TIGHT ANCILLA BOUND FOR QUANTUM AUTOENCODERS IN STATE RECONSTRUCTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Quantum autoencoders offer a principled way to reduce the quantum resources required to store and process quantum data, but the resources needed for reliable reconstruction can be as important as the size of the compressed representation itself. In particular, understanding the fundamental ancilla requirements is essential for determining whether a compressed quantum model can be realized within a given hardware budget. Prior work established that a unitary decoder with n ancilla qubits is universally sufficient for compressing n-qubit states through a k-qubit bottleneck. It remained open, however, whether the full n-ancilla requirement can be *necessary* for arbitrary bottleneck sizes. We resolve this question in the affirmative. Thus, the universal n-ancilla upper bound is worst-case tight for every bottleneck size, revealing a symmetric resource requirement for the encoder and decoder in blind single-copy quantum compression.

1 INTRODUCTION

Quantum information processing is increasingly concerned not only with how quantum states are represented, but also with how much hardware is required to manipulate them (Vartiainen et al., 2004; Shende et al., 2005; Iten et al., 2016; Zhang et al., 2021; 2022; Yuan & Zhang, 2023; Zhang & Yuan, 2024). This distinction is particularly important for quantum data-processing architectures in which a large input system is mapped to a smaller latent register and subsequently reconstructed. A quantum autoencoder provides a natural framework for such a task: information is encoded into a lower-dimensional quantum system and a decoder attempts to recover the original state with high fidelity (Schumacher, 1995; Jozsa et al., 1998; Ding et al., 2019; Lamata et al., 2019; Pepper et al., 2019; Steinbrecher et al., 2019; Bondarenko & Feldmann, 2020; Huang et al., 2020; Bravo-Prieto, 2021; Cao & Wang, 2021; Locher et al., 2023; Ma et al., 2023; 2024). Since the decoder must recreate a state in the larger output space, compression does not automatically translate into an equally small implementation cost. The number of auxiliary qubits required by the decoder can therefore become a fundamental resource constraint, especially in architectures where qubit count, circuit width, or the size of an available environment is limited. The original variational formulation of quantum autoencoders was introduced by Romero et al. (Romero et al., 2017), while the general channel-theoretic language used to describe such compression and reconstruction is widely used in quantum information theory (Barnum, 2000; Horodecki, 2000; Nielsen et al., 2000; Barnum et al., 2001; Devetak & Berger, 2002; Hayden et al., 2002; Datta et al., 2012; Watrous et al., 2018; Wang et al., 2025).

This makes the decoder side of quantum compression a natural question of independent interest. It is tempting to regard the bottleneck dimension as the dominant resource, since the encoder deliberately replaces a larger system by a smaller one. From the standpoint of implementation, however, the reconstruction map has its own irreducible complexity. A decoder acting on a k-qubit latent system and returning an n-qubit state may require an environment of substantial dimension, even when the latent description itself is small. Establishing the relationship between compression ratio and decoder environment is therefore important for deciding whether a compressed architecture can actually be realized within a prescribed hardware budget.

Our work is primarily motivated by the recent analysis of blind, single-copy quantum compression in (Cha et al., 2026), which identifies a universal architecture in which k encoder ancillas and n decoder ancillas are sufficient for optimal compression of n-qubit states through a k-qubit bottleneck. It also

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

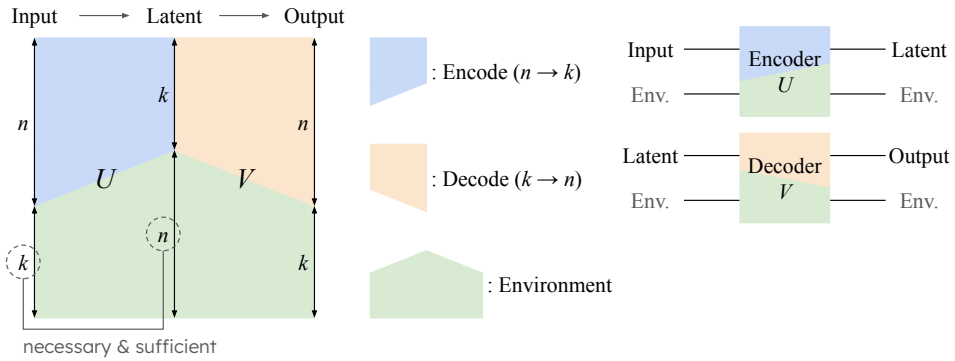

Figure 1: **Schematic of the unitary implementation of the encoder and decoder.** An n-qubit input state is compressed into a k-qubit bottleneck through an encoder unitary U, with the remaining degrees of freedom represented by the environment, and subsequently reconstructed by a decoder unitary V. For both the encoder and decoder, an $(n+k)$-qubit unitary is necessary and sufficient in the worst case.

demonstrates that the decoder requirement is genuinely nontrivial: restricting the decoder to an isometry, corresponding to fewer auxiliary qubits, can reduce the attainable performance, and the corresponding threshold was established tightly in the special case $k = 1$. The remaining question is whether this phenomenon reflects a more general obstruction. In other words, *can one construct source ensembles for which an optimal decoder must use the full n ancillary qubits, no matter how large the bottleneck k is?* Resolving this question is essential for determining whether the previously known universal decoder architecture is genuinely resource-optimal in the worst case.

We answer this question affirmatively by constructing, for every $n > k \geq 1$, a continuous family of pure input states to expose the structure of optimal reconstruction. The family is generated from spin-coherent states of a single qubit (Martin et al., 2010; Giraud et al., 2015; Denis & Martin, 2022). Such states are classical objects in quantum optics and the theory of generalized coherent states (Radcliffe, 1971; Arecchi et al., 1972; Perelomov, 1986; Zhang et al., 1990; Vourdas, 1996), and their connection with symmetric tensor-power spaces is well established (Harrow, 2013). The coherent-state geometry provides a highly structured family whose linear span is substantially larger than the number of parameters used to label the states. That combination gives enough symmetry to construct an optimal compression scheme while retaining enough rigidity to constrain every decoder that attains the optimum. The corresponding unitary implementation of the encoder and decoder is illustrated in Figure 1.

Decoder ancillas are directly tied to Kraus rank in the standard unitary implementation with a pure initialized environment (Shen et al., 2017; Singh & Datta, 2022). Consequently, a lower bound on the Kraus rank becomes a hardware lower bound on the number of decoder ancilla qubits. Our result gives a complete answer to the decoder-side question. In the ancilla-based realization, n decoder ancilla qubits are necessary for optimal performance for this family, while an n-ancilla decoder is sufficient. Thus the universal decoder resource bound established in (Cha et al., 2026) cannot in general be reduced, closing the remaining gap in the theory of blind single-copy quantum compression.

2 Notations and Overview

Throughout, fix integers $n > k \geq 1$ and write $d := 2^n$, $m := 2^k$, $\ell := m - 1$, and $s := 2\ell + 1 = 2^{k+1} - 1$. Thus $H_d \cong \mathbb{C}^d$ is the n-qubit input/output Hilbert space and $H_m \cong \mathbb{C}^m$ is the k-qubit latent Hilbert space. All Hilbert spaces are finite-dimensional and complex, and $\mathcal{L}(H)$ denotes the space of linear operators on H. We write $\mathrm{Tr}_B(X)$ for the partial trace over subsystem B. A quantum state is a positive semidefinite operator of unit trace, and a quantum channel is a completely positive trace-preserving (CPTP) linear map. If Φ is a channel, a Kraus representation is $\Phi(X) = \sum_i K_i X K_i^\dagger$ with $\sum_i K_i^\dagger K_i = I$, and $\mathrm{rank}_{\mathrm{Kraus}}(\Phi)$ denotes the minimum number of

nonzero Kraus operators in such a representation. Further quantum-information preliminaries are provided in Appendix A.

Our proof proceeds by exploiting a coherent-state source. Section 3 develops the coherent-vector identities and splitting isometry used to construct an explicit encoder–decoder pair in Section 4. We then prove in Section 5 a universal fidelity upper bound and analyze its equality conditions, which force every optimal encoder to realize the full latent coherent family. Finally, Section 6 converts the resulting pointwise constraints on decoder Kraus operators into a frame-type identity. Informational completeness and injectivity then force full latent rank, yielding the optimal decoder Kraus-rank lower bound and hence precisely the tight ancilla requirement.

3 Coherent vectors and the splitting isometry

For every nonnegative integer t, let

$$\mathcal{H}_t := \mathbb{C}^{t+1} \tag{1}$$

with orthonormal basis $|0\rangle_t, \ldots, |t\rangle_t$. The subscript will be omitted when no confusion is possible.

Definition 3.1 (Degree-t coherent vector)**.** For $u = (\alpha, \beta) \in \mathbb{C}^2$, define

$$|u; t\rangle := \sum_{a=0}^{t} \sqrt{\binom{t}{a}}\, \alpha^{t-a}\beta^a |a\rangle_t. \tag{2}$$

When $|\alpha|^2 + |\beta|^2 = 1$, this is a unit vector.

Indeed, the binomial theorem gives

$$\||u; t\rangle\|^2 = \sum_{a=0}^{t} \binom{t}{a} |\alpha|^{2(t-a)}\, |\beta|^{2a} = (|\alpha|^2 + |\beta|^2)^t. \tag{3}$$

The overlap of two coherent vectors is

$$\langle v; t \mid u; t\rangle = (\overline{v_0}u_0 + \overline{v_1}u_1)^t. \tag{4}$$

This follows from the same binomial expansion.

The space $\mathcal{H}_t$ is naturally isomorphic to the symmetric subspace of $(\mathbb{C}^2)^{\otimes t}$, and $|u; t\rangle$ corresponds to $|u\rangle^{\otimes t}$ (Bastin et al., 2009; Devi et al., 2012; Mandilara et al., 2014). We do not need to realize those t qubits physically. Only the $(t+1)$-dimensional space and the explicit coordinates in Eq. (2) are used.

3.1 The qubit-uniform measure

We use the following probability distribution on unit qubit vectors. Draw

$$x \sim \mathrm{Unif}[0, 1], \qquad \theta \sim \mathrm{Unif}[0, 2\pi), \tag{5}$$

independently, and set

$$|u(x, \theta)\rangle := \sqrt{1 - x}|0\rangle + e^{i\theta}\sqrt{x}|1\rangle. \tag{6}$$

This is a unitarily invariant distribution on pure qubit states (Zyczkowski & Sommers, 2001). All calculations below use only the integral

$$\int f(u)\, \mathrm{d}\nu(u) := \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(u(x, \theta))\, \mathrm{d}x\, \mathrm{d}\theta. \tag{7}$$

Lemma 3.2 (Integer beta integral)**.** *For nonnegative integers p, q,*

$$\int_0^1 x^p (1 - x)^q\, \mathrm{d}x = \frac{p!\, q!}{(p + q + 1)!}. \tag{8}$$

Lemma 3.3. *For every $t \geq 0$,*

$$\int |u; t\rangle\langle u; t|\, \mathrm{d}\nu(u) = \frac{I_{\mathcal{H}_t}}{t + 1}. \tag{9}$$

Consequently, the coherent vectors $\{|u; t\rangle\}$ span $\mathcal{H}_t$.

Proof. See Appendix B.1. □

3.2 TWO POLYNOMIAL FACTS

The equality analysis later uses polynomial division in the one-variable polynomial ring $\mathbb{C}[z]$. For polynomials f, g, the notation $f \mid g$ means that $g = fh$ for another polynomial h. Two polynomials are *coprime* if their only common divisors are nonzero constants. The field $\mathbb{C}(z)$ consists of ratios of polynomials and is used to express proportionality between polynomial vectors.

Lemma 3.4. *Let $P, Q \in \mathbb{C}[y, z]$ be nonzero polynomials satisfying*

$$P(y, z)Q(y, z) = (1 + yz)^N. \tag{10}$$

Then there are an integer $a \in \{0, \dots, N\}$ and a nonzero constant $c \in \mathbb{C}$ such that

$$P(y, z) = c(1 + yz)^a, \qquad Q(y, z) = c^{-1}(1 + yz)^{N-a}. \tag{11}$$

Proof. See Appendix B.3. □

The key linear map splits one degree-2ℓ coherent vector into a tensor product of two degree-ℓ coherent vectors.

Lemma 3.5 (Vandermonde's identity)**.** *For integers $r, s, q \geq 0$,*

$$\sum_{a+b=q} \binom{r}{a}\binom{s}{b} = \binom{r+s}{q}, \tag{12}$$

where binomial coefficients outside their usual ranges are interpreted as zero.

Define $\Gamma : \mathcal{H}_{2\ell} \to \mathcal{H}_\ell \otimes \mathcal{H}_\ell$ on the standard basis by

$$\Gamma|q\rangle_{2\ell} := \sum_{\substack{0 \leq a,b \leq \ell \\ a+b=q}} \sqrt{\frac{\binom{\ell}{a}\binom{\ell}{b}}{\binom{2\ell}{q}}} |a\rangle_\ell \otimes |b\rangle_\ell. \tag{13}$$

Proposition 3.6 (Coherent splitting)**.** *The map Γ in Eq. (13) is an isometry, and for every $u = (\alpha, \beta) \in \mathbb{C}^2$,*

$$\Gamma|u; 2\ell\rangle = |u; \ell\rangle \otimes |u; \ell\rangle. \tag{14}$$

Proof. See Appendix B.4. □

Let

$$\Pi := \Gamma\Gamma^\dagger \tag{15}$$

be the orthogonal projector onto the $(2\ell + 1)$-dimensional image of Γ inside $\mathcal{H}_\ell \otimes \mathcal{H}_\ell$.

Lemma 3.7. *With $m = \ell + 1$ and $s = 2\ell + 1$,*

$$\mathrm{Tr}_2(\Pi) = \frac{s}{m} I_{\mathcal{H}_\ell}. \tag{16}$$

Proof. See Appendix B.5. □

4 THE SOURCE, ENCODER, AND DECODER

Let $H_d \cong \mathbb{C}^d$ have computational basis $|0\rangle, \dots, |d-1\rangle$. Because $s = 2\ell + 1 \leq d$, define the isometric embedding

$$J : \mathcal{H}_{2\ell} \to H_d, \qquad J|a\rangle_{2\ell} = |a\rangle, \quad 0 \leq a \leq 2\ell. \tag{17}$$

Let

$$P := JJ^\dagger = \sum_{a=0}^{2\ell} |a\rangle\langle a| \tag{18}$$

be the source-support projector.

Definition 4.1. Draw (x, θ) as in Eq. (5), set $u = u(x, \theta)$, and output

$$|\Psi_u\rangle := J|u; 2\ell\rangle. \tag{19}$$

Equivalently,

$$|\Psi_{x,\theta}\rangle = \sum_{a=0}^{2m-2} \sqrt{\binom{2m-2}{a}} (1-x)^{(2m-2-a)/2} x^{a/2} e^{ia\theta} |a\rangle. \tag{20}$$

The induced distribution on pure states in H_d is denoted $\mu_{n,k}$.

Equation (20) makes the construction completely explicit in the computational basis. The source uses only the first $2m-1$ basis states of the d-dimensional n-qubit system.

By Lemma 3.3, the source average state is

$$\overline{\rho} := \int |\Psi_u\rangle\langle\Psi_u| \, \mathrm{d}\nu(u) = \frac{P}{s}. \tag{21}$$

4.1 The encoder

Identify the latent Hilbert space H_m with $\mathcal{H}_\ell$. For $b = 0, \ldots, \ell$, define

$$K_b := (I_{\mathcal{H}_\ell} \otimes \langle b|_\ell) \Gamma J^\dagger : H_d \to \mathcal{H}_\ell. \tag{22}$$

For each computational basis vector $|r\rangle$ with $s \le r \le d-1$, define

$$R_r := |0\rangle\langle r| : H_d \to \mathcal{H}_\ell. \tag{23}$$

Set

$$\mathcal{E}_\star(\rho) := \sum_{b=0}^{\ell} K_b \rho K_b^\dagger + \sum_{r=s}^{d-1} R_r \rho R_r^\dagger. \tag{24}$$

Proposition 4.2. *The map $\mathcal{E}_\star$ in Eq. (24) is a channel. It has a Kraus representation with*

$$m + (d-s) = d - m + 1 \le d \tag{25}$$

operators, and for every unit qubit vector u,

$$\mathcal{E}_\star(|\Psi_u\rangle\langle\Psi_u|) = |u; \ell\rangle\langle u; \ell|. \tag{26}$$

Proof. See Appendix B.6. □

4.2 The rank-m decoder

For $b = 0, \ldots, \ell$, define

$$A_b := \sqrt{\frac{m}{s}} \, J\Gamma^\dagger (I_{\mathcal{H}_\ell} \otimes |b\rangle_\ell) : \mathcal{H}_\ell \to H_d. \tag{27}$$

Set

$$\mathcal{D}_\star(\rho) := \sum_{b=0}^{\ell} A_b \rho A_b^\dagger. \tag{28}$$

Proposition 4.3. *The map $\mathcal{D}_\star$ in Eq. (28) is a channel with Kraus rank at most m. Together with $\mathcal{E}_\star$, it achieves the pointwise fidelity*

$$\langle\Psi_u|(\mathcal{D}_\star \circ \mathcal{E}_\star)(|\Psi_u\rangle\langle\Psi_u|)|\Psi_u\rangle = \frac{m}{s} \tag{29}$$

for every unit qubit vector u. Therefore

$$F_{\mu_{n,k}}(\mathcal{E}_\star, \mathcal{D}_\star) = \frac{m}{s}. \tag{30}$$

Proof. See Appendix B.7. □

5 OPTIMAL FIDELITY AND RIGIDITY OF OPTIMAL PAIRS

Theorem 5.1. *For the source distribution $\mu_{n,k}$,*

$$\max_{\mathcal{E},\mathcal{D}} F_{\mu_{n,k}}(\mathcal{E},\mathcal{D}) = \frac{m}{s} = \frac{2^k}{2^{k+1}-1}, \tag{31}$$

where the maximum is over all CPTP encoders $\mathcal{E} : \mathcal{L}(H_d) \to \mathcal{L}(H_m)$ and decoders $\mathcal{D} : \mathcal{L}(H_m) \to \mathcal{L}(H_d)$.

Proof. See Appendix B.8. □

Fix from now on an arbitrary optimal pair $(\mathcal{E},\mathcal{D})$ for $\mu_{n,k}$. Let

$$\mathcal{D}(X) = \sum_{i=1}^{q} A_i X A_i^\dagger \tag{32}$$

be a minimal Kraus representation, so $q = \mathrm{rank}_{\mathrm{Kraus}}(\mathcal{D})$. Continue to write $\sigma_u = \mathcal{E}(|\Psi_u\rangle\langle\Psi_u|)$.

5.1 IMPLICATIONS OF EQUALITY IN THE UPPER BOUND

Proposition 5.2. *For every optimal pair $(\mathcal{E},\mathcal{D})$, the following statements hold.*

1. *Every decoder Kraus operator has range in the source support:*
$$PA_i = A_i \quad \text{for every } i. \tag{33}$$
2. *For every unit qubit vector u and every i,*
$$A_i^\dagger |\Psi_u\rangle \in \ker(I_m - \sigma_u). \tag{34}$$
3. *There is a nonempty open set of pure qubit states on which σ_u is a pure state.*

Proof. See Appendix B.9. □

5.2 PURE MARGINAL STATES AND PRODUCT VECTORS

Lemma 5.3 (Nielsen et al. (2000))**.** *Let $|w\rangle \in X \otimes Y$ be a unit vector. Its reduced density operator $\mathrm{Tr}_Y(|w\rangle\langle w|)$ is pure if and only if $|w\rangle = |x\rangle \otimes |y\rangle$ for unit vectors $x \in X$, $y \in Y$.*

5.3 A COHERENT-PRODUCT FACTORIZATION THEOREM

For $z \in \mathbb{C}$, define the unnormalized affine coherent vector

$$|v_t(z)\rangle := \sum_{a=0}^{t} \sqrt{\binom{t}{a}}\, z^a |a\rangle_t. \tag{35}$$

It satisfies

$$\langle v_t(w) \mid v_t(z)\rangle = (1 + \overline{w}z)^t. \tag{36}$$

If $u = (1,z)/\sqrt{1+|z|^2}$, then

$$|u;t\rangle = (1+|z|^2)^{-t/2} |v_t(z)\rangle. \tag{37}$$

Proposition 5.4 (Factorization of an isometry on coherent vectors)**.** *Let $W : \mathcal{H}_N \to X \otimes Y$ be an isometry. Suppose $W|v_N(z)\rangle$ is a product vector for every $z \in \mathbb{C}$. Then there are an integer $a \in \{0,\dots,N\}$ and isometries*

$$U_a : \mathcal{H}_a \to X, \qquad V_{N-a} : \mathcal{H}_{N-a} \to Y \tag{38}$$

such that

$$W|v_N(z)\rangle = U_a|v_a(z)\rangle \otimes V_{N-a}|v_{N-a}(z)\rangle \tag{39}$$

for every $z \in \mathbb{C}$. Equivalently, for every unit qubit vector u,

$$W|u;N\rangle = U_a|u;a\rangle \otimes V_{N-a}|u;N-a\rangle. \tag{40}$$

Proof. See Appendix B.10. □

5.4 THE ENCODER MUST OUTPUT THE FULL DEGREE-ℓ COHERENT FAMILY

Proposition 5.5. *For every optimal pair $(\mathcal{E}, \mathcal{D})$, there is a unitary $U : \mathcal{H}_\ell \to H_m$ such that*

$$\mathcal{E}(|\Psi_u\rangle\langle\Psi_u|) = U|u;\ell\rangle\langle u;\ell|U^\dagger \tag{41}$$

for every unit qubit vector u. Consequently, after a fixed unitary change of latent basis, one may assume without loss of generality that

$$\sigma_u = |u;\ell\rangle\langle u;\ell| \quad \text{for every } u. \tag{42}$$

This basis change does not alter decoder Kraus rank or average fidelity.

Proof. See Appendix B.11. □

6 KRAUS RANK LOWER BOUNDS FOR OPTIMAL DECODERS AND THE SHARP DECODER-ANCILLA THEOREM

By Proposition 5.5, we now work in a latent basis satisfying Eq. (42). Define

$$B_i := A_i^\dagger J : \mathcal{H}_{2\ell} \to \mathcal{H}_\ell. \tag{43}$$

Then Eq. (34) says

$$B_i|u;2\ell\rangle \in \operatorname{span}\{|u;\ell\rangle\} \quad \text{for every } u. \tag{44}$$

Lemma 6.1. *Let $B : \mathcal{H}_{2\ell} \to \mathcal{H}_\ell$ be linear and suppose*

$$B|u;2\ell\rangle \in \operatorname{span}\{|u;\ell\rangle\}$$

for every unit qubit vector u. Then there is a vector $|\xi\rangle \in \mathcal{H}_\ell$ such that

$$B|u;2\ell\rangle = \langle \xi \mid u;\ell\rangle|u;\ell\rangle \tag{45}$$

for every unit qubit vector u.

Proof. See Appendix B.12. □

Apply Lemma 6.1 to every B_i. There are vectors $\xi_i \in \mathcal{H}_\ell$ such that

$$B_i|u;2\ell\rangle = \langle \xi_i \mid u;\ell\rangle|u;\ell\rangle. \tag{46}$$

Define

$$R := \sum_{i=1}^{q} |\xi_i\rangle\langle\xi_i| \in \mathcal{L}(\mathcal{H}_\ell). \tag{47}$$

Then

$$\operatorname{rank}(R) \le q. \tag{48}$$

We show that coherent-state expectation values determine an operator uniquely.

Lemma 6.2 (Informational completeness of coherent projectors)**.** *If $X \in \mathcal{L}(\mathcal{H}_\ell)$ satisfies*

$$\langle u;\ell|X|u;\ell\rangle = 0 \quad \text{for every unit qubit vector } u, \tag{49}$$

then $X = 0$.

Proof. See Appendix B.13. □

Define a linear map $\mathcal{T} : \mathcal{L}(\mathcal{H}_\ell) \to \mathcal{L}(\mathcal{H}_\ell)$ by

$$\mathcal{T}(X) := \int \langle u;\ell|X|u;\ell\rangle |u;\ell\rangle\langle u;\ell| \, \mathrm{d}\nu(u). \tag{50}$$

Lemma 6.3. *The map $\mathcal{T}$ is injective and*

$$\mathcal{T}(I_{\mathcal{H}_\ell}) = \frac{I_{\mathcal{H}_\ell}}{m}. \tag{51}$$

Proof. See Appendix B.14. □

Theorem 6.4. *Every decoder occurring in an optimal encoder–decoder pair for $\mu_{n,k}$ has Kraus rank at least*

$$\mathrm{rank}_{\mathrm{Kraus}}(\mathcal{D}) \geq m = 2^k. \tag{52}$$

The decoder $\mathcal{D}_\star$ has Kraus rank m.

Proof. See Appendix B.15. □

We can now state our main result in the circuit language of ancilla qubits.

Corollary 6.5. *For every pair of integers $n > k \geq 1$, the pure-state distribution $\mu_{n,k}$ in Eq. (20) has the following properties.*

1. *Its optimal average reconstruction fidelity through a k-qubit bottleneck is*

$$\max_{\mathcal{E},\mathcal{D}} F_{\mu_{n,k}}(\mathcal{E},\mathcal{D}) = \frac{2^k}{2^{k+1}-1}. \tag{53}$$

2. *Every decoder in every globally optimal encoder–decoder pair has Kraus rank at least 2^k.*

3. *In the unitary decoder architecture with a fixed pure initialized ancilla, every optimal decoder requires at least n decoder ancilla qubits. More explicitly, for every allowed integer $n_E < n$, the best fidelity achievable with n_E decoder ancillas is strictly smaller than the value in Eq. (53).*

4. *There is an optimal decoder using n decoder ancilla qubits. Moreover, the optimal encoder $\mathcal{E}_\star$ can be implemented with k encoder ancilla qubits.*

Corollary 6.6. *For blind single-copy compression of pure n-qubit states through a k-qubit bottleneck under average infidelity, n decoder ancilla qubits are universally sufficient and, in the worst case over source distributions, necessary.*

Remark. The ancilla count in Corollaries 6.5 and 6.6 refers to the standard implementation in which the appended ancilla begins in a fixed pure state, a joint unitary is applied, and an environment is discarded. By Lemma A.9, allowing an arbitrary fixed mixed initial ancilla cannot improve the globally optimized fidelity at a fixed ancilla count. Thus the same worst-case lower bound persists under that relaxation.

7 Low-dimensional examples

Example 7.1 (Two input qubits compressed to one qubit)**.** Take $n = 2$, $k = 1$. Then

$$d = 4, \qquad m = 2, \qquad \ell = 1, \qquad N = 2, \qquad s = 3.$$

The source states are

$$|\Psi_{x,\theta}\rangle = (1-x)|0\rangle + \sqrt{2x(1-x)}e^{i\theta}|1\rangle + xe^{2i\theta}|2\rangle,$$

embedded in a three-dimensional subspace of the four-dimensional two-qubit Hilbert space. The optimal fidelity is

$$\frac{m}{s} = \frac{2}{3}.$$

Every optimal decoder has Kraus rank at least two. A one-ancilla decoder is isometric because it leaves no output environment. Its Kraus rank is one and it is therefore strictly suboptimal. Two decoder ancillas are necessary and sufficient.

This source is different from the equatorial phase family used for the $k = 1$ decoder counterexample in (Cha et al., 2026). Here the parameter u ranges over all pure qubit directions. This is what makes the construction and proof extend uniformly to every k.

Example 7.2 (Three input qubits compressed to two qubits)**.** Take $n = 3$, $k = 2$. Then

$$d = 8, \qquad m = 4, \qquad \ell = 3, \qquad N = 6, \qquad s = 7.$$

The source occupies the first seven computational basis states of an eight-dimensional three-qubit space, and

$$|\Psi_{x,\theta}\rangle = \sum_{a=0}^{6} \sqrt{\binom{6}{a}} (1-x)^{(6-a)/2} x^{a/2} e^{ia\theta} |a\rangle.$$

The optimum is $4/7$, and every optimal decoder has Kraus rank at least four. With two decoder ancillas, the discarded environment has dimension

$$2^{k+n_E-n} = 2^{2+2-3} = 2,$$

so the decoder rank is at most two and cannot be optimal. Three decoder ancillas give environment dimension four and are sufficient.

8 Limitations

Our results provide a rigorous theoretical baseline, while leaving several important directions open for future work. First, the lower bound is a worst-case existence result based on a specially constructed source ensemble, and we do not analyze other physically relevant source distributions. Second, our characterization concerns global optimality under average infidelity and does not provide quantitative guarantees for approximate reconstruction or alternative distortion measures. Third, we optimize over unrestricted noiseless CPTP encoders and decoders rather than practical parameterized circuits, and our ancilla lower bound characterizes circuit width and environment dimension rather than gate count, depth, connectivity, or other implementation costs relevant to near-term quantum hardware.

9 Conclusion

In this work, we established a sharp characterization of the ancillary resources required for blind single-copy quantum compression through a smaller quantum bottleneck. Our main result shows that the apparent asymmetry between compression and reconstruction can, at the level of worst-case unitary implementation, be completely resolved: the same total circuit width is required on both sides of the autoencoder.

On the decoder side, we proved this statement by constructing, for every $n > k$, a continuous family of pure states for which every globally optimal decoder must have sufficiently large Kraus rank to require all n decoder ancilla qubits. At the same time, an optimal decoder with n ancillas exists. Together with the encoder-side result, this gives a symmetric and sharp resource characterization.

Another important aspect of the result is that the obstruction is not simply a consequence of the source occupying a large Hilbert-space support. The source family is described by only a small number of parameters, yet its structured coherent-state geometry imposes strong constraints on every optimal encoder–decoder pair. This illustrates a broader principle that low-dimensional parametrization of quantum data does not by itself guarantee low implementation complexity. The geometry of the family and the way information must be distributed between the latent system and the discarded degrees of freedom can determine the fundamental hardware requirements.

These results also suggest several promising directions for future work. A natural next step is to determine how the $n + k$-qubit threshold changes when the objective is relaxed from exact global optimality to an approximate fidelity target. It would also be valuable to characterize which practically relevant source ensembles permit substantially smaller encoder or decoder unitaries, and to identify structural properties of those ensembles that make such reductions possible. Extending the analysis to mixed-state sources, multiple copies, noisy channels, or alternative distortion measures could reveal whether the symmetry between encoder and decoder persists more broadly. Finally, connecting these worst-case lower bounds to circuit depth, gate complexity, connectivity constraints, and other realistic hardware limitations may help translate the present information-theoretic limits into concrete design principles for near-term quantum autoencoders.

AI USE STATEMENT

In this work, we used generative AI tools for grammar checking, correctness checking, drafting sections, and identifying relevant prior theoretical work. We have not used generative AI tools for the main proof mechanism and figure generation, and data collection and experimental design are not applicable to this work. We have reviewed all AI-assisted work. For example, we verified the correctness and applicability of some previously established lemmas identified by AI as potentially useful for our proof. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

This work is theoretical and does not rely on numerical experiments or external datasets. All assumptions, source distributions, channels, and optimization objectives used in our results are explicitly defined in the text. Complete proofs of the results are provided in the appendix.

REFERENCES

Anurag Anshu, Debbie Leung, and Dave Touchette. Incompressibility of classical distributions. *IEEE Transactions on Information Theory*, 68(3):1758–1771, 2021.

Fortunato Tito Arecchi, Eric Courtens, Robert Gilmore, and Harry Thomas. Atomic coherent states in quantum optics. *Physical Review A*, 6(6):2211, 1972.

Howard Barnum. Quantum rate-distortion coding. *Physical Review A*, 62(4):042309, 2000.

Howard Barnum, Christopher A Fuchs, Richard Jozsa, and Benjamin Schumacher. General fidelity limit for quantum channels. *Physical Review A*, 54(6):4707, 1996.

Howard Barnum, Carlton M Caves, Christopher A Fuchs, Richard Jozsa, and Benjamin Schumacher. On quantum coding for ensembles of mixed states. *Journal of Physics A: Mathematical and General*, 34(35):6767–6785, 2001.

Thierry Bastin, Stéphanie Krins, Pierre Mathonet, Michel Godefroid, Lucas Lamata, and Enrique Solano. Operational families of entanglement classes for symmetric n-qubit states. *Physical Review Letters*, 103(7):070503, 2009.

Fabio Benatti and Roberto Floreanini. Open quantum dynamics: complete positivity and entanglement. *International Journal of Modern Physics B*, 19(19):3063–3139, 2005.

Fabio Benatti, R Floreanini, and R Romano. Complete positivity and entangled degrees of freedom. *Journal of Physics A: Mathematical and General*, 35(23):4955–4972, 2002.

Andreas Bluhm, Lukas Rauber, and Michael M Wolf. Quantum compression relative to a set of measurements. In *Annales Henri Poincare*, volume 19, pp. 1891–1937. Springer, 2018.

Yuri Bolshakov, Cornelis VM van der Mee, André CM Ran, Boris Reichstein, and Leiba Rodman. Extension of isometries in finite-dimensional indefinite scalar product spaces and polar decompositions. *SIAM Journal on Matrix Analysis and Applications*, 18(3):752–774, 1997.

Dmytro Bondarenko and Polina Feldmann. Quantum autoencoders to denoise quantum data. *Physical Review Letters*, 124(13):130502, 2020.

Carlos Bravo-Prieto. Quantum autoencoders with enhanced data encoding. *Machine Learning: Science and Technology*, 2(3):035028, 2021.

Chenfeng Cao and Xin Wang. Noise-assisted quantum autoencoder. *Physical Review Applied*, 15 (5):054012, 2021.

Hyunho Cha, Chae-Yeun Park, and Jungwoo Lee. Toward the Goldilocks blind compression of quantum states. *arXiv preprint arXiv:2605.01258*, 2026.

Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and Its Applications*, 10(3):285–290, 1975.

Nilanjana Datta, Min-Hsiu Hsieh, and Mark M Wilde. Quantum rate distortion, reverse shannon theorems, and source-channel separation. *IEEE Transactions on Information Theory*, 59(1):615–630, 2012.

Nilanjana Datta, Joseph M Renes, Renato Renner, and Mark M Wilde. One-shot lossy quantum data compression. *IEEE Transactions on Information Theory*, 59(12):8057–8076, 2013.

Jérôme Denis and John Martin. Extreme depolarization for any spin. *Physical Review Research*, 4 (1):013178, 2022.

Igor Devetak and Toby Berger. Quantum rate-distortion theory for memoryless sources. *IEEE Transactions on Information Theory*, 48(6):1580–1589, 2002.

AR Usha Devi, Sudha, and AK Rajagopal. Majorana representation of symmetric multiqubit states. *Quantum Information Processing*, 11(3):685–710, 2012.

Yongcheng Ding, Lucas Lamata, Mikel Sanz, Xi Chen, and Enrique Solano. Experimental implementation of a quantum autoencoder via quantum adders. *Advanced Quantum Technologies*, 2 (7-8):1800065, 2019.

Qingxiuxiong Dong, Marco Túlio Quintino, Akihito Soeda, and Mio Murao. Implementing positive maps with multiple copies of an input state. *Physical Review A*, 99(5):052352, 2019.

Olivier Giraud, Daniel Braun, Dorian Baguette, Thierry Bastin, and John Martin. Tensor representation of spin states. *Physical Review Letters*, 114(8):080401, 2015.

Shima Bab Hadiashar, Ashwin Nayak, and Renato Renner. Communication complexity of one-shot remote state preparation. *IEEE Transactions on Information Theory*, 64(7):4709–4728, 2018.

Aram W Harrow. The church of the symmetric subspace. *arXiv preprint arXiv:1308.6595*, 2013.

Masahito Hayashi and Keiji Matsumoto. Quantum universal variable-length source coding. *Physical Review A*, 66(2):022311, 2002.

Patrick Hayden, Richard Jozsa, and Andreas Winter. Trading quantum for classical resources in quantum data compression. *Journal of Mathematical Physics*, 43(9):4404–4444, 2002.

Michał Horodecki. Limits for compression of quantum information carried by ensembles of mixed states. *Physical Review A*, 57(5):3364, 1998.

Michal Horodecki. Optimal compression for mixed signal states. *Physical Review A*, 61(5):052309, 2000.

Chang-Jiang Huang, Hailan Ma, Qi Yin, Jun-Feng Tang, Daoyi Dong, Chunlin Chen, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. Realization of a quantum autoencoder for lossless compression of quantum data. *Physical Review A*, 102(3):032412, 2020.

Raban Iten, Roger Colbeck, Ivan Kukuljan, Jonathan Home, and Matthias Christandl. Quantum circuits for isometries. *Physical Review A*, 93(3):032318, 2016.

Richard Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

Richard Jozsa, Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Universal quantum information compression. *Physical Review Letters*, 81(8):1714, 1998.

Masato Koashi and Nobuyuki Imoto. Compressibility of quantum mixed-state signals. *Physical Review Letters*, 87(1):017902, 2001.

Gerhard Kramer and Serap A Savari. Quantum data compression with commuting density operators. In *Proceedings IEEE International Symposium on Information Theory,*, pp. 44. IEEE, 2002.

Karl Kraus. General state changes in quantum theory. *Annals of Physics*, 64(2):311–335, 1971.

Dennis Kretschmann, Dirk Schlingemann, and Reinhard F Werner. The information-disturbance tradeoff and the continuity of Stinespring's representation. *IEEE Transactions on Information Theory*, 54(4):1708–1717, 2008.

Kohdai Kuroiwa and Debbie Leung. Rate reduction of blind quantum data compression with local approximations based on the unstable structure of quantum states. *Physical Review A*, 107(5): 052421, 2023.

Lucas Lamata, Unai Alvarez-Rodriguez, Jose David Martin-Guerrero, Mikel Sanz, and Enrique Solano. Quantum autoencoders via quantum adders with genetic algorithms. *Quantum Science and Technology*, 4(1):014007, 2019.

David F Locher, Lorenzo Cardarelli, and Markus Müller. Quantum error correction with quantum autoencoders. *Quantum*, 7:942, 2023.

Hailan Ma, Chang-Jiang Huang, Chunlin Chen, Daoyi Dong, Yuanlong Wang, Re-Bing Wu, and Guo-Yong Xiang. On compression rate of quantum autoencoders: Control design, numerical and experimental realization. *Automatica*, 147:110659, 2023.

Hailan Ma, Gary J Mooney, Ian R Petersen, Lloyd CL Hollenberg, and Daoyi Dong. Quantum autoencoders using mixed reference states. *npj Quantum Information*, 10(1):86, 2024.

A Mandilara, T Coudreau, A Keller, and P Milman. Entanglement classification of pure symmetric states via spin coherent states. *Physical Review A*, 90(5):050302, 2014.

John Martin, Olivier Giraud, PA Braun, Daniel Braun, and Thierry Bastin. Multiqubit symmetric states with high geometric entanglement. *Physical Review A*, 81(6):062347, 2010.

Michael A Nielsen, Isaac L Chuang, et al. *Quantum computation and quantum information*, volume 1. Cambridge University Press, 2000.

Alex Pepper, Nora Tischler, and Geoff J Pryde. Experimental realization of a quantum autoencoder: The compression of qutrits via machine learning. *Physical Review Letters*, 122(6):060501, 2019.

Askold Perelomov. *Generalized Coherent States and Their Applications*. Theoretical and Mathematical Physics. Springer-Verlag, Berlin, Heidelberg, 1986.

J Michael Radcliffe. Some properties of coherent spin states. *Journal of Physics A: General Physics*, 4(3):313–323, 1971.

Jonathan Romero, Jonathan P Olson, and Alan Aspuru-Guzik. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2(4):045001, 2017.

Mary Beth Ruskai. Some open problems in quantum information theory. *arXiv preprint arXiv:0708.1902*, 2007.

Pablo Sala, Sarang Gopalakrishnan, Masaki Oshikawa, and Yizhi You. Spontaneous strong symmetry breaking in open systems: Purification perspective. *Physical Review B*, 110(15):155150, 2024.

Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995.

Benjamin Schumacher and Michael A Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629, 1996.

Chao Shen, Kyungjoo Noh, Victor V Albert, Stefan Krastanov, Michel H Devoret, Robert J Schoelkopf, SM Girvin, and Liang Jiang. Quantum channel construction with circuit quantum electrodynamics. *Physical Review B*, 95(13):134501, 2017.

Vivek V Shende, Stephen S Bullock, and Igor L Markov. Synthesis of quantum logic circuits. In *Proceedings of the 2005 Asia and South Pacific Design Automation Conference*, pp. 272–275, 2005.

Satvik Singh and Nilanjana Datta. Detecting positive quantum capacities of quantum channels. *npj Quantum Information*, 8(1):50, 2022.

Gregory R Steinbrecher, Jonathan P Olson, Dirk Englund, and Jacques Carolan. Quantum optical neural networks. *npj Quantum Information*, 5(1):60, 2019.

W Forrest Stinespring. Positive functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

Barbara M Terhal, Isaac L Chuang, David P DiVincenzo, Markus Grassl, and John A Smolin. Simulating quantum operations with mixed environments. *Physical Review A*, 60(2):881, 1999.

Juha J Vartiainen, Mikko Möttönen, and Martti M Salomaa. Efficient decomposition of quantum gates. *Physical Review Letters*, 92(17):177902, 2004.

A Vourdas. Resolutions of the identity in terms of SU(2) coherent states and their use for quantum-state engineering. *Physical Review A*, 54(5):4544, 1996.

Gaoyuan Wang, Jonathan Warrell, Prashant S Emani, and Mark Gerstein. Quantum variational autoencoder utilizing regularized mixed-state latent representations. *Physical Review A*, 111(4): 042416, 2025.

John Watrous et al. *The theory of quantum information*, volume 1. Cambridge University Press, 2018.

Pei Yuan and Shengyu Zhang. Optimal (controlled) quantum state preparation and improved unitary synthesis by quantum circuits with any number of ancillary qubits. *Quantum*, 7:956, 2023.

Christof Zalka and Eleanor Rieffel. Quantum operations that cannot be implemented using a small mixed environment. *Journal of Mathematical Physics*, 43(9):4376–4381, 2002.

Wei-Min Zhang, Robert Gilmore, et al. Coherent states: Theory and some applications. *Reviews of Modern Physics*, 62(4):867, 1990.

Xiao-Ming Zhang and Xiao Yuan. Circuit complexity of quantum access models for encoding classical data. *npj Quantum Information*, 10(1):42, 2024.

Xiao-Ming Zhang, Man-Hong Yung, and Xiao Yuan. Low-depth quantum state preparation. *Physical Review Research*, 3(4):043200, 2021.

Xiao-Ming Zhang, Tongyang Li, and Xiao Yuan. Quantum state preparation with optimal circuit depth: Implementations and applications. *Physical Review Letters*, 129(23):230504, 2022.

Karol Zyczkowski and Hans-Jürgen Sommers. Induced measures in the space of mixed quantum states. *Journal of Physics A: Mathematical and General*, 34(35):7111–7125, 2001.

A Preliminaries

A.1 Hilbert spaces, operators, and states

All Hilbert spaces in this work are finite-dimensional complex inner-product spaces. The inner product $\langle x \mid y \rangle$ is conjugate-linear in x and linear in y. For a Hilbert space H, let $\mathcal{L}(H)$ denote the vector space of linear operators from H to itself. The adjoint of X is denoted $X^\dagger$.

Definition A.1. Let $X : H \to K$ be linear.

- The *image* $\mathrm{im}(X)$ is the set of vectors Xv, the *kernel* $\ker(X)$ is the set of vectors sent to zero, and the *rank* $\mathrm{rank}(X)$ is $\dim \mathrm{im}(X)$.
- The identity operator on H is I_H. The trace of $Y \in \mathcal{L}(H)$ is the sum of its diagonal entries in any orthonormal basis. It is denoted $\mathrm{Tr}(Y)$.
- An operator $P \in \mathcal{L}(H)$ is an *orthogonal projector* if $P = P^\dagger = P^2$. It projects onto $\mathrm{im}(P)$.

- A map $V : H \to K$ is an *isometry* if $V^\dagger V = I_H$. It preserves inner products and norms. A map $U : H \to K$ is *unitary* if $U^\dagger U = I_H$ and $UU^\dagger = I_K$. In finite dimensions this requires $\dim H = \dim K$.
- For a positive semidefinite operator Y, its *support* is the span of the eigenvectors with nonzero eigenvalues. Saying that Y is supported in a projector P means $Y = PYP$.

A qubit has Hilbert space $\mathbb{C}^2$. An n-qubit register has Hilbert space $(\mathbb{C}^2)^{\otimes n}$, of dimension 2^n, with computational basis $|0\rangle, \ldots, |2^n - 1\rangle$. A quantum gate acting on the whole register is represented by a unitary operator.

Definition A.2. An operator $X \in \mathcal{L}(H)$ is *positive semidefinite*, written $X \geq 0$, when

$$\langle v|X|v\rangle \geq 0 \quad \text{for every } |v\rangle \in H.$$

A *density operator* is an operator $\rho \geq 0$ with $\mathrm{Tr}(\rho) = 1$. A unit vector $|\psi\rangle$ represents the pure state $|\psi\rangle\langle\psi|$. A density operator of rank greater than one is called mixed. For Hermitian operators X, Y, the notation $X \leq Y$ means that $Y - X \geq 0$.

A composite system with spaces H_A and H_B has space $H_A \otimes H_B$. If $\{|j\rangle_B\}$ is an orthonormal basis of H_B, the partial trace over B is

$$\mathrm{Tr}_B(X) := \sum_j (I_A \otimes \langle j|_B) X (I_A \otimes |j\rangle_B). \tag{54}$$

It is independent of the chosen orthonormal basis. Operationally, Tr_B means that subsystem B is discarded.

Fact A.3. *Let $X, X_1, \ldots, X_r$ be positive semidefinite operators and let P be an orthogonal projector.*

1. *If $\mathrm{Tr}(X) = 0$, then $X = 0$.*
2. *If $\mathrm{Tr}((I - P)X) = 0$, then $X = PXP$.*
3. *If $X_1 + \cdots + X_r = 0$, then every $X_i = 0$.*
4. *If $\langle v|X|v\rangle = 0$, then $X|v\rangle = 0$.*

A.2 Completely positive and trace-preserving maps

Definition A.4 (Quantum channel)**.** A linear map $\Phi : \mathcal{L}(H_A) \to \mathcal{L}(H_B)$ is

- *trace preserving* if $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$ for every $X \in \mathcal{L}(H_A)$;
- *completely positive* if, for every finite-dimensional auxiliary space H_R, the map $\Phi \otimes \mathrm{id}_R$ sends positive semidefinite operators to positive semidefinite operators.

A *quantum channel* is a completely positive and trace-preserving map, abbreviated CPTP.

Complete positivity rather than ordinary positivity is required because the input system may be entangled with an external reference system (Schumacher & Nielsen, 1996; Benatti et al., 2002; Benatti & Floreanini, 2005; Dong et al., 2019). The finite-dimensional structure theorem below is due to Stinespring, Kraus, and Choi (Stinespring, 1955; Kraus, 1971; Choi, 1975).

Fix orthonormal bases $\{|a\rangle_A\}_{a=1}^{d_A}$ and $\{|b\rangle_B\}_{b=1}^{d_B}$. For an operator $K : H_A \to H_B$, define its vectorization by

$$|K\rangle\!\rangle := \sum_{a=1}^{d_A} |a\rangle_A \otimes K|a\rangle_A \in H_A \otimes H_B. \tag{55}$$

For a linear map $\Phi : \mathcal{L}(H_A) \to \mathcal{L}(H_B)$, define its Choi matrix by

$$J(\Phi) := \sum_{a,a'=1}^{d_A} |a\rangle\langle a'|_A \otimes \Phi(|a\rangle\langle a'|_A). \tag{56}$$

Theorem A.5 (Kraus, Choi, and Stinespring representations (Stinespring, 1955; Kraus, 1971; Choi, 1975))**.** *For a linear map $\Phi : \mathcal{L}(H_A) \to \mathcal{L}(H_B)$, the following are equivalent.*

1. *Φ is completely positive.*
2. *$J(\Phi) \geq 0$.*
3. *There are operators $K_1, \ldots, K_r : H_A \to H_B$ such that*

$$\Phi(X) = \sum_{i=1}^{r} K_i X K_i^\dagger. \tag{57}$$

The map is trace preserving when

$$\sum_{i=1}^{r} K_i^\dagger K_i = I_A. \tag{58}$$

The smallest possible number of nonzero operators in Eq. (57) *equals* $\operatorname{rank} J(\Phi)$*. This number is the* Kraus rank $\operatorname{rank}_{\mathrm{Kraus}}(\Phi)$.

Equivalently, Φ is CPTP when there are a Hilbert space H_E and an isometry $V : H_A \to H_B \otimes H_E$ such that

$$\Phi(X) = \operatorname{Tr}_E(VXV^\dagger). \tag{59}$$

The smallest possible value of $\dim H_E$ *is* $\operatorname{rank}_{\mathrm{Kraus}}(\Phi)$.

Lemma A.6 (Bolshakov et al. (1997))**.** *Let S be a subspace of a finite-dimensional Hilbert space H, and let $W : S \to H'$ be an isometry. If $\dim H = \dim H'$, then there is a unitary $U : H \to H'$ whose restriction to S is W.*

A.3 Blind single-copy compression and average fidelity

Definition A.7 (Compression problem)**.** Let μ be a probability distribution on unit vectors in H_d. An encoder and decoder through an m-dimensional bottleneck are channels

$$\mathcal{E} : \mathcal{L}(H_d) \to \mathcal{L}(H_m), \qquad \mathcal{D} : \mathcal{L}(H_m) \to \mathcal{L}(H_d).$$

Their average reconstruction fidelity is

$$F_\mu(\mathcal{E}, \mathcal{D}) := \mathbb{E}_{\psi \sim \mu}[\langle\psi|(\mathcal{D} \circ \mathcal{E})(|\psi\rangle\langle\psi|)|\psi\rangle]\,. \tag{60}$$

The corresponding average infidelity is

$$L_\mu(\mathcal{E}, \mathcal{D}) := 1 - F_\mu(\mathcal{E}, \mathcal{D}). \tag{61}$$

Thus maximizing fidelity and minimizing infidelity are the same optimization problem.

The word *blind* means that the encoder is not told the classical label of the sampled state (Horodecki, 1998; Koashi & Imoto, 2001; Hayashi & Matsumoto, 2002; Kramer & Savari, 2002; Anshu et al., 2021; Kuroiwa & Leung, 2023). It must apply one fixed channel to every input. The phrase *single-copy* means that only one copy of the sampled state is supplied (Datta et al., 2013; Bluhm et al., 2018; Hadiashar et al., 2018). For a pure target $|\psi\rangle\langle\psi|$, the quantity in brackets in Eq. (60) is the usual fidelity with the reconstructed density operator (Jozsa, 1994; Barnum et al., 1996).

A.4 Decoder ancillas and Kraus rank

An *ancilla* is an auxiliary register introduced to implement a channel and initialized independently of the input (Terhal et al., 1999; Zalka & Rieffel, 2002; Kretschmann et al., 2008; Sala et al., 2024). We write $|0^r\rangle := |0\rangle^{\otimes r}$. The decoder circuit begins with a k-qubit latent register C, appends an n_E-qubit ancilla register E in the fixed pure state $|0^{n_E}\rangle$, applies a unitary, keeps an n-qubit output register F, and discards the remaining register G. Dimension conservation gives

$$\dim H_G = 2^{k+n_E-n}. \tag{62}$$

The architecture is possible only when $n_E \geq n - k$.

Observation A.8. *Consider a decoder implemented as*

$$\mathcal{D}(\rho) = \mathrm{Tr}_G\big[U(\rho_C \otimes |0^{n_E}\rangle\langle 0^{n_E}|_E)U^\dagger\big], \tag{63}$$

where $U : H_C \otimes H_E \to H_F \otimes H_G$ *is unitary,* $\dim H_C = m = 2^k$*, and* $\dim H_F = d = 2^n$*. Then*

$$\mathrm{rank}_{\mathrm{Kraus}}(\mathcal{D}) \le 2^{k+n_E-n}. \tag{64}$$

Conversely, every channel $\mathcal{D} : \mathcal{L}(H_m) \to \mathcal{L}(H_d)$ *of Kraus rank at most* m *can be implemented in this form with* $n_E = n$ *decoder ancilla qubits.*

Proof. Choose an orthonormal basis $\{|j\rangle_G\}_{j=1}^{\dim H_G}$ and define

$$L_j := (I_F \otimes \langle j|_G)U(I_C \otimes |0^{n_E}\rangle_E).$$

Expanding the partial trace in Eq. (63) gives

$$\mathcal{D}(\rho) = \sum_{j=1}^{\dim H_G} L_j \rho L_j^\dagger.$$

Thus the decoder has a Kraus representation with $\dim H_G = 2^{k+n_E-n}$ terms, proving Eq. (64).

Conversely, suppose $\mathcal{D}$ has Kraus operators $L_1, \dots, L_q$, where $q \le m$. Pad the list with zeros to length m, and let $H_G \cong \mathbb{C}^m$. The Stinespring map

$$\widetilde{U}|x\rangle := \sum_{j=1}^{m} L_j|x\rangle \otimes |j\rangle_G$$

is an isometry $H_C \to H_F \otimes H_G$ by Theorem A.5. When $n_E = n$, the two ambient spaces $H_C \otimes H_E$ and $H_F \otimes H_G$ both have dimension md. Identify H_C with the subspace $H_C \otimes \mathrm{span}\{|0^n\rangle_E\}$, and extend $\widetilde{U}$ to a unitary by Lemma A.6. The resulting circuit realizes $\mathcal{D}$. □

Lemma A.9 (Cha et al. (2026), Lemma 3.2)**.** *Fix all register dimensions and the average-fidelity objective in Eq.* (60)*. If the decoder ancilla in Eq.* (63) *is allowed to start in an arbitrary fixed density operator* τ_E*, then the maximum over the unitary and* τ_E *is attained with* τ_E *pure. The same statement holds for the encoder ancilla.*

A.5 Extreme channels and universal decoder sufficiency

The main purpose of this work is to construct a source for which decoder Kraus rank m is necessary. To conclude that n decoder ancillas are the exact *universal* threshold, we also need the complementary fact that some optimal decoder of rank at most m always exists.

Definition A.10. A subset K of a real vector space is *convex* if

$$tx + (1-t)y \in K \quad \text{whenever } x, y \in K \text{ and } 0 \le t \le 1.$$

A point $x \in K$ is an *extreme point* of K if every identity

$$x = ty + (1-t)z, \qquad y, z \in K, \qquad 0 < t < 1,$$

forces $y = z = x$.

Lemma A.11 (Bauer maximum principle)**.** *Let* K *be a nonempty compact convex subset of a finite-dimensional real vector space, and let* f *be a real-valued linear functional. Then* f *attains its maximum at an extreme point of* K*.*

Theorem A.12 (Ruskai (2007), Theorem 1)**.** *Let* H_A *have dimension* m*, and let* $\Phi : \mathcal{L}(H_A) \to \mathcal{L}(H_B)$ *be an extreme point of the convex set of CPTP maps. Then*

$$\mathrm{rank}_{\mathrm{Kraus}}(\Phi) \le m.$$

By Lemma A.11, the infidelity objective, being linear in the decoder channel, attains its optimum at an extreme CPTP map. Theorem A.12 then implies that such an extreme decoder has Kraus rank at most m. Hence an optimal decoder of Kraus rank at most m always exists, establishing the required sufficiency (Cha et al., 2026).

B DEFERRED PROOFS

B.1 PROOF OF LEMMA 3.3

Proof. The (a, b) matrix entry of the integral in the basis of $\mathcal{H}_t$ equals

$$\frac{\sqrt{\binom{t}{a}\binom{t}{b}}}{2\pi}\int_0^{2\pi} e^{i(a-b)\theta}\,\mathrm{d}\theta\int_0^1 (1-x)^{t-(a+b)/2}x^{(a+b)/2}\,\mathrm{d}x.$$

The phase integral is zero unless $a = b$. For $a = b$, Lemma 3.2 gives

$$\binom{t}{a}\int_0^1 (1-x)^{t-a}x^a\,\mathrm{d}x = \frac{t!}{a!(t-a)!}\frac{a!(t-a)!}{(t+1)!} = \frac{1}{t+1}.$$

Thus the integral is $I_{\mathcal{H}_t}/(t+1)$.

If a vector $v \in \mathcal{H}_t$ were orthogonal to every coherent vector, then

$$0 = \int |\langle v \mid u; t\rangle|^2\,\mathrm{d}\nu(u) = \frac{\|v\|^2}{t+1},$$

so $v = 0$. Therefore the coherent vectors span $\mathcal{H}_t$. □

B.2 POLYNOMIAL RANK-ONE FACTORIZATION

Lemma B.1 (Euclid's divisibility lemma for one-variable polynomials)**.** *If $a, b, p \in \mathbb{C}[z]$, the polynomials a and b are coprime, and $b \mid ap$, then $b \mid p$.*

Lemma B.2. *Let $M(z)$ be an $r \times q$ matrix with entries in $\mathbb{C}[z]$. Suppose M is not the zero matrix and every 2×2 minor of $M(z)$ is the zero polynomial. Then there are polynomial column vectors*

$$p(z) \in \mathbb{C}[z]^r, \qquad q(z) \in \mathbb{C}[z]^q$$

such that

$$M(z) = p(z)q(z)^{\mathsf{T}}. \tag{65}$$

Proof. Choose a nonzero column $c(z)$ of $M(z)$. Let $g(z)$ be a greatest common divisor of the entries of $c(z)$, and write $c(z) = g(z)p(z)$, where the entries of $p(z)$ have no nonconstant common divisor.

Let $d(z)$ be any other column. Choose an index j for which the j-th entry of c is nonzero. Over the field $\mathbb{C}(z)$, set $\lambda = d_j/c_j$. The vanishing of the minor formed from rows j, r and the two columns says $c_j d_r - c_r d_j = 0$, hence $d_r = \lambda c_r$ for every row r. Since $c = gp$, we may write $d = h(z)p(z)$ for some $h \in \mathbb{C}(z)$.

Write $h = a/b$ with coprime polynomials a, b. Because every entry of $d = (a/b)p$ is a polynomial, b divides every product ap_r. By Lemma B.1, $b \mid p_r$ for every r. The entries of p have no nonconstant common divisor, so b is constant. Hence h is a polynomial. Applying this argument to each column gives polynomial coefficients $q_j(z)$ and therefore Eq. (65). □

B.3 PROOF OF LEMMA 3.4

Proof. Let $F := \mathbb{C}(y)$, the field of rational functions in y, and regard all polynomials as elements of the one-variable ring $F[z]$. Put $L(z) := 1 + yz$. We first show by induction on N that, in $F[z]$,

$$P = a(y)L^r, \qquad Q = b(y)L^{N-r}$$

for some $r \in \{0, \ldots, N\}$ and nonzero $a(y), b(y) \in F$.

For $N = 0$, the identity $PQ = 1$ forces both polynomials to have degree zero in z. Suppose $N \geq 1$. Substitute $z = -1/y$, which is an element of the field F. Then

$$P(y, -1/y)Q(y, -1/y) = 0.$$

A field has no zero divisors, so one of these two factors is zero. Suppose first that $P(y, -1/y) = 0$. The one-variable factor theorem implies that $z + 1/y$, and therefore also $L(z) = y(z + 1/y)$, divides

P in $F[z]$. Write $P = LP_1$. Then $P_1 Q = L^{N-1}$, and the induction hypothesis applies. The case in which $Q(y, -1/y) = 0$ is identical. This proves the asserted form over $F[z]$.

Finally set $z = 0$. Since $L(0) = 1$,

$$a(y) = P(y, 0), \qquad b(y) = Q(y, 0).$$

Thus a and b are actually polynomials in y, not merely rational functions, and $a(y)b(y) = 1$. The degrees of two nonzero one-variable polynomials add under multiplication, so both must be nonzero constants. Writing $c = a$ proves Eq. (11). □

B.4 Proof of Proposition 3.6

Proof. Vectors $\Gamma|q\rangle$ with different q are supported on disjoint sets of basis vectors $|a\rangle|b\rangle$, so they are orthogonal. By Lemma 3.5,

$$\|\Gamma|q\rangle\|^2 = \frac{1}{\binom{2\ell}{q}} \sum_{a+b=q} \binom{\ell}{a}\binom{\ell}{b} = 1.$$

Hence Γ maps an orthonormal basis to an orthonormal family and is an isometry.

For the coherent-vector identity, insert Eq. (2) and Eq. (13):

$$\begin{aligned}
\Gamma|u; 2\ell\rangle &= \sum_{q=0}^{2\ell} \sqrt{\binom{2\ell}{q}}\, \alpha^{2\ell-q}\beta^q \Gamma|q\rangle \\
&= \sum_{a,b=0}^{\ell} \sqrt{\binom{\ell}{a}\binom{\ell}{b}}\, \alpha^{2\ell-a-b}\beta^{a+b}|a\rangle|b\rangle \\
&= \left(\sum_{a=0}^{\ell} \sqrt{\binom{\ell}{a}}\alpha^{\ell-a}\beta^a|a\rangle\right) \otimes \left(\sum_{b=0}^{\ell} \sqrt{\binom{\ell}{b}}\alpha^{\ell-b}\beta^b|b\rangle\right),
\end{aligned}$$

which is Eq. (14). □

B.5 Proof of Lemma 3.7

Proof. Using Eq. (13), the (a, a') matrix entry of $\mathrm{Tr}_2(\Pi)$ is

$$\sum_{b=0}^{\ell} \langle a, b|\Pi|a', b\rangle.$$

A term can be nonzero only if $a + b = a' + b$, so the matrix is diagonal. Its a-th diagonal entry is

$$D_a = \sum_{b=0}^{\ell} \frac{\binom{\ell}{a}\binom{\ell}{b}}{\binom{2\ell}{a+b}}. \tag{66}$$

By Lemma 3.2,

$$\frac{1}{\binom{2\ell}{a+b}} = (2\ell + 1) \int_0^1 x^{a+b}(1-x)^{2\ell-a-b}\,\mathrm{d}x.$$

Substitute this into Eq. (66) and use the binomial theorem:

$$\begin{aligned}
D_a &= (2\ell+1)\binom{\ell}{a} \int_0^1 x^a(1-x)^{\ell-a} \left[\sum_{b=0}^{\ell} \binom{\ell}{b} x^b(1-x)^{\ell-b}\right] \mathrm{d}x \\
&= (2\ell+1)\binom{\ell}{a} \int_0^1 x^a(1-x)^{\ell-a}\,\mathrm{d}x \\
&= (2\ell+1)\binom{\ell}{a} \frac{a!(\ell-a)!}{(\ell+1)!} \\
&= \frac{2\ell+1}{\ell+1} = \frac{s}{m}.
\end{aligned}$$

Every diagonal entry is therefore s/m, proving Eq. (16). □

B.6 Proof of Proposition 4.2

Proof. Using $\sum_b |b\rangle\langle b| = I_{\mathcal{H}_\ell}$, $\Gamma^\dagger\Gamma = I_{\mathcal{H}_{2\ell}}$, and $JJ^\dagger = P$, we obtain

$$\sum_{b=0}^{\ell} K_b^\dagger K_b = J\Gamma^\dagger\Gamma J^\dagger = P.$$

Also,

$$\sum_{r=s}^{d-1} R_r^\dagger R_r = \sum_{r=s}^{d-1} |r\rangle\langle r| = I_d - P.$$

Thus the Kraus normalization condition in Eq. (58) holds, so $\mathcal{E}_\star$ is CPTP. The displayed Kraus list has $m + d - s = d - m + 1$ terms. Since $m \geq 1$, this count is at most d.

For a source state, the reset Kraus operators vanish, and Proposition 3.6 gives

$$\begin{aligned}\mathcal{E}_\star(|\Psi_u\rangle\langle\Psi_u|) &= \mathrm{Tr}_2\big[\Gamma|u;2\ell\rangle\langle u;2\ell|\Gamma^\dagger\big] \\ &= \mathrm{Tr}_2[|u;\ell\rangle\langle u;\ell| \otimes |u;\ell\rangle\langle u;\ell|] = |u;\ell\rangle\langle u;\ell|.\end{aligned}$$

□

B.7 Proof of Proposition 4.3

Proof. By Lemma 3.7,

$$\begin{aligned}\sum_{b=0}^{\ell} A_b^\dagger A_b &= \frac{m}{s}\sum_b (I \otimes \langle b|)\Gamma\Gamma^\dagger(I \otimes |b\rangle) \\ &= \frac{m}{s}\,\mathrm{Tr}_2(\Pi) = I_{\mathcal{H}_\ell}.\end{aligned}$$

Thus $\mathcal{D}_\star$ is CPTP. Its displayed Kraus representation has $m = \ell + 1$ terms.

By Eq. (26), the latent state is $|u;\ell\rangle\langle u;\ell|$. Moreover, Eq. (14) gives

$$\begin{aligned}A_b^\dagger|\Psi_u\rangle &= \sqrt{\frac{m}{s}}(I \otimes \langle b|)\Gamma J^\dagger J|u;2\ell\rangle \\ &= \sqrt{\frac{m}{s}}(I \otimes \langle b|)\big(|u;\ell\rangle \otimes |u;\ell\rangle\big) \\ &= \sqrt{\frac{m}{s}}\langle b \mid u;\ell\rangle|u;\ell\rangle.\end{aligned}$$

Therefore

$$\begin{aligned}\langle\Psi_u|\mathcal{D}_\star(|u;\ell\rangle\langle u;\ell|)|\Psi_u\rangle &= \sum_b |\langle\Psi_u|A_b|u;\ell\rangle|^2 \\ &= \frac{m}{s}\sum_b |\langle b \mid u;\ell\rangle|^2 = \frac{m}{s},\end{aligned}$$

which proves Eq. (29). Averaging proves Eq. (30). □

B.8 Proof of Theorem 5.1

Proof. Fix arbitrary channels $\mathcal{E}$ and $\mathcal{D}$, and write

$$\sigma_u := \mathcal{E}(|\Psi_u\rangle\langle\Psi_u|). \tag{67}$$

Every density operator satisfies $0 \leq \sigma_u \leq I_m$. Complete positivity implies order preservation, hence

$$0 \leq \mathcal{D}(\sigma_u) \leq \mathcal{D}(I_m).$$

Therefore

$$F_{\mu_{n,k}}(\mathcal{E},\mathcal{D}) = \int \langle \Psi_u | \mathcal{D}(\sigma_u) | \Psi_u \rangle \, \mathrm{d}\nu(u)$$
$$\leq \int \langle \Psi_u | \mathcal{D}(I_m) | \Psi_u \rangle \, \mathrm{d}\nu(u)$$
$$= \mathrm{Tr}\left(\mathcal{D}(I_m) \frac{P}{s} \right) \tag{68}$$
$$\leq \frac{\mathrm{Tr}(\mathcal{D}(I_m))}{s} = \frac{\mathrm{Tr}(I_m)}{s} = \frac{m}{s}. \tag{69}$$

In Eq. (68) we used Eq. (21). In Eq. (69) we used $0 \leq P \leq I_d$, positivity of $\mathcal{D}(I_m)$, and trace preservation. The pair in Propositions 4.2 and 4.3 attains m/s, so the upper bound is tight. □

B.9 Proof of Proposition 5.2

Proof. Optimality means that equality holds throughout Eqs. (68) and (69). Equality in the last inequality gives

$$0 = \mathrm{Tr}\big((I_d - P)\mathcal{D}(I_m)\big).$$

By Fact A.3(ii), $\mathcal{D}(I_m) = P\mathcal{D}(I_m)P$. Since

$$\mathcal{D}(I_m) = \sum_i A_i A_i^\dagger,$$

we have

$$0 = (I-P)\mathcal{D}(I_m)(I-P) = \sum_i (I-P)A_i A_i^\dagger (I-P).$$

Each summand is positive semidefinite, so Fact A.3(iii) yields $(I-P)A_i = 0$, proving (i).

For the first inequality in Eq. (68), define the continuous nonnegative function

$$g(u) := \langle \Psi_u | \mathcal{D}(I_m - \sigma_u) | \Psi_u \rangle.$$

Its integral is zero, so $g(u) = 0$ for every u. Expanding the decoder Kraus representation gives

$$0 = \sum_i \langle A_i^\dagger \Psi_u | (I_m - \sigma_u) | A_i^\dagger \Psi_u \rangle. \tag{70}$$

The operator $I_m - \sigma_u$ is positive semidefinite, and every term in the sum is nonnegative. Each term is therefore zero. By Fact A.3(iv),

$$(I_m - \sigma_u) A_i^\dagger |\Psi_u\rangle = 0,$$

which proves (ii).

Finally, let

$$h(u) := \langle \Psi_u | \mathcal{D}(I_m) | \Psi_u \rangle = \sum_i \left\| A_i^\dagger \Psi_u \right\|^2.$$

The function h is not identically zero. Otherwise positivity would imply $\mathcal{D}(I_m)|\Psi_u\rangle = 0$ for every u. The coherent vectors span PH_d by Lemma 3.3, and $\mathcal{D}(I_m)$ is supported in P, so $\mathcal{D}(I_m) = 0$, contradicting $\mathrm{Tr}(\mathcal{D}(I_m)) = m$. Hence $h(u_0) > 0$ for some u_0, and by continuity $h > 0$ on a nonempty open neighborhood.

For every u in that neighborhood, at least one vector $A_i^\dagger |\Psi_u\rangle$ is nonzero. By (ii), σ_u has eigenvalue one. Since σ_u is a density operator, all other eigenvalues are zero. Thus σ_u is pure on that open set. □

B.10 Proof of Proposition 5.4

Proof. Choose orthonormal bases of X and Y, and reshape the coordinates of $W|v_N(z)\rangle$ into a matrix $M(z)$. Every matrix entry is a polynomial in z, of degree at most N. Product-valuedness means that $M(z)$ has rank one for every z, so every 2×2 minor vanishes for every z and hence is the zero polynomial. By Lemma B.2,

$$M(z) = p(z)q(z)^{\mathsf{T}}$$

for polynomial vectors $p(z) \in X$ and $q(z) \in Y$. In tensor notation,

$$W|v_N(z)\rangle = p(z) \otimes q(z). \tag{71}$$

Because W is an isometry, Eq. (36) and Eq. (71) imply

$$\langle p(w) \mid p(z)\rangle \, \langle q(w) \mid q(z)\rangle = (1 + \overline{w}z)^N. \tag{72}$$

Treat $y = \overline{w}$ as an independent formal variable. Both factors on the left are polynomials in y, z, so Lemma 3.4 gives an integer a and a nonzero constant c such that

$$\langle p(w) \mid p(z)\rangle = c(1 + \overline{w}z)^a, \qquad \langle q(w) \mid q(z)\rangle = c^{-1}(1 + \overline{w}z)^{N-a}. \tag{73}$$

Setting $w = z = 0$ shows $c = \|p(0)\|^2 > 0$. Replace p by $c^{-1/2}p$ and q by $c^{1/2}q$. Then

$$\langle p(w) \mid p(z)\rangle = (1 + \overline{w}z)^a = \langle v_a(w) \mid v_a(z)\rangle. \tag{74}$$

Define a linear map on finite linear combinations of affine coherent vectors by

$$U_a \left(\sum_j c_j |v_a(z_j)\rangle \right) := \sum_j c_j p(z_j).$$

It is well defined and isometric because Eq. (74) implies equality of the squared norms of the two sums. The vectors $v_a(z)$ span $\mathcal{H}_a$, by Lemma 3.3 or by polynomial interpolation, so U_a extends to an isometry on all of $\mathcal{H}_a$. The same argument constructs V_{N-a} from q. This proves Eq. (39). For a unit vector whose first coordinate is nonzero, write

$$u = e^{i\varphi} \frac{(1, z)}{\sqrt{1 + |z|^2}}.$$

By homogeneity, $|e^{i\varphi}u; t\rangle = e^{it\varphi}|u; t\rangle$, so the phases on the two factors add to the phase on the degree-N vector. Together with Eq. (37), this proves Eq. (40). The unit vector with first coordinate zero follows by continuity. □

B.11 Proof of Proposition 5.5

Proof. By Theorem A.5, choose a Stinespring isometry

$$W : H_d \to H_m \otimes H_E$$

for $\mathcal{E}$, and restrict it to the source support:

$$\widetilde{W} := WJ : \mathcal{H}_N \to H_m \otimes H_E.$$

This restriction is still an isometry. By Proposition 5.2(iii), the first marginal of $\widetilde{W}|u; N\rangle\langle u; N|\widetilde{W}^\dagger$ is pure on a nonempty open set of qubit states. By Lemma 5.3, $\widetilde{W}|u; N\rangle$ is a product vector on that open set.

In the affine coordinate z, every 2×2 minor of the polynomial matrix representing $\widetilde{W}|v_N(z)\rangle$ vanishes on a nonempty open subset of $\mathbb{C}$. A one-variable polynomial that vanishes on an open set is identically zero. Hence the vector is a product for every $z \in \mathbb{C}$. Applying Proposition 5.4, there are an integer a and isometries $U_a : \mathcal{H}_a \to H_m$, $V_{N-a} : \mathcal{H}_{N-a} \to H_E$ such that

$$\widetilde{W}|u; N\rangle = U_a|u; a\rangle \otimes V_{N-a}|u; N - a\rangle.$$

Taking the partial trace over E gives

$$\sigma_u = U_a|u;a\rangle\langle u;a|U_a^\dagger. \tag{75}$$

Because U_a is an isometry into the m-dimensional latent space, $a+1 \le m$.

We claim that the vectors $U_a|u;a\rangle$ span all of H_m. Suppose not. Choose a unit vector $|z\rangle \in H_m$ orthogonal to their span. Then $|z\rangle$ is orthogonal to the support of every σ_u, and because σ_u is rank one,

$$|z\rangle\langle z| \le I_m - \sigma_u. \tag{76}$$

The pointwise equality used in Proposition 5.2 gives

$$\langle\Psi_u|\mathcal{D}(I_m - \sigma_u)|\Psi_u\rangle = 0.$$

Positivity and Eq. (76) imply

$$\langle\Psi_u|\mathcal{D}(|z\rangle\langle z|)|\Psi_u\rangle = 0 \quad \text{for every } u. \tag{77}$$

By Eq. (33), the positive operator $\mathcal{D}(|z\rangle\langle z|)$ is supported in P. By Fact A.3(iv), Eq. (77) implies that it annihilates every source coherent vector. Those vectors span PH_d, so $\mathcal{D}(|z\rangle\langle z|) = 0$. This contradicts trace preservation, because

$$\mathrm{Tr}(\mathcal{D}(|z\rangle\langle z|)) = \mathrm{Tr}(|z\rangle\langle z|) = 1.$$

Thus the latent vectors span H_m.

The vectors $|u;a\rangle$ span the $(a+1)$-dimensional space $\mathcal{H}_a$ by Lemma 3.3, and U_a is injective. Therefore their image span has dimension $a+1$. Since it equals H_m,

$$a+1 = m = \ell + 1,$$

so $a = \ell$, and U_a is unitary. This proves Eq. (41).

Finally, replace $\mathcal{E}$ by $\mathrm{Ad}_{U^\dagger} \circ \mathcal{E}$ and $\mathcal{D}$ by $\mathcal{D} \circ \mathrm{Ad}_U$, where $\mathrm{Ad}_U(X) = UXU^\dagger$. Their composition is unchanged, and multiplying every decoder Kraus operator on the right by U preserves the dimension of their linear span and therefore preserves Kraus rank. This establishes Eq. (42). □

B.12 Proof of Lemma 6.1

Proof. Use the affine coherent vectors from Eq. (35). For each finite z, proportionality gives

$$B|v_{2\ell}(z)\rangle = c(z)|v_\ell(z)\rangle \tag{78}$$

for some scalar $c(z)$. The first coordinate of $|v_\ell(z)\rangle$ equals one, so $c(z)$ is the first coordinate of the polynomial vector $B|v_{2\ell}(z)\rangle$. Thus $c(z)$ is a polynomial. The last coordinate of Eq. (78) is $c(z)z^\ell$, while every coordinate of $B|v_{2\ell}(z)\rangle$ has degree at most 2ℓ. Therefore $\deg c \le \ell$.

Every polynomial of degree at most ℓ can be written uniquely in the form

$$c(z) = \langle \xi \mid v_\ell(z)\rangle \tag{79}$$

for a suitable $\xi \in \mathcal{H}_\ell$, because the coordinates of $v_\ell(z)$ are nonzero scalar multiples of $1, z, \dots, z^\ell$. For $u = (1,z)/\sqrt{1+|z|^2}$, multiply Eq. (78) by $(1+|z|^2)^{-\ell}$ and use Eq. (37) to obtain Eq. (45). The remaining unit vector with first coordinate zero follows by continuity. □

B.13 Proof of Lemma 6.2

Proof. It is enough to use affine coherent vectors. Write $X_{ab} = \langle a|X|b\rangle$. By Eq. (35),

$$\langle v_\ell(z)|X|v_\ell(z)\rangle = \sum_{a,b=0}^{\ell} \sqrt{\binom{\ell}{a}\binom{\ell}{b}} X_{ab}\, \overline{z}^{\,a} z^b. \tag{80}$$

The left side vanishes for every $z \in \mathbb{C}$. Put $z = re^{i\theta}$. For each fixed $r \ge 0$, the right side is a finite Fourier series in θ. Hence every Fourier coefficient vanishes. For a fixed difference $b-a=t$, the corresponding coefficient is a polynomial in r whose monomials have distinct powers r^{a+b}. Therefore each coefficient X_{ab} is zero. Thus $X = 0$. □

B.14 Proof of Lemma 6.3

Proof. If $\mathcal{T}(X) = 0$, then

$$\begin{aligned} 0 &= \mathrm{Tr}\big(X^\dagger \mathcal{T}(X)\big) \\ &= \int \overline{\langle u;\ell|X|u;\ell\rangle}\langle u;\ell|X|u;\ell\rangle \,\mathrm{d}\nu(u) \\ &= \int |\langle u;\ell|X|u;\ell\rangle|^2 \,\mathrm{d}\nu(u). \end{aligned}$$

The continuous nonnegative integrand must vanish for every u. By Lemma 6.2, $X = 0$. Thus $\mathcal{T}$ is injective.

For $X = I$, the expectation value in Eq. (50) equals one. Hence Eq. (51) is the resolution identity Eq. (9) with $t = \ell$ and $m = \ell + 1$. □

B.15 Proof of Theorem 6.4

Proof. By Eq. (33), $PA_i = A_i$. Trace preservation of $\mathcal{D}$ therefore implies

$$I_{\mathcal{H}_\ell} = \sum_i A_i^\dagger A_i = \sum_i A_i^\dagger P A_i = \sum_i B_i B_i^\dagger. \tag{81}$$

Fix a unit qubit vector v. Using Eq. (81) and the resolution of the identity on $\mathcal{H}_{2\ell}$,

$$1 = \sum_i \left\| B_i^\dagger |v;\ell\rangle \right\|^2 \tag{82}$$

$$= s \int \sum_i \left| \langle u;2\ell|B_i^\dagger|v;\ell\rangle \right|^2 \mathrm{d}\nu(u). \tag{83}$$

By Eq. (46),

$$\left| \langle u;2\ell|B_i^\dagger|v;\ell\rangle \right|^2 = |\langle \xi_i \mid u;\ell\rangle|^2 \, |\langle u;\ell \mid v;\ell\rangle|^2 .$$

Summing over i and using Eq. (47), Eq. (83) becomes

$$1 = s \int \langle u;\ell|R|u;\ell\rangle \, |\langle u;\ell \mid v;\ell\rangle|^2 \,\mathrm{d}\nu(u) = s\langle v;\ell|\mathcal{T}(R)|v;\ell\rangle. \tag{84}$$

Thus

$$\langle v;\ell|\big(s\mathcal{T}(R) - I\big)|v;\ell\rangle = 0$$

for every coherent vector. By Lemma 6.2,

$$s\mathcal{T}(R) = I. \tag{85}$$

On the other hand, Eq. (51) gives

$$s\mathcal{T}\Big(\frac{m}{s} I\Big) = I.$$

Since $\mathcal{T}$ is injective by Lemma 6.3, comparison with Eq. (85) yields

$$R = \frac{m}{s} I_{\mathcal{H}_\ell}. \tag{86}$$

Therefore $\mathrm{rank}(R) = m$. By Eq. (48),

$$q \geq \mathrm{rank}(R) = m.$$

Because $q = \mathrm{rank}_{\mathrm{Kraus}}(\mathcal{D})$, this proves Eq. (52).

The decoder $\mathcal{D}_\star$ is optimal by Theorem 5.1 and has a Kraus representation with m terms by Proposition 4.3. □