Classical Identification over Quantum Channels

Strong converse bounds for the qubit depolarizing channel

Liuhang Ye · Bjarne Bergh · Nilanjana Datta

arXiv:2603.29987

22. June 2026 · Beyond IID 14 · Shenzhen

UNIVERSITY OF CAMBRIDGE

Outline

1. Message transmission
2. Message identification
3. Result 1: identification converse bound
4. Result 2: simultaneous identification converse bound
5. Conclusion and open questions

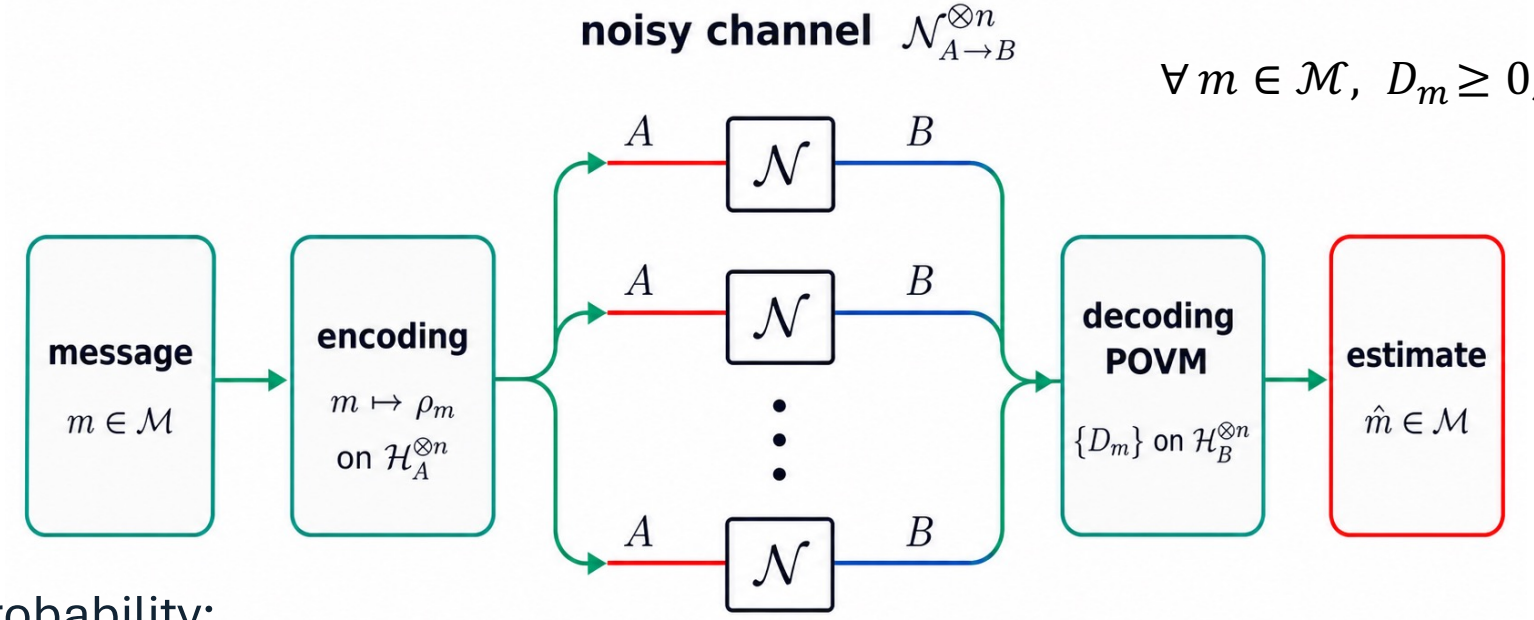
} depolarizing channel

Message Transmission

arXiv:2603.29987

Ordinary message transmission: which message was sent?

Task: the sender chooses one message and the receiver must reconstruct the message.



$$\forall\, m \in \mathcal{M},\ D_m \geq 0,\ \sum_{m\in\mathcal{M}} D_m = \mathbb{1}_{\mathbf{B^n}}.$$

Success probability:

$$\forall\, m \in \mathcal{M}, \qquad \mathrm{Tr}\left[\mathcal{N}_{\mathrm{A}\to\mathrm{B}}^{\otimes n}(\rho_m) D_m\right] \geq 1 - \lambda.$$

An (n, M_n, λ_n) transmission code uses channel n times to distinguish $M_n = |\mathcal{M}|$ messages (more messages with more channel uses) with worst-case error probability λ_n.

n is also referred to as the block length

arXiv:2603.29987

Capacity is the optimal asymptotic rate

- Rate of the code is defined as $R_n = \frac{1}{n}\log M_n$

- Asymptotic achievability

$$R \text{ is achievable} \iff \exists\,(n, M_n, \lambda_n)\text{-codes such that } \lambda_n \to 0 \text{ and } R_n := \frac{1}{n}\log M_n \to R.$$

- Classical capacity of the quantum channel $C(\mathcal{N})$ is the supreme of asymptotically achievable rates.

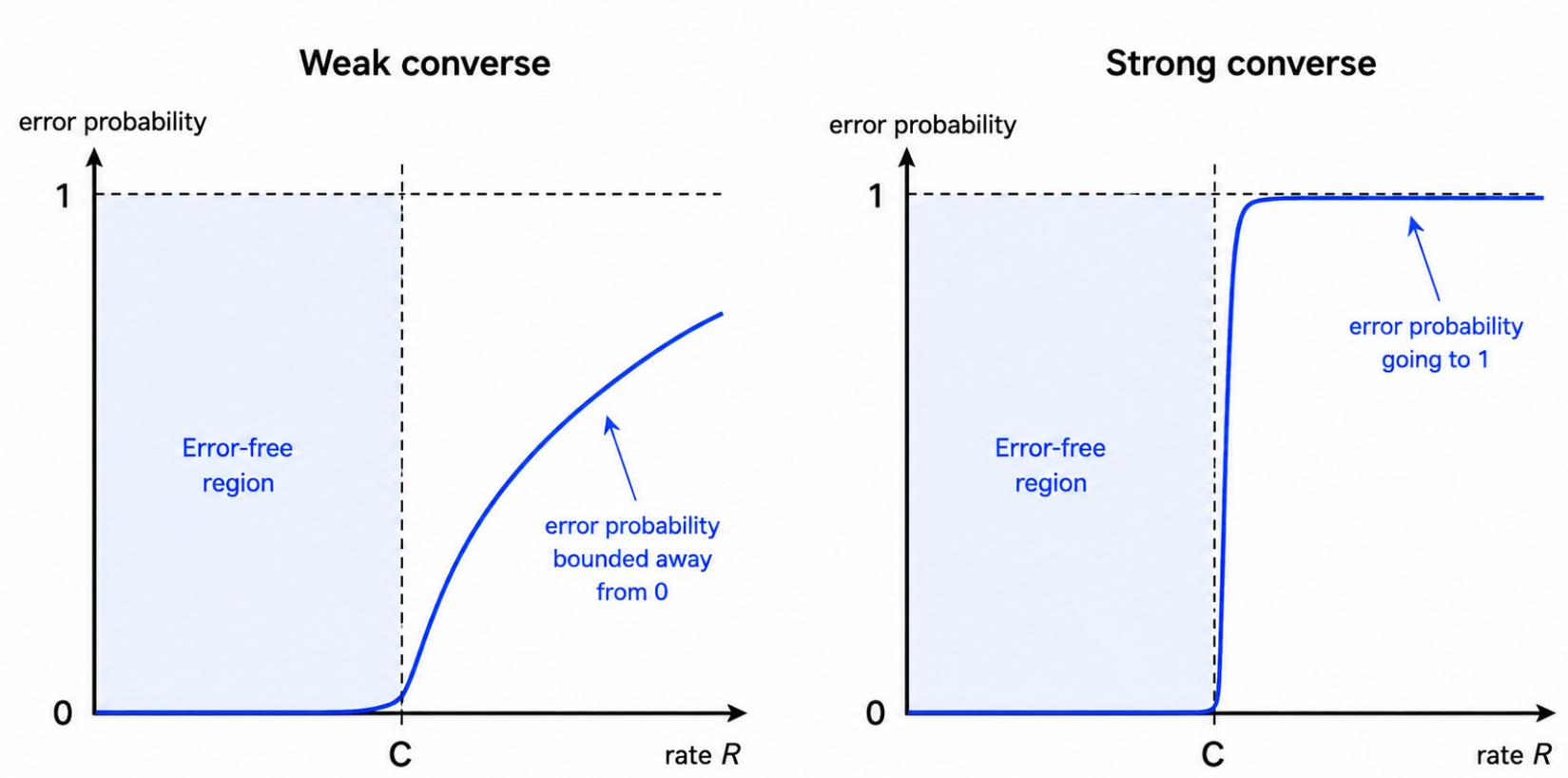
- To prove a capacity theorem (finding a formula for $C(\mathcal{N})$):

  1) show that for any rate below the capacity, there exists code with vanishing error (achievability);
  2) show that for any rate above the capacity, error-free transmission is impossible (weak converse).

- Operationally, capacity is the boundary between possible and impossible rates.

arXiv:2603.29987

Weak versus strong converse



Holevo-Schumacher-Westmoreland (HSW) capacity theorem (97,98): achievability + weak converse

$$C(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n}\chi\left(\mathcal{N}^{\otimes n}\right) \qquad \chi(\mathcal{N}) = \max_{\{p_x,\rho_x\}} \left[H\left(\sum_x p_x \mathcal{N}(\rho_x)\right) - \sum_x p_x H(\mathcal{N}(\rho_x)) \right].$$

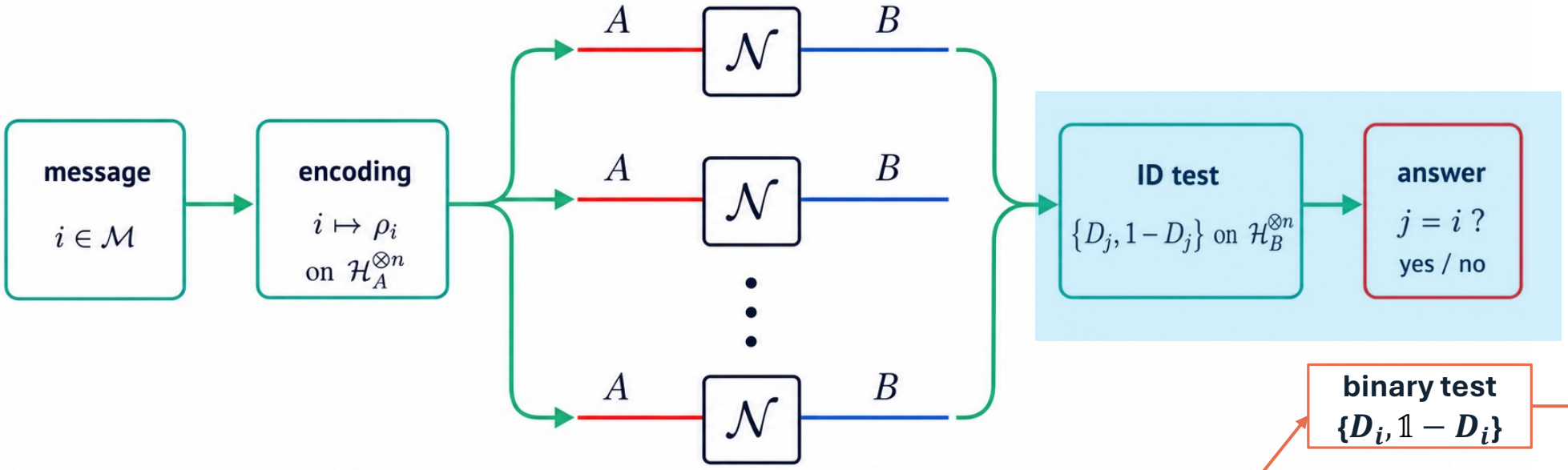
Message Identification

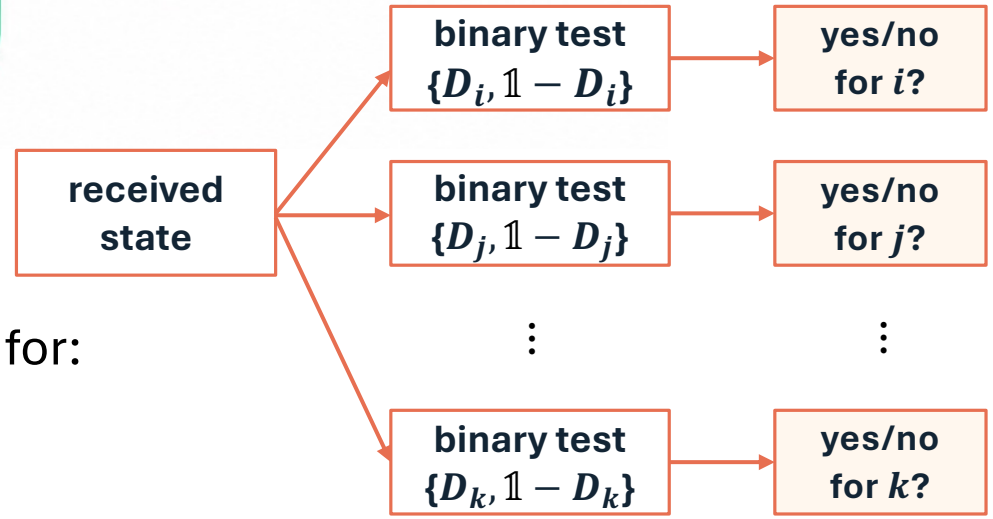
arXiv:2603.29987

Identification: was message i sent?

Task: the sender chooses one message and the receiver answers one yes or no question.

Introduced for classical channels [Ahlswede & Dueck, 89]; extended to quantum channels [Löber, 99].





The receiver has in mind one arbitrary message to test for:
but only one message can be tested.

arXiv:2603.29987

Identification has two error parameters

For each message i, the decoding test is a binary POVM $\{D_i, \mathbb{1} - D_i\}$.
Multiple hypothesis tests: one test for each message against the rest.

Type I error: false negative

$$\mathrm{Tr}\left[\mathcal{N}^{\otimes n}(\rho_i) D_i\right] \geq 1 - \lambda_1$$

Type II error: false positive

$$\mathrm{Tr}\left[\mathcal{N}^{\otimes n}(\rho_j) D_i\right] \leq \lambda_2 \quad (j \neq i)$$

An $(n, N_n, \lambda_{1,n}, \lambda_{2,n})$ identification code uses channel n times to identify from N_n messages with worst-case type-I error $\lambda_{1,n}$ and type-II error $\lambda_{2,n}$.

arXiv:2603.29987

Identification capacity: doubly exponential rate

Identification is an easier task, with message size scaling exponentially faster than transmission!

Transmission

$$M_n \approx 2^{nR} \qquad R = \frac{1}{n}\log M_n$$

Identification

$$N_n \approx 2^{2^{nR}} \qquad R = \frac{1}{n}\log\log N_n$$

- Rate of the code is defined as $R_n := \frac{1}{n}\log\log N_n$
- Asymptotic achievability:

$$R \text{ is achievable} \iff \exists\,(n, N_n, \lambda_{1,n}, \lambda_{2,n})\text{-codes such that } \lambda_{1,n} \to 0,\ \lambda_{2,n} \to 0,\ \text{and } R_n := \frac{1}{n}\log\log N_n \to R.$$

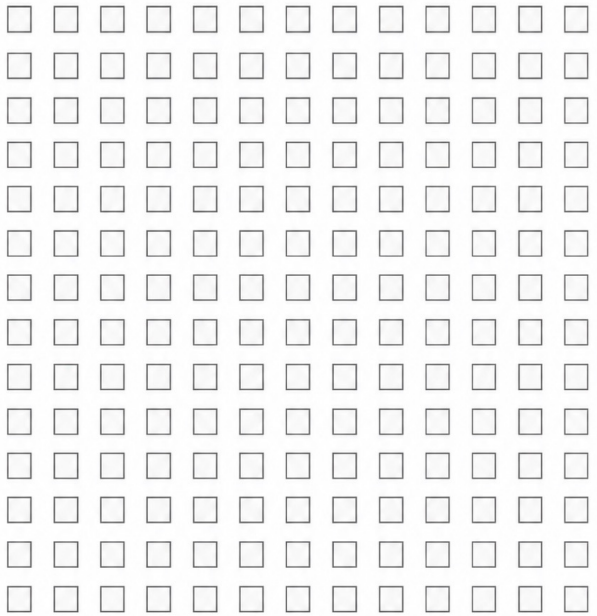
- Classical ID capacity of the quantum channel $C_{\mathrm{ID}}(\mathcal{N})$ is the supreme of asymptotically achievable rates.

arXiv:2603.29987

Identification capacity: doubly exponential rate

Transmission codebook

$\mathcal{M}$, $M \approx 2^{nC}$

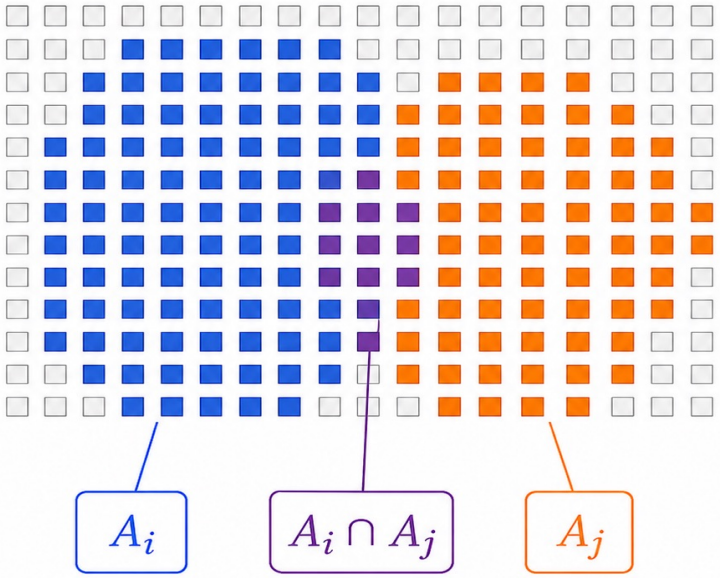


ordinary transmission messages

ID messages are large subsets of $\mathcal{M}$

ID message i ⟷ subset A_i

ID message j ⟷ subset A_j



Large subsets with small overlap

1 **To send i :**

choose $k \in A_i$ at random

2 **To test j :**

accept iff $k \in A_j$

$i = j$:

$k \in A_i$ automatically,

$i \neq j$:

$$\Pr[k \in A_j] = \frac{|A_i \cap A_j|}{|A_i|}.$$

Combinatorics fact: Many large subsets with small pairwise overlaps can be packed inside $\mathcal{M}$, giving $N \approx 2^{O(M)}$. Since $M \approx 2^{nC}$, this yields $N \approx 2^{O(2^{nC})}$.

arXiv:2603.29987

Strong converse for identification capacity

- Identification code has two error parameters, need to impose $\lambda_1 + \lambda_2 < 1$ (otherwise the receiver can always answer yes: $\lambda_1 = 0, \lambda_2 = 1$);

- Strong converse bound for identification:

  for any rate above the bound, not only error-free identification is impossible (weak converse), but also $\lambda_1 + \lambda_2$ converges to a value ≥ 1.

Known results:

- For classical-quantum (cq) channels, $C_{\mathrm{ID}}(W) = C(W)$ and strong converse holds [Ahlswede & Winter, 2000];

- For the quantum identity channel id_{A}, $C_{\mathrm{ID}}(id_{\mathrm{A}}) = 2\log|\mathrm{A}|$ and strong converse holds [Winter, 04];

- For general quantum channels, $C_{\mathrm{ID}}(\mathcal{N}) \leq \log|A| + \hat{Q}(\mathcal{N})$ [Atif, Pradhan & Winter, 24].

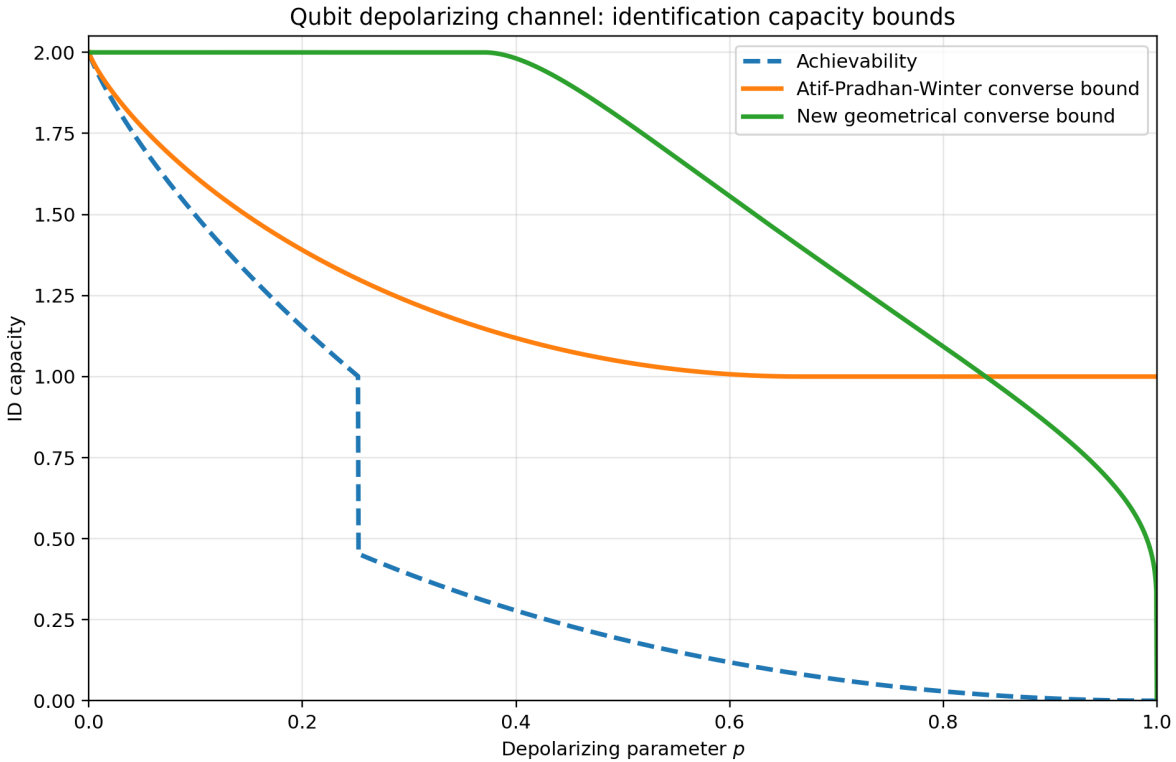
Result 1: strong converse bound for identification via the qubit depolarizing channel

arXiv:2603.29987

Main result 1: a new geometrical strong converse bound

The qubit depolarizing channel is defined as follows, for $0 \leq p \leq 1$

$$\mathcal{N}_p(\rho) = (1-p)\rho + p\,\mathbb{1}/2.$$



$$C_{\mathrm{ID}}(\mathcal{N}_p) \leq \begin{cases} 2, & 0 \leq p \leq 1 - 2^{-2/3}, \\ 2 - D(\gamma(p)\|3/4), & 1 - 2^{-2/3} \leq p < 1, \end{cases} \qquad \gamma(p) = \frac{-1}{2\log(1-p)}$$

Here $D(\cdot\|\cdot)$ is binary relative entropy; logs are base 2.

- The bound is not tight (poor behavior for small p);
- The key feature is the correct completely noisy limit: it goes to 0 as $p \to 1$.

arXiv:2603.29987

Proof strategy: geometrical covering

Binary tests force separation between channel outputs

For each message i, assign the codeword $\rho_i \in \mathcal{H}_A^{\otimes n}$, let

$$\sigma_i := \mathcal{N}^{\otimes n}(\rho_i), \qquad \delta := 1 - \lambda_1 - \lambda_2.$$

Identification requires that for $j \neq i$,

$$\mathrm{Tr}[\sigma_i D_i] \geq 1 - \lambda_1, \qquad \mathrm{Tr}[\sigma_j D_i] \leq \lambda_2.$$

By the variational formula for trace distance,

$$\frac{1}{2}\|\sigma_i - \sigma_j\|_1 = \max_{0 \leq \Lambda \leq \mathbf{1}} \mathrm{Tr}[(\sigma_i - \sigma_j)\Lambda].$$

Apply to $\Lambda = D_i$ gives

$$\frac{1}{2}\|\sigma_i - \sigma_j\|_1 \geq \mathrm{Tr}[(\sigma_i - \sigma_j)D_i] \geq \delta.$$

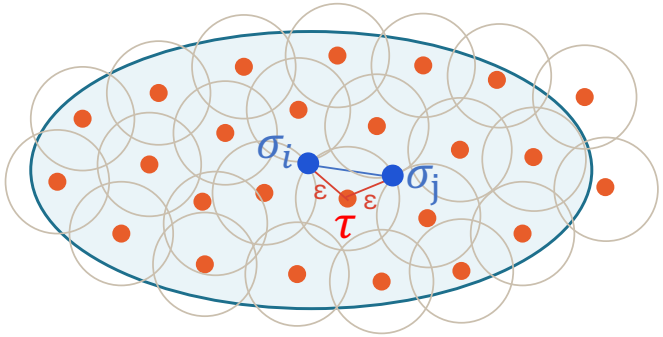
Covering the channel output geometry

Output geometry: $K_n := \{\mathcal{N}^{\otimes n}(\rho) : \rho \in \mathcal{S}(A^n)\}$.

Minimum covering number: minimum number of trace-distance ε-balls covering K_n:

$$M_\varepsilon(K_n) := \min\{|\mathcal{C}| : K_n \subseteq \bigcup_{\tau \in \mathcal{C}} B_\varepsilon(\tau)\}.$$

channel output geometry K_n



Then

$$\text{Number of messages} \leq M_{\delta/2}(K_n)$$

Otherwise

$$\frac{1}{2}\|\sigma_i - \sigma_j\|_1 \leq \frac{1}{2}\|\sigma_i - \omega\|_1 + \frac{1}{2}\|\omega - \sigma_j\|_1 \leq \delta$$

arXiv:2603.29987

Geometry of the qubit depolarizing channel

Represent n-qubit states in the Pauli basis

$$\rho = \frac{\mathbf{1}_d}{d} + \sum_{\alpha \neq 0} r_\alpha \, \widetilde{\sigma}_\alpha, \qquad d = 2^n$$

$$\widetilde{\sigma}_\alpha = 2^{-n/2} \, \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_n}, \qquad \alpha \in \{0,1,2,3\}^n.$$

The Pauli coefficients form a high-dim Bloch vector, contained in an Euclidean ball:

$$r = (r_\alpha)_{\alpha \neq 0} \in \mathbb{R}^{4^n - 1}.$$

$$\|r\|_2^2 = \left\|\rho - \tfrac{\mathbb{1}_d}{d}\right\|_2^2 = \mathrm{Tr}(\rho^2) - \frac{1}{d} \leq 1 - \frac{1}{d}$$

Key contraction rule

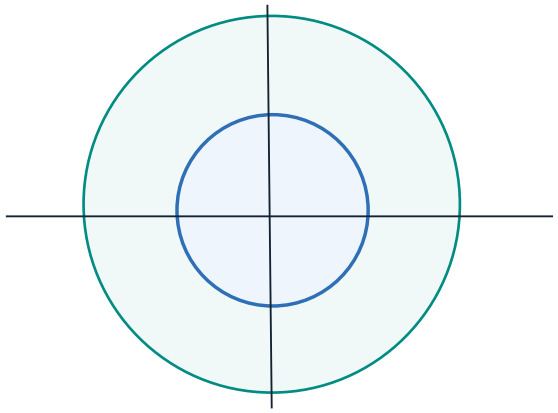
$$r_\alpha \longmapsto r'_\alpha = (1-p)^{w(\alpha)} r_\alpha$$

where w(α) = number of non-zero Pauli operators

Output geometry is contained in a high-dim ellipsoid

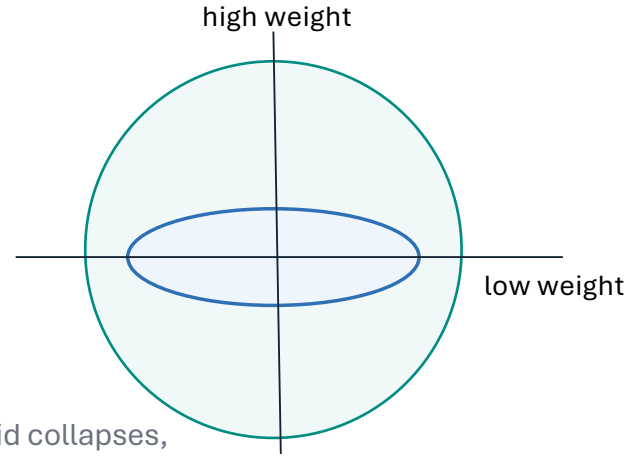
Single qubit:

isotropic shrinking



Multiple qubits:

anisotropic shrinking



Intuition: for p close to 1 the ellipsoid collapses, so the covering number — hence the ID capacity — goes to zero.

arXiv:2603.29987

Strong converse bound for identification

- Standard result on ellipsoid covering provides an upper bound on the minimal covering number of an arbitrary ellipsoid [Dumer, 06];

- Upper bound on the minimal covering number leads to upper bound on the message size of an arbitrary identification code. By taking the asymptotic limit, we obtain the following result:

Theorem 5 (Strong converse bound for the unrestricted classical identification capacity of $\mathcal{N}_p$). *The unrestricted identification capacity of the qubit depolarizing channel with error probability $p \in [0,1]$ satisfies the following strong converse bound: for any $\lambda_1, \lambda_2 > 0$, such that $\lambda_1 + \lambda_2 < 1$,*

$$C_{\mathrm{ID}}(\mathcal{N}_p) \leq \limsup_{n\to\infty} \frac{1}{n} \log\log N(n, \lambda_1, \lambda_2) \leq \begin{cases} 2, & 0 \leq p \leq 1 - 2^{-2/3}, \\ 2 - D\left(\gamma(p) \middle\| \frac{3}{4}\right), & 1 - 2^{-2/3} \leq p < 1, \end{cases} \tag{116}$$

where

$$\gamma(p) := \frac{-1}{2\log(1-p)}, \tag{117}$$

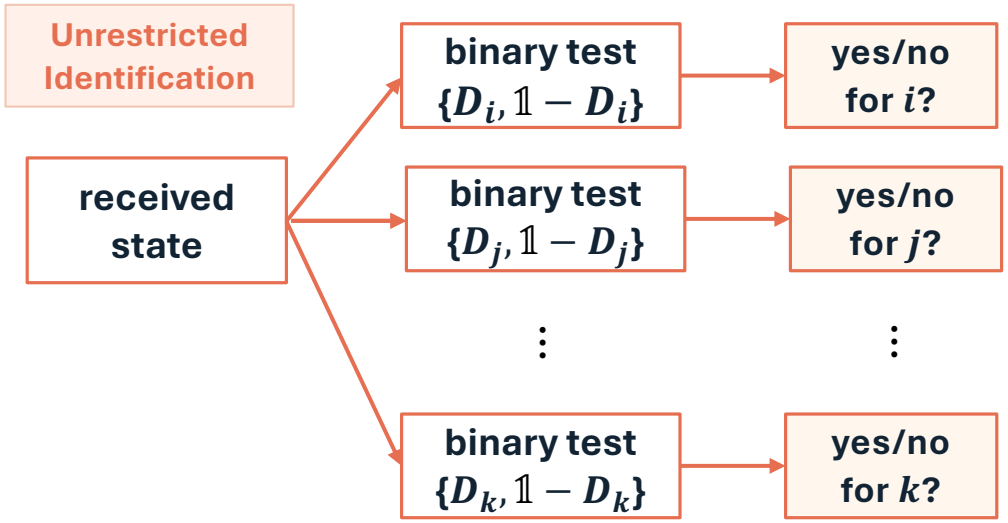
and the binary relative entropy is defined as

$$D(x\|y) := x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}. \tag{118}$$

Result 2: capacity theorem for simultaneous identification under product-basis measurements

arXiv:2603.29987

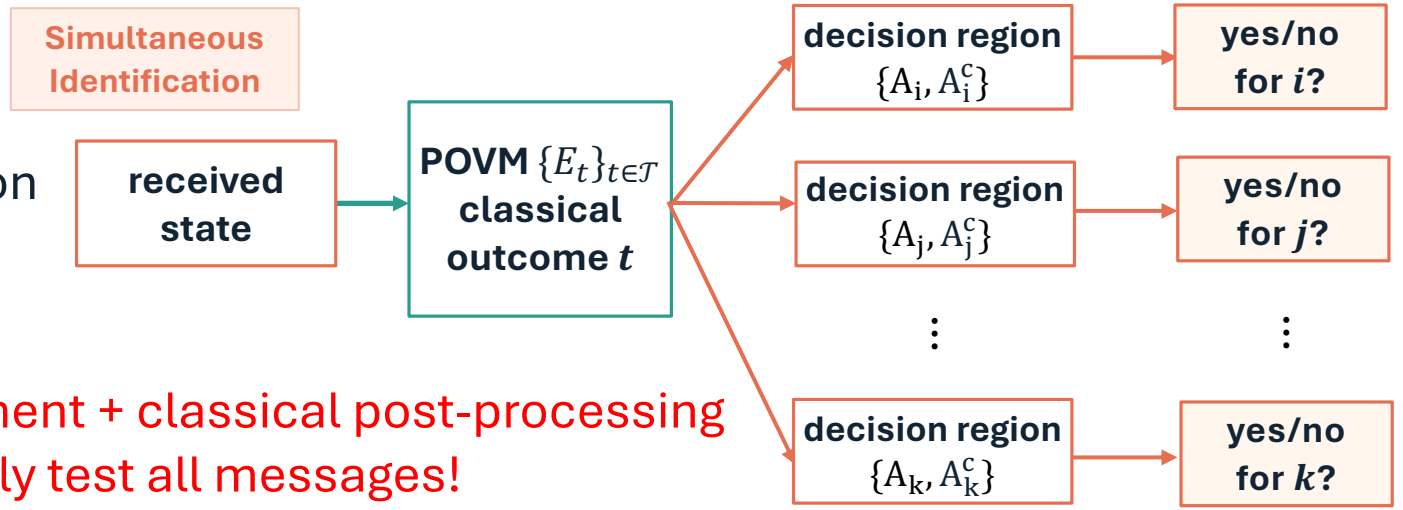
Simultaneous identification [Löber, 99]



Previously (unrestricted identification), the receiver performs one binary POVM $\{D_j, \mathbb{1} - D_j\}$ to test if the message is j, cannot simultaneously test all messages.

In simultaneous identification, the receiver performs a common POVM $\{E_t\}_{t\in\mathcal{T}}$, and associates a decision region $A_j \subseteq \mathcal{T}$ for the message j, equivalently,

$$D_j = \sum_{t\in A_j} E_t$$



one big measurement + classical post-processing can simultaneously test all messages!

arXiv:2603.29987

Simultaneous identification capacity

- Simultaneous identification capacity $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N})$ defined as the supreme of asymptotically achievable rates, but with this extra restriction on the decoding operation (all binary tests originate from the same underlying POVM).

- By definition $C_{\mathrm{ID}}(\mathcal{N}) \geq C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N})$. Moreover, $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}) \geq C(\mathcal{N})$:

  recall the double-exponential argument: the common POVM is the transmission POVM

- We lack an explicit characterization of $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N})$

  For quantum identity channel $C_{\mathrm{ID}}^{\mathrm{sim}}(id_{\mathrm{A}}) = \log|\mathrm{A}|$ [Atif, Pradhan & Winter, 24]

- The only known strong converse bound: $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}) \leq \min\{\log|A|, \log|B|\}$ [Atif, Pradhan & Winter, 24]

- No evidence for separation between $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N})$ and $C(\mathcal{N})$.

arXiv:2603.29987

Further restriction on the identification decoder

- In a general simultaneous-ID, the underlying measurement can be an arbitrary POVM $\{E_t\}_{t\in\mathcal{T}}$;
- We impose further restriction on what measurements the receiver is allowed to perform:
  (recall that the measurement is done on the joint channel output space $\mathcal{H}_B^{\otimes n}$)

Fix an orthonormal basis $\{|\psi_k^{(\mathrm{x_k})}\rangle\}_{\mathrm{x_k}=0}^{d_B-1}$ for the k^{th} tensor factor, and define the product-basis of $\mathcal{H}_B^{\otimes n}$ by taking the tensor product of all the local basis vectors:

$$|\Psi_{x^n}\rangle = \bigotimes_{k=1}^{n} |\psi_k^{(x_k)}\rangle \qquad E_{x^n} = |\Psi_{x^n}\rangle\langle\Psi_{x^n}|$$

We call this a **product-basis measurement**, and study simultaneous identification codes with this kind measurements. The corresponding identification capacity is denoted as $\widetilde{C_{\mathrm{ID}}^{\mathrm{sim}}}(\mathcal{N})$.

arXiv:2603.29987

Main result 2: a new capacity theorem

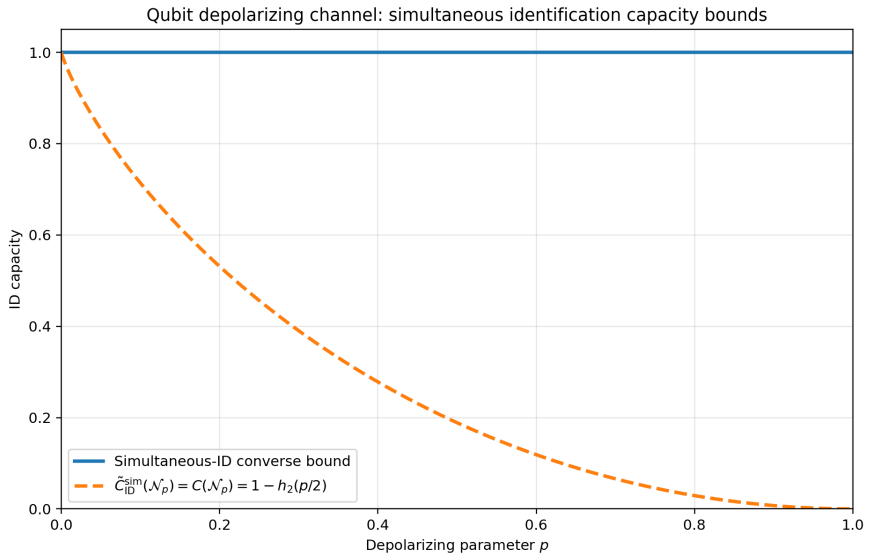
Theorem 3 (Simultaneous identification capacity of $\mathcal{N}_p$ under product-basis measurements). *The simultaneous identification capacity of the qubit depolarizing channel with error probability $p \in [0,1]$, under the constraint of product-basis measurements, is given by*

$$\tilde{C}_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}_p) = C(\mathcal{N}_p) = 1 - h(p/2), \tag{60}$$

and strong converse holds: for any $\lambda_1, \lambda_2 > 0$, such that $\lambda_1 + \lambda_2 < 1$,

$$\tilde{C}_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}_p) \leq \limsup_{n\to\infty} \frac{1}{n} \log\log \tilde{N}_{\mathrm{sim}}(n, \lambda_1, \lambda_2) \leq 1 - h(p/2), \tag{61}$$

where h is the binary entropy function: $h(p) := -p\log p - (1-p)\log(1-p)$.



The bound has the correct completely noisy limit:

$$p = 1 \;\Rightarrow\; \text{zero identification capacity}$$

Conclusion and open questions

arXiv:2603.29987

Take-home messages

1. Identification receiver answers yes/no questions, instead of full decoding.
   Simultaneity means all tests are coarse-graining of a common measurement.

2. Both unrestricted and simultaneous identification exhibit doubly exponential scaling behavior, and one can build identification codes from transmission codes: $C_{\mathrm{ID}} \geq C_{\mathrm{ID}}^{\mathrm{sim}} \geq C$.

3. One way to establish converse bounds for identification is to find minimal coverings of the channel outputs: this can be geometrical (ellipsoid) or information-theoretic (soft-covering).

4. Using the ellipsoid covering technique, we established a strong converse bound for the unrestricted identification capacity for the qubit depolarizing channel.
   Recently, this has been generalized to **arbitrary** quantum channels. [Singh 2606.05032]

5. Under the additional restriction of product-basis measurements, we obtained an exact capacity formula for the qubit depolarizing channel: $\widetilde{C}_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}_p) = C(\mathcal{N}_p) = 1 - h_2(p/2)$

arXiv:2603.29987

Open questions

1. In simultaneous identification for the depolarizing channel, is it possible to lift the product-basis measurement restriction and still obtain the same capacity?

   Can entangled decoding POVM increase identification rate?

   Toy example: for two channel uses ($n = 2$), Bell measurement cannot do better than product-basis measurement; unclear for generic 2-qubit entangled measurements.

2. More broadly, is it true that $C_{\mathrm{ID}}^{\mathrm{sim}}(\mathcal{N}) = C(\mathcal{N})$? If not, find counterexamples.

3. How to characterize the unrestricted ID capacity $\mathrm{C}_{\mathrm{ID}}(\mathcal{N})$?

   Need new information-theoretic covering tools (fully quantum channel resolvability?)

Thank you for listening!

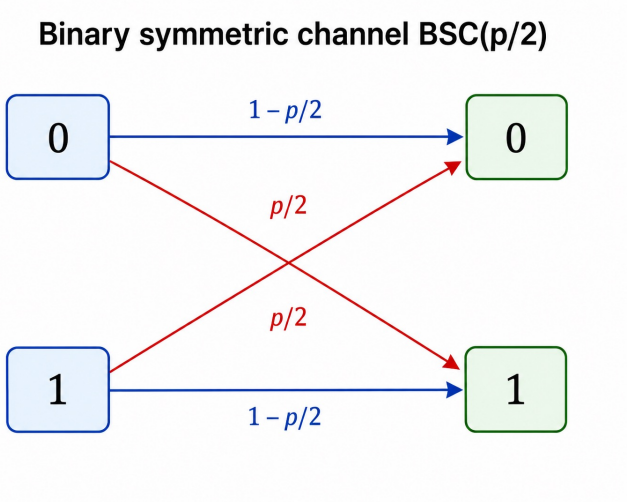
Questions?

Appendix: Proof of Result 2

Proof strategy: achievability

This is straightforward: we have seen that we can always start from a good transmission code and build a good (simultaneous) identification code, where the common POVM is the transmission decoder, with the same asymptotic rate.

We just need to check whether the transmission decoder satisfy the product-basis requirement. To see this, simply notice that $\mathcal{N}_{\mathrm{p}}$ can always be used in a "classical" way, by encoding and decoding in the same orthonormal (e.g., computational) basis: for $\mathrm{x}, \mathrm{y} \in \{0,1\}$,

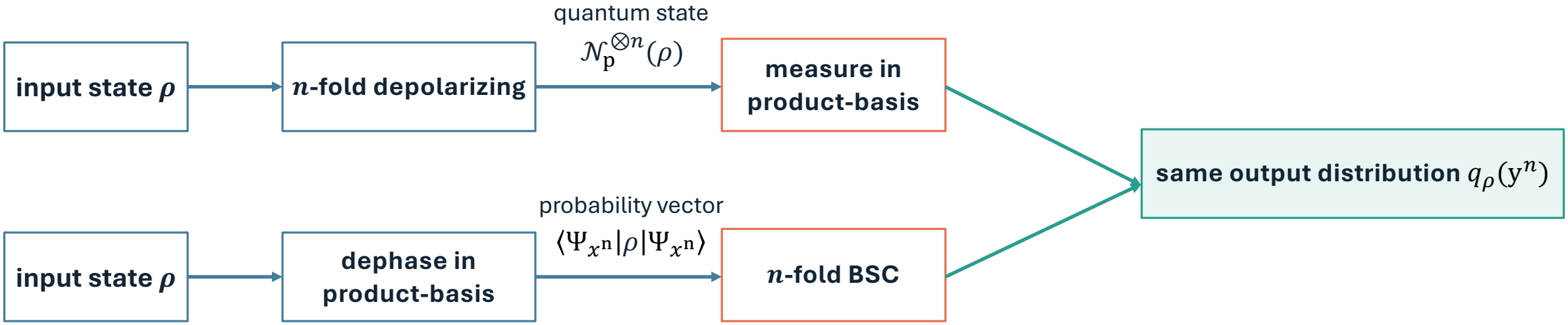


$$W(y|x) = \langle y|\mathcal{N}_{\mathrm{p}}(|x\rangle\langle \mathrm{x}|)|y\rangle = (1-\mathrm{p})\delta_{\mathrm{xy}} + \frac{\mathrm{p}}{2}$$

But the classical capacity of $\mathrm{BSC}(p/2)$ equals to the classical capacity of $\mathcal{N}_{\mathrm{p}}$ [King, 03]: this means that this "classical" way of using the quantum channel is in fact the optimal way! Hence product-basis decoding suffice.

Proof strategy: strong converse

Key observation: first depolarize then perform product-basis measurement is equivalent to first measure (dephasing in the measurement basis) then apply binary symmetric channel,



$$q_\rho(\mathrm{y}^n) = \left\langle \Psi_{\mathrm{y}^{\mathrm{n}}} \middle| \mathcal{N}_{\mathrm{p}}^{\otimes n}(\rho) \middle| \Psi_{\mathrm{y}^{\mathrm{n}}} \right\rangle = \sum_{x^n} \mathrm{W}_{\mathrm{p}/2}^{\otimes n}(y^n|x^n)\ \langle \Psi_{x^{\mathrm{n}}}|\rho|\Psi_{x^{\mathrm{n}}}\rangle$$

$$\mathrm{W}_{\mathrm{p}/2}^{\otimes n}(y^n|x^n) = W_{\mathrm{p}/2}(y_1|x_1) \ldots W_{\mathrm{p}/2}(y_n|x_n)$$

Proof strategy: converse

Find covering of the output space (outputs are now probability distributions q_ρ instead of quantum states σ_ρ), remember:

Previously we require $\frac{1}{2}\|\sigma_i - \sigma_j\|_1 \geq 1 - \lambda_1 - \lambda_2.$ Now we require $\left\|q_{\rho_i} - q_{\rho_j}\right\|_{\mathrm{TV}} \geq 1 - \lambda_1 - \lambda_2.$

The set of all possible output distributions is just the image of the n-fold BSC. Hence the problem reduces to a fully classical problem: covering the output space of the n-fold BSC.

Classical channel resolvability (soft-covering) [Han & Verdú 93, Hayashi, Cheng & Gao 25]:

> Classical channel output $\mathrm{W}^{\otimes n}(P_{\mathrm{X^n}})$ can be approximated by $\mathrm{W}^{\otimes n}(\hat{P}_{\mathrm{X^n}})$ where $\hat{P}_{\mathrm{X^n}}$ is an M-type (empirical distribution) with $\mathrm{M}\sim 2^{\mathrm{nC(W)}}$, so that the channel output space can be covered by the discrete set of empirical distribution outputs whose cardinality is upper bounded by $2^{\mathrm{n}M}\sim 2^{\mathrm{n}2^{\mathrm{nC(W)}}}$.

Taking double logarithms and the asymptotic limit gives $\widetilde{C_{\mathrm{ID}}^{\mathrm{sim}}}(\mathcal{N}) \leq C(\mathrm{BSC}(\frac{\mathrm{p}}{2})) = C(\mathcal{N}_{\mathrm{p}})$.