathcal{L}(\mathcal{H}') \mid \|\mathcal{F}^*(xx^*) - \mathcal{F}^*(x)\mathcal{F}^*(x^*)\| \leq \delta\|x\|^2\},\\
\mathcal{C}_\delta &= \mathcal{C}_\delta^R \cap \mathcal{C}_\delta^L.
\end{aligned}$$

We can show the following theorem.

Theorem 53. *Let $\mathcal{F}^* : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$ be a Hermitian-preserving linear map satisfying the Kadison–Schwarz inequality. Then, for all $\delta > 0$, $\mathcal{F}^*$ is 2δ-multiplicative on $\mathcal{C}_\delta$, i.e.*

$$\forall x, y \in \mathcal{C}_\delta, \|\mathcal{F}^*(xy) - \mathcal{F}^*(x)\mathcal{F}^*(y)\| \leq 2\delta\|x\|\|y\|. \tag{140}$$

More precisely, we have that

$$\forall x \in \mathcal{C}_\delta^L, \forall y \in \mathcal{C}_\delta^R, \|\mathcal{F}^*(xy) - \mathcal{F}^*(x)\mathcal{F}^*(y)\| \leq 2\delta\|x\|\|y\|. \tag{141}$$

Proof of Theorem 53. This proof generalises the ones of [15, Theorems 5.4 and 5.7]. In this proof, we let $\mathbb{1}$ denote the identity operator on $\mathcal{H}$.

We have that $x \in \mathcal{C}_\delta^L$ if and only if $x^* \in \mathcal{C}_\delta^R$ and vice versa. Therefore, it suffices to show that for all $x, y \in \mathcal{C}_\delta^R$,

$$\|\mathcal{F}^*(x^*y) - \mathcal{F}^*(x^*)\mathcal{F}^*(y)\| \leq 2\delta\|x\|\|y\|. \tag{142}$$

Let $x, y \in \mathcal{C}_\delta^R$, $t \in \mathbb{R}$ and $z = tx + y$. As $\mathcal{F}^*$ satisfies the Kadison–Schwarz inequality,

$$0 \leq \mathcal{F}^*(z^*z) - \mathcal{F}^*(z^*)\mathcal{F}^*(z).$$

If we develop the right-hand side, we get:

$$0 \leq t^2Q + tV + W,$$

with

$$\begin{aligned}
Q &= \mathcal{F}^*(x^*x) - \mathcal{F}^*(x^*)\mathcal{F}^*(x),\\
W &= \mathcal{F}^*(y^*y) - \mathcal{F}^*(y^*)\mathcal{F}^*(y),\\
v &= \mathcal{F}^*(x^*y) - \mathcal{F}^*(x^*)\mathcal{F}^*(y),\\
V &= v + v^*.
\end{aligned}$$

Then, as $\mathcal{F}^*$ satisfies the Kadison–Schwarz inequality, $0 \leq Q$ and $0 \leq W$. Furthermore, $Q \leq \|Q\|\mathbb{1} \leq \delta\|x\|^2\mathbb{1}$ and $W \leq \|W\|\mathbb{1} \leq \delta\|y\|^2\mathbb{1}$ by assumption. Since for all $t \in \mathbb{R}$, $t^2 \geq 0$, we have

$$0 \leq t^2Q + tV + W \leq t^2\delta\|x\|^2\mathbb{1} + tV + \delta\|y\|^2\mathbb{1}.$$

As $V = v + v^*$ is Hermitian, we have $V = UDU^*$ with U a unitary matrix and D a diagonal matrix with real coefficients. Conjugating the last inequality by U, we get

$$0 \leq t^2\delta\|x\|^2\mathbb{1} + tD + \delta\|y\|^2\mathbb{1}.$$

The right-hand side of this equation is simply a diagonal operator, whose diagonal coefficients are quadratic polynomials in t. As this equation gives us that this operator is positive semidefinite for all $t \in \mathbb{R}$, each of these quadratic polynomials is always non-negative, and can thus have at most a single root. Therefore, if we write $\{D_i\}_i$ for the diagonal entries of D, the discriminant of the polynomial

$$\delta\|x\|^2t^2 + D_it + \delta\|y\|^2,$$

has to be non-positive for all indices i, that is

$$D_i^2 \leq 4\delta^2 \|x\|^2 \|y\|^2.$$

Thus,

$$|D_i| \leq 2\delta \|x\| \|y\|.$$

As the Schatten ∞-norm of $V = UDU^*$ is $\|V\| = \max_i |D_i|$, we have:

$$\|v + v^*\| = \|V\| \leq 2\delta \|x\| \|y\|.$$

We can apply the same line of argument replacing y with iy to find

$$\|iv - iv^*\| \leq 2\delta \|x\| \|y\|. \tag{143}$$

We can decompose $v = h_1 + ih_2$ with h_1 and h_2 Hermitian operators[5], to find:

$$\|v + v^*\| = 2\|h_1\| \leq 2\delta \|x\| \|y\|,$$
$$\|iv - iv^*\| = 2\|h_2\| \leq 2\delta \|x\| \|y\|.$$

Therefore,

$$\|v\| \leq \|h_1\| + \|h_2\| \leq 2\delta \|x\| \|y\|, \tag{144}$$

which is what we claimed. □

B.3 Proof of Theorem 49

In this subsection, we finally prove Theorem 49, in the same way as we proved the converse of item (i) of Theorem A, that is, by first proving the one-shot case and then using the compatibility of the shape vector and of the ℓ_p-norms with the tensor product to conclude.

Proposition 12. *Let $\mathcal{F} : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{F})$ and $\mathcal{G} : \mathcal{L}(\mathcal{H}_\mathcal{G}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ be two idempotent channels equal to their reduced channels, i.e. $\mathcal{F} = \widehat{\mathcal{F}}$ and $\mathcal{G} = \widehat{\mathcal{G}}$, such that there exist two channels $\mathcal{E}$ and $\mathcal{D}$ satisfying*

$$\|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond \leq \frac{\delta_{\max}\lambda_{\min,1}}{6d}, \tag{145}$$

with $d = \dim(\mathcal{H}_\mathcal{F})$, $\delta_{\max}$ is the constant of Theorem 52 and $\lambda_{\min,1}$ is the minimal non-zero eigenvalue of $\mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F})$. We write $e_{\mathcal{G}\mathcal{E}} = \mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F})^0$ for the orthogonal projector onto the support $\mathrm{supp}(\mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F}))$, and we define

$$\widetilde{\mathcal{G}}^* : x \mapsto e_{\mathcal{G}\mathcal{E}}\mathcal{G}^*(x)e_{\mathcal{G}\mathcal{E}}.$$

Then, $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ is a $\delta_{\max}$-inclusion from the $*$-algebra $\mathrm{Rg}(\mathcal{F}^*)$ into the $*$-algebra $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$. By the Error Reduction Theorem 52, there is therefore an injective unital $*$-homomorphism embedding $\mathrm{Rg}(\mathcal{F}^*)$ into $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$, thus, for all $p \in [1, +\infty]$,

$$\|\lambda(\mathcal{F})\|_p \leq \|\lambda(\mathcal{G})\|_p.$$

We split the proof that the map $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ restricted to $\mathrm{Rg}(\mathcal{F}^*)$ satisfies all the properties of a $\delta_{\max}$-inclusion into $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ into several lemmas, each of them devoted to one of these properties.

We begin with the one related to the almost preservation of the Schatten ∞-norm. Its proof uses the *completely bounded operator norm*, which is defined for all linear maps $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ as

$$\|\mathcal{F}\|_{\mathrm{cb}} = \sup_{n \in \mathbb{N}_{>0}} \|\mathcal{F} \otimes \mathrm{Id}_n\|_{\infty\to\infty},$$

with $\|\cdot\|_{\infty\to\infty}$ the 'infinity-to-infinity' norm, i.e. the map norm of linear maps from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H}')$ when the input and output spaces are endowed with the Schatten ∞-norms. That is:

$$\|\mathcal{F}\|_{\infty\to\infty} = \sup_{\|x\| \leq 1} \|\mathcal{F}(x)\|.$$

Note that for any linear map $\mathcal{F} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, we have

$$\|\mathcal{F}\|_\diamond = \|\mathcal{F}^*\|_{\mathrm{cb}}.$$

[5] Taking $h_1 = \frac{1}{2}(v + v^*)$, $h_2 = \frac{1}{2i}(v - v^*)$.

Lemma 54 (Almost norm preservation)**.** *Let $\mathcal{F}$, $\mathcal{G}$, $\widetilde{\mathcal{G}}$, $\mathcal{E}$ and $\mathcal{D}$ be defined as in the statement of Proposition 12 and let*

$$\delta \geq \|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond = \|\mathcal{F}^* - \mathcal{E}^*\mathcal{G}^*\mathcal{D}^*\|_{\mathrm{cb}}.$$

Then, for all $x \in \mathrm{Rg}(\mathcal{F}^*)$,

$$(1-\delta)\|x\| \leq \|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| \leq \|x\|.$$

Note that this lemma implies that, whenever $\|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond \leq \delta < 1$, the restriction of $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ to $\mathrm{Rg}(\mathcal{F}^*)$ is injective. Indeed, if $x \in \mathrm{Rg}(\mathcal{F}^*)$ and $\widetilde{\mathcal{G}}^*\mathcal{D}^*(x) = 0$, then $(1-\delta)\|x\| \leq 0$, hence $x = 0$.

Proof. The right-hand side inequality follows from the fact that $\widetilde{\mathcal{G}}^*\mathcal{D}^* : \mathcal{L}(\mathcal{H}_\mathcal{F}) \to \mathcal{L}(\mathcal{H}_\mathcal{G})$ is a subunital completely positive map (when the identity considered in the output space of $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ is the identity of $\mathcal{L}(\mathcal{H}_\mathcal{G})$). For the left-hand side inequality, recall that, by Lemma 48, for all $x \in \mathcal{L}(\mathcal{H}_\mathcal{F})$,

$$\mathcal{E}^*\mathcal{G}^*\mathcal{D}^*(x) = \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x).$$

Then, we have for all $x \in \mathrm{Rg}(\mathcal{F}^*)$:

$$\begin{aligned}
\|x\| - \|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| &\overset{(a)}{\leq} \|x\| - \|\mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| \\
&\overset{(b)}{\leq} \|x - \mathcal{E}^*\mathcal{G}^*\mathcal{D}^*(x)\| \\
&\overset{(c)}{=} \|\mathcal{F}^*(x) - \mathcal{E}^*\mathcal{G}^*\mathcal{D}^*(x)\| \\
&\leq \|\mathcal{F}^* - \mathcal{E}^*\mathcal{G}^*\mathcal{D}^*\|_{\mathrm{cb}}\|x\| \\
&\leq \delta\|x\|,
\end{aligned}$$

where (a) follows from the fact that $\|\mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| \leq \|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\|$, since $\mathcal{E}^*\mathcal{G}^*$ is a unital completely positive map. Then, (b) follows from the triangle inequality. Finally, (c) follows from the fact that $\mathcal{F}^*(x) = x$ for all $x \in \mathrm{Rg}(\mathcal{F}^*)$. The last inequality can be rewritten as

$$(1-\delta)\|x\| \leq \|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\|.$$

□

We then move on to the proof that $\widetilde{\mathcal{G}}^*\mathcal{D}^* : \mathrm{Rg}(\mathcal{F}^*) \to e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is unital, where we recall that the unit of the $*$-algebra $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is $e_{\mathcal{G}\mathcal{E}}$, as, for every element $x = e_{\mathcal{G}\mathcal{E}}x'e_{\mathcal{G}\mathcal{E}} \in e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$, $e_{\mathcal{G}\mathcal{E}}x = xe_{\mathcal{G}\mathcal{E}} = x$. Note also that for a general δ-inclusion, only almost unitality is required.

Lemma 55 (Exact unitality)**.** *Let $\mathcal{F}$, $\mathcal{G}$, $\widetilde{\mathcal{G}}$, $\mathcal{E}$ and $\mathcal{D}$ be defined as in Proposition 12. Then:*

$$\widetilde{\mathcal{G}}^*\mathcal{D}^*(\mathbb{1}_\mathcal{F}) = e_{\mathcal{G}\mathcal{E}}.$$

Proof. As $\mathcal{G}$ and $\mathcal{D}$ are assumed to be channels, their adjoints are unital maps, so that

$$\mathcal{G}^*\mathcal{D}^*(\mathbb{1}_\mathcal{F}) = \mathbb{1}_\mathcal{G}.$$

Therefore

$$\widetilde{\mathcal{G}}^*\mathcal{D}^*(\mathbb{1}_\mathcal{F}) = e_{\mathcal{G}\mathcal{E}}\mathcal{G}^*(\mathcal{D}^*(\mathbb{1}_\mathcal{F}))e_{\mathcal{G}\mathcal{E}} = e_{\mathcal{G}\mathcal{E}}\mathbb{1}_\mathcal{G}e_{\mathcal{G}\mathcal{E}} = e_{\mathcal{G}\mathcal{E}}.$$

□

Finally, to prove the almost preservation of the multiplication on $\mathrm{Rg}(\mathcal{F}^*)$ by $\widetilde{\mathcal{G}}^*\mathcal{D}^*$, we will need the following additional lemma.

Lemma 56. *Let $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a channel, write $d = \dim(\mathcal{H})$, $e = \Phi(\mathbb{1})^0$ and $\lambda_{\min,1}$ the minimal non-zero eigenvalue of $\Phi(\mathbb{1})$. Then, for all positive semidefinite $x \in \mathcal{L}(\mathcal{H}')$,*

$$\frac{\lambda_{\min,1}}{d}\|exe\| \leq \|\Phi^*(x)\|. \tag{146}$$

Proof. Let $x \in \mathcal{L}(\mathcal{H}')$ be positive semidefinite, then

$$\begin{aligned}\|\Phi^*(x)\| &\geq \frac{1}{d}\operatorname{tr}(\Phi^*(x)) \\ &= \frac{1}{d}\operatorname{tr}(x\Phi(\mathbb{1})) \\ &\geq \frac{\lambda_{\min,1}}{d}\operatorname{tr}(exe) \\ &\geq \frac{\lambda_{\min,1}}{d}\|exe\|.\end{aligned}$$

□

Note that this lemma implies the second statement of Lemma 48. We can now move on to the proof of the almost preservation of the multiplication, which uses our Theorem 53 on the approximate multiplicative domain of maps satisfying the Kadison–Schwarz inequality.

Lemma 57 (Almost preservation of the multiplication)**.** *Let $\mathcal{F}$, $\mathcal{G}$, $\widetilde{\mathcal{G}}$, $\mathcal{E}$ and $\mathcal{D}$ be defined as in the statement of Proposition 12, let*

$$\delta \geq \|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond,$$

let $\lambda_{\min,1}$ be the minimal non-zero eigenvalue of $\mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F})$, let $d = \dim(\mathcal{H}_\mathcal{F})$ and let $x, y \in \operatorname{Rg}(\mathcal{F}^*)$. We then have:

$$\|\widetilde{\mathcal{G}}^*\mathcal{D}^*(xy) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\widetilde{\mathcal{G}}^*\mathcal{D}^*(y)\| \leq \frac{6\delta d}{\lambda_{\min,1}}\|x\|\|y\|.$$

Proof. We prove that, for all $x \in \operatorname{Rg}(\mathcal{F}^*)$,

$$\begin{aligned}\|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| &\leq \frac{3\delta d}{\lambda_{\min,1}}\|x\|^2, \\ \|\widetilde{\mathcal{G}}^*\mathcal{D}^*(xx^*) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\| &\leq \frac{3\delta d}{\lambda_{\min,1}}\|x\|^2,\end{aligned}$$

and then use Theorem 53 to conclude. First, to simplify notation, we let $\Phi^* = \mathcal{E}^*\mathcal{G}^*\mathcal{D}^* = \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*$. We have:

$$\begin{aligned}\|\Phi^*(x^*x) - \Phi^*(x^*)\Phi^*(x)\| &= \|\Phi^*(x^*x) - \Phi^*(x^*)\Phi^*(x) + x^*x - x^*x\| \\ &\leq \|\Phi^*(x^*x) - x^*x\| + \|x^*x - \Phi^*(x^*)\Phi^*(x) + x^*\Phi^*(x) - x^*\Phi^*(x)\| \\ &\overset{(a)}{\leq} \delta\|x\|^2 + \|x\|\|x - \Phi^*(x)\| + \|x^* - \Phi^*(x^*)\|\|\Phi^*(x)\| \\ &\overset{(b)}{\leq} 3\delta\|x\|^2,\end{aligned}$$

where (a) follows from the fact that as $x \in \operatorname{Rg}(\mathcal{F}^*)$, we have

$$\|\Phi^*(x^*x) - x^*x\| = \|\Phi^*(x^*x) - \mathcal{F}^*(x^*x)\| \leq \|\Phi^* - \mathcal{F}^*\|_{\mathrm{cb}}\|x^*x\| \leq \delta\|x\|^2.$$

By the same argument, we also have that $\|x - \Phi^*(x)\| \leq \delta\|x\|$ and $\|x^* - \Phi^*(x^*)\| \leq \delta\|x\|$, hence (b) follows.

Then, applying the Kadison–Schwarz inequality, we get:

$$\mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) \geq \mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)) \geq \mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\mathcal{E}^*\mathcal{G}^*\widetilde{\mathcal{G}}^*\mathcal{D}^*(x).$$

Therefore:

$$\|\mathcal{E}^*\mathcal{G}^*(\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x))\| \leq \|\Phi^*(x^*x) - \Phi^*(x^*)\Phi^*(x)\| \leq 3\delta\|x\|^2.$$

Thus, by Lemma 56, we obtain

$$\|\widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*x) - \widetilde{\mathcal{G}}^*\mathcal{D}^*(x^*)\widetilde{\mathcal{G}}^*\mathcal{D}^*(x)\| \leq \frac{3d\delta}{\lambda_{\min,1}}\|x\|^2.$$

We can apply the exact same line of reasoning to prove the same inequality but with the roles of x and x^* swapped. □

Gathering all these lemmas, we obtain the proof of Proposition 12.

Proof of Proposition 12. Let $\delta \geq \|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond$. Since we assumed that $\mathcal{F}$ and $\mathcal{G}$ are equal to their reduced channels, i.e. $\mathcal{F} = \widehat{\mathcal{F}}$ and $\mathcal{G} = \widehat{\mathcal{G}}$, $\mathrm{Rg}(\mathcal{F}^*)$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{F})$ and $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$ is a $*$-subalgebra of $\mathcal{L}(\mathcal{H}_\mathcal{G})$. By Lemmas 54, 55 and 57, $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ is a $\max\{\delta, \frac{6d\delta}{\lambda_{\min,1}}\}$-inclusion (we trivially have that $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ is Hermitian preserving). However, as

$$d = \mathrm{tr}(\mathbb{1}_\mathcal{F}) = \mathrm{tr}(\mathcal{G}\mathcal{E}(\mathbb{1}_\mathcal{F})) \geq \lambda_{\min,1},$$

$\frac{d}{\lambda_{\min,1}} \geq 1$ so that $\delta \leq \frac{6d}{\lambda_{\min,1}}\delta$. Finally, $\widetilde{\mathcal{G}}^*\mathcal{D}^*$ is a $\frac{6d\delta}{\lambda_{\min,1}}$-inclusion of $\mathrm{Rg}(\mathcal{F}^*)$ into $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$.

Then, if we suppose that

$$\frac{6d\delta}{\lambda_{\min,1}} \leq \delta_{\max},$$

i.e.

$$\|\mathcal{F} - \mathcal{D}\mathcal{G}\mathcal{E}\|_\diamond \leq \delta \leq \frac{\delta_{\max}\lambda_{\min,1}}{6d},$$

by the Error Reduction Theorem (Theorem 52), there is an injective $*$-homomorphism from $\mathrm{Rg}(\mathcal{F}^*)$ into $e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}}$, thus by Kuperberg's theorem (Theorem 47) and the second part of Proposition 11, for all $p \in [1, +\infty]$,

$$\|\lambda(\mathcal{F})\|_p \leq \|\lambda(e_{\mathcal{G}\mathcal{E}}\,\mathrm{Rg}(\mathcal{G}^*)e_{\mathcal{G}\mathcal{E}})\|_p \leq \|\lambda(\mathcal{G})\|_p.$$

□

We can finally prove Theorem 49.

Proof of Theorem 49. We simply apply Proposition 12 to $\mathcal{F}^{\otimes k}$ and $\mathcal{G}^{\otimes n}$. Note that the dimension of $\mathcal{H}_\mathcal{F}^{\otimes k}$ is $\dim(\mathcal{H}_\mathcal{F})^k = d^k$, with $d = \dim(\mathcal{H}_\mathcal{F})$, so that the hypothesis of Theorem 49 is sufficient. Thus, by Proposition 12, if the hypotheses of Theorem 49 are met, for all $p \in [1, +\infty]$, we have:

$$\|\lambda(\mathcal{F}^{\otimes k})\|_p \leq \|\lambda(\mathcal{G}^{\otimes n})\|_p.$$

If $\mathcal{F}$ is classical, i.e. $\lambda(\mathcal{F}) = (1, \ldots, 1)$, with $\lambda(\mathcal{F}) \neq (1)$ by assumption, then the infimum of the ratios $\log(\|\lambda(\mathcal{G})\|_p)/\log(\|\lambda(\mathcal{F})\|_p)$ for $p \in [1, +\infty]$ is reached at $p = 1$, see Table 2. Since $\|\lambda(\mathcal{F}^{\otimes k})\|_1 \leq \|\lambda(\mathcal{G}^{\otimes n})\|_1$, we have

$$\frac{k}{n} \leq \frac{\log(\|\lambda(\mathcal{G})\|_1)}{\log(\|\lambda(\mathcal{F})\|_1)} = \inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} = C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}).$$

If $\mathcal{F}$ is not classical, then for all $p \in [1, +\infty]$, $\log(\|\lambda(\mathcal{F})\|_p) > 0$. Thus, by the compatibility of the ℓ_p-norms of the shape vector with the tensor product, we have

$$\frac{k}{n} \leq \inf_{p\in[1,+\infty]} \frac{\log(\|\lambda(\mathcal{G})\|_p)}{\log(\|\lambda(\mathcal{F})\|_p)} = C_{\mathrm{UA}}(\mathcal{G} \mapsto \mathcal{F}),$$

where the right-hand side equality is the first statement in item (i) of Theorem A.

□