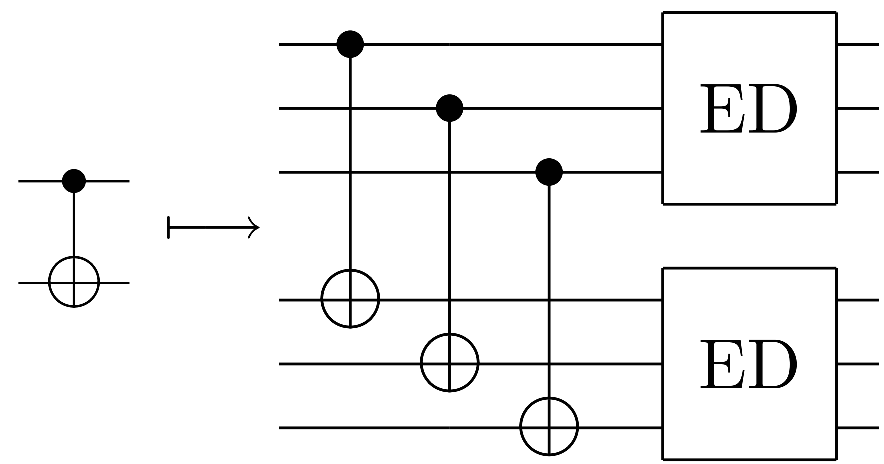
Postselection threshold against biased noise

(A probabilistic mixing lemma and quantum fault tolerance)

Ben W. Reichardt
Caltech

Results

- *Existence* of tolerable noise rates for many fault-tolerance schemes, including:
  - Schemes based on error-*detecting* codes, not just ECCs (Knill-type)
- Tolerable threshold *lower bounds**
  - 0.1% simultaneous depolarization noise†
  - 1.1%, if error model known *exactly*



* Subject to minor numerical caveats

† Versus .02% best lower bound for error-correction-based FT scheme [Aliferis, Cross 2006]

Fault-tolerance motivation:
Fragility of entangled quantum systems

Schrödinger's cat:

$$\frac{1}{\sqrt{2}}\left(|\text{live cat}\rangle + |\text{dead cat}\rangle\right)$$

Fault-tolerance motivation:
Fragility of entangled quantum systems

Schrödinger's cat:

$$\frac{1}{\sqrt{2}}\left(|\text{live cat}\rangle + |\text{dead cat}\rangle\right)$$

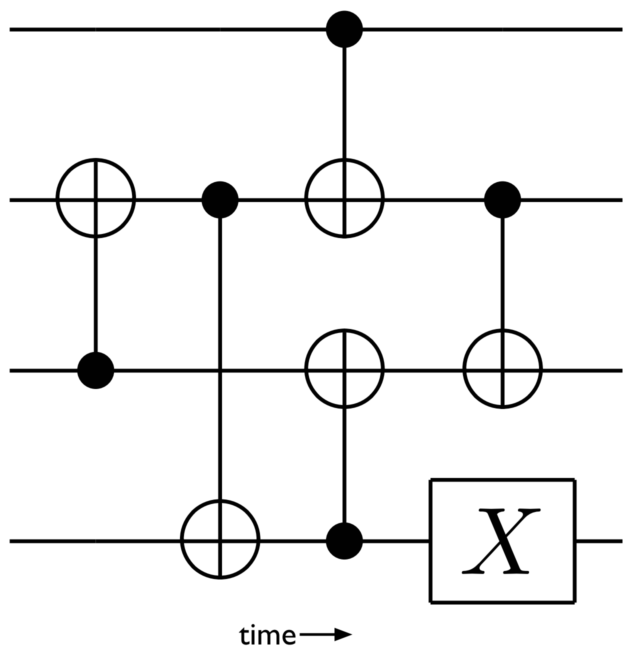
Analogous state in a quantum computer:

$$\frac{1}{\sqrt{2}}\left(|00\cdots 0\rangle + |11\cdots 1\rangle\right)$$

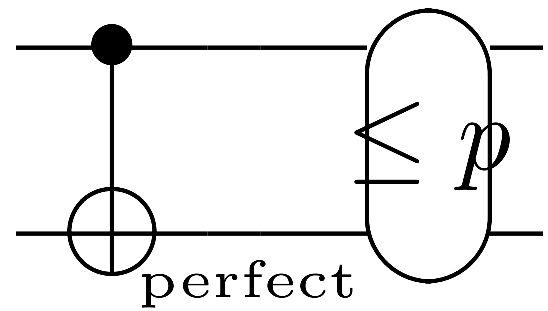
Moral: A single stray photon can collapse whole computer

Fault-tolerance problem

Factor a 1024-bit number using 10^{11} gates



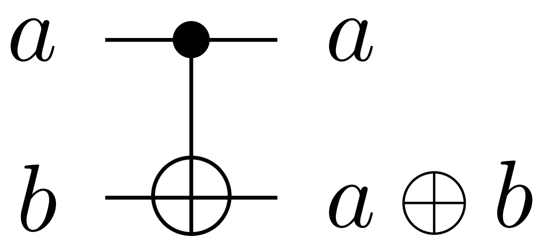
Noise model



For today:

Model a noisy gate as a perfect gate followed by independent *bit-flip* errors — with total error rate at most p

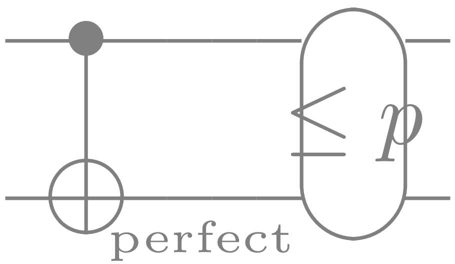
Controlled-NOT gate



flips target if control bit is set

Fault-tolerance intuition

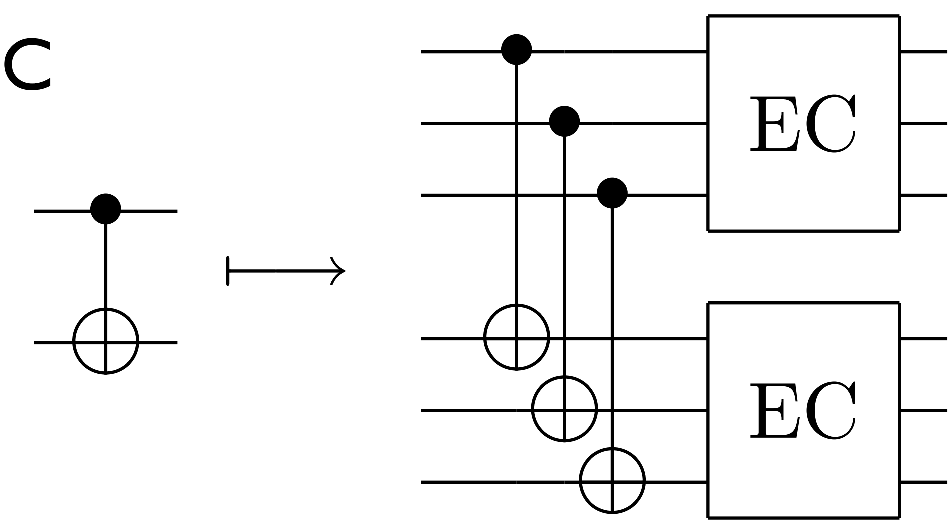
Noise model



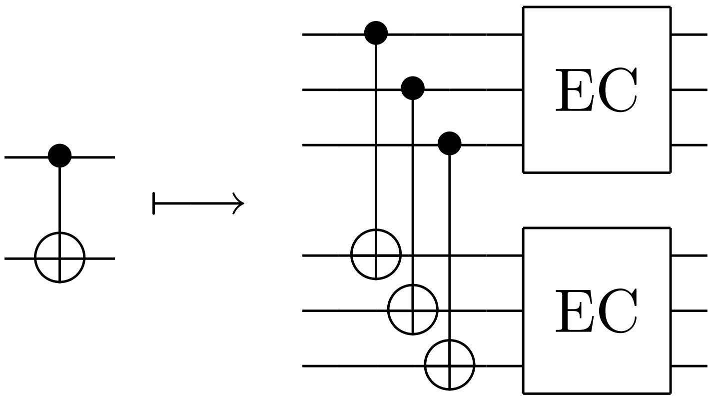
∴ Encode into an error-correcting code

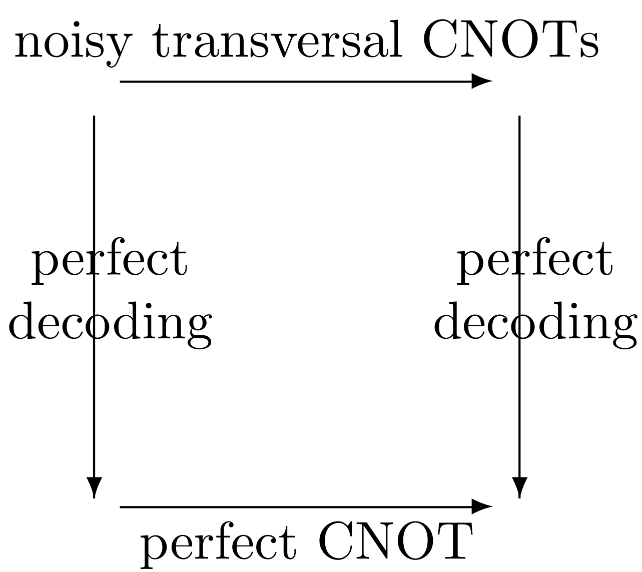
$$
\begin{aligned}
0 &\mapsto 000 \\
1 &\mapsto 111
\end{aligned}
$$

Compute on top of the ECC

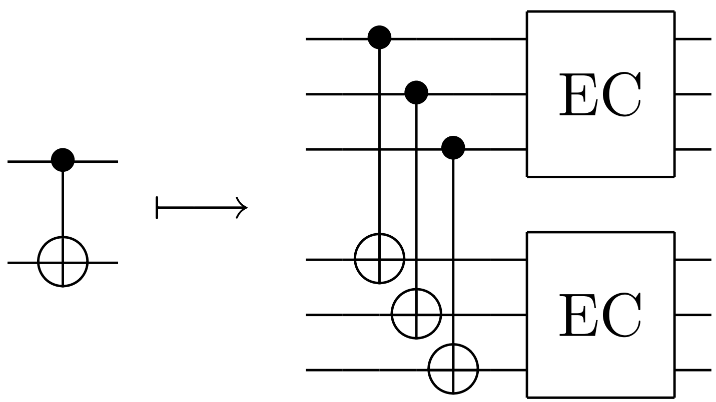


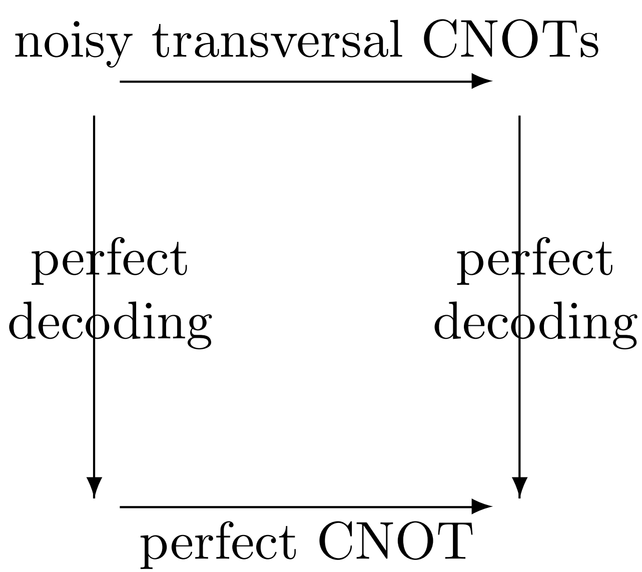
Fault-tolerance intuition



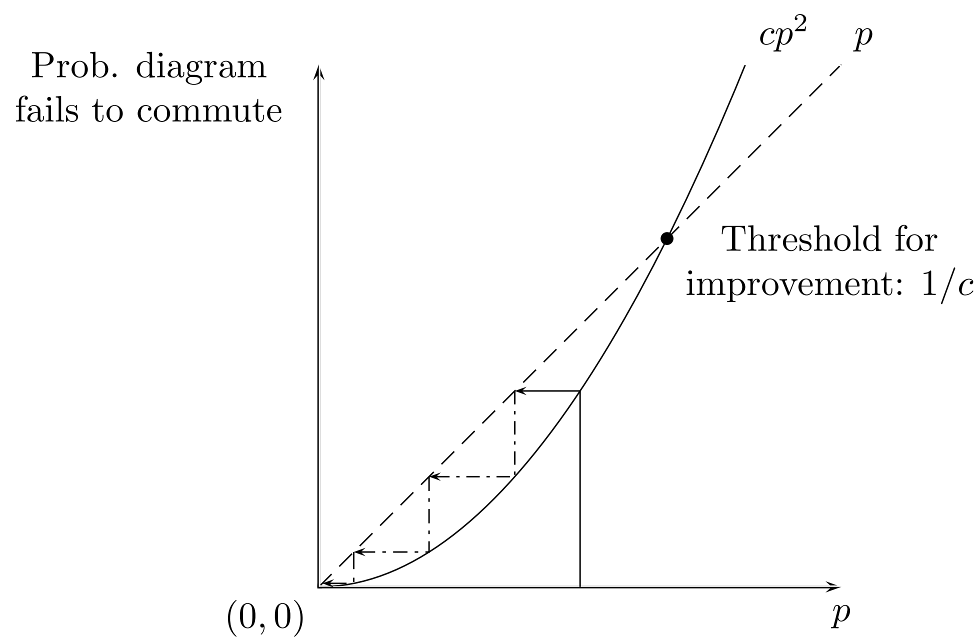


Fault-tolerance intuition





Improved reliability beneath a constant threshold



Tolerable noise threshold results

Shor

Knill/Laflamme/Zurek

Gottesman

Terhal/Burkard

Proofs

Aharonov & Ben-Or * **Kitaev**

prove positive tolerable noise rate (1997)

for codes of distance d≥5

Aliferis/Gottesman/Preskill, Reichardt (2005)

first numerical threshold lower bounds, threshold for distance-3 codes

Estimates & Simulations

Steane ('02-'04)

develops efficient FT scheme, runs extensive simulations

estimates 10^{-3} threshold noise rate with reasonable overhead

Knill ('04-'05) *

developed FT scheme based on postselection — error detection, not error correction (d=2 codes)

estimates 3-6% threshold with high overhead

Zalka

Preskill

Svore/Cross/Chuang/Aho

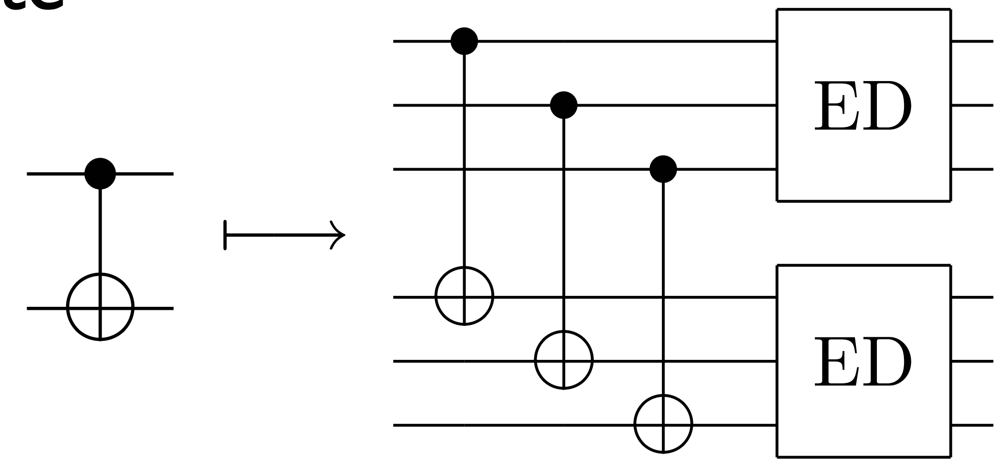
Szkopek et al.

Svore/Terhal/DiVincenzo

Today: Positive threshold for postselection-based FT scheme

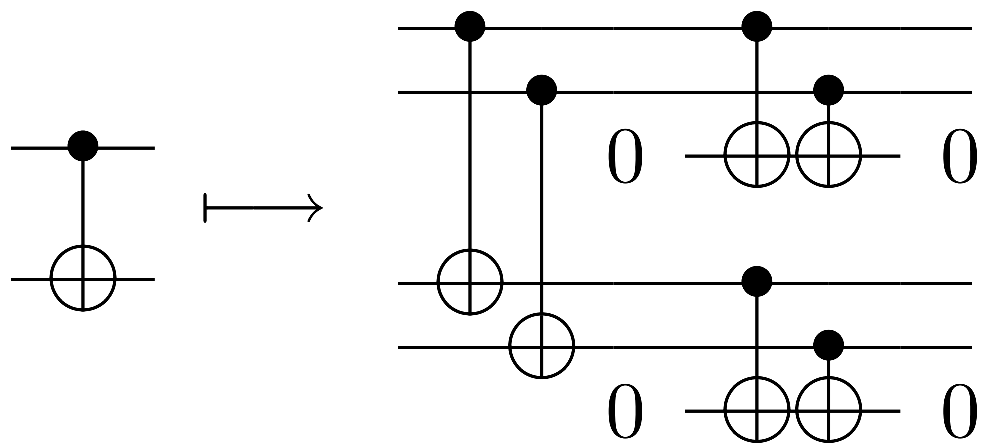
Fault tolerance based on error *detection*

CNOT gate



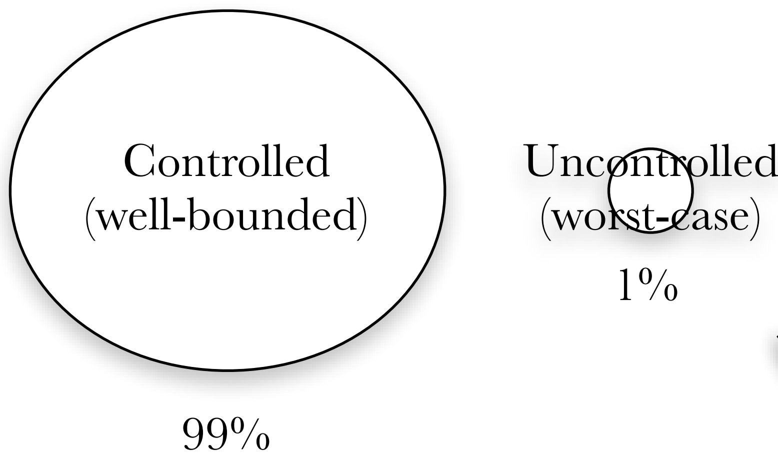
Fault tolerance based on error *detection*

CNOT gate



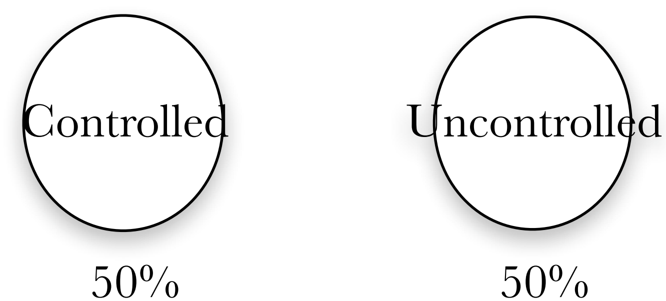
- In simulations, tolerates 10x higher noise rates than error-correction-based FT schemes
- But previously, no proven positive threshold at all!
- Note: Overhead is substantial, but theoretically efficient

Renormalization frustrates previous proofs



Most of the time, errors are detected —
but (counterintuitively) survival probability for
uncontrolled portion could be much *higher*

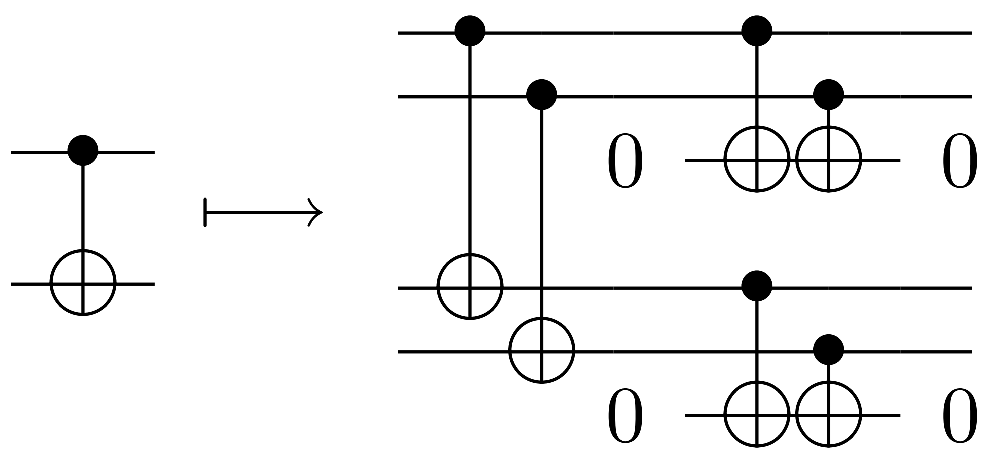
Proofs based on controlling events most of the time,
with occasional failures



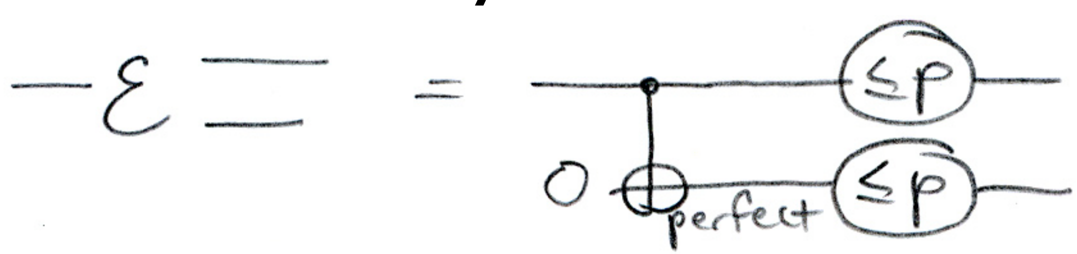
Uncontrolled fraction of probability mass increases exponentially after renormalization!

Proof intuition

CNOT gate

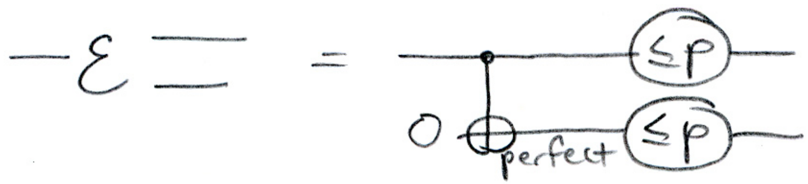


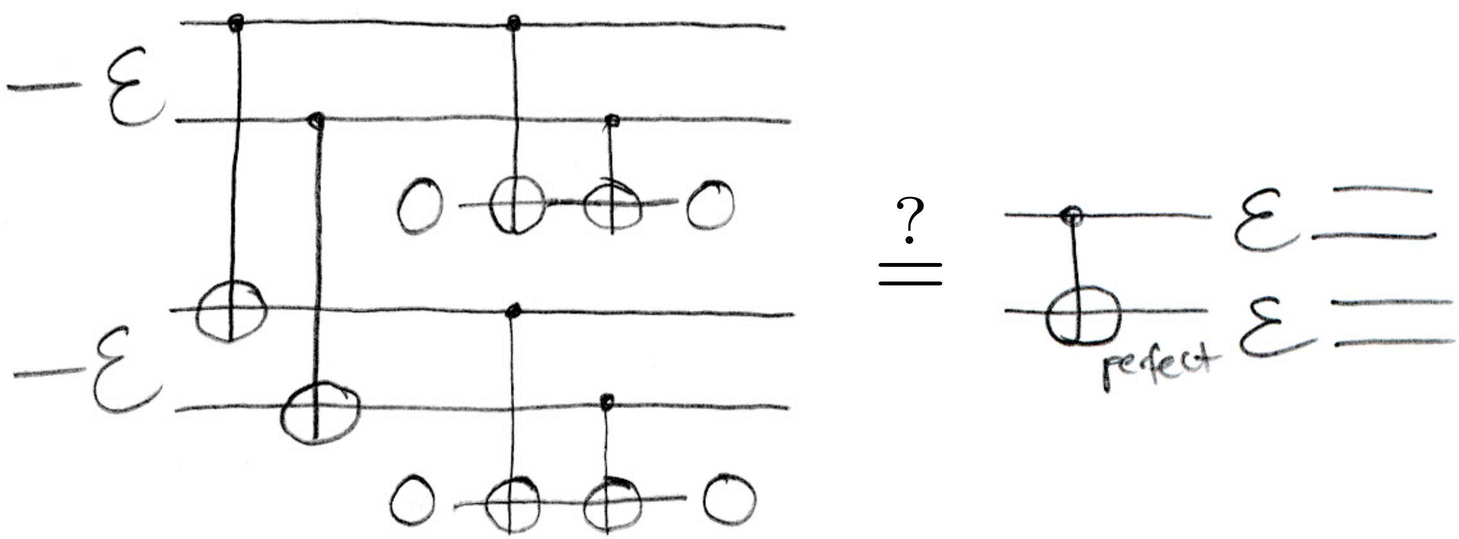
Notation: Noisy encoder



Proof intuition

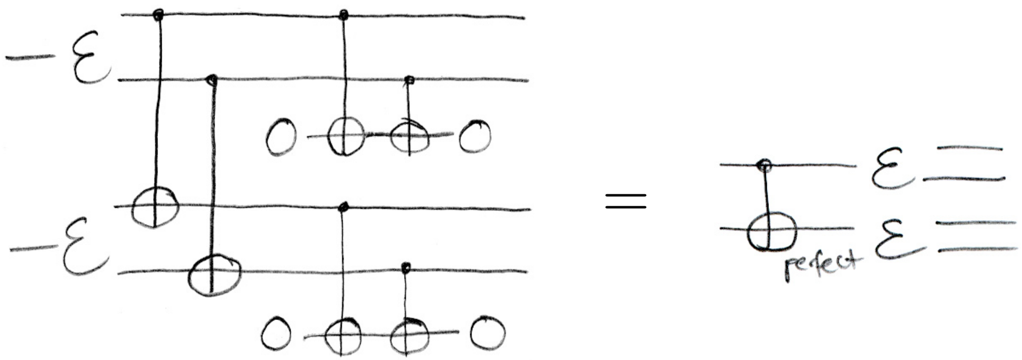
Notation



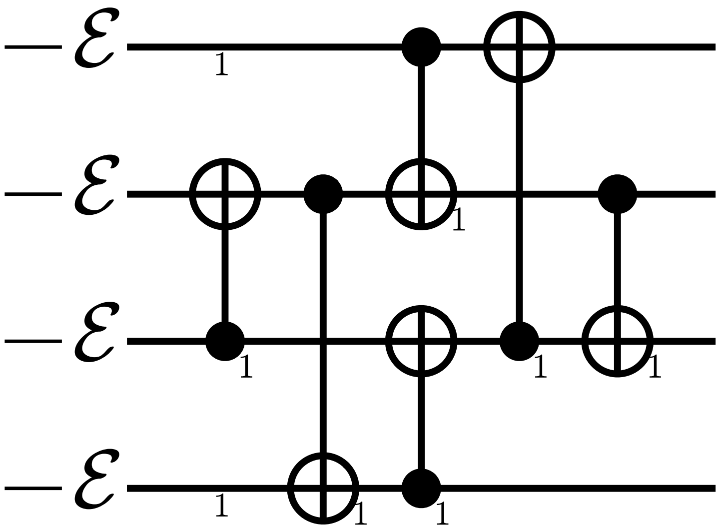


bitwise-independent errors
preceding encoded CNOT gate

bitwise-independent errors
following encoded CNOT gate

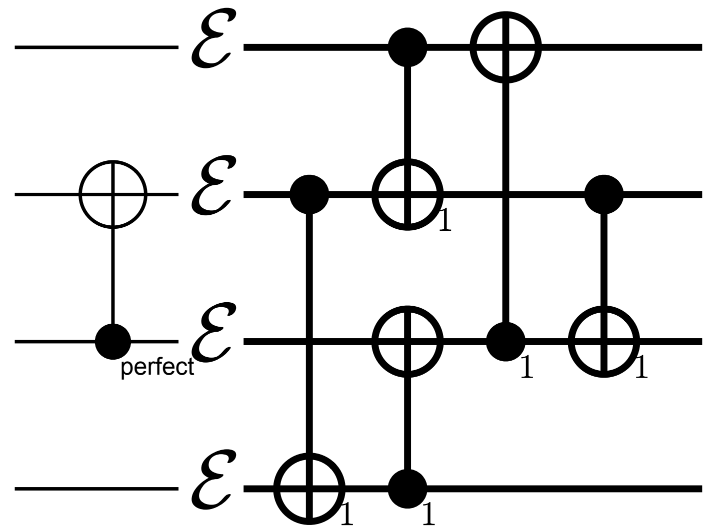


would be nice, since



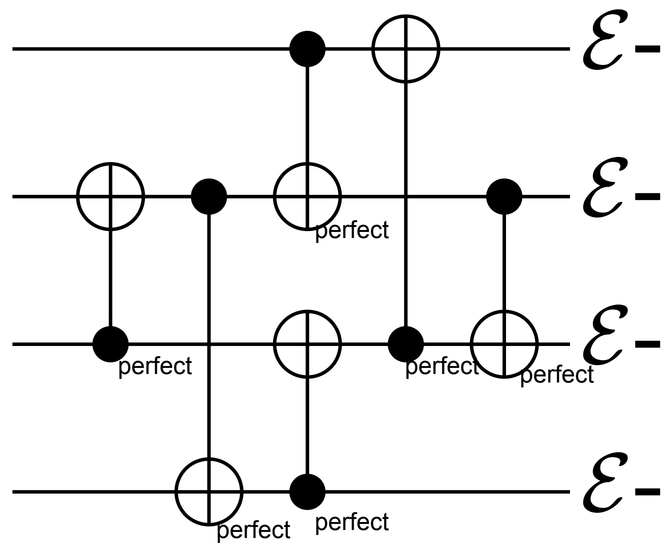
encoded FT circuit

$=$



$= \cdots =$

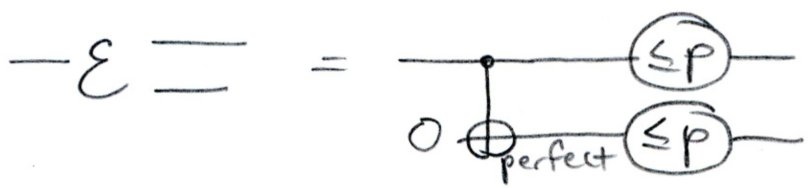
$=$

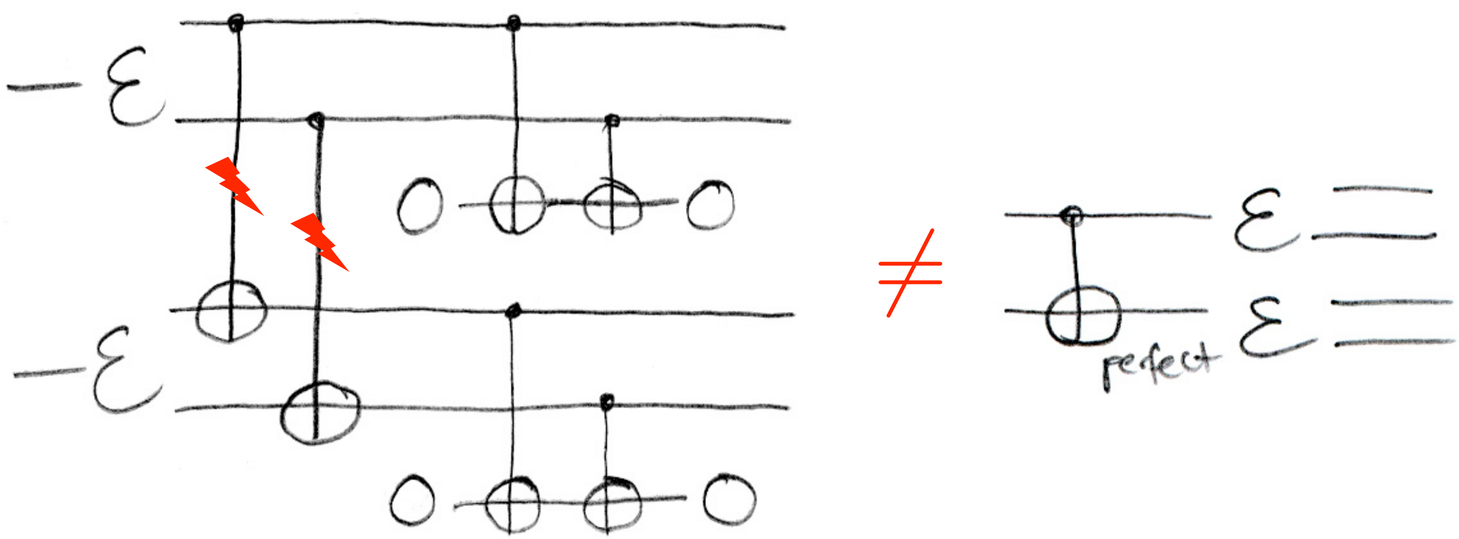


perfect circuit

Proof intuition

Notation



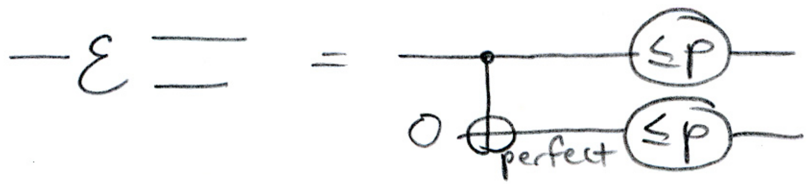


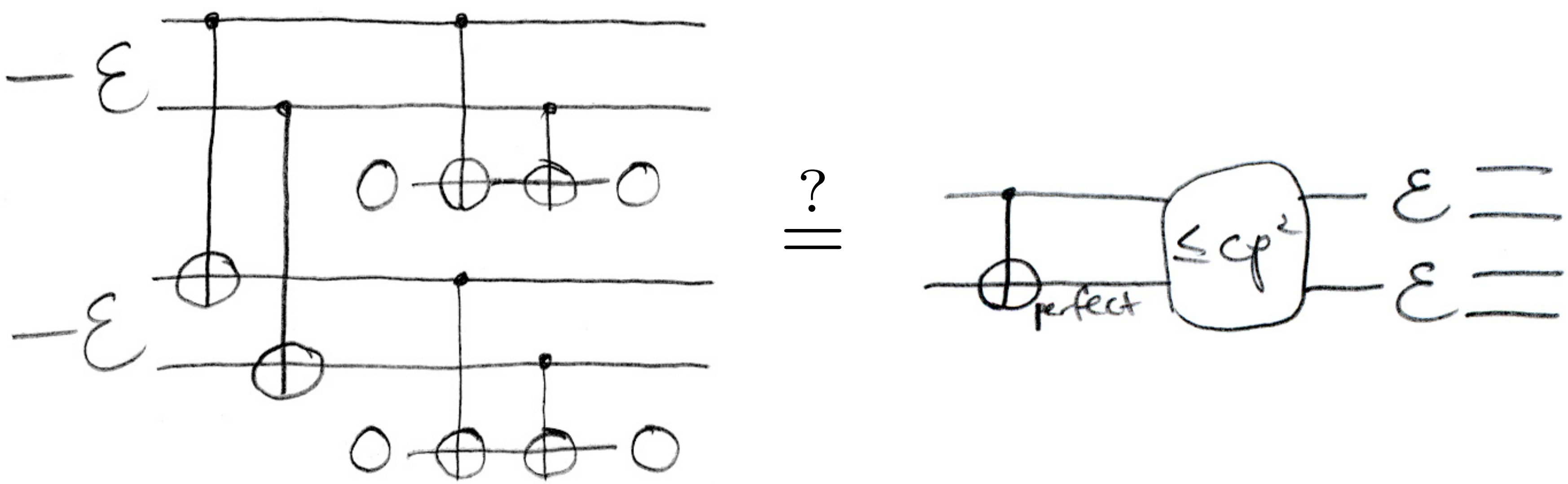
P[XXXX] ~ p^2

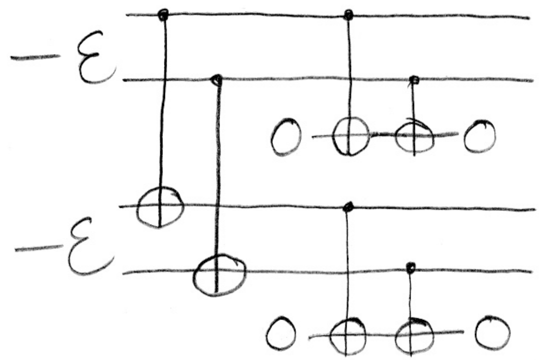
P[XXXX] = p^4

Proof intuition

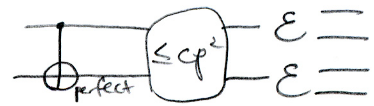
Notation



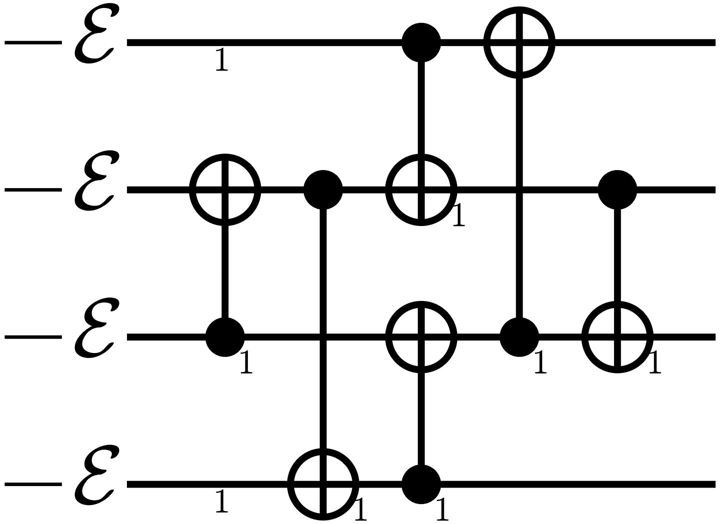




=



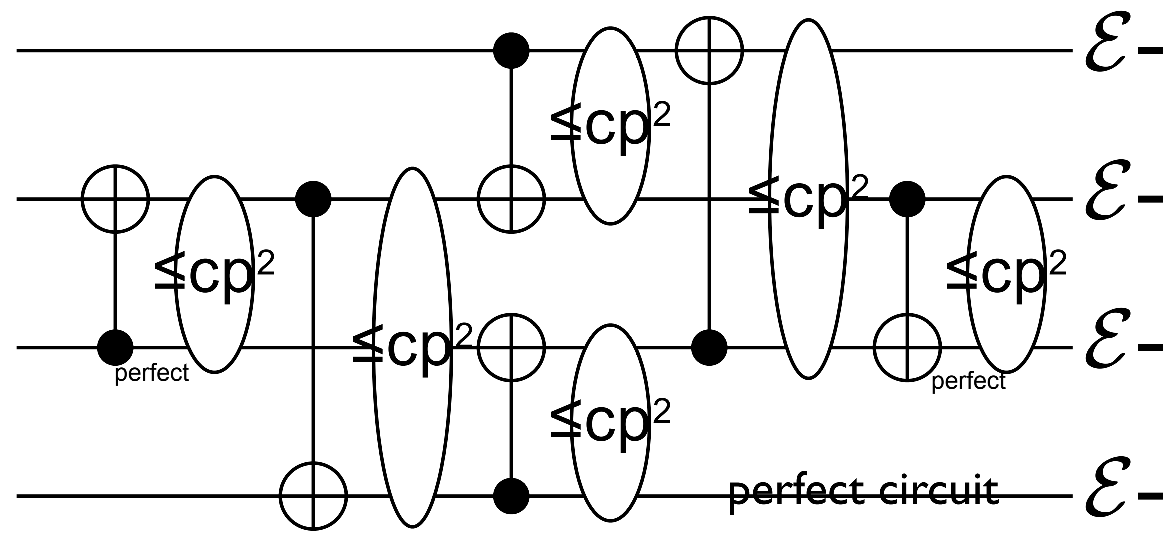
would be nice, since



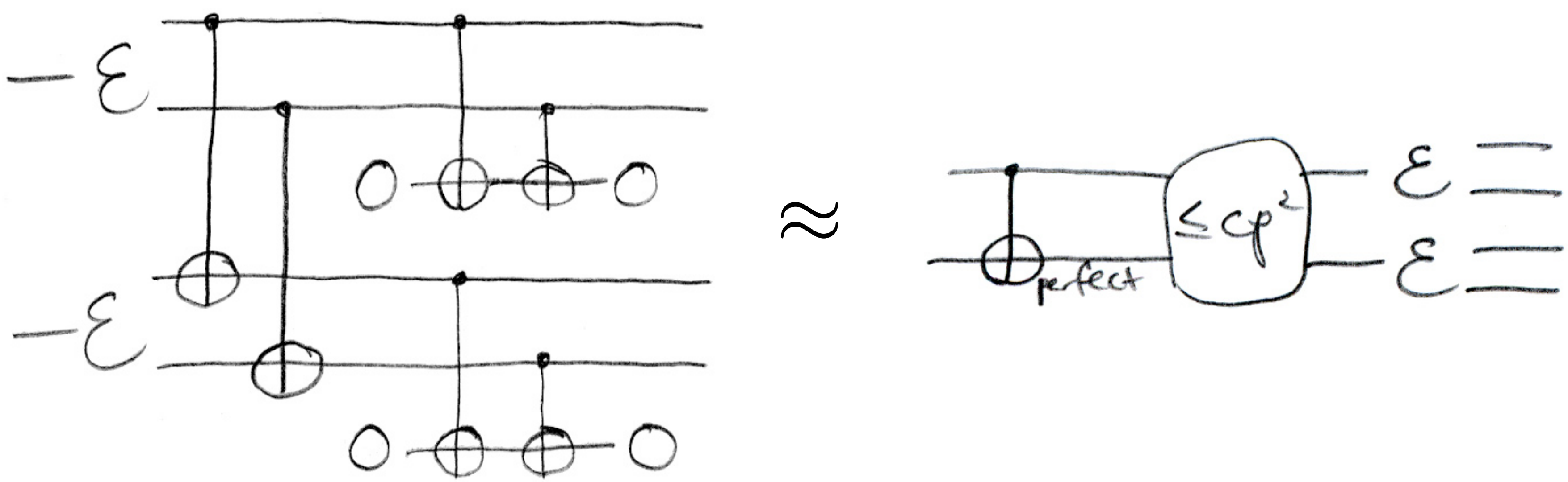
$= \cdots =$

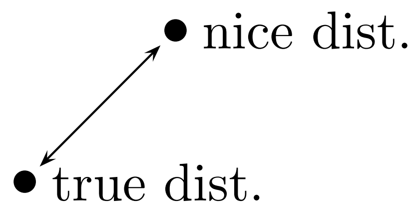
encoded FT circuit

=

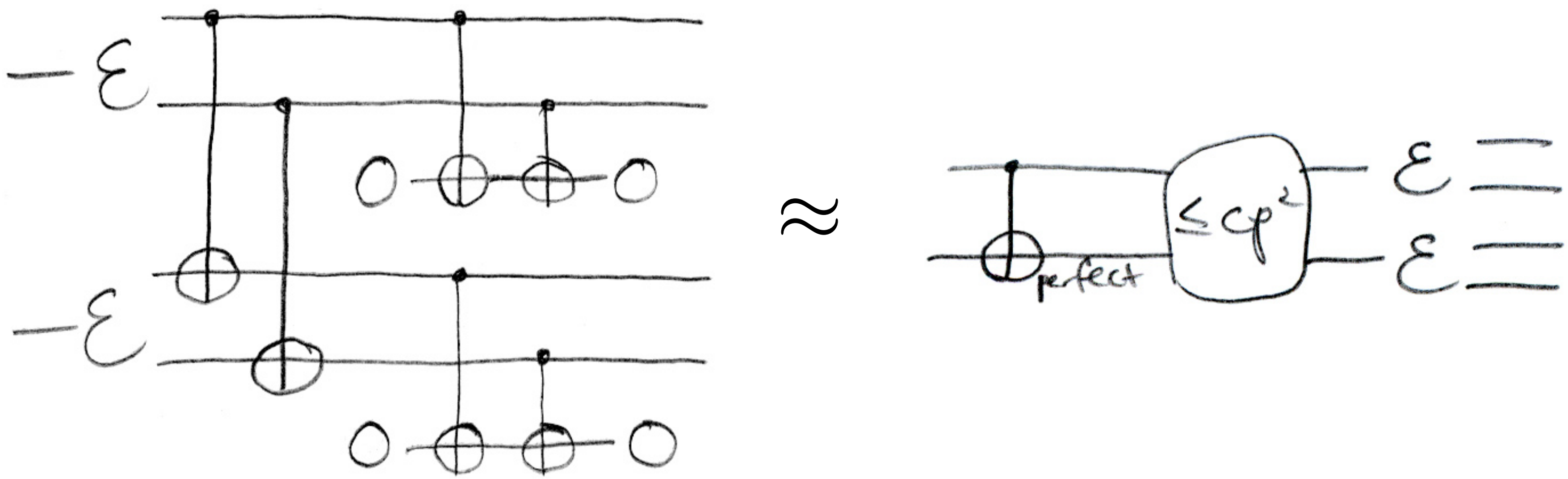


Proof intuition



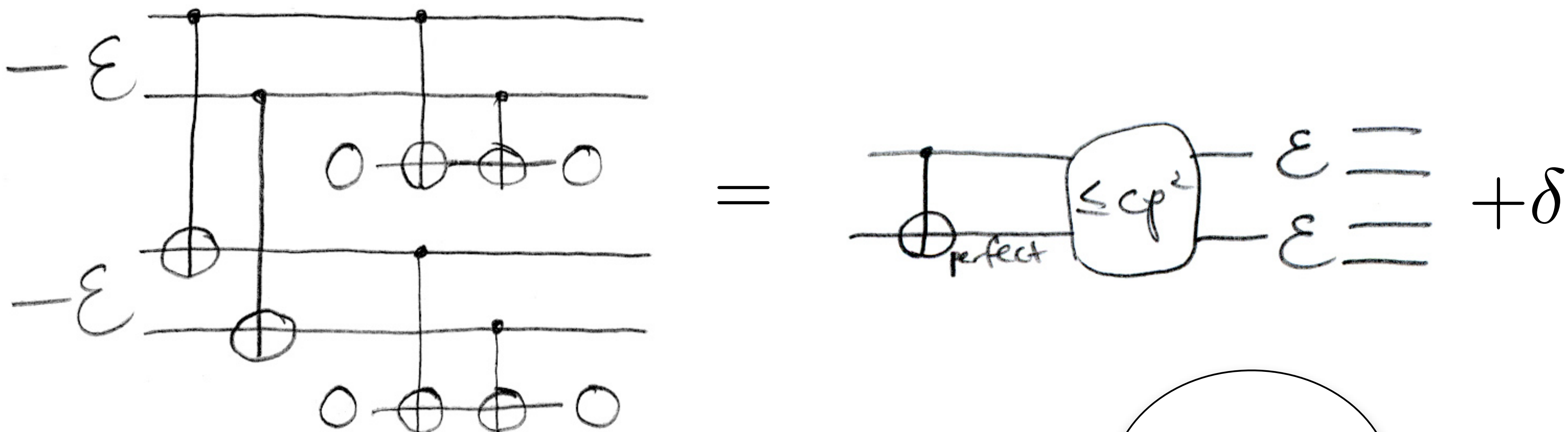


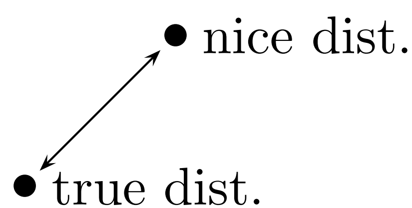
Error orders are correct



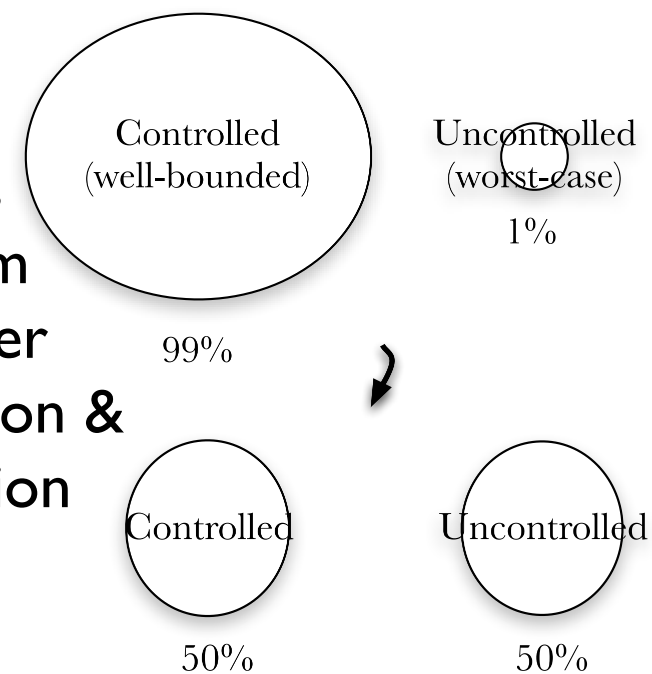
Error	Probability
IIII	$\Theta(1)$
IIIX, IIXI, IXII, XIII	$O(p)$
XXII, IIXX, XXXX, XIXI, XIIX, IXXI, IXIX	$O(p^2)$
IXXX, XIXX, XXIX, XXXI	$O(p^3)$

Proof intuition

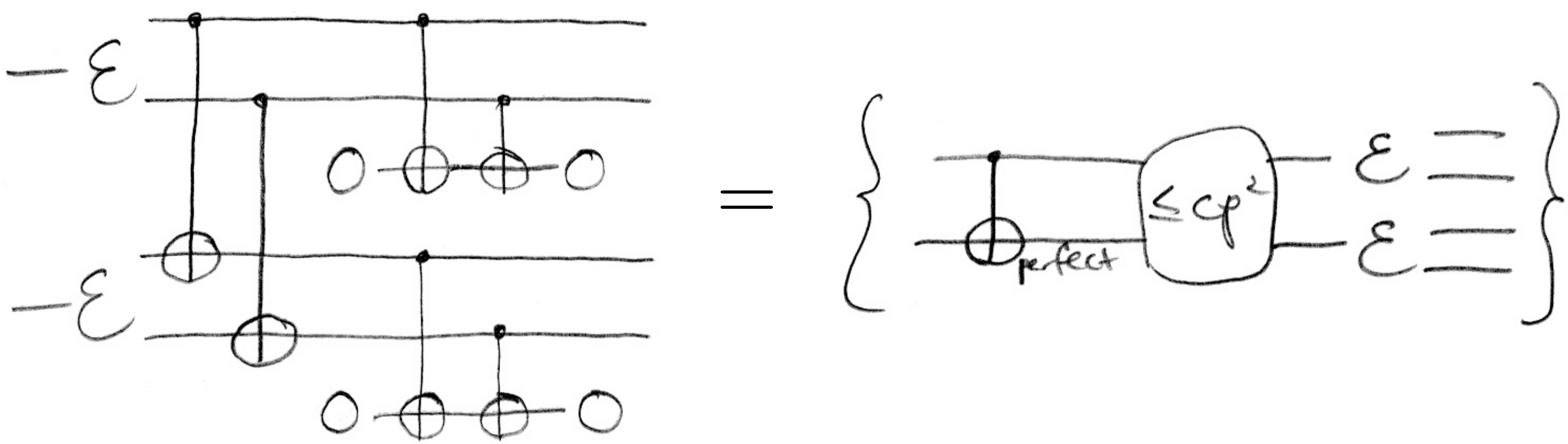


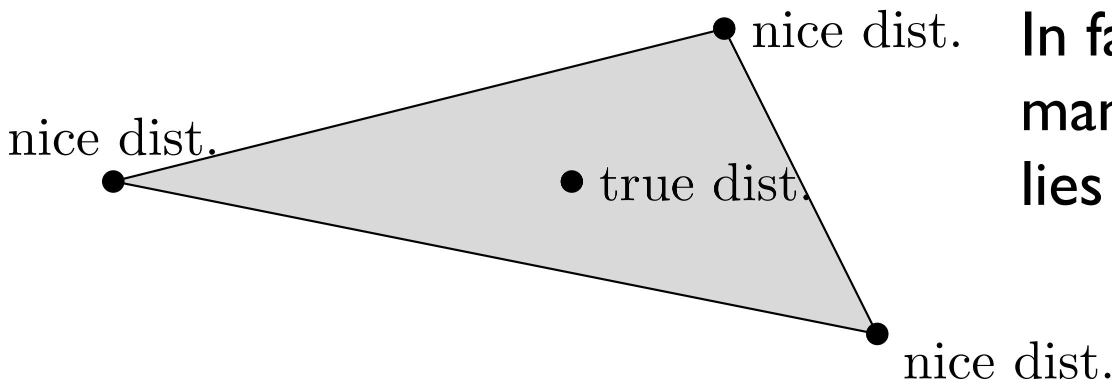


But this gives same problem as before, after error detection & renormalization



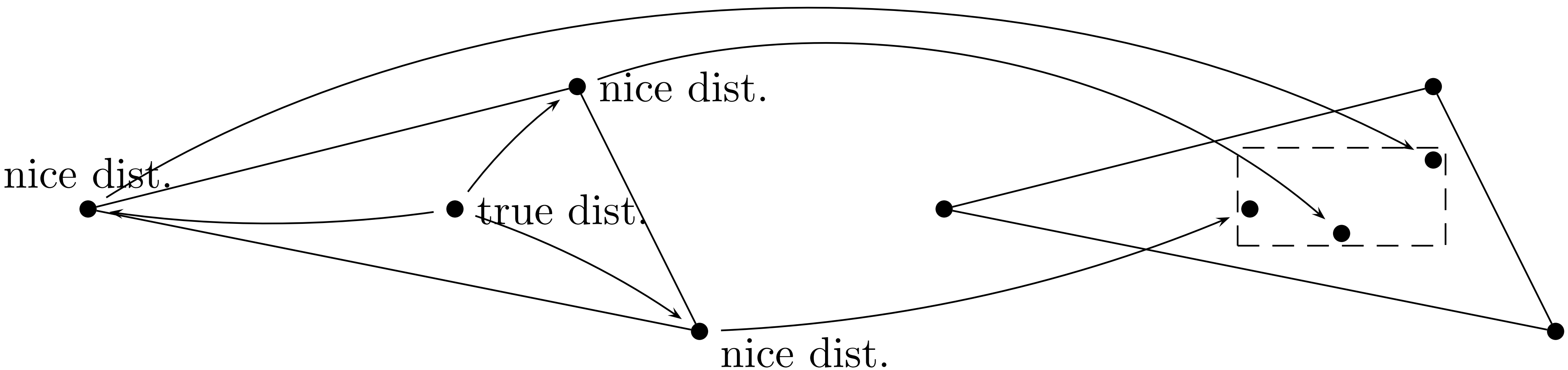
Proof intuition



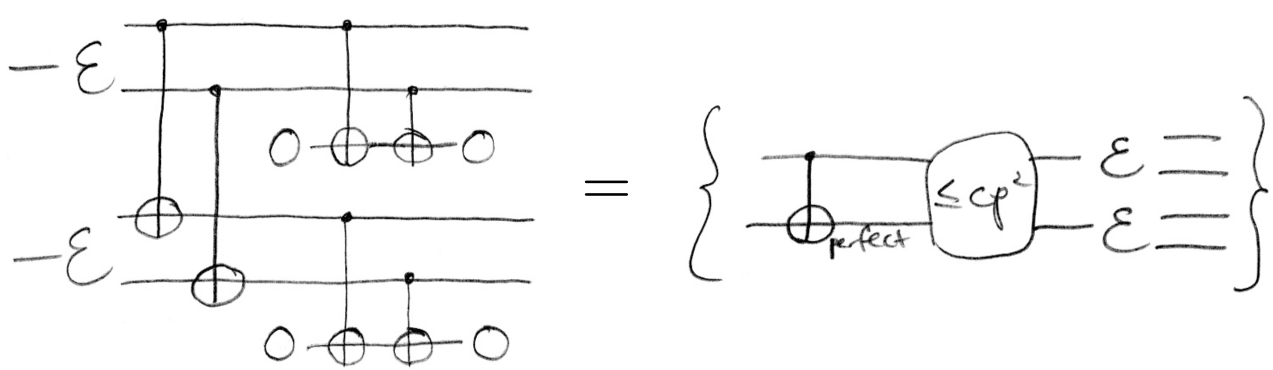


In fact, true distribution is close to many nice (RHS) distributions, and lies in their *convex hull*

Induction step



Analysis of the next encoded CNOT gate proceeds by *picking* one of the vertices — a nice distribution — then applying the CNOT mixing lemma:

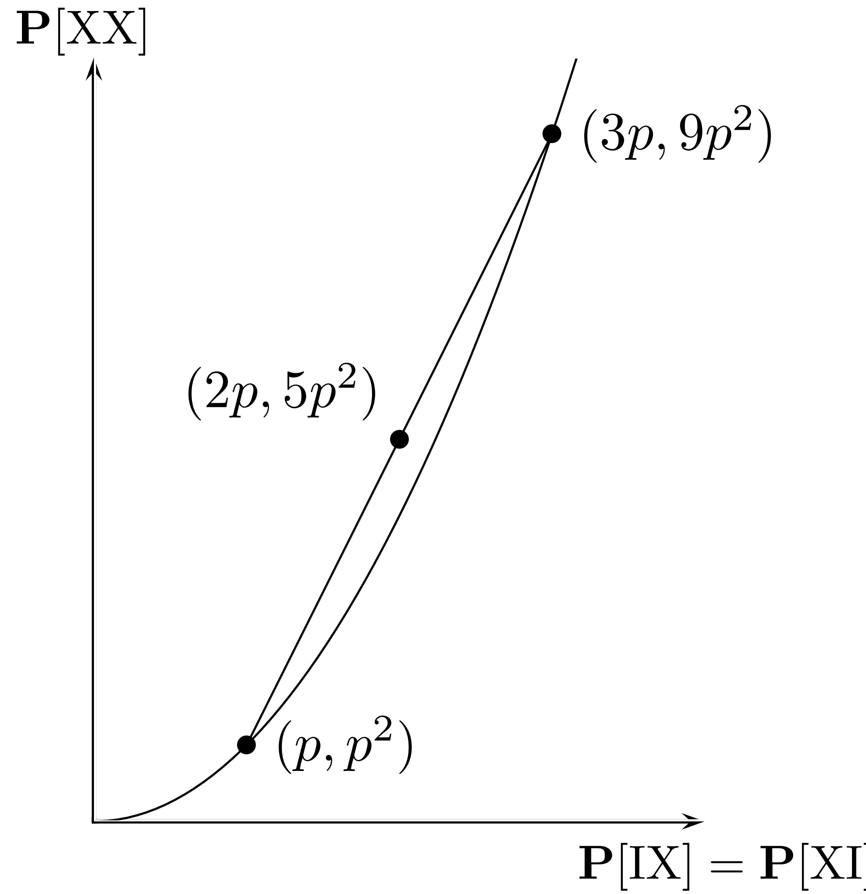


Each output distibution can again be rewritten as mixture of nice distributions, etc.

Two-bit example of distribution mixing

$$\begin{aligned} \mathbf{P}[\mathrm{II}] &= 1 - 4p - 5p^2 \\ \mathbf{P}[\mathrm{IX}] = \mathbf{P}[\mathrm{XI}] &= 2p \\ \mathbf{P}[\mathrm{XX}] &= 5p^2 \end{aligned}$$

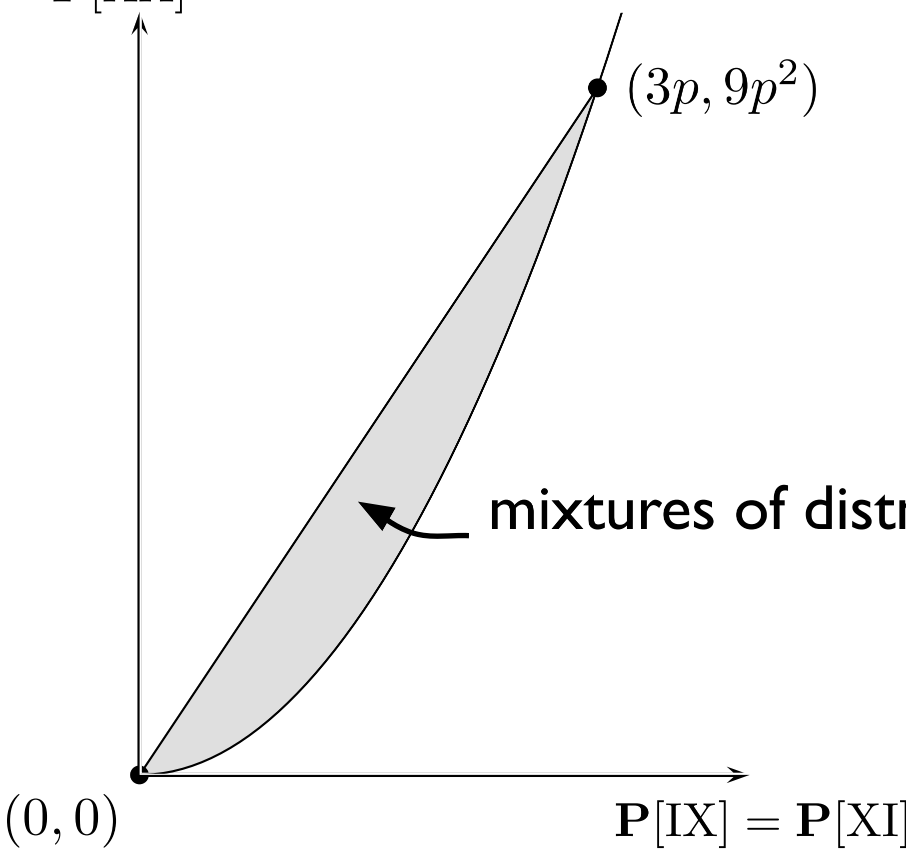
slight positive correlation



$$\begin{pmatrix} 2p \\ 5p^2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} p \\ p^2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 3p \\ 9p^2 \end{pmatrix}$$

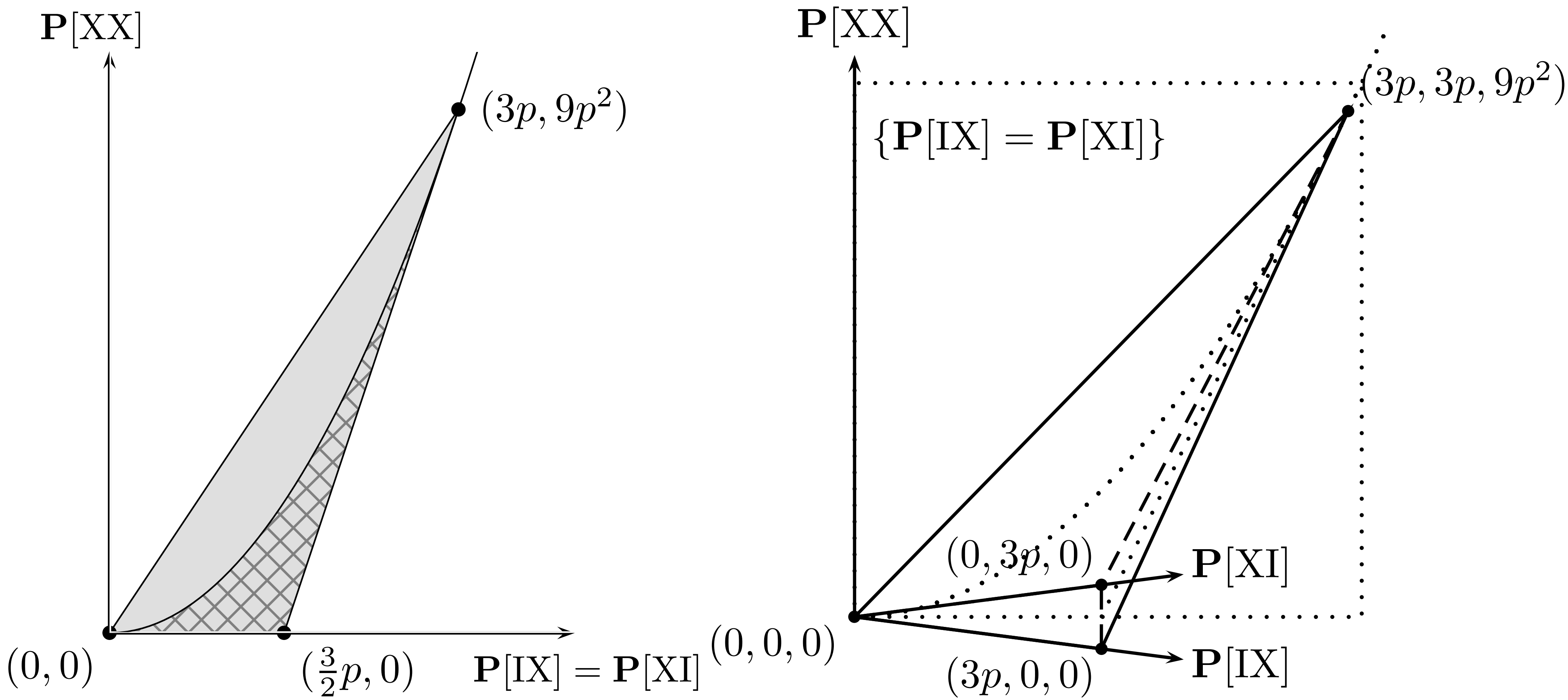
bitwise-independent errors

Two-bit example of distribution mixing



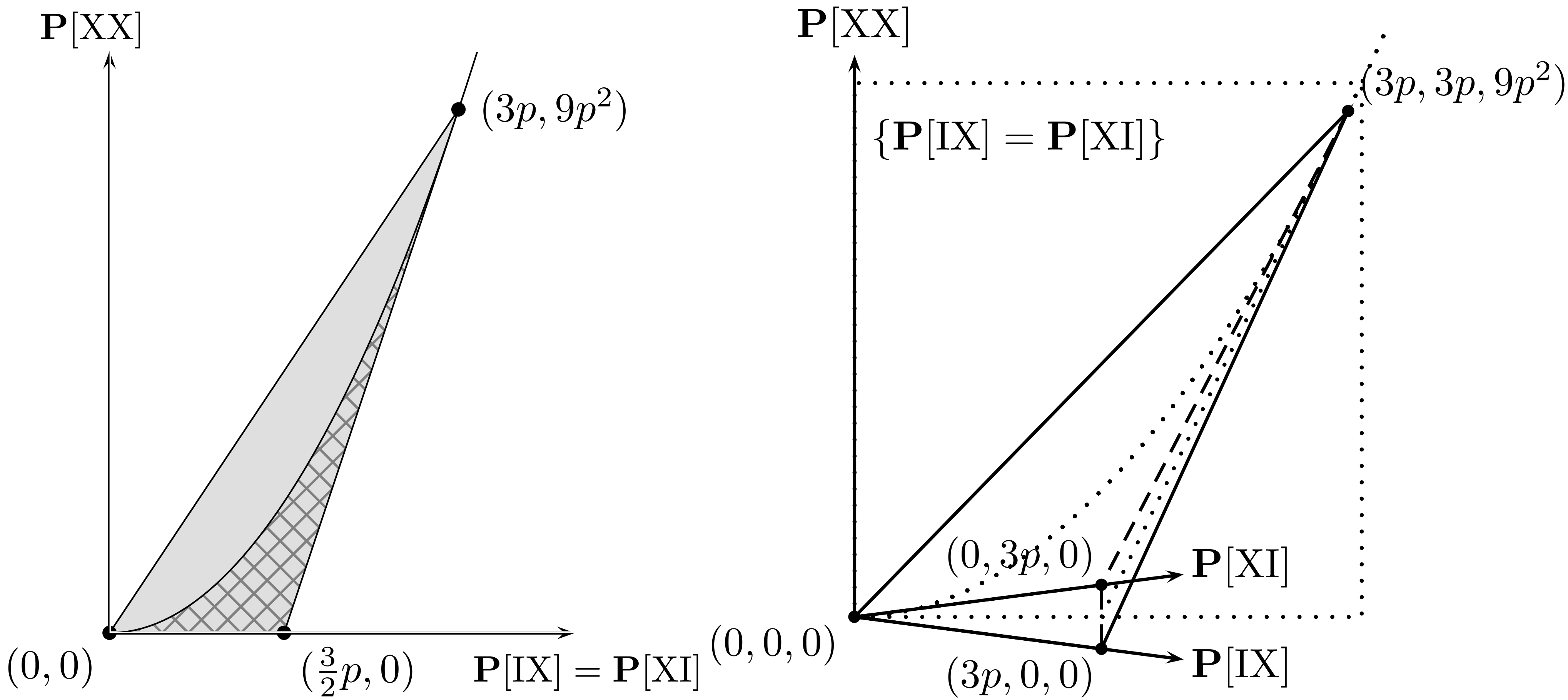
mixtures of distn's having bitwise-independent noise at rate ≤3p

Two-bit example of distribution mixing



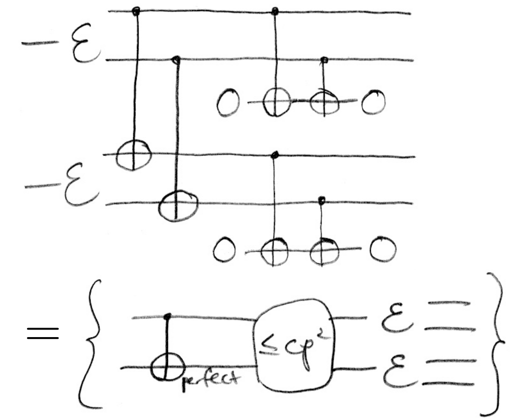
Mixtures (convex combinations) of distributions having bitwise-independent noise at rate ≤3p

Two-bit example of distribution mixing

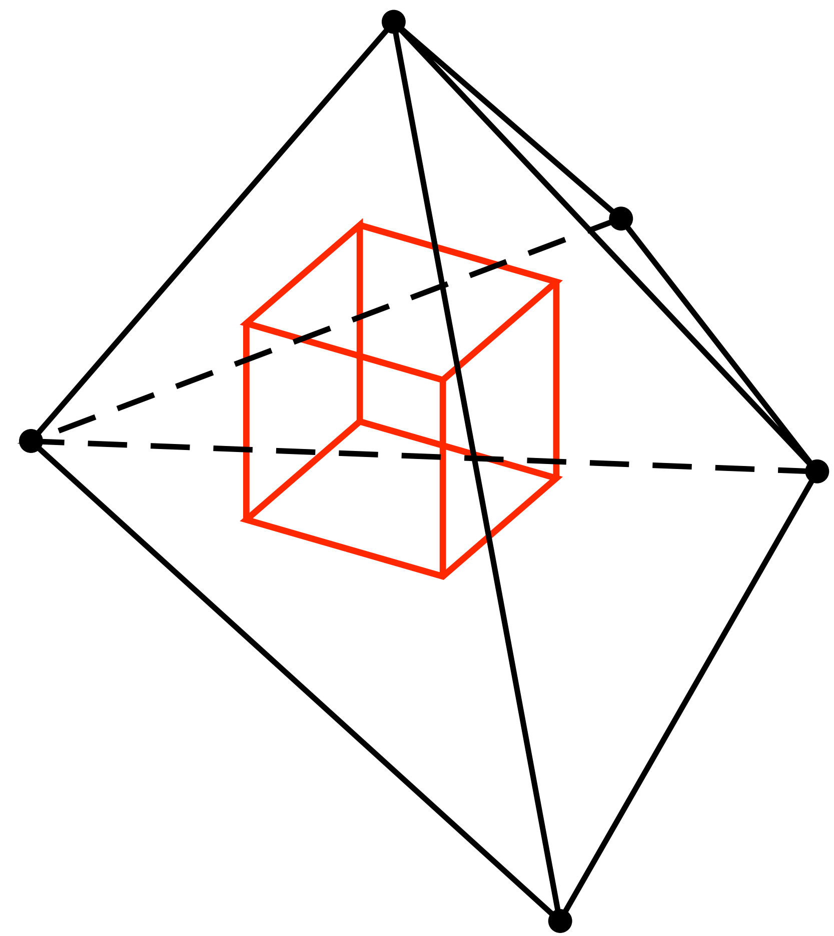


But convex hull of "nice" distributions is much more complicated, in many more dimensions; can't characterize it exactly (even numerically — number of faces is exponential in dimension)

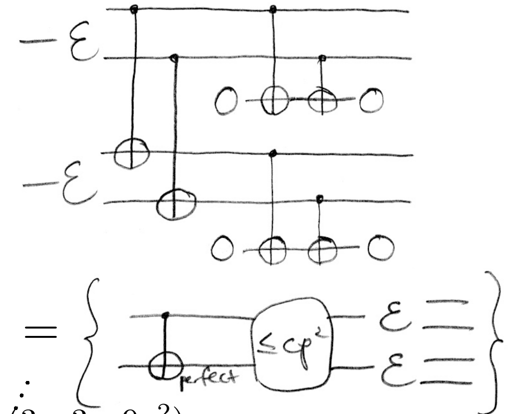
Two methods for showing mixing



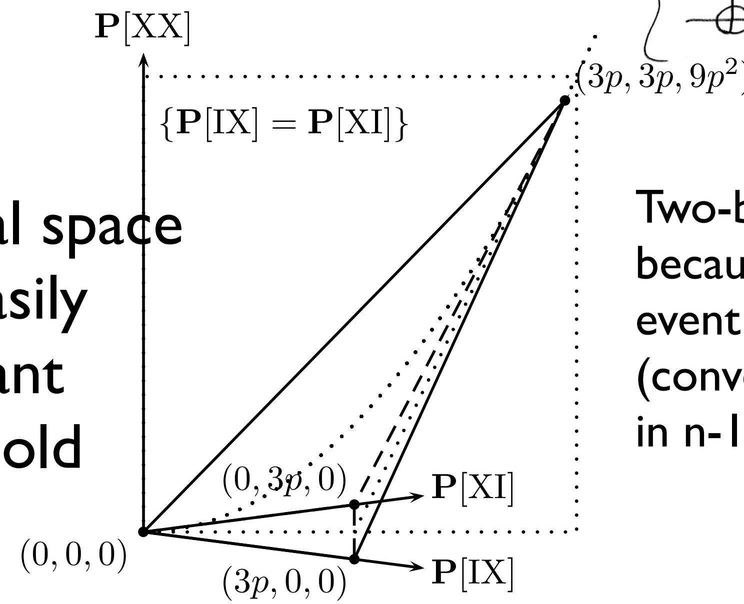
1. We are given upper and lower bounds for each coordinate of the distribution… So use a linear program to check that each vertex of the hypercube lies in the convex hull of extremal "nice" distributions. (Computationally expensive in high dimensions.)



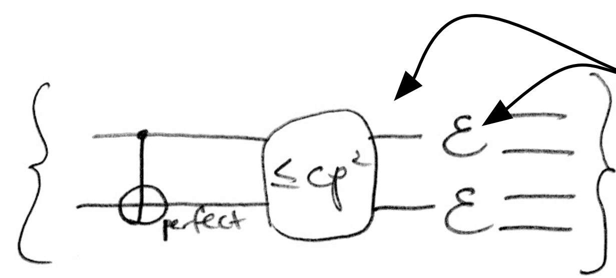
Two methods for showing mixing



2. Map to a higher-dimensional space in which convex hull can be easily characterized. (Loses a constant factor, but sufficient for threshold *existence* proof.)

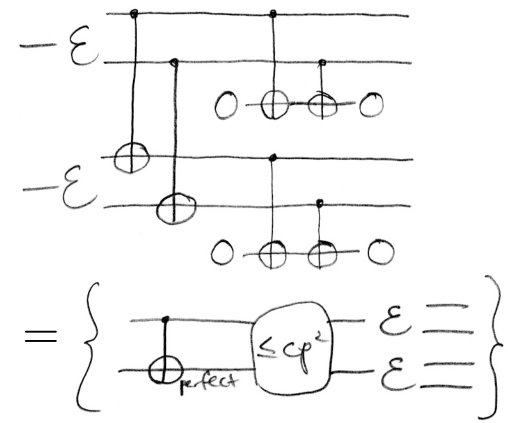


Two-bit case is simple because every error event has distinct effect (convex hull of n points in n-1 dimensions)



Now, different events can lead to same error — convex hull no longer a simplex

Two methods for showing mixing



A point $(q_1, \ldots, q_n) \in [0,1]^n$ corresponds to a bitwise-independent distribution over $\{0,1\}^n$, in which the probability of x is $\prod_{i=1}^n q_i^{x_i}(1-q_i)^{1-x_i}$. Define the lattice ordering $y \preceq x$ for $x, y \in \{0,1\}^n$ if considered as indicators for subsets of $[n]$, $x \subseteq y$.

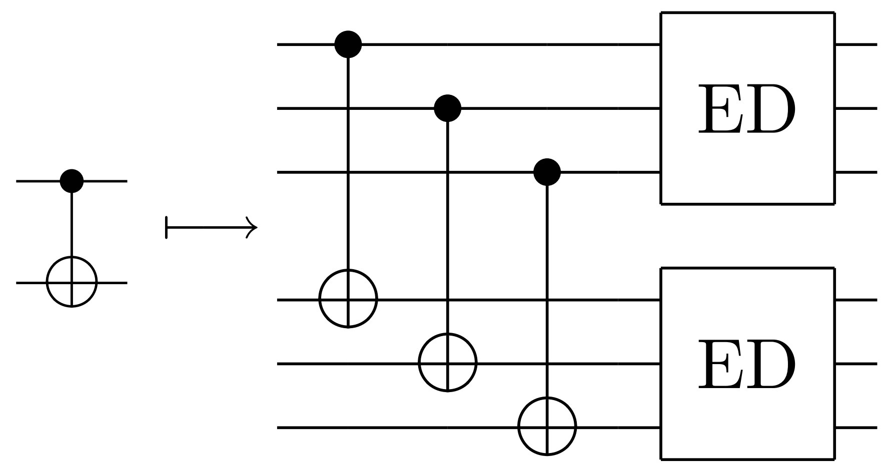
Mixing Lemma. *The convex hull, in the space of distributions over n-bit strings, of the 2^n bitwise-independent distributions $\{0,p_1\} \times \{0,p_2\} \times \cdots \times \{0,p_n\}$ is given exactly by those $\mathbf{P}[\cdot]$ satisfying the inequalities, for each $x \in \{0,1\}^n$:*

$$\sum_{y \preceq x} (-1)^{|x \oplus y|} \frac{\mathbf{P}[\{z \preceq y\}]}{p(\{z \preceq y\})} \geq 0 \quad ,$$

where $p(\{z \preceq y\}) = \prod_{i=1}^n \delta_{y_i,1} p_i$, i.e., the probability of $\{z : z \preceq y\}$ in the distribution $(p_1, \ldots, p_n)$.

Results

- *Existence* of tolerable noise rates for many fault-tolerance schemes, including:
  - Schemes based on error-*detecting* codes, not just ECCs (Knill-type)
- *Numerical* threshold lower bounds*
  - 0.1% simultaneous depolarization noise†
  - 1.1%, if error model known *exactly*



* Subject to minor numerical caveats

† Versus .02% best lower bound for error-correction-based FT scheme [Aliferis, Cross 2006]