EFFICIENT CLASSICAL SIMULATION OF RANDOM SHALLOW 2D QUANTUM CIRCUITS

John Napp (MIT)

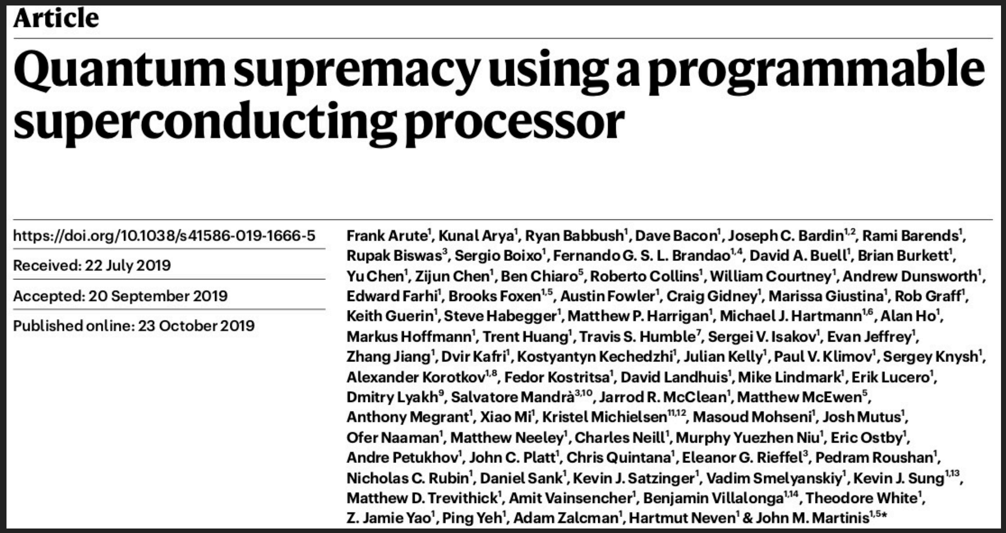
joint work with Rolando La Placa (MIT), Alex Dalzell (Caltech), Fernando Brandão (Caltech, Amazon), Aram Harrow (MIT)

QIP 2021
arXiv:2001.00021

MOTIVATION

"Nature isn't classical... and if you want to make a simulation of nature, you'd better make it quantum mechanical..." (1981)

When does this sentiment *not* hold?



Article

Quantum supremacy using a programmable superconducting processor

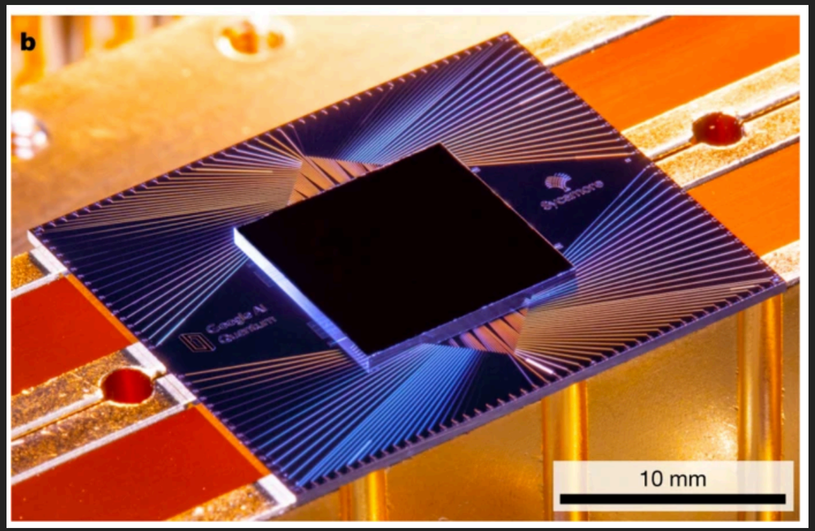
https://doi.org/10.1038/s41586-019-1666-5

Received: 22 July 2019

Accepted: 20 September 2019

Published online: 23 October 2019

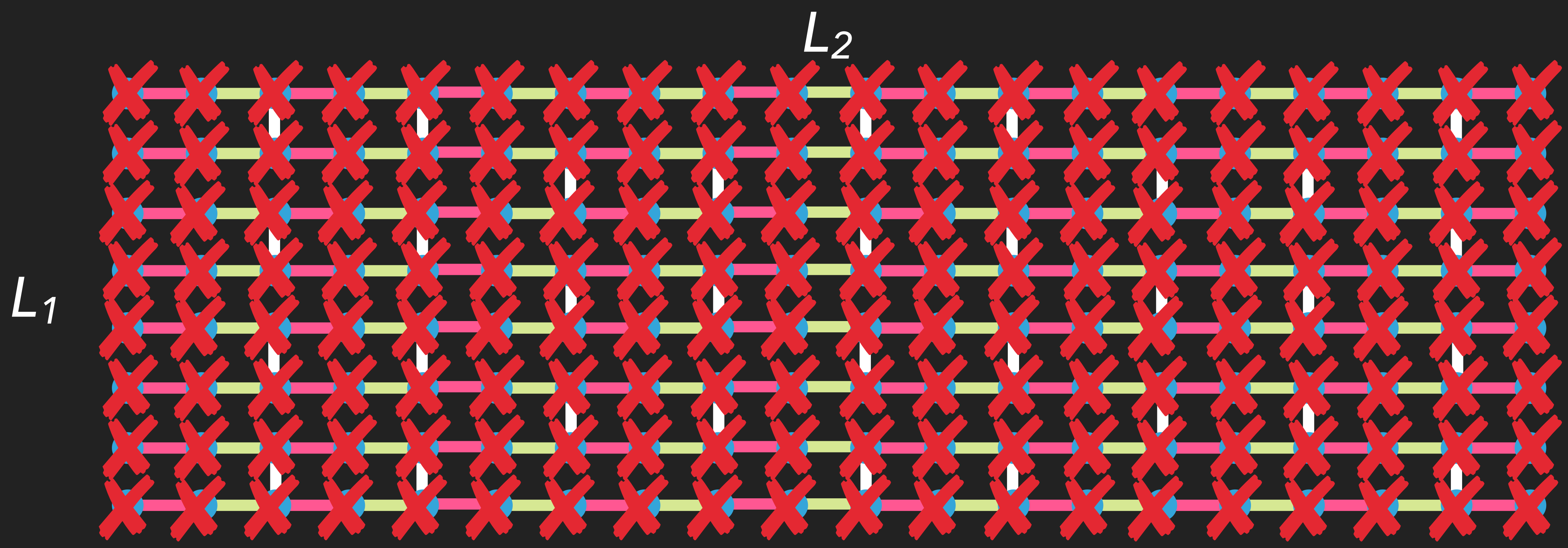
Frank Arute[1], Kunal Arya[1], Ryan Babbush[1], Dave Bacon[1], Joseph C. Bardin[1,2], Rami Barends[1], Rupak Biswas[3], Sergio Boixo[1], Fernando G. S. L. Brandao[1,4], David A. Buell[1], Brian Burkett[1], Yu Chen[1], Zijun Chen[1], Ben Chiaro[5], Roberto Collins[1], William Courtney[1], Andrew Dunsworth[1], Edward Farhi[1], Brooks Foxen[1,5], Austin Fowler[1], Craig Gidney[1], Marissa Giustina[1], Rob Graff[1], Keith Guerin[1], Steve Habegger[1], Matthew P. Harrigan[1], Michael J. Hartmann[1,6], Alan Ho[1], Markus Hoffmann[1], Trent Huang[1], Travis S. Humble[7], Sergei V. Isakov[1], Evan Jeffrey[1], Zhang Jiang[1], Dvir Kafri[1], Kostyantyn Kechedzhi[1], Julian Kelly[1], Paul V. Klimov[1], Sergey Knysh[1], Alexander Korotkov[1,8], Fedor Kostritsa[1], David Landhuis[1], Mike Lindmark[1], Erik Lucero[1], Dmitry Lyakh[9], Salvatore Mandrà[3,10], Jarrod R. McClean[1], Matthew McEwen[5], Anthony Megrant[1], Xiao Mi[1], Kristel Michielsen[11,12], Masoud Mohseni[1], Josh Mutus[1], Ofer Naaman[1], Matthew Neeley[1], Charles Neill[1], Murphy Yuezhen Niu[1], Eric Ostby[1], Andre Petukhov[1], John C. Platt[1], Chris Quintana[1], Eleanor G. Rieffel[3], Pedram Roushan[1], Nicholas C. Rubin[1], Daniel Sank[1], Kevin J. Satzinger[1], Vadim Smelyanskiy[1], Kevin J. Sung[1,13], Matthew D. Trevithick[1], Amit Vainsencher[1], Benjamin Villalonga[1,14], Theodore White[1], Z. Jamie Yao[1], Ping Yeh[1], Adam Zalcman[1], Hartmut Neven[1] & John M. Martinis[1,5]*



Is random circuit sampling truly classically intractable?

RANDOM CIRCUIT SAMPLING

- Start with array of n qudits
- Apply circuit C: d layers of 2-local Haar-random gates, followed by computational basis measurements on all sites
- RCS task: approximately (1/poly(n) additive error) sample from output distribution D_C with high probability over C



Example: depth-3 “brickwork architecture”

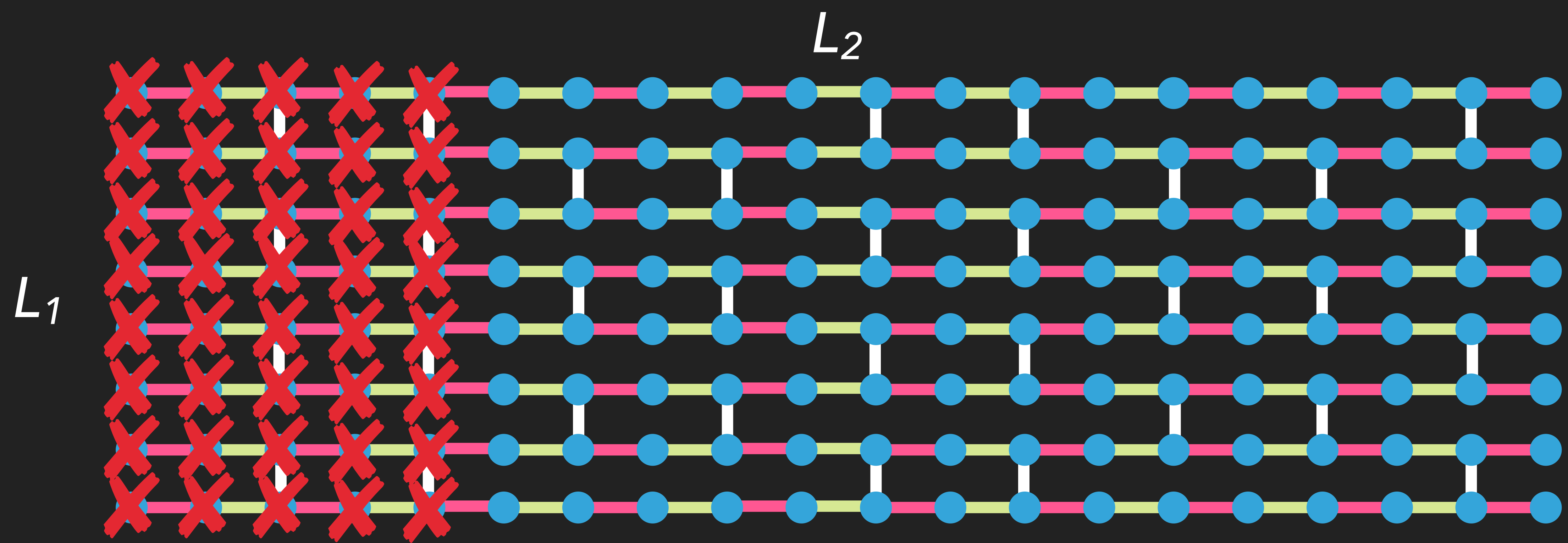
HOW HARD IS RCS?

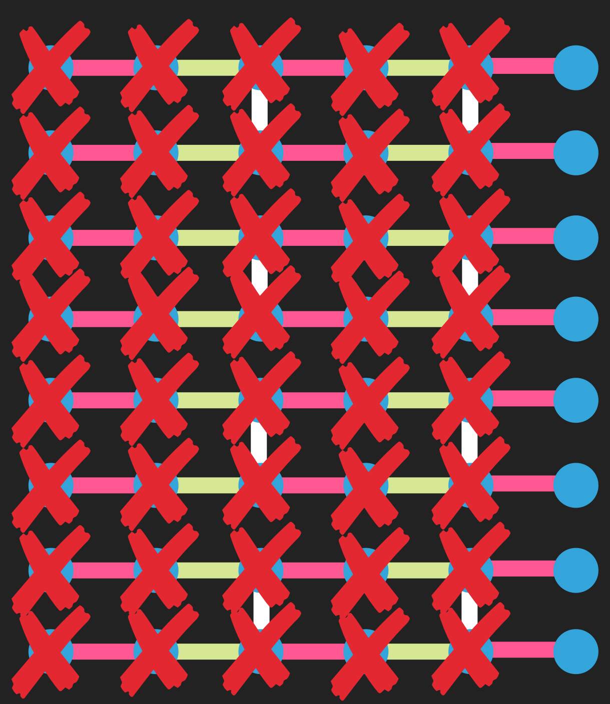
- Assuming non-collapse of PH, for algorithms that make exponentially small error…
  - Sampling is hard in the worst case [Terhal-DiVincenzo `02]
  - Computing *output probabilities* is hard in the average case [Bouland-Fefferman-Nirkhe-Vazirani `18]
- Is this evidence that RCS is hard?

RESULTS SUMMARY

- Propose two classical algorithms for RCS for constant-depth 2D circuits.
- Provable
  - There exists a specific 2D depth-3 random circuit architecture for which the previous hardness results apply, but RCS is classically tractable.
- Conjectured with evidence
  - RCS is classically tractable for sufficiently shallow 2D circuits more generally.
  - Our algorithms transition from poly-time to exponential-time discontinuously when the circuit depth or qudit dimension exceeds some (architecture-dependent) critical value.

SIMULATION VIA A 2D-TO-1D REDUCTION





- Need only simulate an "effective 1D dynamics":

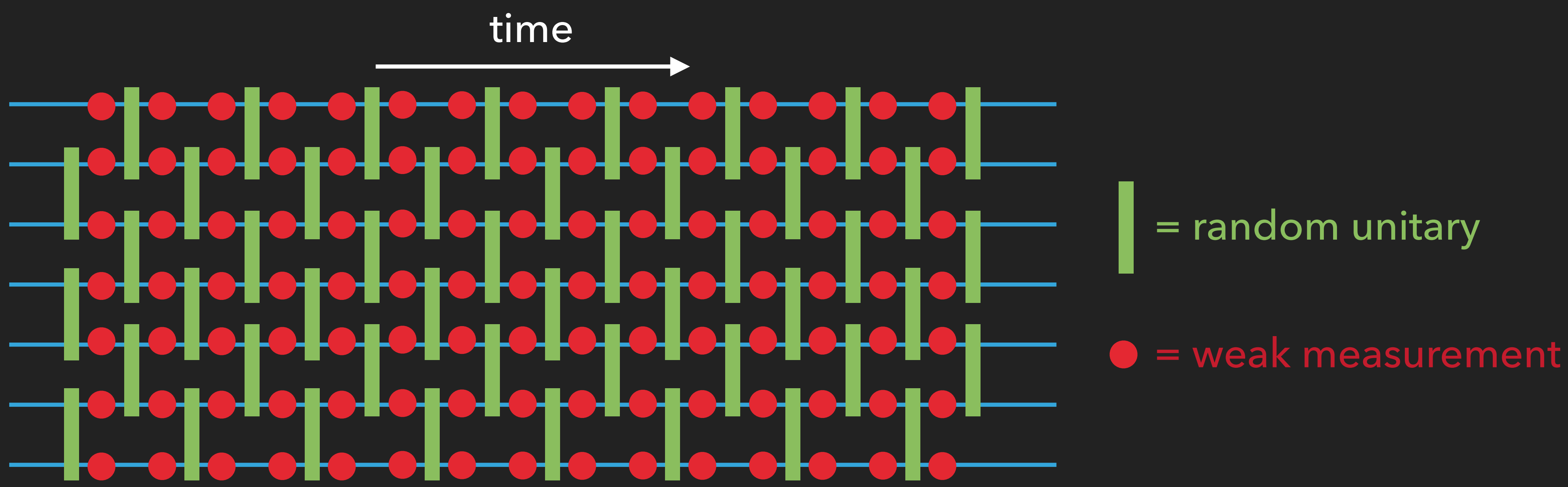
$$|\psi_1\rangle \rightarrow |\psi_2\rangle \rightarrow \cdots \rightarrow |\psi_{L_2}\rangle$$

- If effective 1D process is only lightly entangled, it may be approximately simulated via MPS [Vidal `03]

- Provably works for some architectures; conjectured to work more generally for sufficiently shallow circuits

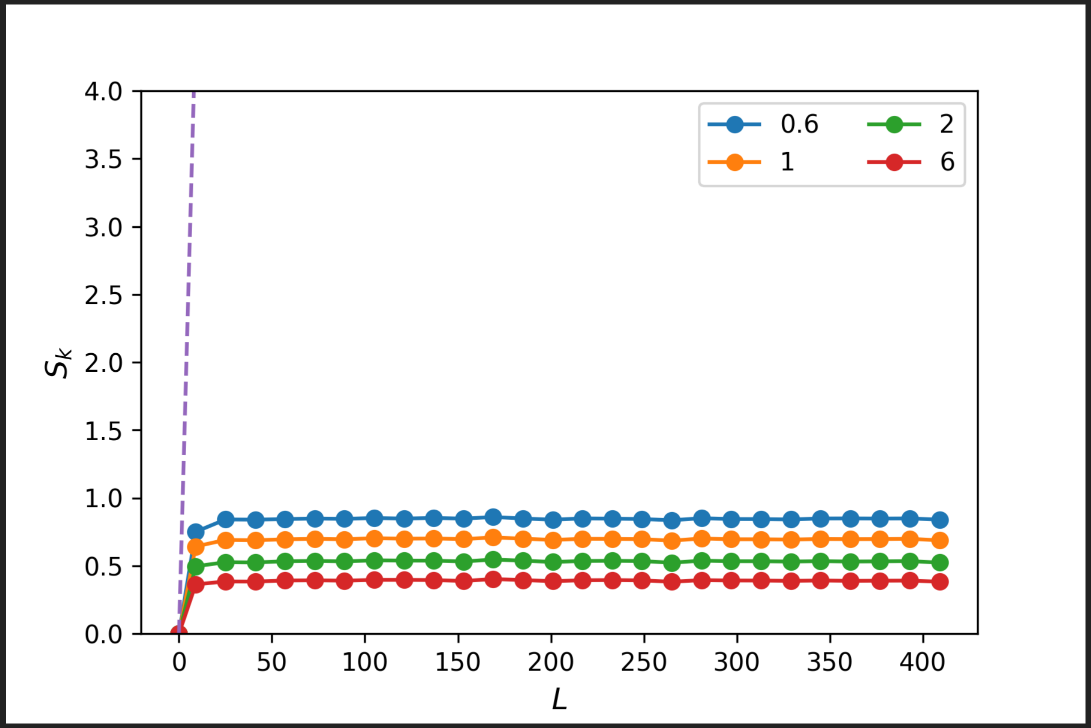
INTUITION: AVERAGE-CASE EFFICIENCY FROM RANDOM-UNITARY-AND-MEASUREMENT CIRCUITS

- Effective 1D dynamics involves alternating rounds of random gates and measurements on subsets of qudits.
- Similar to ``random-unitary-and-measurement'' processes studied in an explosion of recent works (beginning with [Skinner et al., Li et al., Chan et al. 2018])



- Entanglement phase transition from area-law to volume-law phase as measurement strength is decreased or local dimension is increased. Existence of phase transition appears to be robust w.r.t. details of the model.
- Indicates a potential complexity phase transition for classical simulation
  - Heuristically, increasing circuit depth is associated with reducing measurement strength in the associated effective 1D dynamics

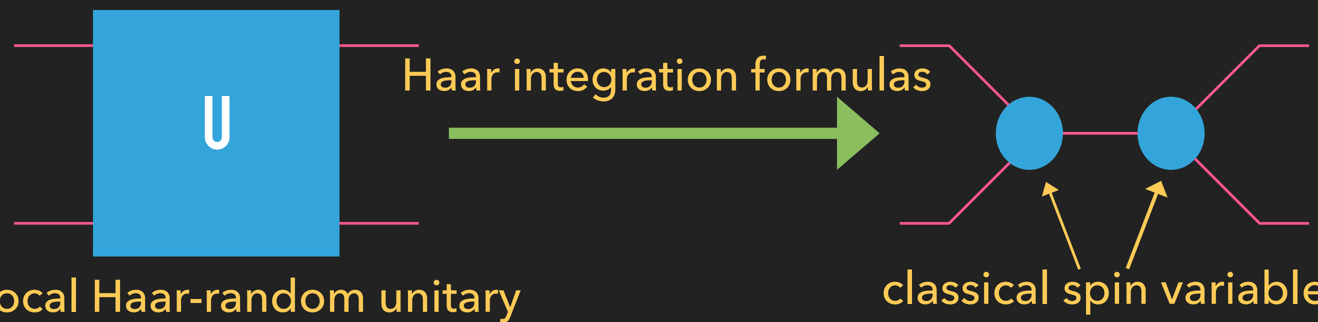
NUMERICAL IMPLEMENTATION



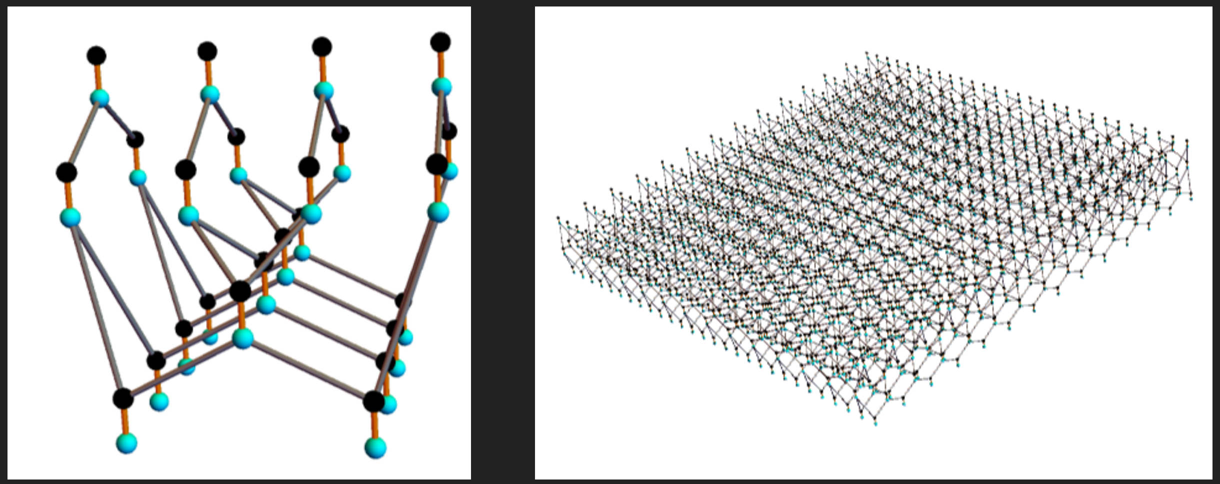
Half-chain Rényi entanglement entropies observed in effective 1D dynamics associated with brickwork architecture, as a function of grid side-length

- Can easily simulate typical instances of the 400 x 400 (ordinary) brickwork architecture with small error on a laptop
  - Algorithm is *self-certifying*: can measure its own sampling error
- Scaling of entanglement spectrum indicates algorithm remains efficient asymptotically for architectures we considered
  - Empirical entanglement spectrum consistent with toy model we analyze

UNDERSTANDING THE ALGORITHM VIA STAT. MECH.



Entanglement properties of the output state of the circuit may be associated with physical properties of an associated classical stat mech model [Hayden et al. `16], [Nahum, Vijay, Haah `17]



mapping applied to a shallow 2D circuit yields quasi-2D Ising-like model

- Disordered (ordered) phase of model corresponds to area law (volume law) for a proxy for entanglement entropy in effective 1D dynamics
- Brickwork architecture's associated stat mech model is *disordered*
- Increasing qudit dimension or depth increases effective interaction strength

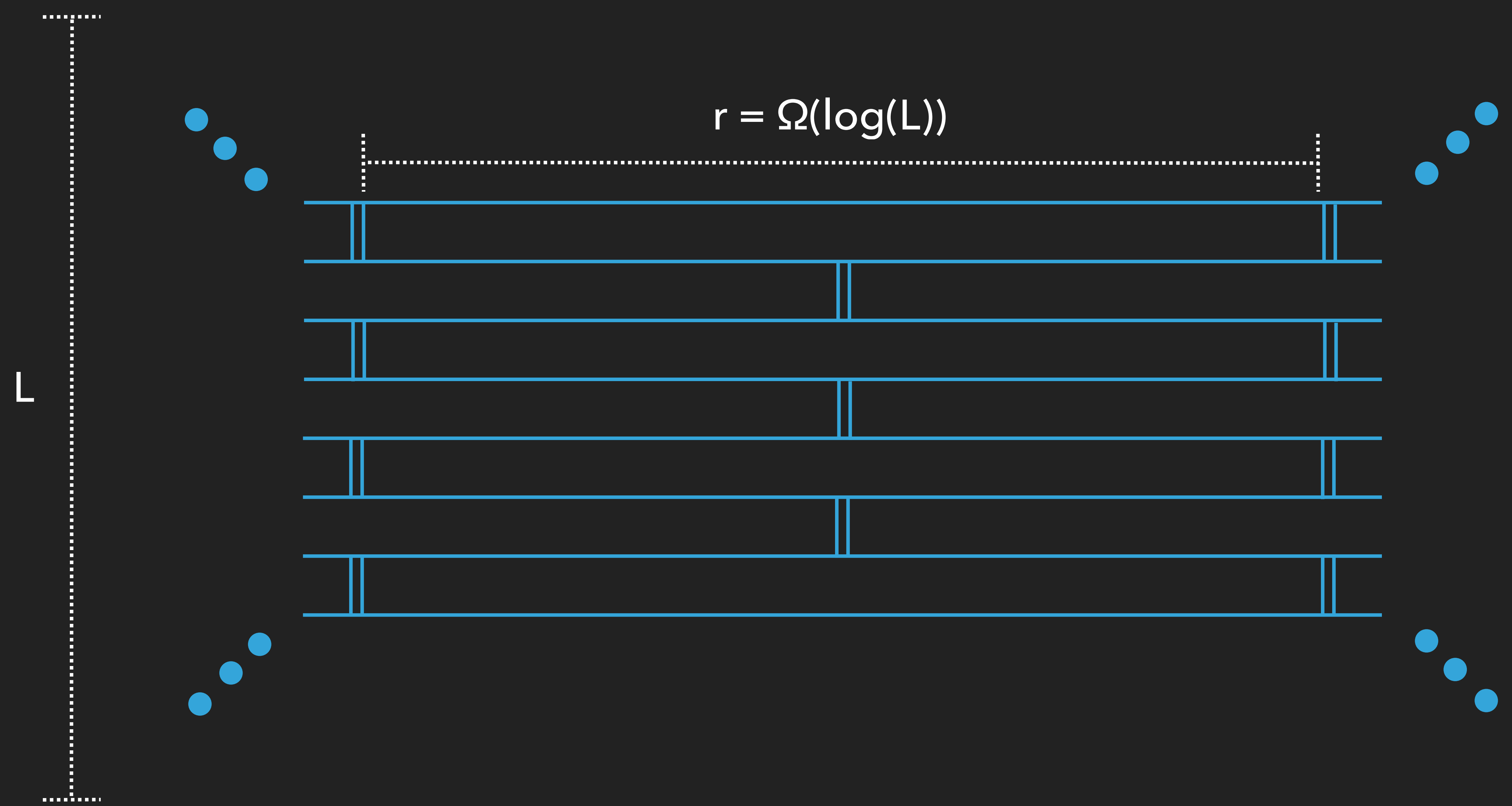
SOME OPEN QUESTIONS

- Can deep 2D random circuits be simulated in poly time?
- Can more general efficiency and phase transition be rigorously proven?
- Can Monte Carlo studies of the associated stat mech model be used to predict the runtime of the algorithm?
- How much easier are noisy circuits?
- For “volume-law” circuit instances, is this algorithm still better than other known exponential-time algorithms?
- Continuous-time versions of results?

Thank you!

EXTENDED BRICKWORK ARCHITECTURE

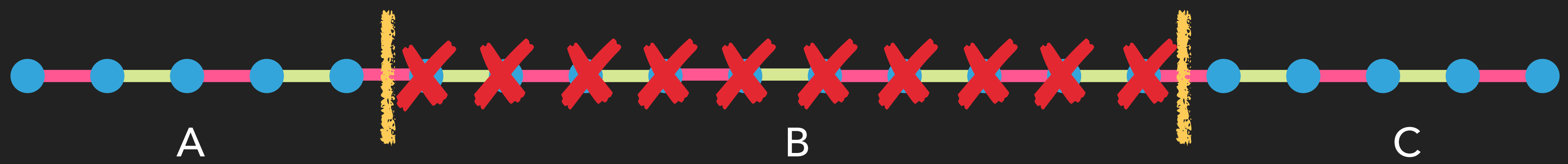
- Modify the ordinary depth-3 brickwork architecture such that vertical gates are "sparse".



- Idea for rigorous efficiency proof: amount of entanglement generated in effective 1D dynamics is exponentially suppressed in *r*
- But, this architecture is non-uniform

ALG. CAN EFFICIENTLY APPROXIMATELY SIMULATE THE E.B. ARCHITECTURE IN THE AVERAGE CASE

- Key ingredient: exponential decay of post-measurement entanglement entropy for 1D random circuits



Measurement outcome: b

Conditional post-measurement state on AC: $\psi_{AC|b}$

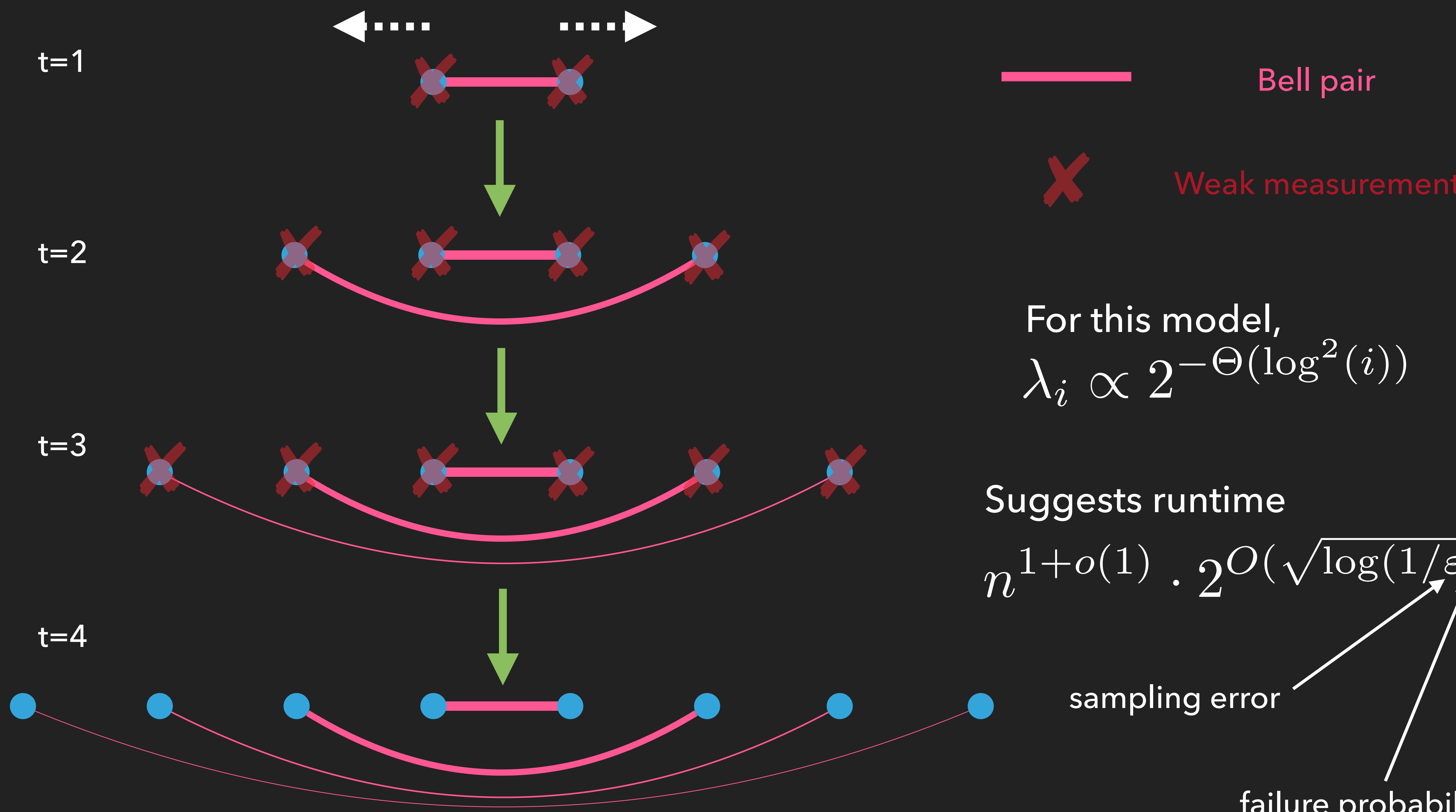
$$\mathbb{E}_{U,b} S(A)_{\psi_{AC|b}} \leq c^{|B|}, \;\; c < 1$$

- As algorithm simulates the 1D dynamics, long stretches without any vertical gates destroy most entanglement in the 1D state with high probability.

- With high probability over circuit realization, algorithm can sample from D_C with small additive error

TOY MODEL FOR AREA LAW PHASE

- To understand the scaling of entanglement spectrum across some cut for area-law dynamics, study simple toy model:



For this model,

$$\lambda_i \propto 2^{-\Theta(\log^2(i))}$$

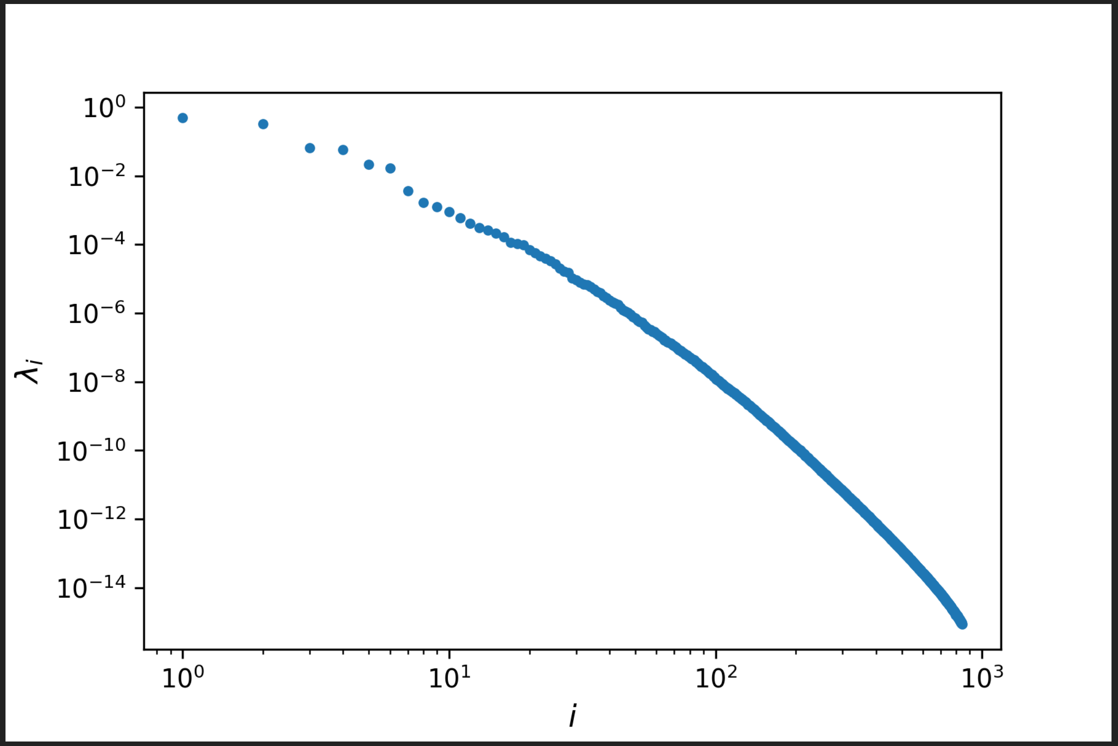
Suggests runtime

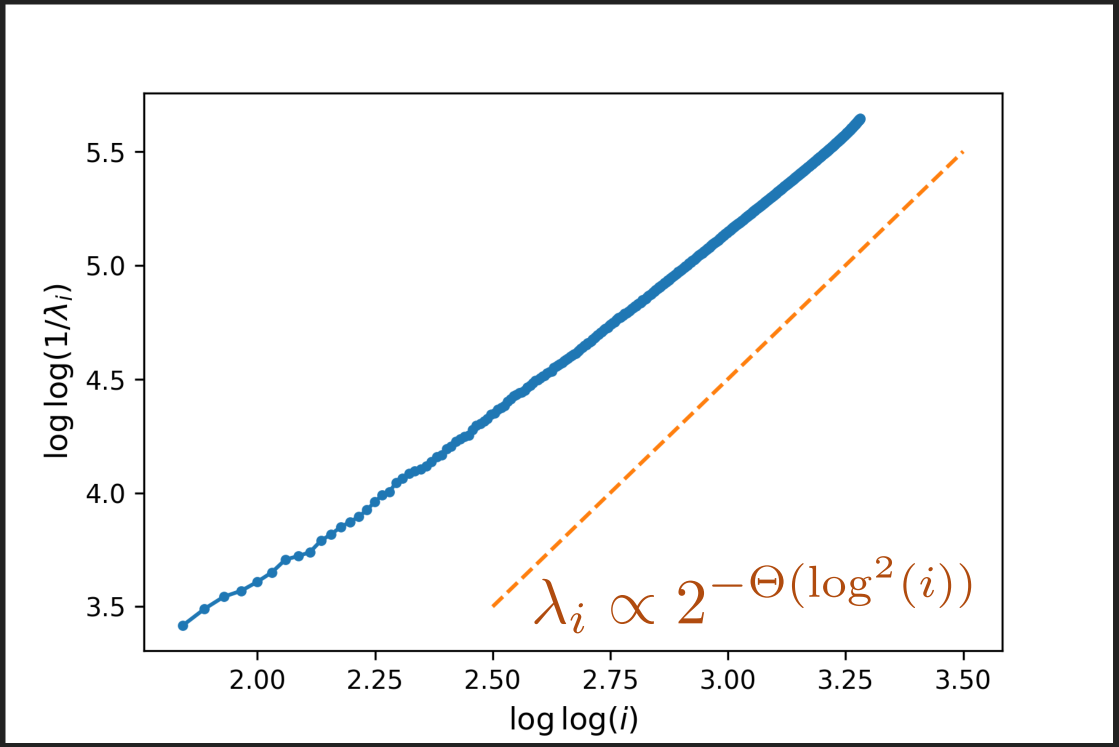
$$n^{1+o(1)} \cdot 2^{O(\sqrt{\log(1/\varepsilon\delta)})}$$

sampling error

failure probability

RAPID DECAY OF SCHMIDT COEFFICIENTS





Typical entanglement spectrum observed in effective 1D dynamics
(for the "cluster state with Haar-random measurements" model)

- Superpolynomial decay suggests efficient, low error MPS compression is possible.
- We study a toy model for which

$$\lambda_i \propto 2^{-\Theta(\log^2(i))}$$

Toy model suggests runtime

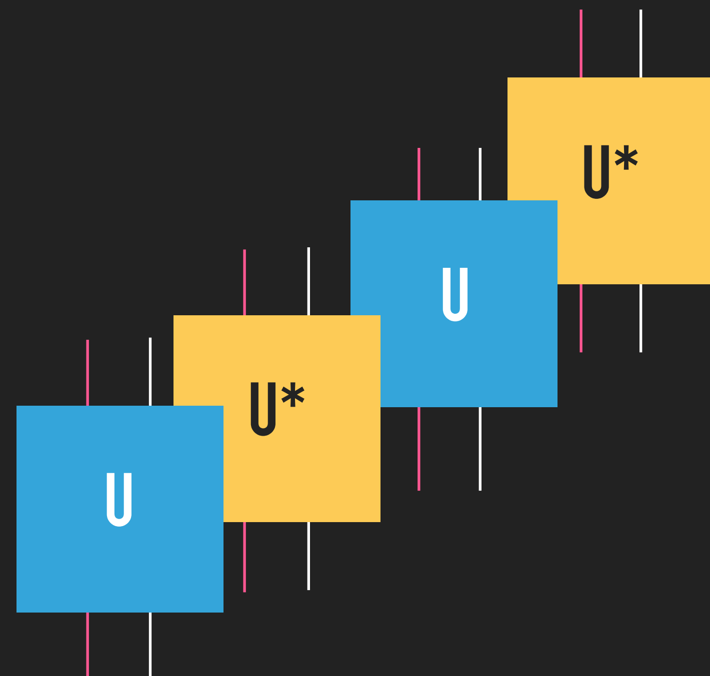
$$n^{1+o(1)} \cdot 2^{O(\sqrt{\log(1/\varepsilon\delta)})}$$

sampling error

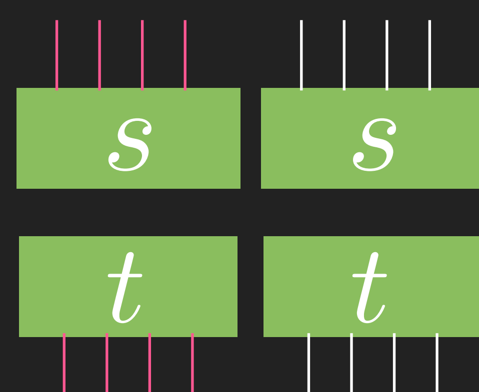
failure probability

RANDOM CIRCUITS AND STATISTICAL MECHANICS

$$\int dU$$



$$= \sum_{s\in\{1,-1\}} \sum_{t\in\{1,-1\}} Wg(q^2; st)$$



"Weingarten function"
[Collins `03]

[Nahum, Vijay, Haah `17]:

1. Perform this integration over all random gates.
2. Interpret s and t as classical spin variables.
3. The expected purity of some subregion A, averaged over random gates, may be expressed as the partition function of an Ising-like model

Previously, mapping from random tensor networks to stat mech models was used to study holographic duality [Hayden et al `16]

SELF-CERTIFICATION

- Fix a circuit realization C
- Fix a truncation error per iteration ε and bond dimension cutoff D_{cutoff}
  - In each iteration, for each bond, discard Schmidt coefficients up to a max error of ε. If a bond dimension > D_{cutoff}, declare failure
  - What is the relation between the true output distribution D_C and the output distribution D'_C sampled?

$$\frac{1}{2}\|\mathcal{D}'_C - \mathcal{D}_C\|_1 \leq n\epsilon + p_{f,C}$$

failure probability

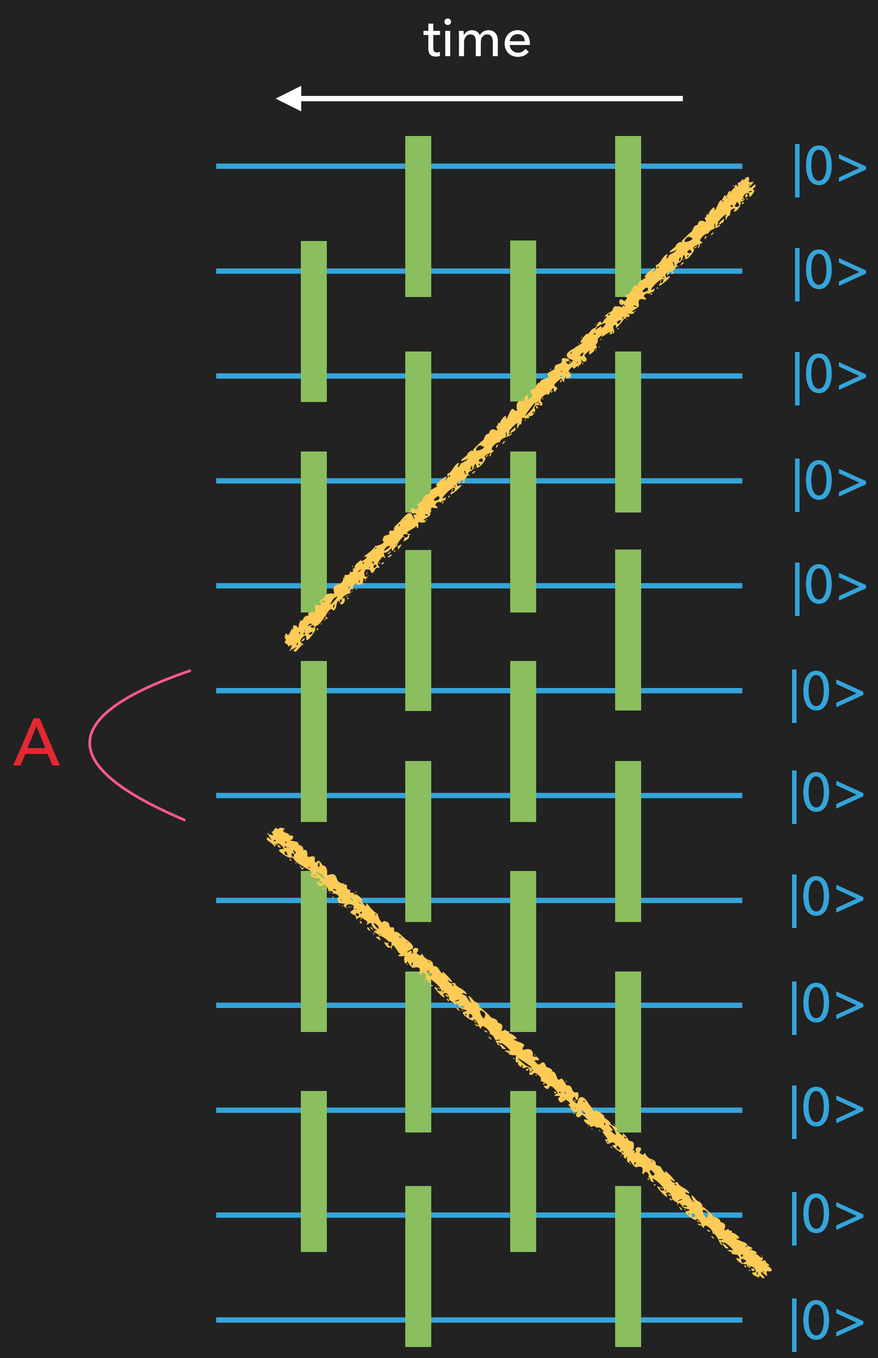
- Algorithm is self-certifying: run it many times to construct a confidence interval for $p_{f,C}$ and therefore the sampling error

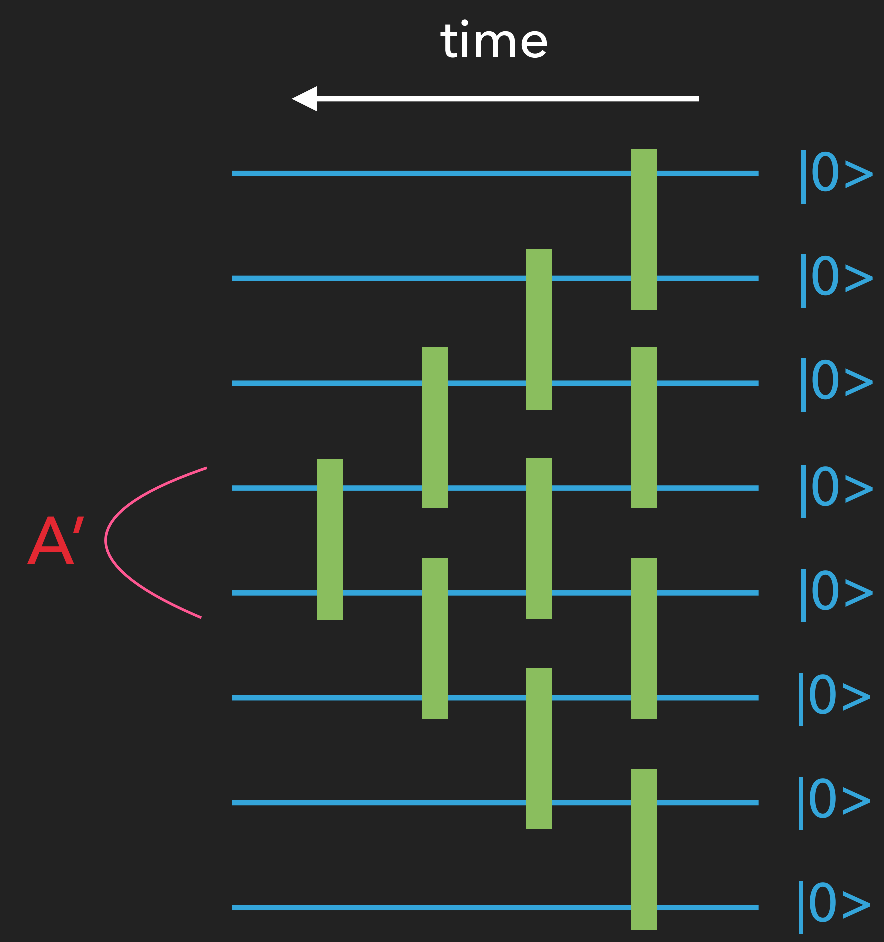
ASYMPTOTIC HARDNESS OF 2D RANDOM CIRCUIT SAMPLING

	sufficiently shallow O(1)-depth	sufficiently deep O(1)-depth	Ω(log n)-depth
noiseless	✔ [this work]	?	? conjectured hard [Boixo et al. \`16]
noisy (e.g. single-qubit depolarizing noise occurring at constant rate)	✔ [this work]	?	✔ [Gao-Duan \`18]

- ✔ = approximate poly-time classical simulation exists. Not necessarily efficient in practice!

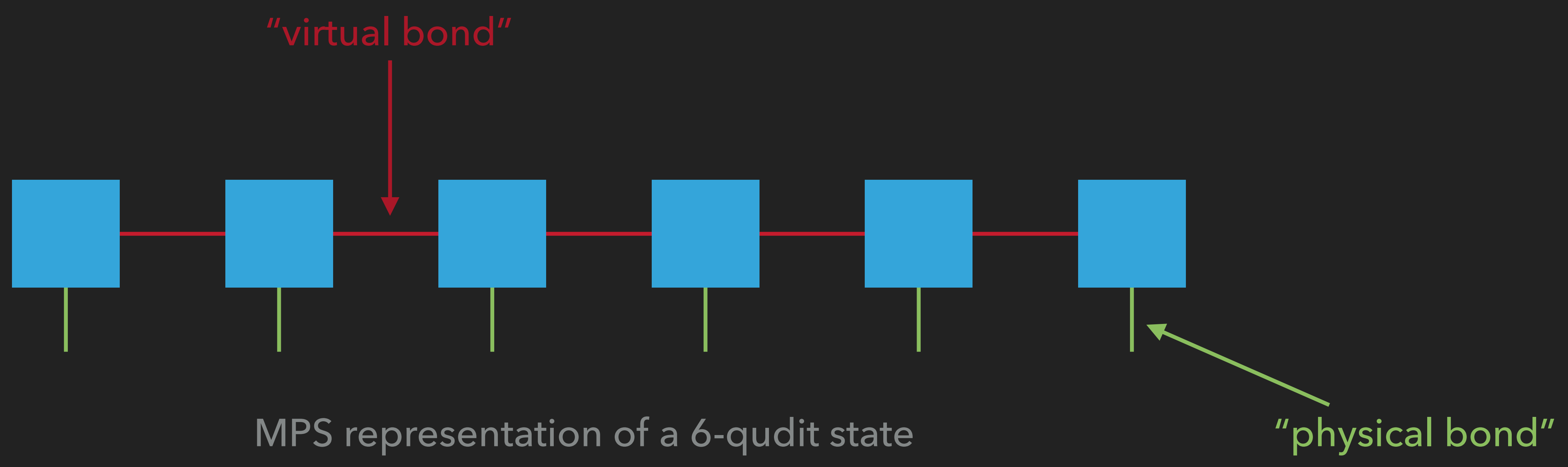
LIGHTCONE ARGUMENT





$$\rho_A = \rho_{A'}$$

MATRIX PRODUCT STATE (MPS) REVIEW

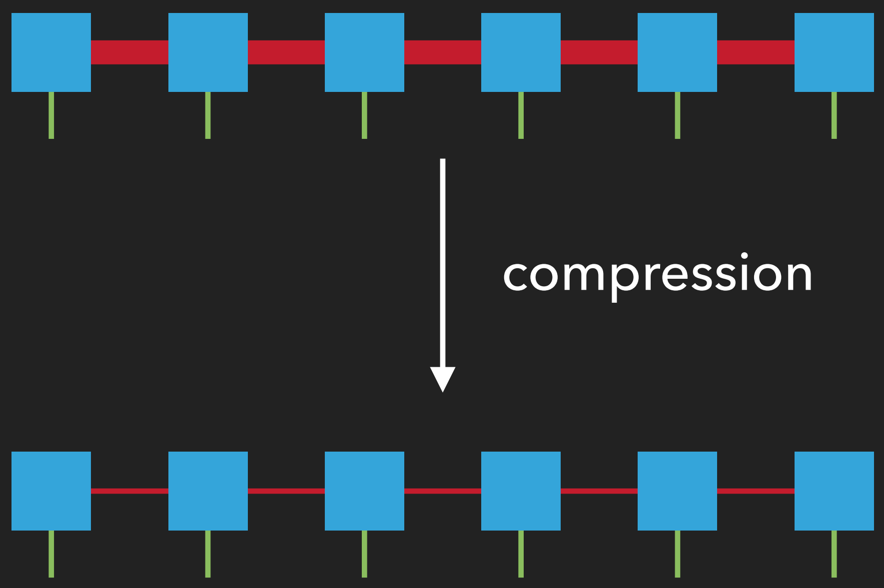


MPS representation of a 6-qudit state

- Bond dimension D: maximal dimension of a virtual bond
- MPS described by a number of parameters polynomial in number of qudits, qudit local dimension, and D
- Can be compressed and manipulated efficiently
- Can only efficiently represent low entanglement states

MATRIX PRODUCT STATE (MPS) REVIEW

- If D is maximum Schmidt rank of a 1D state across any cut, can classically represent with MPS with poly(L, D) parameters.
- Can simulate gate applications & measurements in poly(L, D) time



- Compression: reduce D by truncating smallest Schmidt coefficients across cuts.
  - Error incurred is related to total weight of discarded coefficients
- Efficient approximation by MPS often possible when entanglement entropy obeys area law (bounded by constant across all cuts)

PATCHING ALGORITHM