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

http://mcmahon.aep.cornell.edu


Funding: ImPACT Program, Cabinet Office of Japan
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.
Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

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

\[
\vec{\sigma} \triangleq (\sigma_1, \sigma_2, \ldots, \sigma_N)
\]

\[
\sigma_k \in \{-1, +1\}
\]

\[
J \in \mathbb{R}^{N \times N}
\]
Problem Statement: Given couplings between a set of spins, find the configuration that minimizes the energy function:

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

\[ \vec{\sigma} \triangleq (\sigma_1, \sigma_2, \ldots, \sigma_N) \]
\[ \sigma_k \in \{-1, +1\} \]

- 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:

\[ H(\vec{\sigma}) = - \sum_{1 \leq i < j \leq N} J_{ij}\sigma_i\sigma_j \]
The Ising Problem
(Combinatorial Optimization Version)

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

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

![Graph showing the Ising model with spins and couplings](image)

\[
J = \begin{pmatrix}
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]
**Problem Statement:** Given couplings between a set of spins, find the configuration that minimizes the energy function:

\[ H(\sigma) = - \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j \]
**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:

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

Figure credit: *Science* doi:10.1126/science.354.6310.269
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
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
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
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

MAX-CUT
MAX-CUT

Cut size = 3
MAX-CUT

Cut size = 4
Ising Formulation of MAX-CUT

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

\[ J = - \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]
Ising Formulation of MAX-CUT

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

\[ J = - \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]
Ising Formulation of MAX-CUT

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

\[ J = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]
Overview

• The Ising Problem

• **Ising Machines**
  • Foundations of All-Optical OPO Ising Machines
  • Measurement-Feedback OPO Ising Machines

• Conclusions
Ising Machines

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

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

Quantum Annealing
Ising Machines

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

Quantum Annealing

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

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

Quantum Annealing

CMOS Annealers
*ISSCC* 24.3 (2015)

Credit: dwavesys.com
Overview

• The Ising Problem
• Ising Machines

• Foundations of All-Optical OPO Ising Machines

• Measurement-Feedback OPO Ising Machines
• Conclusions
Optical Parametric Oscillators

Optical Parametric Amplifier

\[ \mathcal{E}_3(0) \rightarrow d_{\text{eff}} \rightarrow \mathcal{E}_3(L) = \mathcal{E}_3(0) \]

\[ \mathcal{E}_2(0) \rightarrow \mathcal{E}_2(L) \]

\[ z = 0 \]

\[ z = L \]

\[ \omega_3 \]

\[ \omega_2 \]

\[ \omega_1 \]
Optical Parametric Oscillators

![Diagram of Optical Parametric Amplifier](image1)

**Optical Parametric Amplifier**

\[ E_3(0) \rightarrow d_{eff} \rightarrow E_3(L) = E_3(0) \]
\[ E_2(0) \rightarrow E_2(L) \]
\[ z = 0 \quad z = L \]

**OPO: Optical Parametric Amplifier in a Cavity**

![Diagram of OPO in a Cavity](image2)

\[ E_3(0) \rightarrow d_{eff} \rightarrow E_{3, out} \]
\[ R(\omega_2) = 1 \]
\[ z = 0 \]
\[ E_{2, out} \]
\[ z = L \]
\[ R(\omega_2) = 1 - T(\omega_2) \]

Diagrams from: EE346 Lecture Notes, Stanford University (M. Fejer)
Optical Parametric Oscillators

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

Diagrams from: EE346 Lecture Notes, Stanford University (M. Fejer)
OPO Phase Properties

Squeezed vacuum state **below threshold**  Squeezed coherent state **above threshold**

![Diagram](image)

OPO Phase Properties

Squeezed vacuum state **below threshold**  

Squeezed coherent state **above threshold**

OPO Phase Properties

Squeezed vacuum state \textit{below threshold} \hspace{2cm} Squeezed coherent state \textit{above threshold}

OPO Phase Properties

Squeezed vacuum state **below threshold**  

Squeezed coherent state **above threshold**

OPO Phase Properties

Squeezed vacuum state below threshold  Squeezed coherent state above threshold


A. Marandi, *et al.* *Optics Express* 20, 1719322 (2012).
Time-Multiplexed OPOs

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

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

$$H = 0$$
From OPOs to Ising Machine

Add **coupling** that connects neighboring pulses.

$N = 4$ Ising Machine

$N = 4$ Ising Machine Results

$N = 4$ Ising Machine Results

$N = 4$ Ising Machine Results

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

\[ T_{rt} = 160 \cdot T_{rep} \]

- Measurement
- Feedback Calculations
- Injection
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

![Graph showing the scaling of cubic graphs](image-url)
Scaling of cubic graphs

![Graph of Success Probability vs Graph Size](image1)

![Graph of Runtime to obtain 99% Success Probability vs Graph Size](image2)
Scaling of cubic graphs

![Graph showing the scaling of cubic graphs](image-url)
Scaling of cubic graphs

![Graph 1: Success Probability vs Graph Size](image1)

![Graph 2: Runtime to obtain 99% Success Probability vs Graph Size](image2)
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

**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

Comparison to Quantum Annealing

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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

**Problem class:** Sherrington-Kirkpatrick spin-glass problems

\((J_{ij} \text{ are } -1 \text{ or } +1 \text{ uniformly at random})\)

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?

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

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

Connectivity makes a big difference!

Why?

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

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

Connectivity makes a big difference!

Why?

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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
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*: $\alpha$ depends on the machine or algorithm

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
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:* $\alpha$ depends on the machine or algorithm

How can we make quantum annealers with better connectivity?

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Comparison to Quantum Annealing

<table>
<thead>
<tr>
<th>SK</th>
<th>MAX-CUT (dense)</th>
<th>MAX-CUT (d = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>( DW2Q )</td>
<td>( CIM )</td>
</tr>
<tr>
<td>10</td>
<td>6.0 ( \mu ) s</td>
<td>25 ( \mu ) s</td>
</tr>
<tr>
<td>20</td>
<td>35 ( \mu ) s</td>
<td>100 ( \mu ) s</td>
</tr>
<tr>
<td>40</td>
<td>6.1 ms</td>
<td>0.4 ms</td>
</tr>
<tr>
<td>60</td>
<td>1.4 s</td>
<td>0.6 ms</td>
</tr>
<tr>
<td>80*</td>
<td>(400 s)</td>
<td>1.8 ms</td>
</tr>
<tr>
<td>100*</td>
<td>(10(^5) s)</td>
<td>3.0 ms</td>
</tr>
</tbody>
</table>

**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

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)

Problem class: Density=50% MAX-CUT
Choosing Optimal Annealing Times

(CIM (simulated))

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)

Problem class: Density=50% MAX-CUT
Choosing Optimal Annealing Times

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)
Choosing Optimal Annealing Times

(Figure from R. Hamerly*, T. Inagaki*, P.L. McMahon*, et al. arXiv:1805.05217)

Problem class: Density=50% MAX-CUT
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: \( \sigma_i^{(P)} = \begin{cases} 1 & \text{P atom at site } i \\ 0 & \text{no P atom at site } i \end{cases} \)


Collaboration with: Alan Aspuru-Guzik (Harvard) and Libor Veis (Czech Acad. Science)
OPO Ising Machine Mechanism

\[ H = - \sum_{ij} J_{ij} \sigma_i \sigma_j \]

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

Simulations: Density Sensitivity

\[ p_{\text{comp}} = 0.6 \ p_{\text{th}} \]

\[ p_{\text{comp}} = 1.3 \ p_{\text{th}} \]
On-Chip All-Optical Ising Machines using Slow Light

\[ T = 450 \text{ nK} \]
\[ \tau_{\text{Delay}} = 7.05 \pm 0.05 \mu \text{s} \]
\[ L = 229 \pm 3 \mu \text{m} \]
\[ v_g = 32.5 \pm 0.5 \text{ m s}^{-1} \]
