

Towards Designing and Deploying Ising Machines

Sachin S. Sapatnekar
ECE Department
University of Minnesota

Grateful acknowledgments: Hüsrev Cılasun, Ahmet Efe, Ulya R. Karpuzcu, Chris H. Kim,
Abhimanyu Kumar, Hao Lo, Will Moy, Nafisa Sadaf Prova, Ziqing Zeng

Funding acknowledgments: DARPA QuICC, NSF ASCENT



UNIVERSITY OF MINNESOTA
Driven to Discover®

ISPD 2025

The concept



UNIVERSITY OF MINNESOTA

Driven to Discover®

A tale of dancing metronomes

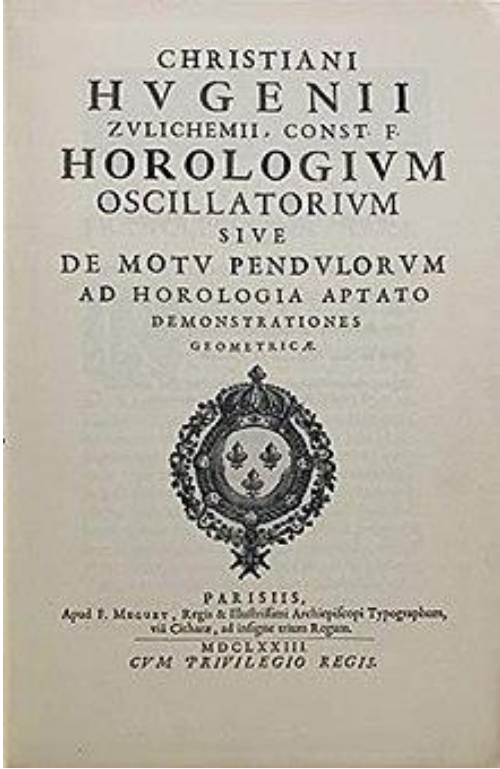


[<https://youtu.be/Aaxw4zbULMs>]

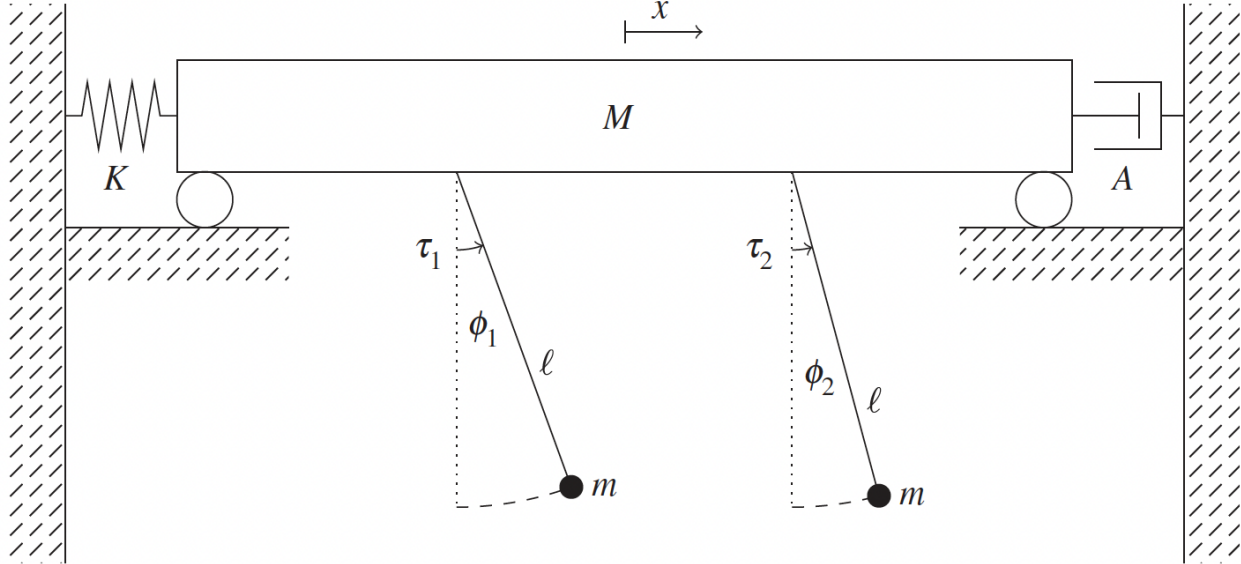
Huygens's clocks (1665)



Christiaan Huygens (1629 – 1695)
by Caspar Netscher (1671), Museum Boerhaave, Leiden



The Pendulum Clock: or
Geometrical Demonstrations
Concerning the Motion of
Pendula as Applied to Clocks
(1673)



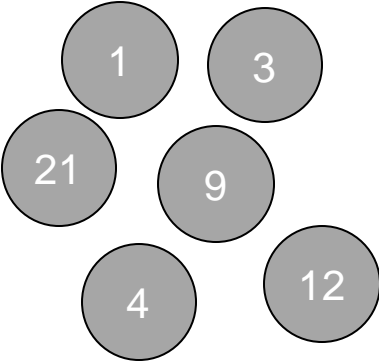
Willms, "Huygens clocks revisited," Royal Society Open Science, 4:170777, 2017.



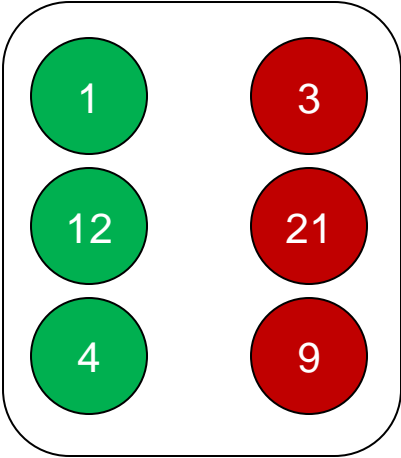
Nice toy, huh?

- But you can also do useful computation with this
 - State = phase of the oscillator
- We can solve optimization problems, e.g., number partitioning (NP-complete)
 - Divide a set of numbers into two subsets whose sum is equal (or as close as possible)

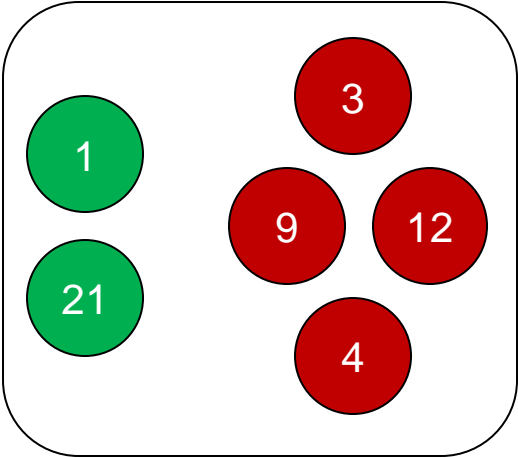
Example



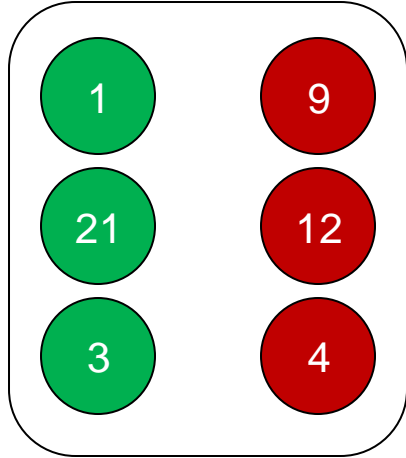
Candidate solutions



17 33



22 28

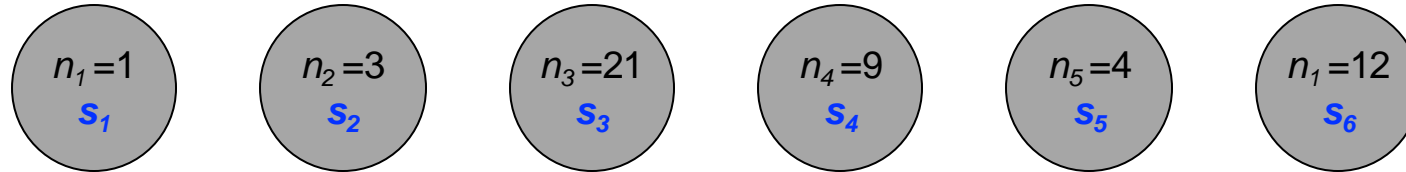


25 25



A mathematical formulation

- Assign variables $s_i \in \{-1, +1\}$ to each number n_i



- Formulate the problem as minimizing a quadratic

$$\text{minimize } (\sum s_i n_i)^2$$

$$\text{e.g., minimize } (1 s_1 + 3 s_2 + 21 s_3 + 9 s_4 + 4 s_5 + 12 s_6)^2$$

- This has terms $s_i^2 = 1$ and $s_i s_j$, and converts to

$$\text{minimize } \sum K_{ij} s_i s_j$$

- Note: Need one reference state! $s_1 = -1 \Rightarrow \text{minimize } (-1 + 3 s_2 + 21 s_3 + 9 s_4 + 4 s_5 + 12 s_6)^2$

What does this have to do with metronomes?

- Assign variables $s_i \in \{-1, +1\}$ to each number n_i

$n_1=1$
 $s_1=-1$

Reference

$n_2=3$
 $s_2=+1$

$n_3=21$
 $s_3=+1$

$n_4=9$
 $s_4=+1$

$n_5=4$
 $s_5=-1$

$n_6=12$
 $s_6=-1$

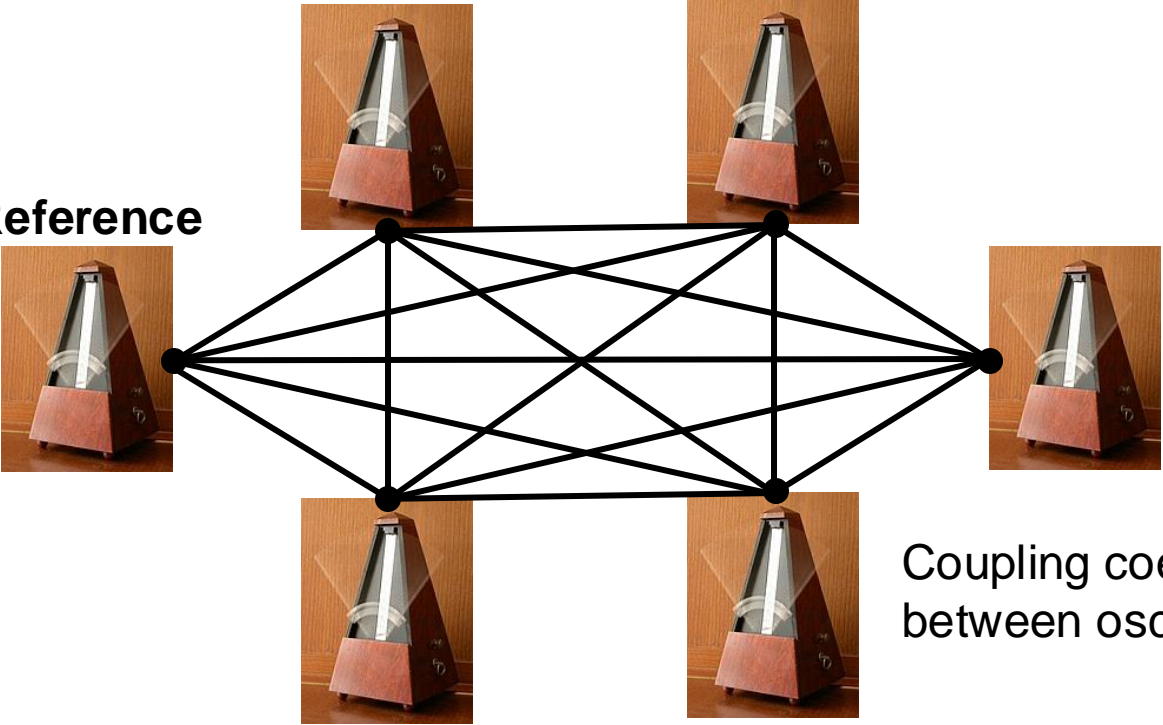
$s_i = -1$

$s_j = +1$



Reference

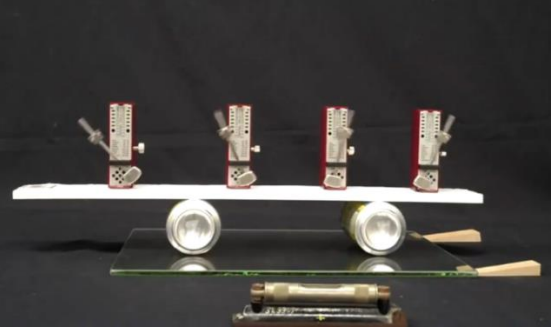
minimize $\sum K_{ij} s_i s_j$



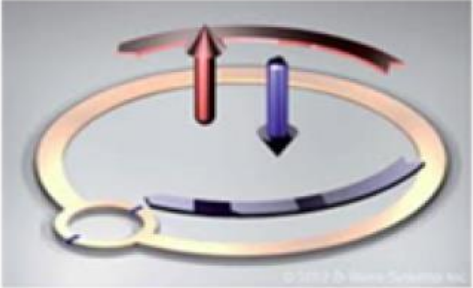
Coupling coefficients K_{ij} between oscillators i and j



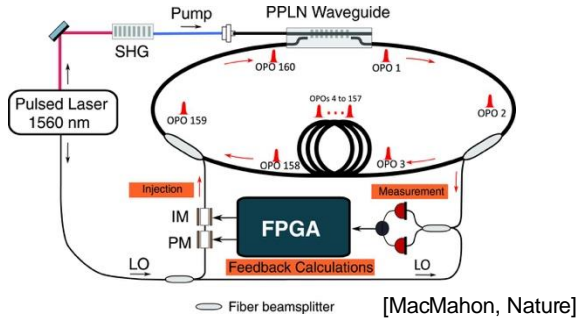
Coupled oscillator technologies



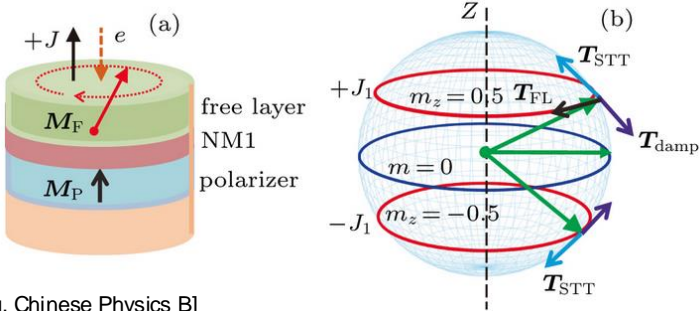
Metronomes



Superconducting

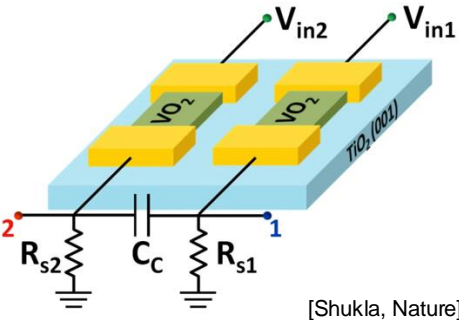


Optical



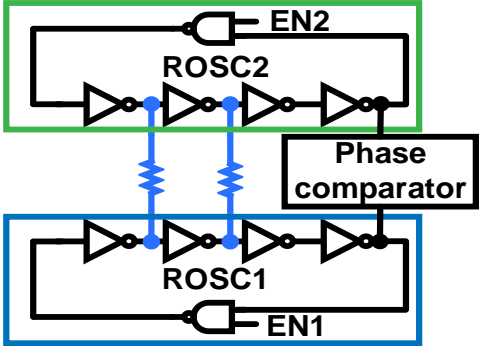
[Zheng, Chinese Physics B]

Spintronic



[Shukla, Nature]

Phase change

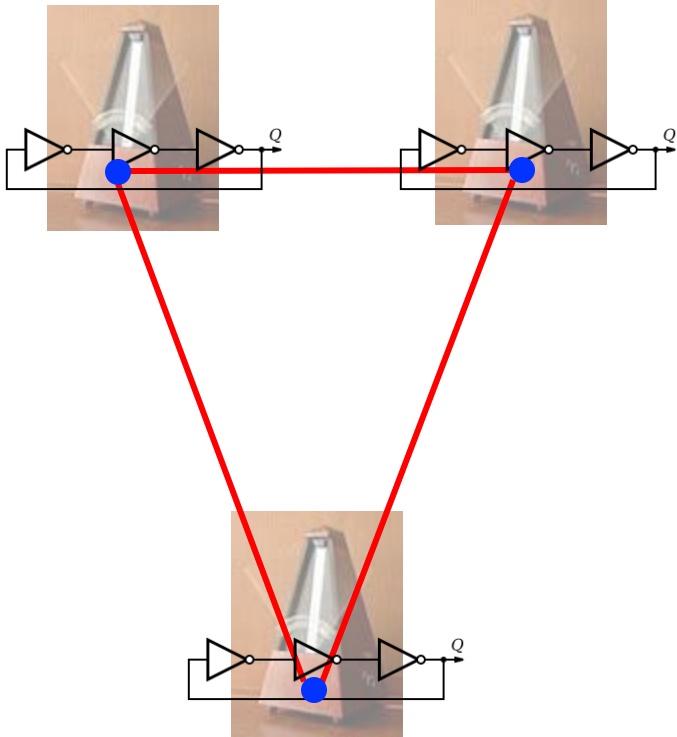


CMOS

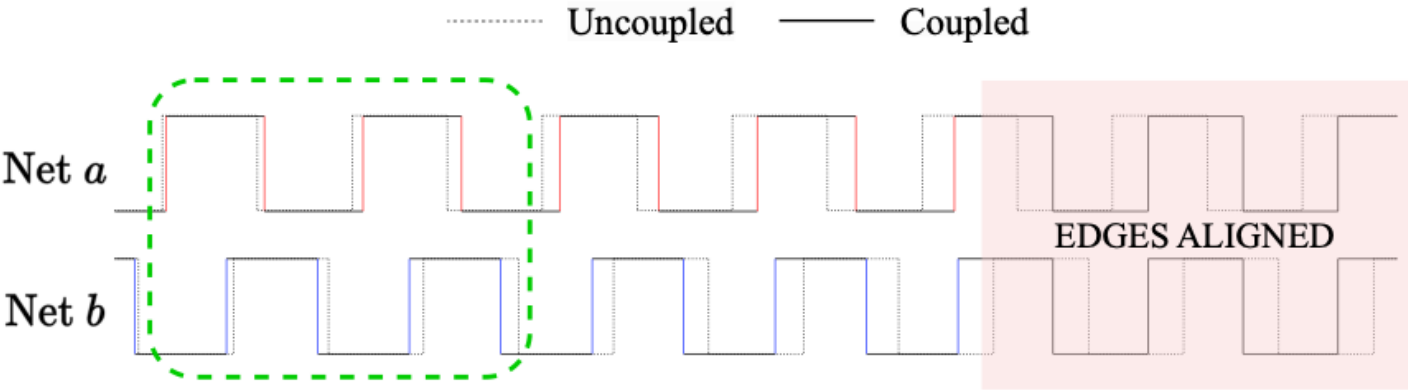
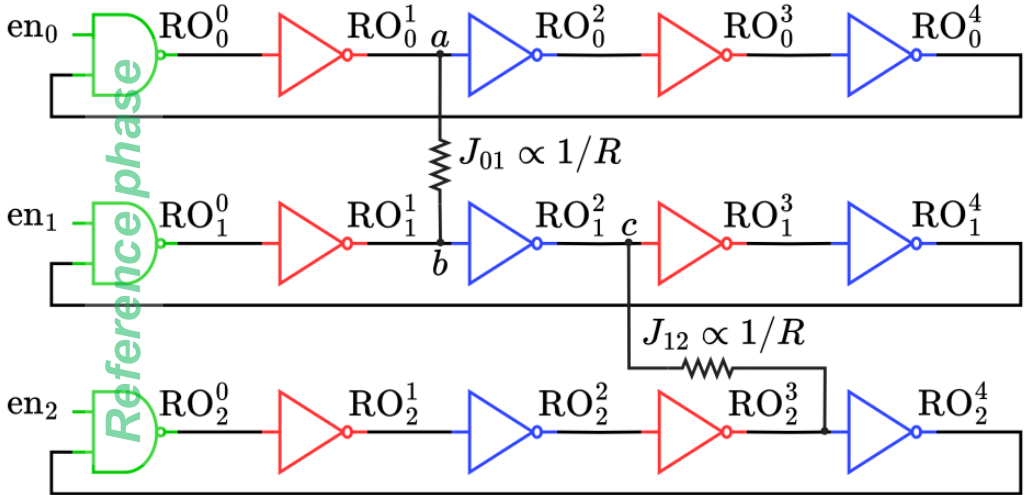


Analog computing with coupled oscillators

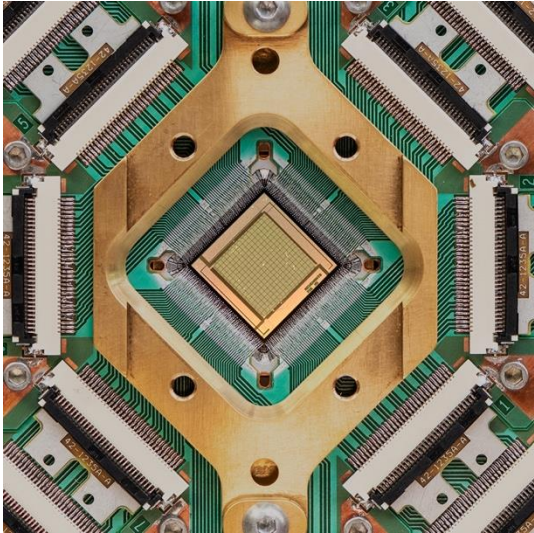
- Using ring oscillators



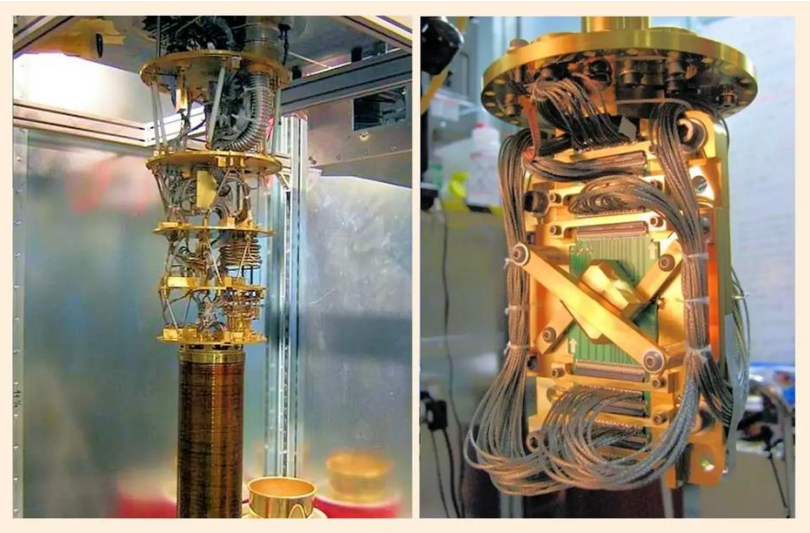
- Could have positive or negative coupling



Analog computing with coupled oscillators



[dwavesys.com]

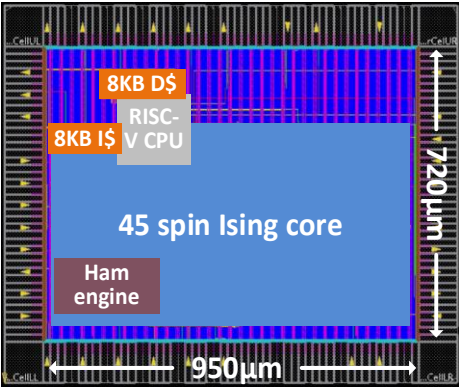


20mK

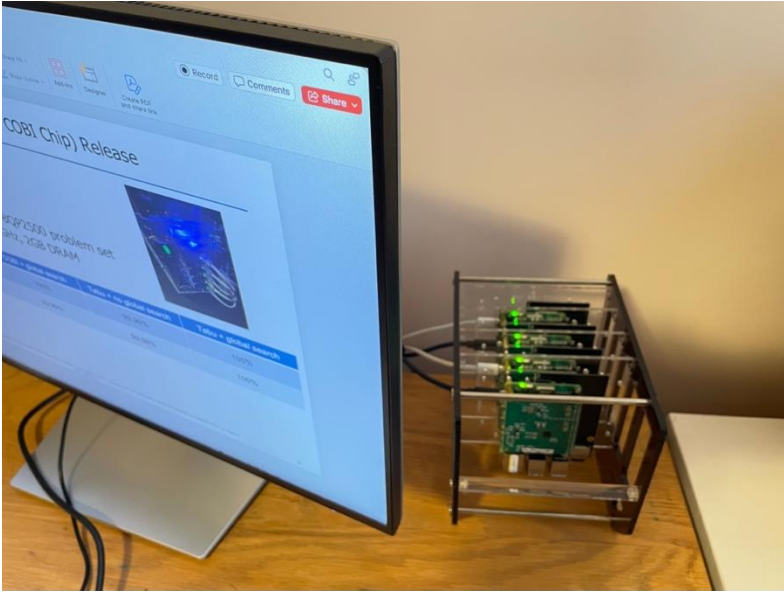
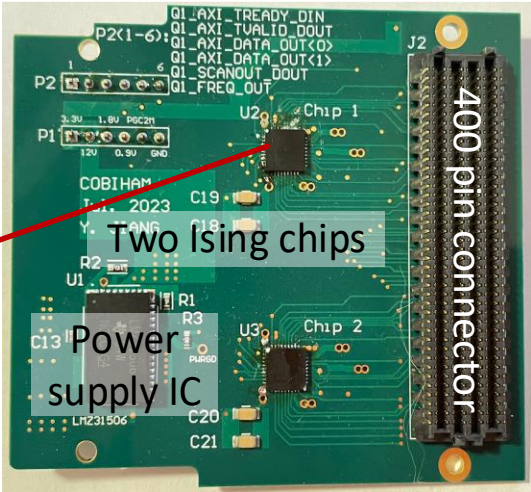
Classical quantum annealer

<https://www.thebrighterside.news/post/in-a-global-first-quantum-computers-crack-rsa-and-aes-data-encryption/>

COBI (Coupled oscillator based Ising)



FPGA Mezzanine Card



Room Temperature

~10mW



More concretely...

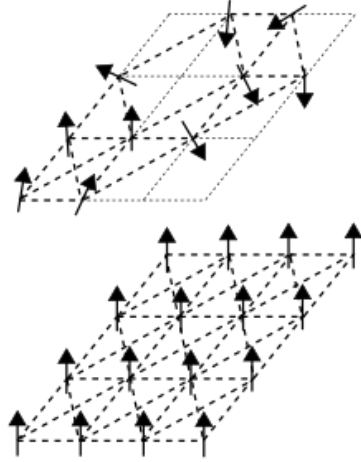


UNIVERSITY OF MINNESOTA

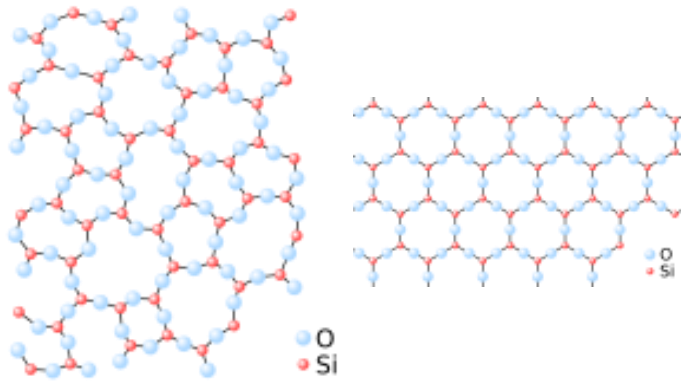
Driven to Discover®

“Ising spin glasses”

- Magnetic systems with potentially disordered magnetization



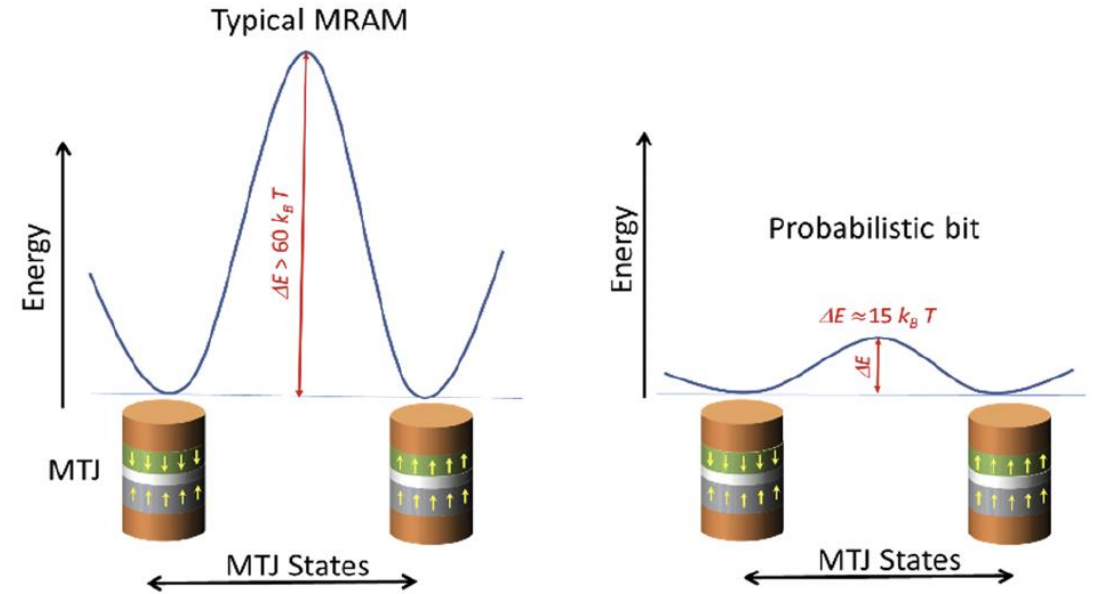
- Analogy with “regular” glass



Glass

Quartz

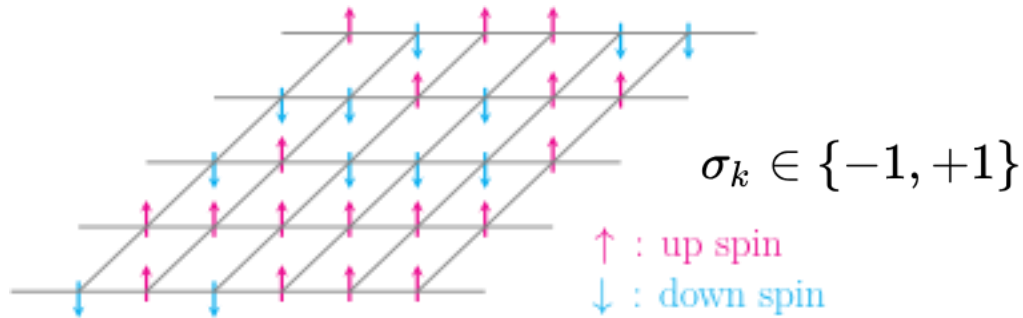
- Has been explored with magnetic tunnel junctions (MTJs)



[Navau and Sort, APL Materials 2021]

Ising computation for dummies (the Wikipedia version)

- Misnomer: formulated by Wilhelm Lenz
- 1D model solved by Ising (Lenz's student)
- 2D model: magnetic dipole moments in a lattice



- Configuration probability

$$P_\beta(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z_\beta},$$

where $\beta = 1/(k_B T)$, and the normalization constant

$$Z_\beta = \sum_{\sigma} e^{-\beta H(\sigma)}$$

- Energy given by a *Hamiltonian* function

$$H(\sigma) = - \sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j;$$

$J_{ij} > 0$, the interaction is called **ferromagnetic**,
 $J_{ij} < 0$, the interaction is called **antiferromagnetic**,
 $J_{ij} = 0$, the spins are *noninteracting*.

$h_j > 0$, the spin site j desires to line up in the positive direction,
 $h_j < 0$, the spin site j desires to line up in the negative direction,
 $h_j = 0$, there is no external influence on the spin site.

- One way to use this: [Lucas, Frontiers in Physics 2024]
 - Map a problem to this coupled system
 - Start from known solution H_0 to the Hamiltonian ; evolve from H_0 to H with temperature T

$$H(t) = \left(1 - \frac{t}{T}\right) H_0 + \frac{t}{T} H_P$$



Timing analysis of coupled oscillators

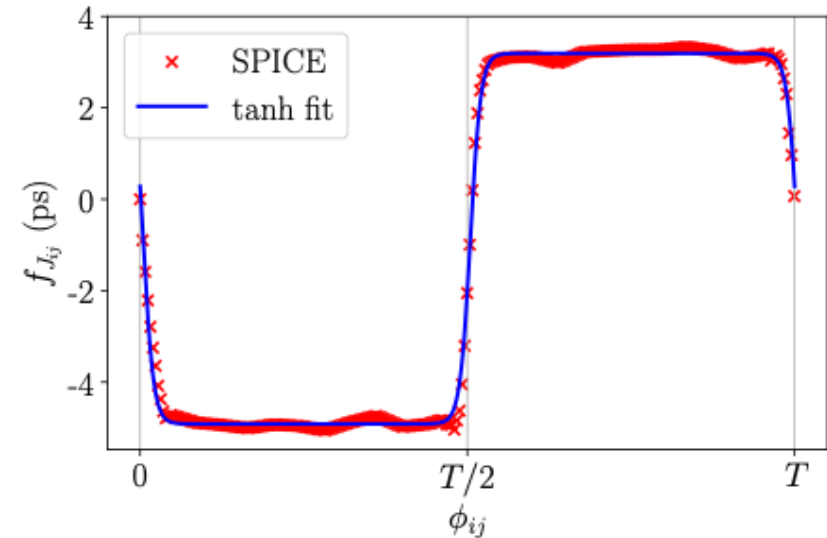
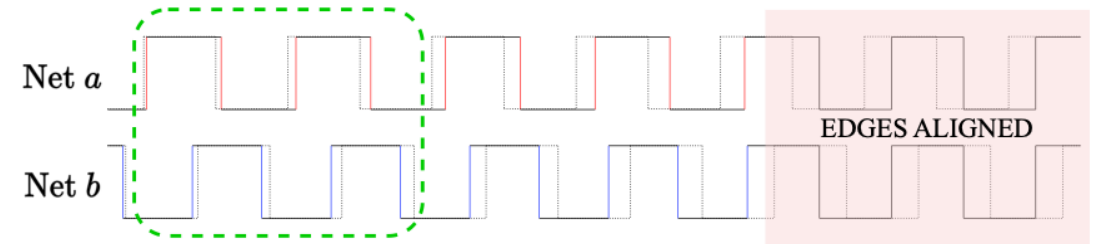
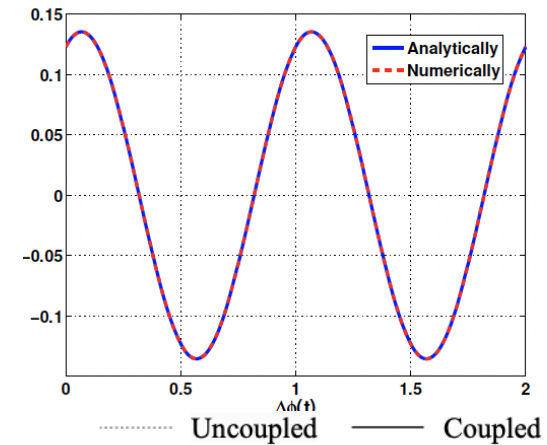
- Generalized Adler equation for N coupled oscillators [Bhansali and Roychowdhury, ASPDAC09]

$$\frac{d\phi_i(t)}{dt} = (\omega_i - \omega^*) + \omega_i \sum_{j=1, j \neq i}^N c_{ij}(\phi_{ij}(t))$$

- We have shown that this is a continuous-time approximation of a series of discrete events

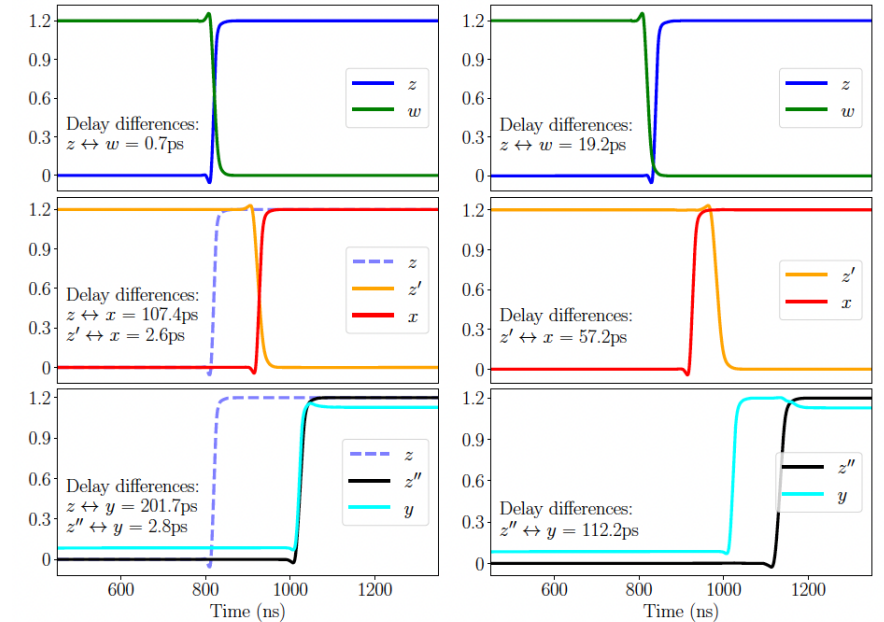
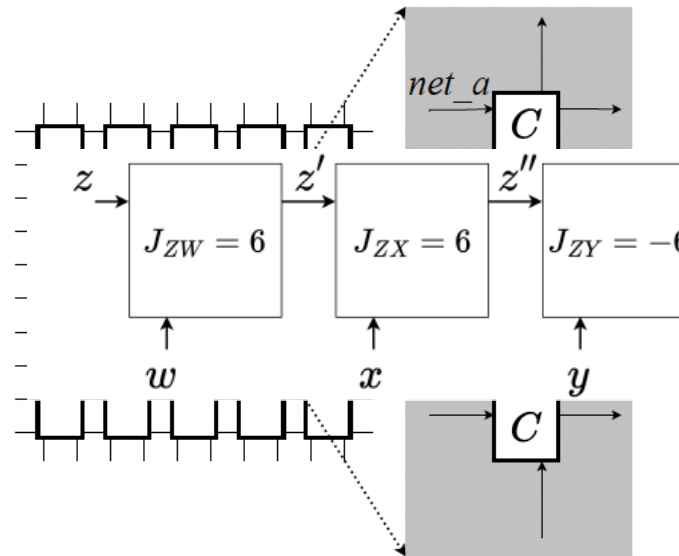
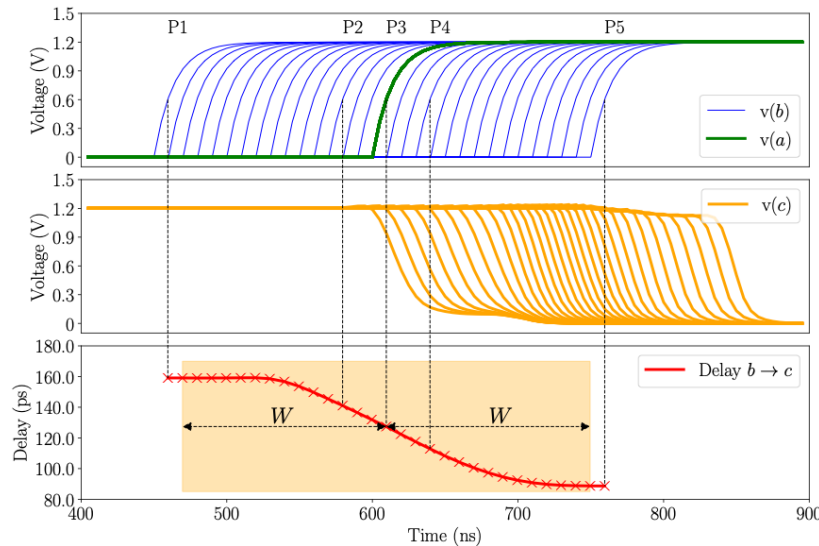
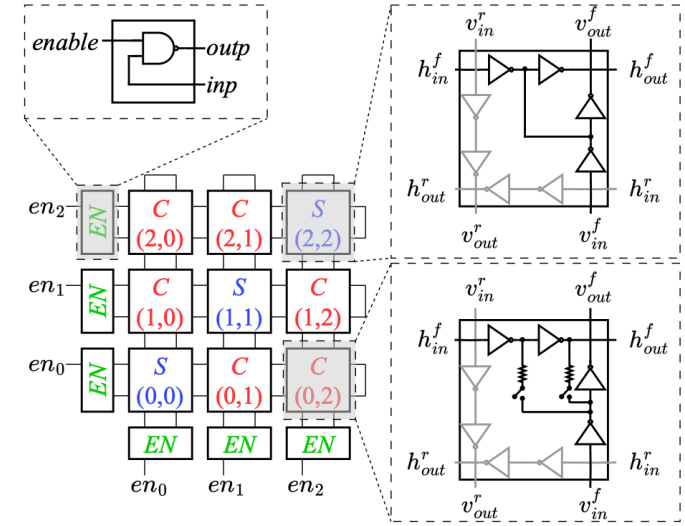
$$c_{ij} = f_{J_{ij}}(\phi_{ij}) / (2\pi)$$

where f is the delay shift per cycle



Timing analysis of coupled oscillators (contd.)

- For RO-based array: can use principles of timing analysis
 - Delay is a function of phase difference
- Event-driven timing
 - Seek events within the window that influences a stage delay

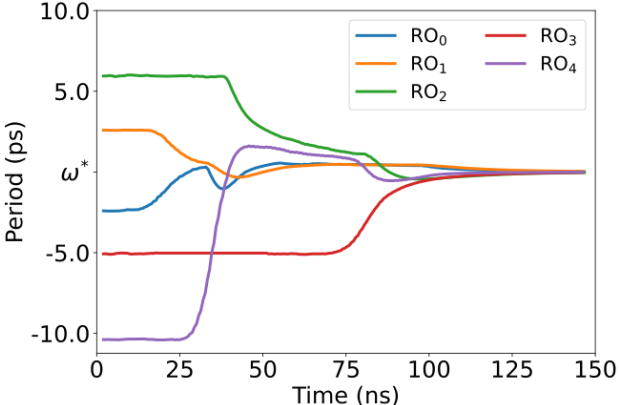


[Kumar et al., arXiv:2502.19399, 2025]

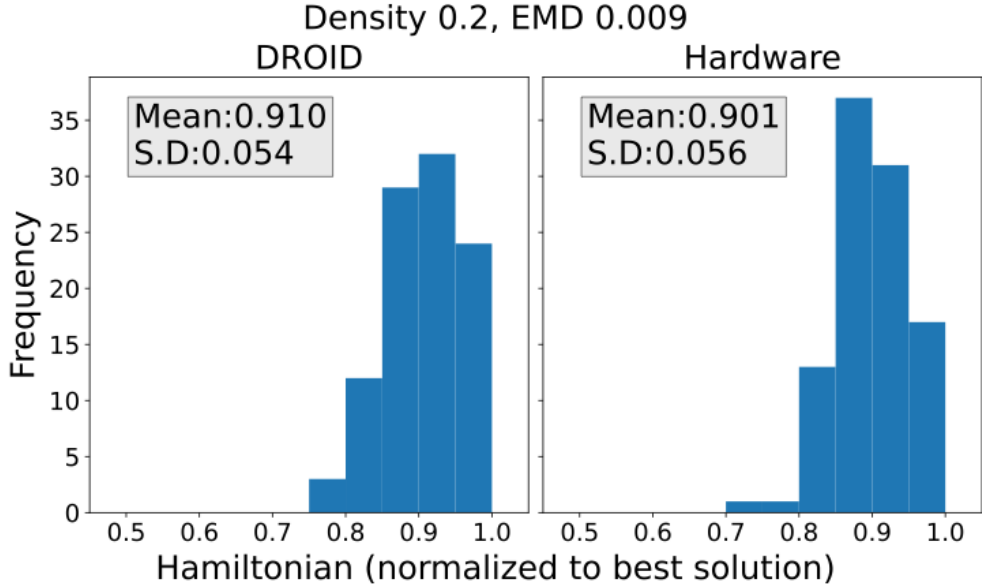


Simulation of all-to-all coupled arrays using DROID

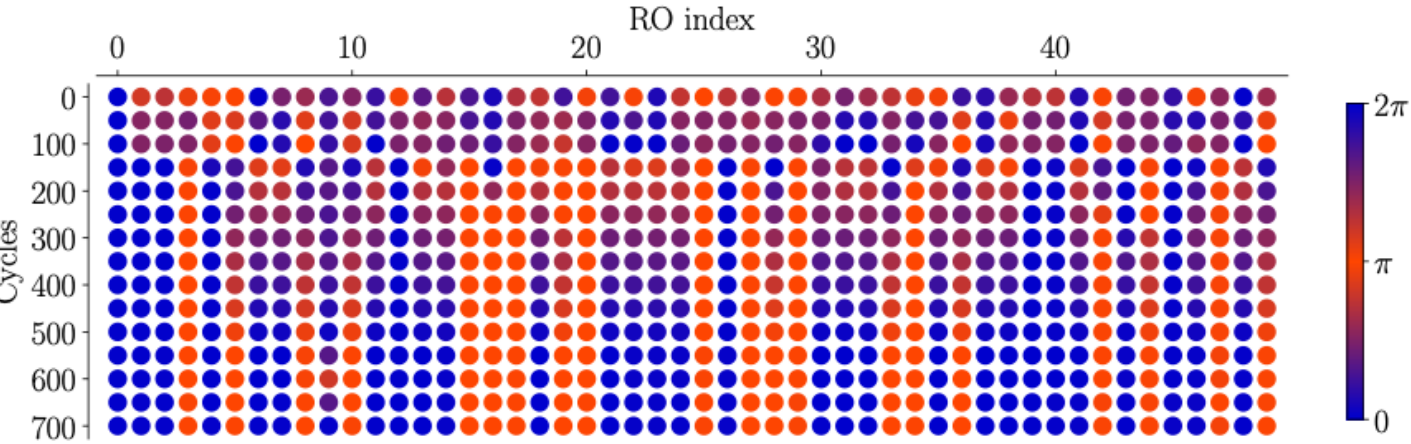
- Evolution of RO periods using DROID



- Comparison of the solution



- Evolution of the solution for 50 coupled ROs



[Kumar et al., arXiv:2502.19399, 2025]



Applications

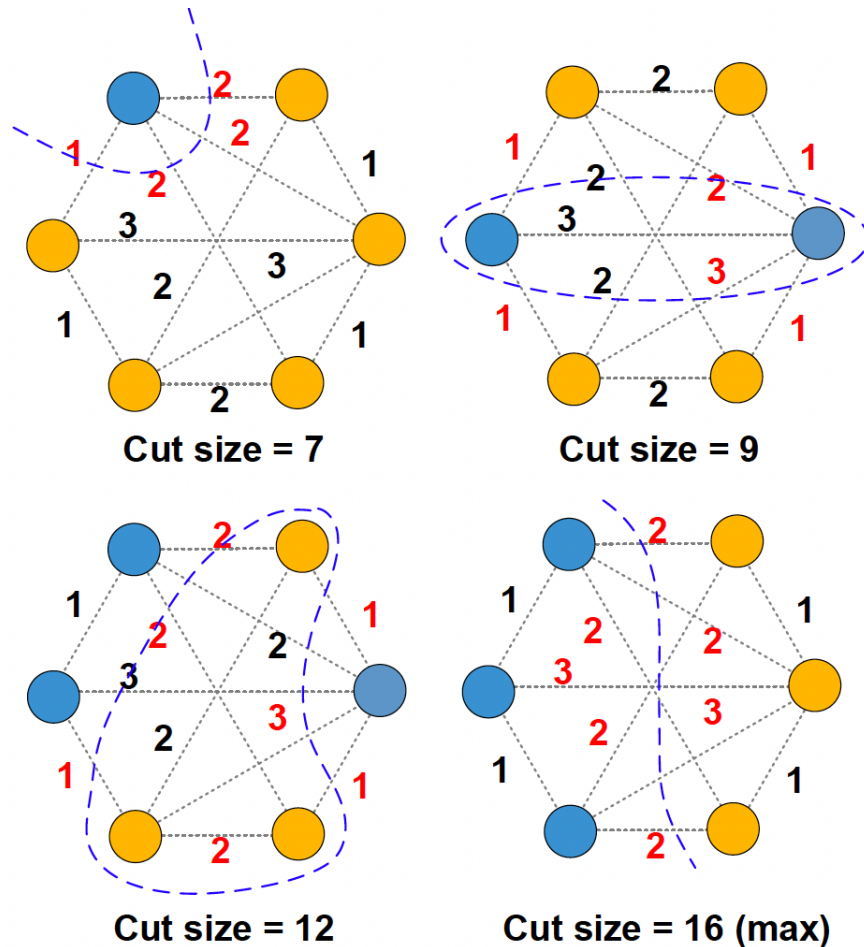


UNIVERSITY OF MINNESOTA

Driven to Discover®

Max-cut on an Ising solver

- NP-hard problem, considered to be the “easy” problem for Ising



$$\begin{aligned}
 H(\sigma) &= \sum_{i,j} w_{ij} \sigma_i \sigma_j \\
 &= \sum_{diff\ group} (-w_{ij}) + \sum_{same\ group} w_{ij} \\
 &= \sum_{diff\ group} (-w_{ij}) + \left[\sum_{i,j} w_{ij} - \sum_{diff\ group} w_{ij} \right] \\
 &= \sum_{all} w_{ij} - 2 \times \underbrace{\sum_{diff\ group} w_{ij}}_{\text{Cut size}}
 \end{aligned}$$

H = Hamiltonian of the system

σ_i = Spin status of magnet i {+1 or -1}

w_{ij} = weight between magnets i and j

Other problems can be modeled in Ising form

0-1 ILP

$$\min \mathbf{c}^T \mathbf{x}$$

subject to $\mathbf{S} \mathbf{x} = \mathbf{b}$, $x_i \in \{0,1\}$

- Minimize unconstrained Hamiltonian

$$H = H_A + H_B$$

$$H_A = A \sum_{j=1}^m \left[b_j - \sum_{i=1}^N S_{ji} x_i \right]^2 \quad H_B = -B \sum_{i=1}^N c_i x_i$$

- Factorizing 15

$$p = (x_1 \ 1)_2, q = (x_2 \ x_3 \ 1)_2$$

$$H = (15 - pq)^2$$

$$H = 128x_1x_2x_3 - 56x_1x_2 - 48x_1x_3 + 16x_2x_3 - 52x_1 - 52x_2 - 96x_3 + 196$$

$$H_{mod} = 200x_1x_2 - 48x_1x_3 - 512x_1x_4 + 16x_2x_3 - 512x_2x_4 + 128x_3x_4 - 52x_1 - 52x_2 - 96x_3 + 768x_4 + 196$$

[Lucas, Frontiers in Physics, 2014] [Jiang et al., Scientific Reports 2018]

Integer knapsack problem

Maximize knapsack cost

$$C = \sum_{\alpha=1}^N c_{\alpha} x_{\alpha}$$

subject to knapsack weight

$$\mathcal{W} = \sum_{\alpha=1}^N w_{\alpha} x_{\alpha} \leq W$$

- Introduce variable $y_n \equiv$ knapsack has weight n
- Maximize unconstrained Hamiltonian

$$H = H_A + H_B$$

$$H_A = A \left(1 - \sum_{n=1}^W y_n \right)^2 + A \left(\sum_{n=1}^W n y_n - \sum_{\alpha} w_{\alpha} x_{\alpha} \right)^2$$

$$H_B = -B \sum_{\alpha} c_{\alpha} x_{\alpha}$$

All-to-all interactions between variables!



QUBO formulations

- QUBO = quadratic unconstrained Boolean optimization
 - Variables $x_i \in \{0,1\}$

$$\min_{\mathbf{x}} F(\mathbf{x}) = \mathbf{x}^T Q \mathbf{x} = \sum_{i=1}^n Q_{ii} x_i + \sum_{i=1}^n \sum_{j=1, j \neq i}^n Q_{ij} x_i x_j$$

- Equivalent Ising formulation
 - Variables $s_i \in \{-1,1\}$

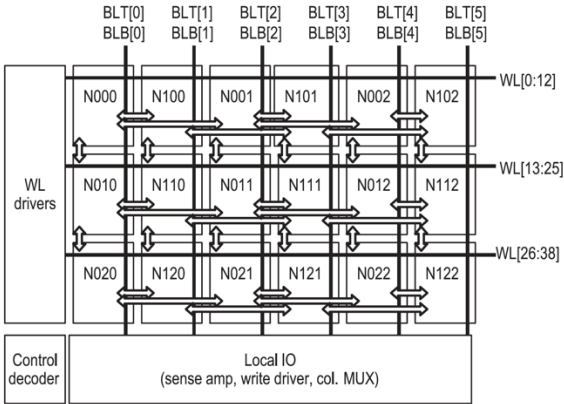
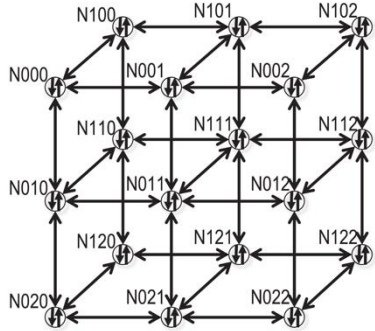


$$x_i = (s_i + 1)/2$$

$$\min_{\mathbf{s}} F(\mathbf{s}) = \underbrace{\sum_i h_i s_i}_{\text{field}} + \underbrace{\sum_{i=1}^n \sum_{j=1, j \neq i}^n J_{ij} s_i s_j}_{\text{coupling}}$$

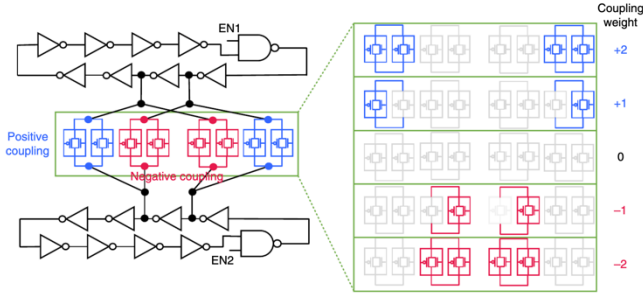
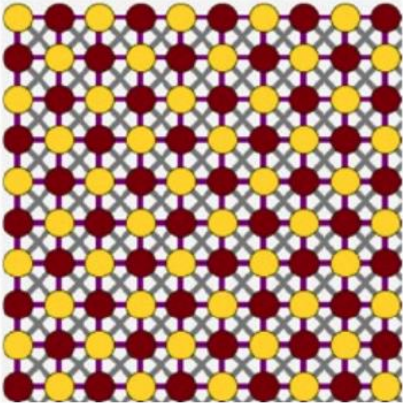
Implementations of Ising machines

- Lattice graph (SRAM-based)



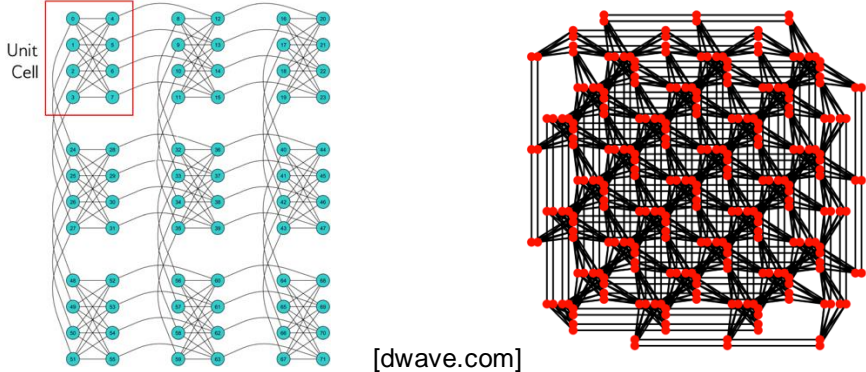
[Yamaoka, JSSC 2016]

- King's graph CMOS ASIC



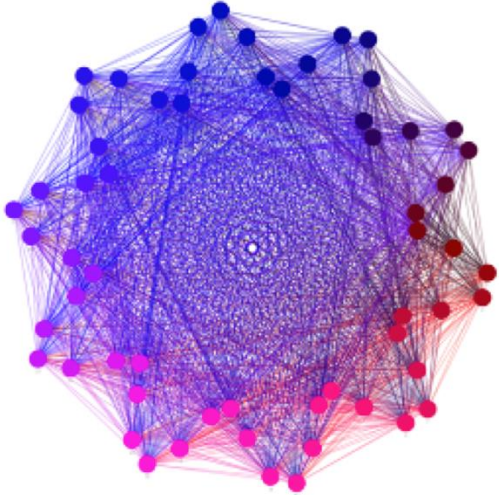
[Moy, Nature Electronics 2022]

- D-Wave



[dwave.com]

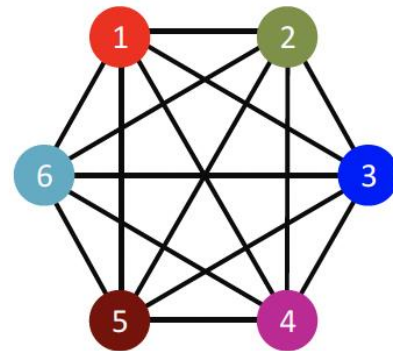
- Ideal structure: "all-to-all connected"



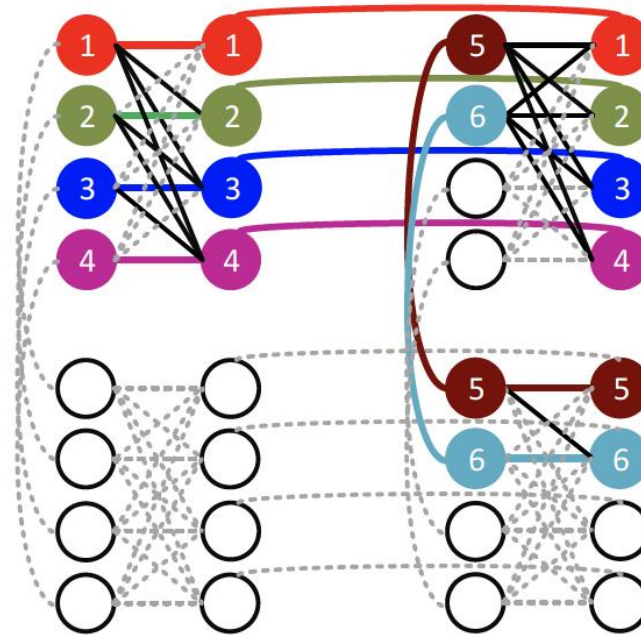
Planar vs. A2A and minor embedding

- Number of connections in lattice-based/King's graph structures = $O(n)$ for n coupled oscillators
- Can program $O(n)$ out of $O(n^2)$ possible connections
- Minor embedding: mapping logical spins to physical spins
- More replication = weaker spin strength

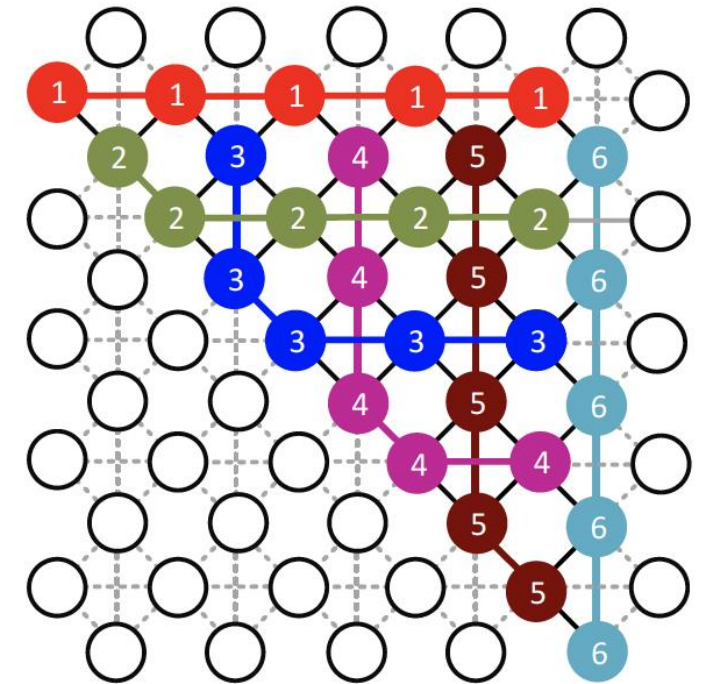
Topology	Connectivity
King's graph (Hitach)	8
Chimera graph (D-Wave)	6
Pegasus graph (D-Wave)	15
Zephyr (D-Wave)	20
All-to-all (This work)	48-1 = 47



All-to-all graph

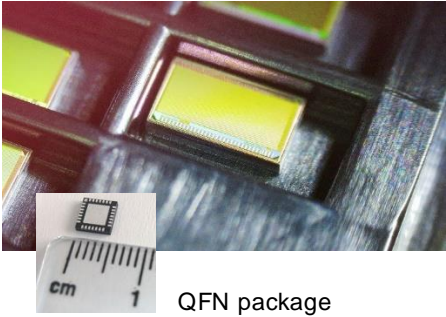
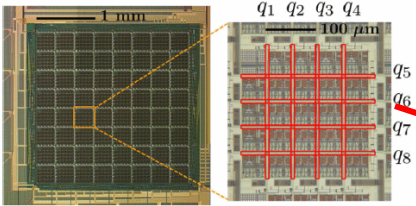


Chimera graph



King's graph

Quantum vs. Quantum-Inspired

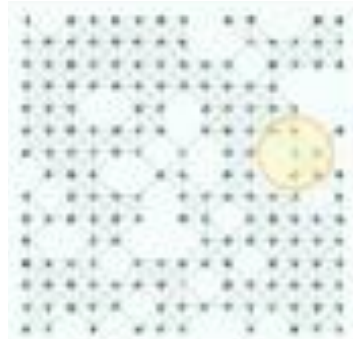
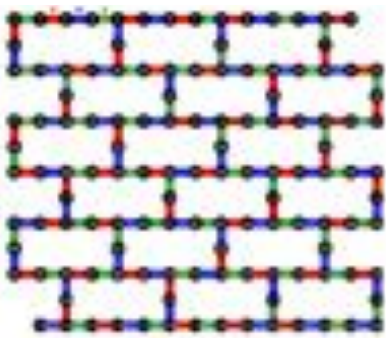
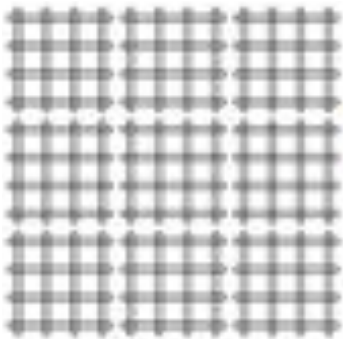



	Quantum Annealer	CMOS Ising Chip
Qubits	5,000+ physical qubits = 47 (all-to-all)	59 physical qubits
Connectivity	15	58 (native all-to-all)
Couplers	35,000+	1,711
Weight resolution	~5 bit	~5 bit
Power consumption	25,000W	0.010W
Technology	Superconducting	Foundry (TSMC)

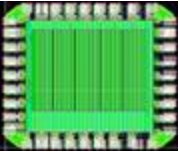
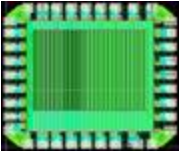
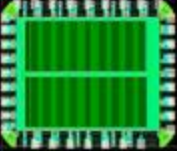
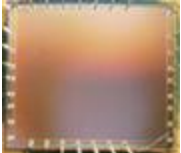
[D-Wave, IOT World Today]

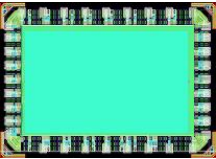
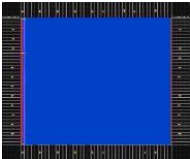
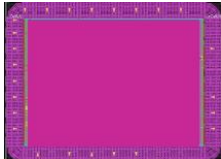
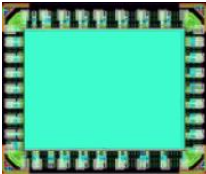


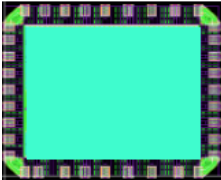
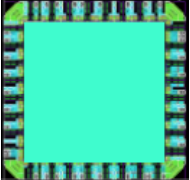
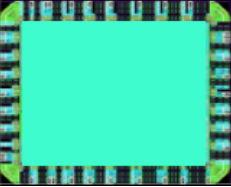
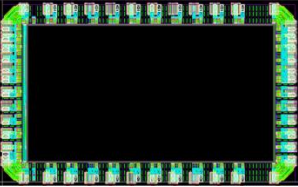
Key challenge: Connectivity

Architecture				
	King's graph	Eagle	Chimera	All-to-all
# of qubits	289 (equivalent to 17 all-to-all)	127	2,000+ (equivalent to ~50 all-to-all)	59
Connectivity	8 (max)	3 (max)	6	58
Weight	0, -1	N/A	N/A	5 bit
Device type	Rydberg atoms	Superconducting	Superconducting	CMOS (room temp.)
Minor embedding	Required	Required	Required	Not required
Power consumption	N/A	N/A	25,000W	0.010W
Reference	Science 2022 [Ebadi]	Nature 2023 [Kim]	D-Wave website	This work

CMOS Ising chip prototypes

Chip name	COBISHIL (28nm, Jan. 2023)	COBITX (28nm, Jan. 2023)	COBIFIXED (28nm, Mar 2023)	COBIFIXED65 (65nm, May 2023)
Chip Layout				

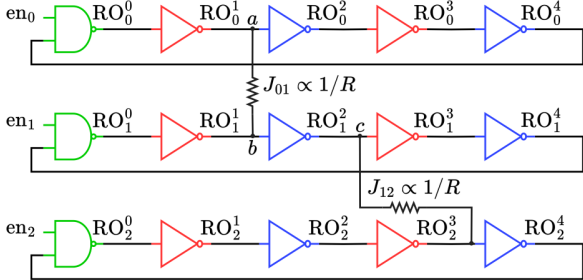
Chip name	COBIHAM (28nm, May 2023)	COBIRISCV2 (28nm, Jun 2023)	COBIGRADIENT (28nm, Sep 2023)	COBITRANSPOSE (28nm, Dec 2023)
Chip Layout				

Chip name	COBIFIXED28 (28nm, Jan. 2024)	COBIDUALRES (28nm, Feb. 2024)	COBIFIXED28P (28nm, Mar. 2024)	COBIFIVE (28nm, May 2024)
Chip Layout				

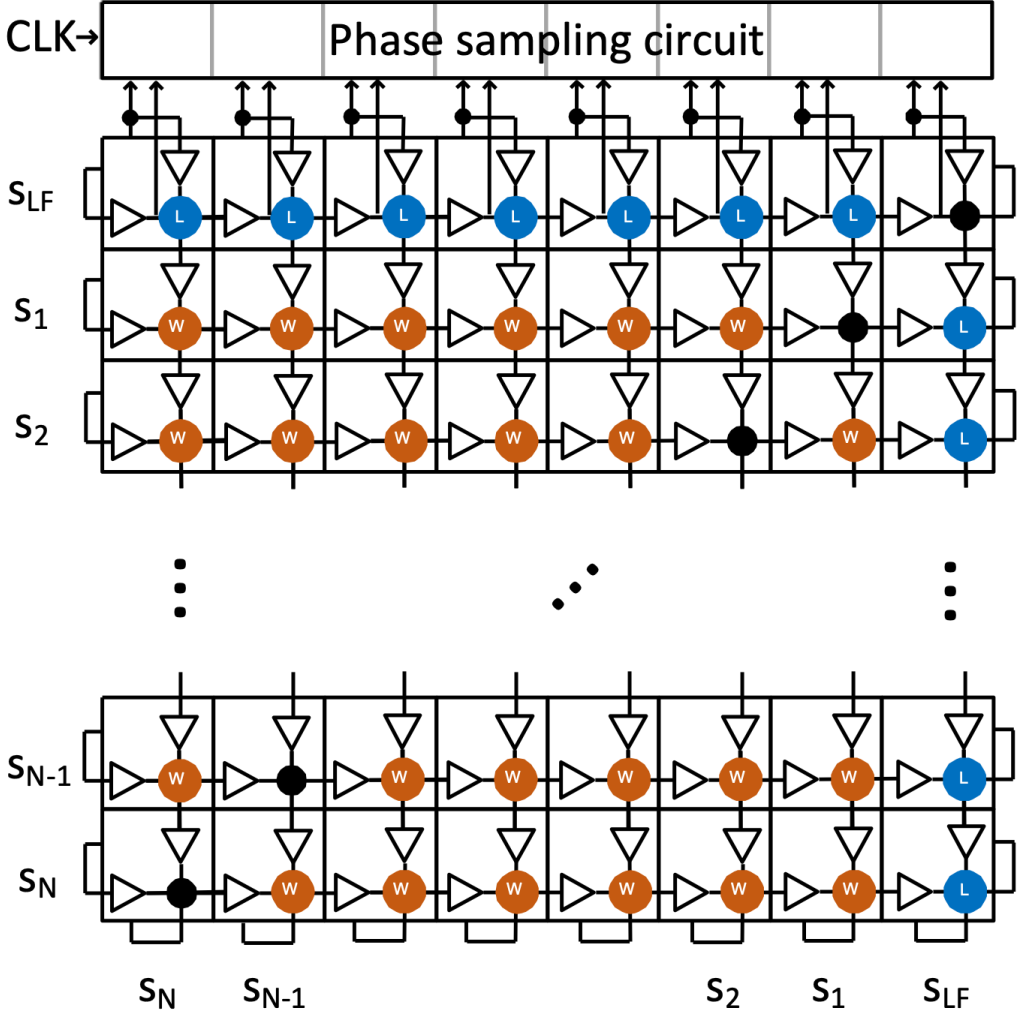
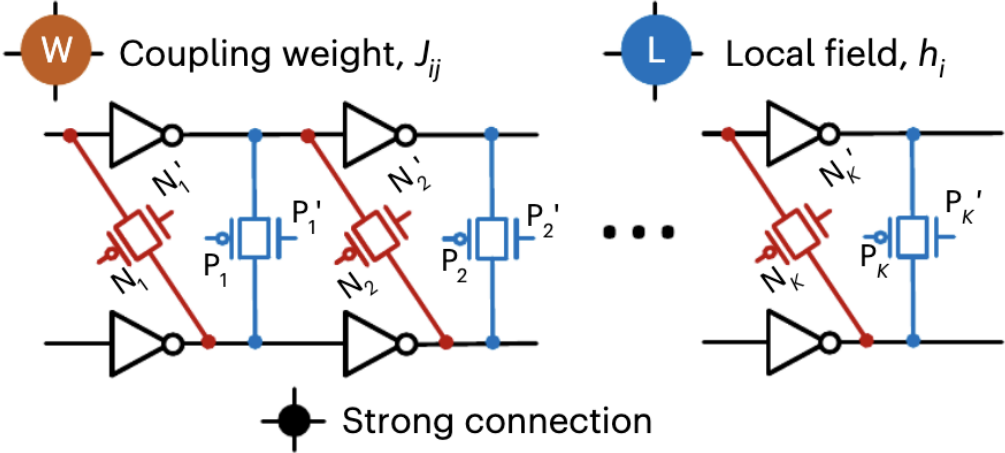


COBI oscillator array

- Coupling strengths depend on wire parasitics

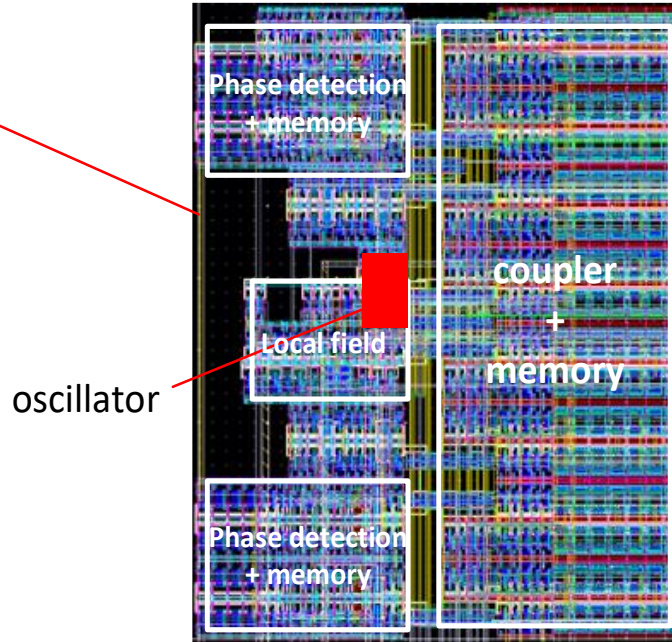
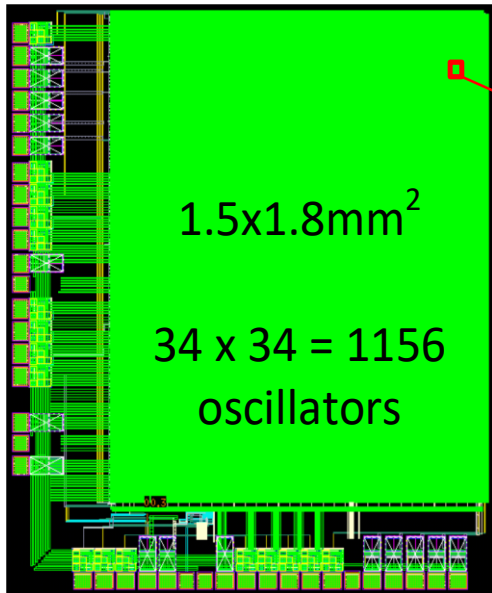


- Need complete symmetry in the structure



COBI oscillator array (contd.)

Chip layout

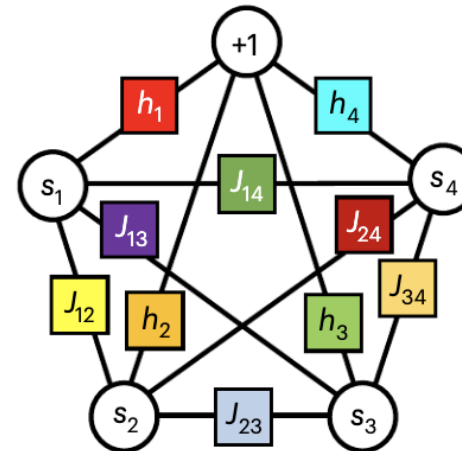


oscillator

Oscillator area < 3% of COBI chip.
Rest is memory, coupler, logic, control, and IO.

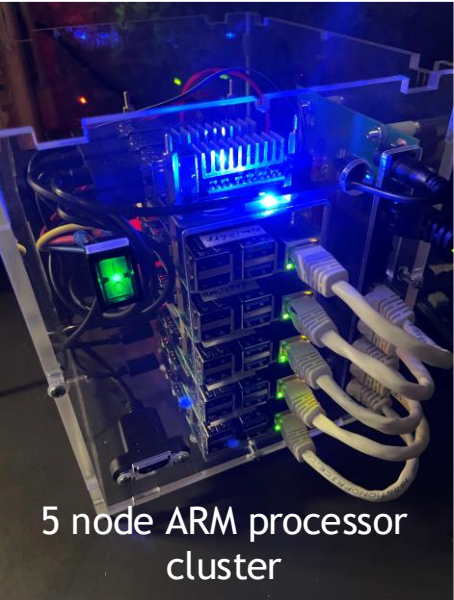
Problem mapping

$$H = h_1 s_1 + h_2 s_2 + h_3 s_3 + h_4 s_4 + J_{12} s_1 s_2 + J_{13} s_1 s_3 + J_{14} s_1 s_4 + J_{23} s_2 s_3 + J_{24} s_2 s_4$$

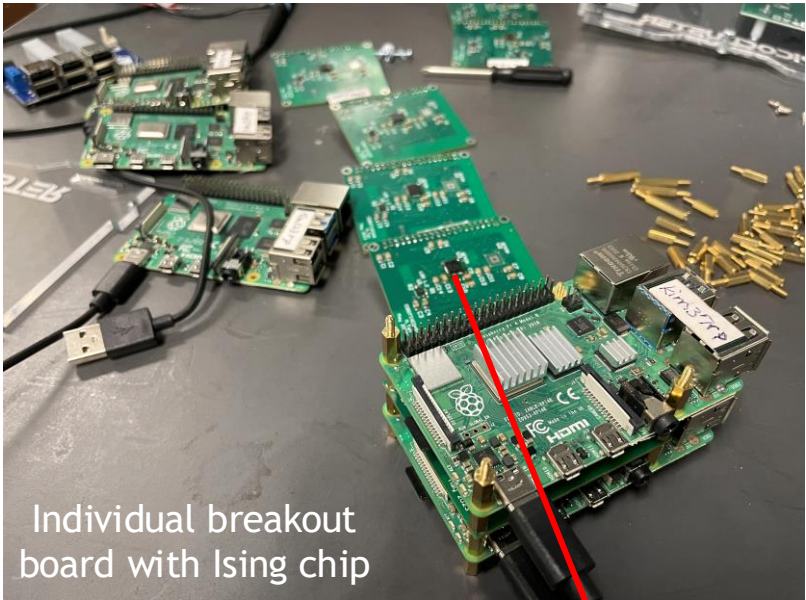


	J_{34}	J_{24}	J_{14}	h_4
J_{34}		J_{23}	J_{13}	h_3
J_{24}	J_{23}		J_{12}	h_2
J_{14}	J_{13}	J_{12}		h_1
h_4	h_3	h_2	h_1	

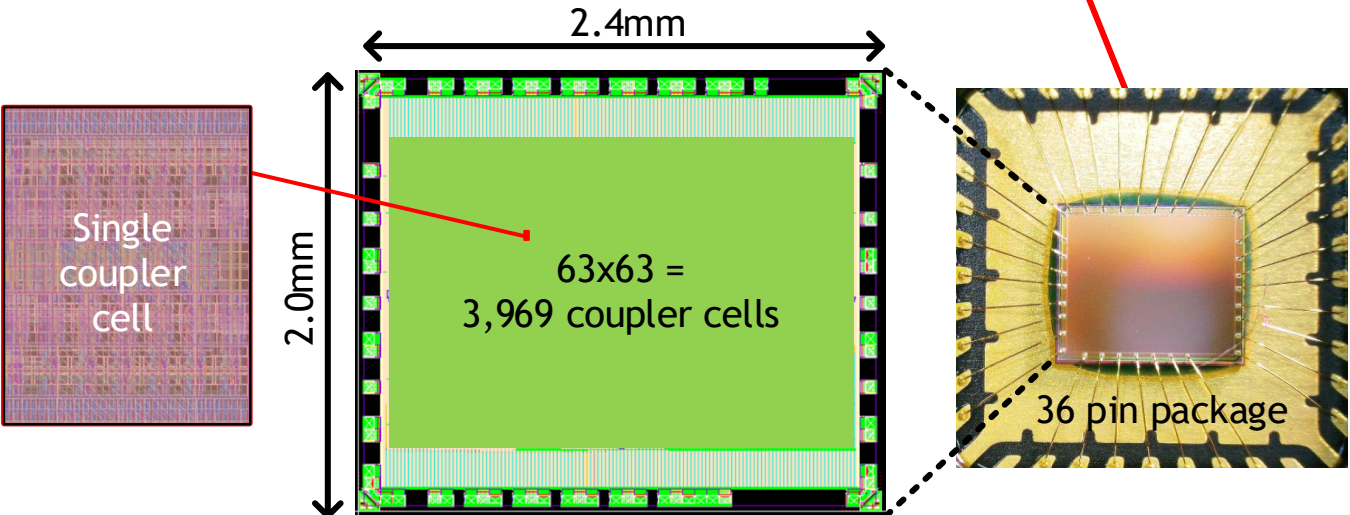
Improved 59-spin A2A Ising solver chip



5 node ARM processor cluster



Individual breakout board with Ising chip

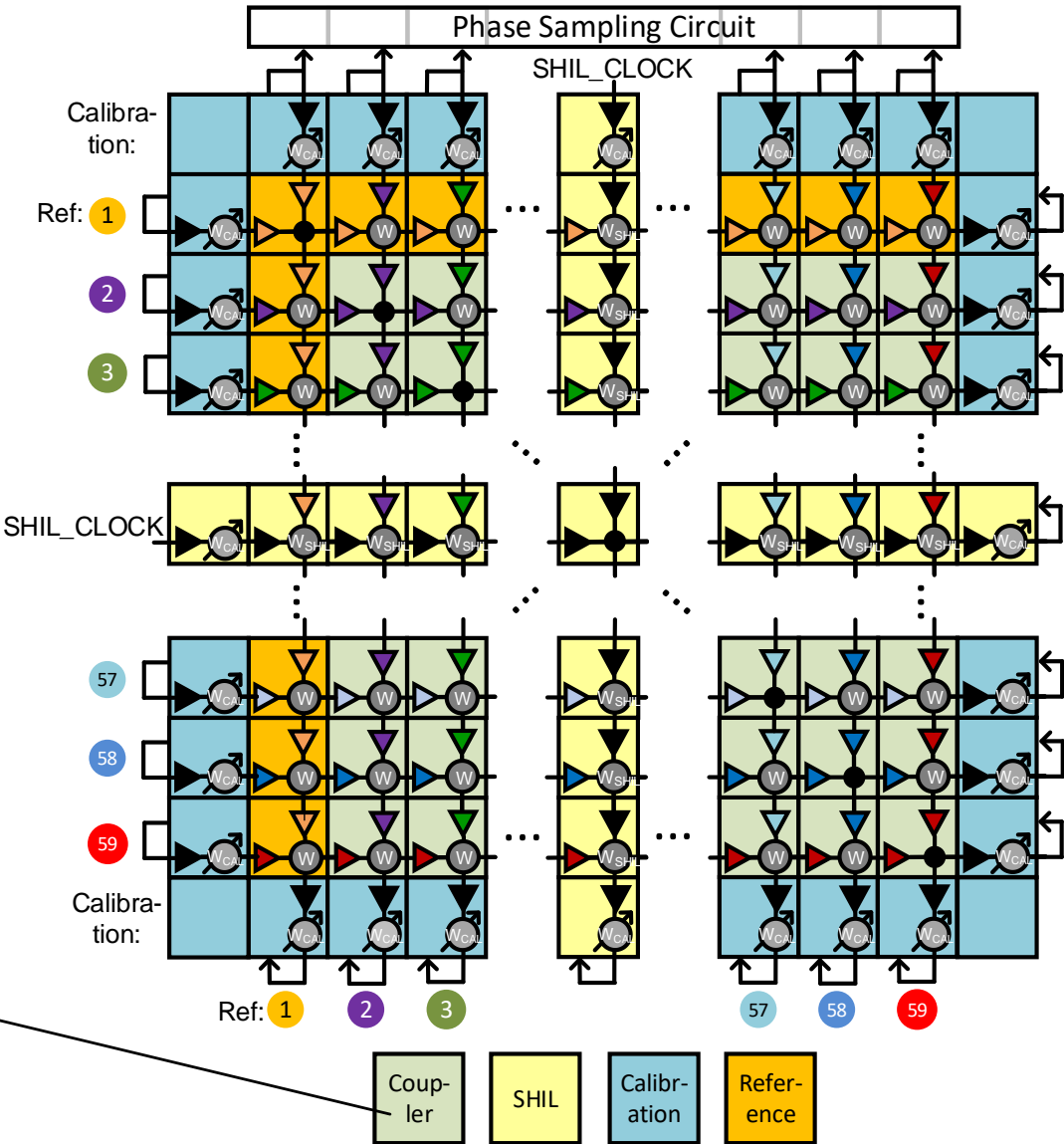
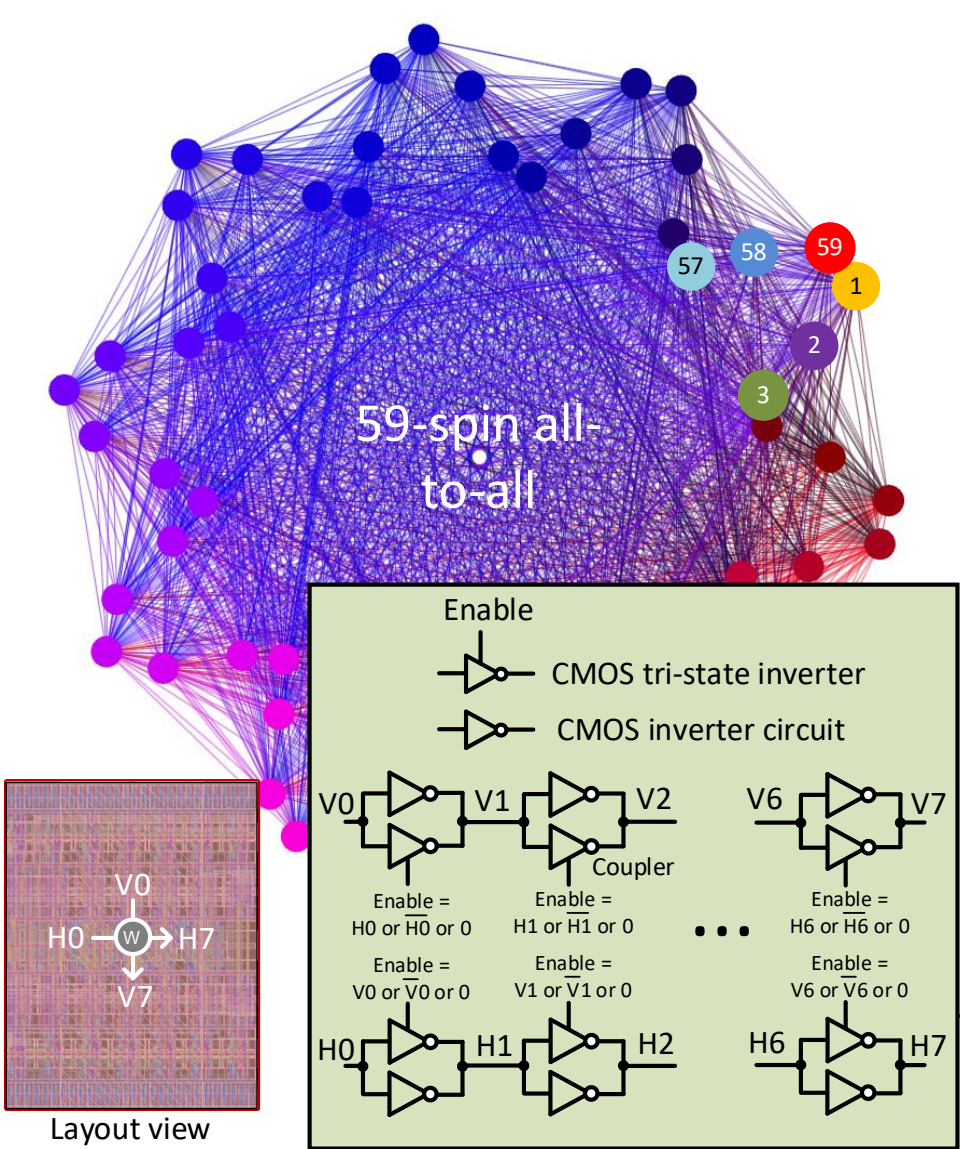


Technology	TSMC 65 nanometer
Supply voltage	1.2V
Die size	2.4 mm x 2.0 mm
Array size	63 x 63
# of spins	59 spins + 2 SHIL + 2 calibration = 63 cells
Connectivity	58 (all-to-all)
Coupler resolution	-14, -13, ..., 13, 14 (29 levels)
# of couplers	$59 \times 58 / 2 = 1,711$
Power consumption	8.2mW (graph density = 0.0) 10.0mW (graph density = 1.0)
Oscillation frequency	5.19 MHz (graph density = 0.0) 5.36 MHz (graph density = 1.0)
Operating temperature	Room temperature
Key features	Zero static power coupler, sub harmonic injection locking (SHIL), individual frequency calibration

[Cilasun et al., <https://doi.org/10.21203/rs.3.rs-4208492/v1>]



Array architecture and coupler design



[Cilasun et al., <https://doi.org/10.21203/rs.3.rs-4208492/v1>]



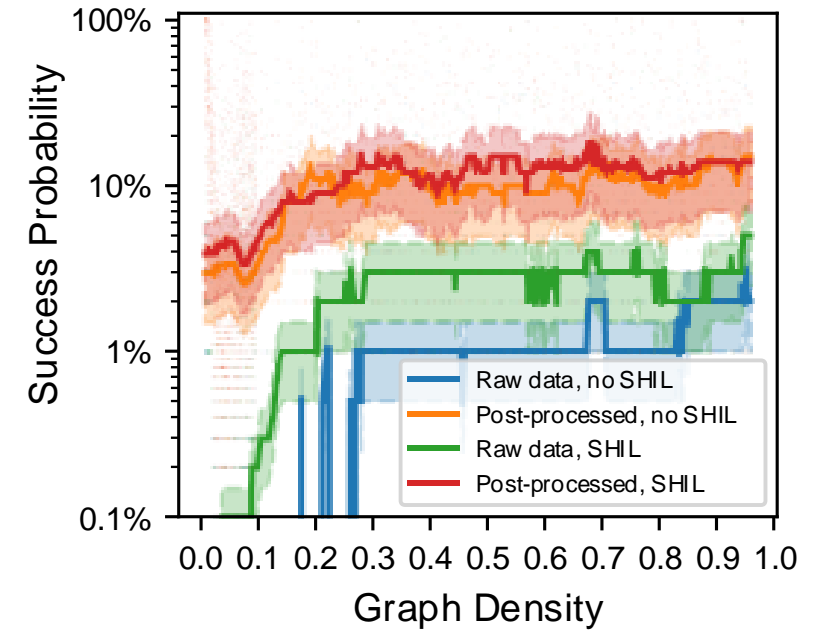
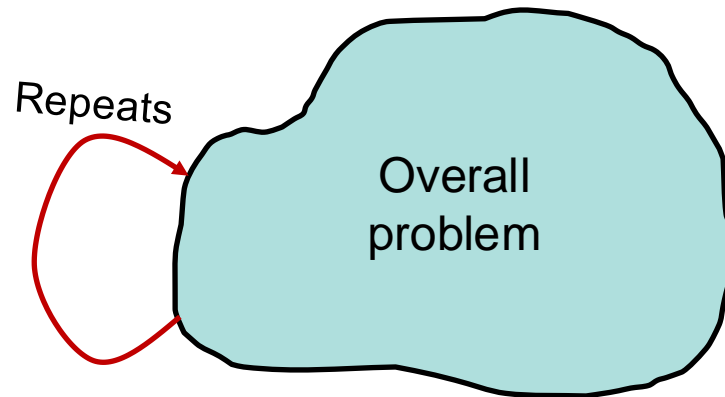
Solving real problems in the real world

The solution is inherently probabilistic!

- Requires repeated solutions – not a problem if each solution is fast
- Let p = probability of achieving a solution
- Over n attempts, probability of success = $1 - (1 - p)^n$
- Time-to-solution (TTS) metrics
 - To achieve p_{target} (e.g., 99%) accuracy,

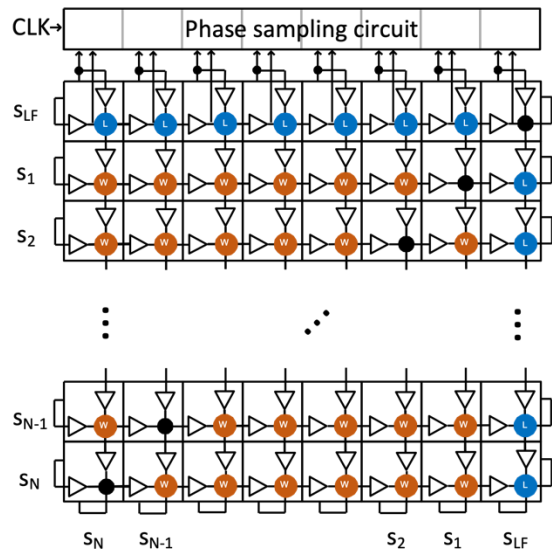
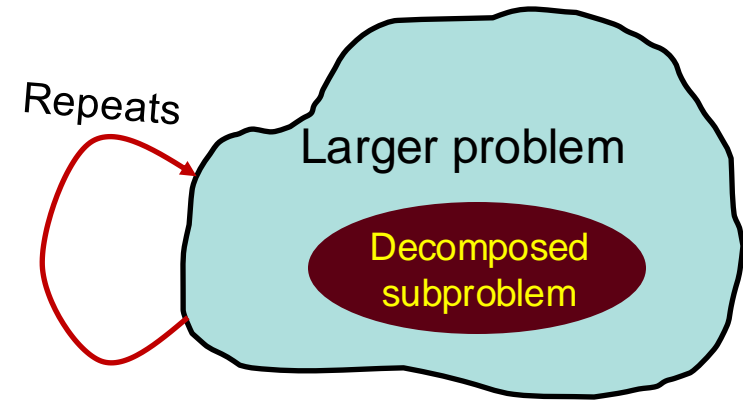
$$p_{target} = 1 - (1 - p)^n$$

$$n = \log(1 - p_{target}) / \log(1 - p)$$

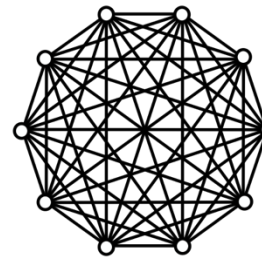


Living with the limitations of the hardware

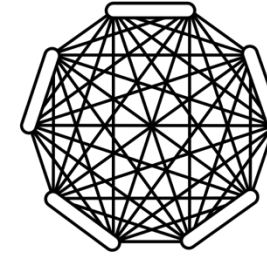
Limitation	Solution
Limited hardware size	<ul style="list-style-type: none"> Problem decomposition
Limited coupling weights	<ul style="list-style-type: none"> Map to multiple spins Truncation
Precision improvement	<ul style="list-style-type: none"> Scaling Truncation



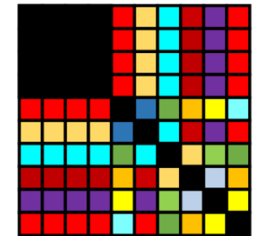
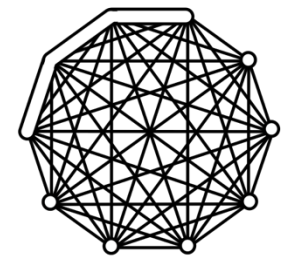
a Default mode (1X weight resolution)



b 2 spin merge (4X weight resolution)

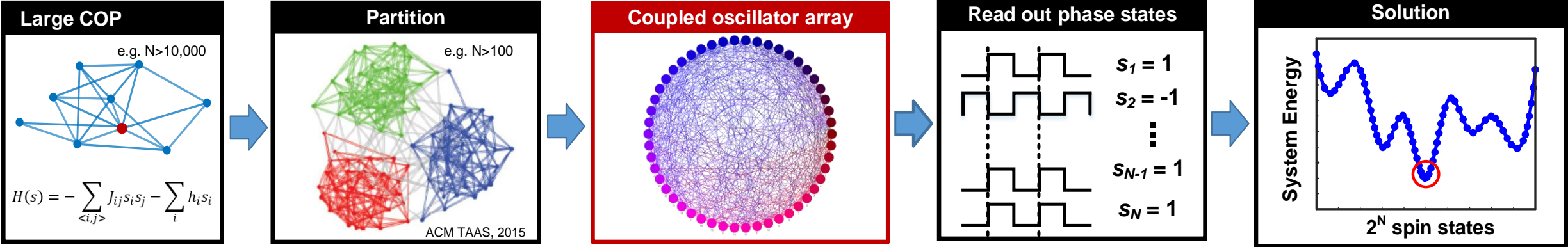


c Asymmetric merge (1X, 4X mixed weight resolution)



[Cilasun *et al.*, Scientific Reports 2024]

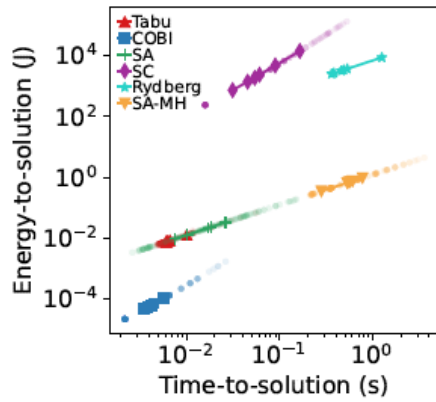
Solving large problems with a decomposer



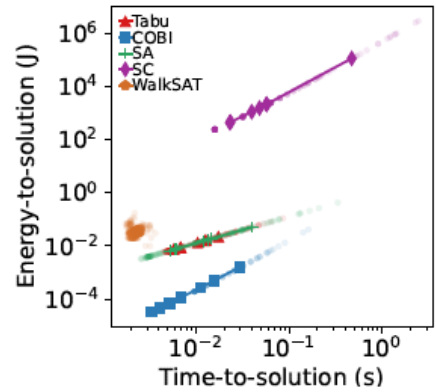
$$H(s) = - \sum_{\langle i,j \rangle} J_{ij} s_i s_j - \sum_i h_i s_i \quad : \text{Ising Hamiltonian}$$

s_i, s_j : Spin state {+1 or -1} J_{ij} : Coupling strength h_i : local field strength

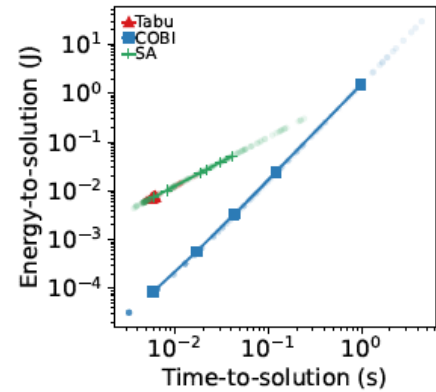




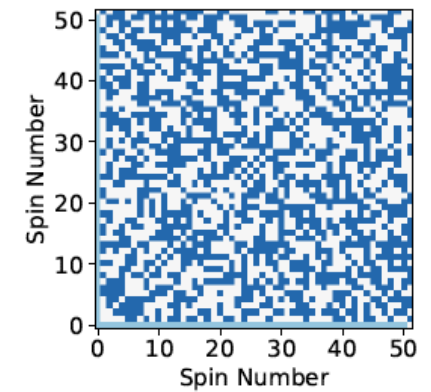
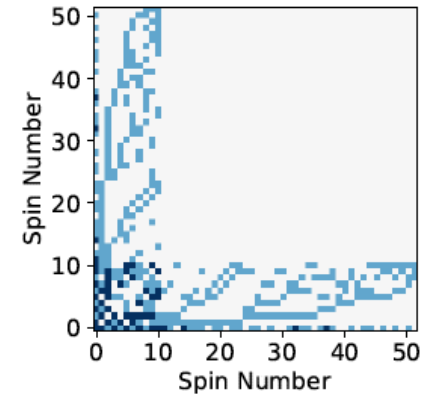
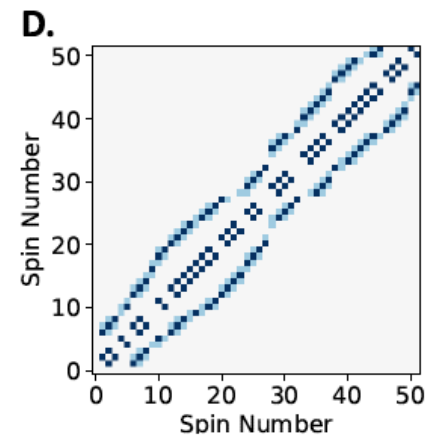
Planar MIS
($n=39, 51$)



SAT
($n=10, 11$)

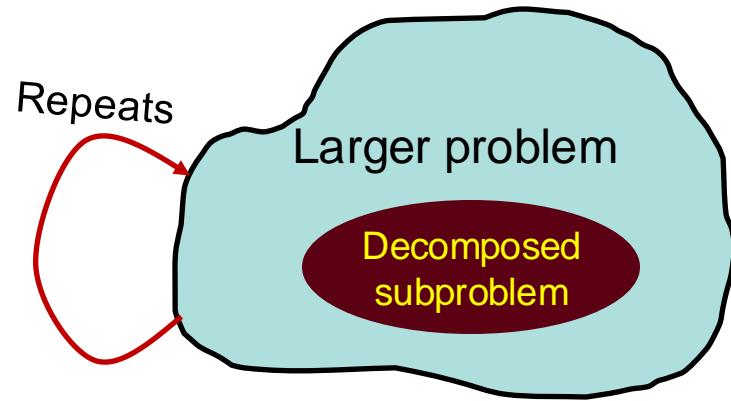


Nonplanar MIS
($n=51$)



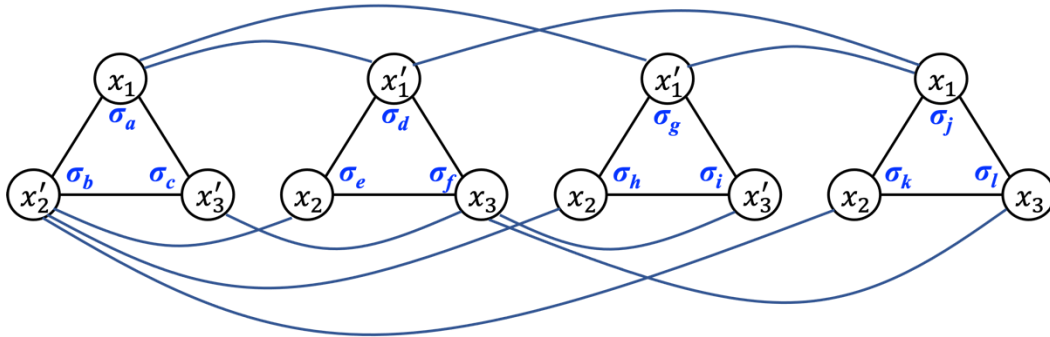
Exploiting parallelism

- Multicore COBI chips can accelerate solution



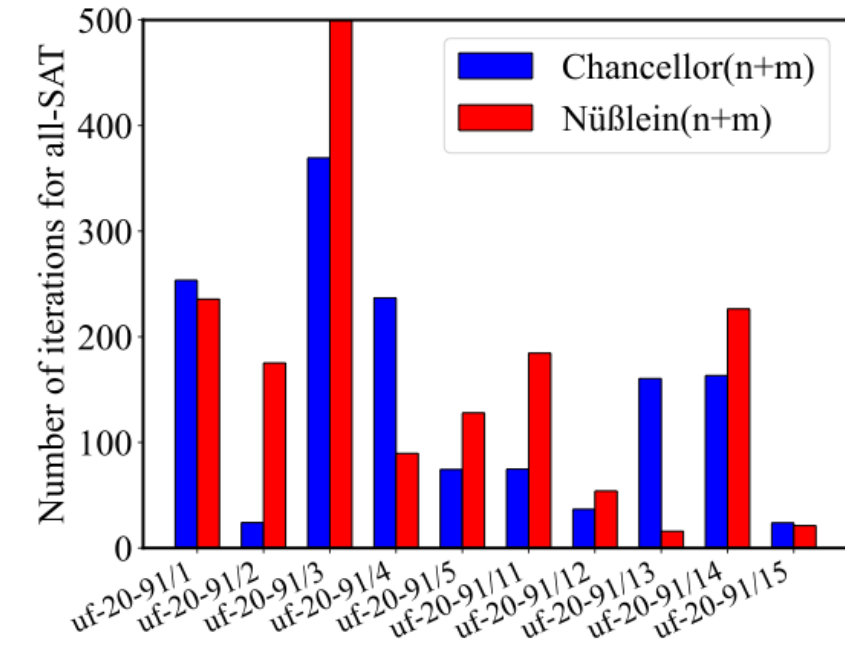
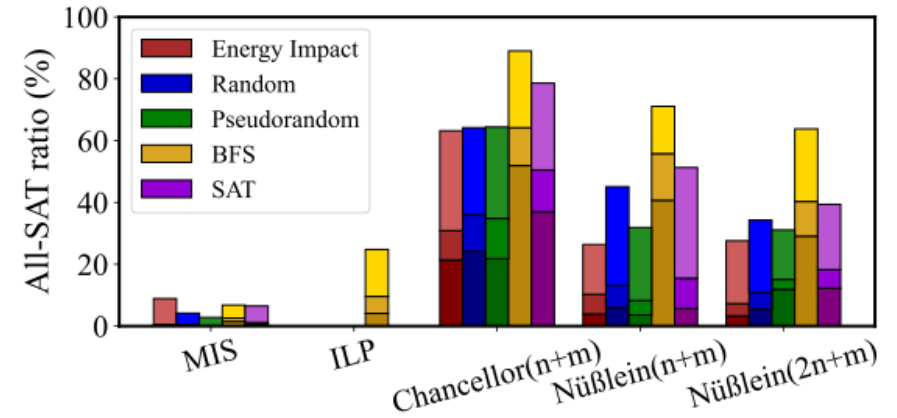
Problem mapping: 3SAT on the COBI Ising chip

- 3SAT parameters: n variables, m clauses
- Various possible mappings
 - Maximal independent set ($3m$ spins)



$$f = (x_1 \vee x'_2 \vee x'_3)(x'_1 \vee x_2 \vee x_3)(x'_1 \vee x_2 \vee x'_3)(x_1 \vee x_2 \vee x_4)(x_4 \vee x'_5 \vee x_6)(x'_4 \vee x_5 \vee x_6)$$

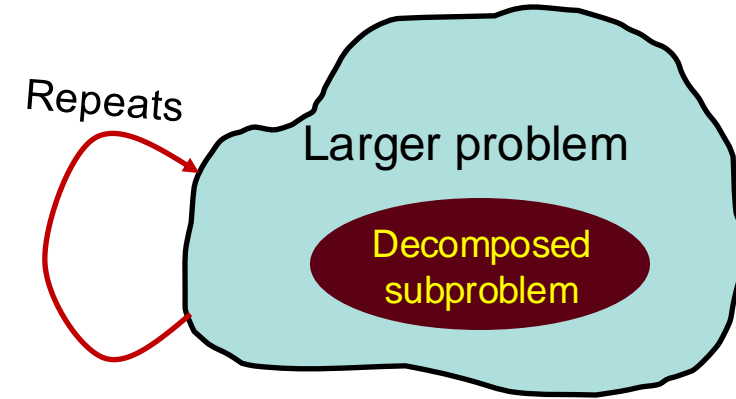
- Chancellor, Nüßlein formulations ($n + m$ spins)
- ILP formulation ($n + 2m$ spins)



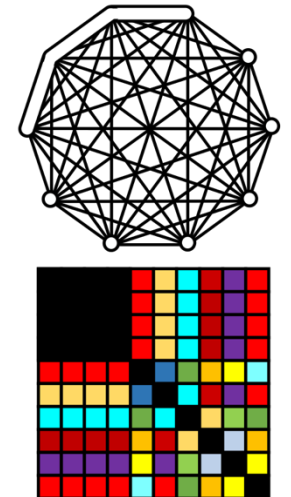
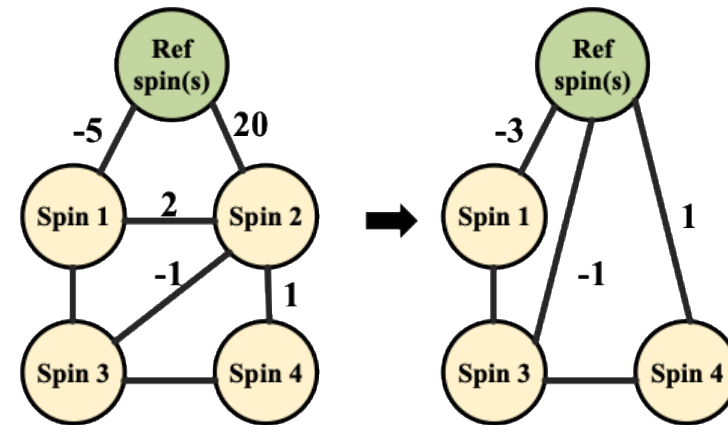
Problem decomposition and frozen spins

- All spins outside the decomposed segment are “frozen” to values in the previous iteration
- Impact of freezing a spin in this expression

$$\min_{\mathbf{s}} F(\mathbf{s}) = \underbrace{\sum_i h_i s_i}_{\text{field}} + \underbrace{\sum_{i=1}^n \sum_{j=1, j \neq i}^n J_{ij} s_i s_j}_{\text{coupling}}$$

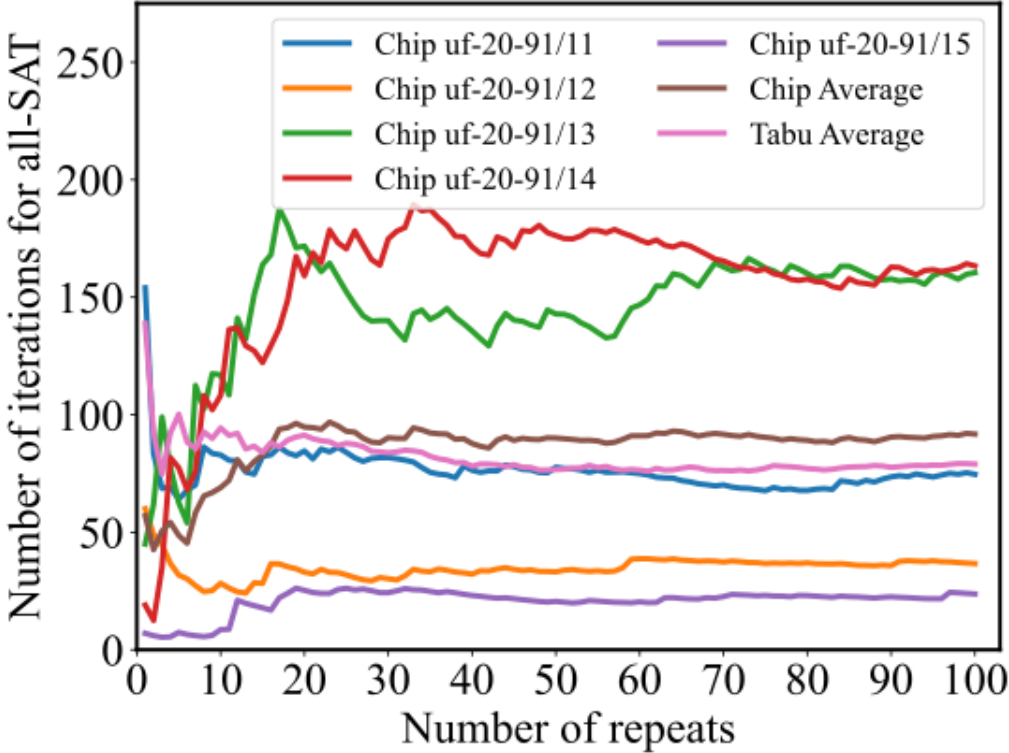


- Frozen $s_i \Rightarrow$ field term constant \Rightarrow removed
- Frozen s_i and $s_j \Rightarrow$ coupling constant \Rightarrow removed
- Unfrozen s_i and $s_j \Rightarrow$ coupling term kept as is
- Frozen s_i but not $s_j \Rightarrow$ coupling \rightarrow field term
 - Field terms can become very large!
 - Theoretical result: if field term $h_i s_i$ satisfies an inequality, it forces the spin s_i value to ± 1



Some early runtime comparisons

Benchmark	Runtime (ms)			Energy (mJ)		
	Tabu	WalkSAT	Chip	Tabu	WalkSAT	Chip
uf20-91-11	4322	2.351	12.83	367,370	199.8	0.128
uf20-91-12	1596	2.337	6.30	135,660	198.6	0.063
uf20-91-13	5095	2.506	27.58	50,950	213.0	0.276
uf20-91-14	7148	2.470	28.07	433,075	210.0	0.281
uf20-91-15	1591	2.377	4.09	607,580	202.0	0.041



Very recent results on ~200-variable SAT problems show great promise

[Cilasun et al., Scientific Reports 2024]



Conclusion

- Quantum-inspired Ising solvers can be manufactured in today's foundry processes
- QUBO challenges and future directions
 - Problem formulation for larger constrained optimization problems
 - Higher-order interactions beyond QUBO
 - Ongoing work on k -SAT shows promise
 - Mundane issues:
 - Huge QUBO file size, memory bottleneck
 - ASIC designs with integrated pre- and post-processing, PCIe, memory interface
- In future: Massively parallel solutions (1000 cores \equiv 10W)

Thanks to

- You, the audience
- Hüsrev Cılasun, Ahmet Efe, Ulya R. Karpuzcu, Chris H. Kim, Abhimanyu Kumar, Hao Lo, Will Moy, Nafisa Sadaf Prova, Ziqing Zeng
- DARPA QuICC, NSF ASCENT

