Title: Neural Belief-Propagation Decoders for Quantum Error-Correcting Codes

Speakers: Yehua Liu

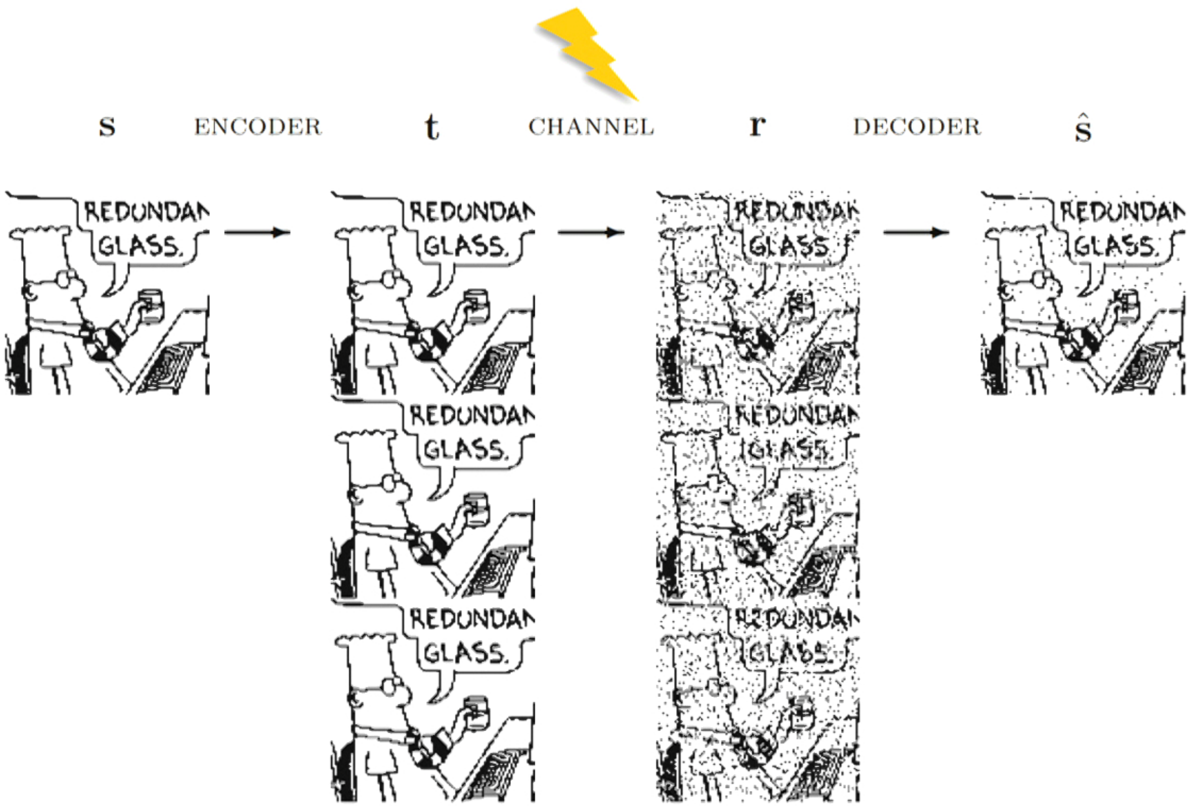
Collection: Machine Learning for Quantum Design

Date: July 11, 2019 - 4:00 PM

URL: http://pirsa.org/19070031

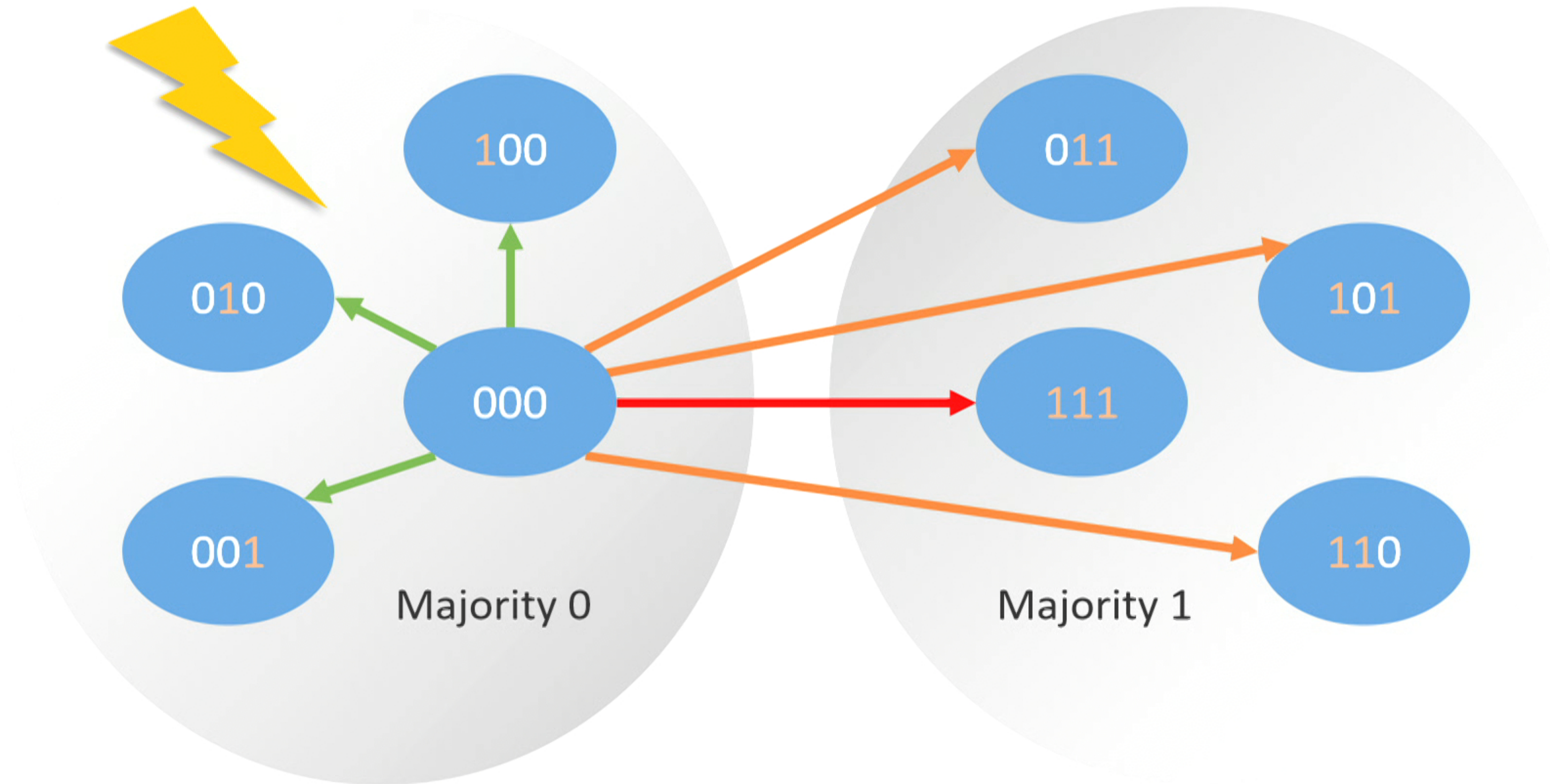
Abstract: Belief-propagation (BP) decoders are responsible for the success of many modern coding schemes. While many classical coding schemes have been generalized to the quantum setting, the corresponding BP decoders are flawed by design in this setting. Inspired by an exact mapping between BP and deep neural networks, we train neural BP decoders for quantum low-density parity-check codes, with a loss function tailored for the quantum setting. Training substantially improves the performance of the original BP decoders. The flexibility and adaptability of the neural BP decoders make them suitable for low-overhead error correction in near-term quantum devices.
Reference: arXiv:1811.07835 (to appear in PRL)

Error correction schematics

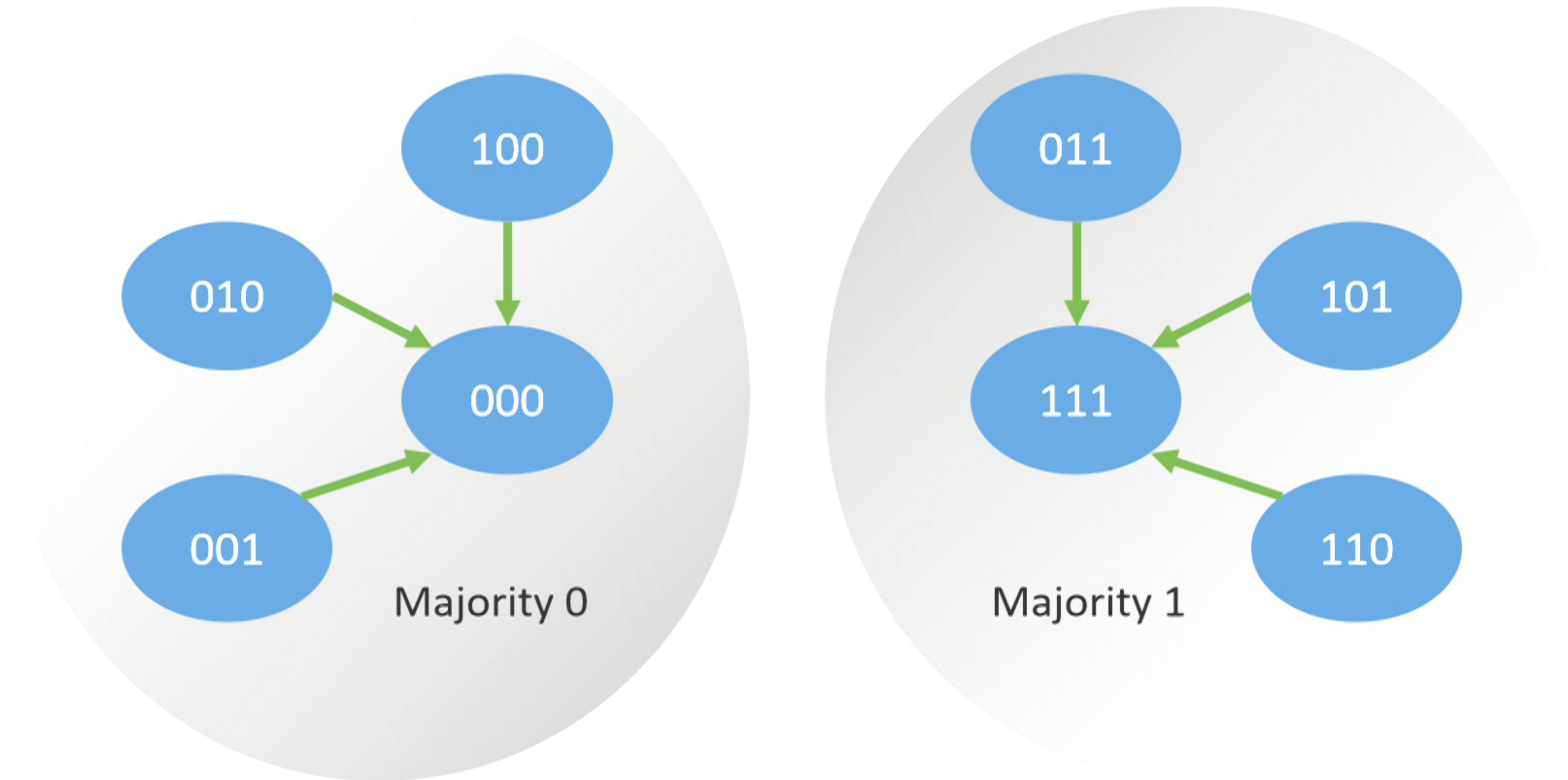


MacKay, *"Information Theory, Inference, and Learning Algorithms"*

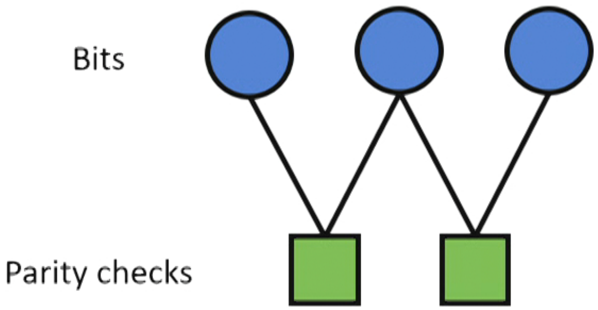
Repetition code: Encoding and transmission

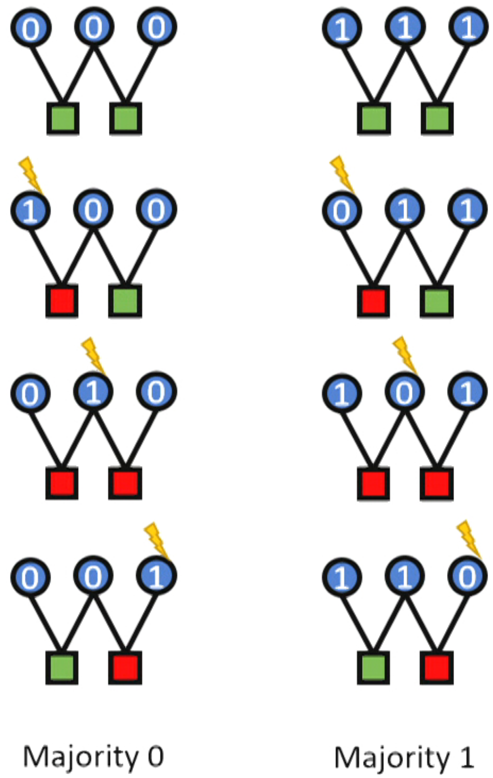


Repetition code: Decoding



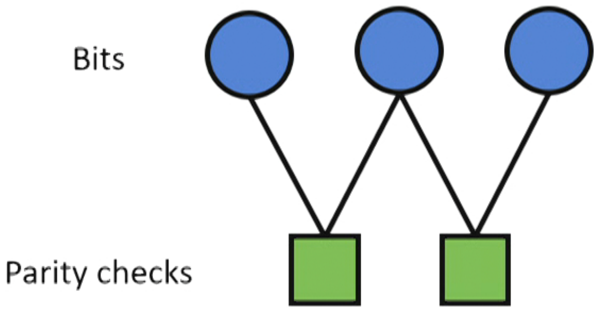
Tanner graph for repetition code



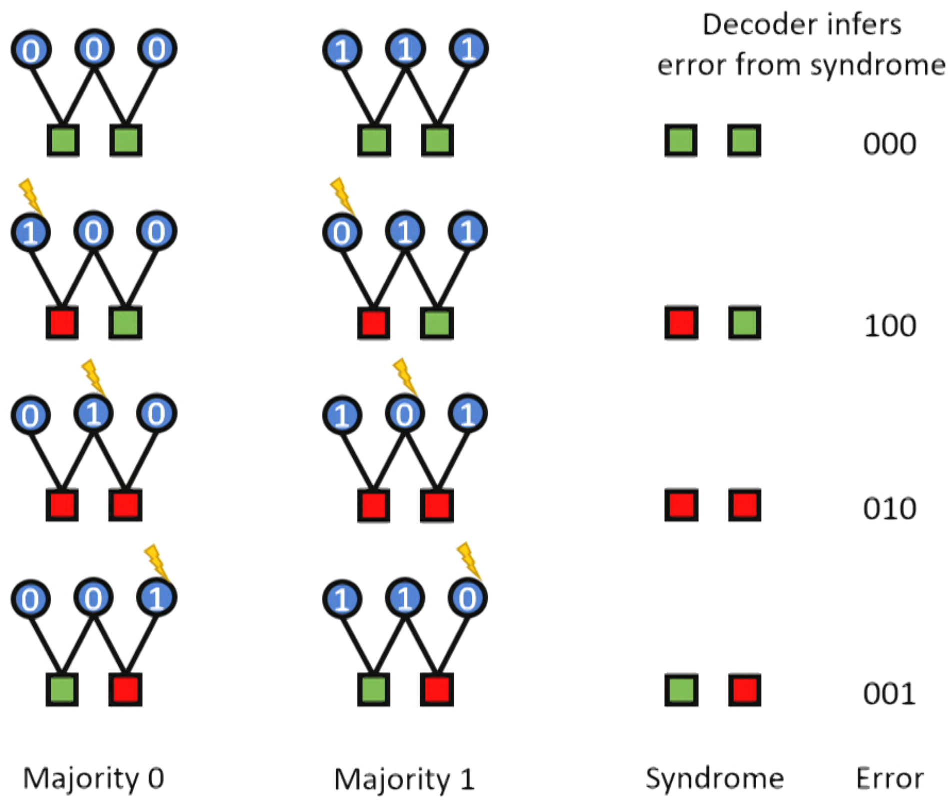


Check is satisfied/unsatisfied:
All incoming bits sum to even/odd.

Tanner graph for repetition code

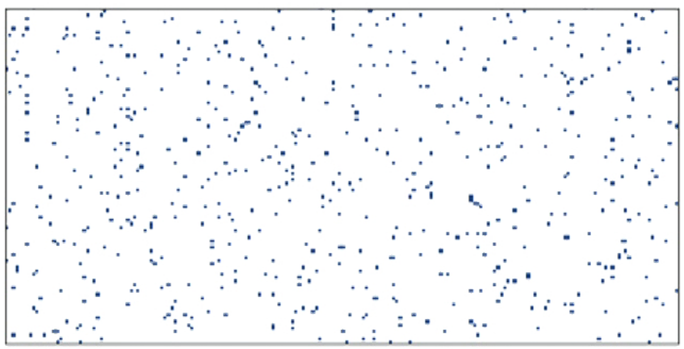


Check is satisfied/unsatisfied:
All incoming bits sum to even/odd.



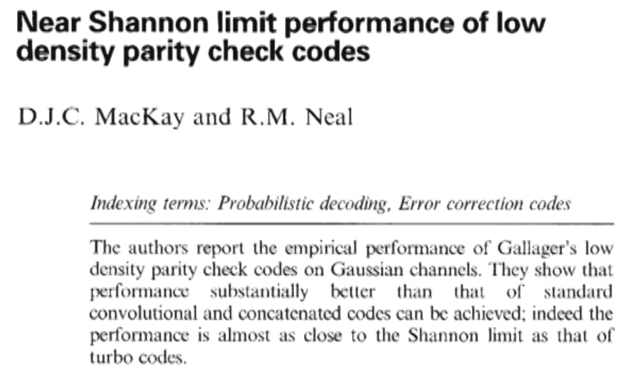
Classical **L**ow-**D**ensity-**P**arity-**C**heck codes

H =



Random sparse parity-check matrix

Robert G. Gallager, 1963



Near Shannon limit performance of low density parity check codes

D.J.C. MacKay and R.M. Neal

Indexing terms: Probabilistic decoding, Error correction codes

The authors report the empirical performance of Gallager's low density parity check codes on Gaussian channels. They show that performance substantially better than that of standard convolutional and concatenated codes can be achieved; indeed the performance is almost as close to the Shannon limit as that of turbo codes.

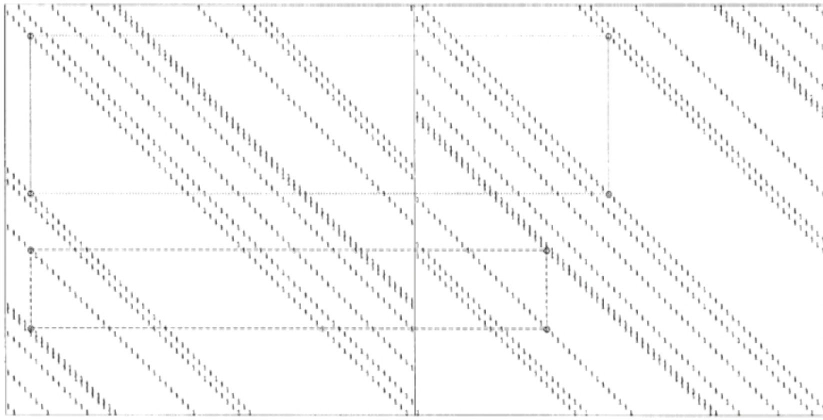
Near optimal performance combined with

belief-propagation decoder

MacKay and Neal, 1997

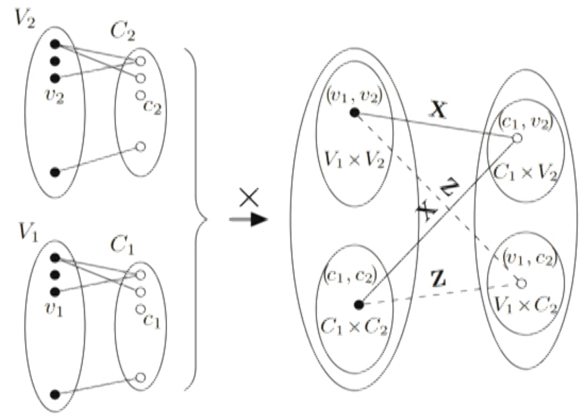
7

Quantum LDPC codes



Quantum bicycle code

MacKay, Mitchison, and McFadden, 2004



Quantum hypergraph-product code

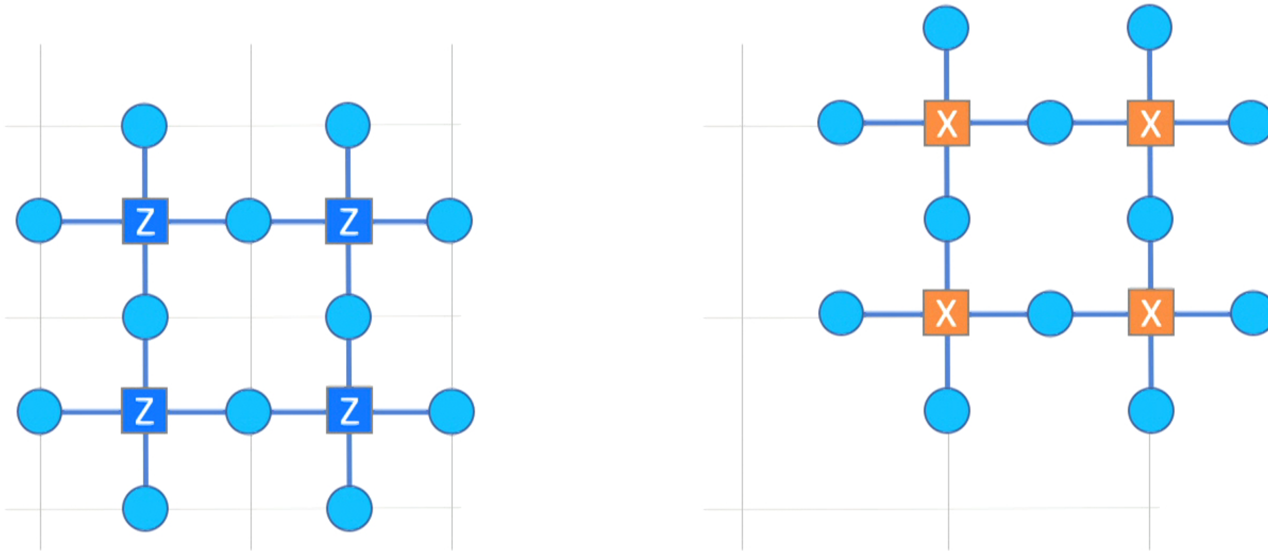
Tillich and Zemor, 2014

logical qubits grows linearly with # physical qubits

Low overhead for error correction

Can we use belief propagation to decode them?

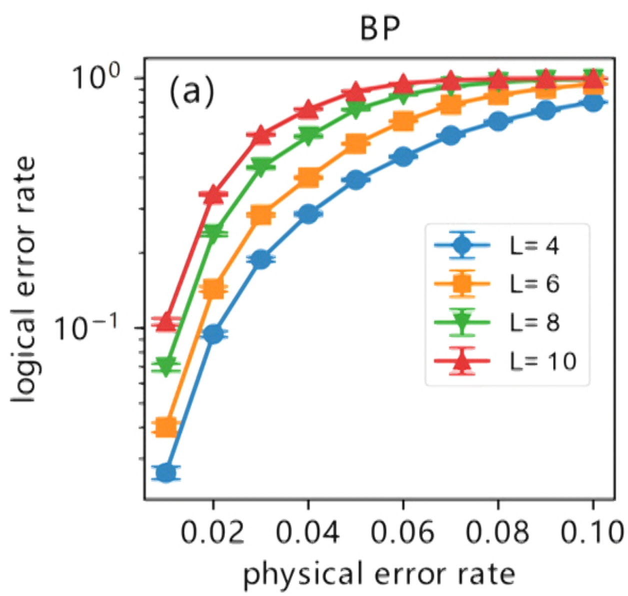
Toric code



Kitaev, 1997

$2L^2$ physical qubits, 2 logical qubits.

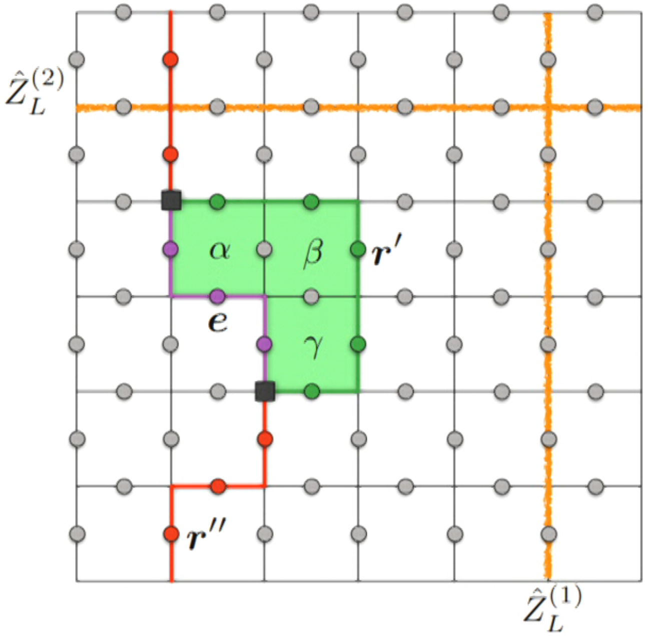
Performance of BP for toric code

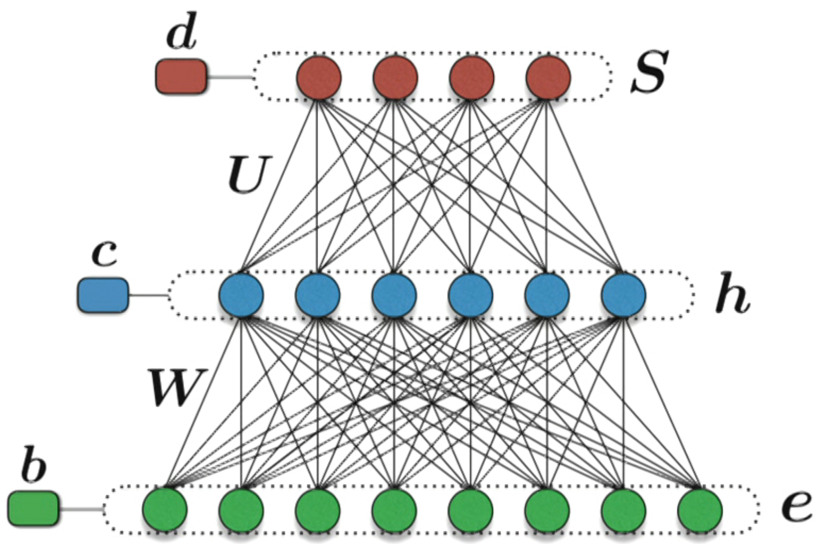


- **Larger** code size leads to **worse** performance.
- **Quantum degeneracy** leads to the bad performance. (Poulin and Chung, 2008)
- Heuristics for improvement. (Babar *et al,* 2015)

Deep learning approach.

Machine learning for quantum design: Neural decoders

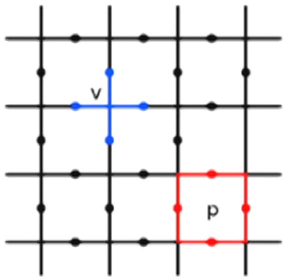




Torlai and Melko, PRL (2017)

Neural decoders for quantum codes: Overview

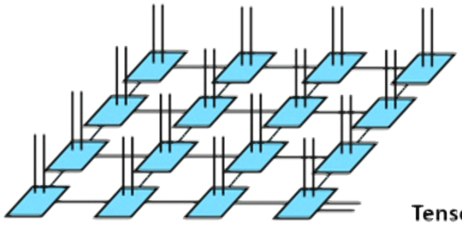
For topological codes **on regular lattices**



Surface code

Torlai and Melko, PRL (2017)
Krastanov and Jiang, Sci. Rep. (2017)
Breuckmann and Xiaotong, Quantum (2017)
Baireuther *et al.*, Quantum (2018)
Varsamopoulos *et al.*, Quantum Sci. Technol. (2018)
Maskara, Kubica, and Jochym-O'Connor, arXiv (2018)
Chamberland and Ronagh, Quantum Sci. Technol. (2018)
Davaasuren *et al.*, arXiv (2018)
Baireuther *et al.*, arXiv (2018)
Xiaotong Ni, arXiv (2018)
Sweke *et al.*, arXiv (2018)
...

Efficient & near-optimal human-designed decoders

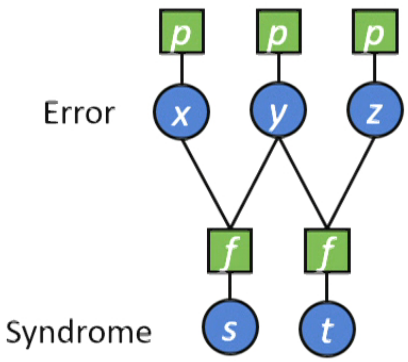


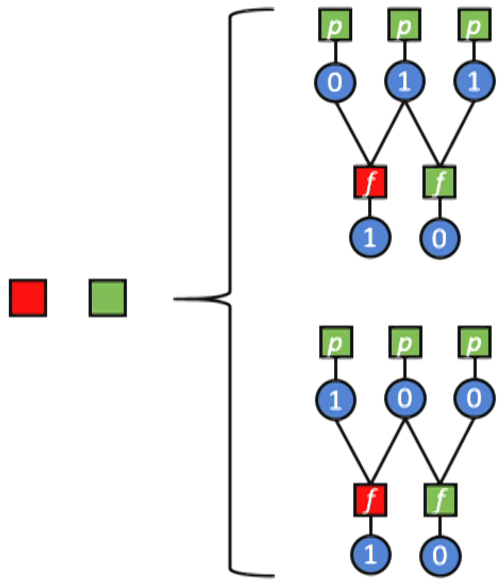
Tensor-net decoder

Bravyi, Suchara, and Vargo, PRA (2014)
Darmawan and Poulin, PRL (2017)
Darmawan and Poulin, PRE (2018)
...

Decoding quantum LDPC codes **on irregular graphs** remains hard, which calls for neural decoders for quantum LDPC codes.

Factor graph





$P(011|10) = p_{err}{}^2(1 - p_{err})$

$P(100|10) \propto p_{err}(1 - p_{err})^2$

- $p(0) = 1 - p_{err}$
- $p(1) = p_{err}$
- $f(x, y|s) = 1[x \oplus y = s]$
- $P(x, y, z|s, t) \propto p(x)p(y)p(z)f(x, y|s)f(y, z|t)$

Statistical inference in graphs

- Machine learning
- Statistical physics
- Information theory

Statistical physics perspective

$$P\,(x, y, z|s, t\,) = Z^{-1}e^{-E(x,y,z|s,t)/T}$$

- The lower the p_{err}, the lower the temperature T.
- Decoding: Finding the ground state, which is NP-complete. (Berlekamp *et al.,* 1978)
- If the ground state is ***not degenerate***, thermal average in low T recovers it:

$$x_{GS} \approx \langle x \rangle_T.$$

Belief propagation (BP): Near-optimal classical decoder

$$l_v = \log \frac{1-p_{err}}{p_{err}} \;\&\; \{s\}$$

BP in

$$\mu_{v\to c}^{(t+1)} = l_v + \sum_{c\prime \in \mathcal{N}(v)\backslash c} \mu_{c\prime \to v}^{(t)}$$

$$\tanh \frac{\mu_{c\to v}^{(t+1)}}{2} = (-1)^{s_c} \prod_{v\prime \in \mathcal{N}(c)\backslash v} \tanh \frac{\mu_{v\prime \to c}^{(t)}}{2}$$

$$\mu_v = l_v + \sum_{c \in \mathcal{N}(v)} \mu_{c\to v}^{(T)}$$

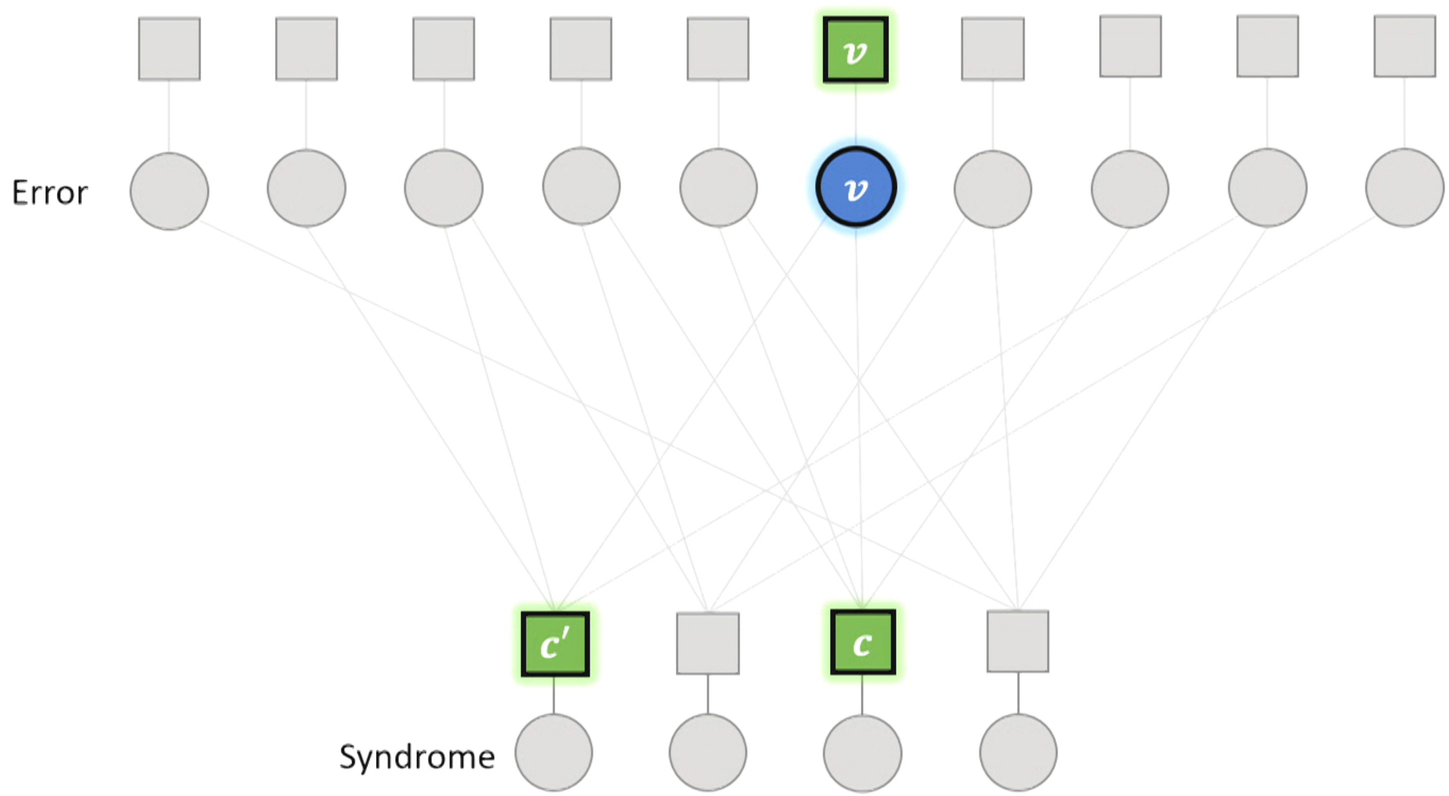
BP out

$$\mu_v = \log \frac{P(e_v = 0|\{s\})}{P(e_v = 1|\{s\})}$$

Sigmoid

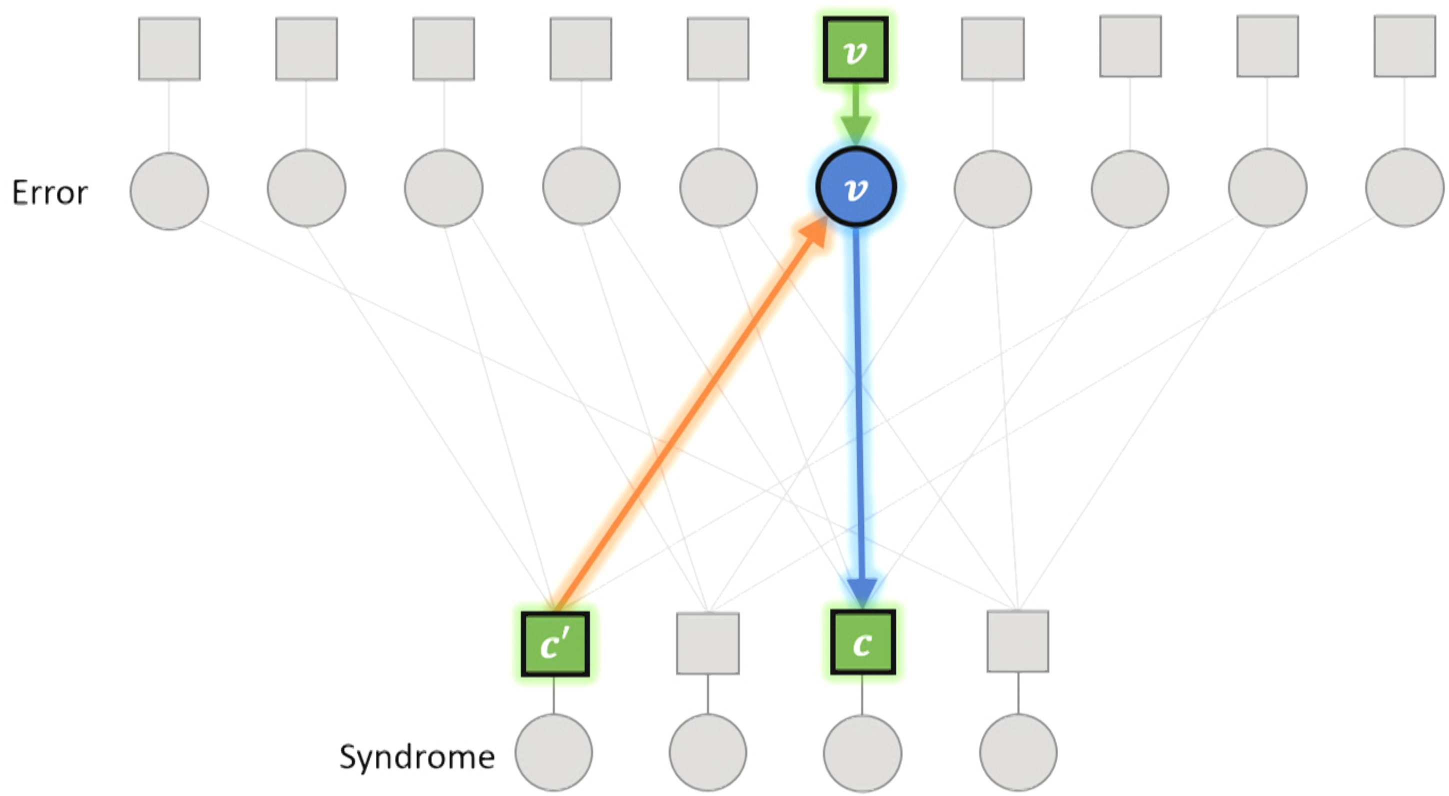
$$P(e_v = 1|\{s\}) = \sigma(-\mu_v) > 0.5?$$

Pearl, 1982 (information theory)
Mezard, Parisi, and Virasoro, 1986 (physics)

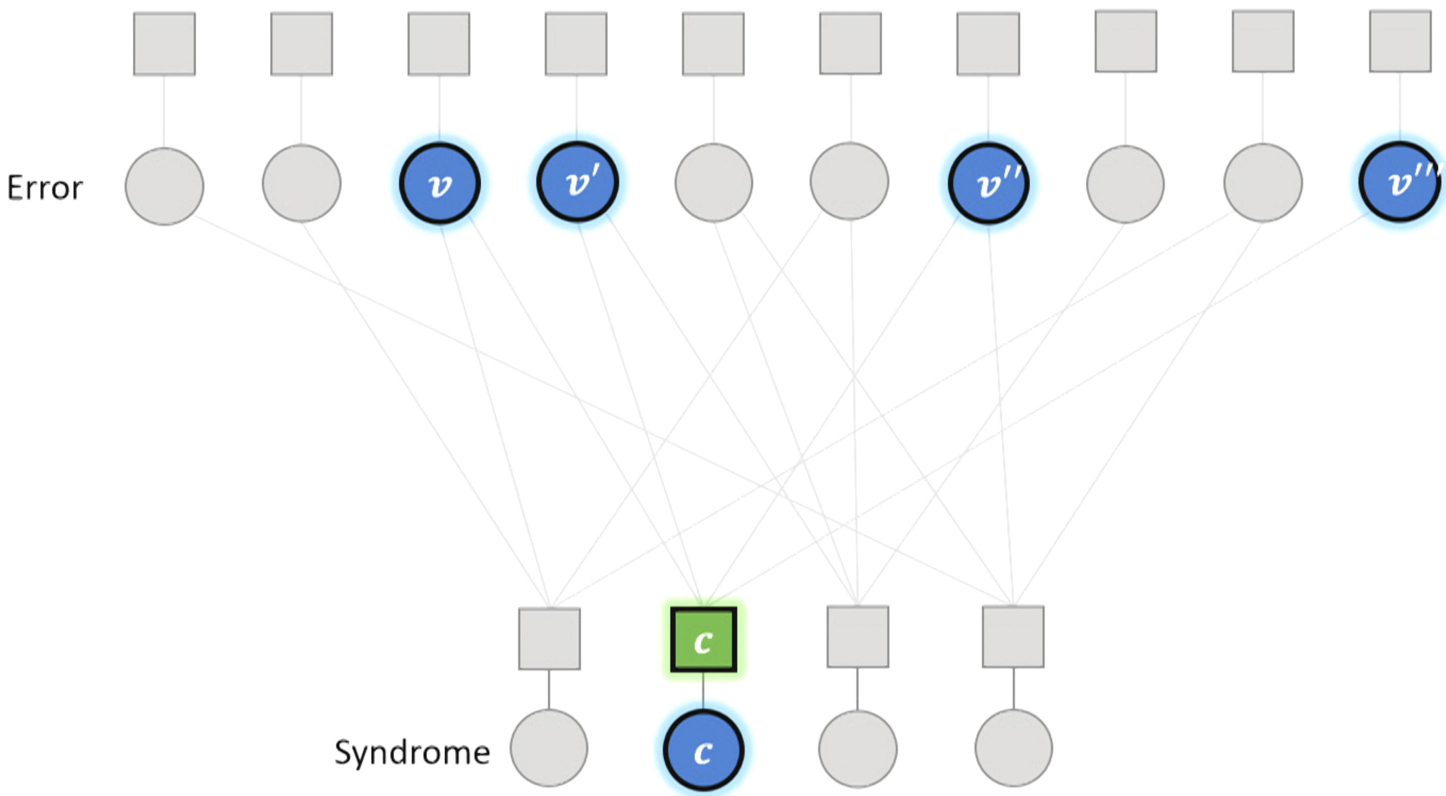


v
Error
v
c'
c
Syndrome

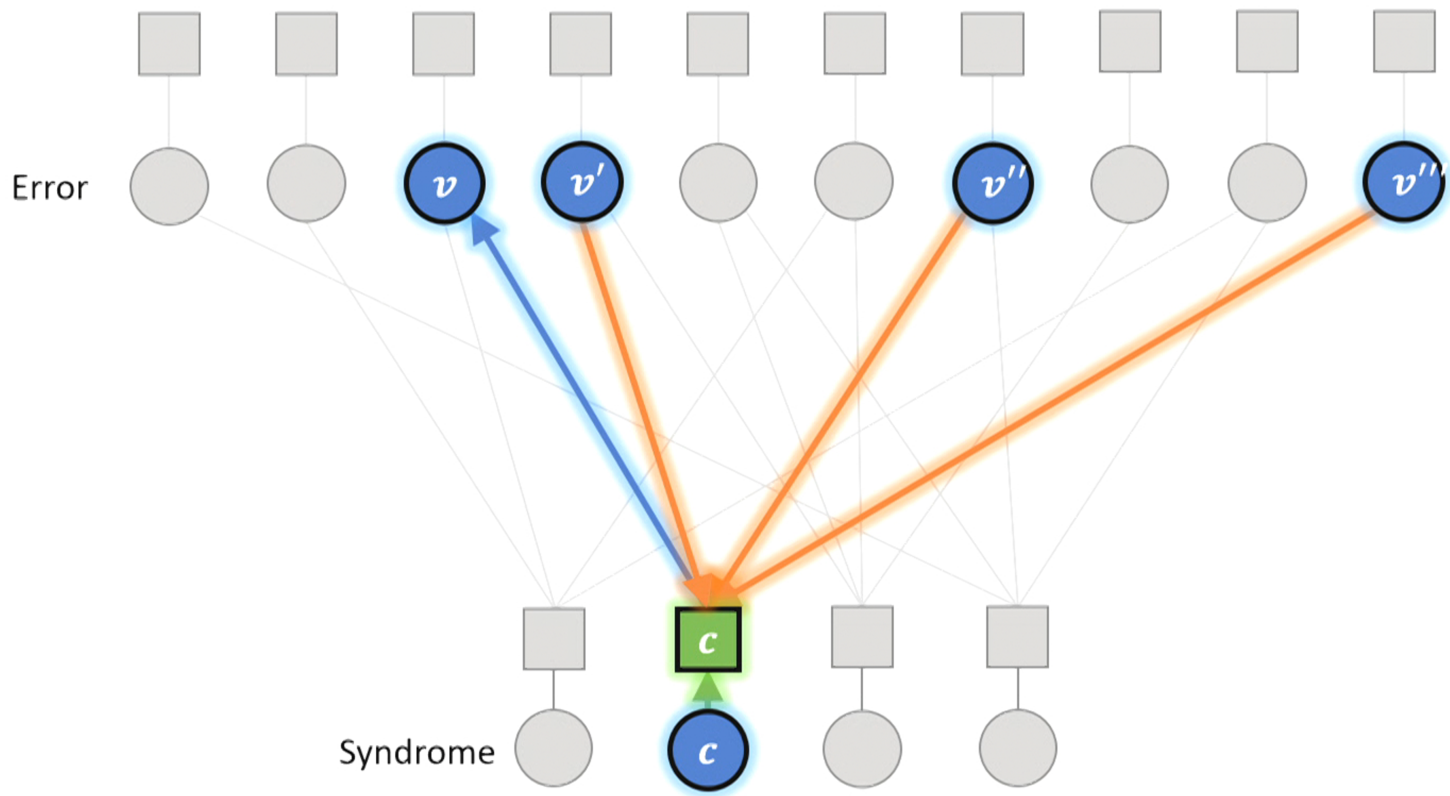
$$\mu_{v \to c}^{(t+1)} = l_v + \sum_{c' \in \mathcal{N}(v) \backslash c} \mu_{c' \to v}^{(t)}$$



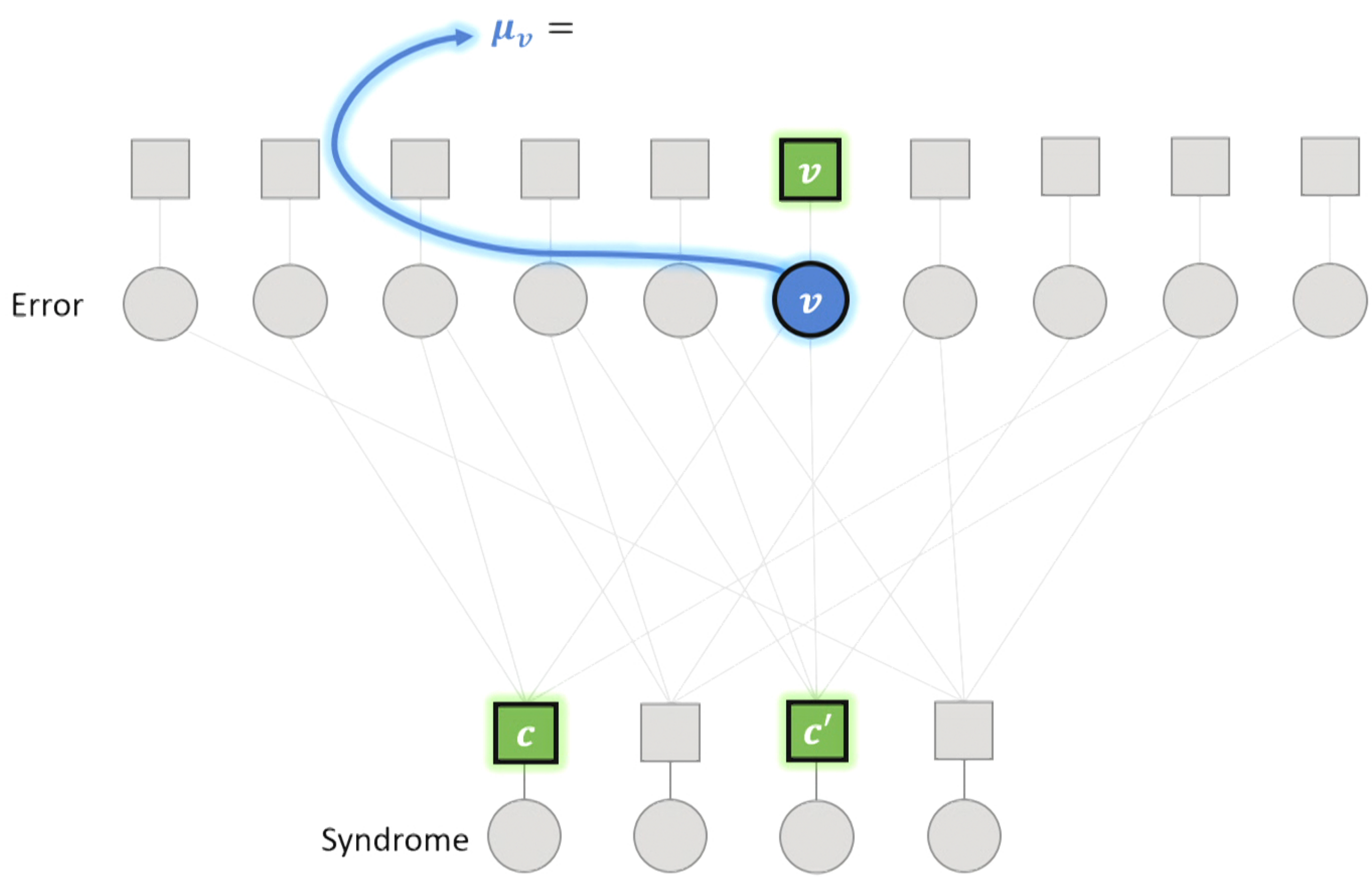
16


Error
v
v'
v''
v'''
c
Syndrome
c

$$\tanh \frac{\mu_{c\to v}^{(t+1)}}{2} = (-1)^{s_c} \prod_{v'\in\mathcal{N}(c)\setminus v} \tanh \frac{\mu_{v'\to c}^{(t)}}{2}$$

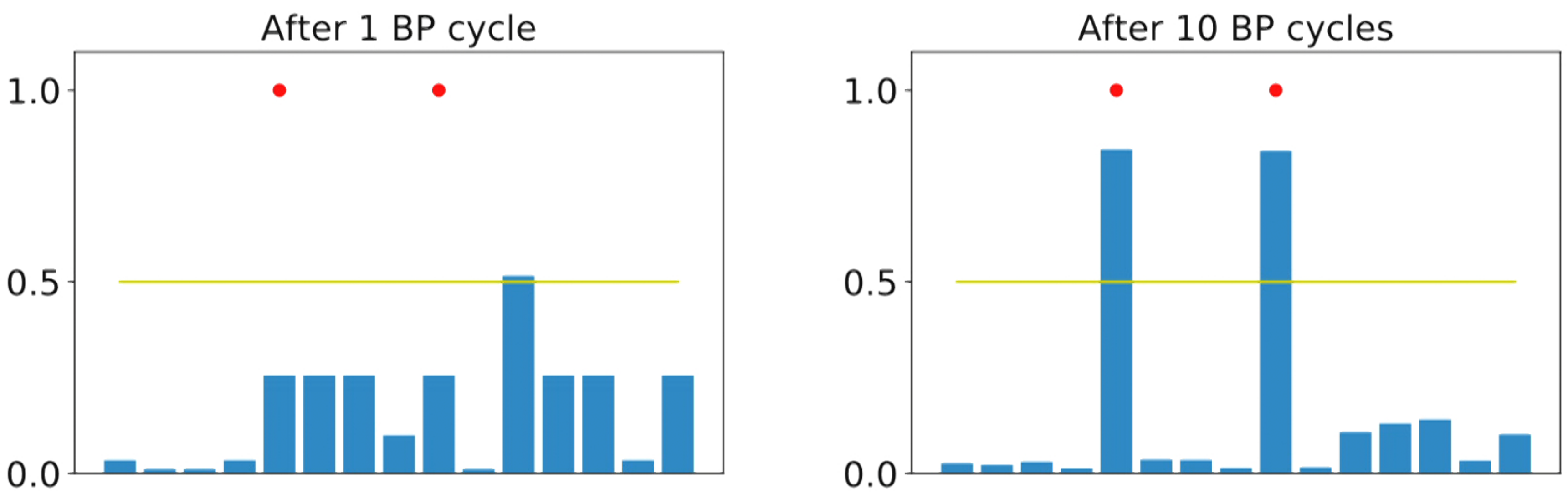


17

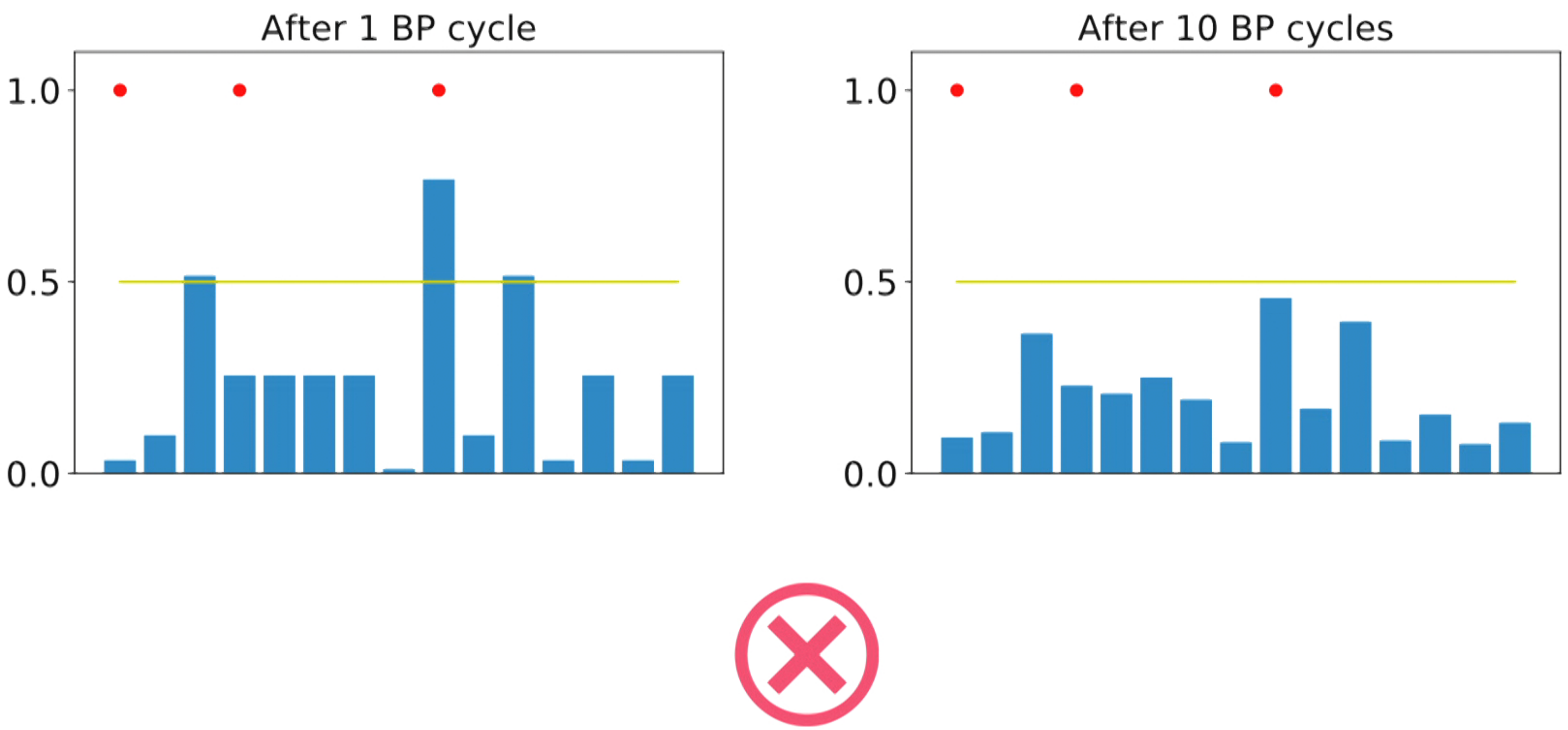


18

Successful decoding



Failed decoding



20

Neural belief propagation

Belief-propagation equations

$$\mu_{v\to c}^{(t+1)} = l_v + \sum_{c'\in\mathcal{N}(v)\setminus c} \mu_{c'\to v}^{(t)}$$

$$\tanh\frac{\mu_{c\to v}^{(t+1)}}{2} = (-1)^{s_c} \prod_{v'\in\mathcal{N}(c)\setminus v} \tanh\frac{\mu_{v'\to c}^{(t)}}{2}$$

$$\mu_v = l_v + \sum_{c\in\mathcal{N}(v)} \mu_{c\to v}^{(T)}$$

Pearl, 1982; Mezard, Parisi, and Virasoro, 1986

Weights, biases, and nonlinearity: **neural network!**

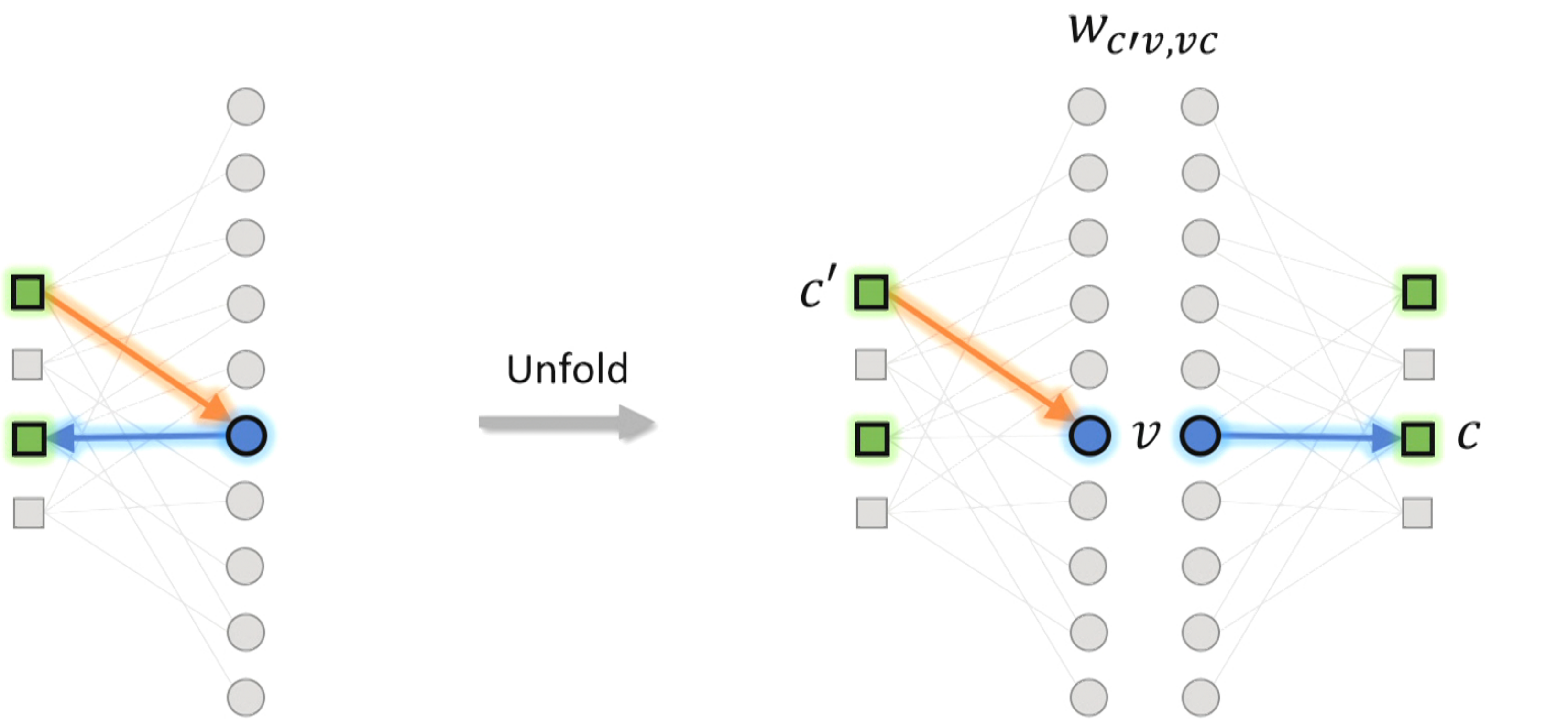
$$\mu_{v\to c}^{(t+1)} = l_v \boldsymbol{b}_v^{(t)} + \sum_{c'\in\mathcal{N}(v)\setminus c} \mu_{c'\to v}^{(t)} \boldsymbol{w}_{c'v,vc}^{(t)}$$

$$\mathbf{log\,tanh}\frac{\mu_{c\to v}^{(t+1)}}{2} = i\pi s_c \boldsymbol{b}_c^{(t)} + \sum_{v'\in\mathcal{N}(c)\setminus v} \mathbf{log\,tanh}\frac{\mu_{v'\to c}^{(t)}}{2} \boldsymbol{w}_{v'c,cv}^{(t)}$$

$$\mu_v = l_v \boldsymbol{b}_v^{(T)} + \sum_{c\in\mathcal{N}(v)} \mu_{c\to v}^{(T)} \boldsymbol{w}_{cv,v}^{(T)}$$

Nachmani, Be'ery, and Burshtein, 2016; Liu and Poulin, 2019

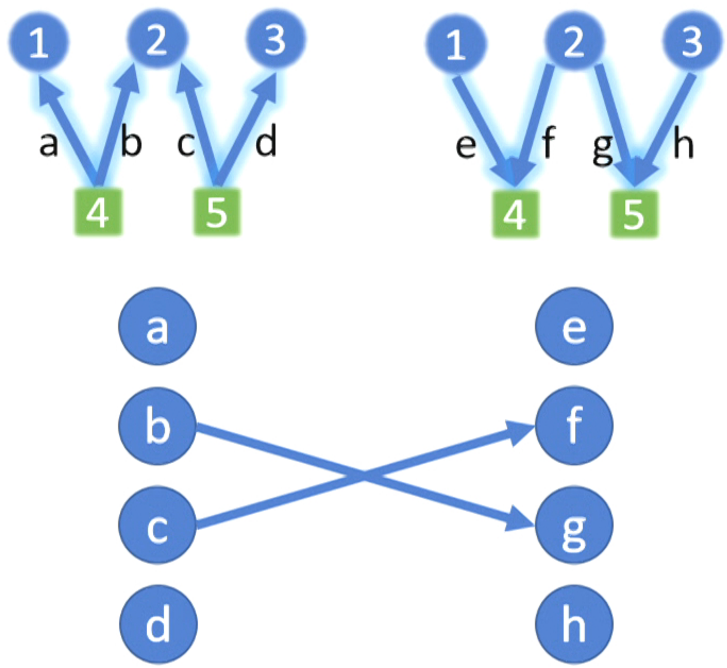
Network architecture: Unfolded message passing



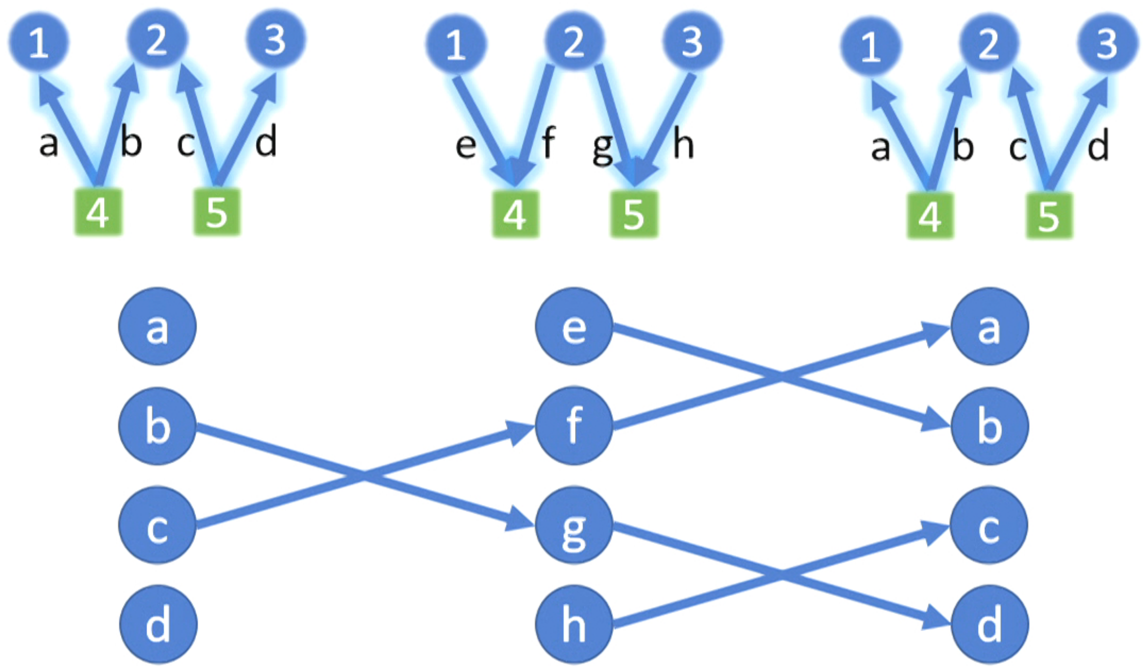
BP: Unweighted, in-place message passing

NBP: Weighted, unfolded message passing

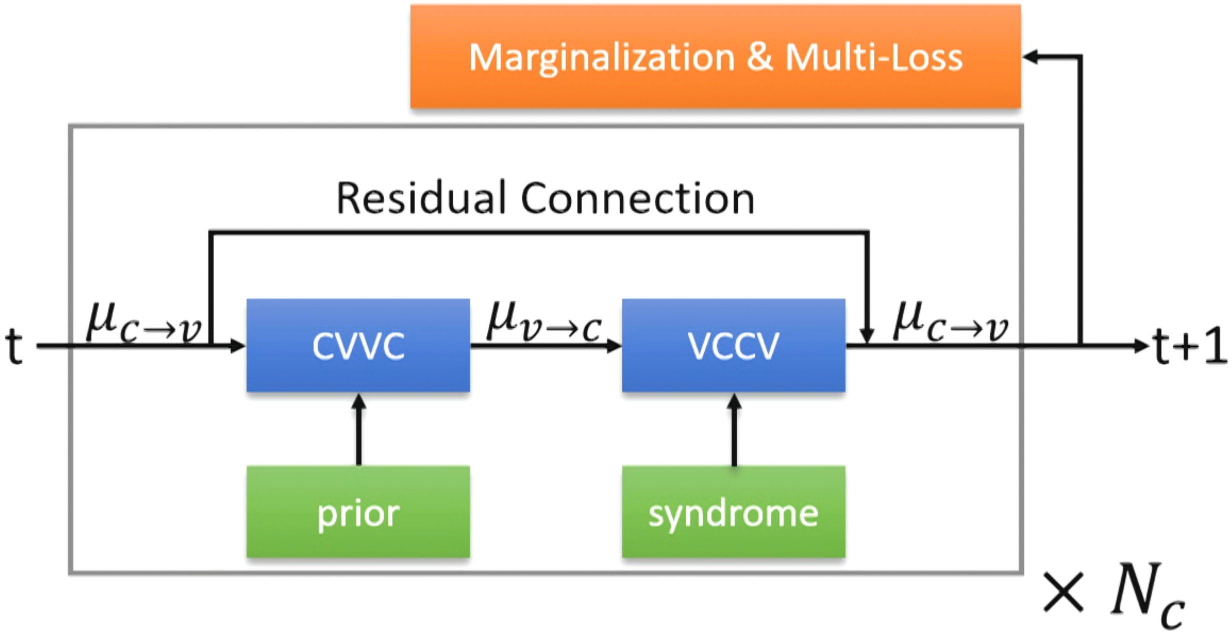
Customized neural network



Customized neural network



Customized neural network



Classical vs. quantum codes

	Classical Codes	Quantum Codes
Error	$\vec{e} = \{e_1, e_2, e_3\}$	$\hat{e} = X_1^{e_1} X_2^{e_2} X_3^{e_3} Z_1^{e_4} Z_2^{e_5} Z_3^{e_6}$
Check	$\vec{H}_1 = \{1,1,0\}, \vec{H}_2 = \{0,1,1\}.$	$\hat{S}_1 = ZZI, \hat{S}_2 = IZZ.$
Syndrome	$s_i = \vec{H}_i \cdot \vec{e}$	$s_i = [\hat{S}_i, \hat{e}] = \vec{S}_i \cdot M\vec{e}$
Successful decoding	$\vec{e}_{\text{dec.}} = \vec{e}$	$[\hat{S}_i, \hat{e}\hat{e}_{\text{dec.}}] = [\hat{L}_i, \hat{e}\hat{e}_{\text{dec.}}] = 0$
Constraint	None	All checks commute

Loss function for quantum error correction

$$\mathcal{L} = \sum_{i=1}^{N_S} f\left(\vec{e}_{\text{tot.}} \cdot M\vec{S}_i\right) + \sum_{i=1}^{N_L} f\left(\vec{e}_{\text{tot.}} \cdot M\vec{L}_i\right)$$

Training: $\Delta\theta = -\alpha \frac{\partial \mathcal{L}}{\partial \theta}, \ \theta = \{w, b\}$

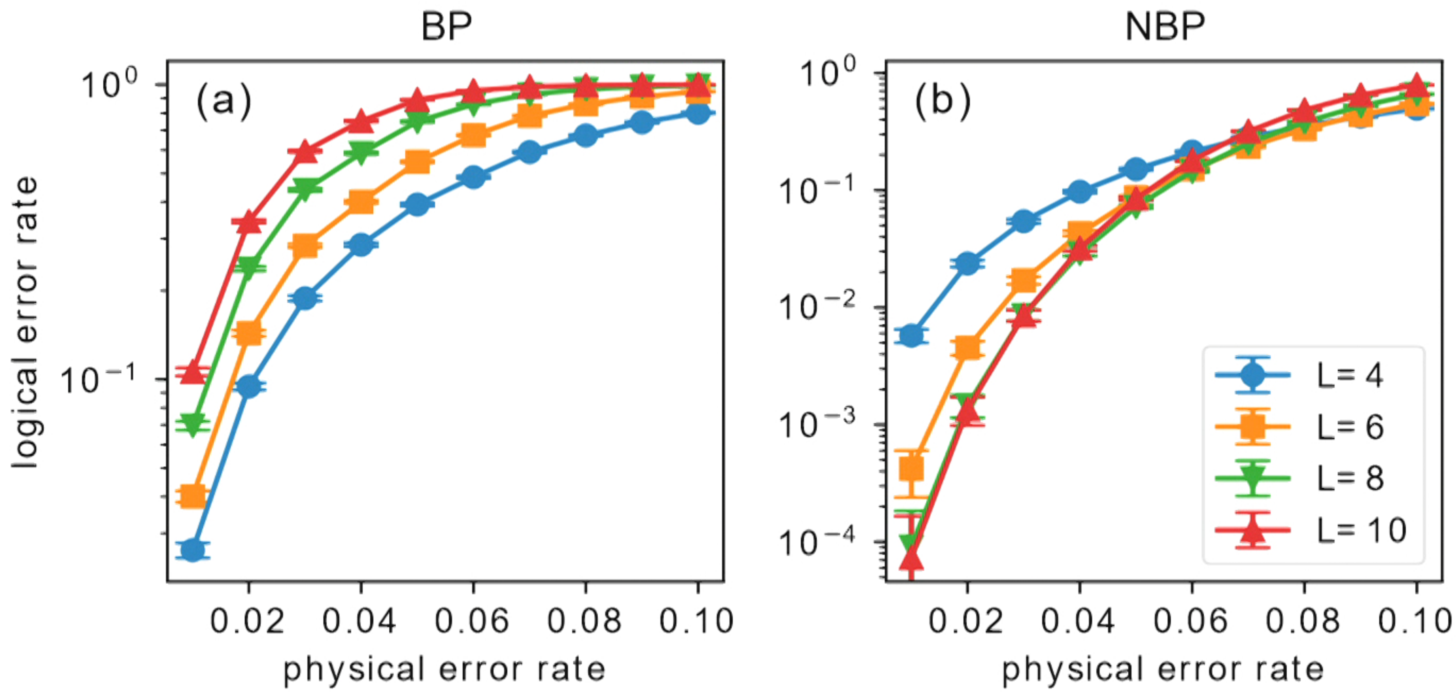
Symbol	Meaning
$\vec{e}_{dec.} = \sigma(-\vec{\mu})$	Decoder output
$\vec{e}_{\text{tot.}} = \vec{e}_{\text{dec.}} + \vec{e}$	Total error
$f(x) = \lvert\sin(\pi x/2)\rvert$	Smooth mod 2

Quantum codes:

- Errors are equivalent up to the **stabilizer group.**
- Stabilizer group elements are orthogonal to all $\left\{\vec{S}, \vec{L}\right\}$.
- When $\vec{e}_{dec.}$ and $\vec{e}$ are equivalent, $\mathcal{L} = 0$.
- Minimizing $\mathcal{L}$ leads to better quantum decoder.

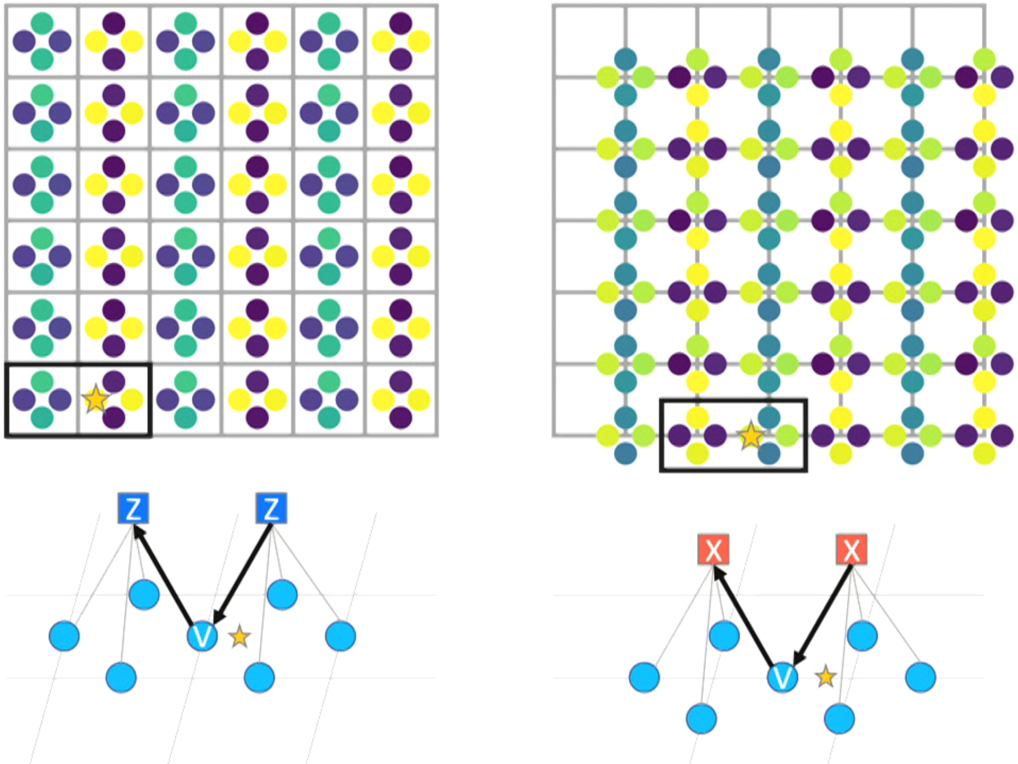
YHL & D. Poulin, Phys. Rev. Lett. **122,** 200501 (2019)

Neural BP decoder for toric codes



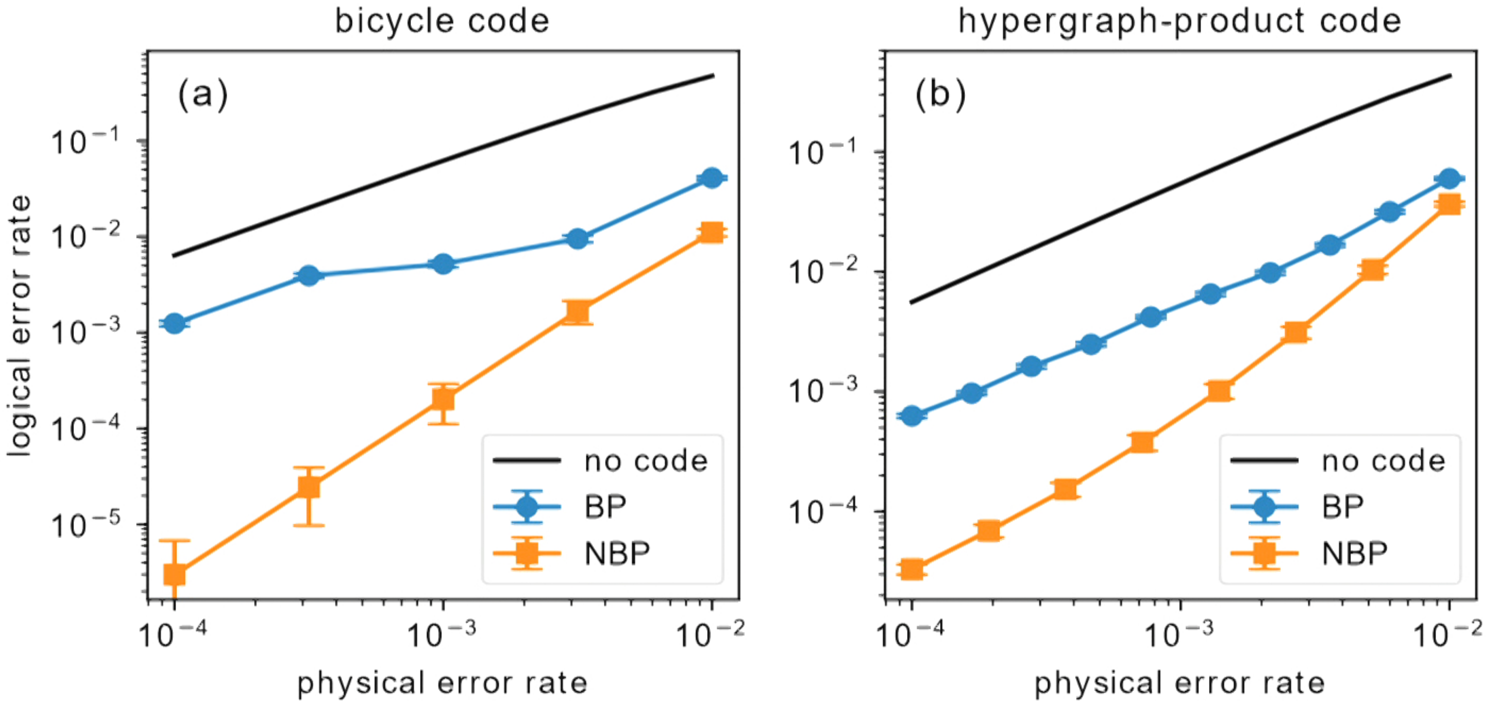
YHL & D. Poulin, Phys. Rev. Lett. **122,** 200501 (2019)

Visualizing learned weights for toric code



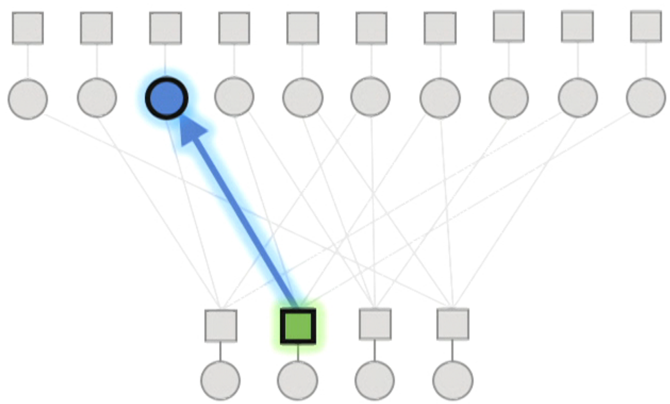
YHL & D. Poulin, Phys. Rev. Lett. **122,** 200501 (2019)

Neural BP decoder for more challenging codes

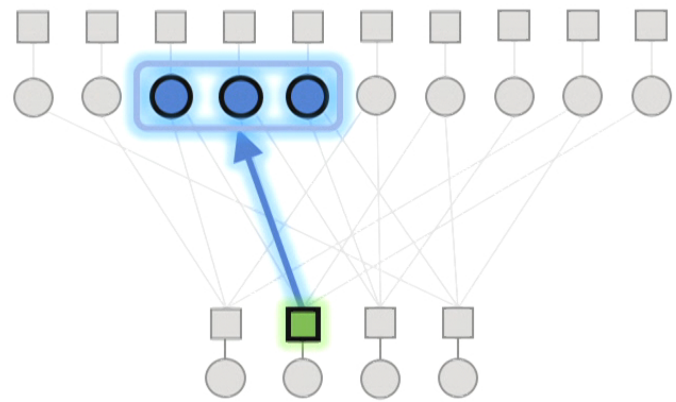


YHL & D. Poulin, Phys. Rev. Lett. **122,** 200501 (2019)

Ongoing work: Cluster neural BP



Bethe single site



Kikuchi cluster variation

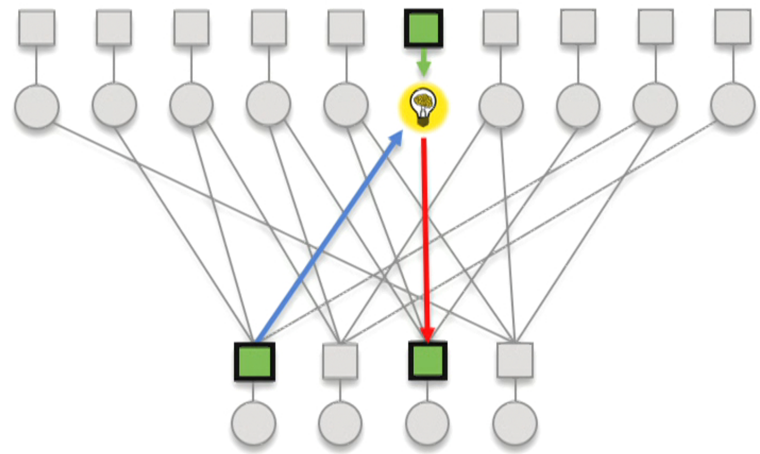
Yedidia, Freeman, and Weiss (2001)

Advantages:

- More accurate inference
- Multiple-point correlations
- More insights into statistical physics
- Also trainable

Summary

- **Belief propagation,** widely used in many research areas, has a **neural-network** representation.
- **Training** greatly improves BP for decoding quantum error-correcting codes.
- Training BP with **"quantum data"** might solve other intricate problems in quantum many-body physics.



YHL & D. Poulin, Phys. Rev. Lett. **122,** 200501 (2019)

Acknowledgement: Pavithran Iyer, Anirudh Krishna, Colin Trout, Liang Jiang, and Stefan Krastanov, Xin Li.