

Finite Alphabet Communication with Noisy Forward Memory in a Normalized Loop

Daniel A. Pyle
Independent Researcher

30 October 2026

Abstract

We study one-use communication of a finite classical message through a stipulated normalized loop containing a backward binary symmetric channel and one noisy forward memory carrier. The carrier is either a classical bit or a qubit; no shared random seed, retained side register, or entanglement is supplied. For every finite message count, we characterize the qubit optimum by a constructive scalar feasibility test. For rational noise parameters and a rational target success, the test uses $O(M^2)$ constant-size rational arithmetic checks. The proof combines an exact decoder reduction, minimum-trace allocation for constant encoders, and a concave geometric cap with a matching polygon construction. For three messages, the qubit and classical-bit optima coincide throughout the allowed noise range; an explicit Bernstein certificate proves the remaining polynomial inequality. For four messages, an exact rational qubit protocol exceeds a bound valid for every classical-bit protocol. These are conditional mathematical results for the declared resource model, not a construction of a physical backward channel. Independent human review and originality assessment remain pending.

1 Introduction

Normalized-loop communication gives a finite-dimensional optimization problem in which both a correct-output weight and its normalization depend on the message. A useful resource question is what changes when a single noisy forward classical bit is replaced by a noisy qubit, while the backward channel, objective, and available side information are kept fixed. This paper gives a complete scalar characterization of that one-use qubit problem for every finite alphabet. It also proves that three messages do not distinguish the two memory carriers, whereas four messages can.

The decoder–memory–encoder architecture is already present in Ji, Lloyd and Wilde [1]. Their discussion of noisy forward memory motivates the restriction considered here, but does not specify all of its operational conventions. The conditional worst-message objective is also established prior art: Ji, Regula and Wilde [2], Eq. (20), use the same type of message-dependent normalization. The present calculations concern a particular feasible resource set, not a new error criterion or a solution of a general capacity problem.

Dimension-matched classical and quantum separations have a substantial earlier literature, including prepare-and-measure dimension witnesses [3]. Independent preparation and measurement devices with local but not shared randomness are studied explicitly by Bowles, Quintino and Brunner [4]. These precedents do not, by themselves, optimize the fixed channel-weighted, normalized objective below. No equivalence with their witnesses is asserted, and no priority claim is made for the exact formulas or characterization in this paper.

The main theorem is stated in Section 5 and proved using the geometric lemma of Section 4. Section 6 proves three-message equality; the complete classical upper bound and the Bernstein coefficient certificate are included in the appendices. Section 7 gives a rational four-message

separation. All results concern a single use of each specified channel and dimension-two forward memory. They do not assert physical time travel or extend to higher-dimensional memory, additional retained registers, shared randomness, entanglement assistance, or multiple channel uses.

2 The normalized loop model

Let $M \geq 1$ and let $m, y \in \{0, \dots, M-1\}$ denote the transmitted and recovered classical messages. The backward input and output are bits a, b . We *assume*, as part of the mathematical model, a backward binary symmetric channel

$$N_p(b | a) = \begin{cases} 1 - p, & b = a, \\ p, & b \neq a, \end{cases} \quad 0 < p < \frac{1}{2}. \quad (1)$$

This stipulation does not establish the physical availability of such a backward channel.

For each b , the decoder prepares subnormalized qubit states $\tau_{yb} \geq 0$ with

$$\sum_y \text{Tr } \tau_{yb} = 1. \quad (2)$$

It outputs y and sends its sole forward carrier through the depolarizing channel

$$\mathcal{L}_\epsilon(\tau) = (1 - \epsilon)\tau + \epsilon \text{Tr}(\tau)I/2, \quad 0 \leq \epsilon \leq 1. \quad (3)$$

Knowing m , the encoder measures a binary POVM $\{E_0^m, E_1^m\}$, where $0 \leq E_0^m \leq I$ and $E_1^m = I - E_0^m$, and produces a . It has no access to b or y other than through this carrier. Local private randomness is allowed. No additional retained memory, shared per-trial random seed, shared entanglement, quantum backward carrier, or other communication is allowed. A fixed public codebook is allowed.

Definition 2.1. The loop weights, normalizations, and worst-message success are

$$w_m(y) = \sum_{a,b} N_p(b | a) \text{Tr}[E_a^m \mathcal{L}_\epsilon(\tau_{yb})], \quad (4)$$

$$Z_m = \sum_y w_m(y), \quad S = \min_m \frac{w_m(m)}{Z_m}. \quad (5)$$

The maximum over all allowed qubit protocols is $S_M^{\text{qu}}(p, \epsilon)$.

The normalization in (5) is performed separately for each message. The weights Z_m are not asserted to be laboratory acceptance probabilities and can exceed one. No extra rejection rule is included. The decoder output y is the recovered record; it is not an encoder-accessible side register.

The corresponding classical model uses a forward BSC(q) with input c and output d , and arbitrary stochastic maps $D(y, c | b)$ and $E(a | m, d)$. Its weights are

$$w_m^{\text{cl}}(y) = \sum_{a,b,c,d} N_p(b | a) D(y, c | b) L_q(d | c) E(a | m, d), \quad (6)$$

with the same normalization and resource restrictions. Write its optimum as $S_M^{\text{cl}}(p, q)$. Diagonal qubit preparations and effects reproduce exactly this classical model at

$$q = \epsilon/2. \quad (7)$$

Indeed, computational-basis preparation followed by (3) and computational-basis readout is BSC($\epsilon/2$). Arbitrary diagonal states and effects give arbitrary stochastic classical maps. More generally, if all encoder effects commute, dephasing the preparations in their common basis preserves all contractions, because depolarization commutes with that dephasing.

Throughout, set

$$A = 1 - p, \quad h = \frac{(1 - 2p)(1 - \epsilon)}{2}, \quad \alpha = \frac{1}{2} + h. \quad (8)$$

Thus $0 \leq h < 1/2$ and $1/2 \leq \alpha \leq A < 1$. For the matched classical channel, its effective bijective-encoder coefficients are α and $1 - \alpha$.

Lemma 2.2 (Positivity and attainment). *Every allowed protocol satisfies $2p \leq Z_m \leq 2A$. Both classical and qubit optima are attained.*

Proof. For each positive state τ , the two numbers $\text{Tr}[E_a^m \mathcal{L}_\epsilon(\tau)]$ are nonnegative and sum to $\text{Tr} \tau$. Weighting them by $N_p(b | a) \in [p, A]$ gives a number in $[p \text{Tr} \tau, A \text{Tr} \tau]$. Summing over y and the two rows b proves the bounds. The allowed states and effects form compact sets in finite dimensions; the classical maps form compact simplices. The objective is continuous because its denominators are uniformly positive. $\square$

The excluded endpoint $p = 0$ may admit zero normalization and is not covered by continuity alone. The endpoint $p = 1/2$ is also outside the stated parameter domain.

3 Exact decoder and encoder reductions

Write the subnormalized states in Bloch form,

$$\tau_{mb} = \frac{t_{mb}I + \mathbf{v}_{mb} \cdot \boldsymbol{\sigma}}{2}, \quad t_{mb} \geq 0, \quad |\mathbf{v}_{mb}| \leq t_{mb}, \quad (9)$$

where $|\cdot|$ is the Euclidean norm. Define

$$T_m = t_{m0} + t_{m1}, \quad u_m = t_{m0} - t_{m1}, \quad \mathbf{d}_m = \mathbf{v}_{m0} - \mathbf{v}_{m1}, \quad \mathbf{D} = \sum_m \mathbf{d}_m. \quad (10)$$

Lemma 3.1 (Reversible decoder reduction). *The reduced variables are feasible if and only if*

$$\sum_m T_m = 2, \quad \sum_m u_m = 0, \quad |u_m| \leq T_m, \quad |\mathbf{d}_m| \leq T_m. \quad (11)$$

In particular the allowed trace imbalance u_m imposes no additional constraint on $\mathbf{d}_m$.

Proof. Necessity follows from (2), (9), and the triangle inequality. Conversely, if (11) holds, set

$$t_{mb} = \frac{T_m + (-1)^b u_m}{2}, \quad \mathbf{v}_{mb} = (-1)^b t_{mb} \frac{\mathbf{d}_m}{T_m} \quad (12)$$

for $T_m > 0$, and set both states to zero for $T_m = 0$. Positivity follows from $|\mathbf{d}_m| \leq T_m$. The two row traces sum to one because $\sum T_m = 2$ and $\sum u_m = 0$, and the construction reproduces $\mathbf{d}_m$. $\square$

Lemma 3.2 (Extreme effects suffice). *Every protocol can be replaced, without lowering any message success, by one whose E_0^m is I , 0 , or a rank-one projector for each m .*

Proof. The extreme points of the effect interval $0 \leq E \leq I$ are orthogonal projectors. To see this, an eigenvalue strictly between zero and one can be perturbed in both directions, so such an effect is not extreme. Conversely a projector cannot be a nontrivial average of two effects: positivity forces both effects to vanish on its kernel, and the upper bound by I forces both to equal I on its range. In dimension two the projectors are 0 , I , and rank-one projectors.

For a fixed decoder, success for message m is a ratio of affine functions of E_0^m , with positive denominator. If an effect is expressed as a convex combination with coefficients γ_j , correct weights v_j , and denominators $z_j > 0$, then

$$\frac{\sum_j \gamma_j v_j}{\sum_j \gamma_j z_j} = \sum_j \frac{\gamma_j z_j}{\sum_k \gamma_k z_k} \frac{v_j}{z_j}.$$

Some extreme component therefore has at least the original success. The effect for one message does not change another message's weights, so these replacements can be made independently. Joint convexity of the decoder–encoder problem is not assumed. $\square$

Call $E_0^m = I$ a constant-0 encoder and $E_0^m = 0$ a constant-1 encoder. Their denominators are one and their successes are, respectively,

$$At_{m0} + pt_{m1}, \quad pt_{m0} + At_{m1}. \quad (13)$$

For a projective encoder $E_0^m = (I + \mathbf{n}_m \cdot \boldsymbol{\sigma})/2$, $|\mathbf{n}_m| = 1$, direct contraction gives

$$S_m = \frac{T_m/2 + h\mathbf{n}_m \cdot \mathbf{d}_m}{1 + h\mathbf{n}_m \cdot \mathbf{D}}. \quad (14)$$

The denominator is positive. Consequently a projective label can reach a target t if and only if

$$\frac{T_m}{2} + h|\mathbf{d}_m - t\mathbf{D}| \geq t. \quad (15)$$

For $h > 0$, a maximizing threshold direction is along $\mathbf{d}_m - t\mathbf{D}$ whenever this vector is nonzero; otherwise any unit vector suffices. The same threshold condition is valid when $h = 0$.

Proposition 3.3 (A universal bound). *For every $M \geq 1$,*

$$S_M^{\text{qu}}(p, \epsilon) \leq \frac{A}{A + (M-1)p}. \quad (16)$$

In particular $S_1^{\text{qu}} = 1$ and $S_2^{\text{qu}} = S_2^{\text{cl}} = A$ at matched forward noise.

Proof. For a label with trace T_m , the correct weight is at most AT_m and the incorrect weight is at least $p(2 - T_m)$, by the argument of Lemma 2.2. Thus

$$S_m \leq \frac{AT_m}{AT_m + p(2 - T_m)}.$$

This expression is increasing in T_m . Some label has $T_m \leq 2/M$, proving (16). A single message is recovered perfectly. For two messages, direct encoding through the backward BSC, ignoring memory, attains A . $\square$

4 The geometric cap

The next lemma controls the joint effect of trace allocation and the normalization vector $\mathbf{D}$. It also supplies the attaining geometry used in the finite-alphabet theorem.

Lemma 4.1 (Exact longitudinal cap). *Fix $0 < h < 1/2$, $0 < t \leq 1$, and $0 < r \leq 2$. Suppose $\mathbf{D} = r\mathbf{e}_z$. For a projective label of trace S , the largest feasible longitudinal coordinate $z = \mathbf{d} \cdot \mathbf{e}_z$ under $|\mathbf{d}| \leq S$ and*

$$S/2 + h|\mathbf{d} - t r \mathbf{e}_z| \geq t$$

is defined precisely when $S \geq L$, where

$$L = \frac{t(1 - hr)}{\alpha}, \quad H = \frac{t(1 + hr)}{\alpha}, \quad \alpha = \frac{1}{2} + h. \quad (17)$$

It equals

$$g(S) = \begin{cases} \frac{S^2 + t^2 r^2 - (t - S/2)^2 / h^2}{2tr}, & L \leq S \leq H, \\ S, & S \geq H. \end{cases} \quad (18)$$

The function g is concave, has slope at least one, and satisfies $g(L) = -L$ and $g(H) = H$. On $[L, H]$, each $z \in [-S, g(S)]$ is feasible with $|\mathbf{d}| = S$.

Proof. The triangle inequality gives

$$S/2 + h|\mathbf{d} - t r \mathbf{e}_z| \leq \alpha S + htr,$$

so feasibility requires $S \geq L$. This bound is attained at $\mathbf{d} = -S\mathbf{e}_z$. Since $r \leq 2$, $H \leq 2t$. For $L \leq S \leq H$, the right side of $|\mathbf{d} - t r \mathbf{e}_z| \geq (t - S/2)/h$ is nonnegative, so squaring is valid. Using $|\mathbf{d}|^2 \leq S^2$ gives

$$2trz \leq S^2 + t^2 r^2 - (t - S/2)^2 / h^2.$$

This is the quadratic branch of (18). It is attained on the sphere $|\mathbf{d}| = S$, with any transverse direction having squared length $S^2 - z^2$. Substitution gives $g(L) = -L$ and $g(H) = H$.

For $S \geq H$, take $\mathbf{d} = S\mathbf{e}_z$. Since

$$H - tr = \frac{t(1 - r/2)}{\alpha} \geq 0,$$

we have $S - tr \geq 0$, and $S/2 + h|S - tr| = \alpha S - htr \geq t$. Thus the trivial bound $z \leq S$ is exact on this branch.

On the quadratic branch, direct differentiation gives

$$g''(S) = \frac{1 - 1/(4h^2)}{tr} < 0, \quad g'(H) - 1 = \frac{2 - r}{2hr} \geq 0. \quad (19)$$

Its derivative decreases to a value at least one; at H the derivative can only decrease when it joins the linear branch of slope one. The joined function is therefore concave with slope at least one. In particular $g(S) - S$ increases to zero at H , while $g(S) + S$ is increasing from zero at L . Hence $-S \leq g(S) \leq S$ on $[L, H]$. Every sphere point with $z \leq g(S)$ satisfies the squared inequality and hence the original unsquared inequality. $\square$

5 A complete finite-alphabet characterization

By Lemma 3.2, it suffices to consider n_0 constant-0 labels, n_1 constant-1 labels, and N projective labels, where

$$n_0 + n_1 + N = M. \quad (20)$$

Grouping equal encoder types is a symmetry reduction, not permission to share a random seed. Labels remain distinct and must individually satisfy the target.

5.1 The trace cost of constant labels

Fix $0 < t \leq 1$ and put $d_0 = n_0 t$, $d_1 = n_1 t$. Let x_i and y_i be the aggregate decoder trace assigned to constant group i in rows $b = 0$ and $b = 1$, respectively. The constraints are

$$Ax_0 + py_0 \geq d_0, \quad px_1 + Ay_1 \geq d_1, \quad (21)$$

$$x_0 + x_1 \leq 1, \quad y_0 + y_1 \leq 1, \quad x_i, y_i \geq 0. \quad (22)$$

Let C denote the minimum total constant trace when these constraints are feasible.

Lemma 5.1 (Exact constant allocation). *The following four cases exhaust (21)–(22).*

- (i) *If $d_0, d_1 \leq A$, then $C = (d_0 + d_1)/A$, attained by $x_0 = d_0/A$, $y_1 = d_1/A$, and $x_1 = y_0 = 0$.*
- (ii) *If $d_0 > A$ and $d_1 \leq A$, then feasibility is equivalent to*

$$C = 1 + \frac{d_0 - A}{p} + \frac{d_1}{A} \leq 2.$$

The minimum is attained by $x_0 = 1$, $x_1 = 0$, $y_0 = (d_0 - A)/p$, and $y_1 = d_1/A$.

- (iii) *If $d_1 > A$ and $d_0 \leq A$, exchange the two groups and the two rows in (ii).*
- (iv) *If both $d_0 > A$ and $d_1 > A$, the class is infeasible.*

Each nonempty group's minimizing row masses may be split equally among its labels, making every constant-label success exactly t .

Proof. Every coefficient in (21) is at most A , so total trace is at least $(d_0 + d_1)/A$. This proves optimality in case (i). In case (ii), let $C' = x_0 + x_1 + y_0 + y_1$, $s_0 = Ax_0 + py_0$, and $s_1 = px_1 + Ay_1$. Then

$$C' - s_0/p - s_1/A + A/p - 1 = (A/p - 1)(1 - x_0) + (1 - p/A)x_1 \geq 0.$$

Consequently $C' \geq d_0/p + d_1/A + 1 - A/p = C$. The displayed minimizing allocation obeys its second-row capacity exactly when $C \leq 2$; its other constraints hold automatically. If $C > 2$, the lower bound contradicts the total row capacity. Case (iii) follows by relabeling. In case (iv), summed demands exceed $2A$, but summed successes are at most $AC' \leq 2A$. Finally the group successes of each minimizing allocation equal $d_i = n_i t$, so equal splitting gives the assertion about individual labels. $\square$

5.2 The exact test

Theorem 5.2 (Finite-alphabet scalar characterization). *Fix $M \geq 1$, $0 < p < 1/2$, $0 \leq \epsilon \leq 1$, and a target $0 < t \leq 1$. The inequality $S_M^{\text{qu}}(p, \epsilon) \geq t$ holds if and only if at least one count triple (20) passes the following test.*

First use Lemma 5.1. Reject a triple with infeasible constant allocation. For each remaining triple let C be its minimum constant trace.

- (i) *If $N = 0$, accept the triple.*
- (ii) *If $N = 1$, put $T = 2 - C$. Accept exactly when*

$$T/2 + h[(1 - t)T + tC] \geq t. \quad (23)$$

- (iii) *If $N \geq 2$, put $T = (2 - C)/N$. If $\alpha T \geq t$, accept. If $\alpha T < t$ and $h = 0$, reject. In the remaining case, put*

$$r_{\min} = \frac{1 - \alpha T/t}{h}, \quad (24)$$

$$P_N(r, t) = t(Nt - 2)r^2 + 2tCr + NT^2 - \frac{N(t - T/2)^2}{h^2}. \quad (25)$$

Accept exactly when

$$r_{\min} \leq 2 \quad \text{and} \quad \max_{r \in [r_{\min}, 2]} P_N(r, t) \geq 0. \quad (26)$$

Every accepted triple gives an explicit valid protocol. For rational p, ϵ, t , all acceptance decisions use rational arithmetic: maximize the quadratic by checking its endpoints and, only when concave with a vertex in the interval, its vertex. The number of count triples is $(M+1)(M+2)/2$.

Proof. We prove necessity and sufficiency for each fixed count triple.

No projective labels. Every feasible protocol must satisfy the constant demands. Conversely take a minimizing constant allocation and split it among labels as in Lemma 5.1. If row trace remains unused, assign it to any existing constant label. This can only increase some correct weights; all constant denominators remain one. There is at least one label because $M \geq 1$.

One projective label. Suppose the actual constant trace is $C' \geq C$ and the projective trace is $T' = 2 - C'$. Write $\mathbf{D} = \mathbf{d} + \mathbf{D}_c$. Since $t \in (0, 1]$,

$$|\mathbf{d} - t\mathbf{D}| = |(1-t)\mathbf{d} - t\mathbf{D}_c| \leq (1-t)T' + tC'.$$

The maximum possible left side of (15), as a function of C' , is therefore bounded by

$$(2 - C')/2 + h[(1-t)(2 - C') + tC'].$$

Its derivative is $-1/2 + h(2t - 1) < 0$. Thus its largest value occurs at the minimum C , proving necessity of (23).

For sufficiency, use the minimizing constant allocation, give the remaining row masses to the projective label, set $\mathbf{d} = T\mathbf{e}_z$, and give each constant label a Bloch difference equal to minus its trace times $\mathbf{e}_z$. Then $\mathbf{D}_c = -C\mathbf{e}_z$ and the triangle bound is equality. Lemma 3.1 reconstructs all states, whatever their assigned trace imbalances. Choosing the encoder direction as in (15) attains the target.

Two or more projective labels with $\alpha T \geq t$. Use the minimizing constant allocation and let its residual row traces be R_0, R_1 . Give every projective label row traces $R_0/N, R_1/N$, hence trace T and imbalance $u = (R_0 - R_1)/N$. Let all constant Bloch differences be zero. Choose projective Bloch differences of norm T in a regular planar N -gon, so their sum is zero; for $N = 2$, use opposite vectors. Then $\mathbf{D} = 0$, and every projective label attains αT . The inequalities $|u| \leq T$ follow directly from $R_0, R_1 \geq 0$.

If $h = 0$ and $\alpha T < t$, arbitrary allocation cannot help: the average projective trace is no greater than T and a projective success is its trace divided by two. Thus this case is infeasible.

Necessity of the quadratic test. Now assume $N \geq 2$, $h > 0$, and $\alpha T < t$. Consider an arbitrary feasible decoder in the count class, with actual constant trace $C' \geq C$, projective traces T_j , and $r = |\mathbf{D}| \leq 2$. If $r = 0$, each projective success is at most αT_j , and some $T_j \leq (2 - C')/N \leq T$, a contradiction. Hence $r > 0$.

Rotate coordinates so $\mathbf{D} = r\mathbf{e}_z$ and apply Lemma 4.1. Each projective trace is at least L and each projective longitudinal coordinate is at most $g(T_j)$. Constant labels contribute at most C' to the longitudinal sum. Jensen's inequality gives

$$r \leq C' + \sum_{j=1}^N g(T_j) \leq C' + Ng((2 - C')/N) \leq C + Ng(T). \quad (27)$$

For the last inequality, increasing the average projective trace from $(2 - C')/N$ to T increases Ng by at least $C' - C$, since g has slope at least one. Both averages are in its domain.

Because $\alpha T < t$, we have $T < t/\alpha \leq H$; hence the quadratic branch applies. Its domain condition $T \geq L$ is precisely $r \geq r_{\min}$. Multiplying the last inequality in (27) by $2tr$ and substituting (18) yields

$$0 \leq t(Nt - 2)r^2 + 2tCr + NT^2 - N(t - T/2)^2/h^2 = P_N(r, t).$$

This proves (26).

Sufficiency of the quadratic test. Choose $r \in [r_{\min}, 2]$ with $P_N(r, t) \geq 0$. Use the minimizing constant allocation, and assign every projective label trace T and imbalance $(R_0 - R_1)/N$ as above. Since $T \geq L$ and $T < H$, Lemma 4.1 permits every common longitudinal coordinate $z \in [-T, g(T)]$ with norm T .

Choose an aggregate constant Bloch difference Be_z , where $B \in [-C, C]$, and use the same z for each projective label. The possible longitudinal sums $B + Nz$ fill the interval

$$[-C - NT, C + Ng(T)] = [-2, C + Ng(T)].$$

The condition $P_N(r, t) \geq 0$ is exactly $r \leq C + Ng(T)$, so that interval contains r . Choose B, z with $B + Nz = r$. If $C > 0$, distribute Be_z among constant labels in proportion to their traces; if $C = 0$, take $B = 0$.

Give the projective differences transverse components of common length $\sqrt{T^2 - z^2}$ arranged as a regular N -gon. Their transverse sum is zero, so their total together with the constants is $D = re_z$. Each has norm T and longitudinal coordinate at most $g(T)$, hence satisfies (15). Select its encoder direction accordingly. Lemma 3.1 gives valid states with the prescribed row traces; the constants still succeed with probability t . No shared randomization is used by this fixed geometric construction.

The count triples exhaust all extreme encoders, completing both directions. All formulas and possible quadratic vertices are rational when the inputs are rational. $\square$

Remark 5.3 (Meaning of the arithmetic bound). The theorem is a complete characterization and decision procedure, not a single elementary expression for every optimum. The $O(M^2)$ count concerns rational arithmetic tests; it is not a bit-complexity claim polynomial in $\log M$. Thresholds $t \leq 0$ are trivially feasible and $t > 1$ are impossible. Downward closure and attainment imply that exact rational bisection gives arbitrarily narrow enclosures of the optimum for rational channel parameters, using 0 and any rational value greater than one as initial bounds. No linear-programming or nonlinear-optimization solver is required.

Corollary 5.4 (Complete erasure). *For $\epsilon = 1$,*

$$S_{2k}^{\text{qu}}(p, 1) = \frac{A}{k} \quad (k \geq 1), \quad S_{2k+1}^{\text{qu}}(p, 1) = \frac{A}{k + A} \quad (k \geq 0). \quad (28)$$

These values also equal the classical optima with erased forward memory.

Proof. After complete depolarization, an effect acts only through the scalar $\text{Tr}(E_0)/2$. Thus every encoder is a mixture of the two constant encoders, and the positive-denominator extremality argument reduces to constants alone. Suppose their multiplicities are $j \geq \ell$ with $j + \ell = M$. If $j = \ell$, Lemma 5.1 gives maximum target A/j . If $j > \ell$, the best target lies in the spillover case and is

$$t = \frac{A}{jA + \ell p}.$$

Indeed, $jt > A$, $\ell t \leq A$, and the spillover formula equals two at this value; larger targets violate it. This also covers $\ell = 0$, giving $t = 1/j$. The denominator $jA + \ell p = Mp + j(A - p)$ is minimized by the most balanced integer split. Substitution gives (28). $\square$

6 Three messages give equal classical and quantum optima

For three messages the general scalar test simplifies to a closed expression.

Theorem 6.1 (Three-message equality). *For every $0 < p < 1/2$ and $0 \leq \epsilon \leq 1$,*

$$S_3^{\text{qu}}(p, \epsilon) = S_3^{\text{cl}}(p, \epsilon/2) = K, \quad (29)$$

where K is the unique root in $(0, A)$ of

$$F(t) = \alpha(A - t)(1 - t) - (1 - \alpha)t^2 = 0. \quad (30)$$

Equivalently,

$$K = \frac{2A\alpha}{\alpha(A + 1) + \sqrt{\alpha^2(A + 1)^2 - 4A\alpha(2\alpha - 1)}}. \quad (31)$$

In particular $A/(A + 1) \leq K < A$, with equality in the first inequality exactly when $\epsilon = 1$.

The full classical statement, including its attaining decoder and upper bounds for every stochastic protocol, is proved in Appendix A. It is used below as a proved ingredient rather than a numerical baseline.

Proof. Classical achievability and the embedding (7) give $S_3^{\text{qu}} \geq K$. If $h = 0$, Corollary 5.4 gives equality. Assume henceforth $h > 0$, so $1/2 < \alpha \leq A < 1$.

With at most one projective encoder, all effects commute; dephasing reduces the protocol exactly to the classical model, whose optimum is K . It remains to bound the two-projector and three-projector classes.

First record the commuting lower bound

$$K \geq L_3 := \frac{2A\alpha}{2A + \alpha} \geq \frac{2\alpha}{3}. \quad (32)$$

For a direct construction, give one constant-0 label trace $C = L_3/A$, imbalance $u = C$, and $\mathbf{d} = 0$. Give two projective labels trace $T = 1 - C/2$, imbalance $u = -C/2$, and opposite Bloch differences $\pm T\mathbf{e}_z$, with matching encoder directions. Then $\mathbf{D} = 0$, the constant succeeds with $AC = L_3$, and each projector succeeds with $\alpha T = L_3$. The trace imbalances are feasible since $C < 1$. All effects commute, so this is a classical protocol. The second inequality in (32) follows from $\alpha \leq A$.

Suppose a protocol succeeds strictly above K . Choose a threshold $t > K$ below its success and with $t < A$; such thresholds can be taken arbitrarily close to K . For $N = 2, 3$, the minimum constant trace and common projective trace are

$$C = (3 - N)t/A, \quad T = (2 - C)/N. \quad (33)$$

The bounds (32) imply $\alpha T < t$ in both cases. Necessity in Theorem 5.2 therefore supplies an $r \in [0, 2]$ with $P_N(r, t) \geq 0$. Dropping the stronger lower bound $r \geq r_{\min}$ only relaxes this necessary condition.

Three projective labels. Here $C = 0$ and $T = 2/3$. Since $t > 2\alpha/3$,

$$P_3(0, t) = 3T^2 - 3(t - T/2)^2/h^2 < 0.$$

If $t \leq 2/3$, the coefficient of r^2 is nonpositive and the maximum for $r \geq 0$ occurs at zero. If $t > 2/3$, the quadratic is convex, so its maximum on $[0, 2]$ occurs at an endpoint. The other endpoint is also strictly negative, by the identity

$$P_N(2, t) = N(4 - h^{-2})(t - T/2)^2, \quad C + NT = 2. \quad (34)$$

This identity follows by expanding (25); its sign is strict because $h < 1/2$ and $t > T/2$. Thus the three-projector class cannot improve on K .

Two projective labels. Evaluate (25) at $t = K$, with $C = t/A$ and $T = 1 - t/(2A)$. Its leading coefficient is $2t(t - 1) < 0$. Its unconstrained maximum Q and vertex are

$$Q = 2 \left[T^2 - \frac{(t - T/2)^2}{h^2} \right] + \frac{tC^2}{2(1 - t)}, \quad r_v = \frac{C}{2(1 - t)}. \quad (35)$$

Introduce

$$v = \frac{t}{A(1 - t)}, \quad w = pv(v + 1). \quad (36)$$

The root equation (30) gives

$$t = \frac{Av}{1 + Av}, \quad \alpha = \frac{Av^2}{v^2 + 1 - w}, \quad r_v = v/2. \quad (37)$$

The strict lower bound $K > A/(A + 1)$ gives $v > 1$, while $p > 0$ and $\alpha \leq A$ give $0 < w \leq 1$.

If $v > 4$, the vertex lies to the right of $[0, 2]$. The constrained maximum is then the endpoint $P_2(2, K) < 0$ from (34). Here $K > T/2$ follows from $K \geq L_3$ and $\alpha > 1/2$.

For $1 < v \leq 4$, substitution in (35) yields the exact identity

$$Q = \frac{Av R(v, w)}{2(v - 1)^2(1 + Av)^4(v + 1)^3}, \quad (38)$$

where the explicit polynomial R is given in Appendix B. Every denominator factor is positive. That appendix exhibits a tensor Bernstein expansion of $R(1 + 3x, w)$ on $[0, 1]^2$ with all coefficients nonpositive and its entire final row strictly negative. It follows that $R(v, w) < 0$ for $1 < v \leq 4$ and $0 < w \leq 1$. Thus $Q < 0$, and the maximum of $P_2(r, K)$ on $[0, 2]$ is strictly negative in this case as well.

In either case, continuity rules out thresholds immediately above K : the maximum on the fixed compact interval $[0, 2]$ of the quadratic $P_2(r, t)$ is continuous in t . A protocol strictly exceeding K would yield $P_2(r, t) \geq 0$ for every sufficiently close threshold $t > K$, contradicting its strictly negative maximum at K . This excludes the final encoder class and proves the theorem. $\square$

7 Four messages admit strict quantum advantage

A short classical ceiling and an explicit rational qubit protocol suffice for a dimension-matched separation. No claim that the following witness is optimal is needed.

Lemma 7.1 (A classical four-message ceiling). *For all $0 < p < 1/2$ and $0 \leq q \leq 1/2$,*

$$S_4^{\text{cl}}(p, q) \leq \frac{1}{2}. \quad (39)$$

Proof. The positive-denominator extremality argument reduces classical message encoders to the four functions C_0, C_1, I, X defined in Appendix A. If two labels have the same encoder, their normalized output distributions coincide. Their two correct probabilities therefore sum to at most one, so at least one is at most $1/2$.

Otherwise all four functions occur once. Their kernels obey $K_{C_0} + K_{C_1} = 1$ and $K_I + K_X = 1$ pointwise, so

$$\sum_m Z_m = \sum_{b, y, c} D(y, c | b) \sum_m K_m(b, c) = 4.$$

Each kernel is at most A , giving

$$\sum_m w_m(m) = \sum_{b, m, c} D(m, c | b) K_m(b, c) \leq 2A.$$

The Z_m -weighted average of message successes is therefore at most $A/2 < 1/2$. In particular their minimum is at most $1/2$. The deterministic reduction covers every stochastic protocol. $\square$

Theorem 7.2 (An exact four-message separation). *At $p = 1/10$, for every $0 \leq \epsilon < 1/76$,*

$$S_4^{\text{qu}}(1/10, \epsilon) > \frac{1}{2} \geq S_4^{\text{cl}}(1/10, \epsilon/2). \quad (40)$$

At $\epsilon = 0$, a rational protocol attains worst-message success $259/512$.

Proof. Use labels $0, 1, +, -$ and set $x = 9/16$. Define the decoder states by

$$\tau_{0,0} = x|1\rangle\langle 1|, \quad \tau_{1,1} = x|0\rangle\langle 0|, \quad \tau_{0,1} = \tau_{1,0} = 0, \quad (41)$$

$$\tau_{\pm,b} = \frac{1-x}{4} [I \pm (-1)^b \sigma_x]. \quad (42)$$

Each state is positive semidefinite, and each row has trace $x + 2(1-x)/2 = 1$. The effects are

$$E_0^0 = I, \quad E_0^1 = 0, \quad E_0^\pm = \frac{I \pm (3/5)\sigma_x + (4/5)\sigma_z}{2}, \quad E_1^m = I - E_0^m. \quad (43)$$

Since $(3/5)^2 + (4/5)^2 = 1$, the nonconstant effects are rank-one projectors. All displayed matrix entries are rational.

Let $k = (1-2p)(1-\epsilon)$. The reduced variables of the constants are $(T_0, u_0, \mathbf{d}_0) = (x, x, -x\mathbf{e}_z)$ and $(T_1, u_1, \mathbf{d}_1) = (x, -x, -x\mathbf{e}_z)$. The other two labels have $T_\pm = 1-x$, $u_\pm = 0$, and $\mathbf{d}_\pm = \pm(1-x)\mathbf{e}_x$. Hence $\mathbf{D} = -2x\mathbf{e}_z$. Equations (13) and (14) give

$$S_0 = S_1 = Ax, \quad Z_\pm = 1 - \frac{4}{5}kx, \quad S_\pm = \frac{(1-x)(1+3k/5)}{2(1-4kx/5)}. \quad (44)$$

At $p = 1/10$ and $x = 9/16$, these simplify exactly to

$$S_0 = S_1 = \frac{81}{160}, \quad Z_\pm = \frac{16+9\epsilon}{25}, \quad S_\pm = \frac{7(37-12\epsilon)}{32(16+9\epsilon)}. \quad (45)$$

In particular,

$$S_\pm - \frac{1}{2} = \frac{3(1-76\epsilon)}{32(16+9\epsilon)}. \quad (46)$$

Both constant successes exceed $1/2$, and (46) is positive precisely on the stated interval. Lemma 7.1 proves the matched classical comparison.

At $\epsilon = 0$, the four successes are

$$\left(\frac{81}{160}, \frac{81}{160}, \frac{259}{512}, \frac{259}{512} \right),$$

whose minimum is $259/512$. The protocol uses only the specified qubit: b and the recovered label are not supplied to the encoder as side information. The two projective effects do not commute, so the common-basis classical reduction does not apply. $\square$

Corollary 7.3 (The smallest separating alphabet). *Within the stated one-use dimension-two model, four is the smallest number of messages for which a qubit can outperform a classical bit at matched depolarizing/BSC noise.*

Proof. One and two messages have equal optima by Proposition 3.3; three have equal optima by Theorem 6.1; Theorem 7.2 supplies a four-message separation. $\square$

8 Scope and verification status

The scalar characterization is global over the stated local operations and finite alphabet, and includes constructive sufficiency. It does not assume a shared seed to convexify the feasible protocol set. The classical upper bounds likewise permit all allowed local stochastic encoders and decoders. The four-message witness proves separation on a positive-noise interval but is not presented as a global optimum at those parameters.

Exact symbolic identities, rational arithmetic checks, and direct matrix contractions were used during development. These checks supplement the proofs printed here. Their independence is limited to separate derivations within an AI-assisted workflow; they are not independent human review or proof-assistant certification. Independent human technical review and a fuller originality assessment remain pending. The conditional normalization is part of an assumed backward-channel model and does not supply a physical mechanism for backward communication.

Author and AI disclosure. Daniel A. Pyle is listed as Independent Researcher. Lyra, an OpenAI AI assistant, carried out the technical derivations, code-based verification, literature checking, and manuscript drafting in the research workflow. Lyra is disclosed as an AI tool and is not listed as an author. Independent human verification of the proofs and a fuller originality assessment remain pending.

Data and code availability. The research materials are available at <https://github.com/dapyle11/normalized-loop-memory>. The paper is self-contained: the proofs and the complete finite Bernstein coefficient certificate do not require external scripts or data files.

Classification. Primary subject: Information Theory (math.IT). Secondary subject: Quantum Physics (quant-ph). Mathematics Subject Classification: 94A40 (primary); 81P47, 81P45 (secondary).

A Complete classical three-message comparison

This appendix proves the classical ingredient in Theorem 6.1 for arbitrary $0 < p < 1/2$ and $0 \leq q \leq 1/2$. In this appendix put

$$A = 1 - p, \quad \alpha = 1 - p - q + 2pq, \quad \beta = 1 - \alpha. \quad (47)$$

Thus $1/2 \leq \alpha \leq A < 1$ and $p \leq \beta \leq 1/2$. At $q = \epsilon/2$, this α agrees with (8).

A.1 Deterministic reduction and a dual bound

The deterministic functions from a memory bit to a backward bit are constants C_0, C_1 , identity $I(d) = d$, and complement $X(d) = 1 - d$. Their loop kernels are

$$K_f(b, c) = \sum_d L_q(d \mid c) N_p(b \mid f(d)). \quad (48)$$

A constant has coefficient A in its matching b row and p in the other. Identity has coefficient α when $b = c$ and β otherwise; complement reverses these values. In particular,

$$K_{C_0} + K_{C_1} = 1, \quad K_I + K_X = 1. \quad (49)$$

The denominator of any constant encoder is one.

Every stochastic encoder is a convex combination of these four functions. For a fixed decoder, its normalized success is a denominator-weighted average of the four deterministic successes,

exactly as in Lemma 3.2. The choices for different messages are independent. Hence it suffices to bound all deterministic encoder triples while allowing an arbitrary stochastic decoder.

For a fixed encoder tuple f_m , positive weights λ_m , and a target t , define a weighted decoder-column expression

$$G(b, y, c) = \lambda_y K_{f_y}(b, c) - t \sum_m \lambda_m K_{f_m}(b, c). \quad (50)$$

If $G(b, y, c) \leq B_b$ for every column and $B_0 + B_1 \leq 0$, then

$$\sum_m \lambda_m [w_m(m) - t Z_m] = \sum_{b, y, c} D(y, c | b) G(b, y, c) \leq B_0 + B_1 \leq 0. \quad (51)$$

Because every weight and denominator is positive, not all message successes can be strictly greater than t . The row bounds B_b may be signed.

Simultaneous flips of the backward input/output bits and simultaneous flips of the forward input/output bits preserve the channel laws. Along with message permutations, these symmetries permit the canonical classes used below.

A.2 The attaining protocol and root properties

Let K be the root of (30). Set $x = K/A$ and choose encoders C_0, I, C_1 for labels 0, 1, 2. In row $b = 0$, the decoder outputs $(y, c) = (0, 1)$ with probability x and $(1, 0)$ with probability $1 - x$. In row $b = 1$, it outputs $(2, 0)$ with probability x and $(1, 1)$ with probability $1 - x$. The two constant successes are Ax , and the identity success is

$$\frac{\alpha(1 - x)}{\alpha(1 - x) + \beta x}. \quad (52)$$

Their equality is $\alpha(A - K)(1 - K) = \beta K^2$, precisely (30). Since $0 < K < A$, this is a valid decoder.

For completeness, $F(0) = \alpha A > 0$, $F(A) = -\beta A^2 < 0$, and on $[0, A]$,

$$F'(t) = -\alpha(A + 1) + 2(\alpha - \beta)t \leq -\alpha(A + 1) + 2\alpha A < 0.$$

Thus there is exactly one root in $(0, A)$. At $t_0 = A/(A + 1)$,

$$F(t_0) = (\alpha - \beta) \frac{A^2}{(A + 1)^2} \geq 0,$$

with equality exactly when $\alpha = \beta = 1/2$. Hence $K \geq A/(A + 1) > A/2$, with strict first inequality when $q < 1/2$. Expanding the root equation gives $(2\alpha - 1)K^2 - \alpha(A + 1)K + A\alpha = 0$. Rationalizing its smaller root gives (31), also valid when $\alpha = 1/2$.

A.3 Three distinct encoders

Up to the stated symmetries there are two possibilities.

Both constants and one bijection. Take $\{C_0, I, C_1\}$. At $t = K$, choose

$$\lambda_{C_0} = \lambda_{C_1} = u, \quad \lambda_I = 1, \quad u = \frac{\beta K}{A - K} = \frac{\alpha(1 - K)}{K} > 0.$$

The equality of the two expressions for u is the root equation. In row $b = 0$, the largest C_0 column in (50) is $u(A - K) - \beta K = 0$, and the largest identity column is $\alpha(1 - K) - uK = 0$. Every C_1 column is no larger than the corresponding C_0 column because $p \leq A$. Row $b = 1$ is symmetric. Thus $B_0 = B_1 = 0$ proves the upper bound K .

One constant and both bijections. Take $\{C_0, I, X\}$. Its exact optimum is

$$U = \frac{2A\alpha}{2A + \alpha}. \quad (53)$$

For the upper bound use weights $\lambda_{C_0} = z = \alpha/A$ and $\lambda_I = \lambda_X = 1$. Write $N_0 = A$, $N_1 = p$. The maximum column in row b is

$$B_b = \max\{zN_b(1 - t) - t, \alpha - t - zN_bt\}.$$

The two expressions tie in row zero; the second dominates in row one. Hence $B_0 + B_1 = 2\alpha - t(2 + \alpha/A)$, which vanishes at $t = U$.

To attain U , let $y = 2\alpha/(2A + \alpha)$. In row zero, put probability $y/2$ on each of $(C_0, 0), (C_0, 1)$ and probability $(1 - y)/2$ on each of $(I, 0), (X, 1)$. In row one, put probability $1/2$ on each of $(I, 1), (X, 0)$. All denominators are one, and the successes are Ay and $\alpha(2 - y)/2$, all equal to U .

Finally $U \leq K$. Direct expansion gives

$$F(U) = \frac{A\alpha}{(2A + \alpha)^2} J, \quad (54)$$

$$J = \alpha^2(2\alpha - 1) + 2\alpha(A - \alpha)(2 - \alpha) + 4(A - \alpha)^2(1 - \alpha) \geq 0.$$

Every displayed summand is nonnegative, and $0 < U < A$. Since F decreases on $[0, A]$ and $F(K) = 0$, the claim follows.

A.4 Repeated encoders

If all three encoders coincide, their normalized output distributions coincide, so the worst success is at most $1/3 < K$.

Otherwise exactly two labels share an encoder. Their decoder probabilities may be symmetrized without reducing the smaller of their successes: this preserves every encoder denominator and every other label's correct weight. Let G be the repeated pair's output group, let x_b be its total decoder mass in row b , and put $X = x_0 + x_1 \in [0, 2]$. Each repeated label has half the group's success. Define the useful bounds

$$u(X) = \frac{\alpha X}{2[\alpha X + \beta(2 - X)]}, \quad v(X) = \frac{\alpha(2 - X)}{\beta X + \alpha(2 - X)}. \quad (55)$$

The first bounds the per-label success of a repeated bijection, and the second bounds the success of a singleton bijection. They follow by bounding a bijection's correct coefficients by α and its incorrect coefficients below by β . The function u increases and is at most $1/2$, while v decreases.

Repeated constant and singleton opposite constant. Orient the repeated constant so its favorable row is zero. The successes are

$$\frac{Ax_0 + px_1}{2}, \quad 1 - px_0 - Ax_1.$$

If both are at least a target t , then $Ax_0 + px_1 \geq 2t$ and $px_0 + Ax_1 \leq 1 - t$. Multiply the first inequality by A and subtract p times the second to obtain

$$(2A + p)t - p \leq (A^2 - p^2)x_0 \leq A - p.$$

Thus $t \leq A/(2A + p) = A/(A + 1) \leq K$.

Repeated constant and singleton bijection. Reaching $K > A/2$ requires $X > 1$. At fixed $X > 1$, repeated-constant success is at most $[A + p(X - 1)]/2$, because $x_0 \leq 1$. The singleton success is at most $v(X)$. By the root identity,

$$v(X) \geq K \implies X \leq \frac{2\alpha(1-K)}{\alpha(1-K) + \beta K} = \frac{2K}{A}. \quad (56)$$

But the repeated messages require

$$X \geq 1 + \frac{2K - A}{p} > \frac{2K}{A},$$

where the strict inequality follows because the difference, multiplied by Ap , is $(A-p)(2K-A) > 0$. This class cannot attain K .

Repeated bijection and singleton constant. The repeated messages are bounded by $u(X)$. For $X \leq 1$ this is at most $\alpha/2 \leq A/2 < K$. For $X > 1$, the singleton constant is at most $A(2-X)$. If $K > 1/2$, the universal bound $u(X) \leq 1/2$ already excludes this class. Otherwise singleton success at least K requires $X \leq 2 - K/A$, and

$$u(2 - K/A) = \frac{\alpha(2A - K)}{2[\alpha(2A - K) + \beta K]} < K.$$

The strict inequality is equivalent to

$$2\beta K^2 - \alpha(2A - K)(1 - 2K) = \alpha K(2A - 1) > 0,$$

using (30). This excludes the class.

Repeated bijection and singleton opposite bijection. The bounds are $u(X)$ and $v(X)$. Again $K > 1/2$ is excluded immediately. For $K \leq 1/2$, equation (56) forces $X \leq 2K/A$. At this endpoint, $u(2K/A) < K$ is equivalent to

$$\Delta := 2A\beta - \alpha + 2(\alpha - \beta)K > 0.$$

Indeed,

$$\begin{aligned} \Delta &= 2K - 1 + \beta(2A + 1 - 4K) \\ &\geq 2K - 1 + p(2A + 1 - 4K) \\ &= (2A - 1)(2K - A) > 0. \end{aligned}$$

The inequality uses $\beta \geq p$ and $2A + 1 - 4K \geq 2A - 1 > 0$. This excludes the last repeated class.

These cases exhaust every deterministic triple, hence every stochastic encoder triple. Together with the attaining decoder they prove $S_3^{\text{cl}}(p, q) = K$ for the full declared classical model.

B An explicit Bernstein sign certificate

This appendix supplies all coefficients needed to prove the strict polynomial sign used in (38). Define

$$\begin{aligned} R(v, w) &= v^9 - 5v^8 - 2v^7w - 7v^7 + 14v^6w - v^6 + v^5w^2 + 24v^5w - 9v^5 \\ &\quad - 9v^4w^2 + 12v^4w - 19v^4 - 25v^3w^2 + 50v^3w - 25v^3 \\ &\quad - 39v^2w^2 + 46v^2w - 7v^2 + 16vw^3 - 8vw^2 - 32vw + 24v \\ &\quad + 16w^2 - 32w + 16. \end{aligned} \quad (57)$$

Substituting (37), $p = w/[v(v+1)]$, $A = 1 - p$, $h = \alpha - 1/2$, $C = t/A$, and $T = 1 - C/2$ into (35), and clearing denominators, gives (38) with (57). Thus the displayed polynomial and rational identity can be checked by finite polynomial multiplication, without a sign oracle or a numerical approximation.

For $n \geq 0$, write the Bernstein basis as

$$B_i^n(x) = \binom{n}{i} x^i (1-x)^{n-i}, \quad 0 \leq i \leq n.$$

The exact degree-(9, 3) expansion is

$$R(1+3x, w) = \sum_{i=0}^9 \sum_{j=0}^3 b_{ij} B_i^9(x) B_j^3(w), \quad (58)$$

where the full coefficient table is

i	b_{i0}	b_{i1}	b_{i2}	b_{i3}
0	-32	-16/3	0	0
1	-368/3	-416/9	-112/9	0
2	-1054/3	-1519/9	-1141/18	-17/2
3	-6493/7	-3667/7	-21037/84	-2109/28
4	-50678/21	-98048/63	-8170/9	-3012/7
5	-132970/21	-288037/63	-197195/63	-13744/7
6	-233775/14	-555439/42	-214544/21	-53376/7
7	-517625/12	-1318745/36	-277048/9	-25536
8	-310000/3	-827200/9	-732416/9	-71680
9	-200000	-544000/3	-163840	-147456

To verify the coefficient conversion directly, if $R(1+3x, w) = \sum_{a=0}^9 \sum_{b=0}^3 c_{ab} x^a w^b$, then

$$b_{ij} = \sum_{a=0}^i \sum_{b=0}^j c_{ab} \frac{\binom{i}{a}}{\binom{9}{a}} \frac{\binom{j}{b}}{\binom{3}{b}}. \quad (59)$$

This follows from the identity

$$x^a = \sum_{i=a}^n \frac{\binom{i}{a}}{\binom{n}{a}} B_i^n(x),$$

which is obtained using $\binom{n}{i} \binom{i}{a} = \binom{n}{a} \binom{n-a}{i-a}$ and the binomial theorem. Thus (57), (59), and the table form a self-contained finite certificate of (58).

Each B_i^n is nonnegative on $[0, 1]$, and the basis sums to one. All b_{ij} are nonpositive. Moreover, for $x > 0$ the final row contributes

$$x^9 \sum_{j=0}^3 b_{9j} B_j^3(w) < 0,$$

because all four coefficients in that row are strictly negative. The remaining rows contribute nonpositive quantities. Therefore

$$R(1+3x, w) < 0 \quad (0 < x \leq 1, 0 \leq w \leq 1), \quad (60)$$

including both w endpoints and $x = 1$. With $x = (v-1)/3$, this proves the sign asserted for $1 < v \leq 4$ and the physical $0 < w \leq 1$.

References

- [1] K. Ji, S. Lloyd, and M. M. Wilde, “Retrocausal capacity of a quantum channel: Communicating through noisy closed timelike curves,” *Physical Review Letters* **136**, 230801 (2026). doi:10.1103/znyd-npk5. arXiv:2509.08965v3.
- [2] K. Ji, B. Regula, and M. M. Wilde, “Postselected communication over quantum channels,” *International Journal of Quantum Information* **22**(05), 2440012 (2024). doi:10.1142/S0219749924400124. arXiv:2308.02583v3.
- [3] R. Gallego, N. Brunner, C. Hadley, and A. Acín, “Device-independent tests of classical and quantum dimensions,” *Physical Review Letters* **105**, 230501 (2010). doi:10.1103/PhysRevLett.105.230501. arXiv:1010.5064.
- [4] J. Bowles, M. T. Quintino, and N. Brunner, “Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices,” *Physical Review Letters* **112**, 140407 (2014). doi:10.1103/PhysRevLett.112.140407. arXiv:1311.1525v3.