Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Optimal entanglement swapping with one auxiliary photon: exact fidelity limit and purity requirements

Annan Chen

annan.chen24@imperial.ac.uk

Imperial College London https://orcid.org/0009-0000-1308-0466

Research Article

Keywords: entanglement swapping, linear optics, partial distinguishability, auxiliary photons, Bell-state measurement, multiphoton interference, source purity

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11714344/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Optimal entanglement swapping with one auxiliary photon: exact fidelity limit and purity requirements

Annan Chen

Department of Physics, Imperial College London, London, United Kingdom

ORCID: 0009-0000-1308-0466 annan.chen24@imperial.ac.uk

Abstract

Wave-packet mismatch in entanglement swapping leaves which-source information and lowers the fidelity of the heralded remote pair. We determine the exact correction achievable with one independent auxiliary photon. For pure data photons with amplitude overlap c, every finite static passive network acting identically on internal states, followed by photon counting and deterministic local recovery, yields conditional Bell fidelity at most $(1 + c + c^2)/(2 - c + 2c^2)$. A five-mode circuit with four two-mode mixings attains this limit and converts first-order mismatch error to second order, whereas circuits with at most three mixings in the logical-mode basis cannot exceed $(1 + c)/2$. At two-photon visibility 0.98, a pure auxiliary reduces the error from 1.0×10^{-2} to 3.4×10^{-5} at heralding probability 0.041. For mixed auxiliary states supported on the span of the data wave packets, quadratic suppression is possible if and only if the purity deficit vanishes at least quadratically in the mismatch. We establish a matching error law at fixed positive heralding probability. With data visibility 0.90 and auxiliary-copy visibility 0.99, the aligned circuit reduces Bell error 6.6-fold. A centered Gaussian auxiliary of optimized width preserves the optimal leading error.

1 Introduction

Entanglement swapping connects remote quantum systems by measuring their travelling photons at a central station. Temporal or spectral mismatch that the detectors do not resolve leaves which-source information and lowers the Bell fidelity of the retained pair. How much error can one independent auxiliary photon remove, and how pure must that photon be? The answers set source requirements for the heralded pairs used in quantum networks [1–4].

We determine the exact answer for pure data photons and passive optics that acts identically on their internal states. For amplitude overlap $0 < c < 1$, the maximum conditional Bell fidelity is

$$F_{\mathrm{opt}}(c) = \frac{1 + c + c^2}{2 - c + 2c^2}. \tag{1}$$

The limit holds for every finite static passive network, every independent auxiliary state, every detection pattern and every deterministic local recovery. A five-mode circuit attains it with four two-mode mixings. Circuits with at most three mixings, counted in the logical-mode basis, cannot exceed $(1+c)/2$ under deterministic LOCC, so four are minimal. Writing $t = 1 - c$, the optimal error is $t^2/3 + O(t^3)$ at a heralding probability tending to $1/24$. For a pure auxiliary and data-photon Hong–Ou–Mandel (HOM) visibility 0.98, the error falls from 1.0×10^{-2} to 3.4×10^{-5}, a factor of 294. Table 1 summarizes the resource hierarchy and its probabilities.

Photon-indistinguishability distillation improves a surviving photon's internal state by interference and heralding [5–9]. Hoch et al. optimize collective overlaps and phases, including quadratic visibility-error suppression [10]. Fusion protocols also transform distinguishability in surviving entangled resources, with a response that depends on the heralding pattern and source model [11]. Here the retained object is an encoded remote Bell pair. A distillation stage followed by a Bell analyzer is bounded by Eq. (1) whenever the entire composition is a static passive network with one auxiliary photon and final photon counting. The one-rail composition in Sec. 5 illustrates the distinction: its wave-packet visibility improves quadratically, while its remote Bell error remains first order. The circuit of Fig. 1(b) instead interferes the auxiliary with modes that already combine both data photons and attains the encoded-pair limit.

The advantage depends on auxiliary purity. For an auxiliary density H supported on the span of the data wave packets, quadratic Bell error is possible precisely when $1 - \mathrm{Tr}\, H^2 = O(t^2)$. At lower purity, tilting the auxiliary's internal Bloch vector away from the midpoint axis and adjusting the optical imbalance can improve the optimal error scaling. At a fixed positive heralding-probability floor, we prove an error law in the mismatch and impurity that is tight up to constant factors. A centered Gaussian photon with optimized width already preserves the leading optimum for symmetric Gaussian data wave packets.

Which correction is preferable depends on the available hardware. In the Gaussian model, selective temporal or spectral filtering can attain unit fidelity at higher probability than the auxiliary circuit (Sec. 5). Our results give the exact benchmark for mode-nonselective correction by path interference and photon counting, including static distillation followed by Bell analysis. The two methods also compose: if a filter leaves pure

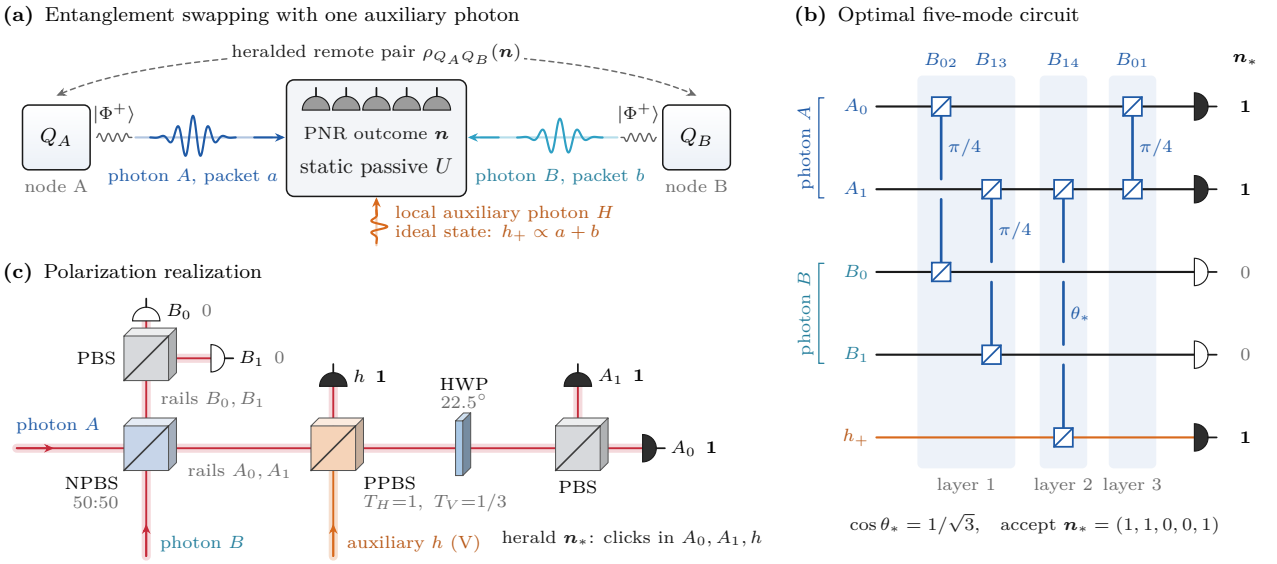


Figure 1: Entanglement swapping with one auxiliary photon. (a) Photons A and B, entangled with remote qubits Q_A and Q_B, carry mismatched wave packets a and b. At the central station they meet a local auxiliary photon H, ideally in the midpoint state $h_+ \propto a + b$, in a fixed passive network U; the photon-number pattern $\boldsymbol{n}$ heralds the remote state $\rho_{Q_AQ_B}(\boldsymbol{n})$. (b) Optimal five-mode circuit $U_{\text{opt}} = B_{01}(\pi/4)B_{14}(\theta_*)B_{13}(\pi/4)B_{02}(\pi/4)$, $\cos\theta_* = 1/\sqrt{3}$, with auxiliary h_+ and accepted pattern $\boldsymbol{n}_* = (1,1,0,0,1)$. Squares mark the two rails coupled by each mixer; breaks mark rails that a mixer crosses without coupling. Mixers in one shaded layer act in parallel. (c) Polarization realization: a 50:50 NPBS (B_{02}, B_{13}), a PPBS with $T_H = 1$, $T_V = 1/3$ (B_{14}) and a HWP at 22.5° (B_{01}), followed by H/V analysis with polarizing beam splitters. In (b,c), filled detectors register one photon each in the accepted pattern; open detectors register none (Appendix D.4).

Table 1: Resource hierarchy for pure data photons ($t = 1 - c$), with calibrated independent auxiliary photons. P is conditional on supplying all required photons and refers to the stated acceptance. Ceilings hold under LU or deterministic LOCC; the one-auxiliary ceiling also holds under deterministic SEP. The mixed row assumes $\operatorname{supp} H \subseteq \operatorname{span}\{a, b\}$ and fixed $0 < p_0 < 1/64$.

Resource	Fidelity limit	Leading Bell error	Attaining P	Acceptance; result
No auxiliary	$(1 + c^2)/2$	t	$1/2$	4 patterns; App. C
One auxiliary, ≤ 3 mixings	$(1 + c)/2$	$t/2$	$(1 + c)/16$	2 patterns; Thm. 2
One auxiliary, any finite network	$F_{\text{opt}}(c)$	$t^2/3$	$P_\star(c)$	1 pattern; Thm. 1
One mixed auxiliary, $P \geq p_0$	—	$\asymp \min\{t^2 + \eta_h, t\sqrt{t + \eta_h}\}$	$\geq p_0$	Thms. 3, 4
Two pure auxiliaries	1	0	$(1 + c)^2/128$	1 pattern; Sec. 5
Gaussian filters[a]	1	0	$P_G(c, 1)$	Bell analyzer; Sec. 5

[a]Internal-mode filters act selectively on internal states, extending the control class of Theorem 1.

wave packets with overlap $c' < 1$, a pure auxiliary matched to them reduces the residual Bell error from first order in $t' = 1 - c'$ to $t'^2/3 + O(t'^3)$. Near unit overlap, the circuit's single-pattern acceptance is about one twelfth that of a conventional Bell analyzer behind the same filters. A useful source regime combines remote-emitter mismatch with high local-photon reproducibility [6, 12, 13]: at data visibility 0.90 and auxiliary-copy visibility 0.99, the aligned circuit reduces Bell error 6.6-fold (Fig. 4).

Auxiliary-assisted Bell-state discrimination addresses the complementary efficiency problem of identifying an ideal logical Bell state [14–18]. Boosted swapping with a two-photon ancillary state raises the theoretical discrimination probability from 1/2 to 5/8 [19]. We instead optimize the fidelity of the heralded remote pair for imperfect input photons. The proof combines partial-distinguishability theory [20–22] with amplitude constraints shared by one optical network.

Theorem 1 gives the exact one-auxiliary fidelity limit and its equality conditions, and Theorem 2 proves that four elementary mixings are minimal for attaining it in the logical-mode basis. Theorems 3 and 4 establish the necessary and sufficient auxiliary-purity criterion and a matching error law at a fixed positive heralding-probability floor. Section 5 translates these results into

measurable source budgets, an optimized Gaussian auxiliary and correction after filtering.

Sections 2–5 develop the model, exact limit, purity laws and source-design implications. Appendices A–F contain the complete proofs; additional calculations are provided in the Supplemental Material [23].

2 Model and objectives

Let Q_A, Q_B be the retained remote qubits. Photon A occupies rails (A_0, A_1) and photon B occupies (B_0, B_1), with ideal input states

$$|\Omega_a\rangle_{Q_A A} = \frac{1}{\sqrt{2}} \sum_{j=0}^{1} |j\rangle_{Q_A} d_j^\dagger(a)|\mathrm{vac}\rangle,$$
$$|\Omega_b\rangle_{Q_B B} = \frac{1}{\sqrt{2}} \sum_{k=0}^{1} |k\rangle_{Q_B} d_{2+k}^\dagger(b)|\mathrm{vac}\rangle. \qquad (2)$$

Here $d_i^\dagger(a)$ creates a photon in spatial mode i and normalized internal state a. Each source has the same internal state on both of its logical rails. We choose wave-packet phases so that $\langle a|b\rangle = c \in (0,1)$ is real. One independent auxiliary photon enters input 4 with density matrix H; every additional input is vacuum. The wave packets define the internal subspace $\mathcal{S} = \mathrm{span}\{a, b\}$ and its orthonormal basis

$$h_\pm = \frac{a \pm b}{\sqrt{2(1 \pm c)}}. \qquad (3)$$

The midpoint h_+ is equivalently the principal eigenvector of $|a\rangle\langle a| + |b\rangle\langle b|$. This fixes h_+ independently of the phase convention.

A fixed finite passive unitary $U \in U(m)$ transforms every internal state in the same way, $d_i^\dagger(\varphi) \mapsto \sum_j U_{ji} d_j^\dagger(\varphi)$. Complete photon-number-resolving (PNR) detection gives the full spatial photon-count outcome $\boldsymbol{n}$, with $\sum_i n_i = 3$, and traces the internal labels. There is no intermediate optical measurement or feed-forward. The unnormalized remote state is $J_{\boldsymbol{n}}$, with

$$p_{\boldsymbol{n}} = \mathrm{Tr}\, J_{\boldsymbol{n}}, \qquad \sum_{\boldsymbol{n}} J_{\boldsymbol{n}} = I_4/4. \qquad (4)$$

We use $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$ and $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. The permitted outcome-dependent remote recovery class $\mathcal{R}$ is local unitaries (LU), deterministic local operations and classical communication (LOCC), or deterministic separable maps (SEP). The last class consists of product Kraus maps

$$\Lambda(J) = \sum_j (L_j \otimes M_j) J (L_j \otimes M_j)^\dagger,$$
$$\sum_j L_j^\dagger L_j \otimes M_j^\dagger M_j = I_4. \qquad (5)$$

For $p = \mathrm{Tr}\, J > 0$, define

$$F_{\mathcal{R}}(J) = \sup_{\Lambda \in \mathcal{R},\, \Phi\,\mathrm{ME}} \frac{\langle \Phi|\Lambda(J)|\Phi\rangle}{p}, \qquad e_{\mathcal{R}} = 1 - F_{\mathcal{R}}, \qquad (6)$$

Table 2: Notation used throughout the main text.

c, t	amplitude overlap and mismatch $t = 1-c$
D	$1 - c^2$
H, r_h, η_h	auxiliary density, Bloch length, impurity $1 - r_h$
h_a, h_b	coefficients of $h = h_a a + h_b b$
r_d	Bloch length of depolarized data photons
η_0	positive constant in $\eta_h = \eta_0 t^\beta$
N_c	$2 - c + 2c^2$
$P_\star(c)$	largest single-pattern probability at F_{opt}
$P_{\mathrm{mid}}(c, r_h)$	probability for the mixed midpoint construction
$J_{\boldsymbol{n}}, p_{\boldsymbol{n}}$	unnormalized remote state and its trace
w_Φ, f_Φ	Bell weight and fixed-target fidelity $w_\Phi / \mathrm{Tr}\, J$
$F_{\mathcal{R}}$	fidelity optimized over deterministic recovery $\mathcal{R}$
$P, \bar{F}$	total accepted probability and average fidelity
$E, E_{p_0,\mathcal{R}}$	best error without / with probability floor p_0

where ME means maximally entangled. Fixed-target fidelity is denoted $f_\Phi(J) = \langle\Phi|J|\Phi\rangle/p$. The LU case is the fully entangled fraction; deterministic LOCC optimization is also the relevant fidelity measure in optimal single-copy teleportation [24].

Probabilities are per supplied photon triple. For one fixed U, H, classical acceptance weights $0 \le \omega_{\boldsymbol{n}} \le 1$ give

$$P = \sum_{\boldsymbol{n}} \omega_{\boldsymbol{n}} p_{\boldsymbol{n}}, \qquad \bar{F} = \frac{\sum_{\boldsymbol{n}} \omega_{\boldsymbol{n}} p_{\boldsymbol{n}} F_{\mathcal{R}}(J_{\boldsymbol{n}})}{P}. \qquad (7)$$

An inequality that holds outcome by outcome also bounds the accepted average. The single-outcome optimum fixes conditional fidelity; optimizing the acceptance weights determines the total probability at a target fidelity.

The exact pure-data ceiling allows H in an arbitrary internal Hilbert space. The purity results impose the additional condition $\mathrm{supp}\, H \subseteq \mathcal{S}$. Our mixing-count statement uses the fixed input logical rails and output counting basis: every elementary two-mode mixing used to change either basis is counted. Fixed phases and relabeling are free.

Rejected optical patterns are discarded; accepted ones are recovered deterministically, so Eq. (6) keeps denominator $\mathrm{Tr}\, J$. Filtering and internal-state-resolved detection are additional capabilities compared in Sec. 5.

3 Fidelity limit and realization

3.1 The sharp all-network bound

Theorem 1 (One-auxiliary fidelity limit)**.** *In the model of Sec. 2, for every independent auxiliary density, every finite passive network and every positive-probability detection pattern,*

$$F_{\mathcal{R}}(J_{\boldsymbol{n}}) \le F_{\mathrm{opt}}(c) = \frac{1 + c + c^2}{2 - c + 2c^2} \qquad (8)$$

for $\mathcal{R} = \mathrm{LU}, \mathrm{LOCC}, \mathrm{SEP}$. Equality requires the pure midpoint auxiliary h_+. At equality, the largest probability of a single detection pattern is

$$P_\star(c) = \frac{2 - c + 2c^2}{72}. \tag{9}$$

A five-mode circuit with LU recovery attains both bounds.

Corollary 1 (Static distillation followed by Bell analysis)**.** *A passive distillation stage and Bell analyzer using one independent auxiliary photon obey Eq. (8) whenever they compose one static network with all photon counts at the end and deterministic remote recovery.*

The composed transfer matrix is a finite passive unitary, so Theorem 1 applies directly.

Any finite network reduces to a small active matrix. Three photons occupy at most three distinct output modes. Only those output rows and the five potentially populated input columns enter their conditional amplitudes. Their active block T obeys $TT^\dagger \preceq I$. Conversely, any such rectangular contraction has a unitary dilation on at most eight modes [25]; its extra inputs are vacuum.

Write each active row as $(\alpha_i^T, \beta_i^T, v_i)$, where α_i, β_i are two-component logical amplitudes and v_i is its auxiliary amplitude. For three distinct occupied outputs, a pure auxiliary supported on $\mathcal{S}$ is $h = h_a a + h_b b$. Repeated internal labels combine the six photon assignments into product vectors

$$\begin{aligned} x_i &= \alpha_i \otimes (v_k\beta_j + v_j\beta_k), \\ y_i &= (v_k\alpha_j + v_j\alpha_k) \otimes \beta_i, \quad \{i,j,k\} = \{0,1,2\}. \end{aligned} \tag{10}$$

They share the same sum, $\sum_i x_i = \sum_i y_i$. This shared-sum constraint sharpens the bound. With $W = (h_b x_0, h_b x_1, h_b x_2, h_a y_0, h_a y_1, h_a y_2)/2$, the state is $J = WGW^\dagger$. The Gram matrix of the internal wave-packet assignments is given in Appendix A. Compression to the shared-sum subspace leaves eigenvalues $\ell, \mu_-, \mu_-, \mu_+, \mu_+$, where

$$\begin{aligned} \mu_\pm &= (1 - c^2)(1 \pm c), \\ \ell &= 1 + 2c^2 + (2c + c^3)\chi_h, \qquad \chi_h = \frac{2\,\mathrm{Re}(\overline{h_a} h_b)}{|h_a|^2 + |h_b|^2}. \end{aligned} \tag{11}$$

Put $p = \mathrm{Tr}\, J$ and $w_\Phi = \langle\Phi|J|\Phi\rangle$. For any maximally entangled target, a product column has squared overlap at most half its squared norm. If the compressed Gram spectrum lies in $[\lambda_{\min}, \lambda_{\max}]$, the part orthogonal to that target has squared coefficient norm at least that of the target component. Weighting these two parts by $\lambda_{\min}$ and $\lambda_{\max}$ gives

$$p - w_\Phi \geq \frac{\lambda_{\min}}{\lambda_{\max}} w_\Phi, \qquad f_\Phi \leq \frac{\lambda_{\max}}{\lambda_{\max} + \lambda_{\min}}. \tag{12}$$

Maximizing this expression over $\chi_h \in [-1, 1]$ selects $\chi_h = 1$. The largest eigenvalue is then $(1 + c)(1 + c + c^2)$ and the smallest is $(1 + c)(1 - c)^2$, yielding Eq. (8). Local product Kraus operators preserve both the product columns and their shared-sum constraint. The homogeneous bound therefore holds branchwise and sums to the same ceiling under deterministic SEP recovery.

Bunched outcomes and auxiliary components orthogonal to $\mathrm{span}\{a, b\}$ have strictly lower ceilings. Appendix A covers all occupations, arbitrary auxiliary mixtures and finite vacuum extensions.

Saturation also fixes the geometry of the logical amplitudes. After local frame alignment, each paired product column is proportional to a rank-one projector P_i, and $\sum_{i=0}^{2} P_i = 3I_2/2$. The three corresponding unit Bloch vectors therefore have pairwise dot product $-1/2$: a trine (three directions at 120°) in the logical Bloch sphere, distinct from the auxiliary's internal Bloch sphere. The resulting conditional state has Bell spectrum $(F_{\mathrm{opt}}, (1 - F_{\mathrm{opt}})/2, (1 - F_{\mathrm{opt}})/2, 0)$. Optimizing the row and source-group gains of this equality family subject to $TT^\dagger \preceq I$ gives Eq. (9). Concurrence rigidity transfers the equality classification and probability bound to deterministic recovery (Appendix B). At $c = 1/2$, the fidelity is $7/8$ and this heralding probability is $1/36$; at $c = 0.9$, they are $271/272$ and $17/450$.

3.2 Four-mixing realization

Define the real two-mode mixing $B_{ij}(\theta)$ by its nontrivial block

$$B_{ij}(\theta)|_{ij} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{13}$$

Products act right to left. The network

$$\begin{aligned} U_{\mathrm{opt}} &= B_{01}(\pi/4) B_{14}(\theta_*) B_{13}(\pi/4) B_{02}(\pi/4), \\ \cos\theta_* &= 1/\sqrt{3}, \end{aligned} \tag{14}$$

is used with auxiliary h_+ and the accepted detection pattern $(1, 1, 0, 0, 1)$. Its unnormalized remote Bell weights, in the order $(\Phi^+, \Phi^-, \Psi^+, \Psi^-)$, are

$$\left(\frac{1 + c + c^2}{72}, \frac{(1 - c)^2}{144}, \frac{(1 - c)^2}{144}, 0 \right). \tag{15}$$

Their sum and largest weight give the two quantities in Theorem 1. The accepted output is already aligned with Φ^+ in this convention. A complete matrix and assignment calculation appear in Appendix D.

The error cancellation can be seen before tracing out any internal label. In the midpoint basis write

$$\begin{aligned} a &= \sqrt{(1 + c)/2}\, h_+ + \sqrt{(1 - c)/2}\, h_-, \\ b &= \sqrt{(1 + c)/2}\, h_+ - \sqrt{(1 - c)/2}\, h_-. \end{aligned}$$

The unwanted Bell amplitudes have zero total coefficient for the all-h_+ word. Pairing assignments that exchange the two data photons also cancels every word with exactly one h_- label. Their first surviving amplitude is proportional to $\langle h_-|a\rangle^2 = (1 - c)/2$, so their weights scale as $(1 - c)^2$. The antisymmetric three-label amplitude vanishes because the occupied internal

space is two dimensional. Writing $e_{\rm opt}(c) = 1 - F_{\rm opt}(c)$ gives

$$e_{\rm opt}(c) = \frac{t^2}{3 - 3t + 2t^2} = \frac{t^2}{3} + O(t^3). \tag{16}$$

The herald postselects a coherent cancellation of the unwanted Bell amplitudes, and Theorem 1 makes the coefficient 1/3 optimal over the network class. The target weight is $\mathrm{Tr}[\Pi_{\rm sym}(\rho_a \otimes \rho_b \otimes H)]/24$, the fully symmetric three-photon internal-state weight. Here $\rho_a = |a\rangle\langle a|$, $\rho_b = |b\rangle\langle b|$, and $\Pi_{\rm sym}$ projects onto the symmetric three-factor subspace. The unassisted ceiling $(1 + c^2)/2$ is the corresponding two-photon symmetric weight (Appendix D.2).

In polarization encoding, one NPBS supplies B_{02}, B_{13}, a PPBS with $T_H = 1, T_V = 1/3$ supplies B_{14}, and a HWP at 22.5° supplies B_{01} before H/V analysis [26]. Figure 1(c) and Appendix D.4 give the implementation. At perfect source overlap, small angle offsets give

$$1 - F = 2\delta_{14}^2 + \frac{4}{3}\delta_{01}^2 + O(\|\boldsymbol{\delta}\|^3). \tag{17}$$

Offsets of 3 mrad in both sensitive angles give an error floor of 3.0×10^{-5}, comparable to the mismatch error $t^2/3$ at $c = 0.99$ (Supplemental Sec. S4.1).

3.3 Minimum mixing count

Theorem 2 (Minimum number of elementary mixings)**.** *In the logical-mode basis defined in Sec. 2, any static circuit using at most three arbitrary two-mode mixings, any independent auxiliary and any finite vacuum extension satisfies*

$$F \le \frac{1 + c}{2} \tag{18}$$

under LU or deterministic single-copy LOCC recovery. This bound is attainable. Four mixings are necessary and sufficient to reach $F_{\rm opt}(c)$ for $0 < c < 1$ in this recovery class.

The proof classifies the interaction graph using photon numbers conserved in each connected component. A nontrivial auxiliary-assisted outcome with only three edges must have a three-mode component and a two-mode component, each connecting the appropriate logical rails. The edges are a tree; phases are absorbed into known input and output frames. Other component patterns isolate the auxiliary, factorize the remote state, or fix a logical bit. Adding vacuum vertices cannot create an uncounted connection.

For the surviving tree, data-first and both data–auxiliary-first orders are covered by the two internal responses $E_\pm = I \pm H$. With pure data, their normalized odd-parity coherence is

$$\xi_\pm = c\frac{|\langle b|E_\pm|a\rangle|}{\sqrt{\langle a|E_\pm|a\rangle\langle b|E_\pm|b\rangle}} \le c. \tag{19}$$

The inequality is Cauchy–Schwarz in the positive response metric. Zero-population responses have zero coherence. Balancing the surviving logical populations gives $(1 + \xi_\pm)/2$; concurrence monotonicity preserves the bound under deterministic LOCC. The difference auxiliary h_- with $B_{04}(\pi/4)B_{13}(\pi/4)B_{02}(\pi/4)$ attains $(1 + c)/2$. Its outcome $(1, 1, 0, 0, 1)$ has probability $(1 + c)/32$; accepting the mirror-image pattern $(1, 0, 0, 1, 1)$ as well doubles this probability. The auxiliary state is optimized independently for each circuit class.

The strict gap

$$F_{\rm opt} - \frac{1 + c}{2} = \frac{c(1 - c)(1 + 2c)}{2(2 - c + 2c^2)} > 0 \tag{20}$$

proves minimality throughout the open interval. Figure 2 shows both the fidelity gap and the change of error order. Appendix C gives the complete ordering and vacuum-extension proof.

4 Auxiliary purity and direction

4.1 The purity needed for quadratic error

In this section, $\mathrm{supp}\, H \subseteq \mathcal{S}$. In the midpoint basis, write

$$H = \tfrac{1}{2}(I + \mathbf{h}\cdot\boldsymbol{\sigma}), \quad \mathbf{h} = (x, y, z), \quad |\mathbf{h}| = r_h, \tag{21}$$

where $\boldsymbol{\sigma}$ is the Pauli vector and $0 \le r_h \le 1$. The eigenvalues are $(1 \pm r_h)/2$. Thus r_h controls purity while the direction controls overlaps with the data wave packets. Define

$$\eta_h = 1 - r_h, \quad D = 1 - c^2, \quad 1 - \mathrm{Tr}\, H^2 = \eta_h(1 - \eta_h/2). \tag{22}$$

Let $E(c, r_h)$ be the infimum conditional Bell error over networks, auxiliary directions at this spectrum, and positive-probability outcomes, with no probability floor. This infimum is the same for LU, LOCC and SEP: each nonzero local product branch is itself an optical active block with the same source, and a deterministic average cannot beat its best branch.

Theorem 3 (Quadratic-error purity criterion)**.** *For pure data and an auxiliary supported on $\mathcal{S}$, along any path $c \to 1^-$,*

$$\begin{aligned} &E(c, r_h) = O((1 - c)^2) \\ \iff\quad &1 - \mathrm{Tr}\, H^2 = O((1 - c)^2). \end{aligned} \tag{23}$$

Necessity holds for every positive-probability outcome. Sufficiency is attained by the midpoint-aligned auxiliary and four-mixing circuit at an outcome probability tending to 1/24. For any fixed $r_h < 1$, the best error is $\Theta(1 - c)$.

The proof extends the compressed-Gram argument to a mixed auxiliary. Its covariance in the nonorthogonal basis (a, b) reweights the two triples of logical product columns. Their shared sum survives this reweighting. After optimizing a scalar spectral envelope over

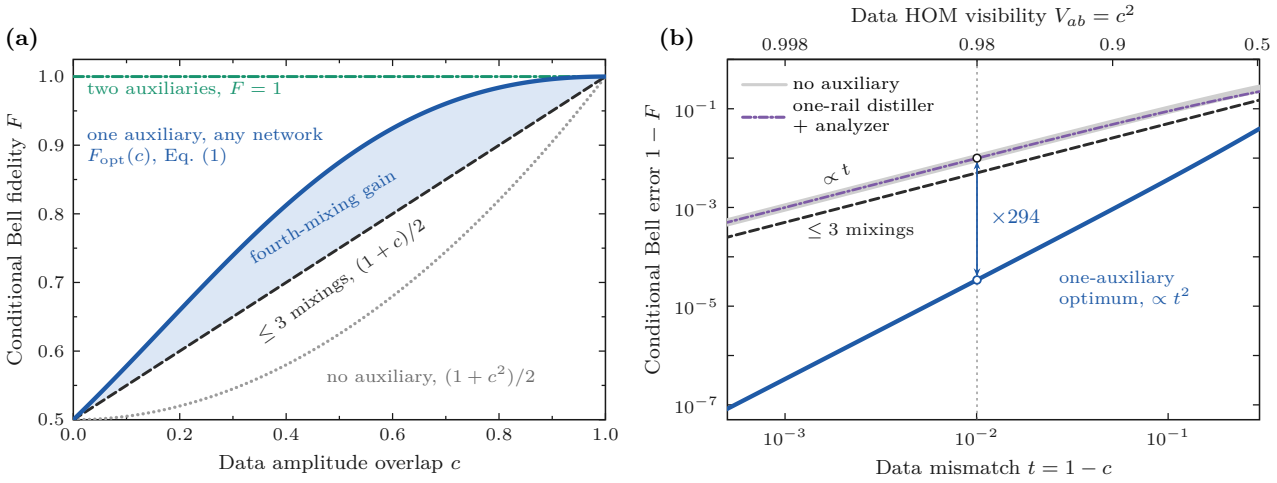


Figure 2: Fidelity limits for pure data photons. (a) Conditional Bell fidelity versus amplitude overlap c without an auxiliary, with one auxiliary and at most three mixings, with one auxiliary and any finite network [Eq. (1)], and with two auxiliaries. Shading: gain from the fourth mixing. (b) Bell error versus mismatch $t = 1 - c$; the top axis gives $V_{ab} = c^2$. The one-rail distiller followed by a Bell analyzer (dash-dot) stays first order and nearly coincides with the unassisted error (wide grey line), whereas the optimum is second order. At $V_{ab} = 0.98$ the error falls 294-fold, from 1.0×10^{-2} to 3.4×10^{-5}.

the Bloch sphere, one obtains the uniform lower bound

$$E(c, r_h) \geq \max\left\{ e_{\rm opt}(c), \min\left[\frac{D}{2(1+cr_h)}, \frac{D\sqrt{1-r_h^2}}{2+c^2} \right] \right\}. \tag{24}$$

The first term controls pure-source mismatch. In the minimum, the first entry is of order t, and the second penalizes source impurity. An error of order t^2 therefore forces $\sqrt{1-r_h^2} = O(t)$, independently of how rare the outcome is. Conversely, the midpoint-aligned source $H_{+r_h} = (I + r_h\sigma_z)/2$ has the exact error and probability

$$e_+ = \frac{t^2 + 2c\eta_h}{2 + 2c^2 - cr_h}, \qquad P_{\rm mid}(c, r_h) = \frac{2 + 2c^2 - cr_h}{72}. \tag{25}$$

As the denominator tends to 3, the leading expression is $(t^2 + 2\eta_h)/3$, separating the mismatch and impurity contributions.

An exchange-invariant auxiliary is unchanged by interchanging a and b. In the midpoint basis that interchange is σ_z, so at fixed spectrum the choices are $H_{\pm r_h}$. The best conditional error over both choices and all networks is exactly

$$e_{\rm axes} = \min\left\{ e_+, \frac{D}{2(1+cr_h)} \right\}. \tag{26}$$

The second branch is attained by the opposite-axis auxiliary and the three-mixing circuit, with heralding probability $(1 + cr_h)/32$. For joint $t, \eta_h \to 0$, these errors have orders $t^2 + \eta_h$ and t, respectively. The exact axis comparison and the spectral inequalities needed for Theorem 3 are given in Appendix E.

Equation (25) also sets finite tolerances: impurity in the orthogonal auxiliary eigenmode adds an error proportional to η_h, although the midpoint mode cancels the first-order mismatch, and misalignment adds a further term. For the same four-mixing instrument and target Φ^+, a Bloch direction error θ relative to the midpoint gives exactly

$$1 - f_{\Phi^+}(\theta) = \frac{t^2 + 2c\eta_h + 2cr_h(1 - \cos\theta)}{2 + 2c^2 - cr_h\cos\theta}. \tag{27}$$

The source probability and target effects are diagonal in the midpoint basis, so this formula includes both transverse directions. The extra term is quadratic in the small Bloch angle. Retaining the aligned $O(t^2 + \eta_h)$ scale requires $\theta = O(\sqrt{t^2 + \eta_h})$; in the quadratic regime this becomes $\theta = O(t)$. This tolerance refers to the fixed four-mixing circuit; Theorem 3 optimizes the source direction freely.

For a desired fractional increase $\delta_{\rm tol}$ above the pure midpoint error, put $N_c = 2 - c + 2c^2$. Requiring $e_+ \leq (1 + \delta_{\rm tol})e_{\rm opt}$ gives the finite budget

$$\eta_h \leq \frac{\delta_{\rm tol} t^2 N_c}{c[2N_c - (1 + \delta_{\rm tol})t^2]}. \tag{28}$$

Here $2N_c > (1 + \delta_{\rm tol})t^2$ and $0 \leq \eta_h \leq 1$. Equation (28) converts an error target into a purity budget.

4.2 Matching error at finite probability

Fix a constant probability floor $p_0 > 0$ and let

$$E_{p_0,\mathcal{R}}(c, r_h) = \inf_{U,H,\boldsymbol{n}:\ {\rm Tr}\, J_{\boldsymbol{n}} \geq p_0} [1 - F_{\mathcal{R}}(J_{\boldsymbol{n}})], \tag{29}$$

where the auxiliary spectrum is fixed and its direction is free. The floor bounds the full outcome probability and is held fixed as $t, \eta_h \to 0$.

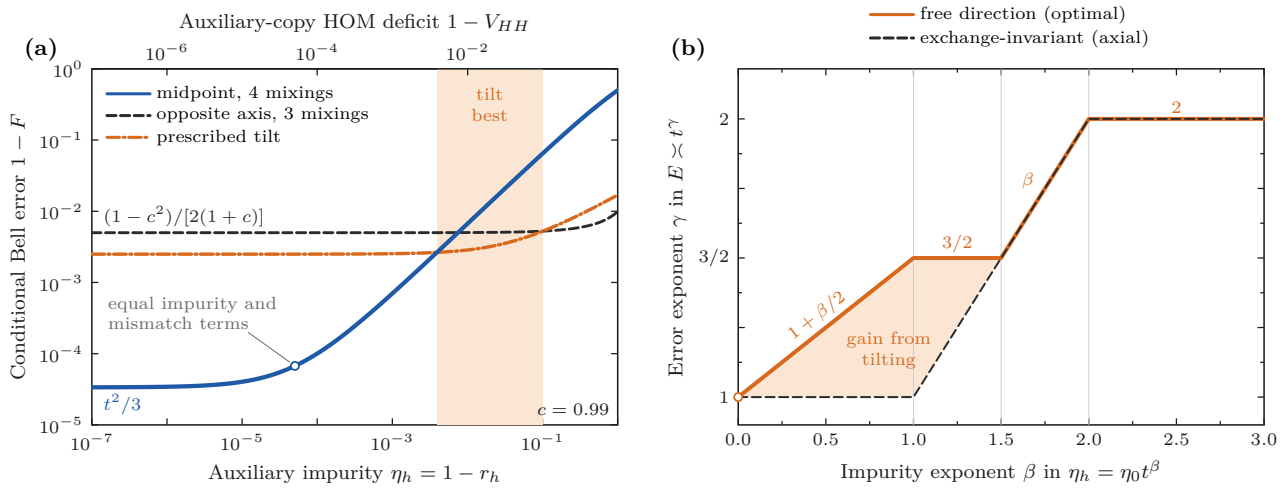


Figure 3: Auxiliary purity and direction for $\operatorname{supp} H \subseteq \operatorname{span}\{a, b\}$. (a) Bell error at $c = 0.99$ versus auxiliary impurity $\eta_h = 1 - r_h$ (top axis: $1 - V_{HH}$) for three controls with the same auxiliary spectrum. Shading: $\eta_h \simeq 0.0039$–0.10, where the prescribed tilt is best (heralding probability 0.013–0.015). Open circle: equal impurity and mismatch terms, $2c\eta_h = t^2$. (b) Optimal error exponent γ, $E \asymp t^\gamma$, for $\eta_h = \eta_0 t^\beta$, $\eta_0 > 0$, $\beta > 0$, at a fixed probability floor $0 < p_0 < 1/64$. Dashed: best exchange-invariant (axial) auxiliary; solid: free direction. The curves coincide for $\beta \geq 3/2$. The open endpoint denotes $\beta \downarrow 0$.

Theorem 4 (Matching fixed-probability error law)**.** *For each fixed $0 < p_0 < 1/64$, and each $\mathcal{R} \in \{\mathrm{LU}, \mathrm{LOCC}, \mathrm{SEP}\}$,*

$$E_{p_0,\mathcal{R}}(c, r_h) \asymp \min\{t^2 + \eta_h,\ t\sqrt{t + \eta_h}\} \tag{30}$$

as $t = 1 - c, \eta_h = 1 - r_h \to 0$. The implied constants depend only on p_0 and hold uniformly in this joint limit. Two real constructions, each accepting one detection pattern, attain the upper order in at most eight modes.

Here $\asymp$ means bounded above and below by positive constants times the displayed function. Constants may differ between recovery classes.

The tilted realization uses a three-row optical contraction with imbalance ν and auxiliary Bloch vector

$$\begin{aligned} s &= \sqrt{D + \eta_h/2}, \qquad \nu = D/s, \\ g_\nu &= \sqrt{D\nu + c^2\nu^2/4}, \\ (x, y, z) &= \frac{r_h}{g_\nu}(\sqrt{D\nu}, 0, c\nu/2). \end{aligned} \tag{31}$$

For small t and η_h, $0 < \nu \leq 1$. The contraction has a real eight-mode unitary dilation, accepting $(1, 1, 1, 0, 0, 0, 0, 0)$ with fixed Bell target Φ^-. Its exact matrix, unnormalized conditional state, and probabilities are given in Appendix F. This direction minimizes the unnormalized Bell loss for the chosen optical block. Uniformly over joint paths $t, \eta_h \to 0$ it gives

$$1 - f_{\Phi^-} = t\sqrt{8t + 2\eta_h}\,[1 + O(s)], \qquad p_\nu \longrightarrow 1/64. \tag{32}$$

The heralding probability stays finite while the tilt and imbalance vary. The limiting probability $1/64$ sets the floor range $p_0 < 1/64$ in Theorem 4.

The two mechanisms cancel different errors. The aligned trine gives $t^2 + \eta_h$; the tilted auxiliary and optical imbalance give $t\sqrt{t + \eta_h}$. A shared-amplitude constraint forces the same network and Bell target to handle both auxiliary eigenmodes, yielding a matching lower bound. Appendix F.3 proves this compatibility condition and its extension to deterministic recovery and a fixed number of accepted patterns. Figure 3 identifies the operating regimes.

For $\eta_h = \eta_0 t^\beta$ with fixed $\eta_0 > 0$, the optimal error exponent is

$$\gamma(\beta) = \begin{cases} 1 + \beta/2, & 0 < \beta \leq 1, \\ 3/2, & 1 \leq \beta \leq 3/2, \\ \beta, & 3/2 \leq \beta \leq 2, \\ 2, & \beta \geq 2. \end{cases} \tag{33}$$

For example, impurity proportional to mismatch leaves exchange-invariant auxiliary states with first-order error, while the tilted construction attains $t^{3/2}$ at finite probability. At very high purity the midpoint construction becomes preferable. Figure 3 compares finite-parameter errors with these asymptotic exponents. The auxiliary tilt has Bloch angle $\tan\theta_H = 2(D + \eta_h/2)^{1/4}/c$, exceeding the short data-angle scale $O(\sqrt{t})$ in the regime where tilt is useful. Purity and direction are consequently separate source controls.

4.3 Stochastic impurity of the data photons

To test the pure-data assumption, consider depolarized data photons with $\rho_a = r_d|a\rangle\langle a| + (1 - r_d)I_2/2$ and $\rho_b = r_d|b\rangle\langle b| + (1 - r_d)I_2/2$, where $0 \leq r_d \leq 1$ is the data Bloch length. The same four-mixing circuit and pure midpoint auxiliary give

$$1 - f_{\Phi^+}(r_d, c) = \frac{(1 - r_d c)^2 + (1 - r_d^2)/2}{3 - r_d^2 + 2r_d^2 c^2 - r_d c}. \tag{34}$$

For $r_d, c \to 1$, this is $[t^2 + (1 - r_d)]/3$ to leading order. Thus stochastic data impurity enters this construction at first order, while coherent mismatch enters at second order. Supplemental Sec. S4.2 gives the derivation and exact probability. This explicit construction distinguishes the two source errors; the all-network optimum for mixed data remains open.

5 Physical source design

5.1 Measurable purity and mismatch budgets

Balanced HOM interference measures the trace overlap of two independent internal densities [27, 28]. For pure data photons and independently prepared copies of the auxiliary density,

$$V_{ab} = c^2, \qquad V_{HH} = \mathrm{Tr}\, H^2 = (1 + r_h^2)/2. \tag{35}$$

Quadratic suppression therefore requires $1 - V_{HH} = O((1 - \sqrt{V_{ab}})^2)$ together with support on $\mathrm{span}\{a, b\}$. Source direction and collective phases require additional characterization [22, 29].

At $V_{ab} = 0.98$, a pure midpoint auxiliary reduces Bell error from 1.0×10^{-2} to 3.4×10^{-5}, with probability 0.041. The two terms of Eq. (25) are equal at $1 - V_{HH} \simeq 5.10 \times 10^{-5}$; below this threshold, impurity contributes less than mismatch. The midpoint four-mixing circuit beats the opposite-axis three-mixing circuit only when $V_{HH} > 0.9925$ at this data visibility. Figure 4 maps the finite-purity tradeoff; tilted controls can further improve an intermediate region.

The auxiliary photon is most useful when remote-source mismatch is large and a local photon is highly reproducible. Values $V_{ab} = 0.90$, $V_{HH} = 0.99$–0.999 yield 6.6- to 31-fold error reductions with the aligned construction. Remote quantum-dot interference has reached visibility near 0.93 [12]; purification experiments already use interference to improve source indistinguishability [6]. Consecutive photons from one resonantly excited quantum-dot source have shown indistinguishability 0.9956 ± 0.0045 after multiphoton correction (0.989 ± 0.004 uncorrected) [13]. These measurements motivate the parameters; a device prediction requires its source model and independent-copy characterization.

5.2 A simple Gaussian auxiliary source

The midpoint is a coherent sum of two data wave packets. For equal-width Gaussian data displaced symmetrically in time or frequency, a centered pure Gaussian auxiliary with optimized width is already nearly optimal. If u is its width relative to the data width and $\kappa_G = -\log c$, the best width for the target Φ^+ satisfies

$$u_*^2 = \kappa_G + \sqrt{1 + \kappa_G^2}. \tag{36}$$

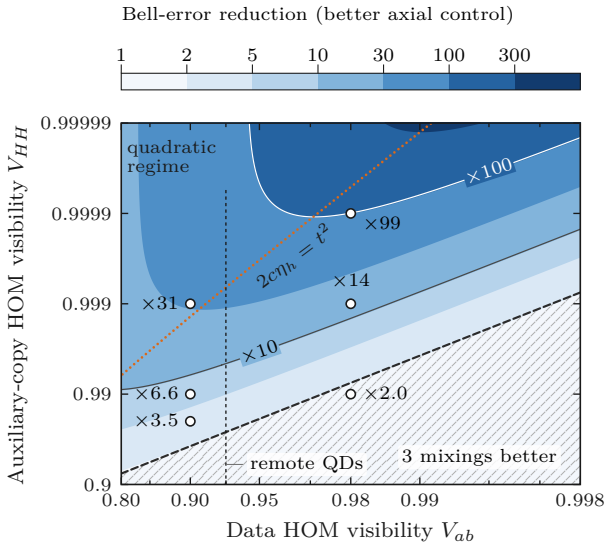


Figure 4: Purity budget: reduction $[(1-c^2)/2]/(1-F)$ from the better axial control (midpoint/four mixings or opposite-axis/three mixings). Axes use logarithmic visibility deficits. In the hatched region below the dashed boundary, three mixings are better. Above the orange dotted line, $2c\eta_h = t^2$, impurity contributes less than mismatch. Thin contours mark 10- and 100-fold reductions. Circles show factors 3.5, 6.6, 31, 2.0, 14 and 99. The dotted vertical line marks remote quantum-dot visibility $V_{ab} \simeq 0.93$ [12]; tilted controls appear separately in Fig. 3.

It tends to the data width as $c \to 1$. The fidelity obeys

$$\begin{aligned} F_{\mathrm{opt}} - f_{\Phi^+} &= t^4/72 + O(t^5), \\ 1 - f_{\Phi^+} &= t^2/3 + O(t^3). \end{aligned} \tag{37}$$

At $c = 0.9$, an ordinary Gaussian photon with $u_* = 1.054$, about 5% wider than the data photons, gives error 0.0036783, compared with 0.0036765 at the all-network optimum. Its support generally extends beyond $\mathrm{span}\{a, b\}$; the direct calculation is in Supplemental Sec. S2.1.

5.3 Extra photons, filtering and heralding probability

With more independent auxiliary photons the fidelity ceiling rises:

$$F_0^\star = (1 + c^2)/2, \qquad F_1^\star = F_{\mathrm{opt}}, \qquad F_2^\star = 1. \tag{38}$$

Two auxiliaries prepared in h_- and four balanced mixings give $J = (1 + c)^2 |\Psi^+\rangle\langle\Psi^+|/128$ for pattern $(1, 1, 0, 0, 1, 1)$. Two antisymmetric data–auxiliary comparisons erase the surviving which-source distinction. Supplemental Sec. S2.2 gives the construction; the data archive includes all 126 detection patterns.

Internal-mode filtering is a different resource. A common rank-one filter matches transmitted data wave packets; frequency-selective attenuation can also achieve exact matching [30–32]. For symmetric Gaus-

sian data, equal-width filters with independently adjusted centers give $F = 1$ and

$$P_G(c,1) = \frac{r_G}{2} c^{2/(1-r_G)}, \qquad (1-r_G)^2 = (-2\log c) r_G. \tag{39}$$

Here $r_G \in (0,1)$ is the transmitted-to-incident intensity-variance ratio, uniquely fixed for $0 < c < 1$ by the second equation (Supplemental Sec. S2.3). The probability includes both filter transmissions and the conventional Bell acceptance $1/2$. At $c = 0.9$, filtering gives $F = 1$ and $P_G \simeq 0.178$, exceeding the fidelity $271/272$ and probability $P = 17/450 \simeq 0.0378$ of the fixed optimal circuit's single pattern $(1,1,0,0,1)$. The auxiliary circuit sets the exact mode-nonselective correction limit and can also follow a filter. If its two transmitted pure wave packets have overlap c' and the filter transmissions are T_A, T_B, the matched pure auxiliary gives

$$1 - F = \frac{t'^2}{3 - 3t' + 2t'^2}, \qquad t' = 1 - c', \qquad P_{\text{combined}} = T_A T_B P_\star(c'). \tag{40}$$

The same filtered photons in a conventional Bell analyzer have acceptance $T_A T_B/2$. Their ratio is $2P_\star(c') \to 1/12$ as $c' \to 1$.

All probabilities condition on supplying the required photons. A common transmission and detection efficiency ζ for each herald photon multiplies an N-photon herald by ζ^N; auxiliary preparation and remote survival contribute separate factors. Supplemental Sec. S1 gives the fixed-circuit all-pattern acceptance optimization.

5.4 Internal-state distillation and encoded swapping

Within our source model, the two-mode conditional operation of Hoch et al. [10], at intensity transmission $1/3$ and midpoint auxiliary, has

$$V_{\text{dist}} = \frac{2 - c^2}{3 - 2c}, \qquad p_{\text{dist}} = \frac{3 - 2c}{9}. \tag{41}$$

Its visibility deficit is $(1-c)^2/(3-2c)$. Applied to one logical rail and followed by a swapping analyzer adjusted to balance the two surviving logical populations, it gives

$$F_{\text{pre}} = \frac{1 + c/\sqrt{3-2c}}{2}, \qquad p_{\text{pre}} = \frac{3-2c}{24(3-c)}. \tag{42}$$

The logical rails undergo different conditional maps, leaving remote Bell error $t - 5t^2/4 + O(t^3)$ despite quadratic wave-packet visibility improvement. Supplemental Sec. S3 derives the complete conditional map. Corollary 1 places all static compositions with the stated resources under the same encoded-pair ceiling.

6 Discussion

For pure but mismatched wave packets, the value of one independent auxiliary photon in entanglement swapping is known exactly and is attainable. The all-network bound identifies the highest conditional Bell fidelity, and equality fixes the midpoint auxiliary, the trine logical geometry and the largest single-pattern probability. The four-mixing realization converts the first-order mismatch error into a quadratic one.

Auxiliary purity determines whether that improvement survives. The measured HOM overlap calibrates a concrete impurity budget. Outside the quadratic regime, source tilt and optical imbalance provide a second cancellation mechanism with a matching finite-probability error law. These controls separate mutual source mismatch, local auxiliary reproducibility and optical calibration as experimentally distinct requirements.

A direct test could combine two independently calibrated remote sources with a high-purity local auxiliary and the NPBS–PPBS–HWP realization. The parameter regime $V_{ab} \simeq 0.90$, $V_{HH} \gtrsim 0.99$ provides a useful finite-purity benchmark. The explicit mixed-data result shows the first-order cost of stochastic source impurity; its all-network optimum remains open. Extending the response constraints to fusion measurements, multi-photon GHZ resources or optical feed-forward would connect these source requirements to broader quantum-network architectures. The present results give an exact fidelity target and explicit source settings for the central station of a photonic quantum network.

Declarations

Funding. The author received no specific funding for this work.

Competing interests. The author declares no competing interests.

Data availability. The relevant data are available on Zenodo [33].

AI-assisted methods. Under the author's direction, OpenAI Codex (GPT-6 Astra) explored candidate circuits and proof strategies and assisted with derivations, code editing, and interpretation of numerical results. OpenAI ChatGPT (GPT-6) and Anthropic Claude Opus 5.5 assisted with critical review and language editing. No generative image model was used. The author takes full responsibility for the accuracy, validation and interpretation of the results and for the content of this article.

Appendices

A: all-network bound; B: equality and probability; C: minimum mixing count; D: attaining circuit and polarization implementation; E: auxiliary purity; F: finite-probability control and converse.

A Universal one-auxiliary limit

Theorem A.1 (One-auxiliary fidelity ceiling). For every model in Sec. 2, every independent auxiliary density, every finite m, every detection pattern of positive probability, and each recovery class $\mathcal{R}$,

$$F_{\mathcal{R}}(J_{\boldsymbol{n}}) \le F_{\mathrm{opt}}(c) = \frac{1+c+c^2}{2-c+2c^2}. \tag{A.1}$$

The value is attained with five spatial modes and local unitaries.

We prove the upper bound for product-Kraus SEP and give an LU attainer. The proof has four steps. A detection pattern uses at most three optical rows, so finite vacuum extensions reduce to a contraction with five potentially populated input columns. Product logical amplitudes bound how much weight a Bell target can collect from the internal Gram matrix. Repeated occupations and auxiliary components outside the data span give strictly lower ceilings. For three distinct outputs, a shared-sum constraint sharpens the Gram spectrum enough to give the stated limit.

A.1 Active optical block and recovery closure

An outcome occupies $k \le 3$ outputs. Only their rows and the five input columns (A_0, A_1, B_0, B_1, h) enter its amplitudes. The resulting $k \times 5$ block T satisfies $TT^\dagger \preceq I_k$. Conversely every such contraction has the unitary dilation (Julia completion) [25]

$$U_T = \begin{pmatrix} T & \sqrt{I_k - TT^\dagger} \\ \sqrt{I_5 - T^\dagger T} & -T^\dagger \end{pmatrix} \in U(5+k). \tag{A.2}$$

The singular-value decomposition gives $Tf(T^\dagger T) = f(TT^\dagger)T$ for the square-root function, making the off-diagonal blocks of $U_T U_T^\dagger$ zero. Its diagonal blocks are identities. The added k inputs are vacuum. This realizes the chosen outcome, its state, and its probability exactly in at most eight modes. For each occupied output i, define its two logical rail vectors and auxiliary amplitude by

$$\alpha_i = (T_{i0}, T_{i1})^T, \qquad \beta_i = (T_{i2}, T_{i3})^T, \qquad v_i = T_{i4}. \tag{A.3}$$

An outcome is *collision-free* when its three photons occupy distinct outputs. For such outcomes $J(\zeta T) = |\zeta|^6 J(T)$. More generally, for a collision-free outcome and diagonal row scaling Δ_{out},

$$J(\Delta_{\mathrm{out}} T \,\mathrm{diag}(g_A I_2, g_B I_2, g_H)) = |\det \Delta_{\mathrm{out}}\, g_A g_B g_H|^2 J(T). \tag{A.4}$$

Every finite block can therefore be attenuated to a contraction at fixed normalized state. A local reference branch $L \otimes M$ is represented by

$$T' = T\,\mathrm{diag}(L^T, M^T, 1). \tag{A.5}$$

For example the first-rail amplitude becomes $\sum_{j=0}^{1} T_{i,A_j} L_{aj}$, the coefficient of remote output a. Uniform positive attenuation of T' changes the three-photon amplitude by its third power and J by its sixth power, leaving the normalized state and H unchanged. This is the homogeneous optical closure used below. For a branch of a trace-preserving product-Kraus channel, the balancing in Proposition F.7 already makes T' a contraction and preserves the branch probability. In particular, rail unitaries allow the Bell target to be aligned with $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$.

For a fixed target and recovery, the Bell numerator is linear in J; its optimized value is convex and positively homogeneous. Since J is linear in H, a mixture is bounded by its best nonzero-probability pure component. An accepted ensemble obeys the analogous probability-weighted bound. These facts justify the pure-auxiliary reduction for Theorem A.1.

A.2 Product columns and the internal-support reduction

Lemma A.2 (Product-column spectral bound). Suppose $J = WGW^\dagger$, with every original column of W a two-qubit product vector. Suppose the conjugate row space of W lies in a subspace on which the compression of G has spectrum in $[\lambda_{\min}, \lambda_{\max}]$, with $\lambda_{\min} > 0$. Then $F_{\mathrm{LU}}(J) \le \lambda_{\max}/(\lambda_{\max} + \lambda_{\min})$ at positive probability. The same bound holds after every deterministic product-Kraus recovery.

Proof. For a maximally entangled target put $p = \mathrm{Tr}\, J$, $w_\Phi = \langle\Phi|J|\Phi\rangle$, $z_{\mathrm{B}} = \Phi^\dagger W$ and $R_\perp = (I - |\Phi\rangle\langle\Phi|)W$. Each product column has squared Bell overlap at most half its squared norm. Consequently

$$\|R_\perp\|_F^2 \ge \|z_{\mathrm{B}}\|^2, \qquad p - w_\Phi = \mathrm{Tr}(R_\perp G R_\perp^\dagger) \ge \lambda_{\min}\|R_\perp\|_F^2 \ge \lambda_{\min}\|z_{\mathrm{B}}\|^2 \ge \frac{\lambda_{\min}}{\lambda_{\max}} w_\Phi. \tag{A.6}$$

It follows that $w_\Phi \le \lambda_{\max} p/(\lambda_{\max}+\lambda_{\min})$. The inequality uses the product character of the original columns. A local Kraus operator $L_j \otimes M_j$ from Eq. (5) multiplies W from the left, preserving product columns and the row constraint. The same bound therefore holds on each local Kraus branch. Summing branches proves the deterministic recovery bound whenever Eq. (5) holds. □

With no auxiliary, an outcome at distinct outputs i, j has

$$J = \frac{1}{4}\,(x,y)\begin{pmatrix}1 & c^2\\ c^2 & 1\end{pmatrix}(x,y)^\dagger, \quad x=\alpha_i\otimes\beta_j, \quad y=\alpha_j\otimes\beta_i. \tag{A.7}$$

Let $F^{(k)}_{\max}(c)$ denote the best fidelity with k independently controllable auxiliary photons in this resource model. The lemma gives $F_{\mathcal{R}} \le F^{(0)}_{\max} = (1+c^2)/2$. A doubly occupied output has a single reference product vector and cannot improve this bound. The two balanced couplers on rail pairs $(0,2)$ and $(1,3)$, accepting one photon in each of outputs $0,1$, produce a balanced odd-parity state with coherence c^2 and attain $F^{(0)}_{\max}$ at probability $1/8$. This also proves the ceiling for a fully distinguishable auxiliary: resolve its output label conceptually and remove it. Conditional on that label, the remaining two-data state is a zero-auxiliary outcome of the same optics; forgetting the label takes a positive mixture with the same fidelity bound.

Let $\mathcal{S} = \mathrm{span}\{a,b\}$. Internal-blind optics conserves the number of photons in $\mathcal{S}^\perp$. Coherence between an auxiliary in $\mathcal{S}$ and one in $\mathcal{S}^\perp$ therefore disappears on tracing the internal state. The orthogonal sector has the ceiling $F^{(0)}_{\max}$. Hence it suffices to bound pure auxiliary states $h \in \mathcal{S}$.

A.3 All bunched outcomes: a sharp bound

For arbitrary independent auxiliary density H, including support outside $\mathcal{S}$, define

$$\begin{aligned} A_H &= \langle a|H|a\rangle \\ B_H &= \langle b|H|b\rangle \\ w &= c\big(c+\langle b|H|a\rangle\big) \\ \xi_b &= \frac{|w|}{\sqrt{(1+A_H)(1+B_H)}}. \end{aligned} \tag{A.8}$$

A $(2,1)$ occupation puts two photons in one output and the third in another; a (3) occupation puts all three in one output. Every complete $(2,1)$ occupation satisfies $F_{\mathcal{R}} \le (1+\xi_b)/2$ for LU and for Eq. (5); every (3) occupation has fidelity $1/2$ after optimal local recovery.

To prove the bound, if i is the doubly occupied output and j the other output, the only reference product vectors are $\alpha_i\otimes\beta_i$, $\alpha_i\otimes\beta_j$ and $\alpha_j\otimes\beta_i$, with auxiliary spatial amplitudes v_j, v_i, v_i. Local unitaries align α_i, β_i to their respective zero axes. The entire state is then supported on $00, 01, 10$, including the coherences involving 00. Write its odd populations as p_o, q_o and odd coherence as z_o. Exact two-boson symmetrization with no auxiliary, an outcome at distinct outputs i, j has

$$p_o = \tfrac{1}{4}|v_i|^2(1+A_H)\|\alpha_i\|^2|\beta_{j,\perp}|^2, \tag{A.9}$$

$$q_o = \tfrac{1}{4}|v_i|^2(1+B_H)\|\beta_i\|^2|\alpha_{j,\perp}|^2, \quad |z_o| = \xi_b\sqrt{p_o q_o}. \tag{A.10}$$

For example the two assignments in the doubly occupied port have inner product $\langle b|a\rangle + \langle b|h\rangle\langle h|a\rangle$; the remaining single-port overlap supplies the other factor c. Their diagonal norm is $1+|\langle a|h\rangle|^2$ or $1+|\langle b|h\rangle|^2$. Averaging gives Eq. (A.10), including the bosonic factorial normalization.

For a positive state supported on $|00\rangle, |01\rangle, |10\rangle$, the spin-flip product $J(\sigma_y\otimes\sigma_y)J^*(\sigma_y\otimes\sigma_y)$ has characteristic polynomial

$$\lambda^2\left[\lambda^2 - 2(p_o q_o + |z_o|^2)\lambda + (p_o q_o - |z_o|^2)^2\right]. \tag{A.11}$$

Thus its homogeneous concurrence is $\widetilde{C}(J) = 2|z_o|$ by the two-qubit concurrence formula [34]; the 00 coherences do not affect this conclusion. For a pure two-qubit state the fully entangled fraction is $(1+C)/2$. Applying convexity to a concurrence-minimizing decomposition gives the corresponding upper bound for mixed states. Therefore

$$w_{\mathrm{LU}}(J) \le \tfrac{1}{2}\,\mathrm{Tr}\,J + |z_o|, \qquad \mathrm{Tr}\,J \ge p_o + q_o \ge 2\sqrt{p_o q_o}, \tag{A.12}$$

which proves the asserted ceiling. Vanishing rows follow by continuity of the homogeneous inequality. A product Kraus operator only replaces all α_j, β_j by new local row vectors, so the same argument applies branchwise and then sums. A (3) event has only $\alpha_i\otimes\beta_i$ as a logical amplitude, proving the product assertion.

To optimize ξ_b over every H, use the midpoint basis of Eq. (3) and write the restricted block of H as $\left(\begin{smallmatrix} p_h & d_h \\ \bar{d}_h & q_h\end{smallmatrix}\right)$. Then $p_h, q_h \ge 0$, $p_h + q_h \le 1$ and $|d_h|^2 \le p_h q_h$. Put $a^2_{\mathrm{mid}} = (1+c)/2$, $b^2_{\mathrm{mid}} = (1-c)/2$, $A = a^2_{\mathrm{mid}} p_h$, $B = b^2_{\mathrm{mid}} q_h$, $\tau_h = A + B$ and $\rho_h = 4a^2_{\mathrm{mid}} b^2_{\mathrm{mid}} (\mathrm{Re}\, d_h)^2$. Direct substitution gives

$$\begin{aligned} \frac{\xi_b^2}{c^2} &= \frac{(c+A-B)^2 + 4a^2_{\mathrm{mid}} b^2_{\mathrm{mid}} (\mathrm{Im}\, d_h)^2}{(1+\tau_h)^2 - \rho_h} \le \frac{N_0 - \rho_h}{D_0 - \rho_h} \\ N_0 &= (c+\tau_h)^2 - 4cB \\ D_0 &= (1+\tau_h)^2. \end{aligned} \tag{A.13}$$

All denominators are positive and $D_0 - N_0 = (1-c)(1+c+2\tau_h) + 4cB > 0$. Hence

$$\frac{\xi_b^2}{c^2} \le \frac{N_0-\rho_h}{D_0-\rho_h} \le \frac{N_0}{D_0} \le \left(\frac{c+\tau_h}{1+\tau_h}\right)^2 \le \left(\frac{1+3c}{3+c}\right)^2. \tag{A.14}$$

The second inequality follows by subtracting, obtaining $\rho_h(D_0 - N_0)/[D_0(D_0 - \rho_h)] \geq 0$; the last uses $\tau_h \leq a_{\mathrm{mid}}^2$ and strict monotonicity of $(c + \tau_h)/(1 + \tau_h)$. Equality requires $p_h = 1, q_h = 0$; positivity then forces the entire auxiliary to be $|h_+\rangle\langle h_+|$.

The sharp bunched-outcome result is consequently

$$F_{\mathrm{bunch}}^{\max}(c) = \frac{3 + 2c + 3c^2}{2(3 + c)}. \tag{A.15}$$

An explicit attainer completes the sharpness claim. Choose h_+ and the two orthonormal detected rows

$$T_{\mathrm{b}} = \begin{pmatrix} 1/\sqrt{3} & 0 & 1/\sqrt{3} & 0 & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \end{pmatrix}. \tag{A.16}$$

They extend to a real orthogonal $U(5)$ by adding $(1, 0, -1, 0, 0)/\sqrt{2}$, $(1, 0, 1, 0, -2)/\sqrt{6}$, and $(0, 1, 0, -1, 0)/\sqrt{2}$. Accept the detection pattern $(2, 1, 0, 0, 0)$. The even logical amplitudes vanish: the second row has only odd rails, while both photons reaching the first row have even rails or the auxiliary. For 01 and 10, normalized bosonic symmetrization gives

$$J_{\mathrm{odd}} = \frac{1}{144}\begin{pmatrix} 3 + c & c(1 + 3c) \\ c(1 + 3c) & 3 + c \end{pmatrix}, \qquad p = \frac{3 + c}{72}. \tag{A.17}$$

For instance Eq. (A.10) uses $A_H = B_H = (1 + c)/2$, $|v_i|^2 = \|\alpha_i\|^2 = \|\beta_i\|^2 = 1/3$, and $|\alpha_{j,\perp}|^2 = |\beta_{j,\perp}|^2 = 1/2$. The positive odd coherence makes Ψ^+ attain Eq. (A.15), at positive probability.

Since $(1 + c)/2 - F_{\mathrm{bunch}}^{\max} = c(1 - c)/(3 + c) > 0$, all bunched outcomes lie strictly below the one-auxiliary maximum proved next.

A.4 Six product coefficients and the shared-sum constraint

It remains to treat three distinct occupied outputs, numbered $0, 1, 2$. Write the auxiliary as $h = h_a a + h_b b$. Let e_i denote the three-port internal tensor with b at port i and a at the other two; let f_i interchange a, b. Define, for $\{i, j, k\} = \{0, 1, 2\}$,

$$x_i = \alpha_i \otimes (v_k\beta_j + v_j\beta_k), \qquad y_i = (v_k\alpha_j + v_j\alpha_k) \otimes \beta_i. \tag{A.18}$$

The accepted joint state of references and internal labels is $\frac{1}{2}\sum_i (h_a y_i \otimes e_i + h_b x_i \otimes f_i)$. Both assignments of the repeated internal state have already been added coherently. Crucially, the same optical matrix enforces

$$\sum_i x_i = \sum_i y_i = \sum_{i \neq j} v_k\alpha_i \otimes \beta_j. \tag{A.19}$$

The six product columns are coupled by this shared-sum identity.

With $W = \frac{1}{2}(h_b x_0, h_b x_1, h_b x_2, h_a y_0, h_a y_1, h_a y_2)$, internal tracing gives $J = WGW^\dagger$ and

$$G = \begin{pmatrix} G_0 & G_1 \\ G_1 & G_0 \end{pmatrix}, \quad G_0 = (1 - c^2)I_3 + c^2\mathbf{1}\mathbf{1}^T, \quad G_1 = (c^3 - c)I_3 + c\mathbf{1}\mathbf{1}^T. \tag{A.20}$$

The entries follow by counting the mismatched internal states: within a group different positions give c^2, while between groups equal positions give c^3 and unequal positions give c. The six tensors are independent for $0 < c < 1$, so this Gram matrix is positive definite.

The full Gram spectrum separates into constant directions with eigenvalues

$$s_+ = (1 + c)(1 + c + c^2), \qquad s_- = (1 - c)(1 - c + c^2), \tag{A.21}$$

and two copies each of the zero-sum eigenvalues

$$\mu_- = (1 - c^2)(1 - c), \qquad \mu_+ = (1 - c^2)(1 + c). \tag{A.22}$$

Using the full spectrum would lose sharpness at small c. The shared sum restricts the two group means to a single direction proportional to (h_b, h_a), including $h_a h_b = 0$. The conjugate row space is contained in a five-dimensional shared-sum subspace whose compressed eigenvalues are

$$\begin{gathered} \ell,\ \mu_-, \mu_-, \mu_+, \mu_+ \\ \ell = 1 + 2c^2 + (2c + c^3)\chi_h \\ \chi_h = \frac{2\,\mathrm{Re}(\overline{h_a}h_b)}{|h_a|^2 + |h_b|^2} \in [-1, 1]. \end{gathered} \tag{A.23}$$

On conjugate rows the constant direction is $(\overline{h_b}, \overline{h_a})$; its Rayleigh value is the same because G is real symmetric.

For every maximally entangled target, the product-column lemma gives

$$f_\Phi(J) \leq B_h(c) := \frac{\max(\ell, \mu_+)}{\max(\ell, \mu_+) + \min(\ell, \mu_-)}. \tag{A.24}$$

For $\ell \in [s_-, s_+]$ this function decreases below μ_-, is constant between μ_- and μ_+, and increases above μ_+. Its only candidate maxima are bounded by

$$\begin{gathered} \frac{s_+}{s_+ + \mu_-} = F_{\mathrm{opt}} = \frac{1 + c + c^2}{2 - c + 2c^2} \\ \frac{\mu_+}{\mu_+ + \mu_-} = \frac{1 + c}{2} \\ \frac{\mu_+}{\mu_+ + s_-} = \frac{(1 + c)^2}{2 + c + 2c^2}. \end{gathered} \tag{A.25}$$

The two required strict comparisons are

$$F_{\mathrm{opt}} - \frac{1 + c}{2} = \frac{c(1 - c)(1 + 2c)}{2(2 - c + 2c^2)} > 0, \tag{A.26}$$

$$F_{\mathrm{opt}} - \frac{(1 + c)^2}{2 + c + 2c^2} = \frac{3c^2}{(2 - c + 2c^2)(2 + c + 2c^2)} > 0. \tag{A.27}$$

Therefore $w_\Phi \leq F_{\mathrm{opt}} p$ for every complex optical block and every auxiliary. The bunched, orthogonal-auxiliary

and mixed-auxiliary reductions complete the proof for arbitrary finite vacuum extension. Applying the same spectral inequality branchwise proves it for Eq. (5).

A.5 A five-mode attainer

Put

$$V = \begin{pmatrix} \sqrt{3}/2 & 1/2 \\ -\sqrt{3}/2 & 1/2 \\ 0 & -1 \end{pmatrix}, \quad V^\dagger V = \tfrac{3}{2} I_2, \quad VV^\dagger = \tfrac{3}{2} I_3 - \tfrac{1}{2}\mathbf{1}\mathbf{1}^T. \tag{A.28}$$

The canonical block $T_0 = 3^{-1/2}[V, V, \mathbf{1}]$ has orthonormal rows. Adding rows $(1, 0, -1, 0, 0)/\sqrt{2}$ and $(0, 1, 0, -1, 0)/\sqrt{2}$ gives an explicit five-mode completion, denoted U_{trine}. With h_+ and the first three outputs each occupied once, Eqs. (A.18)–(A.20) give the unnormalized weights in the order $(\Phi^+, \Phi^-, \Psi^+, \Psi^-)$

$$\left(\frac{1+c+c^2}{72}, \frac{(1-c)^2}{144}, \frac{(1-c)^2}{144}, 0\right). \tag{A.29}$$

For a two-by-two logical matrix X, we use row-ordered vectorization,

$$\operatorname{vec} X = (X_{00}, X_{01}, X_{10}, X_{11})^T, \qquad \operatorname{mat} = \operatorname{vec}^{-1}. \tag{A.30}$$

Thus $\operatorname{vec}(uv^T) = u \otimes v$ for logical two-component vectors u, v, with the first factor referring to Q_A and the second to Q_B. Let V_i denote row i of V. Each product column of the attainer is proportional to $\operatorname{vec}(V_i^T V_i)$, its Bell projection is constant, and the two residual trine directions have equal Gram weight μ_-. This yields F_{opt} and

$$P_\star(c) = \frac{2 - c + 2c^2}{72} > 0. \tag{A.31}$$

The LU-equivalent four-mixing realization is Eq. (14); Appendix D gives its full matrix and conditional-state calculation. Its error is $t^2/(3 - 3t + 2t^2) = t^2/3 + O(t^3)$, establishing the optimal quadratic coefficient at positive probability.

B Geometry and cost of optimal outcomes

Saturation fixes geometry in two different spaces. The internal auxiliary must bisect the two data states, while three rank-one logical projectors form a Bloch trine. The latter is the structure that makes the spectral bound compatible with one optical network.

Theorem B.1 (Endpoint geometry and single-outcome cost). For $0 < c < 1$, equality in Theorem A.1 requires the auxiliary density $H = |h_+\rangle\langle h_+|$. Up to output order and local rail unitaries, its positive-probability optical blocks have the trine form in Eq. (B.4). Their conditional states are LU-equivalent to the Bell-diagonal trine state. At this fidelity ceiling, the largest probability of one detection pattern is

$$P_\star(c) = \frac{2 - c + 2c^2}{72}. \tag{B.1}$$

These statements hold for LU, deterministic LOCC and SEP recovery; five modes attain the probability bound.

B.1 From spectral saturation to the trine

The strict comparisons above show that saturation by a pure in-span auxiliary requires $\chi_h = 1$, hence $h_a = h_b$ up to common phase and $h = h_+$. In this case the compressed maximum $\lambda_{\max} = s_+$ is simple and the compressed minimum $\lambda_{\min} = \mu_-$ has multiplicity two; the remaining eigenvalue μ_+ is strictly intermediate. For the aligned target, write $z_{\mathrm{B}} = \Phi^{+\dagger} W$ and $R_\perp = (I - |\Phi^+\rangle\langle\Phi^+|)W$ as in Lemma A.2. Equality in Eq. (A.6) forces z_{B} to be constant over all six columns and $R_\perp$ to consist of two equal zero-sum triples. Every original product column must saturate its Bell overlap inequality. Paired columns agree, $W_i = W_{i+3} = w_i$, their common Bell projection z_0 is nonzero, and

$$\begin{aligned} \operatorname{mat}(w_i) &= \sqrt{2} z_0 P_i \\ P_i &= |u_i\rangle\langle u_i| \\ \sum_{i=0}^{2} P_i &= \tfrac{3}{2} I_2. \end{aligned} \tag{B.2}$$

Here mat is the inverse vectorization in Eq. (A.30). To see the projector assertion, write a product matrix as uv^T; equality in $|v^T u|^2 \le \|u\|^2 \|v\|^2$ requires v^T proportional to $u^\dagger$. The trace fixes the proportionality and phase. The three unit Bloch vectors $\mathbf{m}_i$ defined by $P_i = (I + \mathbf{m}_i \cdot \boldsymbol{\sigma})/2$ sum to zero; squaring $\mathbf{m}_i + \mathbf{m}_j = -\mathbf{m}_k$ shows each pair has dot product $-1/2$. They form a trine, and its planar isotropy gives normalized Bell spectrum

$$\left(F_{\text{opt}}, \frac{1 - F_{\text{opt}}}{2}, \frac{1 - F_{\text{opt}}}{2}, 0\right). \tag{B.3}$$

The projector structure also fixes the local frames: the state is LU-equivalent to the Bell-diagonal trine state and has maximally mixed local marginals.

All positive-probability LU equality blocks, up to output order and logical input unitaries, are precisely

$$\begin{aligned} T &= \Delta_{\text{out}}[g_A V U_A,\ g_B V U_B,\ \mathbf{1}] \\ \Delta_{\text{out}} &= \operatorname{diag}(d_0, d_1, d_2) \\ d_0 d_1 d_2 g_A g_B &\neq 0 \\ TT^\dagger &\preceq I_3. \end{aligned} \tag{B.4}$$

Here $U_A, U_B \in U(2)$. For fixed target Φ^+ the two local frames can be chosen conjugate; arbitrary U_A, U_B simply allow LU target optimization. To exclude a vanishing auxiliary output amplitude, suppose $v_0 = 0$, then $y_1 = v_2 \alpha_0 \otimes \beta_1$ and $x_0 = \alpha_0 \otimes (v_2 \beta_1 + v_1 \beta_2)$. Either a required column vanishes or two distinct trine

columns share the same first local ray, a contradiction. Relabeling treats every v_i. With all $v_i \neq 0$, set $\widehat{\alpha}_i = \alpha_i/v_i$ and $\widehat{\beta}_i = \beta_i/v_i$. Equality of paired columns implies

$$(\widehat{\alpha}_j + \widehat{\alpha}_k) \otimes \widehat{\beta}_i = \widehat{\alpha}_i \otimes (\widehat{\beta}_j + \widehat{\beta}_k). \tag{B.5}$$

Thus $\sum_i \widehat{\alpha}_i$ is parallel to all three distinct $\widehat{\alpha}_i$ rays and must vanish; likewise $\sum_i \widehat{\beta}_i = 0$. Represent the three trine rays by the rows of V using a local unitary. The relation among these three rows has a one-dimensional complex nullspace generated by $(1, 1, 1)$, so their three row multipliers must be equal, giving Eq. (B.4). Conversely the row and group scaling identity (A.4) proves sufficiency.

To classify the entire source density, first note that once an in-span component saturates, the block is of the trine form. In its canonical scaling the event probability for any auxiliary is

$$p(H) = \frac{2 - c + 2c^2}{72} H_{++} + \frac{2 + c + 2c^2}{72} H_{--} + \frac{4 + c^2}{72} \operatorname{Tr}(H|_{\mathcal{S}^\perp}). \tag{B.6}$$

The cross terms between $\mathcal{S}$ and $\mathcal{S}^\perp$ trace out, and the trine shared-sum symmetry makes the h_+, h_- trace cross term zero. The formula also follows by substituting $h_\pm$ and a unit orthogonal vector in the six-assignment amplitudes. Every displayed coefficient is positive. For a general trine block they are multiplied by the same positive optical factor. All other pure source directions therefore contribute positive weight at strictly smaller fidelity. Convexity forces the entire auxiliary density to be $|h_+\rangle\langle h_+|$ at equality.

B.2 The sharp probability of one complete endpoint outcome

The fidelity is invariant under the gains and row weights in the equality family; their contraction constraint determines the sharp event probability. In Eq. (B.4) let $s_{\mathrm{g}} = |g_A|^2 + |g_B|^2$ and $q_i = |d_i|^2$. The exact optical scaling and row Gram matrix are

$$\begin{aligned} \frac{p}{P_\star} &= 27 |g_A g_B|^2 q_0 q_1 q_2 \\ TT^\dagger &= \Delta_{\mathrm{out}} C(s_{\mathrm{g}}) \Delta_{\mathrm{out}}^\dagger \\ C(s_{\mathrm{g}}) &= \tfrac{3 s_{\mathrm{g}}}{2} I_3 + (1 - \tfrac{s_{\mathrm{g}}}{2}) \mathbf{1}\mathbf{1}^T. \end{aligned} \tag{B.7}$$

The spectrum of $C(s_{\mathrm{g}})$ is $3, 3s_{\mathrm{g}}/2, 3s_{\mathrm{g}}/2$. Its positive definiteness makes the contraction condition equivalent to $C(s_{\mathrm{g}})^{1/2} \operatorname{diag}(q_i) C(s_{\mathrm{g}})^{1/2} \preceq I_3$. This feasible set is convex in q and invariant under all permutations. The average of a feasible point's six permutations is therefore feasible. Writing $\bar{q} = (q_0 + q_1 + q_2)/3$, AM–GM and its spectral constraint give

$$q_0 q_1 q_2 \leq \bar{q}^3 \leq \frac{1}{\max(3, 3s_{\mathrm{g}}/2)^3}, \qquad |g_A g_B|^2 \leq \frac{s_{\mathrm{g}}^2}{4}. \tag{B.8}$$

Consequently

$$\frac{p}{P_\star} \leq \begin{cases} s_{\mathrm{g}}^2/4, & s_{\mathrm{g}} \leq 2, \\ 2/s_{\mathrm{g}}, & s_{\mathrm{g}} \geq 2, \end{cases} \leq 1. \tag{B.9}$$

Equality holds exactly when $s_{\mathrm{g}} = 2$, $|g_A| = |g_B| = 1$ and $|d_i| = 1/\sqrt{3}$. Then $TT^\dagger = I_3$, so five modes attain the bound, which holds for any number of vacuum modes. Five-mode fidelity equality alone forces $s_{\mathrm{g}} = 2$ and $|d_i| = 1/\sqrt{3}$, but permits $|g_A| \neq |g_B|$.

B.3 Deterministic recovery at the endpoint

The endpoint classification extends to deterministic recovery. Equation (A.5) represents each nonzero local branch as an admissible optical outcome up to scale. Concurrence identifies when both that branch and the recovered mixture can saturate the optical ceiling. Use homogeneous concurrence $\widetilde{C}(J) = \operatorname{Tr}(J) C(J/\operatorname{Tr} J)$. For any product filter,

$$\widetilde{C}\big((L \otimes M) J (L \otimes M)^\dagger\big) = |\det L \det M| \, \widetilde{C}(J). \tag{B.10}$$

For pure states this is $2|\det \mathrm{mat}(\psi)|$. For invertible filters the equality follows by mapping convex-roof decompositions in both directions; singular filters give zero on both sides by continuity. From $\sqrt{\det X} \leq \operatorname{Tr} X/2$ for positive 2×2 matrices and Eq. (5),

$$\sum_j |\det L_j \det M_j| \leq \tfrac{1}{4} \sum_j \operatorname{Tr}(L_j^\dagger L_j) \operatorname{Tr}(M_j^\dagger M_j) = 1. \tag{B.11}$$

Convexity and Eq. (B.10) prove concurrence monotonicity under the allowed deterministic maps.

Let ρ_* be an LU equality state classified above. It is LU-equivalent to the Bell-diagonal trine state. Its maximally mixed local marginals follow from that classification, in addition to its spectrum in Eq. (B.3). It has $C(\rho_*) = 2F_{\mathrm{opt}} - 1 > 0$, and a correlation matrix with singular values $F_{\mathrm{opt}}, F_{\mathrm{opt}}, 2F_{\mathrm{opt}} - 1$, so its operator norm is strictly below one. For positive matrices X, Y of determinant one write $X = \cosh u \, I + \sinh u \, \hat{n} \cdot \sigma$ and $Y = \cosh v \, I + \sinh v \, \hat{m} \cdot \sigma$, with $u, v \geq 0$. Then

$$\operatorname{Tr}[(X \otimes Y)\rho_*] \geq \cosh u \cosh v - F_{\mathrm{opt}} \sinh u \sinh v = \cosh(u - v) + (1 - F_{\mathrm{opt}}) \sinh u \sinh v \geq 1. \tag{B.12}$$

Equality requires $u = v = 0$. Any invertible local filtering of ρ_* has normalized concurrence at most $2F_{\mathrm{opt}} - 1$, with equality only if both filters are proportional to unitaries.

Now suppose a raw optical outcome ρ reaches Bell fidelity F_{opt} after a deterministic map. Each nonzero

Kraus branch obeys the already proved optical ceiling. Since their weighted average reaches the ceiling, every one must itself saturate. Pick a branch of positive probability. It is entangled, so both local operators are invertible; its normalized output is an LU equality state by the classification above. Applying its inverse filters and Eq. (B.12) gives $C(\rho) \le 2F_{\mathrm{opt}} - 1$. Conversely the final state's fidelity implies $C \ge 2F_{\mathrm{opt}} - 1$, and deterministic monotonicity gives $C(\rho) \ge 2F_{\mathrm{opt}} - 1$. All inequalities are equalities, so the chosen filters are proportional to local unitaries. The raw outcome already belongs to the LU equality family. This proves both endpoint source rigidity and the sharp heralding probability for the full allowed product-Kraus class.

C Three-mixing ceiling with finite vacuum extensions

Within this section A, B, H denote internal density operators. This section proves the resource theorem for arbitrary finite-dimensional independent internal states. Logical source rails have the same internal state. Detection resolves spatial photon numbers and traces internal labels; reference recovery retains every reference-side outcome. For an unnormalized state write $w_{\mathrm{B}}(J) = \mathrm{Tr}(J) F_{\mathrm{LU}}(J)$; N is the sum of accepted Bell numerators, so the accepted fidelity is N/P.

Theorem C.1 (Three-mixing fidelity ceiling). For fixed A, B, H, put $q_A = \mathrm{Tr}(AH)$, $q_B = \mathrm{Tr}(BH)$, $\tau = \mathrm{Tr}(AB)$, $w = \mathrm{Tr}(HAB)$ and $\xi_\pm = |\tau \pm w|/\sqrt{(1 \pm q_A)(1 \pm q_B)}$. Zero-population responses have zero coherence and are assigned $\xi = 0$. With two dual-rail data photons, one independent auxiliary, any finite number of additional vacuum inputs, a static passive internal-state-blind network of at most three arbitrary pair mixings, every order and known fixed phase, PNR detection of all output modes, fractional policies and classical mixtures of settings, the maximum fidelity under local unitaries or single-copy deterministic LOCC is

$$F^{\max} = \frac{1 + \max(\xi_+, \xi_-)}{2}.$$

A data-first tree attains it. The alternative data–auxiliary-first orderings are AH-first and BH-first, according to which data source meets the auxiliary first. Maximizing this expression over an allowed set of auxiliary states preserves the equality.

Lemma C.2 (Single-parity-block recovery). For J supported on $\{|01\rangle, |10\rangle\}$, with populations p_0, p_1 and off-diagonal entry z, the LU Bell numerator is $w_{\mathrm{B}} = (p_0 + p_1)/2 + |z|$. For $p_0 p_1 > 0$, define $\xi = |z|/\sqrt{p_0 p_1}$. Then

$$F = \frac{1}{2} + \frac{|z|}{p_0 + p_1} \le \frac{1 + \xi}{2},$$

with equality for balanced populations.

Proof. The maximally entangled vector $(|01\rangle + e^{-i \arg z}|10\rangle)/\sqrt{2}$ attains the stated weight; it corresponds to $W = X\,\mathrm{diag}(1, e^{i \arg z})$ for $z \neq 0$ and to $W = X$ for $z = 0$ in the LU Bell numerator, where X is the Pauli X matrix. Any maximally entangled vector has equal squared amplitudes on these two basis states, each at most $1/2$, proving the upper bound. The fidelity inequality follows from $2\sqrt{p_0 p_1} \le p_0 + p_1$. If either population vanishes, positivity gives $z = 0$; a nonzero product-supported outcome has $F = 1/2$. The zero matrix contributes no accepted probability. □

C.1 Six groups of the data-first instrument

Our pair mixing convention is

$$B_{ij}(\theta, \varphi)|_{ij} = \begin{pmatrix} \cos\theta & e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & -\cos\theta \end{pmatrix}.$$

The data-first ordering mixes the data modes first and then sends one output into the final data–auxiliary mixing. Bunching means detecting both photons at one output of that mixing; coincidence means detecting one at each output. A bypass amplitude avoids that mixing. For the real tree $U = B_{04}(\pi/4) B_{13}(\phi) B_{02}(\theta)$, set $x = \cos^2\theta$, $x_R = \cos^2\phi$. For $f \in \{I, +, -\}$ define

f	E_f	k_f	x_f
I	I	$1/4$	$1 - x$
$+$	$I + H$	$1/8$	x
$-$	$I - H$	$1/8$	x

$$a_f = \mathrm{Tr}(E_f A), \quad b_f = \mathrm{Tr}(E_f B), \quad w_f = \mathrm{Tr}(E_f AB).$$

For right output $j = 1, 3$, put $\ell_1 = 1 - x_R$ and $\ell_3 = x_R$. The grouped odd matrix has entries

$$\begin{aligned} (J_{f,j})_{01,01} &= k_f a_f x_f \ell_j, & (J_{f,j})_{10,10} &= k_f b_f (1 - x_f)(1 - \ell_j), \\ (J_{f,j})_{01,10} &= \sigma_{f,j} k_f w_f \sqrt{x(1-x) x_R (1 - x_R)}, \end{aligned} \tag{C.1}$$

where $\sigma_{\pm,1} = 1$, $\sigma_{\pm,3} = -1$, and $\sigma_{I,j} = -\sigma_{\pm,j}$. The other entries vanish. The following mapping specifies every odd raw outcome. Occupations in the middle column refer to left modes $(0, 2, 4)$; append one photon in right mode j and zero in the other right mode.

Group	Left raw occupations	Raw conditional matrix
I, j	$(1,1,0),(0,1,1)$	$J_{I,j}/2$ for each
$+, j$	$(2,0,0),(0,0,2)$	$J_{+,j}/2$ for each
$-, j$	$(1,0,1)$	$J_{-,j}$
forbidden	$(0,2,0)$	0

The other component-count sectors, left photon number 3 and 1, fix respectively $|00\rangle$ and $|11\rangle$. Every raw state in them is a nonnegative multiple of that product projector, and their total (P, N) is $(1/2, 1/4)$. Together these cases account for all 35 occupations, including the impossible component totals.

To derive (C.1), the two odd logical amplitudes have data weights $(\sqrt{x}, \sqrt{1-x})$ when they reach the auxiliary port and $(\sqrt{1-x}, -\sqrt{x})$ when they bypass it. The right weights are $(\sin\phi, \cos\phi)$ at output 1 and $(-\cos\phi, \sin\phi)$ at output 3. The auxiliary comparison resolves the symmetric and antisymmetric internal effects $(I \pm S)/2$, where S swaps the two internal factors. Tracing the independent auxiliary gives $I \pm H$, with half of the symmetric probability in each bunch output. The cross contraction is $\mathrm{Tr}[(I \pm H)AB]$; bypass gives $\mathrm{Tr}(AB)$. Multiplication by the reference factor $1/4$ gives the table and its signs. Proportional raw copies have the same optimal recovery, so grouping preserves all acceptance policies.

Lemma C.3 (Balancing the final auxiliary mixing). In $B_{04}(\gamma)B_{13}(\phi)B_{02}(\theta)$, a balanced final auxiliary mixing $\gamma = \pi/4$ dominates every γ at the same P and at least the same N, under the allowed policies and local corrections above.

Proof. Keep θ, ϕ, A, B, H fixed. Write $\zeta = \sin^2(2\gamma)$ and use the balanced grouped matrices just defined. For each right output,

$$
\begin{aligned}
J_{\mathrm{bunch,total}}(\gamma) &= \zeta J_+, \\
J_{\mathrm{coincidence}}(\gamma) &= J_- + (1-\zeta)J_+, \\
J_{\mathrm{bypass},0}(\gamma) &= \sin^2\gamma\, J_I, \\
J_{\mathrm{bypass},4}(\gamma) &= \cos^2\gamma\, J_I.
\end{aligned}
$$

Each bunch is $\zeta J_+/2$. On the one-data-one-auxiliary input, the coincidence amplitude operator is $\cos^2\gamma I - \sin^2\gamma S$, whose effect is $\Pi_- + \cos^2(2\gamma)\Pi_+$; the two bunch effects are $\sin^2(2\gamma)\Pi_+/2$. These identities hold for the full reference matrices after tracing all internal modes. The single-photon bypass splits one proportional matrix. The product sectors keep their total probabilities. Thus classical coarse graining of the balanced outcomes simulates the unbalanced instrument. Since $w_{\mathrm{B}}(J)$ is convex and positively homogeneous, separate recovery of the finer labels gives at least the simulated N. The matching fractional policy preserves P. The argument is pointwise in the auxiliary, hence survives its optimization and classical mixtures of settings. □

C.2 Classification by conserved component photon numbers

Form the interaction graph including every vacuum input vertex. The total photon number in each connected component is conserved. Group logical inputs by these component photon counts. For logical input (j, k), a component C contains $1_{A_j \in C} + 1_{B_k \in C} + 1_{h \in C}$ photons, where h denotes the auxiliary input mode 4. Let $\mathrm{comp}(i)$ label the connected component containing mode i, and let e_C be the standard basis vector indexed by component C. Put $d_A = e_{\mathrm{comp}(A_0)} - e_{\mathrm{comp}(A_1)}$ and $d_B = e_{\mathrm{comp}(B_0)} - e_{\mathrm{comp}(B_1)}$. Coherence between logical inputs differing in both bits requires $d_A = \pm d_B$.

If both vectors are nonzero, two components must each contain one A rail and one B rail. If the auxiliary belongs to either component, connecting its three required input vertices costs at least two edges; the other pair costs a third. These components use all three edges, leaving none to attach an additional vacuum vertex. The auxiliary-assisted case is a $3+2$ tree with no repeated interaction. If the auxiliary is in a separate component, including possibly vacuum vertices, its outcome multiplies the bare-data instrument by a nonnegative scalar. The arbitrary-vacuum bare bound below applies.

If $d_A = d_B = 0$, the two rails of each data qubit lie in the same component. Separate data components give factorized, hence separable, reference states. Connecting all four data rails into one component uses all three edges, isolating the auxiliary from them and leaving another bare-data case. Any other group fixes at least one logical bit and is separable. This exhausts repeated edges, triangles and finer partitions as well as all finite vacuum extensions.

In a $3+2$ tree that can entangle the references, the two-mode mixing commutes with those of the three-mode component. The first mixing in that component brings together AB, AH or BH. These give the data-first ordering (the auxiliary enters last), AH-first and BH-first, respectively. On a tree one can assign vertex phases so that each edge phase is the difference of its endpoint phases. A common diagonal conjugation makes every mixing real. Output phases vanish under PNR; input data phases are known local reference frames; the auxiliary input phase is global. Because the graph is a tree, no cycle phase arises.

C.3 Bare-data frontier with arbitrary vacuum extensions

Lemma C.4 (Bare-data frontier). With no auxiliary photons, arbitrary passive data optics and vacuum extensions have

$$P^*_{\rm bare}(F_{\rm tar}) = \begin{cases} 1, & F_{\rm tar} \le 1/2 + \tau/4, \\ \tau/[4(F_{\rm tar} - 1/2)], & 1/2 + \tau/4 < F_{\rm tar} \le 1/2 + \tau/2, \\ 0, & F_{\rm tar} > 1/2 + \tau/2. \end{cases}$$

For $\tau = 0$ the rate is one up to $F_{\rm tar} = 1/2$ and zero above.

The following proof uses the standard concurrence convex roof [34]. For distinct output rows r, s, let $a = U_{r,A} \otimes U_{s,B}$ and $b = U_{s,A} \otimes U_{r,B}$, each rail vector having two entries. Their reference state is

$$J_{rs} = \frac{aa^\dagger + bb^\dagger + \tau(ab^\dagger + ba^\dagger)}{4}, \qquad \tau = \mathrm{Tr}(AB). \tag{C.2}$$

Writing $D_{rs} = \det U_{rs,A} \det U_{rs,B}$, the two subnormalized vectors $\sqrt{(1 \pm \tau)/8}(a \pm b)$ have a diagonal spin-flip matrix with singular values $(1 \pm \tau)|D_{rs}|/4$. Thus their unnormalized concurrence is

$$C(J_{rs}) = \tau |D_{rs}|/2. \tag{C.3}$$

For every two-qubit positive matrix, $w_{\rm B}(J) \le (\mathrm{Tr}\, J + C(J))/2$: the pure-state equality and the convex-roof definition of concurrence prove this bound. Let the norms of a, b be a_0, b_0 and their normalized local-overlap magnitudes be $o_A, o_B \in [0, 1]$. Then

$$\begin{aligned} 2|D_{rs}| &= 2a_0 b_0 \sqrt{(1 - o_A^2)(1 - o_B^2)} \\ &\le a_0^2 + b_0^2 - 2a_0 b_0 o_A o_B \le 4\,\mathrm{Tr}\, J_{rs}, \end{aligned} \tag{C.4}$$

using $\sqrt{(1 - o_A^2)(1 - o_B^2)} \le 1 - o_A o_B$ and $\tau \le 1$. Consequently $C(J_{rs}) \le \tau\,\mathrm{Tr}\, J_{rs}$ and each outcome has fidelity at most $(1+\tau)/2$. Bunched outputs are product-supported.

Orthonormal input columns and Cauchy–Binet yield

$$\sum_{r<s} |D_{rs}| \le \sqrt{\sum_{r<s} |\det U_{rs,A}|^2 \sum_{r<s} |\det U_{rs,B}|^2} = 1. \tag{C.5}$$

Therefore every fractional policy has $N - P/2 \le \tau/4$, in addition to $P \le 1$ and the individual fidelity ceiling. All are preserved by classical mixtures of settings. Two balanced data mixings attain probability $1/2$ at fidelity $(1 + \tau)/2$ and probability $1/2$ in product outcomes. Fractional product acceptance saturates the middle segment of the frontier.

C.4 Conditional states for AH-first

On local triple coordinates $(0, 2, 4)$, write $z = \cos^2\gamma$ and $\kappa_\theta = \cos 2\theta$. The matrix is

$$U_{AH} = \begin{pmatrix} \cos\theta\sqrt{z} & \sin\theta & \cos\theta\sqrt{1-z} \\ \sin\theta\sqrt{z} & -\cos\theta & \sin\theta\sqrt{1-z} \\ \sqrt{1-z} & 0 & -\sqrt{z} \end{pmatrix}. \tag{C.6}$$

All component-count sectors except $2+1$ fix the logical state. Within that sector, mode 4 never couples again after the first interaction. Let $n_4 = y_4$ be the number of photons detected in mode 4. If $n_4 = 2$, the A branch is revealed and the state is product-supported. If $n_4 = 0$, a left bunch has normalized odd coherence ξ_+. A coincidence in 0 and 2 has normalized coherence

$$\frac{|\kappa_\theta(\tau + w)|}{\sqrt{(1 + q_A)[(1 + q_B)\kappa_\theta^2 + 1 - q_B]}} \le \xi_+. \tag{C.7}$$

For $n_4 = 1$, expansion of the two creation operators gives A-branch squared norm $z^2 + (1 - z)^2 - 2z(1 - z)q_A$, B-branch squared norm z, and cross coefficient $\sqrt{z}[z\tau - (1 - z)w]$. Dividing by the geometric mean of the branch squared norms gives, when both norms are positive,

$$\chi_A(z) = \frac{|z\tau - (1 - z)w|}{\sqrt{z^2 + (1 - z)^2 - 2z(1 - z)q_A}}. \tag{C.8}$$

The right output mixing changes the two logical populations while preserving this normalized coherence. If a branch norm vanishes, positivity removes its cross term and the outcome is product-supported or zero.

C.5 Domination and attainment

Put $\xi_{\max} = \max(\xi_+, \xi_-)$. Since $2[z\tau - (1 - z)w] = (2z - 1)(\tau + w) + (\tau - w)$, the response bounds give

$$\begin{aligned} |z\tau - (1 - z)w| &\le \frac{\xi_{\max}}{2}\left[|2z - 1|\sqrt{(1 + q_A)(1 + q_B)} + \sqrt{(1 - q_A)(1 - q_B)}\right] \\ &\le \xi_{\max}\sqrt{z^2 + (1 - z)^2 - 2z(1 - z)q_A}. \end{aligned} \tag{C.9}$$

The last step is Cauchy–Schwarz, using $(1+q_B)+(1-q_B) = 2$. For positive branch populations, division by the denominator in Eq. (C.8) yields $\chi_A(z) \le \xi_{\max}$. If either branch population vanishes, positivity removes its coherence; the outcome is product-supported or zero. At $z = 1$, Eq. (C.9) also gives $\tau \le \xi_{\max}$. BH-first interchanges A and B and conjugates w.

For data-first orderings, Lemma C.3 reduces the final auxiliary mixing to a balanced one. The grouped states in Eq. (C.1) then have normalized coherences ξ_+, ξ_- or τ, all at most $\xi_{\max}$. Applying Lemma C.2 to these outcomes and to the auxiliary-first outcomes gives fidelity at most $(1 + \xi_{\max})/2$. For the chosen response $e \in \{+1, -1\}$, balanced auxiliary and right mixings in the data-first ordering, with $\cos^2\theta = (1 + eq_B)/(2 + e(q_A + q_B))$, balance the two populations.

This attains $(1+\xi_e)/2$ at positive probability whenever $\xi_e > 0$. If both response visibilities vanish, a product outcome attains $1/2$. Acceptance policies and classical mixtures preserve the fidelity ceiling.

C.6 Deterministic single-copy LOCC recovery

For an unnormalized two-qubit state use concurrence $C(J) = \mathrm{Tr}(J)C(J/\mathrm{Tr}\,J)$ for $J \neq 0$ and $C(0) = 0$. The preceding proof gives a bound for every raw outcome:

$$C(J_y) \leq \xi_{\max} \mathrm{Tr}\, J_y, \qquad \xi_{\max} = \max(\xi_+, \xi_-). \quad \text{(C.10)}$$

Indeed a nontrivial auxiliary-assisted $3+2$ outcome occupies one parity block, with populations x, y and coherence z. Its concurrence is $2|z| \leq 2\xi_{\max}\sqrt{xy} \leq \xi_{\max}(x+y)$. An auxiliary-isolated outcome obeys the bare-data bound $C(J) \leq \tau\,\mathrm{Tr}\,J$, and $\tau \leq \xi_{\max}$ was proved above. The remaining groups are separable.

For a trace-preserving, outcome-dependent, single-copy LOCC channel Λ_y, concurrence monotonicity [34, 35] gives $C(\Lambda_y(J_y)) \leq C(J_y)$, including all its classical outcomes. The general convex-roof bound $w_{\mathrm{B}}(K) \leq [\mathrm{Tr}\,K + C(K)]/2$ therefore gives

$$w_{\mathrm{B}}(\Lambda_y(J_y)) \leq (1+\xi_{\max})\,\mathrm{Tr}\,J_y/2.$$

The inequality survives acceptance based on the interface outcome and preselected mixtures of settings. The local-unitary attainers remain permitted, so the highest-fidelity equality is unchanged.

C.7 Pure-data optimization and minimum mixing count

For $A = |a\rangle\langle a|$, $B = |b\rangle\langle b|$, each response $E_\pm = I \pm H$ is positive. Its normalized coherence satisfies

$$\xi_\pm = c\frac{|\langle b|E_\pm|a\rangle|}{\sqrt{\langle a|E_\pm|a\rangle\langle b|E_\pm|b\rangle}} \leq c.$$

Choose $H = |h_-\rangle\langle h_-|$ and the balanced data-first tree. The outcomes $(1,1,0,0,1)$ and $(1,0,0,1,1)$ each have probability $(1+c)/32$ and fidelity $(1+c)/2$. This proves attainability of the pure-data three-mixing ceiling. Equation (20) and the four-mixing realization prove Theorem 2.

D The connected four-mixing instrument

D.1 Explicit unitary and six-path calculation

On input columns (A_0, A_1, B_0, B_1, h) the circuit is exactly

$$U_{\mathrm{opt}} = \begin{pmatrix} 1/2 & 1/(2\sqrt{3}) & 1/2 & 1/(2\sqrt{3}) & 1/\sqrt{3} \\ 1/2 & -1/(2\sqrt{3}) & 1/2 & -1/(2\sqrt{3}) & -1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{3} & 0 & 1/\sqrt{3} & -1/\sqrt{3} \end{pmatrix}. \quad \text{(D.1)}$$

Its product decomposition is Eq. (14); multiplying the four real pair matrices gives the displayed unitary. Output modes $(0,1,4)$ are the occupied slots of the outcome $y_* = (1,1,0,0,1)$. For each permutation π of (A,B,H) into those slots, let Y_π be the corresponding tensor product of internal vectors and Z_π the four-component product of spatial amplitudes, in reference order $(00,01,10,11)$. Collect these columns in $Z = (Z_\pi)$. In permutation order $(ABH, AHB, BAH, BHA, HAB, HBA)$,

$$12\sqrt{3}Z = \begin{pmatrix} -3 & 0 & -3 & 0 & 0 & 0 \\ \sqrt{3} & -2\sqrt{3} & -\sqrt{3} & 0 & 2\sqrt{3} & 0 \\ -\sqrt{3} & 0 & \sqrt{3} & -2\sqrt{3} & 0 & 2\sqrt{3} \\ 1 & -2 & 1 & -2 & -2 & -2 \end{pmatrix}. \quad \text{(D.2)}$$

The reference amplitude factor is $1/2$. Tracing the internal slots gives

$$J_{y_*} = \tfrac{1}{4} Z K Z^\dagger, \qquad K_{\pi\sigma} = \langle Y_\sigma | Y_\pi \rangle. \quad \text{(D.3)}$$

For these mixed-source formulas, let A and B denote the data-photon internal density operators, corresponding to ρ_a and ρ_b in the main text. Define $\tau = \mathrm{Tr}(AB)$, $q_A = \mathrm{Tr}(AH)$, $q_B = \mathrm{Tr}(BH)$, and $w = \mathrm{Tr}(HAB)$. Average Eq. (D.3) over the independent spectral ensembles. Equivalently, cycles of the relative permutation give products of density-matrix traces and six invariants $1, \tau, q_A, q_B, w, w^*$. This partial-distinguishability path expansion [20] is specialized here to the four-mixing instrument.

Direct multiplication of (D.2) and (D.3) annihilates the off-diagonal Bell entries and yields

$$\begin{aligned} \ell_{\Phi^+} &= (1+\tau+q+2\,\mathrm{Re}\,w)/144, \\ \ell_{\Phi^-} = \ell_{\Psi^+} &= (2+2\tau-q-2\,\mathrm{Re}\,w)/144, \\ \ell_{\Psi^-} &= (1-\tau-q+2\,\mathrm{Re}\,w)/48. \end{aligned} \quad \text{(D.4)}$$

Here $q = q_A + q_B$. These are unnormalized weights, hence nonnegative on physical source triples. Their sum is $(4+\tau-2q+2\,\mathrm{Re}\,w)/72$. The formula determines the local-unitary Bell fidelity by taking the largest weight and dividing by the trace. A Pauli rotation on one reference maps its Bell state to Φ^+; the Φ^+ weight requires identity recovery. For an entangled

Bell-diagonal conditional state, $C = 2F - 1$, so the concurrence bound and its monotonicity also exclude any improvement by deterministic LOCC [24, 34].

D.2 Symmetric and antisymmetric weights

For two internal factors, let $S_{AB}|i,j\rangle = |j,i\rangle$ and $\Pi_{\rm sym}^{(2)} = (I + S_{AB})/2$. In any finite internal dimension, $\mathrm{Tr}[S_{AB}(A \otimes B)] = \sum_{i,j} A_{ji}B_{ij} = \mathrm{Tr}(AB) = \tau$. Since $\mathrm{Tr}\, A = \mathrm{Tr}\, B = 1$, it follows that $\mathrm{Tr}[\Pi_{\rm sym}^{(2)}(A \otimes B)] = (1 + \tau)/2$, the attainable bare-data fidelity ceiling of Sec. C.

Let V_π permute three internal tensor factors. The projectors $\Pi_{\rm sym} = \sum_{\pi\in S_3} V_\pi/6$ and $\Pi_{\rm anti} = \sum_{\pi\in S_3} \mathrm{sgn}(\pi)V_\pi/6$ give, on $\rho_{\rm int} = A \otimes B \otimes H$,

$$s_{\rm sym} = \mathrm{Tr}(\Pi_{\rm sym}\rho_{\rm int}) = (1 + \tau + q_A + q_B + 2\,\mathrm{Re}\, w)/6, \tag{D.5}$$

$$s_{\rm anti} = \mathrm{Tr}(\Pi_{\rm anti}\rho_{\rm int}) = (1 - \tau - q_A - q_B + 2\,\mathrm{Re}\, w)/6. \tag{D.6}$$

The identity contributes one, the three transpositions contribute τ, q_A, q_B, and the two three-cycles contribute w, w^*. Equation (D.4) therefore gives exactly

$$\ell_{\Phi^+} = s_{\rm sym}/24, \qquad \ell_{\Psi^-} = s_{\rm anti}/8. \tag{D.7}$$

This derivation is linear in each density operator and applies in any finite internal dimension for pure or mixed, commuting or noncommuting source states. For pure wave packets the cyclic product contains the three-photon phase, which pairwise overlap magnitudes alone do not fix [22]. If all occupied internal states lie in a common space of dimension at most two, its third exterior power vanishes, so $s_{\rm anti} = 0$.

D.3 Which amplitudes cancel

Projecting the reference rows of (D.2) onto the Bell basis gives the following internal amplitude coefficient rows, all multiplied by $1/(24\sqrt{6})$ after the reference normalization:

	ABH	AHB	BAH	BHA	HAB	HBA
Φ^+	-2	-2	-2	-2	-2	-2
Φ^-	-4	2	-4	2	2	2
Ψ^+	0	$-2\sqrt{3}$	0	$-2\sqrt{3}$	$2\sqrt{3}$	$2\sqrt{3}$
Ψ^-	$2\sqrt{3}$	$-2\sqrt{3}$	$-2\sqrt{3}$	$2\sqrt{3}$	$2\sqrt{3}$	$-2\sqrt{3}$

(D.8)

For the midpoint auxiliary choose $|h_+\rangle = |0\rangle$ and

$$|a\rangle = \sqrt{(1+c)/2}|0\rangle + \sqrt{(1-c)/2}|1\rangle,$$
$$|b\rangle = \sqrt{(1+c)/2}|0\rangle - \sqrt{(1-c)/2}|1\rangle.$$

The last row is the alternating tensor of three vectors in dimension two and therefore vanishes. In each middle row the coefficients sum to zero, canceling the all-zero internal term. Pairing permutations that exchange A and B cancels every term with exactly one label 1. The remaining terms have coefficient proportional to $(1 - c)/2$. Their squared norms, with the factors in (D.8), are $(1 - c)^2/144$ for each erroneous component. The first row has squared norm $(1 + c + c^2)/72$. This proves the pure-family Bell weights and the quadratic error law.

D.4 Polarization implementation

Use mode order $(a_H, a_V, b_H, b_V, h_V, h_H)$, with h_H vacuum. Internal frequency or time states are independent of logical polarization; each element acts identically on that unresolved degree of freedom. The symmetric beam-splitter convention is

$$S(T) = \begin{pmatrix} \sqrt{T} & i\sqrt{1-T} \\ i\sqrt{1-T} & \sqrt{T} \end{pmatrix}.$$

Set $U_{\rm NPBS} = S_{02}(1/2)S_{13}(1/2)$ and $U_{\rm PPBS} = S_{14}(1/3)$ with $T_H = 1$ [26]. The HWP Jones matrix has diagonal entries $\cos 2\theta, -\cos 2\theta$ and off-diagonal entries $\sin 2\theta$, so $U_{\rm HWP}(\pi/8) = B_{01}(\pi/4)$. Direct multiplication gives

$$U_{\rm pol} = U_{\rm HWP}(\pi/8)U_{\rm PPBS}U_{\rm NPBS} = D_{\rm ph}(U_{\rm opt} \oplus 1)D_{\rm ph}, \quad D_{\rm ph} = \mathrm{diag}(1, 1, i, i, i, 1). \tag{D.9}$$

Input phases are common to the two rails of each data photon and global for the auxiliary; output phases are common to every path of one detection pattern. They leave the entire conditional instrument unchanged. An NPBS supplies two elementary mode mixings, and the PPBS and HWP one each. This distinguishes the mathematical mixing count from the physical element count.

E Purity law and optical realizations

E.1 Source controls at fixed support and purity

Restrict the auxiliary's entire support to $\mathcal{S} = \mathrm{span}\{a, b\}$. In the orthonormal midpoint basis h_+, h_- write

$$H = \tfrac{1}{2}(I + \mathbf{h}\cdot\boldsymbol{\sigma}), \qquad \mathbf{h} = (x, y, z), \quad |\mathbf{h}| = r_h \in [0, 1], \tag{E.1}$$

$$\operatorname{spec} H = \{(1 + r_h)/2, (1 - r_h)/2\}, \qquad \operatorname{Tr} H^2 = (1 + {r_h}^2)/2. \tag{E.2}$$

The length r_h determines the purity, while the direction of $\mathbf{h}$ sets its overlaps with the data states. Define the two axis densities first:

$$H_{\pm r_h} = \tfrac{1}{2}(I \pm r_h \sigma_z). \tag{E.3}$$

The source-control hierarchy is

$$G_{\mathcal{R}}(c; H) = \sup_{U, \boldsymbol{n}: p_{\boldsymbol{n}} > 0} F_{\mathcal{R}}(J_{\boldsymbol{n}}(U, H)), \tag{E.4}$$

$$G_{\mathcal{R},\text{axes}}(c, r_h) = \max_{\zeta = \pm 1} G_{\mathcal{R}}(c; H_{\zeta r_h}), \tag{E.5}$$

$$G_{\mathcal{R},\text{spec}}(c, r_h) = \sup_{|\mathbf{h}| = r_h} G_{\mathcal{R}}(c; H). \tag{E.6}$$

The suprema allow every finite vacuum extension, with no probability floor.

Proposition E.1 (Recovery equivalence without a probability floor). For each specified auxiliary density matrix H, $G_{\text{LU}}(c; H) = G_{\text{LOCC}}(c; H) = G_{\text{SEP}}(c; H)$.

Proof. Each nonzero local branch is realized by an admissible optical outcome with the same H and normalized state. For a trace-preserving channel, balanced local factors make $T \operatorname{diag}(L^T, M^T, 1)$ a contraction without extra attenuation; the explicit construction is proved in Proposition F.7. The recovered target fidelity is a probability-weighted mean of branch fidelities, each at most the global LU supremum. The reverse inequalities follow by inclusion. A separable map in finite dimension admits a finite product Kraus decomposition; closure limits obey the same bound by continuity. □

We henceforth write G, G_{axes} and G_{spec} for these fidelity functions without a probability constraint. Negative-axis H_{-r_h} with the opposite-axis construction gives f_-; positive-axis H_{+r_h} with the trine gives f_+:

$$f_- = \frac{1 + c^2 + 2cr_h}{2 + 2cr_h}, \qquad f_+ = \frac{1 + c^2 + cr_h}{2 + 2c^2 - cr_h}. \tag{E.7}$$

One realization of the first uses $U_{\text{opp}} = B_{04}(\pi/4)B_{13}(\pi/4)B_{02}(\pi/4) \in U(5)$, accepting $(1, 1, 0, 0, 1)$; its probability is $(1+cr_h)/32$. The coupler convention is that of Eq. (14). The canonical U_{trine} of Sec. A with H_{+r_h} has probability $(2 + 2c^2 - cr_h)/72$. Both formulas follow by the six assignments already displayed. The exact two-axis envelope is proved at the end of this section.

For the asymptotic problem set

$$E(c, r_h) = 1 - G_{\text{spec}}(c, r_h), \quad D = 1 - c^2, \quad b_c = 2 + c^2. \tag{E.8}$$

The equality of the unconstrained suprema removes the recovery subscript from E. The converse holds for SEP and the constructions use LU. The impurity parameter satisfies $1 - \operatorname{Tr} H^2 = \eta_h(1 - \eta_h/2)$.

E.2 A source-dependent spectral envelope

The spectral argument extends to a mixed auxiliary density. Its covariance changes the Gram weights, while the optical product columns and their shared sum stay fixed. Reweighting the columns preserves their rank-one character, so the product-column bound applies to the new compressed spectrum. At a collision-free outcome, let the optical rows be $(\alpha_i^T, \beta_i^T, v_i)$, $i = 0, 1, 2$. For $\{i, j, k\} = \{0, 1, 2\}$, put

$$X_i = \tfrac{1}{2}\alpha_i(v_k\beta_j + v_j\beta_k)^T, \qquad Y_i = \tfrac{1}{2}(v_k\alpha_j + v_j\alpha_k)\beta_i^T.$$

All six matrices have rank at most one and $\sum_i X_i = \sum_i Y_i$. Let W_{prod} have their vectorizations as columns, in the order $(X_0, X_1, X_2, Y_0, Y_1, Y_2)$. Use the internal Gram blocks G_0, G_1 of Eq. (A.20). Writing $h = h_a a + h_b b$, the coefficient order (h_b, h_a) associated with (X, Y) gives the covariance

$$C_{11} = \frac{1 - cz}{2D} - \frac{x}{2\sqrt{D}}, \qquad C_{22} = \frac{1 - cz}{2D} + \frac{x}{2\sqrt{D}}, \qquad C_{12} = \frac{z - c}{2D} - \frac{iy}{2\sqrt{D}}. \tag{E.9}$$

Tracing over the internal states yields

$$J = W_{\text{prod}} \begin{pmatrix} C_{11}G_0 & C_{12}G_1 \\ \overline{C_{12}}G_1 & C_{22}G_0 \end{pmatrix} W_{\text{prod}}^{\dagger}. \tag{E.10}$$

The normalization $1/2$ in each product column is the amplitude of the two ideal Bell inputs. For $r_h < 1$, $C_{11}, C_{22} > 0$. Reweight the two triples of columns by $\sqrt{C_{11}}, \sqrt{C_{22}}$, preserving their product character. Their conjugate rows obey the weighted shared-sum constraint $\sum_i X_i'/\sqrt{C_{11}} = \sum_i Y_i'/\sqrt{C_{22}}$, where $X_i' = \sqrt{C_{11}}X_i$ and $Y_i' = \sqrt{C_{22}}Y_i$. The compression to this five-dimensional subspace has eigenvalues

$$D\{\ell, 1 + \xi, 1 + \xi, 1 - \xi, 1 - \xi\}$$

$$\ell = \frac{1 + c^2 + cz}{1 - cz}$$

$$\xi = \frac{c\sqrt{(c - z)^2 + Dy^2}}{\sqrt{(1 - cz)^2 - Dx^2}}.$$

Lemma A.2 applies to these reweighted product columns and this compressed spectrum, giving

$$F_{\mathrm{LU}}(J) \le \max\{F_A, F_B, F_C\}, \qquad F_A = \frac{1+\xi}{2}, \quad F_B = \frac{\ell}{\ell+1-\xi}, \quad F_C = \frac{1+\xi}{\ell+1+\xi}. \tag{E.11}$$

All denominators are positive for $c \in (0,1), r_h < 1$.

At fixed z, r_h, ξ is largest at $x = 0$. This follows by writing its numerator squared before multiplication by c^2 as $N = (c-z)^2 + Dy^2$, and its denominator squared as $N + D(1-{r_h}^2)$. Maximizing N at fixed z, r_h gives $x = 0$. The maximum over $z \in [-r_h, r_h]$ is uniquely at $z = -r_h$ for $0 < r_h < 1$, and $\xi \le c(c+r_h)/(1+cr_h)$. Thus $F_A \le f_-$.

The orientation optimization of this bound reduces to one real variable. Set

$$d_{\mathrm{cov}} = (1-c^2)(1-{r_h}^2)$$
$$q = 1 - cz$$
$$q \in [1-cr_h, 1+cr_h]$$
$$w = \sqrt{q^2 - d_{\mathrm{cov}}}.$$

$$\phi(q) = c(b_c q - d_{\mathrm{cov}}) - b_c\sqrt{q^2 - d_{\mathrm{cov}}}, \qquad \phi'(q) = b_c\left(c - \frac{q}{\sqrt{q^2 - d_{\mathrm{cov}}}}\right) < 0.$$

At the upper endpoint, $\phi(1+cr_h) = -D[c(1-{r_h}^2) + b_c r_h] < 0$. Consequently its only possible interior maximum is at

$$q_0 = \frac{\sqrt{d_{\mathrm{cov}}[(1-c^2)b_c^2 + c^2 d_{\mathrm{cov}}]} - c^2 d_{\mathrm{cov}}}{(1-c^2)b_c}$$
$$\widehat{q} = \max\{1-cr_h, q_0\}.$$

We obtain the spectral upper envelope

$$G_{\mathrm{spec}}(c, r_h) \le \bar{U}(c, r_h) := \max\{f_-(c, r_h), F_B(\widehat{q})\}. \tag{E.12}$$

Whenever its right side equals G_{axes}, the known axial attainers close the global problem at that point.

For bunched outcomes the exchange coefficient from Sec. A.3 becomes

$$\xi_+ = \frac{c\sqrt{(3c+z)^2 + Dy^2}}{\sqrt{(3+cz)^2 - Dx^2}} \le \frac{c(3c+r_h)}{3+cr_h}.$$

The maximal values at fixed z of the three ratios are

$$F_A(q) = \frac{1+cw/q}{2}$$
$$F_B(q) = \frac{b_c - q}{b_c - cw}$$
$$F_C(q) = \frac{q+cw}{b_c + cw}.$$

The first and third are increasing. Their endpoint maxima are f_- and $(1+c^2+2cr_h)/(2+2c^2+cr_h)$, respectively. The latter is at most f_- for $r_h \le 2c$ and below f_+ for $r_h > 2c$. The derivative of the middle ratio has the sign of

The proof is the same convex-quadratic argument, now with denominator excess $D(9-{r_h}^2)$; its maximum is at H_{+r_h}. For $r_h > 0$,

$$\frac{c(c+r_h)}{1+cr_h} - \frac{c(3c+r_h)}{3+cr_h} = \frac{2cr_h(1-c^2)}{(1+cr_h)(3+cr_h)} > 0.$$

Since the fidelity of a bunched outcome is at most $(1+\xi_+)/2$, it is strictly below the achievable f_-. Fully bunched outcomes are separable.

E.3 A universal purity cost

Theorem E.2 (Purity-dependent error bound). For $0 < c < 1$, $0 \le r_h \le 1$, $\operatorname{supp} H \subseteq \mathcal{S}$, and each specified recovery class,

$$E(c, r_h) \ge \max\left\{\frac{t^2}{2-c+2c^2}, \min\left[\frac{D}{2(1+cr_h)}, \frac{D\sqrt{1-{r_h}^2}}{2+c^2}\right]\right\}, \tag{E.13}$$

while the explicit axial sources give

$$E(c, r_h) \le \min\left\{\frac{D}{2(1+cr_h)}, \frac{t^2 + 2c(1-r_h)}{2+2c^2-cr_h}\right\}. \tag{E.14}$$

Proof. The first lower term is Theorem A.1. For $r_h < 1$, the spectral reduction above bounds every collision-free outcome by $\max\{f_-, \sup_q F_B(q)\}$. The F_A branch is at most f_-. The F_C branch is at most f_- when $r_h \le 2c$; for $r_h > 2c$ it is below $f_+ = F_B(1-cr_h)$. Bunched outcomes obey $F_{\mathcal{R}} \le f_-$ by their exchange bound, and fully bunched outcomes are separable. Thus all occupations are covered by this envelope.

Put $d_{\mathrm{cov}} = D(1-{r_h}^2)$ and $w = \sqrt{q^2 - d_{\mathrm{cov}}}$, with $q \in [1-cr_h, 1+cr_h]$. Its radicand is nonnegative because at the lower endpoint $q^2 - d_{\mathrm{cov}} = (c-r_h)^2$. Moreover $q > 0$, $w \le q$, and $b_c - cw \ge b_c - c(1+cr_h) > 0$. The identity

$$(q-cw)^2 - D^2(1-{r_h}^2) = (cq-w)^2 \tag{E.15}$$

and $q - cw > 0$ imply

$$1 - F_B(q) = \frac{q-cw}{b_c - cw} \ge \frac{D\sqrt{1-{r_h}^2}}{b_c}. \tag{E.16}$$

Combining this with $1 - f_- = D/[2(1+cr_h)]$ proves the second lower term. Product-Kraus branches preserve

the source density and the product-column constraints by Eq. (A.5); summing their homogeneous inequalities gives the SEP result. At $r_h = 0$ the formula applies directly to $H = I/2$. At $r_h = 1$ its purity term is zero and Theorem A.1 supplies the bound. Finally, Eq. (E.14) is $1 - \max(f_-, f_+)$ for the two explicit LU constructions. □

Corollary E.1 (Necessary and sufficient purity for quadratic error). For any path $r_h = r_h(c)$ as $c \to 1^-$,

$$\begin{aligned} E(c, r_h(c)) &= O((1-c)^2) \\ &\Longleftrightarrow \\ 1 - r_h(c) &= O((1-c)^2) \\ &\Longleftrightarrow \\ 1 - \operatorname{Tr} H^2 &= O((1-c)^2). \end{aligned} \tag{E.17}$$

The sufficient source orientation has a heralding probability tending to 1/24. For every fixed $r_h < 1$, $E(c, r_h) = \Theta(1-c)$.

Proof. If $E \le Ct^2$, the first entry in the inner minimum of Eq. (E.13) is at least $D/4 \sim t/2$, and exceeds Ct^2 for small t. The second must therefore be at most Ct^2. Since $D \sim 2t$ and $b_c \le 3$, this forces $\sqrt{1 - {r_h}^2} = O(t)$ and $1 - r_h = O(t^2)$. Conversely the aligned trine has exactly

$$\begin{aligned} 1 - f_+ &= \frac{t^2 + 2c\eta_h}{2 + 2c^2 - cr_h} \\ P_{\mathrm{mid}}(c, r_h) &= \frac{2 + 2c^2 - cr_h}{72} \longrightarrow \frac{1}{24}. \end{aligned} \tag{E.18}$$

Its denominator tends to three. Also $1 - \operatorname{Tr} H^2 = (1 - {r_h}^2)/2 = \eta_h(1 + r_h)/2$, which is between $\eta_h/2$ and η_h. For fixed $r_h < 1$, both lower entries are positive constants times t and the opposite-axis construction has error $O(t)$. □

The necessity applies at any positive event probability; sufficiency is attained by orienting the given-purity source along the midpoint.

E.4 Derivation of the exact two-axis benchmark

For H_{-r_h}, the compressed source spectrum is one half times

$$\begin{gathered} \lambda_0 = 1 + c^2 - cr_h \\ \lambda_1 = 1 + c^2 + 2cr_h \\ \text{(twice)} \\ D = 1 - c^2 \\ \text{(twice)}. \end{gathered}$$

When $r_h \le 2c$, the ordering is $\lambda_1 \ge \lambda_0 \ge D$. The product-column inequality then gives $\lambda_1/(\lambda_1 + D) = f_-$. The five-mode circuit U_{opp} defined in Sec. E attains this value on either detection pattern $(1,1,0,0,1)$ or $(1,0,0,1,1)$, each with probability $(1 + cr_h)/32$, using outcome-dependent local corrections. The input order is (A_0, A_1, B_0, B_1, h) and the same coupler convention applies.

For H_{+r_h} and $r_h \ge c$, the extreme spectral values are $(1 + c^2 + cr_h)/2$ and $(1 + c^2 - 2cr_h)/2$. The bound $\lambda_{\max}/(\lambda_{\max} + \lambda_{\min})$ gives f_+, attained by the midpoint trine. If $r_h \le c$, its spectral upper bound is instead $B_+ = (1 + c^2 + cr_h)/(2 + cr_h)$, and

$$f_- - B_+ = \frac{cr_h(1 - c^2)}{(2 + 2cr_h)(2 + cr_h)} > 0 \qquad (r_h > 0).$$

If $r_h > 2c$, the negative-axis upper bound is $\lambda_1/(\lambda_1 + \lambda_0)$, while

$$f_+ - \frac{\lambda_1}{\lambda_1 + \lambda_0} = \frac{3c^2 {r_h}^2}{(2 + 2c^2 - cr_h)(2 + 2c^2 + cr_h)} > 0.$$

Both branches are exact in the overlap. Their equality reduces to $4cr_h^2 + (1 - c^2)r_h - 2c(1 + c^2) = 0$, with the unique positive root $r_{\mathrm{ax}}(c) = [\sqrt{1 + 30c^2 + 33c^4} - (1 - c^2)]/(8c)$. At $r_h = 0$ both choices are the same density. The bunched-outcome bounds of Sec. E.2, applied branchwise under deterministic local recovery, complete the all-network comparison between these two source choices.

F A physical tilted construction and its error

For $0 < \nu \le 1$, define the raw shape and its physical row normalization by

$$T_\nu^{\mathrm{raw}} = \begin{pmatrix} -\sqrt{\nu} & 0 & \sqrt{\nu} & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix}, \tag{F.1}$$

$$\widehat{T}_\nu = \operatorname{diag}\left([2(1+\nu)]^{-1/2}, 8^{-1/2}, 8^{-1/2} \right) T_\nu^{\mathrm{raw}}. \tag{F.2}$$

The raw row Gram matrix is $\left(\begin{smallmatrix} 1+2\nu & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 5 \end{smallmatrix}\right)$. In the symmetric and antisymmetric combinations of the last two physical rows, the antisymmetric eigenvalue is $1/2$ and the other block is

$$\begin{pmatrix} (1+2\nu)/(2+2\nu) & [8(1+\nu)]^{-1/2} \\ [8(1+\nu)]^{-1/2} & 3/4 \end{pmatrix}.$$

Its complement has positive diagonal entries and zero determinant. Thus $\widehat{T}_\nu$ is a contraction and Eq. (A.2) gives a real eight-mode unitary. Inputs $0, 1$ are the first logical rails, $2, 3$ the second, input 4 the auxiliary, and inputs $5, 6, 7$ vacuum. The accepted detection pattern is $(1,1,1,0,0,0,0,0)$.

To display the conditional state, put $k_\nu = \sqrt{\nu}$ and $d = \sqrt{D}$. The source overlaps are $A_H = \langle a|H|a\rangle = (1 + cz + dx)/2$, $B_H = \langle b|H|b\rangle = (1 + cz - dx)/2$ and $K_H = \mathrm{Re}\langle a|H|b\rangle = (c + z)/2$. The last two output rows are exchanged by $\sigma_z \otimes \sigma_z$ on the remote qubits Q_A, Q_B, so the state is an X-shaped matrix, with entries only on the diagonal and antidiagonal. For the real sources used here it is

where

$$J = \frac{1}{256(1+\nu)} \begin{pmatrix} j_{00} & 0 & 0 & -L \\ 0 & O_a & w_{\text{odd}} & 0 \\ 0 & w_{\text{odd}} & O_b & 0 \\ -L & 0 & 0 & 1+c^2 \end{pmatrix}, \quad \text{(F.3)}$$

$$j_{00} = 1 + c^2 + 3D\nu - 2k_\nu dx, \qquad L = 1 + c^2 - k_\nu dx, \quad \text{(F.4)}$$

$$O_a = D + \nu(1 - B_H) - k_\nu(D + dx), \qquad O_b = D + \nu(1 - A_H) + k_\nu(D - dx), \quad \text{(F.5)}$$

$$w_{\text{odd}} = -D - \nu(c^2 - cK_H) + k_\nu dx. \quad \text{(F.6)}$$

These entries are obtained by placing each of (a, b, h) in each of the three distinct outputs, multiplying its three spatial amplitudes and tracing the resulting internal tensor. The Bell-input amplitude is $1/2$. The displayed real source is sufficient for the construction. For the fixed target $|\Phi^-\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$, the probability and Bell numerator are

$$p_\nu = \bar{P}/[128(1+\nu)], \qquad \bar{P} = 2 + \tfrac{\nu}{2}(4 - 3c^2 - cz) - 2\sqrt{D\nu}\, x, \quad \text{(F.7)}$$

$$w_{\Phi^-} = \bar{N}/[128(1+\nu)], \qquad \bar{N} = 1 + c^2 + \tfrac{3}{4}D\nu - \sqrt{D\nu}\, x. \quad \text{(F.8)}$$

For every source direction, $cz \le 1$ and $x \le 1$ give

$$\bar{P} \ge 2 + \frac{3}{2}D\nu - 2\sqrt{D\nu} = \frac{4}{3} + \frac{3}{2}(\sqrt{D\nu} - 2/3)^2 \ge \frac{4}{3}. \quad \text{(F.9)}$$

Consequently $p_\nu \ge 1/192$ throughout $0 < c < 1$, $0 \le r_h \le 1$, $0 < \nu \le 1$. Define

$$g_\nu = \sqrt{D\nu + c^2\nu^2/4}, \qquad (x, y, z) = \frac{r_h}{g_\nu}(\sqrt{D\nu}, 0, c\nu/2). \quad \text{(F.10)}$$

The Bloch length is exactly r_h. This source minimizes the *unnormalized* Bell loss for this optical block:

$$\bar{Q} := \bar{P} - \bar{N} = D + \frac{2+3D}{4}\nu - r_h g_\nu. \quad \text{(F.11)}$$

As $t, \eta_h \to 0$ with $\nu \to 0$, $\bar{P} \to 2$, hence $p \to 1/64$ and $1 - f_{\Phi^-} = \bar{Q}/\bar{P}$. For the positive-$x, z$ source in Eq. (F.10), this limit is approached from below:

$$p_\nu - \frac{1}{64} = -\frac{\nu(3c^2 + cz)/2 + 2\sqrt{D\nu}\, x}{128(1+\nu)} < 0. \quad \text{(F.12)}$$

F.1 One tilted mechanism for arbitrary joint paths

Proposition F.1 (Uniform tilted realization). For $t, \eta_h \to 0$, $t > 0$, $\eta_h \ge 0$, choose

$$s = \sqrt{D + \eta_h/2}, \qquad \nu = D/s, \qquad \mathbf{h} = \frac{r_h}{g_\nu}(\sqrt{D\nu}, 0, c\nu/2). \quad \text{(F.13)}$$

The physical block $\widehat{T}_\nu$ and fixed target Φ^- satisfy, uniformly over these joint paths,

$$1 - f_{\Phi^-} = t\sqrt{8t + 2\eta_h}\,[1 + O(s)], \qquad p \longrightarrow 1/64. \quad \text{(F.14)}$$

The principal mode of H has Bloch polar angle

$$\tan\theta_H = \frac{2}{c}(D + \eta_h/2)^{1/4}. \quad \text{(F.15)}$$

Proof. We have $D \le s^2$, $\eta_h \le 2s^2$ and $\nu \le s$. For small s the Taylor remainder of $\sqrt{1 + 4s - D}$ is bounded uniformly. Thus

$$g_\nu = \frac{\nu}{2}\sqrt{1 + 4s - D} = \frac{\nu}{2} + D - \frac{D\nu}{4} - Ds + O(Ds^2), \quad \text{(F.16)}$$

$$\bar{Q} = D + \frac{2+3D}{4}\nu - (1 - \eta_h)g_\nu = D\nu + Ds + \frac{\eta_h\nu}{2} + O(Ds^2) = 2Ds + O(Ds^2). \quad \text{(F.17)}$$

For example the omitted cubic square-root term is $O(\nu s^3) = O(Ds^2)$; the cross term D^2 and the impurity term $\eta_h D$ have the same bound. The probability expression gives $\bar{P} = 2 + O(s)$ for every direction. Consequently $\bar{Q}/\bar{P} = Ds[1 + O(s)]$. As $D = t(2 - t)$, $Ds = t\sqrt{8t + 2\eta_h}\,[1 + O(t)]$ uniformly, and $t = O(s^2)$. This proves Eq. (F.14), including $\eta_h = 0$ and non-power-law paths. Dividing the positive x and z components in Eq. (F.13) gives Eq. (F.15). For sufficiently small s, $\nu \le 1$, so the contraction and its real eight-mode completion apply. □

The data Bloch directions are at $\pm\arccos c$ about the midpoint axis. Their angle is $O(\sqrt{t})$, whereas θ_H has order $(D + \eta_h/2)^{1/4}$. The prescribed principal

mode is therefore eventually outside their short arc. The ket angles are half the Bloch angles.

For $\eta_h \asymp t$, source coherence and the optical imbalance are both of order $t^{1/4}$, leaving error of order $t^{3/2}$.

F.2 Uniform bounds with a finite probability floor

Proposition F.2 (Uniform resource bounds). For $0 < t < 1$ and $0 \le \eta_h \le 1$, with $\operatorname{supp} H \subseteq \mathcal{S}$,

$$\frac{1}{4}\max\{t^2, t\sqrt{\eta_h}\} \le E(1-t, 1-\eta_h) \le 32\min\{t^2+\eta_h, t\sqrt{t+\eta_h}\}. \tag{F.18}$$

The constructive side uses at most eight modes, a fixed Bell target for each prescribed setting, and one detection pattern with $p \ge 1/192$.

Proof. The pure-source term in Eq. (E.13) is at least $t^2/3$. Its remaining entries are at least $t/4$ and $t\sqrt{\eta_h}/3$, respectively. Since $\sqrt{\eta_h} \le 1$, this proves the lower bound. The trine denominator obeys $2 + 2c^2 - cr_h \ge 15/8$, so its error is at most $(16/15)(t^2+\eta_h)$ and its probability is at least $5/192$.

For the second construction put $\bar{s} = \sqrt{D+\eta_h}$ and $\nu = D/\bar{s} \le 1$. Use $\widehat{T}_\nu$ from Eq. (F.2) and the source in Eq. (F.10). The contraction and probability floor were proved in Sec. F.

Rationalizing the positive loss gives

$$\bar{Q} = \frac{D[D + (3/2)D\nu + \nu^2 + (9/16)D\nu^2]}{D + (2+3D)\nu/4 + g_\nu} + \eta_h g_\nu. \tag{F.19}$$

Let $A_\nu = D + (2+3D)\nu/4$. Since $A_\nu + g_\nu \ge \nu/2$ and $g_\nu \le D + \nu/2$, the six terms obey

$$\bar{Q} \le 2D^2/\nu + 3D^2 + 2D\nu + (9/8)D^2\nu + \eta_h D + \eta_h\nu/2.$$

The bound on g_ν follows by squaring: the difference of squares is $D^2 + D\nu^2/4 \ge 0$. With $\nu = D/\bar{s}$, the terms are at most $2D\bar{s}$, $3D\bar{s}$, $2D\bar{s}$, $(9/8)D\bar{s}$, $D\bar{s}$ and $D\bar{s}/2$, respectively; here $D, \eta_h \le 1$ and $D, \eta_h \le \bar{s}^2$. Thus $\bar{Q} \le (77/8)D\bar{s}$. Equation (F.9) gives $\bar{Q}/\bar{P} \le (231/32)D\sqrt{D+\eta_h} < 21t\sqrt{t+\eta_h}$. Selecting between this construction and the trine proves the stated constant 32, with probability at least $1/192$. □

Proposition F.3 (A sufficient asymptotic probability range). For every fixed $0 < p_0 < 1/64$, the constructive upper bound in Theorem F.4 holds in a sufficiently small neighborhood of $(t, \eta_h) = (0, 0)$.

Proof. Retain the choice $\nu = D/\sqrt{D+\eta_h}$ of Proposition F.2. Then $\nu \le \sqrt{D}$ and, for the whole Bloch ball, $0 \le 4 - 3c^2 - cz \le 5$. Thus

$$|\bar{P} - 2| \le \tfrac{5}{2}\sqrt{D} + 2D^{3/4}, \tag{F.20}$$

$$p_\nu \ge \frac{1 - D^{3/4}}{64(1+\sqrt{D})} \longrightarrow \frac{1}{64}. \tag{F.21}$$

The estimates are uniform in $0 \le \eta_h \le 1$ and in source direction, so some $D_0(p_0) > 0$ ensures $p_\nu \ge p_0$. The trine has $p \ge 5/192 > 1/64$ for all $0 < c < 1$ and $0 \le r_h \le 1$. Both branches of the upper envelope are therefore feasible near the origin. The LU converse in Sec. F.3.2 holds at every fixed positive floor. The finite-branch comparison and fixed-K argument in Sec. F.3.3 preserve the asserted range. □

F.3 Shared-amplitude converse at nonvanishing probability

Use $E_{p_0,\mathcal{R}}$ from Eq. (29) and $m(t,\eta_h) = \min\{t^2 + \eta_h, t\sqrt{t+\eta_h}\}$. The following theorem makes the joint-limit comparison precise.

Theorem F.4 (Purity law at fixed positive probability). Fix $0 < p_0 < 1/64$. There exist $t_0, \eta_{h0}, C_1, C_2 > 0$, depending only on p_0, such that for $0 < t < t_0$ and $0 \le \eta_h < \eta_{h0}$, simultaneously for $\mathcal{R} \in \{\mathrm{LU}, \mathrm{LOCC}, \mathrm{SEP}\}$,

$$C_1 m(t,\eta_h) \le E_{p_0,\mathcal{R}}(1-t, 1-\eta_h) \le C_2 m(t,\eta_h). \tag{F.22}$$

The upper bound uses one detection pattern in at most eight modes. For power-law impurities the exponent is Eq. (33).

The strict range $p_0 < 1/64$ is supplied by Proposition F.3. The lower comparison constants come from compactness; the result is uniform over joint paths of arbitrary relative scale. Proposition F.2 also gives an explicit construction with $p \ge 1/192$ throughout $0 < c < 1$, $0 \le r_h \le 1$.

The proof has four parts. A spectral bound (Sec. E.2 and Theorem E.2) gives the purity cost and the quadratic-error criterion. Two explicit networks (Secs. F–F.1 and Proposition F.2) give the two terms of m. The matching converse uses a second-order constraint on shared optical amplitudes (Sec. F.3.1), and a probability-preserving branch realization extends it to deterministic recovery (Sec. F.3.3). The same finite-branch argument gives the fixed-cardinality accepted-outcome result in Corollary F.1.

F.3.1 Common-source responses and optical compatibility

The converse follows from requiring the same amplitudes to serve every photon assignment. We first compare the two auxiliary eigenmodes, then prove the second-order optical obstruction and finally the matching LU lower bound. The principal mode must produce a nearly Bell state while the weak mode contributes little loss. Their common optical amplitudes constrain these two responses jointly.

Fix one instrument, one detection pattern, and one Bell target Φ. In the midpoint basis, let ψ_+ and ψ_- be the auxiliary eigenvectors with eigenvalues $(1+r_h)/2$ and $(1-r_h)/2$, respectively, and set

$$J_\pm = J(T, |\psi_\pm\rangle\langle\psi_\pm|), \qquad p_\pm = \operatorname{Tr} J_\pm,$$
$$Q_\pm = \operatorname{Tr}[(I - |\Phi\rangle\langle\Phi|)J_\pm], \quad r_{\rm pr} = Q_+/p_+, \quad r_{\rm wk} = Q_-/p_+, \quad q_{\rm resp} = p_-/p_+. \tag{F.23}$$

Here $p_+ > 0$. For a common deterministic recovery, apply it to both $J_\pm$ before evaluating $Q_\pm$. Source effects are the two-dimensional operators M_p, M_Φ satisfying $p(H) = \operatorname{Tr}(M_p H)$ and $w_\Phi(H) = \operatorname{Tr}(M_\Phi H)$; $M_\perp = M_p - M_\Phi \succeq 0$ is the Bell-loss effect. For $H = (1-w)|\psi_+\rangle\langle\psi_+| + w|\psi_-\rangle\langle\psi_-|$, $w = \eta_h/2$, linearity gives

$$e_\Phi(H) = \frac{(1-w)r_{\rm pr} + w r_{\rm wk}}{1 - w + w q_{\rm resp}}. \tag{F.24}$$

Here $r_{\rm pr}$ is the principal-mode Bell error, $r_{\rm wk}$ is the weak mode's Bell loss normalized by the principal-mode event probability, and $q_{\rm resp}$ is the relative event response. All use the same target and outcome.

Varying only an auxiliary mixture weight, while keeping these effects fixed, also yields a source-response product law:

$$\begin{aligned} r_{\rm pr} < t/8 \implies & r_{\rm pr} r_{\rm wk} \geq t^2/64 \\ r_{\rm pr} = o(D) \implies & \liminf \frac{r_{\rm pr} r_{\rm wk}}{D^2} \geq \frac{1}{9}. \end{aligned} \tag{F.25}$$

Indeed, write $u = w/(1-w) \in [0,1]$. Proposition F.2 applied to these auxiliary density eigenvalues gives $r_{\rm pr} + u r_{\rm wk} \geq (t/4)\sqrt{u}$. First use $u = 1$, then $u = (t/(8r_{\rm wk}))^2$, to prove the finite inequality. For the second statement $r_{\rm wk}/D \to \infty$; use $u = [D/((2 + c^2)r_{\rm wk})]^2 \to 0$ in the exact purity term of Eq. (E.13). It gives $r_{\rm pr} r_{\rm wk} \geq D^2(1-u)/[(2+c^2)^2(1+u)]$. The trine attains the limiting constant $1/9$.

The additional physical constraint at nonvanishing probability is the following. Its constants are existence constants obtained by compactness.

Theorem F.5 (Weak-response optical compatibility). For every $p_* > 0$ there are constants $D_*, r_{\rm wk,*}, K_* > 0$ such that, for any finite passive network with complex amplitudes, one detection pattern, a common LU Bell target and $\operatorname{supp} H \subseteq \mathcal{S}$, if $0 < D < D_*$, $p_+ \geq p_*$ and $0 < r_{\rm wk} < r_{\rm wk,*}$, then

$$r_{\rm pr} \geq K_* D r_{\rm wk}. \tag{F.26}$$

We prove this from the shared amplitudes, including blocks in which some auxiliary amplitudes v_i vanish. Write row i of a collision-free active block as $(\alpha_i^T, \beta_i^T, v_i)$. For distinct i, j, k, define two rank-one matrices and an auxiliary-placement matrix

$$\mathcal{X}_i = \alpha_i (v_j \beta_k + v_k \beta_j)^T, \qquad \mathcal{Y}_i = (v_j \alpha_k + v_k \alpha_j)\beta_i^T, \qquad \mathcal{Z}_i = v_i(\alpha_j \beta_k^T + \alpha_k \beta_j^T). \tag{F.27}$$

Here $\mathcal{Z}_i$ collects the amplitudes for detecting the auxiliary at output i. The matrices are $\mathcal{X}_i = \operatorname{mat}(x_i) = 2X_i$ and $\mathcal{Y}_i = \operatorname{mat}(y_i) = 2Y_i$ in the order used throughout; the factor two removes the Bell-input amplitude $1/2$. They obey $\sum \mathcal{X}_i = \sum \mathcal{Y}_i = \sum \mathcal{Z}_i = \mathcal{C}$ and $\mathcal{X}_i + \mathcal{Y}_i + \mathcal{Z}_i = \mathcal{C}$. Put $\mathcal{V}_i = \mathcal{X}_i - \mathcal{Y}_i$, $\xi_c = \sqrt{(1-c)/(1+c)}$ and $\psi_+ = (1, \kappa)^T/\sqrt{1 + |\kappa|^2}$; its orthogonal vector is $\psi_- = (-\bar\kappa, 1)^T/\sqrt{1+|\kappa|^2}$. This chart assumes a nonzero h_+ component of ψ_+; the limiting-source argument in the proof below ensures that condition for the sequences used. Let $B_\Phi = \operatorname{mat}(\Phi)$, with $\|B_\Phi\|_F = 1$, be the common Bell matrix. Define its Frobenius-orthogonal projection by $\Pi M = M - B_\Phi \langle B_\Phi, M\rangle_F$, and write $\mathcal{C}^\perp = \Pi\mathcal{C}$, $\mathcal{Z}_i^\perp = \Pi\mathcal{Z}_i$, $\mathcal{V}_i^\perp = \Pi\mathcal{V}_i$. Expanding the amplitudes of the eight internal basis states gives, with $g = [4(1+\xi_c^2)^2(1+|\kappa|^2)]^{-1}$,

$$Q_+/g = (1 + |\kappa|^2\xi_c^4)\|\mathcal{C}^\perp\|^2 + \sum_i \|\kappa\mathcal{Z}_i^\perp + \xi_c \mathcal{V}_i^\perp\|^2 + \xi_c^2 \sum_i \|\xi_c \mathcal{Z}_i^\perp + \kappa \mathcal{V}_i^\perp\|^2, \tag{F.28}$$

$$Q_-/g = (|\kappa|^2 + \xi_c^4)\|\mathcal{C}^\perp\|^2 + \sum_i \|\mathcal{Z}_i^\perp - \bar\kappa\xi_c \mathcal{V}_i^\perp\|^2 + \xi_c^2 \sum_i \|\mathcal{V}_i^\perp - \bar\kappa\xi_c \mathcal{Z}_i^\perp\|^2. \tag{F.29}$$

The norms are Frobenius norms. These formulas include arbitrary relative complex phases and are sums of positive losses.

The key obstruction is second order. A limit with nonvanishing probability and small loss must make the three auxiliary placements proportional to the same Bell matrix. We choose bounded coordinates around one nonzero placement, then use the shared optical products to relate four residual entries. Their first-order requirements contradict their second-order product constraint, including the boundaries where other auxiliary placements vanish.

Lemma F.6 (Second-order shared-amplitude obstruction). Let a sequence of bounded collision-free active blocks have $\mathcal{C}$ converging to an invertible matrix and each $\mathcal{Z}_i$ converging to a multiple of that limit. If $0 < |\zeta| \to 0$ along the sequence, then not all $\mathcal{W}_i = \mathcal{Z}_i + \zeta\mathcal{V}_i$ lie within $o(|\zeta|^2)$ of scalar multiples of $\mathcal{C}$.

Proof. Pass to a convergent subsequence and select a nonzero limiting $\mathcal{Z}_2$. Set $\mathbf{A} = [\alpha_0\ \alpha_1]$, $\mathbf{B} = [\beta_0\ \beta_1]$

and $S_{\text{swap}} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. The exact identity

$$\mathcal{Z}_2 = v_2 \mathbf{A} S_{\text{swap}} \mathbf{B}^T, \qquad \det \mathcal{Z}_2 = -v_2^2 \det \mathbf{A} \det \mathbf{B} \tag{F.30}$$

shows that v_2, $\det \mathbf{A}$ and $\det \mathbf{B}$ have nonzero limits: their product stays nonzero and all factors are bounded. Thus the changes

$$\alpha_i' = \mathbf{A}^{-1}\alpha_i, \quad \beta_i' = S_{\text{swap}}\mathbf{B}^{-1}\beta_i, \quad v_i' = v_i/v_2$$

and their inverses are bounded. Every assignment matrix undergoes the same map

$$M \mapsto v_2^{-1}\mathbf{A}^{-1} M (S_{\text{swap}}\mathbf{B}^{-1})^T,$$

so its small-residual order is preserved. Dropping the primes gives

$$\alpha_0 = (1,0)^T, \quad \alpha_1 = (0,1)^T, \quad \beta_0 = (0,1)^T, \quad \beta_1 = (1,0)^T,$$
$$\alpha_2 = (\mathfrak{a}_1, \mathfrak{a}_2)^T, \quad \beta_2 = (\mathfrak{b}_1, \mathfrak{b}_2)^T, \quad (v_0, v_1, v_2) = (\mu_0, \mu_1, 1).$$

The bounded coordinate changes preserve residual orders; probabilities are evaluated from the physical contraction. The chart gives exactly

$$\mathcal{C} = \begin{pmatrix} 1 + \mu_0\mathfrak{a}_1 + \mu_1\mathfrak{b}_1 & \mu_1(\mathfrak{a}_1 + \mathfrak{b}_2) \\ \mu_0(\mathfrak{a}_2 + \mathfrak{b}_1) & 1 + \mu_0\mathfrak{b}_2 + \mu_1\mathfrak{a}_2 \end{pmatrix}, \qquad \mathcal{Z}_2 = I. \tag{F.31}$$

Its limit is $\kappa_0 I$, $\kappa_0 \neq 0$. The scalar-matrix conditions on $\mathcal{Z}_0, \mathcal{Z}_1$ give all four cases below; every entry in this table denotes its limiting value:

(μ_0, μ_1)	data-coordinate constraints	$(1 + \mu_1\mathfrak{b}_1, 1 + \mu_0\mathfrak{b}_2)$
$\mu_0 \neq 0,\ \mu_1 \neq 0$	$\mathfrak{a}_1 = \mathfrak{a}_2 = \mathfrak{b}_1 = \mathfrak{b}_2 = 0$	$(1,1)$
$\mu_0 = 0,\ \mu_1 \neq 0$	$\mathfrak{b}_1 = \mathfrak{a}_2,\ \mathfrak{b}_2 = -\mathfrak{a}_1$	$(\kappa_0, 1)$
$\mu_0 \neq 0,\ \mu_1 = 0$	$\mathfrak{b}_1 = -\mathfrak{a}_2,\ \mathfrak{b}_2 = \mathfrak{a}_1$	$(1, \kappa_0)$
$\mu_0 = \mu_1 = 0$	$\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{b}_1, \mathfrak{b}_2$ arbitrary	$(1,1)$

Both factors in the last column therefore have nonzero limits, including the intersections where auxiliary amplitudes vanish.

Suppose $\mathcal{W}_i = \lambda_i \mathcal{C} + R_i$, with $\|R_i\| = o(|\zeta|^2)$. The λ_i are bounded because $\mathcal{W}_i$ is bounded and $\|\mathcal{C}\|$ is bounded below. The off-diagonal entries of $\operatorname{adj}(\mathcal{C})\mathcal{W}_i$ would be $o(|\zeta|^2)$. Direct multiplication gives

$$[\operatorname{adj}(\mathcal{C})\mathcal{W}_0]_{21} = (1-\zeta)\mu_0(\mathfrak{a}_2 + \mathfrak{b}_1)(1 + \mu_1\mathfrak{b}_1), \quad [\operatorname{adj}(\mathcal{C})\mathcal{W}_1]_{12} = (1-\zeta)\mu_1(\mathfrak{a}_1 + \mathfrak{b}_2)(1 + \mu_0\mathfrak{b}_2). \tag{F.32}$$

Hence $\mu_0(\mathfrak{a}_2 + \mathfrak{b}_1), \mu_1(\mathfrak{a}_1 + \mathfrak{b}_2) = o(|\zeta|^2)$, so $\mathcal{C}$ is diagonal to that order. The remaining entries $(\mathcal{W}_0)_{12} = \zeta\mu_1(\mathfrak{b}_2 - \mathfrak{a}_1)$ and $(\mathcal{W}_1)_{21} = \zeta\mu_0(\mathfrak{b}_1 - \mathfrak{a}_2)$ imply $\mu_1\mathfrak{a}_1, \mu_1\mathfrak{b}_2, \mu_0\mathfrak{a}_2, \mu_0\mathfrak{b}_1 = o(|\zeta|)$. Set $u_1 = \mu_0\mathfrak{a}_1$, $u_2 = \mu_1\mathfrak{b}_1$, $u_3 = \mu_0\mathfrak{b}_2$, $u_4 = \mu_1\mathfrak{a}_2$. The shared products then give

$$u_1u_2,\ u_1u_4,\ u_3u_2,\ u_3u_4 = o(|\zeta|^2). \tag{F.33}$$

This polynomial step includes $\mu_0 = 0$ and $\mu_1 = 0$.

Put $A_d = 1 + u_1 + u_2$, $B_d = 1 + u_3 + u_4$. The two diagonal residuals are

$$u_1 - u_3 + u_1u_4 - u_3u_2 + \zeta[(u_2+1)B_d + (u_4+1)A_d] = o(|\zeta|^2), \tag{F.34}$$
$$u_2 - u_4 + u_2u_3 - u_4u_1 - \zeta[(u_1+1)B_d + (u_3+1)A_d] = o(|\zeta|^2). \tag{F.35}$$

All four variables must tend to zero. For example, if u_1 or u_3 has a nonzero limit, Eq. (F.33) forces $u_2, u_4 = o(|\zeta|^2)$; $\mathcal{C} \to \kappa_0 I$ gives $u_1, u_3 \to \kappa_0 - 1$. The second equation would then have leading term $-2\zeta\kappa_0^2$, a contradiction. The opposite case uses the first equation. Consequently the equations give $(u_1 - u_3)/\zeta \to -2$ and $(u_2 - u_4)/\zeta \to 2$. Their product tends to -4, whereas $(u_1 - u_3)(u_2 - u_4) = u_1u_2 - u_1u_4 - u_3u_2 + u_3u_4 = o(|\zeta|^2)$. This contradiction proves the lemma, including every boundary where an auxiliary amplitude vanishes. □

Proof of Theorem F.5. If no stated constants existed at a fixed $p_* > 0$, we could choose a sequence with $D < 1/n$, $0 < r_{\text{wk}} < 1/n$, $p_+ \geq p_*$ and $r_{\text{pr}}/(Dr_{\text{wk}}) < 1/n$. Bunched outcomes have error of order at least D by Eq. (A.15), so this sequence is collision-free. Equation (F.25) gives $r_{\text{wk}}^2/D \to \infty$. The pure-source spectral envelope implies that its principal source mode approaches the positive midpoint,

hence $\kappa \to 0$: if its Bloch z_+ stayed away from 1, the envelope would give a positive multiple of t as its error.

Contractions and Bell targets are compact. On a convergent subsequence, both pure losses tend to zero and $p_+ \geq p_*$ leaves a nonzero Bell matrix $\mathcal{C}$. At $c = 1$, the auxiliary h_- contributes three mutually orthogonal internal states, one for each auxiliary output; their nonnegative losses force $\mathcal{Z}_i \parallel \mathcal{C}$ separately. Moreover the limiting triple $(\mathcal{V}_i^\perp)$ is not zero. Otherwise both $\mathcal{Z}_i$ and $\mathcal{V}_i$ would be parallel to $\mathcal{C}$; from $\mathcal{X}_i + \mathcal{Y}_i = \mathcal{C} - \mathcal{Z}_i$, rank one would force $\mathcal{X}_i = \mathcal{Y}_i = 0$ and $\mathcal{Z}_i = \mathcal{C}$ for every i, contradicting $\sum \mathcal{Z}_i = \mathcal{C} \neq 0$.

Write norms of triples as sums in quadrature. Physical amplitudes are bounded, $p_+ \in [p_*, 1]$, and $g \to 1/4$. The plus identity gives $\|\mathcal{C}^\perp\| = O(\sqrt{r_{\mathrm{pr}}})$. In Eq. (F.29), the last term is $O(\xi_c^2) = o(r_{\mathrm{wk}})$ and the $\mathcal{C}$ term is $O(r_{\mathrm{pr}}) = o(r_{\mathrm{wk}})$. Hence

$$\|(\mathcal{Z}_i^\perp - \bar{\kappa}\xi_c \mathcal{V}_i^\perp)\|^2 = (p_+/g) r_{\mathrm{wk}} + o(r_{\mathrm{wk}})$$
$$\|\bar{\kappa}\xi_c(\mathcal{V}_i^\perp)\| = o(\sqrt{r_{\mathrm{wk}}}).$$

The triangle inequalities and p_+/g bounded above and below give $\|(\mathcal{Z}_i^\perp)\| \asymp \sqrt{r_{\mathrm{wk}}}$. The first sum in Eq. (F.28) then gives

$$\big|\,|\kappa|\|(\mathcal{Z}_i^\perp)\| - \xi_c\|(\mathcal{V}_i^\perp)\|\big| \leq \|(\kappa\mathcal{Z}_i^\perp + \xi_c\mathcal{V}_i^\perp)\| = O(\sqrt{r_{\mathrm{pr}}}) = o(\xi_c).$$

Since $(\mathcal{V}_i^\perp)$ has a nonzero limit, $\kappa \neq 0$ eventually and

$$|\kappa|\sqrt{r_{\mathrm{wk}}} \asymp \xi_c, \qquad \zeta = \xi_c/\kappa \to 0, \qquad |\zeta|^2 \asymp r_{\mathrm{wk}}. \tag{F.36}$$

Thus

$$\frac{\|\Pi(\mathcal{Z}_i + \zeta\mathcal{V}_i)\|}{|\zeta|^2} = O\left(\sqrt{\frac{r_{\mathrm{pr}}}{D r_{\mathrm{wk}}}}\right) \to 0, \qquad \frac{\|\Pi\mathcal{C}\|}{|\zeta|^2} = O(\sqrt{r_{\mathrm{pr}}}/r_{\mathrm{wk}}) \to 0.$$

The second limit uses $r_{\mathrm{pr}}/r_{\mathrm{wk}}^2 = [r_{\mathrm{pr}}/(D r_{\mathrm{wk}})]\,(D/r_{\mathrm{wk}}) \to 0$. Using the normalized Bell matrix B_Φ defined above, put $\mathcal{W}_i = \mathcal{Z}_i + \zeta\mathcal{V}_i$. Since $|\langle B_\Phi, \mathcal{C}\rangle_F|$ stays bounded below,

$$\lambda_i = \frac{\langle B_\Phi, \mathcal{W}_i\rangle_F}{\langle B_\Phi, \mathcal{C}\rangle_F}$$
$$\mathcal{W}_i - \lambda_i\mathcal{C} = \Pi\mathcal{W}_i - \lambda_i\Pi\mathcal{C} = o(|\zeta|^2). \tag{F.37}$$

Here λ_i is bounded. Lemma F.6 excludes these residuals. The active contraction always has size 3×5; compactness of that set and of Bell targets gives constants depending only on p_*, uniformly over paths and finite vacuum extensions. □

F.3.2 The matching lower bound

We now prove the LU part of Theorem F.4.

Proof of the LU part. Use the mixed-auxiliary probability constraint $p(H) = (1 - \eta_h/2)p_+ + (\eta_h/2)p_- \geq p_0$ and $p_- \leq 1$. For $0 \leq \eta_h < \min(p_0, 1)$, this implies $p_+ \geq (p_0 - \eta_h/2)/(1 - \eta_h/2) \geq p_0/2$. Therefore $q_{\mathrm{resp}} \leq 2/p_0$, and Eq. (F.24) compares the mixed error to $r_{\mathrm{pr}} + \eta_h r_{\mathrm{wk}}$ with uniform constants. If $r_{\mathrm{wk}} \geq r_{\mathrm{wk},*}$, the pure-auxiliary bound and this identity give a multiple of $t^2 + \eta_h$. If $r_{\mathrm{wk}} < r_{\mathrm{wk},*}$ and $r_{\mathrm{pr}} < t/8$, Theorem F.5 and Eq. (F.25) give, with constants depending only on p_0,

$$r_{\mathrm{pr}} + \eta_h r_{\mathrm{wk}} \geq K_1 D r_{\mathrm{wk}} + K_2 D^2/r_{\mathrm{wk}} + \eta_h r_{\mathrm{wk}}$$
$$\geq 2\sqrt{K_2}\, D\sqrt{K_1 D + \eta_h}.$$

The product bound ensures $r_{\mathrm{wk}} > 0$ in this case, and half of r_{pr} is used for each lower bound. If $r_{\mathrm{pr}} \geq t/8$, the same required lower order is immediate near the origin. For sufficiently small t, η_h, taking the smaller branch and using $D \asymp t$ proves the uniform LU lower bound. The two physical constructions in Proposition F.2, with Proposition F.3, give the upper bound at each stipulated $p_0 < 1/64$. Substitution of $\eta_h = \eta_0 t^\beta$ gives Eq. (33) for LU. □

F.3.3 Recovery and accepted sets

The optical branch identity also controls probability. This requires a finite representation of the recovery channel and contraction of its local factors. Two-qubit separable channels admit a uniformly finite product-Kraus representation. Selecting a sufficiently probable branch then compares the recovered error with an LU error at a reduced probability floor. The factor 256 comes from this selection; the optical realization of the chosen branch preserves its probability.

Proposition F.7 (Finite-branch probability transfer). Let $N_{\mathrm{K}} = 256$. For the specified source model, at any c, r_h, p_0 where the constrained problems are nonempty,

$$\frac{1}{N_{\mathrm{K}}} E_{p_0/N_{\mathrm{K}},\mathrm{LU}}(c, r_h) \leq E_{p_0,\mathrm{SEP}}(c, r_h)$$
$$\leq E_{p_0,\mathrm{LOCC}}(c, r_h)$$
$$\leq E_{p_0,\mathrm{LU}}(c, r_h). \tag{F.38}$$

The first comparison covers all complete optical occupations.

Proof. The separable-Choi and product-Kraus correspondence is standard [36]. For a two-qubit channel,

divide its 16×16 Choi matrix by its trace four. The resulting state is separable across $(R_{\rm out},R_{\rm in})\mid(S_{\rm out},S_{\rm in})$ and is a convex combination of pure product projectors. Trace-one Hermitian matrices have real affine dimension $16^2-1=255$. Carathéodory's theorem therefore gives at most 256 projectors. Reshaping their two local vectors into 2×2 matrices and absorbing the weights gives

$$\begin{aligned}\Lambda(J)&=\sum_{j=1}^{N_{\rm K}}K_jJK_j^\dagger\\ K_j&=L_j\otimes M_j\\ \sum_j K_j^\dagger K_j&=I_4\end{aligned}$$

with zero terms added if needed. This represents the same channel. It applies also to LOCC channels specified by an infinite outcome tree, since their Choi states lie in the closed finite-dimensional separable set.

For each nonzero term, trace preservation implies $\|L_j\|_{\rm op}\|M_j\|_{\rm op}=\|K_j\|_{\rm op}\le1$. With $\ell_{A,j}=\|L_j\|_{\rm op}$ and $\ell_{B,j}=\|M_j\|_{\rm op}$, replace

$$L_j\longmapsto\sqrt{\ell_{B,j}/\ell_{A,j}}\,L_j,\qquad M_j\longmapsto\sqrt{\ell_{A,j}/\ell_{B,j}}\,M_j. \tag{F.39}$$

Both new norms equal $\sqrt{\ell_{A,j}\ell_{B,j}}\le1$ and K_j is unchanged.

Let T contain each distinct recorded physical output row once, and the five potentially populated input columns. It is an $s\times5$ contraction, $1\le s\le3$. For the balanced factors,

$$\begin{aligned}T_j&=T\ \mathrm{diag}(L_j^T,M_j^T,1)\\ J_{\boldsymbol n}(T_j,H)&=K_jJ_{\boldsymbol n}(T,H)K_j^\dagger.\end{aligned} \tag{F.40}$$

The right multiplier is a contraction, so T_j already is one. The transpose follows from the reference-rail Bell inputs: a data column change sends each logical matrix X to $L_jXM_j^T$. For repeated occupation, the same identity acts on every slot while the factor $1/\sqrt{\prod_i n_i!}$ is unchanged. The repeated-slot matrix is used only to evaluate amplitudes; contractivity is imposed on the distinct-row matrix T_j. Its rectangular Julia completion, Eq. (A.2) with $k=s$, is a finite unitary on $5+s\le8$ modes. Accept the same occupations on those s rows and zero counts in all added outputs. Thus Eq. (F.40) realizes the branch probability without additional attenuation.

Now use the mixed-auxiliary outcome state $J=J_{\boldsymbol n}(T,H)$ and one common Bell target Φ for $\Lambda(J)$. Put

$$\begin{aligned}p_j&=\mathrm{Tr}(K_jJK_j^\dagger)\\ Q_j&=\mathrm{Tr}[(I-|\Phi\rangle\langle\Phi|)K_jJK_j^\dagger]\ge0.\end{aligned}$$

Then $\sum_j p_j=p\ge p_0$ and $\sum_j Q_j=pe_\Phi$. Choosing a largest-probability branch gives

$$p_j\ge p_0/N_{\rm K},\qquad Q_j/p_j\le N_{\rm K}e_\Phi. \tag{F.41}$$

This branch is chosen for the full density H, so it is common to the unlabeled eigencomponents of H. Its optical realization has the same auxiliary density and eigenvalues. Optimizing its LU target can only lower its error, so $E_{p_0/N_{\rm K},{\rm LU}}\le N_{\rm K}e_\Phi$. Taking infima, using approximate optimizers when necessary, proves the first inequality. The others follow from inclusion of the recovery classes. □

Completion of Theorem F.4. The LU lower bound established above applies at the fixed positive floor $p_0/N_{\rm K}$, and the LU constructions attain the upper order at p_0. Proposition F.7 gives the same order for LOCC and SEP. Their common lower constant can be taken from SEP and their common upper constant from LU. This proves the full two-parameter statement. □

Corollary F.1 (A bounded number of accepted outcomes). At each parameter pair (t,η_h), accept at most $K\ge1$ detection patterns from one common instrument (U,H). The network and source may vary with (t,η_h), while $K\ge1$ is independent of them. Use weights $0\le\omega_{\boldsymbol n}\le1$ and outcome-conditioned deterministic SEP recovery. Let $e_{\boldsymbol n}$ be the recovered Bell errors, aligned to a common target by conditional local unitaries, and define

$$P_{\rm acc}=\sum_{\boldsymbol n}\omega_{\boldsymbol n}p_{\boldsymbol n},\qquad E_{\mathcal A}=\frac{\sum_{\boldsymbol n}\omega_{\boldsymbol n}p_{\boldsymbol n}e_{\boldsymbol n}}{P_{\rm acc}}. \tag{F.42}$$

If $P_{\rm acc}\ge p_0>0$, then

$$E_{\mathcal A}\ge\frac{1}{N_{\rm K}K}E_{p_0/(N_{\rm K}K),{\rm LU}}(c,r_h). \tag{F.43}$$

Consequently, for fixed K and $0<p_0<1/64$, optimizing this accepted average error over all allowed networks has order $m(t,\eta_h)$ for each of LU, deterministic LOCC and deterministic SEP. The constants may depend on K,p_0. The same conclusion holds with a fixed mode bound $m_{\max}\ge8$, taking $K=\binom{m_{\max}+2}{3}$.

Proof. Choose an outcome of largest accepted weight $\omega_{\boldsymbol n}p_{\boldsymbol n}$. It satisfies $\omega_{\boldsymbol n}p_{\boldsymbol n}\ge P_{\rm acc}/K$, $p_{\boldsymbol n}\ge p_0/K$ and $e_{\boldsymbol n}\le KE_{\mathcal A}$. Apply the preceding branch argument to that mixed-auxiliary outcome. The resulting LU outcome has probability at least $p_0/(N_{\rm K}K)$ and error at most $N_{\rm K}KE_{\mathcal A}$. Dropping the acceptance weight preserves its state and increases its raw probability. This proves Eq. (F.43). The one-outcome constructions give the reverse order. There are $\binom{m_{\max}+2}{3}$ complete three-photon occupations in $m_{\max}$ modes, and the eight-mode construction is available within the stated bound. □

References

[1] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller. Quantum repeaters: The role of imperfect local operations in quantum communication. *Physical Review Letters*, 81:5932–5935, 1998. doi: 10.1103/PhysRevLett.81.5932.

[2] Nicolas Sangouard, Christoph Simon, Hugues de Riedmatten, and Nicolas Gisin. Quantum repeaters based on atomic ensembles and linear optics. *Reviews of Modern Physics*, 83:33, 2011. doi: 10.1103/RevModPhys.83.33.

[3] Koji Azuma, Sophia E. Economou, David Elkouss, Paul Hilaire, Liang Jiang, Hoi-Kwong Lo, and Ilan Tzitrin. Quantum repeaters: From quantum networks to the quantum internet. *Reviews of Modern Physics*, 95:045006, 2023. doi: 10.1103/RevModPhys.95.045006.

[4] Hans K. C. Beukers, Matteo Pasini, Hyeongrak Choi, Dirk Englund, Ronald Hanson, and Johannes Borregaard. Remote-entanglement protocols for stationary qubits with photonic interfaces. *PRX Quantum*, 5:010202, 2024. doi: 10.1103/PRXQuantum.5.010202.

[5] Jeffrey Marshall. Distillation of indistinguishable photons. *Physical Review Letters*, 129:213601, 2022. doi: 10.1103/PhysRevLett.129.213601.

[6] Carlos F. D. Faurby, Lorenzo Carosini, Huan Cao, Patrik I. Sund, Lena M. Hansen, Francesco Giorgino, Andrew B. Villadsen, Stefan N. van den Hoven, Peter Lodahl, Stefano Paesani, Juan C. Loredo, and Philip Walther. Purifying photon indistinguishability through quantum interference. *Physical Review Letters*, 133:033604, 2024. doi: 10.1103/PhysRevLett.133.033604.

[7] Jason Saied, Jeffrey Marshall, Namit Anand, and Eleanor G. Rieffel. General protocols for the efficient distillation of indistinguishable photons. *Physical Review Applied*, 23:034079, 2025. doi: 10.1103/PhysRevApplied.23.034079.

[8] F. H. B. Somhorst, B. K. Sauër, S. N. van den Hoven, and J. J. Renema. Photon-distillation schemes with reduced resource costs based on multiphoton Fourier interference. *Physical Review Applied*, 23:044003, 2025. doi: 10.1103/PhysRevApplied.23.044003.

[9] F. H. B. Somhorst, J. Saied, N. Kannan, B. Kassenberg, J. Marshall, M. de Goede, H. J. Snijders, P. Stremoukhov, A. Lukianenko, P. Venderbosch, T. B. Demille, A. Roos, N. Walk, J. Eisert, E. G. Rieffel, D. H. Smith, and J. J. Renema. Below-threshold error reduction in single photons through photon distillation, 2026. arXiv:2601.05947.

[10] Francesco Hoch, Anita Camillini, Giovanni Rodari, Eugenio Caruccio, Gonzalo Carvacho, Taira Giordani, Riccardo Albiero, Niki Di Giano, Giacomo Corrielli, Francesco Ceccarelli, Roberto Osellame, Marco Robbio, Leonardo Novo, Nicolò Spagnolo, Ernesto F. Galvão, and Fabio Sciarrino. Optimal distillation of photonic indistinguishability. *Physical Review Research*, 2026. doi: 10.1103/lyw5-qdfx. Accepted 9 October 2026.

[11] S. N. van den Hoven, G. B. Lamers, and J. J. Renema. Photonic fusion operations transform partial distinguishability, 2026. arXiv:2609.01019.

[12] Liang Zhai, Giang N. Nguyen, Clemens Spinnler, Julian Ritzmann, Matthias C. Löbl, Andreas D. Wieck, Arne Ludwig, Alisa Javadi, and Richard J. Warburton. Quantum interference of identical photons from remote GaAs quantum dots. *Nature Nanotechnology*, 17:829–833, 2022. doi: 10.1038/s41565-022-01131-2.

[13] N. Somaschi, V. Giesz, L. De Santis, J. C. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Antón, J. Demory, C. Gómez, I. Sagnes, N. D. Lanzillotti-Kimura, A. Lemaître, A. Auffèves, A. G. White, L. Lanco, and P. Senellart. Near-optimal single-photon sources in the solid state. *Nature Photonics*, 10:340–345, 2016. doi: 10.1038/nphoton.2016.23.

[14] John Calsamiglia and Norbert Lütkenhaus. Maximum efficiency of a linear-optical Bell-state analyzer. *Applied Physics B*, 72:67–71, 2001. doi: 10.1007/s003400000484.

[15] Warren P. Grice. Arbitrarily complete Bell-state measurement using only linear optical elements. *Physical Review A*, 84:042331, 2011. doi: 10.1103/PhysRevA.84.042331.

[16] Fabian Ewert and Peter van Loock. 3/4-efficient Bell measurement with passive linear optics and unentangled ancillae. *Physical Review Letters*, 113:140403, 2014. doi: 10.1103/PhysRevLett.113.140403.

[17] Andrea Olivo and Frédéric Grosshans. Ancilla-assisted linear optical Bell measurements and their optimality. *Physical Review A*, 98:042323, 2018. doi: 10.1103/PhysRevA.98.042323.

[18] Matthias J. Bayerbach, Simone E. D'Aurelio, Peter van Loock, and Stefanie Barz. Bell-state measurement exceeding 50% success probability with linear optics. *Science Advances*, 9(32):eadf4080, 2023. doi: 10.1126/sciadv.adf4080.

[19] Matthias J. Bayerbach, Simone E. D'Aurelio, and Stefanie Barz. Boosted entanglement swapping, 2026. arXiv:2611.05497.

[20] V. S. Shchesnovich. Partial indistinguishability theory for multiphoton experiments in multiport devices. *Physical Review A*, 91:013844, 2015. doi: 10.1103/PhysRevA.91.013844.

[21] Malte C. Tichy. Sampling of partially distinguishable bosons and the relation to the multidimensional permanent. *Physical Review A*, 91:022316, 2015. doi: 10.1103/PhysRevA.91.022316.

[22] Adrian J. Menssen, Alex E. Jones, Benjamin J. Metcalf, Malte C. Tichy, Stefanie Barz, W. Steven Kolthammer, and Ian A. Walmsley. Distinguishability and Many-Particle Interference. *Physical Review Letters*, 118:153603, 2017. doi: 10.1103/PhysRevLett.118.153603.

[23] See the Supplemental Material accompanying this article.

[24] Frank Verstraete and Henri Verschelde. Optimal teleportation with a mixed state of two qubits. *Physical Review Letters*, 90:097901, 2003. doi: 10.1103/PhysRevLett.90.097901.

[25] P. L. Robinson. Julia operators and Halmos dilations, 2018. arXiv:1803.09329.

[26] Nikolai Kiesel, Christian Schmid, Ulrich Weber, Rupert Ursin, and Harald Weinfurter. Linear optics controlled-phase gate made simple. *Physical Review Letters*, 95:210505, 2005. doi: 10.1103/PhysRevLett.95.210505.

[27] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters*, 59:2044, 1987. doi: 10.1103/PhysRevLett.59.2044.

[28] Falk Töppel, Andrea Aiello, and Gerd Leuchs. All photons are equal but some photons are more equal than others. *New Journal of Physics*, 14:093051, 2012. doi: 10.1088/1367-2630/14/9/093051.

[29] Daniel J. Brod, Ernesto F. Galvão, Niko Viggianiello, Fulvio Flamini, Nicolò Spagnolo, and Fabio Sciarrino. Witnessing genuine multiphoton indistinguishability. *Physical Review Letters*, 122:063602, 2019. doi: 10.1103/PhysRevLett.122.063602.

[30] Monika Patel, Joseph B. Altepeter, Yu-Ping Huang, Neal N. Oza, and Prem Kumar. Erasing quantum distinguishability via single-mode filtering. *Physical Review A*, 86:033809, 2012. doi: 10.1103/PhysRevA.86.033809.

[31] Monika Patel, Joseph B. Altepeter, Yu-Ping Huang, Neal N. Oza, and Prem Kumar. Erratum: Erasing quantum distinguishability via single-mode filtering [Phys. Rev. A 86, 033809 (2012)]. *Physical Review A*, 86:069903, 2012. doi: 10.1103/PhysRevA.86.069903.

[32] B. Brecht, Dileep V. Reddy, C. Silberhorn, and M. G. Raymer. Photon temporal modes: A complete framework for quantum information science. *Physical Review X*, 5:041017, 2015. doi: 10.1103/PhysRevX.5.041017.

[33] Annan Chen. Numerical data for "Optimal entanglement swapping with one auxiliary photon: exact fidelity limit and purity requirements". Zenodo, 2026. Data set. https://doi.org/10.5281/zenodo.24141275.

[34] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Physical Review Letters*, 80:2245–2248, 1998. doi: 10.1103/PhysRevLett.80.2245.

[35] Guifré Vidal. Entanglement monotones. *Journal of Modern Optics*, 47:355–376, 2000. doi: 10.1080/09500340008244048.

[36] Scott M. Cohen. Conditions for uniqueness of product representations for separable quantum channels and separable quantum states. *Journal of Mathematical Physics*, 55:062202, 2014. doi: 10.1063/1.4883400.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.pdf