

# Conditional mutual information under independent endpoint channels: exact finite-alphabet reachability

scientific-author-46

September 7, 2026

## Abstract

Two independent channels acting on the two endpoint systems of a tripartite state produce four conditional mutual information values: the one of the input state, the two obtained by acting on a single endpoint, and the one obtained by acting on both. Which quadruples of values occur? The closure of the reachable set was shown earlier to be a polyhedral cone with four extreme rays, but the argument passed through uniformly distributed bits and rational approximation, so it left open whether every point of the cone is reached exactly by a finite system. We settle that question in both the classical and the quantum setting: the reachable set is the cone itself, with no closure, limiting or unbounded-copy argument. The step that achieves exactness is to replace the uniform bit by a biased one, whose entropy takes every value between zero and the logarithm of two, and to act at the two endpoints of each block by either the identity or the erasure map; four blocks, one per extreme ray, realize any prescribed quadruple exactly. Within this block family we determine the least number of product bits, and we show that no realization whatsoever can use an input alphabet smaller than the exponential of the input conditional mutual information, which places the block construction within a factor of sixteen of a bound that holds universally. The block count is nevertheless not minimal over all realizations: one shared input bit and two independent binary symmetric channels realize a quadruple for which the block family needs four bits, and a single bit is the least any realization of that quadruple can use. Every identity and inequality used is recomputed from full joint distributions, at eighty decimal digits and then again by an independent implementation at sixty, on the constructions themselves and on random classical and random quantum inputs.

## 1 Introduction

Take three systems in a joint state, and let two channels act on them independently: one on the first system, one on the third, and nothing on the second. The conditional mutual information of the two endpoint systems given the middle one can then be read off at four stages — before either channel has acted, after the first channel alone, after the second channel alone, and after both. A state together with a pair of channels therefore determines four nonnegative numbers,

$$(x, y, z, w) = (I(A:C | B), I(A':C | B), I(A:C' | B), I(A':C' | B)), \quad (1)$$

where  $A'$  and  $C'$  are the outputs of the two channels. This paper answers the question of which quadruples occur.

Three features of conditional mutual information already restrict the answer strongly. It is nonnegative, by strong subadditivity of the entropy [2]. It cannot increase when one of the two

systems is processed by a channel, so each of the three processed values is at most the input value  $x$ , and the doubly processed value  $w$  is at most each of the two singly processed ones. Finally, the two reductions cannot add up to more than the total reduction: writing the losses as  $x - y$  and  $x - z$  for the single channels and  $x - w$  for both together, the inequality

$$x - w \leq (x - y) + (x - z) \quad \text{equivalently} \quad w \geq y + z - x \quad (2)$$

says that the destruction of correlation by two independent channels is subadditive. The slack in (2),

$$D = x - y - z + w, \quad (3)$$

is the part of the loss that both channels produce in common, and it has an exact identification rather than merely a sign: in the classical setting  $D$  is the conditional mutual information of the two *inputs* given the two outputs and the middle system, and in the quantum setting it is the conditional mutual information of the two channel environments given the two outputs and the middle system (Proposition 7). Correlation is destroyed redundantly, and  $D$  measures the redundancy.

These constraints cut out a closed polyhedral cone in  $\mathbb{R}^4$ ,

$$P = \{(x, y, z, w) \in \mathbb{R}^4 : x, y, z, w \geq 0, y \leq x, z \leq x, w \leq \min(y, z), w \geq y + z - x\}, \quad (4)$$

and the reachable quadruples lie inside it. What is less obvious is that they fill it. The closure of the reachable set was shown to be exactly  $P$  in [5], where the four extreme rays

$$e_0 = (1, 0, 0, 0), \quad e_1 = (1, 1, 0, 0), \quad e_2 = (1, 0, 1, 0), \quad e_3 = (1, 1, 1, 1) \quad (5)$$

were identified and realized. That argument ran through a single *uniform* bit, whose entropy is the fixed value  $\ln 2$ , together with rational combinations of copies of it, so an arbitrary point of the cone was obtained as a limit of reachable points rather than as a reachable point. The additive bound (2) itself, and with it the nonnegativity of the slack  $D$ , is due to [4].

The distinction between a dense subset and the whole set is not a technicality here. A statement about the closure says that a prescribed pattern of correlation loss can be approached; it does not say that any finite system exhibits it, and it leaves the size of the system needed to come within a given distance completely open. Statements about rates and about single-shot tasks are statements about finite systems. The first result of this paper is that no gap remains.

**Theorem 1** (informal form of Theorem 16). *For finite-dimensional systems, the set of quadruples (1) realized by classical states and independent classical channels, and the set realized by quantum states and independent completely positive trace-preserving maps, are both equal to the cone  $P$  itself. Every point of  $P$  is realized exactly, by an explicit construction on finitely many product bits, with no limiting procedure and no unbounded number of copies.*

The step that removes the closure is small and, once seen, unavoidable: replace the uniform bit by a *biased* one. The entropy of a bit with bias  $q$  takes every value between 0 and  $\ln 2$  as  $q$  ranges over  $[0, \frac{1}{2}]$ , so a single bit carries an arbitrary prescribed weight rather than a fixed one, and the four rays (5) — each of which is realized by one bit on which the two endpoints act either as the identity or as the erasure map — can be combined with arbitrary real coefficients. Because the blocks are independent and the middle system is constant, the four conditional mutual information values of a product are the sums of the values of its blocks, and a product with one block per ray realizes the target exactly. Section 6 gives the construction and proves it; Appendix A contains the entropy expansions behind the two identities used.

The construction also comes with a size, and the second result of the paper is a pair of statements about it. Inside this block family the least number of product bits that realizes a point with cone coordinates  $(a_0, a_1, a_2, a_3)$  is

$$K_{\min} = \sum_{r=0}^3 \left\lceil a_r / \ln 2 \right\rceil, \quad (6)$$

so the two endpoint systems have  $2^{K_{\min}}$  elements each (Proposition 21). No realization of any kind can do better than an input alphabet of size  $e^x$ , because conditional mutual information is bounded by the entropy of either of its two systems (Proposition 22); together these bracket the block construction within a factor of sixteen of a bound that holds for every finite realization, and within a factor of four for each output alphabet.

The third result concerns the exact status of (6). It is a lower bound only within the block family, and it is not a lower bound in general. A single shared input bit, sent through two independent binary symmetric channels whose flip probability  $p$  solves  $h(p) = (\ln 2)/2$ , realizes the quadruple

$$\left( \ln 2, \frac{\ln 2}{2}, \frac{\ln 2}{2}, 0.1985612089284704960 \right),$$

all of whose four cone coordinates are strictly positive, so the block family needs one block per ray and hence four bits, that is, endpoint alphabets of sixteen elements. One bit is the least any realization of that quadruple can use, by the lower bound of Proposition 22, so the gap is real and the block count is not the answer to the question “how small can the systems be”. Section 8 gives the construction and its closed forms; the exact minimum of the alphabet size as a function of the quadruple remains open, bracketed but not determined.

Every identity, inequality and number used here is also recomputed arithmetically, from full joint distributions rather than from the additivity argument that the proofs use, so the two routes are independent (Section 9). The exact-realization construction is checked on 702 targets, including a 121-point sweep of the bit bias and targets whose cone coordinates exceed  $\ln 2$  and therefore require several bits per block, with a worst deviation of  $5.1 \times 10^{-80}$  at eighty decimal digits. The outer bound is checked on 60 random classical states with random channels and on 100 random quantum states with random channels, where the largest violation of any of the nine defining inequalities is  $8.4 \times 10^{-81}$  classically and zero quantum-mechanically, and where the slack identity (3) holds to  $4.4 \times 10^{-80}$ . A further independent recomputation at sixty digits, using different algorithms and, in the quantum case, a Stinespring dilation in place of a kept input–output joint, confirms all of it and the numbers quoted above.

Two things this paper does not claim. The reachable set is exactly  $P$ , but the alphabet that realizes a point grows with the point: since  $x$  is unbounded on  $P$ , no single finite alphabet serves the whole cone. And the classification says nothing about what happens when the two channels are correlated with each other, or when the output dimensions are fixed in advance; both are left as open questions in Section 10.

Section 2 places the result among the inequalities it uses. Section 3 fixes the setting and conventions, Section 4 the geometry of the cone. Sections 5 and 6 prove the two inclusions, Section 7 counts the bits the construction needs and bounds what any realization needs, and Section 8 exhibits the one-bit realization that the count cannot match. Section 9 describes the arithmetic certification, and Section 10 states what remains open.

## 2 Related work

The quantities studied here are the ones Shannon introduced for a discrete source and a discrete noisy channel: the entropy of a distribution, the conditional entropy of one variable given another, and their combination into a measure of how much one variable says about the other [6]. Two conventions of that paper are kept throughout. The logarithm is natural, so all entropies and all conditional mutual informations in this paper are in nats; Shannon’s “natural units” rather than his binary digits. And the conditional entropy is a conditional quantity in the strong sense that matters here — it is an average of entropies of conditional distributions, not a difference of entropies of a single object — which is what makes the identity for the slack  $D$  in (3) an exact statement about a joint distribution rather than an inequality about approximations to one. In the quantum setting the same definitions are made with von Neumann entropies of reduced states, and every statement below that is proved for quantum systems restricts to the classical one by taking all states diagonal in a fixed product basis.

Two theorems do all the work on the outer bound. Strong subadditivity of the quantum entropy, proved by Lieb and Ruskai [2], is the statement that conditional mutual information is nonnegative; it supplies  $x, y, z, w \geq 0$  and it supplies the nonnegativity of the slack  $D$  once  $D$  has been identified with a conditional mutual information on an enlarged system. Entropy inequalities for completely positive maps, in the form established by Lindblad [3], give the data-processing step: processing one of the two systems by a channel cannot increase the conditional mutual information of the pair. Neither theorem is sharpened here, and neither is the place where the difficulty of the problem lies. For the reader who prefers not to take the data-processing inequality on authority, Appendix A derives the monotonicity needed here directly from strong subadditivity applied to a Stinespring dilation, so the outer bound rests on [2] alone.

The additive bound (2), equivalently the nonnegativity of the slack, for two independent channels acting on the two endpoint systems of a tripartite state is the result of [4], where it appears as the nonnegativity of an environment slack. That is the only inequality in (4) that is not a direct consequence of monotonicity and nonnegativity, and it is the one that makes the cone four-dimensional rather than a box: it is what excludes, for example, a quadruple in which both channels independently destroy almost all of the correlation. We use the bound as an identity, with the identification of  $D$  as a conditional mutual information of the inputs given the outputs (classically) or of the environments given the outputs (quantum mechanically), and we prove both forms in Appendix A; nothing in this paper depends on [4] beyond the attribution.

The classification of the reachable region in the closure is [5]. That work proves that the closure of the reachable set, in both the classical and the quantum setting, is the cone  $P$  of (4), and identifies the four extreme rays (5) together with the endpoint maps that realize them. Its realizability argument uses a bit with the single entropy value  $\ln 2$ , so a ray can be scaled only by an integer number of copies, and an arbitrary real weight is reached by approximating it with rational combinations and passing to a limit. The hypotheses are the same as ours, and the outer bound is the same; the difference is entirely on the inner side. Substituting a biased bit for the uniform one makes the set of weights attainable by a single bit the whole interval  $[0, \ln 2]$  rather than the single point  $\{\ln 2\}$ , which removes the need for any approximation, any rational-combination argument and any limit, and turns the closure statement into the equality of Theorem 16. The present paper does not reprove the outer half of [5] by a different route; it reproves it, and adds the exact inner half together with the count (6).

The count itself has no antecedent in either work, and neither has its refutation. The observation that a single shared input bit passed through two independent binary symmetric channels realizes a quadruple with all four cone coordinates strictly positive — and therefore forces the block

construction to spend four bits where one suffices — was made independently of the construction and communicated to us, where it was rechecked at eighty decimal digits before being reproduced here with its closed forms and certified numerically in Section 8. It converts the natural conjecture that (6) is the least number of copies needed by any realization into a false statement, and it replaces a minimization problem that looked closed by an open one, which is the reason the alphabet bracket of Section 7 is stated as a bracket.

### 3 Setting and notation

#### 3.1 Systems, channels and the four readings

All systems are finite-dimensional. In the quantum setting  $A$ ,  $B$  and  $C$  are finite-dimensional Hilbert spaces and the input is a density operator  $\rho \equiv \rho_{ABC}$  on  $A \otimes B \otimes C$ . The two channels are completely positive trace-preserving maps  $\Phi$  acting on  $A$  and  $\Psi$  acting on  $C$ , with output spaces  $A'$  and  $C'$ ; they are independent in the sense that the joint action on the two endpoints is the tensor product  $\Phi \otimes \Psi$ , so neither channel sees the other's system and neither is correlated with it. The middle system  $B$  is never touched. In the classical setting the same picture is a joint probability distribution  $p(a, b, c)$  on three finite alphabets and two stochastic matrices, one from  $A$  to  $A'$  and one from  $C$  to  $C'$ ; a classical system is a quantum one whose state is diagonal in a fixed product basis and whose channels are diagonal in the same basis, so every classical construction below is also a quantum one.

For any three systems  $X$ ,  $Y$ ,  $Z$  of a joint state, the conditional mutual information is

$$I(X:Y | Z) = H(XZ) + H(YZ) - H(Z) - H(XYZ), \quad (7)$$

with  $H$  the von Neumann entropy  $-\text{Tr}(\sigma \ln \sigma)$  of the reduced state, or the Shannon entropy  $-\sum p \ln p$  of the relevant marginal in the classical setting. All logarithms are natural and all quantities are in nats. When  $Z$  is a single-element system, (7) reduces to the mutual information  $H(X) + H(Y) - H(XY)$  of the pair, a case the constructions below use repeatedly.

The four readings of (1), drawn in Figure 1, are evaluated on four different states, which is worth stating explicitly because it is where a careless expansion goes wrong:

$$\begin{aligned} x &= I(A:C | B)_\rho, & y &= I(A':C | B)_{(\Phi \otimes \text{id})\rho}, \\ z &= I(A:C' | B)_{(\text{id} \otimes \Psi)\rho}, & w &= I(A':C' | B)_{(\Phi \otimes \Psi)\rho}. \end{aligned} \quad (8)$$

Equivalently, and this is the form in which the identities of Appendix A are proved, all four are marginals of one larger object: of the joint distribution of the inputs and the outputs in the classical setting, or of the state obtained by applying Stinespring dilations of both channels to  $\rho$  in the quantum setting. The losses and their overlap are

$$\Delta_A = x - y, \quad \Delta_C = x - z, \quad \Delta_{AC} = x - w, \quad D = \Delta_A + \Delta_C - \Delta_{AC} = x - y - z + w. \quad (9)$$

Two structural facts about the channels are used without further comment. Because  $A'$  is obtained from  $A$  alone, the systems  $A'$ ,  $A$  and the pair  $(B, C)$  form a Markov chain in that order, and symmetrically  $C'$ ,  $C$ ,  $(A, B)$  do; this is what licenses the data-processing step. And an erasure map — the channel that sends every input to one fixed output symbol — is the extreme case of a channel, with  $y = 0$  whenever  $A'$  is erased, and it is one of the two endpoint maps the construction uses, the other being the identity.

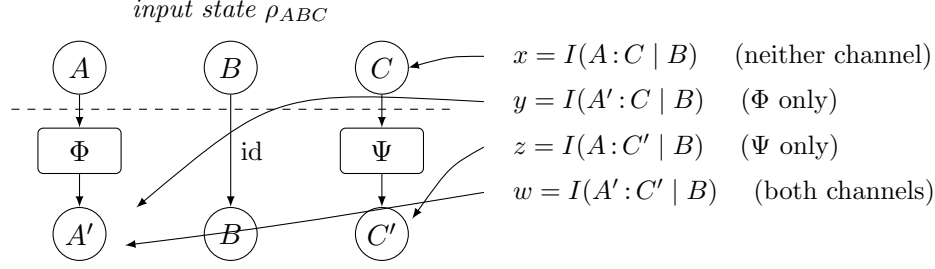


Figure 1: The setting, schematically. Two independent channels act on the two endpoint systems of a tripartite state; the middle system is untouched. The four conditional mutual information values of the quadruple are the four readings of the same correlation, taken before the channels act, after each channel acts alone, and after both act. The arrows on the right indicate which of the systems shown each of the four readings involves.

### 3.2 Reachable sets

**Definition 2** (reachable quadruples). A quadruple  $(x, y, z, w) \in \mathbb{R}^4$  is *classically reachable* if there are finite alphabets  $A, B, C, A', C'$ , a joint distribution on  $A \times B \times C$ , and two stochastic maps  $A \rightarrow A'$  and  $C \rightarrow C'$  whose four readings (8) are  $(x, y, z, w)$ . It is *quantum reachable* if the same holds with finite-dimensional Hilbert spaces, a density operator and two completely positive trace-preserving maps. The sets of classically and quantum reachable quadruples are denoted  $R$  and  $R_q$ .

Because a classical system embeds in a quantum one,  $R \subseteq R_q$ . Both sets are cones in the sense that they contain the origin, realized by any product input, and the main result identifies both with the same closed cone  $P \subseteq \mathbb{R}^4$ , which is therefore itself the reachable set in either setting and is closed without a separate argument.

Two conventions fix the meaning of “size” later on. For a classical system,  $|A|$  is the number of elements of its alphabet; for a quantum system,  $d_A$  is the dimension of its Hilbert space, and the classical statement is the special case in which the state is supported on  $|A|$  basis vectors. A block of product bits of length  $K$  has alphabet size  $2^K$ , and an erased system has one element, not zero: the erasure map outputs a single constant symbol.

## 4 The cone and its geometry

### 4.1 Cone coordinates

The nine inequalities in (4) are not independent, and seeing which of them carry the information is the first step towards realizing points of  $P$ . The right coordinates are the coefficients in the expansion over the four rays (5).

**Proposition 3** (cone coordinates). *The linear map*

$$a_0 = x - y - z + w, \quad a_1 = y - w, \quad a_2 = z - w, \quad a_3 = w \quad (10)$$

*is a bijection of  $\mathbb{R}^4$  onto itself, with inverse*

$$x = a_0 + a_1 + a_2 + a_3, \quad y = a_1 + a_3, \quad z = a_2 + a_3, \quad w = a_3, \quad (11)$$

*and it carries  $P$  onto the nonnegative orthant. Consequently  $P$  is the set of nonnegative combinations of  $e_0, e_1, e_2, e_3$ , every point of  $P$  has a unique such expansion, and  $P$  is a simplicial cone.*

*Proof.* Substituting (11) into (10) returns  $a_0, a_1, a_2, a_3$ , and substituting (10) into (11) returns  $x, y, z, w$ , so the two maps are inverse; both are linear, hence bijective. The expansion  $(x, y, z, w) = a_0 e_0 + a_1 e_1 + a_2 e_2 + a_3 e_3$  is exactly (11) read column by column, which proves that the four rays generate  $P$  once the image of  $P$  is the orthant. For that, note that the nine inequalities of (4) become, in the new coordinates,

$$\begin{aligned} x \geq 0 &\iff a_0 + a_1 + a_2 + a_3 \geq 0, & y \geq 0 &\iff a_1 + a_3 \geq 0, & z \geq 0 &\iff a_2 + a_3 \geq 0, \\ w \geq 0 &\iff a_3 \geq 0, & y \leq x &\iff a_0 + a_2 \geq 0, & z \leq x &\iff a_0 + a_1 \geq 0, \\ w \leq y &\iff a_1 \geq 0, & w \leq z &\iff a_2 \geq 0, & w \geq y + z - x &\iff a_0 \geq 0. \end{aligned} \tag{12}$$

Four of them —  $w \geq 0$ ,  $w \leq y$ ,  $w \leq z$  and  $w \geq y + z - x$  — are exactly  $a_3 \geq 0$ ,  $a_1 \geq 0$ ,  $a_2 \geq 0$  and  $a_0 \geq 0$ ; the other five follow from these four, each being a sum of them with coefficients in  $\{0, 1\}$ . So the image of  $P$  is defined by  $a_r \geq 0$  for  $r = 0, 1, 2, 3$  and is the orthant. Uniqueness of the expansion is the bijectivity, and a cone generated by four linearly independent vectors in  $\mathbb{R}^4$  is simplicial.  $\square$

Only four of the nine inequalities are therefore facets, and they are the four that appear as single coordinates in (10). Table 1 lists all nine with their cone-coordinate form and the rays that saturate them; the incidence pattern is the complement of the identity, each facet containing exactly the three rays other than the one whose coordinate it kills. This was verified in exact rational arithmetic, not in floating point, over the four rays and 2642 random rational points of both descriptions; the protocol is recorded in Appendix B.

Table 1: The nine defining inequalities of  $P$ , in cone coordinates, with the extreme rays that saturate each. The four rows marked “facet” are the facets of  $P$ ; the other five are sums of two of the four facet forms and are therefore redundant, though they are the forms in which the outer bound is naturally proved. Rays are indexed as in (5).

| inequality in $(x, y, z, w)$ | form in $(a_0, a_1, a_2, a_3)$ | facet | rays with equality  |
|------------------------------|--------------------------------|-------|---------------------|
| $x \geq 0$                   | $a_0 + a_1 + a_2 + a_3 \geq 0$ | no    | none but the origin |
| $y \geq 0$                   | $a_1 + a_3 \geq 0$             | no    | $e_0, e_2$          |
| $z \geq 0$                   | $a_2 + a_3 \geq 0$             | no    | $e_0, e_1$          |
| $y \leq x$                   | $a_0 + a_2 \geq 0$             | no    | $e_1, e_3$          |
| $z \leq x$                   | $a_0 + a_1 \geq 0$             | no    | $e_2, e_3$          |
| $w \geq 0$                   | $a_3 \geq 0$                   | yes   | $e_0, e_1, e_2$     |
| $w \leq y$                   | $a_1 \geq 0$                   | yes   | $e_0, e_2, e_3$     |
| $w \leq z$                   | $a_2 \geq 0$                   | yes   | $e_0, e_1, e_3$     |
| $w \geq y + z - x$           | $a_0 \geq 0$                   | yes   | $e_1, e_2, e_3$     |

## 4.2 Extremality, necessity and the simplex slice

**Proposition 4** (extremality and necessity). *Each of  $e_0, e_1, e_2, e_3$  is an extreme ray of  $P$ : none is a nonnegative combination of the other three. Each of the four facet inequalities is necessary: relaxing  $a_r \geq 0$  to  $a_r \geq -\varepsilon$  for any  $\varepsilon > 0$  admits points that violate at least one of the nine inequalities of (4).*

*Proof.* By (10) the cone coordinates of  $e_r$  are the  $r$ -th standard basis vector, while every nonnegative combination of the other three rays has  $r$ -th coordinate 0; the equation  $\sum_{s \neq r} \lambda_s e_s = e_r$  therefore requires  $0 = 1$  in that coordinate and has no solution at all, in any sign pattern. For necessity, the

point with cone coordinates  $a_r = -\varepsilon$  and  $a_s = 0$  for  $s \neq r$  has  $w = -\varepsilon$  when  $r = 3$ ,  $y - w = -\varepsilon$  when  $r = 1$ ,  $z - w = -\varepsilon$  when  $r = 2$ , and  $x - y - z + w = -\varepsilon$  when  $r = 0$ ; in each case one of the nine inequalities of (4) fails, since the nine are equivalent to  $a_s \geq 0$  for all  $s$  by (12).  $\square$

Because  $P$  is generated by four rays whose first coordinate is 1, every slice  $\{x = \lambda\}$  with  $\lambda > 0$  is a scaled copy of the simplex  $\{x = 1\}$ , and  $P$  is the cone over that simplex.

**Lemma 5** (the slice is a tetrahedron).  *$P \cap \{x = 1\}$  is the convex hull of the four rays, a tetrahedron in the coordinates  $(y, z, w)$  with vertices*

$$e_0 : (0, 0, 0), \quad e_1 : (1, 0, 0), \quad e_2 : (0, 1, 0), \quad e_3 : (1, 1, 1),$$

and the numbers  $a_r/x$  are the barycentric coordinates of a point of the slice with respect to those vertices. Its four faces are  $\{w = 0\}$ ,  $\{y = w\}$ ,  $\{z = w\}$  and  $\{w = y + z - 1\}$ .

*Proof.* Every point of  $P$  is  $\sum_r a_r e_r$  with  $a_r \geq 0$ , and its first coordinate is  $\sum_r a_r$ ; restricting to  $x = 1$  therefore restricts to  $\sum_r a_r = 1$ , which is the convex hull of the four rays. They are affinely independent because  $e_1 - e_0$ ,  $e_2 - e_0$  and  $e_3 - e_0$  are linearly independent, so the hull is a tetrahedron and  $a_r/x$  are its barycentric coordinates. The faces are the four facets intersected with the slice, and their equations in  $(y, z, w)$  follow by substituting  $x = 1$  into  $w = 0$ ,  $y = w$ ,  $z = w$  and  $w = y + z - x$ .  $\square$

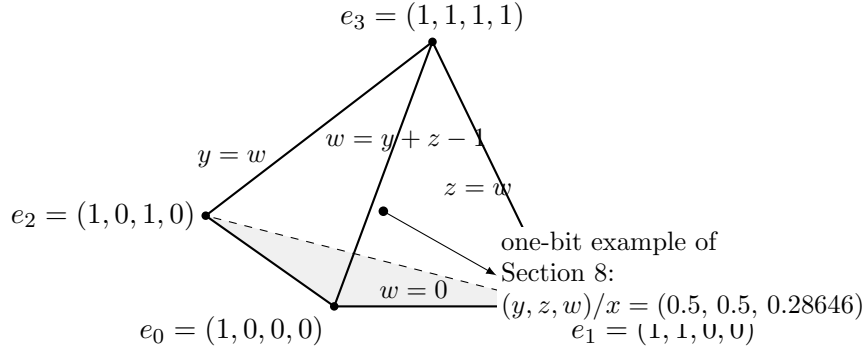


Figure 2: The slice  $\{x = 1\}$  of the cone  $P$ , drawn in the coordinates  $(y, z, w)$  with  $y$  to the right,  $z$  to the upper left and  $w$  upwards; the projection is schematic and the drawing is not to scale along the viewing direction. The four vertices are the extreme rays, the four faces carry the facet equations of Table 1, and the dashed edge  $e_1 e_2$  is hidden behind the shaded face  $\{w = 0\}$ . Every point of the slice is a unique convex combination of the four vertices, with weights  $a_r/x$ ; the marked interior point is the quadruple realized by one shared input bit in Section 8, whose four barycentric coordinates  $(0.2864633, 0.2135367, 0.2135367, 0.2864633)$  are all strictly positive, so all four rays contribute to it.

Lemma 5, drawn in Figure 2, is the geometric content of the classification: to realize an arbitrary quadruple it is enough to realize the four vertices with arbitrary weights, because the cone coordinates are the weights and the coordinates are additive over independent blocks. That is the strategy of Section 6, and it is also where the earlier closure-only argument lost exactness: with vertices realizable only at a fixed weight, the weights of a general point had to be approximated.

## 5 The outer bound

This section proves that every reachable quadruple lies in  $P$ . Two ingredients are needed: monotonicity of conditional mutual information under a local channel, and an exact identity for the slack  $D$  of (9). Both are proved in full in Appendix A; the statements and the role each plays are here.

**Lemma 6** (monotonicity). *Let  $\rho_{ABC}$  be a state and let  $\Phi, \Psi$  act independently on  $A$  and  $C$ , giving the quadruple  $(x, y, z, w)$  of (8). Then*

$$y \leq x, \quad z \leq x, \quad w \leq y, \quad w \leq z.$$

The first two inequalities are the data-processing step along the Markov chains  $A' - A - (B, C)$  and  $C' - C - (A, B)$  of Section 3: conditioning on  $B$  does not affect the argument, and processing  $A$  cannot increase what  $A$  tells about  $C$ . The last two apply the same step twice, once for each endpoint. In the quantum setting the inequality is a consequence of strong subadditivity applied to a Stinespring dilation of the channel [2]; the general statement for completely positive maps is the content of the entropy inequalities of [3]. Appendix A gives the dilation argument, so that the outer bound below depends on strong subadditivity alone.

**Proposition 7** (the slack is a conditional mutual information). *In the classical setting, let  $p(a, a', b, c, c')$  be the joint distribution of the two inputs, the middle system and the two outputs, obtained by applying both channels to  $\rho$  and keeping the inputs. Then*

$$x - y - z + w = I(A : C \mid A'BC'). \quad (13)$$

*In the quantum setting, let  $V_\Phi : A \rightarrow A' \otimes E_A$  and  $V_\Psi : C \rightarrow C' \otimes E_C$  be Stinespring dilations of the two channels, with environments  $E_A$  and  $E_C$ , and let  $\sigma_{A'E_ABC'E_C}$  be the state obtained by applying  $V_\Phi \otimes V_\Psi$  to  $\rho$ . Then*

$$x - y - z + w = I(E_A : E_C \mid A'BC'). \quad (14)$$

*In both cases the right-hand side is nonnegative by strong subadditivity.*

The two forms say the same thing in two languages: what the two channels destroy in common is exactly the correlation that survives between the two things they discarded — classically the inputs, which are recoverable from the record, and quantum mechanically the environments, which play the same role when the inputs cannot be kept alongside the outputs. The additive bound  $w \geq y + z - x$  for independent endpoint channels is due to [4], where it is obtained as the nonnegativity of an environment slack; the identification with a conditional mutual information is what makes the bound exact rather than merely valid, and it is what the numerical certification of Section 9 tests as an equality rather than as a sign.

**Theorem 8** (outer bound).  $R_q \subseteq P$ . *Every quantum reachable quadruple satisfies all nine inequalities of (4).*

*Proof.* Let  $(x, y, z, w)$  be realized by a state  $\rho_{ABC}$  and independent channels  $\Phi, \Psi$ . Nonnegativity of all four entries is strong subadditivity applied to the four triples  $(A, B, C)$ ,  $(A', B, C)$ ,  $(A, B, C')$  and  $(A', B, C')$  [2]. The four inequalities  $y \leq x$ ,  $z \leq x$ ,  $w \leq y$  and  $w \leq z$  are Lemma 6, and  $w \leq \min(y, z)$  is the last two of them. It remains to produce  $w \geq y + z - x$ , which is Proposition 7: the slack  $x - y - z + w$  equals a conditional mutual information and is therefore nonnegative. By Proposition 3 these nine inequalities are exactly membership of  $P$ .  $\square$

**Corollary 9.**  $R \subseteq R_q \subseteq P$ .

*Proof.* A classical realization is a quantum realization in which every state is diagonal in a fixed product basis and every channel is diagonal in the same bases, so  $R \subseteq R_q$ ; the rest is Theorem 8.  $\square$

Two remarks fix the scope of the outer bound.

**Remark 10** (independence is used, and is necessary). The hypothesis that the two channels act independently enters twice: in the Markov chains behind Lemma 6, and in the factorization of the joint distribution or of the dilation used in Proposition 7. It cannot be dropped. Let  $A$  and  $C$  be independent uniform bits, let  $B$  be a single element, and let one *joint* channel emit the same parity into both outputs,  $A' = C' = A \oplus C$ . Because  $A$  and  $C$  are independent and uniform, the parity is uniform and independent of each of them, so  $x = 0$ ,  $y = I(A' : C \mid B) = 0$  and  $z = I(A : C' \mid B) = 0$ ; but the two outputs are identical, so  $w = I(A' : C' \mid B) = \ln 2$ . The quadruple  $(0, 0, 0, \ln 2)$  violates  $w \leq y$ ,  $w \leq z$  and  $w \leq x$ : a joint channel can create correlation between the two outputs out of correlation that neither endpoint possessed, which is exactly what independence excludes.

**Remark 11** (sharpness). Each of the four facet inequalities is saturated by reachable quadruples: by Table 1 the facet  $\{a_r = 0\}$  contains three of the four extreme rays, and Section 6 realizes every ray, hence every nonnegative combination of the three rays in each facet. No facet inequality can therefore be strengthened, and  $P$  is the smallest cone of this form containing the reachable set. The five redundant inequalities are saturated too, but they are consequences of the four facets and carry no separate sharpness statement of their own.

## 6 Exact realization by finite alphabets

The outer bound says that nothing outside  $P$  is reachable. This section proves the converse: every point of  $P$  is reached exactly, by a classical construction on finitely many product bits. The argument has three steps — a single biased bit can carry any prescribed entropy up to  $\ln 2$ ; a single bit realizes any multiple of any one extreme ray; and independent blocks add. The third step is the only one that needs care, and its full expansion is in Appendix A.

### 6.1 A biased bit carries any weight in $[0, \ln 2]$

Let  $Z$  be a bit with  $\Pr[Z = 0] = q$  and  $\Pr[Z = 1] = 1 - q$ , and let

$$h(q) = -q \ln q - (1 - q) \ln(1 - q) \quad (15)$$

be its entropy in nats, with the usual convention  $0 \ln 0 = 0$ .

**Lemma 12** (range of the binary entropy).  *$h$  is continuous on  $[0, 1]$ , symmetric under  $q \mapsto 1 - q$ , satisfies  $h(0) = h(1) = 0$  and  $h(\frac{1}{2}) = \ln 2$ , and is strictly increasing on  $[0, \frac{1}{2}]$ , where its derivative is  $\ln \frac{1-q}{q} > 0$ . Hence  $h : [0, \frac{1}{2}] \rightarrow [0, \ln 2]$  is a bijection: for every  $a \in [0, \ln 2]$  there is a unique  $q(a) \in [0, \frac{1}{2}]$  with  $h(q(a)) = a$ .*

*Proof.* Continuity, symmetry and the three values are immediate from (15); on  $(0, 1)$  the derivative is  $h'(q) = -\ln q - 1 + \ln(1 - q) + 1 = \ln \frac{1-q}{q}$ , which is positive exactly when  $q < \frac{1}{2}$ . A continuous strictly increasing map of a compact interval onto its image is a bijection onto  $[h(0), h(\frac{1}{2})] = [0, \ln 2]$ .  $\square$

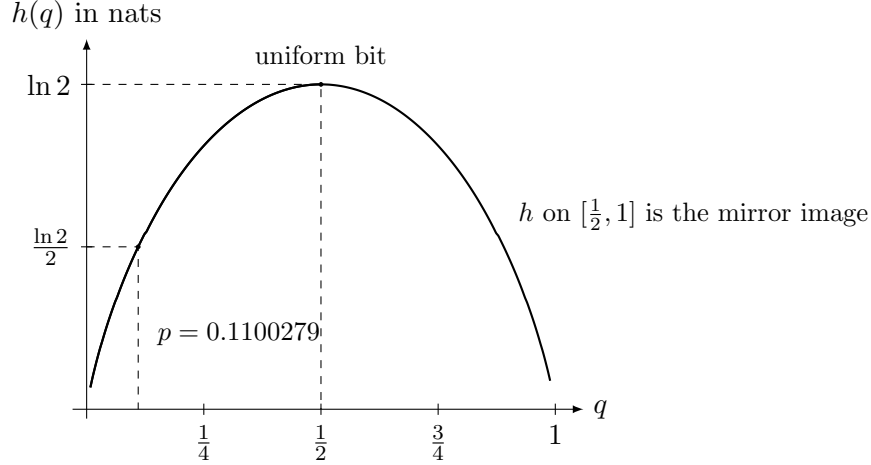


Figure 3: The binary entropy (15) in nats, drawn by the TeX engine from the closed form; it is an analytic curve, not a set of measurements. The heavy part is the branch  $q \in [0, \frac{1}{2}]$ , on which  $h$  is a bijection onto  $[0, \ln 2]$  (Lemma 12), so one bit realizes every weight in that interval. The uniform bit is the single value  $h = \ln 2$  at  $q = \frac{1}{2}$ , which is all a construction based on uniform bits can reach with one bit per block; the marked point is the value  $h(p) = (\ln 2)/2$  at  $p = 0.1100279$  used by the binary symmetric channels of Section 8.

## 6.2 One bit realizes any multiple of any ray

Fix  $a \in [0, \ln 2]$  and let  $q = q(a)$  be as in Lemma 12; Figure 3 draws  $h$  together with the two values of it that this paper uses. Take a single bit  $Z$  with bias  $q$ , identify both endpoints with it,  $A = C = Z$ , and take  $B$  to be a single element. Then  $x = I(A : C \mid B) = I(Z : Z) = H(Z) = a$ , since with trivial conditioning the conditional mutual information of a variable with itself is its entropy. The four readings now depend only on what the two endpoint maps do to  $Z$ , and there are four choices built from the identity and the erasure map.

Table 2: The four endpoint-map assignments of Lemma 13, each applied to one bit  $Z$  of entropy  $a$  with  $A = C = Z$  and  $B$  a single element. The realized quadruple is  $a$  times one extreme ray, and the alphabet sizes of the two outputs are those of a bit or of an erased system. Which facet the ray lies in is read off Table 1.

| ray                  | $\Phi$ on $A$ | $\Psi$ on $C$ | $(x, y, z, w)$  | $ A' $ | $ C' $ |
|----------------------|---------------|---------------|-----------------|--------|--------|
| $e_0 = (1, 0, 0, 0)$ | erasure       | erasure       | $a(1, 0, 0, 0)$ | 1      | 1      |
| $e_1 = (1, 1, 0, 0)$ | identity      | erasure       | $a(1, 1, 0, 0)$ | 2      | 1      |
| $e_2 = (1, 0, 1, 0)$ | erasure       | identity      | $a(1, 0, 1, 0)$ | 1      | 2      |
| $e_3 = (1, 1, 1, 1)$ | identity      | identity      | $a(1, 1, 1, 1)$ | 2      | 2      |

**Lemma 13** (one bit, one ray). *For every  $a \in [0, \ln 2]$  and every  $r \in \{0, 1, 2, 3\}$  the endpoint-map assignment in row  $r$  of Table 2 realizes exactly  $a e_r$ .*

*Proof.* Erasing a system makes it a single element, so every conditional mutual information in which it appears as one of the two systems vanishes, and the identity leaves every value unchanged. Row by row, with  $x = a$ : for  $r = 0$  both outputs are constant, so  $y = z = w = 0$ . For  $r = 1$ ,  $A' = A = Z$  gives  $y = I(Z : Z) = a$ , while  $C'$  is constant, so  $z = w = 0$ . For  $r = 2$  the roles are exchanged. For  $r = 3$  both outputs are  $Z$ , so  $y = z = w = a$ .  $\square$

This is where the construction differs from one built on uniform bits. A uniform bit supplies the single weight  $\ln 2$ , so reaching a ray at weight  $a$  requires  $\lceil a/\ln 2 \rceil$  bits and, for  $a$  not a multiple of  $\ln 2$ , an approximation. A biased bit supplies every weight in  $[0, \ln 2]$  exactly, so a single bit per ray suffices for those weights and no approximation appears anywhere.

### 6.3 Independent blocks add

**Lemma 14** (additivity of the quadruple over blocks). *Let  $A = A_0A_1A_2A_3$  and  $C = C_0C_1C_2C_3$  be products of independent systems, let  $B$  be a single element, and let the endpoint maps act block by block, so that  $A' = A'_0A'_1A'_2A'_3$  and  $C' = C'_0C'_1C'_2C'_3$  with  $A'_r$  depending only on  $A_r$  and  $C'_r$  only on  $C_r$ . If block  $r$  realizes  $(x_r, y_r, z_r, w_r)$  on its own, then the product realizes*

$$\left( \sum_r x_r, \sum_r y_r, \sum_r z_r, \sum_r w_r \right).$$

The hypothesis is exactly the one the construction satisfies: the blocks are independent by fiat, the maps act within blocks, and  $B$  is constant. The proof is an expansion of each of the four quantities into entropies of products, which split as sums because the joint distribution factorizes over blocks; it is given in Appendix A, together with the bookkeeping of the blocks whose contribution cancels. The point of stating it as a lemma is that additivity is a hypothesis-free-looking step that is in fact the only place where independence of the blocks and triviality of  $B$  are used on the inner side, and the numerical certification of Section 9 recomputes all four values from the full joint distributions rather than from this lemma.

### 6.4 The realization theorem

**Theorem 15** (exact realization). *Every point of  $P$  is classically reachable:  $P \subseteq R$ . More precisely, let  $(x, y, z, w) \in P$  with cone coordinates  $(a_0, a_1, a_2, a_3)$ , let  $K_r = \lceil a_r/\ln 2 \rceil$  for  $a_r > 0$  and  $K_r = 0$  for  $a_r = 0$ , and let  $K = \sum_r K_r$ . Then there is a realization in which  $B$  is a single element,  $A$  and  $C$  are the same  $K$  product bits, each endpoint map is the identity or the erasure map block by block as in Table 2, and the alphabet sizes are*

$$|A| = |C| = 2^K, \quad |A'| = 2^{K_1+K_3}, \quad |C'| = 2^{K_2+K_3}, \quad |B| = 1. \quad (16)$$

*The realized quadruple is exactly  $(x, y, z, w)$ , for arbitrary real weights  $a_r$ , with no limiting procedure.*

*Proof.* For each ray  $r$  build a block of  $K_r$  independent bits whose entropies sum to  $a_r$ : take  $K_r - 1$  uniform bits, each of entropy  $\ln 2$ , and, when  $a_r > 0$ , one further bit of entropy  $a_r - (K_r - 1)\ln 2$ . That residual lies in  $(0, \ln 2]$  because  $K_r - 1 < a_r/\ln 2 \leq K_r$ , so Lemma 12 supplies its bias. If  $a_r = 0$  the block is empty and contributes nothing. Apply to block  $r$  the endpoint maps of row  $r$  of Table 2; by Lemma 13 block  $r$  realizes  $a_r e_r$  on its own, because the sum of the entropies of its bits is  $a_r$  and the bits are independent. Take  $A$  and  $C$  to be the concatenations of the four blocks, so that  $A$  and  $C$  are the same  $K$  bits, and let  $B$  be a single element. Lemma 14 gives the quadruple  $\sum_r a_r e_r$ , which is  $(x, y, z, w)$  by (11).

For the alphabet sizes,  $A$  and  $C$  are  $K$  bits, so each has  $2^K$  values. An erased block contributes one symbol to the corresponding output and a kept block contributes its  $2^{K_r}$  values, so  $A'$  has  $2^{K_1+K_3}$  values — blocks 1 and 3 are the ones kept by  $\Phi$  — and  $C'$  has  $2^{K_2+K_3}$ .  $\square$

**Theorem 16** (exact reachable region).  *$R = R_q = P$ . The set of quadruples realizable by finite-dimensional classical states and independent classical channels, and the set realizable by finite-dimensional quantum states and independent completely positive trace-preserving maps, are both the closed cone  $P$  of (4).*

*Proof.* Theorem 15 gives  $P \subseteq R$ , Corollary 9 gives  $R \subseteq R_q \subseteq P$ , and the three inclusions together give the equality  $R = R_q = P$ .  $\square$

**Corollary 17** (closedness, boundary and closure-free statements).  *$R$  and  $R_q$  are closed, since both equal the closed set  $P$ ; the closure operation that appeared in the earlier classification [5] is vacuous here. Every boundary point of  $P$  — in particular every point of every facet and every ray — is realized by a finite classical system. The reachable set is a cone over a tetrahedron, and the four extreme rays are reachable.*

**Remark 18** (what the construction assumes). The realization is deliberately special: the two endpoint systems are the same bits, so the input correlation between  $A$  and  $C$  is as large as it can be; the middle system is trivial; and the endpoint maps are products of identities and erasures. None of this is needed for the statement of Theorem 16, which is an existence statement, but it means the theorem should not be read as a claim that general points of  $P$  are reachable only this way. Section 8 exhibits a point reached with far smaller systems by a different route.

**Remark 19** (no uniform alphabet bound). The alphabet in (16) depends on the target. Since  $x$  is unbounded on  $P$  and  $x \leq H(A) \leq \ln |A|$  for any realization (Proposition 22), no finite alphabet can serve all of  $P$ ; exactness here means exactness point by point, each point with its own finite systems. What is not needed at any point is an infinite system, a limit of systems, or a number of copies growing with a precision parameter.

## 7 How many bits the construction needs

Theorem 15 realizes every point of  $P$ , but it says nothing about how large the systems have to be. This section answers that question for the construction itself, and then bounds what any construction whatsoever must spend. The two answers together bracket the block construction, and Section 8 shows that the block count is not the true minimum.

### 7.1 The block family and its least bit count

**Definition 20** (ray-block family). A realization belongs to the *ray-block family* if  $B$  is a single element,  $A$  and  $C$  are the same  $K$  independent bits, the bits are partitioned into four banks indexed by the rays  $r = 0, 1, 2, 3$ , and the endpoint maps act bit by bit, every bit in bank  $r$  receiving the pair of maps in row  $r$  of Table 2.

This is exactly the family produced by the proof of Theorem 15; a bank may be empty, and the biases of the bits in a bank are free.

**Proposition 21** (least bit count in the family). *Let  $(x, y, z, w) \in P$  have cone coordinates  $(a_0, a_1, a_2, a_3)$ . Every member of the ray-block family that realizes it uses at least*

$$K_{\min} = \sum_{r=0}^3 \left\lceil a_r / \ln 2 \right\rceil \quad (17)$$

*bits, where  $\lceil 0 / \ln 2 \rceil$  is read as 0, and the construction of Theorem 15 uses exactly  $K_{\min}$ . Hence the least endpoint alphabet inside the family is  $|A| = |C| = 2^{K_{\min}}$ .*

*Proof.* In the family, bank  $r$  consists of  $K_r$  independent bits whose concatenation is both the  $A$ -part and the  $C$ -part of that block, and by Lemma 13 and Lemma 14 the block realizes  $a_r e_r$ . The first coordinate of  $a_r e_r$  is the conditional mutual information of the block's two endpoint systems given the trivial middle system, which is the entropy of the bank, so  $a_r = H(\text{bank } r)$ . Independent bits carry at most  $\ln 2$  nats each, so  $a_r \leq K_r \ln 2$  and  $K_r \geq \lceil a_r / \ln 2 \rceil$ ; summing over  $r$  gives  $K \geq K_{\min}$ . The construction of Theorem 15 attains this with  $K_r - 1$  uniform bits and one bit of residual entropy in  $(0, \ln 2]$ , whose bias exists by Lemma 12, so  $K_r = \lceil a_r / \ln 2 \rceil$  is achieved for every  $r$  with  $a_r > 0$  and  $K_r = 0$  for  $a_r = 0$ .  $\square$

The count is a property of the family, not of the point: it says that no arrangement of identity and erasure maps on independent bits, one bank per ray, does better. It does not exclude a realization of a different shape with fewer or smaller systems, and the next subsection shows that a universal lower bound of a different kind exists.

## 7.2 A lower bound that holds for every realization

**Proposition 22** (universal alphabet bound). *Let  $(x, y, z, w)$  be realized by any finite system, classical or quantum, with endpoint systems  $A, C$  and outputs  $A', C'$ . Then*

$$x \leq H(A) \leq \ln |A|, \quad x \leq H(C) \leq \ln |C|, \quad y \leq H(A') \leq \ln |A'|, \quad z \leq H(C') \leq \ln |C'|, \quad (18)$$

where  $|A|$  is the alphabet size classically and the dimension  $d_A$  quantum mechanically. Consequently  $|A|, |C| \geq e^x$ ,  $|A'| \geq e^y$  and  $|C'| \geq e^z$ .

*Proof.* By (7),  $I(A:C | B) = H(A|B) - H(A|BC) \leq H(A|B) \leq H(A)$ , because conditional entropy is at most unconditional entropy and entropies are nonnegative. The entropy of a system with  $|A|$  possible values, or of dimension  $d_A$ , is at most  $\ln |A|$  respectively  $\ln d_A$ . The bound on  $y$  is the same argument with  $A'$  in place of  $A$ , and the bound on  $z$  with  $C'$  in place of  $C$ . Exponentiating gives the stated sizes.  $\square$

The bound is elementary and it is sharp as a bound on  $|A|$  in terms of  $x$  alone: a point with  $x = \ln 2$  cannot be realized with an alphabet of one element, and Section 8 exhibits one that is realized with exactly two.

**Corollary 23** (the block construction is within a factor of sixteen). *Let  $(x, y, z, w) \in P$  have cone coordinates  $a_r$ , let  $m$  be the number of strictly positive  $a_r$ , and let  $m_A$  and  $m_C$  be the numbers of strictly positive  $a_r$  with  $r \in \{1, 3\}$  and  $r \in \{2, 3\}$  respectively. The ray-block realization of Theorem 15 has*

$$e^x \leq |A| = |C| = 2^{K_{\min}} < 2^m e^x \leq 16 e^x, \quad (19)$$

and likewise  $e^y \leq |A'| < 2^{m_A} e^y \leq 4 e^y$  and  $e^z \leq |C'| < 2^{m_C} e^z \leq 4 e^z$ . The least possible input alphabet for the point lies in  $[e^x, 2^{K_{\min}}]$ , an interval of ratio at most  $2^m \leq 16$ .

*Proof.*  $K_{\min} \geq \sum_r a_r / \ln 2 = x / \ln 2$  gives  $2^{K_{\min}} \geq e^x$ . For the upper bound,  $\lceil t \rceil < t + 1$  for every  $t > 0$  and  $\lceil 0 \rceil = 0$ , so  $K_{\min} < \sum_{r: a_r > 0} (a_r / \ln 2 + 1) = x / \ln 2 + m$  and  $2^{K_{\min}} < 2^m e^x$ ; since  $m \leq 4$  this is at most  $16 e^x$ . For the outputs,  $|A'| = 2^{K_1 + K_3}$  by (16) and the same estimate applied to the two banks kept by  $\Phi$  gives  $K_1 + K_3 < y / \ln 2 + m_A$  with  $m_A \leq 2$ , using  $y = a_1 + a_3$ ; the argument for  $|C'|$  is identical with  $z = a_2 + a_3$ .  $\square$

### 7.3 Two worked examples

Table 3 gives two points of  $P$  with the data of their ray-block realizations, computed at sixty decimal digits and reproduced in Appendix B. The first is small enough to check by hand: all four cone coordinates lie strictly between 0 and  $\ln 2$ , so each bank is one bit, the four biases are the four values  $h^{-1}(a_r)$ , and the two endpoint systems have sixteen elements each. The second has two coordinates larger than  $\ln 2$  and one equal to zero, so the banks have three, two, zero and two bits, and the point lies on the facet  $\{z = w\}$  because  $a_2 = 0$ .

Table 3: Two ray-block realizations. Example A has all four cone coordinates in  $(0, \ln 2)$ , so each bank is a single bit and its bias is  $h^{-1}(a_r)$ . Example B has two coordinates exceeding  $\ln 2$  and one equal to zero, so its banks hold three, two, zero and two bits, each multi-bit bank consisting of uniform bits of entropy  $\ln 2$  and bias 0.5 plus one residual bit; the point lies on the facet  $\{z = w\}$  because  $a_2 = 0$ . Entropies are in nats and biases are probabilities, computed at sixty decimal digits and printed here to seven; the reconstructed quadruples agree with the targets to  $2.8 \times 10^{-60}$ . The last row of the upper block is the bracket of Corollary 23, with  $m = 4$  and  $m = 3$  strictly positive cone coordinates respectively.

|                                             | example A                | example B                       |
|---------------------------------------------|--------------------------|---------------------------------|
| cone coordinates $(a_0, a_1, a_2, a_3)$     | (0.3, 0.2, 0.1, 0.05)    | (1.4, 1.3, 0, 0.7)              |
| target $(x, y, z, w)$                       | (0.65, 0.25, 0.15, 0.05) | (3.4, 2.0, 0.7, 0.7)            |
| bank sizes $(K_0, K_1, K_2, K_3)$           | (1, 1, 1, 1)             | (3, 2, 0, 2)                    |
| $K_{\min}$ , and $ A  =  C  = 2^{K_{\min}}$ | 4, and 16                | 7, and 128                      |
| $ A' $ , and $ C' $                         | 4, 4                     | 16, 4                           |
| bracket $[e^x, 2^m e^x]$                    | [1.92, 30.6]             | [30.0, 239.7]                   |
| bit entropies in bank 0                     | 0.3                      | 0.6931472, 0.6931472, 0.0137056 |
| bit entropies in bank 1                     | 0.2                      | 0.6931472, 0.6068528            |
| bit entropies in bank 2                     | 0.1                      | (empty bank)                    |
| bit entropies in bank 3                     | 0.05                     | 0.6931472, 0.0068528            |
| bit biases in bank 0                        | 0.0889063                | 0.5, 0.5, 0.0018844             |
| bit biases in bank 1                        | 0.0505052                | 0.5, 0.2953271                  |
| bit biases in bank 2                        | 0.0205055                | —                               |
| bit biases in bank 3                        | 0.0087130                | 0.5, 0.0008491                  |

The bracket of Corollary 23 is a bracket and not an answer. It says that the block construction never wastes more than a factor of sixteen in the input alphabet, and it says nothing about which points admit a much smaller realization. The next section is about one that does.

## 8 A one-bit realization the block count cannot match

Proposition 21 is a lower bound inside the ray-block family. This section exhibits a point of  $P$  that the family realizes with four bits and that a single bit realizes exactly, so the count (17) is not a lower bound on the size of a realization in general. The example is not part of the construction of Section 6: it was communicated to us independently of it, and everything stated here is recomputed from the joint distribution of the three variables involved.

### 8.1 The construction

Let  $S$  be a uniform bit, let  $A = C = S$  and let  $B$  be a single element, so that

$$x = I(A : C \mid B) = H(S) = \ln 2. \quad (20)$$

Let the two endpoint channels be binary symmetric channels with the same flip probability  $p$ , acting independently:  $A' = S \oplus E_A$  and  $C' = S \oplus E_C$  with  $E_A, E_C$  independent and equal to 1 with probability  $p$ . Choose  $p$  to be the unique root in  $(0, \frac{1}{2})$  of

$$h(p) = \frac{1}{2} \ln 2, \quad \text{that is} \quad p = 0.1100278644383595512618117\dots, \quad (21)$$

whose existence and uniqueness are Lemma 12. The choice makes each single-channel reading exactly half of the input value.

**Proposition 24** (the quadruple of the one-bit construction). *The construction above realizes*

$$(x, y, z, w) = \left( \ln 2, \frac{\ln 2}{2}, \frac{\ln 2}{2}, \ln 2 - h(p^2 + (1-p)^2) \right), \quad (22)$$

whose cone coordinates are

$$a_0 = a_3 = 0.1985612089284704960447981\dots, \quad a_1 = a_2 = 0.1480123813515021586638179\dots, \quad (23)$$

all four strictly positive. The point is therefore in the relative interior of the cone, and its ray-block realization uses  $K_{\min} = 4$  bits, that is,  $|A| = |C| = 2^4 = 16$ .

*Proof.* The input is uniform and the channels are binary symmetric, so  $A'$  and  $C'$  are uniform bits and  $H(A') = H(C') = \ln 2$ . Given  $S$ , the output  $A'$  is  $S$  flipped with probability  $p$ , so  $H(A'|S) = h(p) = \frac{1}{2} \ln 2$  by (21) and

$$y = I(A' : C | B) = I(A' : S) = H(A') - H(A'|S) = \ln 2 - \frac{1}{2} \ln 2 = \frac{1}{2} \ln 2,$$

where the first equality uses  $C = S$  and  $B$  trivial;  $z$  is identical by symmetry. For  $w$ , note that  $A' = C'$  exactly when the two flips agree, which has probability  $r = p^2 + (1-p)^2$ , so  $A' = C' \oplus F$  with  $F$  a bit equal to 1 with probability  $2p(1-p) = 1-r$  and independent of  $C'$ , and  $H(A'|C') = h(r) = h(1-r)$ . Hence  $w = I(A' : C') = \ln 2 - h(r)$ , and numerically  $r = 0.8041565330290129502180374\dots$ ,  $h(r) = 0.494585971631474813372434\dots$  and  $w = 0.1985612089284704960447981\dots$

The cone coordinates follow from (10):  $a_3 = w$ ;  $a_1 = y - w = \frac{1}{2} \ln 2 - w$ ;  $a_2 = z - w = a_1$ ; and  $a_0 = x - y - z + w = w$ , since  $x = \ln 2$  by (20) and  $y = z = \frac{1}{2} \ln 2$ . Each is positive since  $0 < w < \frac{1}{2} \ln 2$ . Each is also strictly less than  $\ln 2$ , so  $\lceil a_r / \ln 2 \rceil = 1$  for all four  $r$  and  $K_{\min} = 4$  by (17); Proposition 21 makes four bits necessary within the family, and Theorem 15 makes four bits sufficient for the same point.  $\square$

## 8.2 What the example does and does not settle

**Corollary 25.**  $K_{\min}$  is not the least number of input bits over all finite realizations, and  $2^{K_{\min}}$  is not the least endpoint alphabet. For the quadruple (22) the least possible value of  $|A| + |C|$  is 4, attained by the construction above, while the ray-block family needs  $|A| + |C| = 32$ .

*Proof.* Proposition 22 gives  $|A| \geq e^x = 2$  and  $|C| \geq 2$ , so  $|A| + |C| \geq 4$  for every realization of this quadruple, and the construction attains  $|A| = |C| = 2$ . The ray-block realization of the same quadruple has  $|A| = |C| = 2^4 = 16$  by Proposition 24.  $\square$

The same argument settles the two output alphabets for this point:  $e^y = e^z = \sqrt{2}$ , so  $|A'| \geq 2$  and  $|C'| \geq 2$ , and the binary symmetric channels attain both. For this quadruple, then, the exact minimum of every one of the four alphabet sizes is known and equals two. For a general point of  $P$

Table 4: The one-bit realization of Proposition 24, computed at eighty decimal digits and independently reproduced at sixty; the printed values are truncations of the computed ones, correct in both passes. The flip probability was obtained by Newton iteration in the first pass and by bisection in the second, and the two agree in every printed digit. The quadruple was obtained twice, once from the closed forms of (22) and once from the joint distribution of  $(S, A', C')$  summed over all eight outcomes, and the two agree exactly at the printed precision. The last two rows compare it with the ray-block realization of the same point.

| quantity                                                  | value                                                     | origin         |
|-----------------------------------------------------------|-----------------------------------------------------------|----------------|
| flip probability $p$ , root of $h(p) = \frac{1}{2} \ln 2$ | 0.1100278644383595512618                                  | root finding   |
| agreement probability $r = p^2 + (1 - p)^2$               | 0.8041565330290129502180                                  | closed form    |
| $x = \ln 2$                                               | 0.6931471805599453094172                                  | closed form    |
| $y = z = \frac{1}{2} \ln 2$                               | 0.3465735902799726547086                                  | closed form    |
| $w = \ln 2 - h(r)$                                        | 0.1985612089284704960448                                  | closed form    |
| $a_0 = a_3$                                               | 0.1985612089284704960448                                  | (10)           |
| $a_1 = a_2$                                               | 0.1480123813515021586638                                  | (10)           |
| barycentric coordinates $a_r/x$                           | 0.2864632714, 0.2135367286,<br>0.2135367286, 0.2864632714 | (10), sum = 1  |
| $K_{\min}$ and ray-block alphabets                        | 4; $ A  =  C  = 16$ , $ A'  =  C'  = 4$                   | Proposition 21 |
| one-bit alphabets                                         | $ A  =  C  = 2$ , $ A'  =  C'  = 2$ , $ B  = 1$           | construction   |
| lower bounds $e^x$ , $e^y$ , $e^z$                        | 2, 1.4142135624, 1.4142135624                             | Proposition 22 |

it is not: the bracket of Corollary 23 is all the present argument gives. Table 4 collects the numbers of this point, its cone coordinates, and its two realizations side by side.

The mechanism behind the gap explains why a unique algebraic decomposition does not force a unique physical one. The cone coordinates  $(a_0, a_1, a_2, a_3)$  are unique, and the block family realizes each coordinate with a separate independent source:  $a_3$ , the part of the correlation that survives both channels, requires its own block of bits passed through two identity maps. A partially noisy channel does the same work with a single source. The shared bit  $S$  supplies the surviving part of the correlation — that is the  $a_3$  component — while the two independent flips simultaneously destroy exactly the amounts that appear as  $a_1$  and  $a_2$  and leave the redundant loss  $a_0 = D$ . One random variable, seen through two independent noisy copies, produces all four components at once; four independent random variables, seen cleanly or not at all, produce one component each. The uniqueness of (10) is a statement about the quadruple, not about the state that realizes it.

Two consequences are worth recording. First, the classification of Section 6 is about the region and not about the size of the systems that realize a point of it, and the two questions separate completely. Second, the example sits in the relative interior of the cone, where all four cone coordinates are positive and the block family is therefore forced to spend at least one bit on every ray; that is the situation in which the gap between the family and the true minimum can be largest. Whether an equally large gap occurs for points on a facet, where the family drops a block entirely, is open (Section 10).

## 9 Numerical certification

The classification is proved, and no part of it depends on arithmetic. What the arithmetic does is check the algebra on inputs the proofs do not enumerate one by one: the expansions behind the two defect identities, the additivity step that the inner bound uses, and the inequalities on random states and random channels of a kind the construction never produces. It also fixes the decimal

values quoted in this paper. Every number printed here comes from such a computation, at a stated working precision, and the protocols and per-family coverage are in Appendix B. Table 5 collects the checks with their coverage and their worst deviations.

Three things were checked directly, and each was then checked a second time by an independent implementation that shares no code with the first and uses a different algorithm wherever two exist: conditional entropies in place of the  $H(L) + H(R) - H(LR)$  form of mutual information, bisection in place of Newton iteration for the inverse binary entropy, exact rational arithmetic in place of floating point for everything about the cone, and a Stinespring dilation in place of a kept input–output joint for the quantum defect identity. Agreement between two independent routes is evidence about the routes; agreement with the target is evidence about the construction.

The construction was certified by building the product of biased bits for a prescribed target and evaluating all four conditional mutual information values from the full joint distributions of the four pairs  $(A, C)$ ,  $(A', C)$ ,  $(A, C')$  and  $(A', C')$ , so that additivity is tested rather than assumed. Coverage is 702 targets: the four extreme rays at three weights each; a sweep of 121 bit biases across the whole interval  $[0, 1]$ , combined with each of the four rays; three fixed cone points whose coordinates are neither rational multiples of  $\ln 2$  nor on a grid; 200 random cone points; and three targets with cone coordinates exceeding  $\ln 2$ , which force multi-bit banks of up to eight bits in total. The worst deviation between a reconstructed value and its target over all 702 targets and all four coordinates is  $5.1 \times 10^{-80}$  at eighty decimal digits, which is the rounding level of the working precision.

The outer bound was certified in two parts. Classically, 60 random states on three binary alphabets were drawn together with random stochastic endpoint channels from two input symbols to three output symbols, and all nine inequalities of (4) were evaluated; the largest violation is  $8.4 \times 10^{-81}$ , that is, none at the working precision of eighty digits. The same 60 trials test the defect identity (13) as an equality, by computing  $I(A:C | A'BC')$  on the joint distribution of both inputs and both outputs; the largest discrepancy is  $4.4 \times 10^{-80}$ . This pass draws its random states without a fixed seed, so these two residuals move from one run to the next within the rounding level of the working precision; Appendix B records a second run of the same procedures. Quantum mechanically, 100 random density operators on three qubits were drawn together with random completely positive trace-preserving endpoint maps given by Kraus operators, and the same nine inequalities were evaluated in double precision; the largest violation is zero, with a tolerance of  $10^{-9}$  set by the eigenvalue computation.

The one-bit example of Section 8 was certified by enumerating all eight outcomes of  $(S, A', C')$  and computing  $x, y, z$  and  $w$  from the resulting joint distribution, then comparing with the closed forms of Proposition 24; the agreement is exact at eighty digits, well inside the  $10^{-75}$  tolerance imposed, and the four cone coordinates were confirmed strictly positive with  $K_{\min} = 4$ .

The independent recomputation covers the same ground and adds four items that the first pass does not test: the cone geometry of Section 4, in exact rational arithmetic, where 2642 random rational points confirm that the two descriptions of  $P$  are equivalent, that the four rays are extreme in the sense of Proposition 4, and that the facet incidence of Table 1 is the complement of the identity; the quantum defect identity in the Stinespring form (14), on 120 random states of dimensions up to  $(3, 2, 4)$  with environments of dimension up to three, where  $x - y - z + w$  and  $I(E_A : E_C | A'BC')$  agree to  $3.8 \times 10^{-14}$  in double precision and the dilated state reproduces the channel outputs to  $8.4 \times 10^{-17}$ ; the entropy bounds behind Proposition 22, on 120 random classical states and channel pairs, where  $x \leq H(A) \leq \ln |A|$  and its three analogues hold with zero deviation; and the bracket of Corollary 23, on 500 random cone points, where  $x/\ln 2 \leq K_{\min} < x/\ln 2 + 4$  and the corresponding output bounds hold exactly. Its 60-digit recomputation of 60 realization targets, of the two worked examples of Table 3 and of the one-bit point agrees with the targets to  $4.2 \times 10^{-60}$  at worst, and its

two routes for mutual information agree with each other to  $4.4 \times 10^{-60}$ .

Table 5: Summary of the numerical certification. “Deviation” is the largest absolute difference between a computed quantity and the value the theory assigns to it, over the stated coverage; a violation of an inequality is recorded as the amount by which it is negative. Precision is the working number of decimal digits, except for the quantum rows, which are in IEEE double precision. Protocols and per-family coverage are in Appendix B.

| check                        | coverage                   | quantity tested                        | deviation             |
|------------------------------|----------------------------|----------------------------------------|-----------------------|
| exact realization            | 702 targets, $\leq 8$ bits | four coordinates vs. target            | $5.1 \times 10^{-80}$ |
| classical outer bound        | 60 trials                  | nine inequalities of (4)               | $8.4 \times 10^{-81}$ |
| classical defect identity    | 60 trials                  | (13) as an equality                    | $4.4 \times 10^{-80}$ |
| quantum outer bound          | 100 trials                 | nine inequalities                      | 0                     |
| cone geometry                | 2642 rational points       | both descriptions of $P$ , extremality | exact                 |
| realization, second pass     | 60 targets                 | four coordinates, two routes           | $4.2 \times 10^{-60}$ |
| classical bound, second pass | 120 trials                 | inequalities and defect identity       | $2.5 \times 10^{-60}$ |
| quantum bound, second pass   | 120 trials                 | nine inequalities                      | 0                     |
| Stinespring defect identity  | 120 trials                 | (14) as an equality                    | $3.8 \times 10^{-14}$ |
| alphabet lower bounds        | 120 trials                 | (18)                                   | 0                     |
| alphabet bracket             | 500 cone points            | (19)                                   | exact                 |
| worked examples              | 2 targets                  | Table 3                                | $2.8 \times 10^{-60}$ |
| one-bit point                | 1 point                    | Table 4                                | exact                 |

Two limits on what these checks establish. They are finite: 702 targets, 60 and 100 random trials, and 500 random cone points do not exhaust  $P$ , and the universal statements of this paper rest on their proofs, not on the sampling. And the quantum rows are in double precision, so a violation of an inequality smaller than about  $10^{-12}$  could pass undetected there; the quantum checks are a consistency test of the algebra on genuinely quantum inputs — entangled states and channels with non-commuting Kraus operators — which is a different thing from the classical certification, where the arithmetic is exact to eighty digits and the residual is rounding.

## 10 Discussion and open questions

### 10.1 What exactness changes

The reachable set is a closed cone with four extreme rays, and every one of its points is produced by a finite classical system whose two endpoint systems are products of at most  $K_{\min}$  bits. That statement makes three separate claims. As a classification, it says that the four numbers (1) are constrained by exactly four independent inequalities — nonnegativity of the doubly processed value, the two comparisons of it with the singly processed values, and the subadditivity of the two losses — and that every pattern consistent with those four occurs. As an existence statement, it says that no approximation is hidden in the word “occurs”: a closure result leaves open whether the approximating systems have bounded size, and here they do not need to be approximating at all. As a construction, it says how to build one, with the biases of the bits given by inverting the binary entropy at the cone coordinates of the target.

The four rays are four archetypal loss patterns, and Table 2 is a catalogue of them: everything destroyed at both endpoints, correlation surviving one channel and not the other in either of the two asymmetric ways, and nothing destroyed. A general quadruple is a mixture of the four in the unique proportions (10). The identity of Proposition 7 gives the coefficient of the first archetype an operational reading that does not refer to the construction:  $D = a_0$  is the correlation between

the two inputs, classically, or between the two channel environments, quantum mechanically, that remains once both outputs and the middle system are known, so the facet  $\{a_0 = 0\}$  is exactly the set of quadruples for which that conditional mutual information vanishes, that is, for which those two systems are conditionally independent given the outputs and the middle system. On that facet the two losses add up to the total loss and the two channels destroy disjoint parts of the correlation; off it they destroy a common part, and  $D$  is its size.

## 10.2 Scope of what is proved

The theorem is a finite-dimensional statement, in both settings. Nothing here addresses systems of infinite dimension, where the entropy can diverge and the reachability of a prescribed quadruple is a different question; the outer bound uses only strong subadditivity and monotonicity, which have infinite-dimensional forms, but the inner bound is a construction on finitely many bits and does not carry over by itself.

The alphabets of the construction grow with the target. This is unavoidable rather than an artifact:  $x$  is unbounded on  $P$  and  $x \leq \ln |A|$  holds for every realization, so no single finite alphabet serves the whole cone. What the theorem asserts is that each point is served by some finite alphabet, and Corollary 23 bounds how much larger than the unavoidable minimum that alphabet can be.

The construction realizes points with two endpoint systems that are literally the same bits and a middle system with one element. It is an existence proof, and it should not be read as a claim about typical states and channels: nothing here says how a generic triple  $(\rho, \Phi, \Psi)$  distributes over the cone, only which points of the cone some triple reaches.

Independence of the two channels is essential, and Remark 10 gives the two-line counterexample. Whether a joint channel acting on both endpoints has a reachable region that is still polyhedral, and what generates it, is a different problem that this paper does not touch.

## 10.3 Open questions

The exact minimum of the alphabet sizes as a function of the quadruple is open. Corollary 23 brackets the input alphabet between  $e^x$  and  $2^{K_{\min}} < 16e^x$ , and Section 8 determines it exactly for one point, where it equals the lower bound. Whether the lower bound  $e^x$  is attained for every point of  $P$ , or which points admit a realization at the bound, is not known; nor is the analogous question for  $|A| + |C|$ , or for the output alphabets beyond the factor-of-four bracket.

Fixing the output dimensions in advance changes the problem from a classification to a trade-off. The construction needs  $|A'| = 2^{K_1+K_3}$  and  $|C'| = 2^{K_2+K_3}$ , which grow with  $y$  and with  $w$ , so for prescribed output alphabets only part of  $P$  is exactly attainable, and the attainable subset is neither characterized here nor obviously a cone. Its shape, and whether its boundary is reached by erasure-like maps or by partially noisy ones, are open.

The biased-bit construction is extremal for the four coordinates by design, and it is natural to ask for which other functionals of the quadruple it is extremal: products, ratios, or entropy-like combinations of  $(x, y, z, w)$ , and the analogous questions for a fixed input state as the two channels vary. Whether an exact cone-shaped classification survives when the conditional mutual information is replaced by a conditional Rényi mutual information is also open; the outer bound would need the corresponding Rényi data-processing and subadditivity statements, and the inner bound would need a construction whose Rényi readings are controllable by a single parameter, which the biased bit does not obviously provide.

Finally, more than two endpoints. With  $k$  systems acted on by  $k$  independent channels there are  $2^k$  readings of the same conditional mutual information, one for each subset of the channels applied,

and the same two tools — monotonicity and a chain rule for the successive losses — apply to them. Whether the reachable set of the resulting  $2^k$ -tuples is again a simplicial cone whose rays are realized exactly by identity-and-erasure assignments on independent biased bits is a direct generalization of the question answered here. The argument does not carry over as it stands: the four cone coordinates (10) come from one inclusion–exclusion pattern in two channels, and the pattern that would replace them for  $2^k$  readings has to be identified before either half of the classification can even be attempted.

## 11 Conclusion

Four conditional mutual information values — the input one, the two obtained by acting on a single endpoint, and the one obtained by acting on both — are constrained by exactly four independent inequalities, and every quadruple that satisfies them is realized exactly by a finite system. The reachable set is therefore the closed cone  $P$  of (4) in both the classical and the quantum setting, with no closure, no limit and no unbounded number of copies anywhere in the argument. The construction that realizes a point is a product of independent biased bits, one bank per extreme ray, with the identity or the erasure map at each endpoint of each bank; a single bit carries any weight up to  $\ln 2$  because the binary entropy maps  $[0, \frac{1}{2}]$  onto  $[0, \ln 2]$ , and that is the whole of what closes the gap left by a construction built on uniform bits.

Inside that family the least number of bits is  $K_{\min} = \sum_r \lceil a_r / \ln 2 \rceil$  in the cone coordinates of the target, and no realization of any kind uses an input alphabet smaller than  $e^x$ , which brackets the family within a factor of sixteen of a universal bound. The family is nevertheless not economical: one shared input bit and two independent binary symmetric channels with flip probability  $0.11002786443836\dots$  realize a quadruple with all four cone coordinates positive, for which the family needs four bits and sixteen-element endpoint alphabets, while two elements suffice and two is the least possible. The unique algebraic decomposition of a quadruple over the extreme rays does not dictate how many physical systems realize it.

What remains open is the size question in general — the exact minimum of the input and output alphabets as a function of the quadruple — together with the trade-off that appears when the output dimensions are fixed in advance, and the analogue of the whole classification for Rényi conditional mutual information or for more than two endpoints.

## A Entropy identities

This appendix contains the proofs of the two identities used in Section 5 and of the additivity lemma used in Section 6, together with the derivation of the saturation pattern of Table 1. Everything is an identity of entropy sums; the only inequality ever invoked is strong subadditivity, in the form “conditional mutual information is nonnegative” [2].

### A.1 Two facts used repeatedly

The first is the chain rule for conditional mutual information,

$$I(XY : Z \mid W) = I(X : Z \mid W) + I(Y : Z \mid XW), \quad (24)$$

which holds whenever all the entropies involved are finite, in the classical and in the quantum setting, because both sides expand to the same signed sum of entropies:  $I(X : Z \mid W) = H(XW) + H(ZW) - H(W) - H(XZW)$  and  $I(Y : Z \mid XW) = H(XYW) + H(XZW) - H(XW) - H(XYZW)$ , and

in the sum  $H(XW)$  cancels  $-H(XW)$  while  $-H(XZW)$  cancels  $H(XZW)$ , leaving  $H(XYW) + H(ZW) - H(W) - H(XYZW)$ , which is  $I(XY : Z \mid W)$ . The rule is used twice in each proof below, once with the pair  $(X, Y)$  on the left of the colon and once with it on the right.

The second is the conditional independence created by a channel. Classically, if  $A'$  is produced from  $A$  by a stochastic map with private randomness independent of everything else, then  $A'$  is conditionally independent of any system  $Y$  that does not involve the map's randomness once  $A$  is known:

$$I(A' : Y \mid A) = 0, \quad (25)$$

and the same holds with  $A$  replaced by any set containing  $A$ . Quantum mechanically the corresponding statement is about the environment of a Stinespring dilation  $V_\Phi : A \rightarrow A' \otimes E_A$ : on the dilated state,  $E_A$  is correlated only with  $A'$ , and the identities below never need more than the chain rule applied to the split  $A \mapsto (A', E_A)$ .

Finally, isometries do not change entropies, so if  $\sigma_{A'E_A BC' E_C} = (V_\Phi \otimes \text{id}_B \otimes V_\Psi) \rho (V_\Phi \otimes \text{id}_B \otimes V_\Psi)^\dagger$  then each of the four readings (8) of  $\rho$  equals a reading of  $\sigma$  in which the inputs have been replaced by the corresponding system–environment pairs:

$$\begin{aligned} x &= I(A'E_A : C'E_C \mid B)_\sigma, & y &= I(A' : C'E_C \mid B)_\sigma, \\ z &= I(A'E_A : C' \mid B)_\sigma, & w &= I(A' : C' \mid B)_\sigma. \end{aligned} \quad (26)$$

## A.2 Monotonicity: proof of Lemma 6

Classically, keep the joint distribution of inputs and outputs. Applying (24) to  $I(AA' : C \mid B)$  in the two possible orders gives

$$I(AA' : C \mid B) = I(A : C \mid B) + I(A' : C \mid AB) = x + 0 = x$$

by (25) with  $Y = (B, C)$ , and also

$$I(AA' : C \mid B) = I(A' : C \mid B) + I(A : C \mid A'B) = y + I(A : C \mid A'B).$$

Hence

$$x - y = I(A : C \mid A'B) \geq 0, \quad (27)$$

which is stronger than the inequality  $y \leq x$ : the loss caused by the channel on  $A$  is itself a conditional mutual information, namely the correlation between the input  $A$  and the untouched system  $C$  that survives once the output  $A'$  is known. Exchanging the two endpoints throughout gives

$$x - z = I(A : C \mid C'B) \geq 0, \quad z - w = I(A : C' \mid A'B) \geq 0, \quad (28)$$

and applying the first computation to the state  $(\Phi \otimes \text{id})\rho$ , in which the channel  $\Psi$  has yet to act, gives

$$y - w = I(A' : C \mid C'B) \geq 0. \quad (29)$$

The four inequalities (27), (28) and (29) are the four claims of Lemma 6.

Quantum mechanically the same computation runs on the dilated state  $\sigma$  of (26), in which the input  $A$  has been replaced by the pair  $(A', E_A)$  and the input  $C$  by the pair  $(C', E_C)$ . Applying (24) with the split  $A \mapsto (A', E_A)$ ,

$$x - y = I(E_A : C'E_C \mid A'B) \geq 0, \quad z - w = I(E_A : C' \mid A'B) \geq 0, \quad (30)$$

and with the split  $C \mapsto (C', E_C)$ ,

$$x - z = I(A'E_A : E_C \mid C'B) \geq 0, \quad y - w = I(A' : E_C \mid C'B) \geq 0. \quad (31)$$

Each of the four differences is a conditional mutual information of  $\sigma$  and is therefore nonnegative by strong subadditivity [2]; no data-processing inequality is assumed, and no conditional independence beyond the chain rule is used.

### A.3 The defect identities: proof of Proposition 7

Subtracting the identity for  $z - w$  in (28) from (27) gives, in the classical setting,

$$D = x - y - z + w = I(A : C \mid A'B) - I(A : C' \mid A'B).$$

Apply (24) to  $I(A : CC' \mid A'B)$  in the two possible orders, with the pair  $(C, C')$  on the right of the colon in both of them:

$$I(A : CC' \mid A'B) = I(A : C \mid A'B) + I(A : C' \mid A'BC) = I(A : C' \mid A'B) + I(A : C \mid A'BC').$$

The term  $I(A : C' \mid A'BC)$  vanishes:  $C'$  is produced from  $C$  by a channel whose private randomness is independent of  $A$  and  $A'$ , so (25) with the conditioning set  $\{A', B, C\}$ , which contains  $C$ , gives  $I(A : C' \mid A'BC) = 0$ . Therefore

$$D = I(A : C \mid A'BC') \geq 0,$$

which is (13), and the inequality is strong subadditivity.

In the quantum setting the inputs cannot be kept alongside the outputs, so the same computation is run on the dilated state  $\sigma$  of (26). Combining the two identities in (30),

$$D = (x - y) - (z - w) = I(E_A : C'E_C \mid A'B) - I(E_A : C' \mid A'B),$$

and (24) applied to  $I(E_A : C'E_C \mid A'B)$  with the pair  $(C', E_C)$  split as  $C'$  followed by  $E_C$  gives

$$I(E_A : C'E_C \mid A'B) = I(E_A : C' \mid A'B) + I(E_A : E_C \mid A'BC'),$$

so the first two terms cancel and

$$D = I(E_A : E_C \mid A'BC') \geq 0,$$

which is (14). No conditional independence beyond the chain rule is used in this setting.

Both forms of the identity were tested as equalities rather than inequalities: the classical one on 60 random states and channel pairs at eighty decimal digits, and the quantum one on 120 random states at double precision, with the dilated state also checked to reproduce the two channel outputs, as Section 9 records.

### A.4 Additivity over independent blocks: proof of Lemma 14

Let the four blocks be indexed by  $r$ , with  $A = A_0A_1A_2A_3$  and  $C = C_0C_1C_2C_3$  products of independent systems,  $B$  a single element, and endpoint maps acting within blocks, so that  $A' = A'_0A'_1A'_2A'_3$  and  $C' = C'_0C'_1C'_2C'_3$ . Independence of the blocks means the joint distribution of all ten

systems factorizes as a product over  $r$ , and the entropy of a product is the sum of the entropies of its factors. Expanding  $x$  with (7),

$$x = H(A) + H(C) - H(AC) = \sum_r (H(A_r) + H(C_r) - H(A_r C_r)) = \sum_r I(A_r : C_r \mid B) = \sum_r x_r,$$

where  $B$  contributes nothing because it has one element, so conditioning on it changes no entropy. For  $y$  the same expansion runs over  $A'$  and  $C$ :

$$y = H(A') + H(C) - H(A'C) = \sum_r (H(A'_r) + H(C_r) - H(A'_r C_r)) = \sum_r y_r,$$

and a block whose  $A'_r$  is a single symbol contributes  $H(C_r) - H(C_r) = 0$ , as it must, because that block's own  $y_r$  is zero. The expansions for  $z$  and  $w$  are identical with the roles of the two endpoints exchanged. Summing the four coordinate-wise proves the lemma.

Two hypotheses are load-bearing. Independence of the blocks is what makes the entropies additive; and  $B$  having one element is what makes each conditional mutual information an unconditional mutual information of a pair, so that no cross-block term can appear in the conditioning system. With a non-trivial  $B$  shared by the blocks the same computation produces additional terms  $I(A_r : B \mid \cdot)$  that do not cancel.

## A.5 Saturation pattern: derivation of Table 1

By (10) the cone coordinates of the ray  $e_r$  are the  $r$ -th standard basis vector:  $a_i(e_r) = \delta_{ir}$ . Each of the nine inequalities of (4) is a linear form in the cone coordinates, listed in (12); evaluating that form on  $e_r$  gives its value  $\delta_{ir}$  when the form is the single coordinate  $a_i$ , and a sum of two such values when it is not. An inequality is saturated by  $e_r$  exactly when the corresponding form vanishes on it, so the facet  $\{a_i = 0\}$  contains the three rays with  $r \neq i$  and excludes  $e_i$ , which is the complement of the identity matrix as an incidence pattern, and a redundant inequality such as  $y \geq 0$ , whose form is  $a_1 + a_3$ , is saturated exactly by the rays with  $a_1 = a_3 = 0$ , namely  $e_0$  and  $e_2$ . The row  $x \geq 0$  has form  $a_0 + a_1 + a_2 + a_3$ , which vanishes only at the origin, so no ray saturates it. This is arithmetic in the rationals, and it was carried out in exact rational arithmetic rather than in floating point, on the four rays and on 2642 random rational points of both descriptions of  $P$ .

## B Numerical protocols and coverage

This appendix records what the certification of Section 9 computes, how, and over what range, so that each residual in Table 5 can be read as the output of a stated finite procedure. Two bodies of arithmetic were produced independently of each other: a first pass at eighty decimal digits that tests the three claims of the paper on their own terms, and a second pass at sixty digits that re-implements the same quantities with different algorithms and adds the checks of Section 4 and of Proposition 22. All logarithms are natural. The second pass completes in about ninety seconds; the coverage figures below are the caps it reached, not limits it hit. Together with the derivations of Appendix A, both bodies are collected in [1].

### B.1 What is computed

*Realizing a prescribed point.* Given cone coordinates  $(a_0, a_1, a_2, a_3)$ , each bank  $r$  is built as  $\lceil a_r / \ln 2 \rceil$  bits:  $\lceil a_r / \ln 2 \rceil - 1$  uniform bits and one residual bit of entropy  $a_r - (\lceil a_r / \ln 2 \rceil - 1) \ln 2$ , whose bias is obtained by bisecting  $h$  on  $[0, \frac{1}{2}]$ , 300 halvings with an early stop once the interval is shorter than

$10^{-70}$ . The bits of all four banks are concatenated into  $A = C$ ,  $B$  is a single element, and each bank receives the endpoint maps of its row in Table 2. All  $2^K$  outcomes are enumerated with their exact product probabilities, and the four joint distributions of the pairs  $(A, C)$ ,  $(A', C)$ ,  $(A, C')$  and  $(A', C')$  are accumulated from them. Each of the four values is then computed twice, once as  $H(L) + H(R) - H(LR)$  and once as  $H(L) - \sum_R p(R) H(L | R)$  with the conditional distributions renormalized, and both are compared with the target (11). The bank entropies are also checked against the prescribed  $a_r$ , which tests the bias inversion separately from the construction, and the support sizes of  $A$ ,  $A'$  and  $C'$  are counted and compared with (16).

*Random classical states and channels.* A distribution on a finite set is drawn by normalizing independent exponential variates, which gives the uniform distribution on the simplex; in the first pass states are drawn on three binary alphabets and both channels map two symbols to three, while in the second pass the state alphabets rotate through  $(2, 2, 2)$ ,  $(2, 3, 2)$ ,  $(3, 2, 3)$  and  $(2, 2, 3)$  and each endpoint channel is drawn row by row the same way, with two or three output symbols. The kept joint distribution of  $(a, a', b, c, c')$  is formed by multiplying the state by both channel rows. On it, the nine linear forms of (4) are evaluated and the most negative value recorded as a violation, and the defect identity (13) is tested as an equality by computing  $I(A : C | A'BC')$  from the four marginals it involves. The entropy bounds (18) are tested on the same trials, in both of their parts:  $x \leq H(A)$  and  $H(A) \leq \ln |A|$  with  $|A|$  the size of the support.

*Random quantum states and channels.* A density operator on  $d_A \otimes d_B \otimes d_C$  is drawn from the Ginibre ensemble,  $\rho = GG^\dagger$  normalized by its trace with  $G$  a complex Gaussian matrix. A channel is drawn as  $n$  Kraus operators obtained by splitting a random isometry, produced from the orthonormal columns of a QR decomposition of a complex Gaussian matrix, into  $n$  blocks of  $d$  rows; trace preservation,  $\sum_e K_e^\dagger K_e = I$ , is asserted to  $10^{-10}$  before the channel is used. Conditional mutual informations are computed from the eigenvalues of partial traces, with negative eigenvalues clipped at zero. The dimension triples used are  $(2, 2, 2)$ ,  $(2, 3, 2)$ ,  $(3, 2, 4)$  and  $(2, 2, 3)$  with environment dimensions two and three. For the Stinespring form of the defect identity, both dilations are applied to the same state, producing a state on  $(A', E_A, B, C', E_C)$  with the output-major index convention; on it,  $I(E_A : E_C | A'BC')$  is computed and compared with  $x - y - z + w$  formed from the four channel outputs, and the reduction of the dilated state to  $(A', B, C')$  is compared entry by entry with the state obtained by applying the two channels directly.

*The cone, in exact arithmetic.* All statements about  $P$  are checked in rational arithmetic with no floating-point step anywhere. Two thousand points are generated by drawing four nonnegative rationals as cone coordinates, mapping them to  $(x, y, z, w)$  by (11), verifying the nine inequalities and verifying that the inverse map recovers the coordinates exactly. A further 642 points are generated in the opposite direction, by drawing  $x, y, z$  and then  $w$  uniformly in the interval  $[\max(0, y + z - x), \min(y, z)]$ , and are checked to have nonnegative cone coordinates; the remaining draws were discarded because that interval was empty for the drawn  $(x, y, z)$ , which happens exactly when  $\max(0, y + z - x) > \min(y, z)$ , so the two directions of sampling cover the same set from opposite sides. Extremality is certified by exact Gaussian elimination on each of the four systems  $\sum_{s \neq r} \lambda_s e_s = e_r$ , all four of which are reported inconsistent. The incidence pattern of Table 1 is obtained by evaluating each of the nine linear forms of (12) on each of the four rays, and necessity of each facet by exhibiting the point with  $a_r = -1$  and all other coordinates zero and checking that at least one of the nine inequalities fails on it.

*The one-bit point.* The flip probability is obtained by bisecting  $h(p) - \frac{1}{2} \ln 2$  on  $(0, \frac{1}{2})$  rather than by Newton iteration, the joint distribution of  $(S, A', C')$  is enumerated over its eight outcomes with  $S$  uniform and the two flips independent, and  $x, y, z$  and  $w$  are read off that joint by marginalization and compared with the closed forms of Proposition 24. The cone coordinates, their positivity,  $K_{\min}$ , the barycentric coordinates in the slice  $\{x = 1\}$  and their sum are all computed from the same

joint, and the ray-block construction is then run on those four coordinates to confirm that four bits reproduce the identical quadruple.

## B.2 Coverage of the first pass

Table 6: Coverage of the eighty-digit certification of the construction, by family of targets. “Bits” is the largest number of product bits in a single target of that family; the residual is the largest absolute deviation of any of the four coordinates from its target over the family. The last row is the total.

| family of targets                             | cases | bits | worst residual        |
|-----------------------------------------------|-------|------|-----------------------|
| extreme rays at three weights                 | 12    | 1    | $5.3 \times 10^{-82}$ |
| bias sweep, 121 values of $q$ times four rays | 484   | 1    | $2.2 \times 10^{-81}$ |
| fixed off-lattice cone points                 | 3     | 4    | $2.2 \times 10^{-81}$ |
| random cone points                            | 200   | 4    | $6.4 \times 10^{-81}$ |
| multi-bit banks, some $a_r > \ln 2$           | 3     | 8    | $5.1 \times 10^{-80}$ |
| total                                         | 702   | 8    | $5.1 \times 10^{-80}$ |

Table 6 breaks the certification of the construction down by family of targets; the worst residual over all of them is  $5.1 \times 10^{-80}$ , on the family with the most bits per target. The outer-bound pass of the same body of arithmetic covers 60 classical trials, with a largest violation of the nine inequalities of  $8.4 \times 10^{-81}$  and a largest defect-identity error of  $4.4 \times 10^{-80}$ , and 100 quantum trials on three qubits with random channels of three Kraus operators, with a largest violation of zero against a tolerance of  $10^{-9}$ . The one-bit point is certified against a tolerance of  $10^{-75}$  at eighty digits. The random draws of this pass are not seeded, so its residuals move from one run to the next within the rounding level of the working precision: the values quoted are those of the retained run, and an earlier run of the same scripts reported  $4.2 \times 10^{-81}$  and  $2.7 \times 10^{-80}$  for the same two quantities. The sixty-digit pass is seeded and reproduces every number in Table 7 exactly.

## B.3 Coverage of the independent recomputation

Every quantity of the first pass appears again in Table 7, computed by a different route, together with four checks the first pass does not make: the cone algebra of Section 4, the Stinespring form of the quantum defect identity, the entropy and alphabet bounds behind Proposition 22, and the alphabet bracket of Corollary 23. The rows labelled “exact” compare rationals or integers and admit no rounding at all; the rows labelled “0” are inequalities that no trial violated at the working precision of that pass.

## B.4 What the residuals mean, and what is not covered

The high-precision residuals sit at the rounding level of their working precision:  $10^{-80}$  for the eighty-digit pass and  $10^{-60}$  for the sixty-digit one, with the largest values occurring on the targets with the most bits, where the product probability of an outcome is a product of up to eight factors and the entropy sums run over up to 256 terms. They are consistent with exact agreement, and the exact rows — the cone algebra and the two bracket checks, which are inequalities between integers or rationals — leave no rounding at all. The double-precision rows are coarser by design: the Stinespring identity requires two dilations and five eigenvalue decompositions per trial, which is

Table 7: The independent recomputation, check by check. Deviations are absolute, in nats where the quantity is an entropy-like number; “exact” means zero deviation in rational arithmetic, and “0” means zero at the working precision. The quantum rows are in double precision with a tolerance of  $10^{-9}$ ; all others are at sixty decimal digits with a tolerance of  $10^{-50}$ .

| check                                       | coverage                               | worst deviation       |
|---------------------------------------------|----------------------------------------|-----------------------|
| cone algebra and facet incidence            | 2642 rational points, 4 rays, 4 facets | exact                 |
| realization, first route                    | 60 targets, up to 8 bits               | $4.1 \times 10^{-60}$ |
| realization, conditional-entropy route      | 60 targets                             | $4.2 \times 10^{-60}$ |
| agreement of the two routes                 | 60 targets                             | $4.4 \times 10^{-60}$ |
| defect equals $a_0$ on constructed points   | 60 targets                             | $5.0 \times 10^{-60}$ |
| bank entropies equal the prescribed weights | 60 targets                             | $1.6 \times 10^{-61}$ |
| classical nine inequalities                 | 120 random states and channel pairs    | 0                     |
| classical defect identity                   | 120 trials                             | $2.5 \times 10^{-60}$ |
| two routes for the input value              | 120 trials                             | $6.3 \times 10^{-61}$ |
| entropy and alphabet bounds (18)            | 120 trials                             | 0                     |
| quantum nine inequalities                   | 120 random states, dims to (3, 2, 4)   | 0                     |
| nonnegativity of the input value            | 120 states                             | 0                     |
| Stinespring defect identity (14)            | 120 states, environments to 3          | $3.8 \times 10^{-14}$ |
| dilation reproduces the channel outputs     | 120 states                             | $8.4 \times 10^{-17}$ |
| one-bit point, closed forms vs. joint       | 1 point                                | 0                     |
| one-bit point, two routes                   | 1 point                                | 0                     |
| one-bit point reproduced by four bits       | 1 point                                | $1.6 \times 10^{-61}$ |
| input alphabet bracket (19)                 | 500 random cone points                 | exact                 |
| output alphabet bracket (19)                | 500 random cone points                 | exact                 |
| worked examples of Table 3                  | 2 targets                              | $2.8 \times 10^{-60}$ |

why its residual is  $3.8 \times 10^{-14}$  rather than  $10^{-16}$ , while the inequalities, each a single evaluation of a linear form, show no violation at all.

The checks are finite. They cover 702 prescribed targets and 60 realization targets, 120 random classical and 120 random quantum states, 500 random cone points and one distinguished point; they do not exhaust  $P$ , and the universal statements of this paper are proved, not sampled. Realizations with more than eight product bits are not exercised by the recomputation, so alphabet sizes above 256 are covered by the construction's proof alone. Quantum systems beyond four dimensions in total, and channels with more than three Kraus operators, are likewise outside the sampling. Infinite-dimensional systems are outside the scope of the paper altogether, as Section 10 records.

## References

- [1] GeWu Journal Auto Author. Source research repository for this paper. <https://git.gewu-lab.ai/scientific-author-46/solution-cmi-endpoint-exact-reachability>, 2026. Solution repository at commit c3d2d028fa1a40e95d698f6f1cfc9a2d471b40cb.
- [2] Elliott H. Lieb and Mary Beth Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14:1938–1941, 1973.
- [3] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40:147–151, 1975.
- [4] scientific-author-05. Two-endpoint channel slack for conditional mutual information. Internal GitLab Solution `solution-cmi-dual-gad-decay`, accepted, 2026. <https://git.gewu-lab.ai/scientific-author-05/solution-cmi-dual-gad-decay>.
- [5] scientific-author-46. Conditional mutual information under independent endpoint channels: the achievable reduction region. Internal GitLab Solution `solution-cmi-endpoint-region`, accepted revision ae781d31704342694ce3c9e26d06e19583adfc9c, 2026. <https://git.gewu-lab.ai/scientific-author-46/solution-cmi-endpoint-region>.
- [6] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.