Mathematical proof note

Strict Nonbinary Rate Bounds Beyond Optimized Classical and Mixed-Channel Envelopes

Yang Shifu
National Mobile Communications Research Laboratory and the Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing 211189, China

Version 1.0 — October 8, 2026

Result overview

For every fixed integer alphabet size $q \geq 3$, we strictly improve the joint flagged envelope of the full-domain Aaltonen/product/shortening family specified here and the unmasked q-ary mixed-channel family near the Plotkin endpoint. Cyclic Gram channels represent the classical envelope exactly. Conditional support expansion improves boundary generators, while independent coherence variations improve full-rank mixed-channel generators outside a finite stationary set. A positive information slope for the remaining generators forces every sufficiently low-rate optimizer to contain an improvable component. The improved channels have output dimension at most $2q^2 + q$, and the resulting rate bound applies to linear and nonlinear codes.

1 Statement and Auxiliary-Channel Criterion

Let $q \geq 3$ be a fixed integer, $k = q - 1$, and $\theta = k/q$. For the maximum cardinality $A_q(n, d)$ of an arbitrary q-ary code of minimum Hamming distance at least d, put

$$R_q(\delta) = \limsup_{n\to\infty} n^{-1} \log_q A_q(n, \lceil \delta n \rceil).$$

The benchmark jointly optimizes all parameters of the named classical components, all unmasked q-MC inputs, and all finite orthogonal flags.

For a finite cyclically output-symmetric channel $x \mapsto \sigma_x$, write

$$\begin{aligned} P &= q^{-1} \sum_x \sigma_x, \\ \chi &= S(P) - q^{-1} \sum_x S(\sigma_x), \\ p &= 1 - q^{-2} \sum_x \mathrm{Tr}(P^{-1/2}\sigma_x P^{-1/2}\sigma_x). \end{aligned} \tag{1}$$

Here $S(A) = -\operatorname{Tr} A \ln A$, and inverses act on mean support. Cyclic symmetry means $U\sigma_x U^* = \sigma_{x+1}$, with indices modulo q. The q-ary pretty-good-measurement criterion [1, Prop. 27] gives

$$p < \delta \quad \Longrightarrow \quad R_q(\delta) \le \chi/\ln q. \tag{2}$$

The channel is fixed before the ordinary blocklength limit. For input-independent orthogonal flags, $\sigma_x = \bigoplus_i a_i \sigma_{i,x}$, both p and χ are affine in the weights a_i.

Sections 2 and 4 define two compact generator families $\mathfrak{A}_q, \mathfrak{M}_q$. Define their full flagged envelope by

$$\begin{aligned} B_q(\delta) &= \min_{\substack{(p,r)\in \operatorname{conv}\mathcal{K}_q \\ p\le\delta}} r, \\ \mathcal{K}_q &= \{(p_\sigma, \chi_\sigma/\ln q) : \sigma \in \mathfrak{A}_q \cup \mathfrak{M}_q\}. \end{aligned} \tag{3}$$

Theorem 1. *For every fixed integer $q \ge 3$, there is $\eta_q > 0$ such that, for every $\theta - \eta_q \le \delta < \theta$, a cyclic channel ρ of output dimension at most $2q^2 + q$ satisfies*

$$p_\rho < \delta, \qquad R_q(\delta) \le \chi_\rho/\ln q < B_q(\delta). \tag{4}$$

At each interior point of this interval, one fixed channel retains a uniform positive gap on a compact neighborhood.

The proof gives $\eta_q = \theta(\beta_q/(4q))^2$, with a strictly positive β_q defined from a finite stationary set in (29). No uniformity in growing q is asserted.

2 The Complete Named Classical Envelope

We first specify the comparison domain. Put

$$H_q(x) = \frac{-x\ln x - (1-x)\ln(1-x) + x\ln k}{\ln q}.$$

Aaltonen's scalar source [2, Eq. (9) and concluding theorem] has triples (ω, ξ, η) satisfying

$$\begin{aligned} 0 \le \omega \le 1, \quad 0 \le \eta \le \frac{q-2}{q-1}\omega, \\ 0 \le \xi - \eta \le \min(\omega - \eta, 1 - \omega), \\ \beta_* \le \omega - \frac{q-1}{q-2}\eta. \end{aligned} \tag{5}$$

For $s = (\omega - \eta)/(1-\eta)$ and $t = (\xi - \eta)/(1-\eta)$, the threshold and objective are

$$\begin{aligned} \beta_* &= (1-\eta)\frac{s(1-s) - t(1-t)}{1 + 2\sqrt{t(1-t)}}, \\ e_A &= 2\beta_* + (\omega - \beta_*)k_{q-1}\left(\frac{\eta}{\omega - \beta_*}\right), \\ f_q(\xi,\eta) &= \frac{H(\eta, \xi-\eta, 1-\xi) + \eta\ln(q-2)}{\ln q}, \\ k_m(x) &= \frac{m - 1 - (m-2)x - 2\sqrt{(m-1)x(1-x)}}{m}. \end{aligned} \tag{6}$$

Here H is Shannon entropy in natural units; collapsed expressions use continuous values. The whole-space objective is $1 - H_q(\omega) + f_q(\xi, \eta)$, with $e_A \le \delta$. The original free parameter $\beta \ge \beta_*$ can be eliminated without loss: the derivative of $2\beta + (\omega - \beta)k_{q-1}(\eta/(\omega - \beta))$ is

$$\frac{q + \sqrt{(q-2)\eta/(\omega - \beta - \eta)}}{q-1} > 0.$$

2.1 A compact Gram parameterization

Let e_0 be the first basis vector, $e_+ = k^{-1/2} \sum_{j>0} e_j$, and $Q_\perp = I - e_0 e_0^* - e_+ e_+^*$. Use the compact domain

$$\mathcal{Z} = \{(s, u, v) : 0 \le s \le 1,\ 0 \le u \le \sqrt{s(1-s)},\ 0 \le v \le s + u\}.$$

For $A = s(1-s) - u^2$ and $D = 1 + 2u + (q-2)v^2$, set

$$T = D^{-1/2} \left[vQ_\perp + \begin{pmatrix} 1 - s + u & \sqrt{A} \\ \sqrt{A} & s + u \end{pmatrix}_{e_0, e_+} \right], \qquad G = T^2. \tag{7}$$

The two-dimensional block has determinant $u(1 + 2u)$ before normalization and nonnegative diagonal, so $T \succeq 0$. Expansion gives $\operatorname{Tr} T^2 = 1$.

Lemma 2. *Every source triple in* (5) *is represented by* (7)*, and conversely. With*

$$\begin{aligned} \eta &= (q-2)v^2/D, & \alpha &= (1 + 2u)/D, \\ t &= (1 - \sqrt{1 - 4u^2})/2, & \omega &= \eta + \alpha s, \quad \xi = \eta + \alpha t, \end{aligned} \tag{8}$$

the source pair is

$$F = \ln q - H(\operatorname{diag} G) + S(G), \tag{9}$$

$$p = 1 - \sum_j T_{jj}^2 = \frac{2A + (q-2)(s + u - v)^2/(q-1)}{D} = e_A. \tag{10}$$

Its rate coordinate is $F/\ln q = 1 - H_q(\omega) + f_q(\xi, \eta)$.

Proof. Squaring the block gives

$$\begin{aligned} \operatorname{diag} G &= (1 - \omega, \omega/k, \ldots, \omega/k), \\ \operatorname{spec} G &= \{\alpha(1-t), \alpha t, (\eta/(q-2))^{[q-2]}\}. \end{aligned}$$

These identities prove (9). Moreover

$$T_{00} = (1 - s + u)/\sqrt{D}, \qquad T_{jj} = \frac{s + u + (q-2)v}{(q-1)\sqrt{D}} \quad (j > 0),$$

giving (10). For the inverse use

$$\begin{aligned} s &= \frac{\omega - \eta}{1 - \eta}, & t &= \frac{\xi - \eta}{1 - \eta}, \\ u &= \sqrt{t(1-t)}, & v &= \sqrt{\frac{\eta(1 + 2u)}{(q-2)(1-\eta)}}. \end{aligned}$$

The degree constraint is exactly $u^2 \le s(1-s)$. Also

$$\beta_* = A/D, \quad \omega - \frac{q-1}{q-2}\eta = \frac{(1+2u)s - v^2}{D},$$

so the last source constraint is $v \le s+u$. It implies the remaining constraint because $v^2 \le (s+u)^2 \le s(1+2u)$. Finally $\omega - \beta_* - \eta = (s+u)^2/D$; insertion into (6) proves its threshold identity. Here $\eta < 1$ throughout the source domain, and all zero-prior endpoints are retained. □

Writing $v_j = Te_j$, define the classical cyclic generator

$$\sigma_x = \bigoplus_{c=0}^{q-1} v_{x+c} v_{x+c}^*. \tag{11}$$

This is a density matrix in dimension q^2, and shifting the blocks implements cyclic symmetry. Its mean is $\bigoplus_c G/q$; the output eigenvalues are diag G, so its information is F. Its PGM success is $\sum_j (v_j^* G^{-1/2} v_j)^2 = \sum_j T_{jj}^2$, since $TG^{-1/2}T = T$ on support. Thus (11) realizes the exact source pair, even for singular G. Let $\mathfrak{A}_q$ contain all these generators.

2.2 Products, shortening, and endpoint components

Let $W_q(\omega, d)$ be the infimum of f_q at fixed ω and $e_A \le d$, with the included zero-rate constant-weight branch at $d \ge D_0(\omega) = 2\omega - q\omega^2/(q-1)$. Define B_{cl} as the closure of the source bound under the full two-block construction [3, Thms. 6–7], ordinary shortening, and the Laihonen–Litsyn operation [4, Thm. 1]

$$\tau[1 - H_q(w)] + (1-\tau)B(d'), \quad \delta = 2\tau w + (1-\tau)d'. \tag{12}$$

Here $0 \le w \le \theta$, $0 < \tau < 1$, and $0 \le d' \le 1$. Include the Hamming, Elias, Singleton, Plotkin and first-LP components, finite recursive uses with their extraction terms, and their closed limiting values.

Lemma 3. *This named classical closure is exactly the lower error-budget boundary of the convex hull of the pairs of* $\mathfrak{A}_q$.

Proof. The rank-one point $(s,u,v) = (\omega, 0, 0)$ has pair $(D_0(\omega), 1 - H_q(\omega))$; its actual error is at most the Hamming charge 2ω. The endpoints $(0,1)$ and $(\theta, 0)$ occur in the family. The first-LP pair is also included: for $f = q^{-1/2}\sum_j e_j$, use

$$G_x = \frac{x}{k}(I - ff^*) + (1-x)ff^*, \qquad 0 \le x \le \theta.$$

It corresponds to $\omega = \theta, \eta = (q-2)x/k, \xi = x$. Its information is $(\ln q)H_q(x)$ and its error is

$$\gamma_q(x) = 1 - \frac{(\sqrt{1-x} + \sqrt{kx})^2}{q}.$$

The source constraints follow from $x \le \theta$ and

$$\beta_* = \frac{\theta}{q}\left(\sqrt{1-x} - \sqrt{x/k}\right)^2 \le \theta - x,$$

using $(\sqrt{a} - \sqrt{b})^2 \le a - b$ for $a \ge b$. Since γ_q is an involution on $[0, \theta]$, this gives $H_q(\gamma_q(\delta))$.

For the two-block construction, the constant-weight bound is $\nu W_q(\omega_1, d_1) + (1-\nu)W_q(\omega_2, d_2)$. The extraction terms are $H_q(\omega) - \nu H_q(\omega_1) - (1-\nu)H_q(\omega_2)$ and $1 - H_q(\omega)$. They cancel to give

$$\begin{aligned}\nu[1 - H_q(\omega_1) + W_q(\omega_1, d_1)] \\ + (1-\nu)[1 - H_q(\omega_2) + W_q(\omega_2, d_2)].\end{aligned} \tag{13}$$

The constraints are $\omega = \nu\omega_1 + (1-\nu)\omega_2$, $\delta = \nu d_1 + (1-\nu)d_2$, and $0 \le d_i \le \min(2\omega_i, 1)$. Each component is a source pair with actual error at most its charged distance, so a flag realizes a pair no worse than (13). The same argument treats (12) and ordinary shortening. The endpoint chord $1 - \delta/\theta$ dominates Singleton. Rank-one points handle Elias and Hamming. The LL first-LP straight line is the shortening chord through $\delta = (q-2)^2/q^2$, where $\gamma_q(\delta) = 1/q$. Induction and compactness handle recursion and limits.

For the reverse inclusion, a lower-hull minimizer at an interior error uses at most two source points. Indeed zero information occurs only at $(\theta, 0)$; hence the minimum is positive and has actual error δ, and an exposed supporting face has dimension at most one. Given the two compositions and errors, set $\varphi = \nu\omega_1$ in the original product domain and take $d_i = e_i$. Its constraints hold because $e_i \le \theta$ and $e_i \le 2\omega_i$. The latter follows by expanding (6):

$$e_A = 2\omega - \frac{qy + (q-3)\eta + 2\sqrt{(q-2)\eta(y-\eta)}}{q-1}, \quad y = \omega - \beta_*.$$

Each selected source is feasible for $W_q(\omega_i, e_i)$, so (13) attains a value no larger than the convex boundary. Endpoints follow by continuity. □

The closed zero-rate and shortening conventions used here incur only a polynomial endpoint correction. At $d = nD_0(\omega)$, centered simplex images of constant-weight words have pairwise nonpositive inner products in dimension $n(q-1)$, hence at most $2n(q-1)$ words: their unit Gram matrix K has $\operatorname{Tr} K^2 \le 2 \operatorname{Tr} K$, so $\operatorname{rank} K \ge |C|/2$. The zero-norm case has one word. Small positive source degrees approach the same zero-rate value from below the threshold. For LL ball selection at equality $d = 2r > 0$, a puncturing fiber has disjoint weight-r supports and at most $\lfloor t/r \rfloor \le n$ elements on the t deleted coordinates. This polynomial correction disappears in the exponent; radius zero is injective. Thus all closed branches in the stated envelope are retained.

3 Improvement at a Conditional Support Boundary

For a finite ensemble with mean P, let Π_x project onto $\operatorname{supp} \sigma_x$, and put $\kappa = q^{-1} \sum_x \operatorname{Tr}[(I - \Pi_x)P]$.

Lemma 4. *Every cyclic channel with $0 < p < \theta$ and $\kappa > 0$ has a cyclic replacement with strictly smaller error and information, using at most q additional dimensions.*

Proof. Write $h(a) = -a\ln a - (1-a)\ln(1-a)$ and $J_\varepsilon(a) = h(\varepsilon a) - \varepsilon h(a)$. Encoding a binary label by σ_x and P, with probabilities $1-\varepsilon, \varepsilon$, and measuring $\{\Pi_x, I-\Pi_x\}$, gives information $J_\varepsilon(\mathrm{Tr}[(I-\Pi_x)P])$. The Holevo measurement inequality [5] therefore gives, for $\sigma_x^\varepsilon = (1-\varepsilon)\sigma_x + \varepsilon P$,

$$\chi_\varepsilon \le (1-\varepsilon)\chi - J_\varepsilon(\kappa). \tag{14}$$

Here Jensen applies because

$$J_\varepsilon''(a) = \frac{\varepsilon(1-\varepsilon)}{(1-a)(1-\varepsilon a)} > 0, \qquad J_\varepsilon(a) \ge \varepsilon a[\ln(1/\varepsilon) - 1].$$

The latter follows by expanding the two binary entropies and using $(1-a)\ln(1-a) \ge -a$.

The mean stays P. Expanding (1) gives the exact error $p_\varepsilon = p + (2\varepsilon - \varepsilon^2)(\theta - p)$. For $0 < \varepsilon < \min(1, \sqrt{p})$, use the noiseless flag

$$\begin{aligned} t_\varepsilon &= \frac{(2\varepsilon-\varepsilon^2)(\theta-p) + \varepsilon^2}{p_\varepsilon}, \\ \tau_x &= (1-t_\varepsilon)\sigma_x^\varepsilon \oplus t_\varepsilon |x\rangle\langle x|. \end{aligned} \tag{15}$$

Then $0 < t_\varepsilon < 1$, $p_\tau = p - \varepsilon^2$, and

$$\chi - \chi_\tau \ge \varepsilon\chi + J_\varepsilon(\kappa) - t_\varepsilon \ln q > 0$$

for sufficiently small ε, since $t_\varepsilon = O(\varepsilon)$ and $\kappa > 0$. □

For a classical generator (11), $\kappa = 0$ implies commuting outputs. If some column $v_j = 0$, a positive-trace mean block is absent from an output support, giving $\kappa > 0$. Otherwise zero defect requires range $G \subseteq \mathrm{span}\, v_j$ for every j, hence $\mathrm{rank}\, G = 1$. All blocks are then multiples of the same projection and commute.

4 The Full-Rank Mixed-Channel Generators

The unmasked q-MC family of [1, Sec. 7], in Fourier coordinates, is the phase orbit of

$$\begin{aligned} A_t &= (1-t)P + tww^*, \qquad w_j = \sqrt{P_{jj}}, \\ P &= \mathrm{diag}(1-\alpha, \alpha/k, \ldots, \alpha/k), \end{aligned} \tag{16}$$

with $0 \le \alpha \le \theta$, $0 \le t \le 1$, and phase unitary $U = \mathrm{diag}(1, \zeta, \ldots, \zeta^{q-1})$, $\zeta = e^{2\pi i/q}$. Let $\mathfrak{M}_q$ include this entire compact domain. For $0 < \alpha < \theta$, use

$$\begin{aligned} r &= \sqrt{k(1-\alpha)/\alpha} > 1, \qquad Z = r^2 + k, \\ L &= 2r + k - 1, \qquad A = (r-1)(r+k). \end{aligned}$$

The diagonal is $(b, B, \ldots, B) = (r^2/Z, 1/Z, \ldots, 1/Z)$, and the two coherence classes are $c = tr/Z$ and $d = t/Z$. Set $u = 1-t$, $v = 1+kt$, and $w_0 = 1 + (k-1)t$. The block

$$K = \begin{pmatrix} r^2 & \sqrt{k}tr \\ \sqrt{k}tr & w_0 \end{pmatrix}$$

has eigenvalues $\nu_+ > \nu_- > 0$, with $\nu_- = r^2uv/\nu_+$. Direct diagonalization gives

$$p = \theta(1 - t^2 L/Z), \tag{17}$$

$$\chi_{\mathrm{MC}} = Z^{-1}\left[\nu_+\ln\nu_+ + \nu_-\ln\nu_- + (k-1)u\ln u - 2r^2\ln r\right]. \tag{18}$$

4.1 Two independent fixed-error directions

Define

$$\begin{aligned} m &= \frac{\ln(\nu_+/\nu_-)}{\nu_+ - \nu_-}, \\ L_{00} &= \ln \nu_- + m(r^2 - \nu_-), \qquad L_{++} = \ln \nu_- + m(w_0 - \nu_-), \\ D_\perp &= \ln u - L_{++} + ktrm. \end{aligned} \tag{19}$$

Keep the diagonal fixed and vary the two positive coherences independently. The seed is positive definite exactly when

$$B - d > 0, \qquad b[B + (k-1)d] - kc^2 > 0,$$

and its error is

$$p = \theta\left(1 - \frac{2c^2}{\sqrt{bB}} - \frac{(k-1)d^2}{B}\right). \tag{20}$$

At the original point use

$$d_h = d + h, \qquad c_h = \sqrt{c^2 - (k-1)r(d_h^2 - d^2)/2}.$$

This preserves error exactly and has $c_0' = -(k-1)/2$. Matrix differentiation gives

$$\left.\frac{d\chi}{dh}\right|_0 = -(k-1)D_\perp. \tag{21}$$

Indeed the first coherence of $\ln A_t$ is trm, while the other is $(L_{++} - \ln u)/k$; contraction with the two perturbations proves (21). If $D_\perp \neq 0$, choose the decreasing sign of h, then slightly multiply both new coherences by a factor greater than one. Positive definiteness is open, error decreases strictly by (20), and continuity preserves the information gain.

Along a fixed-error curve inside the original family, $t' = tA/(ZL)$. Set

$$D_\parallel = \partial_r \chi_{\mathrm{MC}} + \frac{tA}{ZL}\partial_t \chi_{\mathrm{MC}}.$$

Using (18) gives

$$\begin{aligned} \partial_r \chi_{\mathrm{MC}} &= \frac{2r}{Z}(L_{00} - 2\ln r + kt^2 m - \chi_{\mathrm{MC}}), \\ \partial_t \chi_{\mathrm{MC}} &= \frac{2ktr^2 m + (k-1)(L_{++} - \ln u)}{Z}. \end{aligned} \tag{22}$$

Introduce

$$\begin{aligned} T_0 &= \ln(v/u) - tm(qr + tA/2), \\ C_0 &= \frac{q + (k-1)t}{k} - \frac{(k-1)tA}{2krL}. \end{aligned}$$

Then

$$D_\parallel = \frac{2kr}{Z^2}(T_0 + C_0 D_\perp). \tag{23}$$

To verify this identity substitute $L_{00} + L_{++} = \ln(r^2 uv)$ and $L_{++} = \ln u + ktrm - D_\perp$ into (22). If $D_\parallel \neq 0$, move in the decreasing direction along the fixed-error curve, then increase t slightly. This also strictly improves both coordinates. Only simultaneous zeros can remain.

5 There Are Only Finitely Many Exceptional Points

Lemma 5. *For each $r > 1$, $D_\perp(r,t)$ has exactly one zero $t_*(r) \in (0,1)$, is negative below it and positive above it, and t_* is real analytic.*

Proof. Put $a = t/(1-t)$ and

$$H_0(y) = (y + r^2)(y + 1),$$
$$H_1(y) = (r^2 + k)y + qr^2, \qquad H_a = H_0 + aH_1.$$

The matrix-log resolvent, followed by $z = (1-t)y$, gives

$$D_\perp = ka(r-1)I(a,r), \qquad I(a,r) = \int_0^\infty \frac{y-r}{(y+1)H_a(y)}\,dy. \tag{24}$$

For example the unscaled integrand for $D_\perp/[kt(r-1)]$ is $(z - r(1-t))/[(z+1-t)\det(K+zI)]$, which verifies the normalization.

Normalize $W_a(y) = 1/[(y+1)H_a(y)]$ to a probability density. Its first moment is finite and $I = 0$ means $\mathbb{E}_a Y = r$. For $h_0 = H_0/H_1$,

$$H_1^2 h_0' = (r^2 + k)y^2 + 2qr^2 y + r^2(kr^2 + 1) > 0.$$

Thus

$$\frac{d}{da}\mathbb{E}_a Y = \mathrm{Cov}_a\left(Y, -\frac{1}{h_0(Y) + a}\right) > 0.$$

The two functions increase strictly; the bounded score and finite first moment justify differentiation and the strict covariance inequality. The spectral formulas give

$$\lim_{t\downarrow 0} \frac{D_\perp}{t} = k\left(\frac{2r\ln r}{r^2-1} - 1\right) < 0,$$
$$\lim_{t\uparrow 1} \frac{D_\perp}{\ln(1/(1-t))} = \frac{k(r-1)}{r^2+k} > 0. \tag{25}$$

For the first sign use $r - r^{-1} > 2\ln r$; the second uses $\nu_+ \to Z$, $\nu_-/(1-t) \to qr^2/Z$. Existence and uniqueness follow. At a zero $I_a > 0$, so the analytic implicit-function theorem applies. □

Lemma 6. *The set $\mathcal{R}_q = \{r > 1 : D_\parallel(r, t_*(r)) = 0\}$ is finite and nonempty.*

Proof. First examine $r = 1 + \eta$. Differentiation of the rational integral in (24) gives

$$I(0,1) = 0, \qquad I_r(0,1) = -1/6, \qquad I_a(0,1) = q/6.$$

Locally it extends analytically to negative a: $H_0/H_1 \geq 1/q$, so $|a| < 1/(2q)$ keeps $H_a \geq H_0/2$. Consequently $t_*(1+\eta) = \eta/q + O_q(\eta^2)$. For $t = c\eta + d\eta^2$, expansion of $\ln(I+E)$ in the two-dimensional block gives

$$D_\perp = \frac{kc(qc-1)}{6}\eta^3 + O(\eta^4),$$
$$T_0 = -\frac{cq(cq+c-1)}{6}\eta^3 + O(\eta^4). \tag{26}$$

The cubic coefficients do not depend on d. Substituting $c = 1/q$ into (23) yields

$$D_{\|}(1+\eta, t_*(1+\eta)) = -\frac{k}{3q^3}\eta^3 + O_q(\eta^4) < 0. \tag{27}$$

At the other endpoint, uniformly for $0 < t < 1$, $r^2 \le \nu_+ \le r^2 + k$ and $0 < \nu_- \le k$. For any fixed $t_0 \in (0,1)$, $\nu_-/(1-t_0) \to 1 + kt_0$ and $m = O(\ln r/r^2)$, so the stable expression

$$D_\perp = \ln((1-t)/\nu_-) + m(ktr - w_0 + \nu_-)$$

tends to $-\ln(1+kt_0) < 0$. Lemma 5 forces $t_*(r) \to 1$. Along this curve $\nu_-/(1-t_*) \to q$, and its root equation is

$$m(kt_*r - w_0 + \nu_-) = \ln(\nu_-/(1-t_*)).$$

Therefore $rm \to \ln q/k$. Since $(\nu_+ - \nu_-)/r^2 \to 1$,

$$\frac{\ln(1/(1-t_*))}{r} \longrightarrow \frac{\ln q}{k}, \qquad \frac{T_0}{r} \longrightarrow \frac{\ln q}{2k}.$$

The first limit also follows by subtracting $\ln \nu_+$ from $\ln(\nu_+/\nu_-)$ and adding $\ln(\nu_-/(1-t_*))$. Equation (23) now gives

$$r^2 D_{\|}(r, t_*(r)) \longrightarrow \ln q > 0.$$

The function is analytic and has opposite strict signs at its two ends. Its zeros lie in a compact interior interval; being isolated there, they are finite, and the sign change gives at least one. □

6 The Joint Optimizer Must Be Improvable

Call a generator improvable if some finite cyclic replacement has strictly smaller p and χ. Every generator is improvable unless it is commuting, is a zero-error or zero-information endpoint, or is one of the finite points of Lemma 6. For the classical family this follows from Lemma 4 and the rank-one argument following it. For q-MC, $\alpha = 0$ or $t = 0$ is uninformative; at $\alpha = \theta$ the pure orbit is an orthonormal Fourier basis, so its mixtures commute. At $t = 1$, $0 < \alpha < \theta$, the output is pure and $\kappa = 1 - \mathrm{Tr}(PA_1) > 0$, so Lemma 4 applies. All remaining points have $r > 1, 0 < t < 1$, and the two coherence arguments apply unless both derivatives vanish.

We need only a simple slope bound for commuting channels. In a common eigenbasis let a_y be the output distribution and Q_y the input posterior under the uniform prior u_q. Then

$$\theta - p = \sum_y a_y \|Q_y - u_q\|_2^2, \qquad \chi = \sum_y a_y D(Q_y \| u_q).$$

Pinsker's inequality gives

$$\chi \ge \tfrac{1}{2} \sum_y a_y \|Q_y - u_q\|_1^2 \ge \tfrac{1}{2}(\theta - p). \tag{28}$$

Every exceptional interior point has $0 < p < \theta$ and positive information. Hence the finite set defines

$$\beta_q = \min\left\{\frac{1}{2}, \min_{r \in \mathcal{R}_q} \frac{\chi_{\mathrm{MC}}(r, t_*(r))}{\theta - p(r, t_*(r))}\right\} > 0. \tag{29}$$

A zero-error channel has $\chi = \ln q$ by the measurement-information inequality, and a zero-information channel has $p = \theta$. Consequently every nonimprovable generator satisfies

$$\chi \geq \beta_q(\theta - p). \tag{30}$$

Proof of Theorem 1. The pair set in (3) is compact and contains $(0, 1)$ and $(\theta, 0)$. Information zero forces every output to equal its mean, hence $p = \theta$. Conversely, with $Z_x = P^{-1/4}\sigma_x P^{-1/4}$,

$$\theta - p = q^{-2} \sum_x \mathrm{Tr}[(Z_x - P^{1/2})^2],$$

so $p = \theta$ also forces equal outputs. Thus B_q is positive below θ, convex and nonincreasing. Its minimum has actual error δ, since otherwise mixing with $(\theta, 0)$ would strictly lower its value. A supporting line shows that at most two original generators attain it, as in Lemma 3. Finite convexity gives continuity in the interior.

Write $\delta = \theta(1 - s)$. The included first-LP source satisfies

$$\gamma_q(\delta) = \frac{ks^2}{(\sqrt{1 + ks} + \sqrt{1 - s})^2} \leq ks^2.$$

Using $h(x) + x \ln k \leq x \ln(ek/x)$ yields, for $0 < s \leq (\beta_q/(4q))^2$,

$$(\ln q)B_q(\delta) \leq ks^2[1 + 2\ln(1/s)] < \beta_q(\theta - \delta). \tag{31}$$

Indeed division by θs gives $qs[1 + 2\ln(1/s)]$, increasing in this range. At its upper endpoint, put $x = \beta_q/(4q) \leq 1/24$; this ratio is $(\beta_q/4)x[1 + 4\ln(1/x)] < \beta_q/4$, since the increasing factor is less than $17/24$.

Let an attaining mixture have weights $\lambda, 1 - \lambda$. If all positive-weight components were nonimprovable, their affine pairs and (30) would contradict (31). Replace an improvable component of positive weight by its improving channel. The new mixture has strictly smaller error and information, proving (4) via (2). Each original component has dimension at most q^2; a boundary repair adds at most q, and a full-rank coherence improvement uses dimension q. The two-component mixture therefore has dimension at most $2q^2 + q$.

At an interior δ_0, retain this single channel. Its strict error margin persists nearby, and continuity of B_q preserves a positive information gap on a smaller compact neighborhood. All channel parameters are chosen before taking the blocklength limit. □

Sources

[1] O. Alrabiah and V. Guruswami, "Binary code rate bounds via classical–quantum channels," Aug. 2026. arXiv:2608.09347v1. doi: 10.48550/arXiv.2608.09347.

[2] M. Aaltonen, "A new upper bound on nonbinary block codes," *Discrete Math.*, vol. 83, nos. 2–3, pp. 139–160, 1990. doi: 10.1016/0012-365X(90)90002-Y.

[3] Y. Ben-Haim and S. Litsyn, "A new upper bound on the rate of non-binary codes," *Adv. Math. Commun.*, vol. 1, no. 1, pp. 83–92, 2007. doi: 10.3934/amc.2007.1.83.

[4] T. Laihonen and S. Litsyn, "On upper bounds for minimum distance and covering radius of non-binary codes," *Des. Codes Cryptogr.*, vol. 14, pp. 71–80, 1998. doi: 10.1023/A:1008260505585.

[5] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Probl. Inf. Transm.*, vol. 9, no. 3, pp. 177–183, 1973. Journal archive.