

Coherent Ising Machines: non-von Neumann computing using networks of optical parametric oscillators

Peter McMahon

Cornell University

CMU 47-779 *Quantum Integer Programming*

6 October 2020

Credits

Alireza Marandi, Zhe Wang, Kenta Takata, Robert L. Byer, Yoshihisa Yamamoto. “Network of time-multiplexed optical parametric oscillators as a coherent Ising machine.” *Nature Photonics* **8**, 937-942 (2014).

Peter L. McMahon*, Alireza Marandi*, Yoshitaka Haribara, Ryan Hamerly, Carsten Langrock, Shuhei Tamate, Takahiro Inagaki, Hiroki Takesue, Shoko Utsunomiya, Kazuyuki Aihara, Robert L. Byer, M. M. Fejer, Hideo Mabuchi, Yoshihisa Yamamoto. “**A fully programmable 100-spin coherent Ising machine with all-to-all connections.**” *Science* **354**, No. 6312, 614 - 617 (2016).

Ryan Hamerly*, Takahiro Inagaki*, Peter L. McMahon* *et al.* “**Experimental investigation of performance differences between coherent Ising machines and a quantum annealer.**” *Science Advances* **5**, 5, eaau0823 (2019).

The topic of this talk:

A physics-based **computing**
machine that provides a novel and
scalable approach to **solving**
difficult **optimization problems**.

Overview

• The Ising Problem

- Ising Machines
- Foundations of All-Optical OPO Ising Machines
- Measurement-Feedback OPO Ising Machines
- Conclusions

The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$

The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$

This is an **NP-hard problem***, and is **difficult to solve in practice** for moderate-size N .

→ Approximate (heuristic) solvers take hours when $N=10,000$.

* Relation to MAX-CUT will be shown later.

The Ising Problem

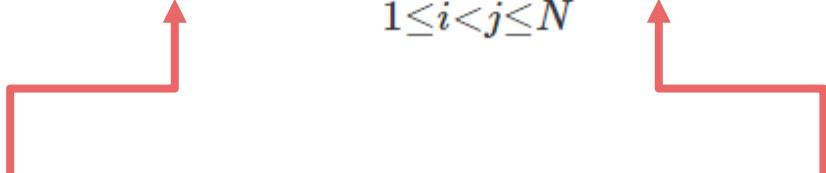
(Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

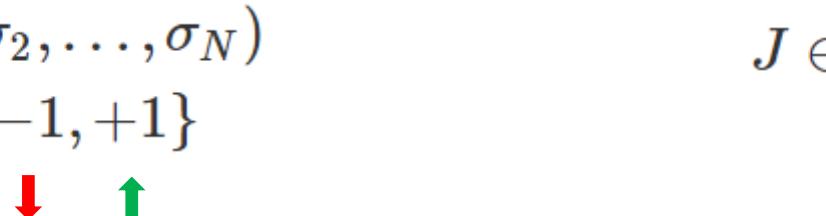
$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$

$\vec{\sigma} \triangleq (\sigma_1, \sigma_2, \dots, \sigma_N)$

$\sigma_k \in \{-1, +1\}$



$J \in \mathbb{R}^{N \times N}$



The Ising Problem

(Combinatorial Optimization Version)

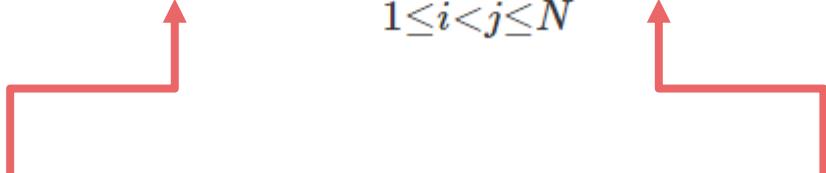
Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$

$\vec{\sigma} \triangleq (\sigma_1, \sigma_2, \dots, \sigma_N)$

$\sigma_k \in \{-1, +1\}$

$J \in \mathbb{R}^{N \times N}$

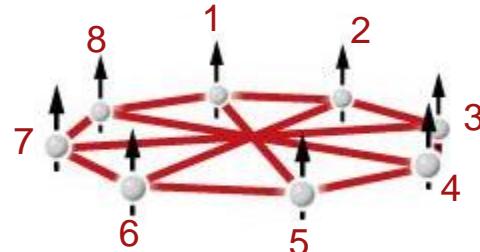
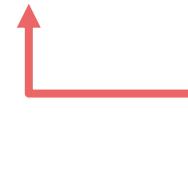


- 1D Ising model with nearest-neighbor connections
- 2D Ising model with nearest-neighbor connections
- Generalized Ising model with arbitrary connections

The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$



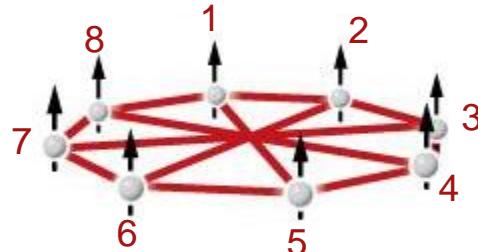
$$J = - \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The Ising Problem

(Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$

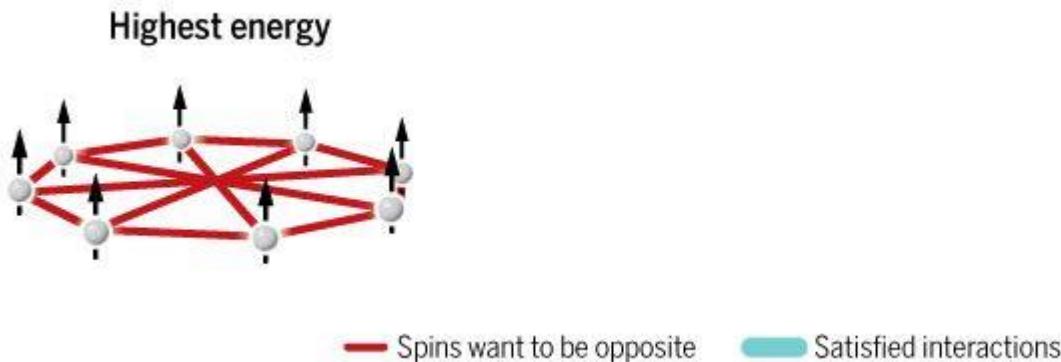


$$J = - \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

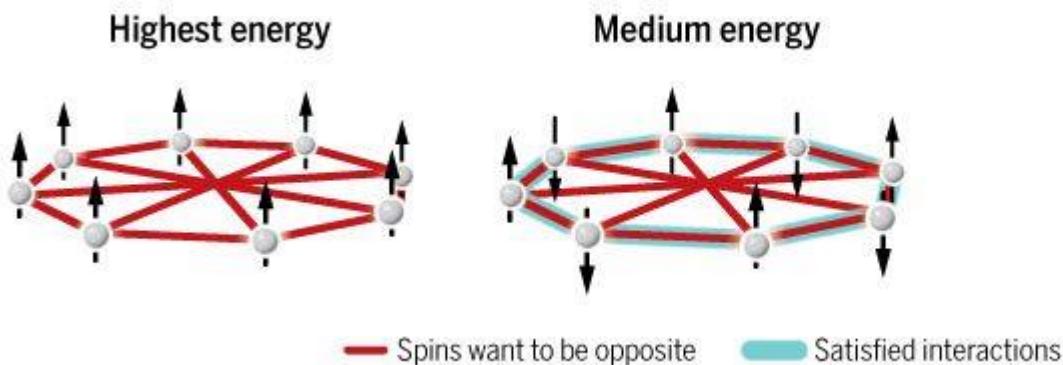
$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$



The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

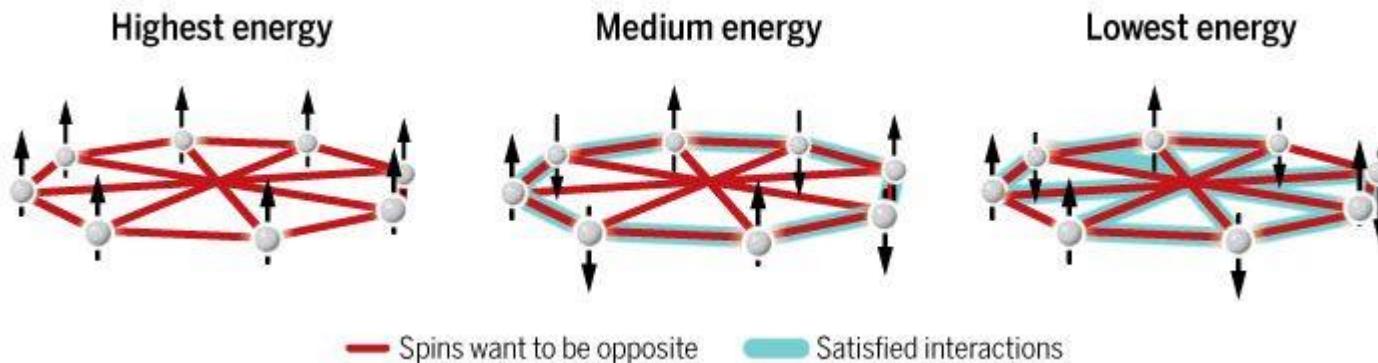
$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$



The Ising Problem (Combinatorial Optimization Version)

Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j$$



Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:

Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:
 - Planning/scheduling problems

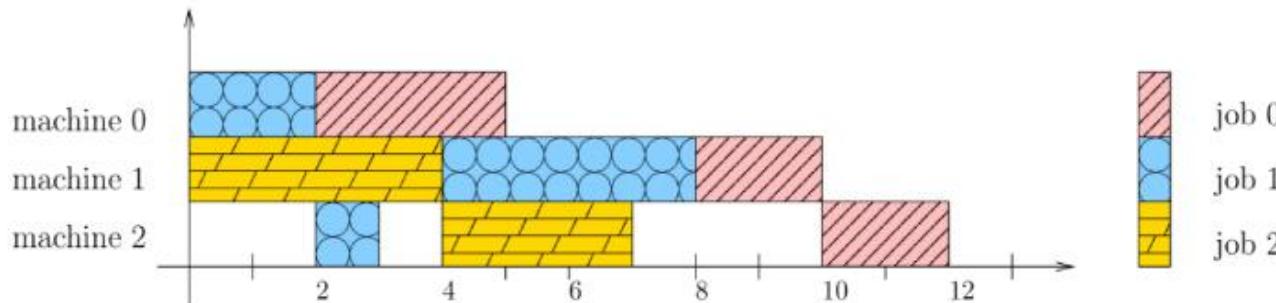
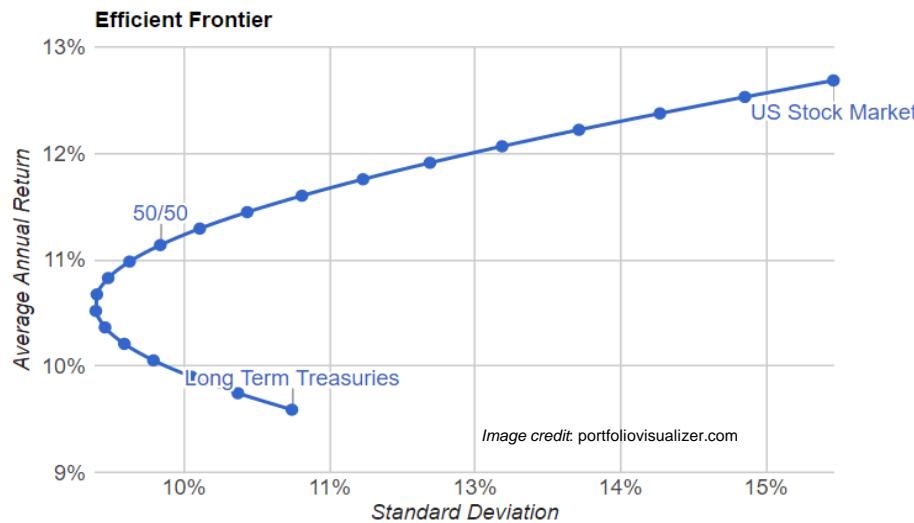


Image credit: developer.google.com

Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:
 - Planning/scheduling problems
 - Portfolio optimization



Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:
 - Planning/scheduling problems
 - Portfolio optimization
 - Protein folding

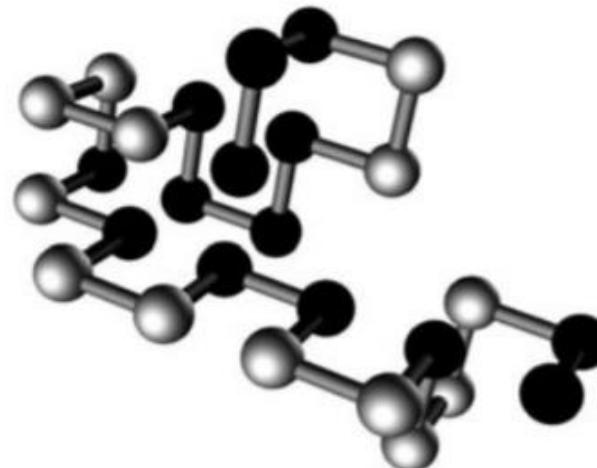


Image credit: J. Comp. Bio. 21, 11, pp. 823 (2014)

Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:
 - Planning/scheduling problems
 - Portfolio optimization
 - Protein folding
 - Graph problems

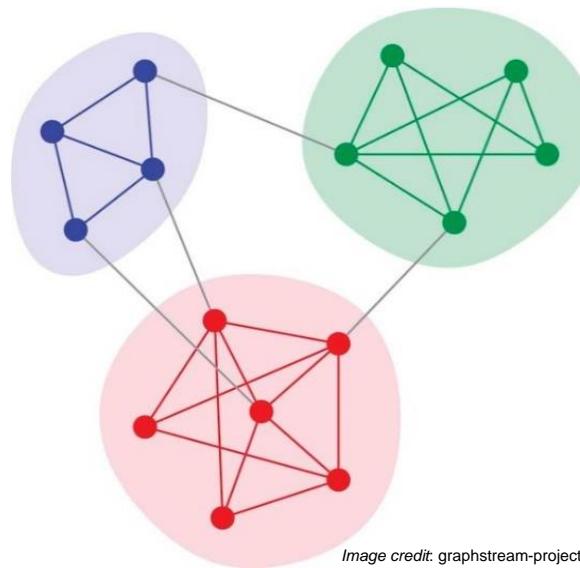
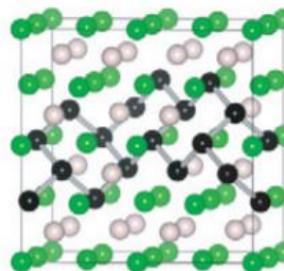


Image credit: graphstream-project.org

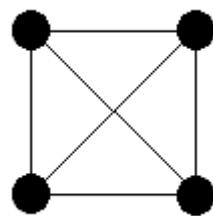
Why is solving the Ising problem interesting?

- Many interesting discrete optimization problems can be framed as Ising problems:
 - Planning/scheduling problems
 - Portfolio optimization
 - Protein folding
 - Graph problems
 - Materials design



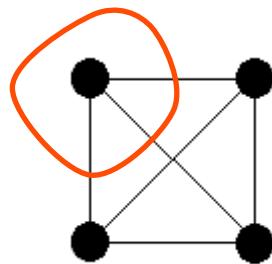
Yuge, et al. *Phys. Rev. B* **77**, 094121 (2008)

MAX-CUT



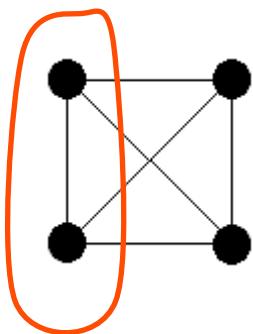
MAX-CUT

Cut size = 3

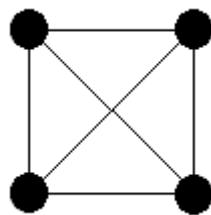


MAX-CUT

Cut size = 4

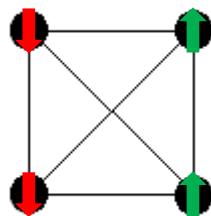


Ising Formulation of MAX-CUT



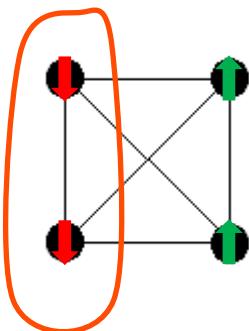
$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j \quad \mathcal{J} = - \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Ising Formulation of MAX-CUT



$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j \quad J = - \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Ising Formulation of MAX-CUT

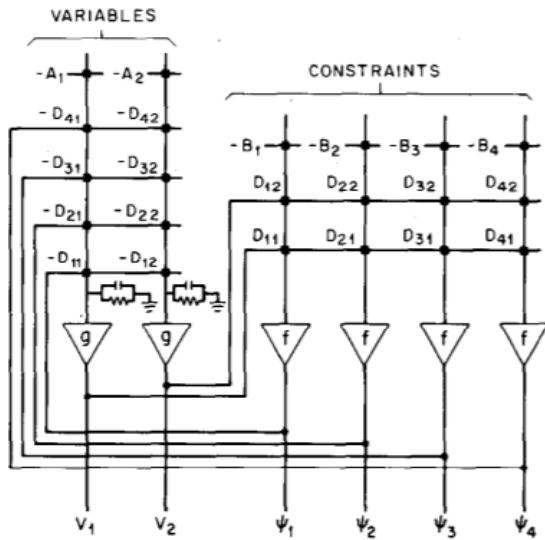


$$H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j \quad J = - \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Overview

- The Ising Problem
- **Ising Machines**
- Foundations of All-Optical OPO Ising Machines
- Measurement-Feedback OPO Ising Machines
- Conclusions

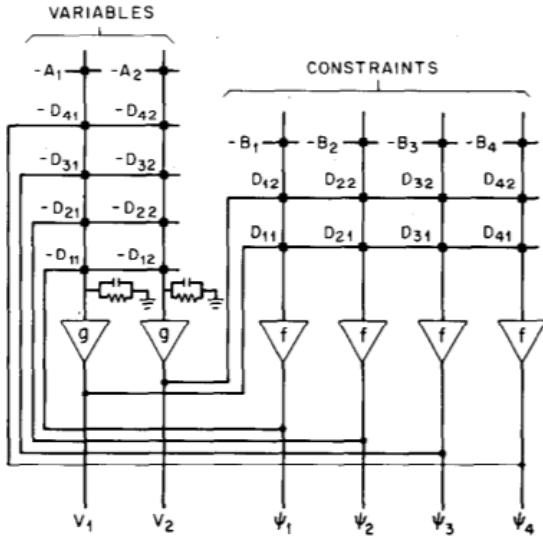
I sing Machines



Neural Networks

Biol. Cybern. 52, 141-152
(1985)

Ising Machines



Neural Networks

Biol. Cybern. 52, 141-152
(1985)

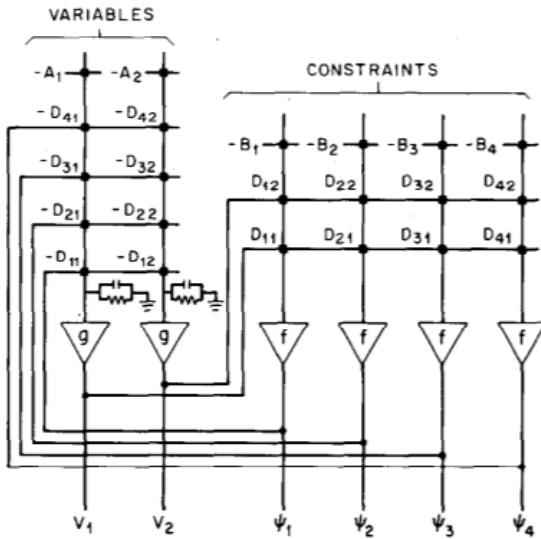
Quantum Annealing

Nature 473, 194-198 (2011)



Credit: dwavesys.com

Ising Machines



Neural Networks

Biol. Cybern. 52, 141-152
(1985)

Quantum Annealing

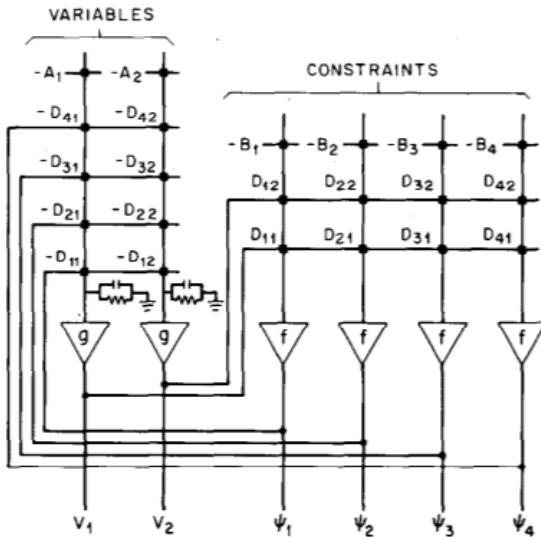
Nature 473, 194-198 (2011)

Uses *quantum* dynamics
to solve a ***classical*** spin
Hamiltonian!



Credit: dwavesys.com

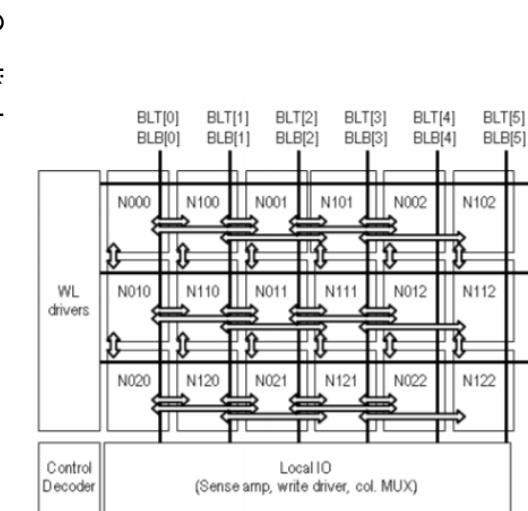
Ising Machines



Neural Networks
Biol. Cybern. 52, 141-152
(1985)



Quantum Annealing
Nature 473, 194-198 (2011)

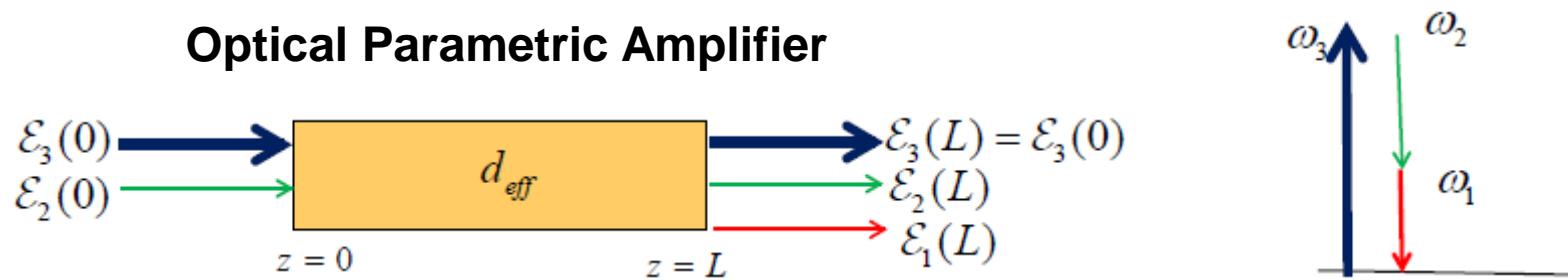


CMOS Annealers
ISSCC 24.3 (2015)

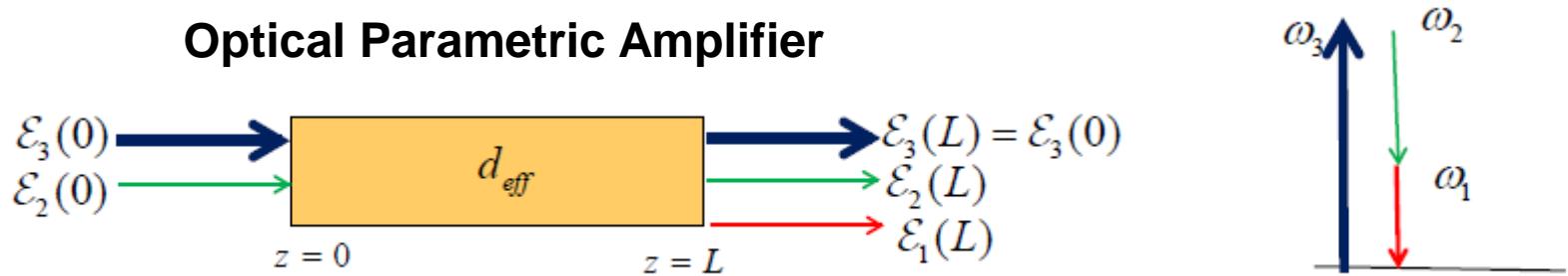
Overview

- The Ising Problem
- Ising Machines
- **Foundations of All-Optical OPO Ising Machines**
- Measurement-Feedback OPO Ising Machines
- Conclusions

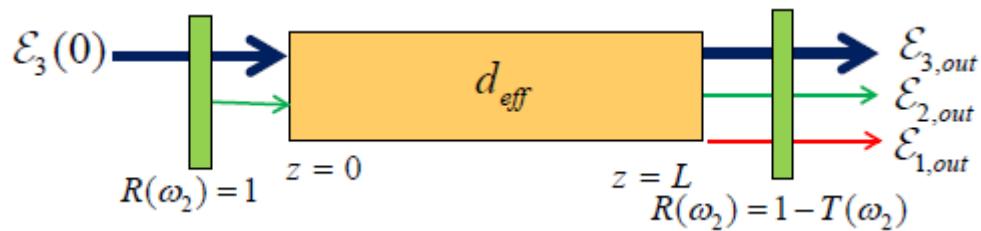
Optical Parametric Oscillators



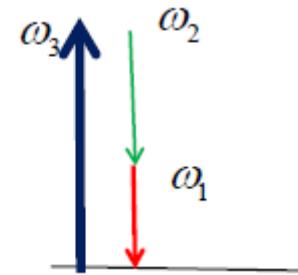
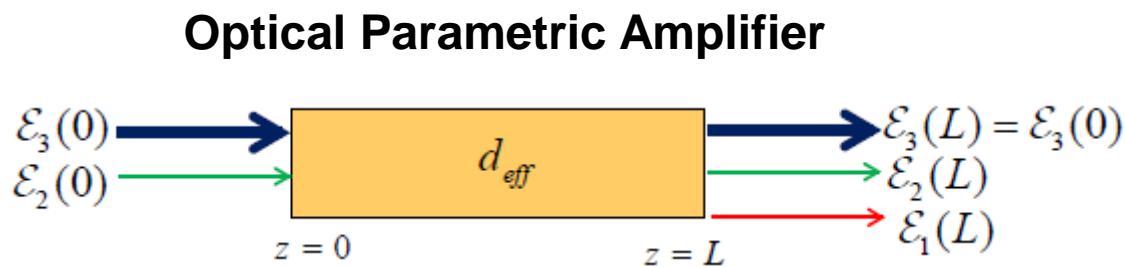
Optical Parametric Oscillators



OPO: Optical Parametric Amplifier in a Cavity



Optical Parametric Oscillators



OPO: Optical Parametric Amplifier in a Cavity

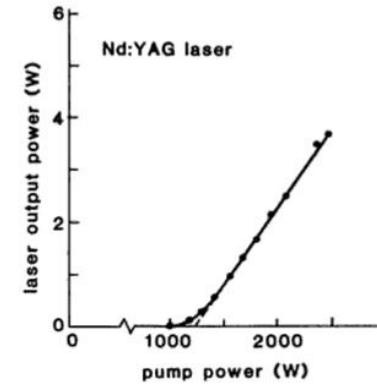
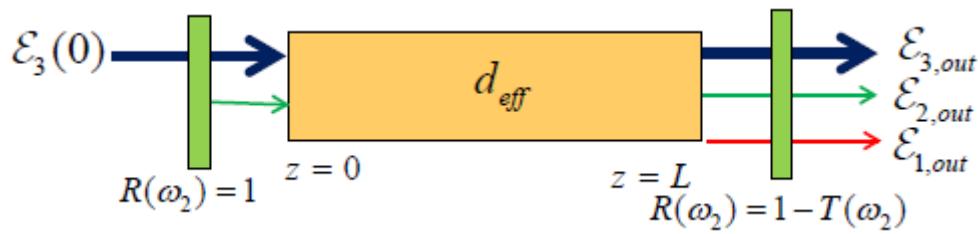
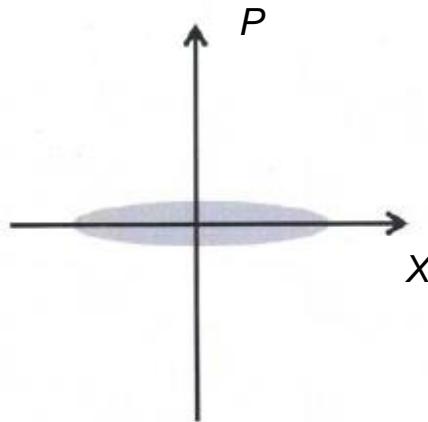


Diagram credit:
A. E. Siegman, "Lasers" (1986)

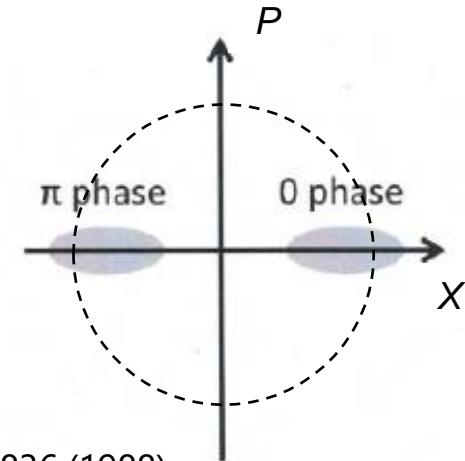
An OPO has a **threshold**, just like a laser does.

OPO Phase Properties

Squeezed vacuum state **below threshold**



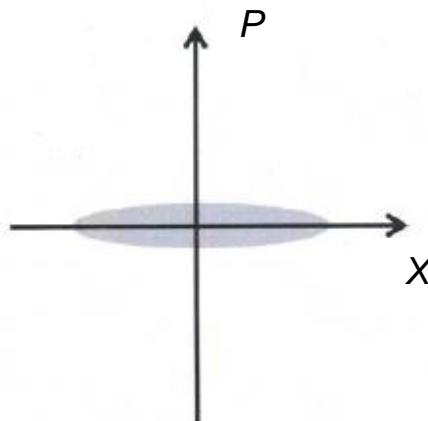
Squeezed coherent state **above threshold**



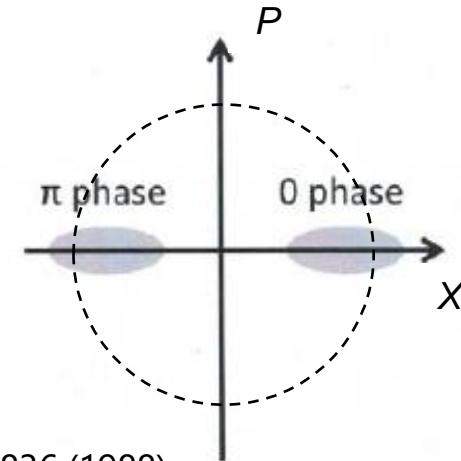
M. Wolinski and H. J. Carmichael. *PRL* **60**, 1836 (1988)

OPO Phase Properties

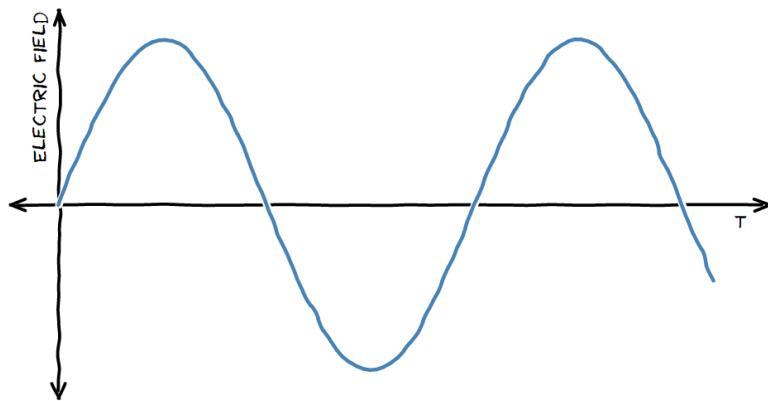
Squeezed vacuum state **below threshold**



Squeezed coherent state **above threshold**

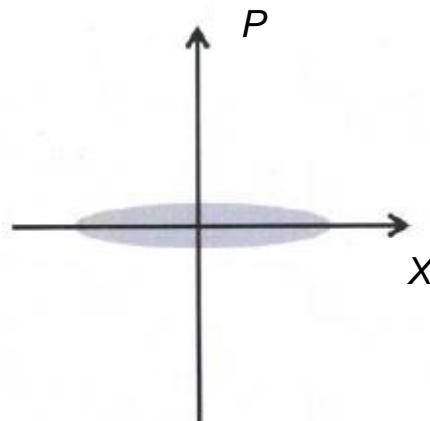


M. Wolinski and H. J. Carmichael. *PRL* **60**, 1836 (1988)

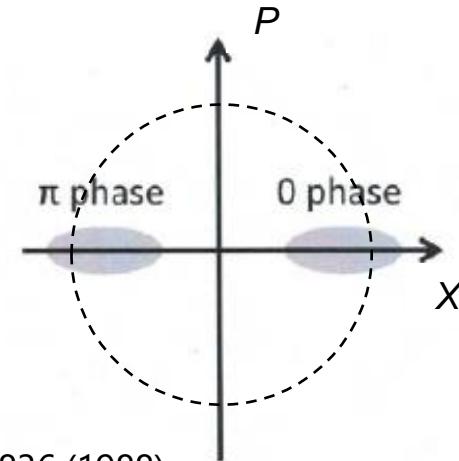


OPO Phase Properties

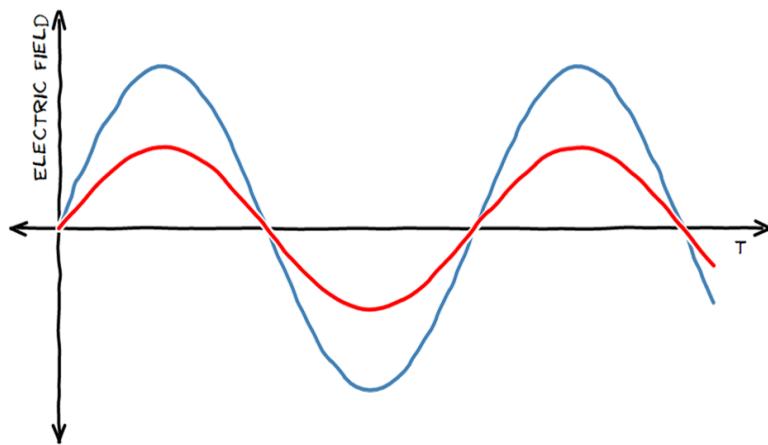
Squeezed vacuum state **below threshold**



Squeezed coherent state **above threshold**

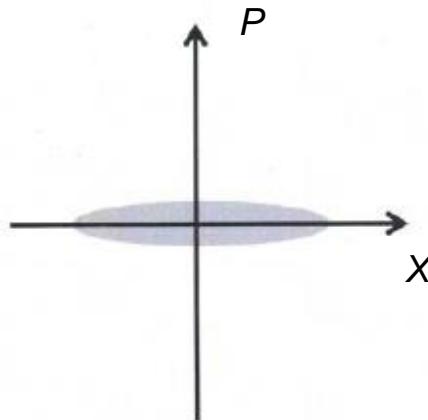


M. Wolinski and H. J. Carmichael. *PRL* **60**, 1836 (1988)

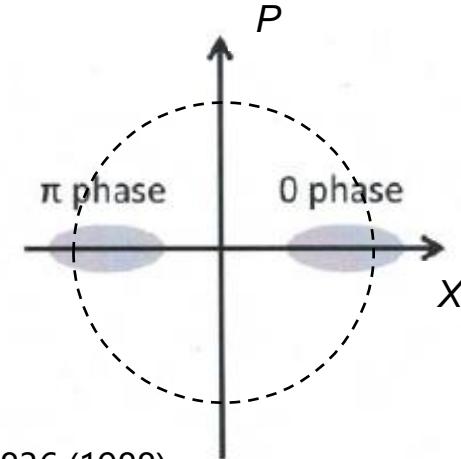


OPO Phase Properties

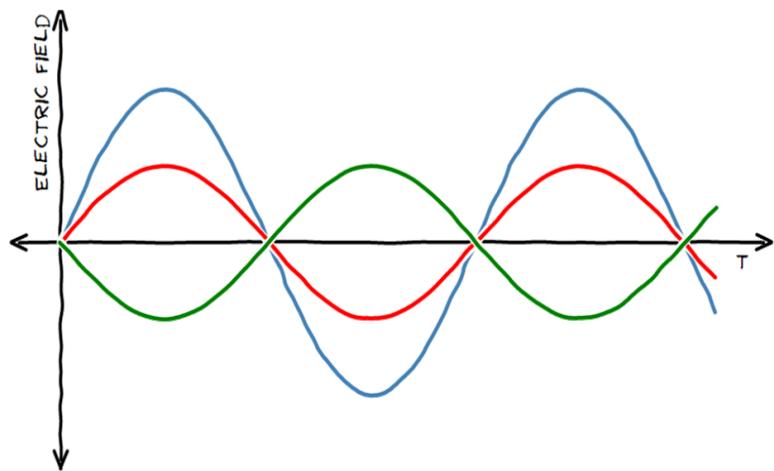
Squeezed vacuum state **below threshold**



Squeezed coherent state **above threshold**

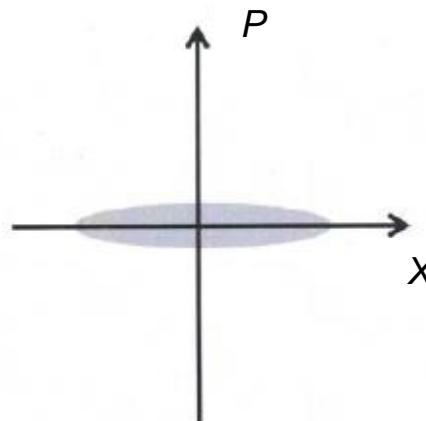


M. Wolinski and H. J. Carmichael. *PRL* **60**, 1836 (1988)

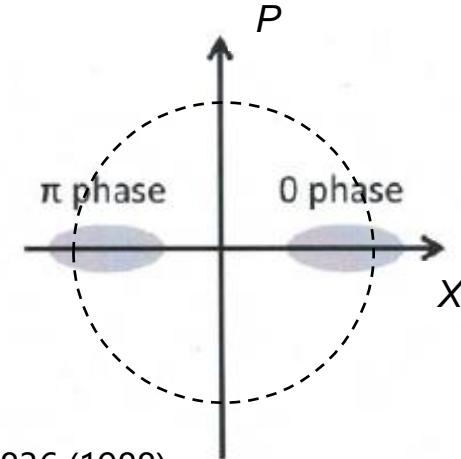


OPO Phase Properties

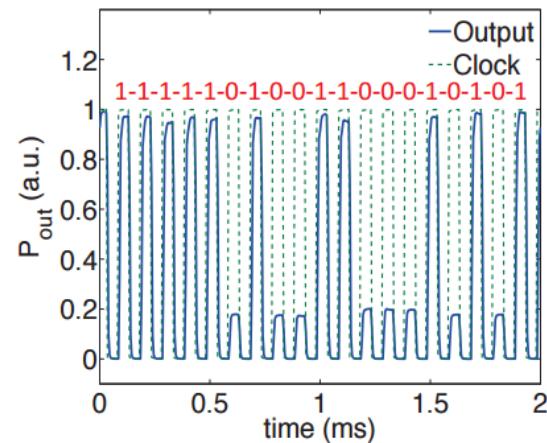
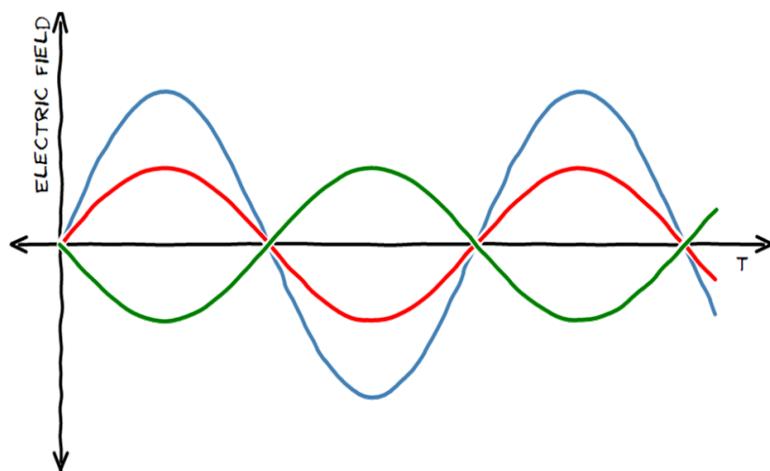
Squeezed vacuum state **below threshold**



Squeezed coherent state **above threshold**

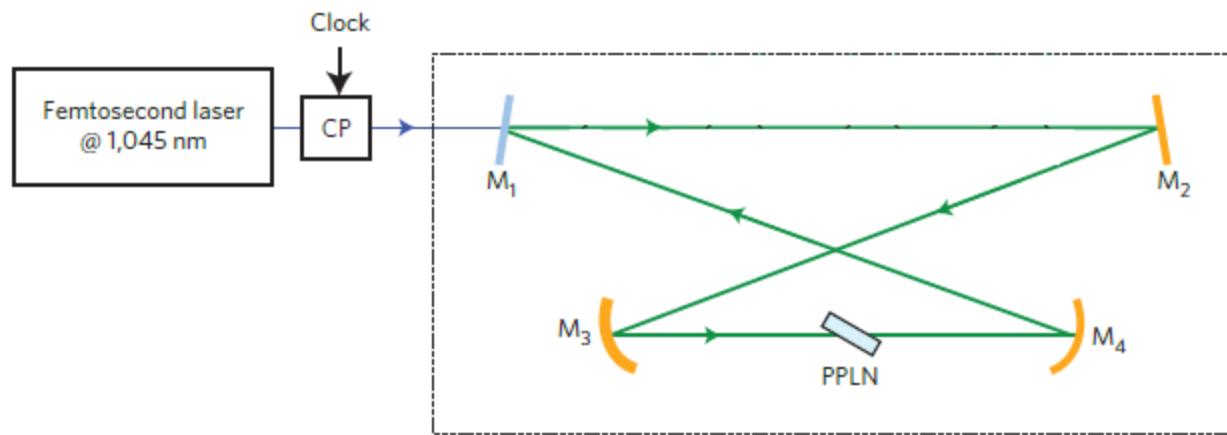


M. Wolinski and H. J. Carmichael. *PRL* **60**, 1836 (1988)



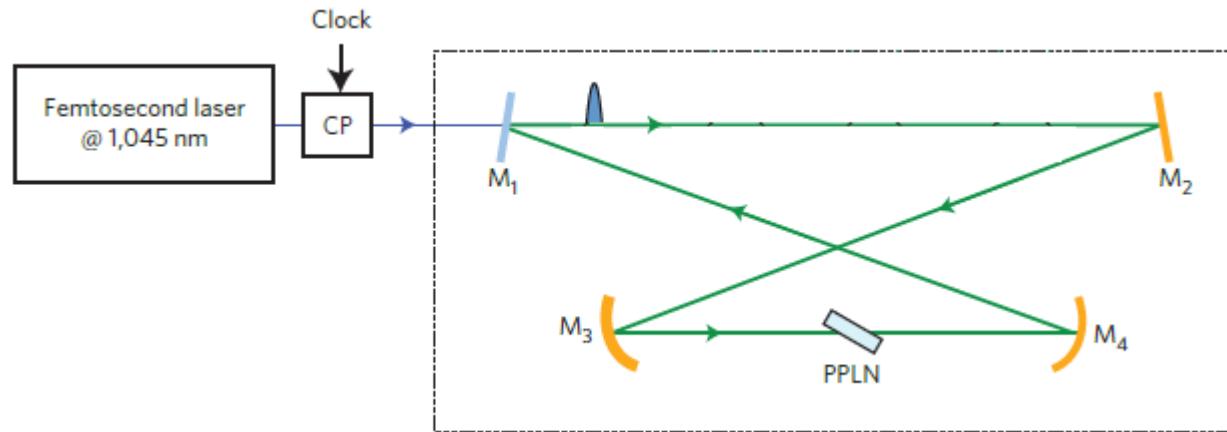
A. Marandi, et al. *Optics Express* 20, 17 19322 (2012).

Time-Multiplexed OPOs



Figures adapted from: A. Marandi, *et al.* *Nature Photonics* 8, 937 (2014).

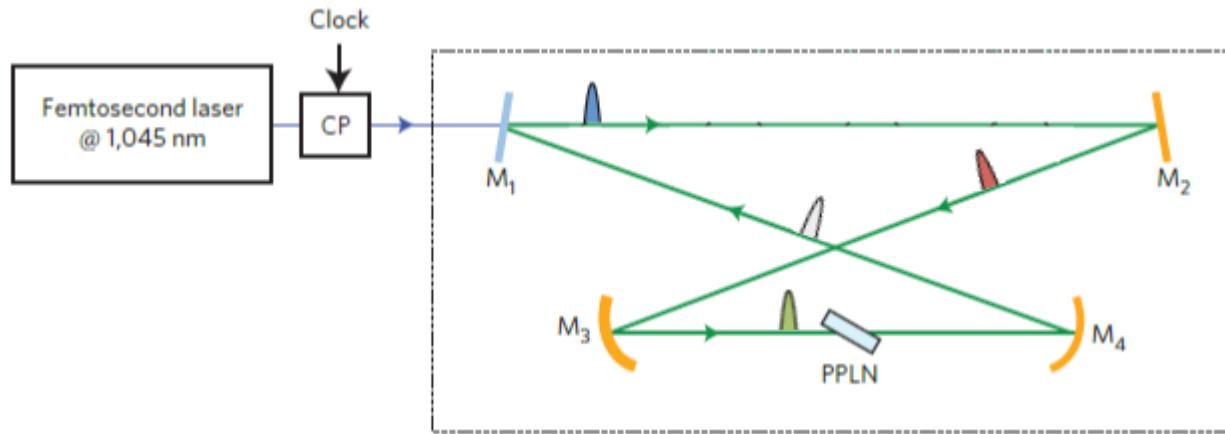
Time-Multiplexed OPOs



Typical pulsed OPO case:

$$\frac{1}{\text{Laser pulse repetition rate}} = \text{Cavity roundtrip time}$$

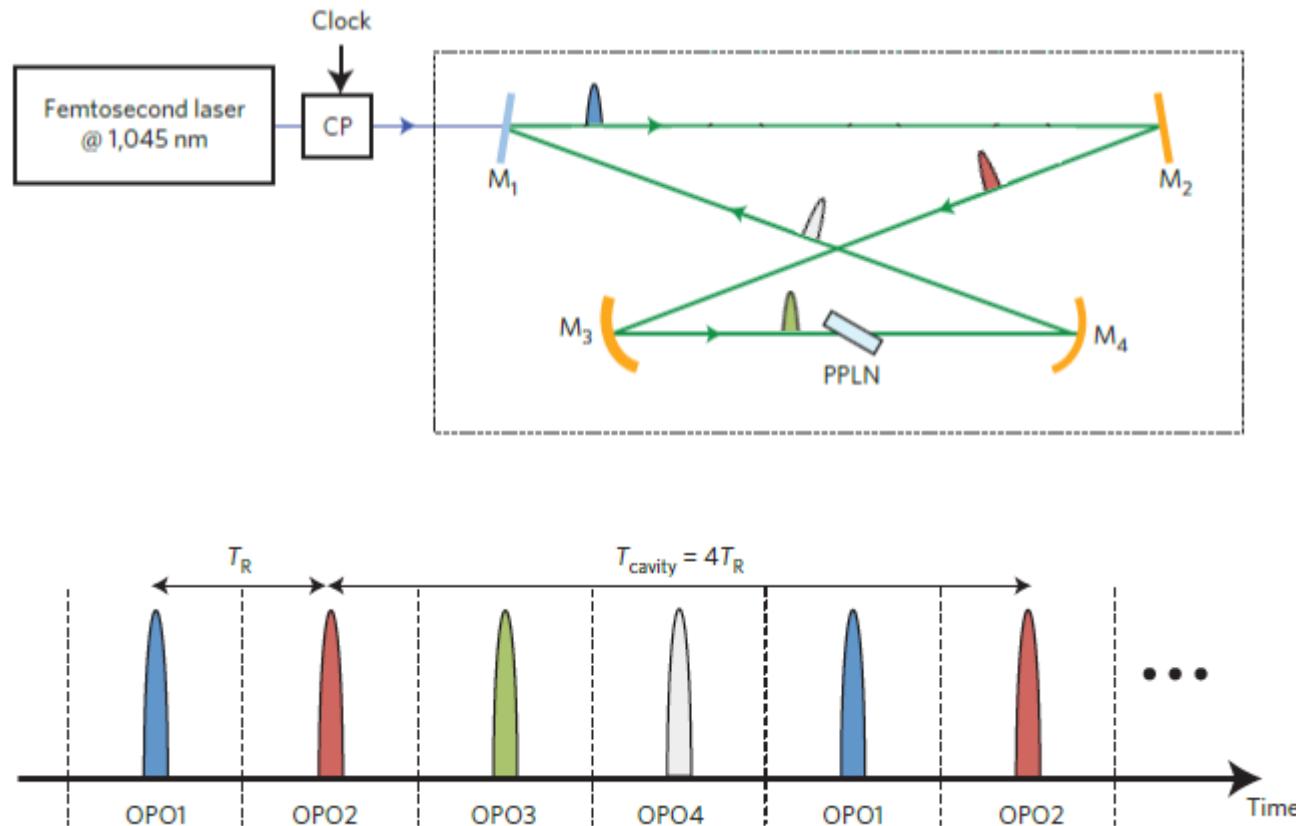
Time-Multiplexed OPOs



Time-multiplexed OPO case:

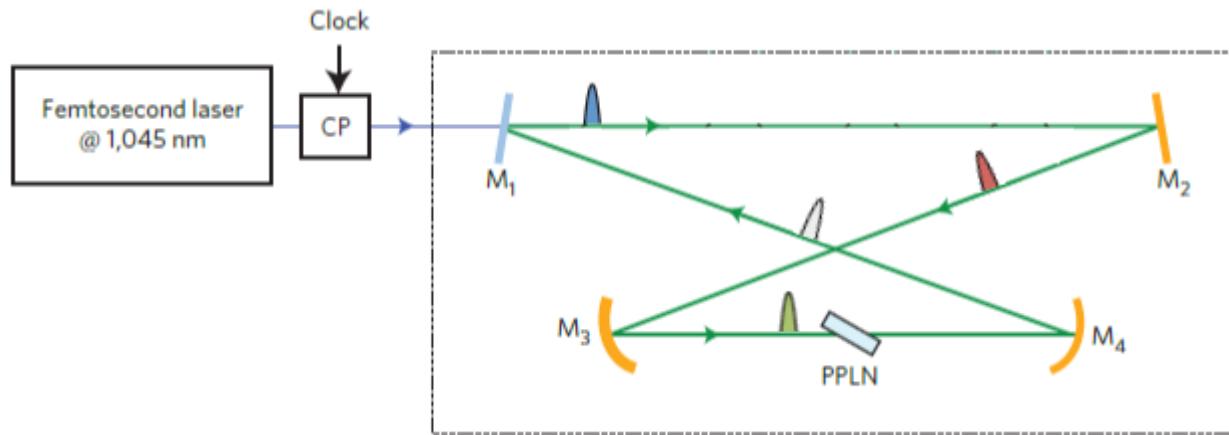
$$\frac{1}{\text{Laser pulse repetition rate}} = \frac{\text{Cavity roundtrip time}}{N_{\text{pulses}}}$$

Time-Multiplexed OPOs



Figures adapted from: A. Marandi, et al. *Nature Photonics* 8, 937 (2014).

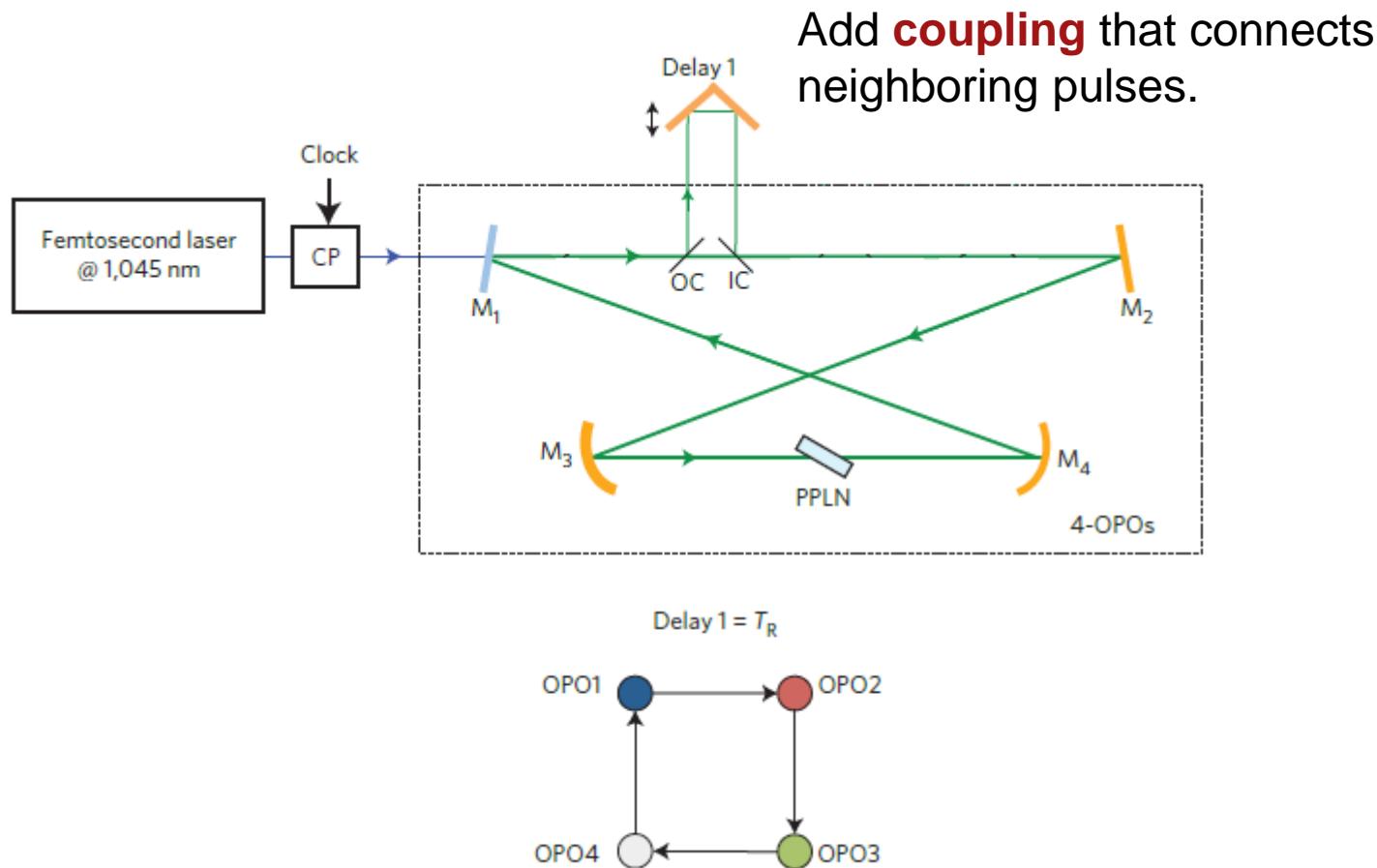
From OPOs to Ising Machine



So far we just have N uncoupled “spins”.
(Recall that each pulse has phase 0 or π).

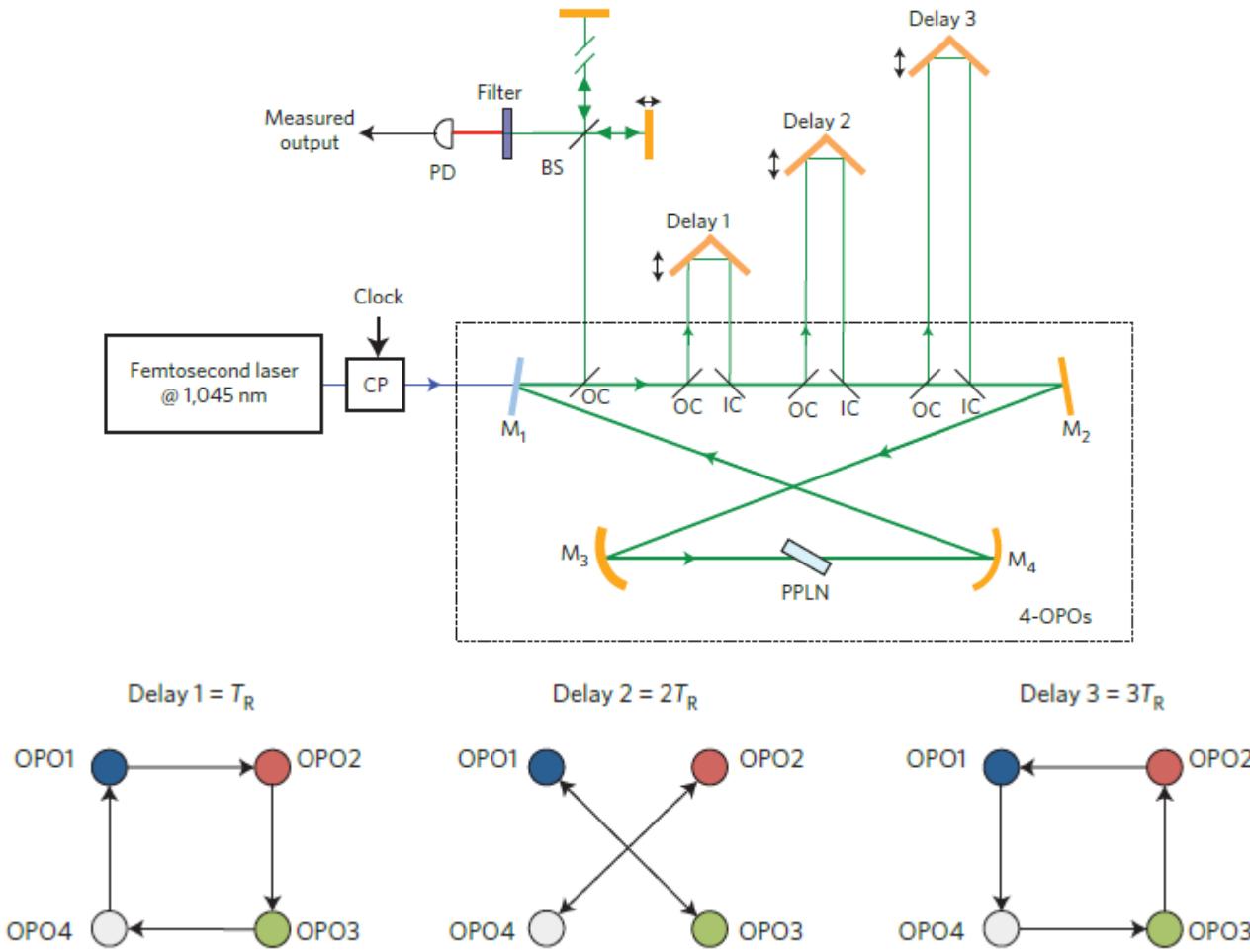
$$H = 0$$

From OPOs to Ising Machine



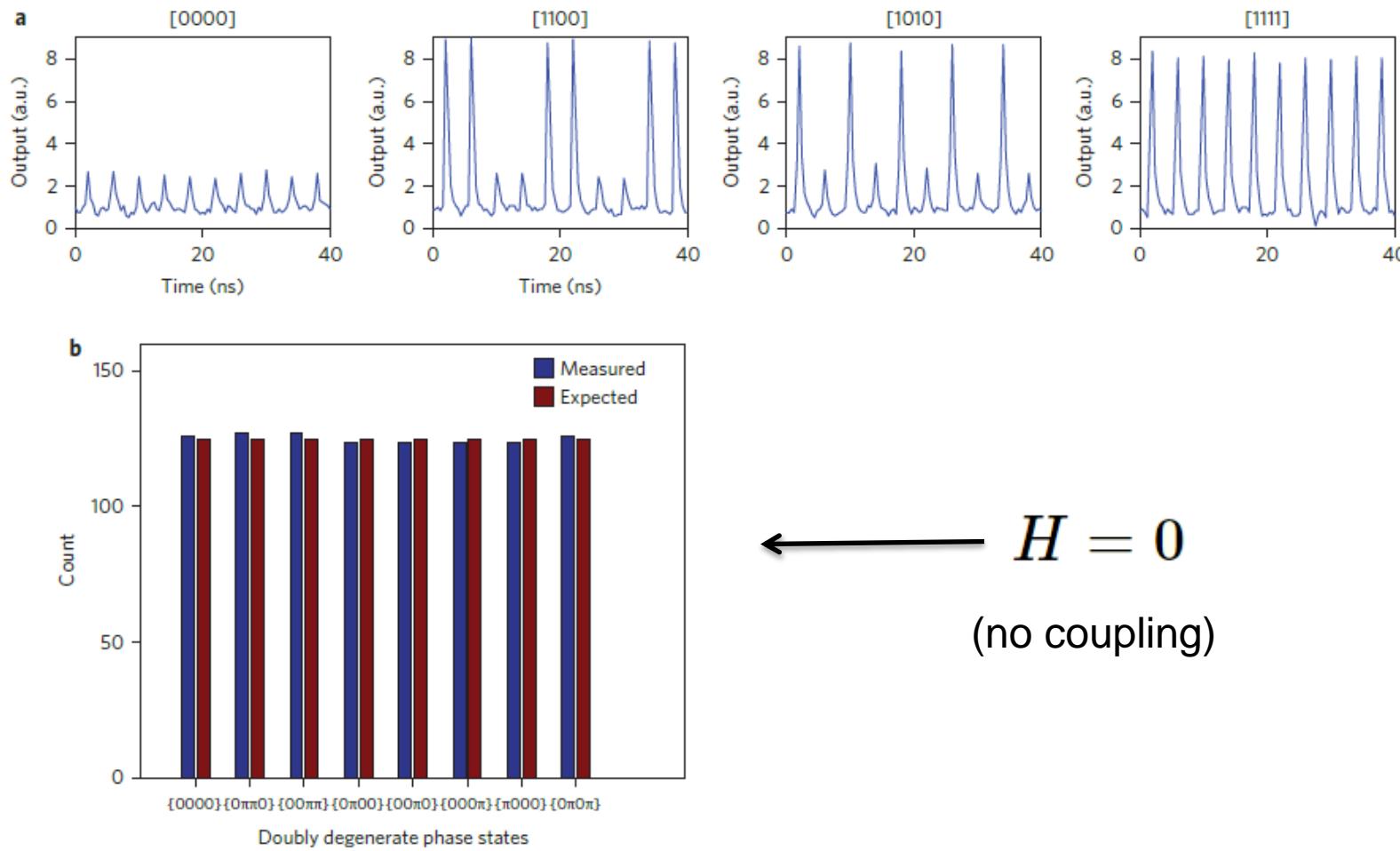
Figures adapted from: A. Marandi, *et al. Nature Photonics* 8, 937 (2014).

$N = 4$ Ising Machine



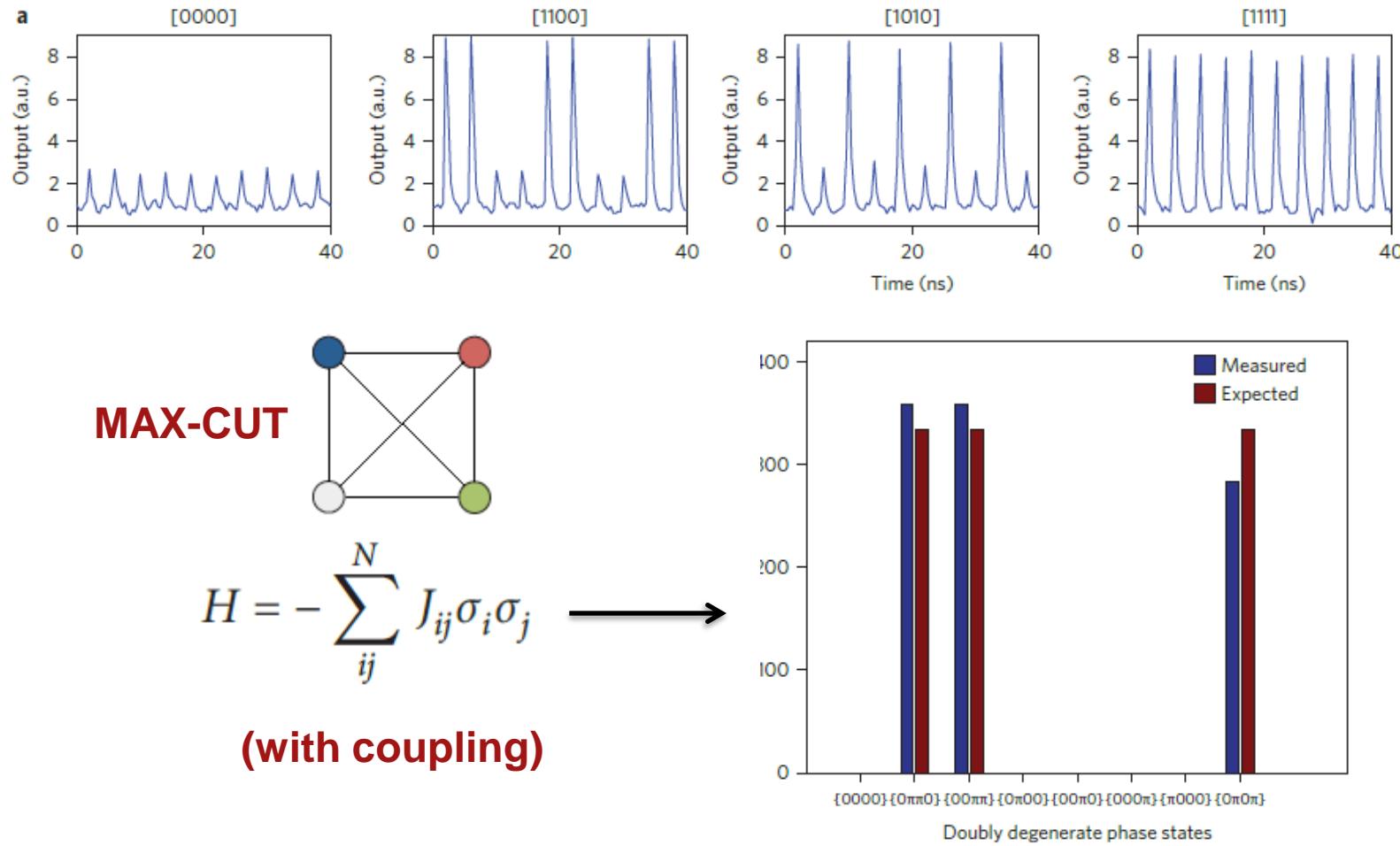
Figures adapted from: A. Marandi, et al. *Nature Photonics* 8, 937 (2014).

$N = 4$ Ising Machine Results



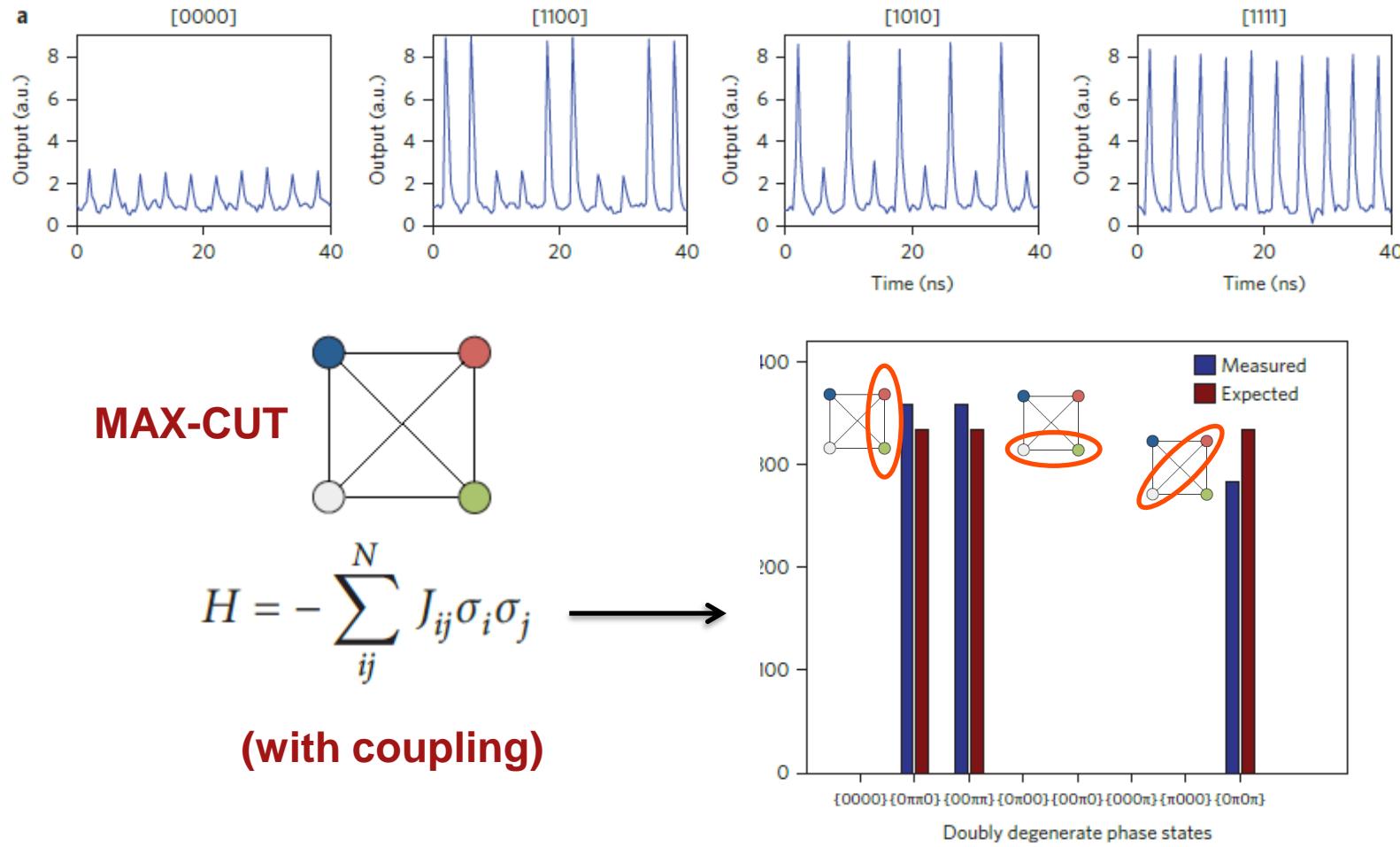
Figures adapted from: A. Marandi, *et al. Nature Photonics* 8, 937 (2014).

$N = 4$ Ising Machine Results



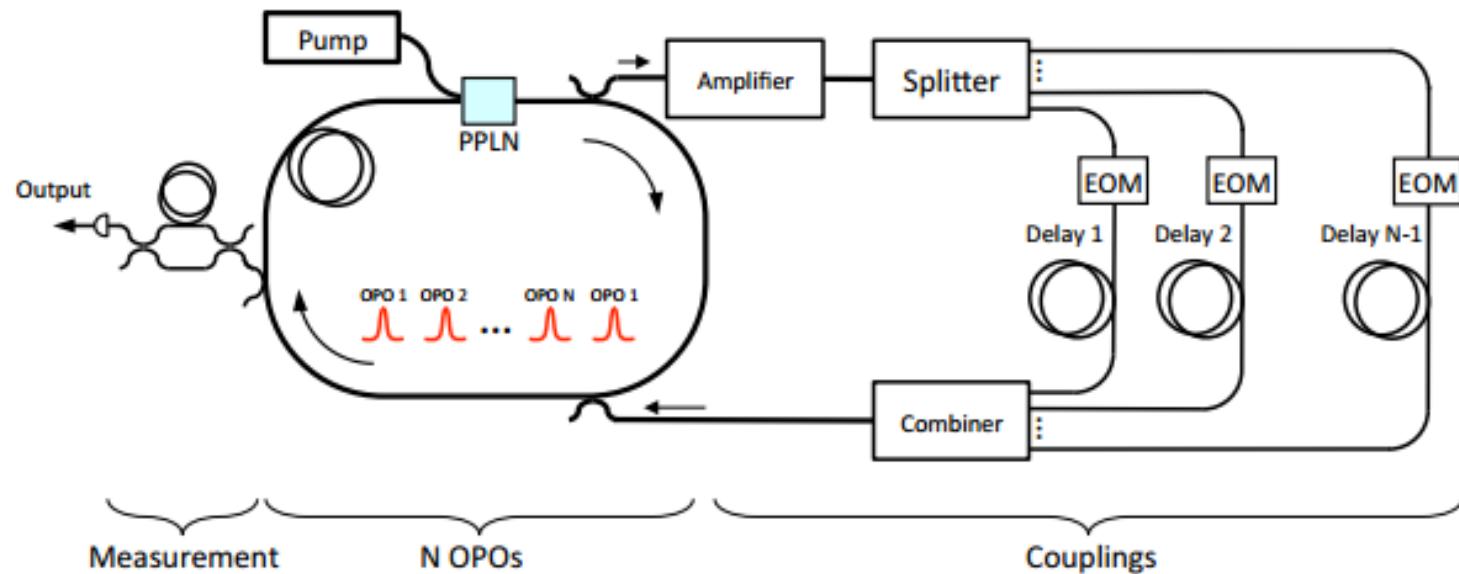
Figures adapted from: A. Marandi, et al. *Nature Photonics* 8, 937 (2014).

$N = 4$ Ising Machine Results



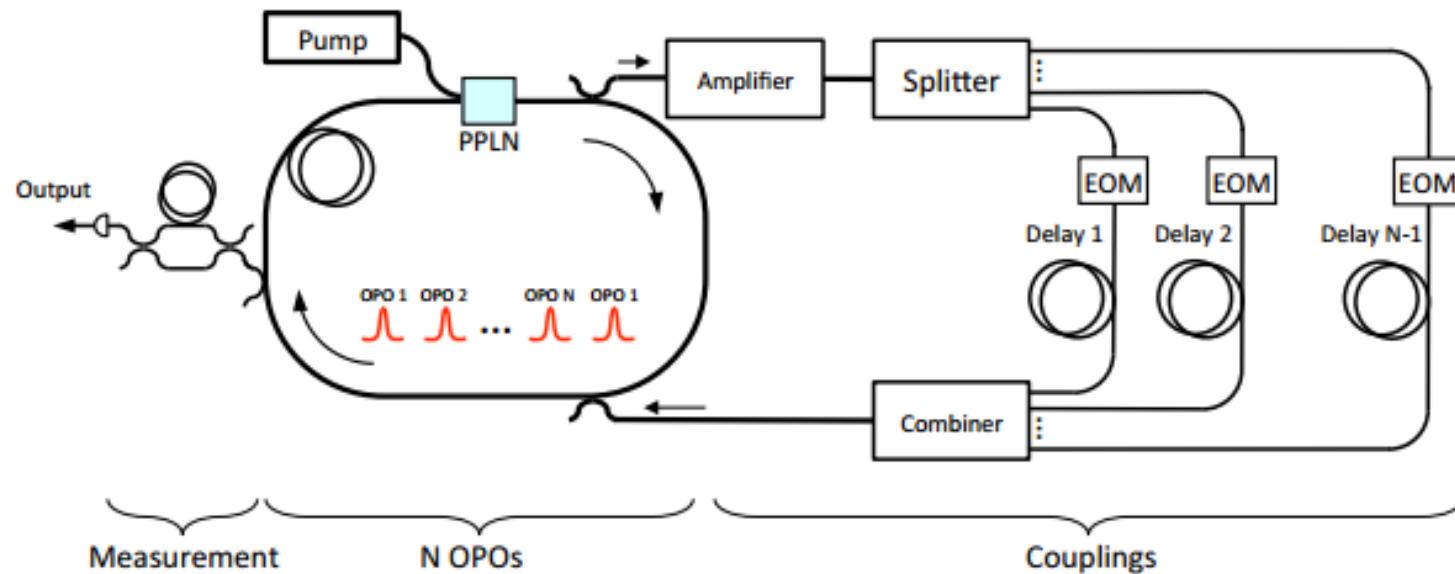
Figures adapted from: A. Marandi, et al. *Nature Photonics* 8, 937 (2014).

Achieving Arbitrary Connectivity



Note: with temporal control of couplings, coupling strength can change during computation (similar to annealing schedule in AQC)

Achieving Arbitrary Connectivity

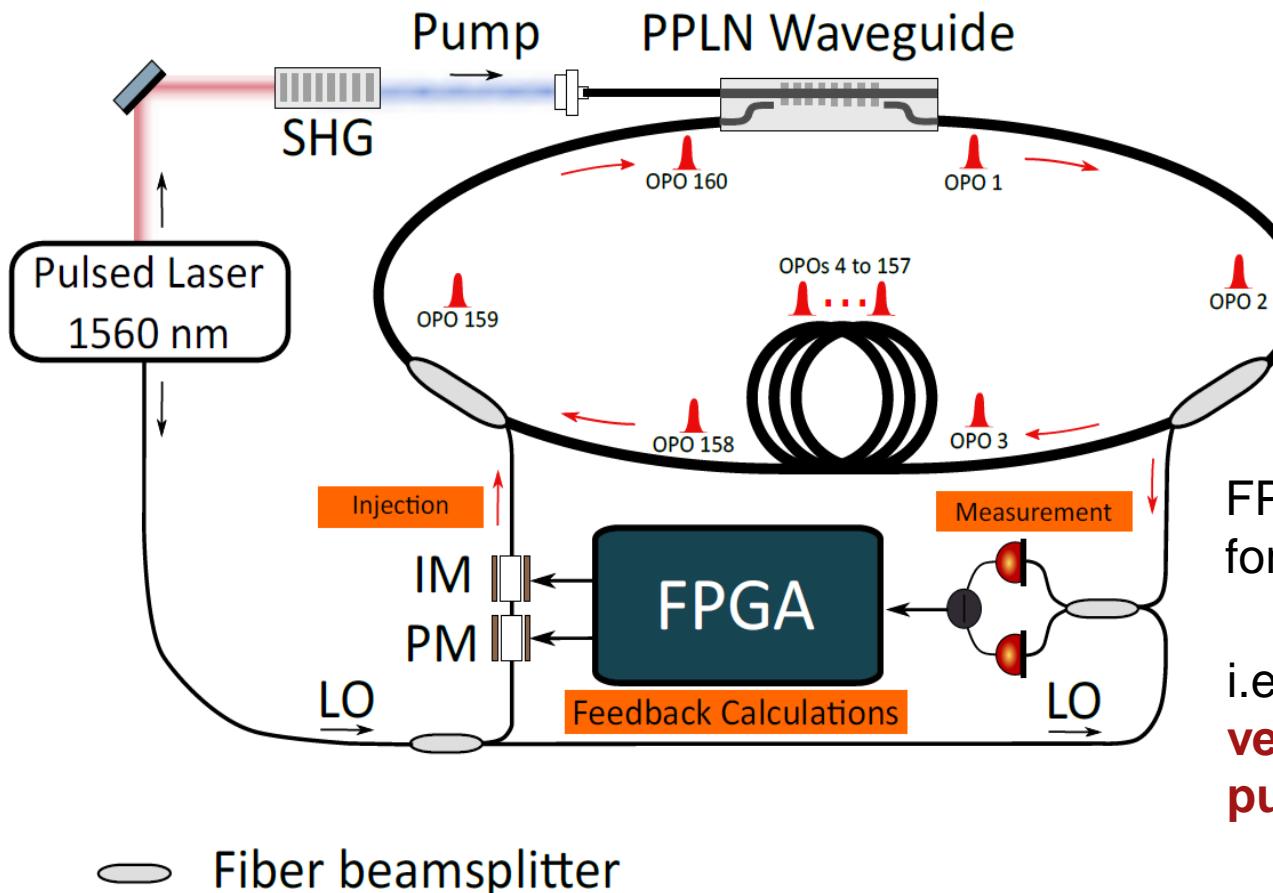


Want $N > 128$, preferably $N \gg 1000$.
→ Cost, loss, and phase stabilization issues!

Overview

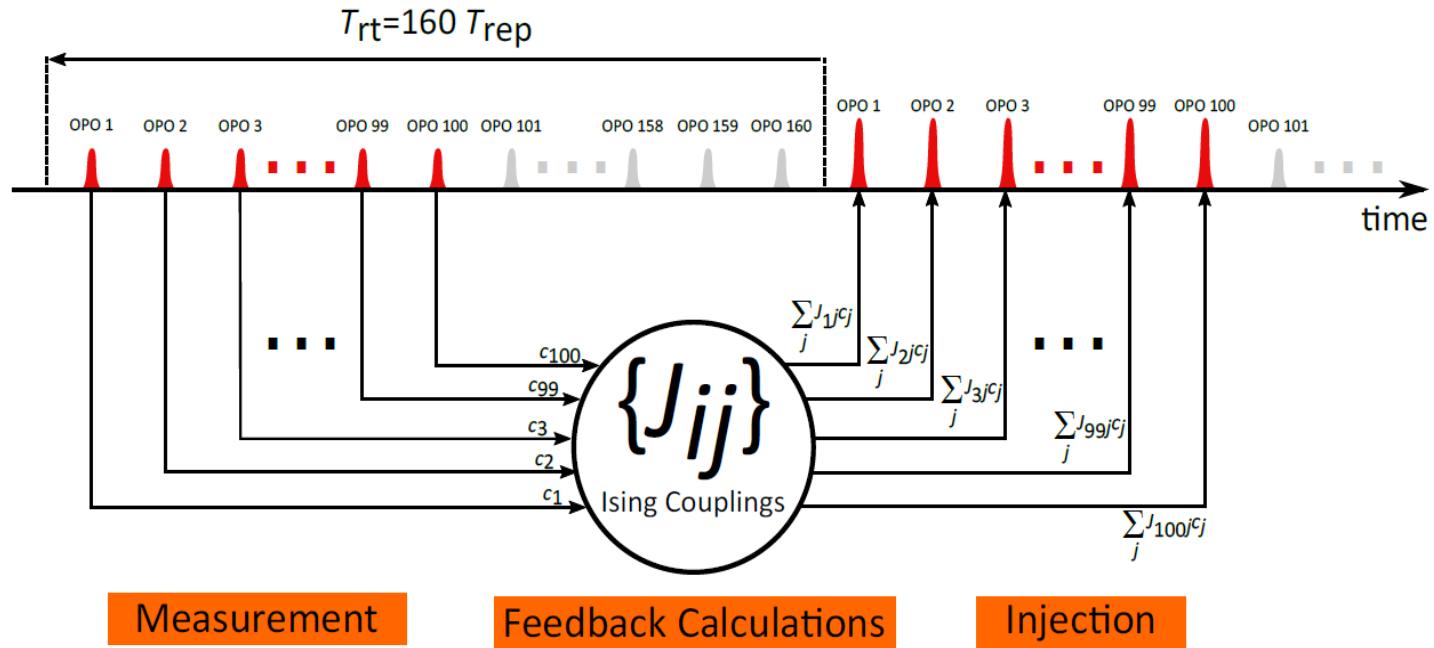
- The Ising Problem
- Ising Machines
- Foundations of All-Optical OPO Ising Machines
- **Measurement-Feedback OPO Ising Machines**
- Conclusions

Measurement-Feedback OPO Ising Machine

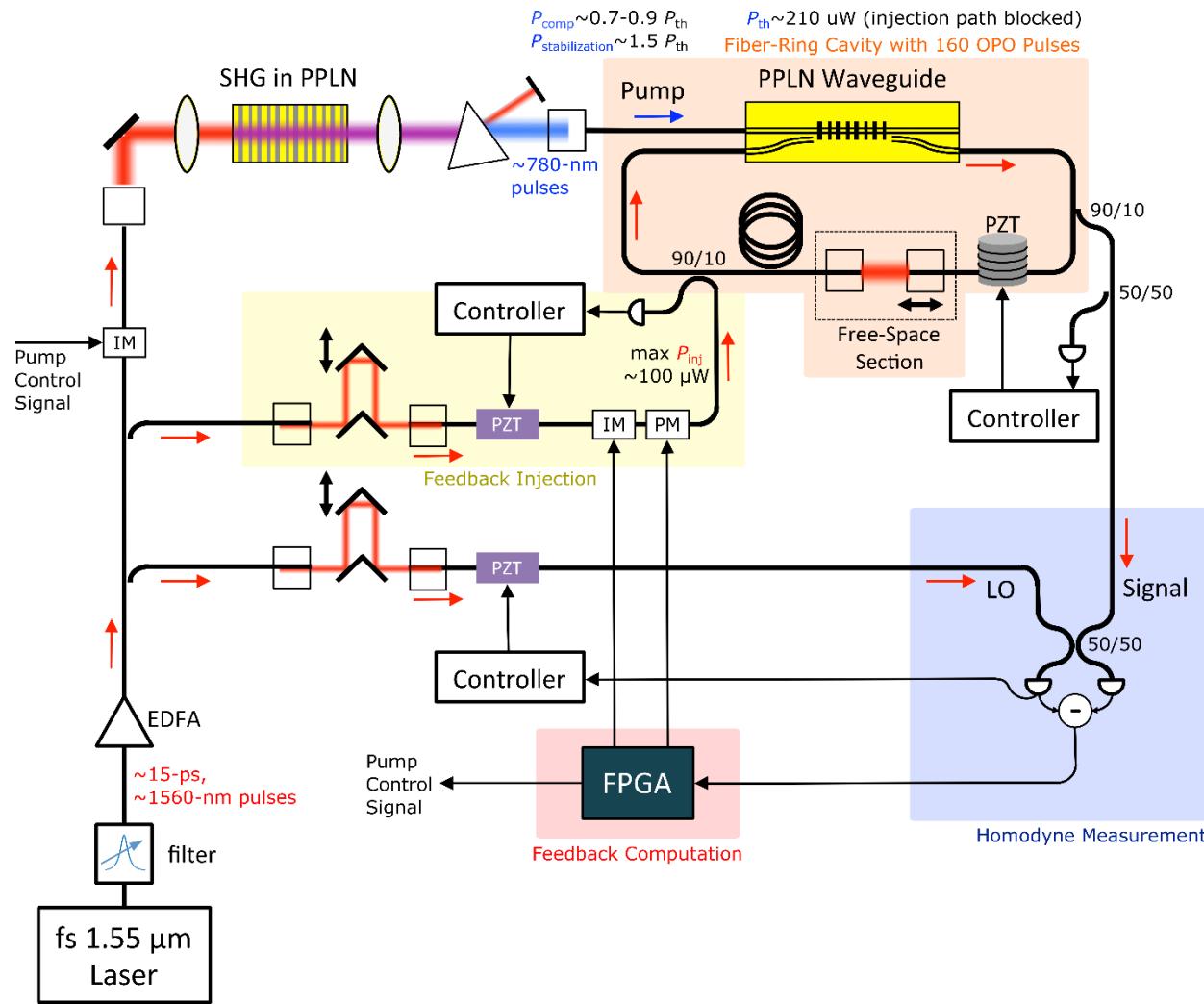


FPGA computes feedback for i th pulse: $\sum_{j=1}^{N=100} J_{ij} c_j$
i.e., **one N -dim vector-vector dot product per pulse**

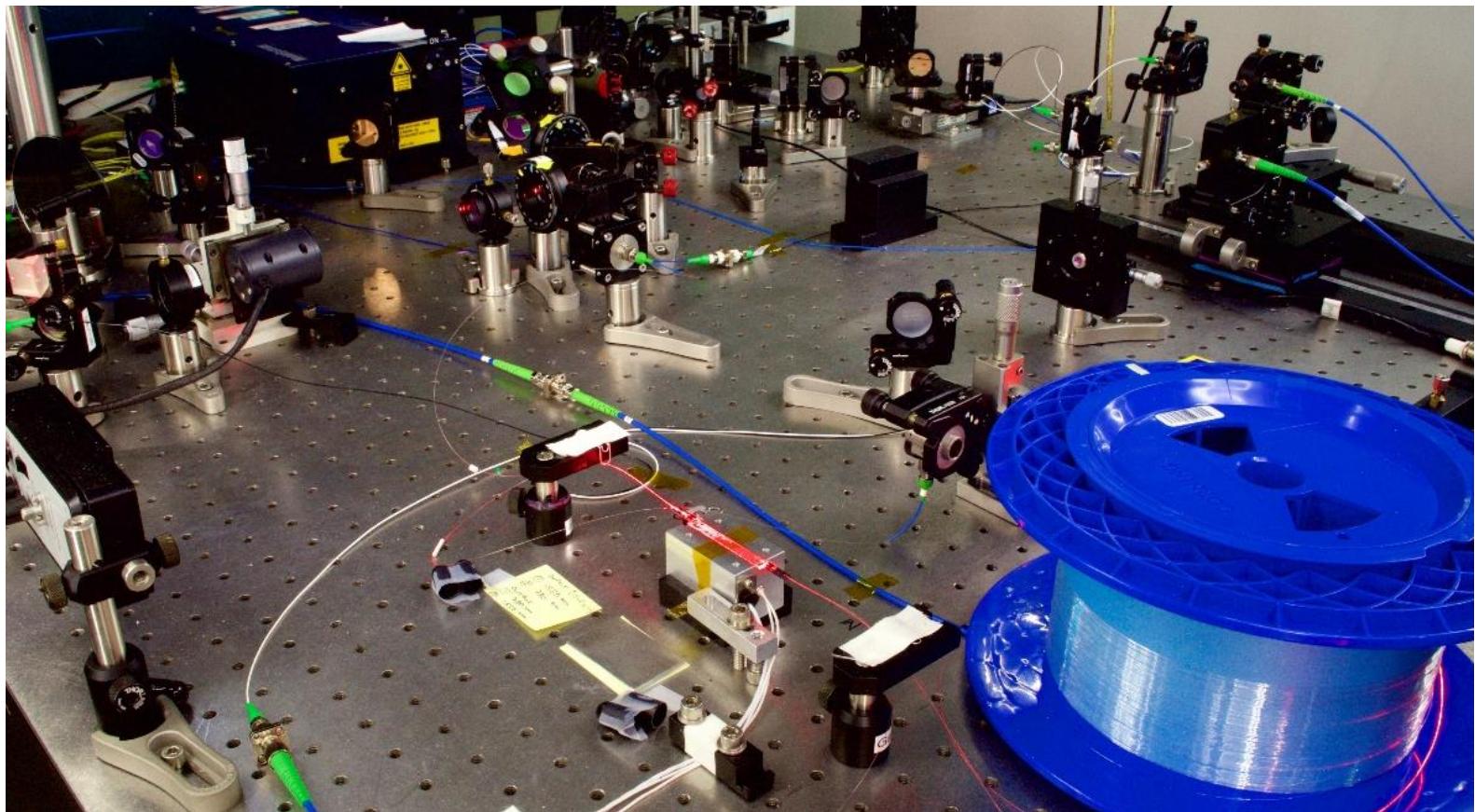
Feedback Calculations



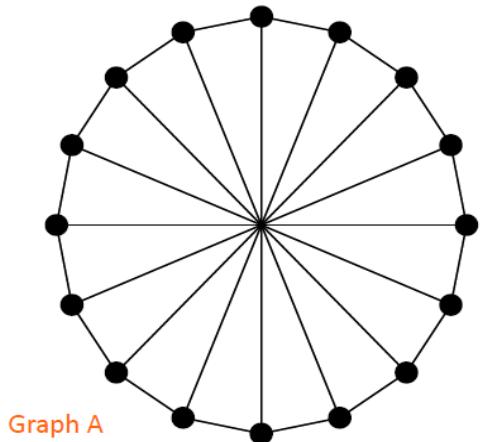
Experimental Realization



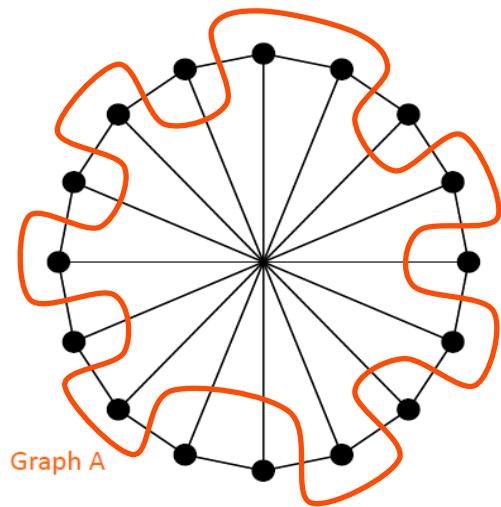
Experimental Realization



MAX-CUT on small graphs



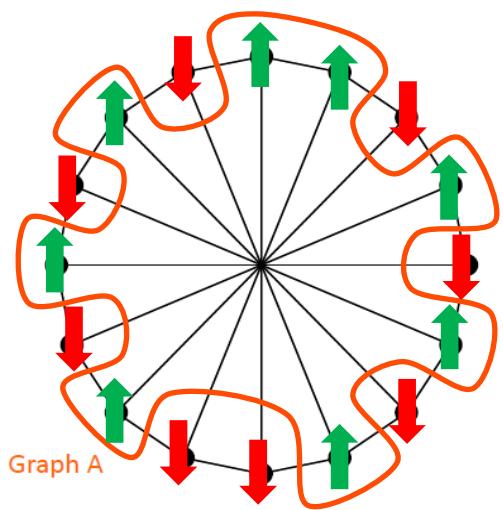
MAX-CUT on small graphs



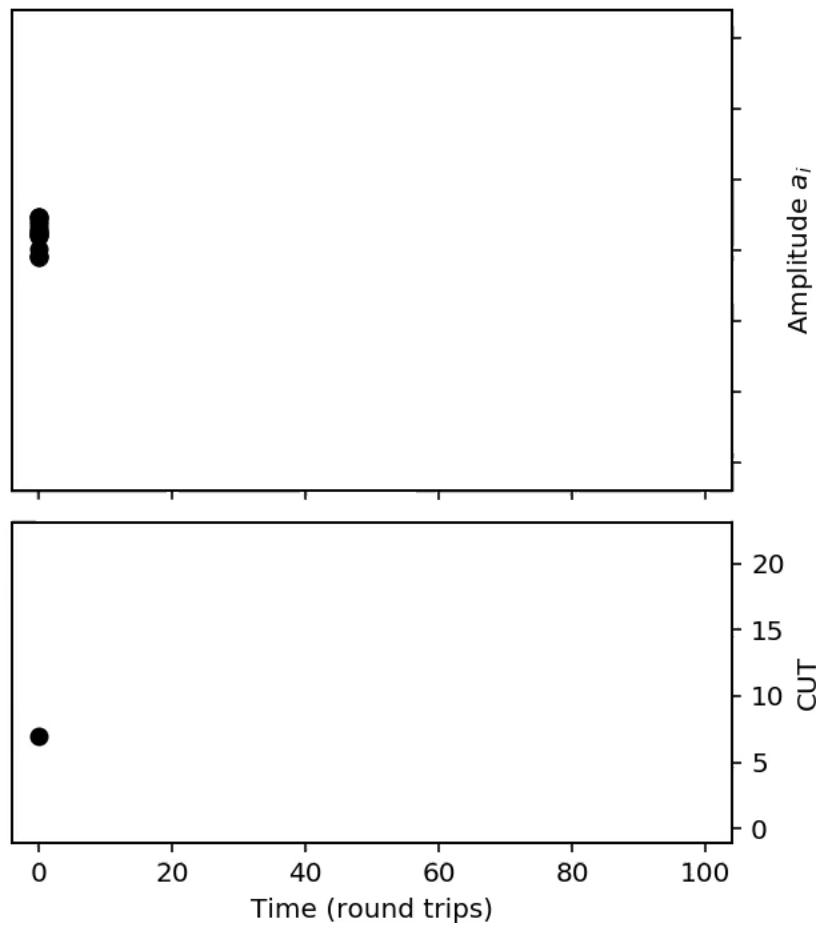
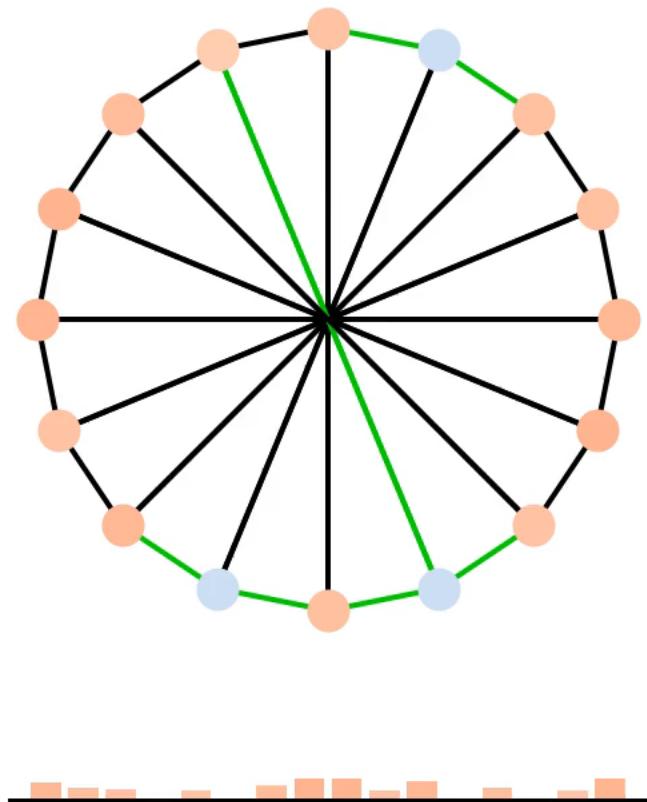
22 edges are crossed

→ Size of “maximum cut” is 22

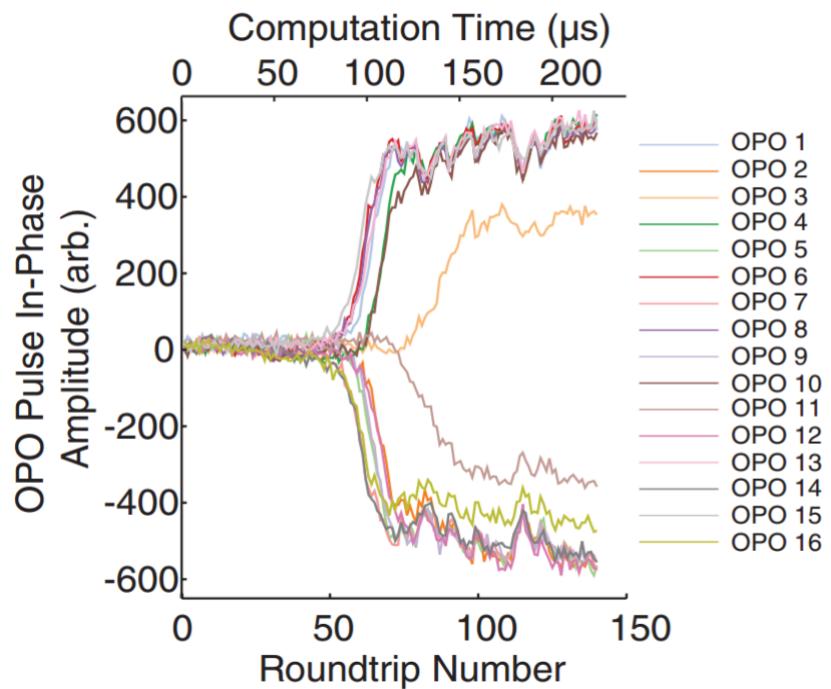
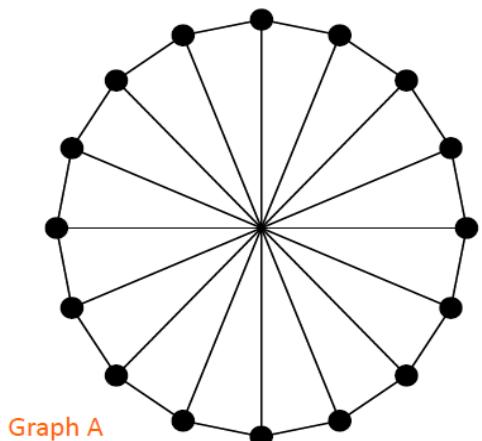
MAX-CUT on small graphs



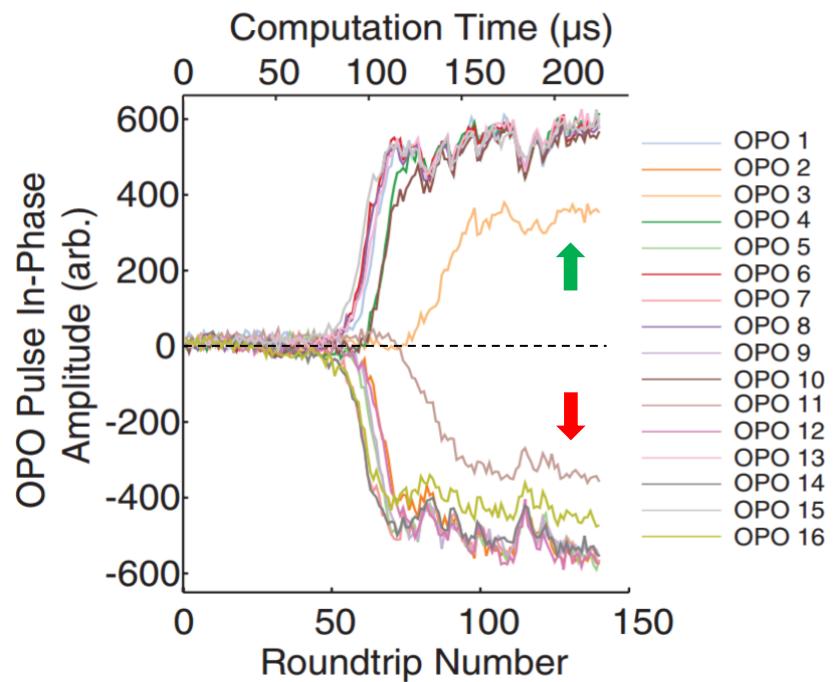
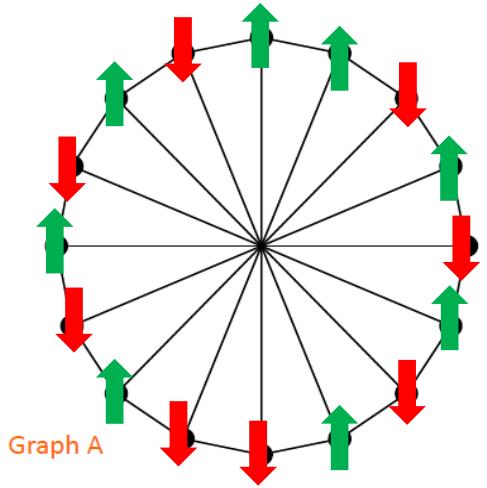
MAX-CUT on small graphs



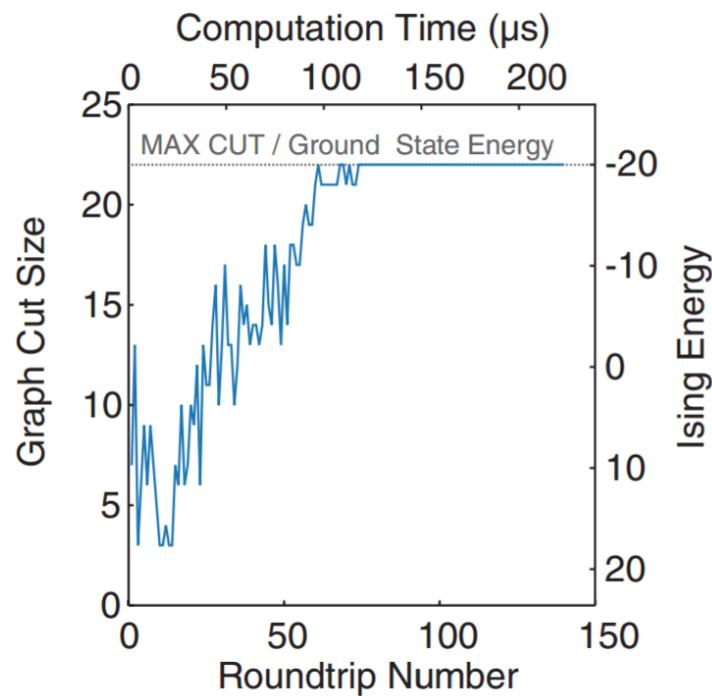
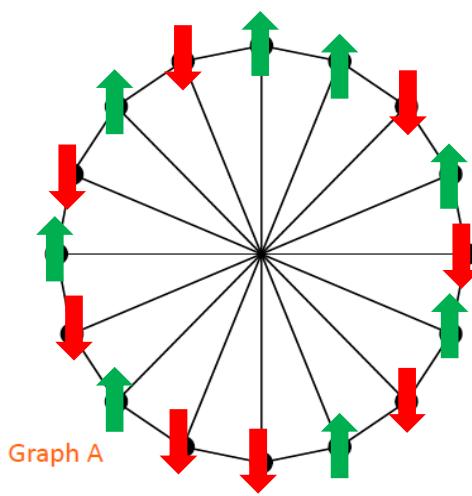
MAX-CUT on small graphs



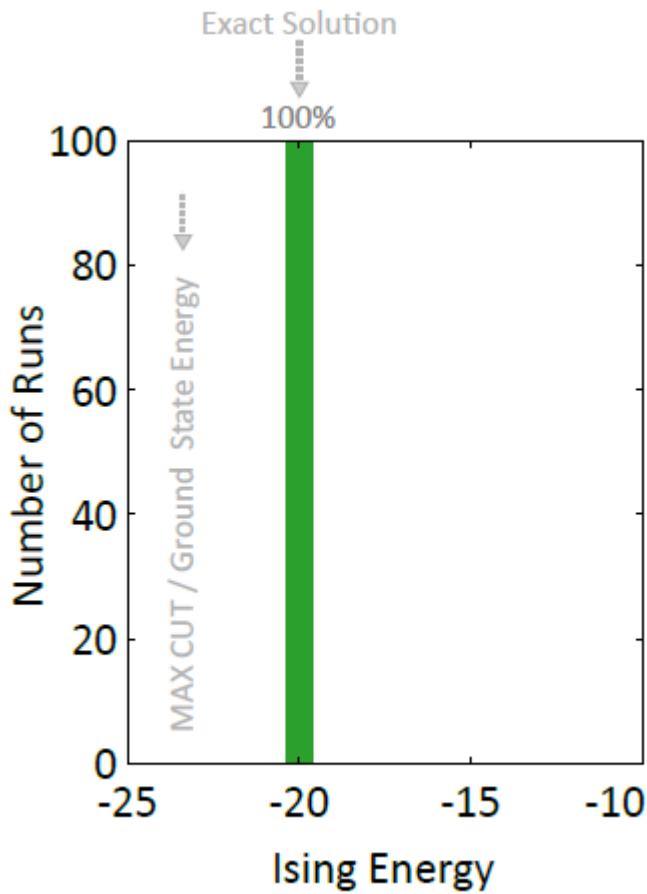
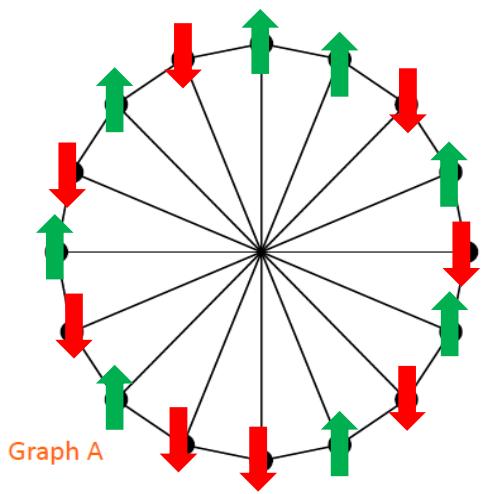
MAX-CUT on small graphs



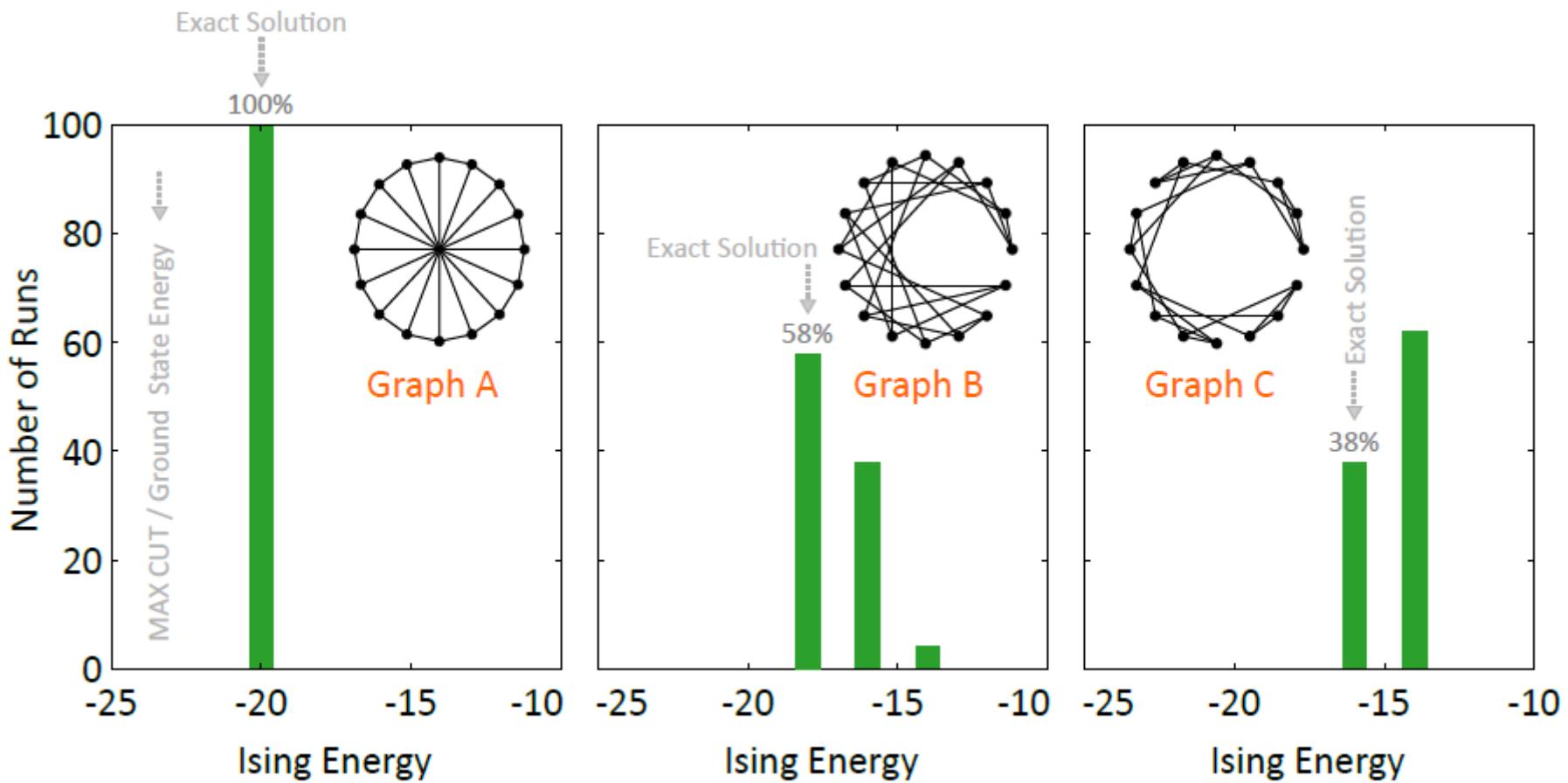
MAX-CUT on small graphs



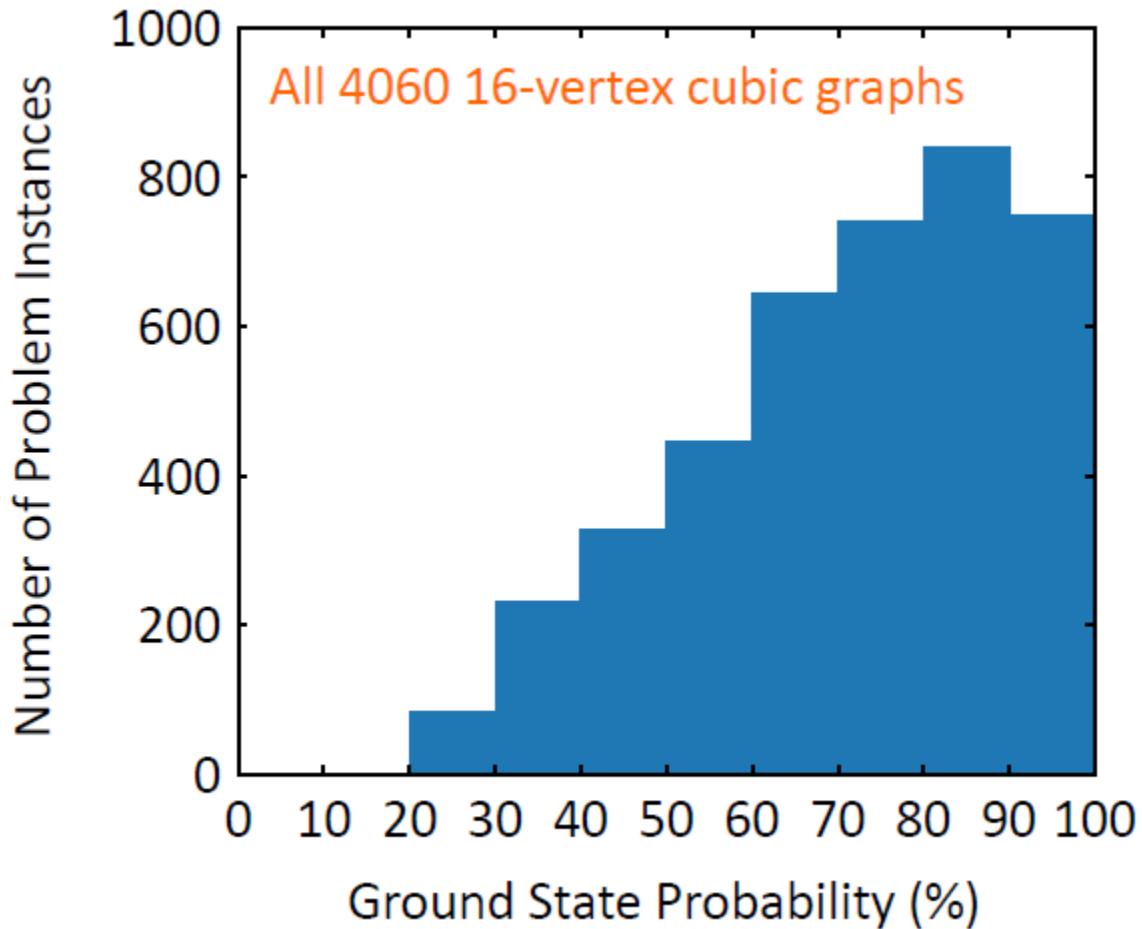
MAX-CUT on small graphs



MAX-CUT on small graphs

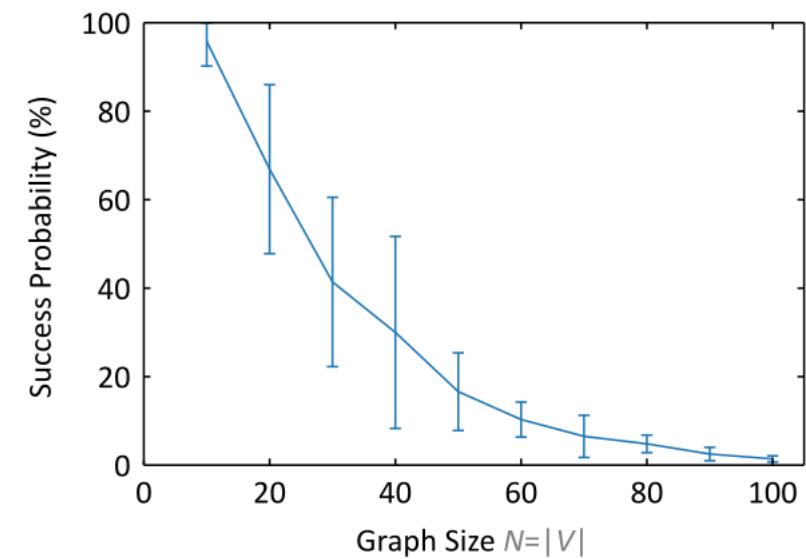


MAX-CUT on small graphs

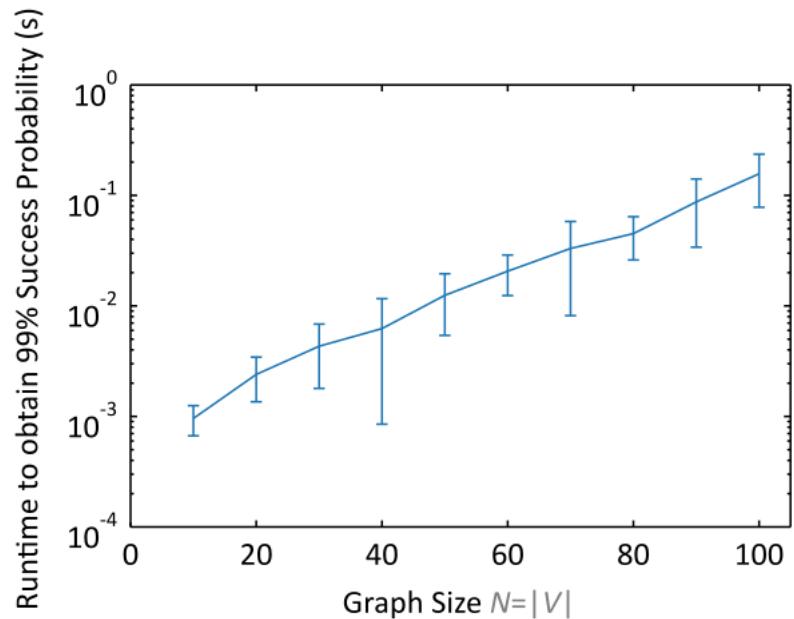
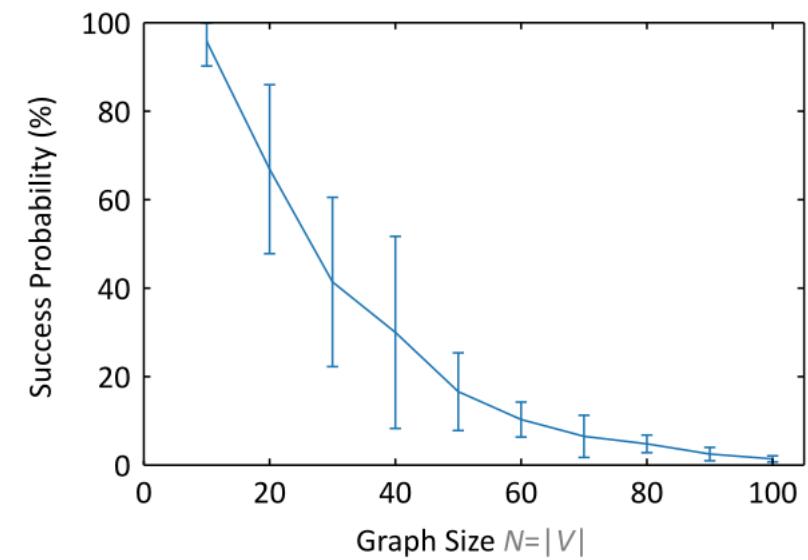


Main point: we can solve **every** $N=16$ cubic graph instance, so the previous success probabilities we showed were not just lucky data points.

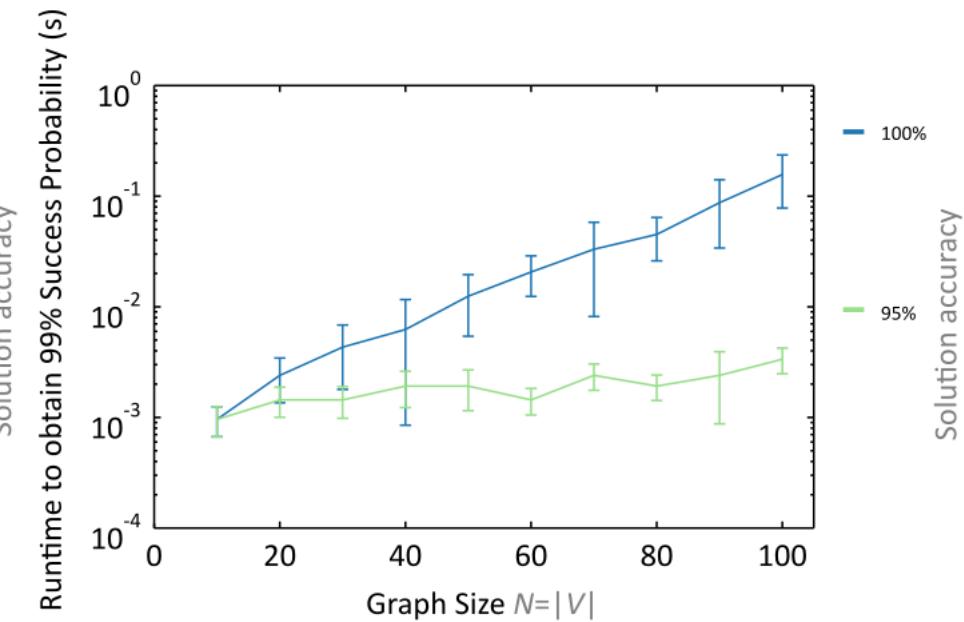
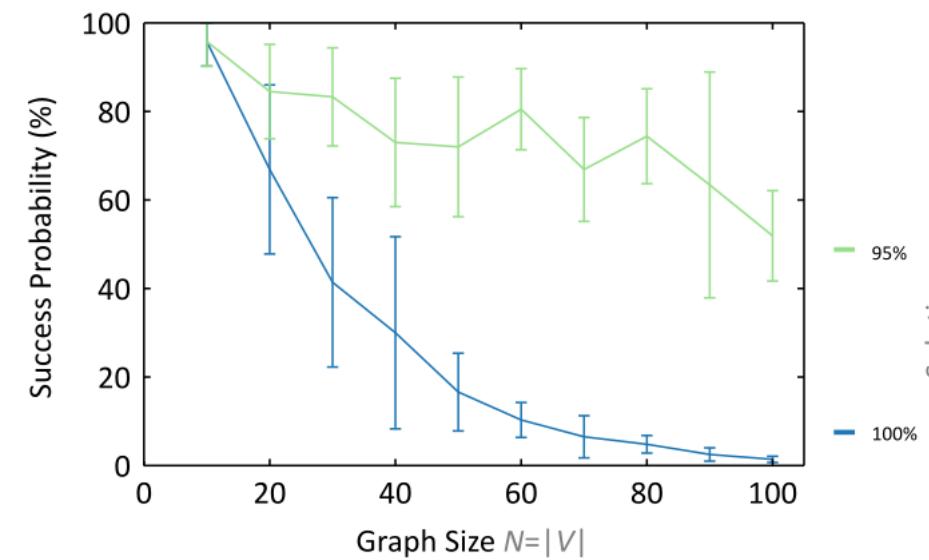
Scaling of cubic graphs



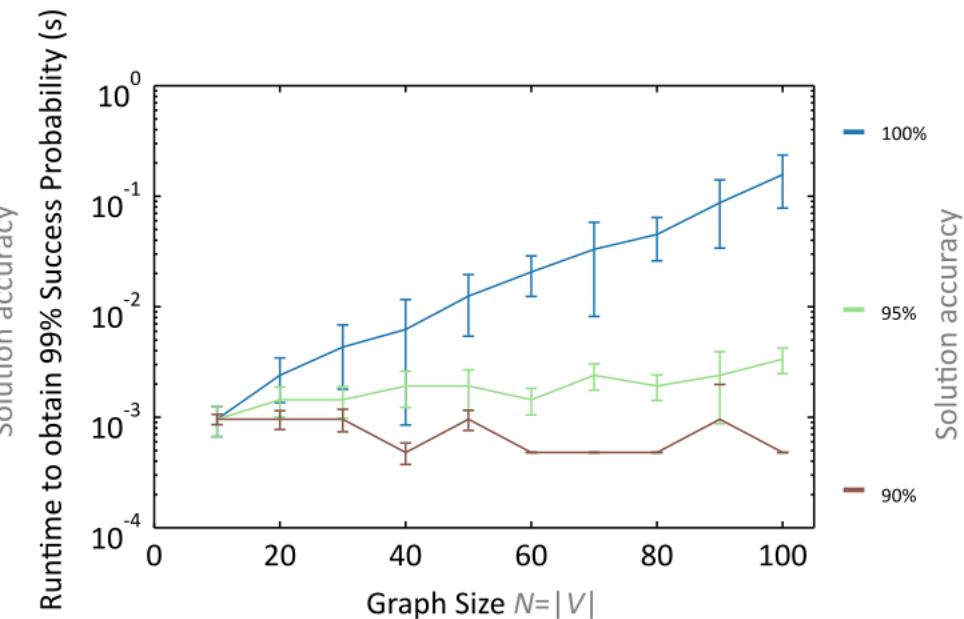
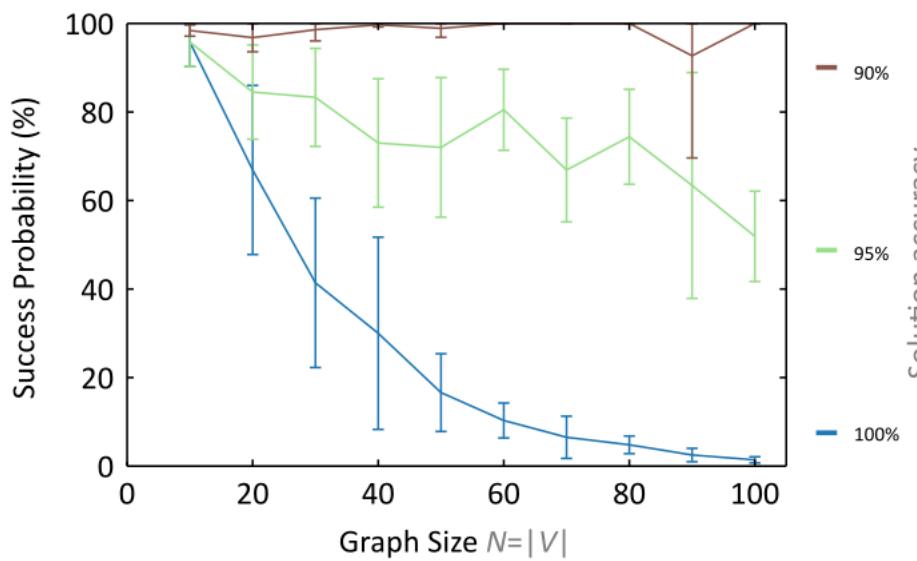
Scaling of cubic graphs



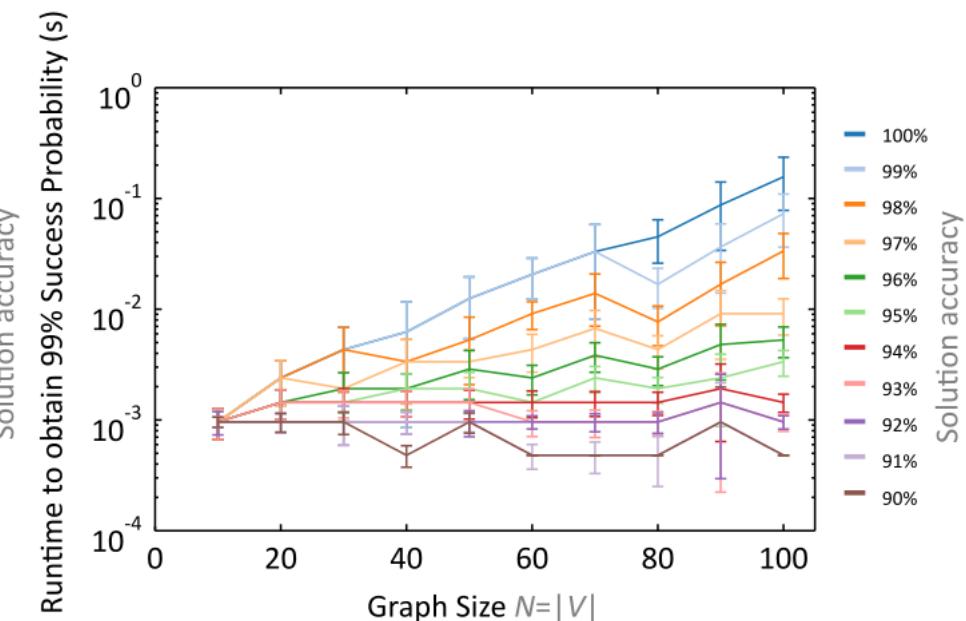
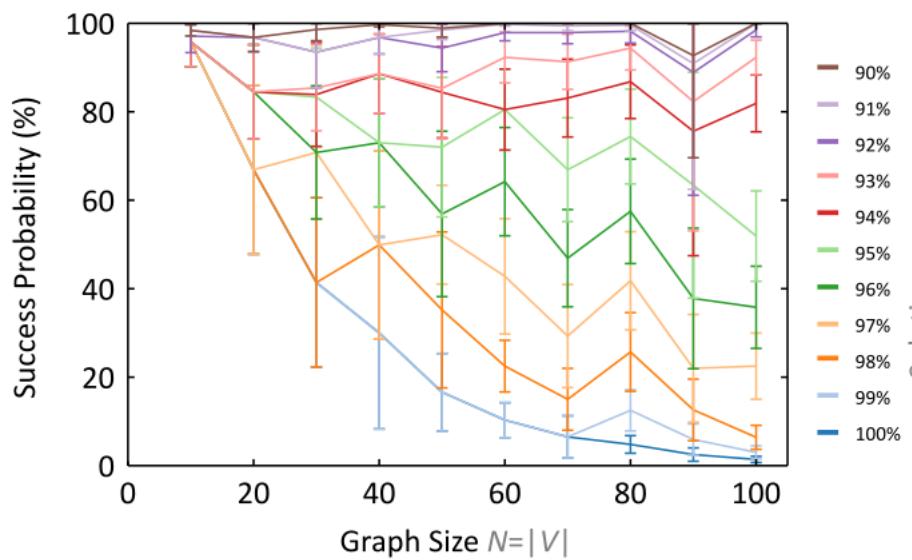
Scaling of cubic graphs



Scaling of cubic graphs

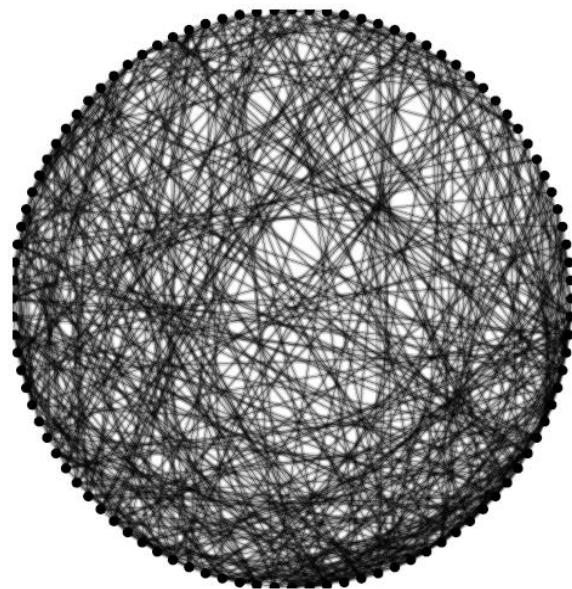


Scaling of cubic graphs



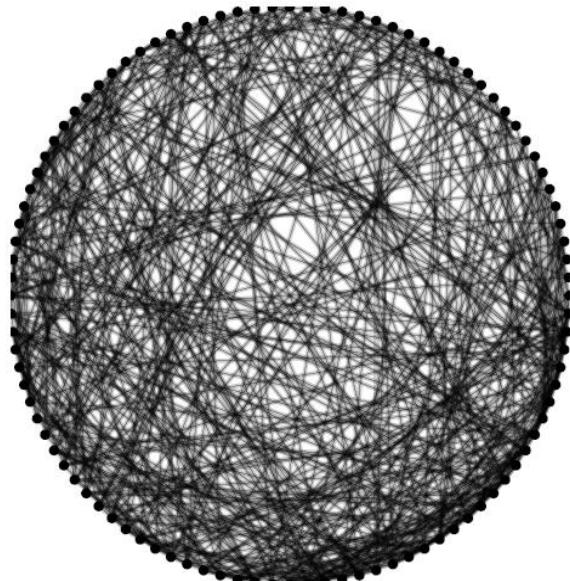
What about non-cubic graphs?

Dense(r) Random Graph

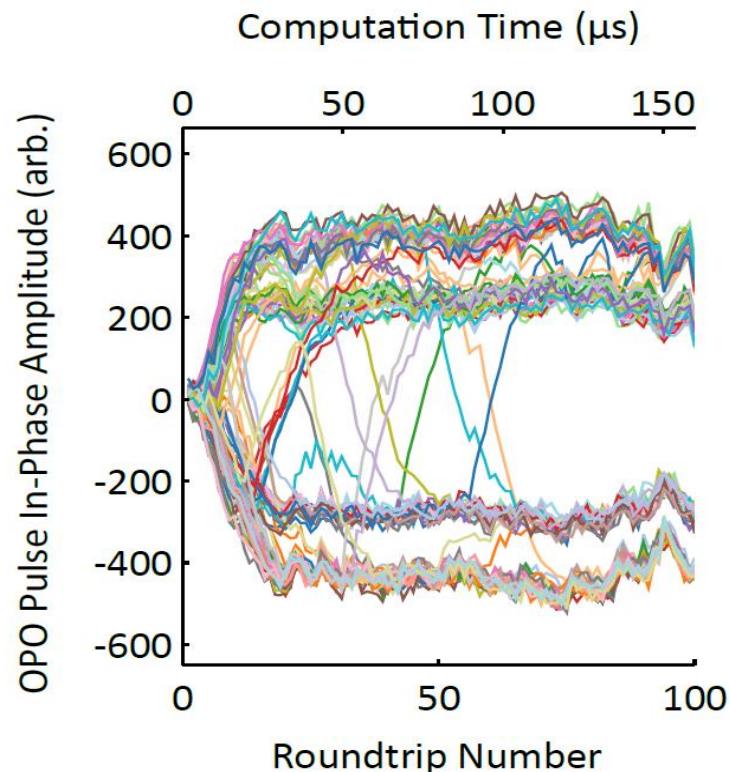


100 vertices; 495 edges

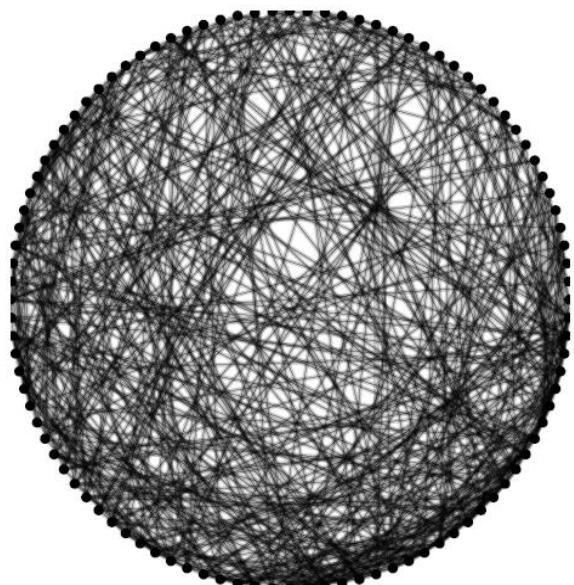
Dense(r) Random Graph



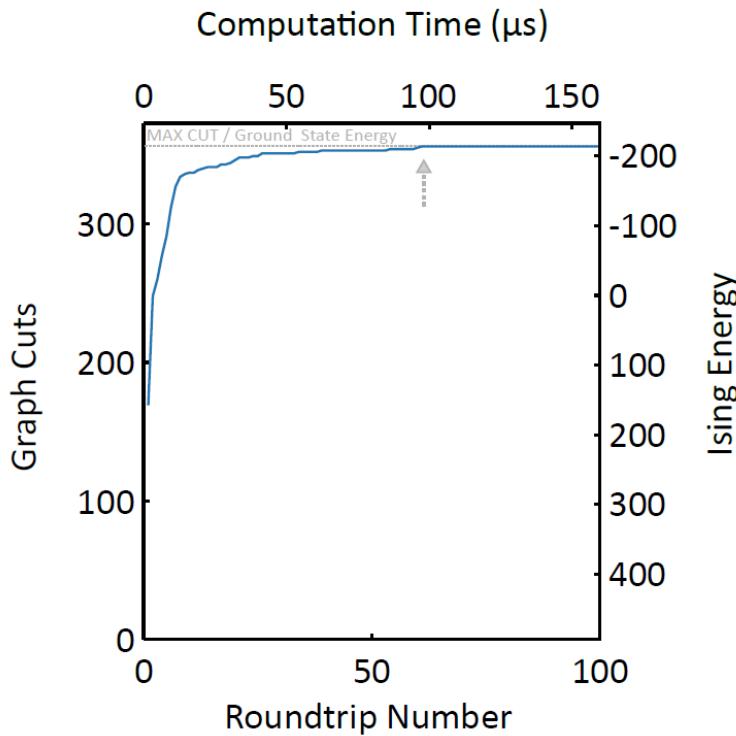
100 vertices; 495 edges



Dense(r) Random Graph



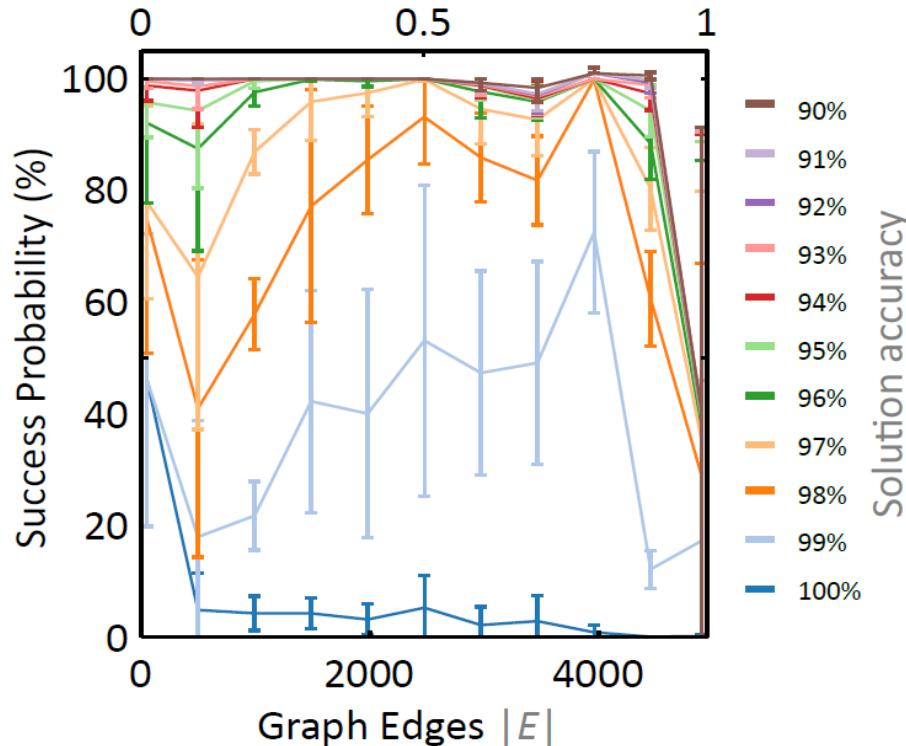
100 vertices; 495 edges



Sensitivity to Edge Density

Random graphs with $N=100$ vertices and $|E|$ edges

Graph Edge Density $d=2|E|/(|V|(|V|-1))$



Main point: system seems to be able to find exact and approximate solutions for a large range of graphs, including both sparse and dense graphs.

Comparison to Quantum Annealing

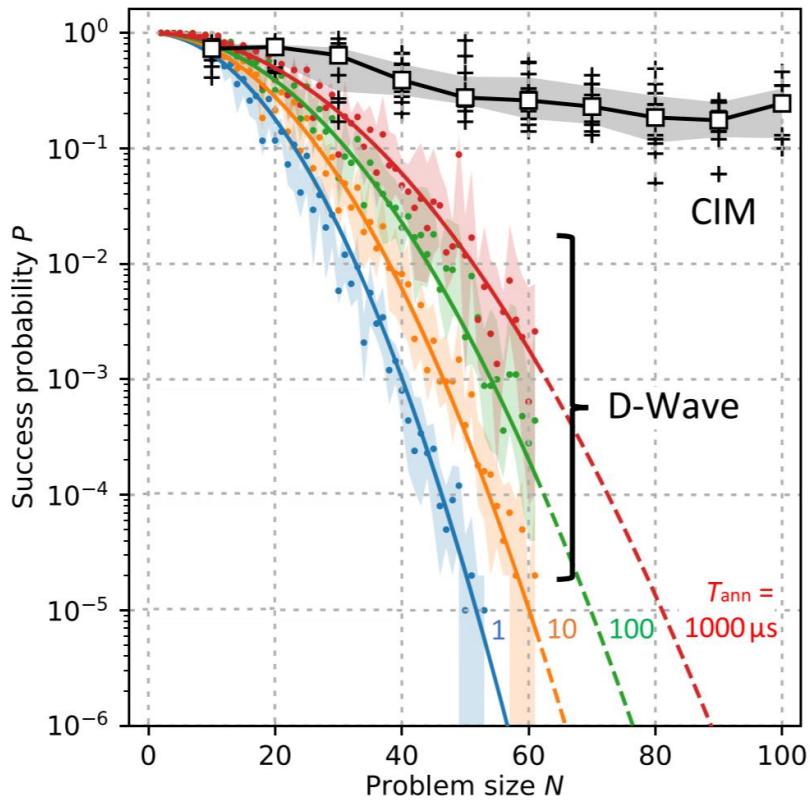


Credit: dwavesys.com

- M.W. Johnson, et al. *Nature* **473**, 194-198 (2011)
T.F. Ronnow, et al. *Science* **345**, 6195, 420-424 (2014)
S. Boixo, et al. *Nature Comm.* **7**, 10327 (2016)

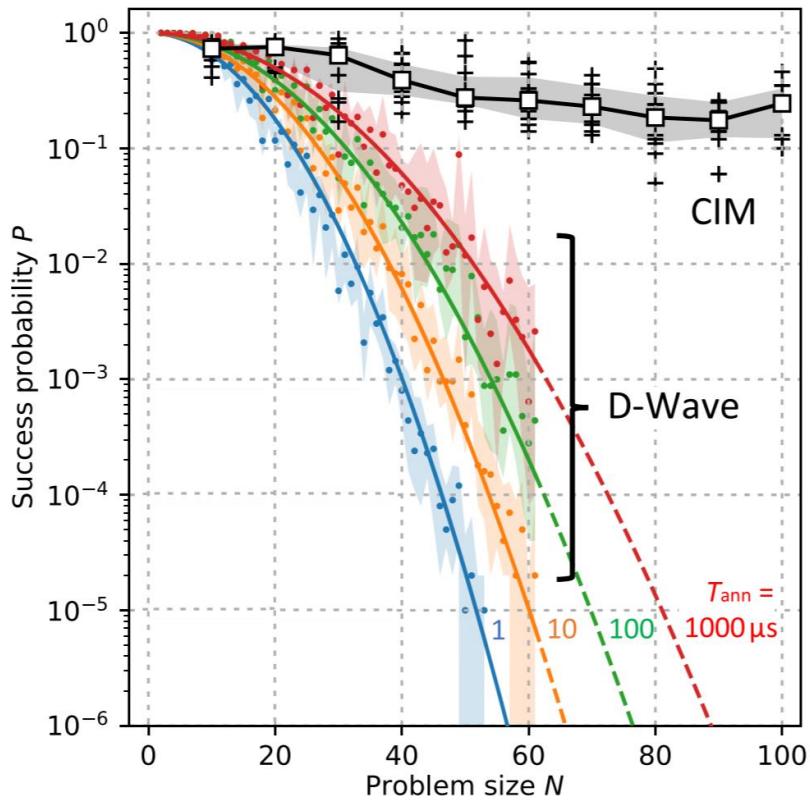
Comparison to Quantum Annealing

Problem class: Sherrington-Kirkpatrick spin-glass problems
(J_{ij} are -1 or +1 uniformly at random)



Comparison to Quantum Annealing

Problem class: Sherrington-Kirkpatrick spin-glass problems
(J_{ij} are -1 or +1 uniformly at random)

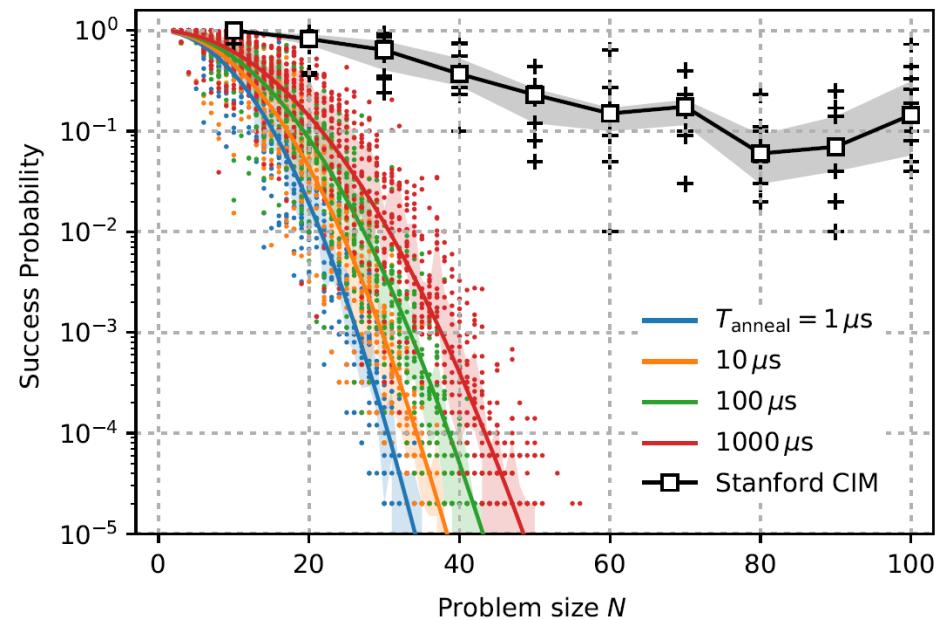


For any given size N , which annealing time should we choose?

If we want to predict performance for larger problem sizes than we can currently solve, which annealing time should we choose?

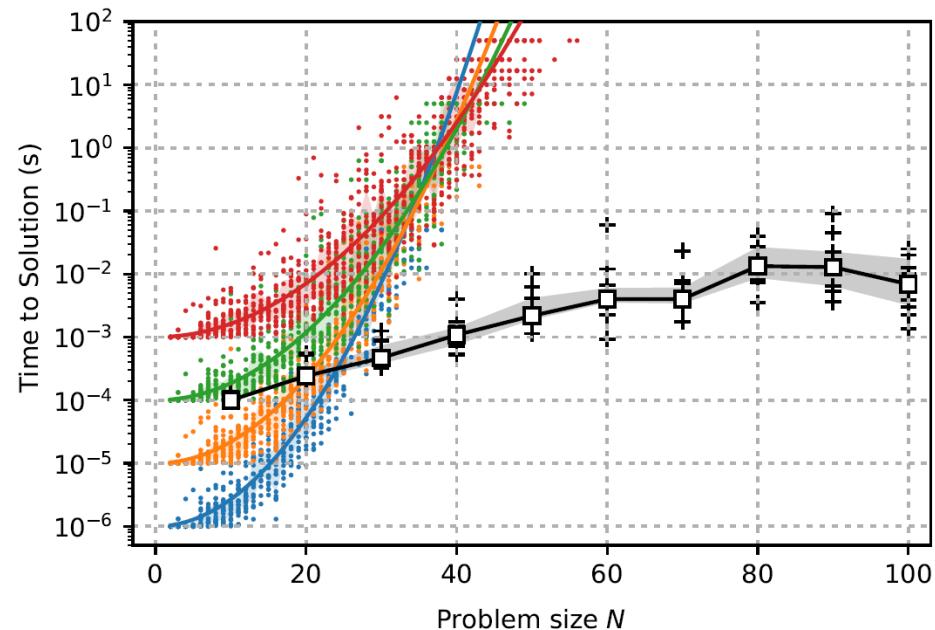
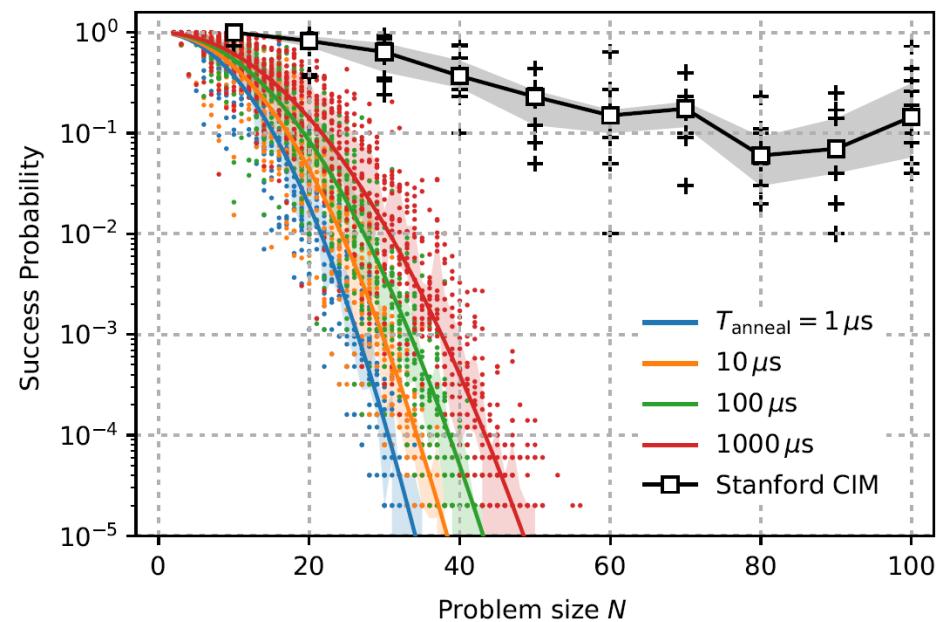
Comparison to Quantum Annealing

Problem class: MAX-CUT on unweighted graphs with edge density = 50%



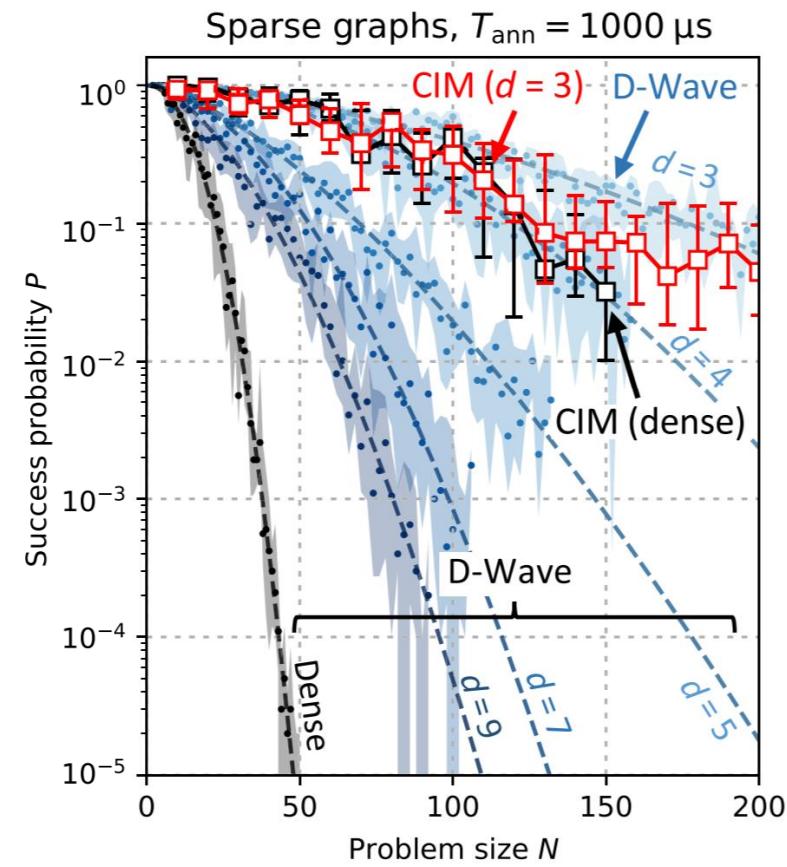
Comparison to Quantum Annealing

Problem class: MAX-CUT on unweighted graphs with edge density = 50%



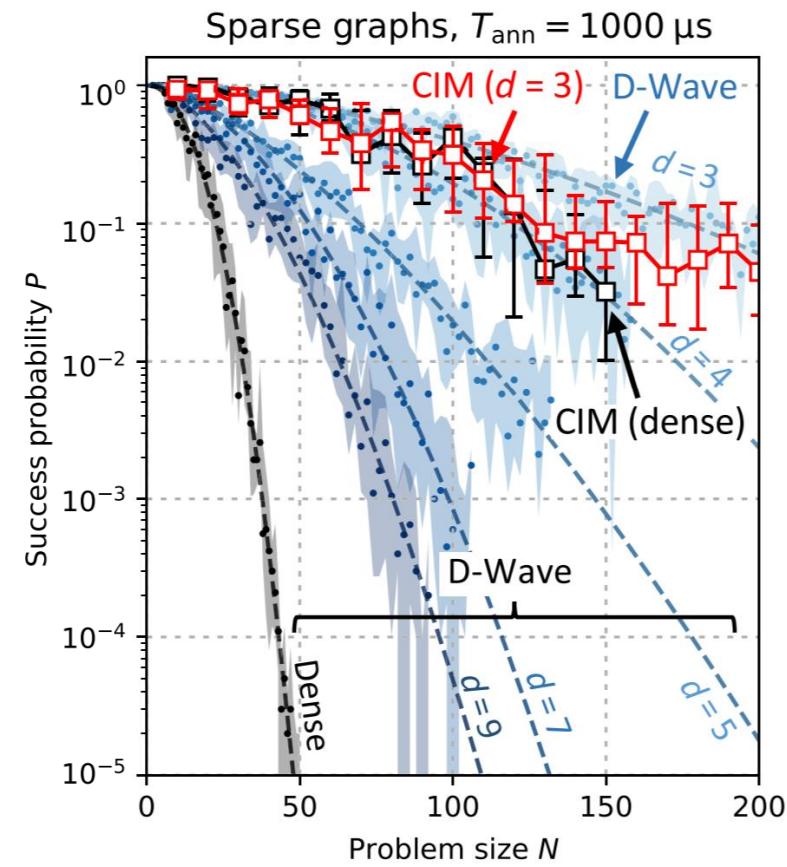
Comparison to Quantum Annealing

Problem class: MAX-CUT on regular graphs with degree d



Comparison to Quantum Annealing

Problem class: MAX-CUT on regular graphs with degree d

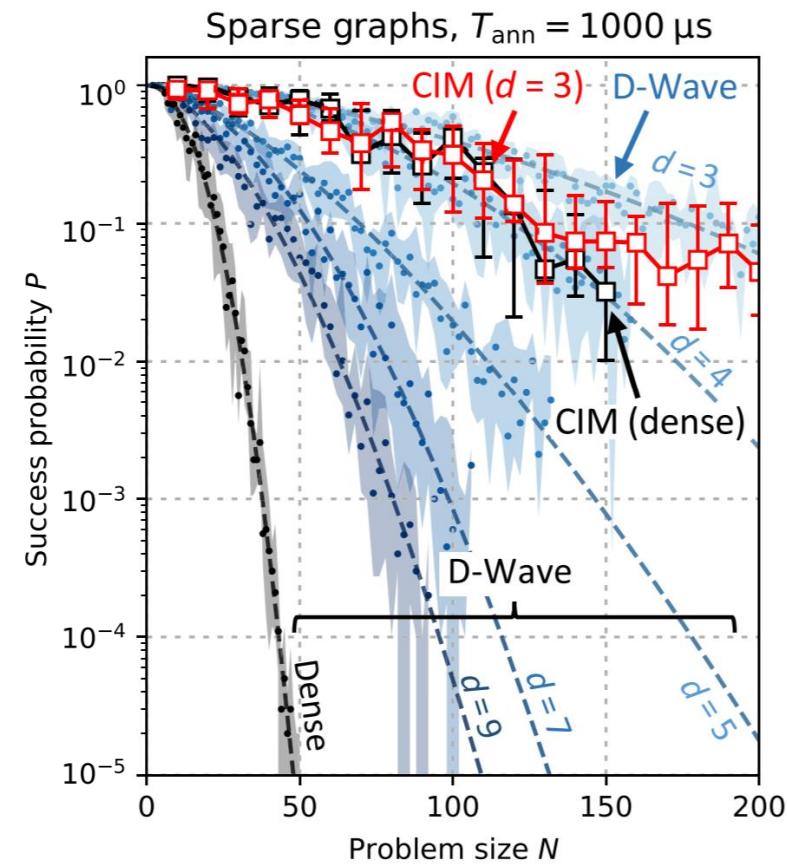


Connectivity makes a big difference!

Why?

Comparison to Quantum Annealing

Problem class: MAX-CUT on regular graphs with degree d



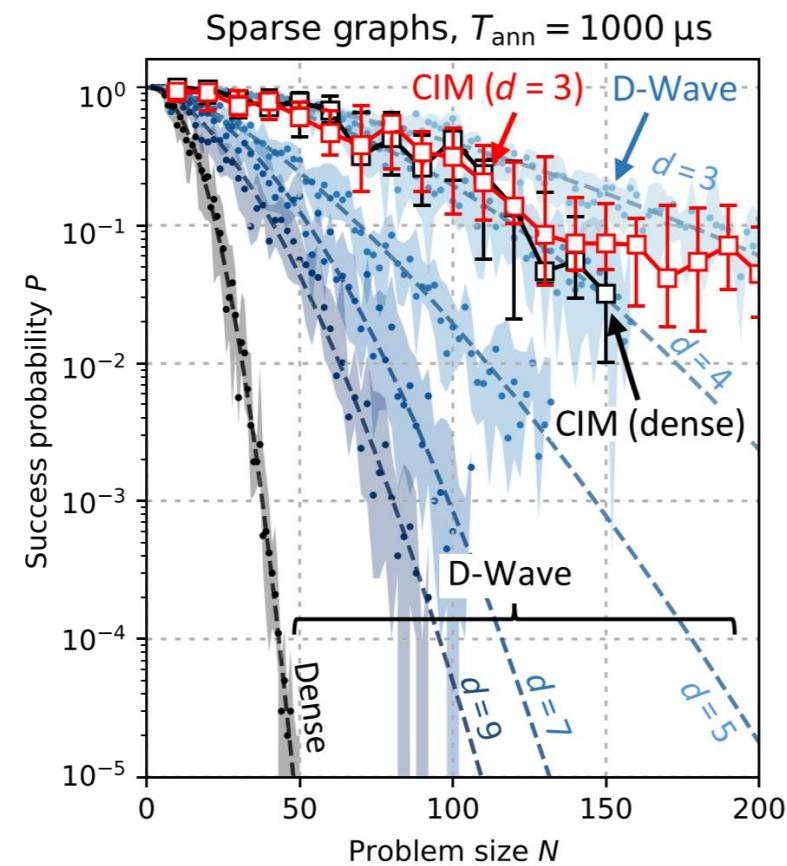
Connectivity makes a big difference!

Why?

Success probability $P \sim \exp(-\alpha N_{\text{physical}})$

Comparison to Quantum Annealing

Problem class: MAX-CUT on regular graphs with degree d



Connectivity makes a big difference!

Why?

Success probability $P \sim \exp(-\alpha N_{\text{physical}})$

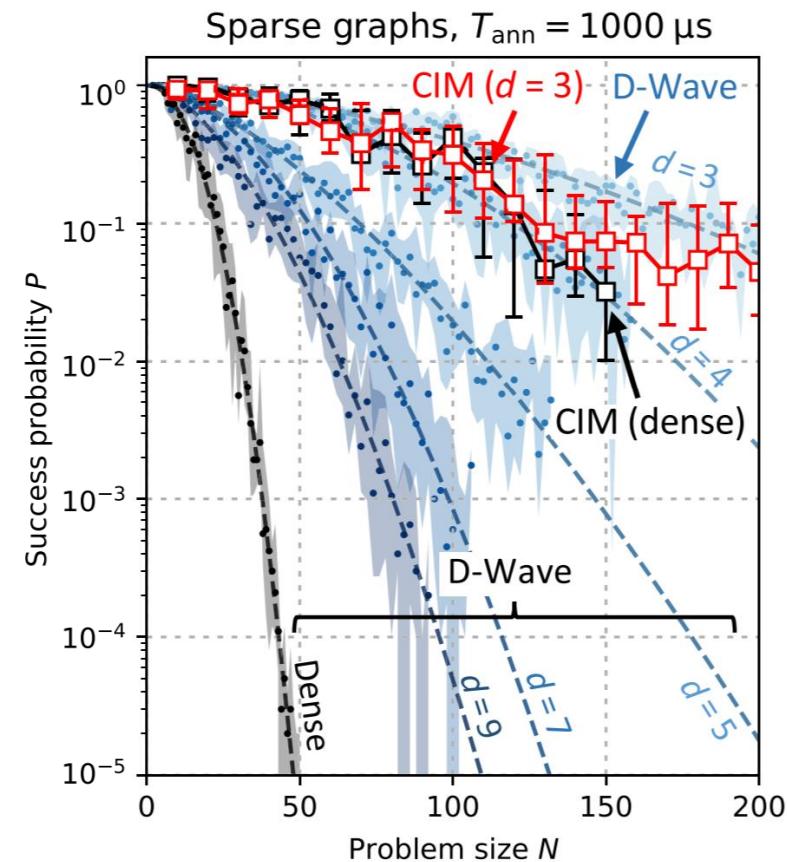
For **dense** problems: $N_{\text{physical}} \sim N^2$

For **sparse** problems: $N_{\text{physical}} \sim N$

Note: α depends on the machine or algorithm

Comparison to Quantum Annealing

Problem class: MAX-CUT on regular graphs with degree d



Connectivity makes a big difference!

Why?

Success probability $P \sim \exp(-\alpha N_{\text{physical}})$

For **dense** problems: $N_{\text{physical}} \sim N^2$

For **sparse** problems: $N_{\text{physical}} \sim N$

Note: α depends on the machine or algorithm

How can we make quantum annealers with better connectivity?
See: Onodera et al. *npj Quantum Information* **6**, 48 (2020).

Comparison to Quantum Annealing

SK				MAX-CUT (dense)				MAX-CUT ($d = 3$)			
N	DW2Q	CIM	Factor	N	DW2Q	CIM	Factor	N	DW2Q	CIM	Factor
10	6.0 μ s	25 μ s	0.2	10	6.0 μ s	25 μ s	0.2	10	1.0 μ s	50 μ s	0.02
20	35 μ s	100 μ s	0.3	20	0.4 ms	100 μ s	4	20	3.0 μ s	100 μ s	0.03
40	6.1 ms	0.4 ms	15	40	6.1 s	0.4 ms	10^4	50	12 μ s	0.4 ms	0.03
60	1.4 s	0.6 ms	2000	55	10^4 s	1.2 ms	10^7	100	100 μ s	3.3 ms	0.03
80*	(400 s)	1.8 ms	(10^5)	80*	(10^{11} s)	1.8 ms	(10^{13})	150	2.8 ms	22 ms	0.1
100*	(10^5 s)	3.0 ms	(10^7)	100*	(10^{19} s)	2.3 ms	(10^{21})	200	11 ms	51 ms	0.2

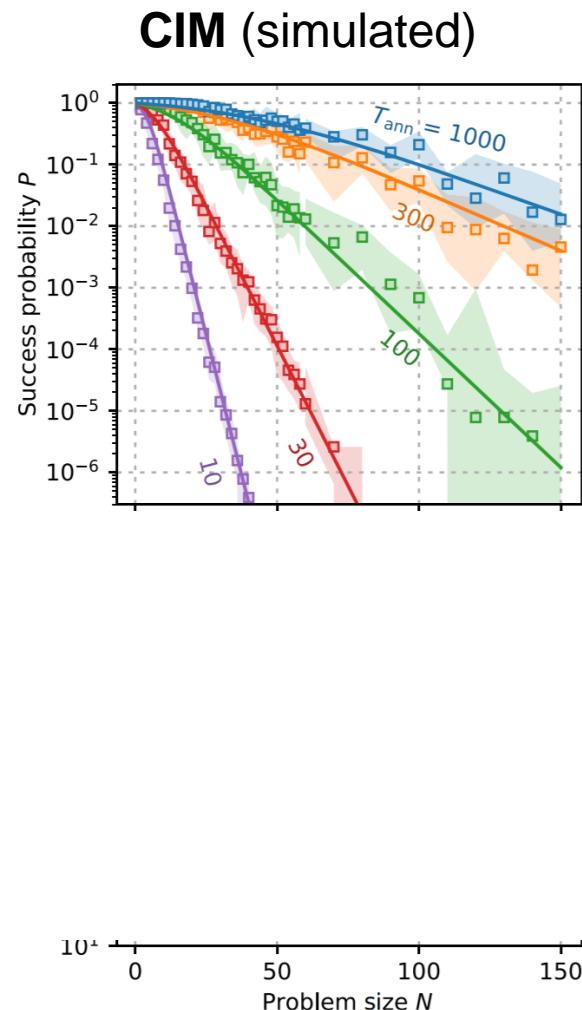
Fully-connected graphs

50%-density graphs

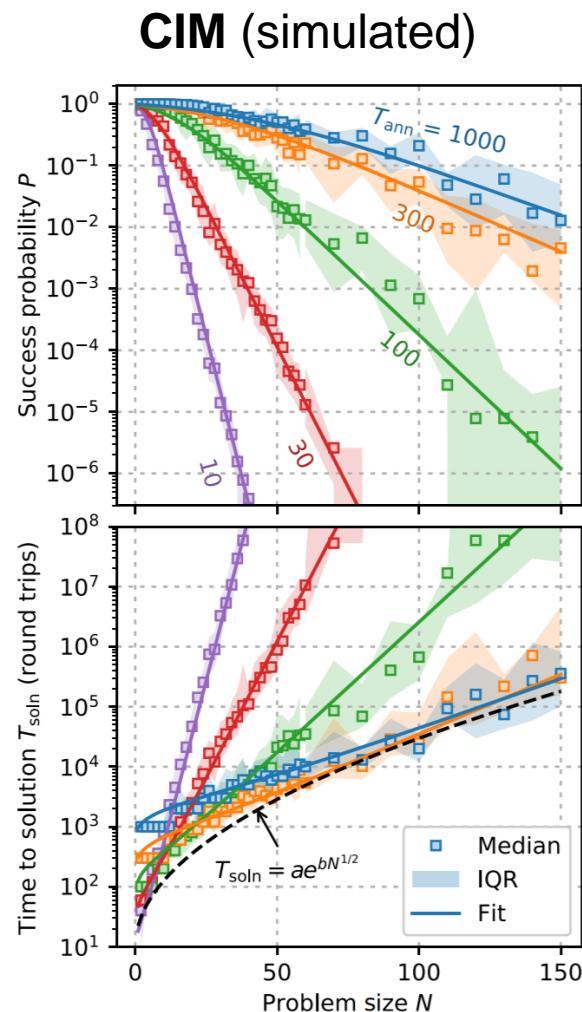
Cubic graphs

Note: CIM runtimes are comparable to those of a laptop running a state-of-the-art solver
→ no speedup from use of optics yet!

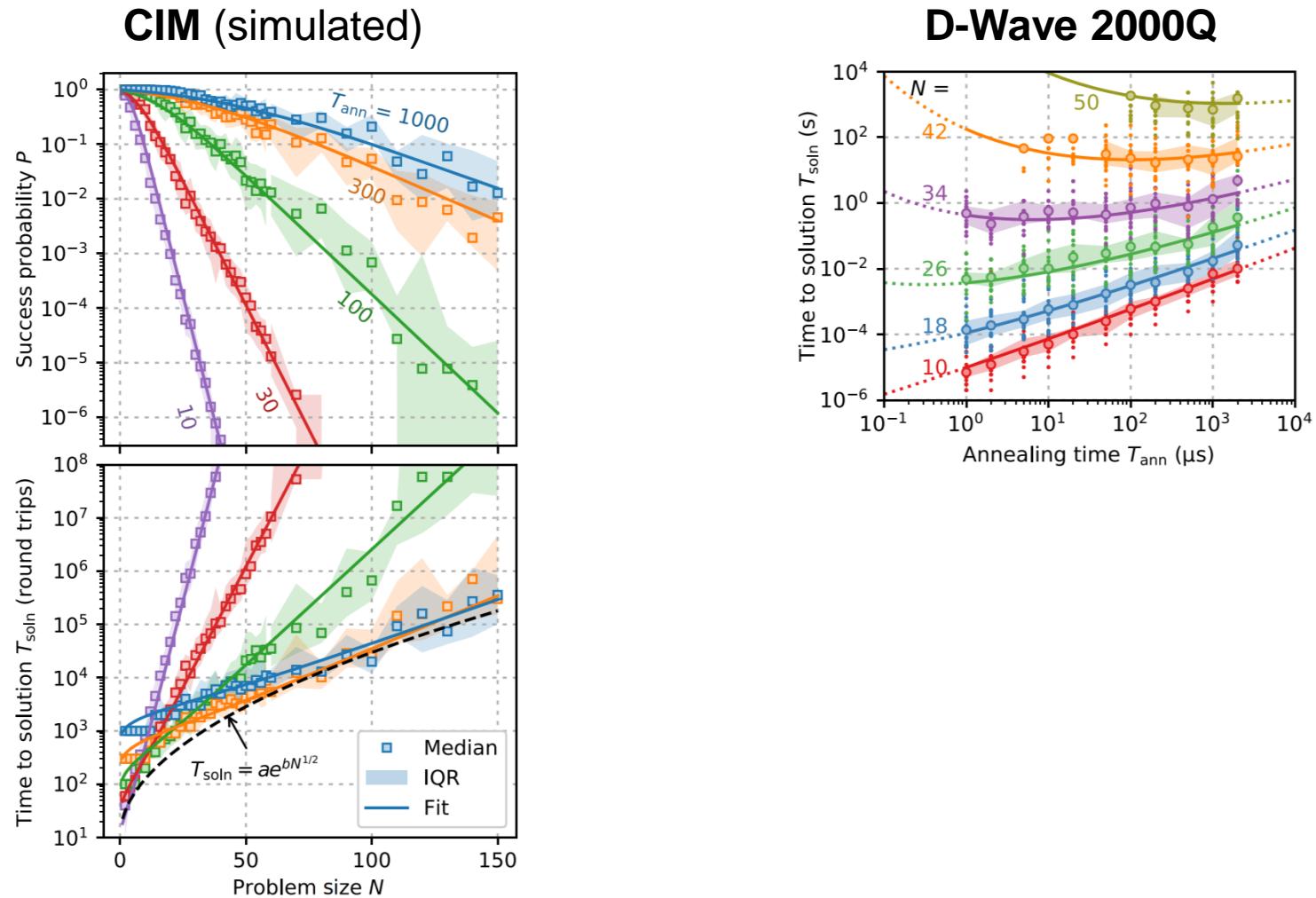
Choosing Optimal Annealing Times



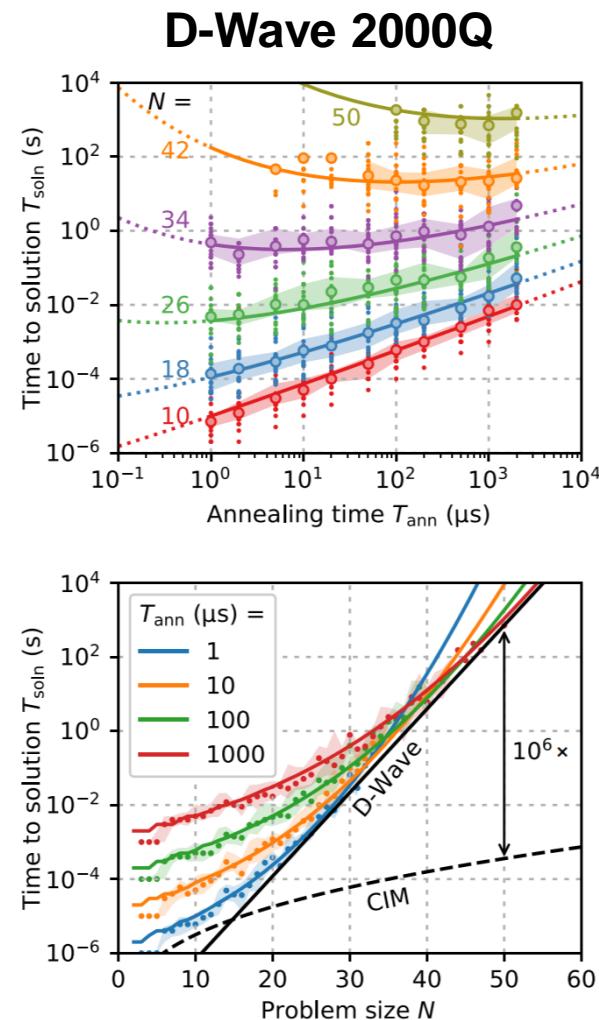
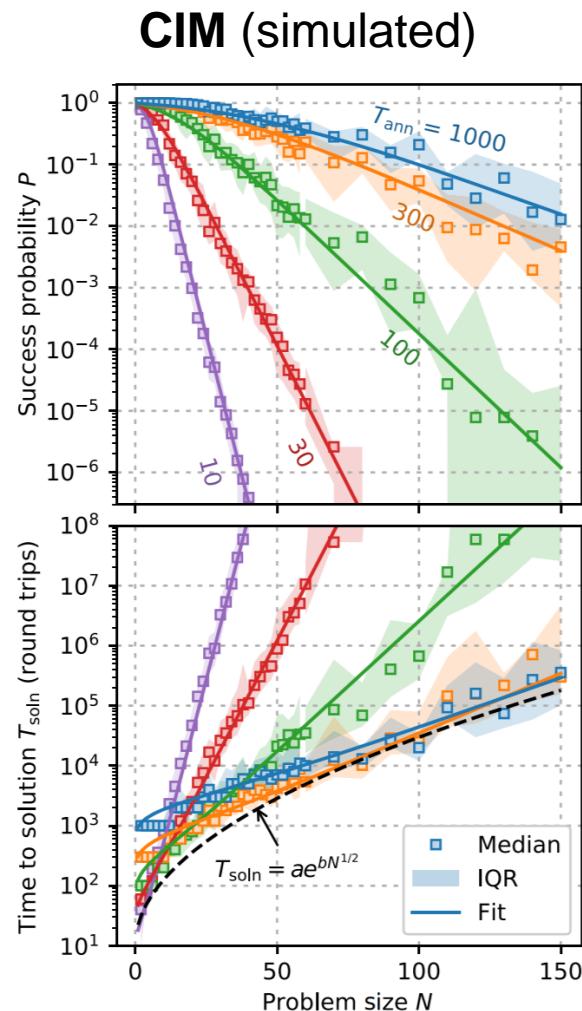
Choosing Optimal Annealing Times



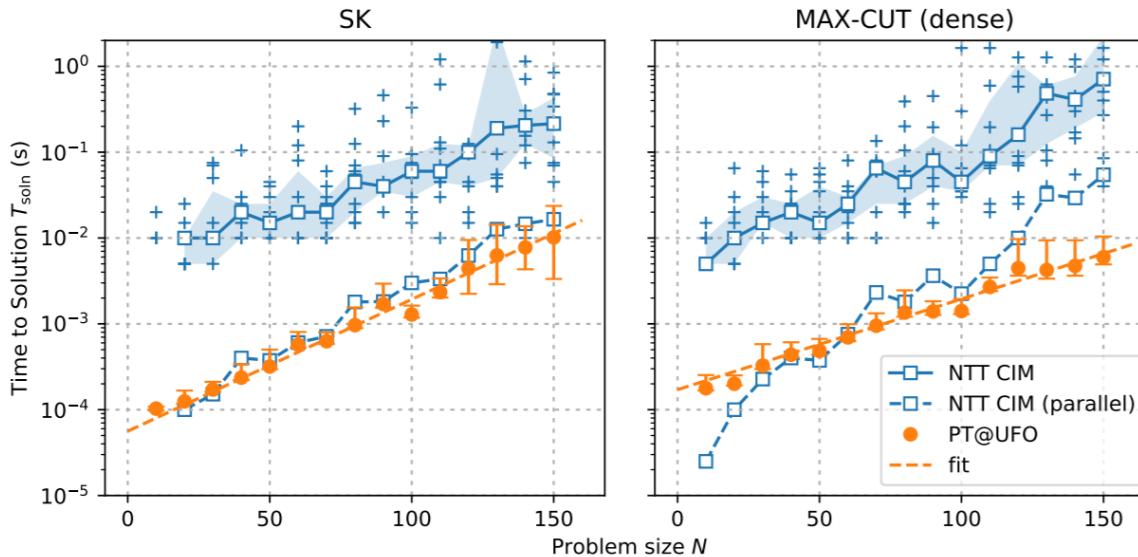
Choosing Optimal Annealing Times



Choosing Optimal Annealing Times



Comparison to Classical State-of-the-Art



Overview

- The Ising Problem
- Ising Machines
- Foundations of All-Optical OPO Ising Machines
- Measurement-Feedback OPO Ising Machines
- Conclusions

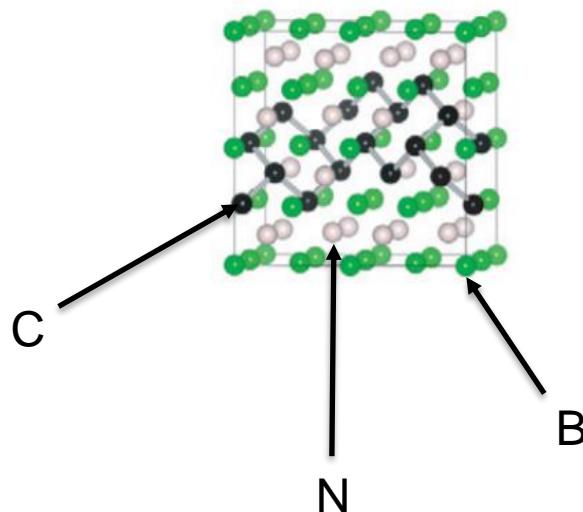
Summary

- Networks of **coupled OPOs** provide an alternative platform for physically emulating Classical **Ising Spin Hamiltonians**.
- $N = 100$ spin system with **10,000 spin-spin connections** (all-to-all) has been implemented.
- System can find **exact** and **approximate** solutions for a large range of graphs.
- **Time-division multiplexing** and **measurement-feedback** provide the tools to allow scalable all-to-all connectivity in optical systems, and may have some relevance to AQC.

Backup Slides

Example Application: Cluster Expansion for Materials

Cubic Carbon Boron Nitride (c-BNC)



Ising problem with spins: $\sigma_i^{(P)} = \begin{cases} 1 & P \text{ atom at site } i \\ 0 & \text{no } P \text{ atom at site } i \end{cases}$

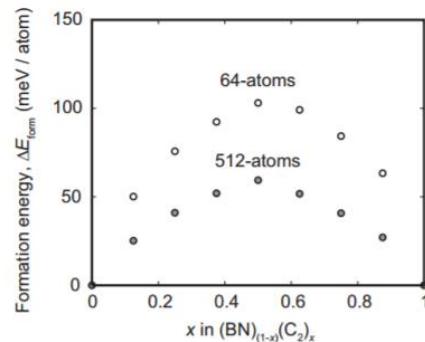


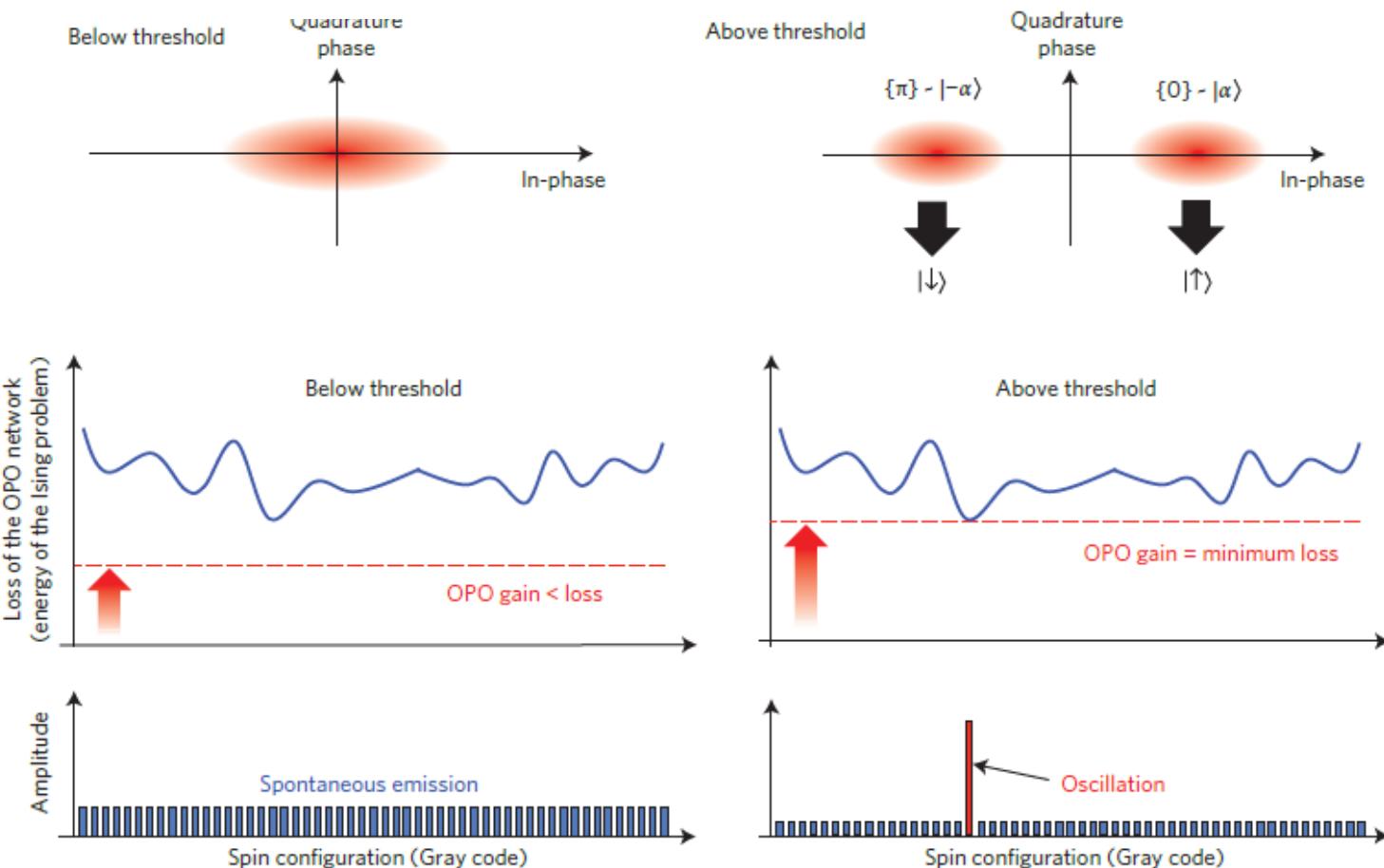
FIG. 4. Ground-state convex hull for c-BNC along the composition range of $(BN)_{(1-x)}(C_2)_x$ ($0 \leq x \leq 1$). Open and closed circles denote the atomic arrangements in the lowest formation energy within the supercells consisting of 64 and 512 atoms, obtained from the MC simulation with simulated annealing algorithm.

Yuge, et al. *Phys. Rev. B* **77**, 094121 (2008)

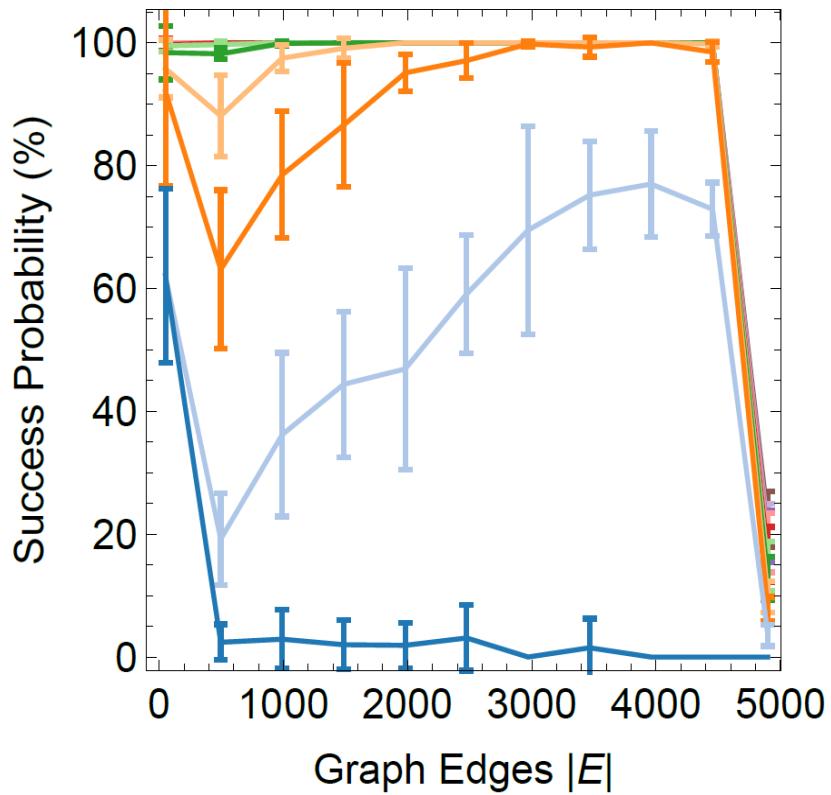
OPO Ising Machine Mechanism

$$H = - \sum_{ij}^N J_{ij} \sigma_i \sigma_j$$

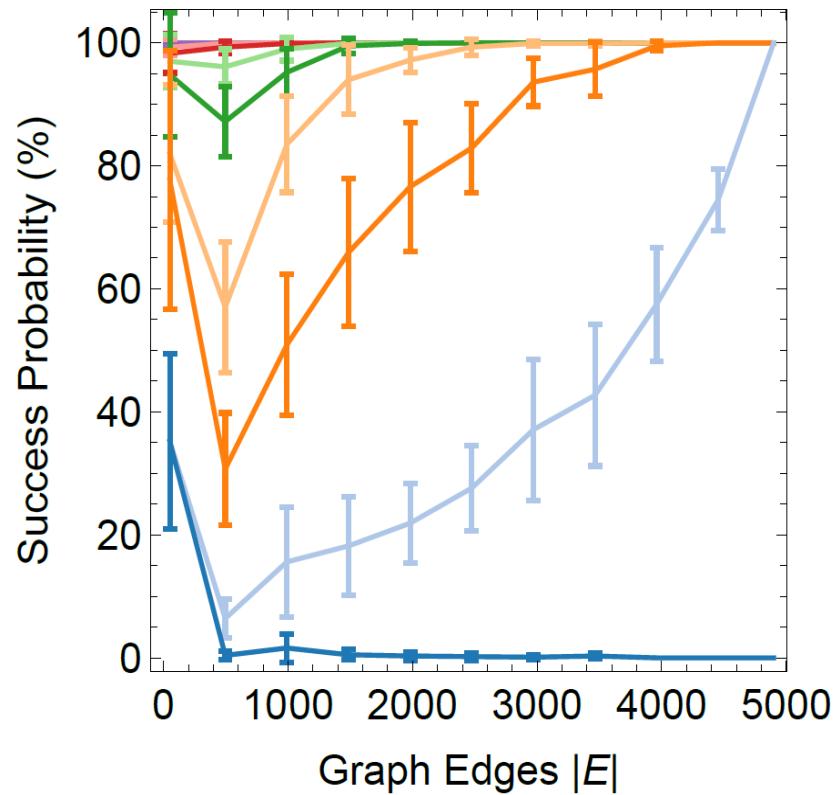
→ design interactions in a network of OPOs to realize this Hamiltonian



Simulations: Density Sensitivity

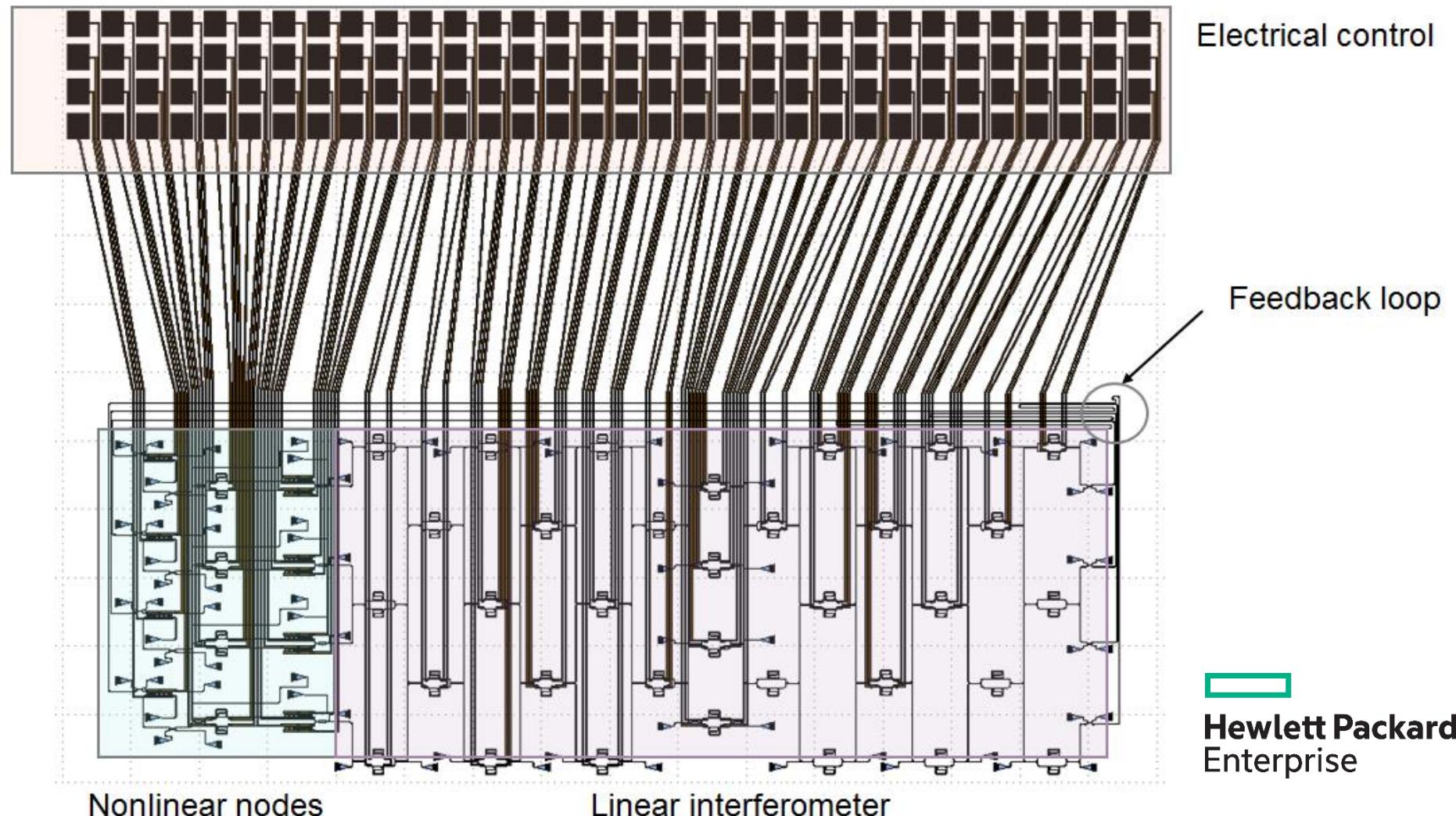


$$p_{\text{comp}} = 0.6 p_{\text{th}}$$

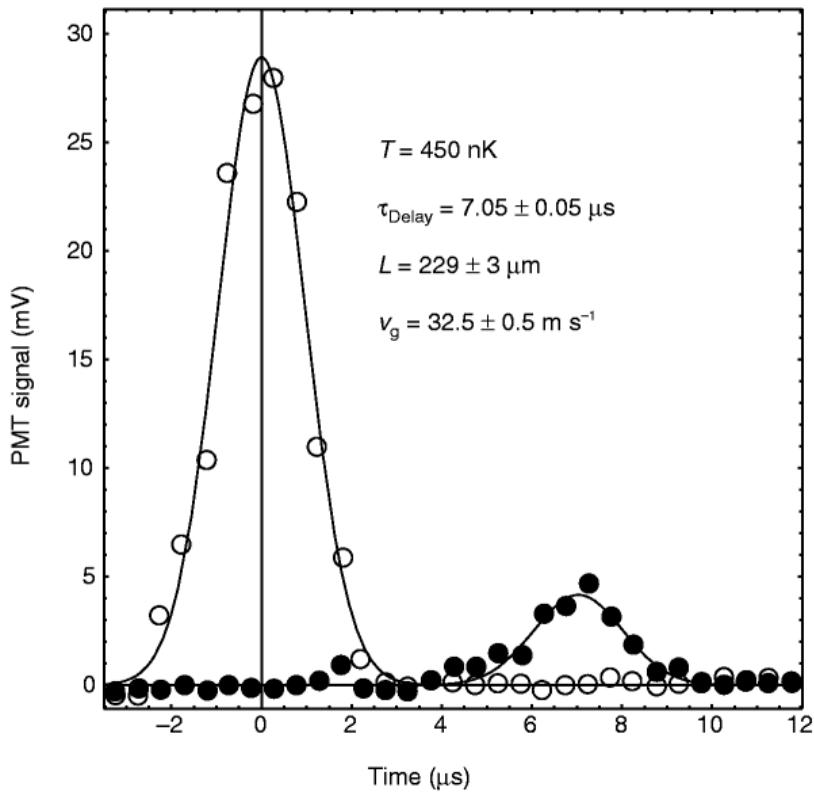


$$p_{\text{comp}} = 1.3 p_{\text{th}}$$

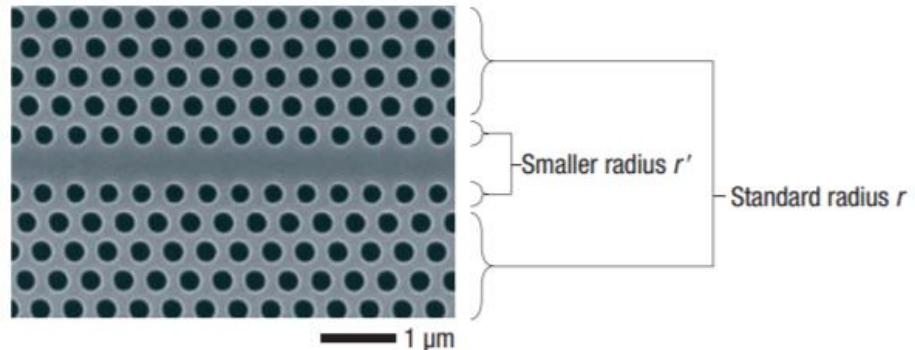
Microring Resonator Network



On-Chip All-Optical Ising Machines using Slow Light



L. Hau, et al. *Nature* **397**, 594-598 (1999)



T. Baba. *Nature Photonics* **2**, 465 (2008)

