QEC & ML

Advantages of versatile neural network decoding for topological codes

Aleksander Kubica

work w/ N. Maskara and T. Jochym-O'Connor

arXiv: 1802.08680

TOWARD QUANTUM COMPUTATION

- Promises of quantum computation:
  — simulations of many-body systems,
  — quantum algorithms, …

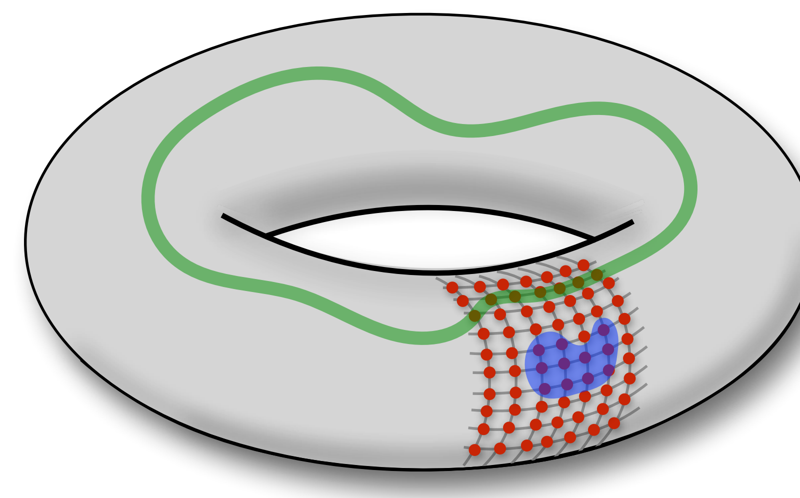
- So let's build a quantum computer!
- To operate quantum computer we need to reliably store & process quantum information.
- Interactions with environment causes errors. Use error-correcting codes!
- **Threshold theorem**: scalable quantum computation possible given sufficiently weak and uncorrelated noise *[KLZ98,ABO98,AGP06,…]*!

Knill&Laflamme&Żurek'98; Aharonov&Ben-Or'98; Aliferis&Gottesman&Preskill'06

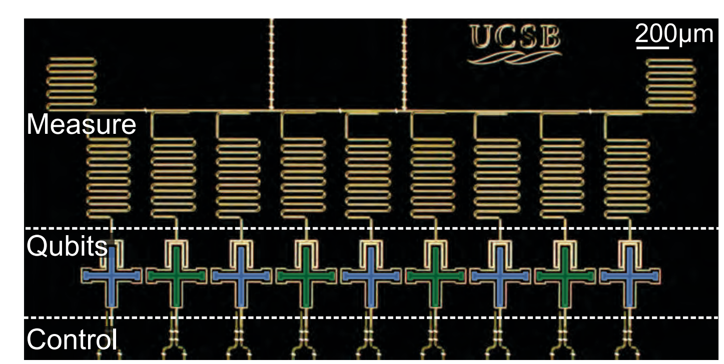
QUANTUM ERROR CORRECTING CODES

- Protect information by encoding into a quantum code *[S95]:*

$$|\psi\rangle \in \mathbb{C}^2 \xrightarrow{\text{encode}} |\overline{\psi}\rangle \in (\mathbb{C}^2)^{\otimes n}$$



- **Topological stabilizer codes** *[DKLP03]*: local stabilizers, logical information encoded non-locally.
- Desired properties: fault-tolerant logical gates, efficient decoders, high error-correction thresholds.



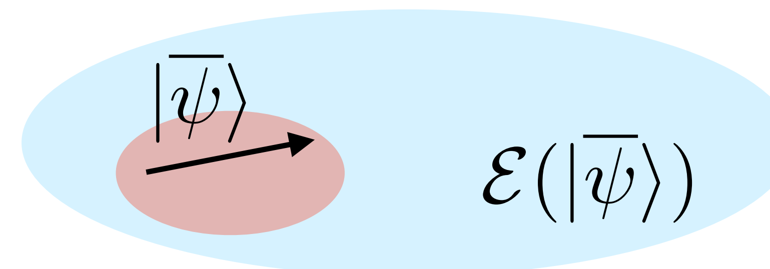
Kelly et al., Nature 519, (2015)

- Locality comes with a price *[BPT09, JKY18, …]* — no-go theorems for storage and computation!
- **Side remark**: topological codes as toy models of (exotic) quantum phases of matter, e.g. 3D fractons *[H11]*.

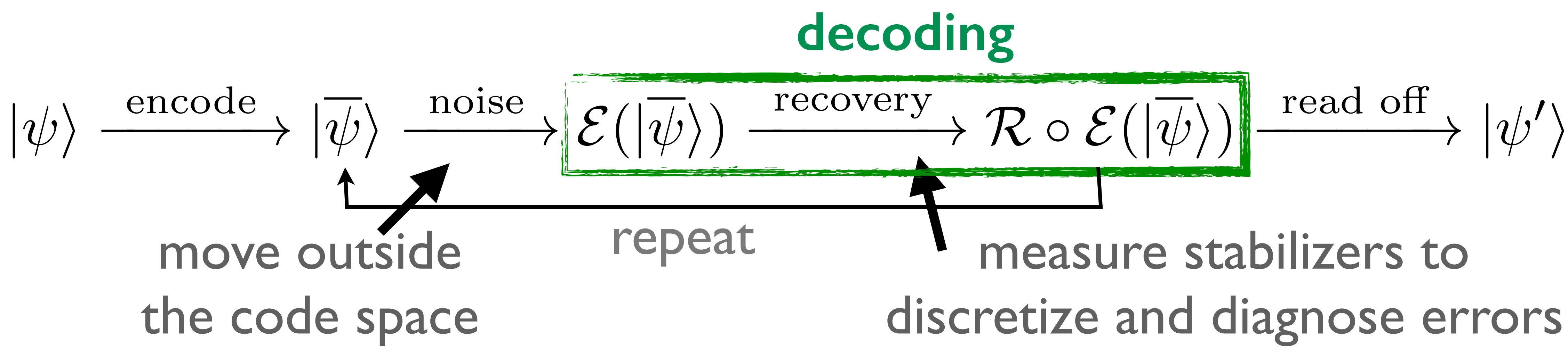
Shor'95; Dennis et al.'03; Bravyi&Poulin&Terhal'09; Jochym&K.&Yoder'18; Haah'11

DECODING PROBLEM FOR STABILIZER CODES

- **stabilizer codes** *[G96]*: commuting Pauli operators
  code space = (+1)-eigenspace of stabilizers



- **quantum error-correction game**



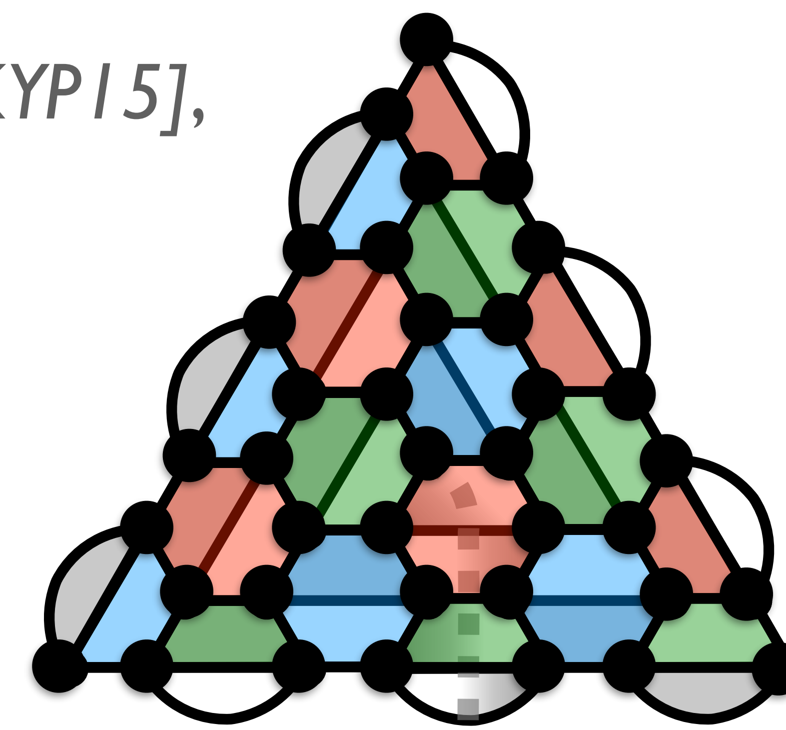
- **decoder:** algorithm to find a Pauli recovery from stabilizer measurements

- **successful decoding** iff recovery returns the state to code space <u>**AND**</u> error + recovery do not implement a non-trivial logical operator

Gottesman'96

CHALLENGES IN DESIGNING DECODERS

- Decoding of generic stabilizer codes is computationally hard *[HL11,IP13]*.
- Dominant **sources of errors** not known/depend heavily on the device.
- Many results on decoders, usually designed and analyzed for **simplistic noise models**, e.g. the bit-flip. Correlations: X/Z or spatial *[BN17]*?
- Codes may be related, e.g. color and toric codes *[KYP15]*, but decoders **difficult to adapt**!
- Desirable decoding methods should:
  — minimize human input,
  — be easily adaptable to different noise/code,
  — be efficient and have good performance.



Hsieh&LeGall'11; Iyer&Poulin'13; Brown&Nickerson'17; K.&Yoshida&Pastawski'15

OUTLINE

- **Goal:** explore adaptability of (vanilla) neural-network decoding for various codes and (correlated) noise models

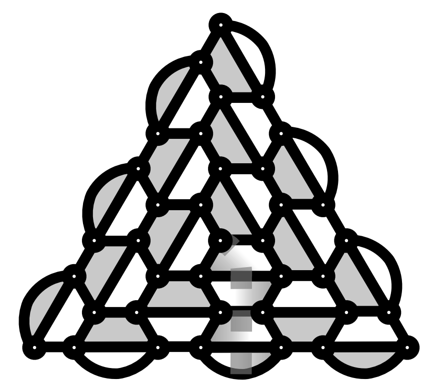
- **Previously** *[TM16,…]:* surface code, small distance (*d*=*3*–*7*), simple noise

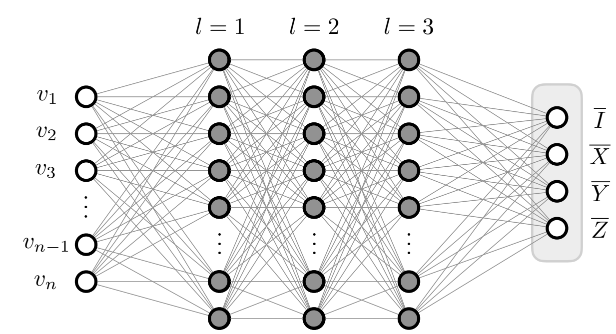
1. 2D toric code with a twist

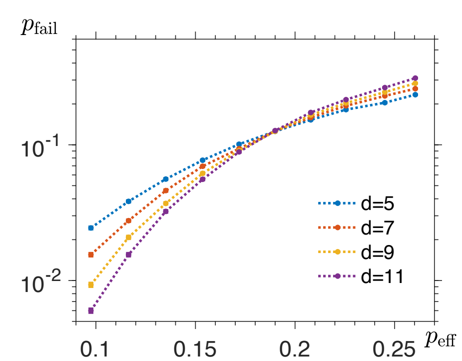
2. neural-network decoding

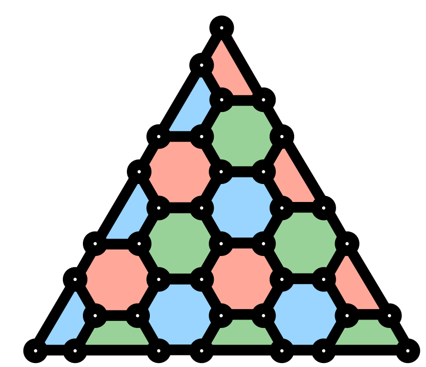
3. benchmarking performance

4. 2D triangular color code







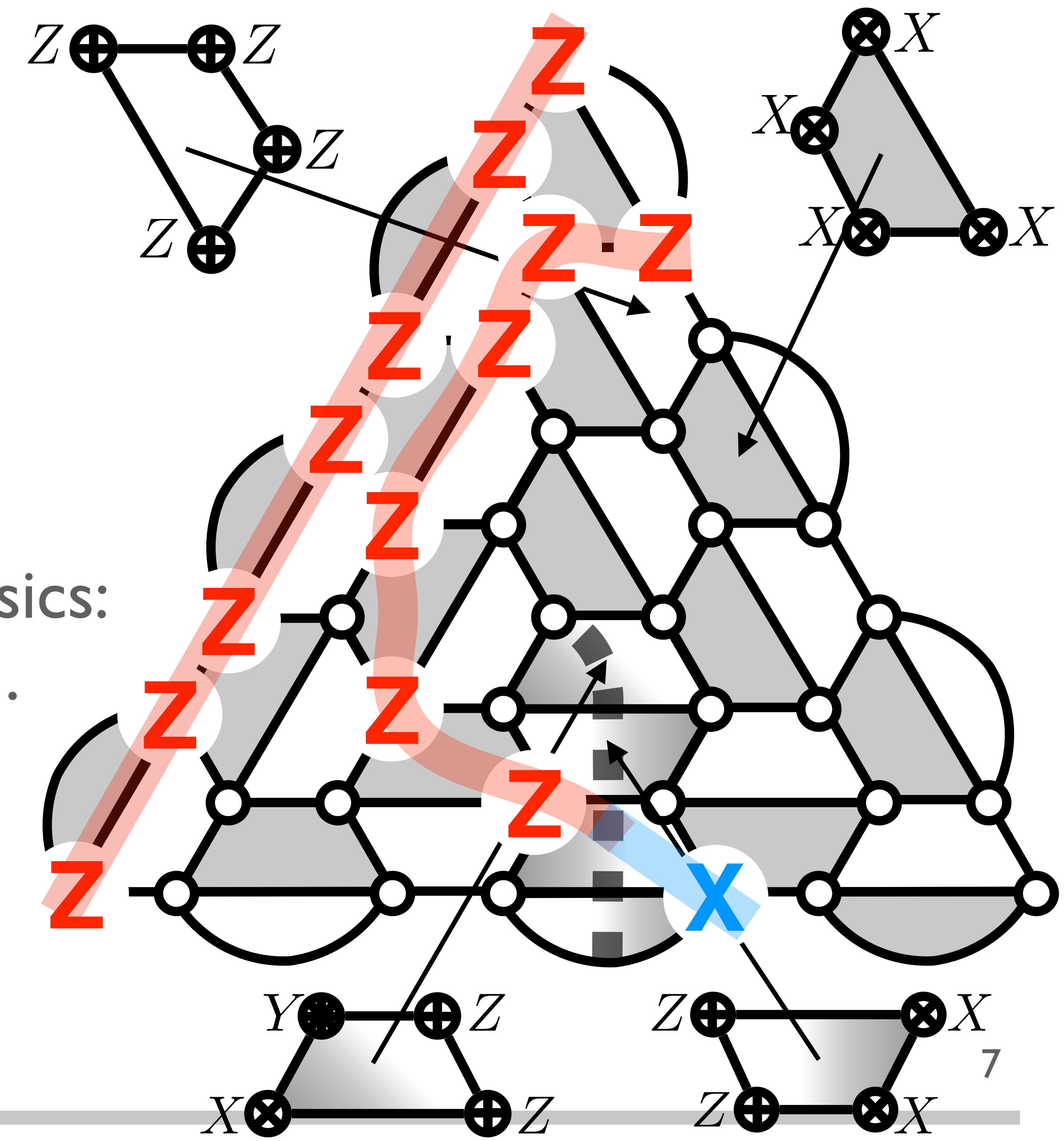


Torlai&Melko'16; Baireuther et al.'17; Varsamopoulos et al.'18;…

arXiv: **1802.08680**

2D TORIC CODE WITH A TWIST

- many versions *[K97,…]*:
  lattices, boundaries, twists, …
- 2D toric code with a twist *[YK16]*:
  — 2-colorable faces
  — 4-valent vertices
- stabilizers = X-/Z-faces, mixed faces
- simple model capturing interesting physics:
  anyons, condensation on boundaries, …
- high error-correction threshold,
  local stabilizers of weight ≤ 4
- logical Pauli operators = 1D strings



Kitaev'97; Yoder&Kim'16

STABILIZER HAMILTONIAN AND EXCITATIONS

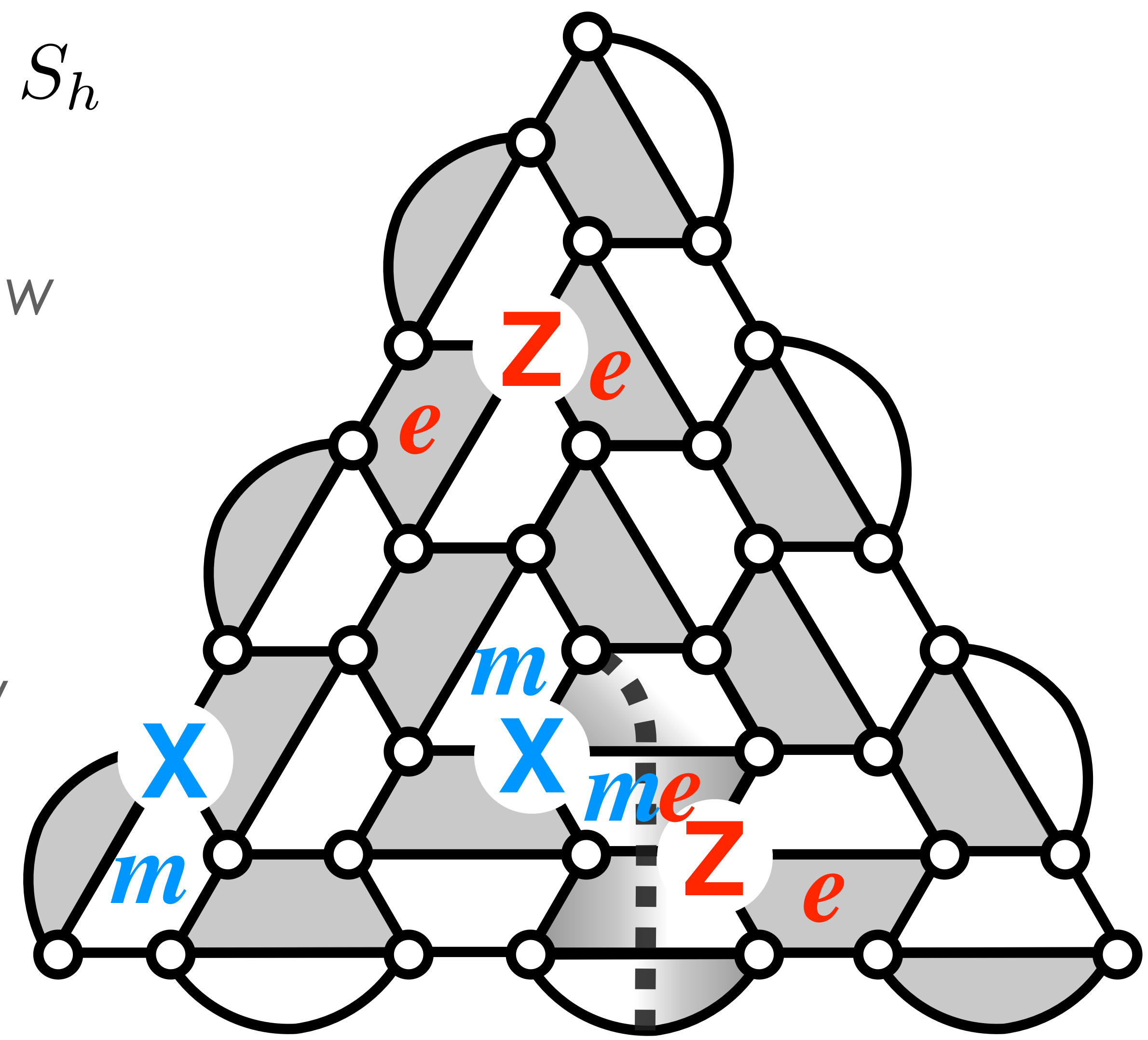
- find Hamiltonian whose ground space = code space

$$H_{TC} = -\sum_{f\in F_D} X_f - \sum_{g\in F_W} Z_g - \sum_{h\in F_M} S_h$$

- violated stabilizers = excitations e_D & m_W
- **bulk:** excitations created in pairs

$$e_D \times e_D = m_W \times m_W = 1$$

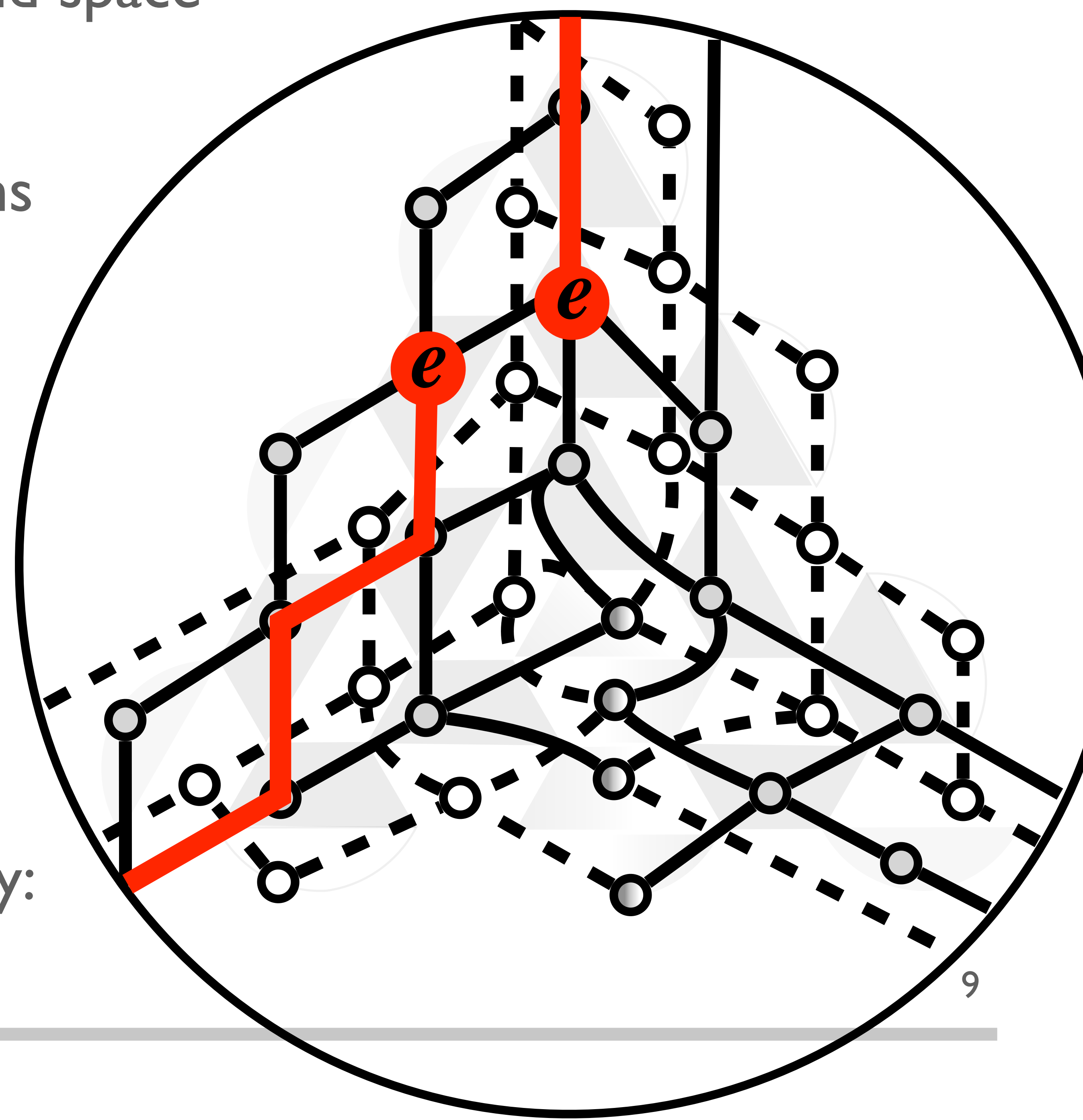
- **boundary:** can create a single e_D or m_W
- logical operator = create an excitation and move to the other boundary
- **defect line** *[B10]*: e_D and m_W are swapped



Bombin'10

HOW TO REMOVE EXCITATIONS

- **decoding** = returning to code/ground space
  = removing excitations
- always possible to remove excitations by pushing them to the boundary!
- **excitation graph:**
  vertices = excitations
  edges = local Pauli operators
- easy to construct (fusion rules) and use to find a removal operator
- do not need to find the error exactly: success iff up to a stabilizer!



DECODING AS CLASSIFICATION PROBLEM

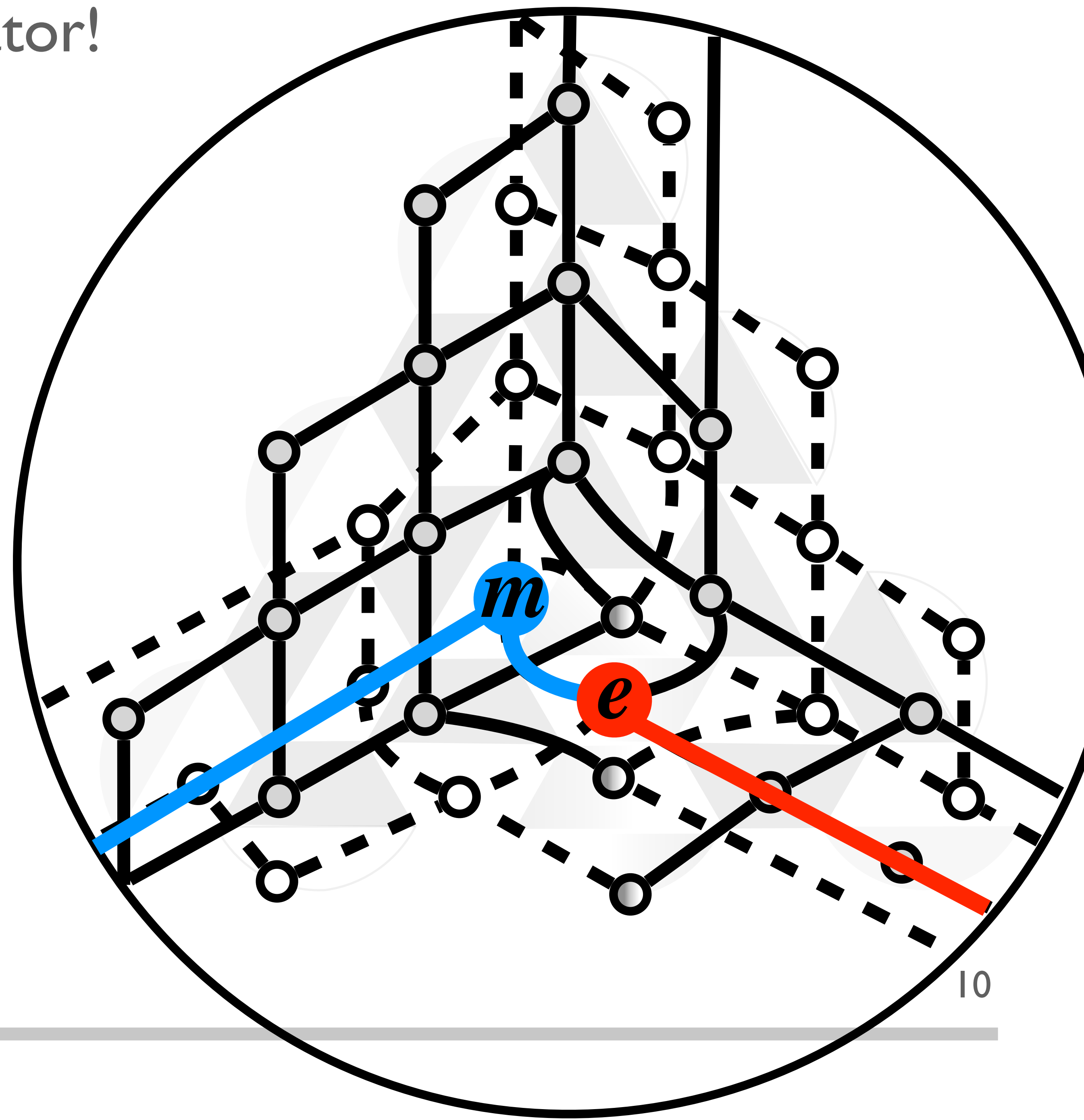
- Pushing excitations to boundary returns to code space but likely to introduce logical* operator!

- Pauli errors = Q (not known!)
  excitations/syndrome = U
  removal operator = R_U

- $R_U\, Q \sim L$ — if only we knew L…

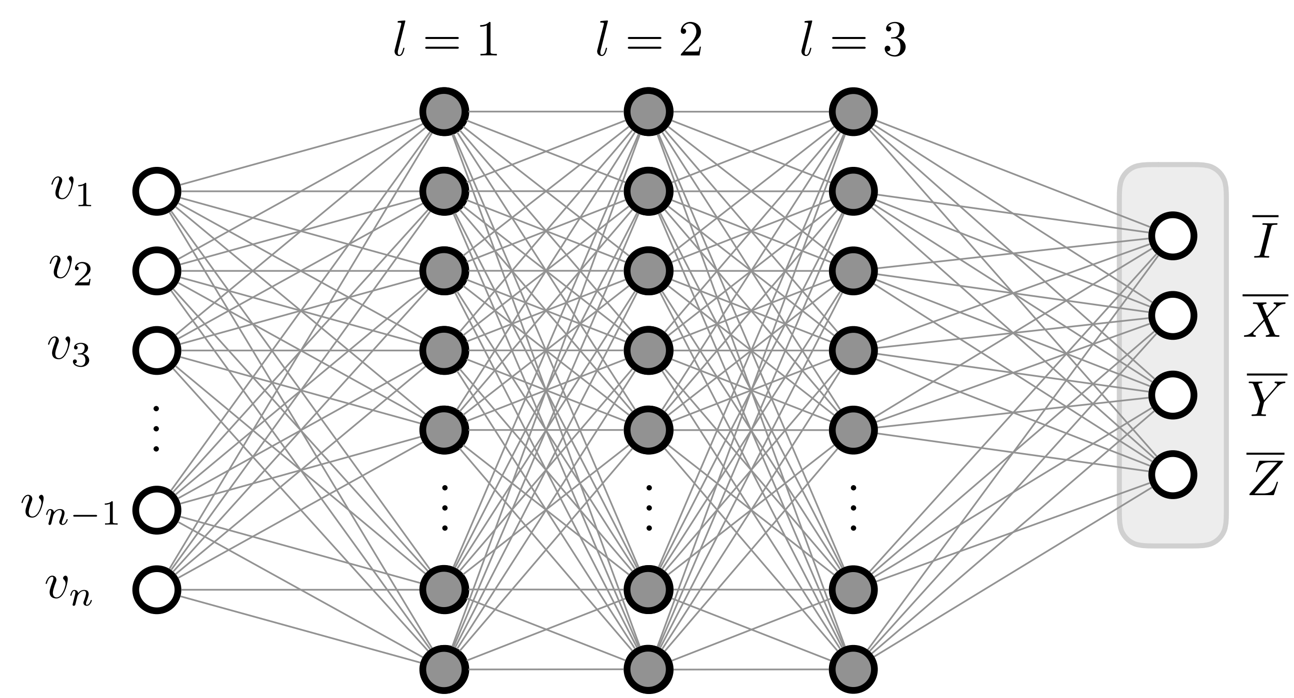
- This is a classification problem!
  (excitations U, logical L)

- Many errors Q w/ the same U!
  Find the most likely equivalence class of errors (labeled by $L \sim R_U\, Q$).



NEURAL-NETWORK DECODING

- Feedforward neural networks: layers, nodes, activation function.
- **Neural decoder:**
  (1) excitation removal: $U \longrightarrow R_U$
  (2) neural net to classify: $U \longrightarrow L$
  output = recovery $R_U L$



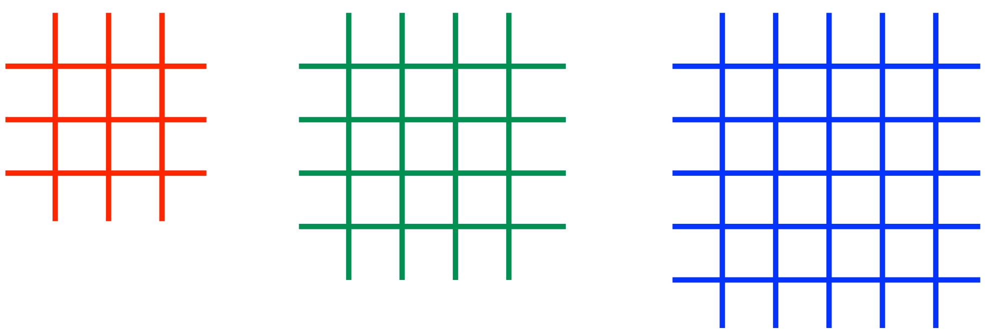
- Details of (1) excitation removal not important; usually easy to figure out.
- Training neural net = minimization problem (cross entropy) for specified **code**, **noise model**, **removal algorithm,** but can use different error rates!
- Standard neural net optimizations: Adam, mini-batch, He (for ReLU),…

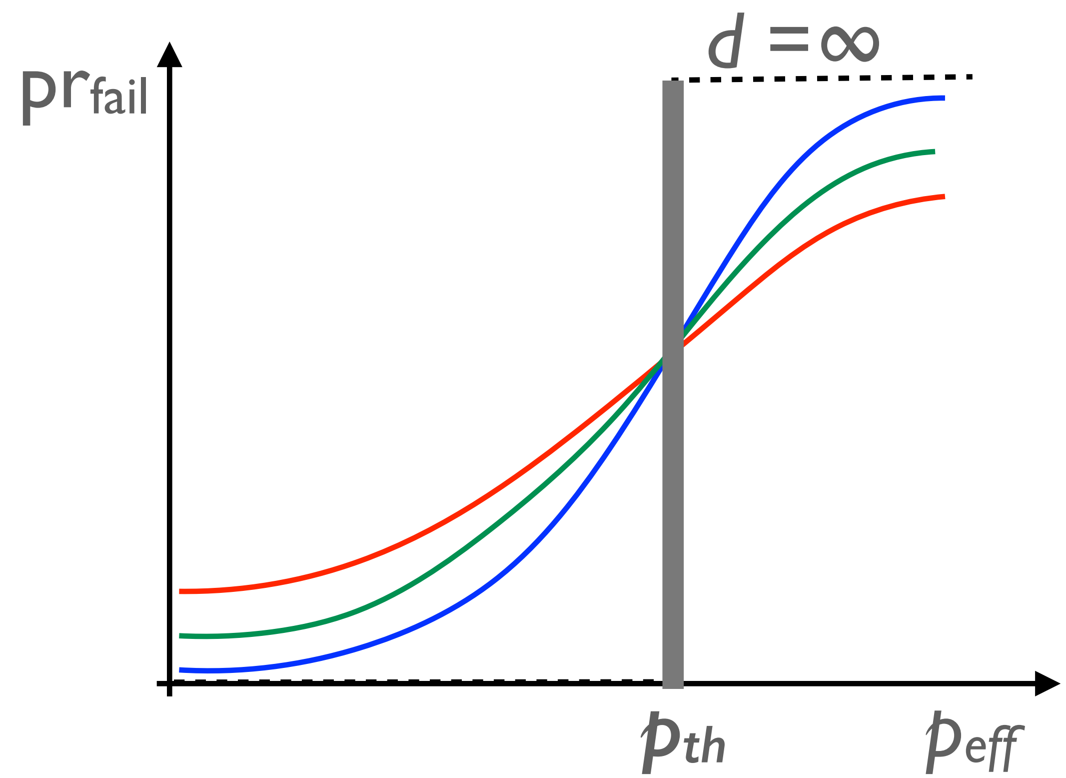
(CORRELATED) NOISE MODELS

- Three simple Pauli error models with the error rate *p*:
  - **bit-/phase-flip** — every qubit independently affected by X and Z
    $p_{\text{eff}} = 2p - p^2$ pr(X) = pr(Z) = *p*
  - **depolarizing** — every qubit independently affected by X, Y, or Z
    $p_{\text{eff}} = p$ pr(X) = pr(Y) = pr(Z) = *p/3*
  - **NN-depolarizing** — every pair of nearest-neighbor qubits affected by non-trivial Pauli $P_1P_2 \neq II$ w/ $pr(P_1P_2) = p/15$
    $p_{\text{eff}}^{(n)} = \frac{4}{5}np + o(p^2)$
- easy to specify/simulate; capture realistic noise features (correlations)
- <u>**effective error rate**</u> = probability of any non-trivial error on the qubit

ERROR-CORRECTION THRESHOLDS

- Consider a family of codes with growing distance d.





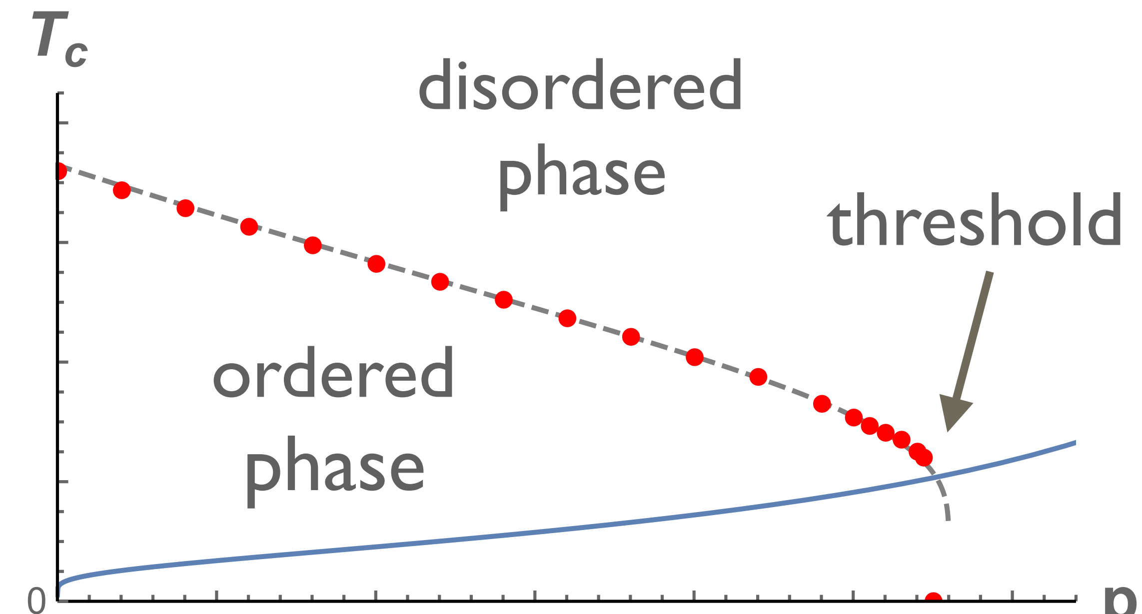
- What is the probability $\mathrm{pr}_{\mathrm{fail}}(p_{\mathit{eff}}, d)$ of unsuccessful decoding?

- **Threshold** p_{th} = "max error rate", i.e., if error rate $p_{\mathit{eff}} \leq p_{th}$, then $\mathrm{pr}_{\mathrm{fail}}(p_{\mathit{eff}}, d) \longrightarrow 0$ as the distance $d \longrightarrow \infty$.

- Non-zero threshold is a non-trivial property: guaranteed # errors ~ d, but w.h.p. correct # errors ~ d^2!

INTERLUDE — OPTIMAL THRESHOLDS FROM STAT-MECH

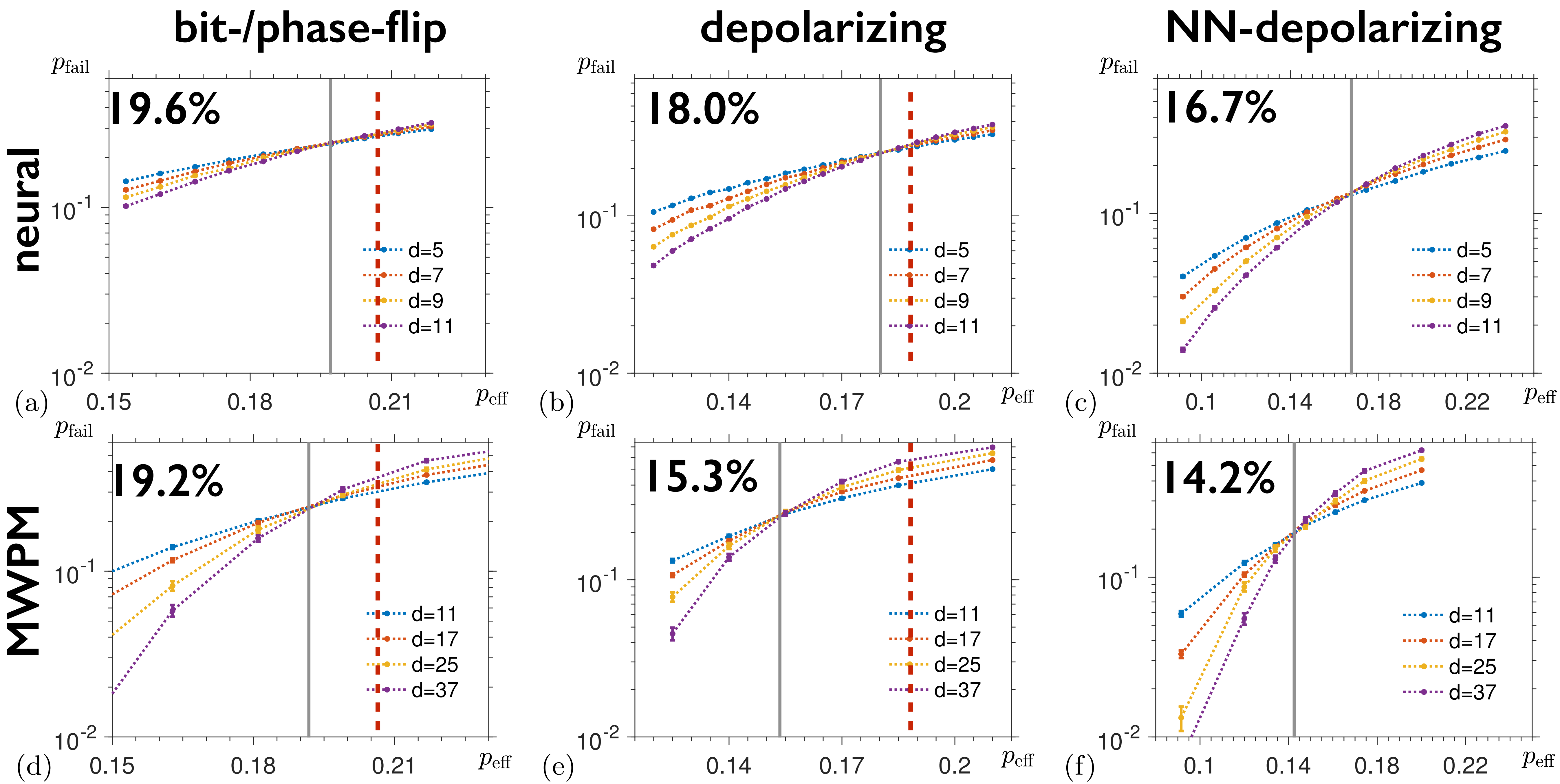
- Values of thresholds relevant for: comparing codes and decoders, overhead estimates, experiment, …
- *[DKLP03]:* connection between toric code decoding and a classical spin model (random-bond Ising)
- ordered phase = successful correction
  critical point = optimal threshold



- Other models *[KBMD09,KBBSP17,LMNWB18]*: 2D color code (3-body Ising), 3D toric and codes (Ising gauge theory), 2D Bacon-Shor-type,…

Dennis et al.'03; Katzgraber&Bombin&Martin-Delgado'09; K. et al.'17; Li et al.'18

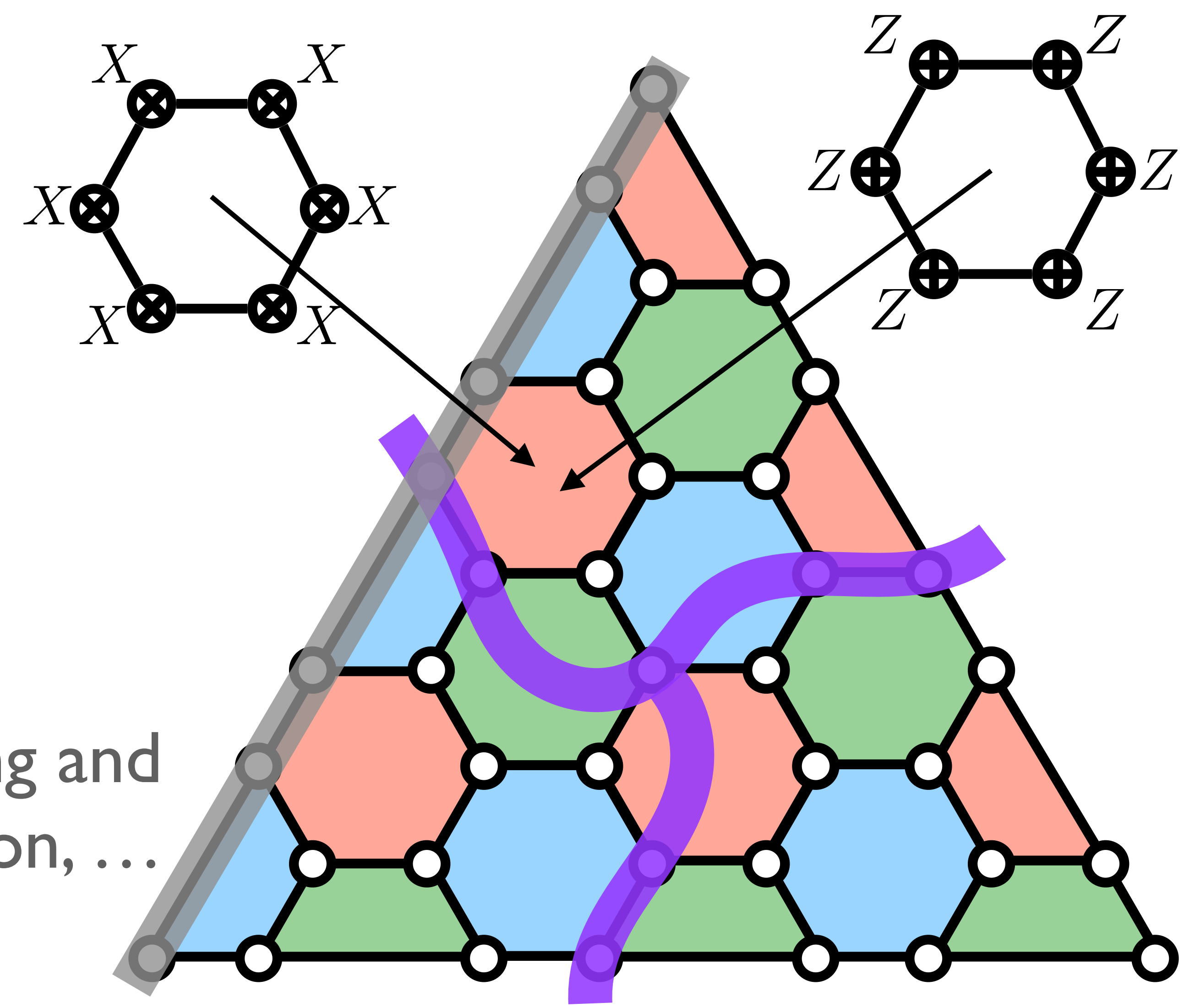
THRESHOLDS FOR TORIC CODE



arXiv: 1802.08680

2D COLOR CODE

- 2D color code *[BMD06]* lattice:
  — 3-colorable faces
  — 3-valent vertices
- qubits = vertices (same positions!)
  stabilizers = X-face and Z-face
- logical Clifford gates are transversal!
- other ideas *[B15,B16]*: code switching and dim-jump, single-shot error correction, …
- decoding seems to be challenging, thus worse performance?!



Bombin&Martin-Delgado'06; Bombin'15,'16

COLOR CODE EXCITATIONS

- ground space of stabilizer Hamiltonian = code space

$$H_{CC} = -\sum_{f \in F} X_f - \sum_{f \in F} Z_f$$

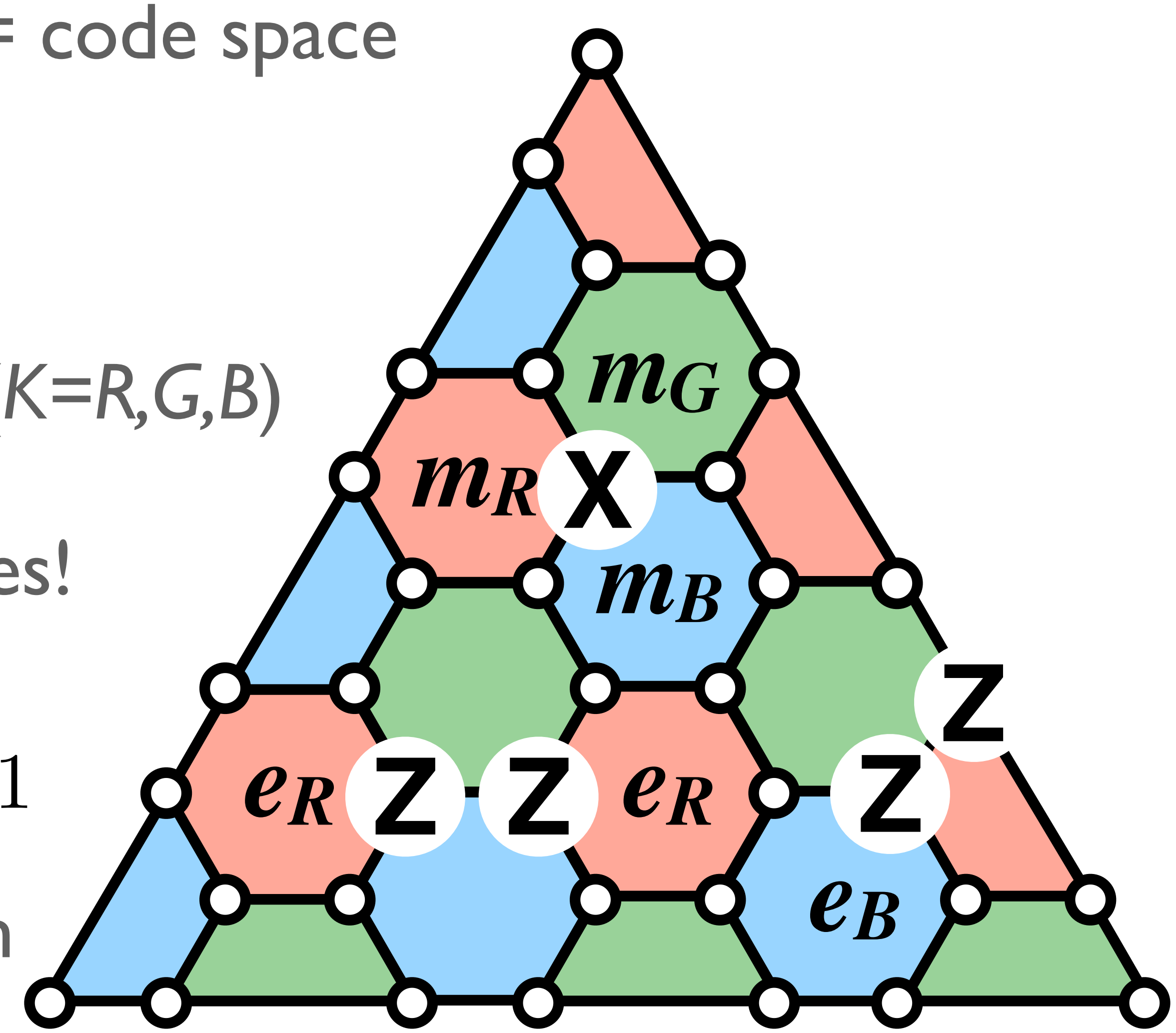
- violated stabilizers = excitations e_K, m_K (K=R,G,B)

- **bulk:** excitations can be created in triples!

$$e_K \times e_K = m_K \times m_K = 1$$
$$e_R \times e_G \times e_B = m_R \times m_G \times m_B = 1$$

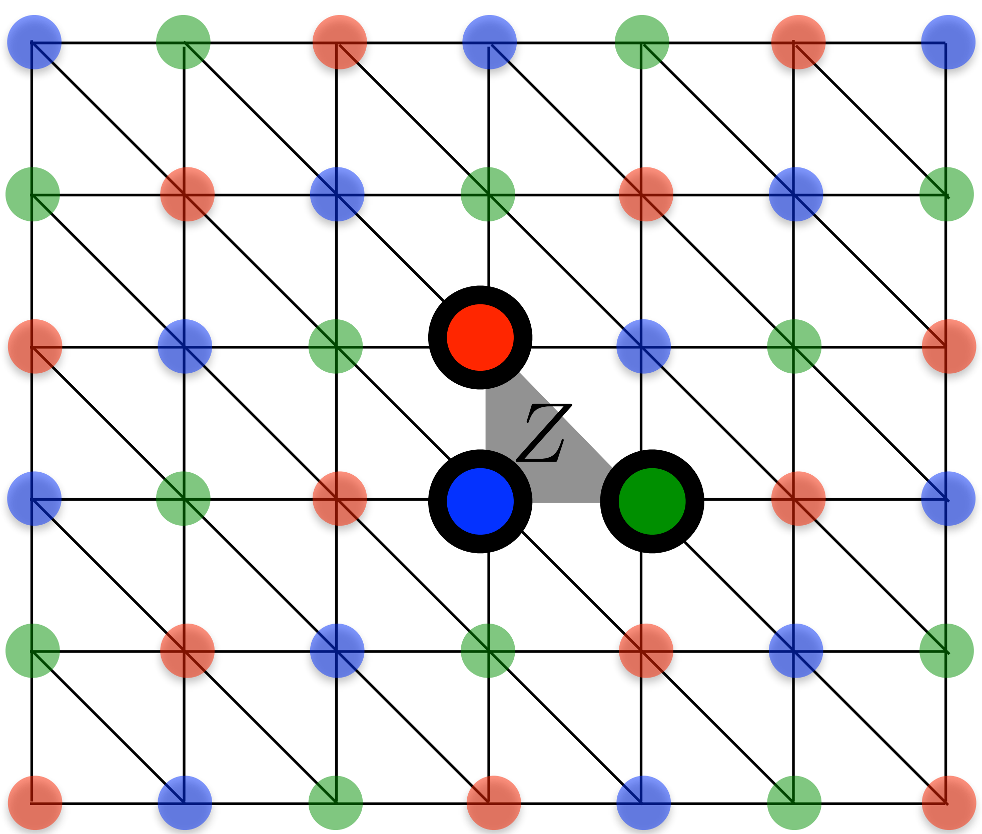
- **boundary:** can create a single excitation

- more boundaries and defect lines than in toric code *[Y15, KBPE18]*!



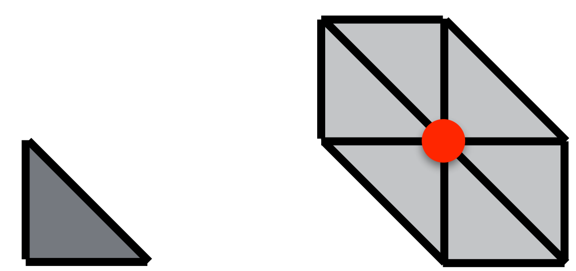
Yoshida'15; Kesselring et al.'18

2D COLOR CODE REDEFINED



- **(Dual) lattice:** made of triangles and vertices are 3-colorable.

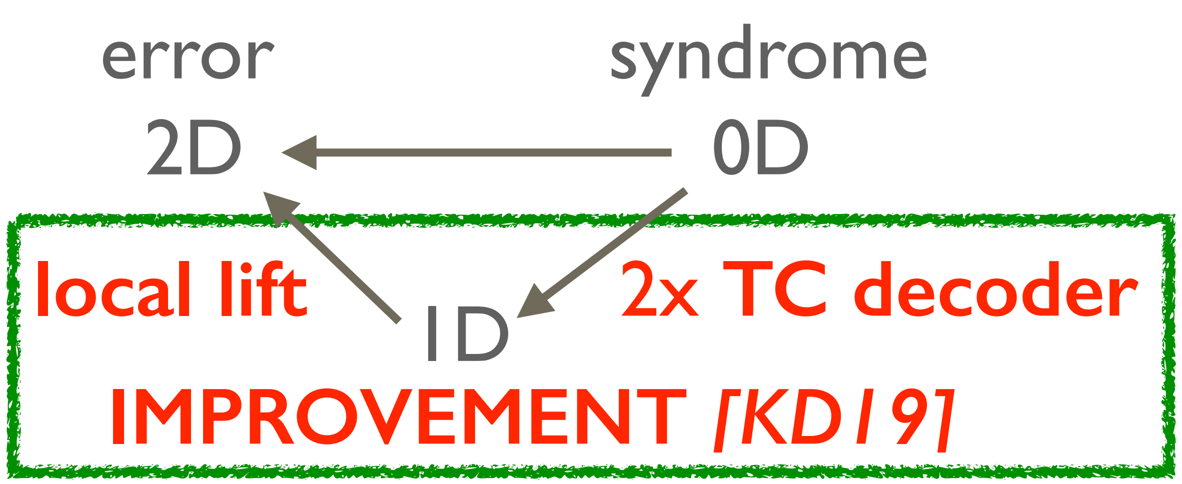
- **2D color code** redefined:
  - qubits = triangles,
  - stabilizers = X- & Z-vertices.

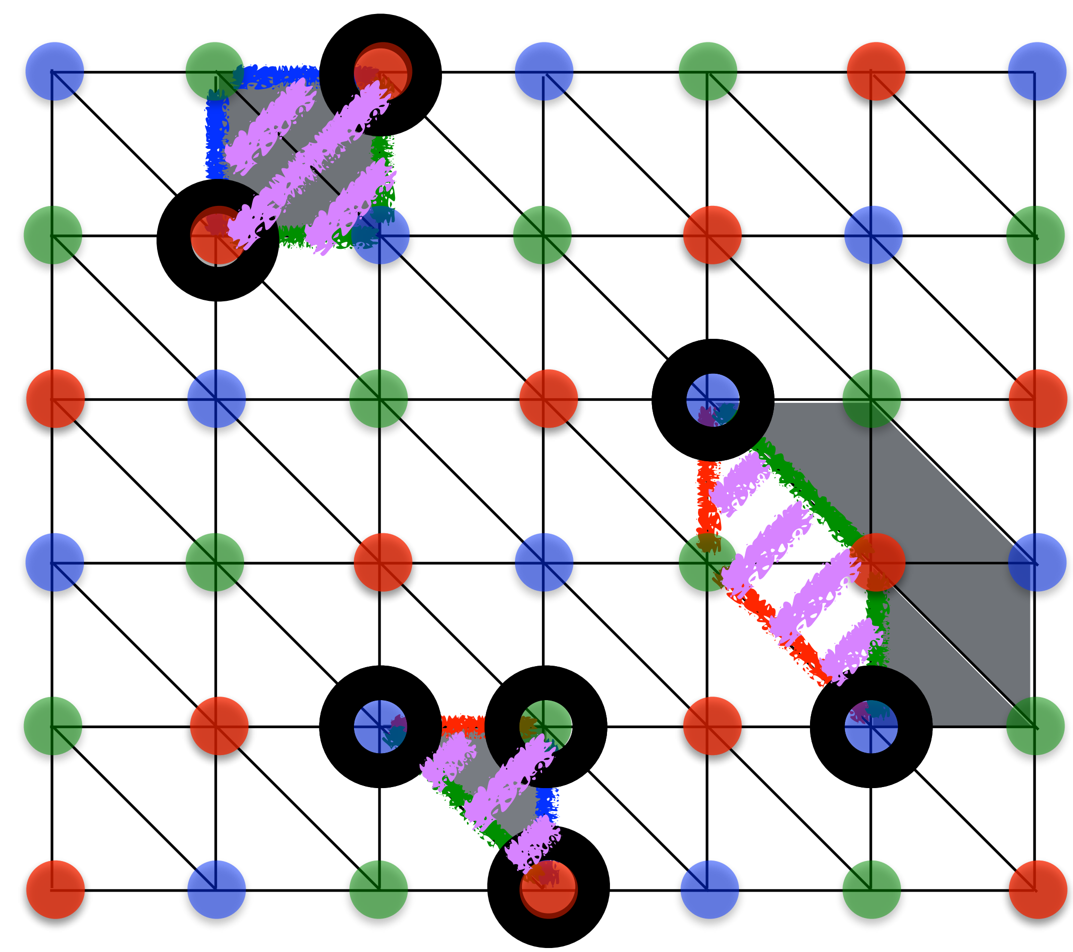


- Decoding seems to be more challenging: excitations created in triples, thus not only pairing!

HOW TO DECODE COLOR CODES?

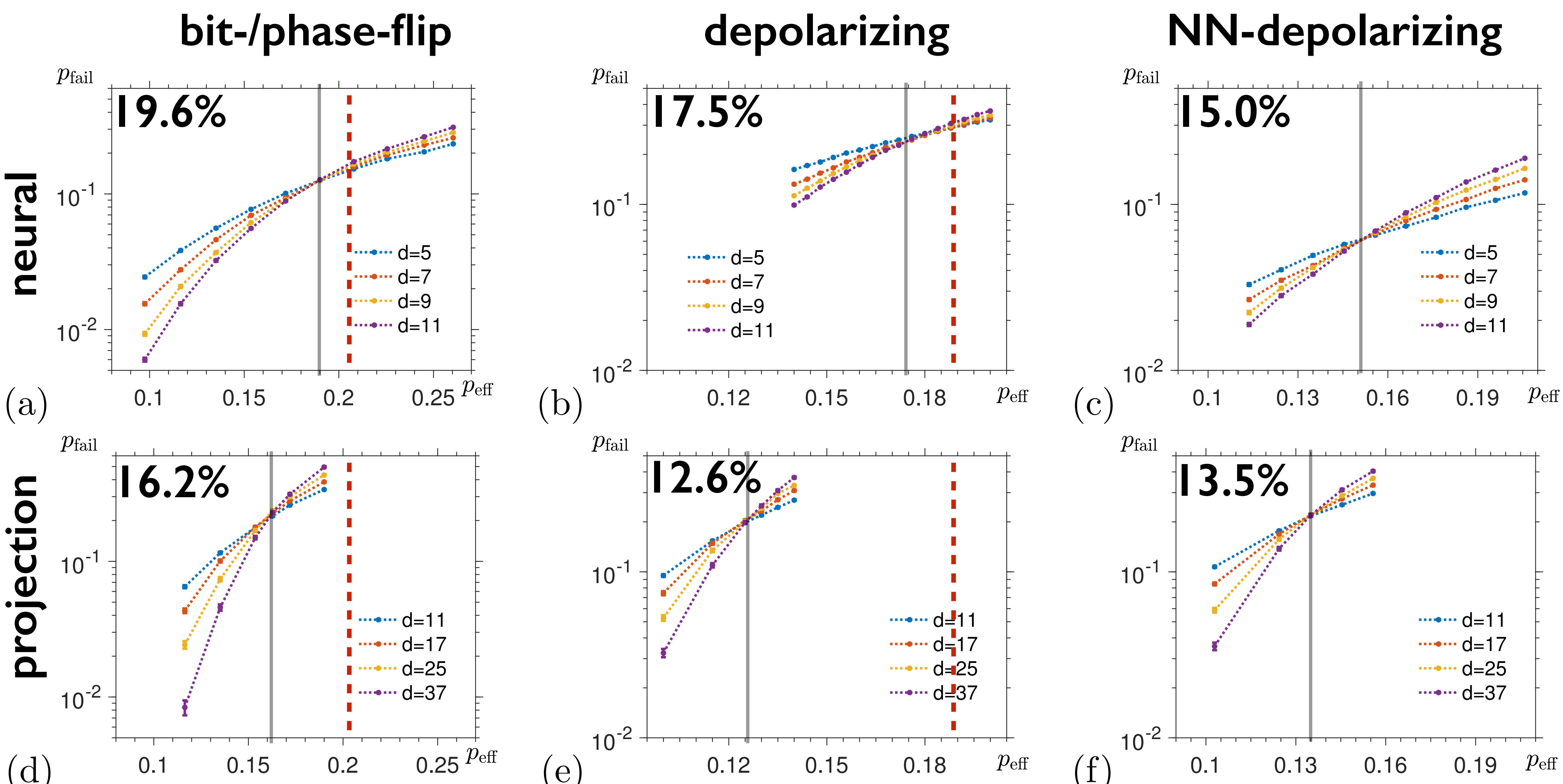
- **Idea:** color and toric codes are related *[KYP15]* — can we use existing toric code decoders?
- Noise changes — correlated errors!
- 2D projection decoder *[D14]*:
  - TC decoder on three sublattices,
  - global filling.





K.&Yoshida&Pastawski'15; Delfosse'14; K.'18; K.&Delfosse'19 (in preparation)

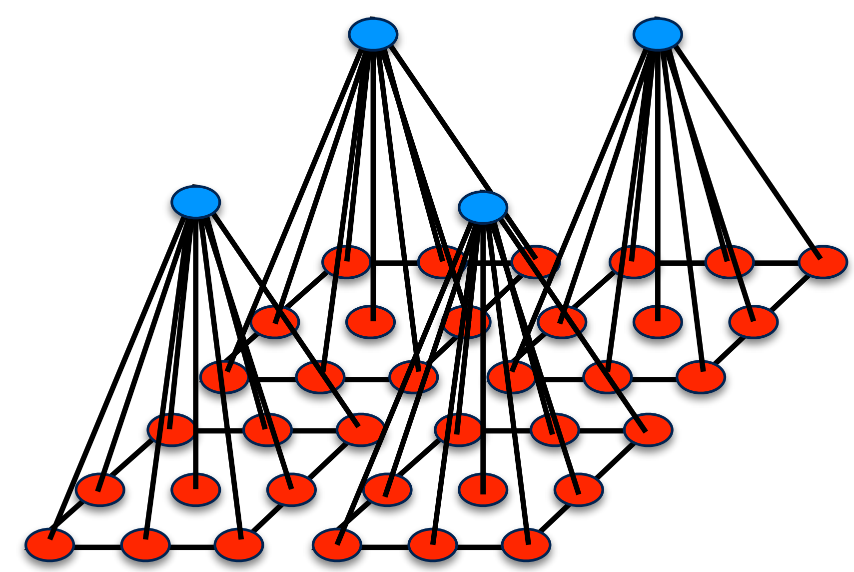
THRESHOLDS FOR COLOR CODE

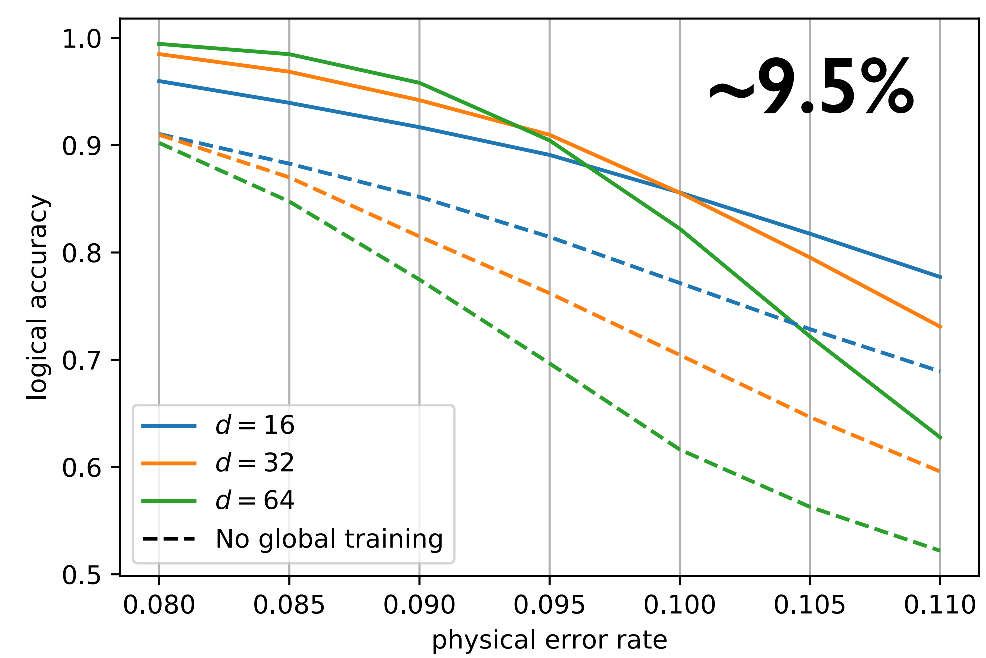


arXiv: 1802.08680

SCALING UP CODE DISTANCE

- So far we haven't used any knowledge about the system. Geometric locality of stabilizers!
- *[H04, KP18, DCP10, BH13,…]*: decoders based on cellular automata, renormalization group, … Provable thresholds!
- Translational invariance and RG ideas: convolutional neural networks?
- *[N18]*: large-distance toric code ($d \leq 64$)



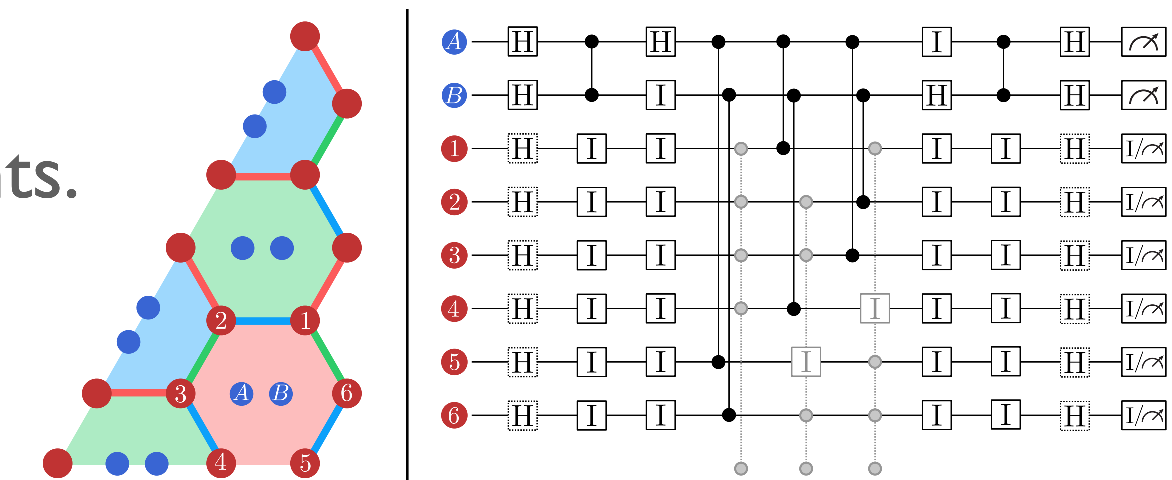


Ni'18

Harrington'04; K.&Preskill'18; Duclos-Cianci&Poulin'10; Bravyi-Haah'13; Ni'18

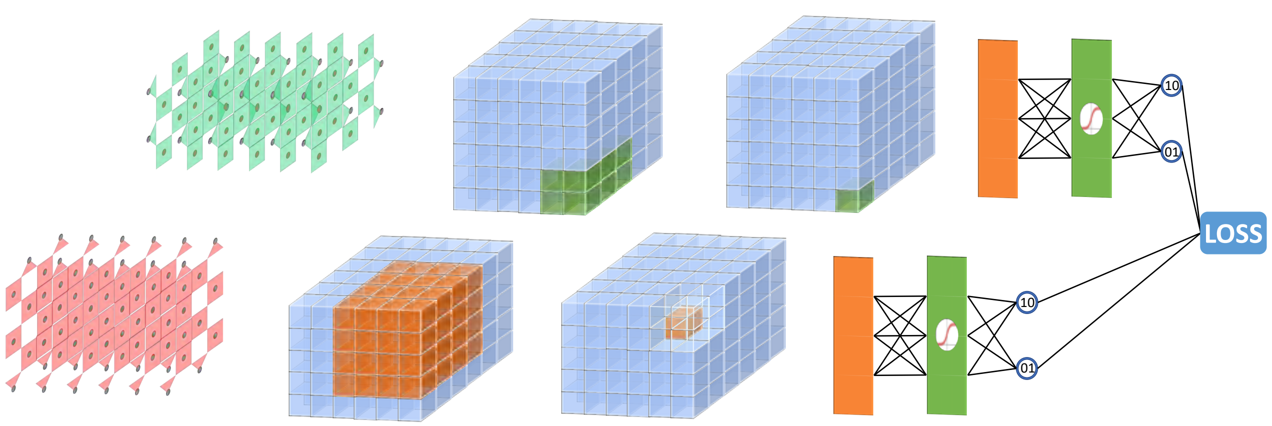
REALISTIC SCENARIO: CIRCUIT-LEVEL NOISE

- Syndrome extraction is far from perfect! Need: ancillas and repeated measurements.
- *[CR18, BCCBO18]:* small-distance toric/surface and color codes ($d \leq 7$).

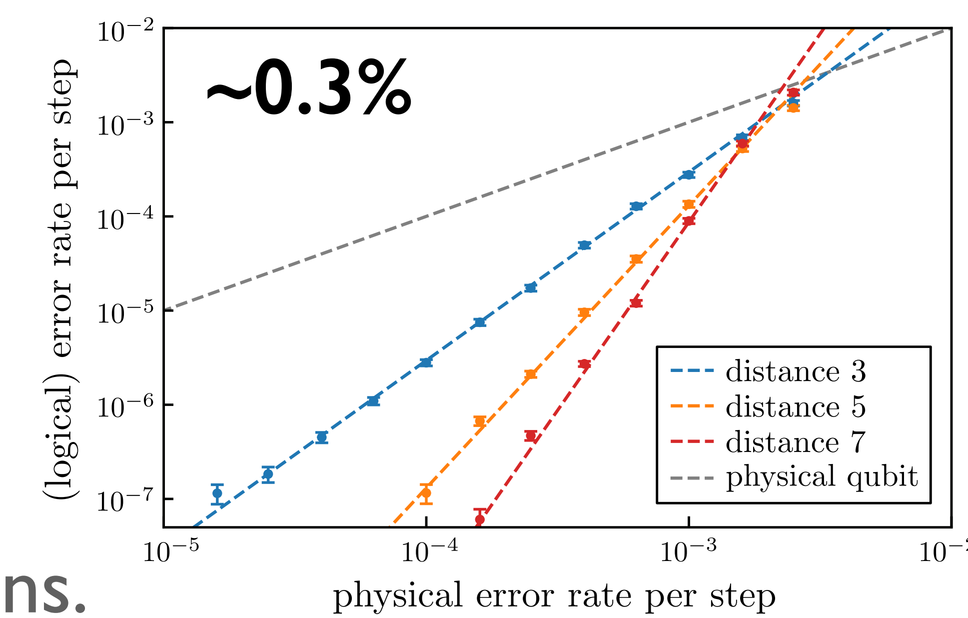


Baieruther et al.'18

- Convolutional and recurrent neural networks with internal memory.



Chamberland&Ronagh'18



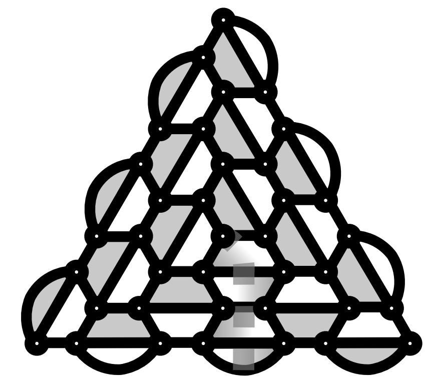
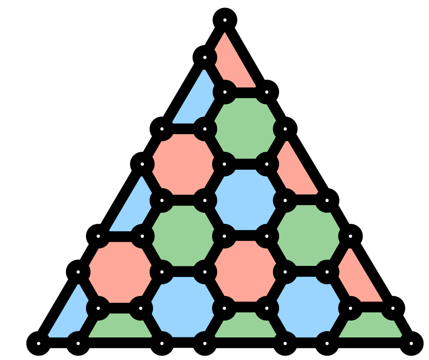
Baieruther et al.'18

- Decoding runtime of a trained network ~10 ns.

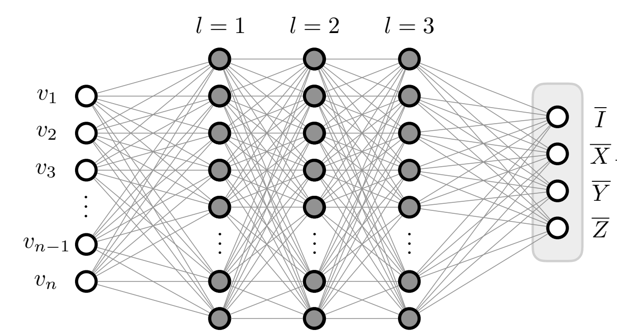
Chamberland&Ronagh'18; Baieruther et al.'18

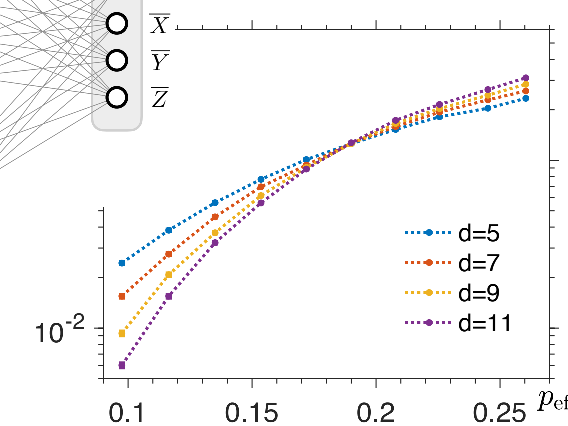
DISCUSSION

- We studied decoding of toric/color codes for (correlated) noise: bit-/phase-flip, depolarizing, NN-depolarizing.

- **Our results:**
  - neural-network decoding is **<u>versatile</u>** and **<u>outperforms</u>** efficient decoders
  - 2D color code threshold **<u>significantly improved</u>**





- Future: transferability, real-experiment data and training in low error-rate regime, certifying performance, interpretability, …

THANK YOU!

arXiv: 1802.08680