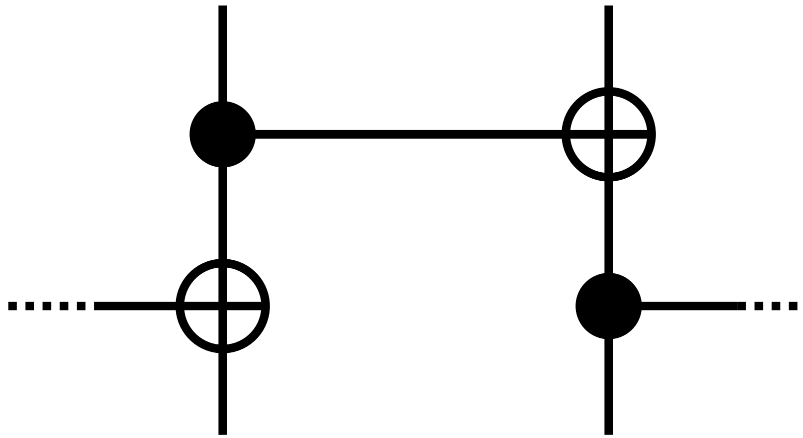
Tensor networks and coding theory: Polar and branching-MERA codes



Andy Ferris

ICFO – Institute of Photonic Science (Spain)

Max Planck Institute for Quantum Optics (Germany)

with

David Poulin

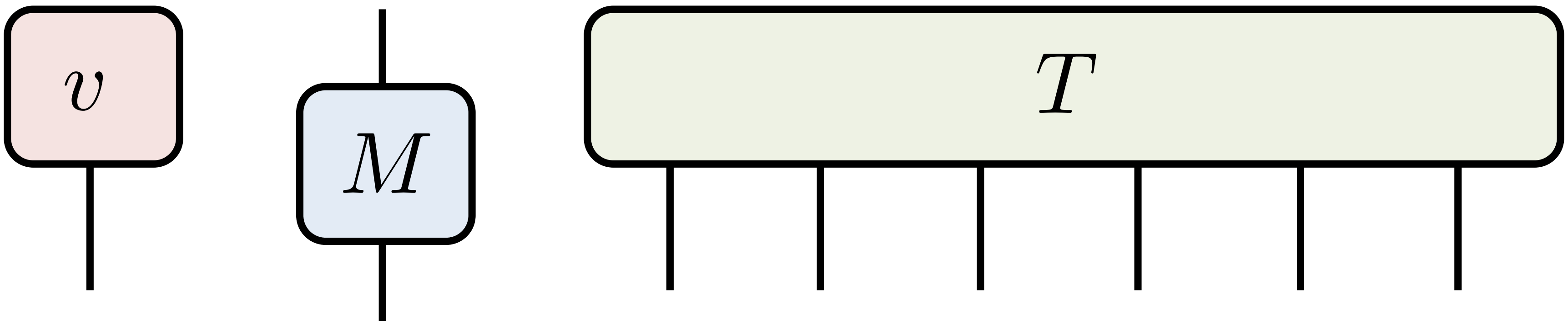
Université de Sherbrooke (Canada)

Why tensor networks?

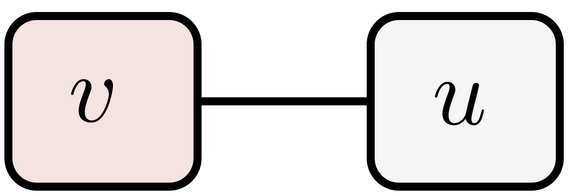
- Tensors networks are tools for efficiently representing large objects
  - Example 1: Can **decompose a large object**, like a quantum many-body wavefunction, as a product of small tensors (e.g. MPS/DMRG, PEPS, MERA).
  - Example 2: Tensor products are *linear* and are perfect for **representing circuits and channels**

Tensor network diagrams

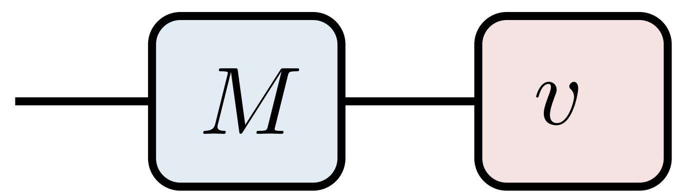
Tensor with n indices is drawn with n "legs"



Tensor contractions are represented by joining legs



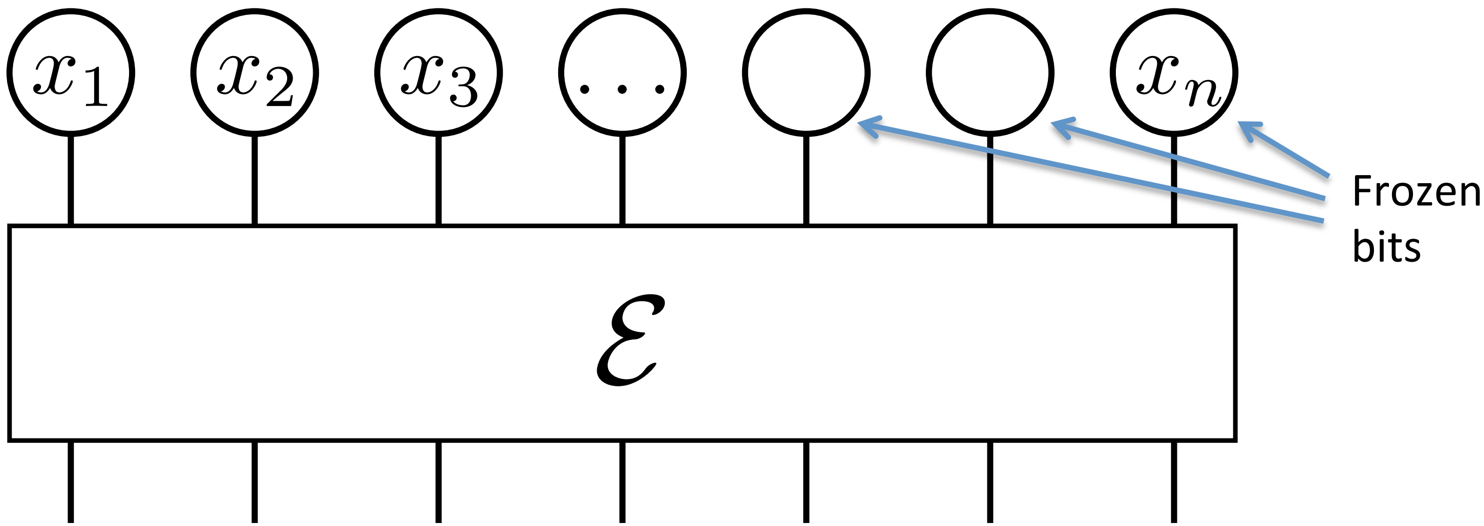
dot-product



matrix-vector product

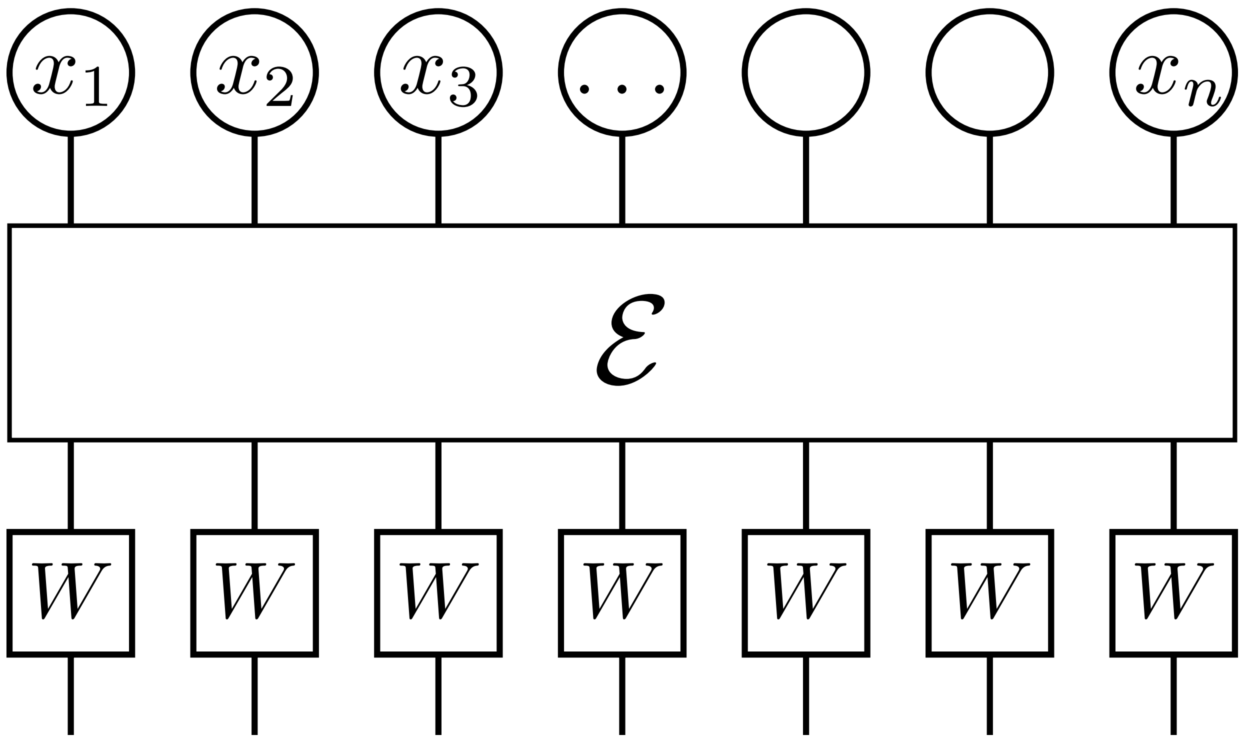
Encoding

- Encoding: reversible/unitary circuit with k data bits and n – k “frozen” bits (set to 0, say)



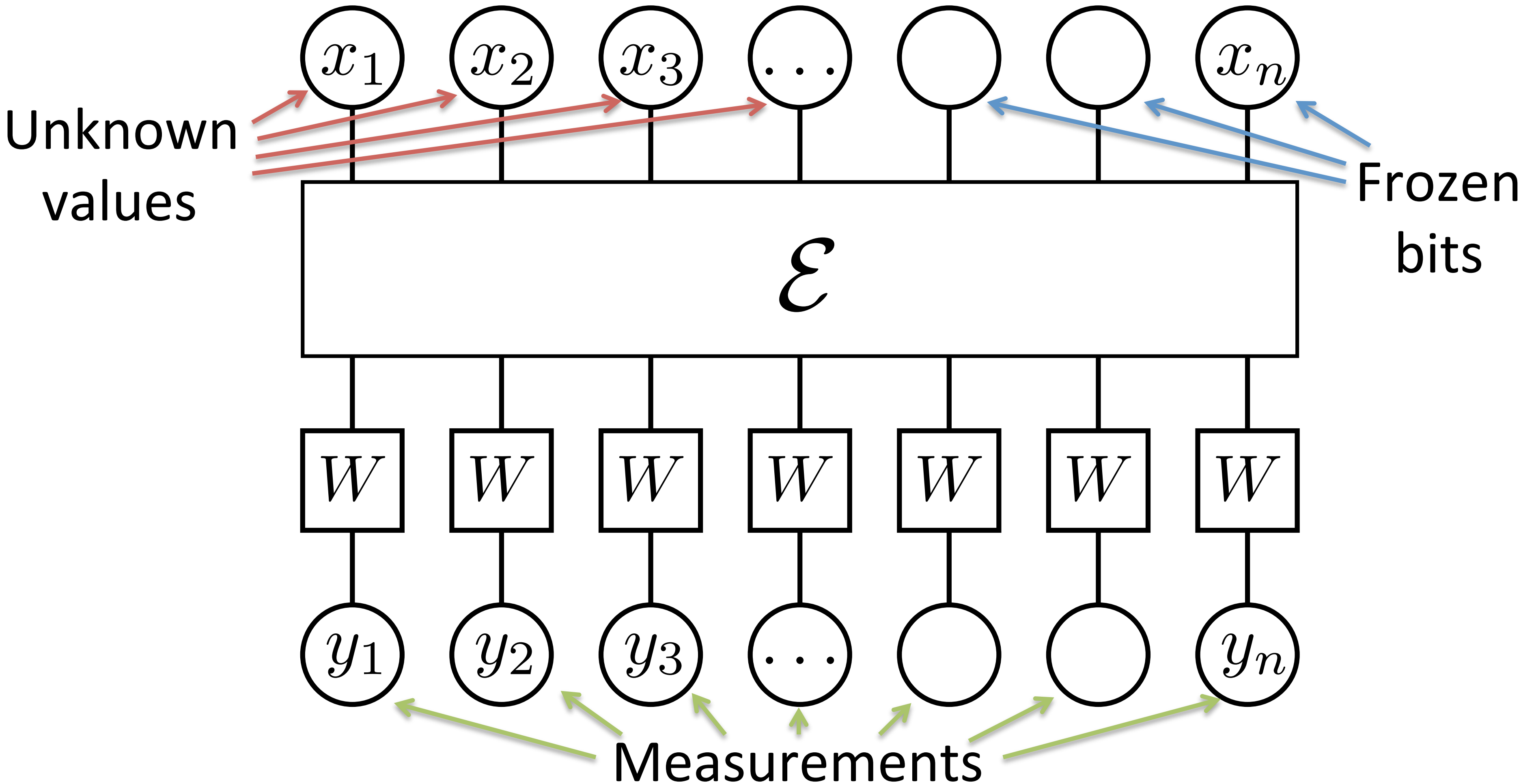
Noisy channels

Each bit (qubit) channel will undergo uncorrelated noise.



The classical decoding problem

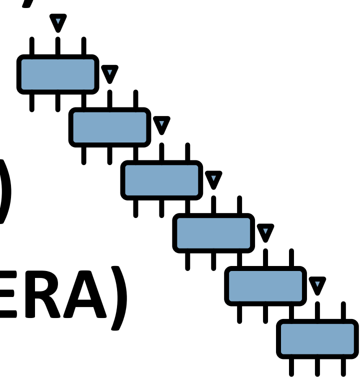
Find the inputs x_i given the measurements y_i



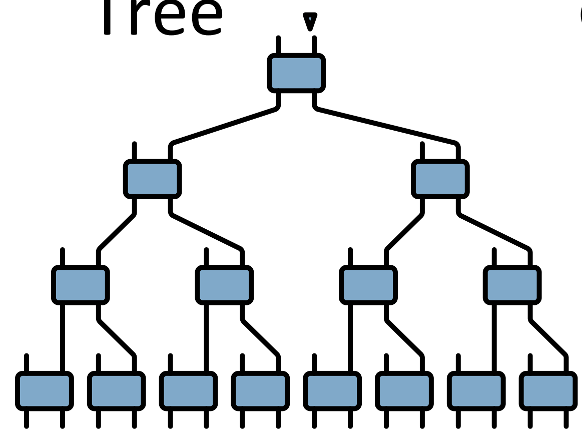
Many types of codes and circuits…

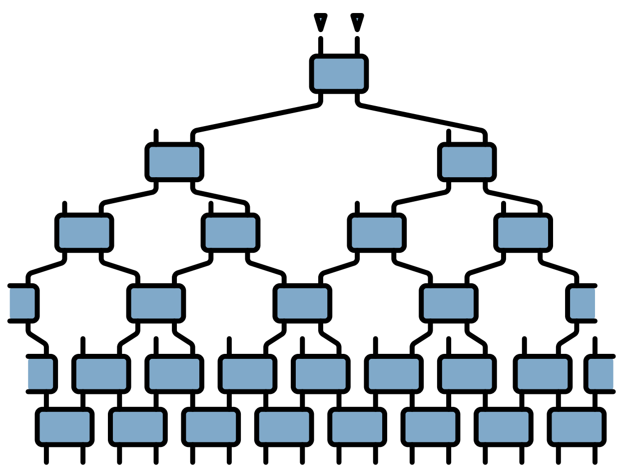
- Repetition codes
- **Convolutional codes (MPS)**
- **Concatenated codes (tree)**
- LDPC codes
- **Topological (MERA/PEPS)**
- **Polar code (Branching MERA)**

MPS



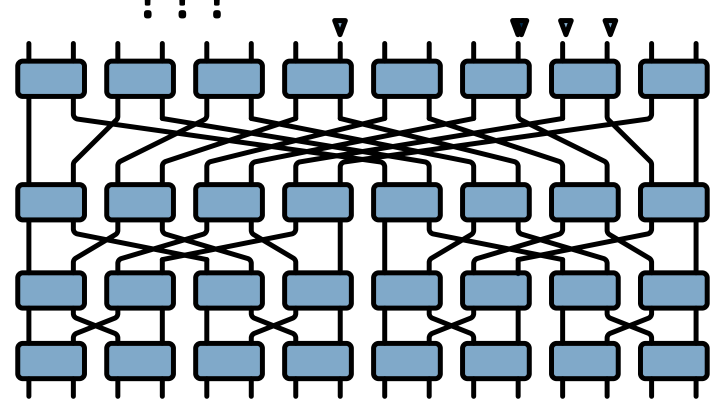
Tree





MERA

???



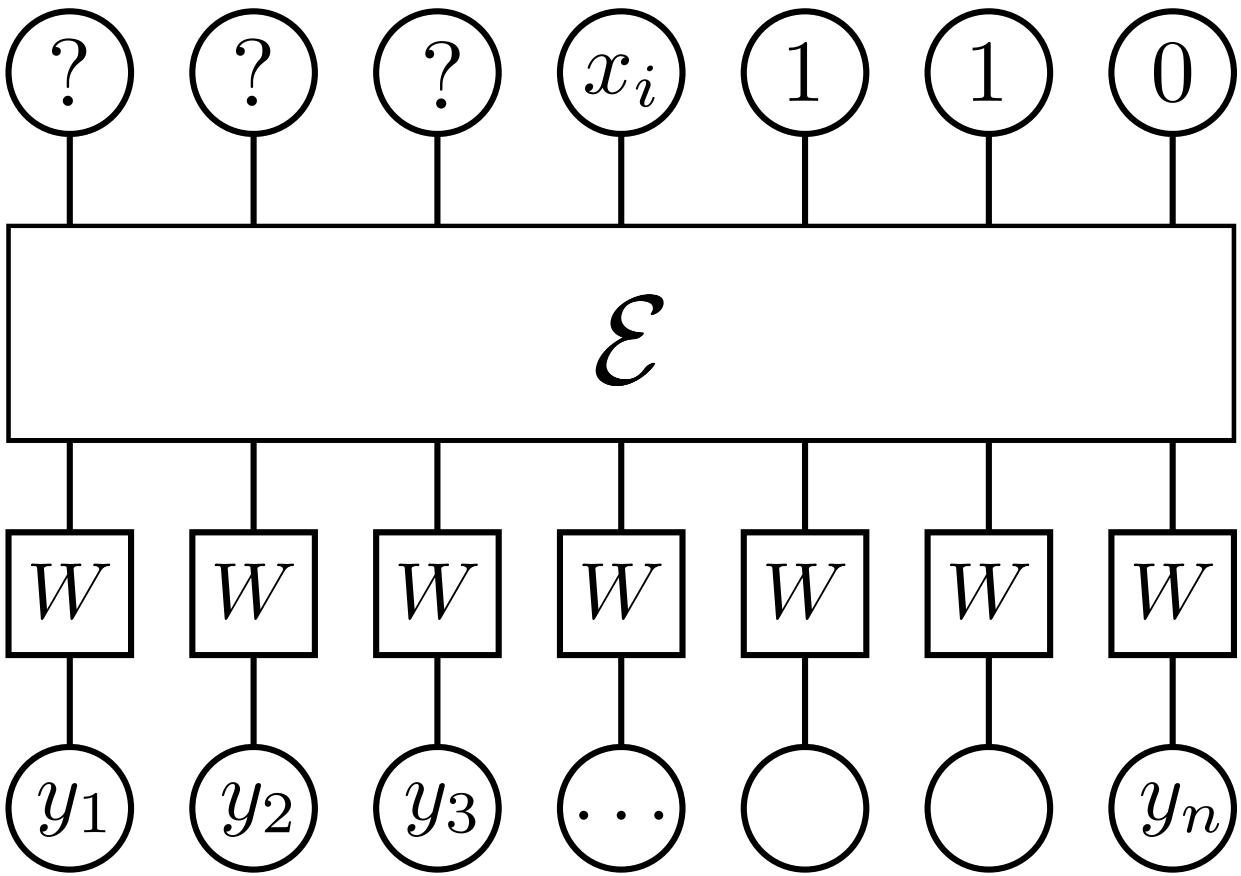
Many types of decoders…

- The "optimal" decoder is maximum likelihood
  - Given the measurements and error model, what is the most likely message sent? (in general: HARD)

- Many other "approximate" decoders
  - Bitwise sequential / successive cancelation
    - Find most likely first bit, then second given first, etc…
  - List decoding, belief propagation, renormalization group, etc.

Successive cancelation

- One bit at a time, from-right-to-left.



The Polar Code

Erdal Arikan
(Bilkent, Turkey)

- Arikan introduced the "polar" code in 2008/2009.
  - "polar" for *channel polarization*
- First code that is both:
  - Provably capacity achieving for generic symmetric channels
  - Efficient to decode (cost = $n \log n$)

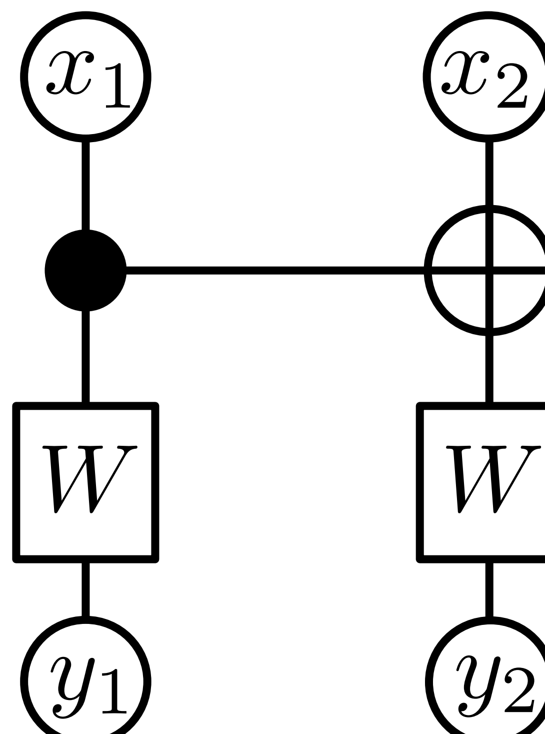
E. Arikan, IEEE Trans. on Inf. Theory **55**, 3051 (2009).

Channel Polarization

- Using successive cancelation (sequential) decoding, the polar code polarizes each input channel to be very "good" or very "bad"
- Take 2-bit CNOT

First decode x_2

Then decode x_1
given knowledge of x_2

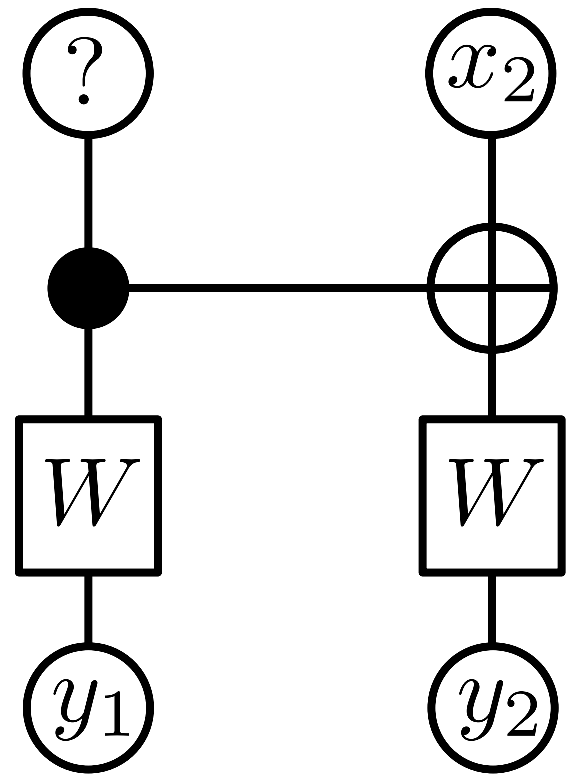


Channel Polarization

Step 1: determine whether x_2 = 0 or 1 is more likely, given y_1, y_2 and W.

$$P(x_2|y_1, y_2)$$

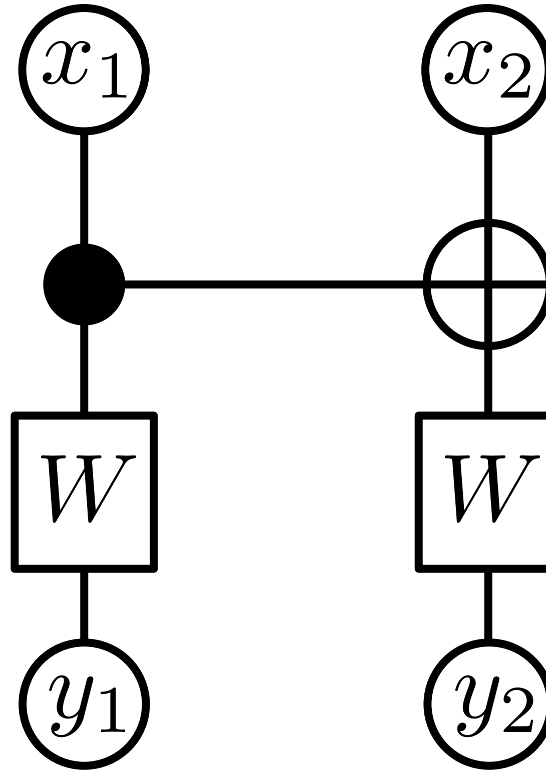
(x_1 is unknown, assumed to be in 50/50 mixture)



CNOT acts to copy noise from unknown x_1.

Channel Polarization

Step 2: determine whether x_1 = 0 or 1 is more likely, given y_1, y_2, W and previously determined x_2.



$$P(x_1|x_2, y_1, y_2)$$

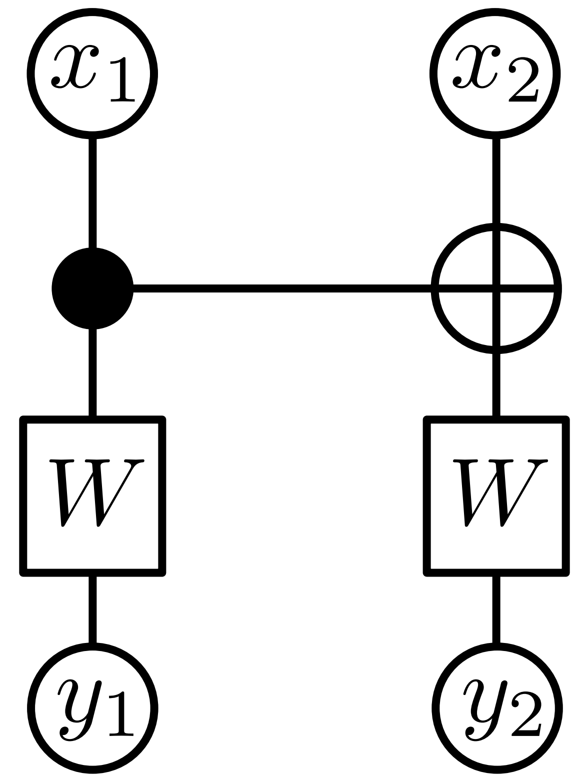
CNOT creates two copies of x_1 to protect from noise

Channel Polarization

Step 2 is always more accurate than step 1.

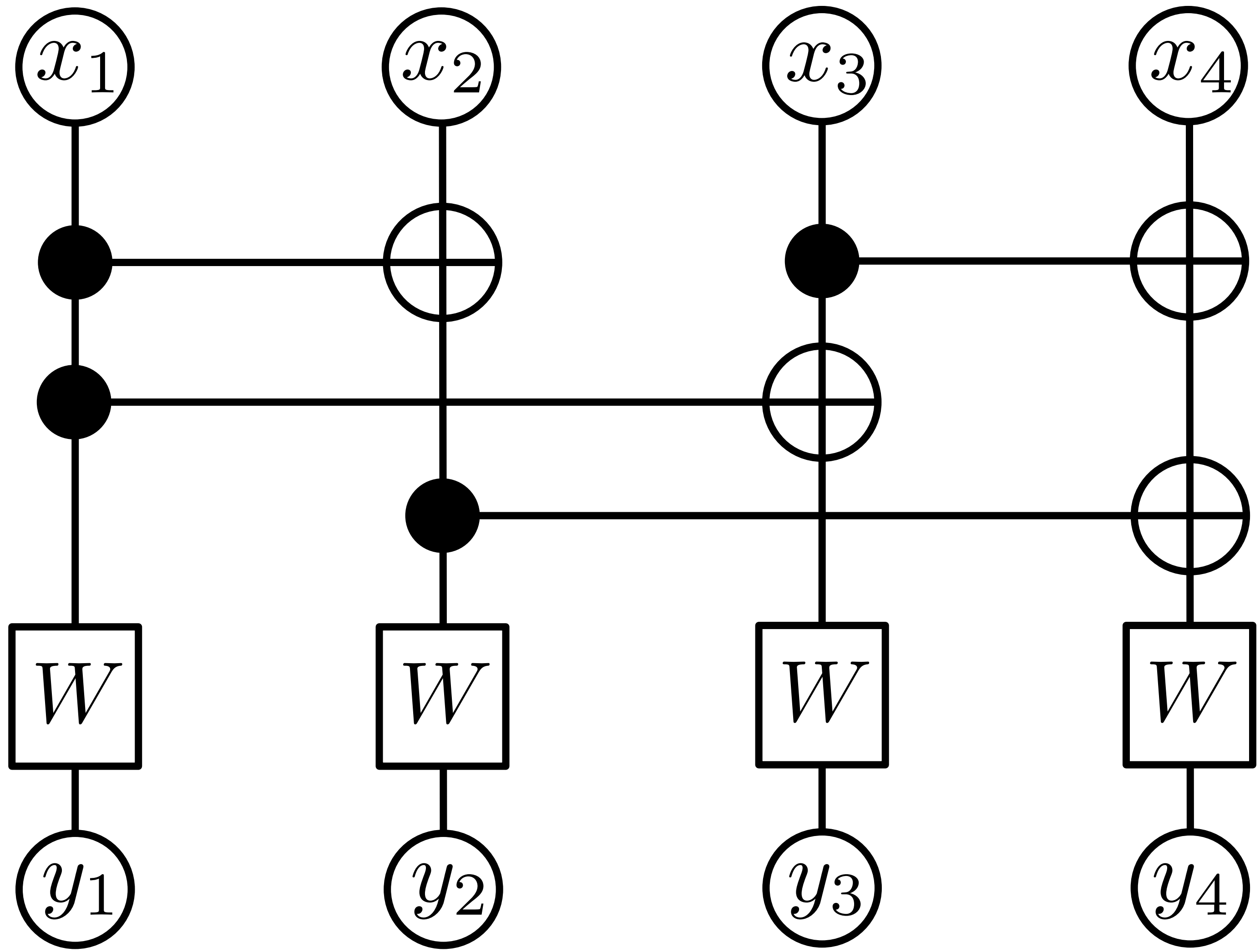
$$I(x_1|x_2) \geq I(x_2)$$

BUT error in step 1 could cause an error in step 2. Freeze the value of x_2 and use x_1 as data.

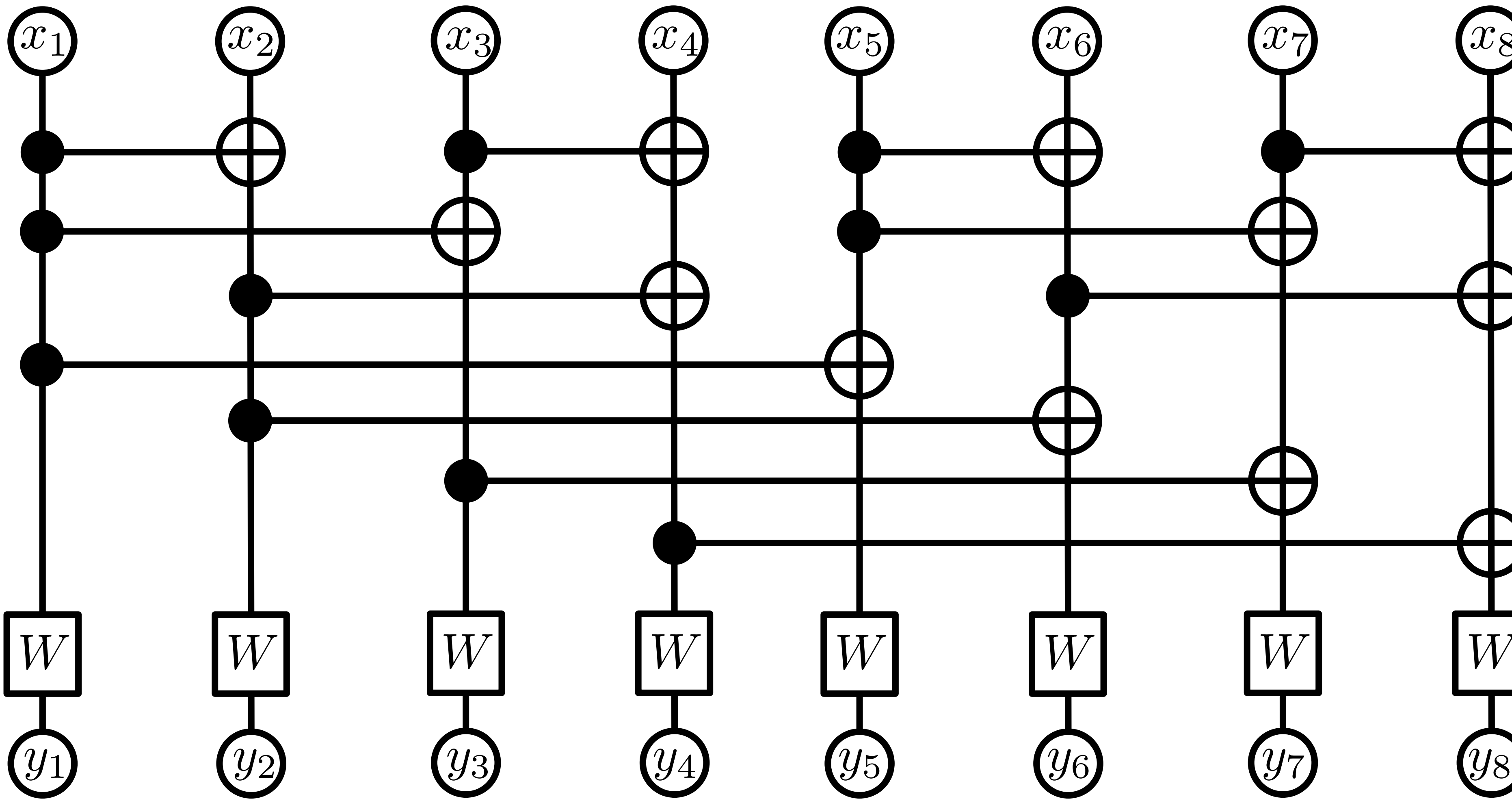


Summary: x_1 is a “good” channel, x_2 is “bad” channel

Channel Polarization

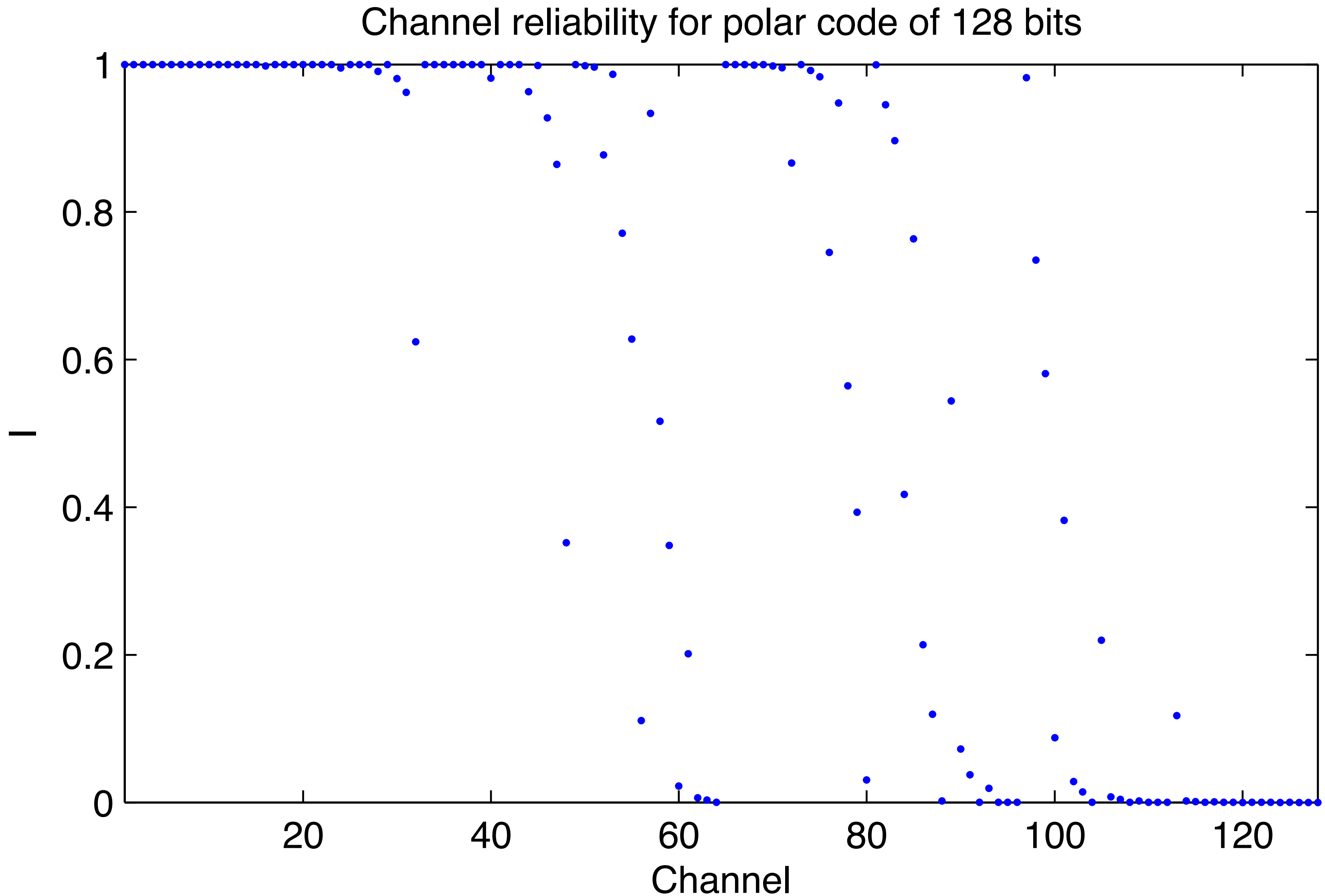


Channel Polarization



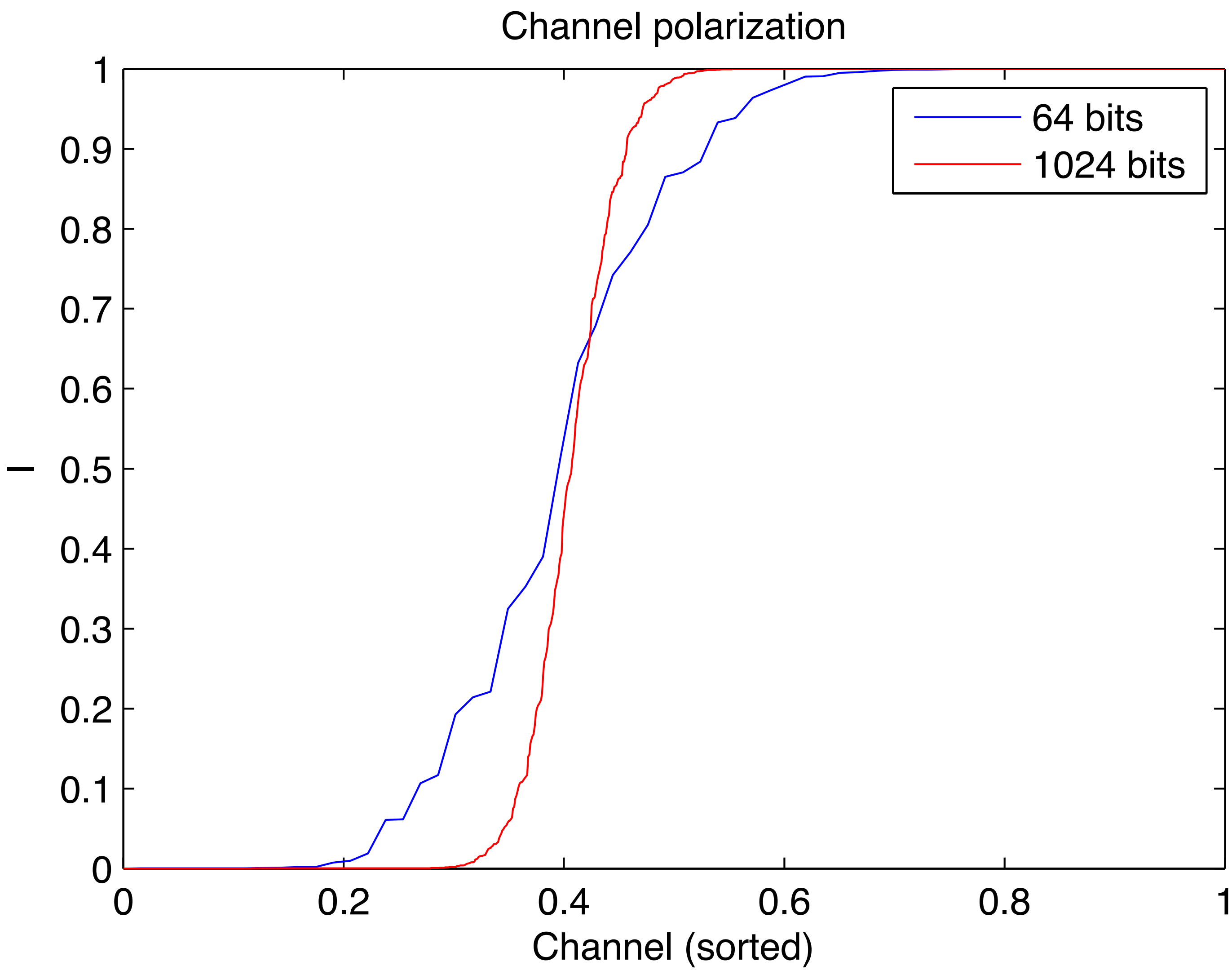
Channel Polarization

Channels either become “good” or “bad”. Rate = 0.5



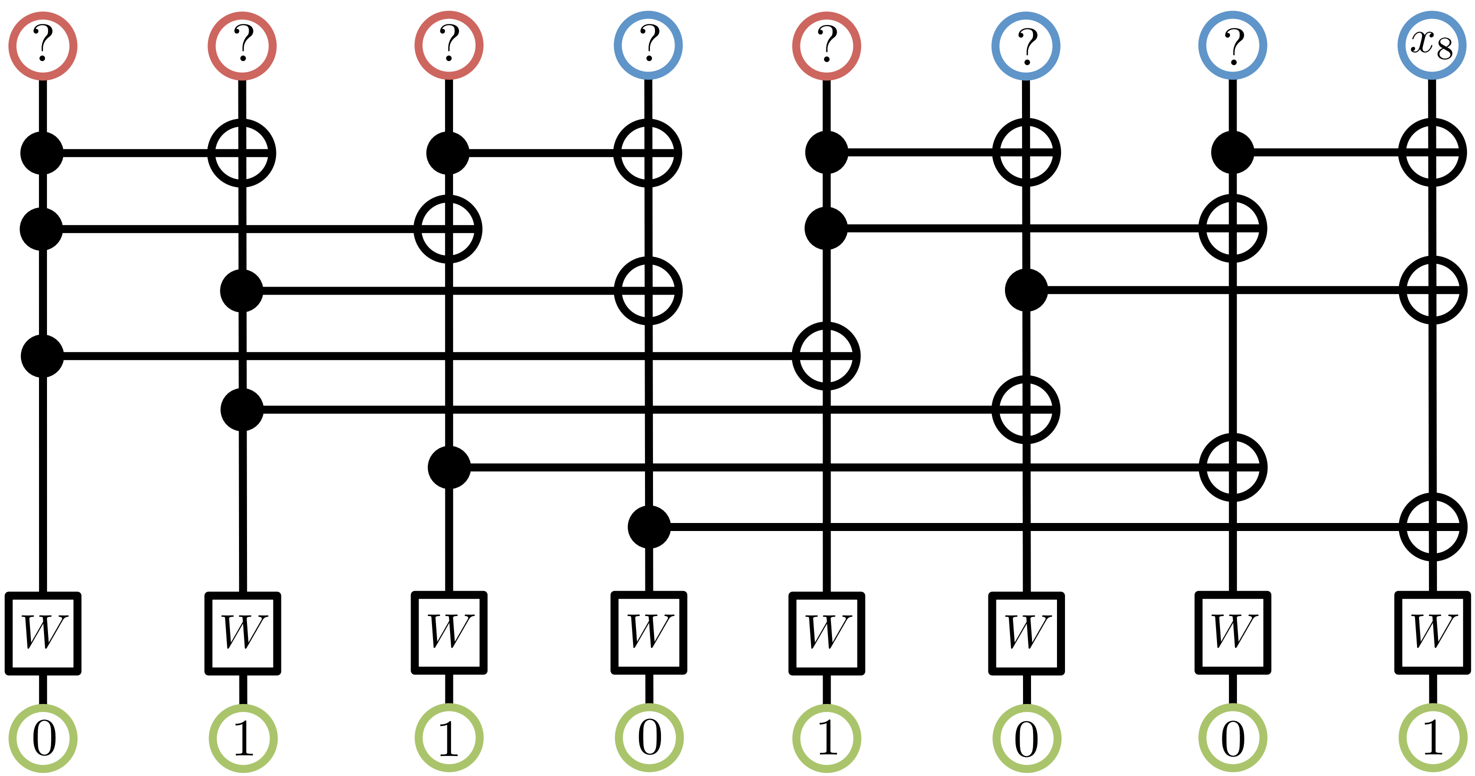
Channel Polarization

Channels either become "good" or "bad". Rate = 0.5



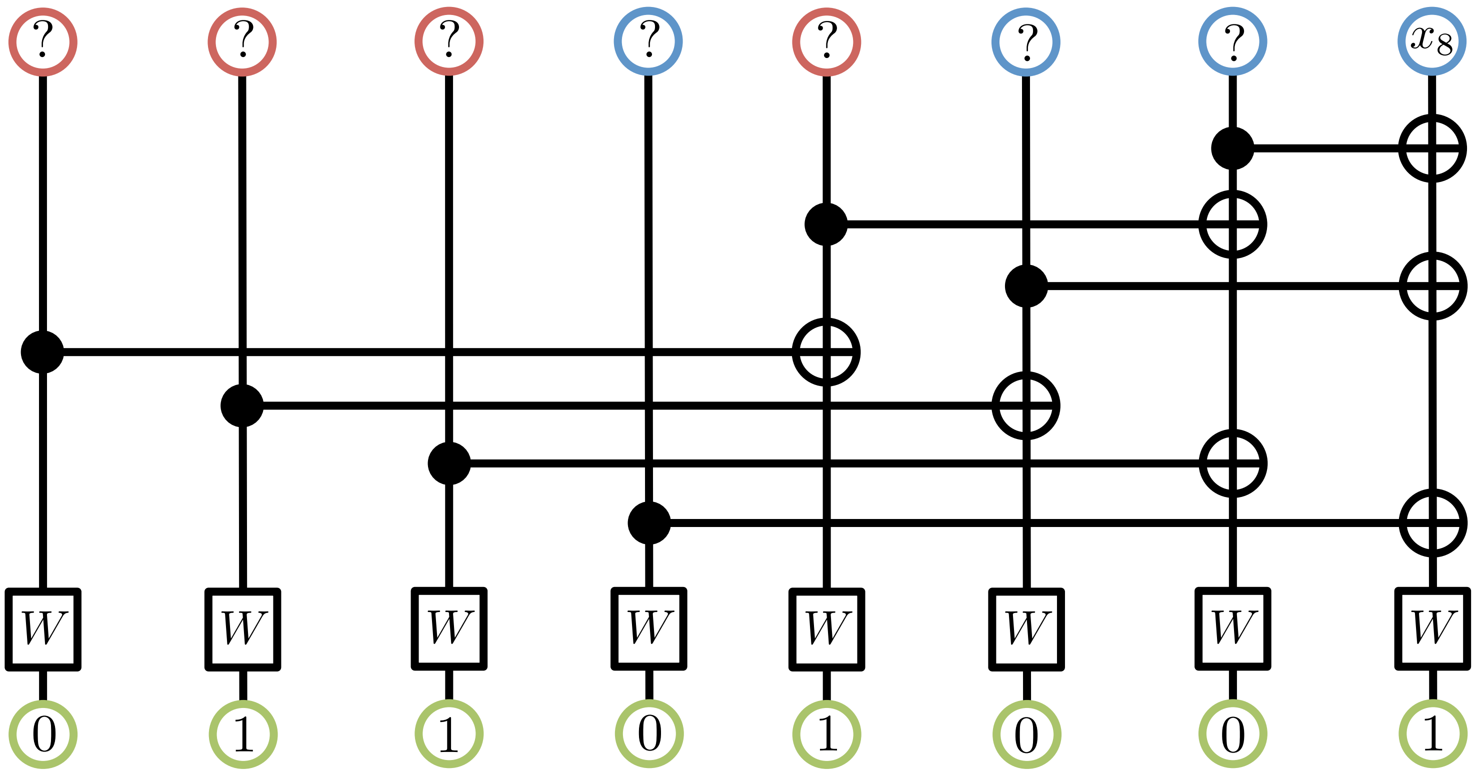
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



Decoding the Polar code

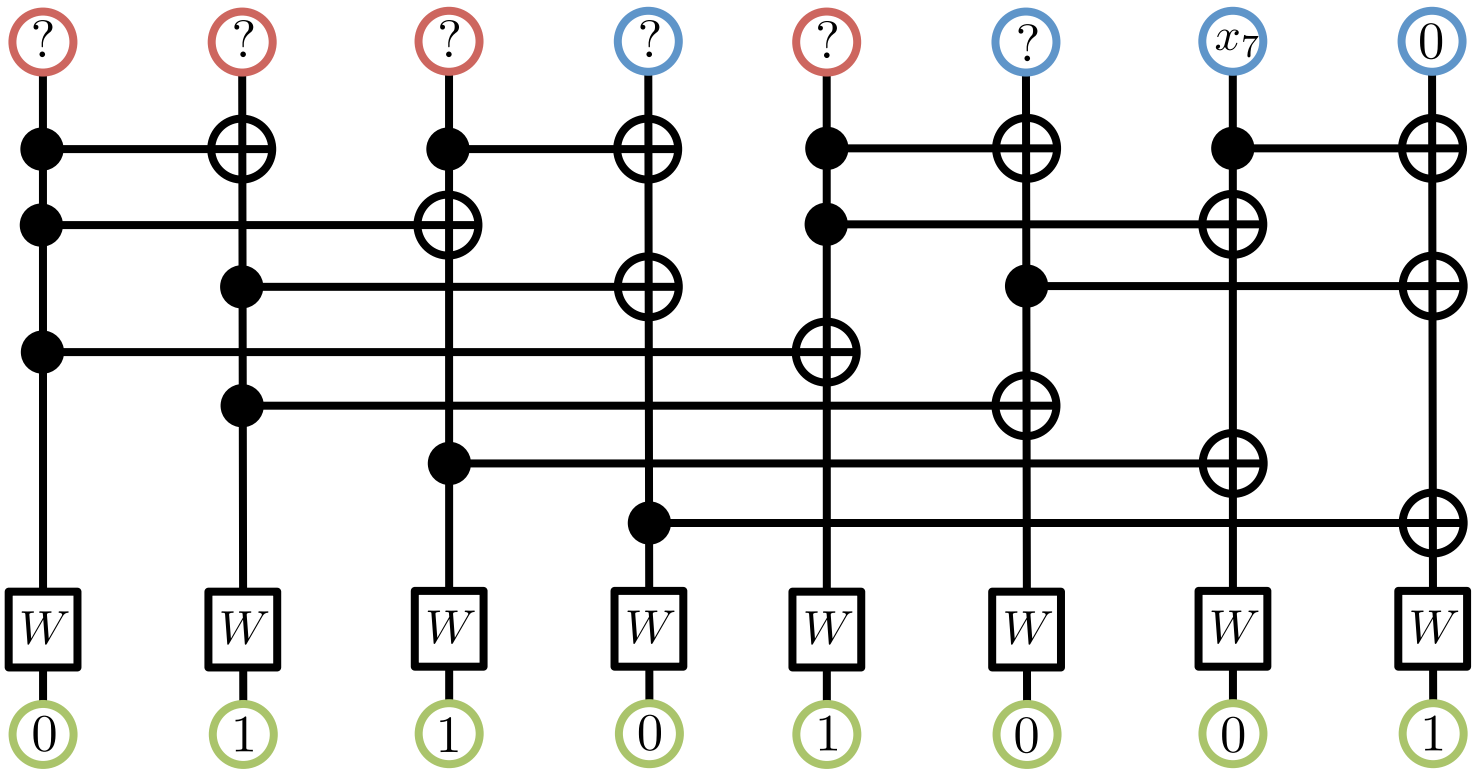
- Most the tensors cancel under successive cancellation decoding.



- Cost to calculate is $\mathcal{O}(n)$

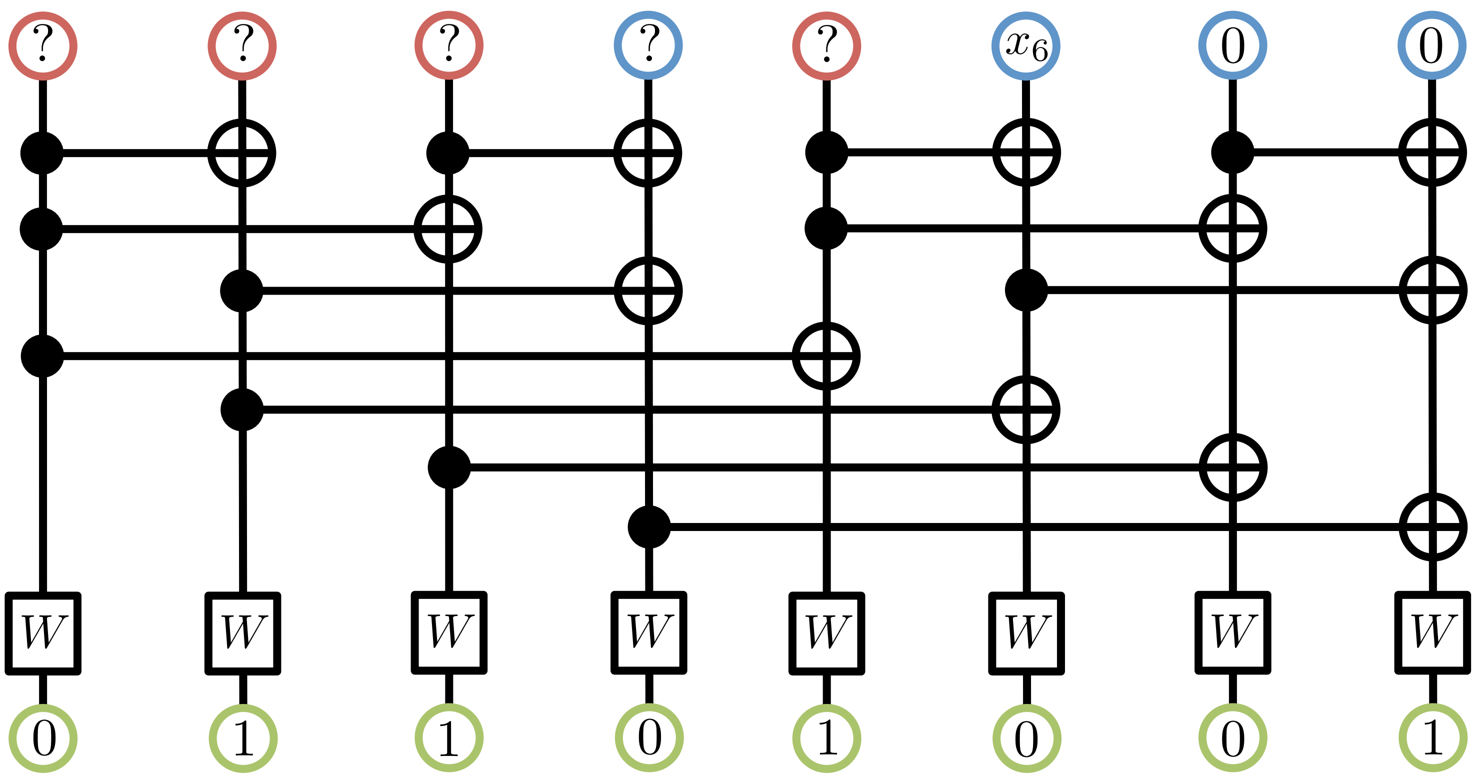
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



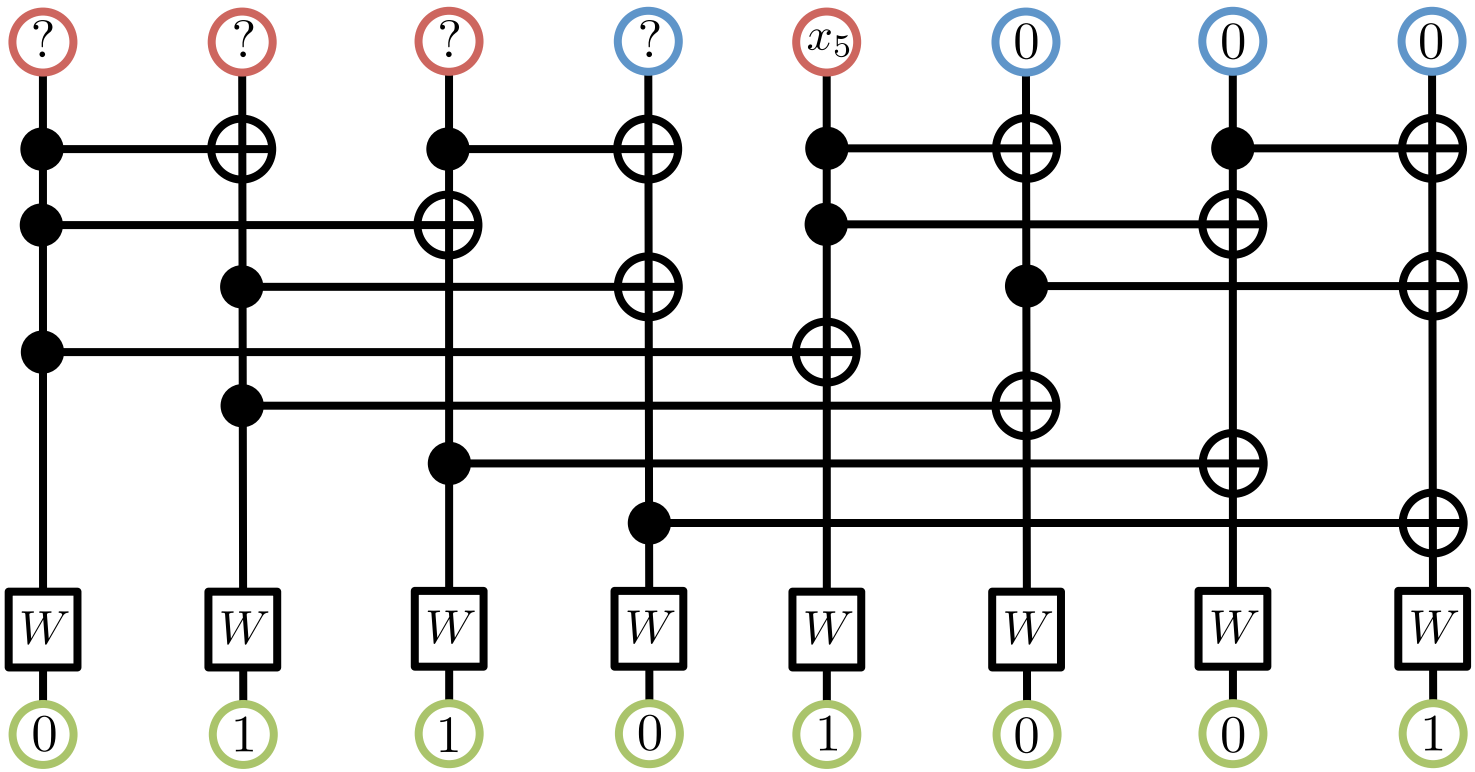
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



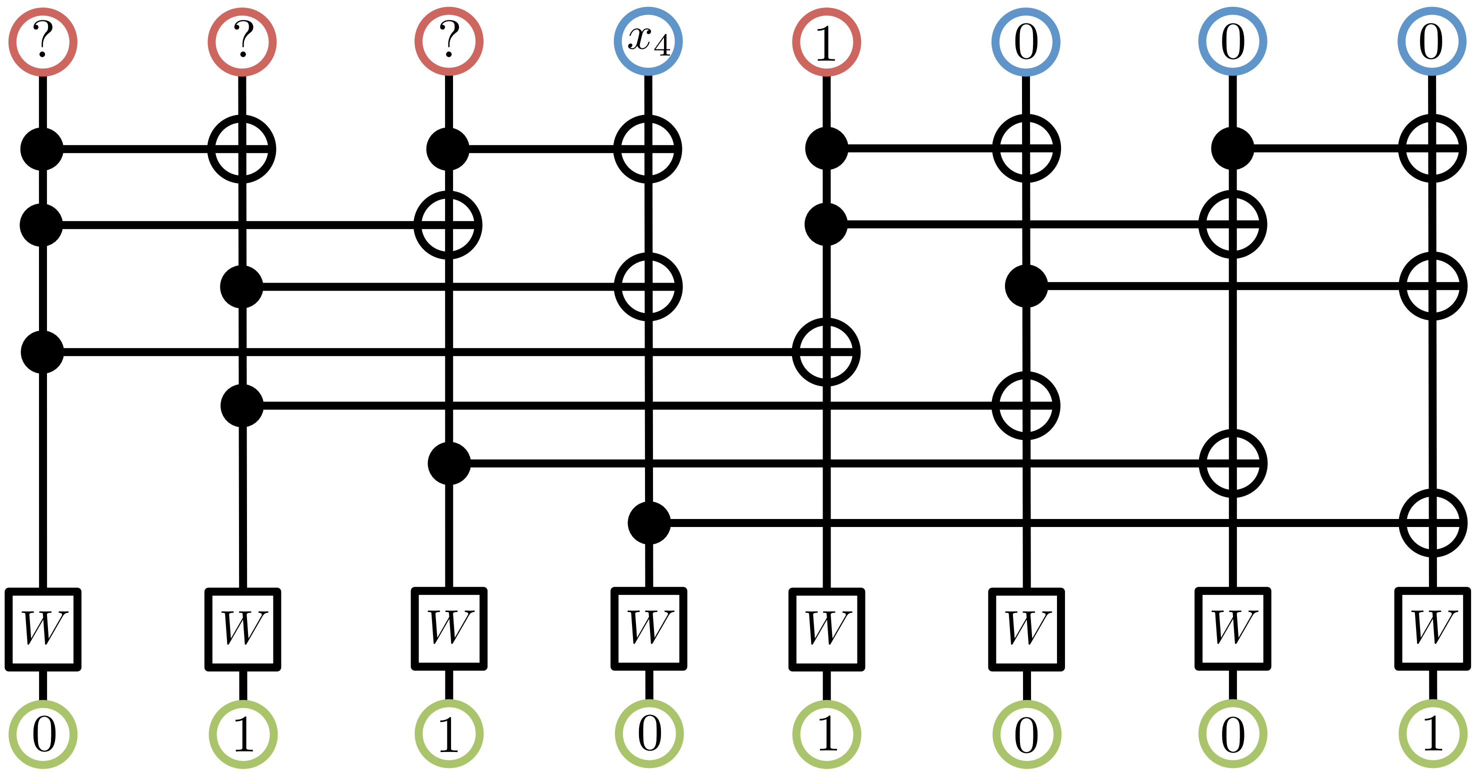
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



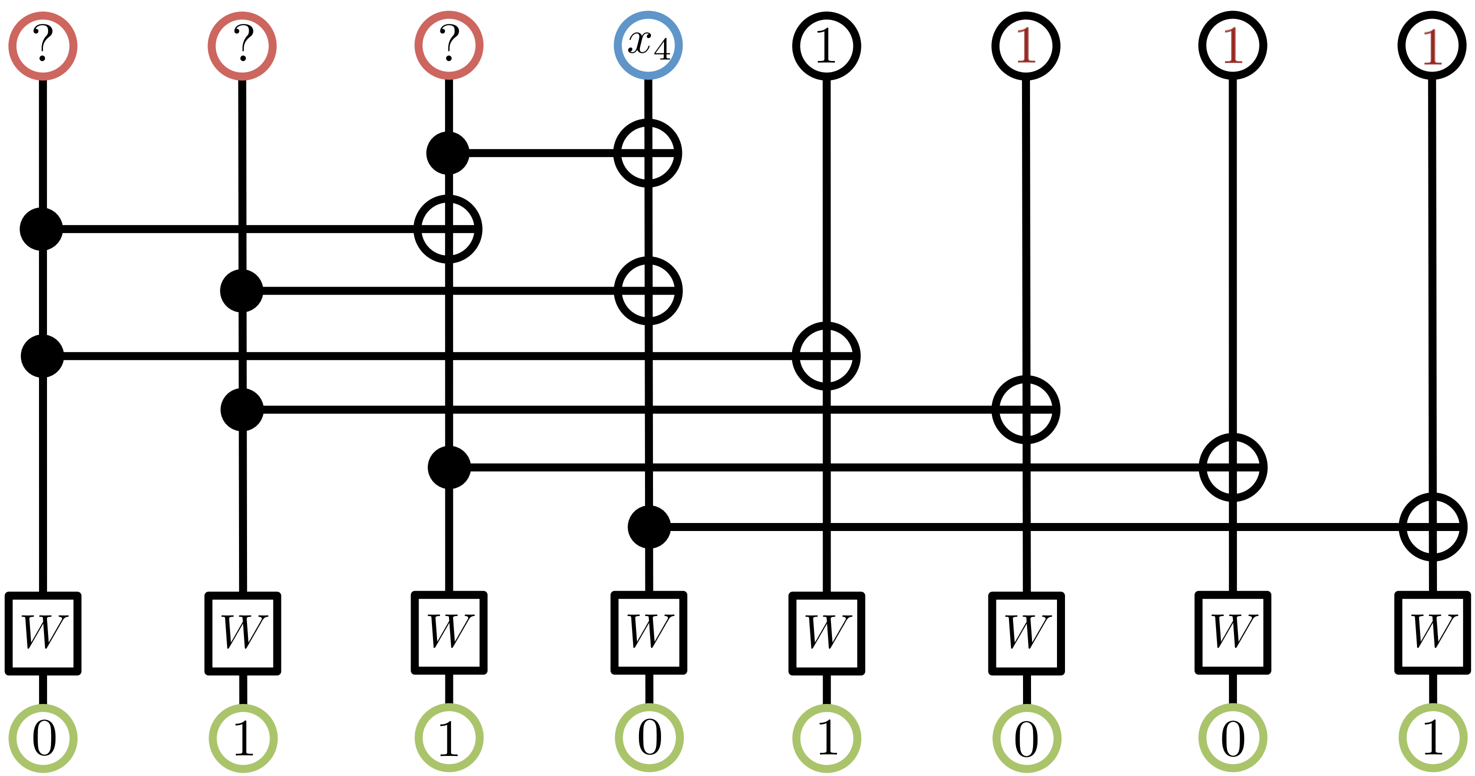
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



Decoding the Polar code

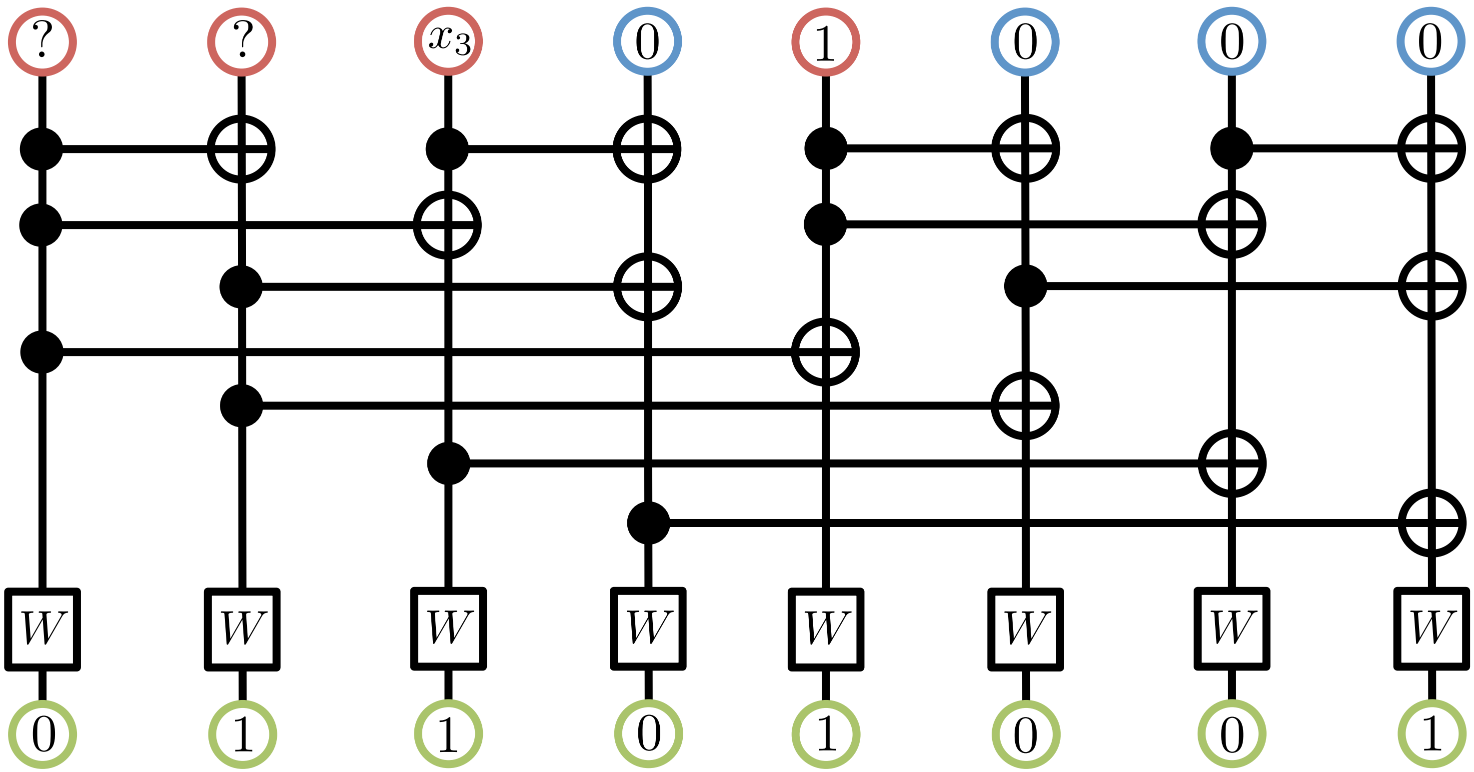
- Most the tensors cancel under successive cancellation decoding.



- Cost to calculate is $\mathcal{O}(n)$

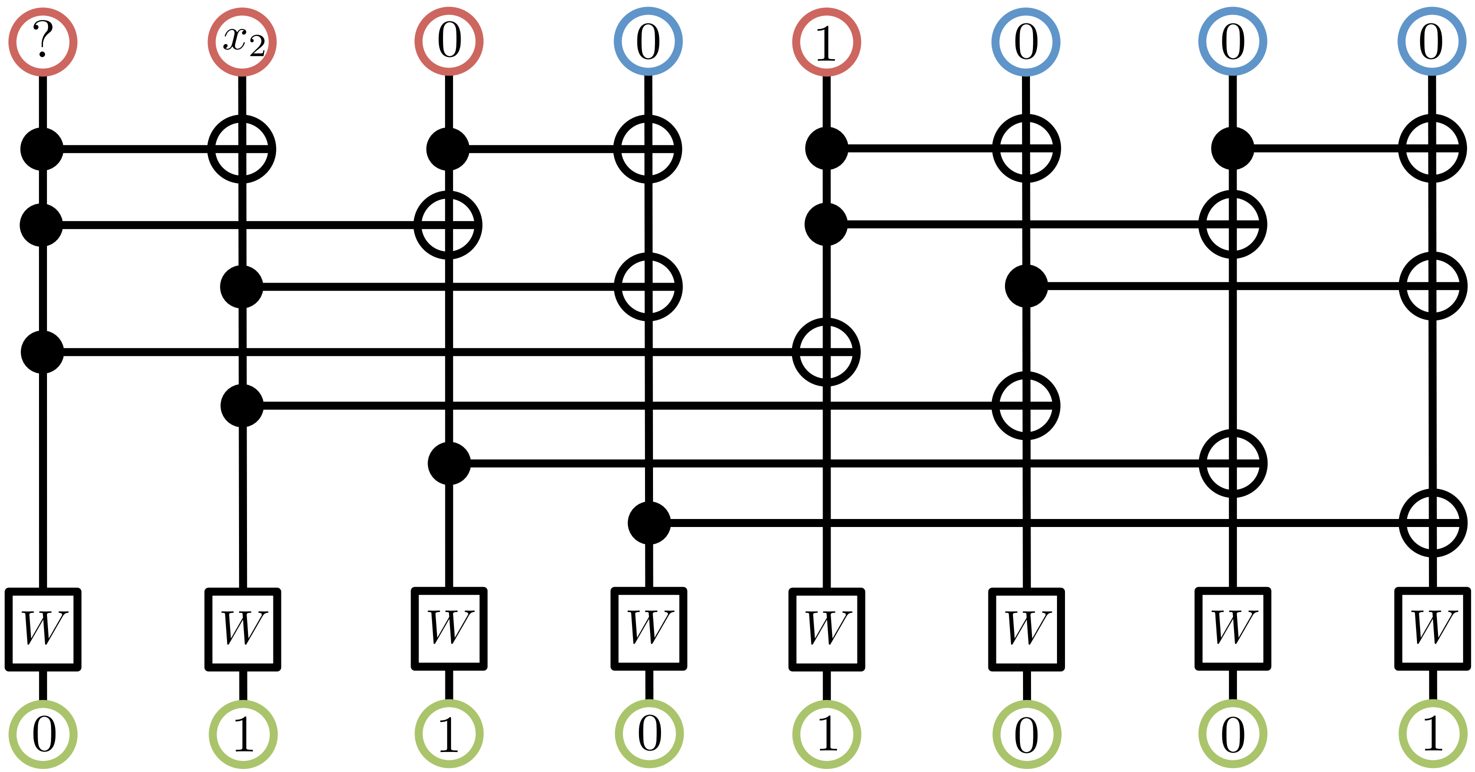
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



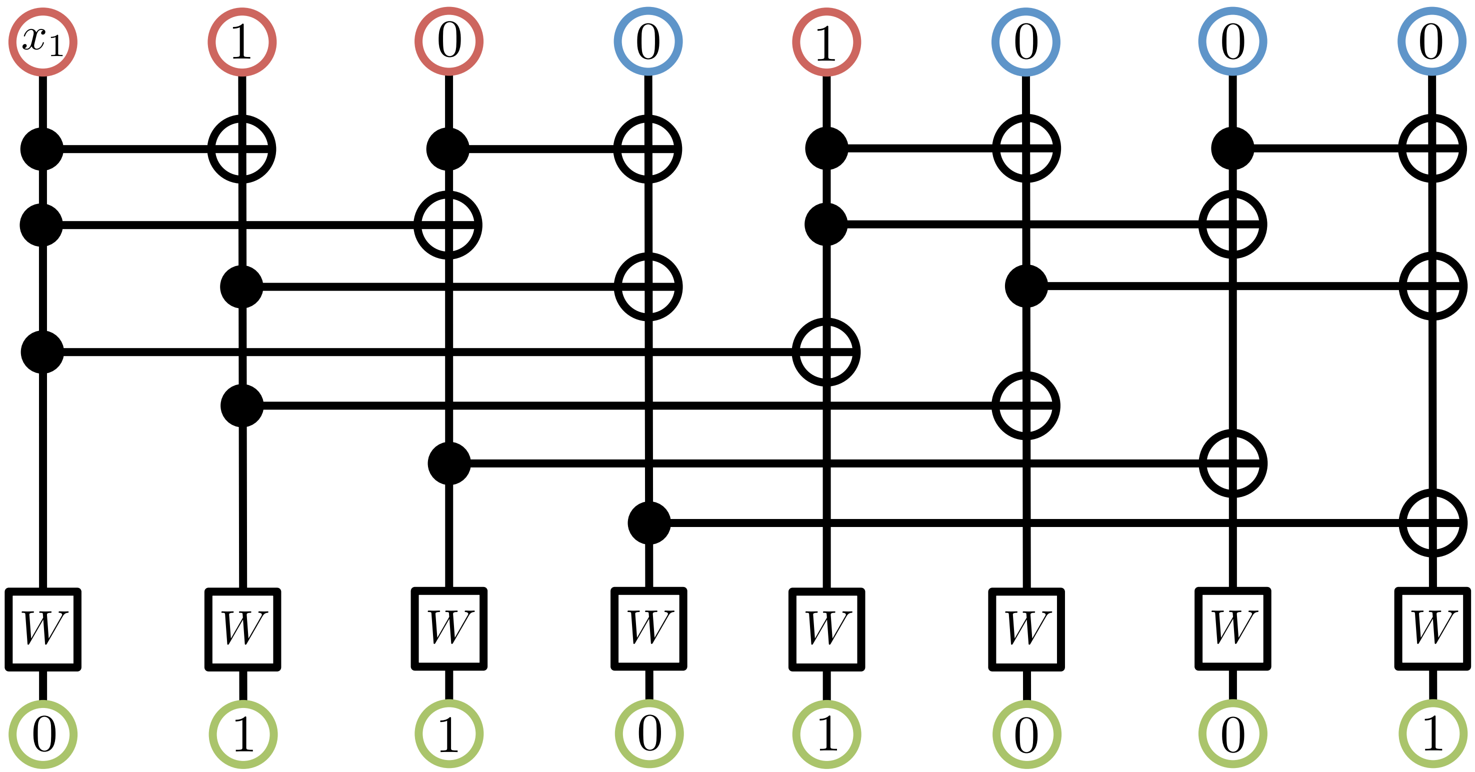
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



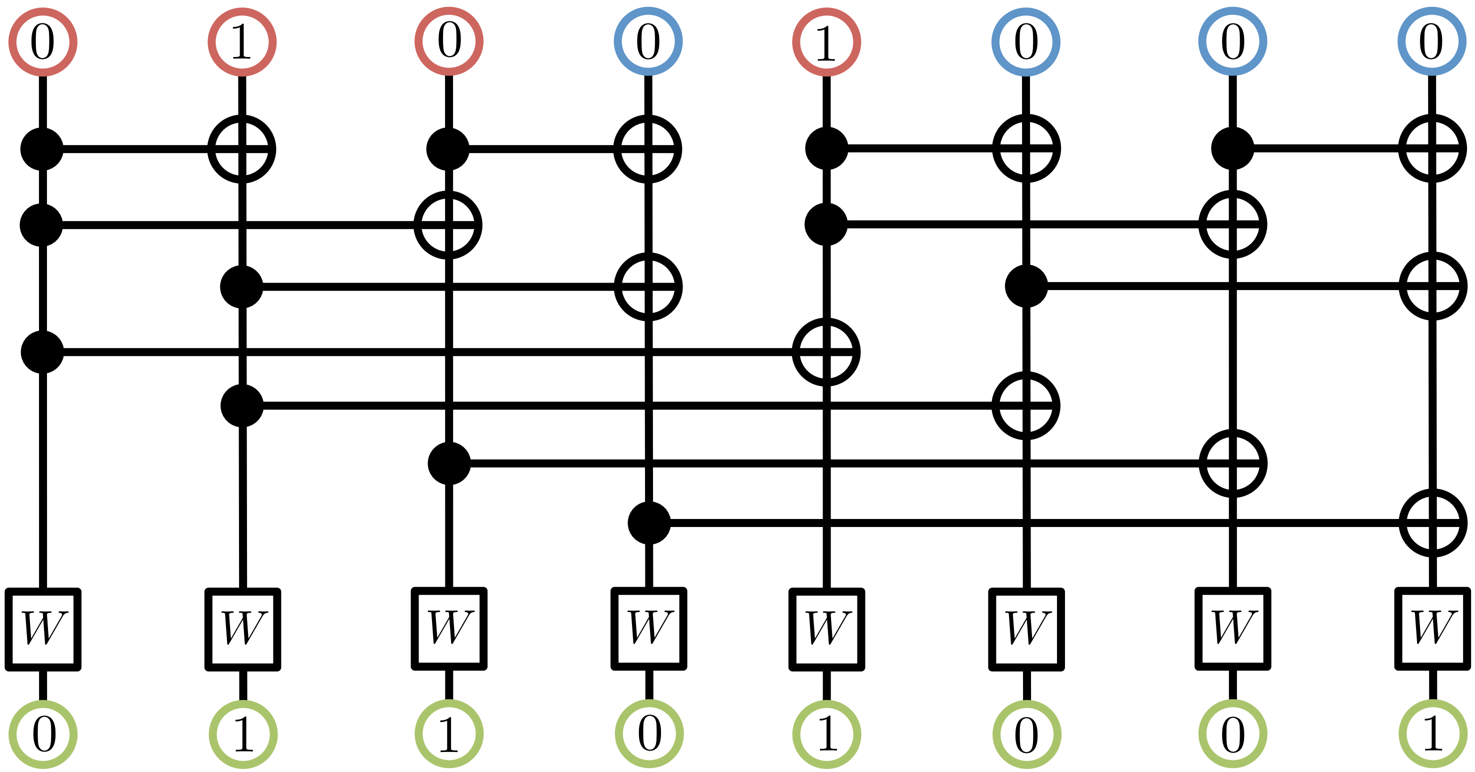
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



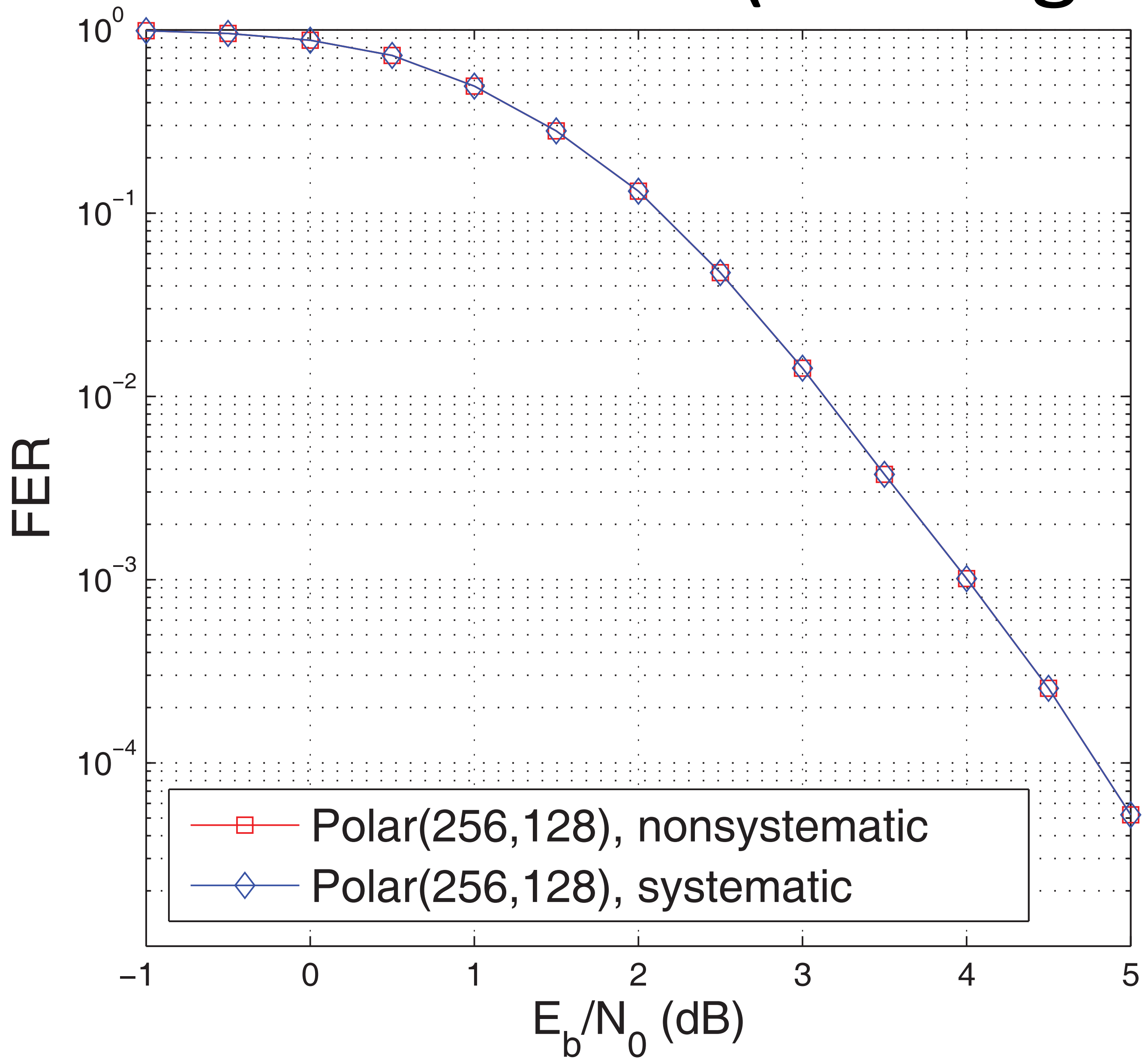
Decoding the Polar code

- Most the tensors cancel under successive cancellation decoding.



- Single bit costs $\mathcal{O}(n)$, but all bits in $\mathcal{O}(n \log n)$!

Performance (existing results)



AWGN channel,
256 bits,
rate 1/2

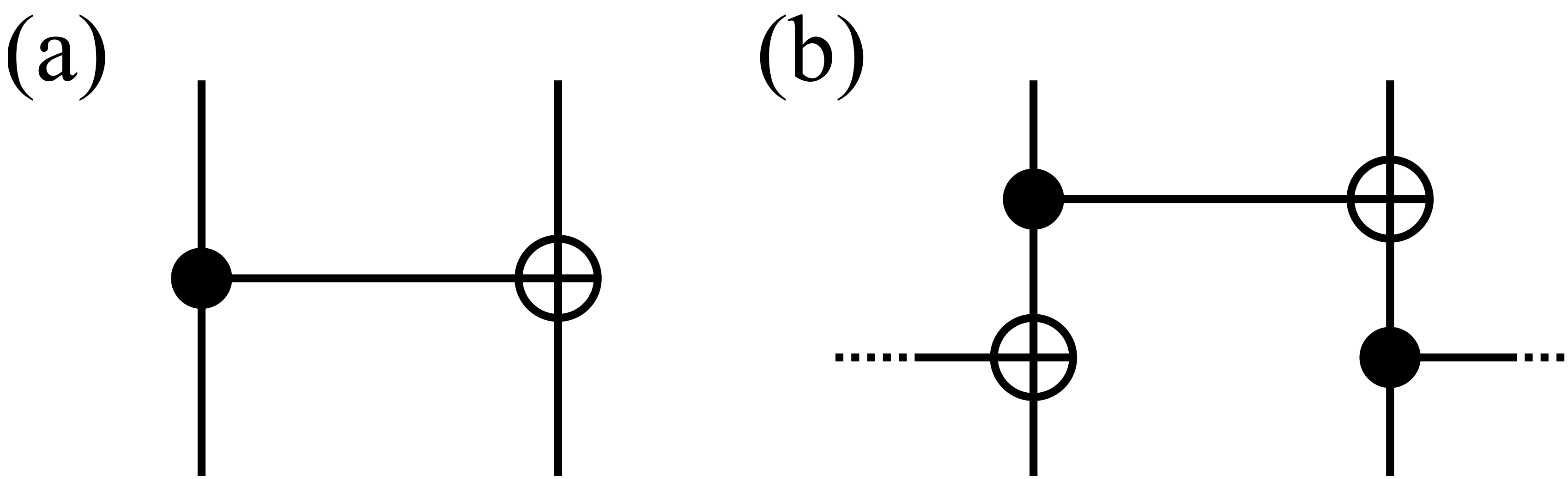
E. Arikan, IEEE Communications Letters **15**, 860 (2011).

Going beyond the Polar code

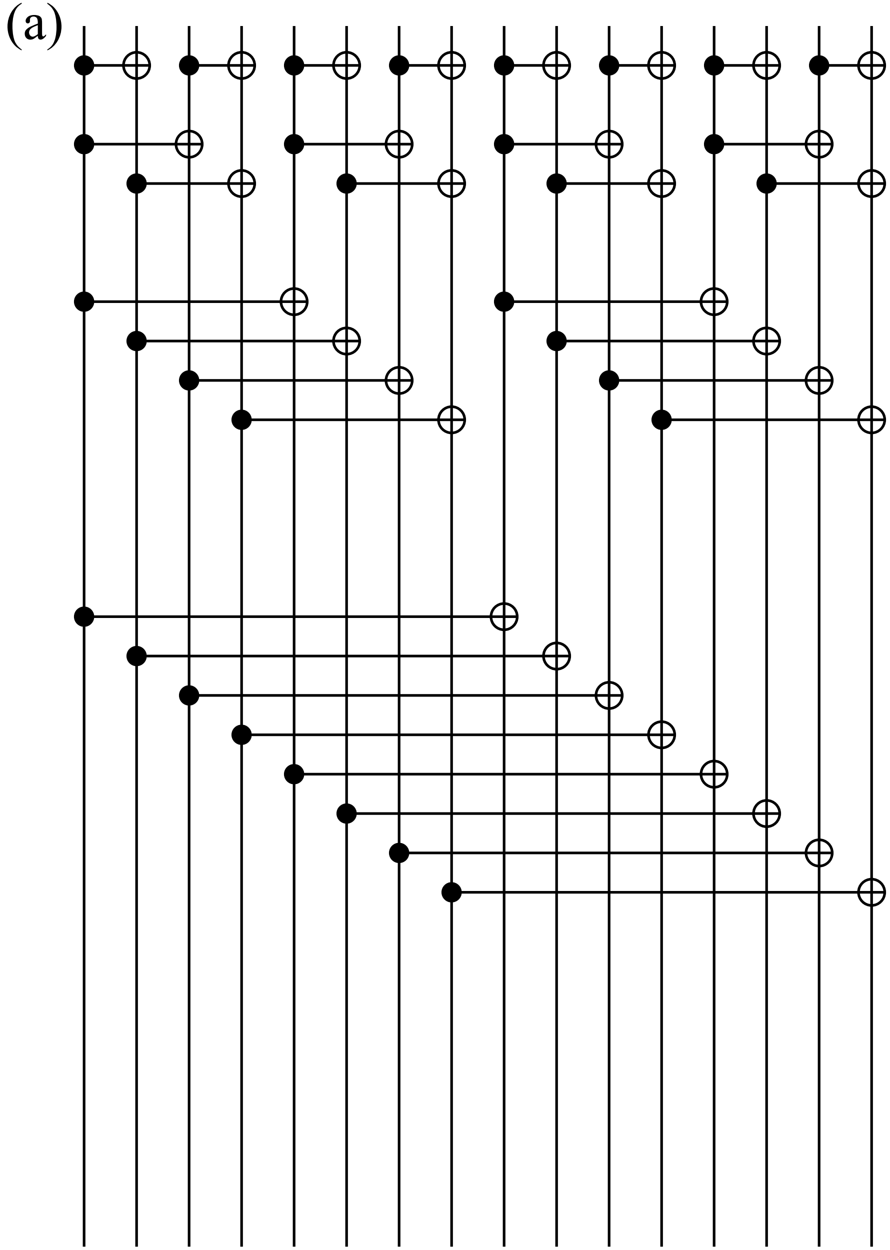
- Tensor networks studied extensively in many-body physics.
- Recently, a new tensor network was proposed called the “branching-MERA” that has a similar multi-scale structure to the Polar code.
- It contains twice the tensors (gates) and is able to better “scramble” the data.
- Successive cancelation remains $\mathcal{O}(n \log n)$.

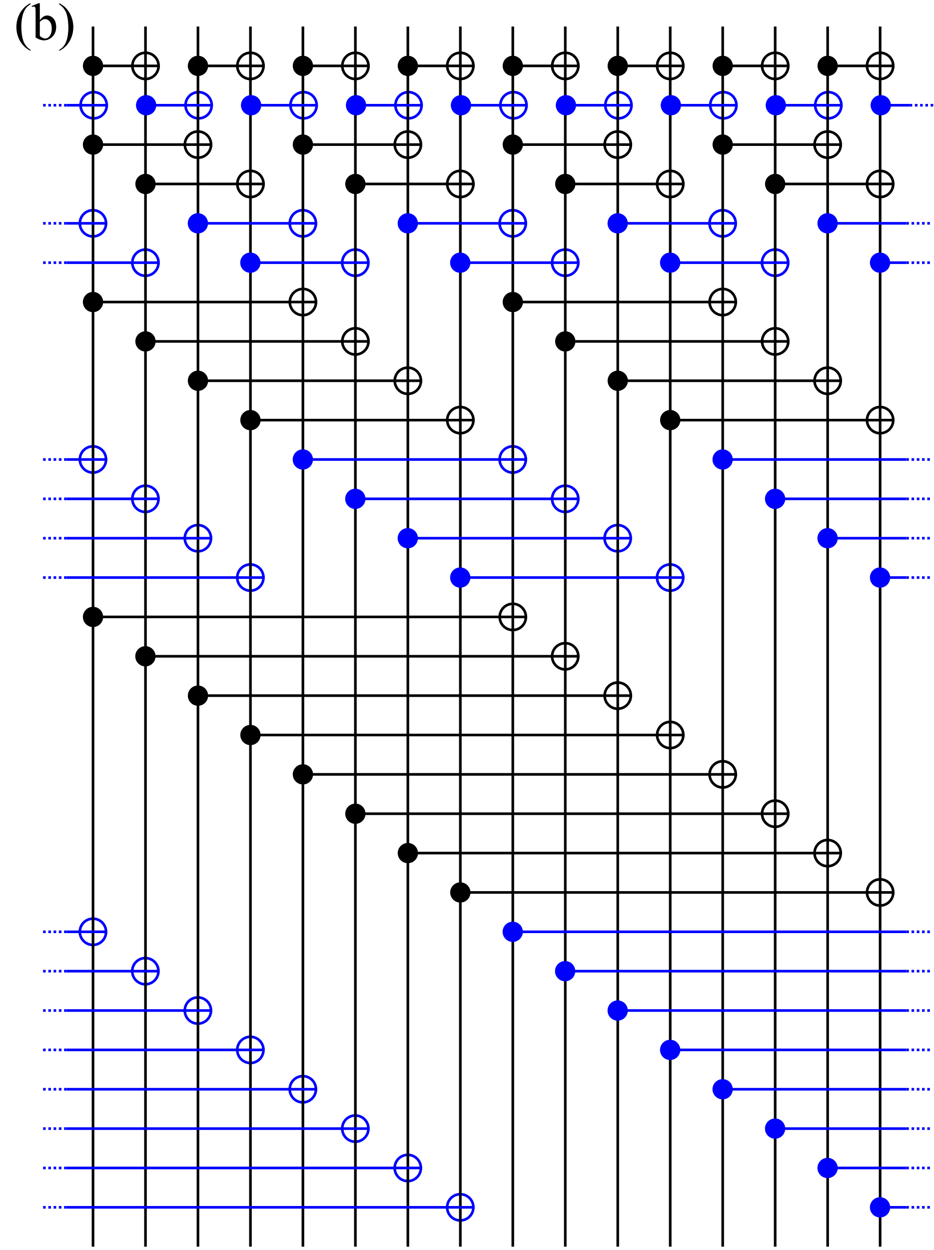
Branching-MERA code

- The primitive piece of the polar code is CNOT.
- The branching-MERA code adds another CNOT to the other neighbour:

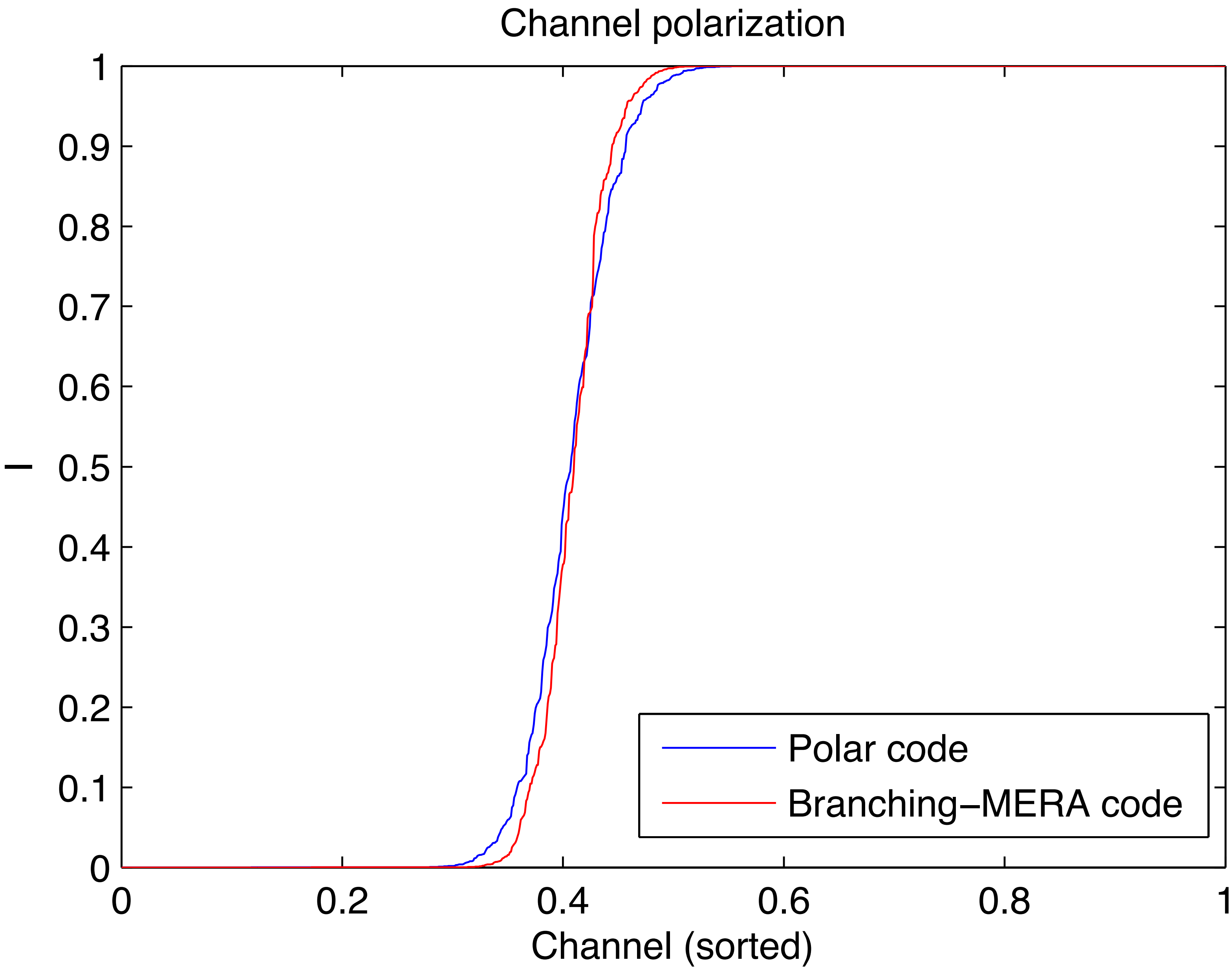


Polar vs. branching-MERA code

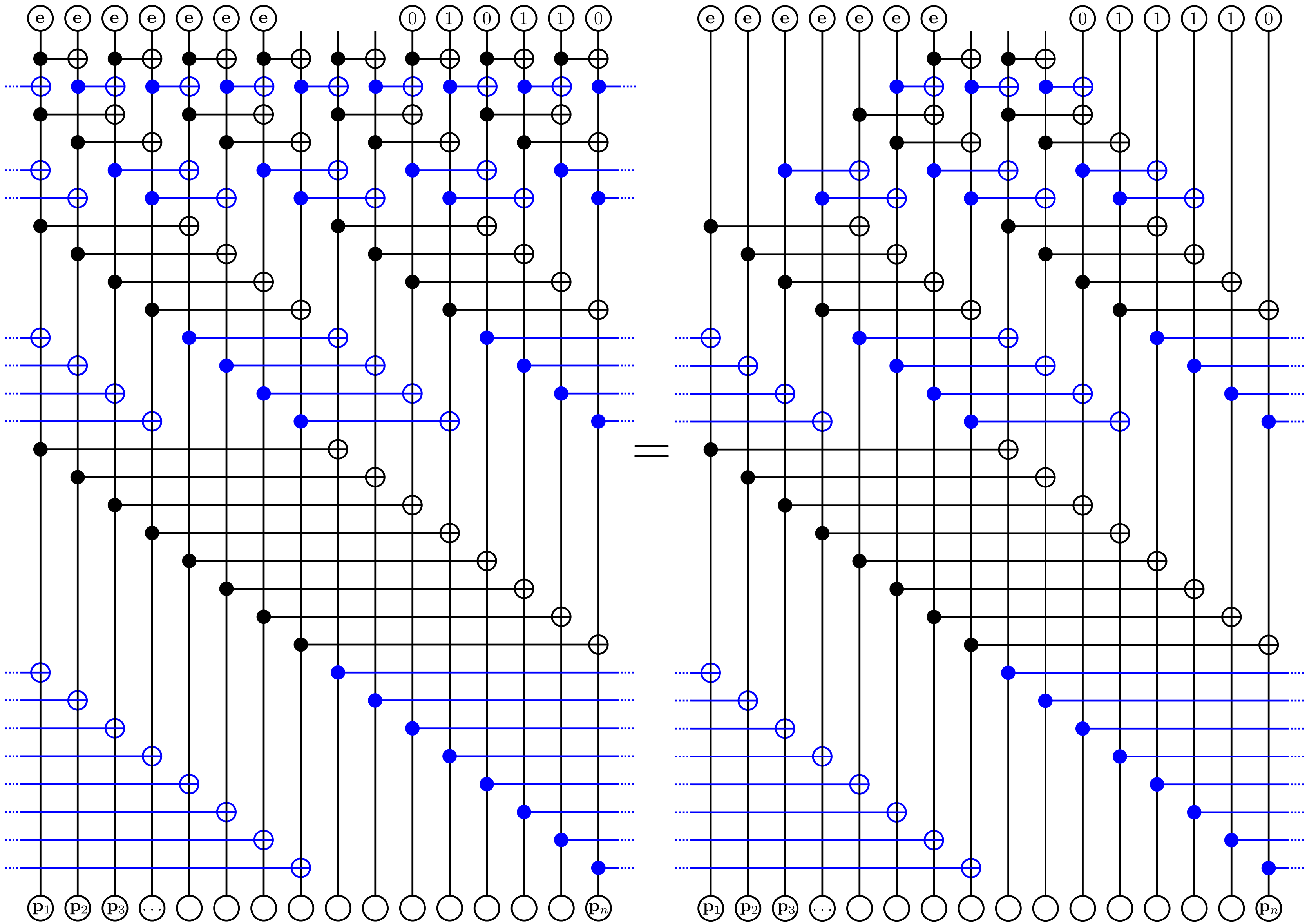




Better polarization than Polar code

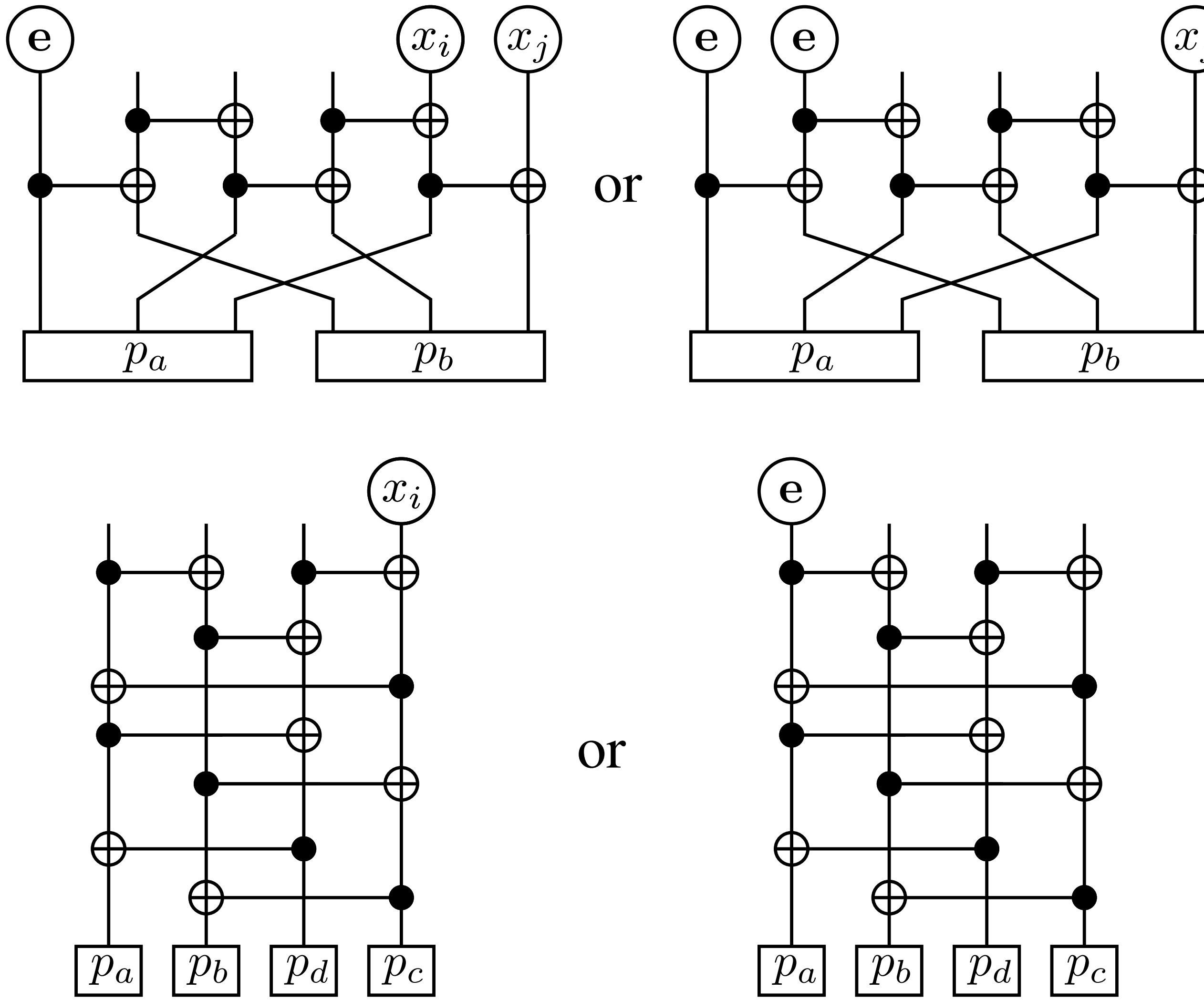


Successive cancelation



Decoding

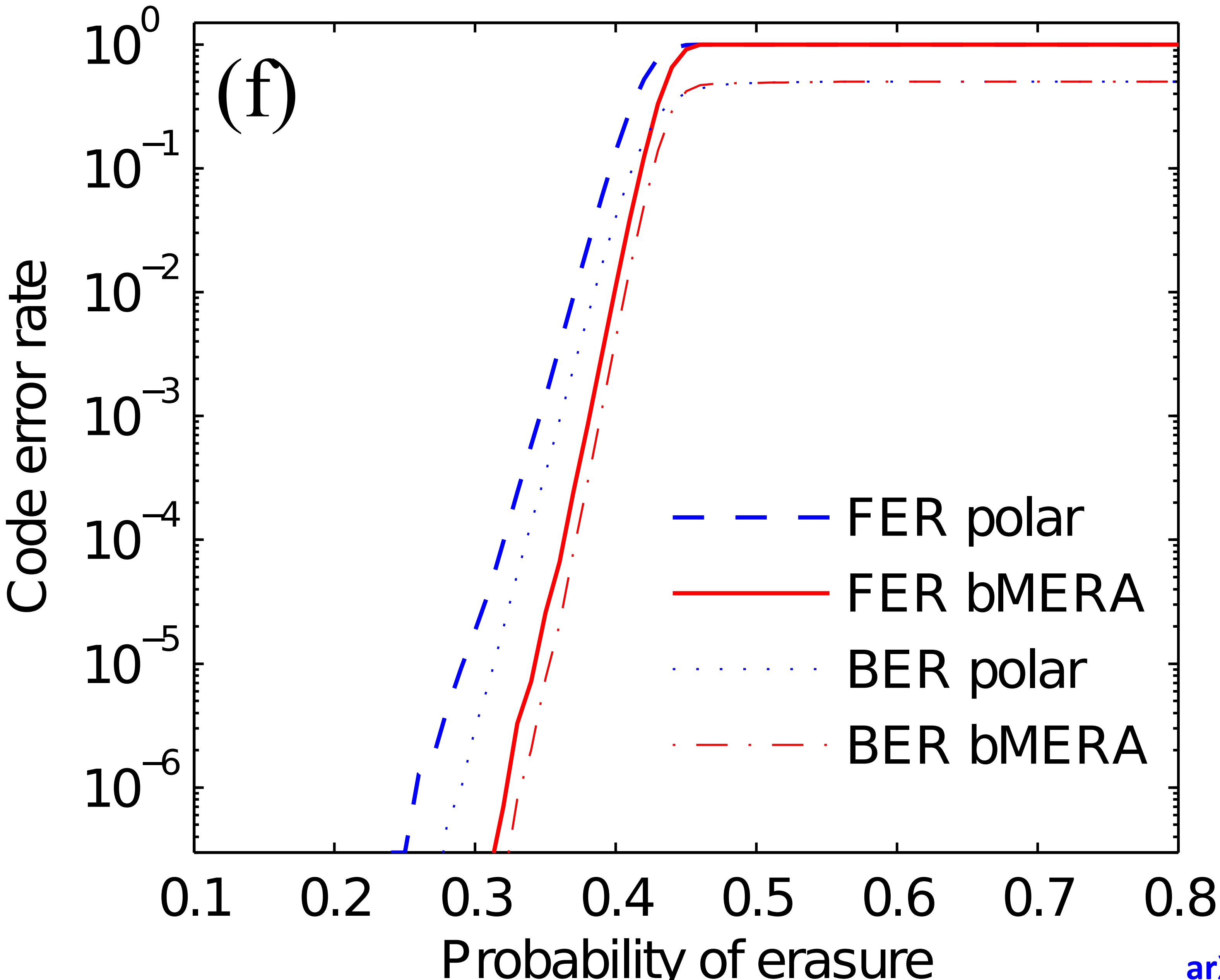
Decoding is similar to MERA/ branching-MERA contractions, done “level-by-level”.



Performance

- Numerical cost is about twice that of polar code
- However, error-correction performance is improved and can be significantly better in some regimes

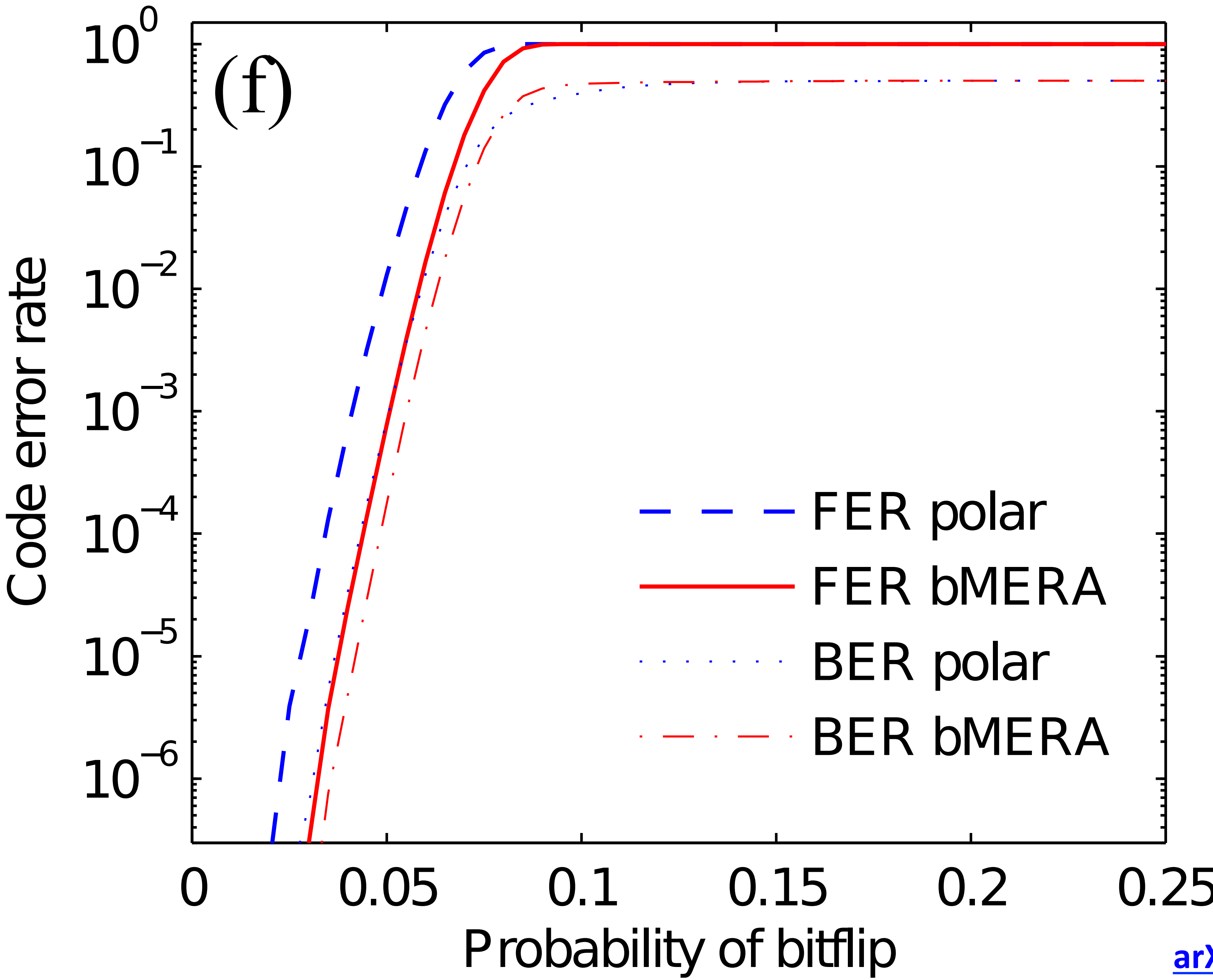
Erasure channel



8192 bits
code rate 1/2

arXiv:1312.4575

Bit-flip channel



8192 bits
code rate 1/2

arXiv:1312.4575

Classical Results

- The polar code is “almost perfect”, except for finite-code performance could be a bit better

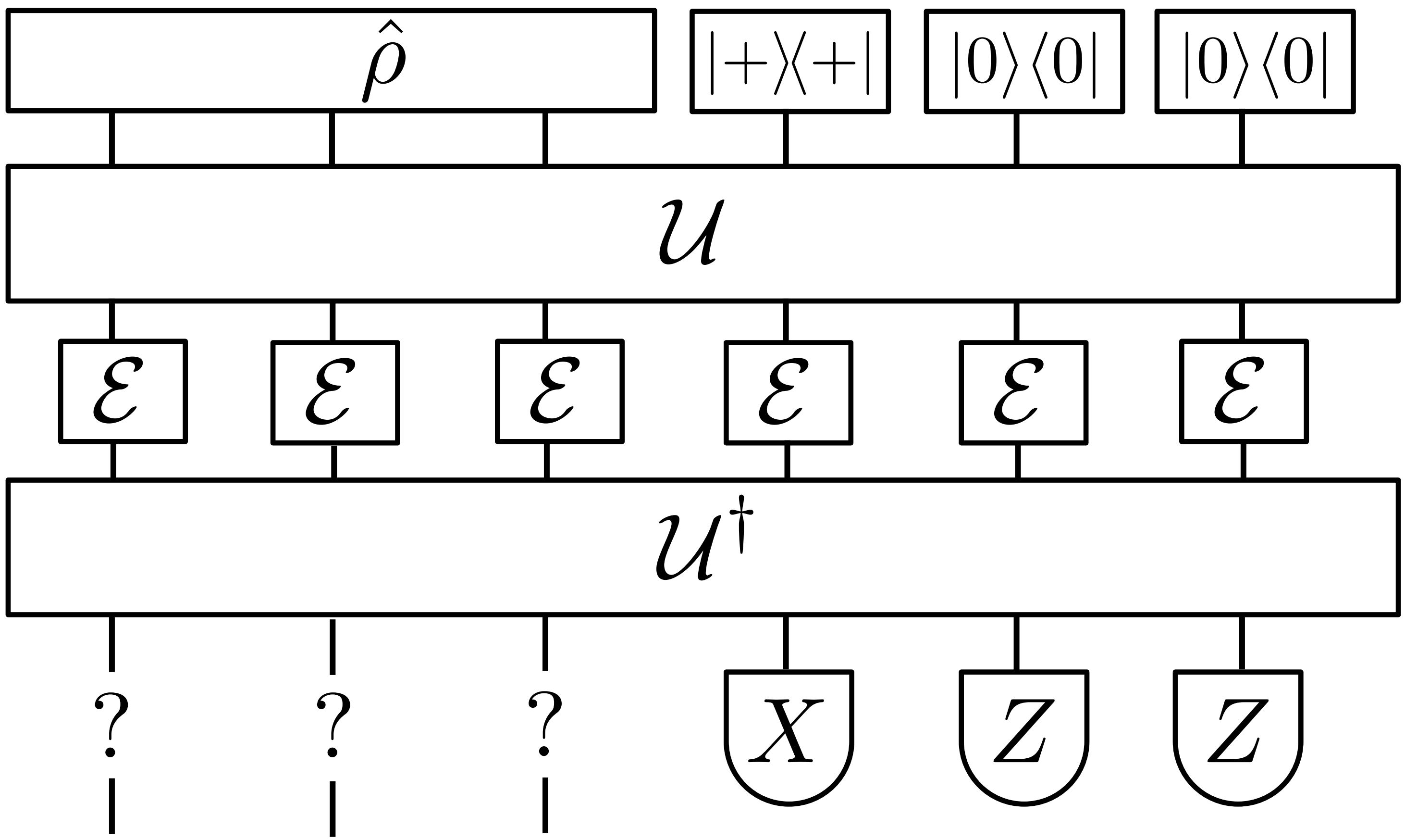
- Branching-MERA code does improve this somewhat, bringing “waterfall” region closer to the Shannon limit and significantly reducing the error-rate in the low-error regime.
  - No “error-floor” in either case, unlike some commonly-used LDPC codes.
  - Open question on scaling exponents, etc.

Quantum codes

Work on quantum polar codes:
Renes, Depuis, Renner, Wilde,
Guha, Sutter, Dutton……..

Quantum decoding problem

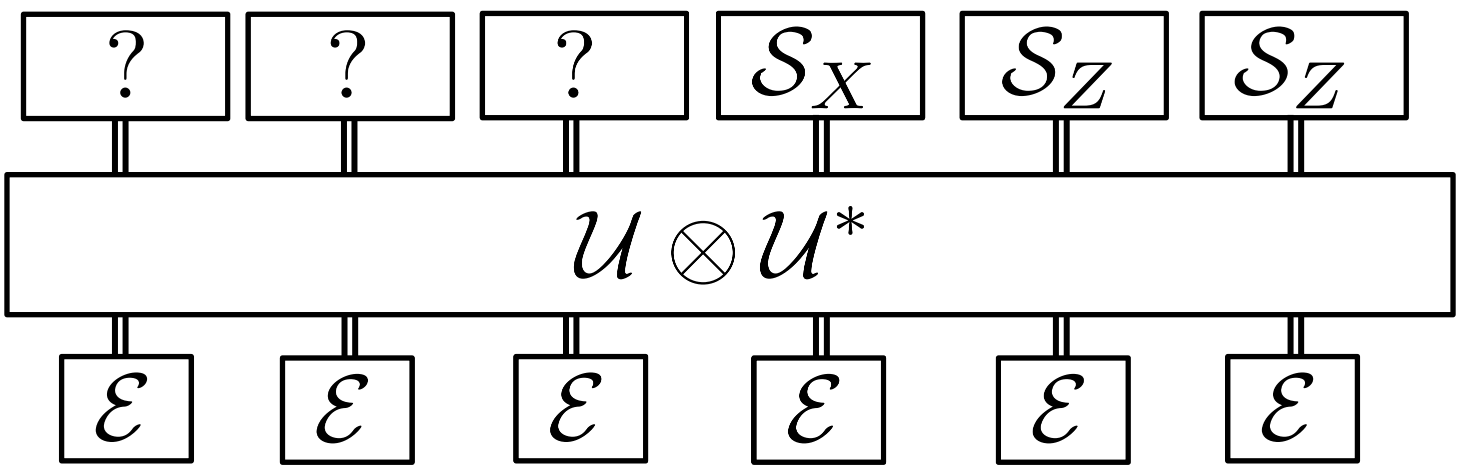
Given stabilizer measurements, what is the most likely Pauli operation which recovers the data?



Quantum decoding problem

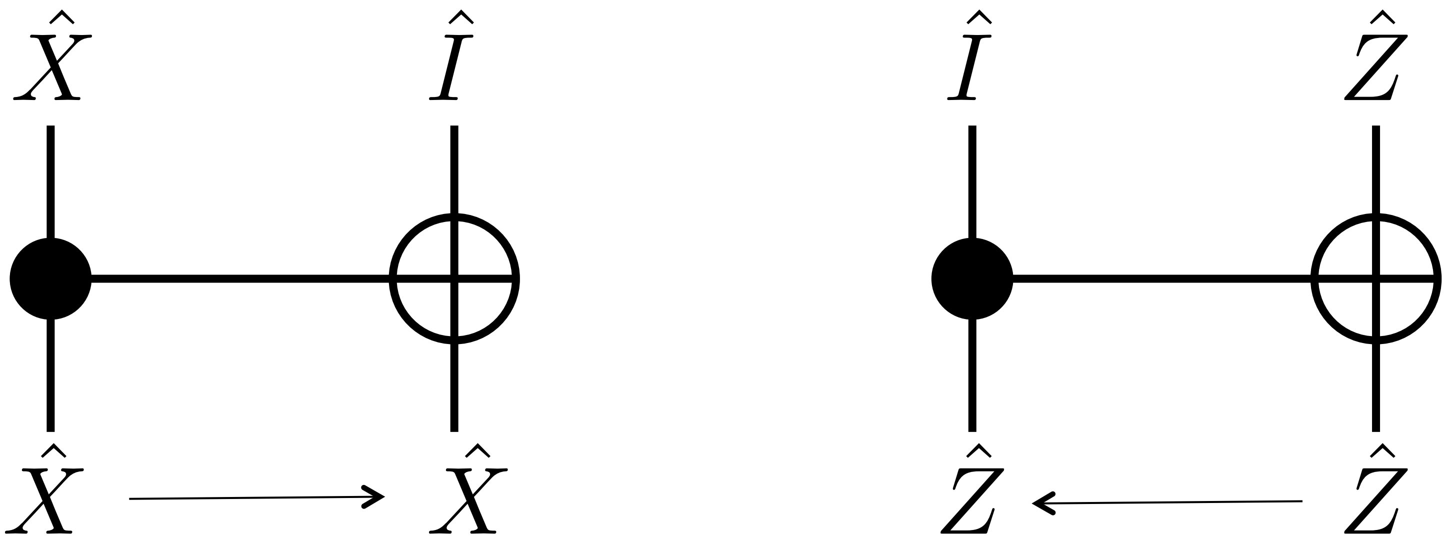
Given stabilizer measurements, what is the most likely Pauli operation which recovers the data?

Rather, given some knowledge of some (stabilizer) channels, what were the logical channels?

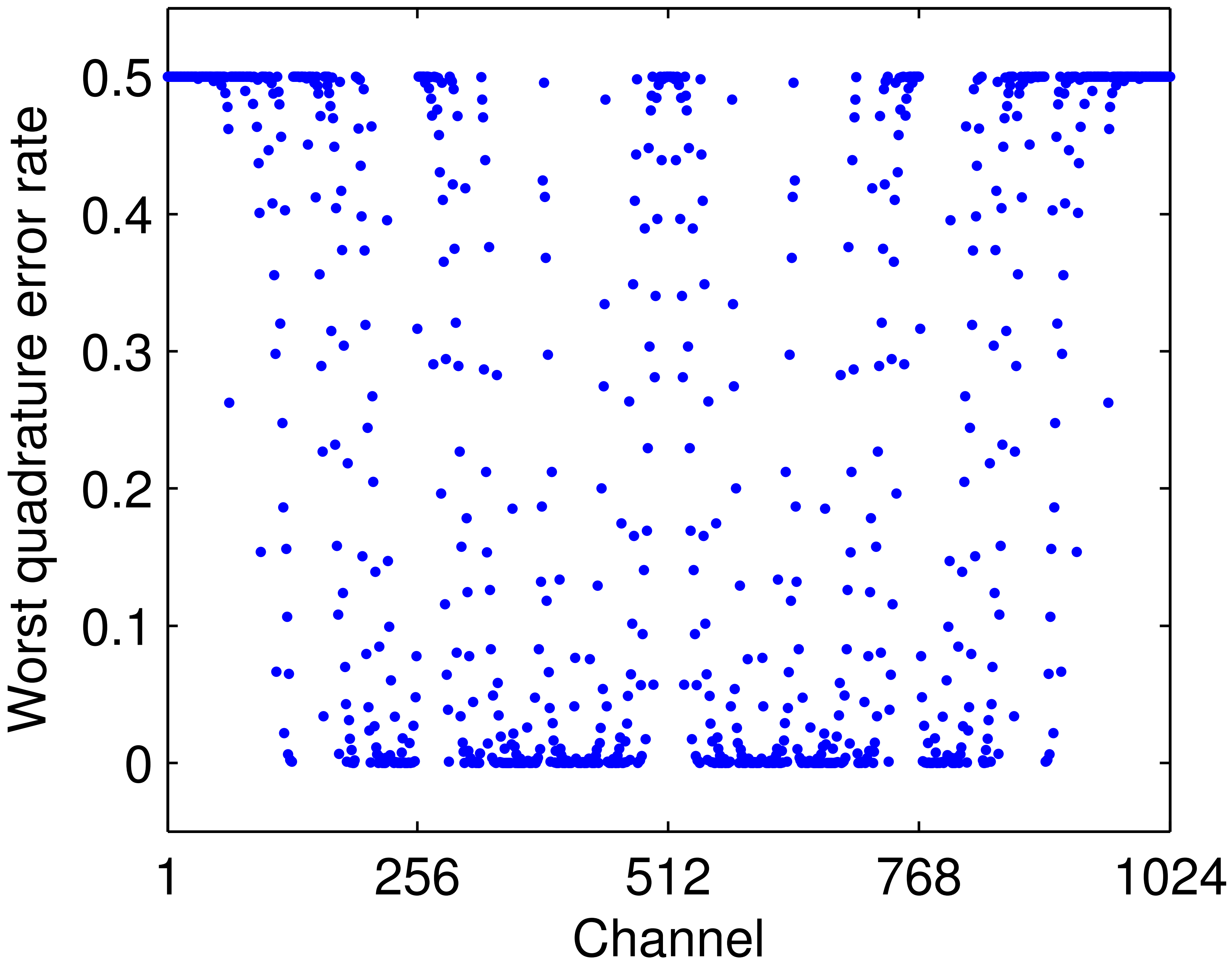


Quantum polar code

- Exactly same encoding circuit, with CNOTs



- Amplitude protected on left, while phase is protected on the right
  - Put data in the center, best of both worlds



Worst quadrature error rate
0.5
0.4
0.3
0.2
0.1
0
1
256
512
768
1024
Channel

Decoding

Decoding of $\hat{X}$ and $\hat{Z}$ errors done sequentially.

PHASE 1: Decode amplitude (right-to-left)

PHASE 2: Decode phase (left-to-right)

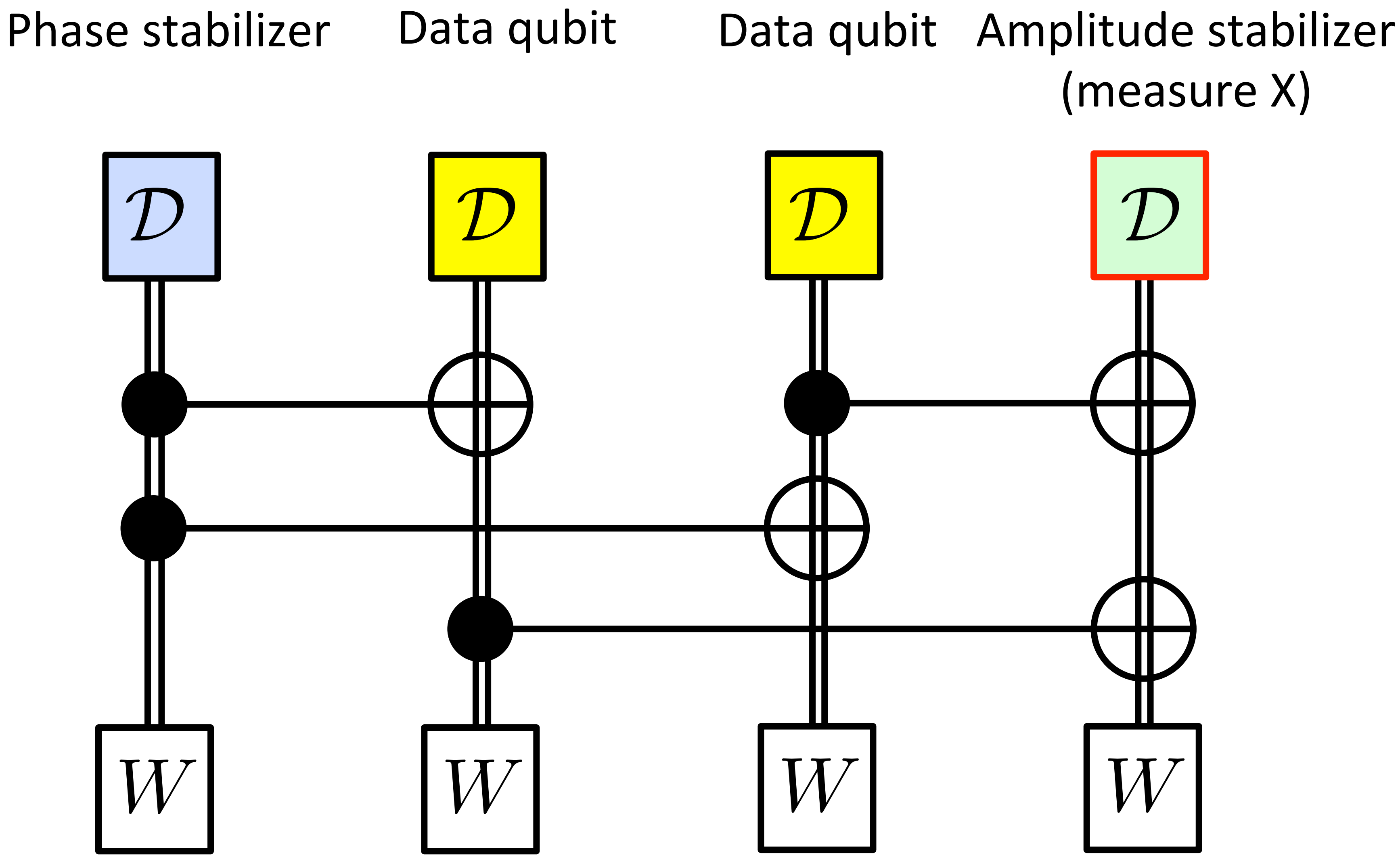
Dutton, Guha, Wilde (2012)

Task is to determine what the error channel on each bit, using successive cancelation:

Initial state: $\hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$

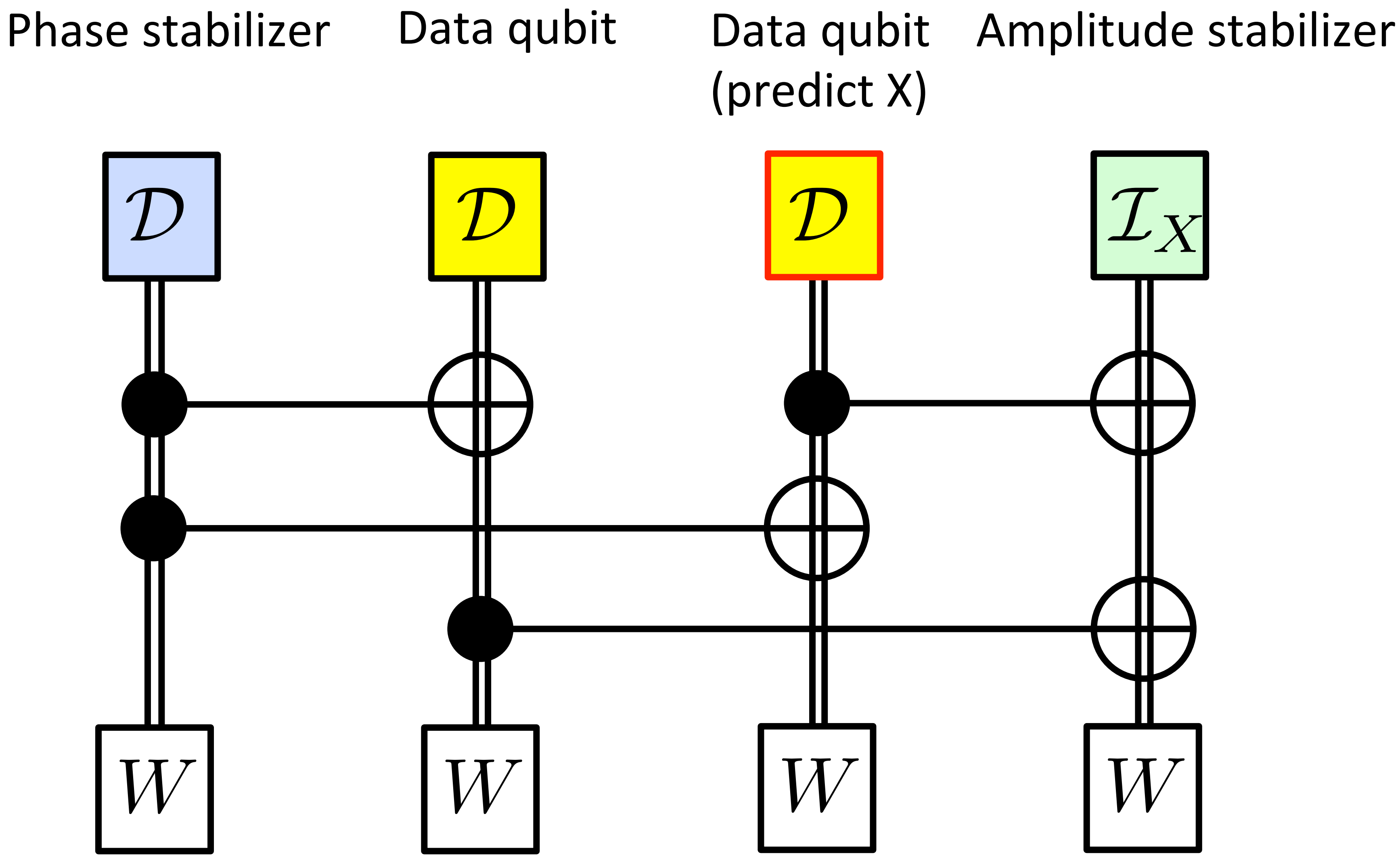
After phase 1: $\hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z}$ or $\hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$

After phase 2: $\hat{I} \otimes \hat{I}$ or $\hat{X} \otimes \hat{X}$ or $\hat{Y} \otimes \hat{Y}$ or $\hat{Z} \otimes \hat{Z}$



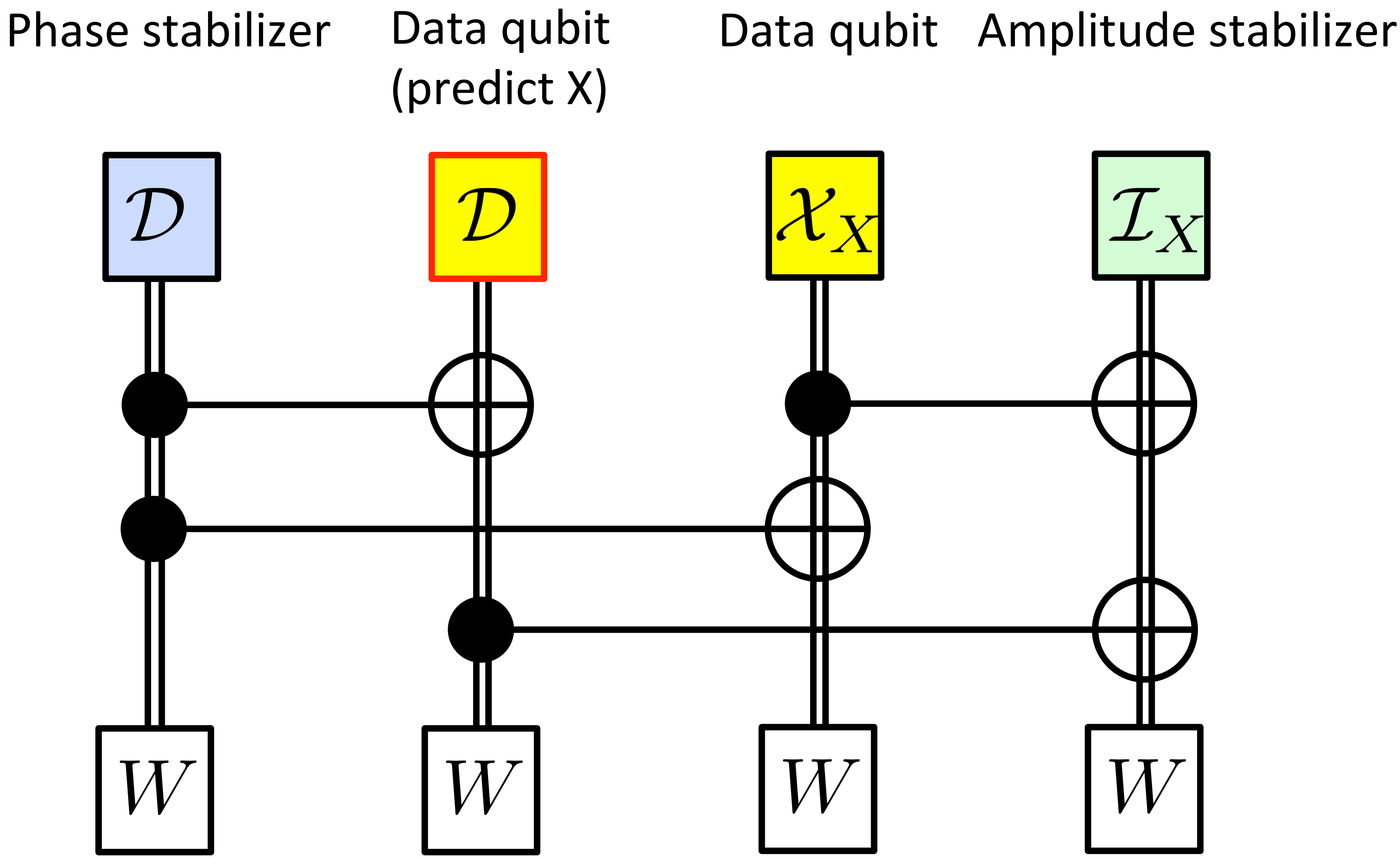
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



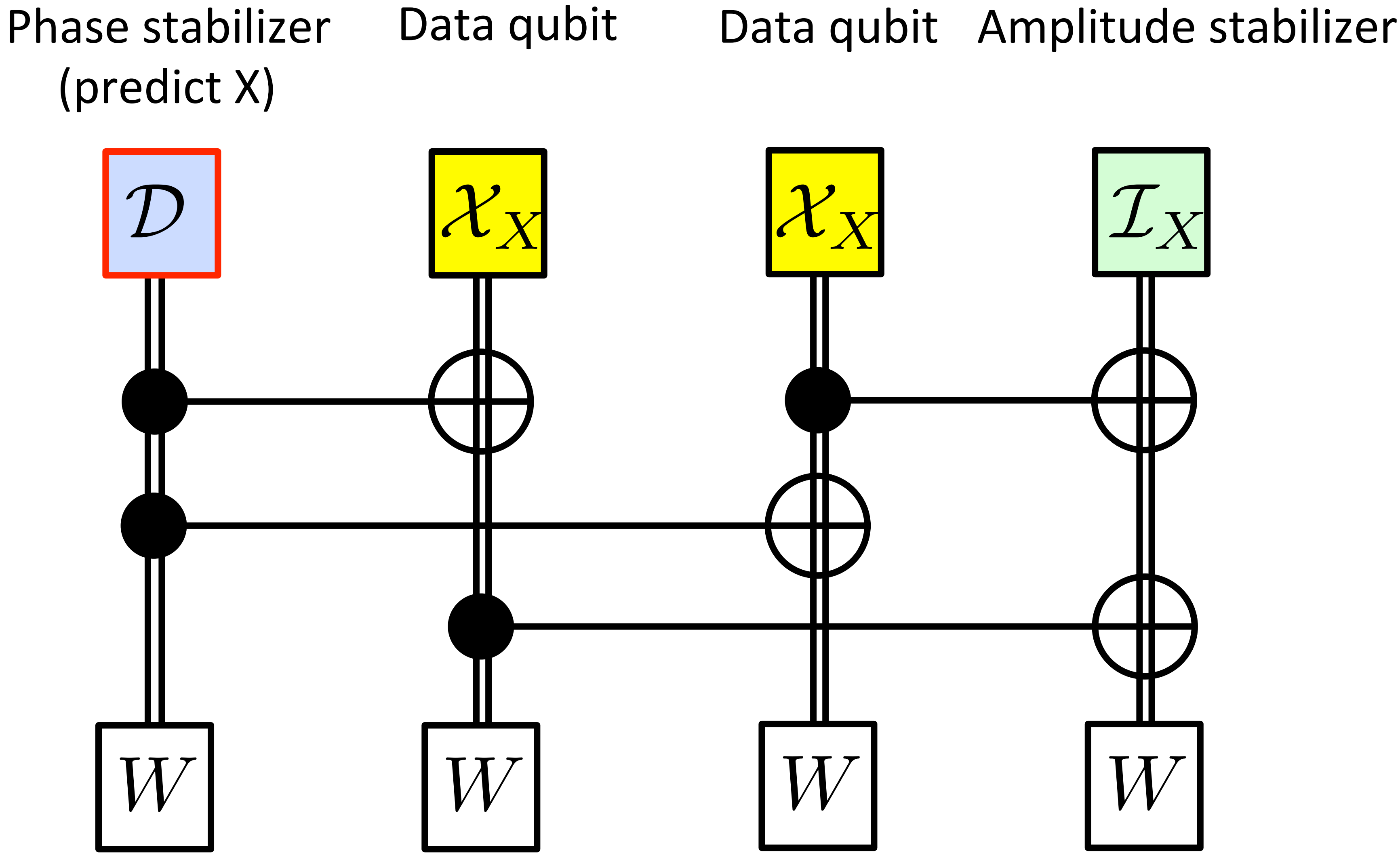
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



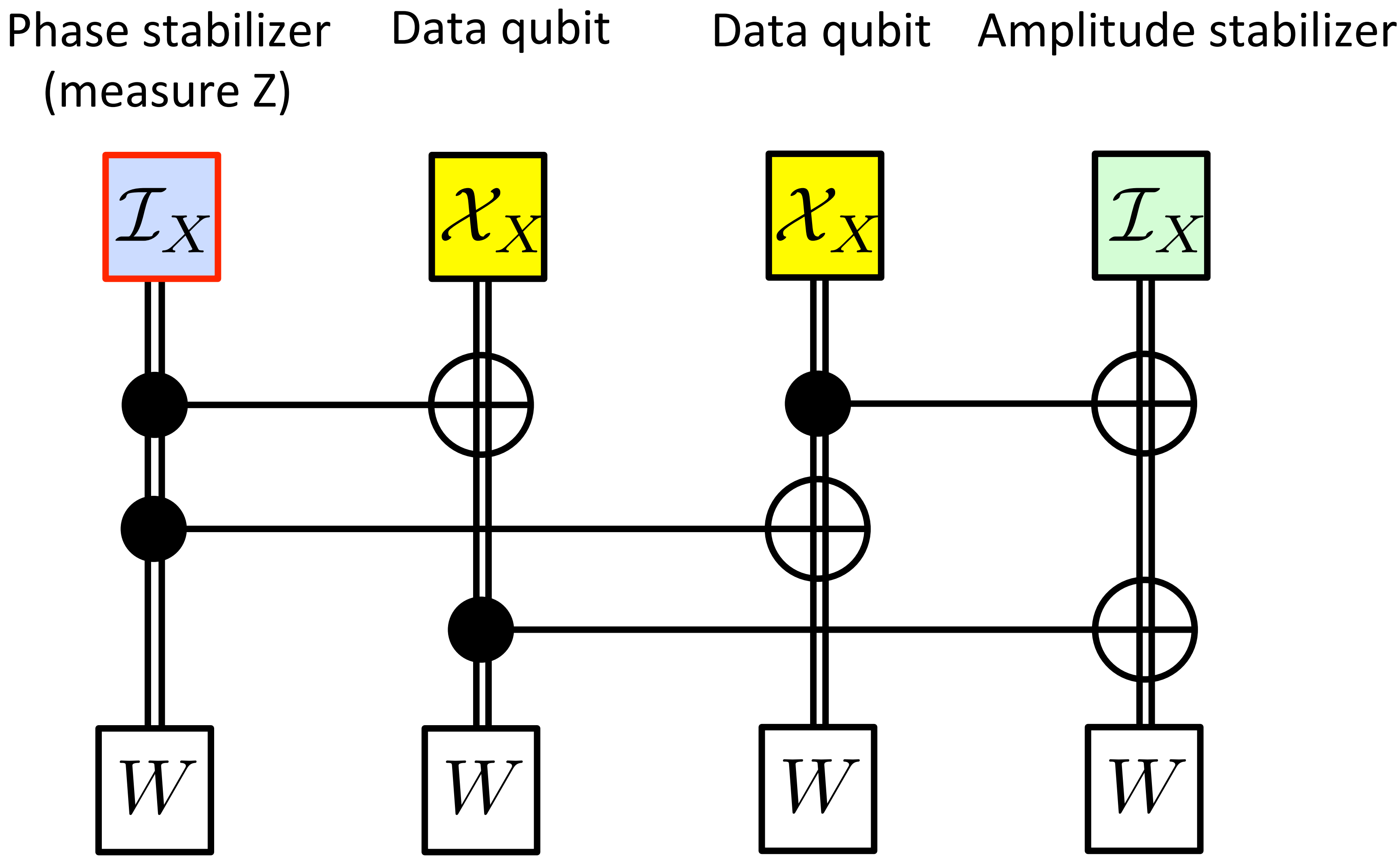
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



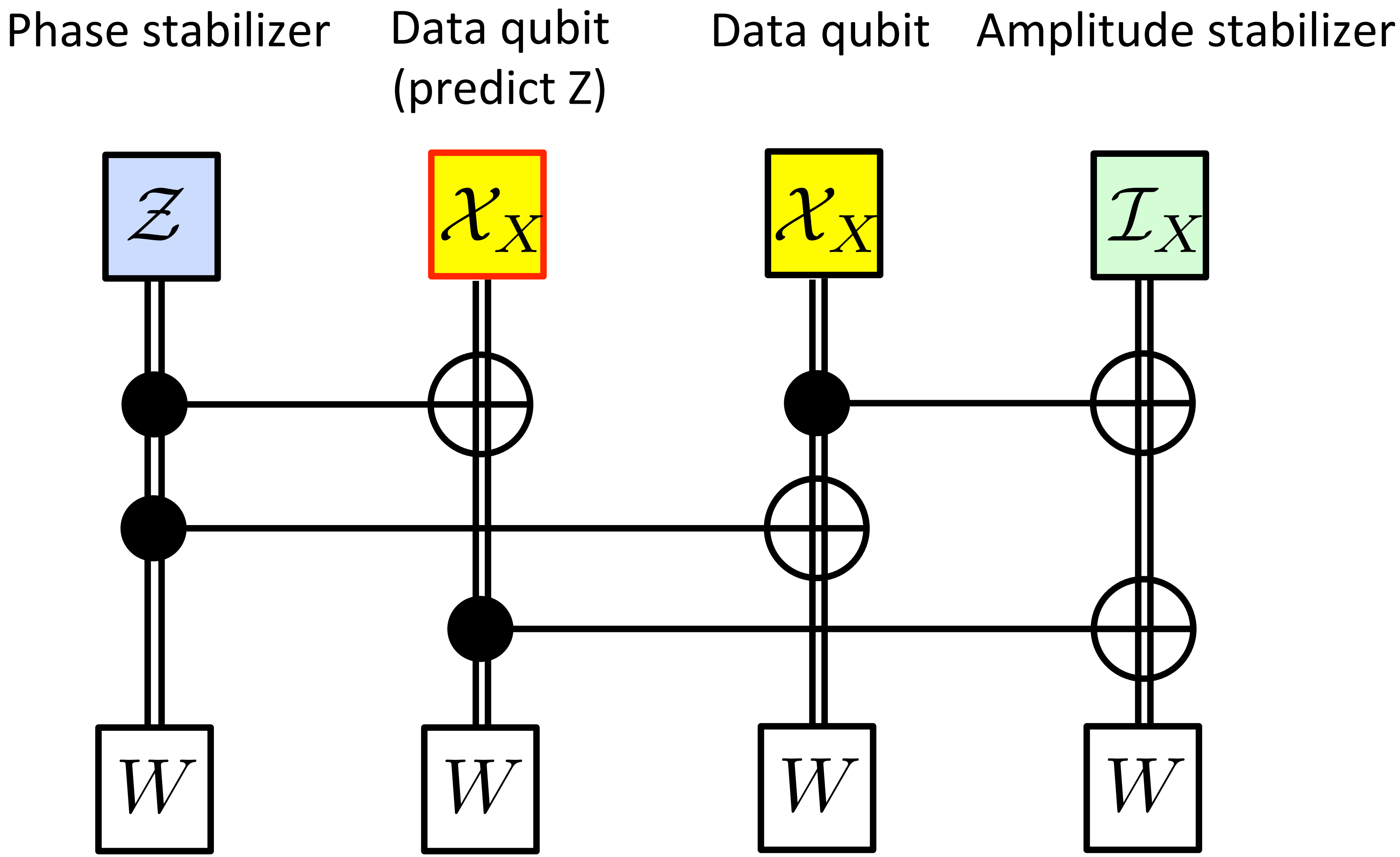
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



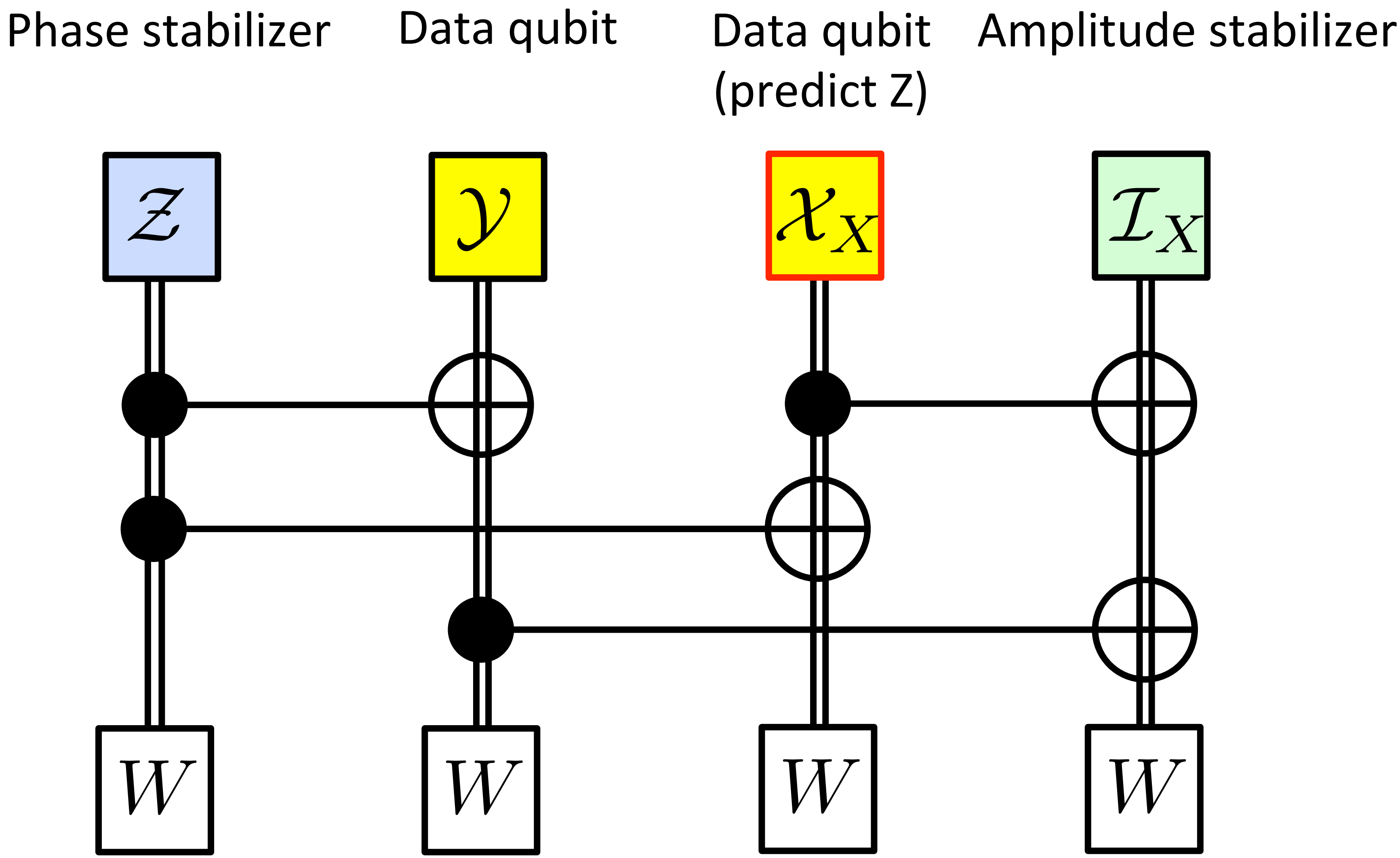
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



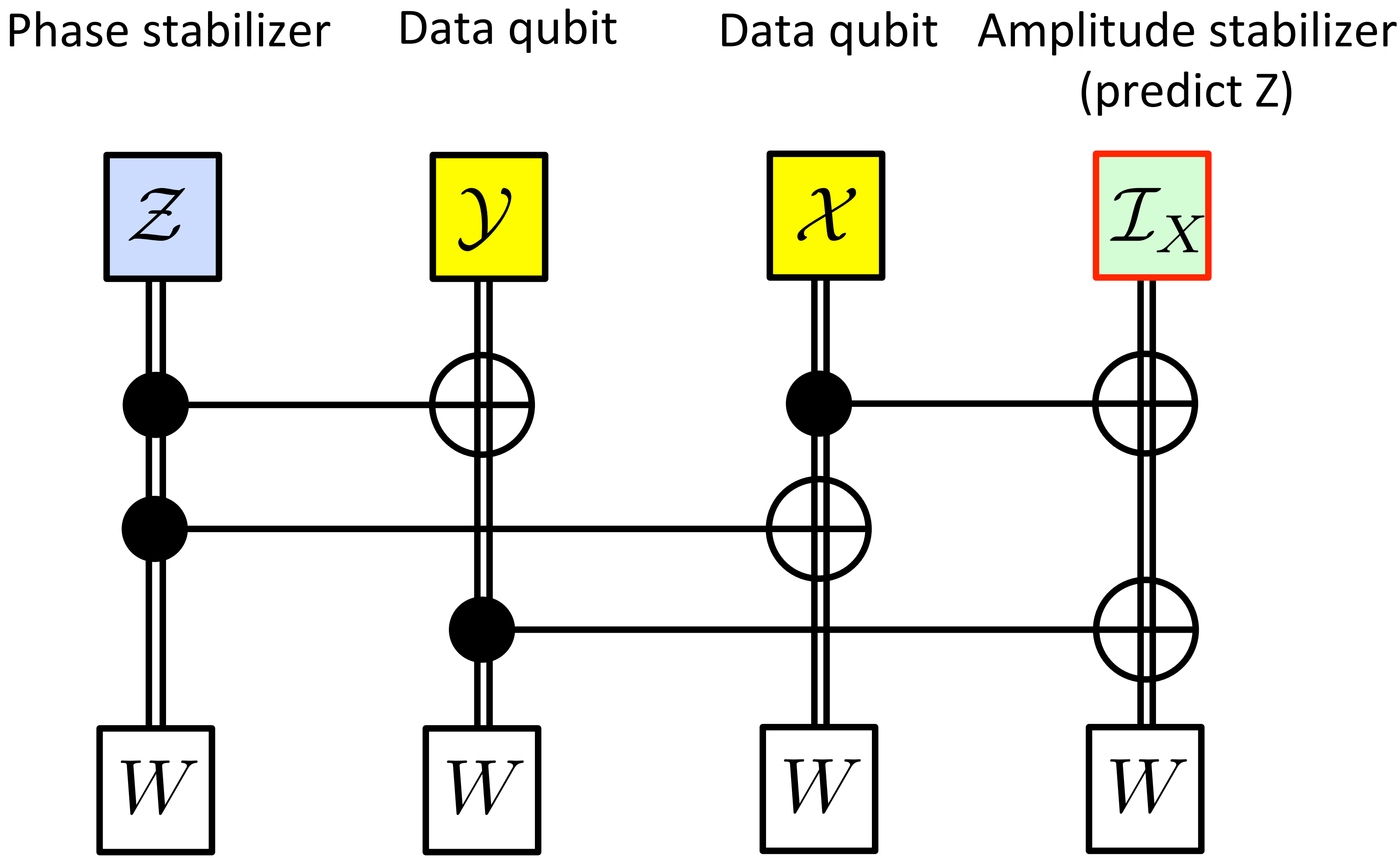
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

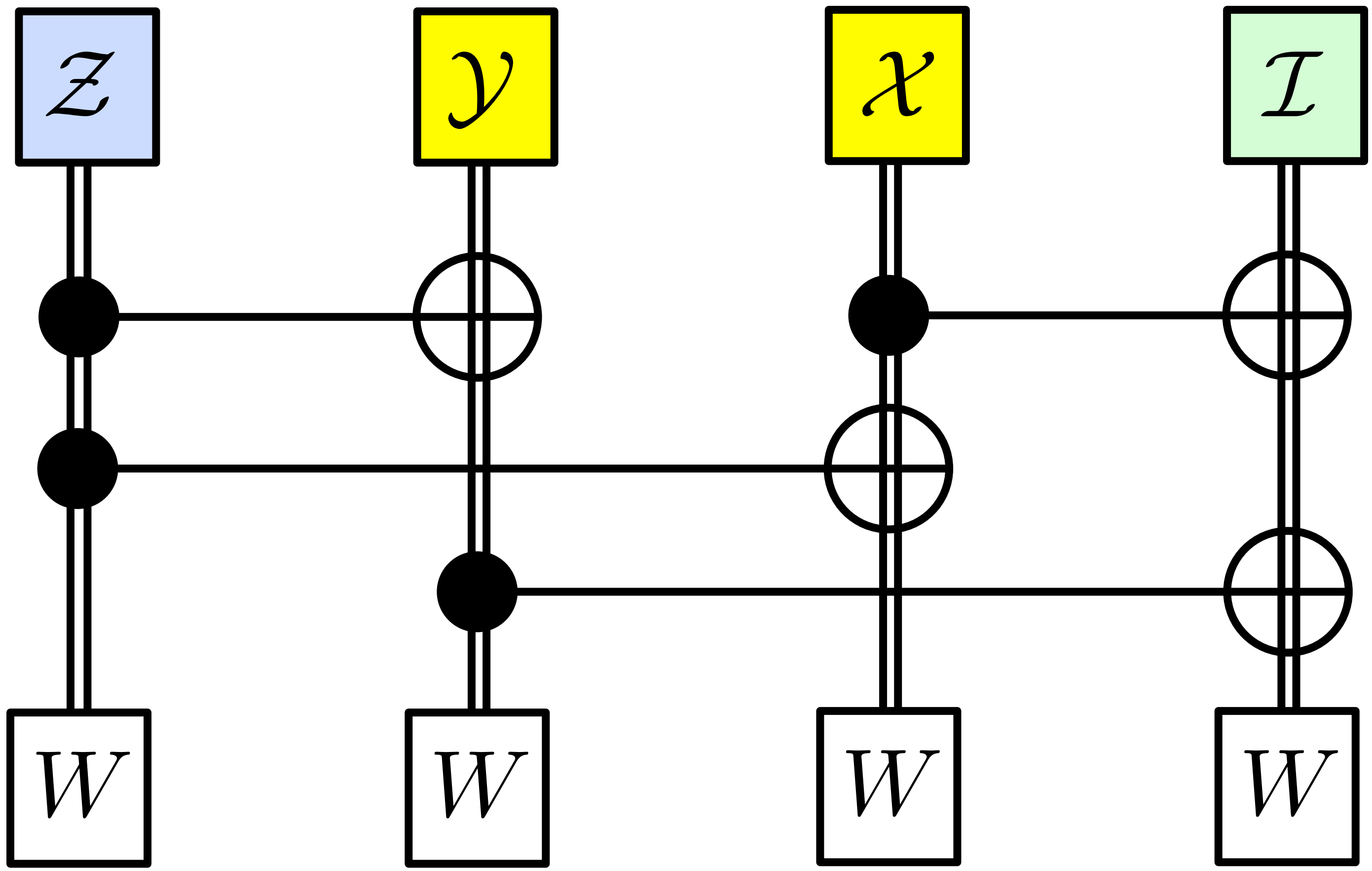
$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$

DONE!



$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

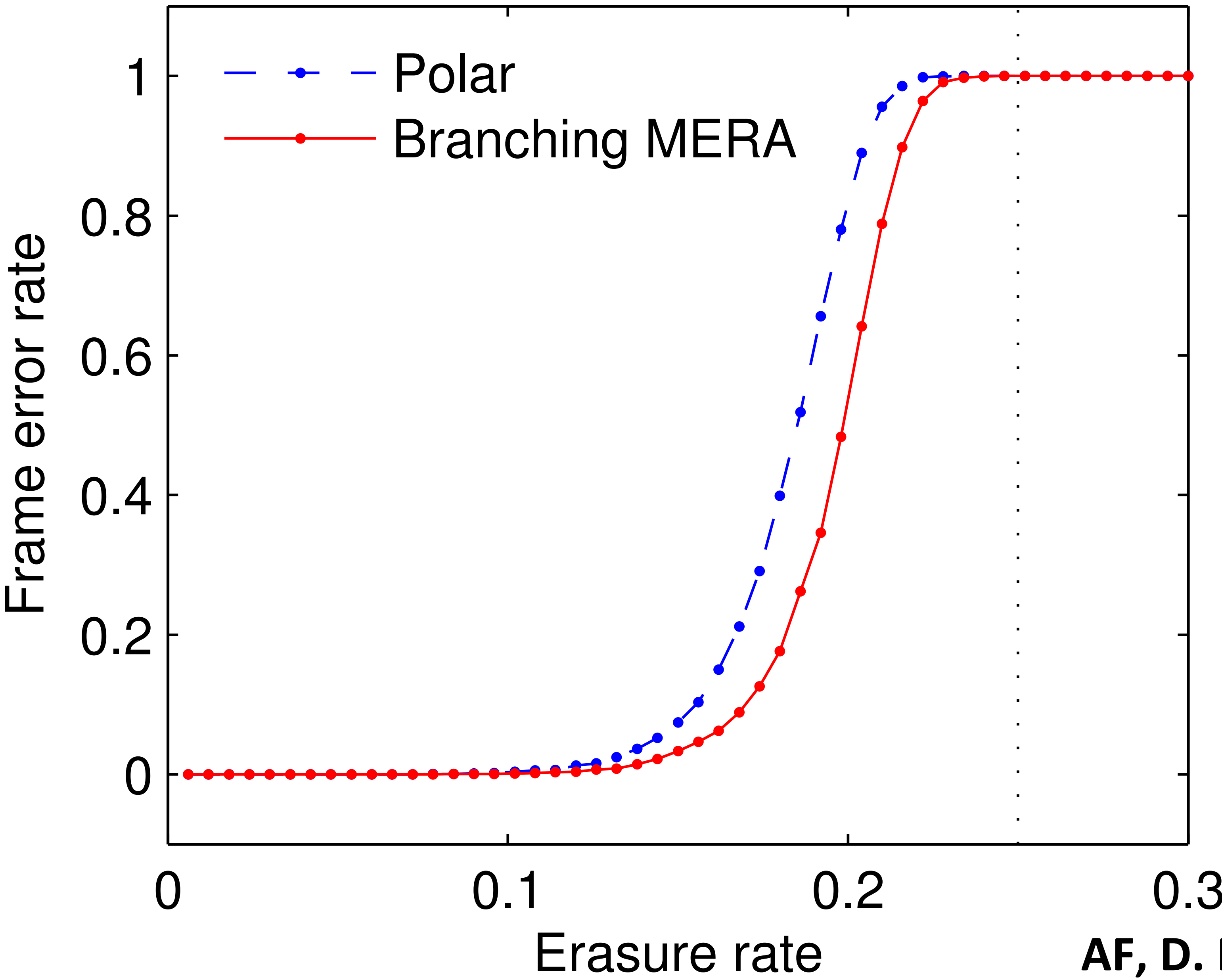
$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$

Numerical results

- Can simulate this with erasure channel, Pauli channel, etc.

- The branching-MERA code is also a good quantum code!

Quantum Erasure Channel

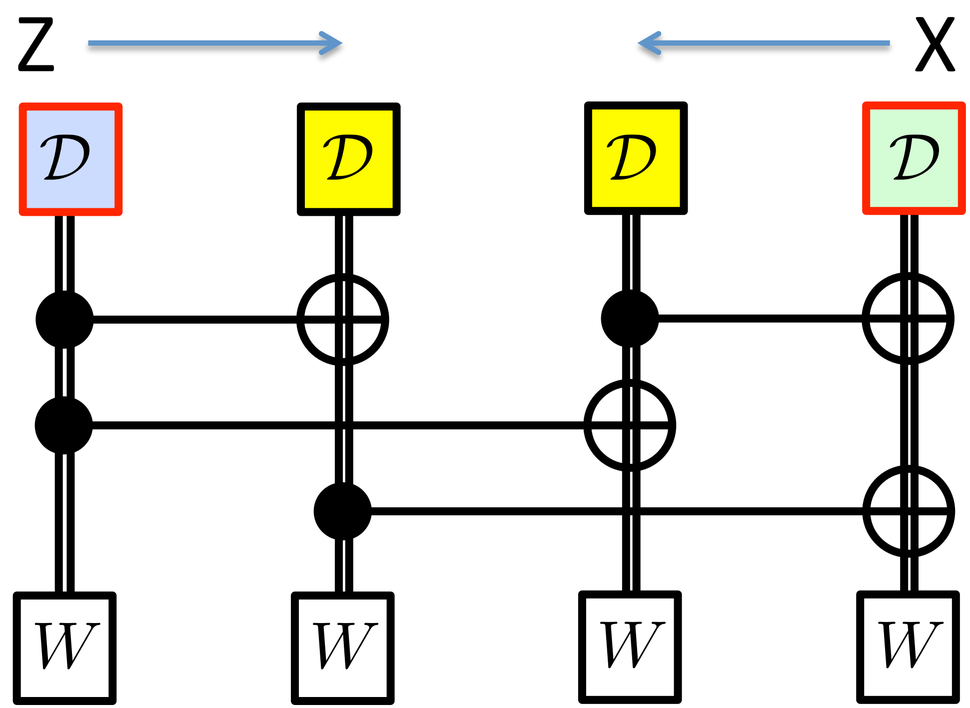


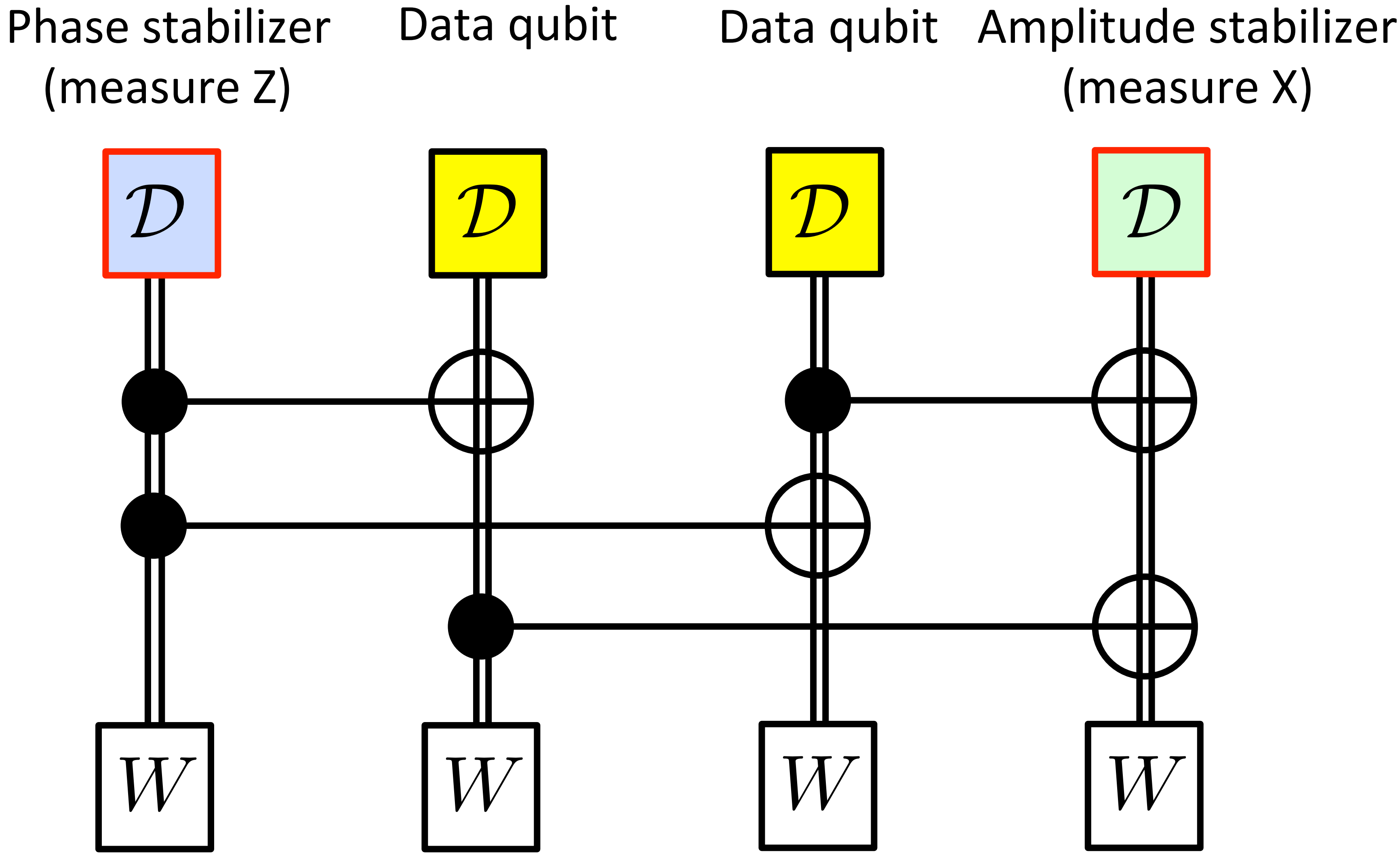
4096 qubits with 2048 data qubits

AF, D. Poulin (2014)

X/Z correlations

- Second step: Decoding Z knows about X
- First step: Decoding X doesn't know about Z…
- Can decode the problem symmetrically
  - X from right
  - Z from left
  - Simultaneous

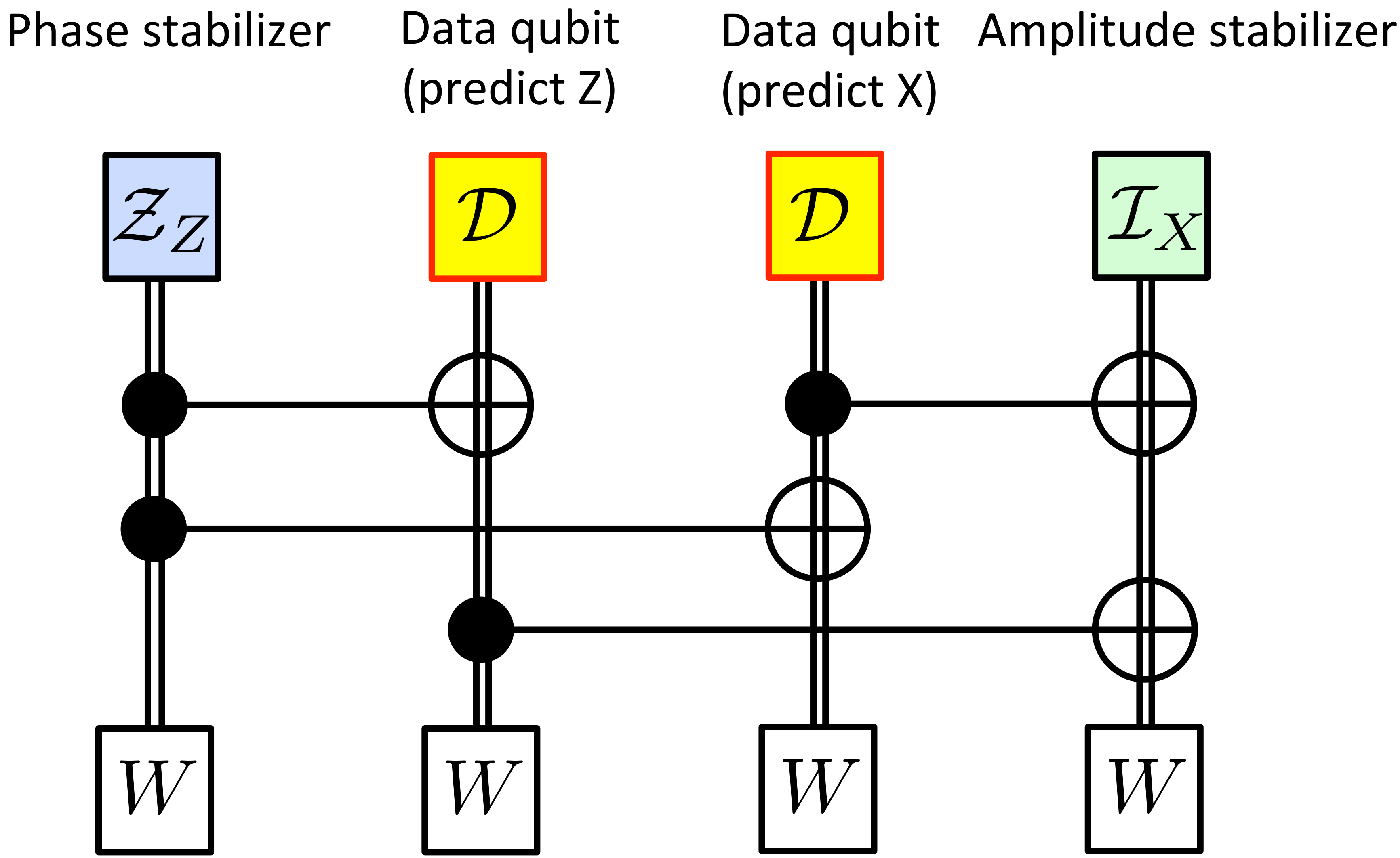




$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{I}_Z = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} \qquad \mathcal{Z}_Z = \hat{Z} \otimes \hat{Z} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



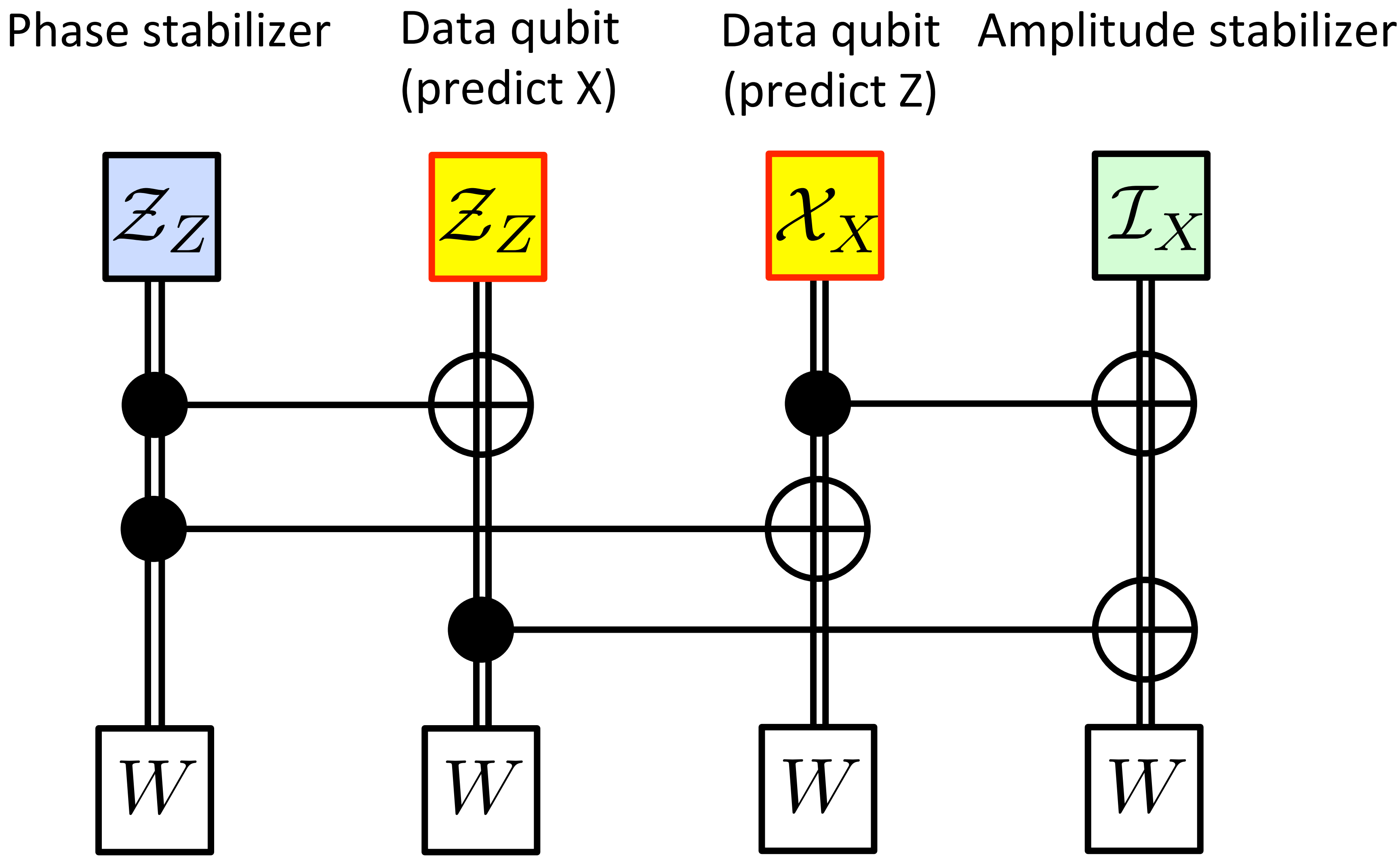
$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z}$$

$$\mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{I}_Z = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X}$$

$$\mathcal{Z}_Z = \hat{Z} \otimes \hat{Z} + \hat{Y} \otimes \hat{Y}$$

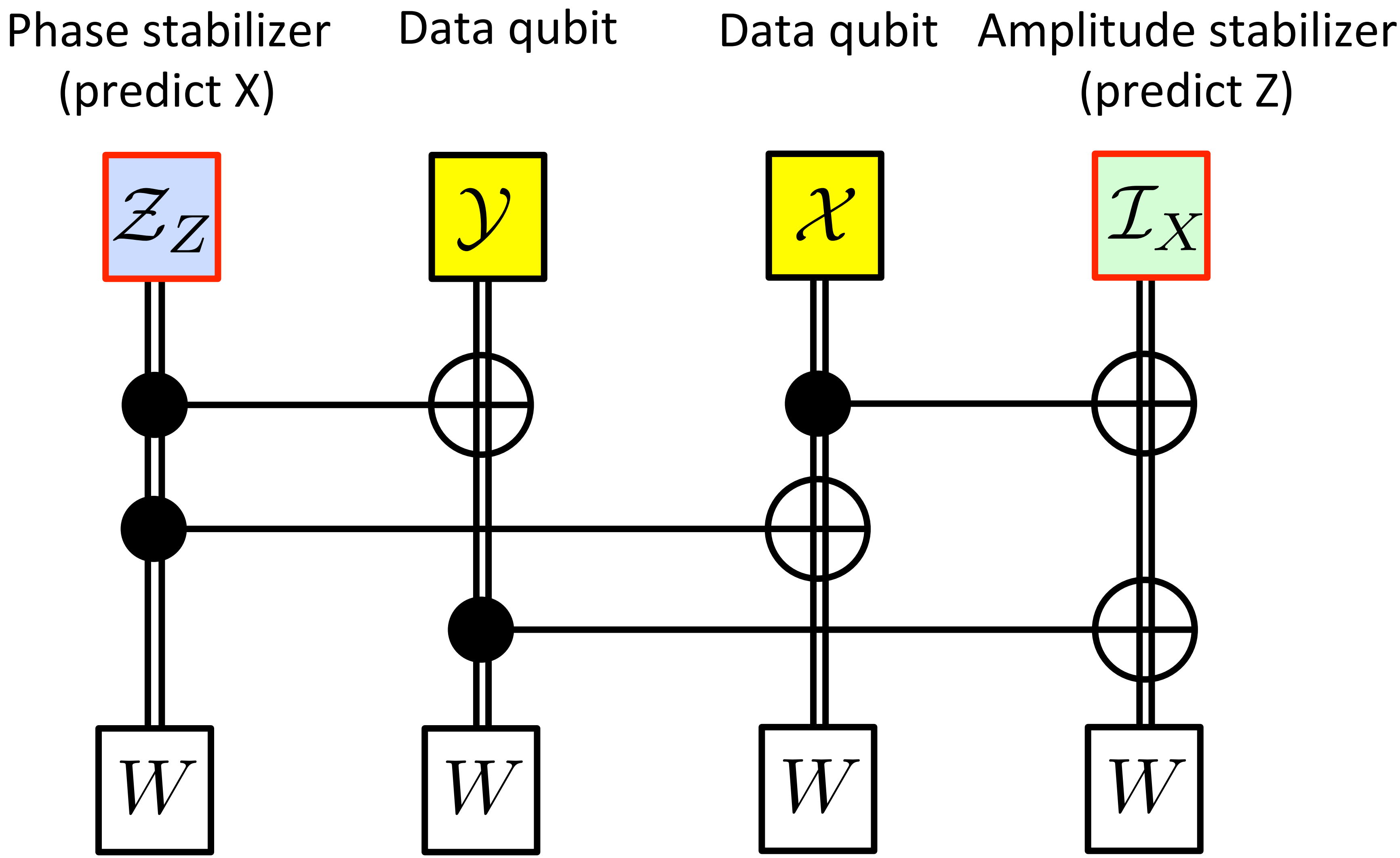
$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$



$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{I}_Z = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} \qquad \mathcal{Z}_Z = \hat{Z} \otimes \hat{Z} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$

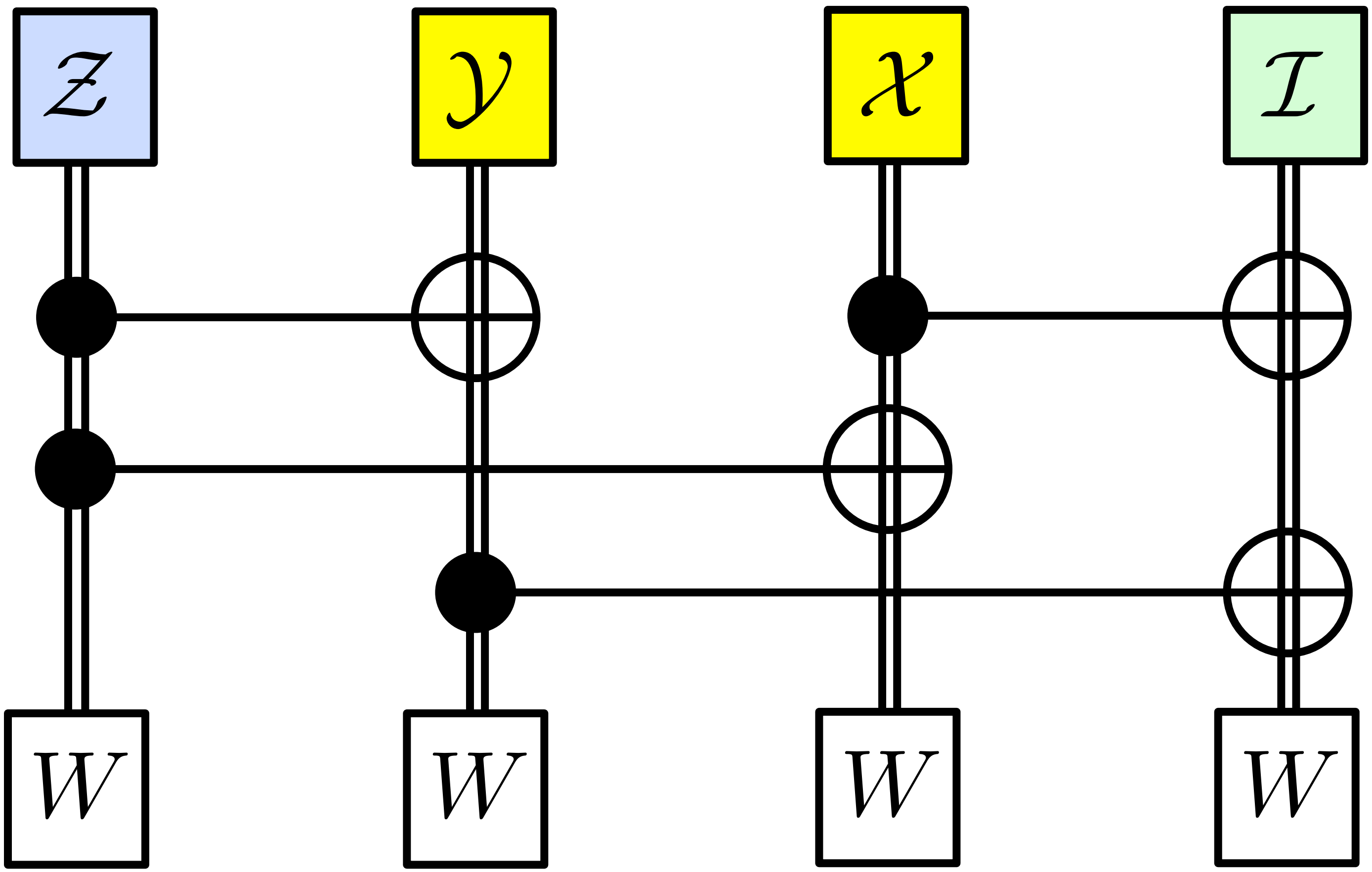


$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{I}_Z = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} \qquad \mathcal{Z}_Z = \hat{Z} \otimes \hat{Z} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$

DONE!

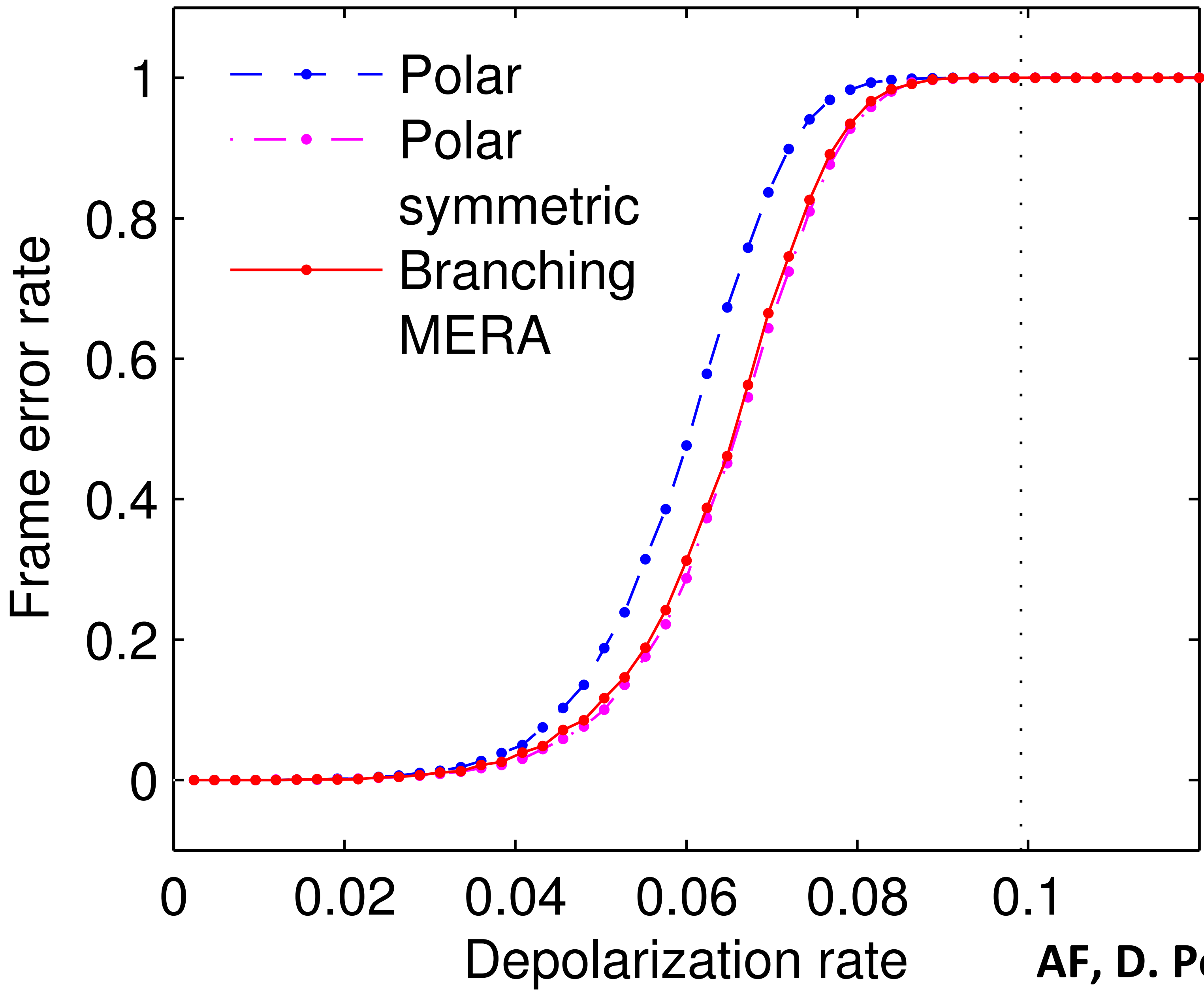


$$\mathcal{I}_X = \hat{I} \otimes \hat{I} + \hat{Z} \otimes \hat{Z} \qquad \mathcal{X}_X = \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{I}_Z = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} \qquad \mathcal{Z}_Z = \hat{Z} \otimes \hat{Z} + \hat{Y} \otimes \hat{Y}$$

$$\mathcal{D} = \hat{I} \otimes \hat{I} + \hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y} + \hat{Z} \otimes \hat{Z}$$

Quantum Depolarizing Channel



4096 qubits with 2048 data qubits

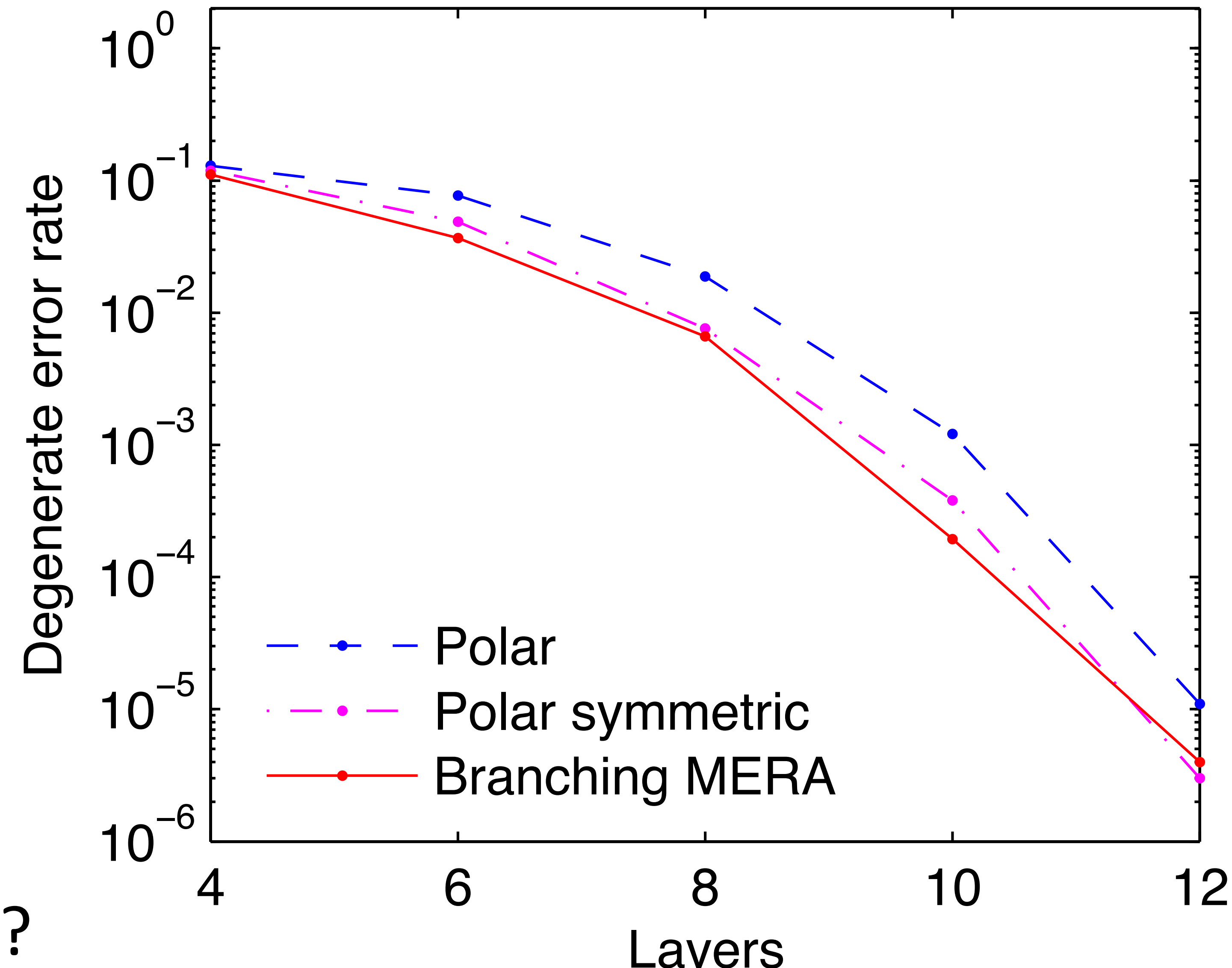
AF, D. Poulin (2014)

Non-degenerate behaviour

The code almost never has degenerate errors!

Hashing bound only

Other constructions?



Wait a minute…

Everything seems great, but we haven't talked about:

Fault tolerance

E.g. all stabilizers are large

Idea: use purified ancilla codes to perform correction, measurement, gates, etc…
(*a la* Todd Brun, yesterday)

Conclusion

- Tensor networks provide insight into coding!

- Many extensions to this work is possible
  - Use more complicated tensors than CNOT
    - gates on 2 qudits
    - Do bosonic/fermionic gates make sense?
  - Better decoders
  - Use for code-based encryption, source-coding
  - **Fault tolerance?**

Thank you!