

Constant-gap entanglement distillation and a spectral criterion for noise-universal communication

Lluis Eriksson

18 October 2026 — Working paper

Abstract

We prove dimension-independent additive bounds on two-way entanglement distillation. For every conjugate local diagonal unitary invariant state, distillable entanglement, secret key and relative entropy of entanglement lie in an explicitly computable interval of width at most $\log_2 5$ bits per copy. For the broader class invariant under local diagonal signs, the diagonal-relative-entropy reference exceeds distillable entanglement by at most $\log_2 9$. The proof uses established partial breeding with a stopped classification whose entropy is bounded independently of dimension; a reduction-based seed pays for all initial assistance. The resulting recovery bound also implies at most $\log_2 6$ ebits of irreversibility for any mixture of a pure state and separable noise, without a marginal constraint, and controls mixtures with a mixed maximally correlated component. For an identity component mixed with arbitrary channel noise, we give six capacity guarantees and reduce their simultaneous worst cases to depolarizing channels. An explicit isotropic interval has absolute width tending to $2(1-f)$ at fixed entangled fraction f , refining the known relative large-dimension approximation. For general discrete Hamiltonians with a finite partition sum at some temperature, we identify the exact spectral boundary of absolute entanglement-assisted recovery: for each fixed reset state, the loss vanishes precisely when the partition sum diverges at its critical inverse temperature. In that regime all six capacities attain their erasure benchmarks with absolute $o(1)$ loss. In the complementary regime every fixed reset has a strictly positive limiting assisted deficit, given by a critical-Gibbs variational formula. An explicit integer spectrum has an exact constant-loss plateau. The reset may have infinite energy or entropy. Oscillator consequences include reversible distributed states and a vanishing classical arbitrary-partner additivity defect. Optimal constants, efficient implementations and general exact finite-energy capacities remain unresolved.

1 Introduction and scope

We give an explicit, dimension-independent additive approximation to entanglement distillation for two established symmetry classes of bipartite quantum states. The computable lower endpoint uses the coherent block on $\text{span}\{|jj\rangle\}$; a stopped classification isolates that block at bounded entanglement cost. For conjugate local diagonal unitary invariant (CLDUI) states, the distillable entanglement, secret key and relative entropy of entanglement lie in one interval of width at most $\log_2 5$ bits per copy. For local diagonal orthogonal invariant (LDOI) states, an explicit diagonal-relative-entropy reference is within $\log_2 9$ of distillable entanglement. These are additive bounds, rather than approximations normalized by $\log d$.

The invariant classes and their block characterizations are established [21, 22]. Recent semidefinite hierarchies study their separability structure [23]; our claim concerns operational distillation bounds, not the introduction or separability characterization of those classes. The underlying parity measurements, partial breeding, adaptive discarding and removal of assistance also have established

antecedents [12, 13, 17, 18]. Vollbrecht–Wolf [14, Section VI.B–C] obtained optimal large-dimension rates for isotropic states after normalization by $\log d$. We use those protocol primitives, hashing [10], reduction-based distillability [15] and the logarithmic reduction argument [20] with explicit attribution. The contribution is the bounded-entropy stopped classification and its absolute-loss consequences, not those primitives.

The same argument proves that mixing any pure state with arbitrary separable noise creates at most $\log_2 6$ ebits of irreversibility, with no matching-marginal assumption. It also treats a mixed maximally correlated useful component, for which the controlled quantities are $E_D, K_D, E_R^\infty, E_R$, rather than entanglement cost. Corollary 7 applies the pure-state result to every identity/entanglement-breaking channel mixture. Section 3 then allows arbitrary input-dependent noise: standard twirling organizes six simultaneous worst-case capacities, while the classification theorem provides a new absolute two-way interval for the isotropic minimizer. The minimax reduction does not solve its unknown capacity functions. The final part applies the recovery bound to oscillator reset channels, obtaining absolute vanishing gaps for six energy-constrained capacities for any fixed reset state, including infinite energy or entropy. We give the finite-dimensional operational proof first, then the oscillator construction and its separate channel-theoretic dependencies.

Section 10 determines when the oscillator mechanism extends to a general discrete Hamiltonian. The critical Gibbs entropy description is established [28, Proposition 1a]; we derive an operational dichotomy from it. Divergence of the partition sum at the critical inverse temperature permits entropy-preserving escape and six absolute recovery laws. If the critical sum is finite, every fixed reset state instead has a positive entanglement-assisted deficit, characterized by relative entropy to the critical Gibbs state plus a continuous binary-flag loss. The proof uses infinite-dimensional relative-entropy sufficiency [30] and finite-flag continuity [29]. A constructed spectrum makes the loss constant at every sufficiently large energy. The scalar vacuum-reset decomposition is credited to [7, Appendix C]; the new claim is the capacity criterion and defect, not that decomposition or the critical entropy description.

The constants are not claimed optimal, and the protocols are asymptotic existence results, not efficient coding algorithms. We do not claim general mixed-state reversibility, resolution of NPT bound entanglement, general exact finite-energy reset capacities, or exhaustive priority certification.

2 Entanglement recovery with a constant absolute loss

All systems in this section are finite dimensional. E_D is two-way distillable entanglement, E_C is standard LOCC entanglement cost, and K_D is distillable secret key, with no initial entanglement or key and free authenticated public classical communication. Rates are per input copy. Security includes a purification and the full public transcript. States, bases and decompositions are known. E_R denotes relative entropy of entanglement and E_R^∞ its regularization. We use the established inequalities

$$E_D \leq K_D \leq E_R^\infty \leq E_R,$$

with the key upper bound from [19, Theorem 9]. Logarithms are base two. Write

$$h(x) = -x \log_2 x - (1-x) \log_2 (1-x),$$

with zero terms interpreted by continuity, and set

$$b(x, y) = (x+y)h\left(\frac{y}{x+y}\right), \quad b(0, 0) = 0. \quad (1)$$

Lemma 1 (A seed from a dominated component). *Let $\rho = a\theta + (1-a)\chi$, where $0 < a \leq 1$, and put $u_A = S(\theta_A) - S(\theta)$. If $a2^{u_A} > 1$, then ρ has a strictly positive two-way distillation rate.*

Proof. Put $X = \rho_A$. If the reduction criterion held, then $a\theta \leq \rho \leq X \otimes I$. Restrict to $\text{supp } X$; this contains the support of θ_A , since $X \geq a\theta_A$. Operator monotonicity of the logarithm, applied with a positive regularization when θ is singular, gives

$$\log_2 a - S(\theta) \leq \text{Tr } \theta_A \log_2 X, \quad \log_2 a + u_A + D(\theta_A \| X) \leq 0.$$

Nonnegativity of relative entropy contradicts the hypothesis. Violation of reduction implies a positive finite-state distillation rate [15, Section VI]. This uses the established logarithmic method relating reduction to conditional entropy [20, Theorem 2, equations (21)–(24)], rather than a new general distillability criterion. The finite-dimensional filtering and isotropic-twirling protocol supplies the seed. Its twirl has an exact finite convex-unitary realization by Carathéodory's theorem, so no uncountable outcome convention is needed. The seed rate may be small; the state is fixed before the coding block length tends to infinity. $\square$

Theorem 2 (Stopped classification). *For any state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$ in a specified paired product basis, let*

$$B_{jk} = \langle jj | \rho | kk \rangle, \quad w = \text{Tr } B, \quad v = w \{S(\text{diag}(B/w)) - S(B/w)\},$$

setting $v = 0$ when $w = 0$. Then

$$E_D(\rho) \geq \max\{0, v - h(w) - 2(1 - w)\}. \quad (2)$$

If ρ is invariant under every $U \otimes \bar{U}$ with U diagonal unitary (the established CLDUI class), then

$$\max\{0, v - h(w) - 2(1 - w)\} \leq E_D(\rho) \leq K_D(\rho) \leq E_R^\infty(\rho) \leq E_R(\rho) \leq v. \quad (3)$$

The interval has width at most $\log_2 5$, independently of d .

Proof. Finite phase averaging. Apply shared phases $U = \text{diag}(\zeta^{a_j})$, $a_j \in \{0, 1, 2\}$ and $\zeta = e^{2\pi i/3}$, as $U \otimes \bar{U}$. With $\Pi_{\text{eq}} = \sum_j |jj\rangle\langle jj|$, the averaged state is

$$T(\rho) = \Pi_{\text{eq}} \rho \Pi_{\text{eq}} + \sum_{j \neq k} \langle jk | \rho | jk \rangle |jk\rangle\langle jk|. \quad (4)$$

The characters $a_j - a_k$ are distinct for ordered pairs $j \neq k$, whereas all $|jj\rangle$ have zero character. This finite LOCC map preserves the entire equality block, including mixed maximally correlated states. The averaged AB marginal is block diagonal in $z = j \oplus k$ after padding labels to $m = \lceil \log_2 d \rceil$ bits. This is not an assertion about a state conditioned on the public phase record or on a purifying environment. The cases $d = 1$ or $w = 0$ have a trivial lower bound. When $w = 1$, ordinary hashing gives $E_D \geq v$.

A fixed low-entropy classification. For invertible binary $m \times m$ matrices A , let J_A be the first nonzero coordinate of Az , or zero for $z = 0$. Put $q = 1 - w$. Uniformly averaging A sends each nonzero z uniformly over nonzero vectors, so

$$\Pr(J = \ell \mid J > 0) = \frac{2^{m-\ell}}{2^m - 1}, \quad 1 \leq \ell \leq m.$$

Its entropy is $2 - m/(2^m - 1) + \log_2(1 - 2^{-m}) \leq 2$. Concavity gives $\mathbb{E}_A H(J_A) \leq h(q) + 2q$. Thus one fixed A satisfies this bound and can be used on all copies. No efficient search for this matrix is assumed.

Coherent queries and their cost. For binary local controls $a(j), b(k)$, append a qubit Bell pair and apply controlled X on each party's half. Computational-basis outcomes u, v on that pair induce the source Kraus operator

$$K_{uv} = 2^{-1/2} \Pi_{u \oplus v}^{\text{par}}, \quad \Pi_y^{\text{par}} = \sum_{a(j) \oplus b(k) = y} |jk\rangle \langle jk|. \quad (5)$$

Indeed $(X^{a(j)} \otimes X^{b(k)})|\Phi_2\rangle$ depends only on $a(j) \oplus b(k)$. The redundant measured bit reveals no Schmidt index. The same formula holds for parities over subsets of source copies. Each query consumes one ebit while preserving coherences inside a parity eigenspace. This is the established appended-pair/partial-breeding primitive [12, 13].

At stage ℓ , learn the next bit only on copies with previous bits zero, discarding a copy as soon as a one is learned. Let $\pi_\ell = \Pr(J_A = \ell)$ and $r_\ell = \Pr(J_A = 0 \text{ or } J_A \geq \ell)$. Binary typical-set hashing of these copies costs $r_\ell h(\pi_\ell/r_\ell)$ ebits per original copy, and the stopped-prefix chain rule gives

$$C := \sum_{\ell=1}^m r_\ell h(\pi_\ell/r_\ell) = H(J_A) \leq h(w) + 2(1-w). \quad (6)$$

Zero-weight terms are omitted. For clarity, a typical set of binary strings of length k has at most $2^{k(H+\eta)}$ members; a random linear hash with $k(H+2\eta)$ rows collides with another typical string with probability at most $2^{-k\eta}$, and atypical probability vanishes. If that row count exceeds k , measure all k bits instead. Every row is implemented by (5). Typical budgets cover random sub-sample sizes, with abort on atypical counts. There are finitely many stages at fixed d . On the averaged block-diagonal marginal, correct classification acts exactly as the ideal block projectors; the trace-norm error is at most twice the classification error. No individual Schmidt label is measured.

Survivors and seed. The all-zero projector is Π_{eq} . Survivors occur with probability w and have state $\theta = \Pi_{\text{eq}}\rho\Pi_{\text{eq}}/w$, whose correlated-basis matrix is B/w . Write $u = S(\theta_A) - S(\theta) = v/w$; this is nonnegative because dephasing increases entropy. Hashing [10, Theorem 10] achieves gross yield $G = wu = v$. Correct classification leaves independent identical survivor states conditional on their positions. A fixed output length below their typical count handles random numbers.

If $v - h(w) - 2(1-w) \leq 0$, the asserted lower bound is trivial. Otherwise $wu > h(w) \geq -w \log_2 w$, so $w2^u > 1$. The averaged state is $w\theta + (1-w)\chi_{\text{diag}}$ with a separable mismatch state. Lemma 1 therefore supplies a positive unassisted seed rate. This establishes the seed from the actual survivor state, including when it is mixed.

Paying for and recycling the reserve. Fix η with $0 < 2\eta < G - C$. A sufficiently long block of length ℓ consumes at most $(C + \eta)\ell$ ebits and returns at least $(G - \eta)\ell$, with joint error $\varepsilon_\ell \rightarrow 0$. Generate the first reserve from $O(\ell)$ source copies at the positive fixed seed rate. Run $t_\ell \rightarrow \infty$ blocks, retaining a reserve from each output, and choose t_ℓ slowly enough that $t_\ell \varepsilon_\ell \rightarrow 0$. The initial fraction is $O(1/t_\ell)$; contractivity and telescoping bound the joint error by the seed error plus $t_\ell \varepsilon_\ell$. Approximate reserves are included in that error. A diagonal choice of lengths followed by $\eta \downarrow 0$ achieves net rate $G - C$. This standard removal of assistance [13, Section V] needs no hashing error exponent. Final fidelity to pure Bell pairs implies decoupling from any purification and the complete public transcript. No free entanglement or restricted Eve has been introduced.

For a CLDUI input, (4) already equals ρ . Its fully dephased state $\Delta\rho$ is separable, and the direct-sum entropy formula gives $D(\rho||\Delta\rho) = S(\Delta\rho) - S(\rho) = v$. Thus $E_R(\rho) \leq v$, giving (3). The width is bounded by $h(w) + 2(1-w) \leq \log_2 5$, including when the lower endpoint is zero. $\square$

Corollary 3 (Diagonal orthogonal symmetry). *Suppose ρ is invariant under every $O \otimes O$ with O a diagonal sign matrix (the established LDOI class). Define B, w, v as above and $\mathcal{C}_\Delta(\rho) = S(\Delta\rho) - S(\rho)$.*

Then

$$\max\{0, v - h(w) - 2(1 - w)\} \leq E_D(\rho) \leq K_D(\rho) \leq E_R^\infty(\rho) \leq E_R(\rho) \leq \mathcal{C}_\Delta(\rho), \quad (7)$$

$$0 \leq \mathcal{C}_\Delta(\rho) - E_D(\rho) \leq h(w) + 3(1 - w) \leq \log_2 9. \quad (8)$$

In particular the same constant bounds $E_R - E_D$ and $K_D - E_D$.

Proof. The known block characterization [22, Proposition 2.2 and equations (5)–(8)] consists of B on the equality subspace and positive 2×2 blocks R_{ij} on $\text{span}\{|ij\rangle, |ji\rangle\}$, $i < j$. These subspaces are globally orthogonal; we do not measure them by a presumed free nonlocal operation. Set $p_{ij} = \text{Tr } R_{ij}$. The entropy identity for this direct sum gives

$$\mathcal{C}_\Delta(\rho) = v + c, \quad c = \sum_{i < j: p_{ij} > 0} p_{ij} \{S(\text{diag}(R_{ij}/p_{ij})) - S(R_{ij}/p_{ij})\}.$$

Each bracket lies in $[0, 1]$, hence $0 \leq c \leq 1 - w$. Theorem 2 gives the lower bound by the finite local phase twirl, which preserves B and removes the swap coherences. The separable reference $\Delta\rho$ gives the upper bound. The interval width is at most $h(w) + 2(1 - w) + c \leq h(w) + 3(1 - w)$, whose maximum is $\log_2 9$ at $w = 1/9$. $\square$

The bounds involve the spectrum of a $d \times d$ block and, for the sharper upper endpoint, 2×2 blocks. Evaluating these quantities is distinct from implementing an efficient distillation protocol. Neither constant is claimed optimal. The upper endpoints apply to the stated invariant classes; for an arbitrary state only (2) is asserted. No general bound on $E_C - E_D$ or solution of the NPT bound-entanglement problem follows.

Theorem 4 (An arbitrary noisy maximally correlated component). *Let $\theta = \sum_{jk} t_{jk} |jj\rangle\langle kk|$ be a state, $\sigma = \theta_A$, and $u = S(\sigma) - S(\theta)$. Let χ be any state on the same $d \times d$ space, with no marginal condition. For $\lambda, \delta \geq 0$, $\lambda + \delta \leq 1$, define*

$$\omega = \lambda\theta + \delta\chi + (1 - \lambda - \delta)\sigma \otimes |e\rangle\langle e|,$$

where e is an orthogonal receiver flag. Then

$$E_D(\omega) \geq \max\{0, \lambda u - b(\lambda, \delta) - \delta \log_2 5\}. \quad (9)$$

If $\theta = \psi$ is pure, write $s = S(\psi_A)$. For separable χ , $E_C(\omega) \leq \lambda s$ and

$$0 \leq E_C(\omega) - E_D(\omega) \leq b(\lambda, \delta) + \delta \log_2 5. \quad (10)$$

In particular, every mixture $\rho = \lambda\psi + (1 - \lambda)\chi$ of one pure state and an arbitrary separable state satisfies

$$E_C(\rho) - E_D(\rho) \leq h(\lambda) + (1 - \lambda) \log_2 5 \leq \log_2 6. \quad (11)$$

Proof. Bob announces the code/flag outcome; put $P = \lambda + \delta$. The case $P = 0$ is trivial. Apply Theorem 2 to the accepted normalized state. Put $\kappa = \text{Tr } \Pi_{\text{eq}} \chi$, $P_0 = \lambda + \delta\kappa$ and $\Gamma = \lambda\theta + \delta\Pi_{\text{eq}}\chi\Pi_{\text{eq}}$. When $\kappa > 0$, the normalized projected noise ξ is maximally correlated and has $I_c(\xi) \geq 0$. Entropy concavity and the entropy-of-mixture upper bound yield

$$P_0 I_c(\Gamma/P_0) \geq \lambda u + \delta\kappa I_c(\xi) - b(\lambda, \delta\kappa) \geq \lambda u - b(\lambda, \delta\kappa).$$

The physical rate is the accepted rate multiplied by P . The classification cost is at most $Ph(\delta(1 - \kappa)/P) + 2\delta(1 - \kappa)$. The three-branch entropy identity gives

$$b(\lambda, \delta\kappa) + Ph\left(\frac{\delta(1 - \kappa)}{P}\right) = b(\lambda, \delta) + \delta h(\kappa).$$

Since $h(\kappa) + 2(1 - \kappa) \leq \log_2 5$, this proves (9); zero-weight branches are interpreted directly. The seed and reserve are already paid for by Theorem 2, which uses the actual survivor. For pure θ and separable χ , formation by typical branch strings uses pure-state dilution [11] at fraction λ and free separable preparation otherwise, giving $E_C \leq \lambda s$. Combine with $E_D \leq E_C$. The last maximum is $\log_2 6$ at $\lambda = 1/6$. These are upper bounds, not attained gaps established here. $\square$

Corollary 5 (Mixed good component and separable contamination). *For $\rho = \lambda\theta + (1 - \lambda)\chi$ with maximally correlated θ and separable χ , each of $E_D, K_D, E_R^\infty, E_R$ lies between $\max\{0, \lambda u - h(\lambda) - (1 - \lambda)\log_2 5\}$ and λu . Consequently these rates converge absolutely to $(1 - q_n)u_n$ for $\lambda_n = 1 - q_n \rightarrow 1$, even in growing dimensions.*

Proof. The lower bound is (9) with $\delta = 1 - \lambda$. For θ , its diagonal separable reference gives $E_R(\theta) \leq u$. Convexity of E_R gives $E_R(\rho) \leq \lambda u$, completing the sandwich. No reversibility of E_C is implied for a mixed θ . $\square$

Corollary 6. *Let $\rho_n = (1 - q_n)\psi_n + q_n\chi_n$, where ψ_n is pure and χ_n is separable, in possibly growing finite dimensions. If $q_n \rightarrow 0$, then $E_C(\rho_n) - E_D(\rho_n) \rightarrow 0$ without requiring $q_n \log d_n \rightarrow 0$. Each of E_D, E_C lies between $(1 - q_n)S((\psi_n)_A) - h(q_n) - q_n \log_2 5$ and $(1 - q_n)S((\psi_n)_A)$.*

For example, take $\psi_n = \Phi_{d_n}$, $q_n = (\log_2 d_n)^{-1/2}$, and $d_n \rightarrow \infty$. The absolute irreversibility vanishes although $q_n \log_2 d_n$ diverges, for every choice of separable χ_n .

Corollary 7. *For any entanglement-breaking channel $\mathcal{B}_d$ from $\mathbb{C}^d$ to itself, let $\mathcal{N}_d = \lambda \text{id}_d + (1 - \lambda)\mathcal{B}_d$. With free two-way classical communication and no initial entanglement or key, each of $Q_2(\mathcal{N}_d)$ and $K_2(\mathcal{N}_d)$ lies in*

$$[\max\{0, \lambda \log_2 d - h(\lambda) - (1 - \lambda)\log_2 5\}, \lambda \log_2 d]. \quad (12)$$

The gap bound is uniform over d and $\mathcal{B}_d$, and at most $\log_2 6$. It tends to zero as $\lambda \rightarrow 1$, with no relation required between $1 - \lambda$ and d .

Proof. Transmit halves of Φ_d . The noisy Choi branch is separable with Alice marginal I/d , so Theorem 4 proves the lower bound. Ebts yield quantum communication and secret key. For any purified input, an extension that records the branch and a separable decomposition of the entanglement-breaking output has conditional mutual information $2\lambda S(\rho)$. Thus channel squashed entanglement is at most $\lambda \log_2 d$. Its adaptive upper bound on K_2 , and $Q_2 \leq K_2$, follow from [9] in finite dimension. The channel convexity and vanishing property on entanglement-breaking channels are also established in [16, Sections III.B–C]. No covariance or reset form of $\mathcal{B}_d$ is required. $\square$

3 Arbitrary noise and worst-case communication

Let $U : \mathbb{C}^d \rightarrow \mathcal{H}_B$ be an isometry, $d \geq 2$, with finite-dimensional output, and consider the fixed memoryless channel

$$\mathcal{N}_{\mathcal{M}} = \lambda \text{Ad}_U + (1 - \lambda)\mathcal{M}, \quad q = 1 - \lambda,$$

where $\mathcal{M}$ is any CPTP map. It need not be separable, unital, covariant or a replacement channel. We use unassisted classical, quantum and private capacities C, Q, P , entanglement-assisted classical capacity C_{EA} , and two-way quantum and key capacities Q_2, K_2 . Only C_{EA} receives free initial

entanglement; only Q_2, K_2 receive free authenticated public communication. Each block has finite memories, alphabets and rounds, with no uniform bound across blocklengths. All private tasks include the full channel environment and public record. The infima below optimize over fixed channels, with codes allowed to depend on the channel. They are not compound-channel or adversarial-noise capacities.

Corollary 8 (Six noise-universal bounds). *Define the benchmarks and losses by*

X	r_X	c_X
C	$\lambda \log_2 d$	$h(\lambda)$
C_{EA}	$2\lambda \log_2 d$	$h(\lambda)$
Q, P	$(2\lambda - 1)_+ \log_2 d$	$h(\lambda)$
Q_2, K_2	$\lambda \log_2 d$	$h(\lambda) + q \log_2 5$

For every $\mathcal{M}$, $X(\mathcal{N}_{\mathcal{M}}) \geq \max\{0, r_X - c_X\}$, and

$$\max\{0, r_X - c_X\} \leq \inf_{\mathcal{M}} X(\mathcal{N}_{\mathcal{M}}) \leq r_X. \quad (13)$$

For $\lambda \leq 1/2$, the Q and P infima are exactly zero. The losses are at most one bit for the first four capacities and $\log_2 6$ for the last two; all tend to zero as $\lambda \rightarrow 1$, uniformly in dimension and noise.

Proof. The first four lower bounds are standard entropy-of-mixture estimates. For a uniform orthonormal input ensemble, put $\omega_i = \mathcal{M}(|i\rangle\langle i|)$ and $\bar{\omega} = d^{-1} \sum_i \omega_i$. Concavity and the mixture upper bound give Holevo information at least $\lambda \log_2 d - h(\lambda) + q[S(\bar{\omega}) - d^{-1} \sum_i S(\omega_i)]$. For the Bell input write the output as $\rho = \lambda \Phi_U + q\chi$, where $\chi_A = I/d$. Then

$$S(\rho_B) \geq \lambda \log_2 d + qS(\chi_B), \quad S(\rho_{AB}) \leq h(\lambda) + q[\log_2 d + S(\chi_B)].$$

Coherent-information coding and $P \geq Q$ give the Q, P bounds; adding $S(\rho_A) = \log_2 d$ and the entanglement-assisted coding theorem gives the C_{EA} bound [3, 26].

For Q_2, K_2 , Bob projects onto $\text{ran } U$. The retained unnormalized state is $\lambda \Phi_d + \delta \chi'$ for some $0 \leq \delta \leq q$. Apply Theorem 4 to its normalized state and multiply the achievable rate by $\lambda + \delta$. This gives $\lambda \log_2 d - b(\lambda, \delta) - \delta \log_2 5$, at least the asserted bound. The rejected state's Alice marginal is irrelevant: those copies are discarded. Distributing Choi copies and distilling them is allowed for Q_2 , and $K_2 \geq Q_2$.

For the upper bound on the infimum, choose any replacement noise. The resulting channel is receiver postprocessing of the erasure channel with survival λ , whose six capacities are r_X [1, 26, 9]. For completeness, erasure has coherent-information rate $(2\lambda - 1) \log_2 d$ when degradable and has $Q = P = 0$ when antidegradable; its two-way quantum/key rate is $\lambda \log_2 d$, achieved on successful transmissions and bounded by the erasure squashed-entanglement bound. This upper bound concerns the infimum, not every noise: $\mathcal{M} = \text{Ad}_U$ is noiseless. $\square$

The worst noise admits a standard symmetry reduction. We give the resource accounting because private and unassisted capacities must not silently receive free shared secret randomness. Define the depolarizing channel $\mathcal{D}_f$ through its normalized Choi state

$$J_f = f\Phi_d + \frac{1-f}{d^2-1}(I - \Phi_d), \quad \mathcal{D}_f(\rho) = p_f \rho + (1-p_f) \frac{I}{d} \text{Tr } \rho, \quad p_f = \frac{d^2 f - 1}{d^2 - 1}. \quad (14)$$

Proposition 9 (Simultaneous minimax reduction). *For each of the six capacities above,*

$$\inf_{\mathcal{M}} X(\mathcal{N}_{\mathcal{M}}) = X(\mathcal{D}_{f_*}), \quad f_* = \max\{\lambda, d^{-2}\}. \quad (15)$$

One noise channel attains all six infima simultaneously. This reduces the problem to depolarizing capacities; it does not evaluate the generally unknown quantum, private or two-way capacities.

Proof. First suppose $U = I$. Conjugation twirling sends any channel $\mathcal{N}$ to $\mathcal{D}_f$, where $f = \langle \Phi_d | J_{\mathcal{N}} | \Phi_d \rangle$. The Haar average is a finite convex combination of unitary conjugations by finite-dimensional convexity. Twirling cannot increase any of the six capacities. For C, Q, C_{EA} , expand an n -use code for the twirled channel over the independent unitary seeds. Average decoding error or entanglement infidelity tends to zero, so one deterministic sequence works; absorb it into the encoder and decoder. This is the usual coding use of twirling [24, Appendix B].

For P , a Stinespring dilation of the twirl retains a coherent seed register. Dephasing that register in Eve's output gives the actual environment of $\mathcal{N}$ with the public seed. Contractivity gives security for this degraded environment. Since the seed is independent of the message, the block diagonal secrecy distance is the average conditional secrecy distance. A deterministic sequence has vanishing sum of average decoding and secrecy errors. Thus no random resource is added to the unassisted code. For adaptive Q_2, K_2 , fresh public seeds implement each twirled use. Measuring the coherent environment seed gives precisely this implementation, with every seed retained by Eve. The original protocol ignores those extra labels after the required conjugations; the full output security is preserved by contractivity.

For $\mathcal{N} = \lambda \text{id} + q\mathcal{M}$ we have $f \geq \lambda$. When $\lambda \geq d^{-2}$, $p_f \geq p_\lambda \geq 0$; receiver postprocessing by a depolarizing channel of parameter p_λ/p_f reduces $\mathcal{D}_f$ to $\mathcal{D}_\lambda$. The case $p_f = 0$ is immediate. Hence $X(\mathcal{N}) \geq X(\mathcal{D}_\lambda)$. Equality is attained by the CPTP noise

$$\mathcal{M}_0(\rho) = \frac{dI \text{Tr} \rho - \rho}{d^2 - 1}, \quad J_{\mathcal{M}_0} = \frac{I - \Phi_d}{d^2 - 1}.$$

When $\lambda \leq d^{-2}$, the CPTP map $\mathcal{M} = (\mathcal{D}_{d^{-2}} - \lambda \text{id})/(1 - \lambda)$ makes the mixture completely depolarizing. Its Choi eigenvalues are $(d^{-2} - \lambda)/(1 - \lambda)$ on Φ_d and $[d^2(1 - \lambda)]^{-1}$ on its orthogonal complement. All six capacities then vanish.

For general U , decode by $\mathcal{R}(Y) = U^* Y U + \text{Tr}[(I - U U^*)Y] \tau_0$, with any state τ_0 . This CPTP map satisfies $\mathcal{R} \text{Ad}_U = \text{id}$, so the lower bound reduces to the square case. Composing its attaining noises with Ad_U gives equality. $\square$

The classical formulas in this reduction are established results. For $\lambda \geq d^{-2}$, put $a = (d\lambda + 1)/(d + 1)$. King [25, Theorem 1] and Bennett et al. [26] give, respectively,

$$\begin{aligned} \inf_{\mathcal{M}} C(\mathcal{N}_{\mathcal{M}}) &= \log_2 d - h(a) - (1 - a) \log_2(d - 1), \\ \inf_{\mathcal{M}} C_{\text{EA}}(\mathcal{N}_{\mathcal{M}}) &= 2 \log_2 d - h(\lambda) - (1 - \lambda) \log_2(d^2 - 1). \end{aligned}$$

Below d^{-2} both are zero. The new classification bound instead controls the unresolved two-way rates with an absolute interval.

Corollary 10 (An absolute isotropic interval). *For $f \geq 1/d$, set*

$$q = 1 - f, \quad t = \frac{q}{d + 1}, \quad w = f + t, \quad v = w \log_2 d - w h(f/w) - t \log_2(d - 1),$$

and define

$$L = \max\{0, v - h(w) - 2(1 - w)\}, \quad U = \log_2 d - h(f) - q \log_2(d - 1).$$

Both $E_D(J_f)$ and $K_D(J_f)$ lie in $[L, U]$, whose width is exactly

$$U - L = \min\{U, qA_d\}, \quad A_d = \frac{2d}{d + 1} + \log_2(d + 1) - \frac{d}{d + 1} \log_2(d - 1). \quad (16)$$

The same interval bounds the two minimax capacities in (15) with $f = \lambda \geq 1/d$. For fixed $0 < f < 1$, its width tends to $2(1 - f)$ as $d \rightarrow \infty$. The two-way minimax capacities vanish for $\lambda \leq 1/d$ and are strictly positive for $\lambda > 1/d$.

Proof. The correlated block of J_f has eigenvalues f and $q/(d^2 - 1)$ with multiplicity $d - 1$, total weight w and uniform diagonal. Thus Theorem 2 gives L . The known isotropic relative entropy is $E_R(J_f) = U$ [27, Supplementary Note 5, Eqs. (S195)–(S196)]; it bounds both rates. The entropy chain rule

$$h(w) + wh(f/w) = h(f) + qh\left(\frac{1}{d+1}\right)$$

gives $U - [v - h(w) - 2(1 - w)] = qA_d$, proving (16). Depolarizing channels are teleportation covariant, so their two-way capacities equal $E_D(J_f)$ and $K_D(J_f)$ respectively: distribution gives one direction and exact teleportation simulation gives the other [24, 27]. Isotropic states are separable up to $f = 1/d$; above it, reduction violation gives positive distillability [15, 14]. This proves the threshold, even where the displayed quantitative lower bound is zero. $\square$

Vollbrecht–Wolf [14, Section VI.B–C, Eqs. (40)–(47)] already obtained the optimal normalized isotropic rate $E_D(J_f)/\log_2 d \rightarrow f$. The assertion here is the finite-dimensional absolute interval (16), not that prelog, the symmetry reduction, or the established classical and entanglement-assisted formulas. No equality $E_D = K_D$, optimal constant, or exact positive depolarizing quantum/private capacity is claimed.

4 Oscillator model, prior work and conventions

Let $\mathcal{H} = \ell^2(\mathbb{N}_0)$ and $H = \sum_{j \geq 0} j|j\rangle\langle j|$. Fix $0 \leq \lambda \leq 1$ and a density operator τ , initially without energy or entropy restrictions. When finite, write $M = \text{Tr } H\tau$. The memoryless reset channel is

$$\mathcal{R}_{\lambda, \tau}(\rho) = \lambda\rho + (1 - \lambda)\tau \text{Tr } \rho. \quad (17)$$

Here λ is a survival probability, not a beam-splitter transmissivity. This is not the Gaussian thermal-loss channel. No assumption $[H, \tau] = 0$ is made. The comparison erasure channel is

$$\mathcal{E}_\lambda(\rho) = \lambda\rho \oplus (1 - \lambda) \text{Tr } \rho |e\rangle\langle e|,$$

with an additional orthogonal flag dimension. Entropies are in bits, except for an explicit $\ln$. Put

$$g(N) = (N + 1) \log_2(N + 1) - N \log_2 N,$$

interpreting zero terms by continuity. The Gibbs variational principle gives $\sup_{\text{Tr } H\rho \leq N} S(\rho) = g(N)$, attained by the geometric distribution.

Bennett, DiVincenzo and Smolin established finite-dimensional erasure capacities [1]. Zhong, Oh and Jiang studied continuous-variable erasure with an external flag, explicitly distinguishing it from replacement by the vacuum without a flag [6]. Catalano et al. recently studied mixtures of bosonic pure-loss channels [7]. Their Section IV, equations (37)–(42), treats exactly the identity/vacuum-reset case and optimizes its entanglement-assisted capacity. Section VI.C, equation (48), gives coherent information $(2\lambda - 1)S(\rho)$ on inputs orthogonal to the vacuum. That formula, together with the erasure upper bound, already yields the vacuum case of our quantum asymptotic statement. Their structural results also include non-entanglement-breaking behavior of nontrivial fading channels. These special cases are not claimed as new. Strong classical additivity of the finite-dimensional erasure channel is also established, for example by Amosov–Mancini [8, Theorem 9]; our partner statement transfers an erasure upper bound to reset with a vanishing defect. Catalano et al. also study unconstrained two-way capacities of finite pure-loss mixtures [7, Section III.D, Theorem 1], including their divergence in the presence of an identity component. We do not claim that divergence or the vacuum-reset mechanism as new.

The hidden-branch estimate below proves four absolute-gap limits without reset energy or entropy assumptions. Theorem 4 then gives two further assisted limits and state-entanglement reversibility under the same unrestricted fixed-state assumption. The coding, hashing and pure-state dilution dependencies are established [4, 5, 3, 10, 9, 11].

Operational conventions. For n uses let $H^{(n)} = \sum_i H_i$. The input allowance is mean total energy at most nN , averaged over encoded inputs. $C(\mathcal{N}, N)$ is classical capacity with vanishing decoding error, no prior entanglement and no feedback. $C_{\text{EA}}(\mathcal{N}, N)$ allows unlimited prior entanglement, with the cost on channel inputs. $Q(\mathcal{N}, N)$ is entanglement-transmission capacity with the cost on the encoded maximally mixed logical state. $P(\mathcal{N}, N)$ is unassisted private classical transmission, with no preshared key or free public discussion. Messages are equiprobable; decoding error is averaged over messages. For Q the entanglement fidelity tends to one. For P the average trace distance of each full-environment state from a common fixed state tends to zero. The security environment is that of a full dilation of the channel, not just the mathematical reset-branch register. Wilde–Qi’s operational equivalences [3, Theorem 1] apply to a general channel and energy observable; they identify these quantum and private conventions with the corresponding uniform-energy quantum and secret-key transmission tasks. The finite-output-entropy coding formulas in their Theorems 2–3 will be used only on the finite decoded channel or the erasure comparison channel. Their secret-key transmission task does not allow free two-way communication. The allowance is across the block; there is no peak constraint on each individual use. We consider vanishing-error capacities, not strong-converse rates at a fixed nonzero error.

Theorem 11. *For every fixed λ and τ as above, as $N \rightarrow \infty$,*

$$C(\mathcal{R}_{\lambda,\tau}, N) = \lambda g(N) + o(1), \quad (18)$$

$$C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, N) = 2\lambda g(N) + o(1), \quad (19)$$

$$Q(\mathcal{R}_{\lambda,\tau}, N) = (2\lambda - 1)_+ g(N) + o(1), \quad (20)$$

$$P(\mathcal{R}_{\lambda,\tau}, N) = (2\lambda - 1)_+ g(N) + o(1). \quad (21)$$

Each leading term is an upper bound. The remainders for Q and P can differ; no equality at finite energy is claimed. For $\lambda \leq 1/2$, $Q = P = 0$ exactly for every N . No finiteness of $S(\tau)$ or $\text{Tr } H\tau$ is required. The convergence is uniform over τ with $\text{Tr } H\tau \leq M_0$ for each fixed finite M_0 , but no uniformity over all states is claimed.

The absolute error is stronger than matching only a coefficient of $\log N$. One family of finite-rank inputs proves all four limits. We obtain achievable rates through coding theorems, not an efficient finite-block quantum coding algorithm.

5 An exact erasure approximation

For integers $K \geq 1, d \geq 2$, let $V|j\rangle = |K + j\rangle$ for $0 \leq j < d$, and $\Pi = VV^*$. Define the deterministic decoder

$$D(X) = V^*\Pi X \Pi V \oplus \text{Tr}[(I - \Pi)X]|e\rangle\langle e|. \quad (22)$$

It measures membership in the code subspace, preserving quantum information inside it, without postselecting away failed uses. Set

$$\tau_\Pi = V^*\tau V, \quad \epsilon = \text{Tr } \Pi\tau, \quad \delta = (1 - \lambda)\epsilon, \quad \mathcal{V}(\rho) = V\rho V^*. \quad (23)$$

Denote the d -dimensional erasure channel by $\mathcal{E}_\lambda^{(d)}$.

Lemma 12. *The decoded channel satisfies*

$$D\mathcal{R}_{\lambda,\tau}\mathcal{V}(\rho) = [\lambda\rho + (1-\lambda)\tau_\Pi \text{Tr}\rho] \oplus (1-\lambda)(1-\epsilon) \text{Tr}\rho |e\rangle\langle e|,$$

and

$$\frac{1}{2}\|D\mathcal{R}_{\lambda,\tau}\mathcal{V} - \mathcal{E}_\lambda^{(d)}\|_\diamond = \delta, \quad \epsilon \leq \text{Tr}[\mathbf{1}_{\{H \geq K\}}\tau]. \quad (24)$$

Proof. Substitute (17) in (22). The difference from erasure is $X \mapsto \text{Tr}(X)A$, where $A = (1-\lambda)(\tau_\Pi \oplus -\epsilon|e\rangle\langle e|)$. This replacement map has diamond norm $\|A\|_1$: partial trace has completely bounded trace norm one and a density operator attains the norm. The positive and negative blocks have equal absolute trace, so $\|A\|_1 = 2\delta$. Since $\Pi \leq \mathbf{1}_{\{H \geq K\}}$, the tail bound follows. When $M < \infty$, $H \geq K\Pi$ also gives $\epsilon \leq M/K$. $\square$

The distance includes arbitrary references. Reset coherences between Π and its complement are removed by D , so no square-root loss from gentle measurement is needed.

6 A dimension-free information-loss bound

Recall $b(\lambda, \delta)$ from (1), for $0 \leq \delta \leq 1 - \lambda$. It is the remaining uncertainty about a binary branch after observing whether the output is in the code subspace.

Lemma 13. *Let σ be a d -dimensional state of entropy s , let ψ_{RB} purify it, and let $\omega_{RB} = (\text{id} \otimes D\mathcal{R}_{\lambda,\tau}\mathcal{V})(\psi)$. Then*

$$I(R : B)_\omega \geq 2\lambda s - b(\lambda, \delta), \quad I_c(\sigma, D\mathcal{R}_{\lambda,\tau}\mathcal{V}) \geq (2\lambda - 1)s - b(\lambda, \delta). \quad (25)$$

For a diagonal input ensemble and a number measurement after D , the classical mutual information is at least $\lambda s - b(\lambda, \delta)$.

Proof. Introduce a classical register F recording transmit or reset:

$$\omega_{RBF} = \lambda\psi_{RB} \otimes |T\rangle\langle T| + (1-\lambda)\sigma_R \otimes D(\tau)_B \otimes |Z\rangle\langle Z|.$$

The reference is independent of F , so $I(R : BF) = 2\lambda s$. The chain rule gives $I(R : B) = 2\lambda s - I(R : F|B)$. Since F is classical, $S(F|RB) \geq 0$, and hence $I(R : F|B) \leq S(F|B)$. Measure only code membership at B , producing the binary variable Y . Data processing gives $S(F|B) \leq H(F|Y)$. An erasure output identifies reset; an accepted output has probability $p = \lambda + \delta$ and conditional reset probability $q = \delta/p$. Thus $H(F|Y) = ph(q) = b(\lambda, \delta)$. Finally $I_c = I(R : B) - S(R)$. For classical symbols X , the branch register is again independent of the input and $I(X : Y_{\text{full}}, F) = \lambda s$. The same chain rule, using membership as a coarsening of the full number outcome, proves the classical bound. All registers in this proof are finite dimensional. $\square$

The bound does not scale with $\log d$. In particular, $0 \leq b(\lambda, \delta) \leq h(\delta)$: the rule that guesses transmission on acceptance has error probability δ , so its error bit bounds $H(F|Y)$. The function is nondecreasing in δ , since $\partial_\delta b = \log_2(1 + \lambda/\delta)$ when $\delta > 0$. Ordinary conditional-entropy continuity [2] would instead introduce a term proportional to $\delta \log d$.

Proposition 14. *For a diagonal $\sigma = \sum_{j=0}^{d-1} p_j |j\rangle\langle j|$ of entropy s , suppose $K + \sum_j j p_j < N$. Then*

$$C(\mathcal{R}_{\lambda,\tau}, N) \geq \lambda s - b(\lambda, \delta), \quad (26)$$

$$Q(\mathcal{R}_{\lambda,\tau}, N) \geq \max\{0, (2\lambda - 1)s - b(\lambda, \delta)\}, \quad (27)$$

$$P(\mathcal{R}_{\lambda,\tau}, N) \geq Q(\mathcal{R}_{\lambda,\tau}, N), \quad (28)$$

$$C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, N) \geq 2\lambda s - b(\lambda, \delta). \quad (29)$$

Proof. Apply Lemma 13 and the coding theorems with strict cost slack to the finite input and finite decoded output, with bounded cost $V^*HV = \text{diag}(K, \dots, K + d - 1)$. The finite-alphabet classical coding theorem proves (26); coherent-information achievability [3, Theorem 2] proves (27); and the entanglement-assisted theorem [5, Theorem 1] proves (29). A quantum code yields a private code. Every such code is implemented on the physical channel by V and D . This proves the lower bounds even when the physical reset output has infinite entropy, without invoking a finite-output-entropy formula on that channel itself. $\square$

Proposition 15. *For every $N \geq 0$,*

$$C(\mathcal{R}_{\lambda, \tau}, N) \leq \lambda g(N), \quad C_{\text{EA}}(\mathcal{R}_{\lambda, \tau}, N) \leq 2\lambda g(N), \quad (30)$$

$$Q(\mathcal{R}_{\lambda, \tau}, N) \leq (2\lambda - 1)_+ g(N), \quad P(\mathcal{R}_{\lambda, \tau}, N) \leq (2\lambda - 1)_+ g(N). \quad (31)$$

Proof. Reset is a postprocessing of $\mathcal{E}_\lambda$: keep the transmitted block and replace the orthogonal flag by τ . Every capacity is therefore bounded by that of $\mathcal{E}_\lambda$ with the same input cost.

For a classical n -use ensemble, condition on the independent surviving subset J . The Holevo information in that branch is at most $\sum_{i \in J} S(\bar{\rho}_i)$. Its average is at most $\lambda \sum_i S(\bar{\rho}_i) \leq n\lambda g(N)$. Geometric number ensembles approach equality. For entanglement assistance, $I(\rho, \mathcal{E}_\lambda) = 2\lambda S(\rho)$; apply [5, Theorem 1].

For $\lambda \geq 1/2$, erasure is degradable and has coherent information $(2\lambda - 1)S(\rho)$. Wilde–Qi’s Theorem 3 [3] makes its energy-constrained quantum and private capacities equal to optimized coherent information. The Gibbs condition holds since $\text{Tr } e^{-\beta H} = (1 - e^{-\beta})^{-1} < \infty$ for $\beta > 0$. Erasure output entropy is bounded by $h(\lambda) + \lambda g(N)$. This verifies the finite-output-entropy condition on erasure, the channel to which that theorem is applied. Its complementary output entropies are finite by purification and subadditivity. No such condition is required of the postprocessed reset channel. For $\lambda \leq 1/2$, erasure is antidegradable: the environment can simulate the receiver, forcing its unassisted private and quantum capacities to zero. Postprocessing preserves this conclusion. Both arguments agree at $\lambda = 1/2$. $\square$

7 Entropy-preserving escape

A uniform distribution on a cutoff subspace loses a nonvanishing constant relative to $g(N)$. Instead, for sufficiently large N choose

$$\begin{aligned} K &= \lceil \sqrt{N} \rceil, \quad m = N - K > 0, \quad r = \frac{m}{m+1}, \\ d &= \lceil 2(m+1) \ln(N+1) \rceil, \quad t = r^d, \quad p_j = \frac{(1-r)r^j}{1-r^d} \quad (0 \leq j < d). \end{aligned} \quad (32)$$

Lemma 16. *For this distribution,*

$$K + \sum_j j p_j = K + m - \frac{dt}{1-t} < N, \quad (33)$$

$$s_N = g(m) + \log_2(1-t) + \frac{dt}{1-t} \log_2 r. \quad (34)$$

Also $t \leq (N+1)^{-2}$ and $a_N := g(N) - s_N = O(N^{-1/2}) + O(\log N/N^2) = o(1)$.

Proof. Differentiate the finite geometric sum to obtain the mean. Substitute it into $-\sum_j p_j [\log_2(1-r) - \log_2(1-t) + j \log_2 r]$ to obtain the entropy. Since $\ln r = -\ln(1+1/m) \leq -1/(m+1)$, the stated tail bound follows. The shift loss is at most

$$g(N) - g(N-K) = \int_{N-K}^N \log_2(1+1/x) dx \leq K \log_2(1+1/(N-K)) = O(N^{-1/2}).$$

Finally $d|\ln r| = O(\ln N)$, so the other entropy terms have absolute sum $O(t(1+d|\ln r|))$. $\square$

Proof of Theorem 11. With $\delta_N = (1-\lambda) \text{Tr} \Pi_N \tau$, the preceding propositions give

$$0 \leq \lambda g(N) - C(\mathcal{R}_{\lambda, \tau}, N) \leq \lambda a_N + b(\lambda, \delta_N), \quad (35)$$

$$0 \leq 2\lambda g(N) - C_{\text{EA}}(\mathcal{R}_{\lambda, \tau}, N) \leq 2\lambda a_N + b(\lambda, \delta_N). \quad (36)$$

For $\lambda > 1/2$, the gap of either Q or P below the common erasure upper bound is at most

$$(2\lambda - 1)a_N + b(\lambda, \delta_N). \quad (37)$$

Since τ is trace class and $K \rightarrow \infty$, $\delta_N \leq (1-\lambda) \text{Tr}[\mathbf{1}_{\{H \geq K\}} \tau] \rightarrow 0$. The binary loss $b(\lambda, \delta_N)$ therefore vanishes without a dimension-dependent multiplier. If $M < \infty$, then $\delta_N \leq (1-\lambda)M/K = O(M/\sqrt{N})$. Monotonicity of b permits the replacement $\delta_N \mapsto (1-\lambda) \min\{1, M/K\}$. The resulting bounds are $O_M(\log N/\sqrt{N})$ for fixed positive M ; when $M = 0$ the branch loss is zero. The constants are uniform for $M \leq M_0$. For $\lambda \leq 1/2$ the quantum and private claims are already exact. The endpoints $\lambda = 0, 1$ are the replacer and identity channels. $\square$

Specific reset tails improve the rate. For thermal τ of mean $M > 0$, $\text{Tr} \Pi \tau \leq [M/(M+1)]^K$. Taking $K = \lceil c \ln(N+1) \rceil$ with $c > 2/\ln(1+1/M)$, and the other choices in (32) with $m = N-K$, gives all four gaps $O_M(\log N/N)$; (40) below also gives this rate for both two-way gaps. This follows from the same estimates; optimal rates are not asserted.

For each fixed N we first take coding block length to infinity. Only afterwards do we take $N \rightarrow \infty$; no coding theorem uniform in the growing dimension $d(N)$ is required.

8 Two-way reset capacities and reversible distributed states

Q_2 and K_2 for the oscillator use the countably decomposable LOCC convention of [9], with free authenticated public two-way classical discussion and no initial entanglement or secret key. For adaptive n -use protocols the energy constraint is $n^{-1} \sum_i \text{Tr} H \rho_{\text{input}}^{(i)} \leq N$, averaging each marginal over previous outcomes [9, equation (101)]. Security includes the full channel environment and all public messages.

Theorem 17. *For every fixed oscillator density operator τ , with no entropy or energy restriction,*

$$Q_2(\mathcal{R}_{\lambda, \tau}, N) = \lambda g(N) + o(1), \quad K_2(\mathcal{R}_{\lambda, \tau}, N) = \lambda g(N) + o(1). \quad (38)$$

Each leading term is an upper bound; the remainders need not agree. For every finite encoding in Proposition 14,

$$Q_2(\mathcal{R}_{\lambda, \tau}, N) \geq \max\{0, \lambda s - b(\lambda, \delta) - \delta \log_2 5\}, \quad K_2 \geq Q_2. \quad (39)$$

Proof. For the upper bound use reset as receiver postprocessing of erasure. Davis–Shirokov–Wilde [9, Theorem 4 and equation (141)] bound both two-way capacities of erasure by its energy-constrained squashed entanglement. Its Gibbs and finite-output-entropy hypotheses were verified in Proposition 15. For a purified input ρ , retaining the erasure branch in a classical extension gives conditional mutual information $2\lambda S(\rho)$; hence this squashed entanglement is at most $\lambda g(N)$. This covers adaptive average-cost protocols and is applied to erasure even when reset has infinite output entropy.

For the lower bound, transmit halves of the finite input purification and apply D . The resulting state has the form of Theorem 4, with $\chi = \sigma_A \otimes (\tau_\Pi/\epsilon)$ if $\epsilon > 0$, and a pure useful component with $u = s$. This yields (39); when $\epsilon = 0$ it is exactly flagged erasure. The seed also comes from copies of this same finite-energy input, so no uncounted input cost or initial entanglement is introduced. For (32), $s_N = g(N) - a_N$, $a_N \rightarrow 0$ and $\delta_N \rightarrow 0$ for every fixed trace-class τ . Both loss terms in (39) vanish independently of its entropy, proving the two limits separately. The endpoints $\lambda = 0, 1$ are respectively replacement and identity. $\square$

For $M < \infty$, put $\bar{\epsilon}_N = \min\{1, M/K\}$ and $\bar{\delta}_N = (1 - \lambda)\bar{\epsilon}_N$. Each gap is bounded by

$$\lambda a_N + b(\lambda, \bar{\delta}_N) + \bar{\delta}_N \log_2 5, \quad (40)$$

since b is increasing in its second argument. This is $O_{M_0}(\log N/\sqrt{N})$ uniformly for $M \leq M_0 < \infty$. No rate uniform over all reset states is claimed.

Corollary 18. *Let ψ_N purify the encoded input in (32) and $\Omega_N = (\text{id} \otimes \mathcal{R}_{\lambda, \tau})(\psi_N)$. For every fixed τ ,*

$$E_D(\Omega_N) = \lambda g(N) + o(1), \quad E_C(\Omega_N) = \lambda g(N) + o(1), \quad E_C(\Omega_N) - E_D(\Omega_N) = o(1).$$

The same holds after applying D locally to the receiver.

Proof. Local decoding and Theorem 4 give $E_D(\Omega_N) \geq \lambda s_N - b(\lambda, \delta_N) - \delta_N \log_2 5$. The decomposition $\Omega_N = \lambda \psi_N + (1 - \lambda)\sigma_A \otimes \tau$ gives $E_C(\Omega_N) \leq \lambda s_N$ by the same random-branch preparation and pure-state dilution. A density on a separable Hilbert space has a countable spectral decomposition, so local preparation of τ is allowed even at infinite entropy or energy. Decoding preserves this upper bound. Use $E_D \leq E_C$ and let $N \rightarrow \infty$. $\square$

Local operations and preparation are not energy constrained in this LOCC cost. This is a statement about the specified distributed states, not the cost of simulating the channel. For each N the operational block length tends to infinity first. No finite- N reversibility is asserted.

9 Uniform classical additivity defect

For finite-dimensional Ψ , let $C(\mathcal{R}_{\lambda, \tau} \otimes \Psi, N)$ mean n parallel pairs, with mean oscillator energy at most N per pair and no extra constraint on the finite-dimensional input.

Theorem 19. *For every finite-dimensional Ψ ,*

$$\begin{aligned} 0 &\leq C(\mathcal{R}_{\lambda, \tau} \otimes \Psi, N) - C(\mathcal{R}_{\lambda, \tau}, N) - C(\Psi) \\ &\leq \lambda g(N) - C(\mathcal{R}_{\lambda, \tau}, N). \end{aligned}$$

The supremum of this defect over all finite-dimensional partners tends to zero as $N \rightarrow \infty$.

Proof. It suffices to show $C(\mathcal{E}_\lambda \otimes \Psi, N) \leq \lambda g(N) + C(\Psi)$. For finite-dimensional $\mathcal{M}$, define its convex output-entropy roof

$$\widehat{S}_{\mathcal{M}}(\omega) = \inf_{\omega = \sum_x p_x |\psi_x\rangle\langle\psi_x|} \sum_x p_x S(\mathcal{M}(|\psi_x\rangle\langle\psi_x|)).$$

If B is finite dimensional and A is possibly infinite dimensional,

$$S((\text{id}_A \otimes \mathcal{M})(\omega_{AB})) \geq \widehat{S}_{\mathcal{M}}(\omega_B). \quad (41)$$

For pure ω_{AB} , the left side is $S(\mathcal{M}^c(\omega_B))$. Concavity on any pure-state decomposition of ω_B , and equality of output and complementary entropies on pure inputs, give the lower bound. For mixed ω_{AB} , use its countable spectral decomposition, entropy concavity and convexity of the finite-dimensional roof. The application below has finite entropies by the energy bound.

Fix n and $\mathcal{M} = \Psi^{\otimes n}$, always finite dimensional. For an ensemble $\{p_x, \rho_{A^n B}^x\}$ and surviving subset J , the output is $(\text{id}_{A_J} \otimes \mathcal{M})(\rho_{A_J B}^x)$, with entropy at least $\widehat{S}_{\mathcal{M}}(\rho_B^x)$. Hence each full output has entropy at least $nh(\lambda) + \widehat{S}_{\mathcal{M}}(\rho_B^x)$. Subadditivity bounds the average output entropy by

$$nh(\lambda) + \lambda \sum_i S(\bar{\rho}_{A_i}) + S(\mathcal{M}(\bar{\rho}_B)).$$

Subtract the conditional entropies and use roof convexity to obtain

$$\begin{aligned} \chi &\leq \lambda \sum_i S(\bar{\rho}_{A_i}) + S(\mathcal{M}(\bar{\rho}_B)) - \widehat{S}_{\mathcal{M}}(\bar{\rho}_B) \\ &\leq n\lambda g(N) + \chi(\Psi^{\otimes n}). \end{aligned}$$

The constrained coding theorem and regularization [4] yield the desired upper bound. This direct subset argument avoids iteration over infinite-dimensional partner channels. Postprocessing gives the reset upper bound, and independent product coding gives the lower bound $C(\mathcal{R}_{\lambda, \tau} \otimes \Psi, N) \geq C(\mathcal{R}_{\lambda, \tau}, N) + C(\Psi)$. Subtract and use Theorem 11. $\square$

This is an asymptotic classical statement, not exact additivity at finite energy. No quantum or private arbitrary-partner additivity is asserted.

10 A sharp spectral criterion for absolute recovery

The oscillator construction extends to a larger class of Hamiltonians, but not to every Hamiltonian with finite entropy at finite energy. In this section H has a countably infinite eigenbasis, eigenvalues $0 = e_0 \leq e_1 \leq \dots \rightarrow \infty$, finite multiplicities, and $Z(\beta) := \text{Tr} 2^{-\beta H} < \infty$ for at least one $\beta > 0$. Define

$$\beta_c = \inf\{\beta \geq 0 : Z(\beta) < \infty\}, \quad Z_c = \lim_{\beta \downarrow \beta_c} Z(\beta), \quad F_H(E) = \sup_{\text{Tr } H\rho \leq E} S(\rho).$$

Thus $\beta_c < \infty$ and $1 < Z_c \leq \infty$. The base of the exponential fixes the units of β ; all entropies remain in bits. The input, error and security conventions are those above, with H replacing the number operator. In particular, the two-way convention is countably decomposable LOCC, with expected block-average input cost. Fix $0 < \lambda < 1$ and any density operator τ ; neither its energy nor its entropy is required to be finite.

Theorem 20 (Critical-partition dichotomy). *If $Z_c = \infty$, then for every fixed τ ,*

$$X(\mathcal{R}_{\lambda,\tau}, H, E) = r_X F_H(E) + o(1), \quad E \rightarrow \infty, \quad (42)$$

simultaneously for

$$(X, r_X) = (C, \lambda), (C_{\text{EA}}, 2\lambda), (Q, (2\lambda - 1)_+), (P, (2\lambda - 1)_+), (Q_2, \lambda), (K_2, \lambda).$$

Each benchmark is an upper bound. If $Z_c < \infty$, put $\gamma_c = 2^{-\beta_c H}/Z_c$. Then, for every fixed τ ,

$$2\lambda F_H(E) - C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, E) \rightarrow \Delta_{\lambda,\tau}(H) > 0, \quad (43)$$

$$\Delta_{\lambda,\tau}(H) = \inf_{\rho: \text{Tr } H\rho < \infty} \{2\lambda D(\rho \| \gamma_c) + e_{\lambda,\tau}(\rho)\}, \quad (44)$$

$$0 < \Delta_{\lambda,\tau}(H) \leq e_{\lambda,\tau}(\gamma_c) \leq h(\lambda). \quad (45)$$

Here $e_{\lambda,\tau}$ is the finite binary-flag loss defined in Lemma 23. In particular, for each fixed reset state, absolute entanglement-assisted recovery holds if and only if $Z_c = \infty$. No finite- Z_c formula for Q, P, Q_2, K_2 is asserted.

The critical Gibbs entropy description used below is established: Shirokov [28, Proposition 1a] gives the variational formula, critical slope, linear regime and possible nonattainment. We use that description to distinguish two operational capacity regimes. The general energy-capacity formulas of [3, 9] with an all-temperature Gibbs hypothesis cannot simply be applied here, since β_c may be positive. We give the erasure converse directly.

10.1 Entropy escape and the erasure benchmark

Lemma 21. *Under the stated spectral assumptions,*

$$\frac{F_H(E)}{E} \rightarrow \beta_c, \quad F_H(E) - \beta_c E \nearrow \log_2 Z_c. \quad (46)$$

If $Z_c = \infty$, there exist diagonal finite-rank states σ_E and their support projections P_E such that

$$\text{Tr } H\sigma_E < E, \quad S(\sigma_E) = F_H(E) - o(1), \quad \text{Tr } P_E \tau \rightarrow 0 \quad (47)$$

for every fixed trace-class density τ .

Proof. For completeness, the relevant entropy argument can be expressed by the Gibbs duality

$$\log_2 Z(\beta) = \sup_{E \geq 0} \{F_H(E) - \beta E\}.$$

The finite concave nondecreasing function F_H has an asymptotic slope. Gibbs' bound makes it at most β_c . A strictly smaller slope would make the displayed supremum finite for some $\beta < \beta_c$, a contradiction. All secant slopes dominate the asymptotic slope, so $F_H(E) - \beta_c E$ is nondecreasing. Its supremum is $\log_2 Z_c$, proving (46). This reasoning does not assume that the entropy supremum is attained.

Remove the first m energy eigenvectors, and write $Z_{>m}$ and $F_{>m}$ for the partition sum and entropy supremum on the remaining space. For fixed m and sufficiently large E , its Gibbs distribution has mean E at an inverse temperature $\beta_m(E) > \beta_c$. Its mean tends to infinity as $\beta \downarrow \beta_c$: otherwise integrating the bounded mean, the negative derivative of $\log Z_{>m}$, would contradict the divergence of $Z_{>m}$. Thus $\beta_m(E) \downarrow \beta_c$. Evaluating the full Gibbs bound at that temperature gives

$$0 \leq F_H(E) - F_{>m}(E) \leq \log_2 \frac{Z(\beta_m(E))}{Z_{>m}(\beta_m(E))} \rightarrow 0.$$

The removed partition sum is bounded and the retained sum diverges. Choose $m = m(E) \rightarrow \infty$ slowly by a diagonal argument. Upper truncation of each tail Gibbs distribution preserves its entropy in the limit and lowers its mean energy. If necessary, first decrease the chosen mean by a positive amount tending sufficiently quickly to zero; continuity of the finite concave entropy supremum supplies strict energy slack with $o(1)$ entropy loss. The resulting finite spectral window lies above its first $m(E)$ vectors. Consequently $\text{Tr } P_E \tau \leq \text{Tr } \mathbf{1}_{\{j > m(E)\}} \tau \rightarrow 0$ for every fixed τ . There is no uniform rate over all density operators. $\square$

Lemma 22 (Erasure with a single finite partition sum). *For the six operational conventions of Theorem 20,*

$$X(\mathcal{E}_\lambda, H, E) = r_X F_H(E). \quad (48)$$

Proof. For $E > 0$, finite-support entropy approximants with strict cost slack, followed by finite-dimensional erasure coding, give achievability. The order is: fix the finite approximation, take block length to infinity, then improve the approximation. For two-way transmission the receiver announces the surviving copies and distills their pure entanglement. No entropy-maximizing infinite-dimensional code state is required. At $E = 0$, use the finite ground eigenspace directly.

For the converse, any admissible n -use average input has $S(\rho_{A^n}) \leq \sum_i S(\rho_{A_i}) \leq nF_H(E) < \infty$. The erasure outputs and complementary outputs also have finite entropy, bounded by this sum plus $nh(\lambda)$. Conditioning on the independent erasure pattern bounds the output Holevo quantity by $\lambda \sum_i S(\rho_{A_i})$. Fano's inequality on the finite message alphabet proves the C converse. For assistance, $I(\rho, \mathcal{E}_\lambda) = 2\lambda S(\rho)$; mutual-information subadditivity gives the corresponding n -use bound. The general constrained entanglement-assisted formula [5, Theorem 1] applies to the present positive observable.

For $\lambda \leq 1/2$ erasure is antidegradable, so $Q = P = 0$. For $\lambda > 1/2$ let B^n and W^n be its output and full environment. Product degradability and data processing of total correlation give

$$\begin{aligned} I_c(\rho, \mathcal{E}_\lambda^{\otimes n}) &= \sum_i I_c(\rho_{A_i}, \mathcal{E}_\lambda) - \text{TC}(B^n) + \text{TC}(W^n) \\ &\leq (2\lambda - 1) \sum_i S(\rho_{A_i}). \end{aligned}$$

Here $\text{TC}(B^n) = \sum_i S(B_i) - S(B^n)$, and similarly for W^n . All these entropy expressions are finite. If a quantum encoder has an additional purifying system F , then $S(F) \leq S(R) + S(A^n) < \infty$ for its finite logical reference R . The complementary channel is antidegradable, so FW^n has a two-extension on W^n . Weak monotonicity gives $S(F|W^n) \geq 0$, and purity of the full dilation gives $I(R)B^n = I_c(\rho, \mathcal{E}_\lambda^{\otimes n}) - S(F|W^n)$. Coherent-information monotonicity under decoding and continuity on the finite logical reference prove the quantum converse. For any finite private ensemble, its private information is $I_c(\bar{\rho}) - \sum_x p_x I_c(\rho_x)$. Degradability makes each subtracted coherent information nonnegative; classical Fano and secrecy continuity on the finite message register give the same private converse. None of these arguments uses uniform entropy continuity on the whole energy ball.

For two-way protocols consider one input system A held by Alice, with her remaining memory L and Bob's memory B . Transferring A to Bob can increase squashed entanglement by at most $S(A)$. Indeed the extended conditional-mutual-information chain rule [29, equation (14)] gives, for any extension W ,

$$I(L; AB|W) \leq I(LA; B|W) + 2S(A).$$

Choose an extension within any positive tolerance of the pre-use squashed entanglement. Starting from zero, this quantity remains finite by induction, and $S(A)$ is finite by the input cost, so there is no subtraction of infinite quantities. Failure is a local discard and cannot increase squashed

entanglement. Store the independent success flag in the extension; the expected increment is at most $\lambda S(A)$. Average over public histories and apply entropy concavity and Jensen's inequality to obtain

$$E_{\text{sq}}(\text{final state}) \leq \lambda n F_H(E).$$

We use squashed entanglement in infinite dimensions [31, Proposition 2B] and the approximate-private-state bound [32, Theorem 2]. Retain the local purifiers as shields when representing the key protocol by an approximate private state; Eve includes the public transcript. The bound is applied to that state, before measuring its key registers. Its error terms depend on the finite key size, not on shield dimension. With vanishing error it bounds the secret-key rate, and therefore the entanglement rate, by $\lambda F_H(E)$. The operations retain the stated countably decomposable convention. $\square$

Proof of (42). Decode the finite spectral window of Lemma 21 by projection and an orthogonal failure flag. The accepted reset weight is $\delta_E = (1 - \lambda) \text{Tr } P_E \tau \rightarrow 0$ and its binary ambiguity cost is $b_E = (\lambda + \delta_E) h(\lambda/(\lambda + \delta_E)) \rightarrow 0$. The finite-channel results proved above give respectively the lower bounds

$$\lambda S(\sigma_E) - b_E, \quad 2\lambda S(\sigma_E) - b_E, \quad [(2\lambda - 1)S(\sigma_E) - b_E]_+,$$

for $C, C_{\text{EA}}, Q/P$, and $[\lambda S(\sigma_E) - b_E - \delta_E \log_2 5]_+$ for Q_2/K_2 . The paid seed uses copies of the same strictly admissible input. Reset is postprocessing of erasure, so Lemma 22 supplies all upper bounds. Substitution of $S(\sigma_E) = F_H(E) - o(1)$ proves the result. For $\lambda \leq 1/2$, the quantum and private capacities are identically zero. $\square$

10.2 A finite flag and a positive variational defect

The necessity statement is stronger than a counterexample using a vacuum reset: every fixed reset state has a positive limiting entanglement-assisted deficit when Z_c is finite.

Lemma 23. *For any density operators ρ, τ , choose a purification ψ_ρ with B marginal ρ , and let*

$$\Omega_{JRB} = \lambda |0\rangle\langle 0|_J \otimes \psi_\rho + (1 - \lambda) |1\rangle\langle 1|_J \otimes \rho_R \otimes \tau_B.$$

The quantity $e_{\lambda, \tau}(\rho) := I(J; R|B)_\Omega$ is well defined, continuous in trace norm in ρ and τ , and lies in $[0, h(\lambda)]$. If $S(\rho) < \infty$, then

$$I(\rho, \mathcal{R}_{\lambda, \tau}) = 2\lambda S(\rho) - e_{\lambda, \tau}(\rho). \quad (49)$$

For every faithful mixed state γ , $e_{\lambda, \tau}(\gamma) > 0$, even if $S(\gamma) = \infty$.

Proof. Equivalently, e is the difference of the Holevo quantities of the two binary ensembles $\{\lambda, \psi_\rho; 1 - \lambda, \rho_R \otimes \tau\}$ and $\{\lambda, \rho; 1 - \lambda, \tau\}$. Define those quantities through relative entropy. Both are finite and at most $h(\lambda)$, regardless of the state entropies. Data processing gives the asserted bounds. Keeping J in the channel output gives $I(R; BJ) = 2\lambda S(\rho)$; the chain rule then proves (49) when $S(\rho)$ is finite. Canonical purifications vary continuously in trace norm, by the square-root continuity of positive trace-class operators. Conditional mutual information is uniformly continuous when one unconditioned system is the two-dimensional J [29, Corollary 1, equation (25)]. This proves continuity of e without continuity of $S(\rho)$.

Suppose $e_{\lambda, \tau}(\gamma) = 0$. The binary Holevo data processing inequality under Tr_R is then an equality. Each component has finite relative entropy to its ensemble average, because that average dominates it by its positive prior weight. Equality in relative-entropy data processing therefore supplies one normal recovery channel $T : B \rightarrow RB$ with

$$T(\gamma) = \psi_\gamma, \quad T(\tau) = \gamma_R \otimes \tau.$$

Here we use the infinite-dimensional sufficiency theorem [30, Section 2.3, Theorem 2], restricted to the support of the ensemble average; no bounded density inverse is assumed. A channel mapping a faithful state to a pure state is constant: if $Q = I - \psi_\gamma$, then $\text{Tr } \gamma T^*(Q) = 0$ and positivity and faithfulness imply $T^*(Q) = 0$. Every output is consequently supported on that same one-dimensional subspace. Thus $T(\tau) = \psi_\gamma$, contrary to the mixed product state $\gamma_R \otimes \tau$. $\square$

Proof of (43)–(45). If $Z_c < \infty$, then $\beta_c > 0$ and γ_c is faithful. For every finite-energy input, entropy is finite and the Gibbs identity gives

$$D(\rho \| \gamma_c) = \beta_c \text{Tr } H\rho + \log_2 Z_c - S(\rho) < \infty.$$

The constrained entanglement-assisted formula [5, Theorem 1] and Lemma 23 imply

$$2\lambda F_H(E) - h(\lambda) \leq C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, E) \leq 2\lambda F_H(E).$$

Write this finite concave nondecreasing capacity as C_E . By (46), $C_E/E \rightarrow 2\lambda\beta_c$. Concavity implies that $C_E - 2\lambda\beta_c E$ is nondecreasing, so its limit equals its supremum. Optimizing the energy allowance down to the mean cost of each input yields

$$\begin{aligned} \sup_E (C_E - 2\lambda\beta_c E) &= \sup_{\text{Tr } H\rho < \infty} [I(\rho, \mathcal{R}_{\lambda,\tau}) - 2\lambda\beta_c \text{Tr } H\rho] \\ &= 2\lambda \log_2 Z_c - \Delta_{\lambda,\tau}(H). \end{aligned}$$

This identity uses suprema, not existence of an optimizing state. Combining it with (46) gives the limit.

If the infimum in (44) were zero, both its nonnegative summands would vanish along a sequence. Pinsker's inequality would give $\rho_n \rightarrow \gamma_c$ in trace norm. Continuity would then give $e_{\lambda,\tau}(\rho_n) \rightarrow e_{\lambda,\tau}(\gamma_c) > 0$, a contradiction. For the upper bound take normalized finite energy truncations $\gamma_{c,m}$. Their relative entropy is $-\log_2 \text{Tr}(P_m \gamma_c) \rightarrow 0$, and their flag loss converges to that of γ_c . This remains valid when the critical state has infinite mean energy or infinite entropy. Neither attainment of the variational infimum nor a positive lower bound uniform over all reset states is claimed. $\square$

10.3 An explicit permanent loss at arbitrarily large energy

For a unique ground state and the reset $\tau = |0\rangle\langle 0|$, the defect is a scalar optimization. Put $q = 1 - \lambda$ and define

$$\begin{aligned} a &= \lambda + qx, & \mu_\pm &= \frac{1}{2}(a \pm \sqrt{a^2 - 4\lambda qx(1-x)}), \\ f_\lambda(x) &= h(x) + h(\lambda(1-x)) - H(\mu_+, \mu_-, q(1-x)), \\ e_\lambda(x) &= 2\lambda h(x) - f_\lambda(x). \end{aligned} \tag{50}$$

For a diagonal input $x|0\rangle\langle 0| + (1-x)\sigma$, $I(\rho, \mathcal{R}_{\lambda,\tau}) = 2\lambda(1-x)S(\sigma) + f_\lambda(x)$. These eigenvalues and this decomposition are the established vacuum-reset calculation of Catalano et al. [7, Section VI.B and Appendix C, equations (C1)–(C11)]. For $E > 0$, applying it to the excited-space entropy supremum $F_{H,+}$ gives

$$C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, E) = \sup_{0 \leq x < 1} \{2\lambda(1-x)F_{H,+}(E/(1-x)) + f_\lambda(x)\}, \tag{51}$$

where x must meet the minimum excited-energy constraint; the actual $x = 1$ input gives $C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, 0) = 0$ at $E = 0$. To justify diagonal restriction without uniform energy-ball entropy continuity, average successively over phases on the first finitely many energy vectors. Covariance and

concavity do not decrease mutual information. The pinched inputs converge to the diagonal input, and their entropies increase to its finite diagonal entropy, by lower semicontinuity and the entropy upper bound on that diagonal. The same pinching argument applies to output and complement; the latter retains the displayed two-dimensional block and the remaining eigenvalues $q(1-x)\text{spec}(\sigma)$. Their limiting entropies are finite. Passing to the limit preserves the inequality.

If $Z_c < \infty$, let $Z_+ = Z_c - 1$ and $p_c = 1/Z_c$. The excited-space version of (46), substituted into (51), gives

$$\Delta_{\lambda,|0\rangle}(H) = \min_{0 \leq x \leq 1} \{2\lambda D_{\text{Bern}}(x||p_c) + e_\lambda(x)\}. \quad (52)$$

The remainder is uniform because $E/(1-x) \geq E$; the $x \rightarrow 1$ endpoint is approached through admissible $x < 1$, and is not a claim that the vacuum carries energy. Strict positivity can also be checked directly: dephase and coarsen the reference to its ground/excited label in the flag-loss expression. This gives

$$e_\lambda(x) \geq (q + \lambda x)h\left(\frac{q}{q + \lambda x}\right) - xh(\lambda) > 0 \quad (0 < x < 1).$$

The last expression is a classical conditional mutual information: at the vacuum output an excited reference forces failure, whereas a ground reference allows either branch. Compactness and the strictly interior p_c prove positivity in (52).

Proposition 24 (An exact constant-loss plateau). *Take one ground vector and, for each integer $n \geq 1$, a shell of energy $n^2 + n$ with multiplicity 2^{n^2} . Then $\beta_c = 1$, $Z_c = 2$, and, for the ground-state reset and every $E \geq 8$,*

$$C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, E) = 2\lambda E + \max_{0 \leq x \leq 1} f_\lambda(x), \quad (53)$$

$$2\lambda F_H(E) - C_{\text{EA}}(\mathcal{R}_{\lambda,\tau}, H, E) = 2\lambda - \max_x f_\lambda(x) > 0. \quad (54)$$

The scalar maximum is unique and interior. At $\lambda = 1/2$ the gap is numerically $0.409719932065909\dots$ bits per use. The capacity equality is a supremum over operational codes; an entropy-maximizing finite-energy input need not exist.

Proof. Here $Z(\beta) = 1 + \sum_{n \geq 1} 2^{(1-\beta)n^2 - \beta n}$, which establishes the critical parameters. The normalized excited critical Gibbs law has shell probabilities 2^{-n} , mean energy $\sum (n^2 + n)2^{-n} = 8$ and entropy 8. The full critical state has ground probability $1/2$, mean energy 4 and entropy 5. The Gibbs entropy suprema [28, Proposition 1a] are therefore

$$F_H(E) = E + 1 \quad (E \geq 4), \quad F_{H,+}(E) = E \quad (E \geq 8).$$

One can see the nonattained linear part explicitly: truncate the critical Gibbs law, then mix in a uniform state on a much higher shell m . Its weight is of order the extra energy divided by $m^2 + m$, while its entropy per added energy tends to $m^2/(m^2 + m) \rightarrow 1$. Let the upper shell tend to infinity and then improve the finite base truncation. The displayed upper bounds are approached. Above the critical mean, equality in the Gibbs bound would force the critical Gibbs state itself, which has the wrong energy; thus these are suprema rather than maxima. Substitution into (51) proves (53), and (52) proves the positive gap.

Finally $f_\lambda(x)$ is the mutual information of a two-level diagonal input. For any nontrivial mixture of two different such inputs, its concavity gap is $\chi_{\text{input}} + \chi_{\text{output}} - \chi_{\text{environment}} \geq \chi_{\text{output}} > 0$, by data processing and $\lambda > 0$. Hence it is strictly concave. It vanishes at both endpoints and is positive in the interior, as is already seen by measuring the reference in the binary basis. The reported numerical root is a diagnostic evaluation of this unique maximum, not a proof by sampling. $\square$

For contrast, a shell of energy n and multiplicity 2^n for each $n \geq 0$ has $\beta_c = 1$ but $Z_c = \infty$, and $F_H(E) = E + g(E)$. Removing the first m shells leaves the entropy supremum $E + g(E - m)$ for $E \geq m$; its deficit tends to zero. Thus a positive critical inverse temperature is not itself the obstruction. Normalizability at that temperature is decisive. These deliberately constructed spectra establish the mathematical boundary of the capacity law; no experimental Hamiltonian or thermodynamic phase transition is claimed.

11 Consequences and quantitative checks

Corollary 25. *For every $N > 0$, $Q(\mathcal{R}_{\lambda,\tau}, N) > 0$ and $P(\mathcal{R}_{\lambda,\tau}, N) > 0$ if and only if $\lambda > 1/2$.*

Proof. Only positivity needs proof. If $\lambda > 1/2$, the main theorem gives a finite E_0 and a positive quantum rate. Use those codes on a fixed fraction $0 < \theta < \min\{1, N/E_0\}$ of block positions and send vacuum on the rest. The rate remains positive and mean energy is below N . A quantum code yields a private code. $\square$

This uses block-average energy and does not assert a peak-energy result.

Corollary 26. *For every oscillator density operator τ , even without a finite mean energy assumption, $\mathcal{R}_{\lambda,\tau}$ is not entanglement breaking if $\lambda > 0$.*

Proof. Fix $d = 2$ and let $K \rightarrow \infty$. Trace class implies $\epsilon \rightarrow 0$. Apply the decoded channel to half of a normalized Bell pair. In the transmitted block, the expectation of the partial transpose on the normalized antisymmetric vector is $-\lambda/2 + \delta/4$, which is negative for large K . Thus the decoded channel has an entangled Choi state, impossible if the original channel were entanglement breaking. $\square$

The vacuum versions overlap [7]; these structural consequences are not presented as the main contribution.

Numerical checks. The original accompanying script checks 200 random finite-matrix cases, including reset coherences and entangled references. The largest discrepancies in the distance identity and erasure coherent information were $4.45 \cdot 10^{-16}$ and $2.34 \cdot 10^{-15}$. No tested finite-energy, continuity or classical lower bound failed. A separate new-seed script checks 500 finite-matrix cases for the hidden-branch, forward/reverse hashing and pinching inequalities, including pure reset states, zero overlap and survival endpoints. The largest negative residual in those lower-bound inequalities has magnitude below $1.2 \cdot 10^{-15}$, consistent with floating-point error. The partial-breeding diagnostics add 120 cases checking parity Kraus operators, the fixed-matrix entropy bound and survivor hashing. A further 240 cases test the full phase average and general-noise bounds, including 120 separable states with matching marginals. Their algebraic errors are below $4.1 \cdot 10^{-15}$ and no lower-bound violation exceeds rounding error. A further 1000 cases check the dominated-component seed algebra without matching marginals, including singular supports; 480 cases check mixed maximally correlated components and survivor entropies. Another 264 cases check invariant-state block entropies and finite phase averages, including explicit control families. These finite-matrix checks are not proofs of asymptotic coding, security, capacity, or exhaustive counterexample searches. The minimax supplement adds 65 exact rational checks and 45 numerical parameter cases, including an independent nonidentity Weyl construction, correlated-block spectra and the interval-width identity; its largest algebraic residual is below $2 \cdot 10^{-14}$. It does not test coding or privacy by simulation.

Table 1 gives conservative analytic gap bounds from (32) for $\lambda = 0.7$ and only the assumption $M \leq 10$. They bound absolute losses from the erasure benchmarks, not measured rates or optimized capacity estimates.

N	Classical	Quantum/private	EA classical	Two-way Q_2, K_2
10^2	0.988103	0.942327	1.094915	1.684682
10^3	0.448692	0.434617	0.481535	0.666373
10^4	0.190675	0.186325	0.200824	0.260333
10^5	0.075710	0.074336	0.078917	0.097685
10^6	0.028947	0.028514	0.029957	0.035913
10^8	0.003891	0.003848	0.003992	0.004587

Table 1: Analytic upper bounds on gaps, rounded upward to six decimals. A loose bound can exceed the benchmark itself, in which case a trivial zero-rate lower bound is stronger.

Limitations. The reset state is fixed; no uniform convergence over all density operators is claimed. The pure-component irreversibility theorem requires separable noise for its cost upper bound, but no matching marginal. A mixed maximally correlated component can have intrinsic cost irreversibility. The CLDUI/LDOI upper bounds apply to those invariant classes; only the distillation lower bound applies before twirling an arbitrary state. The argument does not cover reset states changing arbitrarily with N , general infinite-dimensional input-dependent noise, partial-loss fading components or Gaussian thermal attenuation. The separate finite-dimensional arbitrary-noise result in Section 3 concerns fixed memoryless channels and codes that can depend on them. It does not provide adversarial noise protection or evaluate the unknown depolarizing capacity functions. The oscillator result determines neither general finite- N capacities nor optimal error rates or second-order coding costs. The growing code dimension is an asymptotic construction, not an efficient experimental receiver. Exact finite-energy classical additivity for arbitrary reset states would require a separate argument.

Spectral checks. The spectral checks independently evaluate the vacuum joint-state spectrum on 200 diagonal inputs, test phase pinching on 200 complex inputs, and check 2,807 scalar positivity cases. Twenty-five parameter choices compare the two defect expressions; five plateau maxima are evaluated with 80-digit arithmetic. Forty-two finite high-shell mixtures check the nonattained entropy construction and fifteen cases check the contrasting divergent-partition example. These are finite numerical diagnostics, not proofs of the spectral dichotomy, capacity converses, infinite-dimensional recovery or priority. The analytic arguments above supply those mathematical claims, subject to their cited dependencies.

Provenance and reproducibility. This working paper was developed with AI assistance in exploration, drafting and computational checking. A model review is not human peer review or formal proof verification. The supplement contains the verification scripts and numerical records; external source PDFs and private review records are excluded. Erasure coding, the energy-constrained coding theorems, hashing and pure-state dilution are cited dependencies, distinct from the hidden-branch estimate, stopped-classification bound and their numerical checks. No model score is evidence of a theorem.

References

- [1] C. H. Bennett, D. P. DiVincenzo and J. A. Smolin, Capacities of quantum erasure channels, *Phys. Rev. Lett.* **78**, 3217–3220 (1997). [arXiv:quant-ph/9701015](https://arxiv.org/abs/quant-ph/9701015).

- [2] A. Winter, Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints, *Commun. Math. Phys.* **347**, 291–313 (2016). [arXiv:1507.07775](#).
- [3] M. M. Wilde and H. Qi, Energy-constrained private and quantum capacities of quantum channels, *IEEE Trans. Inf. Theory* **64**, 7802–7827 (2018). [arXiv:1609.01997](#).
- [4] A. S. Holevo, Entanglement-Assisted Capacities of Constrained Quantum Channels, *Theory Probab. Appl.* **48**(2), 243–255 (2004), [doi:10.1137/S0040585X97980415](#). [arXiv:quant-ph/0211170](#).
- [5] A. S. Holevo and M. E. Shirokov, On classical capacities of infinite-dimensional quantum channels, *Problems of Information Transmission* **49**, 15–31 (2013). [doi:10.1134/S003294601301002X](#). Preprint title: On the entanglement-assisted classical capacity of infinite-dimensional quantum channels, [arXiv:1210.6926](#).
- [6] C. Zhong, C. Oh and L. Jiang, Information transmission with continuous variable quantum erasure channels, [arXiv:2205.09711](#) (2022).
- [7] G. Catalano, M. Fanizza, F. A. Mele, G. De Palma and V. Giovannetti, Convex combinations of bosonic pure-loss channels, [arXiv:2604.26874v2](#) (30 September 2026).
- [8] G. G. Amosov and S. Mancini, The decreasing property of relative entropy and the strong superadditivity of quantum channels, [arXiv:0803.0452](#) (2008).
- [9] N. Davis, M. E. Shirokov and M. M. Wilde, Energy-constrained two-way assisted private and quantum capacities of quantum channels, *Phys. Rev. A* **97**, 062310 (2018). [arXiv:1801.08102v3](#).
- [10] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, [arXiv:quant-ph/0306078](#) (2003).
- [11] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, Concentrating partial entanglement by local operations, *Phys. Rev. A* **53**, 2046–2052 (1996). [arXiv:quant-ph/9511030](#).
- [12] E. Hostens, J. Dehaene and B. De Moor, Asymptotic adaptive bipartite entanglement-distillation protocol, *Phys. Rev. A* **73**, 062337 (2006). [arXiv:quant-ph/0602205v2](#).
- [13] K. G. H. Vollbrecht and F. Verstraete, Interpolation of recurrence and hashing entanglement distillation protocols, *Phys. Rev. A* **71**, 062325 (2005). [arXiv:quant-ph/0404111v2](#).
- [14] K. G. H. Vollbrecht and M. M. Wolf, Efficient distillation beyond qubits, *Phys. Rev. A* **67**, 012303 (2003). [arXiv:quant-ph/0208152](#).
- [15] M. Horodecki and P. Horodecki, Reduction criterion of separability and limits for a class of distillation protocols, *Phys. Rev. A* **59**, 4206–4216 (1999). [arXiv:quant-ph/9708015v3](#).
- [16] K. Goodenough, D. Elkouss and S. Wehner, Assessing the performance of quantum repeaters for all phase-insensitive Gaussian bosonic channels, *New J. Phys.* **18**, 063005 (2016). [arXiv:1511.08710v2](#).
- [17] E. N. Maneva and J. A. Smolin, Improved two-party and multi-party purification protocols, [arXiv:quant-ph/0003099](#) (2000).

- [18] A. W. Leung and P. W. Shor, Adaptive Entanglement Purification Protocols with Two-way Classical Communication, [arXiv:quant-ph/0702156v3](#) (2007).
- [19] K. Horodecki, M. Horodecki, P. Horodecki and J. Oppenheim, General paradigm for distilling classical key from quantum states, [arXiv:quant-ph/0506189](#) (2005).
- [20] K. G. H. Vollbrecht and M. M. Wolf, Conditional entropies and their relation to entanglement criteria, [arXiv:quant-ph/0202058](#) (2002).
- [21] N. Johnston and O. MacLean, Pairwise completely positive matrices and conjugate local diagonal unitary invariant quantum states, *Electronic Journal of Linear Algebra* **35**, 156–180 (2019). [arXiv:1807.06897](#).
- [22] S. Singh and I. Nechita, Diagonal unitary and orthogonal symmetries in quantum theory, *Quantum* **5**, 519 (2021). [doi:10.22331/q-2021-08-09-519](#).
- [23] J. Britz and M. Laurent, Semidefinite hierarchies for diagonal unitary invariant bipartite quantum states, [arXiv:2512.06551v2](#) (2026).
- [24] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824–3851 (1996). [arXiv:quant-ph/9604024](#).
- [25] C. King, The capacity of the quantum depolarizing channel, *IEEE Trans. Inf. Theory* **49**, 221–229 (2003). [arXiv:quant-ph/0204172](#).
- [26] C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal, Entanglement-Assisted Classical Capacity of Noisy Quantum Channels, *Phys. Rev. Lett.* **83**, 3081–3084 (1999). [arXiv:quant-ph/9904023](#).
- [27] S. Pirandola, R. Laurenza, C. Ottaviani and L. Banchi, Fundamental limits of repeaterless quantum communications, *Nature Communications* **8**, 15043 (2017). [arXiv:1510.08863](#).
- [28] M. E. Shirokov, Entropic characteristics of subset of states, [arXiv:quant-ph/0510073v2](#) (2005), Proposition 1a.
- [29] M. E. Shirokov, Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of quantum channels, [arXiv:1512.09047v7](#) (2017).
- [30] A. Jenčová, Recoverability of quantum channels via hypothesis testing, *Lett. Math. Phys.* **114**, 31 (2024), [doi:10.1007/s11005-024-01775-2](#). [arXiv:2303.11707](#).
- [31] M. E. Shirokov, Squashed entanglement in infinite dimensions, [arXiv:1507.08964v5](#) (2015), Proposition 2B.
- [32] M. M. Wilde, Squashed entanglement and approximate private states, [arXiv:1606.08028v3](#) (2018).