A Generalized Stein Lemma for Quantum Channels

Minbo Gao*[1,2], Zhengfeng Ji†[3], and Chenghua Liu‡[1,2]

[1]Institute of Software, Chinese Academy of Sciences, Beijing, China
[2]University of Chinese Academy of Sciences, Beijing, China
[3]Department of Computer Science and Technology, Tsinghua University, Beijing, China

Abstract

We prove a generalized quantum Stein lemma for parallel discrimination of an arbitrary finite-dimensional channel against families of alternative channels. The alternatives are compact and convex, closed under tensor products, and contain a faithful replacer. At every fixed type-I error tolerance, the optimal type-II error exponent equals the regularized Umegaki channel relative entropy minimized over the alternatives, and both defining limits exist. Inputs may be entangled across channel uses and with a reference system. The main technical result constructs exactly trace-preserving approximations with exponentially small diamond error, dominated in completely positive order by free channels at every rate above the common exponent. The proof combines uniform auxiliary map approximation, iterative reduction of the domination rate, and tensor amplification. It yields asymptotic equipartition with trace-preserving smoothing at every subexponentially vanishing error, an exponential strong converse, and stability under vanishing diamond-norm perturbations. Under additional permutation invariance and closure under insertion of a fixed faithful input state and discarding of its output, we also prove an explicit finite-block completion bound. This quantitative construction uses weighted discarding, local operator corrections, and comparison with auxiliary extensions. Applications include state-preserving, coherence-restricted, covariant, entanglement-breaking, and positive-partial-transpose channels. For isometric targets against entanglement-breaking or positive-partial-transpose alternatives, we obtain exact finite-block testing and smoothing formulas.

1 Introduction

Quantum hypothesis testing gives an operational meaning to relative entropy. Consider two states ρ and σ, and let the effect $0 \le Q_n \le I$ represent acceptance of $\rho^{\otimes n}$. The two error probabilities are

$$\alpha_n = 1 - \operatorname{Tr} Q_n \rho^{\otimes n}, \qquad \beta_n = \operatorname{Tr} Q_n \sigma^{\otimes n}.$$

In the asymmetric problem, one bounds α_n by a fixed $\varepsilon \in (0,1)$ and minimizes β_n. The quantum Stein lemma identifies the limiting exponent of this minimum with the Umegaki relative entropy $D(\rho\|\sigma) = \operatorname{Tr} \rho(\log_2 \rho - \log_2 \sigma)$, with value $+\infty$ when the support of ρ is not contained in that of σ [1, 2]. Thus relative entropy measures the asymptotic rate at which an alternative can be excluded while the target is accepted with a prescribed probability.

For a quantum process, the experimenter must choose an input as well as a measurement. An input may be entangled with a reference that does not pass through the process; with several channel uses, it may also be entangled across those uses. A further difficulty arises when the alternative is a family of processes rather than a single known channel. Such families occur naturally in resource theories, where membership in a specified set expresses the absence of the resource under consideration. Detecting a resource then requires one experiment that rejects every member of the alternative family.

*gmb17@tsinghua.org.cn
†jizhengfeng@tsinghua.edu.cn
‡liuch.russell@gmail.com

This paper treats parallel access to a target channel $\mathcal{N} : \mathrm{L}(A) \to \mathrm{L}(B)$. At blocklength n, the target is $\mathcal{N}^{\otimes n}$, whereas an alternative is an arbitrary member of $\mathfrak{F}_n \subseteq \mathrm{CPTP}(A^{\otimes n} \to B^{\otimes n})$. Here A and B are finite-dimensional input and output Hilbert spaces, and CPTP denotes completely positive trace-preserving maps. We write $\mathfrak{F} = (\mathfrak{F}_n)_{n\geq 1}$ and call its members free channels. The alternatives may correlate all inputs and outputs within the block. The tester prepares one reference-assisted input and measures the joint reference and output after all channel uses. There is no restriction to classical inputs or to product input states. The access model is nevertheless parallel: it does not insert intermediate operations between uses of the unknown process.

1.1 Results and scope

Let $\beta_{\varepsilon,n}$ be the smallest worst-case alternative acceptance probability among testers whose target acceptance is at least $1 - \varepsilon$. Write D_{ch} for Umegaki relative entropy optimized over a reference-assisted input shared by the target and the specified alternative. Our main result is

$$\lim_{n\to\infty} -\frac{1}{n} \log_2 \beta_{\varepsilon,n} = \lim_{n\to\infty} \frac{1}{n} \inf_{\mathcal{M}\in\mathfrak{F}_n} D_{\mathrm{ch}}(\mathcal{N}^{\otimes n} \| \mathcal{M}). \tag{1}$$

Both limits exist, and their value is independent of $\varepsilon \in (0, 1)$. The three assumptions in Definition 1 are compactness and convexity of the alternative sets, closure under tensor products, and membership of a faithful replacer. These assumptions suffice for all the asymptotic results below.

The main technical result gives exponential approximation at any domination rate S strictly above the common rate in (1). We write $\Phi \geq_{\mathrm{CP}} \Psi$ when $\Phi - \Psi$ is completely positive, and $\|\cdot\|_\diamond$ for the diamond norm, defined in Section 2. For all sufficiently large n, there are a free channel $\mathcal{S}_n \in \mathfrak{F}_n$ and a completely positive trace-preserving map $\mathcal{L}_n$ such that

$$\mathcal{L}_n \leq_{\mathrm{CP}} 2^{nS}\mathcal{S}_n, \qquad \|\mathcal{L}_n - \mathcal{N}^{\otimes n}\|_\diamond \leq K_S e^{-\gamma_S n},$$

for constants $K_S, \gamma_S > 0$. The same approximation and dominator work for every input, reference, and acceptance effect. This uniform control yields an exponential strong converse: above the Stein rate, the target acceptance probability is at most $K2^{-cn}$ for suitable $K, c > 0$. It also gives asymptotic equipartition (AEP) for trace-preserving smoothing with $0 < \delta_n \leq \bar{\delta} < 2$ and $\log_2(1/\delta_n) = o(n)$, including every fixed diamond radius in $(0, 2)$. The Stein exponent is stable under vanishing diamond perturbations of the target sequence.

A second contribution is an explicit finite-block completion bound. For this bound we additionally assume permutation invariance and marginal closure: inserting one fixed faithful input state and discarding the corresponding output preserves membership in the family. These conditions are stated in Definition 14. For every fixed $\varepsilon \in (0, 1)$ and $0 < \delta \leq 1/16$, there are a free channel $\mathcal{S}_n$ and a CPTP map $\mathcal{L}_n$ with

$$\begin{aligned} \mathcal{S}_n \geq_{\mathrm{CP}} \beta_{\varepsilon,n} 2^{-g_\varepsilon(n,\delta)}\mathcal{L}_n, \quad \|\mathcal{L}_n - \mathcal{N}^{\otimes n}\|_\diamond \leq \delta, \\ g_\varepsilon(n, \delta) = K_\varepsilon n^{2/3} \log_2 \frac{n+1}{\delta}. \end{aligned}$$

The constant K_ε depends only on the local dimensions, the minimum eigenvalues of the witnessing states, and ε. Thus the bound controls the domination cost at a specified blocklength and accuracy relative to the finite-block testing coefficient. The overhead is sublinear for fixed or inverse-polynomial accuracy.

All five conditions hold for the resource families in Table 1, including state-preserving, maximally incoherent, dephasing-covariant, symmetry-covariant, entanglement-breaking, and positive-partial-transpose channels. For isometric targets against the last two families, we compute testing and smoothing exactly at every blocklength.

1.2 Relation to earlier work

The generalized state Stein problem was formulated by Brandão and Plenio [3]. Following the identification of a gap in the original proof [4], proofs were obtained by Hayashi and Yamasaki [5] and Lami [6]. Families of state hypotheses built from mixtures of tensor powers were studied by Berta, Brandão, and Hirche [7]. Related developments include generalized asymptotic equipartition [8] and error exponents for families of correlated state hypotheses [9]. With a one-dimensional input, our asymptotic theorem reduces to the generalized state problem without permutation or partial-trace closure, as already established for states by Hayashi and Yamasaki [5].

For discrimination between two quantum channels, Gao, Ji, and Liu [10] established the channel Stein lemma and an exponential strong converse under arbitrary adaptive strategies. Bergh, Datta, and Salzmann characterized the parallel Stein exponent for composite hypotheses generated by independent, not necessarily identically distributed channel uses and their convex hulls, with $\varepsilon \downarrow 0$ taken after the blocklength $\liminf$ or $\limsup$ [11, Thm. 10]. Our alternatives include correlated block channels subject to Definition 1. Generalized Stein lemmas for classical–quantum channels were established by Hayashi and Yamasaki [12] and, independently, by Bergh, Datta, and Khaitan [13]. A classical–quantum channel measures in a fixed input basis and prepares an output state depending on the result. The present theorem allows general quantum inputs and outputs, under the structural assumptions above. In particular, the marginal axiom used in our finite-block completion bound is not required for our asymptotic theorem and should not be read as a hypothesis shared by all of the classical–quantum results.

Exact trace preservation creates a separate issue from the testing problem. Gour [14] gives channel pairs for which max-relative entropy smoothed over trace-preserving maps does not converge to the regularized relative entropy when the dominator is prescribed to be the tensor power of one fixed channel. Our smoothing instead optimizes the dominator over $\mathfrak{F}_n$. Convexity, tensor closure, and the free replacer allow the dominator to change during amplification and completion. A singleton family $\mathfrak{F}_n = \{\mathcal{M}^{\otimes n}\}$ satisfies the faithful-replacer assumption only if $\mathcal{M}$ itself is a faithful replacer. Thus the present theorem does not establish a general fixed-dominator channel equipartition theorem.

1.3 Structure of the argument

The proof adapts the uniform Stinespring approximation and tensor amplification methods of [10] to families of alternative channels. Testing minimax supplies an alternative that is hardest to distinguish from the target at a fixed error tolerance. A semidefinite representation of the testing quantity then produces a uniform auxiliary map approximation for prescribed Stinespring dilations. Keeping their auxiliary spaces fixed allows two approximations relative to a common free dominator to be subtracted and tensorized.

The amplification argument has two stages. First, a sufficiently accurate fixed block can be repeated and its tensor expansion truncated to give exponential accuracy, at an arbitrarily small increase in domination rate. Second, a rough approximation from testing is combined with an exponentially accurate approximation at a higher rate. This produces accurate blocks at an improved rate. Starting from exact domination by the faithful replacer, finitely many improvements reach every rate above the lower limiting testing exponent.

The resulting approximations compare the upper and lower limiting testing exponents at every pair of fixed error tolerances. This proves ordinary convergence and error independence. A direct entropy comparison and the weak converse identify the common limit with the regularized Umegaki channel relative entropy. The proof establishes these limits without an additivity assumption.

The quantitative finite-block bound uses a separate construction. A positive-overlap branch is corrected by local operator expansions obtained from permutation averaging and weighted discarding. The additional family conditions convert these corrections into free domination, and a replacer restores trace preservation and the discarded sites. Comparison with auxiliary extensions reduces a general target to the isometric construction.

Section 2 states the model and main results. Section 3 develops testing minimax, auxiliary map approximation, and trace-defect completion. Section 4 proves exponential approximation and the Stein identity, and Section 5 derives their asymptotic consequences. Section 6 states the additional assumptions and the quantitative finite-block results. Resource families and exact benchmarks appear in Section 7, followed by the discussion in Section 8. Appendix A proves quantitative completion; Appendix B gives symmetric entropy estimates and an alternative limit proof under all five conditions.

2 Model and main results

2.1 Operator and channel conventions

All Hilbert spaces considered in this paper are finite dimensional. We write $\mathrm{L}(H,K)$ for the linear operators from H to K, and $\mathrm{L}(H)=\mathrm{L}(H,H)$. The adjoint of an operator or map is denoted by a dagger; for maps it is taken in the Hilbert–Schmidt inner product. We denote the identity operator by I_H and the identity map by id_H, omitting subscripts when the space is clear. For a Hermitian operator X, $\lambda_{\min}(X)$ is its smallest eigenvalue, and $\operatorname{supp} X$ is its range support. The norm of a vector is its Hilbert-space norm. Partial traces are denoted by Tr_K; for a state ρ on $A\otimes B$, write $\rho_B=\mathrm{Tr}_A\,\rho$. All transposes are taken in fixed orthonormal bases, using product bases on tensor powers. Our operator and channel conventions follow [15], with the reference factor written first in Choi operators. Write $\mathrm{D}(H)$ for the density operators on H and $\mathrm{CPTP}(A\to B)$ for the CPTP maps from $\mathrm{L}(A)$ to $\mathrm{L}(B)$. Pure states are represented by unit vectors $|\psi\rangle$, with density operator $|\psi\rangle\langle\psi|$. Mixed states and reduced states are written as density operators, typically denoted by ρ, σ, τ, or ω. A state is called faithful if it is positive definite, equivalently full rank. In finite dimensions its smallest eigenvalue is then strictly positive. The operator and trace norms are denoted by $\|\cdot\|_\infty$ and $\|\cdot\|_1$, respectively. We use $\log_2$ for information quantities and ln for the natural logarithm. For a linear operator K, let $\mathrm{Ad}_K(X)=KXK^\dagger$. In asymptotic estimates, subscripts on $O(\cdot)$ indicate the parameters on which the implicit constant may depend; $o(n)$ denotes a quantity whose ratio to n tends to zero. For any square operator X, its Hermitian part is $\operatorname{Re} X=(X+X^\dagger)/2$; an inequality involving $\operatorname{Re} X$ is understood in the Loewner order. For Hermiticity-preserving maps, $\Phi\geq_{\mathrm{CP}}\Psi$ means that $\Phi-\Psi$ is completely positive (CP). Pre- and postcomposition with a CP map, and tensoring with a CP map, preserve this order, as do positive linear combinations. In particular, if $\Phi_i\geq_{\mathrm{CP}}\Psi_i\geq_{\mathrm{CP}}0$ for $i=1,2$, then $\Phi_1\otimes\Phi_2\geq_{\mathrm{CP}}\Psi_1\otimes\Psi_2$.

For states ρ,σ on the same space, the relative entropy is

$$D(\rho\|\sigma)=\mathrm{Tr}\,\rho(\log_2\rho-\log_2\sigma),$$

with value $+\infty$ unless $\operatorname{supp}\rho\subseteq\operatorname{supp}\sigma$. Functions of positive operators are defined spectrally, with $0\log_2 0=0$ in entropy expressions. When σ is singular and $\operatorname{supp}\rho\subseteq\operatorname{supp}\sigma$, the term $\mathrm{Tr}\,\rho\log_2\sigma$ is evaluated on $\operatorname{supp}\sigma$. We use the data-processing and joint-convexity properties of relative entropy. For a linear map $\Phi:\mathrm{L}(H)\to\mathrm{L}(K)$, the diamond norm is the completely bounded trace norm,

$$\|\Phi\|_\diamond=\sup_R\sup_{0\neq Z\in\mathrm{L}(R\otimes H)}\frac{\|(\mathrm{id}_R\otimes\Phi)(Z)\|_1}{\|Z\|_1}.$$

A reference of dimension $\dim H$ suffices. The diamond norm bounds the trace-norm difference of outputs for every reference-assisted input, is nonincreasing under pre- and postcomposition by channels, and equals one on channels. Consequently a diamond distance at most δ changes any tester's acceptance probability by at most δ. We use this bound throughout; for two channels the sharper bound $\delta/2$ also holds because their output difference has trace zero. For a CP map Φ, one has $\|\Phi\|_\diamond = \|\Phi^\dagger(I)\|_\infty$. The diamond norm is multiplicative under tensor products; in particular, tensoring with a channel preserves it. For operators K and L with the same domain and codomain,

$$\|\mathrm{Ad}_K - \mathrm{Ad}_L\|_\diamond \le (\|K\|_\infty + \|L\|_\infty)\|K - L\|_\infty. \tag{2}$$

To see this, let $\widetilde{K} = I_R \otimes K$ and $\widetilde{L} = I_R \otimes L$. For an arbitrary, not necessarily positive, operator Z,

$$\widetilde{K} Z \widetilde{K}^\dagger - \widetilde{L} Z \widetilde{L}^\dagger = (\widetilde{K} - \widetilde{L}) Z \widetilde{K}^\dagger + \widetilde{L} Z (\widetilde{K} - \widetilde{L})^\dagger,$$
$$\|\widetilde{K} Z \widetilde{K}^\dagger - \widetilde{L} Z \widetilde{L}^\dagger\|_1 \le (\|K\|_\infty + \|L\|_\infty)\|K - L\|_\infty \|Z\|_1.$$

Here we used $\|UZV\|_1 \le \|U\|_\infty \|Z\|_1 \|V\|_\infty$ and invariance of operator norm under tensoring with an identity. Taking the supremum proves (2). This estimate will convert an operator approximation of a Kraus branch into a diamond approximation during trace completion.

We use the unnormalized Choi operator

$$C_\Phi = (\mathrm{id}_H \otimes \Phi)(|\Gamma_H\rangle\langle\Gamma_H|), \qquad |\Gamma_H\rangle = \sum_{j=1}^{\dim H} |j\rangle|j\rangle. \tag{3}$$

For an operator $L : H \to K$, we write $|L\rangle = (I_H \otimes L)|\Gamma_H\rangle$ for its vectorization. Thus $|L\rangle$ is a vector in $H \otimes K$, whereas $L|j\rangle$ is a vector in K. One has $C_{\mathrm{Ad}_L} = |L\rangle\langle L|$ and $\langle L \mid M\rangle = \mathrm{Tr}\, L^\dagger M$. The Choi criterion gives $\Phi \ge_{\mathrm{CP}} \Psi$ if and only if $C_\Phi \ge C_\Psi$. A map $\Phi : \mathrm{L}(H) \to \mathrm{L}(K)$ is CPTP if and only if $C_\Phi \ge 0$ and $\mathrm{Tr}_K C_\Phi = I_H$.

2.2 Alternative families

Fix A, B, put $a = \dim A$, $b = \dim B$, and fix a channel $\mathcal{N} : \mathrm{L}(A) \to \mathrm{L}(B)$. Abbreviate $\mathcal{N}_n = \mathcal{N}^{\otimes n}$ and $A^n = A^{\otimes n}$, with analogous notation for other systems. The symbol A_i denotes the ith tensor factor of A^n, and $[n] = \{1, \dots, n\}$. For a state $\omega \in \mathrm{D}(B)$, the replacer channel is $\mathcal{R}_\omega(X) = \mathrm{Tr}(X)\omega$.

A channel is *free* when it belongs to the family specified below. The axioms relate these sets across blocklengths and allow each member of $\mathfrak{F}_n$ to correlate all n uses.

Definition 1 (Admissible free families)**.** A sequence $\mathfrak{F} = (\mathfrak{F}_n)_{n\ge1}$, with $\mathfrak{F}_n \subseteq \mathrm{CPTP}(A^n \to B^n)$, is admissible if:

(F1) Each $\mathfrak{F}_n$ is nonempty, compact, and convex.

(F2) $\mathcal{M} \otimes \mathcal{L} \in \mathfrak{F}_{n+m}$ whenever $\mathcal{M} \in \mathfrak{F}_n$ and $\mathcal{L} \in \mathfrak{F}_m$.

(F3) There is a fixed positive-definite $\omega \in \mathrm{D}(B)$ with $\mathcal{R}_\omega \in \mathfrak{F}_1$.

Tensor closure implies $\mathcal{R}_\omega^{\otimes n} \in \mathfrak{F}_n$.

The word free refers only to membership in these sets, and admissible means that the sequence satisfies these three conditions. In particular, no closure under arbitrary input or output processing is assumed. The state ω appearing below is always the witness in F3; write $w = \lambda_{\min}(\omega) > 0$. Different witnessing states can change the constants in our estimates, but the testing problem itself is determined by $\mathcal{N}$ and $\mathfrak{F}$. Unless stated otherwise, $\mathfrak{F}$ is assumed admissible throughout. The quantitative results in Section 6 impose two additional conditions, stated there as F4 and F5.

2.3 Testing and relative entropy

We may take the reference-assisted input to be a pure state $|\psi\rangle \in R \otimes A^n$. Indeed, any mixed input can be purified by enlarging the reference. Extending the acceptance effect by the identity preserves all testing probabilities, and data processing shows that purification cannot decrease the output relative entropy. Schmidt compression then allows $R \simeq A^n$. All input state vectors in the optimizations below have unit norm. Write $\varrho_R = \mathrm{Tr}_{A^n} |\psi\rangle\langle\psi|$ for the reference marginal.

For a unit vector $|\psi\rangle \in R \otimes A^n$, define the output density operators

$$\begin{aligned} \rho_n^\psi &= (\mathrm{id}_R \otimes \mathcal{N}_n)(|\psi\rangle\langle\psi|), \\ \sigma_{\mathcal{M}}^\psi &= (\mathrm{id}_R \otimes \mathcal{M})(|\psi\rangle\langle\psi|). \end{aligned}$$

The target output is denoted by ρ and an alternative output by σ; superscripts and subscripts specify the input and channel when needed. The stabilized channel relative entropy and its distance to the free family are

$$\begin{aligned} D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}) &= \sup_{R, |\psi\rangle} D(\rho_n^\psi \| \sigma_{\mathcal{M}}^\psi), \\ E_n &= \inf_{\mathcal{M} \in \mathfrak{F}_n} D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}). \end{aligned}$$

The supremum is over normalized input vectors; as explained above, $R \simeq A^n$ suffices. In $D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M})$ the optimizing input may depend on $\mathcal{M}$. The testing problem below instead chooses one input and one measurement for the entire alternative family. For $0 < \varepsilon < 1$, define

$$\beta_{\varepsilon,n} = \inf_{\substack{R, |\psi\rangle,\ 0 \le Q \le I \\ \mathrm{Tr}\, Q \rho_n^\psi \ge 1-\varepsilon}} \sup_{\mathcal{M} \in \mathfrak{F}_n} \mathrm{Tr}\, Q \sigma_{\mathcal{M}}^\psi. \tag{4}$$

A tester is the pair consisting of the input state $|\psi\rangle$ and the acceptance effect Q, together with its reference space R. The two measurement effects are Q and $I - Q$: the outcome associated with Q declares the channel to be the target, and the other outcome rejects the target. Its type-I and type-II errors are, respectively,

$$\begin{aligned} \alpha(|\psi\rangle, Q) &= 1 - \mathrm{Tr}\, Q \rho_n^\psi, \\ \beta(|\psi\rangle, Q; \mathcal{M}) &= \mathrm{Tr}\, Q \sigma_{\mathcal{M}}^\psi. \end{aligned}$$

Feasibility means $\langle\psi|\psi\rangle = 1$, $0 \le Q \le I$, and $\alpha(|\psi\rangle, Q) \le \varepsilon$. Thus (4) minimizes the largest acceptance probability over free alternatives, subject to accepting $\mathcal{N}_n$ with probability at least $1-\varepsilon$. The target channel determines which testers satisfy this constraint through the output state ρ_n^ψ. For a simple feasible tester, choose any normalized input and set $Q = (1-\varepsilon)I$. Operationally, this tester ignores the output and declares the channel to be the target with probability $1 - \varepsilon$. Since every output state has trace one,

$$\mathrm{Tr}\, Q \rho_n^\psi = \mathrm{Tr}\, Q \sigma_{\mathcal{M}}^\psi = 1 - \varepsilon \qquad (\mathcal{M} \in \mathfrak{F}_n).$$

Its type-I error is therefore ε, and its worst-case type-II error is $1 - \varepsilon$, proving the elementary bound $\beta_{\varepsilon,n} \le 1 - \varepsilon$. There is one input state $|\psi\rangle$ and one effect Q for all alternatives; we call such a tester *uniform*, meaning independent of the alternative. The purification argument above ensures that these optimizations cover all physically admissible input states and effects.

To relate this composite testing problem to testing against a single alternative, we introduce the corresponding single-alternative optimum. For a target channel $\mathcal{T} : \mathrm{L}(H) \to \mathrm{L}(K)$ and a single alternative channel $\mathcal{M} : \mathrm{L}(H) \to \mathrm{L}(K)$, write $b_\varepsilon(\mathcal{T}, \mathcal{M})$ for the optimal reference-assisted parallel type-II error at target channel $\mathcal{T}$ and type-I error at most ε. The infimum below is over finite-dimensional reference systems R, unit vectors $|\psi\rangle \in R \otimes H$, and acceptance effects Q on $R \otimes K$:

$$b_\varepsilon(\mathcal{T}, \mathcal{M}) = \inf_{\substack{R, |\psi\rangle,\ 0 \le Q \le I \\ \operatorname{Tr} Q(\mathrm{id}_R \otimes \mathcal{T})(|\psi\rangle\langle\psi|) \ge 1-\varepsilon}} \operatorname{Tr} Q(\mathrm{id}_R \otimes \mathcal{M})(|\psi\rangle\langle\psi|).$$

In $b_\varepsilon(\mathcal{T}, \mathcal{M})$, the tester may depend on the specified alternative $\mathcal{M}$, whereas $\beta_{\varepsilon,n}$ requires one tester for all alternatives in $\mathfrak{F}_n$. The minimax theorem proved in Lemma 5 establishes the relation

$$\beta_{\varepsilon,n} = \max_{\mathcal{M} \in \mathfrak{F}_n} b_\varepsilon(\mathcal{N}_n, \mathcal{M}).$$

For the asymptotic analysis, we write $a_{\varepsilon,n} = -\log_2 \beta_{\varepsilon,n}$. The faithful replacer ensures $\beta_{\varepsilon,n} > 0$; the quantitative bounds are given in Lemma 4, with the constant $C = \log_2 b - \log_2 \lambda_{\min}(\omega)$.

2.4 Main theorems

Theorem 2 (Generalized Stein identity for quantum channels)**.** *Let $\mathcal{N} : \mathrm{L}(A) \to \mathrm{L}(B)$ be arbitrary and let $\mathfrak{F}$ satisfy F1, F2, and F3. Then there is a finite $R_{\mathfrak{F}}(\mathcal{N}) \ge 0$ such that*

$$\lim_{n\to\infty} -\frac{1}{n} \log_2 \beta_{\varepsilon,n} = R_{\mathfrak{F}}(\mathcal{N}) = \lim_{n\to\infty} \frac{E_n}{n} \qquad (0 < \varepsilon < 1).$$

Both limits are ordinary limits over all positive integers. Writing $w = \lambda_{\min}(\omega)$, one has $R_{\mathfrak{F}}(\mathcal{N}) \le \log_2 b - \log_2 w$.

The rate need not be strictly positive when $\mathcal{N} \notin \mathfrak{F}_1$. Example 24 gives an admissible family whose correlated alternatives contain a fixed target component at every blocklength, forcing $R_{\mathfrak{F}}(\mathcal{N}) = 0$.

The main technical result constructs exactly trace-preserving approximations with exponential accuracy at every rate above the Stein rate. Their domination by a free channel controls all reference-assisted tests simultaneously.

Theorem 3 (Exponential free-channel approximation)**.** *Under the assumptions of Theorem 2, for every $S > R_{\mathfrak{F}}(\mathcal{N})$ there are $K_S, \gamma_S > 0$ such that, for all sufficiently large n, one can find $\mathcal{S}_n \in \mathfrak{F}_n$ and $\mathcal{L}_n \in \mathrm{CPTP}(A^n \to B^n)$ satisfying*

$$\begin{aligned} \mathcal{L}_n &\le_{\mathrm{CP}} 2^{nS} \mathcal{S}_n, \\ \|\mathcal{L}_n - \mathcal{N}_n\|_\diamond &\le K_S e^{-\gamma_S n}. \end{aligned} \tag{5}$$

The dominator is chosen independently of the input, reference, and test. The constants and the blocklength threshold may depend on $\mathcal{N}$, $\mathfrak{F}$, and S.

Section 3 develops the testing and approximation tools; Section 4 uses them to prove Theorems 2 and 3 jointly. The proof starts from a lower limiting testing exponent and establishes its convergence before identifying the entropy rate. Under two additional conditions, Theorem 16 gives an explicit finite-block domination bound in terms of $\beta_{\varepsilon,n}$. Its constant is uniform over targets and families with the specified dimensions and witnessing-state eigenvalue bounds.

3 Testing and uniform channel approximation

In this section, we provide the four ingredients of the main proof. First, we apply replacer domination to obtain uniform bounds and a weak converse (Lemma 4). Second, testing minimax allows us to select the hardest alternative (Lemma 5). Third, auxiliary map approximation converts its testing bounds into a single operator approximation valid for every input (Lemma 6). Finally, a completion step restores exact trace preservation at a cost controlled by the trace defect (Lemma 7). For all family-dependent statements, we rely only on Definition 1.

3.1 Uniform bounds and testing minimax

Lemma 4. *Assume only F1, F2, and F3. Set* $C = \log_2 b - \log_2 w$*, where* $w = \lambda_{\min}(\omega) > 0$*. For every* n *and* $0 < \varepsilon < 1$*,*

$$0 \le E_n \le nC, \tag{6}$$
$$(1-\varepsilon)2^{-nC} \le \beta_{\varepsilon,n} \le 1-\varepsilon, \tag{7}$$
$$(1-\varepsilon)(-\log_2 \beta_{\varepsilon,n}) \le E_n + 1. \tag{8}$$

Proof. We first prove the dimension bound

$$\rho_{RB} \le b\,\rho_R \otimes I_B, \qquad \rho_{RB} \in \mathrm{D}(R \otimes B). \tag{9}$$

For a pure state, write a Schmidt decomposition $|\phi\rangle = \sum_{j=1}^{r} \sqrt{p_j}|u_j\rangle|v_j\rangle$, where $r \le b$. For any vector $|x\rangle \in R \otimes B$, Cauchy–Schwarz gives

$$\begin{aligned} |\langle x|\phi\rangle|^2 &\le r \sum_{j=1}^{r} p_j |\langle x|u_j, v_j\rangle|^2 \\ &\le b\,\langle x|(\rho_R \otimes I_B)|x\rangle. \end{aligned}$$

In the last step, each $|v_j\rangle\langle v_j|$ is bounded by I_B. For a mixed state, decompose it into pure states, apply this bound to each component, and sum with the mixture weights. The reference marginals sum to ρ_R, proving (9).

We now apply this inequality to the channel outputs. For an arbitrary unit input vector $|\psi\rangle \in R \otimes A^n$, set

$$\begin{aligned} \rho &= (\mathrm{id}_R \otimes \mathcal{N}_n)(|\psi\rangle\langle\psi|), \\ \sigma &= (\mathrm{id}_R \otimes \mathcal{R}_\omega^{\otimes n})(|\psi\rangle\langle\psi|) \\ &= \varrho_R \otimes \omega^{\otimes n}. \end{aligned}$$

Because the joint map is $\mathrm{id}_R \otimes \mathcal{N}_n$ and $\mathcal{N}_n$ is trace preserving, the reference marginal is unchanged. That is, $\rho_R = \varrho_R$. Apply (9) with the whole output B^n, whose dimension is b^n, and then use $\omega^{\otimes n} \ge w^n I_{B^n}$. Equation (9) yields

$$\rho \le b^n w^{-n} \sigma = 2^{nC}\sigma. \tag{10}$$

The inequality in (10) implies $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$. Restrict to $\operatorname{supp}\sigma$, where σ is positive definite, and let I denote the identity on this subspace. For every $t > 0$, $\rho + tI \le 2^{nC}\sigma + tI$, so operator monotonicity of the logarithm gives

$$\log_2(\rho + tI) \le \log_2(2^{nC}\sigma + tI).$$

Multiplying by ρ and taking the trace preserves this inequality. Letting $t \downarrow 0$, using $x \log_2 x \to 0$ as $x \downarrow 0$, yields

$$\operatorname{Tr}\rho\log_2\rho \le \operatorname{Tr}\rho\log_2(2^{nC}\sigma) = nC + \operatorname{Tr}\rho\log_2\sigma,$$

where we used $\operatorname{Tr}\rho = 1$. Rearranging proves $D(\rho\|\sigma) \le nC$. This bound holds for every reference-assisted input, hence $D_{\mathrm{ch}}(\mathcal{N}_n\|\mathcal{R}_\omega^{\otimes n}) \le nC$. Since $\mathcal{R}_\omega^{\otimes n} \in \mathfrak{F}_n$, minimizing over the free alternatives gives $E_n \le nC$. Nonnegativity of relative entropy gives $E_n \ge 0$, proving (6).

For the testing lower bound, let Q be any feasible acceptance effect for the chosen input, so that $\operatorname{Tr} Q\rho \ge 1 - \varepsilon$. Equation (10) implies

$$\operatorname{Tr} Q\sigma \ge 2^{-nC}\operatorname{Tr} Q\rho \ge (1-\varepsilon)2^{-nC}.$$

The worst-case acceptance probability over $\mathfrak{F}_n$ is at least the acceptance probability of this free replacer. Taking the infimum over feasible testers therefore gives $\beta_{\varepsilon,n} \ge (1-\varepsilon)2^{-nC}$. For the upper bound, choose any unit input vector and $Q = (1-\varepsilon)I$. Every channel output has trace one, so this effect accepts both the target and every alternative with probability $1-\varepsilon$. It is thus feasible and gives $\beta_{\varepsilon,n} \le 1-\varepsilon$, completing (7).

For any feasible tester with input state $|\psi\rangle$ and effect Q, set $\rho = (\mathrm{id}_R \otimes \mathcal{N}_n)(|\psi\rangle\langle\psi|)$ and $\sigma_{\mathcal{M}} = (\mathrm{id}_R \otimes \mathcal{M})(|\psi\rangle\langle\psi|)$. Write $p = \operatorname{Tr} Q\rho \ge 1-\varepsilon$ and $t = \sup_{\mathcal{M}\in\mathfrak{F}_n} \operatorname{Tr} Q\sigma_{\mathcal{M}}$. Let $h_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$ denote the binary entropy, which is at most one. For each $\mathcal{M} \in \mathfrak{F}_n$, put $q = \operatorname{Tr} Q\sigma_{\mathcal{M}}$. The binary measurement $\{Q, I-Q\}$ sends ρ and $\sigma_{\mathcal{M}}$ to the probability distributions $(p, 1-p)$ and $(q, 1-q)$. The definition of channel relative entropy and data processing give

$$\begin{aligned}
D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}) &\ge D(\rho \| \sigma_{\mathcal{M}}) \\
&\ge D((p, 1-p) \| (q, 1-q)) \\
&= p\log_2 \frac{1}{q} + (1-p)\log_2 \frac{1}{1-q} - h_2(p) \\
&\ge p\log_2 \frac{1}{q} - h_2(p).
\end{aligned}$$

The last step drops the nonnegative term $(1-p)\log_2(1/(1-q))$. At boundary probabilities these expressions are interpreted by continuity; if $q = 0$, then $p > 0$ makes the relative entropy infinite and the desired lower bound is immediate. Next, $q \le t \le 1$ and $h_2(p) \le 1$ imply

$$\begin{aligned}
p\log_2 \frac{1}{q} - h_2(p) &\ge p\log_2 \frac{1}{t} - 1 \\
&\ge (1-\varepsilon)\log_2 \frac{1}{t} - 1.
\end{aligned}$$

For the fixed feasible tester, t is its worst-case acceptance probability over the entire family $\mathfrak{F}_n$. Thus the lower bound $(1-\varepsilon)\log_2(1/t) - 1$ is independent of the individual alternative $\mathcal{M}$. Taking the infimum over $\mathcal{M} \in \mathfrak{F}_n$ gives

$$E_n = \inf_{\mathcal{M}\in\mathfrak{F}_n} D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}) \ge (1-\varepsilon)\log_2 \frac{1}{t} - 1.$$

By the definition of $\beta_{\varepsilon,n}$, there is a sequence of feasible testers whose worst-case acceptance probabilities t_j converge to $\beta_{\varepsilon,n}$. Applying the preceding bound to each tester and taking $j \to \infty$ yields

$$E_n \ge (1-\varepsilon)\log_2 \frac{1}{\beta_{\varepsilon,n}} - 1,$$

which is (8). The lower bound $\beta_{\varepsilon,n} \ge (1-\varepsilon)2^{-nC} > 0$ ensures that the logarithm is finite and continuous at this limit. □

The tester representation makes the worst-case testing optimization a compact convex minimax problem.

Lemma 5. *For a channel $\mathcal{T} : \mathrm{L}(H) \to \mathrm{L}(K)$, a nonempty compact convex set $\mathcal{F} \subseteq \mathrm{CPTP}(H \to K)$, and $0 < \varepsilon < 1$, fix a reference system $R \simeq H$. Write $\rho^{\psi}_{\mathcal{T}} = (\mathrm{id}_R \otimes \mathcal{T})(|\psi\rangle\langle\psi|)$ and $\sigma^{\psi}_{\mathcal{M}} = (\mathrm{id}_R \otimes \mathcal{M})(|\psi\rangle\langle\psi|)$. Then*

$$\begin{aligned}
&\inf_{\substack{|\psi\rangle, Q:\ \langle\psi|\psi\rangle = 1,\ 0\le Q\le I \\ \operatorname{Tr} Q\rho^{\psi}_{\mathcal{T}} \ge 1-\varepsilon}} \sup_{\mathcal{M}\in\mathcal{F}} \operatorname{Tr} Q\sigma^{\psi}_{\mathcal{M}} \\
&\quad = \max_{\mathcal{M}\in\mathcal{F}} b_\varepsilon(\mathcal{T}, \mathcal{M}).
\end{aligned} \tag{11}$$

The infimum ranges over input states and acceptance effects satisfying the target acceptance constraint. The same pair $(|\psi\rangle, Q)$ is used for every $\mathcal{M} \in \mathcal{F}$, and an optimal pair exists.

Proof. Write the input state as $|\psi\rangle = (I_H \otimes \sqrt{\rho})|\Gamma_H\rangle$ for some $\rho \in \mathrm{D}(H)$, with the reference isometry absorbed into Q. Its physical input marginal is ρ. Using $(I_H \otimes \sqrt{\rho})|\Gamma_H\rangle = (\sqrt{\rho}^T \otimes I_H)|\Gamma_H\rangle$, we can move the input operator to the reference factor. For any channel Φ, the output is $(\sqrt{\rho}^T \otimes I_K)C_\Phi(\sqrt{\rho}^T \otimes I_K)$. Define

$$T = (\sqrt{\rho}^T \otimes I_K)Q(\sqrt{\rho}^T \otimes I_K).$$

The bounds $0 \leq Q \leq I$ imply $0 \leq T \leq \rho^T \otimes I_K$, and cyclicity of the trace gives the acceptance probability $\operatorname{Tr} TC_\Phi$.

Conversely, given $\rho \in \mathrm{D}(H)$ and $0 \leq T \leq \rho^T \otimes I_K$, use the input above and set $Q = ((\rho^T)^{-1/2} \otimes I_K)T((\rho^T)^{-1/2} \otimes I_K)$, where the inverse is taken on the support and Q is extended by zero. The bound on T ensures that it is supported on $\operatorname{supp}(\rho^T) \otimes K$ and that $0 \leq Q \leq I$, so this choice recovers T and realizes the same acceptance probabilities. Thus testers are equivalently represented by

$$\begin{gathered} \rho \in \mathrm{D}(H), \qquad 0 \leq T \leq \rho^T \otimes I_K, \\ \Pr(\text{accept} \mid \Phi) = \operatorname{Tr} TC_\Phi. \end{gathered} \tag{12}$$

The feasible set $\mathcal{X}_\varepsilon$ of pairs (ρ, T) satisfying (12) and $\operatorname{Tr} TC_{\mathcal{T}} \geq 1 - \varepsilon$ is nonempty, compact, and convex. Nonemptiness follows by taking $T = (1-\varepsilon)\rho^T \otimes I_K$ for any $\rho \in \mathrm{D}(H)$. The constraints are closed and convex, and $0 \leq T \leq I_{H\otimes K}$ ensures boundedness. Since $\operatorname{Tr} TC_{\mathcal{M}}$ is continuous and bilinear, Sion's minimax theorem [16] gives

$$\min_{(\rho,T)\in\mathcal{X}_\varepsilon} \max_{\mathcal{M}\in\mathcal{F}} \operatorname{Tr} TC_{\mathcal{M}} = \max_{\mathcal{M}\in\mathcal{F}} \min_{(\rho,T)\in\mathcal{X}_\varepsilon} \operatorname{Tr} TC_{\mathcal{M}}.$$

The inner minimum on the right is $b_\varepsilon(\mathcal{T}, \mathcal{M})$, proving (11). Compactness and continuity ensure attainment of the tester minimum; the maximum over $\mathcal{F}$ is attained because the inner minimum is upper semicontinuous in $\mathcal{M}$. The realization above gives an optimal input and effect with $R \simeq H$. □

3.2 Uniform auxiliary map approximation

For channels $\mathcal{T}, \mathcal{M} : H \to K$ and $t \geq 0$, put

$$h_t(\mathcal{T}\|\mathcal{M}) = \sup_{|\psi\rangle,\, 0\leq Q\leq I} (p - tq),$$

where p, q are the two acceptance probabilities for a common reference-assisted input and effect. The following one-shot statement keeps both auxiliary spaces fixed. This is essential for subtracting the auxiliary maps and representing the residual. The approximation bound was established in [10, Prop. 4.2]; we include its proof and the domination consequences used below for completeness.

Lemma 6 (Testing and auxiliary map approximation)**.** *Fix Stinespring isometries $V : H \to K \otimes E$ and $W : H \to K \otimes F$ of channels $\mathcal{T}$ and $\mathcal{M}$. For every $t \geq 0$ there is an auxiliary linear map $A : F \to E$ with*

$$\|A\|_\infty \leq \sqrt{t}, \qquad \|V - (I_K \otimes A)W\|_\infty^2 \leq h_t(\mathcal{T}\|\mathcal{M}). \tag{13}$$

In particular, if a channel $\mathcal{L} \leq_{\mathrm{CP}} t\mathcal{M}$ satisfies $\|\mathcal{L} - \mathcal{T}\|_\diamond \leq \delta$, the squared error in (13) is at most $\delta/2$. If $\mathcal{T} \leq_{\mathrm{CP}} t\mathcal{M}$, one can take zero error. Conversely, the CP map obtained from $(I_K \otimes A)W$ by tracing out the auxiliary system E is dominated by $\|A\|_\infty^2\mathcal{M}$.

Proof. The tester representation (12) and semidefinite duality give, with $\Delta = C_{\mathcal{T}} - tC_{\mathcal{M}}$,

$$\begin{aligned} h_t(\mathcal{T}\|\mathcal{M}) &= \max_{\rho\in\mathrm{D}(H),\, 0\leq T\leq\rho^T\otimes I_K} \operatorname{Tr} \Delta T \\ &= \min_{Y\geq 0,\, Y\geq\Delta} \|\operatorname{Tr}_K Y\|_\infty. \end{aligned}$$

This is an instance of the semidefinite approach to completely bounded norms [17]. For completeness, a positive multiplier Y for the tester upper bound and a scalar μ for $\operatorname{Tr}\rho = 1$ give the dual constraints $Y \geq \Delta$ and $(\operatorname{Tr}_K Y)^T \leq \mu I$. Strict feasibility follows by choosing a faithful ρ, a sufficiently small positive scalar tester, and, on the dual side, sufficiently large scalar Y, μ. A dual sublevel set is compact because $\operatorname{Tr} Y \leq (\dim H)\|\operatorname{Tr}_K Y\|_\infty$, so a minimizing Y exists.

Let F_V, F_W be the matrices whose columns are the vectorized Kraus operators of the prescribed dilations. Thus $F_V F_V^\dagger = C_{\mathcal{T}}$ and $F_W F_W^\dagger = C_{\mathcal{M}}$. Set $G = tC_{\mathcal{M}} + Y$ and let G^+ denote its Moore–Penrose inverse. Define

$$F_0 = tC_{\mathcal{M}} G^+ F_V, \qquad F_1 = Y G^+ F_V.$$

Since $C_{\mathcal{T}} \leq G$, every column of F_V lies in $\operatorname{supp} G$. Thus $F_0 + F_1 = GG^+ F_V = F_V$. Put $Z_0 = tC_{\mathcal{M}}$ and $Z_1 = Y$. For $j = 0, 1$,

$$F_j F_j^\dagger = Z_j G^+ C_{\mathcal{T}} G^+ Z_j \leq Z_j G^+ Z_j \leq Z_j.$$

The first inequality uses $C_{\mathcal{T}} \leq G$; the second follows from $0 \leq Z_j \leq G$. To justify the latter even for singular G, the operator $(G^+)^{1/2} Z_j^{1/2}$ is a contraction, because its product with its adjoint is at most the projection onto $\operatorname{supp} G$. Hence $Z_j^{1/2} G^+ Z_j^{1/2} \leq I$, and congruence by $Z_j^{1/2}$ gives $Z_j G^+ Z_j \leq Z_j$. In particular, $F_0 F_0^\dagger \leq tF_W F_W^\dagger$ and $F_1 F_1^\dagger \leq Y$.

The first bound implies that the columns of F_0 lie in the range of F_W. Taking $D = F_W^+ F_0$ therefore gives $F_0 = F_W D$, and

$$DD^\dagger \leq tF_W^+ F_W F_W^\dagger (F_W^+)^\dagger = tP_{(\ker F_W)^\perp} \leq tI_F.$$

Here $P_{(\ker F_W)^\perp}$ is the orthogonal projection onto the indicated subspace of F. With the column convention above, $A = D^T : F \to E$ is the auxiliary map and $\|A\|_\infty = \|D\|_\infty \leq \sqrt{t}$. Let $R = V - (I_K \otimes A)W$. Its Kraus-column matrix is F_1, so $\operatorname{Tr}_K F_1 F_1^\dagger = (R^\dagger R)^T$ in our Choi convention. Consequently,

$$\|R\|_\infty^2 = \|\operatorname{Tr}_K F_1 F_1^\dagger\|_\infty \leq \|\operatorname{Tr}_K Y\|_\infty = h_t(\mathcal{T}\|\mathcal{M}).$$

This proves (13) in the prescribed auxiliary spaces, with one auxiliary map A for all inputs.

If $\mathcal{L} \leq_{\mathrm{CP}} t\mathcal{M}$, every tester satisfies $p - tq \leq p - p_{\mathcal{L}} \leq \delta/2$, since the two target outputs have equal trace. For exact domination take $\mathcal{L} = \mathcal{T}$. Finally, the Kraus-column matrix of $(I_K \otimes A)W$ is $F_W A^T$; its product with its adjoint is bounded by $\|A\|_\infty^2 C_{\mathcal{M}}$, proving the converse assertion. □

3.3 Completion of a trace defect

We will also use a completion estimate in which the free mixture is weighted by the actual trace defect.

Lemma 7 (Completion at the cost of the trace defect)**.** *Assume only that $\mathfrak{F}_n$ is convex and contains $\mathcal{R}_\omega^{\otimes n}$. Let $\mathcal{Q}$ be CP and trace nonincreasing, and suppose $\mathcal{Q} \leq_{\mathrm{CP}} \lambda\mathcal{M}$ for $\mathcal{M} \in \mathfrak{F}_n$ and $\lambda \geq 1$. Put $D = I - \mathcal{Q}^\dagger(I)$ and $d = \|D\|_\infty$. Then there are a channel $\mathcal{L}$ and $\mathcal{S} \in \mathfrak{F}_n$ with*

$$\mathcal{L} \leq_{\mathrm{CP}} (\lambda + d)\mathcal{S}, \qquad \|\mathcal{L} - \mathcal{N}_n\|_\diamond \leq 2\|\mathcal{Q} - \mathcal{N}_n\|_\diamond.$$

Here $0 \leq d \leq \min\{1, \|\mathcal{Q} - \mathcal{N}_n\|_\diamond\}$.

Proof. Set $\mathcal{C}_D(X) = \operatorname{Tr}(DX)\omega^{\otimes n}$ and $\mathcal{L} = \mathcal{Q} + \mathcal{C}_D$. Then $\mathcal{L}$ is CPTP, $\mathcal{C}_D \leq_{\mathrm{CP}} d\mathcal{R}_\omega^{\otimes n}$, and $\|\mathcal{C}_D\|_\diamond = d$. The free channel

$$\mathcal{S} = \frac{\lambda\mathcal{M} + d\mathcal{R}_\omega^{\otimes n}}{\lambda + d}$$

gives the claimed domination by convexity. For every input state ρ, $\operatorname{Tr} D\rho = \operatorname{Tr}(\mathcal{N}_n - \mathcal{Q})(\rho)$ is at most $\|\mathcal{N}_n - \mathcal{Q}\|_\diamond$. Maximizing over ρ bounds d; the triangle inequality proves the error estimate. □

4 Exponential approximation and the Stein identity

We prove Theorems 2 and 3 under the three conditions in Definition 1, using the testing and approximation tools of Section 3. For a fixed testing tolerance, we start from the lower limit $b_\varepsilon = \liminf_n a_{\varepsilon,n}/n$, without assuming convergence of either the testing exponent or E_n/n.

The proof has two amplification steps. First, Proposition 8 turns any unbounded sequence of approximations with vanishing error into exponentially accurate approximations at every sufficiently large blocklength, with an arbitrarily small increase in rate. Next, Proposition 9 uses testing to construct such a sequence at a rate below a previously available approximation rate. Starting from exact replacer domination and iterating this improvement gives exponential approximation at every rate above b_ε. These approximations force the testing upper limits at all tolerances to be at most b_ε. The entropy comparison in Lemma 10 then identifies the common ordinary limit.

4.1 Amplification and exact padding

The following proposition starts from CPTP approximations to $\mathcal{N}^{\otimes k}$ at an unbounded sequence of blocklengths k, with vanishing diamond error and a bounded domination rate, as specified in (14). Under convexity, tensor closure, and the faithful-replacer assumption, it produces exponentially accurate approximations at every sufficiently large blocklength. The proof of Proposition 9 in the next subsection constructs the required sequence by combining an auxiliary map approximation obtained from testing with an approximation already available at a higher domination rate. That construction uses only Definition 1. The tensor truncation below adapts the fixed-block amplification of [10, Lemma 4.5] to free dominators and exact trace preservation.

Proposition 8 (Amplification from vanishing error)**.** *Suppose each $\mathfrak{F}_n$ is convex and $\mathfrak{F}$ satisfies F2 and F3. Fix a finite $r_0 \geq 0$. Assume that, along an unbounded sequence of blocklengths k, there are CPTP maps $\mathcal{L}_k$, free channels $\mathcal{M}_k \in \mathfrak{F}_k$, and numbers $\lambda_k \geq 1$ such that*

$$\begin{gathered} \mathcal{L}_k \leq_{\mathrm{CP}} \lambda_k \mathcal{M}_k, \qquad \delta_k := \|\mathcal{L}_k - \mathcal{N}_k\|_\diamond \longrightarrow 0, \\ \limsup_{k\to\infty} \frac{\log_2 \lambda_k}{k} \leq r_0. \end{gathered} \tag{14}$$

The limits in (14) are taken along that sequence. Then for every $S > r_0$ there are free dominators and CPTP approximations satisfying (5).

Proof. Write $C = \log_2 b - \log_2 w$. Equation (10), applied to a maximally entangled input, gives $\mathcal{N}_k \leq_{\mathrm{CP}} 2^{kC}\mathcal{R}_\omega^{\otimes k}$. If $S \geq C$, the exact target and the free replacer prove the claim. Otherwise choose $r_0 < r < S < C$ and set $L = C + 1$.

Step 1: Two auxiliary maps between the same spaces. Fix a Stinespring isometry $V : A \to B \otimes E$ of $\mathcal{N}$. For each sufficiently large k in the given sequence, form the free channel

$$\widehat{\mathcal{M}}_k = \tfrac{1}{2}\mathcal{M}_k + \tfrac{1}{2}\mathcal{R}_\omega^{\otimes k}$$

and fix a dilation $W_k : A^k \to B^k \otimes F_k$ of it. By Lemma 6, there are auxiliary maps $A_k, B_k : F_k \to E^{\otimes k}$ such that

$$\begin{gathered} X_k = (I \otimes A_k)W_k, \qquad V^{\otimes k} = (I \otimes B_k)W_k, \\ \|V^{\otimes k} - X_k\|_\infty \leq e_k := \sqrt{\delta_k/2}, \\ \|A_k\|_\infty \leq 2^{kr/2}, \qquad \|B_k\|_\infty \leq 2^{kL/2}. \end{gathered}$$

For A_k use $\mathcal{L}_k \leq_{\mathrm{CP}} 2\lambda_k\widehat{\mathcal{M}}_k$ and $\log_2(2\lambda_k)/k < r$ eventually. For B_k use $\mathcal{N}_k \leq_{\mathrm{CP}} 2^{kC+1}\widehat{\mathcal{M}}_k$ and $kC + 1 \leq kL$. Thus $B_k - A_k$ represents the exact residual in the same codomain, even though the dominator may vary with k.

Step 2: Fix the block and truncate its exact expansion. Choose $0 < \eta < 1$ such that $\bar{S} = r + \eta(L - r) < S$, and put $z = 4^{1/\eta}$. Choose and then fix k in that sequence large enough that

$$e_k \leq \frac{1}{1+z}, \qquad v := \bar{S} + \frac{2\log_2 3}{k} < S.$$

Keep this k fixed for the remainder of the proof; the number m of repeated blocks will tend to infinity. Abbreviate $X = X_k$ and $H = V^{\otimes k} - X$. The identity $(X + H)^{\otimes m} = V^{\otimes km}$ is exact. Let Z_m be its subset expansion restricted to terms with at most $\lfloor \eta m \rfloor$ factors H. Since $\|X\|_\infty \leq 1 + e_k$ and $\|H\|_\infty \leq e_k$, the omitted tail obeys

$$\begin{aligned} \|Z_m - V^{\otimes km}\|_\infty &\leq \sum_{j > \eta m} \binom{m}{j} (1 + e_k)^{m-j} e_k^j \\ &\leq z^{-\eta m}(1 + e_k + z e_k)^m \leq 2^{-m}. \end{aligned}$$

The second inequality follows by multiplying each omitted term by $z^{j-\eta m} \geq 1$ and including all terms in the binomial sum.

Replace X and H in the same subset sum by A_k and $B_k - A_k$. This defines $D_m : F_k^{\otimes m} \to E^{\otimes km}$ with $Z_m = (I \otimes D_m) W_k^{\otimes m}$ and

$$\begin{aligned} \|D_m\|_\infty &\leq \sum_{j \leq \eta m} \binom{m}{j} 2^j 2^{k[(m-j)r + jL]/2} \\ &\leq 3^m 2^{km\bar{S}/2} = 2^{kmv/2}. \end{aligned}$$

We used $\|B_k - A_k\|_\infty \leq 2\, 2^{kL/2}$ and $r < L$. The tensor-power dominator $\widehat{\mathcal{M}}_k^{\otimes m}$ belongs to $\mathfrak{F}_{km}$, although it may contain arbitrary correlations inside each block.

Step 3: Pad exactly and restore trace preservation. For $n = km + \ell$, $0 \leq \ell < k$, use

$$\mathcal{T}_n = \widehat{\mathcal{M}}_k^{\otimes m} \otimes \mathcal{R}_\omega^{\otimes \ell} \in \mathfrak{F}_n.$$

For $\ell = 0$ the last factor is omitted. The exact case of Lemma 6 applied to the replacer gives a representation of $V^{\otimes \ell}$ with auxiliary map norm at most $2^{\ell C/2}$. Tensoring D_m with this auxiliary map yields a representation of $\widehat{V}_n$ satisfying

$$\|\widehat{V}_n - V^{\otimes n}\|_\infty \leq e_n := 2^{-m}.$$

The corresponding auxiliary map relative to $\mathcal{T}_n$ has squared operator norm at most $\lambda_n := 2^{kmv + \ell C}$. Exact padding does not increase the error. The contraction $T_n = \widehat{V}_n/(1 + e_n)$ defines the subchannel $\mathcal{Q}_n(X) = \mathrm{Tr}_{E^n} T_n X T_n^\dagger$. It satisfies $\mathcal{Q}_n \leq_{\mathrm{CP}} \lambda_n \mathcal{T}_n$ and

$$\|T_n - V^{\otimes n}\|_\infty \leq 2e_n, \qquad \|\mathcal{Q}_n - \mathcal{N}_n\|_\diamond \leq 4e_n,$$

by (2) and contraction of the partial trace. Lemma 7 gives CPTP $\mathcal{L}_n$ and free $\mathcal{S}_n$ with

$$\mathcal{L}_n \leq_{\mathrm{CP}} (\lambda_n + 1)\mathcal{S}_n, \qquad \|\mathcal{L}_n - \mathcal{N}_n\|_\diamond \leq 8e_n.$$

Since k is fixed and $v < S$, $\log_2(\lambda_n + 1) \leq nv + \ell(C - v) + 1 \leq nS$ eventually. Finally $8e_n = 8\, 2^{-\lfloor n/k \rfloor} \leq 16 e^{-(\ln 2)n/k}$. This proves the claim for every sufficiently large integer n, with $K_S = 16$ and $\gamma_S = (\ln 2)/k$. $\square$

4.2 Iterative reduction of the approximation rate

Testing gives a uniform auxiliary map approximation with error strictly below one, which need not tend to zero. The next argument combines it with an exponentially accurate approximation at a higher rate. The result improves that higher rate by a fixed fraction. Iteration reaches every rate above the lower limiting testing exponent. This adapts the two-rate tensor amplification of [10, Lemma 4.4] to the present composite setting.

Proposition 9 (From testing to exponential approximation)**.** *Assume F1, F2, and F3. For any fixed $0<\varepsilon<1$, put $b_\varepsilon=\liminf_n a_{\varepsilon,n}/n$. Then for every $S>b_\varepsilon$ there are free dominators and CPTP approximations satisfying (5).*

Proof. Let $P(s)$ denote the existence of the approximations in (5) at rate s, for all sufficiently large blocklengths. Exact replacer domination proves $P(C)$. If $S\geq C$ the conclusion is immediate. Otherwise fix $b_\varepsilon<r<S<C$ and an increasing sequence of integers k along which $a_{\varepsilon,k}/k\to b_\varepsilon$.

Step 1: A rough approximation from testing. Choose a hardest $\mathcal{M}_k\in\mathfrak{F}_k$ from Lemma 5, and set $t_k=\varepsilon/\beta_{\varepsilon,k}$. Every tester with target acceptance $p\geq 1-\varepsilon$ has alternative acceptance $q\geq\beta_{\varepsilon,k}$, hence $p-t_kq\leq 1-\varepsilon$. For $p<1-\varepsilon$ the same upper bound holds because $q\geq 0$. Thus

$$h_{t_k}(\mathcal{N}_k\|\mathcal{M}_k)\leq 1-\varepsilon. \tag{15}$$

Write $d=\sqrt{1-\varepsilon}<1$ and fix a dilation V of $\mathcal{N}$.

Step 2: One improvement of a higher rate. Suppose $r<s\leq C$ and $P(s)$ holds. Let $\mathcal{L}_k\leq_{\mathrm{CP}}2^{ks}\mathcal{T}_k$ be its CPTP approximations, with $\mathcal{T}_k\in\mathfrak{F}_k$ and diamond error at most $Ke^{-\gamma k}$. For each sufficiently large k in the selected sequence, fix a dilation W_k of the common free dominator

$$\mathcal{U}_k=\tfrac{1}{2}\mathcal{M}_k+\tfrac{1}{2}\mathcal{T}_k.$$

Since $\mathcal{U}_k\geq_{\mathrm{CP}}\mathcal{M}_k/2$, the testing tail at coefficient $2t_k$ is at most the tail in (15). Apply Lemma 6 twice, with the same prescribed $V^{\otimes k}$ and W_k, to obtain

$$\begin{gathered}X_k=(I\otimes A_k)W_k,\quad Y_k=(I\otimes B_k)W_k,\\ \|V^{\otimes k}-X_k\|_\infty\leq d,\qquad \|V^{\otimes k}-Y_k\|_\infty\leq e_k,\\ e_k:=\sqrt{K/2}e^{-\gamma k/2},\\ \|A_k\|_\infty\leq\sqrt{2t_k}\leq 2^{kr/2},\quad \|B_k\|_\infty\leq\sqrt{2}\,2^{ks/2}.\end{gathered}$$

The bound on A_k holds eventually by the choice of r. Put $H_k=Y_k-X_k$, $A_0=1+d$, and $B_0=(1+d)/2<1$. For sufficiently large k,

$$\|X_k\|_\infty\leq A_0,\qquad \|H_k\|_\infty\leq d+e_k\leq B_0.$$

Choose $0<\eta<1$ and $\kappa>0$, depending only on d, so that

$$h_2(x)\ln 2+(1-x)\ln A_0+x\ln B_0\leq-\kappa\quad(\eta\leq x\leq 1). \tag{16}$$

Such a choice exists by continuity, since the expression equals $\ln B_0<0$ at $x=1$.

Take $m=k$ and truncate $(X_k+H_k)^{\otimes m}=Y_k^{\otimes m}$ to terms with at most $\lfloor\eta m\rfloor$ residual factors. Denote the resulting operator by Z_k. The bound $\binom{m}{j}\leq 2^{mh_2(j/m)}$ gives

$$\begin{aligned}\|Z_k-V^{\otimes km}\|_\infty&\leq(m+1)e^{-\kappa m}+me_k(1+e_k)^{m-1}\\&=:u_k\longrightarrow 0.\end{aligned}$$

The second term is the telescoping bound for $Y_k^{\otimes m}-V^{\otimes km}$; it vanishes because $m=k$ and e_k decays exponentially.

Replacing X_k,H_k by A_k,B_k-A_k in this same subset sum gives an auxiliary map D_k relative to $W_k^{\otimes m}$. Set $\bar{s}=(1-\eta)r+\eta s$. Since $r<s$ and $\|B_k-A_k\|_\infty\leq 3\,2^{ks/2}$,

$$\begin{aligned}\|D_k\|_\infty&\leq\sum_{j\leq\eta m}\binom{m}{j}3^j2^{k[(m-j)r+js]/2}\\&\leq 4^m2^{km\bar{s}/2}.\end{aligned}$$

Tensor closure gives $\mathcal{U}_k^{\otimes m} \in \mathfrak{F}_{km}$. Scale Z_k by $1 + u_k$ to get a contraction, then apply Lemma 7, exactly as in the proof of Proposition 8. This gives CPTP approximations at lengths $N = k^2$, with error at most $8u_k \to 0$ and domination cost at most

$$\lambda_N = 2^{N\bar{s}+4k} + 1.$$

In particular $\limsup(\log_2 \lambda_N)/N \le \bar{s}$ along this unbounded sequence. By Proposition 8, $P(s')$ holds for every $s' > \bar{s}$. Here both k and $m = k$ grow, and the truncation alone supplies vanishing error only along the lengths $N = k^2$. The application of Proposition 8 then fixes one accurate block and supplies exponential accuracy at all sufficiently large lengths, including those outside this sequence.

Step 3: Iterate a fixed contraction of the rate gap. The choice of η in (16) depends only on $d = \sqrt{1-\varepsilon}$, so the same η works at every iteration. Set $\theta = (1+\eta)/2 \in (\eta, 1)$ and $s_j = r + \theta^j (C - r)$. Starting from $P(s_0) = P(C)$, Step 2 proves $P(s_{j+1})$ because

$$s_{j+1} > (1-\eta)r + \eta s_j.$$

After finitely many steps, $s_j < S$, proving $P(S)$. All constants may depend on the chosen rate margin. No quantitative completion estimate, permutation average, or marginal reduction is used in this argument. $\square$

4.3 Entropy from an accurate free-dominated channel

The following comparison identifies the entropy rate without symmetry or an entropy minimax argument. Mixing the dominator with the replacer gives a bound on $\log_2 \sigma - \log_2(\mu \otimes I)$, where σ is the resulting reference–output state and μ its reference marginal. On $\operatorname{supp} \mu \otimes B^n$, this bound is uniform over input states.

Lemma 10 (Entropy comparison)**.** *Suppose $\mathcal{L} \le_{\mathrm{CP}} 2^z \mathcal{S}$ for a CPTP map $\mathcal{L}$ and $\mathcal{S} \in \mathfrak{F}_n$, and $\|\mathcal{L} - \mathcal{N}_n\|_\diamond \le \delta \le 1$. Then*

$$E_n \le z + 1 + \frac{\delta}{2}[n(C + \log_2(ab)) + 1] + h_2(\delta/2).$$

Only convexity of $\mathfrak{F}_n$ and membership of $\mathcal{R}_\omega^{\otimes n}$ in $\mathfrak{F}_n$ are needed.

Proof. Set $\mathcal{M} = (\mathcal{S} + \mathcal{R}_\omega^{\otimes n})/2 \in \mathfrak{F}_n$. On a common pure input with $R \simeq A^n$, let ρ, ν, σ be the outputs of $\mathcal{N}_n, \mathcal{L}, \mathcal{M}$, respectively, and let μ be their common reference marginal. The domination $\nu \le 2^{z+1}\sigma$ gives $D(\nu\|\sigma) \le z + 1$. On $\operatorname{supp} \mu \otimes B^n$, both logarithms in $H = \log_2 \sigma - \log_2(\mu \otimes I)$ are well-defined. The replacer component and (9) imply

$$\tfrac{1}{2} w^n \mu \otimes I \le \sigma \le b^n \mu \otimes I.$$

Operator monotonicity of the logarithm therefore gives

$$(n \log_2 w - 1)I \le H \le n \log_2 b\, I.$$

Put $t = \|\rho - \nu\|_1/2 \le \delta/2$. Since $\rho - \nu$ has zero reference marginal and zero trace,

$$|\operatorname{Tr}(\rho - \nu)\log_2 \sigma| = |\operatorname{Tr}(\rho - \nu)H| \le t(nC + 1).$$

Here a trace-zero Hermitian matrix of trace norm $2t$ pairs with an operator of spectral width $nC+1$ by at most $t(nC+1)$. Entropy continuity in dimension at most $(ab)^n$ gives $|S(\rho) - S(\nu)| \le tn \log_2(ab) + h_2(t)$ [15, Thm. 5.26]. This also holds trivially when $ab = 1$. Subtracting the two relative entropies and using $t \le \delta/2 \le 1/2$ proves the displayed bound uniformly over the input. Take the supremum over inputs and then use $E_n \le D_{\mathrm{ch}}(\mathcal{N}_n\|\mathcal{M})$. $\square$

4.4 Identification of the asymptotic rate

Proof of Theorems 2 and 3. Step 1: From a lower limit to a common ordinary limit. Fix $0 < \varepsilon < 1$ and put $b_\varepsilon = \liminf_n a_{\varepsilon,n}/n \in [0, C]$. For every $S > b_\varepsilon$, Proposition 9 gives free $\mathcal{S}_n$ and CPTP $\mathcal{L}_n$, for all sufficiently large n, with domination 2^{nS} and error at most $d_n = K_S e^{-\gamma_S n}$. Fix $0 < \varepsilon' < 1$ and a tester $(R, |\psi\rangle, Q)$ with target acceptance at least $1 - \varepsilon'$. Let $\rho = \rho_n^\psi$ be its target output and let $\widetilde{\rho} = (\mathrm{id}_R \otimes \mathcal{L}_n)(|\psi\rangle\langle\psi|)$. The diamond bound gives $\|\widetilde{\rho} - \rho\|_1 \le d_n$. Both outputs have trace one, so their difference is traceless. For an acceptance effect $0 \le Q \le I$, this gives the sharper probability bound

$$\operatorname{Tr} Q\widetilde{\rho} \ge \operatorname{Tr} Q\rho - \tfrac{1}{2}\|\widetilde{\rho} - \rho\|_1 \ge 1 - \varepsilon' - d_n/2.$$

Moreover, $\mathcal{L}_n \le_{\mathrm{CP}} 2^{nS}\mathcal{S}_n$ implies $\widetilde{\rho} \le 2^{nS}\sigma^\psi_{\mathcal{S}_n}$ for this reference-assisted input. Since $\mathcal{S}_n \in \mathfrak{F}_n$, the tester's largest acceptance probability over free alternatives obeys

$$\begin{aligned} q_n &:= \sup_{\mathcal{M} \in \mathfrak{F}_n} \operatorname{Tr} Q\sigma^\psi_{\mathcal{M}} \\ &\ge \operatorname{Tr} Q\sigma^\psi_{\mathcal{S}_n} \ge 2^{-nS} \operatorname{Tr} Q\widetilde{\rho} \ge 2^{-nS}(1 - \varepsilon' - d_n/2). \end{aligned}$$

Minimizing this uniform bound over feasible testers gives

$$\beta_{\varepsilon',n} \ge 2^{-nS}(1 - \varepsilon' - d_n/2).$$

For fixed S and $0 < \varepsilon' < 1$, we have $d_n \to 0$, so $1 - \varepsilon' - d_n/2 \to 1 - \varepsilon' > 0$. The factor is therefore positive for all sufficiently large n, and its logarithm converges to the finite constant $\log_2(1 - \varepsilon')$. Dividing this bounded logarithm by n gives $n^{-1}\log_2(1 - \varepsilon' - d_n/2) \to 0$. To apply this observation, take negative logarithms of the testing bound and use $a_{\varepsilon',n} = -\log_2 \beta_{\varepsilon',n}$:

$$\frac{a_{\varepsilon',n}}{n} \le S - \frac{1}{n}\log_2(1 - \varepsilon' - d_n/2).$$

The second term on the right tends to zero, so $\limsup_n a_{\varepsilon',n}/n \le S$. Since this holds for every $S > b_\varepsilon$, letting $S \downarrow b_\varepsilon$ yields

$$\limsup_n \frac{a_{\varepsilon',n}}{n} \le \liminf_n \frac{a_{\varepsilon,n}}{n} \qquad (0 < \varepsilon, \varepsilon' < 1).$$

Taking $\varepsilon' = \varepsilon$ proves ordinary convergence at each tolerance; interchanging the two tolerances shows that all limits coincide. Hence

$$\frac{a_{\varepsilon,n}}{n} \longrightarrow R \qquad (0 < \varepsilon < 1).$$

The same construction already gives exponential free approximation at each $S > R$. Neither this convergence argument nor the approximation construction assumes additivity of channel relative entropy or near subadditivity of the testing quantity.

Step 2: Identification of the entropy limit. The weak converse and Lemma 10 identify the common limit. The weak converse (8) gives $\liminf_n E_n/n \ge (1 - \varepsilon)R$ for every fixed ε. Letting $\varepsilon \downarrow 0$ after this limit yields $\liminf_n E_n/n \ge R$. Conversely, for each $S > R$ the approximations from Step 1 have $z = nS$ and exponentially vanishing $\delta = d_n$ in Lemma 10. Consequently $\limsup_n E_n/n \le S$. Let $S \downarrow R$. Thus the ordinary entropy limit exists and equals R. The bound $0 \le R \le C$ follows from Lemma 4.

Theorem 3 follows from the approximation already obtained in Step 1 and the identification in Step 2. □

5 Asymptotic consequences

The Stein identity and exponential approximation give three consequences under the basic assumptions in Definition 1: asymptotic equipartition for trace-preserving smoothing, an exponential strong converse, and stability under vanishing diamond-norm perturbations.

5.1 CPTP-smoothed asymptotic equipartition

Classically, the smallest constant λ such that $P(x) \le \lambda Q(x)$ for every x is the largest likelihood ratio. Its logarithm is max-relative entropy. For quantum states, positive operator order replaces pointwise order:

$$D_{\max}(\rho\|\sigma) = \log_2 \inf\{\lambda \ge 1 : \rho \le \lambda\sigma\}.$$

The factor $2^{D_{\max}(\rho\|\sigma)}$ bounds every measurement acceptance ratio [18]. For channels, complete positivity imposes the same bound simultaneously on every reference-assisted input. Define the optimized max-relative entropy of a channel by

$$D_{\max}(\mathcal{L}\|\mathfrak{F}_n) = \inf_{\mathcal{S}\in\mathfrak{F}_n} \log_2 \inf\{\lambda \ge 1 : \mathcal{L} \le_{\mathrm{CP}} \lambda\mathcal{S}\}. \tag{17}$$

The restriction $\lambda \ge 1$ follows already by taking traces on a normalized input. If $\mathcal{L} \le_{\mathrm{CP}} \lambda\mathcal{S}$ and $\lambda > 1$, the map $\mathcal{R} = (\lambda\mathcal{S} - \mathcal{L})/(\lambda - 1)$ is a channel, so

$$\mathcal{S} = \lambda^{-1}\mathcal{L} + (1 - \lambda^{-1})\mathcal{R}.$$

Thus $2^{-D_{\max}}$ describes the largest possible channel weight inside a free dominator; the residual channel need not be free. For $\delta \ge 0$ define its CPTP-smoothed value

$$Z_n^\delta = \inf_{\substack{\mathcal{L}\in\mathrm{CPTP}(A^n\to B^n)\\ \|\mathcal{L}-\mathcal{N}_n\|_\diamond\le\delta}} D_{\max}(\mathcal{L}\|\mathfrak{F}_n). \tag{18}$$

Smoothing means allowing approximation in the specified metric, not averaging channels or smoothing each output state separately. The same map $\mathcal{L}$ must approximate $\mathcal{N}_n$ on every reference-assisted input. All feasible approximations here are trace preserving. The quantities in (17) and (18) are nonnegative and finite: the Choi operator of every channel is positive with trace a^n, so $C_{\mathcal{L}} \le a^n I$, whereas $C_{\mathcal{R}_\omega^{\otimes n}} \ge w^n I$. Thus $\mathcal{L} \le_{\mathrm{CP}} (a/w)^n \mathcal{R}_\omega^{\otimes n}$. These infima are attained. Indeed, the domination constraint is closed in Choi operators, and the displayed uniform bound allows λ to be restricted to the compact interval $[1, (a/w)^n]$. The free-channel set and the diamond ball of CPTP approximations are compact as well.

The terminology asymptotic equipartition originates in classical typicality: for independent samples from P, the quantity $-n^{-1}\log_2 P(X_1, \ldots, X_n)$ converges in probability to the Shannon entropy. In relative-entropy formulations, smoothing removes the effect of exceptional likelihood ratios, and the normalized smoothed maximum converges to an entropy rate. The next theorem establishes this form of equipartition for the channel resource quantity in (18).

Theorem 11 (CPTP-smoothed resource AEP)**.** *Under the assumptions of Theorem 2, if* $0 < \delta_n \le \bar{\delta} < 2$ *and* $\log_2(1/\delta_n) = o(n)$*, then*

$$\lim_{n\to\infty} \frac{Z_n^{\delta_n}}{n} = R_{\mathfrak{F}}(\mathcal{N}). \tag{19}$$

This includes all fixed radii in $(0, 2)$*, inverse-polynomial errors, and* $\delta_n = 2^{-n^\theta}$ *for every* $0 < \theta < 1$.

Proof. For $0 < \varepsilon < 1$ and $0 \le \delta < 2(1 - \varepsilon)$, suppose $\mathcal{L} \le_{\mathrm{CP}} 2^z\mathcal{S}$ is feasible in (18). Every test with target acceptance at least $1 - \varepsilon$ accepts $\mathcal{L}$ with probability at least $1 - \varepsilon - \delta/2$. The factor $1/2$ follows from exact trace preservation and the use of the full diamond norm. Its worst-case alternative acceptance is therefore at least $2^{-z}(1 - \varepsilon - \delta/2)$. Minimizing over tests gives $\beta_{\varepsilon,n} \ge 2^{-z}(1-\varepsilon-\delta/2)$. Taking logarithms and minimizing over the feasible approximations gives the general testing lower bound

$$a_{\varepsilon,n} + \log_2(1 - \varepsilon - \delta/2) \le Z_n^\delta. \tag{20}$$

For the limit choose $\varepsilon = (1-\bar{\delta}/2)/2 > 0$. The lower logarithmic term is bounded uniformly in n, so Theorem 2 gives $\liminf_n Z_n^{\delta_n}/n \geq R_{\mathfrak{F}}(\mathcal{N})$. For any $S > R_{\mathfrak{F}}(\mathcal{N})$, Theorem 3 gives CPTP approximations at cost at most nS and error at most $K_S e^{-\gamma_S n}$. The condition on δ_n ensures that $K_S e^{-\gamma_S n} \leq \delta_n$ eventually. These approximations are feasible, giving $\limsup_n Z_n^{\delta_n}/n \leq S$. Let $S \downarrow R_{\mathfrak{F}}(\mathcal{N})$. □

Three optimization problems should be distinguished. A prescribed dominator $\mathcal{M}^{\otimes n}$ with trace-preserving smoothing, the same dominator with trace-nonincreasing smoothing, and the free-family optimization (18) have different feasible sets. Gour's counterexample [14] concerns the first of these; allowing a subchannel changes the normalization constraint. Here we keep exact trace preservation and optimize the free dominator. In the proof of this AEP, the optimization over the free dominator in (17) permits tensor amplification and completion of the trace defect. The closure axioms construct a suitable dominator at each blocklength, while the smoothing preserves the channel's exact trace constraint. Equation (19) identifies the asymptotic resource cost of CP domination with the Stein rate.

5.2 Operational threshold

For an arbitrary n-use parallel tester, let p_n denote its acceptance probability on $\mathcal{N}_n$ and let q_n denote its largest acceptance probability over $\mathfrak{F}_n$. The pair (p_n, q_n) describes its performance at any type-I error.

Theorem 12 (Operational threshold)**.** *Under the assumptions of Theorem 2, there exists a sequence of uniform testers with*

$$p_n \longrightarrow 1, \qquad -\frac{1}{n}\log_2 q_n \longrightarrow R_{\mathfrak{F}}(\mathcal{N}).$$

Conversely, fix $r > R_{\mathfrak{F}}(\mathcal{N})$. There are $K_r, c_r > 0$ such that every parallel n-use tester satisfies

$$q_n \leq 2^{-nr} \quad \Longrightarrow \quad p_n \leq K_r 2^{-c_r n}.$$

The constants are independent of the tester and its reference. In particular, exceeding the Stein rate forces the type-I error to tend to one.

Proof. For each $j \geq 2$, Theorem 2 supplies an N_j such that $|a_{1/j,n}/n - R_{\mathfrak{F}}(\mathcal{N})| \leq 1/j$ for all $n \geq N_j$. Increase the N_j so that they are strictly increasing and $N_j \geq j$. For $n \geq N_2$ use an optimal uniform tester at error $1/j(n)$, where $j(n) = \max\{j : N_j \leq n\}$. The optimum exists by Lemma 5; $p_n \geq 1 - 1/j(n) \to 1$ and its exponent converges to $R_{\mathfrak{F}}(\mathcal{N})$.

For the converse, fix $R_{\mathfrak{F}}(\mathcal{N}) < S < r$ and apply Theorem 3. A diamond error at most $d_n = K_S e^{-\gamma_S n}$ changes acceptance by at most $d_n/2$, since both maps are trace preserving. CP domination therefore gives

$$p_n \leq d_n/2 + 2^{nS} q_n.$$

If $q_n \leq 2^{-nr}$, this is at most $(K_S/2+1)2^{-c_r n}$ for all sufficiently large n, with $c_r = \min\{\gamma_S/\ln 2, r - S\} > 0$. Increasing the prefactor covers the finitely many remaining n. □

5.3 Stability under target perturbations

The same exponent persists under asymptotically vanishing perturbations of the target sequence.

Corollary 13 (Diamond-norm stability)**.** *Assume the hypotheses of Theorem 2. Let $\mathcal{T}_n : \mathrm{L}(A^n) \to \mathrm{L}(B^n)$ be arbitrary CPTP maps with $\|\mathcal{T}_n - \mathcal{N}_n\|_\diamond \to 0$. Let $\widehat{\beta}_{\varepsilon,n}$ be defined as in (4), with target $\mathcal{T}_n$ and the same alternatives $\mathfrak{F}_n$. For every fixed $0 < \varepsilon < 1$,*

$$\lim_n -\frac{1}{n}\log_2 \widehat{\beta}_{\varepsilon,n} = R_{\mathfrak{F}}(\mathcal{N}).$$

The approximating targets $\mathcal{T}_n$ may have arbitrary correlations across uses.

Proof. Choose any $u > 0$ with $u < \min\{\varepsilon, 1-\varepsilon\}$. Eventually the diamond distance is at most u, so transfer of the same input state and effect between the two targets gives

$$\beta_{\varepsilon+u,n} \le \widehat{\beta}_{\varepsilon,n} \le \beta_{\varepsilon-u,n}.$$

For the first inequality every test feasible for $\mathcal{T}_n$ is feasible for $\mathcal{N}_n$ at error $\varepsilon + u$; the other direction uses a test feasible for $\mathcal{N}_n$ at error $\varepsilon - u$. Both bounding exponents tend to $R_{\mathfrak{F}}(\mathcal{N})$ by Theorem 2. ☐

6 Quantitative finite-block completion

The asymptotic results give exponential approximation at any rate strictly above $R_{\mathfrak{F}}(\mathcal{N})$. We now prove an explicit bound at a specified blocklength and accuracy, with domination cost measured against the finite-block testing quantity $a_{\varepsilon,n}$. This refinement uses permutation invariance and marginal closure.

6.1 Additional assumptions

Write S_n for the permutation group of $[n]$. For $\pi \in S_n$, let U_π^H permute the factors of H^n according to π.

Definition 14 (Additional conditions for quantitative completion)**.** In addition to Definition 1, consider the following conditions:

(F4) For every $\pi \in S_n$, $\mathrm{Ad}_{U_\pi^B} \circ \mathcal{M} \circ \mathrm{Ad}_{(U_\pi^A)^\dagger} \in \mathfrak{F}_n$ whenever $\mathcal{M} \in \mathfrak{F}_n$.

(F5) There is a fixed positive-definite $\tau \in \mathrm{D}(A)$ such that, for every $n \ge 2$ and $\mathcal{M} \in \mathfrak{F}_n$,

$$X \longmapsto \mathrm{Tr}_{B_n} \mathcal{M}(X \otimes \tau) \quad \text{belongs to } \mathfrak{F}_{n-1}.$$

Iteration of F5 gives

$$\begin{gathered} X \longmapsto \mathrm{Tr}_{B_{m+1}\cdots B_n} \mathcal{M}(X \otimes \tau^{\otimes(n-m)}) \in \mathfrak{F}_m, \\ 1 \le m < n. \end{gathered} \tag{21}$$

Together with F4, this gives the analogous statement for every discarded subset, after relabeling the remaining sites in their original order. The input state τ and output state ω are fixed across blocklengths and may be different.

Remark 15 (The marginal assumption)*.* Condition F5 permits reduction of a free block channel by inserting τ and discarding its output. We use it for weighted discarding and to replace selected inputs and outputs in the free multiplier construction. The asymptotic results in Sections 4 and 5 hold under the three basic assumptions alone.

6.2 Free-channel domination

Theorem 16 (Free-channel domination)**.** *Let $\mathfrak{F}$ be admissible and also satisfy F4 and F5. For every fixed $0 < \varepsilon < 1$ there is a constant $0 \le K_\varepsilon < \infty$ such that, for every $n \ge 1$ and $0 < \delta \le 1/16$, there are $\mathcal{S}_n \in \mathfrak{F}_n$ and $\mathcal{L}_n \in \mathrm{CPTP}(A^n \to B^n)$ with*

$$\begin{aligned} \mathcal{S}_n &\ge_{\mathrm{CP}} \beta_{\varepsilon,n}\, 2^{-g_\varepsilon(n,\delta)} \mathcal{L}_n, \\ \|\mathcal{L}_n - \mathcal{N}_n\|_\diamond &\le \delta. \end{aligned} \tag{22}$$

Here $g_\varepsilon(n,\delta) = K_\varepsilon n^{2/3} \log_2((n+1)/\delta)$. The constant depends only on the local dimensions, the minimum eigenvalues of τ and ω, and ε. It is independent of n, δ, the target channel, and the particular free family.

The usefulness of Theorem 16 is that a free channel $\mathcal{S}_n$ contains an accurate approximation to the target as a component. Indeed, set $p_n = \beta_{\varepsilon,n}2^{-g_\varepsilon(n,\delta)}$. Since both $\mathcal{S}_n$ and $\mathcal{L}_n$ are trace preserving, the CP order in (22) gives the convex decomposition

$$\mathcal{S}_n = p_n\mathcal{L}_n + (1-p_n)\mathcal{Q}_n, \qquad \mathcal{Q}_n \in \mathrm{CPTP}(A^n \to B^n).$$

Here $0 < p_n < 1$ by Lemma 4, and the residual channel $\mathcal{Q}_n = (\mathcal{S}_n - p_n\mathcal{L}_n)/(1-p_n)$ need not be free. Thus a free channel can contain a component that is δ-close to $\mathcal{N}_n$ in diamond norm, with a weight controlled by the testing coefficient. For polynomially decreasing δ, the loss $g_\varepsilon(n,\delta)$ is $o(n)$, so this component weight has the same exponential rate as $\beta_{\varepsilon,n}$. This connects the testing problem to an explicit CP domination of an approximation to the target by a free channel.
Proof outline. For an isometric target, testing supplies a CP branch with positive operator overlap in every input direction. Permutation averaging and weighted discarding yield a compressed operator with a local expansion. Polynomial approximations to its inverse and to the target image projector correct the branch, while the additional family conditions turn these corrections into free CP domination. A free replacer restores trace preservation and pads the discarded sites. Comparing testing against all auxiliary extensions reduces an arbitrary target to the isometric case. The full proof is given in Appendix A.

6.3 Finite-block smoothing bounds

Corollary 17 (Finite-block smoothing bounds)**.** *Assume the hypotheses of Theorem 16. For every $0 < \varepsilon < 1$, $0 < \delta \le 1/16$, and $\varepsilon + \delta/2 < 1$,*

$$\begin{aligned} a_{\varepsilon,n} + \log_2(1-\varepsilon-\delta/2) &\le Z_n^\delta \\ &\le a_{\varepsilon,n} + K_\varepsilon n^{2/3}\log_2\frac{n+1}{\delta}. \end{aligned}$$

The lower bound requires only Definition 1 and holds for every $0 \le \delta < 2(1-\varepsilon)$.

Proof. The lower bound is (20). For the upper bound, Theorem 16 supplies a channel within diamond distance δ of $\mathcal{N}_n$ whose domination cost is at most $a_{\varepsilon,n} + K_\varepsilon n^{2/3}\log_2((n+1)/\delta)$. This channel is feasible in (18). □

6.4 Comparison of testing tolerances

Lemma 18. *Assume the hypotheses of Theorem 16. For any $0 < \varepsilon_- \le \varepsilon_+ < 1$, there is a finite constant K such that*

$$\sup_{\varepsilon,\varepsilon' \in [\varepsilon_-,\varepsilon_+]} |a_{\varepsilon,n} - a_{\varepsilon',n}| \le K n^{2/3}\log_2(n+1) \qquad (n \ge 1).$$

In particular, the normalized difference tends to zero for every two fixed errors in $(0,1)$.

Proof. For fixed $\varepsilon,\varepsilon' \in (0,1)$ choose $\delta = \min\{1/16, (1-\varepsilon')/2\}$ in Theorem 16. Every tester feasible at error ε' accepts $\mathcal{L}_n$ with probability at least $1-\varepsilon'-\delta > 0$. Since $\mathcal{S}_n$ is a free alternative, its CP domination gives

$$\beta_{\varepsilon',n} \ge \beta_{\varepsilon,n}2^{-g_\varepsilon(n,\delta)}(1-\varepsilon'-\delta).$$

Taking negative logarithms yields the upper bound

$$a_{\varepsilon',n} - a_{\varepsilon,n} \le g_\varepsilon(n,\delta) - \log_2(1-\varepsilon'-\delta) = O_{\varepsilon,\varepsilon'}(n^{2/3}\log_2(n+1)).$$

Interchanging the errors gives the reverse bound. For the uniform statement it suffices to apply this result to the endpoints: $a_{\varepsilon,n}$ is nondecreasing in ε, so any difference within the interval is bounded by $a_{\varepsilon_+,n} - a_{\varepsilon_-,n}$. □

Table 1: Channel families satisfying the basic and additional conditions in Definitions 1 and 14. Each row refers to all block channels satisfying the indicated constraint. The witnessing states are those in the marginal and replacer conditions.

Channel family	Witnessing states	Verification
Preserving $\tau^{\otimes n} \mapsto \omega^{\otimes n}$	Faithful τ, ω	Proposition 19
Unital ($A = B = \mathbb{C}^d$)	$\tau = \omega = I/d$	Proposition 19
Maximally incoherent or dephasing-covariant	$\tau = I/a$, $\omega = I/b$	Proposition 20
Symmetry-covariant	Faithful invariant τ, ω	Proposition 21
Entanglement-breaking or positive-partial-transpose	Any faithful τ, ω	Proposition 22

7 Resource families and exact benchmarks

7.1 Admissible resource families

If $\dim A = 1$, channels are states and Theorem 2 recovers the generalized state Stein identity under F1, F2, and F3. For the quantitative completion bound, the additional marginal axiom becomes ordinary partial-trace closure. The examples below satisfy both the basic conditions in Definition 1 and the additional conditions F4 and F5. We verify these directly and show why the witnessing input state for the quantitative bound should be adapted to the resource family. The exact examples give finite-block testing and smoothing formulas. Table 1 summarizes the channel families.

Proposition 19 (Preservation of a faithful reference state)**.** *Fix faithful states $\tau \in \mathrm{D}(A)$ and $\omega \in \mathrm{D}(B)$. The family*

$$\mathfrak{F}_n^{\tau \to \omega} = \{\mathcal{M} \in \mathrm{CPTP}(A^n \to B^n) : \mathcal{M}(\tau^{\otimes n}) = \omega^{\otimes n}\} \tag{23}$$

is admissible and also satisfies F4 and F5, with the displayed τ and ω. For a replacer target $\mathcal{N} = \mathcal{R}_\rho$, one has the exact identities

$$E_n = nD(\rho\|\omega), \qquad \beta_{\varepsilon,n} = \beta_\varepsilon^{\mathrm{st}}(\rho^{\otimes n}\|\omega^{\otimes n}), \tag{24}$$

where $\beta_\varepsilon^{\mathrm{st}}(\varrho\|\sigma) = \min_{0 \le Q \le I,\ \mathrm{Tr}\, Q\varrho \ge 1-\varepsilon} \mathrm{Tr}\, Q\sigma$. In particular $R_{\mathfrak{F}}(\mathcal{R}_\rho) = D(\rho\|\omega)$.

Proof. The preservation constraint is affine and closed inside the compact convex set of channels. The replacer $\mathcal{R}_\omega$ belongs to $\mathfrak{F}_1^{\tau \to \omega}$. Permutation and tensor closure follow because the prescribed reference states are tensor powers. For the one-site marginal $\mathcal{M}'$,

$$\mathcal{M}'(\tau^{\otimes(n-1)}) = \mathrm{Tr}_{B_n} \mathcal{M}(\tau^{\otimes n}) = \omega^{\otimes(n-1)}.$$

This proves all axioms.

Using the input $\tau^{\otimes n}$ without a reference gives outputs $\rho^{\otimes n}$ and $\omega^{\otimes n}$ for the target and every alternative, respectively. It follows that $E_n \ge nD(\rho\|\omega)$ and $\beta_{\varepsilon,n} \le \beta_\varepsilon^{\mathrm{st}}$. For the reverse bounds choose the free alternative $\mathcal{R}_\omega^{\otimes n}$. On any reference-assisted input state, the target and alternative outputs are $\varrho_R \otimes \rho^{\otimes n}$ and $\varrho_R \otimes \omega^{\otimes n}$. Their relative entropy is $nD(\rho\|\omega)$. Every joint effect is equivalent on these two outputs to the single-system effect

$$Q_B = \mathrm{Tr}_R[(\varrho_R^{1/2} \otimes I)Q(\varrho_R^{1/2} \otimes I)], \qquad 0 \le Q_B \le I.$$

Thus no tester improves on the displayed state-testing optimum. This proves (24). □

For finite-dimensional Hamiltonians at finite inverse temperature, take τ and ω to be the corresponding Gibbs states. Equation (23) then gives the full Gibbs-preserving channel family [19].

When $A = B = \mathbb{C}^d$ and $\tau = \omega = I/d$, this is the family of all unital block channels. In particular a pure-state reset target has rate $\log_2 d$ against unital alternatives.

The fixed-state marginal condition is strictly weaker than closure under every auxiliary input state. For $A = B$ and $\tau = \omega$, the two-site SWAP channel belongs to $\mathfrak{F}_2^{\tau\to\tau}$. Inserting a state η in the second input and tracing its corresponding output yields $X \mapsto \mathrm{Tr}(X)\eta$, which preserves τ only when $\eta = \tau$. If $\tau \neq I/d$, even insertion of the maximally mixed state fails this test. Allowing a general faithful τ in the quantitative completion theorem therefore enlarges its scope beyond a maximally-mixed-insertion formulation.

Coherence constraints give a second class of admissible families. Fix product reference bases and let Δ_A, Δ_B denote complete dephasing in those bases. The following are standard coherence-operation classes [20]. At blocklength n, maximally incoherent operations (MIO) are channels mapping every diagonal input state to a diagonal output state, equivalently

$$\Delta_B^{\otimes n} \circ \mathcal{M} \circ \Delta_A^{\otimes n} = \mathcal{M} \circ \Delta_A^{\otimes n}.$$

Dephasing-covariant incoherent operations (DIO) are channels satisfying $\Delta_B^{\otimes n} \circ \mathcal{M} = \mathcal{M} \circ \Delta_A^{\otimes n}$.

Proposition 20 (Coherence-restricted channels)**.** *The full MIO and DIO block-channel families are admissible and also satisfy F4 and F5, with $\tau = I_A/a$ and $\omega = I_B/b$. Consequently all results of Theorems 2, 11 and 12 apply to these families and arbitrary target channels.*

Proof. Each defining equality is linear, so both sets are compact and convex inside the set of channels. Both contain the maximally mixed replacer. Permutations commute with product dephasing. Tensor products preserve the DIO equality. They preserve MIO because every diagonal state on two blocks is a convex combination of product basis-state projections; each such projection is sent to a diagonal state. Finally, let $\mathcal{P}(X) = X \otimes I_A/a$ and $\mathcal{T} = \mathrm{Tr}_{B_n}$. These maps intertwine dephasing at the two blocklengths:

$$\Delta_A^{\otimes n} \circ \mathcal{P} = \mathcal{P} \circ \Delta_A^{\otimes(n-1)}, \qquad \Delta_B^{\otimes(n-1)} \circ \mathcal{T} = \mathcal{T} \circ \Delta_B^{\otimes n}.$$

Combining these identities with the DIO equality for $\mathcal{M}$ shows that $\mathcal{T} \circ \mathcal{M} \circ \mathcal{P}$ remains DIO. For MIO, a diagonal input on the remaining sites together with the prepared maximally mixed state is diagonal; its output and output marginal remain diagonal. This proves the marginal axiom as well. □

Covariance provides a further common structure, compatible with intersections of resource constraints.

Proposition 21 (Covariance and intersections)**.** *Let G be a group with finite-dimensional unitary representations $U_A(g)$ and $U_B(g)$. Suppose τ and ω are faithful invariant states. The family of all channels obeying*

$$\mathcal{M} \circ \mathrm{Ad}_{U_A(g)^{\otimes n}} = \mathrm{Ad}_{U_B(g)^{\otimes n}} \circ \mathcal{M} \quad (g \in G)$$

is admissible and also satisfies F4 and F5, with τ, ω. Moreover, any intersection of families satisfying all five conditions with the same witnessing states τ, ω also satisfies all five.

Proof. The covariance constraints are closed and linear. The invariant-state replacer is covariant and supplies nonemptiness. Collective tensor representations commute with site permutations, and tensor products preserve the covariance equation. Preparation of invariant τ intertwines the reduced and full input actions; partial trace intertwines the corresponding output actions. Thus the resulting marginal is covariant.

For an intersection, compactness, convexity, permutation invariance, and tensor and marginal closure are inherited. Nonemptiness is guaranteed by the common faithful replacer and its tensor powers. All five axioms therefore hold. □

For Hamiltonians H_A, H_B, take the time-translation representations $U_A(s) = e^{-isH_A}$ and $U_B(s) = e^{-isH_B}$ for $s \in \mathbb{R}$. With Gibbs reference states, the proposition and Proposition 19 cover the intersection of time-covariant and Gibbs-preserving channels.

7.2 Exact formulas and correlated alternatives

A channel is entanglement breaking (EB) if its output is separable from every reference for every input. Equivalently its Choi operator is separable across reference and output, or it has a measure-and-prepare representation [21]. A positive-partial-transpose (PPT) channel here means a channel whose Choi operator is positive after transposition on the entire output system. Equivalently $T_B \circ \mathcal{M}$ is CP, where T_B is transposition. At blocklength n, the relevant cut is $A^n : B^n$.

Proposition 22 (Entanglement-breaking and PPT alternatives)**.** *The families of all EB block channels and all PPT block channels are admissible and also satisfy F4 and F5, for any faithful input state τ and any faithful output replacer state ω. For any isometry $J : A \to B$ and target $\mathcal{N} = \mathrm{Ad}_J$, both families have the exact finite-block values*

$$E_n = n \log_2 a, \qquad \beta_{\varepsilon,n} = (1-\varepsilon)a^{-n}, \qquad a = \dim A.$$

Proof. Normalized EB Choi operators form the intersection of the compact convex separable-state set with the channel marginal constraint. For PPT replace separability by positivity of the output partial transpose, also a closed convex constraint. Both families contain every replacer. Tensor products and simultaneous permutations preserve their respective Choi properties.

An EB channel remains EB after any input or output channel composition: use reference separability after the input composition and preservation of separability by a local output channel. This proves its marginal closure. For a PPT channel the reduced map $\mathcal{M}'$ satisfies

$$T_{B^{n-1}} \circ \mathcal{M}' = \mathrm{Tr}_{B_n} \circ T_{B^n} \circ \mathcal{M} \circ (X \mapsto X \otimes \tau).$$

The right-hand side is CP because preparation and partial trace are CP and $T_{B^n} \circ \mathcal{M}$ is CP by the PPT assumption. The displayed identity uses $\mathrm{Tr}_{B_n} \circ T_{B^n} = T_{B^{n-1}} \circ \mathrm{Tr}_{B_n}$: transposition on the discarded factor leaves its trace unchanged, while transposition on the remaining factors passes through the partial trace. This proves the PPT marginal axiom. Every reference-assisted PPT-channel output is PPT because $T_{B^n} \circ \mathcal{M}$ is CP. Every EB output is PPT as well.

Put $D = a^n$, $J_n = J^{\otimes n}$, and use the normalized maximally entangled input $|\Phi_D\rangle = D^{-1/2} \sum_{j=1}^{D} |j\rangle_R |j\rangle_{A^n}$. The target output is the pure state $\rho_J = |\phi_J\rangle\langle\phi_J|$ with $|\phi_J\rangle = (I_R \otimes J_n)|\Phi_D\rangle$. All its nonzero Schmidt coefficients equal $D^{-1/2}$. Write X^{T_B} for partial transposition on the entire B^n output in this calculation. In Schmidt bases, $\rho_J^{T_B}$ is $1/D$ times the swap operator on a D-by-D subspace and zero on its orthogonal complement. Consequently $\|\rho_J^{T_B}\|_\infty = 1/D$. For every alternative output σ from either family, σ^{T_B} is positive with trace one. Therefore

$$\langle \phi_J|\sigma|\phi_J\rangle = \mathrm{Tr}\, \rho_J^{T_B} \sigma^{T_B} \le 1/D. \tag{25}$$

The effect $(1-\varepsilon)|\phi_J\rangle\langle\phi_J|$ proves $\beta_{\varepsilon,n} \le (1-\varepsilon)/D$. Binary measurement onto ρ_J also gives $D(\rho_J\|\sigma) \ge \log_2 D$, so $E_n \ge \log_2 D$.

For both reverse bounds, let Δ be complete dephasing in any basis of A^n and take the EB alternative $\mathcal{M}_* = \mathrm{Ad}_{J_n} \circ \Delta$. Its Kraus operators $K_j = J_n|j\rangle\langle j|$ satisfy $\sum_j K_j = J_n$. Cauchy–Schwarz on their Choi vectors gives

$$\sum_j |K_j\rangle\langle K_j| \ge D^{-1}|J_n\rangle\langle J_n|, \qquad \mathcal{M}_* \ge_{\mathrm{CP}} D^{-1}\,\mathrm{Ad}_{J_n}. \tag{26}$$

Thus every feasible tester has alternative acceptance at least $(1-\varepsilon)/D$, and operator monotonicity of logarithm yields $D_{\mathrm{ch}}(\mathrm{Ad}_{J_n} \|\mathcal{M}_*) \le \log_2 D$. Since EB channels are PPT, the same alternative proves both upper bounds on E_n and both lower bounds on $\beta_{\varepsilon,n}$. □

Corollary 23 (Exact smoothed isometric benchmark)**.** *For either family in Proposition 22, an isometric target, and every $0 \le \delta \le 2$,*

$$Z_n^\delta = \log_2 \max\{1, a^n(1-\delta/2)\}. \tag{27}$$

The factor $1/2$ here corresponds to our use of the full diamond norm in the smoothing constraint.

Proof. Put $D = a^n$. If $D = 1$, every channel with this input is a replacer and the assertion is immediate. Assume $D > 1$. Let $\mathcal{L}$ be a feasible smoothing channel and let σ be its output on the maximally entangled input state used above. Since σ and ρ_J are states,

$$\langle \phi_J | \sigma | \phi_J \rangle \geq 1 - \tfrac{1}{2} \|\sigma - \rho_J\|_1 \geq 1 - \delta/2.$$

If $\mathcal{L} \leq_{\mathrm{CP}} \lambda \mathcal{S}$ with $\mathcal{S}$ free, (25) gives $1 - \delta/2 \leq \lambda/D$. Together with $\lambda \geq 1$, this proves the lower bound in (27).

For the upper bound use the same $\mathcal{M}_* = \mathrm{Ad}_{J_n} \circ \Delta$ as in Proposition 22. In a basis indexed by $j = 0, \ldots, D-1$, let $U|j\rangle = e^{2\pi i j/D}|j\rangle$. Then $\Delta = D^{-1} \sum_{l=0}^{D-1} \mathrm{Ad}_{U^l}$ and

$$\|\mathcal{M}_* - \mathrm{Ad}_{J_n}\|_\diamond \leq \|\Delta - \mathrm{id}\|_\diamond \leq 2(1 - 1/D).$$

Indeed, the $l = 0$ term is $D^{-1}\mathrm{id}$, and the other terms form $(1 - 1/D)$ times a channel. Set

$$s = \min\{1, \delta/[2(1 - 1/D)]\}, \qquad \mathcal{L}_s = (1 - s)\,\mathrm{Ad}_{J_n} + s\mathcal{M}_*.$$

This is CPTP and within diamond distance δ of the target. Equation (26) gives

$$\begin{gathered} \mathcal{L}_s \leq_{\mathrm{CP}} [(1 - s)D + s]\mathcal{M}_*, \\ (1 - s)D + s = \max\{1, D(1 - \delta/2)\}. \end{gathered}$$

The same EB dominator is available in both families, completing the proof. □

The preceding exact formulas describe exponential discrimination. The following construction exhibits how persistent correlations can instead produce a zero rate for a target outside the one-use free set.

Example 24 (Correlated alternatives with zero Stein rate)**.** Work in the state specialization $A = \mathbb{C}$. Choose a faithful state ω, a unit vector $|\phi\rangle$ with $\rho = |\phi\rangle\langle\phi| \neq \omega$, and $0 < c < 1$. On a block of size k put $\zeta_k = c\rho^{\otimes k} + (1 - c)\omega^{\otimes k}$. Let $\mathfrak{F}_n$ be the convex hull of the states obtained by partitioning the n labeled sites into blocks and assigning to each block of size k either ζ_k or $\omega^{\otimes k}$. For fixed n this is a finite convex hull and hence compact. All partitions ensure permutation invariance; joining partitions ensures tensor closure. Tracing a site from a block replaces ζ_k by ζ_{k-1} or removes a singleton, and preserves the analogous product state. Thus marginal closure holds. The faithful product state is included, so the family satisfies both the basic and the additional conditions. Writing conv for the convex hull, the one-use set is $\mathfrak{F}_1 = \mathrm{conv}\{\omega, c\rho + (1 - c)\omega\}$, so $\rho \notin \mathfrak{F}_1$. At blocklength n, $\zeta_n \geq c\rho^{\otimes n}$ gives, identifying ρ with its preparation channel,

$$0 \leq E_n \leq \log_2(1/c), \qquad \beta_{\varepsilon,n} \geq c(1 - \varepsilon), \qquad R_{\mathfrak{F}}(\rho) = 0.$$

The correlated alternative ζ_n contains a target component of fixed weight c at every blocklength. This keeps the worst-case type-II error bounded away from zero and explains the vanishing rate.

8 Discussion

For parallel discrimination against correlated alternative channels, the regularized Umegaki channel relative entropy of resource is the fixed-error Stein exponent, the CPTP-smoothed max-relative entropy rate, and the threshold above which target acceptance decays exponentially. Compactness and convexity, tensor closure, and a faithful free replacer suffice for all these asymptotic statements.

Uniform auxiliary map approximation provides the operator control needed to prove this correspondence. Iterative reduction of the domination rate and amplification of a fixed accurate block produce exponential diamond accuracy. These approximations establish the ordinary testing limits directly, and an entropy comparison identifies their common value. The same construction gives equipartition whenever $0 < \delta_n \le \bar{\delta} < 2$ and $\log_2(1/\delta_n) = o(n)$, as well as an exponential strong converse.

The quantitative completion theorem controls a different aspect of the problem: its bound applies at a specified blocklength and accuracy with a constant uniform over the stated class of targets and families. Permutation invariance and fixed-input marginal closure permit weighted discarding and local corrections, yielding an explicit overhead $O_\varepsilon(n^{2/3} \log_2((n+1)/\delta))$ relative to the testing coefficient. The faithful input state can be chosen to respect a preferred reference state of the resource family. This construction also gives an alternative proof of the Stein identity through near subadditivity.

The applications include state preservation, coherence restrictions, covariance, entanglement breaking, and positive partial transpose. The isometric benchmarks resolve testing and smoothing exactly at every blocklength. The zero-rate construction shows how persistent correlations can suppress exponential discrimination even when the target is outside the one-use free set.

The access model is parallel, with arbitrary correlated block alternatives. Extensions to sequential access require a corresponding causal model for those alternatives. Further questions include the optimal strong-converse exponent, equipartition at fixed positive exponential smoothing rates, and resource-conversion rates under specified classes of superchannels.

A Proof of quantitative completion

We prove Theorem 16 under Definition 1 and the additional conditions F4 and F5. We first extract a branch with uniform positive overlap and compare testing against auxiliary extensions. Local operator estimates then give quantitative completion for isometric targets, followed by the reduction for general targets.

A.1 Branch extraction and auxiliary extensions

A completely positive branch of a channel $\mathcal{M}$ is a CP map $\mathcal{B}$ for which $\mathcal{M} - \mathcal{B}$ is CP. Positivity of both maps and trace preservation of $\mathcal{M}$ imply $0 \le \mathcal{B}^\dagger(I) \le \mathcal{M}^\dagger(I) = I$ and $0 \le (\mathcal{M} - \mathcal{B})^\dagger(I) \le I$. Thus both maps are trace nonincreasing; together they can be viewed as the two outcomes of an instrument. For $\mathcal{B} = p \operatorname{Ad}_V$, the probability of the branch on an input state ρ is $p \operatorname{Tr} V\rho V^\dagger$, which need not equal p. In particular, the branch coefficient and its input-dependent probability must be distinguished.

We first extract a branch with positive overlap from a testing lower bound. We then compare discrimination of a channel with discrimination of a Stinespring isometry against all extensions of the original alternatives to an auxiliary output system. An isometric target produces a pure output on a purified input, making the branch argument available. Access to its auxiliary output can improve discrimination, however, so an explicit comparison is needed before using that argument for a noisy target.

A.1.1 A branch with uniform positive overlap

Lemma 25. *Let $J : H \to K$ be an isometry and let $\mathcal{M} : \mathrm{L}(H) \to \mathrm{L}(K)$ be CPTP. Fix $0 < \delta < 1$, and suppose $\beta = b_\delta(\operatorname{Ad}_J, \mathcal{M}) > 0$. Set $r = \sqrt{1-\delta}$ and $c = (1-r)/(1+r)$. Then there is $V : H \to K$ such that*

$$\|V\|_\infty \le 1, \qquad \operatorname{Re}(J^\dagger V) \ge cI, \qquad \mathcal{M} \ge_{\mathrm{CP}} \beta \operatorname{Ad}_V . \tag{28}$$

Proof. We construct V by finding a Kraus combination A close in operator norm to a positive multiple of J, and then rescaling A. The Kraus representation will give CP domination, while closeness to J will give the norm and overlap bounds in (28). To obtain this approximation, we show that its failure would produce a feasible test with alternative acceptance strictly below β.

Choose Kraus operators K_i for $\mathcal{M}$ and define

$$\mathcal{E} = \left\{ \sum_i z_i K_i : \sum_i |z_i|^2 \leq 1 \right\}.$$

This set is compact and convex because it is the linear image of the Euclidean unit ball of Kraus coefficients. We first check that every $A \in \mathcal{E}$ gives a CP map dominated by $\mathcal{M}$. Fix $A = \sum_i z_i K_i \in \mathcal{E}$, so that $\sum_i |z_i|^2 \leq 1$. By linearity of vectorization, $|A\rangle = \sum_i z_i |K_i\rangle$. The Kraus representation of $\mathcal{M}$ and the definition of Ad_A give

$$C_{\mathcal{M}} = \sum_i |K_i\rangle\langle K_i|, \qquad C_{\mathrm{Ad}_A} = |A\rangle\langle A|.$$

For any vector $|v\rangle \in H \otimes K$, Cauchy–Schwarz therefore yields

$$\begin{aligned} |\langle v \mid A\rangle|^2 &= \left| \sum_i z_i \langle v \mid K_i\rangle \right|^2 \\ &\leq \Big(\sum_i |z_i|^2 \Big) \sum_i |\langle v \mid K_i\rangle|^2 \\ &\leq \sum_i |\langle v \mid K_i\rangle|^2 = \langle v | C_{\mathcal{M}} | v\rangle. \end{aligned}$$

Thus $\langle v|(C_{\mathcal{M}} - C_{\mathrm{Ad}_A})|v\rangle \geq 0$ for every $|v\rangle$, which means $C_{\mathcal{M}} - C_{\mathrm{Ad}_A} \geq 0$. By the Choi criterion, $\mathcal{M} - \mathrm{Ad}_A$ is completely positive, or equivalently $\mathcal{M} \geq_{\mathrm{CP}} \mathrm{Ad}_A$.

We now seek $A \in \mathcal{E}$ such that

$$\|uJ - A\|_\infty \leq ru, \qquad u = \frac{\sqrt{\beta}}{1-r} > 0.$$

This is the approximation needed for the final rescaling. To prove that it exists, we express the distance from uJ to $\mathcal{E}$ as an optimization over operators X that can be converted into tests. For $X \in \mathrm{L}(H, K)$, the support function of $\mathcal{E}$ with respect to the real trace pairing is defined by

$$h_{\mathcal{E}}(X) = \max_{A \in \mathcal{E}} \operatorname{Re} \operatorname{Tr} X^\dagger A.$$

It is the largest value of the real linear functional $A \mapsto \operatorname{Re} \operatorname{Tr} X^\dagger A$ on $\mathcal{E}$. Writing $c_i = \operatorname{Tr} X^\dagger K_i$, we obtain

$$h_{\mathcal{E}}(X) = \max_{\sum_i |z_i|^2 \leq 1} \operatorname{Re} \sum_i z_i c_i = \left(\sum_i |c_i|^2 \right)^{1/2}.$$

Indeed, Cauchy–Schwarz gives the upper bound, and for $c \neq 0$ it is attained by $z_i = \overline{c_i}/\|c\|_2$; when $c = 0$, both sides vanish.

To express the operator-norm distance from uJ to $\mathcal{E}$, use duality between the operator and trace norms:

$$\|Y\|_\infty = \max_{\|X\|_1 \leq 1} \operatorname{Re} \operatorname{Tr} X^\dagger Y.$$

Applying this identity to $Y = uJ - A$ gives

$$\min_{A \in \mathcal{E}} \|uJ - A\|_\infty = \min_{A \in \mathcal{E}} \max_{\|X\|_1 \leq 1} \operatorname{Re} \operatorname{Tr} X^\dagger (uJ - A).$$

Both feasible sets are compact and convex. For each fixed u, the objective is continuous and affine in each variable, so minimax exchanges the extrema. The inner minimum over A then equals $u \operatorname{Re} \operatorname{Tr} X^\dagger J - h_{\mathcal{E}}(X)$. Substituting the support function computed above yields

$$\begin{aligned} &\min_{A \in \mathcal{E}} \|uJ - A\|_\infty \\ &= \max_{\|X\|_1 \le 1} \left\{ u \operatorname{Re} \operatorname{Tr} X^\dagger J - \left(\sum_i |\operatorname{Tr} X^\dagger K_i|^2 \right)^{1/2} \right\}. \end{aligned} \tag{29}$$

Suppose, for a contradiction, that $\min_{A \in \mathcal{E}} \|uJ - A\|_\infty > ru$. By (29), there is an operator X with $\|X\|_1 \le 1$ such that

$$u \operatorname{Re} \operatorname{Tr} X^\dagger J - \left(\sum_i |\operatorname{Tr} X^\dagger K_i|^2 \right)^{1/2} > ru.$$

Denote the left-hand side by $F(X)$, so that $F(X) > ru > 0$. In particular, $X \neq 0$, and hence $0 < \|X\|_1 \le 1$. Both terms in $F(X)$ scale by s when X is replaced by sX, for $s > 0$. Therefore

$$F\left(\frac{X}{\|X\|_1} \right) = \frac{F(X)}{\|X\|_1} \ge F(X) > ru,$$

where the middle inequality uses $F(X) > 0$ and $\|X\|_1 \le 1$. The normalized operator thus still satisfies the strict inequality, and we may replace X by it and assume $\|X\|_1 = 1$. Define

$$\alpha = \operatorname{Re} \operatorname{Tr} X^\dagger J, \qquad q_X = \sum_i |\operatorname{Tr} X^\dagger K_i|^2.$$

Rearranging the strict inequality gives $\sqrt{q_X} < u(\alpha - r)$. Since its left-hand side is nonnegative and $u > 0$, we obtain $\alpha > r$. On the other hand, trace-norm duality and the isometry property give

$$\alpha \le |\operatorname{Tr} X^\dagger J| \le \|X\|_1 \|J\|_\infty = 1.$$

Consequently,

$$\sqrt{q_X} < u(\alpha - r) \le u(1 - r) = \sqrt{\beta},$$

so $q_X < \beta$. We now use X to construct a test whose target acceptance exceeds $r^2 = 1 - \delta$ and whose alternative acceptance is q_X. This will contradict the definition of β. Write a singular-value decomposition $X = \sum_j x_j |u_j\rangle\langle v_j|$ with $\sum_j x_j = 1$. The normalized input and measurement vectors

$$|\psi\rangle = \sum_j \sqrt{x_j} |j\rangle_R |v_j\rangle_H, \qquad |\phi\rangle = \sum_j \sqrt{x_j} |j\rangle_R |u_j\rangle_K$$

define a feasible test. Indeed, its effect $Q = |\phi\rangle\langle\phi|$ has target acceptance

$$\operatorname{Tr} Q(\mathrm{id}_R \otimes \mathrm{Ad}_J)(|\psi\rangle\langle\psi|) = |\operatorname{Tr} X^\dagger J|^2 \ge \alpha^2 > 1 - \delta.$$

For the alternative channel, the Kraus representation gives

$$\begin{aligned} &\operatorname{Tr} Q(\mathrm{id}_R \otimes \mathcal{M})(|\psi\rangle\langle\psi|) \\ &\quad = \sum_i |\langle\phi|(I_R \otimes K_i)|\psi\rangle|^2 \\ &\quad = \sum_i |\operatorname{Tr} X^\dagger K_i|^2 = q_X < \beta. \end{aligned}$$

This contradicts $\beta = b_\delta(\mathrm{Ad}_J, \mathcal{M})$, the minimum alternative acceptance among tests with target acceptance at least $1 - \delta$. Hence the distance is at most ru. Compactness of $\mathcal{E}$ ensures that some $A \in \mathcal{E}$ attains it.

It remains to obtain the three bounds in (28). Write $A = uJ + D$, where $\|D\|_\infty \le ru$, and set $V = A/[u(1+r)]$. The triangle inequality gives

$$\|V\|_\infty \le \frac{u\|J\|_\infty + \|D\|_\infty}{u(1+r)} \le 1.$$

Since $J^\dagger J = I$ and $\mathrm{Re}(J^\dagger D) \ge -\|J^\dagger D\|_\infty I \ge -ruI$, we also have

$$\mathrm{Re}(J^\dagger V) = \frac{uI + \mathrm{Re}(J^\dagger D)}{u(1+r)} \ge \frac{1-r}{1+r} I = cI.$$

Finally, the CP domination established for every member of $\mathcal{E}$ yields

$$\mathcal{M} \ge_{\mathrm{CP}} \mathrm{Ad}_A = u^2(1+r)^2 \,\mathrm{Ad}_V = \frac{\beta}{c^2}\,\mathrm{Ad}_V \ge_{\mathrm{CP}} \beta\,\mathrm{Ad}_V\,.$$

The last inequality uses $0 < c < 1$ and complete positivity of Ad_V, completing the proof. □

A.1.2 Comparison with auxiliary extensions

Lemma 26. *Let $\mathcal{N} : \mathrm{L}(A) \to \mathrm{L}(B)$ be CPTP, let $\mathcal{F} \subseteq \mathrm{CPTP}(A \to B)$ be nonempty, compact, and convex, and fix a Stinespring isometry $V : A \to B \otimes E$. Define*

$$\mathcal{G} = \{\mathcal{L} \in \mathrm{CPTP}(A \to B \otimes E) : \mathrm{Tr}_E \circ \mathcal{L} \in \mathcal{F}\}.$$

Write β_ε for worst-case testing of $\mathcal{N}$ against $\mathcal{F}$ and γ_δ for testing of Ad_V against $\mathcal{G}$. For $0 < \varepsilon < 1$, set

$$s = \sqrt{1-\varepsilon}, \quad r = (1+s)/2, \quad \delta = 1 - r^2, \quad t = (1-s)^2/4.$$

Then $0 < \delta < 1$, $t > 0$, and

$$t\beta_\varepsilon \le \gamma_\delta, \qquad \gamma_\varepsilon \le \beta_\varepsilon. \tag{30}$$

The parameters are independent of the dimensions.

Proof. The set $\mathcal{G}$ is nonempty: append a fixed state on E to any channel in $\mathcal{F}$. It is compact and convex, being the inverse image of $\mathcal{F}$ under the continuous linear partial-trace map, intersected with the compact convex set of channels. The second inequality follows by lifting any original effect Q to $Q \otimes I_E$, preserving all relevant acceptance probabilities.

For the first inequality, we must rule out an excessive advantage from access to E. A lifted effect induces, through a dual optimization and spectral thresholding, a feasible effect on the original output. The constants in this construction depend only on the acceptance tolerance, not on $\dim E$.

Use Lemma 5 to choose a hardest $\mathcal{M} \in \mathcal{F}$ with $b_\varepsilon(\mathcal{N}, \mathcal{M}) = \beta_\varepsilon$. Fix any feasible lifted tester. Purify its input and extend its effect by the identity on the additional reference, so that its input state is represented by a unit vector $|\psi\rangle$. Set

$$|\phi\rangle = (I_R \otimes V)|\psi\rangle, \qquad \rho = \mathrm{Tr}_E\, |\phi\rangle\langle\phi|,$$
$$\sigma = (\mathrm{id} \otimes \mathcal{M})(|\psi\rangle\langle\psi|).$$

For this fixed input state, the outputs of channels extending $\mathcal{M}$ are exactly the states τ_{RBE} with $\mathrm{Tr}_E\, \tau = \sigma$. To verify the nontrivial direction, apply a Stinespring dilation of $\mathcal{M}$ to $|\psi\rangle$, obtaining a purification of σ with auxiliary system Z. Purify the desired τ further. The Schmidt decompositions provide an isometry from the support of the first purification's Z marginal into the second purifying space. Tracing out the extra register gives a channel on that support from Z to E; extend it to the whole Z by sending the orthogonal complement to a fixed state. Acting with this channel on the dilation of $\mathcal{M}$ gives an extension of $\mathcal{M}$ on *all* inputs and produces the required τ on $|\psi\rangle$.

Let Q be the lifted effect. Its largest acceptance over these extensions is the following semidefinite program (SDP), with dual variable $H = H^\dagger$ acting on $R \otimes B$:

$$\begin{aligned} c_Q &= \max_{\tau \geq 0,\ \operatorname{Tr}_E \tau = \sigma} \operatorname{Tr} Q\tau \\ &= \inf_{\substack{H = H^\dagger \\ H \otimes I_E \geq Q}} \operatorname{Tr} H\sigma. \end{aligned} \tag{31}$$

The primal is nonempty and compact, and $H = 2I$ is strictly dual feasible; strong duality applies. We use a minimizing sequence of dual feasible operators, which also covers singular σ. Every feasible H is positive, since $H \otimes I_E \geq Q \geq 0$.

Fix such H and let $P = \mathbf{1}_{[t,\infty)}(H)$ be its spectral projection onto eigenvalues at least t. Put $p = \operatorname{Tr} P\rho$. Split $|\phi\rangle$ into its P and $I - P$ components. Using $Q \leq I$ and $Q \leq H \otimes I_E$ gives

$$\begin{aligned} \sqrt{\langle \phi | Q | \phi \rangle} &\leq \| Q^{1/2} (P \otimes I_E) |\phi\rangle \| \\ &\quad + \| Q^{1/2} ((I - P) \otimes I_E) |\phi\rangle \| \\ &\leq \sqrt{p} + \sqrt{t}. \end{aligned}$$

The first component has norm at most $\sqrt{p}$ because $Q \leq I$. On the second component, H has eigenvalues below t, so $Q \leq H \otimes I_E$ bounds its squared norm by $t(1-p) \leq t$. The target acceptance is at least $1 - \delta = r^2$. Since $r - \sqrt{t} = s$, we obtain $p \geq s^2 = 1 - \varepsilon$. Thus P, with the same original input state, is a feasible test of $\mathcal{N}$ against $\mathcal{M}$. Its alternative acceptance is bounded by $\operatorname{Tr} P\sigma \leq \operatorname{Tr} H\sigma / t$, whence

$$\beta_\varepsilon = b_\varepsilon(\mathcal{N}, \mathcal{M}) \leq \operatorname{Tr} H\sigma / t.$$

Taking the infimum in (31) yields $c_Q \geq t\beta_\varepsilon$. The lifted tester's worst-case alternative acceptance is at least c_Q. Taking the infimum over lifted testers proves the first inequality in (30). Here the auxiliary projector estimates the single-alternative testing value at the fixed hardest $\mathcal{M}$. The original lifted tester remains common to all extensions. $\square$

A.2 Local approximation and free domination

Three lemmas control the local corrections used in completion. After discarding a sublinear number of sites, the first lemma approximates a permutation-invariant operator by terms supported on few sites. The second converts an expansion of this kind into a free CP dominator. The third constructs a compatible approximation to the projector onto an isometry's product image.

An operator on a tensor product is *supported on* a set of sites S if it equals an operator on S tensored with the identity elsewhere. A finite local-support expansion is an expression $O = \sum_i c_i F_i$ with $\|F_i\|_\infty \leq 1$. Its coefficient cost is $\sum_i |c_i|$. We track both the maximum support size and this cost. In this section $\operatorname{supp} F_i$ denotes a chosen set of sites supporting F_i, and $|\operatorname{supp} F_i|$ is its cardinality. The expansions allow arbitrary complex coefficients and supported contractions. In particular, the F_i need not be Pauli operators or elements of any fixed operator basis. Their site support is distinct from the range support of a positive operator. For products, support sizes add as an upper bound and coefficient costs multiply: if $O_1 = \sum_i c_i F_i$ and $O_2 = \sum_j d_j G_j$, then $F_i G_j$ is a contraction on the union of the two supports and the resulting cost is at most $(\sum_i |c_i|)(\sum_j |d_j|)$.

A.2.1 Weighted discarding

Permutation symmetry controls the algebra in which an operator lives, but does not itself make that operator local. Weighted discarding provides the additional regularization. The following lemma quantifies both the approximation error and the cost of a local expansion; both estimates are needed for free domination.

Lemma 27. *Fix a faithful $\tau \in \mathrm{D}(\mathbb{C}^d)$ and put $t = \lambda_{\min}(\tau) > 0$. There are constants n_0 and $C_{d,t}$, depending only on d and t, with the following property. For $n \geq n_0$ and $0 < \xi \leq 1/16$, set*

$$\ell_n = \ln \frac{n+1}{\xi}, \qquad m = \lceil n^{2/3} \ell_n \rceil, \qquad k = n - m,$$

and assume $m \leq n/2$. For every permutation-invariant contraction W on $(\mathbb{C}^d)^{\otimes n}$, define

$$X = \mathbb{E}_{\tau,m}(W) := \mathrm{Tr}_m[(I \otimes (\tau^{1/2})^{\otimes m}) W (I \otimes (\tau^{1/2})^{\otimes m})].$$

Here Tr_m traces out the last m sites, leaving the first k. There is a finite expansion $P = \sum_i c_i F_i$ on k sites such that

$$\|P - X\|_\infty \leq \xi, \tag{32}$$

$$|\operatorname{supp} F_i| \leq C_{d,t} n^{2/3},$$

$$\sum_i |c_i| \leq \exp(C_{d,t} n^{2/3}), \qquad \|F_i\|_\infty \leq 1. \tag{33}$$

The constants are uniform in ξ and W. The map $\mathbb{E}_{\tau,m}$ is unital and completely positive.

Proof. Step 1: A controlled integral representation. The last assertion follows directly from the displayed sandwich formula and $\mathrm{Tr}\, \tau^{\otimes m} = 1$. For $d = 1$ take $P = X = W$, supported on the empty set. For $d > 1$, let $U(d)$ denote the unitary group on $\mathbb{C}^d$. Index partitions of n by $\lambda \vdash n$, and write $\ell(\lambda)$ for the number of their nonzero parts. Schur–Weyl duality [22] gives

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash n,\ \ell(\lambda) \leq d} Q_\lambda \otimes P_\lambda, \qquad W = \bigoplus_\lambda W_\lambda \otimes I_{P_\lambda}.$$

Here Q_λ carries an irreducible representation of $U(d)$, and P_λ carries one of the permutation group. The two commuting actions, denoted by $D_\lambda(g)$ on Q_λ and $\Pi_\lambda(\pi)$ on P_λ, have the forms

$$g^{\otimes n} = \bigoplus_\lambda D_\lambda(g) \otimes I_{P_\lambda}, \qquad U_\pi = \bigoplus_\lambda I_{Q_\lambda} \otimes \Pi_\lambda(\pi).$$

Commutation of W with every U_π gives the displayed block form of W. This is the precise commutant consequence of Schur–Weyl duality used here. Put $r_\lambda = \dim Q_\lambda$. Schur orthogonality, with normalized Haar measure, yields

$$W = \int_{U(d)} f_W(g) g^{\otimes n}\, dg, \tag{34}$$

$$f_W(g) = \sum_\lambda r_\lambda \,\mathrm{Tr}[W_\lambda D_\lambda(g)^\dagger].$$

Indeed, the integral of $\mathrm{Tr}[Z D_\lambda(g)^\dagger] D_\mu(g)$ is zero for $\lambda \neq \mu$ and equals Z/r_λ otherwise, by entrywise orthogonality. Since $\|W_\lambda\|_\infty \leq 1$,

$$|f_W(g)| \leq v_n := \sum_\lambda r_\lambda^2 \leq \binom{n + d^2 - 1}{d^2 - 1} \leq (n + d^2 + 1)^{d^2 - 1}.$$

The sum is the dimension of the permutation commutant. To bound it, expand operators in tensor products of the d^2 matrix units. Invariance under site permutations makes each coefficient constant on an orbit. Each orbit is determined by d^2 multiplicities summing to n, so there are $\binom{n+d^2-1}{d^2-1}$ such orbits.

Step 2: Damping under weighted discarding. The function f_W may be complex valued; this representation is a linear integral, not a probability mixture. Its useful feature is the polynomial bound on its absolute coefficient mass. On a unitary tensor power, weighted discarding acts as

$$\mathbb{E}_{\tau,m}(g^{\otimes n}) = [\mathrm{Tr}(\tau g)]^m g^{\otimes k}.$$

Each discarded site contributes one scalar factor. Applying $\mathbb{E}_{\tau,m}$ to (34) therefore gives

$$X = \int f_W(g)\,[\mathrm{Tr}(\tau g)]^m g^{\otimes k}\,dg. \tag{35}$$

This formula explains the regularization: if $|\mathrm{Tr}(\tau g)|$ is appreciably below one, its mth power damps the contribution. If it is close to one, faithfulness of τ forces g close to a scalar unitary, as quantified below. Choose

$$\begin{aligned} h &= (d^2+2)\ln(n+d^2+1)+\ln(4/\xi), \\ u &= \sqrt{\frac{2h}{tm}}, \qquad R = \lceil e^2 ku + 2h \rceil. \end{aligned}$$

Uniformly in the permitted ξ, $h = O_d(\ell_n)$ and $u = O_{d,t}(n^{-1/3})$. The active-regime assumption $m \le n/2$ also gives $\ell_n \le n^{1/3}/2$, hence $R = O_{d,t}(n^{2/3})$. Increasing n_0 ensures $u < 1$ and $R < k$.

Step 3: Concentration near scalar unitaries. Divide the integral into two regions, according to this damping. Call g good when $|\mathrm{Tr}(\tau g)| \ge 1 - tu^2/2$, and put $z = \mathrm{Tr}(\tau g)/|\mathrm{Tr}(\tau g)|$ on that set. Positivity of $\tau - tI$ implies

$$\begin{aligned} t\|g - zI\|_\infty^2 &\le t\,\mathrm{Tr}[(g-zI)^\dagger(g-zI)] \\ &\le \mathrm{Tr}[\tau(g-zI)^\dagger(g-zI)] \\ &= 2(1-|\mathrm{Tr}(\tau g)|) \le tu^2. \end{aligned}$$

For bad g, $|\mathrm{Tr}(\tau g)|^m \le e^{-h}$. Its contribution to (35) therefore has norm at most $v_n e^{-h}$. The estimate applies to arbitrary faithful τ and unitary g, including the noncommuting case.

Step 4: Truncation and a finite local expansion. On the good set write $g = z(I+H)$, where $\|H\|_\infty \le u$, and expand $g^{\otimes k} = z^k \sum_{S\subseteq[k]} H_S$. Here H_S acts as H on S and as the identity elsewhere. Truncate to $|S| \le R$ and integrate to define P. The omitted tail obeys

$$\sum_{r>R} \binom{k}{r} u^r \le \sum_{r>R} \left(\frac{eku}{r}\right)^r \le 2e^{-R},$$

because $R \ge e^2 ku$. As $R \ge 2h$ and $v_n e^{-h} \le \xi/4$, it follows that

$$\|P - X\|_\infty \le v_n e^{-h} + 2v_n e^{-R} \le \xi/4 + 2v_n e^{-2h} \le \xi.$$

Each $H_S/u^{|S|}$ is a supported contraction, and the integral expansion has absolute coefficient mass at most

$$v_n \sum_{r\le R} \binom{k}{r} u^r \le v_n e^{ku} \le \exp[O_{d,t}(n^{2/3})].$$

Absorb the complex phases into the contractions, normalize by this upper bound, and add a zero contraction if necessary. The resulting barycenter lies in the convex hull of the compact union of contraction balls supported on at most R sites. Carathéodory's theorem in a finite-dimensional real vector space gives a finite convex representation without increasing its cost. This proves (32)–(33). □

A.2.2 Free multipliers

The next lemma turns a local-support expansion into CP domination by a member of the free family. The correction operator itself need not represent a free operation. Instead, replacing inputs and outputs on its support gives a free channel that dominates the corrected branch at a cost exponential only in that support. Convexity then combines the separate corrections.

Lemma 28. *Let an admissible family also satisfy F4 and F5, with local input A and output D, $a = \dim A$, $d = \dim D$, faithful input state τ, and faithful replacer state ω. Put $t = \lambda_{\min}(\tau)$ and $w = \lambda_{\min}(\omega)$. Suppose $\mathcal{M} \in \mathfrak{F}_n$ and $\mathcal{M} \geq_{\mathrm{CP}} p \operatorname{Ad}_K$, where $p > 0$ and $K : A^n \to D^n$. Let $O = \sum_i c_i F_i$ be a finite expansion into output contractions of support sizes r_i. Define*

$$h_r = (ad)^r [(at)(dw)]^{-r/2}, \qquad H = \sum_i |c_i| h_{r_i}. \tag{36}$$

If $H > 0$, there is $\mathcal{M}' \in \mathfrak{F}_n$ such that

$$\mathcal{M}' \geq_{\mathrm{CP}} \frac{p}{H^2} \operatorname{Ad}_{OK} . \tag{37}$$

The qualification $H > 0$ excludes the zero expansion. Since every h_r is strictly positive, $H = 0$ means that all coefficients vanish and $O = 0$; in that case domination is vacuous and no division by H is needed.

Proof. For a subset S of r sites, first replace its input by $\tau^{\otimes r}$, apply $\mathcal{M}$, discard its output, and insert $\omega^{\otimes r}$. The resulting channel $\mathcal{M}_S$ is free by F5, F2, and F4. For the empty subset use $\mathcal{M}$; for the full subset use the free replacer. As a map on n sites, this construction is

$$\mathcal{M}_S = \mathcal{R}_{\omega,S} \circ \mathcal{M} \circ \mathcal{R}_{\tau,S},$$

with identities on the remaining sites understood.

On one site, $\mathcal{R}_\tau \geq_{\mathrm{CP}} (at)\mathcal{R}_{I/a}$ and $\mathcal{R}_\omega \geq_{\mathrm{CP}} (dw)\mathcal{R}_{I/d}$. Using these CP inequalities and the unitary Weyl expansions of the completely depolarizing input and output channels on S gives the inequality below. Here an s-dimensional unitary Weyl basis consists of s^2 unitaries W_j, including the identity, with $\operatorname{Tr} W_i^\dagger W_j = s\delta_{ij}$ and $s^{-2} \sum_j W_j Y W_j^\dagger = \operatorname{Tr}(Y) I/s$; δ_{ij} is one if $i = j$ and zero otherwise. We use their tensor products on S. Each replaced input site contributes a factor at/a^2, and each output site contributes dw/d^2. Thus

$$\mathcal{M}_S \geq_{\mathrm{CP}} p[(at)(dw)]^r d^{-2r} a^{-2r} \sum_{P,Q} \operatorname{Ad}_{PKQ} . \tag{38}$$

Here P and Q range over output and input Weyl bases on S, respectively, and are extended by the identity. For an output contraction $F = F_S \otimes I$ supported on S, write $F = \sum_P f_P P$. Weyl orthogonality implies $\sum_P |f_P|^2 = d^{-r} \operatorname{Tr}(F_S^\dagger F_S) \leq 1$. Keeping $Q = I$ in (38) is allowed because all discarded Choi summands are positive. For any Choi-space vector $|v\rangle$,

$$\begin{aligned} |\langle v \mid FK\rangle|^2 &= \left| \sum_P f_P \langle v \mid PK\rangle \right|^2 \\ &\leq \Big(\sum_P |f_P|^2 \Big) \sum_P |\langle v \mid PK\rangle|^2. \end{aligned}$$

Thus Cauchy–Schwarz on vectorized operators yields

$$\sum_P |PK\rangle\langle PK| \geq |FK\rangle\langle FK|.$$

The Choi criterion for CP order therefore gives $\mathcal{M}_S \geq_{\mathrm{CP}} p h_r^{-2} \operatorname{Ad}_{FK}$. The Weyl expansion is an algebraic decomposition of the depolarizing map. Freeness of $\mathcal{M}_S$ follows from marginal closure, tensor closure, and permutation invariance.

Omit zero c_i and let $u_i = |c_i| h_{r_i}/H$. Choose the corresponding free $\mathcal{M}_i$ above. For arbitrary vectors $|v_i\rangle$, weighted Cauchy–Schwarz gives

$$\left| \sum_i c_i v_i \right\rangle \left\langle \sum_i c_i v_i \right| \leq H^2 \sum_i u_i h_{r_i}^{-2} |v_i\rangle\langle v_i|,$$

Indeed, with $a_i = u_i h_{r_i}^{-2} > 0$ and any test vector $|v\rangle$,

$$\left|\sum_i c_i \langle v \mid v_i\rangle\right|^2 \le \left(\sum_i \frac{|c_i|^2}{a_i}\right) \sum_i a_i |\langle v \mid v_i\rangle|^2,$$
$$\sum_i \frac{|c_i|^2}{a_i} = H \sum_i |c_i| h_{r_i} = H^2.$$

This proves the operator inequality, including arbitrary complex phases in the coefficients. Apply this to the vectorizations of $F_i K$. The convex mixture $\mathcal{M}' = \sum_i u_i \mathcal{M}_i$ proves (37). □

A.2.3 Approximation of a product-image projector

The image projector of a tensor-power isometry is itself a product, but as a single supported operator it generally acts on every site. Applying the multiplier estimate directly would therefore incur a linear logarithmic cost. We instead approximate it by a polynomial of low degree in the number of sites outside the image. A degree ℓ polynomial has a supported-contraction expansion on at most ℓ sites, even though the polynomial acts on the whole block.

Lemma 29. *Let the local space be $A \otimes C_0$, fix a unit vector $|0\rangle \in C_0$, and let $P_k = I_{A^k} \otimes (|0\rangle\langle 0|)^{\otimes k}$. For $k \ge 2$ and $0 < \xi \le 1/16$, there is an operator $H_{k,\xi}$ such that*

$$\|H_{k,\xi}\|_\infty \le 1, \qquad \|H_{k,\xi} - P_k\|_\infty \le \xi. \tag{39}$$

It has a contraction expansion of support size at most $\min\{k, \ell\}$ and coefficient cost at most 15^ℓ, where $\ell = \lceil \sqrt{k} \ln(2/\xi) \rceil$.

Proof. If $\dim C_0 = 1$, then $P_k = I$; take $H_{k,\xi} = I$, which satisfies all the claimed bounds. Otherwise let q_i be the one-site projection onto $|0\rangle^\perp$ in C_0, including the identity on A, and put $Z = \sum_{i=1}^k q_i$. Its spectrum is contained in $\{0, 1, \ldots, k\}$, and P_k is the projection onto its zero eigenspace. Let T_j be the Chebyshev polynomial, with $T_0 = 1$, $T_1 = x$, and $T_{j+1} = 2xT_j - T_{j-1}$. Set

$$H_{k,\xi} = \frac{T_\ell((k+1-2Z)/(k-1))}{T_\ell((k+1)/(k-1))}. \tag{40}$$

At $Z = 0$ this equals one. For $1 \le Z \le k$ the argument in the numerator is in $[-1, 1]$, so its absolute value is at most one. Moreover $\operatorname{arcosh}((k+1)/(k-1)) \ge 1/\sqrt{k}$. For example, $\cosh(1/\sqrt{k}) \le 1 + 1/k \le 1 + 2/(k-1)$ for $k \ge 2$, using the power series on $[0, 1]$. Thus the denominator in (40) is at least $\frac{1}{2} e^{\ell/\sqrt{k}}$, proving (39).

The operator $X = (k+1-2Z)/(k-1)$ has a one-site contraction expansion, also allowing the identity, of coefficient cost $(3k+1)/(k-1) \le 7$. Products of contractions are contractions supported on the union of their supports. The recurrence gives a support-j expansion of $T_j(X)$ of cost at most 15^j: the induction step costs at most $14 \cdot 15^j + 15^{j-1} < 15^{j+1}$. Dividing by the denominator, which is at least one, does not increase the cost. □

A.3 Channel completion

We first correct a positive-overlap branch for an isometric target, then use auxiliary extensions to treat an arbitrary target. This proves Theorem 16, whose quantitative consequences are derived in Section 6. Appendix B.2 also uses the resulting domination bound to give an alternative proof of the Stein identity.

A.3.1 Completion for an isometric target

The core construction corrects a branch with positive overlap with an isometric target. Its local corrections are controlled by three estimates: weighted local-support approximation, free multiplier completion, and a product-image projector (Lemmas 27 to 29). Their quantitative forms were established in Appendix A.2.

The purpose of the next proposition is to repair a branch that has nonzero overlap with the target in every input direction. Its conclusion preserves the original coefficient p, apart from a controlled loss that is subexponential at fixed or inverse-polynomial accuracy, while replacing that branch by a channel close to the target. The first three steps of the proof obtain symmetry, discard sites, and correct the reduced operator. Step 4 establishes free domination and restores trace preservation; Step 5 restores the discarded sites. Step 6 handles the regime in which the local approximation is not used.

Proposition 30 (Isometric target)**.** *Let $J : A \to D$ be an isometry, assume $\dim D$ is a multiple of $a = \dim A$, and let $\mathfrak{F}$ be admissible and also satisfy F4 and F5, with input state τ and output replacer state ω. Fix $c > 0$. There is a constant K_c with the following property. If*

$$\mathcal{M} \in \mathfrak{F}_n, \qquad \mathcal{M} \geq_{\mathrm{CP}} p\,\mathrm{Ad}_V, \qquad 0 < p \leq 1,$$
$$\|V\|_\infty \leq 1, \qquad \mathrm{Re}(J_n^\dagger V) \geq cI, \qquad J_n = J^{\otimes n},$$

then, for every $0 < \delta \leq 1/16$, there are $\mathcal{S}_n \in \mathfrak{F}_n$ and a CPTP map $\mathcal{L}_n$ such that

$$p\,2^{-K_c n^{2/3}\log_2((n+1)/\delta)}\mathcal{L}_n \leq_{\mathrm{CP}} \mathcal{S}_n,$$
$$\|\mathcal{L}_n - \mathrm{Ad}_{J_n}\|_\diamond \leq \delta.$$

The constant depends only on the local dimensions, c, $\lambda_{\min}(\tau)$, and $\lambda_{\min}(\omega)$. It is independent of n, δ, p, V, J, and the free family.

Proof. Step 1: Symmetrization and precision parameters. For any unit vector $|x\rangle \in A^n$, the assumptions give

$$c \leq \mathrm{Re}\langle x|J_n^\dagger V|x\rangle \leq \|J_n|x\rangle\|\,\|V|x\rangle\| \leq 1.$$

Thus $0 < c \leq 1$.

Since $\dim D$ is a multiple of $\dim A$, choose C_0 with $\dim C_0 = \dim D/\dim A$ and an output unitary $U : D \to A \otimes C_0$ such that $UJ|x\rangle = |x\rangle \otimes |0\rangle$. Such a unitary exists by extending orthonormal bases of the two isometric images. Applying U at every output site transforms $\mathfrak{F}_r$ into $\{\mathrm{Ad}_{U^{\otimes r}} \circ \mathcal{M} : \mathcal{M} \in \mathfrak{F}_r\}$ and ω into $U\omega U^\dagger$. This change preserves all family axioms, τ, the eigenvalues of ω, and the norm, overlap, CP-order, and diamond-distance bounds used here.

Keeping the same notation, identify $D^r \simeq A^r \otimes C_0^r$ and write $J_r = J^{\otimes r}$, so that $J_r|x\rangle = |x\rangle \otimes |0\rangle^{\otimes r}$. For $\pi \in S_n$, define

$$V_\pi = U_\pi^D V (U_\pi^A)^\dagger, \qquad \overline{V} = \frac{1}{n!}\sum_\pi V_\pi,$$
$$\mathcal{M}_\pi = \mathrm{Ad}_{U_\pi^D} \circ \mathcal{M} \circ \mathrm{Ad}_{(U_\pi^A)^\dagger}.$$

Let $\overline{\mathcal{M}} = n!^{-1}\sum_\pi \mathcal{M}_\pi$. This channel is free by permutation invariance and convexity. Since J_n intertwines the input and output permutations, each V_π has norm at most one and $\mathrm{Re}(J_n^\dagger V_\pi) \geq cI$. Their average satisfies both properties. The required CP order follows from the covariance identity

$$\frac{1}{n!}\sum_\pi |V_\pi\rangle\langle V_\pi| - |\overline{V}\rangle\langle\overline{V}| = \frac{1}{n!}\sum_\pi |V_\pi - \overline{V}\rangle\langle V_\pi - \overline{V}| \geq 0.$$

Consequently $\overline{\mathcal{M}} \geq_{\mathrm{CP}} p\,\mathrm{Ad}_{\overline{V}}$. This compares Choi operators, not the generally non-Hermitian operators V_π themselves. Thus $W = J_n^\dagger\overline{V}$ is a permutation-invariant contraction with $\mathrm{Re}\,W \geq cI$.

We choose the approximation precision explicitly. Put

$$\lambda = c/2, \qquad q_0 = \sqrt{1 - 3c^2/4},$$
$$q_1 = (1 + q_0)/2, \quad M_c = \frac{\lambda}{1 - q_1}.$$

and fix

$$\kappa_c = \min\left\{\frac{1}{16}, \frac{q_1 - q_0}{2\lambda}, \frac{1}{16(2M_c + 1)}\right\}, \qquad \xi = \kappa_c \delta. \tag{41}$$

Let $\ell_n = \ln((n+1)/\xi)$, $m = \lceil n^{2/3}\ell_n \rceil$, and $k = n - m$. First treat the active regime where n exceeds the uniform threshold in Lemma 27 and $m \leq n/2$; increase the threshold if necessary to ensure $k \geq 2$. In the remaining regime the faithful replacer gives an exact approximation with a cost covered by the same estimate.

Step 2: Weighted discarding. Keep the first k input-output pairs and discard the last m. The reduced channel is

$$\mathcal{M}_k(Y) = \mathrm{Tr}_{D_{k+1}\cdots D_n} \overline{\mathcal{M}}(Y \otimes \tau^{\otimes m}), \qquad Y \in \mathrm{L}(A^k).$$

It belongs to $\mathfrak{F}_k$ by (21). Applying this preparation and partial trace to the CP inequality for $\overline{\mathcal{M}}$ gives a reduced CP branch. We next select a single coherent Kraus combination from that branch. Diagonalize $\tau^{\otimes m} = \sum_j u_j |j\rangle\langle j|$. It follows that $\mathcal{M}_k \geq_{\mathrm{CP}} p\,\mathrm{Ad}_K$, where

$$K = \sum_j u_j (I_{D^k} \otimes \langle j|J_m^\dagger)\overline{V}(I_{A^k} \otimes |j\rangle). \tag{42}$$

To verify this claim, choose an orthonormal basis $\{|e_l\rangle\}$ of D^m containing the orthonormal vectors $J_m|j\rangle$. The marginal of $\mathrm{Ad}_{\overline{V}}$ has Kraus operators

$$A_{l,j} = \sqrt{u_j}(I_{D^k} \otimes \langle e_l|)\overline{V}(I_{A^k} \otimes |j\rangle).$$

For the basis index $l(j)$ with $|e_{l(j)}\rangle = J_m|j\rangle$, set $A_j = A_{l(j),j}$. Equation (42) reads $K = \sum_j \sqrt{u_j} A_j$. There are two square-root weights: one comes from preparing the mixed input and one from the coherent Kraus combination. Since $\sum_j u_j = 1$, for every vector $|v\rangle$ in the reduced Choi space,

$$|\langle v \mid K\rangle|^2 \leq \sum_j |\langle v \mid A_j\rangle|^2 \leq \sum_{l,j} |\langle v \mid A_{l,j}\rangle|^2.$$

The Choi criterion proves $\mathcal{M}_k \geq_{\mathrm{CP}} p\,\mathrm{Ad}_K$. No measurement outcome j is recorded in the final branch; the combination is justified by CP order. Equation (42) is a convex average of contractions, so $\|K\|_\infty \leq 1$. Its compression satisfies

$$X = J_k^\dagger K = \mathbb{E}_{\tau,m}(W), \qquad \|X\|_\infty \leq 1, \qquad \mathrm{Re}\, X \geq cI. \tag{43}$$

For clarity, the compression is

$$J_k^\dagger K = \sum_j u_j (I_{A^k} \otimes \langle j|) W (I_{A^k} \otimes |j\rangle),$$

which is exactly the weighted expectation defined in Lemma 27. That expectation is unital and completely positive, hence contractive in operator norm and order preserving on Hermitian parts. This proves both bounds in (43).

Step 3: Local correction of the reduced branch. There are two errors to correct. The component $(I - J_k J_k^\dagger)K$ lies outside the target image, while the component inside that image is $J_k X$, which is distorted by X. The inequality $\mathrm{Re}\, X \geq cI$ implies

$$c\langle v|v\rangle \leq \mathrm{Re}\langle v|X|v\rangle \leq \||v\rangle\| \, \|X|v\rangle\|,$$

so X is invertible and $\|X^{-1}\|_\infty \le c^{-1}$. An exact correction would be

$$O_{\text{exact}} = (X^{-1} \otimes I_{C_0^k}) J_k J_k^\dagger, \qquad O_{\text{exact}} K = J_k.$$

The issue is to control its free-domination cost. We approximate the inverse and the image projector separately by operators with controlled local expansions.

Apply Lemma 27 with precision ξ. It provides P with $\|P - X\|_\infty \le \xi$, support size $R = O_{a,t}(n^{2/3})$, and coefficient cost $M_P \le \exp[O_{a,t}(n^{2/3})]$, where $t = \lambda_{\min}(\tau)$. For $a = 1$, use the scalar $P = X$ directly. The positive real part of X implies

$$\begin{aligned}(I - \lambda X)^\dagger (I - \lambda X) &= I - 2\lambda \operatorname{Re} X + \lambda^2 X^\dagger X \\ &\le (1 - 2\lambda c + \lambda^2) I = (1 - 3c^2/4) I.\end{aligned}$$

By (41), $\|I - \lambda P\|_\infty \le q_0 + \lambda\xi \le q_1 < 1$. Choose an integer $L \ge 0$ with $q_1^{L+1} \le \xi$ and $L = O_c(\ln(1/\xi))$. Define

$$G = \lambda \sum_{j=0}^{L} (I - \lambda P)^j. \tag{44}$$

Then

$$\|G\|_\infty \le M_c, \qquad \|GX - I\|_\infty \le M_c \xi + q_1^{L+1} \le (M_c + 1)\xi. \tag{45}$$

Here P approximates the compressed operator X and is not a projection. The finite geometric-series identity gives $GP = I - (I - \lambda P)^{L+1}$, so

$$GX - I = G(X - P) - (I - \lambda P)^{L+1}.$$

Bounding these two summands proves (45). Expanding (44) gives a contraction expansion of support at most LR and cost at most $\lambda(L+1)(1 + \lambda M_P)^L$. Both the support bound and the logarithm of this cost are $O_{a,t,c}(n^{2/3}\ell_n)$, uniformly in the active regime.

Lift G to $\widehat{G} = G \otimes I_{C_0^k}$ on D^k. Let $P_k = J_k J_k^\dagger$ and choose $H = H_{k,\xi}$ from Lemma 29. Its support and logarithmic cost are $O(\sqrt{k}\ln(2/\xi)) = O(n^{2/3}\ell_n)$. Thus $O = \widehat{G}H$ also has a contraction expansion with support and logarithmic cost $O_{a,d,t,c}(n^{2/3}\ell_n)$, where $d = \dim D$. The operator P_k is the exact image projection, whereas H is its local polynomial approximation. The identities $P_k K = J_k X$ and $\widehat{G} J_k = J_k G$ give

$$OK - J_k = \widehat{G}(H - P_k)K + J_k(GX - I).$$

The leakage and distortion errors are therefore bounded separately:

$$\begin{aligned}\|OK - J_k\|_\infty &\le \|G\|_\infty \|H - P_k\|_\infty \|K\|_\infty \\ &\quad + \|GX - I\|_\infty \\ &\le (2M_c + 1)\xi =: \zeta.\end{aligned}$$

The error is controlled by the bounded operator norm of G; its expansion cost enters the CP-domination estimate below.

Step 4: Free domination and trace-preserving completion. Put $w = \lambda_{\min}(\omega)$. Write the expansion of O just obtained as $O = \sum_i c_i F_i$, let r_i be its support sizes, and set $H_* = \sum_i |c_i| h_{r_i}$ using (36). Since $\log_2 h_r = O_{a,d,t,w}(r)$, the support and coefficient bounds from Step 3 control this weighted cost. Since $\zeta \le \delta/16 < 1$, the corrected branch is nonzero and $H_* > 0$. By Lemma 28 there is a free $\mathcal{K}_k$ with $\mathcal{K}_k \ge_{\text{CP}} p H_*^{-2} \operatorname{Ad}_{OK}$, where

$$\log_2 \max\{1, H_*\} = O_{a,d,t,c,w}(n^{2/3}\ell_n).$$

Set $T = OK/(1+\zeta)$, $h_* = \max\{1, H_*\}$, and $p' = p/h_*^2$. Then $\|T\|_\infty \le 1$, $\|T - J_k\|_\infty \le 2\zeta$, $0 < p' \le 1$, and $\mathcal{K}_k \ge_{\mathrm{CP}} p' \operatorname{Ad}_T$. Complete this branch to the channel

$$\mathcal{L}_T(Y) = TYT^\dagger + \operatorname{Tr}[(I - T^\dagger T)Y]\omega^{\otimes k}.$$

The defect satisfies $0 \le D_T = I - T^\dagger T \le I$, since T is a contraction. The second term, denoted by $\mathcal{C}_T$, has Choi operator $D_T^T \otimes \omega^{\otimes k} \ge 0$, so it is CP, and for every Y, $\operatorname{Tr}\mathcal{L}_T(Y) = \operatorname{Tr} T^\dagger T Y + \operatorname{Tr} D_T Y = \operatorname{Tr} Y$. Thus $\mathcal{L}_T$ is exactly trace preserving, not merely close to a trace-preserving map. The bound $D_T \le I$ also gives $D_T^T \otimes \omega^{\otimes k} \le I \otimes \omega^{\otimes k}$, and hence $0 \le_{\mathrm{CP}} \mathcal{C}_T \le_{\mathrm{CP}} \mathcal{R}_\omega^{\otimes k}$. Hence the free mixture $\mathcal{S}_k = (\mathcal{K}_k + \mathcal{R}_\omega^{\otimes k})/2$ dominates $(p'/2)\mathcal{L}_T$. Indeed,

$$\mathcal{S}_k - \frac{p'}{2}\mathcal{L}_T = \frac{1}{2}(\mathcal{K}_k - p' \operatorname{Ad}_T) + \frac{1}{2}(\mathcal{R}_\omega^{\otimes k} - p'\mathcal{C}_T)$$

is CP because $p' \le 1$. Moreover,

$$\begin{aligned}\|\mathcal{L}_T - \operatorname{Ad}_{J_k}\|_\diamond &\le 2\|T - J_k\|_\infty + \|I - T^\dagger T\|_\infty \\ &\le 4\|T - J_k\|_\infty \le 8\zeta \le \delta.\end{aligned}$$

We used (2), $\|\mathcal{C}_T\|_\diamond = \|I - T^\dagger T\|_\infty$, and $I - T^\dagger T = J_k^\dagger(J_k - T) + (J_k^\dagger - T^\dagger)T$.

Step 5: Restoring the discarded sites. The local replacer satisfies $\mathcal{R}_\omega \ge_{\mathrm{CP}} (w/a)\operatorname{Ad}_J$: its Choi operator is $I_A \otimes \omega \ge wI$ and the Choi vector of J has squared norm a. Pad the discarded sites by setting $\mathcal{S}_n = \mathcal{S}_k \otimes \mathcal{R}_\omega^{\otimes m}$ and $\mathcal{L}_n = \mathcal{L}_T \otimes \operatorname{Ad}_{J_m}$. Then

$$\mathcal{S}_n \ge_{\mathrm{CP}} p\, 2^{-[1+2\log_2 h_* + m\log_2(a/w)]}\mathcal{L}_n.$$

Tensoring with the channel Ad_{J_m} preserves the diamond error, so it stays at most δ. Each term in the exponent is bounded by a constant times $n^{2/3}\ell_n$. The constant depends only on a, d, t, c, w. Finally, $\xi = \kappa_c\delta$ implies $\ell_n = O_c(\log_2((n+1)/\delta))$, proving the claimed bound.

Step 6: The remaining blocklengths and accuracies. For the complementary regime take the exact approximation $\mathcal{L}_n = \operatorname{Ad}_{J_n}$ and the free $\mathcal{S}_n = \mathcal{R}_\omega^{\otimes n}$. It suffices to pay exponent $n\log_2(a/w)$. If $n \ge 4$ and $m > n/2$, then $n^{2/3}\ell_n > n/2 - 1 \ge n/4$. This absorbs the exact exponent into the claimed bound. The finitely many smaller n, including the fixed threshold, are absorbed by a further uniform increase of K_c. The proof therefore covers every n and every $0 < \delta \le 1/16$ with the same constant. □

Remark 31 (The scale of the completion cost)*.* The active regime is a case of the proof, not a restriction on the statement. Its choice of m balances the cost of restoring discarded sites against the cost of correcting the reduced branch. Before substituting m, the local approximation has support scale $O(n\sqrt{h/m} + h)$, with $h = O(\log((n+1)/\xi))$. Ignoring logarithmic factors suggests balancing m against $n/\sqrt{m}$, giving $m \asymp n^{2/3}$. This is a heuristic explanation of the power. The uniform bound proved above uses $m = \lceil n^{2/3}\ell_n\rceil$, the inverse degree $O_c(\log(1/\xi))$, and the explicit complementary-regime estimate. No optimality of the power $2/3$ is asserted.

A.3.2 Completion for a general target

To pass from this construction to an arbitrary target, we use the lemmas on branch extraction and comparison with auxiliary extensions (Lemmas 25 and 26), established in Appendix A.1. Branch extraction supplies the positive overlap from a testing lower bound. The extension comparison makes the resulting bound applicable to a Stinespring dilation with constants independent of the blocklength.

Proof of Theorem 16. Fix a desired accuracy $0 < \nu \le 1/16$. Choose an auxiliary space E of dimension ab and a Stinespring isometry $V : A \to B \otimes E$ for $\mathcal{N}$; the Kraus-rank bound allows this choice [15, Thm. 2.22]. The local lifted output dimension is ab^2, a multiple of a. For each n define

$$\mathcal{G}_n = \{\mathcal{T} \in \mathrm{CPTP}(A^n \to (B \otimes E)^n) : \mathrm{Tr}_{E^n} \circ \mathcal{T} \in \mathfrak{F}_n\}. \tag{46}$$

We check every axiom. The family is nonempty because any $\mathcal{M} \in \mathfrak{F}_n$ can be extended by a fixed auxiliary state. Taking the B marginal is a continuous linear map on Choi operators. Its inverse image of the closed convex set $\mathfrak{F}_n$, intersected with the compact set of channels, is therefore compact and convex. Its permutation and tensor-product closures follow by taking B marginals. It contains the tensor powers of the replacer channel with faithful output state $\omega \otimes I_E/(ab)$. Finally, insert the same fixed τ at the last input and discard the corresponding BE output. The B marginal of the remaining channel equals the τ-insertion marginal of an element of $\mathfrak{F}_n$, so it lies in $\mathfrak{F}_{n-1}$ by F5. Thus the lifted family satisfies the one-site marginal axiom with exactly the original input state τ.

Let $\gamma_{\delta,n}$ be the worst-case testing value for target $\mathrm{Ad}_{V^{\otimes n}}$ against $\mathcal{G}_n$, at type-I error at most δ. This δ is a testing tolerance; ν remains the desired diamond accuracy. Apply Lemma 26 to each block of n uses, choosing δ, t from the fixed ε. Then

$$\gamma_{\delta,n} \geq t\beta_{\varepsilon,n}. \tag{47}$$

These constants do not depend on n. By Lemma 5, a hardest lifted alternative $\widetilde{\mathcal{M}}_n \in \mathcal{G}_n$ exists with $b_\delta(\mathrm{Ad}_{V^{\otimes n}}, \widetilde{\mathcal{M}}_n) = \gamma_{\delta,n}$. The faithful replacer and Lemma 4 imply $\gamma_{\delta,n} > 0$. By Lemma 25, this alternative CP-dominates $\gamma_{\delta,n}\,\mathrm{Ad}_{W_n}$ for an operator satisfying $\|W_n\|_\infty \leq 1$ and $\mathrm{Re}[(V^{\otimes n})^\dagger W_n] \geq c_\delta I$, where $c_\delta = (1-\sqrt{1-\delta})/(1+\sqrt{1-\delta}) > 0$ is independent of n.

Apply Proposition 30 in the lifted family at accuracy ν. It supplies a free $\widetilde{\mathcal{S}}_n$ and a channel $\widetilde{\mathcal{L}}_n$ with

$$\widetilde{\mathcal{S}}_n \geq_{\mathrm{CP}} \gamma_{\delta,n}\, 2^{-h_\delta(n,\nu)} \widetilde{\mathcal{L}}_n, \qquad \|\widetilde{\mathcal{L}}_n - \mathrm{Ad}_{V^{\otimes n}}\|_\diamond \leq \nu,$$

where $h_\delta(n,\nu) = O_\delta(n^{2/3}\log_2((n+1)/\nu))$, with a constant uniform in n and ν. Trace out E^n. The resulting $\mathcal{S}_n$ belongs to $\mathfrak{F}_n$ by (46), the resulting $\mathcal{L}_n$ is CPTP, and postcomposition with the CPTP partial-trace map preserves CP order and cannot increase the diamond distance. Use (47) and absorb $\log_2(1/t)$ into $K_\varepsilon n^{2/3}\log_2((n+1)/\nu)$ for all $n \geq 1$. This gives the required CP domination and diamond-norm approximation on the original output space. □

B Symmetric entropy estimates and an alternative limit proof

The entropy minimax identity below holds under the basic assumptions. Permutation invariance gives a further finite-block entropy estimate. With marginal closure as well, quantitative completion and near subadditivity yield an alternative proof of Theorem 2. This proof supplies a second route to the ordinary limits through the explicit completion bound of Theorem 16.

B.1 Entropy minimax and a large-error bound

We next use concavity in the input marginal to exchange the input and alternative optimizations. For $\rho \in \mathrm{D}(A^n)$, let $|\chi_\rho\rangle = (I \otimes \sqrt{\rho})|\Gamma_{A^n}\rangle$. For a channel $\mathcal{M} : \mathrm{L}(A^n) \to \mathrm{L}(B^n)$, define

$$f(\rho,\mathcal{M}) = D\big((\mathrm{id}\otimes\mathcal{N}_n)(|\chi_\rho\rangle\langle\chi_\rho|)\big\|(\mathrm{id}\otimes\mathcal{M})(|\chi_\rho\rangle\langle\chi_\rho|)\big).$$

The state $|\chi_\rho\rangle$ is the canonical purification of the physical input marginal ρ. In the vectorization convention of (3), $|\chi_\rho\rangle = |\sqrt{\rho}\rangle$ and its reference marginal is ρ^T, consistently with the tester convention in (12).

Lemma 32. *The function f is concave in ρ and convex in $\mathcal{M}$, allowing extended values. Moreover,*

$$E_n = \max_{\rho\in\mathrm{D}(A^n)} \inf_{\mathcal{M}\in\mathfrak{F}_n} f(\rho,\mathcal{M}). \tag{48}$$

Here permutation covariance means $\mathcal{M} \circ \mathrm{Ad}_{U_\pi^A} = \mathrm{Ad}_{U_\pi^B} \circ \mathcal{M}$ for every π. If $\mathcal{M}$ is permutation-covariant and has a positive replacer component, a permutation-invariant ρ attains $D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M})$. Having a positive replacer component means that $\mathcal{M} \geq_{\mathrm{CP}} \eta \mathcal{R}_\omega^{\otimes n}$ for some $\eta > 0$.

Proof. Step 1: Concavity and convexity. Purification and Schmidt compression give $D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}) = \sup_\rho f(\rho, \mathcal{M})$. For $\rho = \sum_i p_i \rho_i$, the vector $\sum_i \sqrt{p_i} |i\rangle |\chi_{\rho_i}\rangle$ is another purification of ρ. All purifications of a fixed input marginal are related by isometries on a sufficiently enlarged reference. The flagged vector may therefore be used to compute $f(\rho, \mathcal{M})$. Dephasing its flag produces a direct sum of the output pairs for ρ_i, with common weights p_i. Dephasing is a reference channel. Data processing and the direct-sum identity therefore give $f(\rho, \mathcal{M}) \geq \sum_i p_i f(\rho_i, \mathcal{M})$. Convexity in $\mathcal{M}$ follows from joint convexity of relative entropy.

Step 2: Faithful regularization. For $0 < \delta < 1$ put $\mathcal{M}_\delta = (1-\delta)\mathcal{M} + \delta \mathcal{R}_\omega^{\otimes n}$ and $f_\delta(\rho, \mathcal{M}) = f(\rho, \mathcal{M}_\delta)$. Denote the target output on $|\chi_\rho\rangle$ by ϱ and the $\mathcal{M}_\delta$ output by σ. These states obey $\varrho \leq K_\delta \sigma$ with $K_\delta = 2^{nC}/\delta$, by (10). This makes f_δ finite and jointly continuous, including at singular input marginals. Here is a useful direct verification. On pairs of states of a fixed dimension D_0 satisfying $\varrho \leq K\sigma$, replace $\log_2 \sigma$ by $\log_{2,t} \sigma$, defined by replacing each eigenvalue λ of σ by $\max\{\lambda, t\}$ before taking the logarithm, where $t > 0$. Write $\lambda_j(\sigma)$ for the eigenvalues of σ. Then

$$\begin{aligned} 0 &\leq \operatorname{Tr} \varrho(\log_{2,t} \sigma - \log_2 \sigma) \\ &\leq K \sum_{0<\lambda_j(\sigma)<t} \lambda_j(\sigma) \log_2 \frac{t}{\lambda_j(\sigma)} \leq \frac{K D_0 t}{e \ln 2}. \end{aligned}$$

Zero eigenvalues contribute nothing because $\varrho \leq K\sigma$. Thus the cross-entropy is a uniform limit of continuous functions; $\operatorname{Tr} \varrho \log_2 \varrho$ is continuous as well. Canonical purification is continuous in ρ.

Step 3: Passage to the unregularized problem. Minimax now applies to f_δ. In the rest of this argument, every infimum over $\mathcal{M}$ is over $\mathfrak{F}_n$. If $g(\rho) = \inf_{\mathcal{M}} f(\rho, \mathcal{M})$ and $g_\delta(\rho) = \inf_{\mathcal{M}} f_\delta(\rho, \mathcal{M})$, then

$$0 \leq g_\delta(\rho) - g(\rho) \leq -\log_2(1-\delta). \tag{49}$$

The lower bound holds because all $\mathcal{M}_\delta$ belong to $\mathfrak{F}_n$; the upper bound follows from $\mathcal{M}_\delta \geq_{\mathrm{CP}} (1-\delta)\mathcal{M}$ and operator monotonicity of logarithm. The same bounds hold between $\inf_{\mathcal{M}} \max_\rho f_\delta(\rho, \mathcal{M})$ and E_n. Letting $\delta \downarrow 0$ proves (48). Each g_δ is continuous by compactness and joint continuity; (49) implies uniform convergence to g. Consequently the maximum is attained.

Step 4: Symmetry. For a fixed permutation-covariant $\mathcal{M}$ with a positive replacer component, $f(\cdot, \mathcal{M})$ is continuous by the same argument. Both $\mathcal{M}$ and the tensor-power target $\mathcal{N}_n$ are permutation-covariant. Conjugating the input marginal by U_π^A therefore conjugates both output states by the same reference-output unitary. Unitary invariance of relative entropy gives $f(U_\pi^A \rho (U_\pi^A)^\dagger, \mathcal{M}) = f(\rho, \mathcal{M})$. Its concavity shows that averaging a maximizing ρ over permutations preserves maximality. □

Permutation symmetry now gives a finite-block lower bound on the testing exponent in terms of the optimized entropy.

Lemma 33. *Let $\mathfrak{F}$ be admissible and also satisfy F4. Set $h = ab$ and $v_n = \binom{n+h^2-1}{h^2-1}$. For $0 < \varepsilon < 1$,*

$$-\log_2 \beta_{\varepsilon,n} \geq \frac{E_n - (1-\varepsilon)(nC+1) - \log_2 v_n}{\varepsilon} - 1. \tag{50}$$

Proof. Step 1: A symmetric hardest alternative. As a function of $\mathcal{M}$, $b_\varepsilon(\mathcal{N}_n, \mathcal{M})$ is concave, since it is the infimum of linear acceptance probabilities over a fixed feasible tester set. Permuting the input and output of a tester preserves its target acceptance because $\mathcal{N}_n$ is permutation-covariant. This bijection of feasible testers makes $b_\varepsilon(\mathcal{N}_n, \mathcal{M})$ invariant under simultaneous input-output permutations of $\mathcal{M}$. Average a hardest alternative from Lemma 5 over these permutations. Convexity and permutation closure keep the average in $\mathfrak{F}_n$, while concavity preserves maximality. The resulting hardest alternative is permutation-covariant.

Fix any such $\mathcal{M}$ and set $\mathcal{M}' = (\mathcal{M} + \mathcal{R}_\omega^{\otimes n})/2$. Choose a permutation-invariant maximizing marginal for $D_{\mathrm{ch}}(\mathcal{N}_n \| \mathcal{M}')$, using Lemma 32. Let ρ, σ be the target and $\mathcal{M}'$ outputs on its canonical purification. Then $D(\rho\|\sigma) \geq E_n$ and $\rho \leq 2^{nC+1}\sigma$. Both outputs are invariant under pair permutations on $(\mathbb{C}^{ab})^{\otimes n}$. To see this, the maximizing input marginal and its square root commute with every input permutation. Permutation matrices are real in the product basis, so the canonical purification is fixed by the same permutation on its reference and input factors. Covariance of both $\mathcal{N}_n$ and $\mathcal{M}'$ then transfers this invariance to their reference-output states.

Step 2: Polynomial spectral complexity. Every permutation-invariant operator on $(\mathbb{C}^h)^{\otimes n}$ lies in an algebra of dimension at most v_n: expand in tensor products of the h^2 matrix units and group terms into permutation orbits. Each orbit is determined by h^2 multiplicities summing to n. A Hermitian element of this algebra has at most v_n distinct eigenvalues, since its spectral projectors, equivalently suitable polynomials in that element, are linearly independent.

Step 3: Reduction to a classical likelihood ratio. Write the distinct-eigenvalue decomposition as $\sigma = \sum_{j=1}^{v} \mu_j P_j$, with $v \leq v_n$, and define

$$\mathcal{P}_\sigma(Z) = \sum_{j=1}^{v} P_j Z P_j, \qquad \rho' = \mathcal{P}_\sigma(\rho).$$

Then ρ' commutes with σ, and pinching preserves traces against every operator commuting with the P_j. Pinching into v orthogonal blocks is an average of v unitary conjugations, so its entropy increase is at most $\log_2 v$. Indeed, the entropy of a mixture is at most the average entropy plus the Shannon entropy of its mixing weights, while each unitary conjugate has entropy $S(\rho) = -\operatorname{Tr}\rho\log_2\rho$. Moreover $\operatorname{Tr}\rho'\log_2\sigma = \operatorname{Tr}\rho\log_2\sigma$, so $D(\rho\|\sigma) - D(\rho'\|\sigma) = S(\rho') - S(\rho) \leq \log_2 v$. Since pinching fixes σ,

$$D(\rho'\|\sigma) \geq D(\rho\|\sigma) - \log_2 v_n \geq E_n - \log_2 v_n.$$

Also $\rho' \leq 2^L\sigma$, where $L = nC + 1$. Diagonalize the commuting states ρ', σ, writing their diagonal entries as ρ'_j and σ_j, respectively. Under the probability distribution given by ρ', the log-likelihood variable $Z = \log_2(\rho'_j/\sigma_j)$ is bounded above by L and has expectation at least $E_n - \log_2 v_n$. Zero-probability coordinates are omitted.

Step 4: A feasible threshold test. For $t < L$, write $p = \Pr(Z \geq t)$. The bound $\mathbb{E}Z \leq t + (L-t)p$ yields

$$p \geq \frac{E_n - \log_2 v_n - t}{L - t}. \tag{51}$$

Choose a common eigenbasis of ρ' and σ and let Q_t project onto the coordinates with $\rho'_j/\sigma_j \geq 2^t$. Coordinates outside the support of ρ' may be excluded. Then

$$\begin{aligned}
&\operatorname{Tr} Q_t\rho = \operatorname{Tr} Q_t\rho' = p,\\
&\operatorname{Tr} Q_t\sigma = \sum_{Z_j \geq t}\sigma_j \leq 2^{-t}\sum_{Z_j\geq t}\rho'_j \leq 2^{-t}.
\end{aligned}$$

The first equality follows because Q_t is fixed by the pinching. The output σ_0 of the original $\mathcal{M}$ satisfies $\sigma_0 \leq 2\sigma$, so its acceptance is at most 2^{1-t}. Choose

$$t = \frac{E_n - \log_2 v_n - (1-\varepsilon)L}{\varepsilon}.$$

One has $t < L$ by Lemma 4, and substituting this t in (51) gives $p \geq 1 - \varepsilon$. The resulting bound on $b_\varepsilon(\mathcal{N}_n, \mathcal{M})$ is uniform in the permutation-covariant alternative. Minimax and the hardest-alternative symmetry prove (50). $\square$

B.2 Convergence by near subadditivity

Throughout this subsection, assume Definition 1 and both F4 and F5. The quantitative completion and error-comparison results of Section 6, together with Lemma 33, give the following alternative limit argument.

The next elementary sequence estimate converts the completion overhead into convergence of the normalized testing quantity. Sublinearity makes a single block's loss negligible per use, but the concatenation argument must also control the accumulated loss over many scales. What we use is the summability of the dyadic errors below. An arbitrary $o(n)$ remainder is not a substitute for this estimate.

Lemma 34. *Suppose* $0 \leq a_n \leq C_0 n + C_1$ *and*

$$a_{n+m} \leq a_n + a_m + F(n+m) \tag{52}$$

for all positive n, m*, where* F *is nonnegative, nondecreasing, and* $F(s) = O(s^{2/3}\log_2(s+1))$*. Then* a_n/n *has a finite ordinary limit.*

Proof. Fix $k \geq 1$. Merge m blocks of length k in balanced rounds, carrying an unpaired block to the next round. At round j, each merged block has size at most $2^j k$. For every nonempty round, there are at most $2m/2^j$ merges. Iterating (52) therefore gives

$$\frac{a_{mk}}{mk} \leq \frac{a_k}{k} + 2\sum_{j\geq 1} \frac{F(2^j k)}{2^j k}.$$

The infinite sum is an upper bound on the finite list of merge costs, all of which are nonnegative. Its terms are bounded by a constant times $k^{-1/3}2^{-j/3}[\log_2(k+1)+j]$. The series converges, and its sum is $O(k^{-1/3}\log_2(k+1))$, hence tends to zero as $k \to \infty$.

For general $n = mk+r$, $0 \leq r < k$, an additional merge when $r > 0$ yields $a_n \leq a_{mk}+a_r+F(n)$. For fixed k the remainder contribution divided by n vanishes. Since $mk/n \to 1$ and $a_{mk}/(mk)$ is uniformly bounded, we obtain

$$\limsup_n \frac{a_n}{n} \leq \frac{a_k}{k} + 2\sum_{j\geq 1} \frac{F(2^j k)}{2^j k}.$$

Choose a sequence of integers k along which a_k/k tends to its lower limit. The series on the right tends to zero along this sequence. The upper limit of a_n/n is therefore no larger than its lower limit, which proves convergence. $\square$

Alternative proof of Theorem 2 under all five conditions. Step 1: Existence of the operational limit. Set $\varepsilon_0 = 1/2$ and let $a_n = -\log_2 \beta_{\varepsilon_0,n}$. The uniform bounds imply $0 \leq a_n \leq nC+1$. Set $\eta_n = \min\{1/16, n^{-2}\}$. Since $\log_2((n+1)/\eta_n) = O(\log_2(n+1))$, applying Theorem 16 at tolerance ε_0 and accuracy η_n gives channels $\mathcal{S}_n \in \mathfrak{F}_n$ and $\mathcal{L}_n$ satisfying

$$\mathcal{S}_n \geq_{\mathrm{CP}} \beta_{\varepsilon_0,n}2^{-f(n)}\mathcal{L}_n, \qquad \|\mathcal{L}_n - \mathcal{N}_n\|_\diamond \leq \eta_n,$$
$$f(n) = K'_{\varepsilon_0} n^{2/3}\log_2(n+1) = o(n).$$

Tensor closure and CP order give

$$\mathcal{S}_n \otimes \mathcal{S}_m \geq_{\mathrm{CP}} \beta_{\varepsilon_0,n}\beta_{\varepsilon_0,m}2^{-f(n)-f(m)}(\mathcal{L}_n \otimes \mathcal{L}_m).$$

The identity

$$\mathcal{L}_n \otimes \mathcal{L}_m - \mathcal{N}_n \otimes \mathcal{N}_m = (\mathcal{L}_n - \mathcal{N}_n)\otimes \mathcal{L}_m + \mathcal{N}_n \otimes (\mathcal{L}_m - \mathcal{N}_m)$$

and the fact that $\mathcal{L}_m$ and $\mathcal{N}_n$ are channels give diamond error at most $\eta_n + \eta_m \leq 1/8$ from $\mathcal{N}_{n+m}$. Thus every $(n+m)$-use tester of target acceptance at least $1/2$ accepts $\mathcal{L}_n \otimes \mathcal{L}_m$ with probability at least $1/4$. This statement includes inputs entangled across the two blocks, because the bound is in diamond norm. It follows that

$$\begin{aligned} a_{n+m} &\le a_n + a_m + f(n) + f(m) + 2 \\ &\le a_n + a_m + F(n+m), \end{aligned}$$

where $F(s) = 2f(s) + 2$. The hypotheses of Lemma 34 hold. Hence a_n/n has a finite limit R, and Lemma 18 gives

$$-\frac{1}{n}\log_2 \beta_{\varepsilon,n} \longrightarrow R \quad \text{for every fixed } 0 < \varepsilon < 1. \tag{53}$$

Step 2: Identification of the entropy limit. The finite-block entropy bounds now identify the common limit. The weak converse (8) gives $\liminf_n E_n/n \ge (1-\varepsilon)R$ for every fixed ε. Letting $\varepsilon \downarrow 0$ after this limit yields $\liminf_n E_n/n \ge R$. On the other hand, Lemma 33 rearranges to

$$E_n \le \varepsilon(-\log_2 \beta_{\varepsilon,n}) + (1-\varepsilon)(nC+1) + \log_2 v_n + \varepsilon.$$

Since $\log_2 v_n = O(\log n)$, (53) implies $\limsup_n E_n/n \le \varepsilon R + (1-\varepsilon)C$. Now let $\varepsilon \uparrow 1$ to obtain $\limsup_n E_n/n \le R$. Thus the ordinary entropy limit exists and equals R. The bound $0 \le R \le C$ follows from Lemma 4. □

Acknowledgment

OpenAI Codex was used in proof development, mathematical checking, literature work, and preparation of the source manuscript, and in expanding, reorganizing, and typesetting the present exposition.

References

[1] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, pp. 99–114, 1991. https://doi.org/10.1007/BF02100287.

[2] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2428–2433, 2000. https://doi.org/10.1109/18.887855. arXiv:quant-ph/9906090.

[3] F. G. S. L. Brandão and M. B. Plenio, "A generalization of quantum Stein's lemma," *Commun. Math. Phys.*, vol. 295, pp. 791–828, 2010. https://doi.org/10.1007/s00220-010-1005-z.

[4] M. Berta, F. G. S. L. Brandão, G. Gour, L. Lami, M. B. Plenio, B. Regula, and M. Tomamichel, "On a gap in the proof of the generalised quantum Stein's lemma and its consequences for the reversibility of quantum resources," *Quantum*, vol. 7, Art. no. 1103, 2023. https://doi.org/10.22331/q-2023-09-07-1103.

[5] M. Hayashi and H. Yamasaki, "The generalized quantum Stein's lemma and the second law of quantum resource theories," *Nature Phys.*, vol. 21, pp. 1988–1993, 2025. https://doi.org/10.1038/s41567-025-03047-9.

[6] L. Lami, "A solution of the generalized quantum Stein's lemma," *IEEE Trans. Inf. Theory*, vol. 71, pp. 4454–4484, 2025. https://doi.org/10.1109/TIT.2025.3543610.

[7] M. Berta, F. G. S. L. Brandão, and C. Hirche, "On composite quantum hypothesis testing," *Commun. Math. Phys.*, vol. 385, pp. 55–77, 2021. https://doi.org/10.1007/s00220-021-04133-8.

[8] K. Fang, H. Fawzi, and O. Fawzi, "Generalized quantum asymptotic equipartition," *Commun. Math. Phys.*, vol. 407, Art. no. 208, 2026. https://doi.org/10.1007/s00220-026-05719-w.

[9] K. Fang and M. Hayashi, "Error exponents of quantum state discrimination with composite correlated hypotheses," *IEEE Trans. Inf. Theory*, vol. 72, no. 6, pp. 4140–4165, 2026. https://doi.org/10.1109/TIT.2026.3684314.

[10] M. Gao, Z. Ji, and C. Liu, "Quantum channel Stein's lemma with an exponential strong converse," 2026. arXiv:2609.27196.

[11] B. Bergh, N. Datta, and R. Salzmann, "Composite classical and quantum channel discrimination," 2023. arXiv:2303.02016.

[12] M. Hayashi and H. Yamasaki, "Generalized quantum Stein's lemma for classical-quantum dynamical resources," 2025. arXiv:2509.07271.

[13] B. Bergh, N. Datta, and A. Khaitan, "Generalized quantum Stein's lemma and reversibility of quantum resource theories for classical-quantum channels," 2025. arXiv:2509.13280.

[14] G. Gour, "Failure of the asymptotic equipartition property for quantum channels," 2026. arXiv:2609.15004.

[15] J. Watrous, *The Theory of Quantum Information*. Cambridge, U.K.: Cambridge University Press, 2018. Author's book page.

[16] M. Sion, "On general minimax theorems," *Pacific J. Math.*, vol. 8, pp. 171–176, 1958. https://doi.org/10.2140/pjm.1958.8.171.

[17] J. Watrous, "Semidefinite programs for completely bounded norms," *Theory Comput.*, vol. 5, pp. 217–238, 2009. https://doi.org/10.4086/toc.2009.v005a011.

[18] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2816–2826, 2009. https://doi.org/10.1109/TIT.2009.2018325.

[19] P. Faist, J. Oppenheim, and R. Renner, "Gibbs-preserving maps outperform thermal operations in the quantum regime," *New J. Phys.*, vol. 17, Art. no. 043003, 2015. https://doi.org/10.1088/1367-2630/17/4/043003.

[20] E. Chitambar and G. Gour, "Critical examination of incoherent operations and a physically consistent resource theory of quantum coherence," *Phys. Rev. Lett.*, vol. 117, Art. no. 030401, 2016. https://doi.org/10.1103/PhysRevLett.117.030401.

[21] M. Horodecki, P. W. Shor, and M. B. Ruskai, "Entanglement breaking channels," *Rev. Math. Phys.*, vol. 15, pp. 629–641, 2003. https://doi.org/10.1142/S0129055X03001709. arXiv:quant-ph/0302031.

[22] W. Fulton and J. Harris, *Representation Theory: A First Course*, Graduate Texts in Mathematics, vol. 129. New York, NY, USA: Springer, 1991. https://doi.org/10.1007/978-1-4612-0979-9.