Capacities, strong-converse exponents and complementary simulation for qubit Davies channels

Lluis Eriksson

Working paper — Version 3, 25 October 2026

Abstract

We extend the thermal block entropy estimate for generalized amplitude damping to the full qubit Davies region, including additional dephasing. The proof combines the triangular inequality of Tang, Zhu, Bai and Wang with a controlled-purification consequence of strong subadditivity: a third vertex completes the fixed-entropy decomposition used for the thermal boundary. The resulting population-resolved bound holds with every finite-dimensional partner. It implies strong superadditivity of the convex roof of output entropy, and hence arbitrary-partner additivity of Holevo information and classical capacity for both the Davies channel and its complementary channel. We give scalar convex-envelope formulas for both capacities. Every bipartite state supported on the two-dimensional Stinespring range has strongly additive entanglement of formation, so its state entanglement cost requires no regularization. A scalar moment program evaluates that formation functional for arbitrary input coherence. The state/channel correspondence, complement duality and existing axial convex-roof geometry are explicitly attributed to prior work. A normalized reverse channel stays within the Davies region, allowing us also to remove regularization from the parallel entanglement cost of the complementary channel. We obtain an elementary rate in a known flat-roof region and exhibit a nonunital example whose simulation cost strictly exceeds the cost of its Choi state. Finite-block and numerical bounds support the results. The triangular theorem remains an external dependency; the original channel's simulation cost, adaptive simulation and quantum capacity are not determined. We further prove multiplicativity of all maximum output Schatten norms for diagonal-output-filtered Davies maps with arbitrary finite completely positive partners. The proof combines Fang's triangular norm compression with an Araki–Lieb–Thirring bound at the third vertex. Consequently all sandwiched Renyi Holevo orders are strongly additive, and the exact unassisted classical strong-converse exponent has a single-letter scalar evaluation. The general exponent and coding theorems remain external dependencies.

1 Channel family, antecedents and result

A qubit can relax toward a thermal population while its coherence decays faster than relaxation alone requires. Write the corresponding channel as

$$\Phi_{a,b,\lambda}\begin{pmatrix}1-p & z\\ \bar{z} & p\end{pmatrix}=\begin{pmatrix}1-t(p) & \lambda z\\ \lambda\bar{z} & t(p)\end{pmatrix},\qquad t(p)=a+kp,\quad k=1-a-b, \tag{1}$$

where

$$a,b\geq 0,\qquad a+b\leq 1,\qquad 0\leq\lambda^2\leq k,\qquad \lambda\geq 0. \tag{2}$$

A fixed phase of the coherence multiplier can be removed by an output unitary. The map is completely positive: it is generalized amplitude damping followed by dephasing with multiplier $\lambda/\sqrt{k}$ when $k > 0$. The limiting case $k = 0$ is a constant diagonal-output channel.

The region (2) is the closure of the usual dissipative qubit Davies family, allowing degenerate limits and exchange of the energy labels. Roga, Fannes and Życzkowski characterized these maps and their single-use minimum output entropy [1]. At finite positive relaxation times their parameters satisfy

$$k = e^{-t/T_1}, \qquad \lambda = e^{-t/T_2}, \qquad T_2 \le 2T_1.$$

Here t in the exponents is time, whereas $t(p)$ in (1) is an output population. We use (1) as the mathematical definition throughout. General phase-covariant qubit channels allow the larger complete-positivity region $\lambda^2 \le k + ab$; that entire region is not claimed here.

The proof builds on two recent results. Tang et al. established strong additivity for qubit channels admitting a pure output, using a triangular block entropy inequality [2]. OpenAI extended that inequality to the thermal equality case and obtained strong additivity for generalized amplitude damping [3]. In our parameters that boundary is $\lambda^2 = k$. We add a third vertex to their fixed-entropy decomposition, proving the inequality for $\lambda^2 < k$ as well. The new vertex follows directly from strong subadditivity, and is not a new entropy axiom. The scalar convex envelope is necessary because extra dephasing need not preserve convexity of the pure-output entropy.

Concurrent work on generalized amplitude damping. Fang [4] (arXiv:2609.14621, first posted 13 September 2026, revised 6 October 2026) studies generalized amplitude-damping channels with damping probability γ and equilibrium occupation N, both in $[0, 1]$: the output population is $\gamma N + (1 - \gamma)q$ and the coherence is multiplied by $\sqrt{1-\gamma}$. In the notation of (1) this is $a = \gamma N$, $b = \gamma(1 - N)$, $k = \lambda^2 = 1 - \gamma$, that is, the boundary $\lambda^2 = k$ of (2). According to its abstract and its Corollary 2, the classical capacity of these channels equals the known one-copy Holevo information for all $\gamma, N \in [0, 1]$, so no regularization is needed and the capacity is a scalar optimization. The paper also determines the strong-converse exponent over the full parameter range through the sandwiched Rényi Holevo information, and proves additivity of that quantity at every order above one with a triangular Schatten-norm inequality, which it describes as extending the log-determinant method of Tang et al. from the entropy to trace powers. For version 3 we additionally read its triangular norm-compression theorem and displayed proof, and the complete displayed proof of its diagonal-filter multiplicativity theorem. We use [4, Theorem 3] as an external result. Its appendix convexity dependencies have not been independently formalized here. Section 7 extends the norm argument to the full Davies region using the same third-vertex geometry as the entropy proof, with Araki–Lieb–Thirring replacing strong subadditivity at that vertex.

Consequently, on the boundary $\lambda^2 = k$ the equality $C(\Phi) = \chi(\Phi)$ of Theorem 1, with the one-copy maximization (4), is not new relative to [4] (nor to [3], dated 24 September 2026). What Theorem 1 adds, as far as the statements of [4] show, is the interior $0 \le \lambda^2 < k$ with extra dephasing, where the channel is generally not a generalized amplitude-damping channel; the additivity statements (5)–(7) with an arbitrary finite-dimensional partner channel, which are also implied on its GAD boundary by the arbitrary-CP-partner norm theorem of [4]; and the complementary-channel and simulation results of Sections 5 and 6. The all-order extension and its remaining dependencies are stated separately in Section 7; no priority claim is made for the already covered GAD boundary.

Several overlaps are already understood: unital qubit channels [9], entanglement-breaking channels [10], the pure-output boundary [2], and the generalized amplitude-damping boundary $\lambda^2 = k$ [3, 4]. Earlier work treated approximate Holevo additivity for damping combined with dephasing [11]; its ordinary damping examples now also fall within the pure-output theorem. Our contribution is the stated block extension and its additivity consequences across the thermal interior. Roga explicitly

listed Davies capacity additivity as open in 2011 [15, Section 6, p. 78]; this historical statement is not a certificate of present priority. We do not claim a new description of Davies dynamics or a new single-qubit minimum-entropy optimization. The state formation and complementary-channel consequences below use the established Matsumoto–Shimono–Winter correspondence [16], complement duality [19], and the regularized formation characterization of state cost [17]. Hellmund and Uhlmann already analyzed the axial one-use convex roofs [20]; the new operational step is the stated strong additivity, not rediscovery of those roofs.

The simulation result concerns the complementary channel and the parallel diamond-norm task of Berta et al. [23]. Wilde already determined the cost of epolarizing channels, the complements of depolarizing channels, even under sequential tests [24, Section IV.D]. More generally, the unital qubit case is covered by the unital additivity and covariant channel/state machinery [9, 16, 24]. We do not claim these cases as new. The additional nonunital region is useful precisely because the optimizing reference input need not be a Bell pair; Section 6 gives an exact example. The scalar roof and the parallel simulation theorem are antecedents; closure of the explicit reverse parameters lets us apply the Davies range theorem to this different bipartition.

All spaces below are finite dimensional. Let $S(\rho) = -\operatorname{Tr}\rho\ln\rho$ be entropy in nats and put

$$h(p) = -p\ln p - (1-p)\ln(1-p), \qquad g(u) = h\left(\frac{1+\sqrt{1-4u}}{2}\right).$$

Zero terms are interpreted by continuity. Define

$$v(p) = t(p)(1-t(p)) - \lambda^2 p(1-p), \quad e(p) = g(v(p)), \quad f = \operatorname{co} e, \tag{3}$$

where $\operatorname{co} e$ is the largest convex function below e on $[0,1]$.

For a channel $\mathcal{N}$, $\chi(\mathcal{N})$ denotes the maximum ensemble Holevo information in bits, $S_{\min}(\mathcal{N}) = \min_\rho S(\mathcal{N}(\rho))$, and

$$C(\mathcal{N}) = \sup_{n\geq 1}\frac{1}{n}\chi(\mathcal{N}^{\otimes n})$$

is its unassisted memoryless classical capacity, with arbitrary entangled block encodings and collective output measurements [12, 13]. No preshared entanglement or feedback is assumed.

Theorem 1. *For every channel $\Phi = \Phi_{a,b,\lambda}$ satisfying (2),*

$$C(\Phi) = \chi(\Phi) = \frac{1}{\ln 2}\max_{0\leq p\leq 1}\{h(a+kp) - f(p)\}. \tag{4}$$

An ensemble with at most four pure signals, consisting of at most two equally weighted opposite-phase pairs within their respective populations, attains the one-use maximum. For every finite-dimensional channel $\mathcal{N}$,

$$\chi(\Phi\otimes\mathcal{N}) = \chi(\Phi) + \chi(\mathcal{N}), \tag{5}$$

$$S_{\min}(\Phi\otimes\mathcal{N}) = \min_p e(p) + S_{\min}(\mathcal{N}), \tag{6}$$

$$C(\Phi\otimes\mathcal{N}) = \chi(\Phi) + C(\mathcal{N}). \tag{7}$$

In particular $\chi(\Phi^{\otimes n}) = n\chi(\Phi)$ for every n.

Theorem 1 states unrestricted Holevo additivity. The stronger fixed-marginal convex-roof statement is proved in Section 4. The partner's classical capacity retains its regularization; we do not replace $C(\mathcal{N})$ by $\chi(\mathcal{N})$. Formula (4) is a scalar convex-envelope optimization, not a claimed elementary closed form.

2 The block entropy extension

We first state precisely the external inequality used in the proof.

Lemma 2 (Triangular entropy inequality, Tang et al.)**.** *Let A, B, C have the same finite rectangular shape, with squared Hilbert–Schmidt norms x, y, z summing to one. Then*

$$S\left(\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}^*\right) \geq xS(AA^*/x) + yS(BB^*/y) + zS(CC^*/z) + g(yz).$$

A zero-weight term is omitted.

This is the same-shape rectangular case of [2, Theorem 4.2], after permutation of block rows, also stated and proved in [3, Theorem 2.1]. We invoke that result; it is not proved anew or attributed to this paper.

The additional vertex uses only data processing for conditional entropy. Its two-label case also follows from Amosov's phase-damping estimate [14, Proposition 5], with partial trace as the partner channel. We give the short purification proof to fix the direction of the entropy inequality; the vertex estimate itself is not claimed as new.

Lemma 3 (Controlled purifications)**.** *Let $\rho = (\rho_{ij})$ be a density matrix and let X_i be matrices of the same shape with $\operatorname{Tr} X_i X_i^* = 1$. Then*

$$S\left(\sum_{ij} \rho_{ij}|i\rangle\langle j| \otimes X_i X_j^*\right) \geq S(\rho) + \sum_i \rho_{ii} S(X_i X_i^*).$$

Proof. Vectorize X_i as a normalized vector $|X_i\rangle_{BE}$. The map $V|i\rangle = |i\rangle_A|X_i\rangle_{BE}$ is an isometry: explicitly,

$$(V^*V)_{ij} = \langle i|j\rangle\langle X_i|X_j\rangle = \delta_{ij} \operatorname{Tr}(X_i X_i^*) = \delta_{ij}.$$

The retained label A supplies orthogonality; no Hilbert–Schmidt orthogonality among the X_i is assumed. Thus $\Omega = V\rho V^*$ has $S(ABE)_\Omega = S(\rho)$, and its AB marginal is the matrix in the statement. Dephase A, obtaining Ω'. Strong subadditivity implies that applying a channel to the conditioning system AB cannot decrease $S(E|AB)$ [8]. Consequently

$$S(AB)_\Omega - S(ABE)_\Omega \geq S(AB)_{\Omega'} - S(ABE)_{\Omega'}.$$

The right side is $\sum_i \rho_{ii} S(X_i X_i^*)$: the classical entropy of i cancels, and each conditional BE vector is pure. Rearrangement proves the claim. □

Theorem 4 (Thermal inequality with additional dephasing)**.** *Let $X, Y \in \mathbb{C}^{d\times r}$ satisfy $\operatorname{Tr} XX^* = \operatorname{Tr} YY^* = 1$. Let $\alpha, \beta, \eta, \delta, K \geq 0$ satisfy*

$$\alpha + \beta + \eta + \delta = 1, \qquad \alpha\delta - \beta\eta \geq K^2.$$

For

$$T = \begin{pmatrix} \alpha XX^* + \beta YY^* & KXY^* \\ KYX^* & \eta XX^* + \delta YY^* \end{pmatrix}, \quad U_0 = \alpha + \beta, \quad U_1 = \eta + \delta, \quad u = U_0U_1 - K^2,$$

one has

$$S(T) \geq (\alpha + \eta)S(XX^*) + (\beta + \delta)S(YY^*) + g(u). \tag{8}$$

Proof. The assumptions ensure positivity of T: its coherent part with diagonal blocks $\alpha XX^*, \delta YY^*$ is positive because $\alpha\delta \geq K^2$, and the remaining diagonal blocks are positive.

Assume first $U_0U_1 > 0$ and $u > 0$. With U_0, U_1, K fixed, the constraint becomes

$$\beta U_1 + \eta U_0 \leq u.$$

In the coefficient order $(\alpha, \beta, \eta, \delta)$ the feasible triangle has vertices

$$\begin{aligned} A &= (K^2/U_1, u/U_1, 0, U_1), \\ B &= (U_0, 0, u/U_0, K^2/U_0), \\ C &= (U_0, 0, 0, U_1). \end{aligned}$$

The original point has respective weights

$$w_A = \beta U_1/u, \qquad w_B = \eta U_0/u, \qquad w_C = 1 - w_A - w_B.$$

They are nonnegative. Each vertex retains the same row masses and off-diagonal scalar K, hence the same u.

At A, write

$$T_A = M_A M_A^*, \qquad M_A = \begin{pmatrix} \sqrt{\alpha}X & \sqrt{\beta}Y \\ \sqrt{\delta}Y & 0 \end{pmatrix}.$$

Here $\sqrt{\alpha\delta} = K$ and $\beta\delta = u$, so Lemma 2 gives (8) at this vertex. At B use

$$M_B = \begin{pmatrix} \sqrt{U_0}X & 0 \\ \sqrt{\delta}Y & \sqrt{\eta}X \end{pmatrix}.$$

It satisfies $M_B M_B^* = T_B$; interchange its block rows to apply the triangular lemma, with corner-mass product $U_0\eta = u$.

At C, apply Lemma 3 to $\rho = \left(\begin{smallmatrix} U_0 & K \\ K & U_1 \end{smallmatrix}\right)$ and the pair X, Y. The scalar state has determinant u and entropy $g(u)$, giving the desired bound. Finally, entropy concavity applied to $T = w_A T_A + w_B T_B + w_C T_C$ preserves $g(u)$ and the affine weights of the two block entropies.

If a row mass vanishes, $K = 0$ and the result is ordinary entropy concavity in the remaining block. If $u = 0$ and both row masses are positive, the constraint forces $\beta = \eta = 0$; Lemma 3 applies with a pure ρ. This also establishes all boundaries. □

The estimate is sharp for every allowed scalar coefficient tuple: setting $X = Y$ gives $T = \rho \otimes XX^*$, so equality holds. The equality boundary $\alpha\delta - \beta\eta = K^2$ is exactly the two-vertex thermal argument in [3, Theorem 3.1]; the third vertex is needed to cover the inequality region.

3 Population-resolved bound and capacity

Lemma 5. *Let $\mathcal{N}$ be any finite-dimensional channel and $|\psi\rangle = \sqrt{1-p}|0\rangle|\psi_0\rangle + \sqrt{p}|1\rangle|\psi_1\rangle$, where the nonzero branches are normalized. Then*

$$S((\Phi \otimes \mathcal{N})(\psi)) \geq e(p) + (1-p)S(\mathcal{N}(\psi_0)) + pS(\mathcal{N}(\psi_1)). \tag{9}$$

No orthogonality of the conditional vectors is required.

Proof. For Kraus operators L_j of $\mathcal{N}$, form the matrices with columns $X_j = L_j\psi_0$ and $Y_j = L_j\psi_1$. Their squared Hilbert–Schmidt norms are one. The output has the form in Theorem 4, with

$$\alpha = (1-p)(1-a), \quad \beta = pb, \quad \eta = (1-p)a, \quad \delta = p(1-b), \quad K = \lambda\sqrt{p(1-p)}.$$

Here $\alpha\delta - \beta\eta = kp(1-p) \geq K^2$, $\alpha + \eta = 1 - p$, and $u = v(p)$. Apply (8). At $p = 0, 1$ only one product branch remains and the formula follows directly. □

For clarity, the convex envelope can be written as

$$f(p) = \min_{\substack{0\leq q,r,s\leq 1 \\ (1-s)q+sr=p}} \{(1-s)e(q) + se(r)\}. \tag{10}$$

Indeed the minimum of $\int e\, d\mu$ over probability measures with mean p is attained by compactness. This minimum is the lower convex envelope. An extreme minimizing measure has at most two support points: weights on three distinct support points admit a nonzero signed perturbation preserving total mass and mean. Equivalently, each lower supporting line of the convex hull of the graph needs at most two contact points. Continuity of e includes endpoint and one-point cases.

Proof of Theorem 1. Refining any signal ensemble into pure states does not decrease its Holevo information. A pure qubit with population p has output entropy $e(p)$, independently of its phase. If an ensemble has mean population $\bar{p}$, dephasing its average output increases entropy, and convexity of $f \leq e$ gives

$$(\ln 2)\chi(\text{ensemble}) \leq h(a + k\bar{p}) - \sum_\ell w_\ell e(p_\ell) \leq h(a + k\bar{p}) - f(\bar{p}).$$

Conversely choose q, r, s attaining (10). At population q use $\sqrt{1-q}|0\rangle \pm \sqrt{q}|1\rangle$, each with probability $(1-s)/2$; at population r use the analogous pair, each with probability $s/2$. The average is diagonal with population p, and the average output entropy is $f(p)$. Maximizing over p proves the single-use formula and the four-signal bound.

For a pure ensemble $\{w_\ell, \psi_\ell\}$ on the input of $\Phi \otimes \mathcal{N}$, apply Lemma 5 to every signal. The branch outputs of $\mathcal{N}$ with probabilities $w_\ell(1-p_\ell), w_\ell p_\ell$ form a valid ensemble of states of the output of $\mathcal{N}$; write $\psi_{0\ell}, \psi_{1\ell}$ for the conditional branches of ψ_ℓ and $\bar{p} = \sum_\ell w_\ell p_\ell$. Its average is exactly the marginal $\bar{\omega}_{B'}$ on the output of $\mathcal{N}$ of the average output state $\bar{\omega}$. The marginal of $\bar{\omega}$ on the output of Φ is a qubit state with diagonal $(1-t(\bar{p}), t(\bar{p}))$, so its entropy is at most $h(a+k\bar{p})$. Subadditivity therefore bounds the entropy of the full average output by $h(a+k\bar{p}) + S(\bar{\omega}_{B'})$. Subtracting the average of (9), and using $\sum_\ell w_\ell e(p_\ell) \geq f(\bar{p})$ together with $S(\bar{\omega}_{B'}) - \sum_\ell w_\ell[(1-p_\ell)S(\mathcal{N}(\psi_{0\ell})) + p_\ell S(\mathcal{N}(\psi_{1\ell}))] \leq (\ln 2)\chi(\mathcal{N})$, yields

$$(\ln 2)\chi(\text{joint ensemble}) \leq h(a + k\bar{p}) - f(\bar{p}) + (\ln 2)\chi(\mathcal{N}).$$

The single-use formula and product ensembles prove (5). The same one-step inequality bounds every pure output entropy below by $\min_p e(p) + S_{\min}(\mathcal{N})$; entropy concavity handles mixed inputs, and products attain the bound. This proves (6).

Iteration of (5) gives $\chi(\Phi^{\otimes n} \otimes \mathcal{N}^{\otimes n}) = n\chi(\Phi) + \chi(\mathcal{N}^{\otimes n})$. Taking the regularized capacity proves (7), including $C(\Phi) = \chi(\Phi)$. □

4 Entanglement formation in the dilation range

For any finite channel $\mathcal{M}$ define its convex roof of output entropy, in nats, by

$$\widehat{S}_{\mathcal{M}}(\rho) = \min_{\rho = \sum_j w_j \psi_j} \sum_j w_j S(\mathcal{M}(\psi_j)),$$

where the ψ_j are pure states. If $V : \mathbb{C}^2 \to B \otimes E$ is a Stinespring isometry of Φ, then

$$F_\Phi(\sigma) := \widehat{S}_\Phi(\sigma) = E_F(V\sigma V^*). \tag{11}$$

Here E_F denotes bipartite state entanglement of formation in nats. Equation (11) is the standard channel/state correspondence [16]. Each pure component of a state supported on $\operatorname{ran} V$ lies in that range and therefore pulls back to a pure qubit; this also gives a direct proof of the equality.

An elementary observation strengthens the population argument. Every qubit state $\sigma = \left(\begin{smallmatrix} 1-p & z \\ \bar{z} & p \end{smallmatrix}\right)$ is a mixture of pure states with the same population p: two opposite phases, with suitably unequal weights, realize every $|z| \le \sqrt{p(1-p)}$. Their output entropies are all $e(p)$. Thus

$$f(p) \le F_\Phi(\sigma) \le e(p). \tag{12}$$

The upper bound, rather than equality with f, is what is needed next.

Theorem 6 (Strong convex-roof superadditivity)**.** *For every finite-dimensional channel $\mathcal{N}$ and every joint input state $\rho_{AA'}$,*

$$\widehat{S}_{\Phi\otimes\mathcal{N}}(\rho) \ge \widehat{S}_\Phi(\rho_A) + \widehat{S}_\mathcal{N}(\rho_{A'}). \tag{13}$$

The same assertion holds with Φ replaced by any complementary channel Φ^c.

Proof. For a pure $\rho = \psi$, Lemma 5 bounds its output entropy by $e(p)$ plus the average output entropies of the conditional A' branches. By (12), the first term is at least $\widehat{S}_\Phi(\psi_A)$. Those branches decompose $\psi_{A'}$, so their average is at least $\widehat{S}_\mathcal{N}(\psi_{A'})$. For mixed ρ, apply this to a minimizing pure decomposition and use convexity of both marginal roofs.

For the complement, let $\mathcal{N}^c$ be a complementary channel to $\mathcal{N}$. For pure ψ, purity of the complete Stinespring output gives

$$S((\Phi^c \otimes \mathcal{N})(\psi)) = S((\Phi \otimes \mathcal{N}^c)(\psi)).$$

Apply Lemma 5 to the right side. Its conditional branches are pure, so $S(\mathcal{N}^c(\psi_j)) = S(\mathcal{N}(\psi_j))$. Thus the same one-step inequality holds for Φ^c. Its pure-output entropy is again $e(p)$, and the preceding argument applies. This use of complement duality is standard [18, Theorem 1, equations (6) and (9)]; see also [19]. □

For a fixed input σ, the maximal ensemble Holevo information is $(S(\mathcal{M}(\sigma)) - \widehat{S}_\mathcal{M}(\sigma))/\ln 2$. Subadditivity of the mean output entropy and (13) therefore give its subadditivity at arbitrary joint input marginals, with equality at product input states. This is stronger than the unconstrained maximum in Theorem 1; it uses the established convex-roof framework of [16], not an equivalence asserted without its needed hypotheses.

Corollary 7 (State entanglement cost)**.** *Let $\omega = V\sigma V^*$. For every finite-dimensional bipartite state $\Omega_{BB'EE'}$ with $\Omega_{BE} = \omega$,*

$$E_F(\Omega_{BB':EE'}) \ge E_F(\omega_{B:E}) + E_F(\Omega_{B':E'}).$$

Consequently, for every finite bipartite state τ,

$$\begin{aligned} E_F(\omega \otimes \tau) &= E_F(\omega) + E_F(\tau), \\ E_C(\omega) &= F_\Phi(\sigma)/\ln 2, \\ E_C(\omega \otimes \tau) &= F_\Phi(\sigma)/\ln 2 + E_C(\tau). \end{aligned}$$

Here E_C is asymptotic LOCC state *entanglement cost in ebits.*

Proof. Positivity and the support of the first marginal imply that Ω is supported on $\operatorname{ran} V \otimes B' \otimes E'$. Indeed its expectation against $(I - VV^*) \otimes I$ is zero; positivity forces both the corresponding diagonal block and its cross blocks to vanish. Put $\rho = (V^* \otimes I)\Omega(V \otimes I)$, a density operator on $AB'E'$ with marginal σ on A. Pure decompositions correspond bijectively under the isometry $V \otimes I$. Thus

$$E_F(\Omega_{BB':EE'}) = \widehat{S}_{\Phi \otimes \mathrm{Tr}_{E'}}(\rho), \qquad E_F(\omega_{B:E}) = \widehat{S}_{\Phi}(\sigma), \qquad E_F(\Omega_{B':E'}) = \widehat{S}_{\mathrm{Tr}_{E'}}(\rho_{B'E'}).$$

These equalities use nats consistently with F_Φ. Apply Theorem 6 to these three roofs. The displayed superadditivity statement follows. For product states, product pure decompositions attain the reverse inequality. Iteration gives $E_F(\omega^{\otimes n}) = nE_F(\omega)$, and more generally $E_F(\omega^{\otimes n} \otimes \tau^{\otimes n}) = nE_F(\omega) + E_F(\tau^{\otimes n})$. The state-cost theorem $E_C(\rho) = \lim_n E_F(\rho^{\otimes n})/(n \ln 2)$ [17] proves both cost statements. □

These are states in a Stinespring range, not generally Choi states. The result does not evaluate entanglement cost for simulating the channel. At the GAD boundary the same corollary follows from OpenAI's one-step inequality and the standard correspondence; our block extension supplies the remaining Davies region.

For completeness the required one-use formation functional has a scalar moment description. This is a convenient version of the known axial convex-roof problem [20], not a claim that its geometry is new.

Proposition 8. *With* $r(q) = \sqrt{q(1-q)}$ *and* $\sigma = \left(\begin{smallmatrix} 1-p & z \\ \bar{z} & p \end{smallmatrix}\right)$,

$$F_\Phi(\sigma) = \min_{\mu} \left\{ \int e(q)\, d\mu(q) : \int q\, d\mu = p, \quad \int r(q)\, d\mu \geq |z| \right\}, \tag{14}$$

where μ *ranges over probability measures on* $[0,1]$. *In particular* $F_\Phi(\mathrm{diag}(1-p,p)) = f(p)$.

Proof. A pure decomposition of σ induces a population measure; the triangle inequality for its coherences gives the moment constraint. Conversely, given such a measure, put $R = \int r\, d\mu$. At every population choose the same average phase factor z/R, with the sign convention matching the upper-right matrix entry. Two opposite phases with probabilities $(1 + |z|/R)/2, (1 - |z|/R)/2$ realize that factor and leave the entropy integral unchanged. If $R = 0$, necessarily $z = 0$ and the endpoint mixture suffices. Compactness gives a minimizing measure. An extreme minimizer can have at most three populations: four supported pieces admit a nonzero signed weight perturbation preserving normalization, the population moment and the radius moment. This also supplies a finite ensemble. For $z = 0$ the constraint is vacuous, reducing the program to the convex envelope. □

Formula (14) is an optimization formula, not an elementary closed form. For example, the sufficient condition $D \geq 0$, $4AD \geq B^2$, with $v(q) = A + Bq + Dq^2$, makes e convex and then $F_\Phi(\sigma) = e(p)$ for every coherence. This is the flat-roof regime already identified by Hellmund and Uhlmann [20, Theorem 2 and Sections IV.C–D]. We do not claim this criterion or the flat roof as new. For completeness, here is a direct verification of the convexity of e (it is also what Corollary 11 uses). If $D > 0$, then $v(q) = D\big(q + \frac{B}{2D}\big)^2 + \frac{4AD - B^2}{4D}$ is the squared Euclidean norm of the affine map $q \mapsto \big(\sqrt{D}\, q + \frac{B}{2\sqrt{D}},\ \sqrt{(4AD - B^2)/(4D)}\big)$, so $\sqrt{v}$ is convex. If $D = 0$, then $4AD \geq B^2$ forces $B = 0$ and $v \equiv A$ is constant. Next, $G(s) = g(s^2)$, which is Wootters' function of the concurrence at $2s$ [21], is convex and nondecreasing on $[0, \frac{1}{2}]$. We verify this directly: with $g'(u) = 2\,\mathrm{atanh}(w)/w$, $w = \sqrt{1-4u}$ (differentiate $g(u) = \varphi(w)$, $\varphi(w) = h(\frac{1+w}{2})$, $\varphi'(w) = -\,\mathrm{atanh}\, w$), and $w = \sqrt{1-4s^2} = \tanh y$ with $y \geq 0$, one finds $G'(s) = 2s\, g'(s^2) = 2y/\sinh y \geq 0$. As s increases, w and hence y decrease, and $y/\sinh y$ is decreasing in y, so G' increases. Therefore $e = G \circ \sqrt{v}$ is convex (nondecreasing convex function of a convex function), and $f = e$. Outside that regime, (14) retains the coherence constraint even though additivity still holds.

5 Classical capacity of the complementary channel

The complementary output is the environment output of a pure Stinespring dilation, including purification of a mixed reservoir when necessary. It is not automatically the experimentally accessible part of a thermal bath. Its minimal dimension is at most four. To compute its entropy at a diagonal input, choose unit vectors u, v in an environment sector of dimension two, and orthogonal vectors e_+, e_- orthogonal to that sector, with $\langle v, u\rangle = \lambda/\sqrt{(1-a)(1-b)}$. A dilation is

$$V|0\rangle = \sqrt{1-a}|0\rangle|u\rangle + \sqrt{a}|1\rangle|e_+\rangle,$$
$$V|1\rangle = \sqrt{b}|0\rangle|e_-\rangle + \sqrt{1-b}|1\rangle|v\rangle.$$

The positivity constraint makes the prescribed inner product feasible. If its denominator vanishes then $\lambda = 0$ and the irrelevant overlap may be chosen arbitrarily. Tracing the environment gives (1). For input $\mathrm{diag}(1-p, p)$ the environment spectrum is

$$a(1-p), \quad bp, \quad \frac{w(p) + \sqrt{w(p)^2 - 4d(p)}}{2}, \quad \frac{w(p) - \sqrt{w(p)^2 - 4d(p)}}{2}, \tag{15}$$

where

$$w(p) = (1-a)(1-p) + (1-b)p, \qquad d(p) = p(1-p)((1-a)(1-b) - \lambda^2).$$

Indeed the remaining block is the mixture of $|u\rangle\langle u|$ and $|v\rangle\langle v|$ with the indicated trace and determinant. Let $H_E(p)$ be the entropy in nats of the four probabilities in (15).

Theorem 9. *For every qubit Davies channel in (2),*

$$C(\Phi^c) = \chi(\Phi^c) = \frac{1}{\ln 2} \max_{0 \le p \le 1} \{H_E(p) - f(p)\}.$$

At most two opposite-phase population pairs attain the one-use maximum. For every finite-dimensional $\mathcal{N}$,

$$\chi(\Phi^c \otimes \mathcal{N}) = \chi(\Phi^c) + \chi(\mathcal{N}),$$
$$C(\Phi^c \otimes \mathcal{N}) = \chi(\Phi^c) + C(\mathcal{N}),$$
$$S_{\min}(\Phi^c \otimes \mathcal{N}) = \min_p e(p) + S_{\min}(\mathcal{N}).$$

Proof. For $U_\theta = \mathrm{diag}(1, e^{i\theta})$, let W_θ act as the identity on u, v and with phases $e^{-i\theta}, e^{i\theta}$ on e_+, e_-. Then $VU_\theta = (U_\theta \otimes W_\theta)V$, so $\Phi^c(U_\theta \sigma U_\theta^*) = W_\theta \Phi^c(\sigma) W_\theta^*$: Φ^c is phase covariant. Let $\{w_\ell, \psi_\ell\}$ be a pure-state ensemble with mean σ and mean population p, and let $Z = U_\pi = \mathrm{diag}(1, -1)$. Then $\Phi^c(Z\sigma Z)$ is unitarily equivalent to $\Phi^c(\sigma)$, so by concavity of entropy $S(\Phi^c(\bar\sigma)) \ge S(\Phi^c(\sigma))$ for $\bar\sigma = \frac{1}{2}(\sigma + Z\sigma Z) = \mathrm{diag}(1-p, p)$, and $S(\Phi^c(\bar\sigma)) = H_E(p)$. For a pure input ψ_ℓ of population p_ℓ the vector $V\psi_\ell$ on BE is pure, hence $S(\Phi^c(\psi_\ell)) = S(\Phi(\psi_\ell)) = e(p_\ell)$, and $\sum_\ell w_\ell e(p_\ell) \ge f(p)$ because $\sum_\ell w_\ell p_\ell = p$. Thus the Holevo quantity of the ensemble is at most $H_E(p) - f(p)$ nats. The same opposite-phase ensemble used in Theorem 1 has diagonal mean and attains both bounds.

The proof of Theorem 6 established the full one-step inequality for Φ^c. Repeating the branch-ensemble proof of Theorem 1, with H_E replacing $h \circ t$, gives the arbitrary-partner Holevo and minimum-entropy identities. Regularization gives the capacity identity. □

The function H_E is concave, being entropy along an affine family of states; therefore $H_E - f$ is concave. A sampled lower hull f_N gives the same one-sided discretization error as Proposition 20, since only the error in f enters that argument. The limits provide useful checks: an identity channel has a constant complement of capacity zero, whereas a constant qubit-output channel has complementary capacity one bit.

6 Parallel simulation of the complementary channel

This section concerns the resources needed to simulate $\mathcal{N} = \Phi^c$, with its full Stinespring output. Let

$$G_{\mathcal{M}}(\sigma) = E_F((\mathrm{id}_R \otimes \mathcal{M})(\psi_\sigma)), \qquad E_F^{\mathrm{ch}}(\mathcal{M}) = \max_\sigma G_{\mathcal{M}}(\sigma),$$

where ψ_σ is a purification and E_F uses nats. Reference isometries do not affect this definition. The parallel simulation cost E_C^{par} is the asymptotic number of shared ebits per use required by LOCC protocols whose simulated tensor-power channel converges in diamond norm; classical communication is free. The theorem of Berta et al. gives

$$E_C^{\mathrm{par}}(\mathcal{M}) = \lim_{n\to\infty} \frac{E_F^{\mathrm{ch}}(\mathcal{M}^{\otimes n})}{n \ln 2}. \tag{16}$$

This operational theorem is external [23, Theorem 12]. Sequential or adaptive tests constitute a different task [24].

We need two facts about $G_{\mathcal{M}}$: its concavity and continuity, and the subadditivity (18) below. Both are written out in full.

Kraus formula, concavity and continuity of $G_{\mathcal{M}}$. For $X \geq 0$ put $\widetilde{S}(X) = \operatorname{Tr} X\, S(X/\operatorname{Tr} X)$ and $\widetilde{S}(0) = 0$. This function is positively homogeneous. For nonzero $X, Y \geq 0$, concavity of S applied to $\frac{X+Y}{\operatorname{Tr} X+\operatorname{Tr} Y} = \frac{\operatorname{Tr} X}{\operatorname{Tr} X+\operatorname{Tr} Y}\frac{X}{\operatorname{Tr} X} + \frac{\operatorname{Tr} Y}{\operatorname{Tr} X+\operatorname{Tr} Y}\frac{Y}{\operatorname{Tr} Y}$ gives $\widetilde{S}(X+Y) \geq \widetilde{S}(X)+\widetilde{S}(Y)$; together, these say that $\widetilde{S}$ is concave. A Kraus representation $\mathcal{K} = \{K_i\}$ of $\mathcal{M}$ (finitely many operators with $\sum_i K_i^* K_i = 1$) defines

$$f_{\mathcal{K}}(\sigma) = \sum_i \widetilde{S}(K_i \sigma K_i^*),$$

which is concave in σ, being a sum of concave functions composed with linear maps. Here $\widetilde{S}(K_i\sigma K_i^*) = \operatorname{Tr}(K_i\sigma K_i^*)\, S(\cdot/\operatorname{Tr})$ is exactly the weighted term of the Kraus characterization.

(i) Kraus formula. For every state σ, singular ones included,

$$G_{\mathcal{M}}(\sigma) = \inf_{\mathcal{K}} f_{\mathcal{K}}(\sigma), \tag{17}$$

the infimum being over all finite Kraus representations of $\mathcal{M}$; this is the Kraus freedom used in [2, (C.1)]. Proof: fix a purification ψ_σ on RA, a Stinespring isometry $V : A \to B \otimes E$ of $\mathcal{M}$, and put $\Psi = (1_R \otimes V)\psi_\sigma$, so that $(\mathrm{id}_R \otimes \mathcal{M})(\psi_\sigma) = \operatorname{Tr}_E \Psi\Psi^*$. For a pure vector on RB the two marginals have the same nonzero spectrum, so the entanglement entropy of a component may be computed on B. If $\mathcal{K}$ is a Kraus representation, the unnormalized vectors $\phi_i = (1_R \otimes K_i)\psi_\sigma$ satisfy $\sum_i \phi_i\phi_i^* = (\mathrm{id} \otimes \mathcal{M})(\psi_\sigma)$, and the marginal of ϕ_i on B is $K_i\sigma K_i^*$; so this pure-state decomposition has cost $f_{\mathcal{K}}(\sigma)$, which gives "$\leq$". Conversely, let $\sum_j \phi_j\phi_j^*$ be any pure-state decomposition of the same state. By the Schrödinger–Hughston–Jozsa–Wootters theorem [5] there are vectors $e_j \in E$ with $\phi_j = (1_{RB} \otimes \langle e_j|)\Psi$ and $\sum_j |e_j\rangle\langle e_j| = \Pi$, the projector onto the support of the E-marginal of Ψ. Add rank-one terms resolving $1_E - \Pi$; they produce zero vectors, which are omitted. Then $K_j = (1_B \otimes \langle e_j|)V$ satisfy $\sum_j K_j^* K_j = V^*(1 \otimes 1_E)V = 1$ and $\phi_j = (1_R \otimes K_j)\psi_\sigma$, so the decomposition has cost $f_{\mathcal{K}}(\sigma)$ for this $\mathcal{K}$; this gives "$\geq$". Because we purify σ itself, no extension from the support of σ is needed.

(ii) Concavity. For states σ_1, σ_2, $0 \leq \mu \leq 1$, and every $\mathcal{K}$, $f_{\mathcal{K}}(\mu\sigma_1 + (1-\mu)\sigma_2) \geq \mu f_{\mathcal{K}}(\sigma_1) + (1-\mu) f_{\mathcal{K}}(\sigma_2) \geq \mu G_{\mathcal{M}}(\sigma_1) + (1-\mu) G_{\mathcal{M}}(\sigma_2)$. Taking the infimum over $\mathcal{K}$ and using (17) shows that $G_{\mathcal{M}}$ is concave.

(iii) Continuity. Take $|\psi_\sigma\rangle = (1_R \otimes \sqrt{\sigma}) \sum_i |i\rangle_R |i\rangle_A$. The map $\sigma \mapsto \sqrt{\sigma}$ is continuous on positive matrices (indeed $\|\sqrt{X} - \sqrt{Y}\| \leq \|X - Y\|^{1/2}$), so $\sigma \mapsto (\mathrm{id} \otimes \mathcal{M})(\psi_\sigma)$ is continuous. Entanglement of

formation is continuous on finite-dimensional bipartite states [6, 7], and does not depend on the chosen purification (purifications differ by an isometry on R). Hence $G_{\mathcal{M}}$ is continuous on the compact set of input states, and the maximum $E_F^{\mathrm{ch}}(\mathcal{M})$ exists. This uses the continuity of E_F from the literature, as an external fact.

(iv) Diagonal maximizer for Φ^c. With V, W_θ as in the proof of Theorem 9 and $Z = \mathrm{diag}(1,-1)$, we have $\Phi^c(Z\sigma Z) = W_\pi \Phi^c(\sigma) W_\pi^*$. If ψ_σ purifies σ, then $(1_R \otimes Z)\psi_\sigma$ purifies $Z\sigma Z$ and $(\mathrm{id}\otimes\Phi^c)$ of it equals $(1\otimes W_\pi)(\mathrm{id}\otimes\Phi^c)(\psi_\sigma)(1\otimes W_\pi)^*$; local unitaries leave E_F unchanged, so $G_{\Phi^c}(Z\sigma Z) = G_{\Phi^c}(\sigma)$. By (ii), $G_{\Phi^c}(\frac{1}{2}(\sigma + Z\sigma Z)) \geq G_{\Phi^c}(\sigma)$, and $\frac{1}{2}(\sigma + Z\sigma Z)$ is the diagonal part of σ. Hence a maximizer can be taken diagonal.

The other fact is subadditivity of channel formation, [2, Appendix C, Lemma C.2]:

$$E_F^{\mathrm{ch}}(\mathcal{M}\otimes\mathcal{L}) \leq E_F^{\mathrm{ch}}(\mathcal{M}) + E_F^{\mathrm{ch}}(\mathcal{L}). \tag{18}$$

We reproduce its proof in full. It is the argument of [2, Lemma C.2]; it gives the stronger pointwise statement $G_{\mathcal{M}\otimes\mathcal{L}}(\rho_{AA'}) \leq G_{\mathcal{M}}(\rho_A) + G_{\mathcal{L}}(\rho_{A'})$ for every joint input $\rho_{AA'}$, correlated or entangled.

Proof of (18). Fix $\rho_{AA'}$ and $\varepsilon > 0$. By (17) choose Kraus representations $\mathcal{K} = \{K_i\}$ of $\mathcal{M}$ with $f_{\mathcal{K}}(\rho_A) \leq G_{\mathcal{M}}(\rho_A) + \varepsilon$, and $\mathcal{L} = \{L_j\}$ of $\mathcal{L}$ with $f_{\mathcal{L}}(\rho_{A'}) \leq G_{\mathcal{L}}(\rho_{A'}) + \varepsilon$. Only the two marginals of $\rho_{AA'}$ enter this choice.

Step 1 (product representation). The operators $K_i \otimes L_j$ form a Kraus representation of $\mathcal{M}\otimes\mathcal{L}$, so by the easy half of (17)

$$G_{\mathcal{M}\otimes\mathcal{L}}(\rho_{AA'}) \leq \sum_{ij} \widetilde{S}(\omega_{ij}), \qquad \omega_{ij} = (K_i\otimes L_j)\rho_{AA'}(K_i\otimes L_j)^*.$$

Step 2 (classical registers). Put $\omega_{IJBB'} = \sum_{ij} |ij\rangle\langle ij| \otimes \omega_{ij}$ with outcome registers I, J and outputs B, B'. With $q_{ij} = \mathrm{Tr}\,\omega_{ij}$, $S(BB'|IJ)_\omega = S(IJBB') - S(IJ) = \sum_{ij} q_{ij} S(\omega_{ij}/q_{ij}) = \sum_{ij} \widetilde{S}(\omega_{ij})$.

Step 3 (marginals). Because $\sum_j L_j^* L_j = 1$, $\mathrm{Tr}_{B'} \sum_j (K_i\otimes L_j)\rho_{AA'}(K_i\otimes L_j)^* = K_i \rho_A K_i^*$. Hence $\omega_{IB} = \sum_i |i\rangle\langle i| \otimes K_i\rho_A K_i^*$ and $S(B|I)_\omega = f_{\mathcal{K}}(\rho_A)$. Symmetrically $S(B'|J)_\omega = f_{\mathcal{L}}(\rho_{A'})$ (trace preservation of the other channel: the local outcome statistics do not depend on the other instrument).

Step 4 (entropy chain).

$$S(BB'|IJ) \;\leq\; S(B|IJ) + S(B'|IJ) \;\leq\; S(B|I) + S(B'|J).$$

The first inequality is $I(B : B'|IJ) \geq 0$, i.e. strong subadditivity; the second is monotonicity of conditional entropy under discarding conditioning systems ($S(B|IJ) \leq S(B|I)$ and $S(B'|IJ) \leq S(B'|J)$), again strong subadditivity [8].

Combining Steps 1–4, $G_{\mathcal{M}\otimes\mathcal{L}}(\rho_{AA'}) \leq G_{\mathcal{M}}(\rho_A) + G_{\mathcal{L}}(\rho_{A'}) + 2\varepsilon$. Let $\varepsilon \to 0$ and bound the right side by $E_F^{\mathrm{ch}}(\mathcal{M}) + E_F^{\mathrm{ch}}(\mathcal{L})$. Maximizing the left side over $\rho_{AA'}$ gives (18). □

Theorem 10 (Complementary simulation)**.** *For every Davies channel Φ in* (2) *and every finite channel $\mathcal{M}$,*

$$\begin{aligned} E_F^{\mathrm{ch}}(\Phi^c\otimes\mathcal{M}) &= E_F^{\mathrm{ch}}(\Phi^c) + E_F^{\mathrm{ch}}(\mathcal{M}), \\ E_C^{\mathrm{par}}(\Phi^c) &= E_F^{\mathrm{ch}}(\Phi^c)/\ln 2, \\ E_C^{\mathrm{par}}(\Phi^c\otimes\mathcal{M}) &= E_F^{\mathrm{ch}}(\Phi^c)/\ln 2 + E_C^{\mathrm{par}}(\mathcal{M}). \end{aligned}$$

An exact scalar formula is

$$E_C^{\mathrm{par}}(\Phi^c) = \frac{1}{\ln 2} \max_{0\leq p\leq 1} f_p(t(p)), \tag{19}$$

where f_p is the lower convex envelope of the pure-output entropy for the reversed Davies parameters

$$a_p = \frac{pb}{1-t}, \qquad b_p = \frac{(1-p)a}{t}, \qquad k_p = \frac{kp(1-p)}{t(1-t)}, \qquad \lambda_p^2 = \frac{\lambda^2 p(1-p)}{t(1-t)}, \quad t = t(p). \tag{20}$$

Endpoint and constant pure-output exceptions are specified in the proof.

Proof. For $0 < p < 1$ and $0 < t < 1$, apply the dilation of Section 5 to half of $\sqrt{1-p}|00\rangle + \sqrt{p}|11\rangle$. Grouping the resulting pure vector by its B output gives an isometry W_p from a qubit to RE:

$$W_p|0\rangle = \frac{\sqrt{(1-p)(1-a)}|0\rangle|u\rangle + \sqrt{p\bar{b}}|1\rangle|e_-\rangle}{\sqrt{1-t}},$$
$$W_p|1\rangle = \frac{\sqrt{(1-p)a}|0\rangle|e_+\rangle + \sqrt{p(1-b)}|1\rangle|v\rangle}{\sqrt{t}}.$$

Its columns are orthonormal. Tracing E gives a qubit channel Ψ_p with parameters (20); in particular $1 - a_p - b_p = k_p$ and $0 \le \lambda_p^2 \le k_p$. This is the key closure property. The reference–environment output state is exactly

$$\omega_p = (\mathrm{id}_R \otimes \Phi^c)(\psi_p) = W_p \,\mathrm{diag}(1-t, t) W_p^*.$$

Corollary 7 therefore gives strong formation additivity for ω_p with every finite bipartite partner, and $G_{\Phi^c}(\mathrm{diag}(1-p, p)) = f_p(t)$.

Choose a diagonal input maximizing G_{Φ^c} and an optimizing purified input for $\mathcal{M}$, with output state τ. Product inputs, together with the established state additivity, give

$$E_F^{\mathrm{ch}}(\Phi^c \otimes \mathcal{M}) \ge E_F(\omega_p \otimes \tau) = E_F^{\mathrm{ch}}(\Phi^c) + E_F^{\mathrm{ch}}(\mathcal{M}).$$

Equation (18) proves equality. Iteration and (16) prove both operational identities and (19).

At $p = 0, 1$ the reference is pure and $G = 0$; a zero maximizing formation value also gives the product lower bound by local discard. If $t = 0$ or 1 at an interior p, the original channel is constant with pure output. Its complement is isometric, $G(\mathrm{diag}(1-p, p)) = h(p)$ and its cost is one ebit. Pure maximal entanglement gives its strong product lower bound. These cases define the missing values of (19); all other boundary cases follow from the displayed construction or continuity. □

This proof uses an optimizing reference–output state in a different Stinespring range; complement duality alone does not transfer simulation cost. The construction is the standard normalized reverse-channel (weighted-transpose or Petz) construction [25], with its explicit Davies parameters verified here. At the GAD boundary the same deduction uses the antecedent one-step inequality [3]. The added dephasing and nonunital parameters are the extension.

Corollary 11 (An elementary rate)**.** *Put $x = \lambda^2$ and*

$$c(a, b) = k + 2ab - 2\sqrt{ab(k + ab)}.$$

If $c(a, b) \le x \le k$, then

$$E_C^{\mathrm{par}}(\Phi^c) = h\left(\frac{1+\lambda}{2}\right) / \ln 2.$$

A Bell input maximizes channel formation in this region.

Proof. Step 1: the region is the convexity region. In the notation $v(q) = A + Bq + Dq^2$ of Section 4 (with $A = a(1-a)$, $B = k(1-2a) - x$, $D = x - k^2$) put $Q(x) = 4AD - B^2$. A direct expansion gives $Q(x) = -x^2 + \big(2k(1-2a) + 4A\big)x - k^2\big(4A + (1-2a)^2\big) = -(x-c)(x-c')$ with $c' = k + 2ab + 2\sqrt{ab(k+ab)}$: the sum of the roots is $2k(1-2a) + 4A = 2k + 4ab$ (use $1 - a - b = k$) and their product is k^2 (since $4a(1-a) + (1-2a)^2 = 1$), which are $c + c'$ and cc'. Moreover $c \le k \le c'$ (because $\sqrt{k+ab} \le \sqrt{k} + \sqrt{ab}$), and $c' = \big(\sqrt{(1-a)(1-b)} + \sqrt{ab}\big)^2 \le 1$ by Cauchy–Schwarz (use $k + ab = (1-a)(1-b)$), so $c = k^2/c' \ge k^2$. Hence for $c \le x \le k$ we have $D \ge 0$ and $4AD \ge B^2$, and by the verification in Section 4 the pure-output entropy e is convex, i.e. its roof is flat: $f = e$. This coincides with the Hellmund–Uhlmann threshold [20, Equation (29) and Section V.B], which is the antecedent.

Step 2: the reversed channel stays in the region. With $L = p(1-p)/(t(1-t))$ the reverse parameters (20) satisfy $k_p = Lk$, $a_p b_p = Lab$, $\lambda_p^2 = Lx$. The function c is positively homogeneous in (k, ab), so $c(a_p, b_p) = Lc(a, b)$, and $c(a, b) \le x \le k$ gives $c(a_p, b_p) \le \lambda_p^2 \le k_p$. By Step 1 applied to (a_p, b_p, λ_p^2), the reversed roof is flat: $f_p = e_p$, and $f_p(t) = e_p(t)$.

Step 3: the substitution. Write $e_p(q) = g(v_p(q))$ with $v_p(q) = t_p(q)(1 - t_p(q)) - \lambda_p^2 q(1-q)$ and $t_p(q) = a_p + k_p q$. At $q = t$,

$$t_p(t) = \frac{pb + kp(1-p)}{1-t} = \frac{p\,(b + k(1-p))}{1-t} = p,$$

because $1 - t = 1 - a - kp = b + k(1-p)$. Hence $v_p(t) = p(1-p) - Lx\,t(1-t) = p(1-p) - xp(1-p) = (1-x)p(1-p)$, that is,

$$f_p(t) = e_p(t) = g\big((1-x)p(1-p)\big).$$

Step 4: the maximum. As g is increasing and $(1-x)p(1-p) \le (1-x)/4$, with equality at $p = 1/2$ (uniquely when $x < 1$), the maximum in (19) is $g((1-x)/4) = h\big(\frac{1+\sqrt{x}}{2}\big) = h\big(\frac{1+\lambda}{2}\big)$, since $\sqrt{1-(1-x)} = \lambda$. At $x = 1$, the original channel is the identity and its complement has zero cost for every input population. The exceptional constant pure-output case has the same value one (there $x = 0$); continuity covers the remaining endpoints. □

The threshold itself is an antecedent, not a new convex-roof criterion or a necessary boundary for entropy flatness or for this simulation formula. For example, $a = 0.1, b = 0.3, \lambda^2 = 0.5$ has complementary parallel cost about 0.600876037 ebits per use. This determines the complementary simulation rate, not the simulation cost of Φ.

There is also a sufficient low-coherence regime with a simpler optimization. Define the classical conditional entropy

$$J(p) = h(p) + (1-p)h(a) + ph(b) - h(a + kp).$$

Chord identity. For $0 < p < 1$ and $0 < t = t(p) < 1$,

$$(1-t)\,h(a_p) + t\,h(b_p) = J(p). \tag{21}$$

Indeed, let X, Y be bits with $P(X=1) = p$, $P(Y=1|X=0) = a$, $P(Y=1|X=1) = 1-b$. Then $P(Y=1) = a(1-p) + (1-b)p = t$, and the posterior probabilities are $P(X=1|Y=0) = pb/(1-t) = a_p$ and $P(X=0|Y=1) = (1-p)a/t = b_p$, so $H(X|Y) = (1-t)h(a_p) + th(b_p)$. The chain rule $H(X|Y) = H(X) + H(Y|X) - H(Y)$ gives $H(X|Y) = h(p) + (1-p)h(a) + ph(b) - h(t) = J(p)$. Since $e_p(0) = g(a_p(1-a_p)) = h(a_p)$ and $e_p(1) = g(b_p(1-b_p)) = h(b_p)$, the left side of (21) is the value at t of the chord of e_p over $[0,1]$. Therefore $f_p(t) \le J(p)$ for every p, and $f_p(t) = J(p)$ whenever e_p is concave on $[0,1]$ (the convex envelope of a concave function is its chord).

If p_0 maximizes J, is interior, and has $0 < t_0 = t(p_0) < 1$, then

$$\lambda^2 \le \frac{k^2 p_0(1-p_0)}{t_0(1-t_0)} \quad \Longrightarrow \quad E_C^{\mathrm{par}}(\Phi^c) = J(p_0)/\ln 2. \tag{22}$$

Indeed, the quadratic coefficient of the reversed determinant at p_0 is $D_{p_0} = \lambda_{p_0}^2 - k_{p_0}^2 = L(\lambda^2 - Lk^2) \le 0$ with $L = L(p_0)$, so v_{p_0} is concave on $[0,1]$. Since g is increasing and concave (see Section 9: $g'(u) = 2\,\mathrm{atanh}(w)/w$ is positive and nonincreasing), $e_{p_0} = g \circ v_{p_0}$ is concave, so $f_{p_0}(t_0) = J(p_0)$ by (21). Together with $f_p(t(p)) \le J(p) \le J(p_0)$ for all p, formula (19) gives $E_C^{\mathrm{par}}(\Phi^c) = J(p_0)/\ln 2$. Equation (22) is sufficient; no complete transition diagram is asserted.

Proposition 12 (Strict separation from Choi-state cost)**.** *For $a = 1/10, b = 3/10, \lambda^2 = 1/10$, let $\mathcal{N} = \Phi^c$ and let $J_{\mathcal{N}}$ be its normalized Choi state. Then*

$$E_C^{\mathrm{par}}(\mathcal{N}) > E_C(J_{\mathcal{N}}).$$

In particular a Bell input does not maximize $E_F^{\mathrm{ch}}(\mathcal{N})$.

Proof. Write $G_{\mathcal{N}}(p) = G_{\mathcal{N}}(\mathrm{diag}(1-p,p)) = f_p(t(p))$, as in the proof of Theorem 10. Here $k = 3/5$ and $t(p) = \frac{1}{10} + \frac{3}{5}p \in (0,1)$ for $0 < p < 1$. Let $L(p) = p(1-p)/(t(1-t))$.

Step 1: $G_{\mathcal{N}} = J$ on an explicit interval. At $p = 1/2$ the reversed parameters have $a_p = 1/4$, $b_p = 1/8$, $k_p = 5/8$ and $\lambda_p^2 = 5/48$. The quadratic coefficient of the reversed determinant is $D_p = \lambda_p^2 - k_p^2 = L(\lambda^2 - Lk^2) = 5/48 - 25/64 = -55/192 < 0$. For general p, $D_p < 0$ if and only if $\lambda^2 t(1-t) < k^2 p(1-p)$, which for these parameters reads $324p^2 - 312p + 9 < 0$, that is, $p \in I = \left(\frac{13}{27} - \frac{\sqrt{595}}{54},\ \frac{13}{27} + \frac{\sqrt{595}}{54}\right) \approx (0.0298, 0.9332)$, an open interval containing $1/2$. For $p \in I$ the function v_p is concave on $[0,1]$, hence $e_p = g \circ v_p$ is concave (g is increasing and concave, Section 9), so $f_p(t)$ is the value at t of the chord of e_p. By the chord identity (21) this is $J(p)$. Thus $G_{\mathcal{N}}(p) = J(p)$ for all $p \in I$; this is the lower bound $G_{\mathcal{N}} \ge J$ (the upper bound $G_{\mathcal{N}} \le J$ holds for every p).

Step 2: the derivative. From $J'(p) = \ln\frac{1-p}{p} - h(a) + h(b) - k\ln\frac{1-t}{t}$ we get, at $p = 1/2$ where $t = 2/5$,

$$J'(1/2) = h(3/10) - h(1/10) - \tfrac{3}{5}\ln(3/2).$$

Writing $\ln\frac{3}{10} = \ln 3 - \ln 10$ and similarly for the other logarithms, the terms in $\ln 10$ cancel and $h(3/10) = \ln 10 - \frac{3}{10}\ln 3 - \frac{7}{10}\ln 7$, $h(1/10) = \ln 10 - \frac{9}{5}\ln 3$, so

$$J'(1/2) = \tfrac{9}{10}\ln 3 - \tfrac{7}{10}\ln 7 + \tfrac{3}{5}\ln 2 = \tfrac{1}{10}\ln\left(\frac{2^6 3^9}{7^7}\right) > 0,$$

because $2^6 3^9 = 1259712 > 823543 = 7^7$. Since $1/2 \in I$ and J is differentiable there, there is $p_1 \in I$, $p_1 > 1/2$, with $J(p_1) > J(1/2)$. By Step 1, $G_{\mathcal{N}}(p_1) > G_{\mathcal{N}}(1/2)$.

Step 3: the comparison. The Choi state is $J_{\mathcal{N}} = \omega_{1/2} = W_{1/2}\,\mathrm{diag}(1-t,t)W_{1/2}^*$, supported on the range of the Davies Stinespring isometry $W_{1/2}$; by Corollary 7, $E_C(J_{\mathcal{N}}) = G_{\mathcal{N}}(1/2)/\ln 2$. On the other side, by Corollary 7 applied to W_{p_1} and induction on n, $G_{\mathcal{N}^{\otimes n}}(\sigma^{\otimes n}) = E_F(\omega_{p_1}^{\otimes n}) = nG_{\mathcal{N}}(p_1)$ for $\sigma = \mathrm{diag}(1-p_1,p_1)$, so $E_F^{\mathrm{ch}}(\mathcal{N}^{\otimes n}) \ge nG_{\mathcal{N}}(p_1)$, and (16) gives $E_C^{\mathrm{par}}(\mathcal{N}) \ge G_{\mathcal{N}}(p_1)/\ln 2 > G_{\mathcal{N}}(1/2)/\ln 2 = E_C(J_{\mathcal{N}})$. In particular $E_F^{\mathrm{ch}}(\mathcal{N}) \ge G_{\mathcal{N}}(p_1) > G_{\mathcal{N}}(1/2)$, so a Bell input does not maximize. □

The strict inequality thus uses only the product-input lower bound, strong additivity of E_F on the Davies Stinespring range (Corollary 7), and the external operational formula (16). It does not use the subadditivity (18), nor the identity $E_C^{\mathrm{par}}(\Phi^c) = E_F^{\mathrm{ch}}(\Phi^c)/\ln 2$ of Theorem 10, whose upper bound is where (18) enters. Consequently this channel cannot be simulated exactly by LOCC using one copy of

its own Choi state per call: preparing many such states at their state cost would contradict the strict inequality. Other resource states are not excluded. The example has $\lambda^2 > ab$ and is nonunital, so it does not merely restate the covariant epolarizing example of [24]. Approximate values are 0.704192652 ebits at the Bell input and 0.704712865 near $p \approx 0.5169491$ (the grid value is 0.516949108; a 75-digit computation gives the maximizer 0.516949105); the strict proof does not use these values.

7 Filtered output norms and strong-converse exponents

This section strengthens the order-one conclusions for the original Davies channel. It does not assert the corresponding filtered-norm or exponent identities for its complementary channel. The triangular Schatten inequality below is Fang's theorem, not a result proved here. The coefficient triangle is the one used in Section 2. The additional ingredient at its third vertex is a standard consequence of Araki–Lieb–Thirring.

For a finite-dimensional completely positive map Ψ, not necessarily trace preserving, define

$$\nu_p(\Psi) = \max_{\rho \geq 0,\, \operatorname{Tr}\rho = 1} \|\Psi(\rho)\|_p, \qquad 1 \leq p \leq \infty.$$

The maximum exists by compactness. For $H \geq 0$ write $\Phi_H(X) = H\Phi(X)H$.

Lemma 13 (External triangular norm compression)**.** *For every $r \geq 2$ and compatible rectangular blocks,*

$$\left\| \begin{pmatrix} U & 0 \\ V & W \end{pmatrix} \right\|_r \leq \left\| \begin{pmatrix} \|U\|_r & 0 \\ \|V\|_r & \|W\|_r \end{pmatrix} \right\|_r .$$

This is [4, Theorem 3], including its rectangular and operator-norm endpoints. Its trace-power convexity argument is an external dependency of the present extension.

Lemma 14 (Controlled block factors)**.** *Let $C \geq 0$ be an $n \times n$ scalar matrix and let X_i be $d \times r$ matrices with $\|X_i\|_{2p} = 1$, where $1 \leq p < \infty$. Then*

$$\left\| \sum_{ij} C_{ij} |i\rangle\langle j| \otimes X_i X_j^* \right\|_p \leq \|C\|_p .$$

The analogous statement holds for $p = \infty$ with $\|X_i\|_\infty = 1$.

Proof. Set $D = \operatorname{diag}(X_1, \ldots, X_n)$, $A = C \otimes I_r$ and $B = D^*D$. The matrix on the left is DAD^*, whose nonzero eigenvalues are those of $B^{1/2}AB^{1/2}$. The standard Araki–Lieb–Thirring inequality [26, Theorem 1] gives

$$\begin{aligned} \operatorname{Tr}\big[(DAD^*)^p\big] &\leq \operatorname{Tr} B^{p/2} A^p B^{p/2} \\ &= \sum_i (C^p)_{ii} \operatorname{Tr}\big[(X_i^* X_i)^p\big] = \operatorname{Tr} C^p . \end{aligned}$$

Rectangular factors add zero eigenvalues and do not change this argument. For $p = \infty$, D is a contraction, so $\|DAD^*\| \leq \|A\| = \|C\|$. □

Theorem 15 (Three-vertex block norm bound)**.** *Let $\|X\|_{2p} = \|Y\|_{2p} = 1$, $1 \leq p < \infty$, and let $\alpha, \beta, \eta, \delta, K \geq 0$ satisfy $\alpha\delta - \beta\eta \geq K^2$. No trace normalization is imposed. For*

$$T = \begin{pmatrix} \alpha XX^* + \beta YY^* & KXY^* \\ KYX^* & \eta XX^* + \delta YY^* \end{pmatrix}, \quad C = \begin{pmatrix} \alpha + \beta & K \\ K & \eta + \delta \end{pmatrix},$$

one has $\|T\|_p \leq \|C\|_p$. The corresponding statement holds at $p = \infty$ with operator-normalized X, Y. The constant one is sharp.

Proof. Write $U_0 = \alpha + \beta$, $U_1 = \eta + \delta$, and $u = U_0U_1 - K^2$. For positive row masses and $u > 0$, use precisely the vertices A, B, C and weights w_A, w_B, w_C in the proof of Theorem 4. The condition is equivalent to $\beta U_1 + \eta U_0 \le u$, so all three weights are nonnegative. Their convex combination is the original coefficient tuple, while the scalar comparator C is unchanged at every vertex.

At vertex A, its factor is

$$M_A = \begin{pmatrix} (K/\sqrt{U_1})X & \sqrt{u/U_1}\,Y \\ \sqrt{U_1}\,Y & 0 \end{pmatrix}.$$

The scalar matrix of its block $2p$-norms has Gram matrix C. Permuting a block row or column reduces it to Lemma 13. Since $\|M_A M_A^*\|_p = \|M_A\|_{2p}^2$, the desired bound follows. At vertex B use

$$M_B = \begin{pmatrix} \sqrt{U_0}\,X & 0 \\ (K/\sqrt{U_0})Y & \sqrt{u/U_0}\,X \end{pmatrix};$$

its compressed Gram matrix is again C. At the third vertex, $T_C = \operatorname{diag}(X, Y)(C \otimes I_r)\operatorname{diag}(X, Y)^*$, so Lemma 14 applies. The triangle inequality for the Schatten norm combines these three bounds.

If $u = 0$ and both row masses are positive, the coefficient condition forces $\beta = \eta = 0$ and the third-vertex argument applies directly. If a row mass vanishes, $K = 0$ and the remaining row is bounded by the triangle inequality. These arguments include zero scalar coefficients. The same proof uses the norm endpoints for $p = \infty$. Taking $X = Y$ to be a rank-one matrix of unit Schatten norm gives $T = C \otimes XX^*$ and equality, proving sharpness. □

Theorem 16 (Diagonal-filter multiplicativity)**.** *Let $\Phi = \Phi_{a,b,\lambda}$ satisfy (2). For every nonnegative diagonal H, every finite-dimensional completely positive map Ψ, and every $1 \le p \le \infty$,*

$$\nu_p(\Phi_H \otimes \Psi) = \nu_p(\Phi_H)\nu_p(\Psi).$$

The partner may have arbitrary input/output dimensions and need not preserve trace.

Proof. For finite p, norm convexity reduces to a normalized pure input $|0\rangle|u\rangle + |1\rangle|v\rangle$. Complete positivity yields a factorization

$$(\mathrm{id} \otimes \Psi)(\psi) = \begin{pmatrix} XX^* & XY^* \\ YX^* & YY^* \end{pmatrix}.$$

Put $x = \|X\|_{2p}^2$, $y = \|Y\|_{2p}^2$, $m = x + y$ and $q = y/m$. Suppose first $xy > 0$, and normalize X and Y by $\sqrt{x}$ and $\sqrt{y}$. The output before filtering, divided by m, has the coefficients

$$\alpha = (1-q)(1-a), \quad \beta = qb, \quad \eta = (1-q)a, \quad \delta = q(1-b), \quad K = \lambda\sqrt{q(1-q)}.$$

They obey $\alpha\delta - \beta\eta - K^2 = (k - \lambda^2)q(1-q) \ge 0$. For $H = \operatorname{diag}(h_0, h_1)$, multiply α, β by h_0^2, η, δ by h_1^2, and K by h_0h_1. The determinant gap is multiplied by $h_0^2h_1^2$ and remains nonnegative, including singular filters. Theorem 15 therefore gives

$$\|(\Phi_H \otimes \Psi)(\psi)\|_p \le m\|\Phi_H(\psi_q^+)\|_p, \quad |\psi_q^+\rangle = \sqrt{1-q}|0\rangle + \sqrt{q}|1\rangle.$$

Finally,

$$\begin{aligned} m &= \|\Psi(uu^*)\|_p + \|\Psi(vv^*)\|_p \\ &\le (\|u\|^2 + \|v\|^2)\nu_p(\Psi) = \nu_p(\Psi). \end{aligned}$$

If $m = 0$ the output vanishes; if exactly one block norm is zero, the input after $\mathrm{id} \otimes \Psi$ has a single basis branch and the same estimate follows directly. Maximization proves the upper bound and product inputs prove the reverse bound. The finite-dimensional limit $p \to \infty$ gives the remaining endpoint. □

7.1 Renyi radii and their scalar evaluation

In the rest of this section logarithms are to base two. For $\alpha > 1$ define

$$\widetilde{\chi}_\alpha(\mathcal{N}) = \inf_\sigma \sup_\rho \frac{1}{\alpha - 1} \log_2 \operatorname{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \mathcal{N}(\rho) \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right].$$

The standard support convention gives infinity if an output is not supported in σ. The ensemble-radius identity and product-ensemble superadditivity used below are external results [30, 31].

Corollary 17. *For every channel $\mathcal{N}$ and every $\alpha > 1$,*

$$\widetilde{\chi}_\alpha(\Phi \otimes \mathcal{N}) = \widetilde{\chi}_\alpha(\Phi) + \widetilde{\chi}_\alpha(\mathcal{N}).$$

In particular, the n-use quantity is exactly $n\widetilde{\chi}_\alpha(\Phi)$ for every positive integer n.

Proof. Phase covariance and data processing under discarding a uniform binary phase flag show that dephasing a reference state cannot increase its worst-output radius; this is the argument in [4, Eq. (37)]. Thus diagonal references suffice for Φ. Strictly positive references suffice by approximation with $(1-\epsilon)\sigma + \epsilon I/d$; the uniform radius penalty is at most $-\log_2(1-\epsilon)$. For a fixed full-rank reference, the supremum is $\alpha/(\alpha-1)$ times the logarithm of the maximum output α-norm of the corresponding sandwiched map. Use a product reference for $\Phi \otimes \mathcal{N}$ and apply Theorem 16 with the sandwiched partner, which is a completely positive map. Optimizing the two references proves the upper bound. Product-ensemble superadditivity [31], with the ensemble-radius identity [30], proves the reverse bound. □

For a direct scalar formula set $\sigma_s = \operatorname{diag}(1-s, s)$, $0 < s < 1$, $h_0 = (1-s)^{(1-\alpha)/(2\alpha)}$, $h_1 = s^{(1-\alpha)/(2\alpha)}$, and $t = a + kq$. Define

$$B_{\alpha,s}(q) = \begin{pmatrix} h_0^2(1-t) & h_0 h_1 \lambda \sqrt{q(1-q)} \\ h_0 h_1 \lambda \sqrt{q(1-q)} & h_1^2 t \end{pmatrix}.$$

If $L_\pm(\alpha, s, q)$ are its two eigenvalues, then

$$\widetilde{\chi}_\alpha(\Phi) = \inf_{0<s<1} \max_{0 \le q \le 1} \frac{\log_2(L_+^\alpha + L_-^\alpha)}{\alpha - 1}. \tag{23}$$

Indeed pure inputs suffice by norm convexity, and phase covariance makes their eigenvalues depend only on q. No minimax interchange or single-orbit optimal-ensemble assertion is made in this formula.

7.2 The exact exponent above capacity

Let $P_{\mathrm{s}}^\star(n, R; \Phi)$ be the best average success probability for unassisted classical communication using n copies, at least $\lceil 2^{nR} \rceil$ messages, arbitrary entangled input codewords and a collective output measurement. The channel is memoryless; feedback and preshared entanglement are excluded.

Corollary 18. *For every $R \ge 0$ the limit defining the optimal strong-converse exponent exists and equals*

$$\lim_{n\to\infty} -\frac{1}{n} \log_2 P_{\mathrm{s}}^\star(n, R; \Phi) = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} [R - \widetilde{\chi}_\alpha(\Phi)]. \tag{24}$$

The right side is positive for $R > C(\Phi)$. Every finite-block code obeys the upper success bound obtained by exponentiating $-n$ times the right side of (24). Product input codewords with collective decoding attain the optimal asymptotic exponent.

Proof. The classical Arimoto converse bounds every code at each $\alpha > 1$ by $2^{-n(\alpha-1)(R-\widetilde{\chi}_\alpha(\Phi))/\alpha}$ after Corollary 17. Optimizing α gives the finite-block statement. Equality of the asymptotic exponent follows from the general regularized exponent theorem [29, Theorem 5 and Appendix E], since Corollary 17 removes its regularization at every order. For product-codeword attainment, apply the classical-quantum coding exponent theorem [30, Section VI] to the channel with alphabet all pure qubit input states. Its optimized radius equals that of Φ by norm convexity; the finite-dimensional output permits finite approximating alphabets. This gives the same exponent with product codewords. This coding theorem is an external operational dependency, not a construction of a particular finite-block decoder in this paper. □

The new norm proof extends Fang's GAD boundary to the full region $0 \le \lambda^2 \le k$. The standard unital, entanglement-breaking and zero-temperature overlaps are not claimed as new. For example, $a = 1/10, b = 3/10, \lambda^2 = 2/5$ is nonunital, has strict extra dephasing because $2/5 < k = 3/5$, and is not entanglement breaking because $\lambda^2 > ab$. The CP-allowed strip $k < \lambda^2 \le k + ab$ is not covered. Neither the general two-row norm-compression conjecture nor any qutrit-ladder capacity statement follows from these results. Integer-order filtered multiplicativity is already covered by the entrywise-positive-map theorem of King, Nathanson and Ruskai [27]. Without filtering, the general qubit results of King and Koldan [28] already cover $p = 2$ and $p \ge 4$. The extension here must therefore be read with these overlaps, in particular the real noninteger orders and the filtered nonunital, non-entanglement-breaking interior, rather than as a new result for every individual parameter and order in the theorem.

8 Finite-block consequences

The preceding result also applies to nonidentical channels in (2). Write e_i, f_i, t_i for the corresponding functions of channel Φ_i.

Corollary 19. *For every state ρ on n qubits, with marginal excitation probabilities p_i,*

$$S((\Phi_1 \otimes \cdots \otimes \Phi_n)(\rho)) \ge \sum_{i=1}^{n} f_i(p_i).$$

Proof. For a pure input split at the first site and use (9). Apply induction on each conditional remaining state and combine its populations using convexity of each f_i. At the first site $e_1 \ge f_1$. A pure-state decomposition, entropy concavity and convexity of the f_i give the mixed-input statement. □

This yields an exact accounting of the Holevo deficit. Let a pure ensemble have output states ω_ℓ, mean $\bar{\omega}$, populations $p_{i\ell}$ and means $\bar{p}_i$. Put

$$c_i = (\ln 2)\chi(\Phi_i), \quad \sigma = \bigotimes_i \operatorname{diag} \bar{\omega}_i, \quad R_\ell = S(\omega_\ell) - \sum_i f_i(p_{i\ell}) \ge 0.$$

Using quantum relative entropy in nats, direct subtraction gives

$$\begin{aligned} \sum_i c_i - (\ln 2)\chi(\text{ensemble}) = &\sum_i [c_i - h(t_i(\bar{p}_i)) + f_i(\bar{p}_i)] \\ &+ D(\bar{\omega}\|\sigma) + \sum_\ell w_\ell R_\ell \\ &+ \sum_i \left[\sum_\ell w_\ell f_i(p_{i\ell}) - f_i(\bar{p}_i) \right]. \end{aligned} \tag{25}$$

Every term is nonnegative. In particular, $D(\bar{\omega}\|\sigma) = \sum_i h(t_i(\bar{p}_i)) - S(\bar{\omega})$. The support convention is harmless: a zero marginal basis probability gives zero support of $\bar{\omega}$ on that sector. Equation (25) is a consequence of the population bound and standard relative-entropy identities. It extends the usual finite-block deficit method, including that used for amplitude damping in [2]; it does not imply uniqueness of optimizers or a uniform positive gap.

9 Scalar computation and examples

The envelope in (10) can be evaluated by the lower convex hull of the sampled curve. The next bound makes the discretization error explicit, separately from floating-point error.

Proposition 20. *Suppose e is L-Lipschitz on $[0,1]$. For a uniform grid of mesh $1/N$, let f_N be the lower convex envelope of its sampled points, linearly extended between hull vertices. If*

$$C_N = \frac{1}{\ln 2} \max_p \{h(a+kp) - f_N(p)\},$$

then

$$C_N \le C(\Phi) \le C_N + \frac{L}{2N \ln 2}.$$

Proof. The piecewise-linear interpolant $\widetilde{e}_N$ satisfies $|\widetilde{e}_N - e| \le L/(2N)$: at relative position θ in a cell the error is at most $2L\theta(1-\theta)/N$. Taking lower convex envelopes preserves uniform error bounds. Moreover $f_N \ge f$, because mixtures of sampled populations form a subset of all mixtures; and $\operatorname{co} \widetilde{e}_N = f_N$. Thus $0 \le f_N - f \le L/(2N)$, which gives the claim. □

For parameters with $m = \min_p v(p) > 0$, write

$$v(p) = A + Bp + Dp^2, \quad A = a(1-a), \quad B = k(1-2a) - \lambda^2, \quad D = \lambda^2 - k^2.$$

The minimum is found from the endpoints and any interior vertex of this quadratic. A valid Lipschitz constant is

$$L = \max\{|B|, |B+2D|\}\, g'(m), \qquad g'(m) = \frac{2 \operatorname{atanh} \sqrt{1-4m}}{\sqrt{1-4m}},$$

with limit 2 at $m = 1/4$. The function g' is nonincreasing; this follows by differentiating $\operatorname{atanh} x/x$. When $m = 0$ one can instead use uniform continuity; the displayed finite-L bound is not asserted at that boundary.

For $k > 0$, on a hull segment of slope s, the maximizing population is the projection onto that segment of

$$p = \frac{(1+\exp(s/k))^{-1} - a}{k}.$$

This follows by setting $k \ln((1-t)/t) = s$ and using concavity. The case $k = 0$ has zero capacity.

A strict interior example is

$$a = \frac{1}{10}, \quad b = \frac{3}{10}, \quad k = \frac{3}{5}, \quad \lambda^2 = \frac{1}{2}.$$

It is nonunital, has additional dephasing, and is not entanglement breaking. In fact the Choi partial transpose is positive exactly when $\lambda^2 \le ab$, which fails here. Its Choi rank is four, and $\min v = 5/56 > 0$, so it admits no pure output. The example is therefore not covered merely by the unital, entanglement-breaking, pure-output or $\lambda^2 = k$ boundary cases.

a	b	λ^2	capacity (bits, approximate)	mean population
0.10	0.30	0.60	0.503934607	0.468332
0.10	0.30	0.50	0.407849860	0.470350
0.10	0.30	0.10	0.296671803	0.471876
0.02	0.30	0.514624	0.436551160	0.441065

These are reproducible floating-point illustrations on a grid with $N = 200000$, not machine-certified intervals. Proposition 20's exact-arithmetic error bounds are below 3.8×10^{-6} bits for these cases, before rounding error. The last example's sampled hull optimizer mixes population 0 with approximately 0.958865, illustrating why a single population pair need not suffice. No exact three-signal necessity or global transition boundary is inferred from that numerical observation. For the same four channels, a separate computation with $N = 60000$ gives complementary capacities approximately 0.492274059, 0.562095952, 0.906804474 and 0.469017072 bits. Direct Stinespring matrices reproduce (15) and the phase covariance in 500 random cases, with maximum entry or eigenvalue error below 5.0×10^{-16}. These are floating-point illustrations, not machine-certified intervals. For $a = 0.02, b = 0.30, \lambda^2 = 0.514624$, a sampled version of (14) at $p = 0.44$ gives state formation costs about 0.466960 ebits at $z = 0$, 0.467338 at $z = 0.25$, and 0.469937 at $z = 0.45$. The coherence constraint matters; the scalar roof $f(p)$ alone does not describe all states. The linear program uses 20001 population samples and reconstructs a feasible phase ensemble up to numerical tolerance; no continuous optimization certificate is claimed.

10 Verification, dependence and scope

The proof uses Lemma 2 as an external result and gives all subsequent matrix and channel arguments explicitly. Some minimum-output-entropy subregions are also covered by Fukuda's composition criterion [22]; that criterion is not a generic closure theorem for unrestricted Holevo additivity under composition. A root-agent exploratory screen tested 16000 random rectangular block instances without a violation. Separately, an Astra candidate audit tested 9000 block instances, their 27000 vertex instances, 1200 channel outputs with arbitrary finite partner samples, and 300 ensembles. Its smallest residual was of order 10^{-14} in equality cases. These floating-point tests are error-finding aids, not proofs.

A tempting alternative argument—that extra dephasing must increase the block entropy by at least the scalar entropy increase—failed a numerical check and is not used. The valid argument preserves the scalar entropy across the three vertices and invokes concavity only after their separate estimates.

The present family excludes the permitted phase-covariant region $\lambda^2 > 1-a-b$ and channels with negative longitudinal multiplier. We determine the state costs in Corollary 7 and the complementary parallel simulation cost in Theorem 10. We do not determine the original channel's simulation cost, adaptive simulation cost, quantum or private capacity, interacting multiqubit Davies dynamics, or finite-error coding rates. Memoryless use is essential to the operational interpretation. No claim of exhaustive literature coverage or worldwide priority accompanies the stated extension; its mathematical scope is explicit and its immediate antecedents are identified.

AI assistance and review. The author directed an AI-assisted research process. Codex developed the extension, proofs, exposition and scalar calculations. A separately dispatched reviewer requested as GPT-6 Astra with xhigh effort audited frozen proof dossiers and an earlier capacity-only manuscript. The reviewer has prior involvement and is not blind to the work. A subsequent

integrated version was also reviewed. Earlier integrated versions added complementary parallel simulation, reverse-channel closure and the strict Choi-cost separation; every manuscript assessment must identify its exact input version, and an earlier assessment does not automatically apply to these additions. Model runtime identity was not independently attested. No human peer review or predetermined numerical grade is claimed.

Version 2. Version 2 (21 October 2026) was prepared by a Claude Sonnet 5.5 session acting as author-reviser, after three automated assessments of version 1 (two by Claude Sonnet 5.5, one by Claude Haiku 5.5; none human, none independent of the model family used for drafting). Changes: the concurrent paper [4] is cited and compared with Theorem 1; the Kraus formula with concavity and continuity of $G_{\mathcal{M}}$, and the proof of (18), are written out; the chord identity (21), the proof of Proposition 12 and the flat-roof identity in Corollary 11 are written out; the phase-averaging step of Theorem 9 is made explicit; Appendix A lists what each script reproduces. No theorem, hypothesis or numerical value of version 1 was changed. External dependencies that remain: Lemma 2, the continuity of E_F [6, 7], the Hughston–Jozsa–Wootters theorem [5], and (16).

Version 3. Version 3 adds Section 7, including diagonal-filter multiplicativity, all-order Renyi Holevo additivity and the exact classical strong-converse exponent. It expands the label-isometry calculation and the Stinespring-range lift in Corollary 7; it does not delete or replace the public assessments of version 2. The new extension was developed with OpenAI Codex. A separately dispatched GPT-6 Astra xhigh audit examined a frozen analytical dossier with shared filesystem access; its runtime identity and context isolation were not independently attested. That targeted audit is not a full assessment of this combined PDF. Any full-manuscript assessment must identify its exact PDF; earlier scores do not transfer to this version. The new computational controls are symbolic coefficient identities and bounded floating-point rectangular block tests; neither is a formal proof of the universal matrix inequalities.

Availability. The version 3 supplement additionally supplies the symbolic coefficient checks and rectangular Schatten-norm diagnostics for Section 7. The root checks cover 560 block tuples and 560 controlled vertices across seven orders, while the targeted audit reports 192 block tuples, 96 controlled vertices and 384 joint outputs with CP partners. These are bounded floating-point diagnostics apart from the explicit coefficient identities. They do not prove the external triangular norm theorem, certify global numerical radii, or replace the analytical proof. The supplement contains the proof dossier, numerical scripts and reports, and source-retrieval metadata. Original implementation code is released under CC0; the manuscript is offered under CC BY 4.0. External papers retain their authorship and licences. The scalar scripts do not implement directed-rounding interval certification. The version 2 supplement adds the two exact-arithmetic check scripts listed in Appendix A.

A Reproducibility

The supplement contains the scripts below (Python with NumPy, SciPy and mpmath). The historical version 2 runs were performed on 21 October 2026 in fresh writable copies, as `python -B <script>.py`, with Python 3.12.6, NumPy 2.2.6, SciPy 1.18.0 and mpmath 1.3.0 on Windows 11. The two "independent" scripts both write `independent-results.json`; they must be run in separate copies. Rational and integer comparisons and SymPy identities are exact algebraic checks; high-precision evaluations and floating-point diagnostics are distinguished in the table. Numerical optimization and sampling are error-finding aids, not universal proofs, and are not certified by

directed-rounding interval arithmetic. The version 3 controls were run separately on 25 October 2026; their logs and any later reproductions retain their own dates. The table lists what each script computes and how its output compares with the figures in the text.

Script	Computes	Output compared with the text
`test_thermal_extension.py`	16000 random complex block instances of (8) (shapes 2×2 to 4×4, random admissible coefficients, $0\le K\le\sqrt{\alpha\delta}-\beta\eta$)	Minimum of LHS$-$RHS is $+8.0\times10^{-7}$: no violation. It also reports the minimum of the "dephasing growth" gap, $-0.495<0$: this is the failed alternative argument of Section 10.
`capacity_scalar_screen.py`	Hull evaluation of (4) for the four parameter sets of the table in Section 9, $N=200000$	0.503934607, 0.407849860, 0.296671803, 0.436551160 bits, and mean populations 0.468332, 0.470350, 0.471876, 0.441065, as in the table; the discretization bounds are 2.6×10^{-6} to 3.8×10^{-6} bits ($<3.8\times10^{-6}$ as stated). The last example mixes populations 0 and 0.958865.
`complement_formation_examples.py`	500 random Stinespring matrices (isometry, channel, spectrum (15), covariance); complementary capacities with $N=60000$; sampled version of (14)	Maximum errors $\le 5.0\times10^{-16}$. Capacities 0.492274059, 0.562095952, 0.906804474, 0.469017072 bits; formation 0.466960, 0.467338, 0.469937 ebits at $z=0, 0.25, 0.45$. All as in the text.
`check_reverse_channel.py`	500 random cases of the reverse isometry W_p and of the closure $1-a_p-b_p=k_p$, $0\le\lambda_p^2\le k_p$; sampled values of (19)	Errors $\le 6.0\times10^{-15}$; smallest Davies slack $8.7\times10^{-5}>0$. Costs 0.600876037 (for $\lambda^2=0.5$) and 0.704712865 ebits (for $\lambda^2=0.1$, $p\approx0.5169491$).
`check_closed_form.py`	60 random parameter sets in the region $c(a,b)\le x\le k$: scaling $c(a_p,b_p)=Lc(a,b)$, the determinant identity $v_p(t)=(1-x)p(1-p)$, and the sampled roof against $g((1-x)p(1-p))$	Errors 3.3×10^{-16}, 1.2×10^{-16} and 2.0×10^{-10} nats. These are the numerical counterparts of Steps 2–3 of the proof of Corollary 11.
`check_nonbell_exact.py`	Proposition 12 in exact rational arithmetic and 75-digit `mpmath`	Reversed parameters $1/4, 1/8, 5/8, 5/48$; $D=-55/192$; $J'(1/2)=\frac{1}{10}\ln(1259712/823543)$; 0.704192652 and 0.704712865 bits; unique maximizer 0.516949105.
`independent_audit.py`	9000 block instances, 27000 vertex instances, 1200 pure inputs with arbitrary partners, 300 Holevo ensembles	No failure at tolerance 2×10^{-11}; smallest residual -1.2×10^{-14} (equality cases), as stated in Section 10.
`independent_cost_audit.py`, `independent_entanglement_audit.py`	Further diagnostics (reverse channel, Petz form, moment programs, boundary cases, complement capacities)	Run to completion without failure. Their numbers are not quoted in the text.
`checks/verify_v2_steps.py` (new)	Exact (SymPy) and numerical checks of the steps added in version 2	$J'(1/2)$ equals $\frac{1}{10}\ln(2^6 3^9/7^7)$ to 10^{-170}; $D(1/2)=-55/192$; roots $\frac{13}{27}\pm\frac{\sqrt{595}}{54}$; $t_p(t)=p$; chord $=J$ to 10^{-16}; sampled $f_p(t)-J(p)$ below 2×10^{-15} at $p=0.1,\dots,0.9$ for $(a,b,\lambda^2)=(0.1,0.3,0.1)$; $f_p(t)=g((1-x)p(1-p))$ for $x=0.5$; cost 0.600876037 bits.
`checks/verify_flat_region.py` (new)	Vieta identities of Step 1 of the proof of Corollary 11	$4AD-B^2+(x-c)(x-c')=0$ and $cc'=k^2$ symbolically; $k^2\le c\le k\le c'$ in 20000 random (a,b).

The only discrepancy found between the text of version 1 and the scripts was in the digits of the maximizer in Section 9: 0.516949108 is the grid value and 0.516949105 the high-precision value. Not reproducible by the supplement: the numerical results are not interval-certified, and the exploratory screens do not establish any theorem. The inequality of Lemma 2 is an external dependency that

none of these scripts proves.

References

[1] W. Roga, M. Fannes and K. Życzkowski, *Davies maps for qubits and qutrits*, Reports on Mathematical Physics 66, 311–329 (2010). arXiv:0911.5607v2, Sections 3.1–3.3.

[2] Z. Tang, C. Zhu, G. Bai and X. Wang, *Classical capacity and entanglement cost of the amplitude damping channel*, arXiv:2609.28592v1 (2026), Theorem 4.2, Section 5 and Appendix C (Lemmas C.1 and C.2).

[3] OpenAI, *Classical capacity and entropy inequalities for generalized amplitude damping*, 24 September 2026, Theorems 2.1 and 3.1, Corollary 1.2. OpenAI mathematics repository. Retrieved 17 October 2026; PDF SHA256 recorded in the supplement.

[4] K. Fang, *The classical capacity of generalized amplitude-damping channels and its strong-converse exponent*, arXiv:2609.14621v2 (6 October 2026), Theorems 1, 3 and 4. Theorem 3 is an external dependency; its appendix convexity proof is not formally verified here.

[5] L. P. Hughston, R. Jozsa and W. K. Wootters, *A complete classification of quantum ensembles having a given density matrix*, Physics Letters A 183, 14–18 (1993).

[6] M. A. Nielsen, *Continuity bounds for entanglement*, Physical Review A 61, 064301 (2000). arXiv:quant-ph/9908086.

[7] A. Winter, *Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints*, Communications in Mathematical Physics 347, 291–313 (2016). arXiv:1507.07775.

[8] E. H. Lieb and M. B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, Journal of Mathematical Physics 14, 1938–1941 (1973). doi:10.1063/1.1666274.

[9] C. King, *Additivity for unital qubit channels*, Journal of Mathematical Physics 43, 4641–4653 (2002). doi:10.1063/1.1500791.

[10] P. W. Shor, *Additivity of the classical capacity of entanglement-breaking quantum channels*, Journal of Mathematical Physics 43, 4334–4340 (2002). arXiv:quant-ph/0201149.

[11] F. Leditzky, E. Kaur, N. Datta and M. M. Wilde, *Approaches for approximate additivity of the Holevo information of quantum channels*, Physical Review A 97, 012332 (2018). arXiv:1709.01111.

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

[13] B. Schumacher and M. D. Westmoreland, *Sending classical information via noisy quantum channels*, Physical Review A 56, 131–138 (1997). doi:10.1103/PhysRevA.56.131.

[14] G. G. Amosov, *On estimating the output entropy of a tensor product of the quantum phase-damping channel with an arbitrary channel*, Problems of Information Transmission 49, 224–231 (2013). arXiv:1301.2886, Proposition 5.

[15] W. Roga, *Entropy of quantum channel in the theory of quantum information*, PhD thesis, arXiv:1108.5065v3 (2011), Sections 5.6 and 6.

[16] K. Matsumoto, T. Shimono and A. Winter, *Remarks on additivity of the Holevo channel capacity and of the entanglement of formation*, Communications in Mathematical Physics 246, 427–442 (2004). arXiv:quant-ph/0206148v4, Sections 4–5.

[17] P. M. Hayden, M. Horodecki and B. M. Terhal, *The asymptotic entanglement cost of preparing a quantum state*, Journal of Physics A 34, 6891–6898 (2001). arXiv:quant-ph/0008134, Theorem 1.

[18] A. S. Holevo, *On complementary channels and the additivity problem*, arXiv:quant-ph/0509101v3 (2005), Theorem 1.

[19] C. King, K. Matsumoto, M. Nathanson and M. B. Ruskai, *Properties of conjugate channels with applications to additivity and multiplicativity*, Markov Processes and Related Fields 13, 391–423 (2007). arXiv:quant-ph/0509126v3.

[20] M. Hellmund and A. Uhlmann, *Concurrence and entanglement entropy of stochastic 1-qubit maps*, Physical Review A 79, 052319 (2009). arXiv:0903.1340, Sections IV.C–D and V.

[21] W. K. Wootters, *Entanglement of formation of an arbitrary state of two qubits*, Physical Review Letters 80, 2245–2248 (1998). arXiv:quant-ph/9709029.

[22] M. Fukuda, *Extending additivity from symmetric to asymmetric channels*, Journal of Physics A 38, L753–L758 (2005). arXiv:quant-ph/0505022v4.

[23] M. Berta, F. G. S. L. Brandão, M. Christandl and S. Wehner, *Entanglement cost of quantum channels*, IEEE Transactions on Information Theory 59, 6779–6795 (2013). arXiv:1108.5357v3, Theorem 12.

[24] M. M. Wilde, *Entanglement cost and quantum channel simulation*, Physical Review A 98, 042338 (2018). arXiv:1807.11939v3, Sections II–IV.

[25] H. Barnum and E. Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, Journal of Mathematical Physics 43, 2097–2106 (2002). arXiv:quant-ph/0004088v1, Equation (4), for the normalized reversal construction.

[26] K. M. R. Audenaert, *On the Araki–Lieb–Thirring inequality*, arXiv:math/0701129, Theorem 1 (standard ALT inequality).

[27] C. King, M. Nathanson and M. B. Ruskai, *Multiplicativity properties of entrywise positive maps on matrix algebras*, arXiv:quant-ph/0409181v1, Theorem 3.

[28] C. King and N. Koldan, *New multiplicativity results for qubit maps*, arXiv:quant-ph/0512185v2, Theorem 1.

[29] H.-C. Cheng and M. Tomamichel, *No information transmission through quantum channels above capacity*, arXiv:2609.08998v1 (2026), Theorem 5 and Appendix E.

[30] M. Mosonyi and T. Ogawa, *Strong converse exponent for classical-quantum channel coding*, Communications in Mathematical Physics 355, 373–426 (2017), arXiv:1409.3562v8, Sections IV and VI.

[31] S. Beigi, *Sandwiched Renyi divergence satisfies data processing inequality*, Journal of Mathematical Physics 54, 122202 (2013), arXiv:1306.5920v6, Theorems 11–12.