QUANTUM

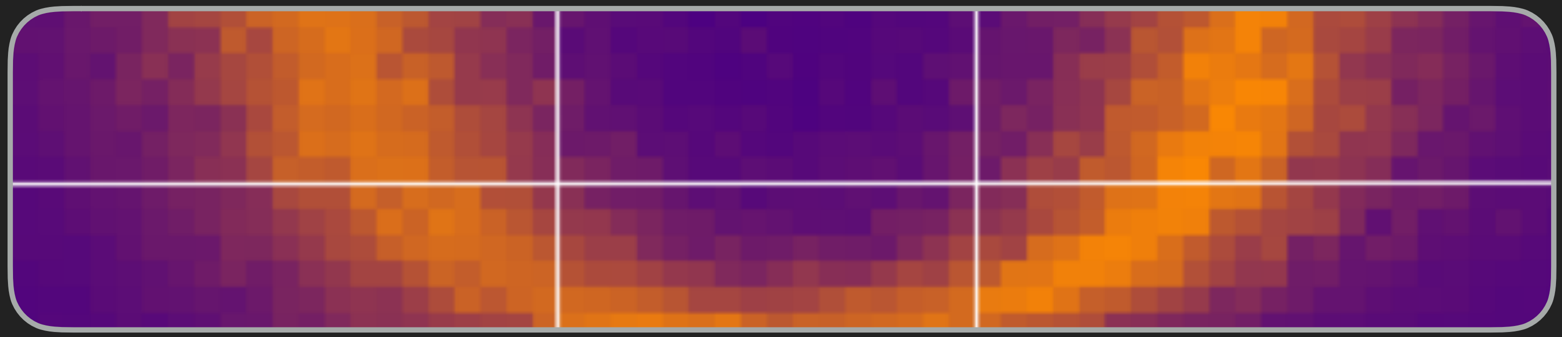
MACHINE LEARNING

JENS EISERT, FU BERLIN

JOINT WORK WITH RYAN SWEKE, MARKUS KESSELRING,EVERT P. L. VAN NIEUWENBURG AND OTHERS,

MACHINE LEARNING → QUANTUM PHYSICS

- Learning quantum phases of matter

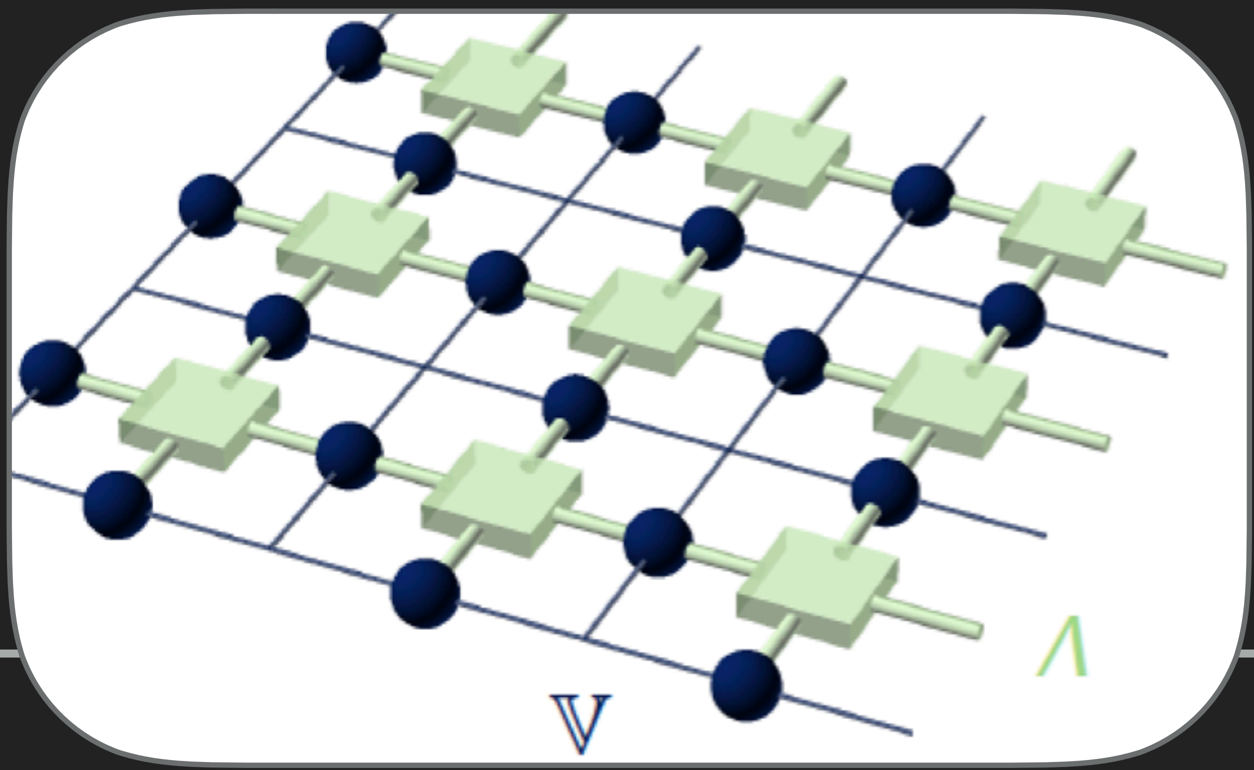


- Tomographic recovery
- Learning descriptors in quantum materials

Carrasquilla, Melko, Nature Phys 13, 431 (2017)
Huembeli, Dauphin, Wittek, Phys Rev B 97, 134109 (2018)
van Nieuwenburg, Bairey, Refael, Phys Rev B 98, 060301 (2018)
Schuett, Arbabzadah, Chmiela, Mueller, Tkatchenko, Nature Comm 8, 13890 (2017)
Broecker, Carrasquilla, Melko, Trebst, Scientific Rep 7, 8823 (2017)

MACHINE LEARNING ⟶ QUANTUM PHYSICS

- Storage of quantum information needs a memory

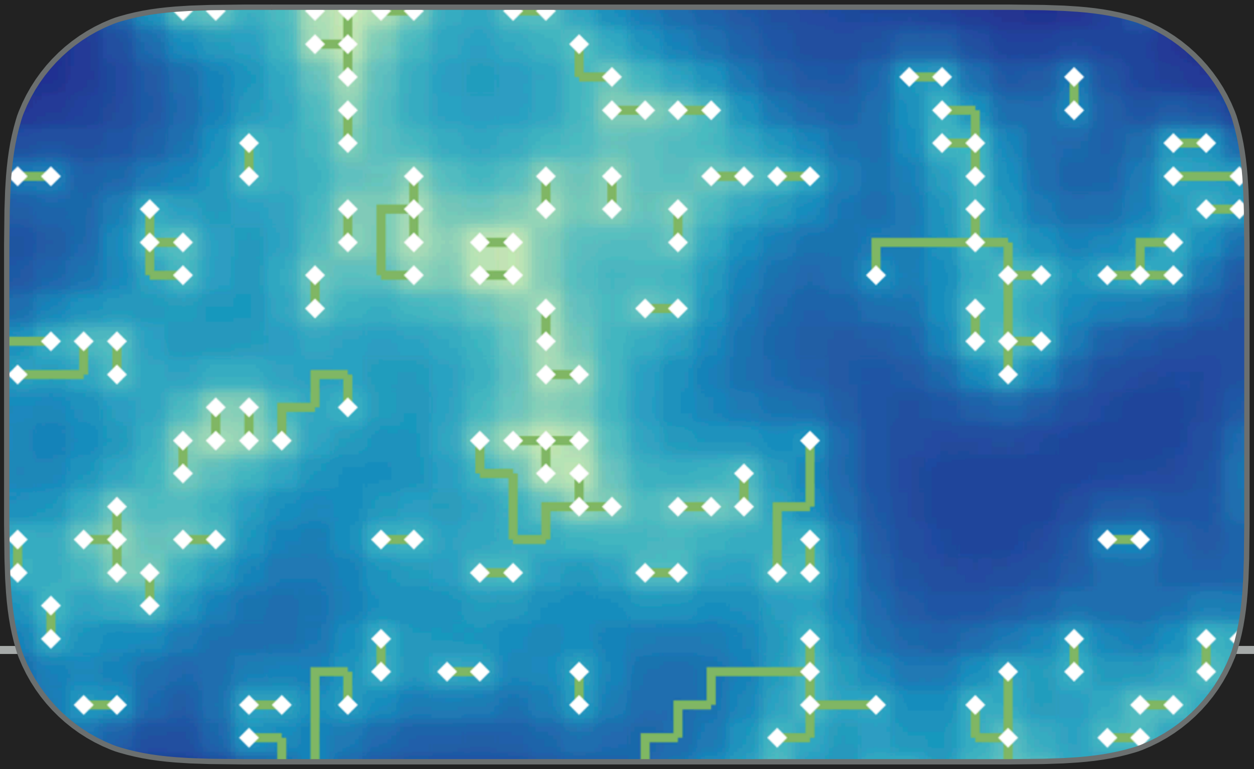


- This will have to be topological

Sweke, Kesselring, van Niewenburg, Eisert,arXiv:1810.07207

MACHINE LEARNING → QUANTUM PHYSICS

- Bottleneck: The "*decoder problem*"



- Main part: *Fault-tolerant setting* natural problem for reinforcement learning

Sweke, Kesselring, van Niewenburg, Eisert,arXiv:1810.07207

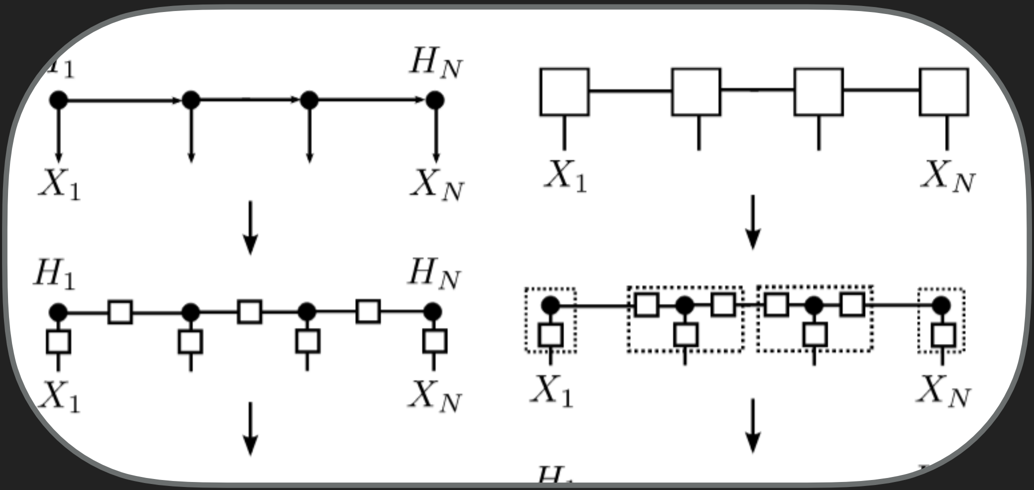
MACHINE LEARNING → QUANTUM PHYSICS

- Improved machine learning using near term quantum circuits
- Quantum inspired tensor network learning

Wittek, Quantum machine learning, Academic press (2014)
Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, Lloyd, Seth,Nature, 549, 195–202 (2017)
Farhi, Neven, arXiv:1802.06002
Melnikov, Nautrup, Hendrik Poulsen, Krenn, Dunjko, Tiersch, Zeilinger, Briegel, PNAS 115, 1221 (2018)
Glasser, Pancotti, Cirac, arXiv:1806.05964
Wiebe, Braun, Lloyd, Seth, Phys Rev Lett 109, 050505 (2012)

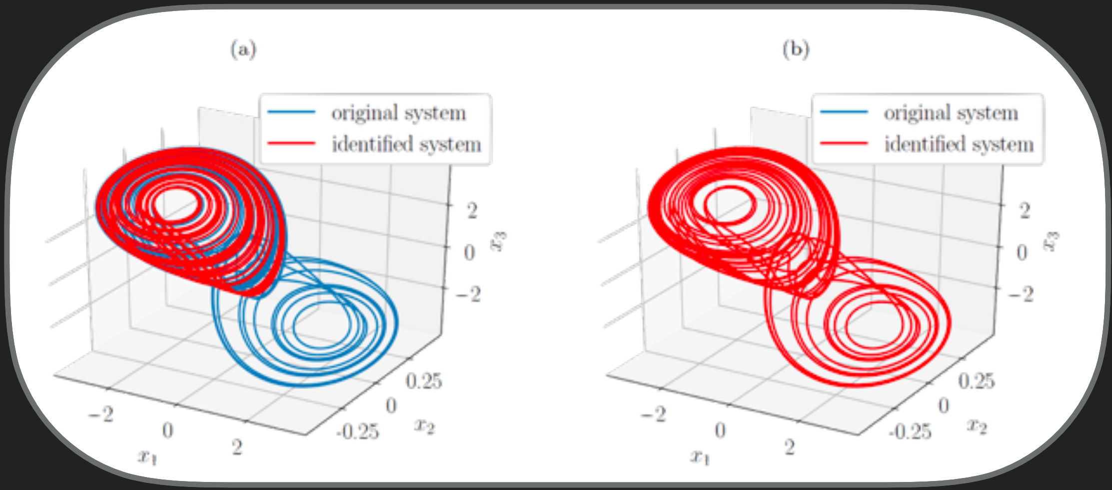
MACHINE LEARNING ← QUANTUM PHYSICS

- Outlook: *Tensor networks* and probabilistic graphical models



- Comprehensive analysis of *expressivity*

Glasser, Pancotti, Cirac, Eisert, Sweke, in preparation



- Quantum-inspired data-driven *learning of dynamical laws*

Gelss, Klus, Eisert, Schuette, J Comput Nonlinear Dynam 14, 061006 (2019)

MACHINE LEARNING — QUANTUM PHYSICS

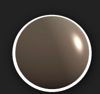
- **What is missing?**
- New and convincing applications of (Q)ML
- A mathematical understanding

MACHINE LEARNING FOR TOPOLOGICAL QUANTUM MEMORIES

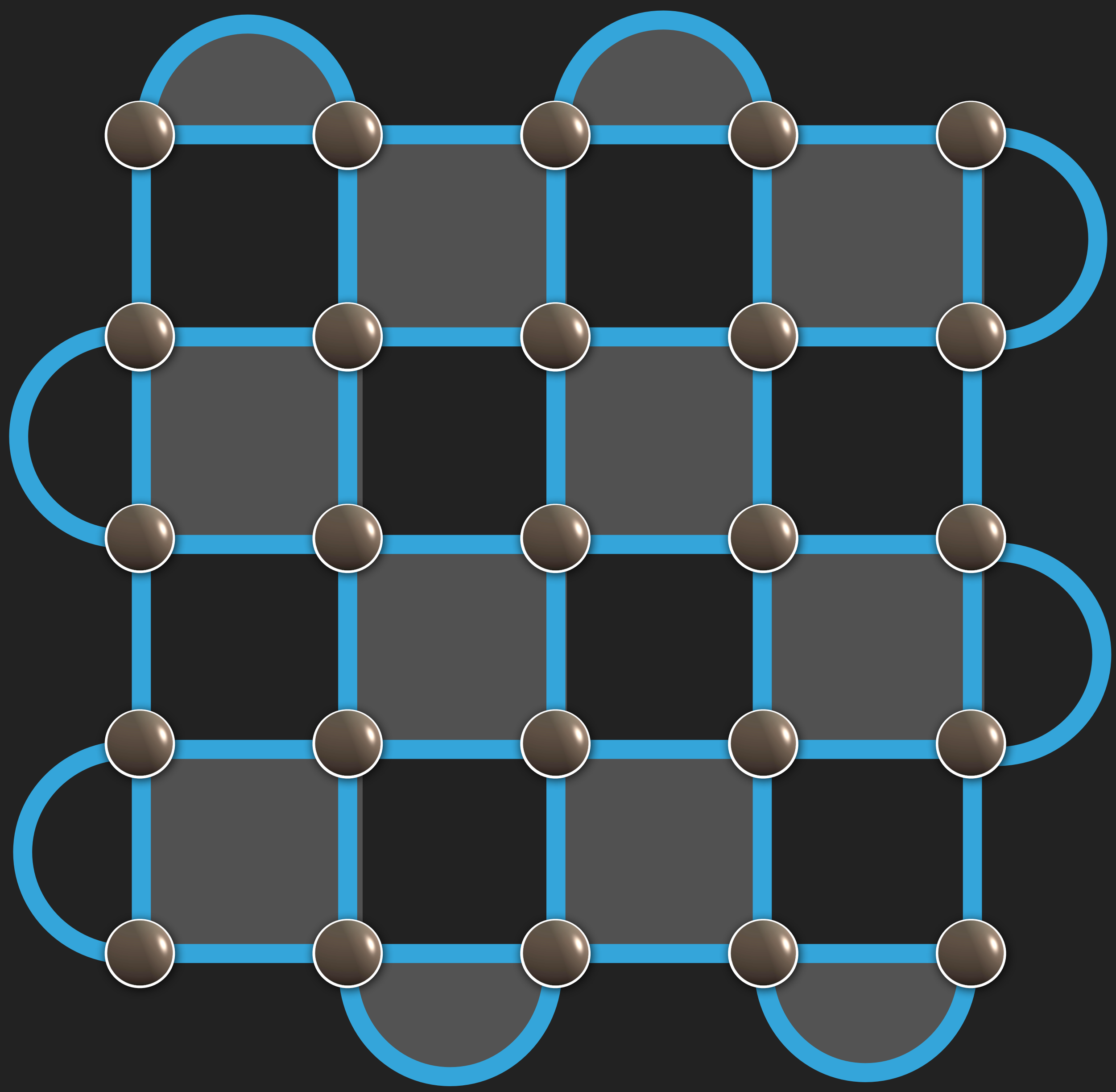
arXiv:1810.07207 and in progress

With Ryan Sweke, Markus Kesselring, Evert van Niewenburg

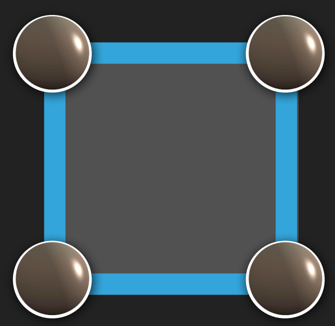
- Storage of quantum information

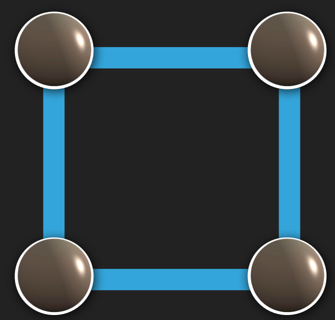


- Storage of quantum information has to be **non-local** to be protected



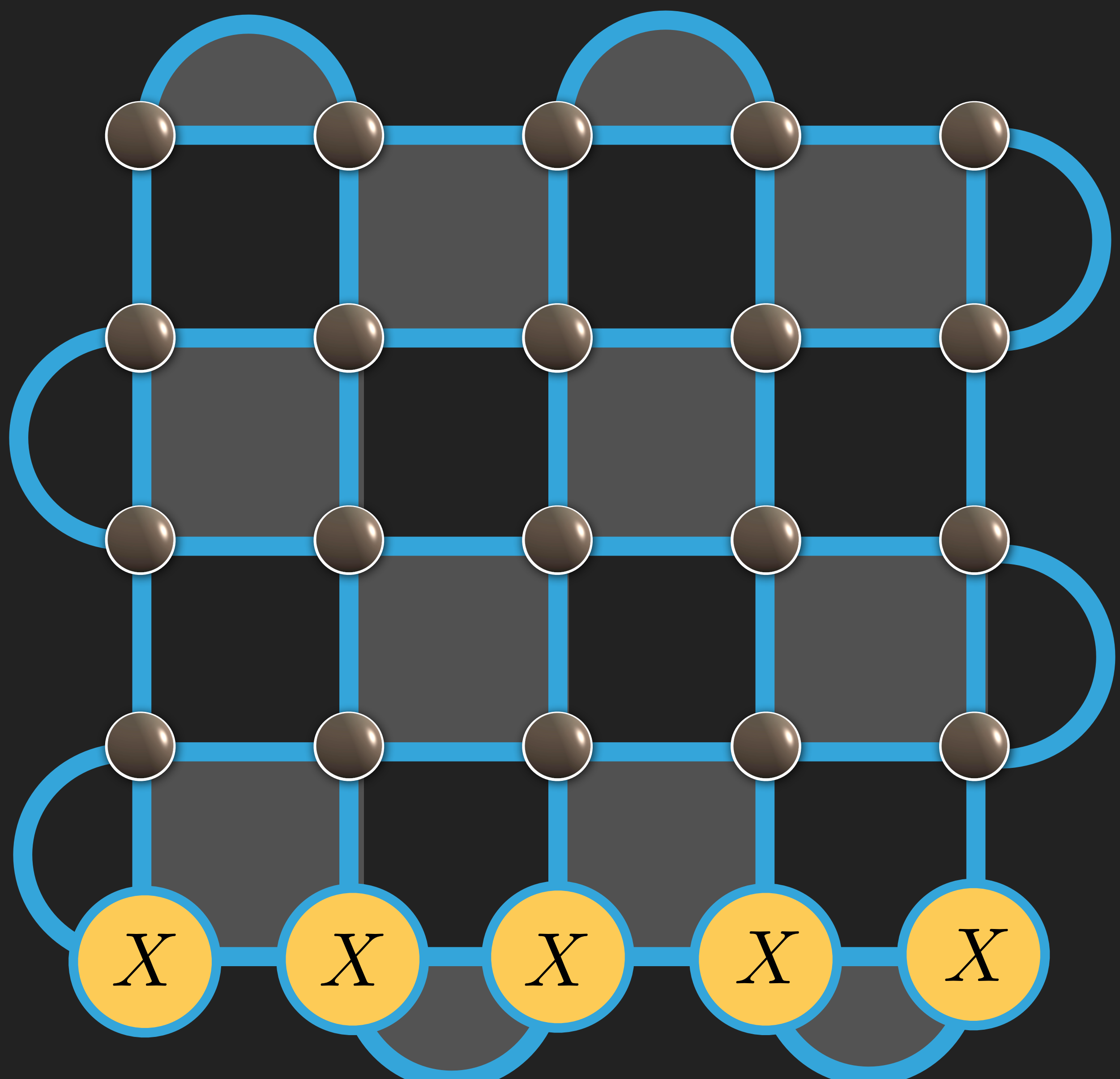
- **Topological surface code**
- Two types of stabilizers





Dennis, Kitaev, Preskill, JMP 43, 4452 (2002)
Kitaev, Ann Phys 303, 2 (2003)

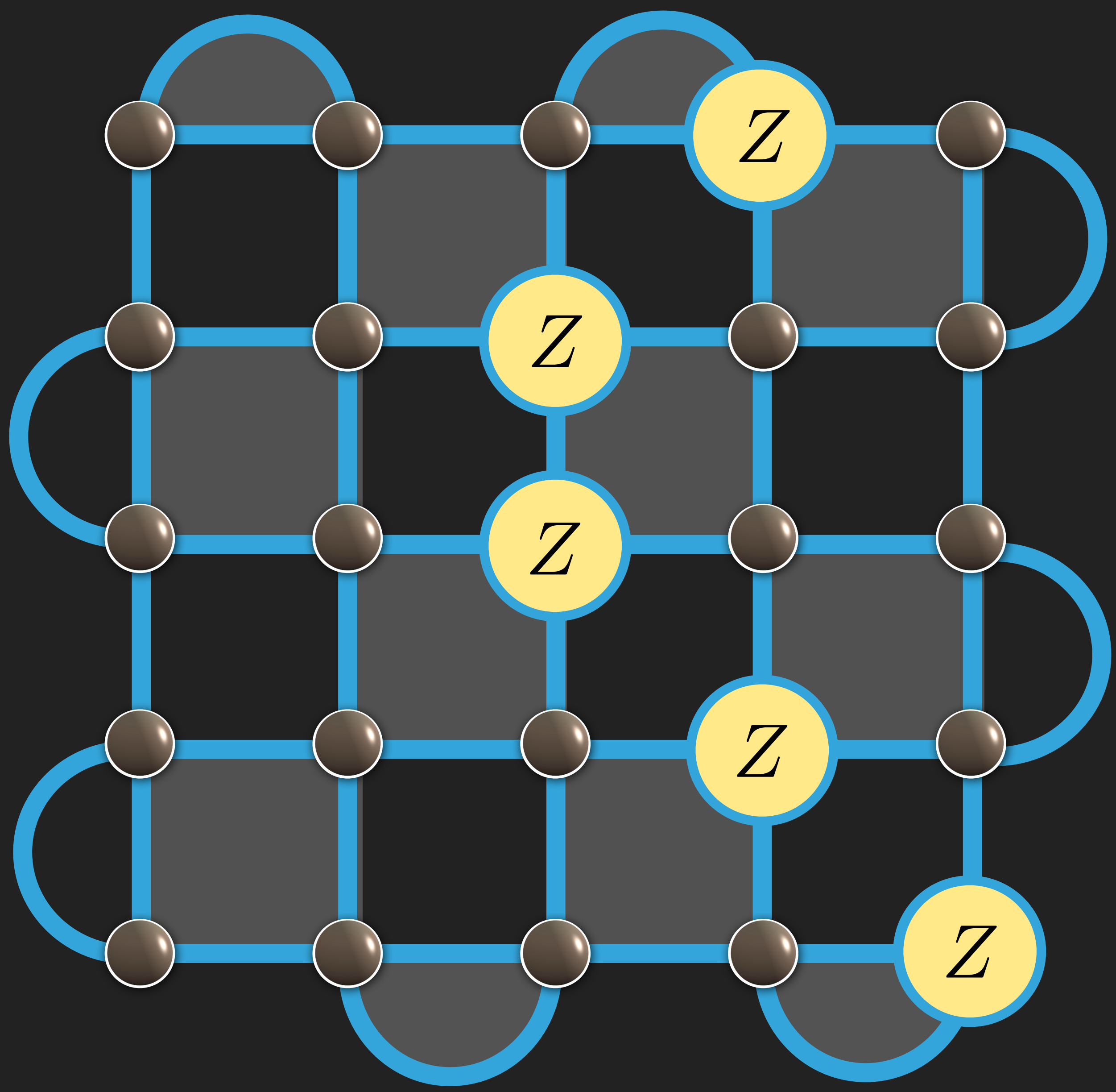
- Storage of quantum information has to be **non-local** to be protected



- **Topological surface code**
- Logical operators span code
- From left to right X_L

Dennis, Kitaev, Preskill, JMP 43, 4452 (2002)
Kitaev, Ann Phys 303, 2 (2003)

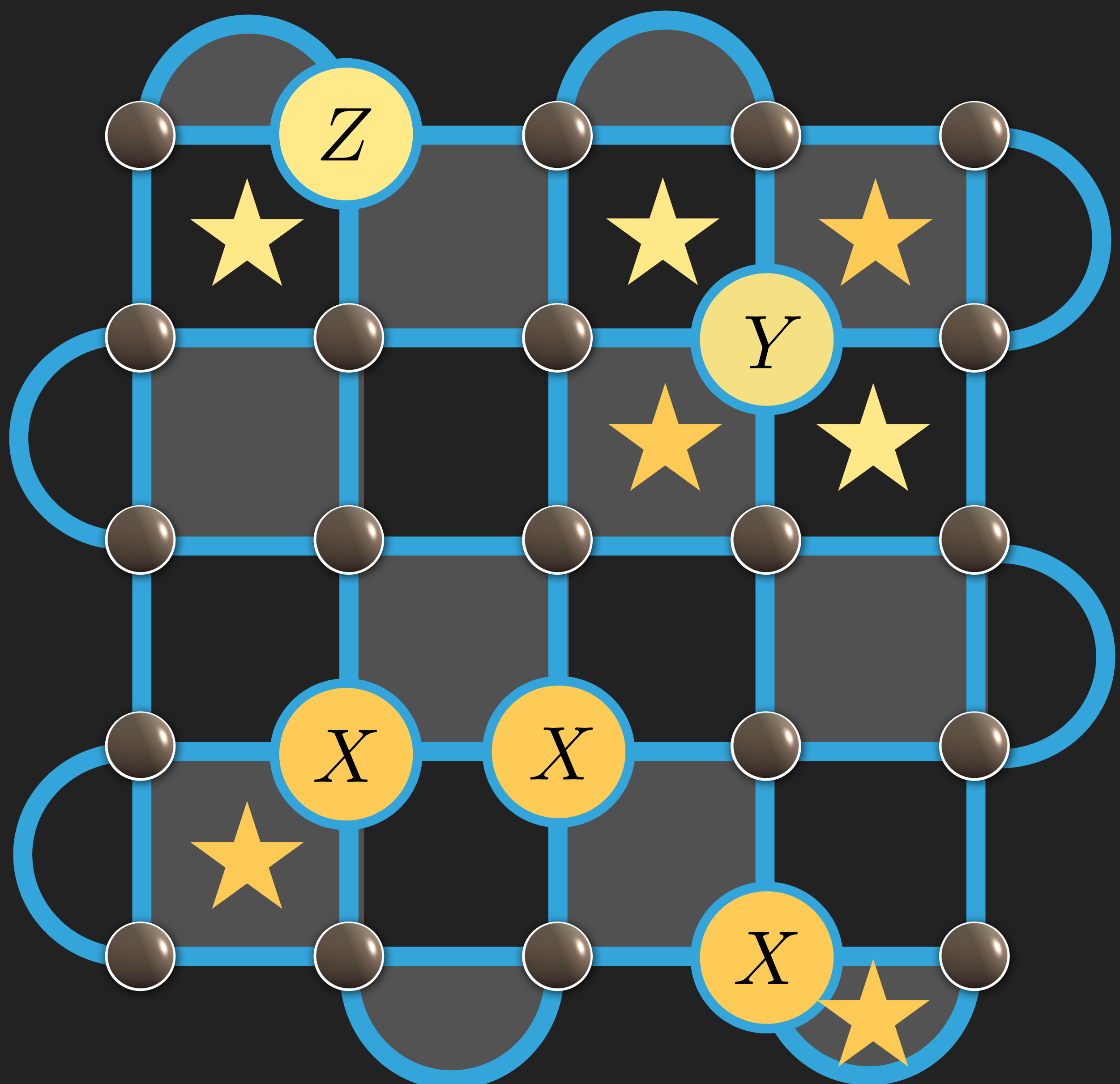
- Storage of quantum information has to be **non-local** to be protected



- **Topological surface code**
- Logical operators span code
- From top to bottom Z_L

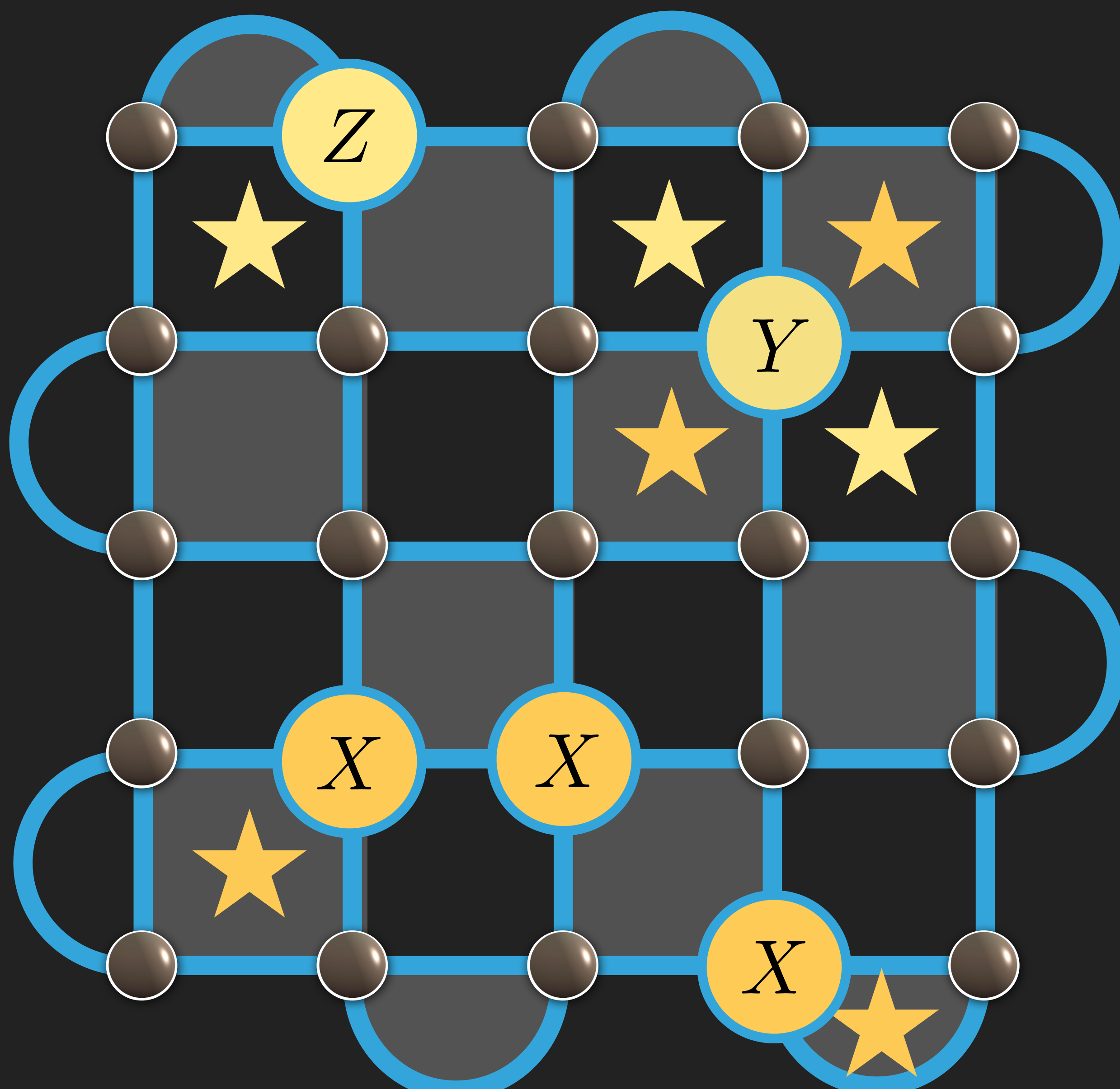
Dennis, Kitaev, Preskill, JMP 43, 4452 (2002)
Kitaev, Ann Phys 303, 2 (2003)

- Costs energy to create pairs of anyons, but then they diffuse



Dennis, Kitaev, Preskill, JMP 43, 4452 (2002)
Kitaev, Ann Phys 303, 2 (2003)

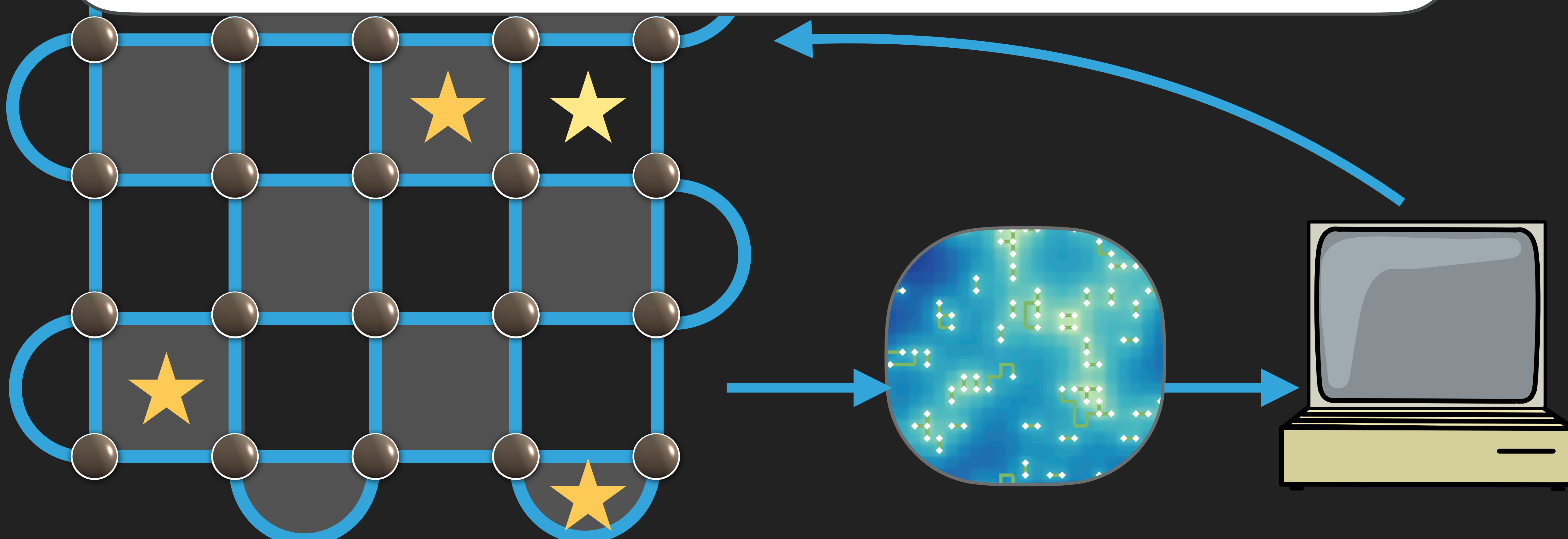
- Need to **correct** (or track)



- (Imperfect) measurement of all stabilizers generates **syndrome**
- Many errors result in same syndrome - paths are not visible

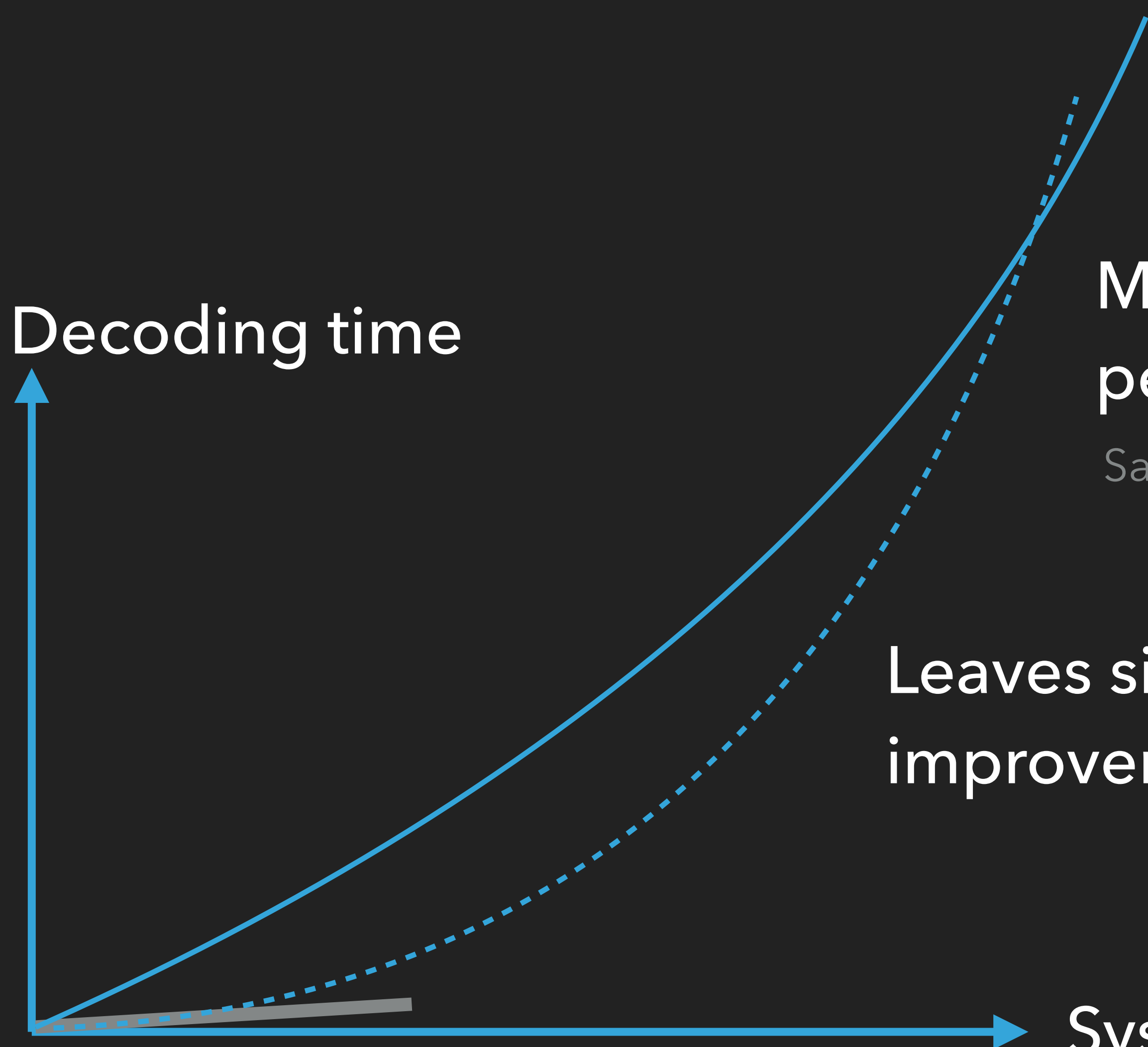
Dennis, Kitaev, Preskill, JMP 43, 4452 (2002)
Kitaev, Ann Phys 303, 2 (2003)

- Need to **correct** (or track)
  - **Decoding problem:** Given the syndrome ($O(L^2)$ bits) - from faulty measurements - identify the **homology class** and find operations to bring the state back to the code space



Fowler, Goyal, Q Inf Comp 9, 721 (2009)
Wootton, Loss PRL 109, 160503 (2012)
Bravyi, Haah, PRL 111, 200501 (2013)
Duclos-Cianci, Poulin, PRL 104, 050504 (2010)
Herold, Campbell, Eisert, Kastoryano, Nature PJ Quant Inf 1, 15010 (2015)

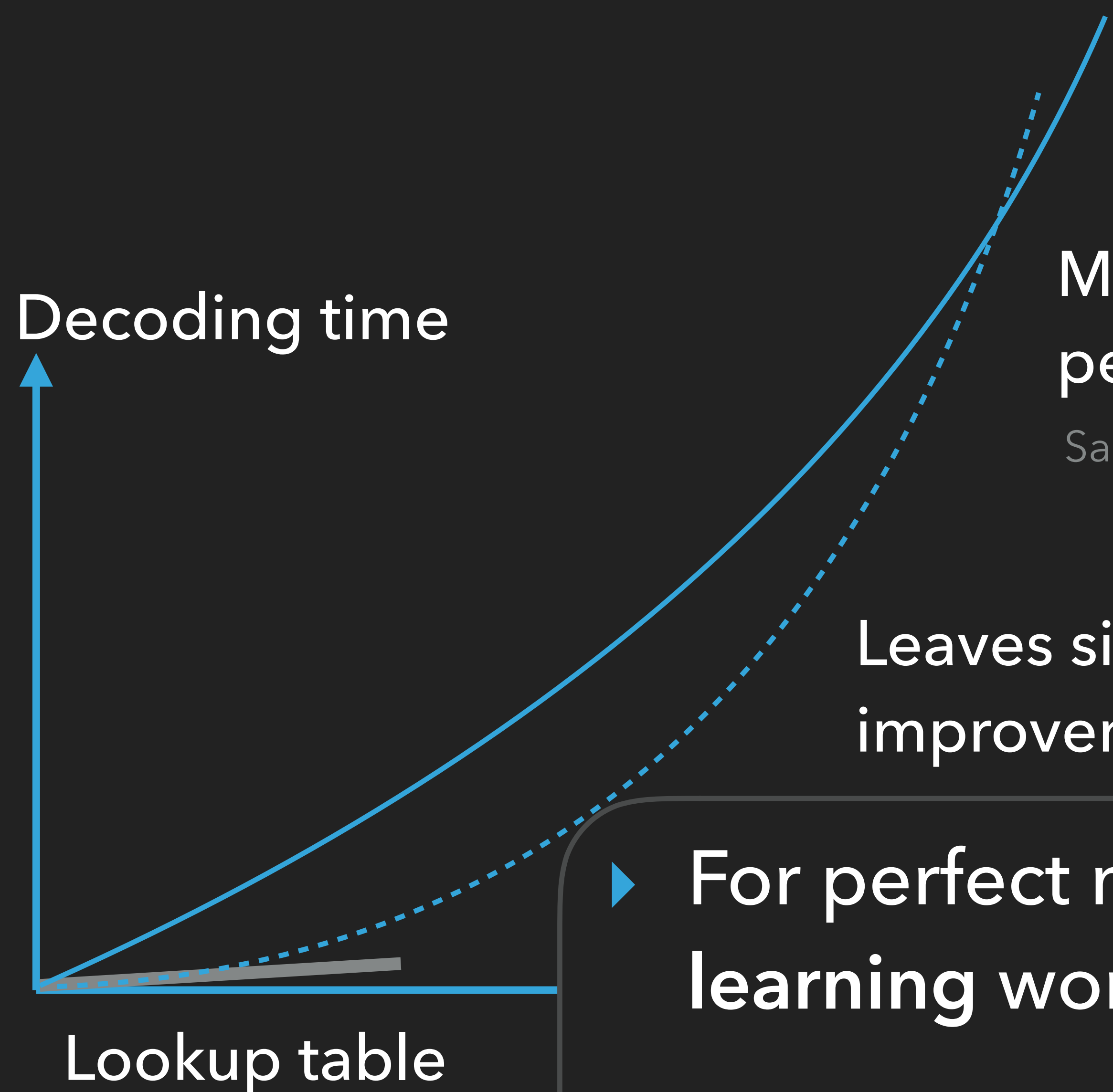
MACHINE LEARNING?



Minimum weight perfect matching $O(L^{4.7})$

Sankiwski, Theo Comp Sc 410, 4480 (2009)

Leaves significant space for improvement with machine learning?



Minimum weight perfect matching $O(L^{4.7})$

Sankiwski, Theo Comp Sc 410, 4480 (2009)

Leaves significant space for improvement with machine learning?

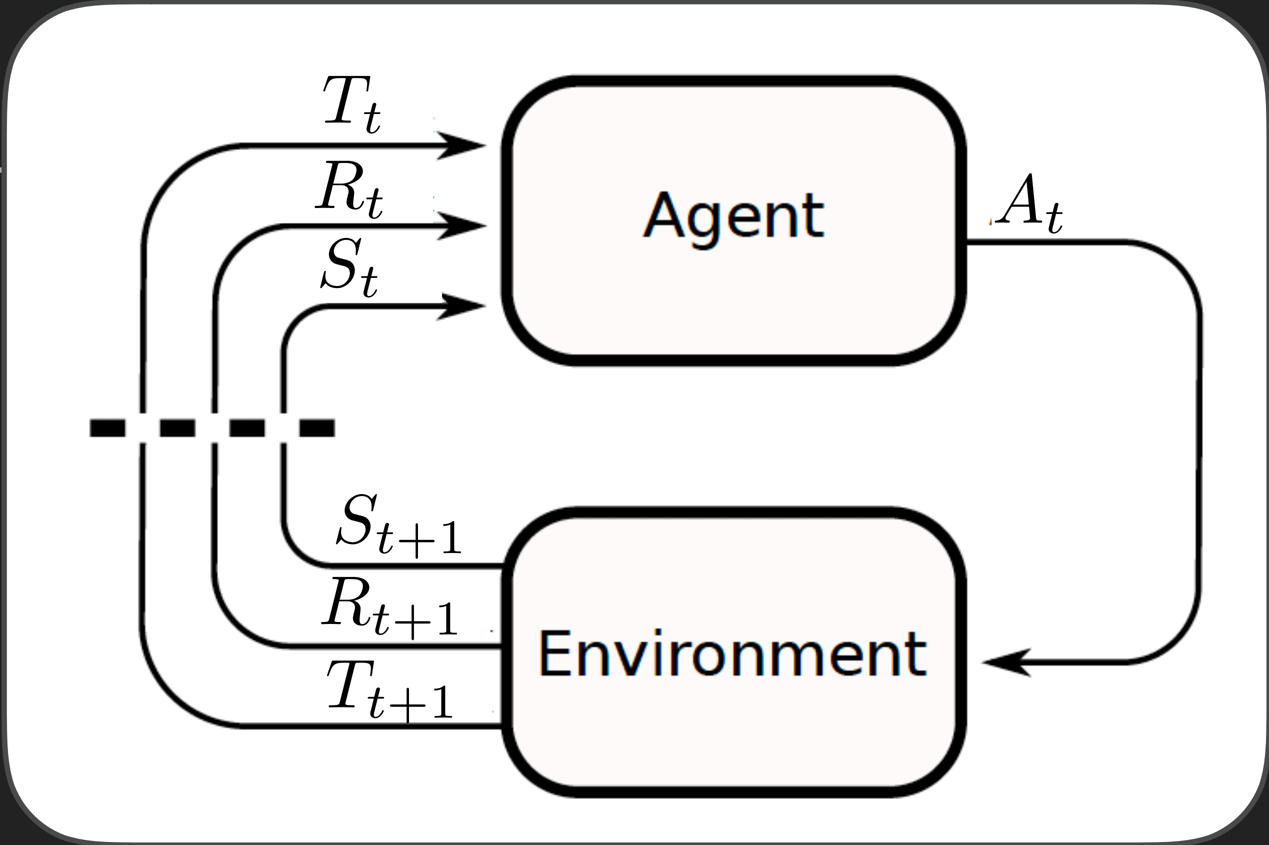
- For perfect measurements, **supervised learning** works well

Torlai, Melko, Phys Rev Lett 119, 030501 (2017)
Krastanov, Jiang, Scientific Reports 7, 11003 (2017)
Varsamopoulos, Criger, Bertels, Quant Sc Tech 3, 1 (2017)
Breuckmann, Ni, arXiv:1710.09489

- **BUT: FAULT TOLERANCE INTRINSICALLY NOT CAPTURED**

REINFORCEMENT DECODING

- Reinforcement learning approach



- State $S_t = \text{syndrome}$
- Action $A_t \in \{1, X_1, \ldots, X_{d^2}\}$
- Reward $R_t = \begin{cases} 1, & \text{if no anyons } \wedge \text{ no logical errors} \\ 0, & \text{otherwise} \end{cases}$
- Game over $T_t = \begin{cases} \text{True}, & \text{if } S_t \text{ can be decoded}, \\ \text{False}, & \text{otherwise} \end{cases}$

Sweke, Kesselring, van Nieuwenburg, Eisert, arXiv:1810.07207

- Environment described by discrete Markov chain

$$p(s', r'|s, a) := \Pr(S_t = s', R_t = r'|S_{t-1} = s, A_{t=1} = a)$$

- Goal is to learn a policy $\pi(a|s)$
- *Value* functions are of key importance, for $\gamma \in (0, 1)$

$$v_\pi(s) = \mathbb{E}_\pi(G_t|S_t = s) = \mathbb{E}_\pi\left(\sum_{k=0}^{\text{terminal}} \gamma^k R_{t+k+1}|S_t = s\right)$$

$$q_\pi(s, a) = \mathbb{E}_\pi(G_t|S_t = s, A_t = a)$$

- Can place order over policies, $\pi > \pi' \Leftrightarrow v_\pi(s) > v_{\pi'}(s) \, \forall s \in \mathcal{S}$
- Want to learn optimal policy π^*

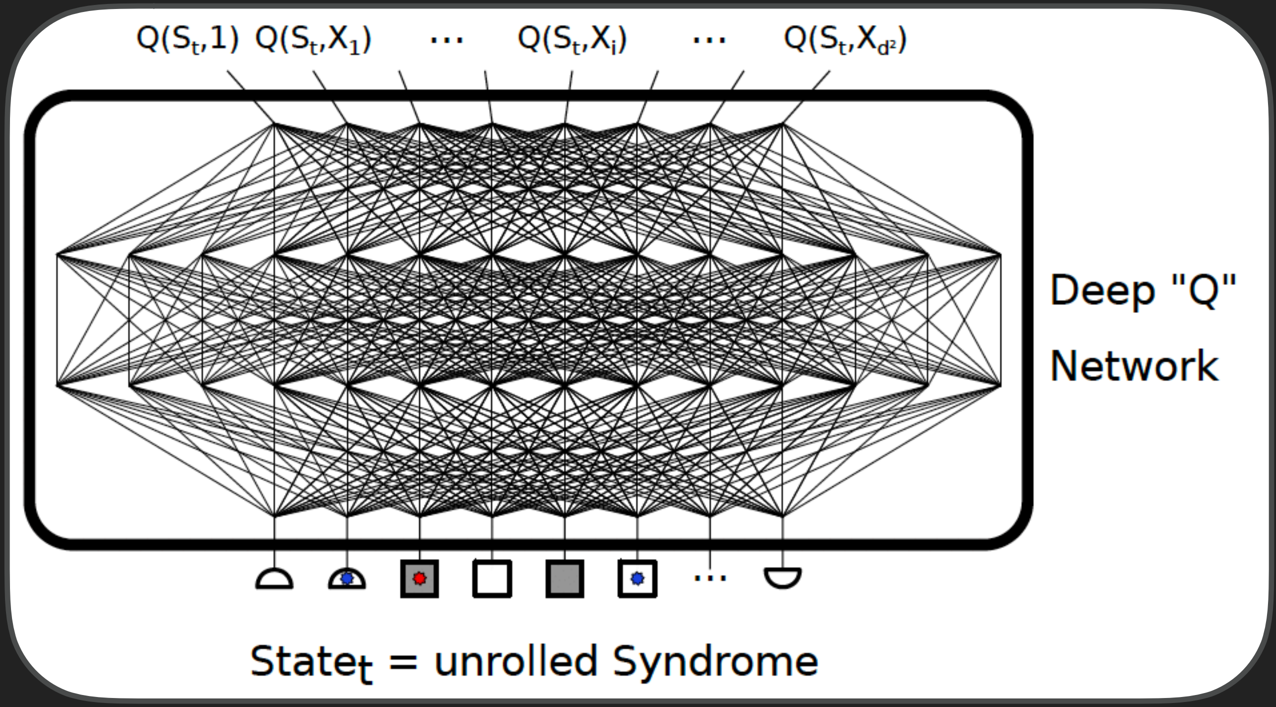
- Optimal Q function is fixed point of Bellman's equation

$$q_*(s,a) = \mathbb{E}(R_{t+1} + \gamma \max_{a'} q_*(S_{t+1},a')|S_t = s, A_t = a)$$

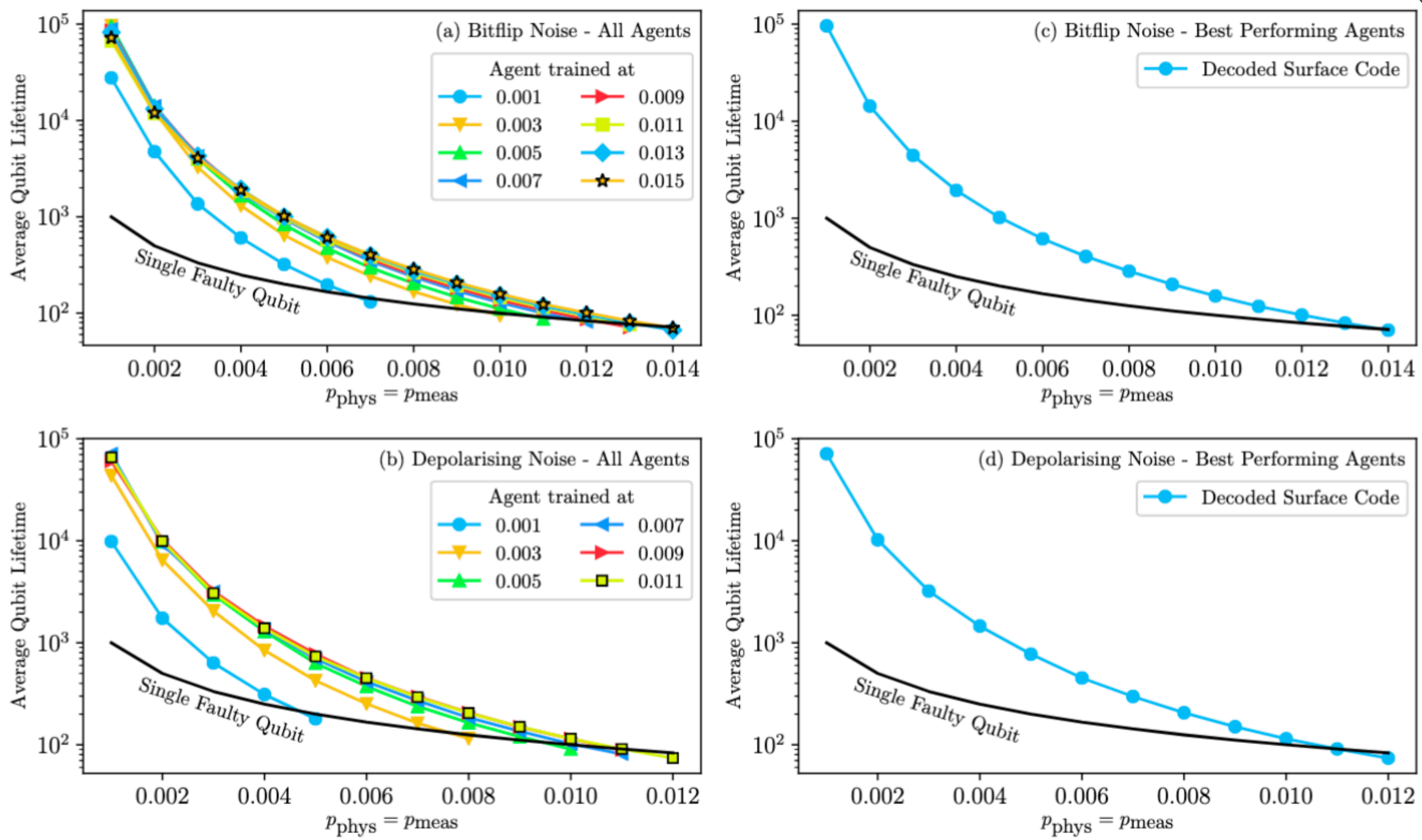
- Can create a loss function for iterative *Q-learning*

$$\begin{aligned}\text{loss} &= y_{\text{pred}} - y_{\text{true}} \\ &= q(S_t,A_t) - (R_{t+1} + \gamma \max_{a'} q(S_{t+1},a'))\end{aligned}$$

- *Deep* Q-learning seeks to parameterize Q with a neural network



- Results are very encouraging (bit-flip and depolarizing noise)



Sweke, Kesselring, van Nieuwenburg, Eisert, arXiv:1810.07207
See also Baireuther, Caio, Criger, Beenakker, O'Brien, New J Phys 21, 013003 (2019)

- **Lesson:** The fault tolerant decoding problem is much suitable to be tackled by reinforcement learning

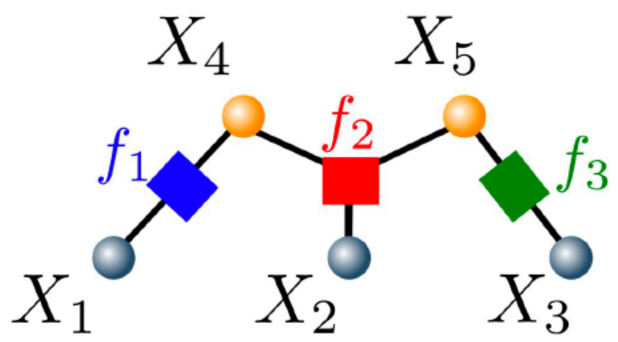
- Now: Multi-agent setting and RG decoders
- Gate-level noise
- Machine-learning for lattice surgery and fault tolerant computing

Sweke, Kesselring, van Nieuwenburg, Eisert, arXiv:1810.07207
Trotta, Sweke, Kesselring, van Nieuwenburg, Eisert, in preparation

OUTLOOK: TENSOR NETWORKS IN QUANTUM INSPIRED LEARNING

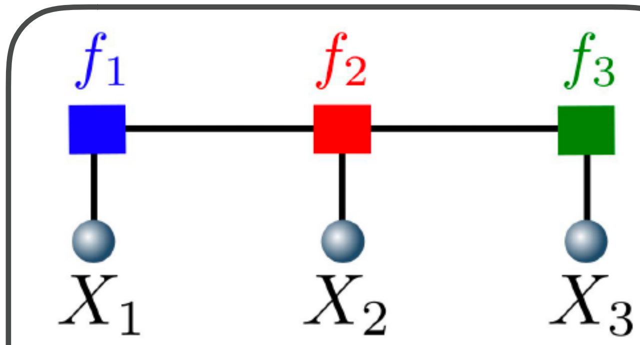
J Comput Nonlinear Dynam 14, 061006 (2019) and in progress

- Probabilistic graphical models



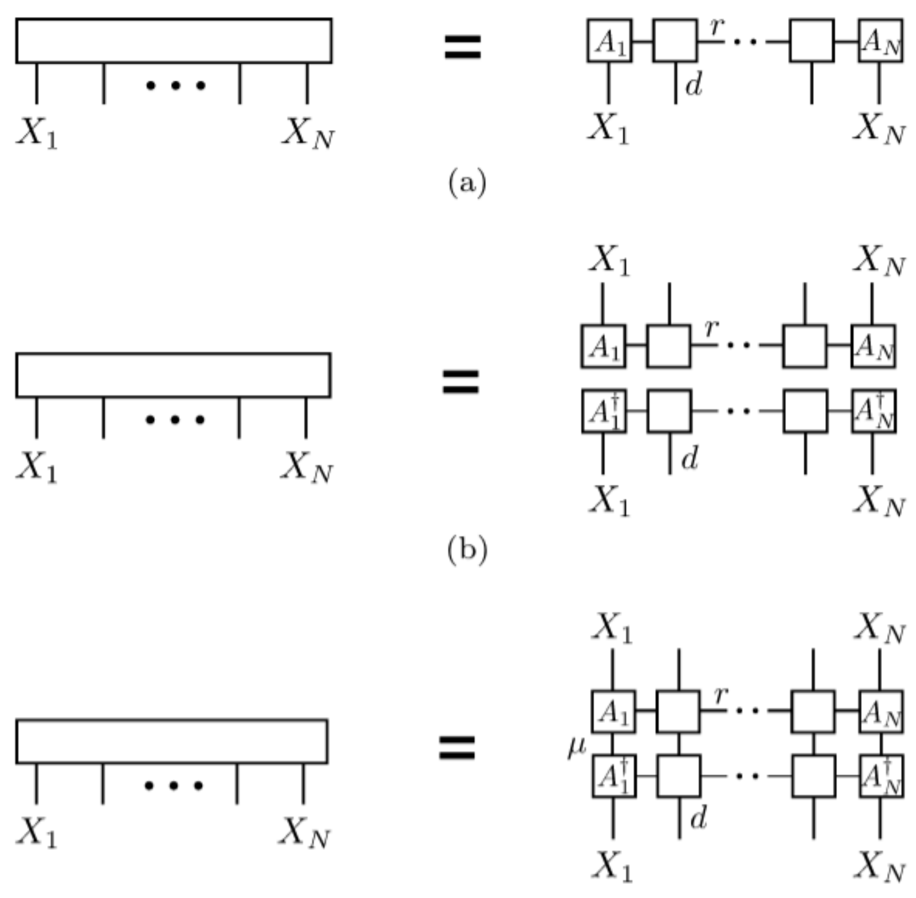
- Bayesian network/directed acyclic graphical model

- Tensor networks



Glasser, Pancotti, Cirac, arXiv:1806.05964
Many others

- **HOW CAN PROBABILITY DISTRIBUTIONS BE CAPTURED AS MPS?**



- MPS of bond dimension (BD) r, square of MPS of BD r, purification of BD r, lots ot other "ranks"

	$r_{\mathbb{R}}$	$r_{\mathbb{R}_{\geq 0}}$	$r_{\mathbb{R}^2}$	$r_{\mathbb{C}^2}$	$r_{\mathbb{R}_\infty}$	$r_{\mathbb{C}_\infty}$
$r_{\mathbb{R}}$	$=$	$\leq x$	$\leq x^2$	$\leq x^2$	$\leq x^2$	$\leq x^2$
$r_{\mathbb{R}_{\geq 0}}$	No	$=$	No	No	No	No
$r_{\mathbb{R}^2}$	No	No	$=$	No	No	No
$r_{\mathbb{C}^2}$	No	?	$\leq x$	$=$	?	?
$r_{\mathbb{R}_\infty}$	No	$\leq x$	$\leq x$	$\leq 2x$	$=$	$\leq 2x$
$r_{\mathbb{C}_\infty}$	No	$\leq x$	$\leq x$	$\leq x$	$\leq x$	$=$

- Comprehensive analysis of efficient embeddings into each other

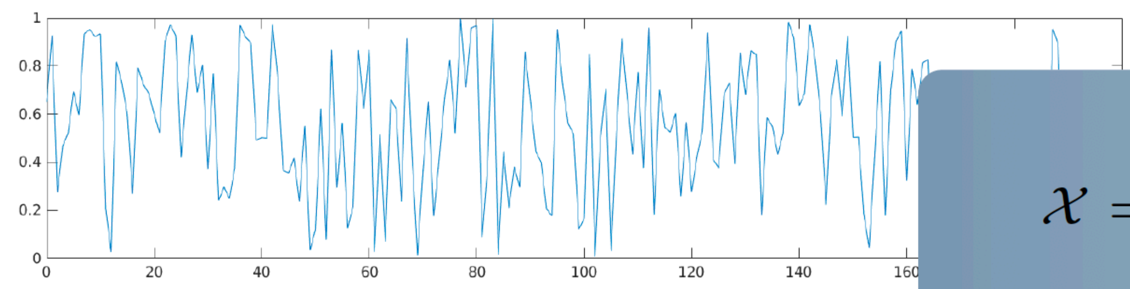
- **Lesson:** Expressivity of probability distributions with MPS can be largely addressed

Glasser, Pancotti, Cirac, Eisert, Sweke, in preparation

- **HOW CAN ONE LEARN DYNAMICAL LAWS FROM DATA?**

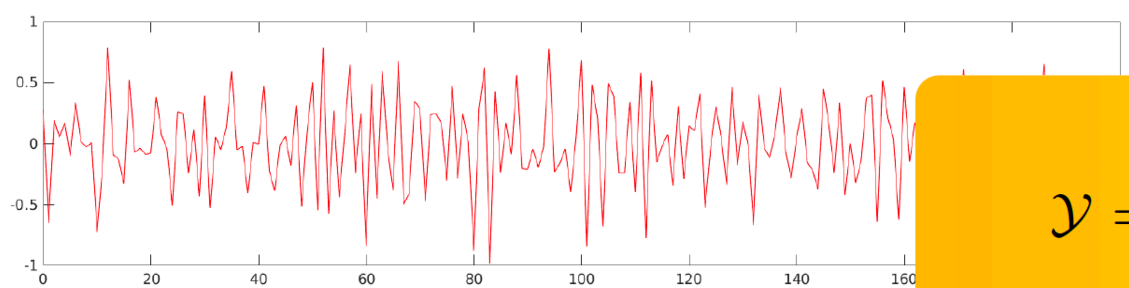
LEARNING DYNAMICAL LAWS FROM DATA

- Learn **dynamical laws** from data



$$\mathcal{X} = \begin{bmatrix} | & | & | & & | \\ X_1 & X_2 & X_3 & \dots & X_m \\ | & | & | & & | \end{bmatrix}$$

$$\frac{d}{dt} X(t) = F(X(t)) = Y(t)$$



$$\mathcal{Y} = \begin{bmatrix} | & | & | & & | \\ Y_1 & Y_2 & Y_3 & \dots & Y_m \\ | & | & | & & | \end{bmatrix}$$

- Express functions in terms of a **dictionary** $\{\psi_1, \ldots, \psi_p\}$

$$\Psi(X) = [\psi_1(X) \ldots \psi_p(X)]^T$$

- Transformed **data matrix**

$$\Psi(\mathcal{X}) = [\Psi_1(X_1) \ldots \Psi(X_m)]$$

- Determine the **coefficient matrix**

$$\Xi = [\xi_1 \ldots \xi_d]$$

such that the **cost function** is minimized

$$\|\mathcal{Y} - \Xi^T \Psi(\mathcal{X})\|_2 + \lambda \|\Xi\|_1$$

Brunton, Proctor, Kutz, Proc Natl Ac Sc 113, 3932 (2016)

LEARNING DYNAMICAL LAWS FROM DATA

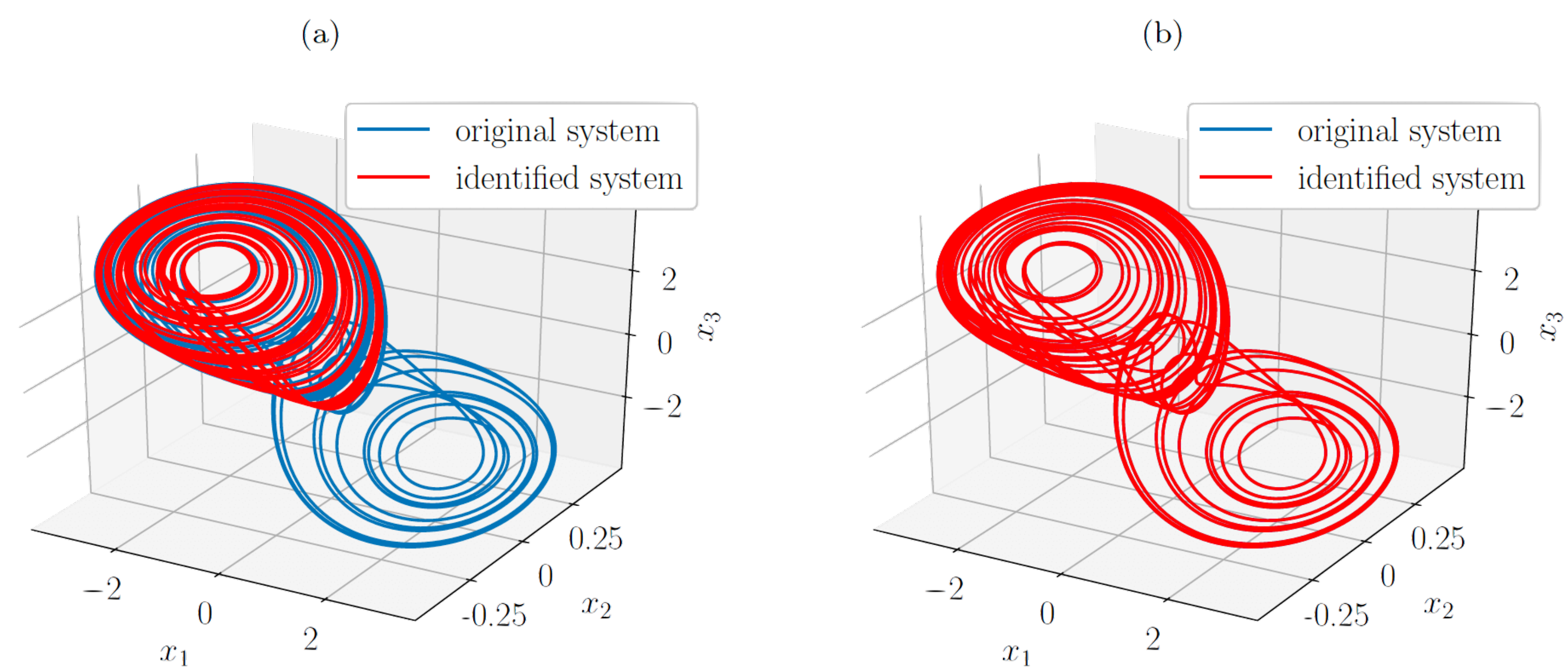
- Express
- Transfo
- Determ

such th

is minir

- **SINDy** finds a **sparse** coefficient matrix
- E.g., recovers Chua's circuit well

$$\dot{x}_1 = \alpha(x_2 - x_1 - g(x_1))$$
$$\dot{x}_2 = x_1 - x_2 + x_3$$
$$\dot{x}_3 = -\beta x_2$$

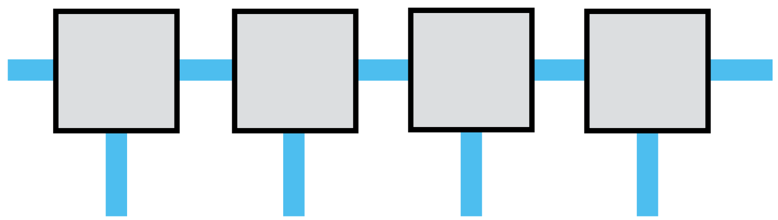


- H

CAN TENSOR NETWORK METHODS HELP?

Brunton, Proctor, Kutz, Proc Natl Ac Sc 113, 3932 (2016)

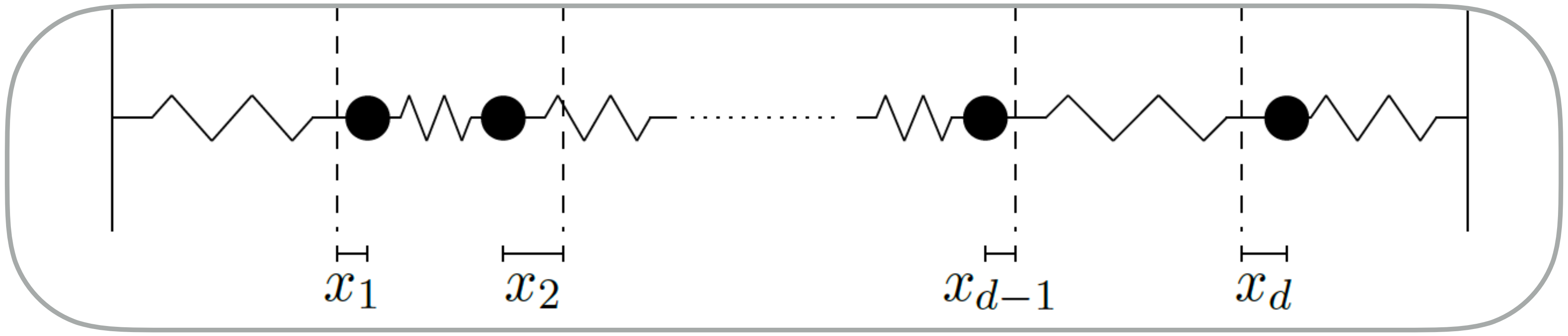
- **Idea:** Represent Ξ and $\Psi(\mathcal{X})$ as (real) **matrix-product states**



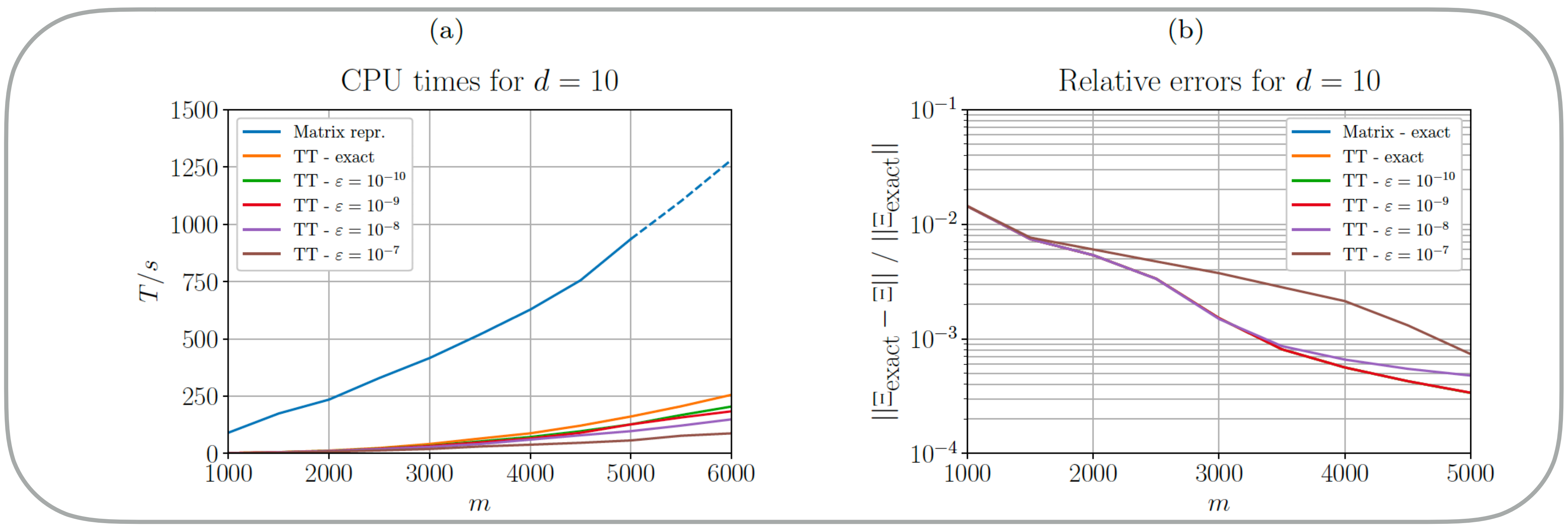
- Solve $\min_{\Xi} \|\mathcal{Y} - \Xi^T \Psi(\mathcal{X})\|_2$ directly in MPS space
- Steps such as **pseudo-inverse** can be efficiently computed, $O(mD^2)$ for m data points

Gelss, Klus, Eisert, Schuette, J Comput Nonlinear Dynam 14, 061006 (2019)

- **Works well in practice:** E.g., Fermi-Pasta-Ulam-Tsingou problem



with dictionary $\{1, x, x^2, x^3\}$ per site



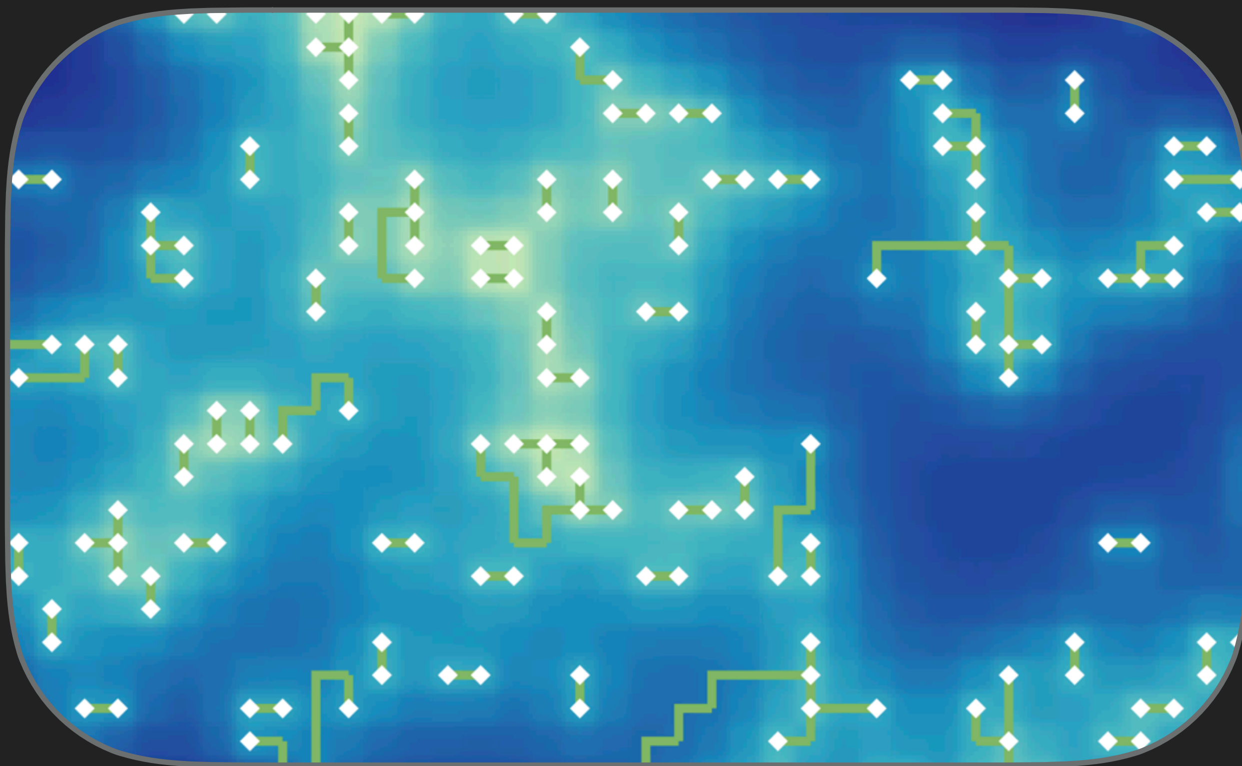
Gelss, Klus, Eisert, Schuette, J Comput Nonlinear Dynam 14, 061006 (2019)

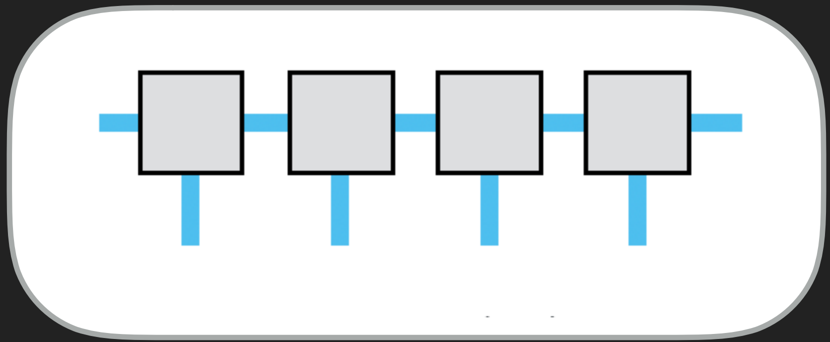
- **Lesson:** Tensor networks give encouraging results when learning dynamical laws

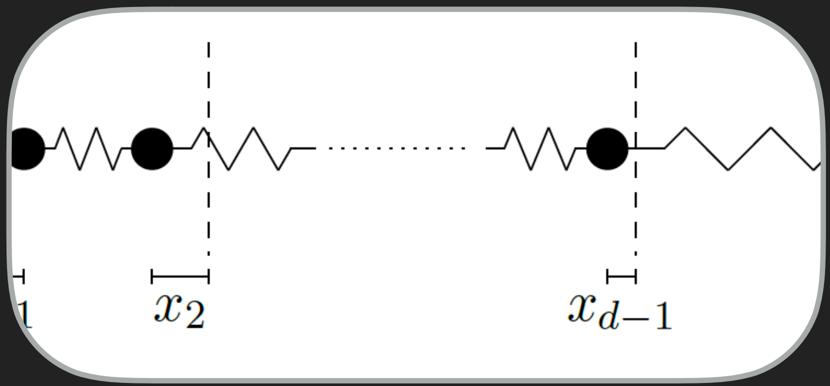
- **Trees**, **PEPS**, other tensor network dictionaries?
- Recovery **guarantees**, correlation measures?
- In progress with **Google**: *Quantum dynamical learning*

Gelss, Klus, Eisert, Schuette, J Comput Nonlinear Dynam 14, 061006 (2019)

SUMMARY







- Reinforcement learning can decode fault tolerant memories
- Tensor networks in learning

Thanks for your attention

http://www.physik.fu-berlin.de/en/einrichtungen/ag/ag-eisert/