



 **Los Alamos**
NATIONAL LABORATORY
— EST. 1943 —

Delivering science and technology
to protect our nation
and promote world stability



Unraveling Quantum Annealers with Machine Learning

Marc Vuffray, PIML 2018



**Joint work with A. Lokhov, Y. Kharkov,
C. Coffrin, M. Chertkov, S. Misra**

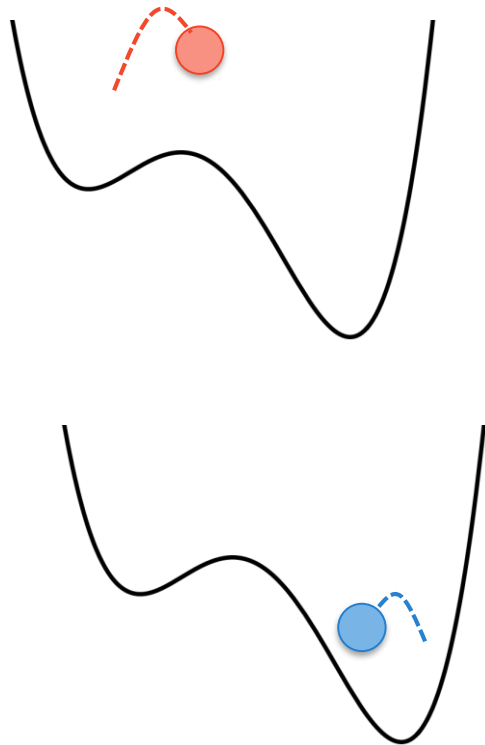


Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Quantum Annealing

Quantum Annealing is a Global Optimization Heuristic

Classical Annealing:



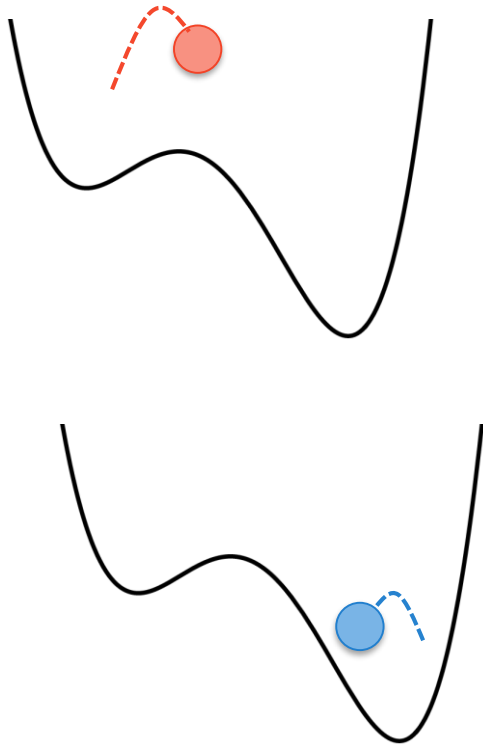
Interpolation Process

Temperature

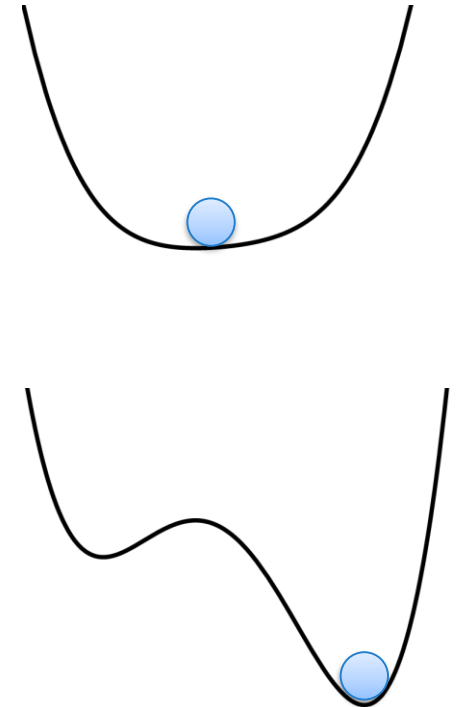


Quantum Annealing is a Global Optimization Heuristic

Classical Annealing:



Quantum Annealing:



Interpolation Process

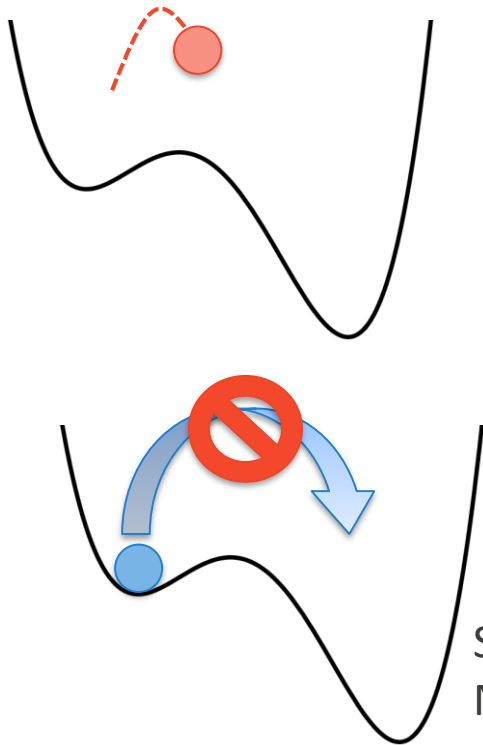


Temperature

Objective

Quantum Annealing is a Global Optimization Heuristic

Classical Annealing:



Stuck in Local Minimum

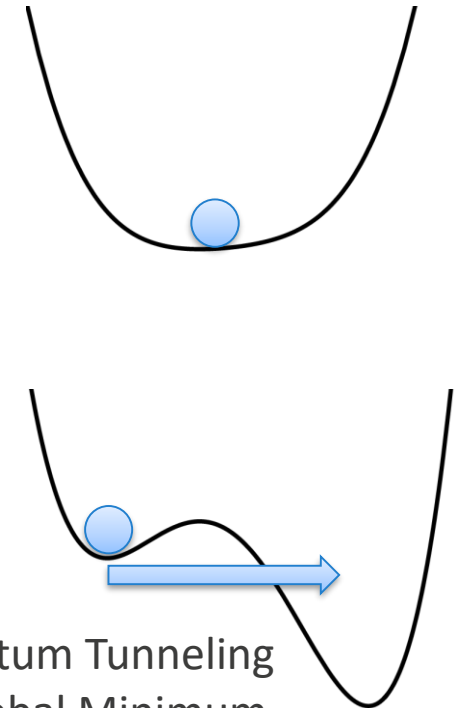
Interpolation Process



Temperature

Objective

Quantum Annealing:



Quantum Tunneling to Global Minimum

Quantum Annealing for Solving QUBO

Quadratic
Unconstrained
Binary
Optimization

Quantum Annealing for Solving QUBO

Quadratic
Unconstrained
Binary
Optimization

$$\arg \min_{\underline{x}} \sum_{i=1}^N b_i x_i + \sum_{i=1}^N \sum_{j=1}^i a_{ij} x_i x_j$$

s.t. $\forall i, x_i \in \{0,1\}$

Input: $a_{ij}, b_i \in \mathbb{R}$

Output: $x_i \in \{0,1\}$

Quantum Annealing for Solving QUBO

Quadratic
Unconstrained
Binary
Optimization

$$\begin{aligned} \arg \min_{\underline{x}} \quad & \sum_{i=1}^N b_i x_i + \sum_{i=1}^N \sum_{j=1}^i a_{ij} x_i x_j \\ \text{s.t.} \quad & \forall i, x_i \in \{0,1\} \end{aligned}$$

Input: $a_{ij}, b_i \in \mathbb{R}$

Output: $x_i \in \{0,1\}$

NP Hard

General class of combinatorial problems

$$E(\underline{x}) = \underline{b}^T \underline{x} + \underline{x}^T \underline{A} \underline{x}$$

QUBO = Finding Ground States of Ising Models

Quadratic
Unconstrained
Binary
Optimization

$$\arg \min_{\underline{x}} \sum_{i=1}^N b_i x_i + \sum_{i=1}^N \sum_{j=1}^i a_{ij} x_i x_j$$

s.t. $\forall i, x_i \in \{0,1\}$

Input: $a_{ij}, b_i \in \mathbb{R}$

Output: $x_i \in \{0,1\}$

QUBO = Finding Ground States of Ising Models

Quadratic
Unconstrained
Binary
Optimization

$$\arg \min_{\underline{x}} \sum_{i=1}^N b_i x_i + \sum_{i=1}^N \sum_{j=1}^i a_{ij} x_i x_j$$

s.t. $\forall i, x_i \in \{0,1\}$

Input: $a_{ij}, b_i \in \mathbb{R}$

Output: $x_i \in \{0,1\}$

Ground States
of Ising Models

$$\arg \min_{\underline{\sigma}} \sum_{i=1}^N h_i \sigma_i + \sum_{i=1}^N \sum_{j=1}^i J_{ij} \sigma_i \sigma_j$$

s.t. $\forall i, \sigma_i \in \{-1, +1\}$

Input: $J_{ij}, h_i \in \mathbb{R}$

Output: $\sigma_i \in \{-1, +1\}$

QUBO = Finding Ground States of Ising Models

Ground States
of Ising Models

$$\arg \min_{\underline{\sigma}} \sum_{i=1}^N h_i \sigma_i + \sum_{i=1}^N \sum_{j=1}^i J_{ij} \sigma_i \sigma_j$$

s.t. $\forall i, \sigma_i \in \{-1, +1\}$

Input: $J_{ij}, h_i \in \mathbb{R}$

Output: $\sigma_i \in \{-1, +1\}$

QUBO = Finding Ground States of Ising Models

Ground States
of Ising Models

$$\arg \min_{\underline{\sigma}} \sum_{i=1}^N h_i \sigma_i + \sum_{i=1}^N \sum_{j=1}^i J_{ij} \sigma_i \sigma_j$$

s.t. $\forall i, \sigma_i \in \{-1, +1\}$

Input: $J_{ij}, h_i \in \mathbb{R}$

Output: $\sigma_i \in \{-1, +1\}$

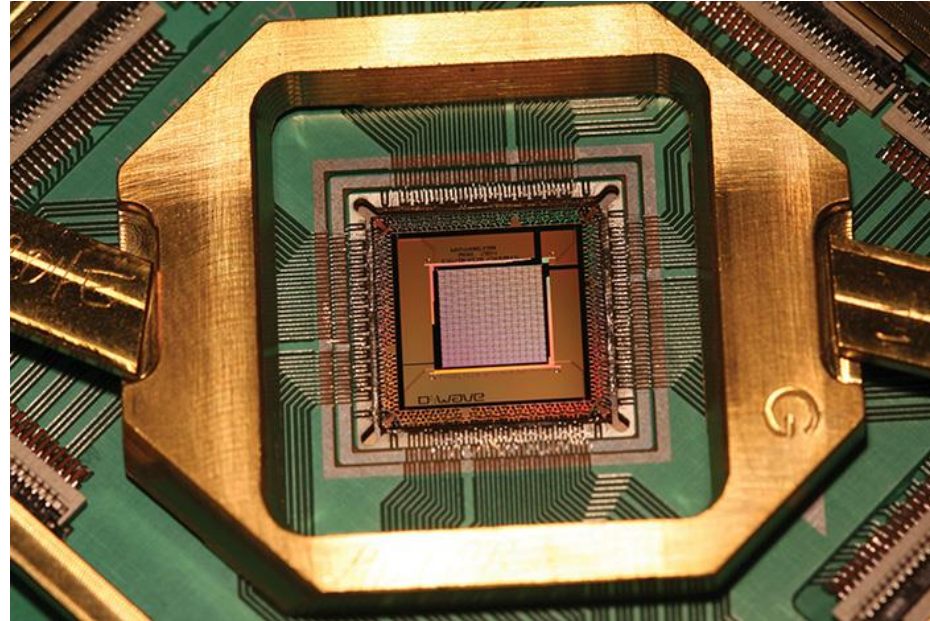
Magnetic field = h_i

Coupler = J_{ij}

Spin = σ_i

The D-Wave Implementation

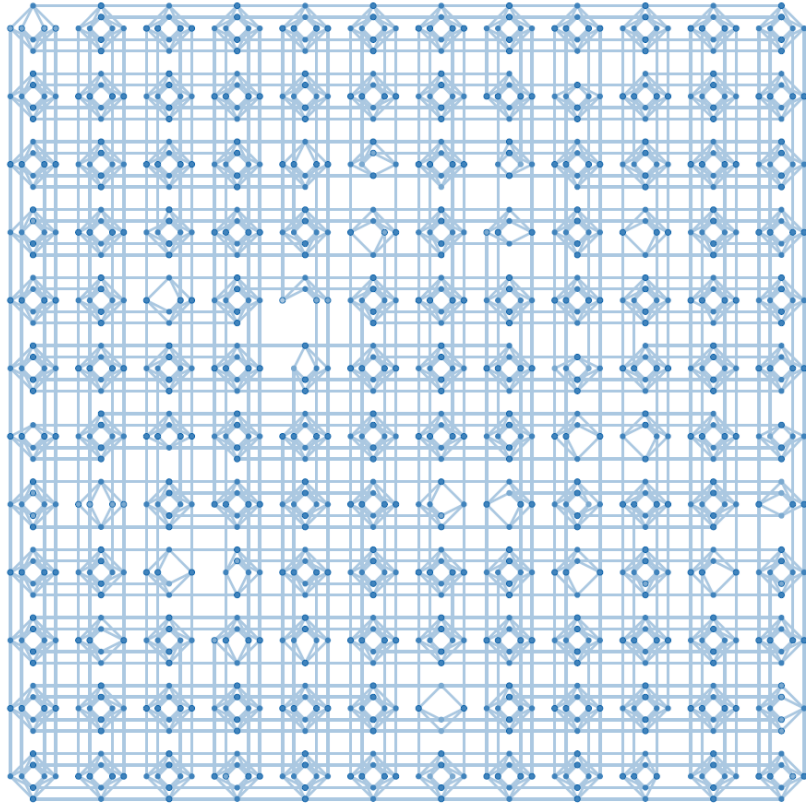
D-Wave: A Quantum Annealer for Solving QUBO



# of qubits	1,095 (95.1%)
# of couplers	3,061 (91.1%)
Temperature	10.45 mK

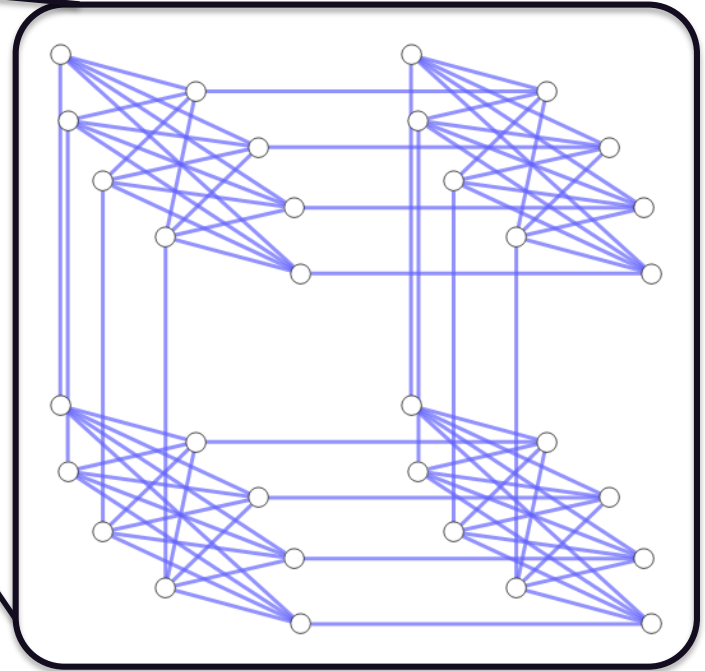
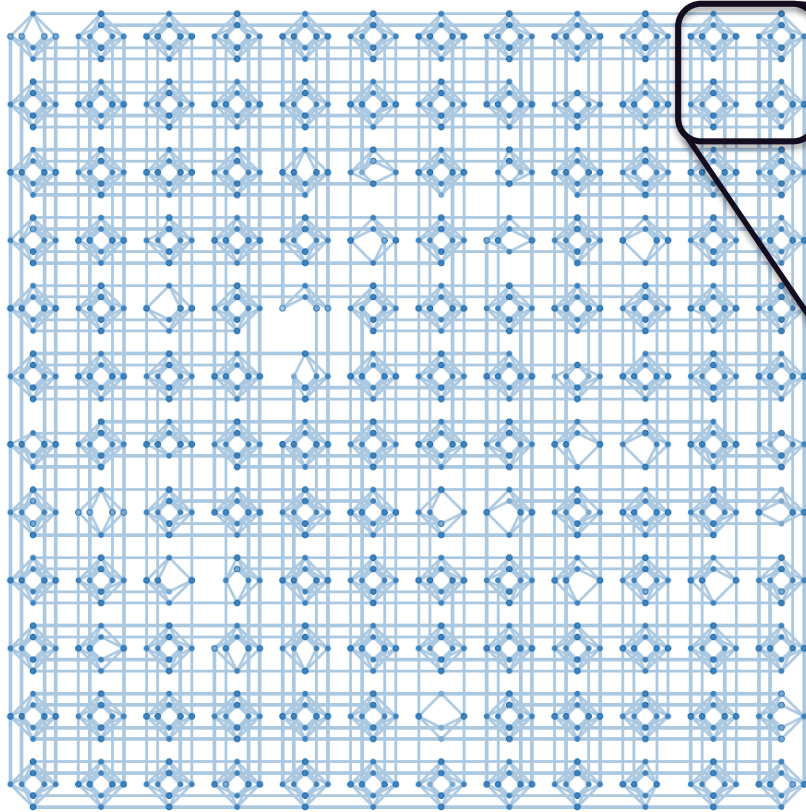
Annealing time	5–2000 μ s
h range	[-2, +2]
J range	[-1, +1]

The Set of Possible Couplings: A Bipartite Graph



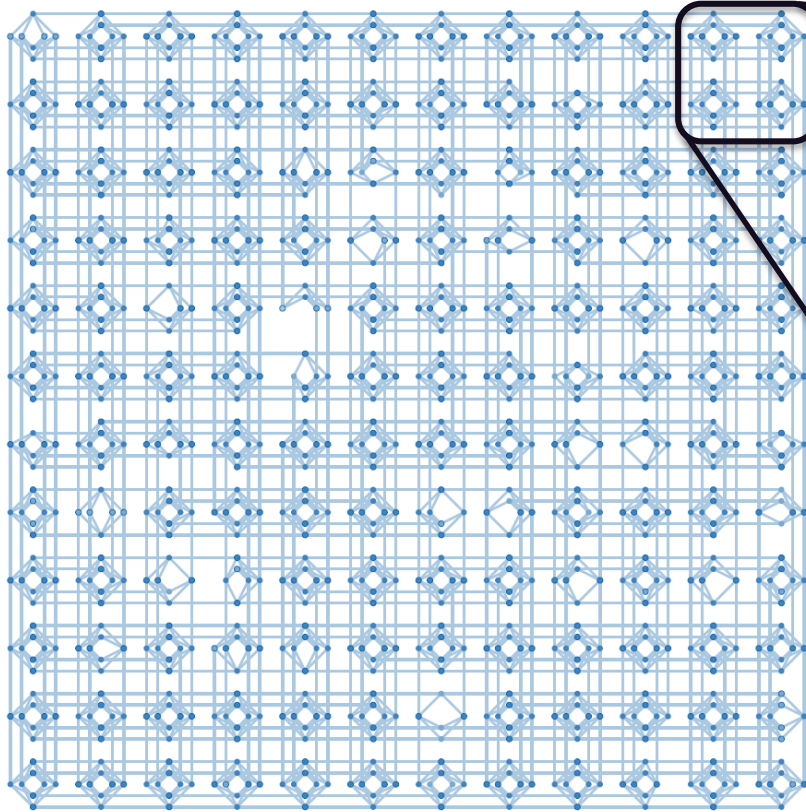
Ensemble of Spins and Couplers

The Set of Possible Couplings: A Bipartite Graph

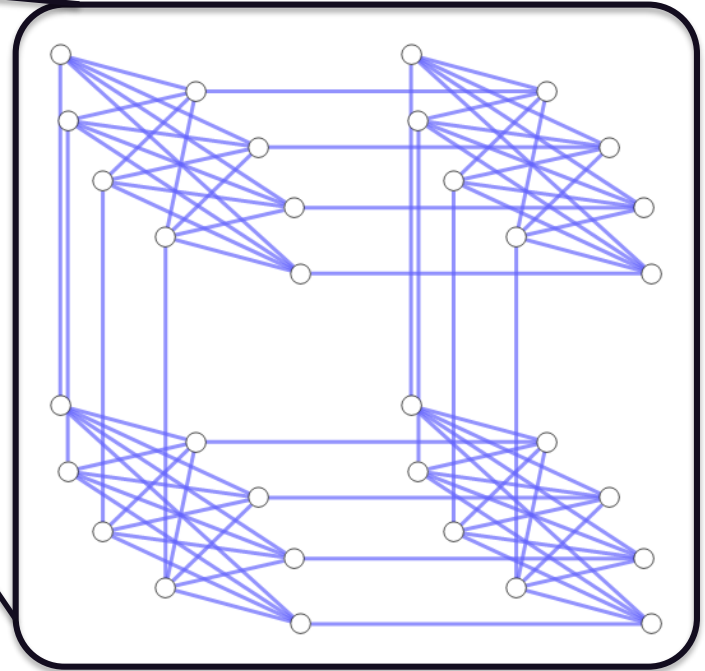


Ensemble of Spins and Couplers

The Set of Possible Couplings: A Bipartite Graph



Ensemble of Spins and Couplers



$$\sigma_i \text{ --- } J_{ij} \sigma_i \sigma_j \text{ --- } \sigma_j$$

Programming the D-Wave



Programming the D-Wave



Question 1: Can we characterize the probability distribution on $\underline{\sigma}$?

Programming the D-Wave



Question 1: Can we characterize the probability distribution on $\underline{\sigma}$?

Question 2: What can we learn about the machine from the probability distribution?

Programming the D-Wave



Question 1: Can we characterize the probability distribution on $\underline{\sigma}$?

Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

Interaction Screening for Solving Inverse Problems

The Generalized Inverse Ising Problem

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

The Generalized Inverse Ising Problem

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp\left(\sum_i h_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j + \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k + \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l + \dots\right)$$

The Generalized Inverse Ising Problem

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

The Generalized Inverse Ising Problem

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

Inverse Ising = Learning \mathbf{h} & \mathbf{J} from i.i.d. samples $\underline{\sigma}^{(1)}, \underline{\sigma}^{(2)}, \dots, \underline{\sigma}^{(M)}$

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

✓ **Exact**, tractable & parallelizable algorithm

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

- ✓ **Exact**, tractable & parallelizable algorithm
- ✓ Works for **arbitrary models**: low temperature, spin glass, dense, ...

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

- ✓ **Exact**, tractable & parallelizable algorithm
- ✓ Works for **arbitrary models**: low temperature, spin glass, dense, ...
- ✓ Requires **IT optimal** number of samples

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

- ✓ **Exact**, tractable & parallelizable algorithm
- ✓ Works for **arbitrary models**: low temperature, spin glass, dense, ...
- ✓ Requires **IT optimal** number of samples



Interaction Screening: Efficient and Sample-Optimal Learning of Ising Models

M. Vuffray, S. Misra, A. Lokhov, M. Chertkov

(2016)



Optimal Structure and Parameter Learning of Ising Models

A. Lokhov, M. Vuffray, S. Misra, M. Chertkov

(In press)

The Restricted Interaction Screening Estimator (RISE)

Estimated couplings are the results of a convex optimization:

$$\arg \min_{h,J} \mathbb{E} \left[\exp \left(- \sum_i h_i \sigma_i - \sum_{ij} J_{ij} \sigma_i \sigma_j - \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k - \sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l - \dots \right) \right]$$

- ✓ **Exact**, tractable & parallelizable algorithm
- ✓ Works for **arbitrary models**: low temperature, spin glass, dense, ...
- ✓ Requires **IT optimal** number of samples



Interaction Screening: Efficient and Sample-Optimal Learning of Ising Models

M. Vuffray, S. Misra, A. Lokhov, M. Chertkov

(2016)



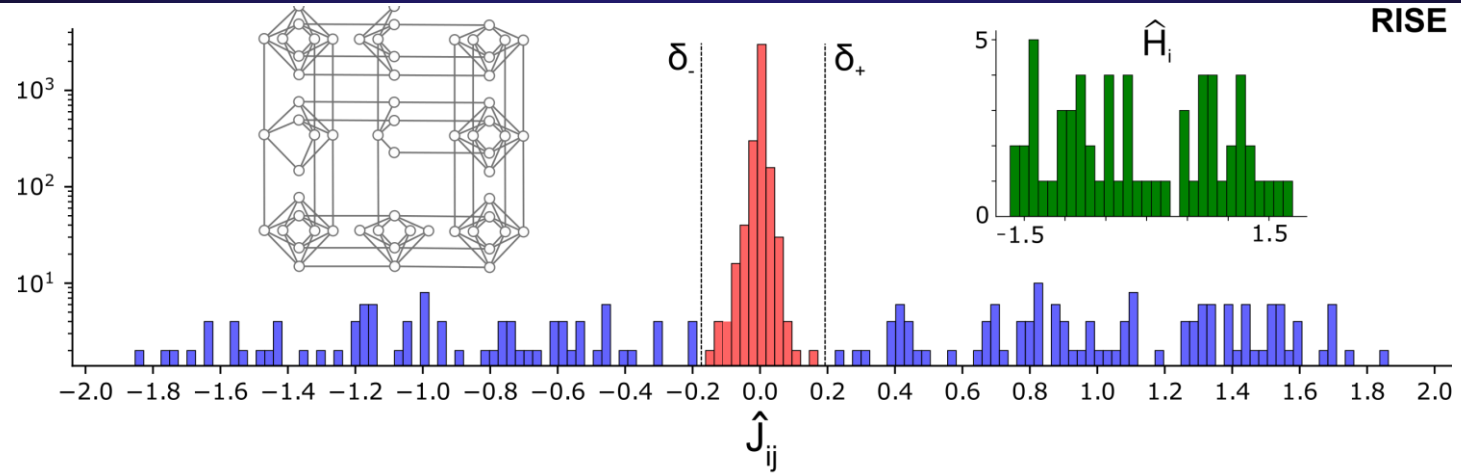
Optimal Structure and Parameter Learning of Ising Models

A. Lokhov, M. Vuffray, S. Misra, M. Chertkov

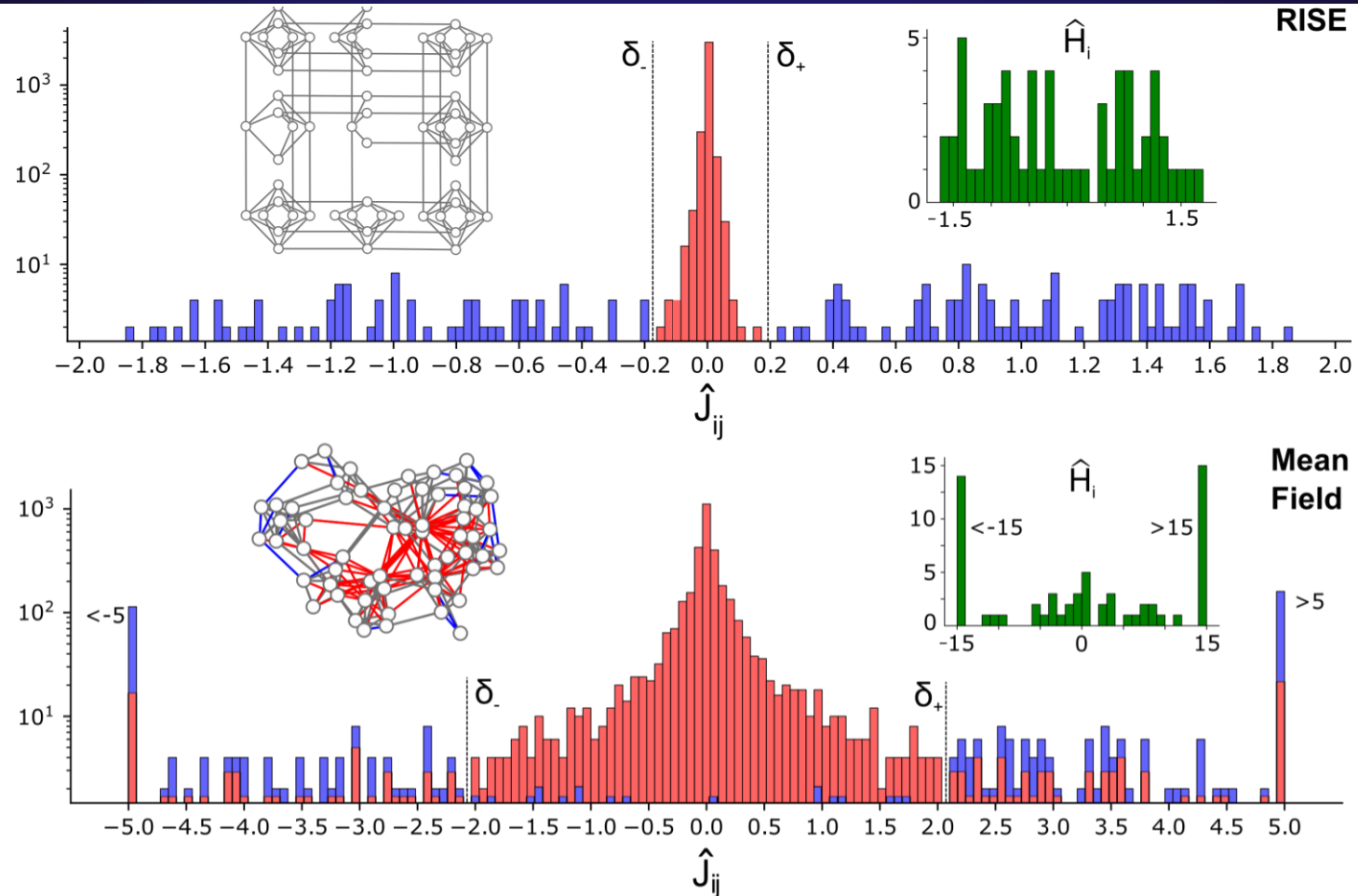
(In press)

GitHub https://github.com/lanl-ansi/inverse_ising

Comparing RISE with Mean Field Methods



Comparing RISE with Mean Field Methods



Going Back to Studying D-Wave



✓ **Question 1:** Can we characterize the probability distribution on $\underline{\sigma}$?

Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

Going Back to Studying D-Wave



✓ **Question 1:** Can we characterize the probability distribution on $\underline{\sigma}$? Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

Going Back to Studying D-Wave



✓ **Question 1:** Can we characterize the probability distribution on $\underline{\sigma}$? Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

- **Multi-Body Interactions? (orders >2?)**

Going Back to Studying D-Wave



✓ **Question 1:** Can we characterize the probability distribution on $\underline{\sigma}$? Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

- **Multi-Body Interactions? (orders >2?)**
- **Structure of Interactions?**

Going Back to Studying D-Wave



✓ **Question 1:** Can we characterize the probability distribution on $\underline{\sigma}$? Inverse Problem

Question 2: What can we learn about the machine from the probability distribution?

- **Multi-Body Interactions? (orders >2?)**
- **Structure of Interactions?**
- **Couplings Input-Output Response?**

Type of Multi-Body Interactions

What are Multi-Body Interactions?

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

What are Multi-Body Interactions?

Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

↑
Single Spin Term

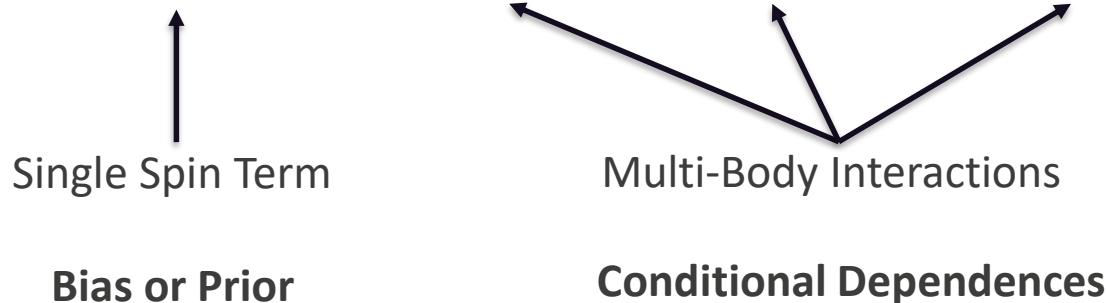
↙ ↘
Multi-Body Interactions

What are Multi-Body Interactions?

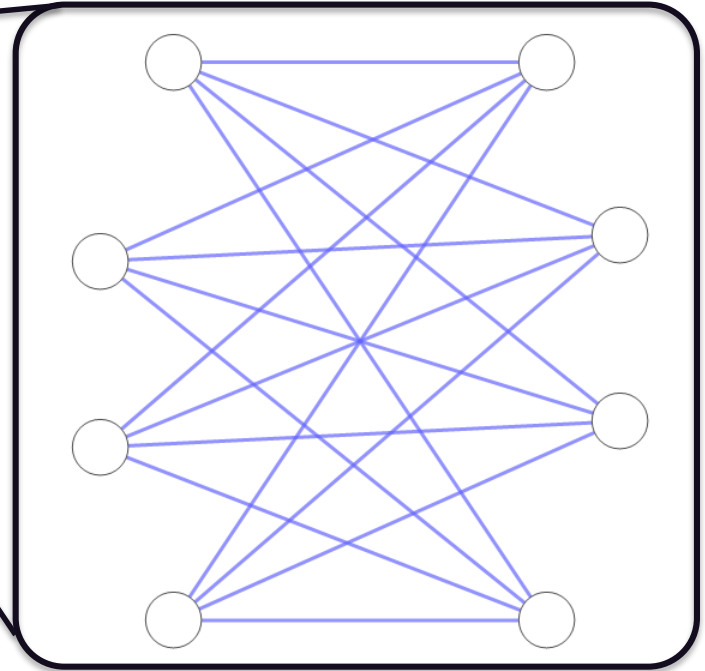
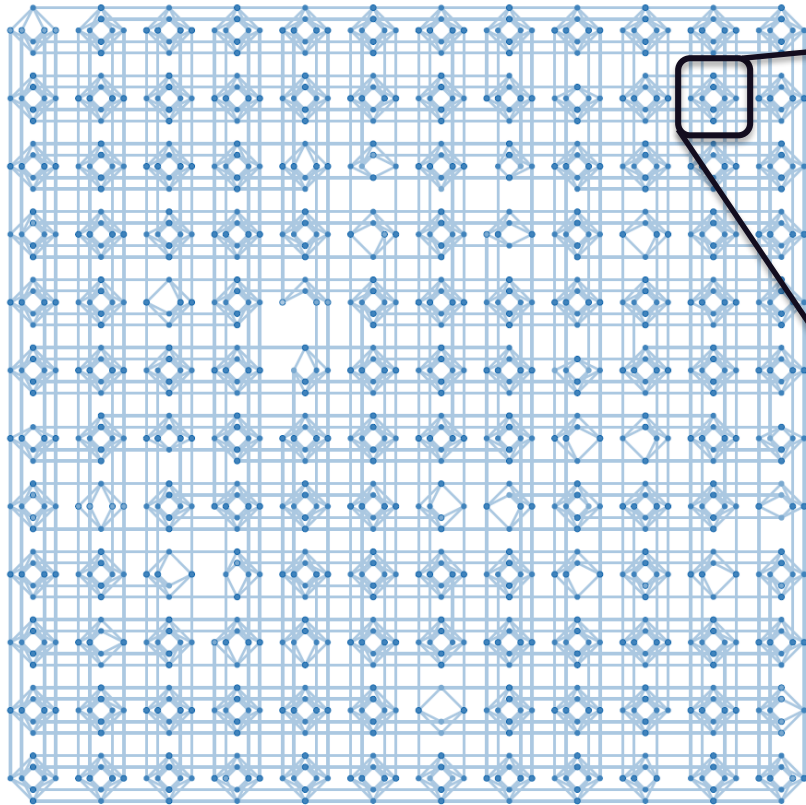
Random spin configurations: $\underline{\sigma} \in \{-1, +1\}^N$

with probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

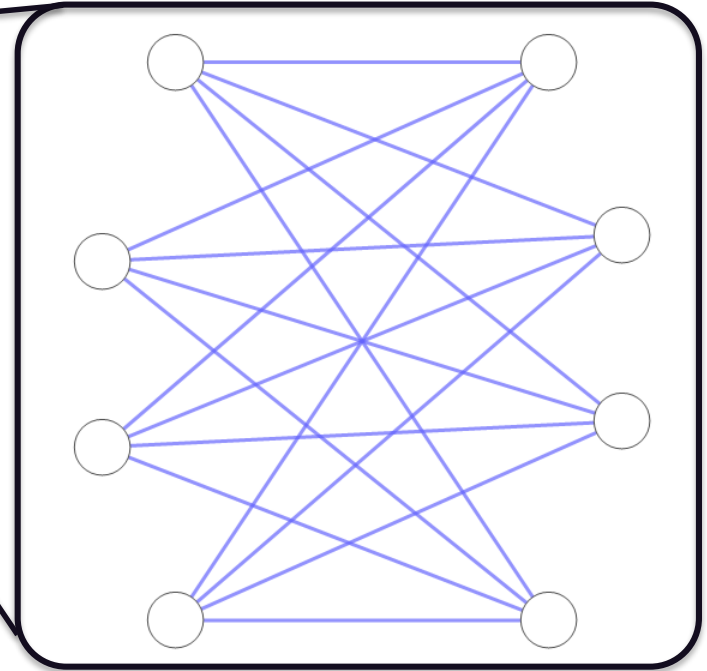
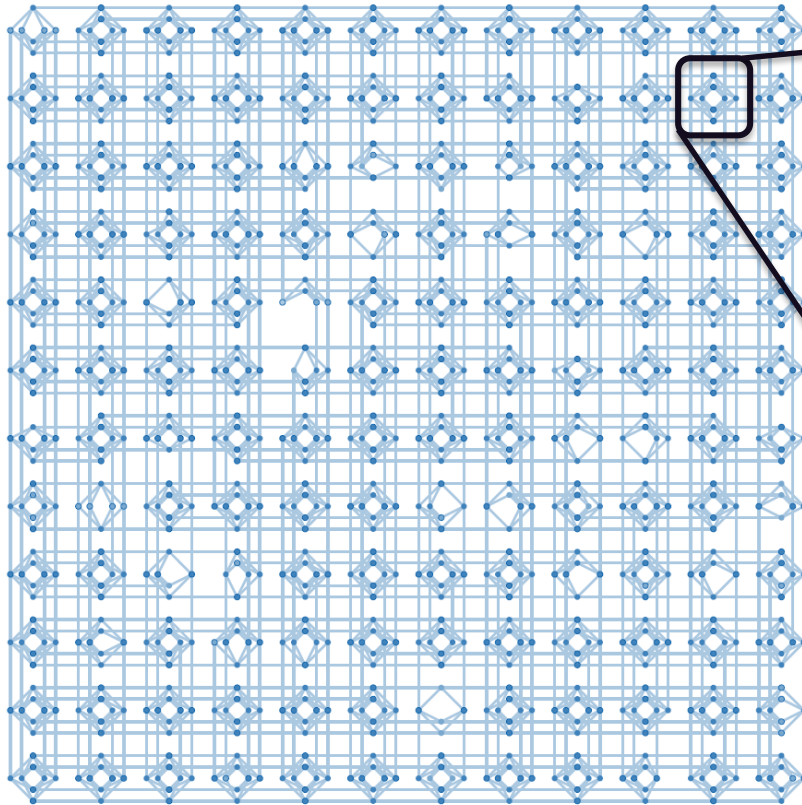


Multi-Body Interactions: Results on Chimera Cell



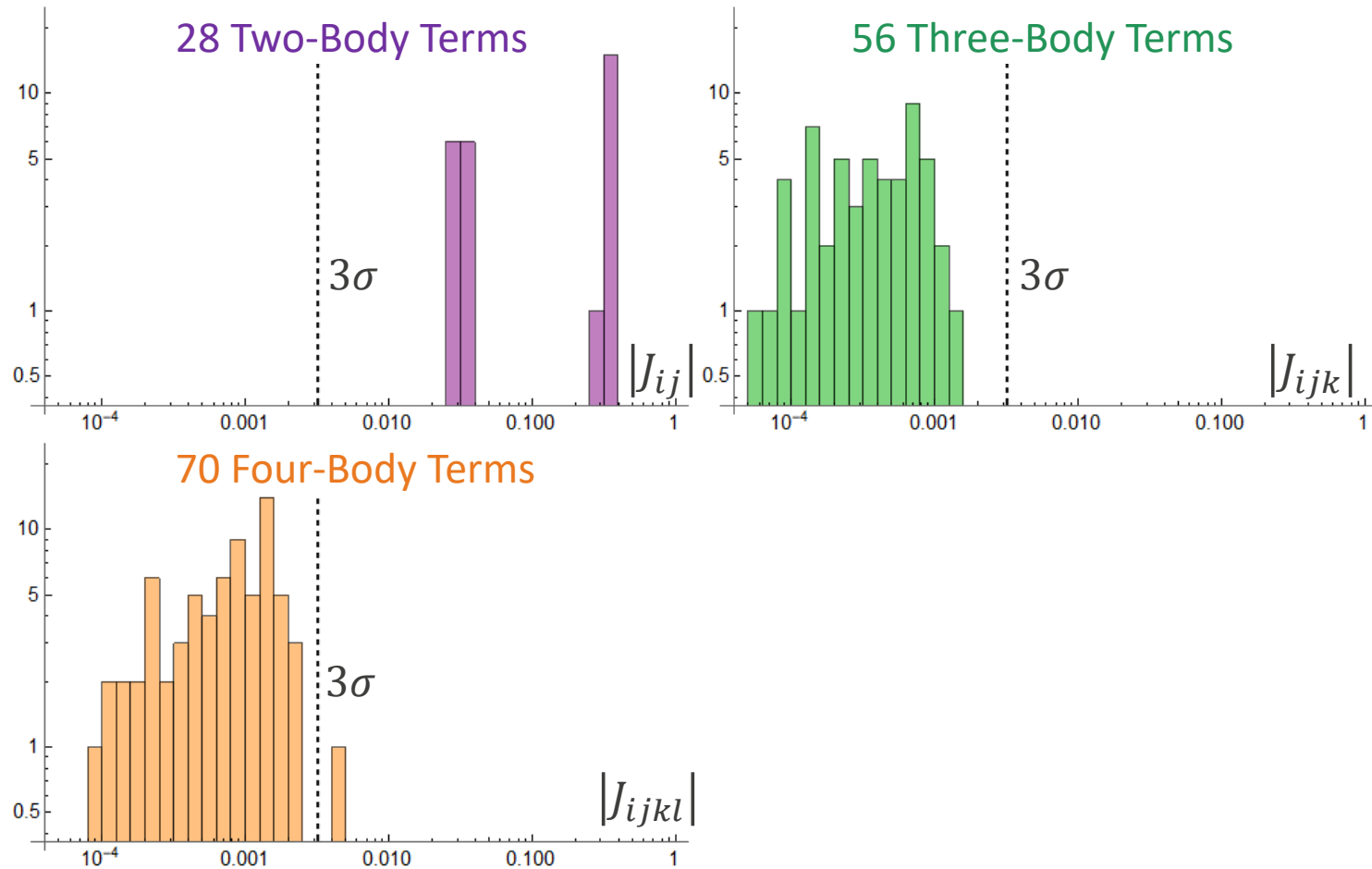
1 Chimera Cell: - 8 Spins

Multi-Body Interactions: Results on Chimera Cell

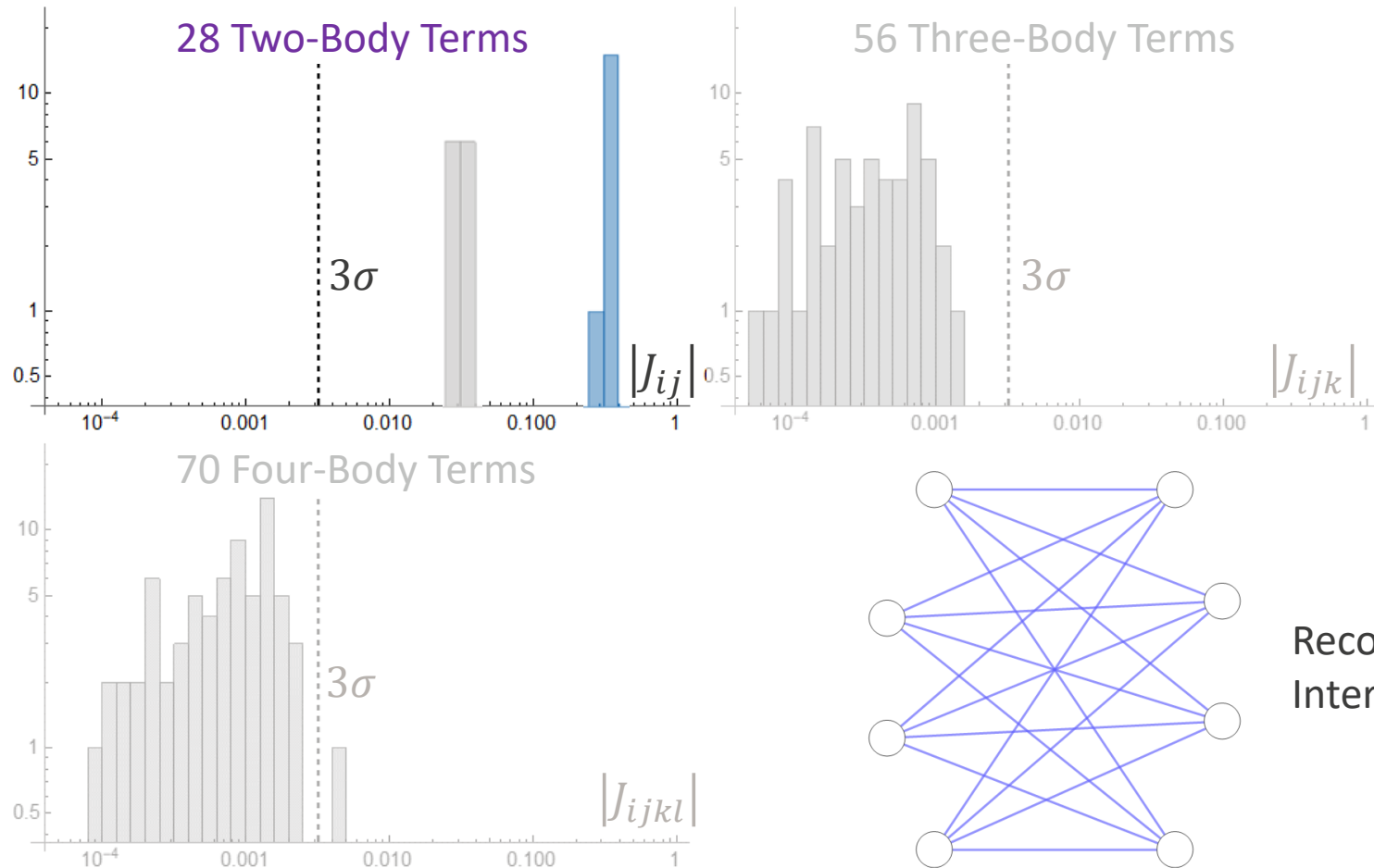


- 1 Chimera Cell: - 8 Spins
- 16 **Input** Couplers set to 0.025
- 0 **Input** Magnetic Fields

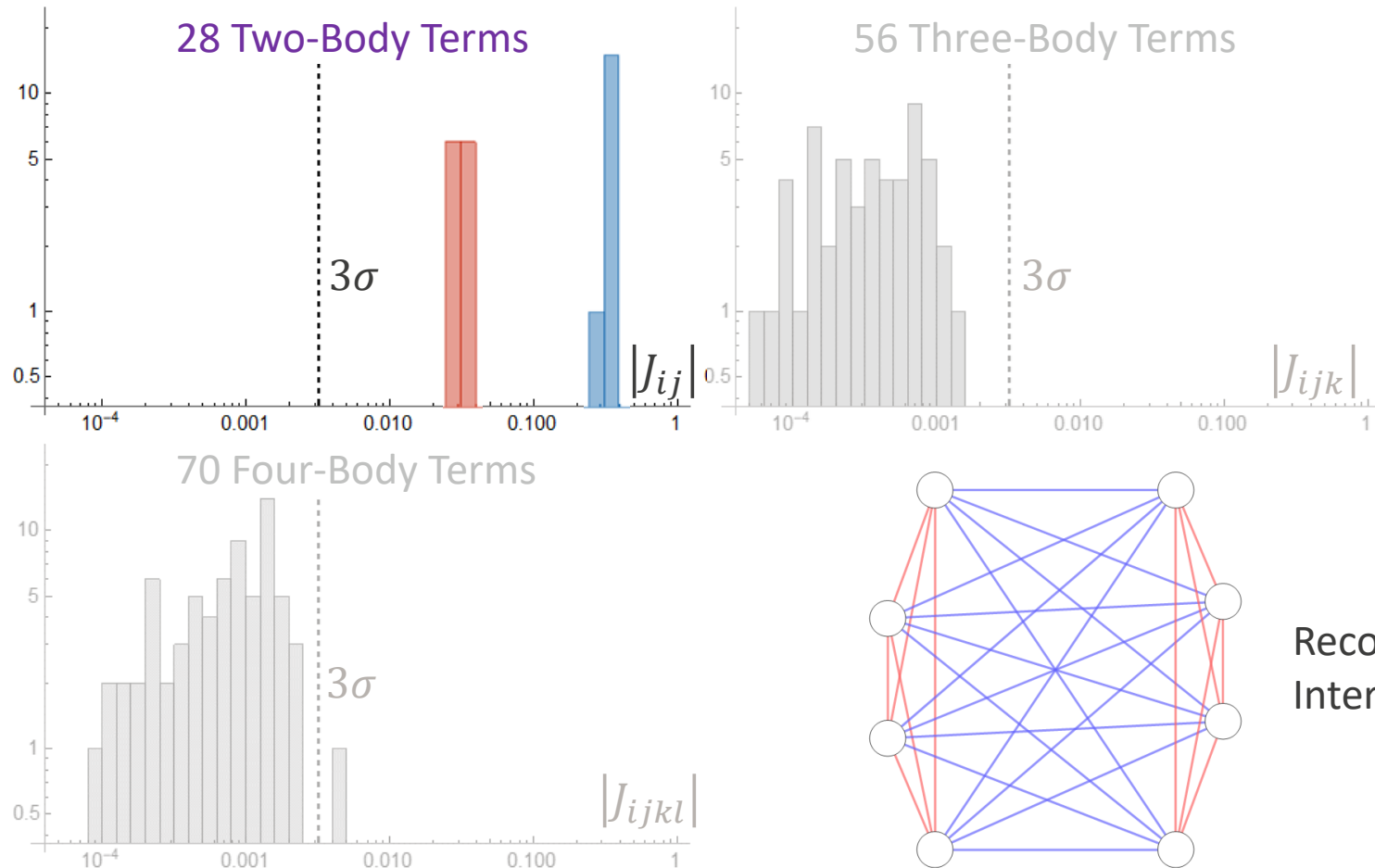
Log-Histogram of Reconstructed Multi-Interactions



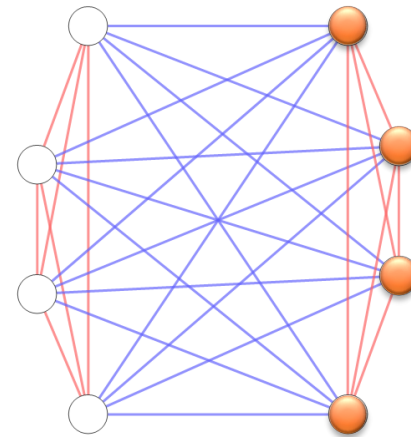
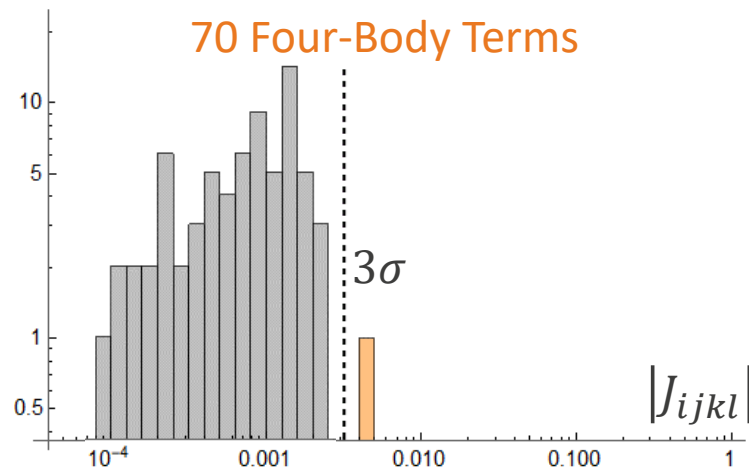
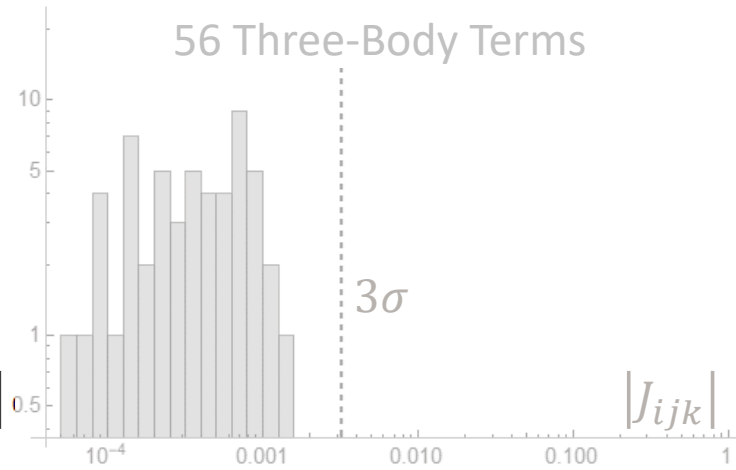
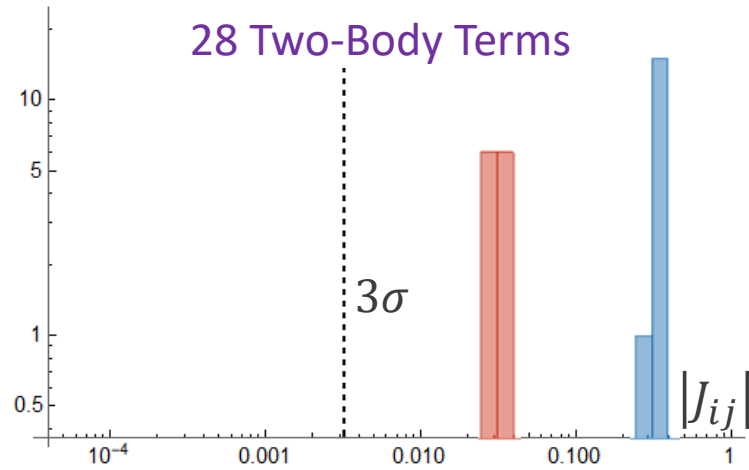
Log-Histogram of Reconstructed Multi-Interactions



Log-Histogram of Reconstructed Multi-Interactions



Log-Histogram of Reconstructed Multi-Interactions



Reconstructed Interactions

Conclusion: Pair Interactions are Dominant

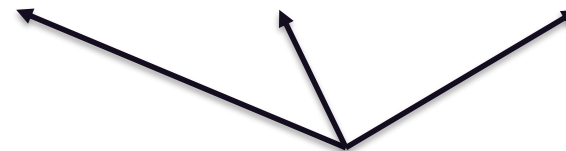
D-Wave Random Configurations: $\underline{\sigma} \in \{-1, +1\}^N$

have probability distribution:

$$\mu(\underline{\sigma}) \propto \exp \left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}} + \underbrace{\sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k}_{3^{\text{rd}} \text{ Order}} + \underbrace{\sum_{ijkl} J_{ijkl} \sigma_i \sigma_j \sigma_k \sigma_l}_{4^{\text{th}} \text{ Order}} + \dots \right)$$

↑
Single Spin Term

Bias or Prior



Multi-Body Interactions

Conditional Dependences

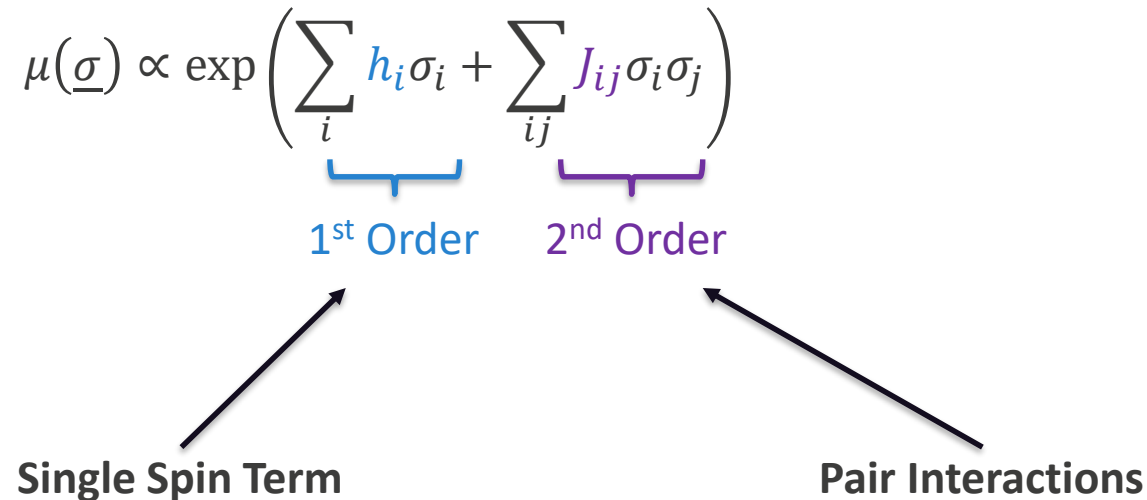
Conclusion: Pair Interactions are Dominant

D-Wave Random Configurations: $\underline{\sigma} \in \{-1, +1\}^N$

have probability distribution:

$$\mu(\underline{\sigma}) \propto \exp\left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}}\right)$$

Single Spin Term Pair Interactions



Conclusion: Pair Interactions are Dominant

D-Wave Random Configurations: $\underline{\sigma} \in \{-1, +1\}^N$

have probability distribution:

$$\mu(\underline{\sigma}) \propto \exp\left(\underbrace{\sum_i h_i \sigma_i}_{1^{\text{st}} \text{ Order}} + \underbrace{\sum_{ij} J_{ij} \sigma_i \sigma_j}_{2^{\text{nd}} \text{ Order}}\right)$$

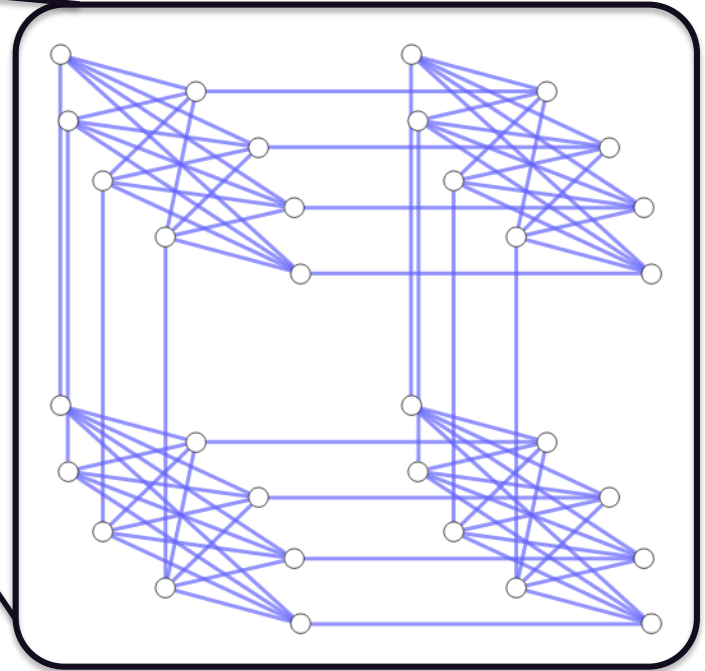
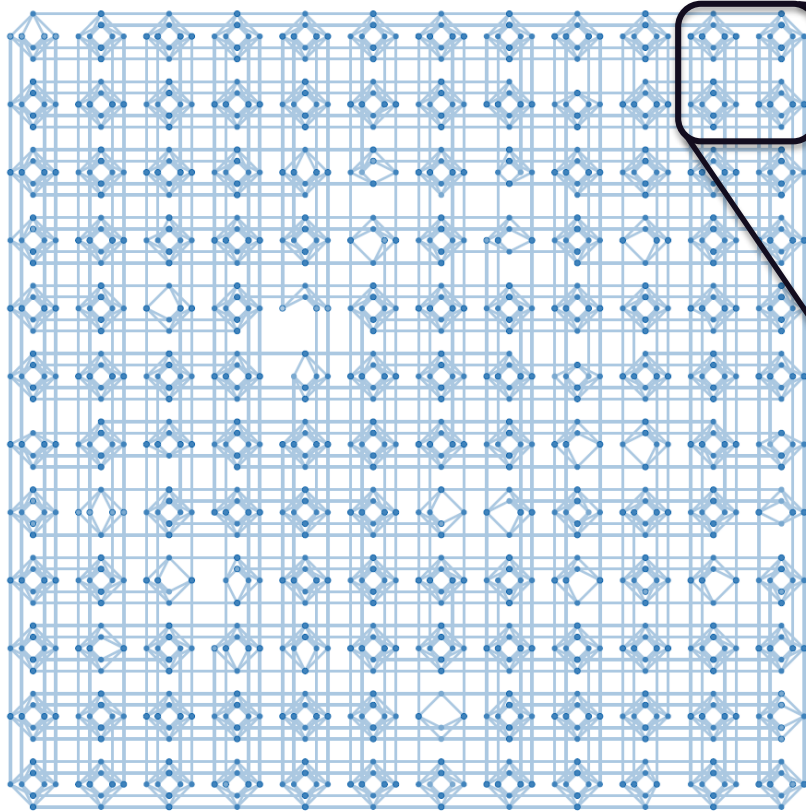
Single Spin Term

Pair Interactions

Question: What is the structure spanned by pair interactions?

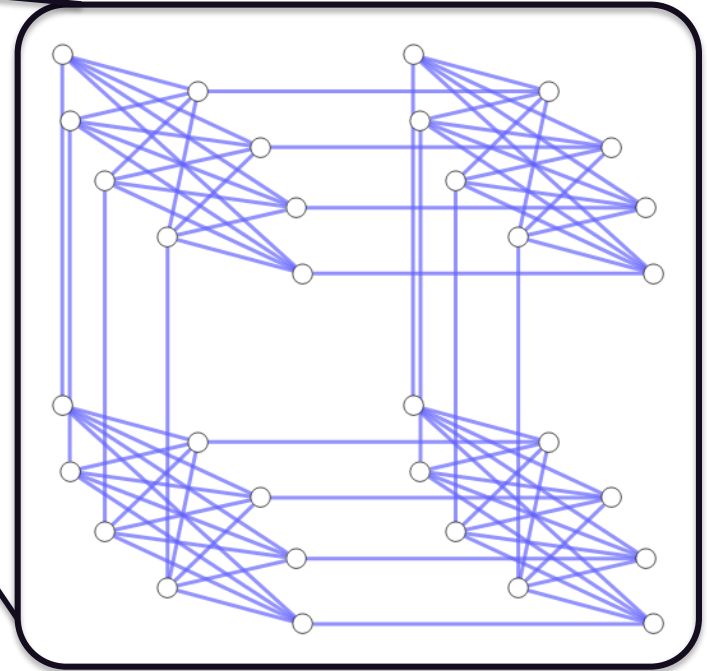
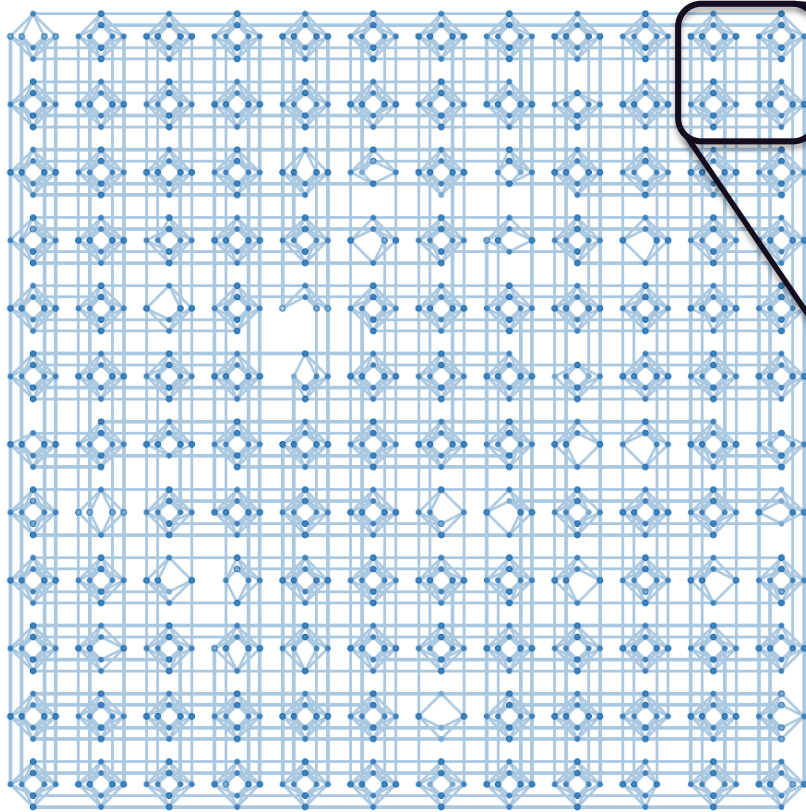
The Structure of the Two-Body Interactions

Pair Interactions: Results on 4 Chimera Cells



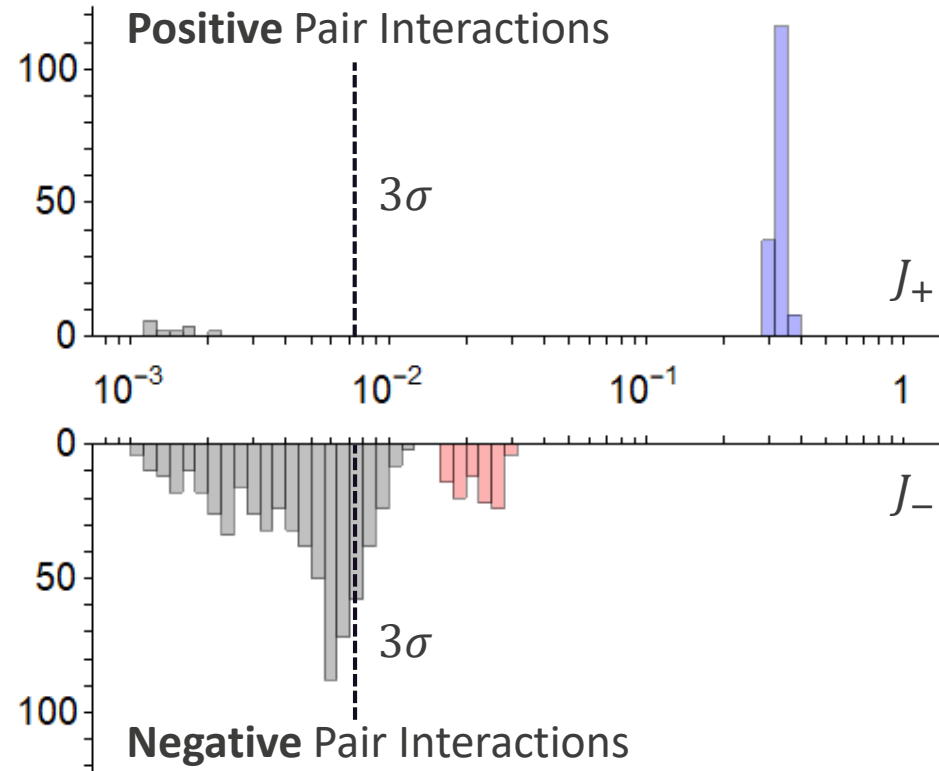
4 Chimera Cells: - 32 Spins

Pair Interactions: Results on 4 Chimera Cells

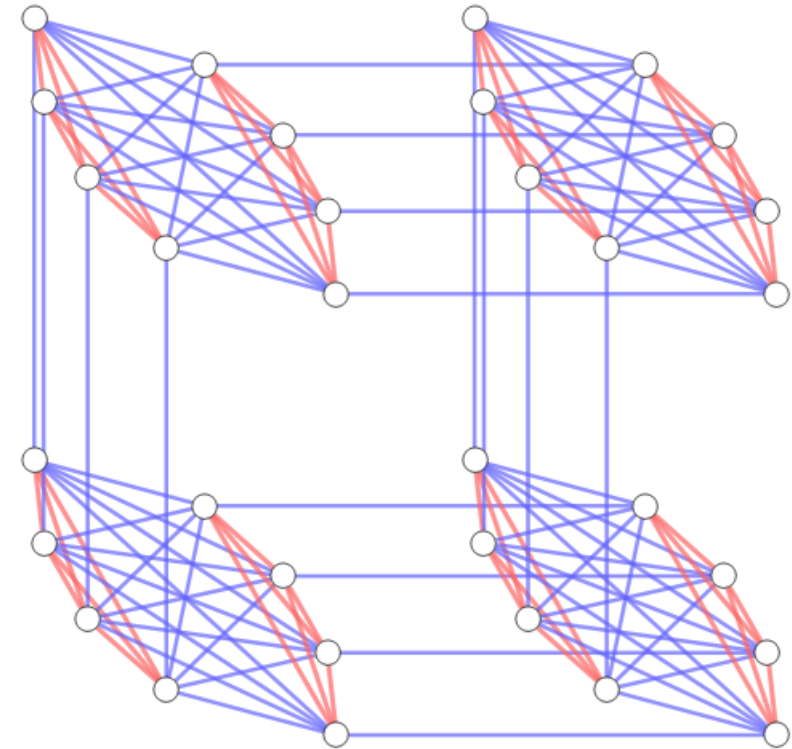
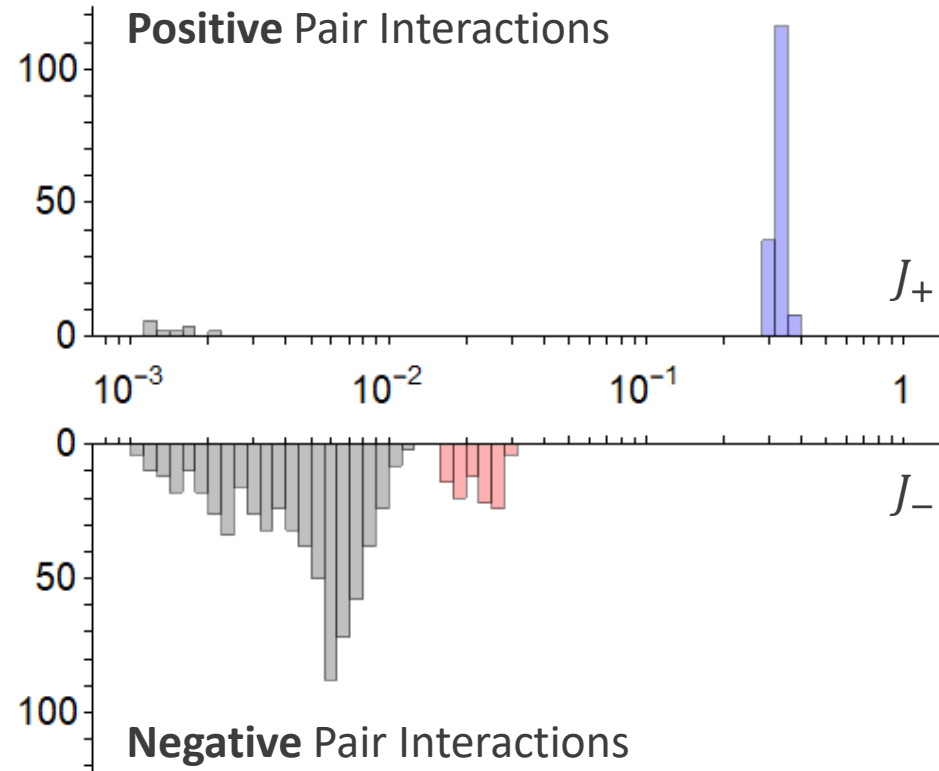


4 Chimera Cells: - 32 Spins
- 80 **Input** Couplers set to 0.025
- 0 **Input** Magnetic Fields

The Structure of Pair-Interactions

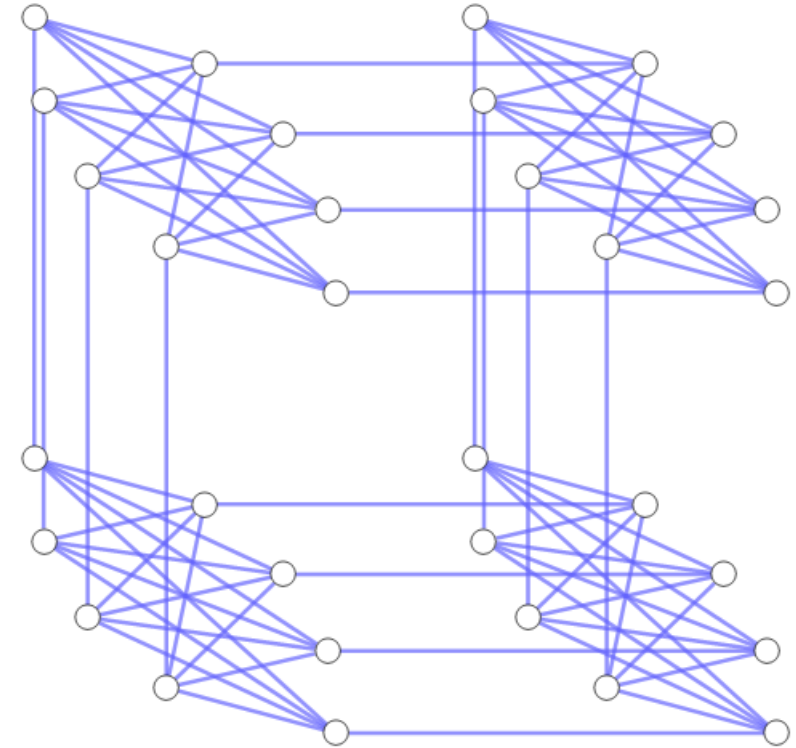
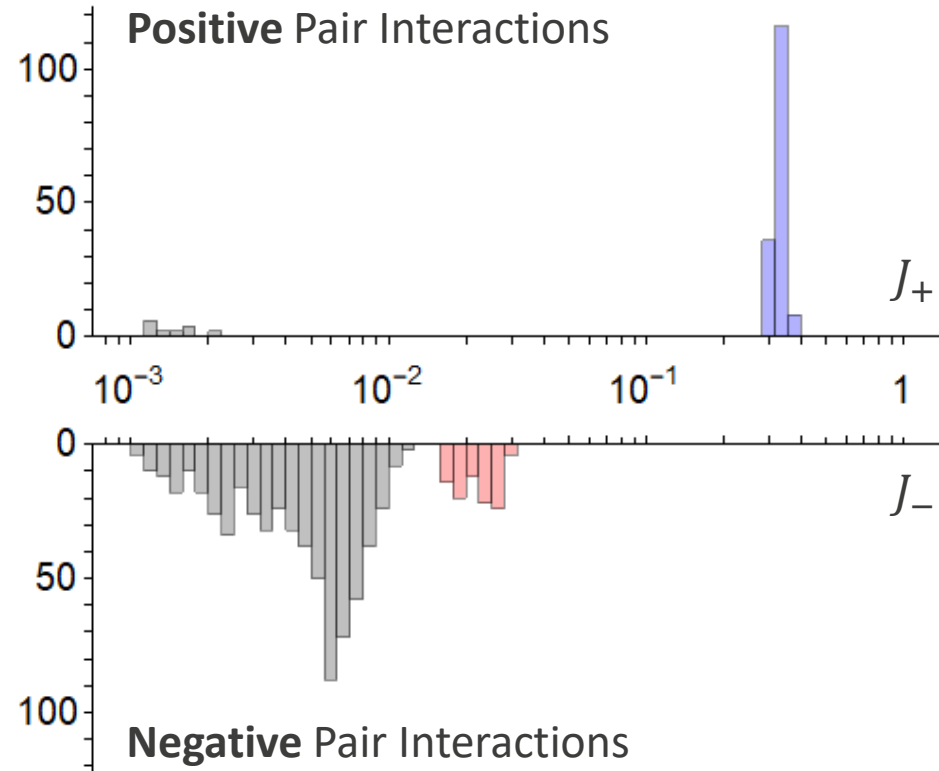


The Structure of Pair-Interactions



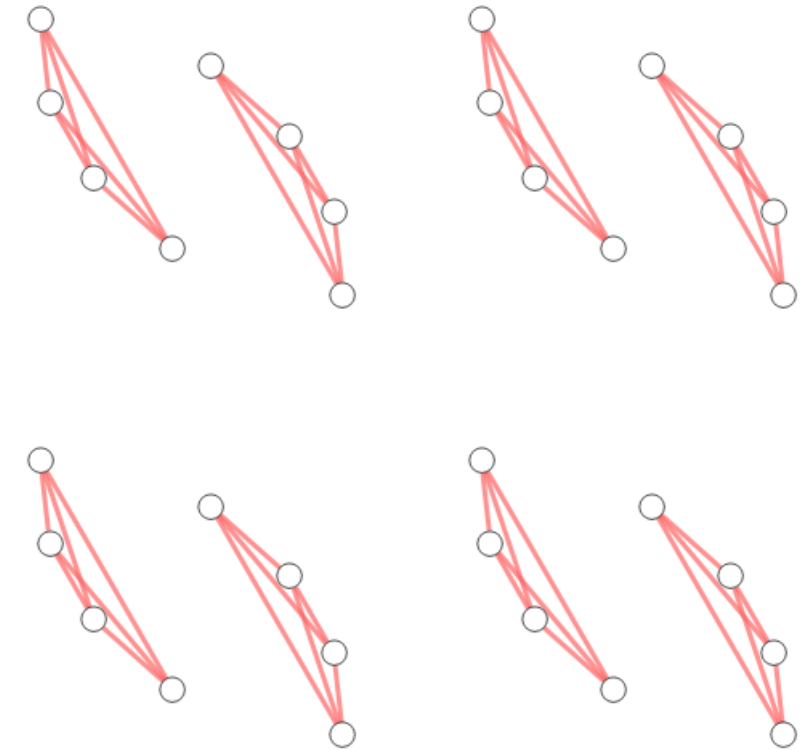
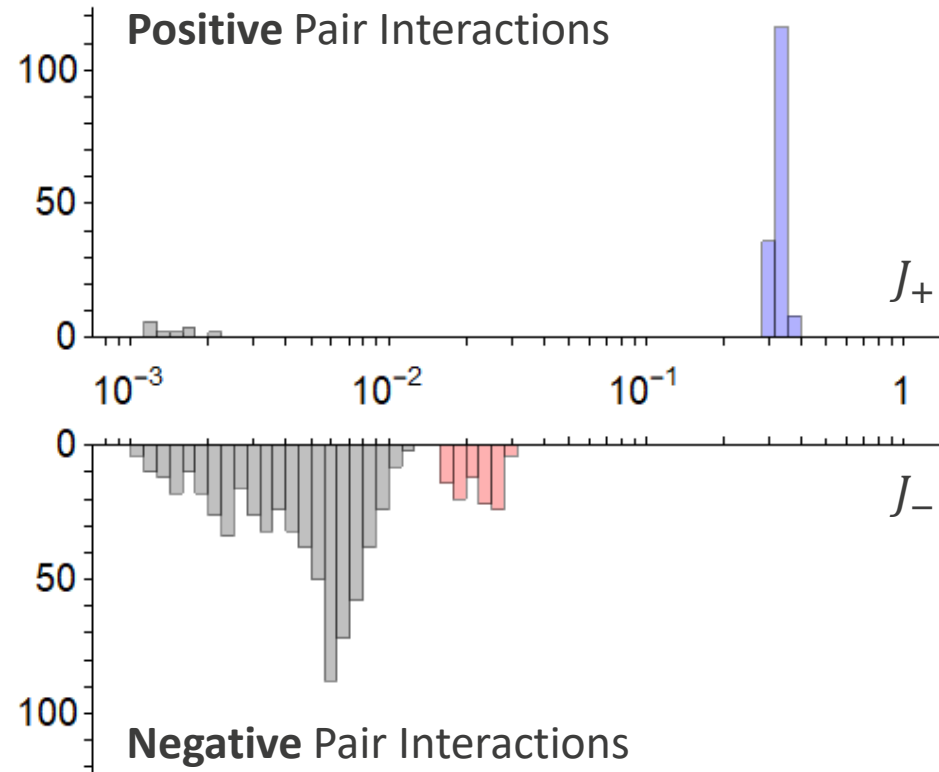
Reconstructed Structure

The Structure of Pair-Interactions



Most **Significant** Interactions \equiv D-Wave **Couplers**

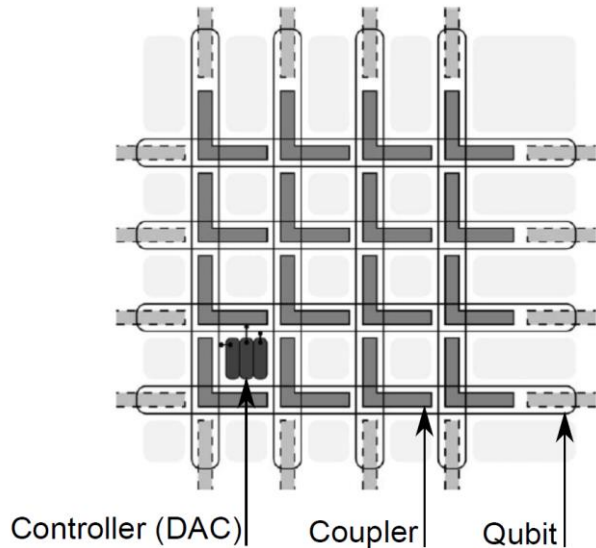
The Structure of Pair-Interactions



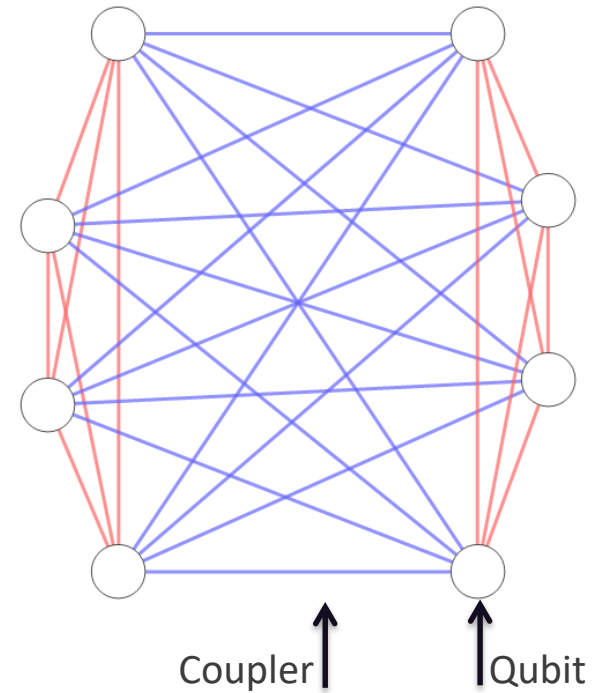
Weak **Spurious** Interactions between Spins

Spurious Interactions & Echo of the Chip Architecture

Chimera Cell Controller



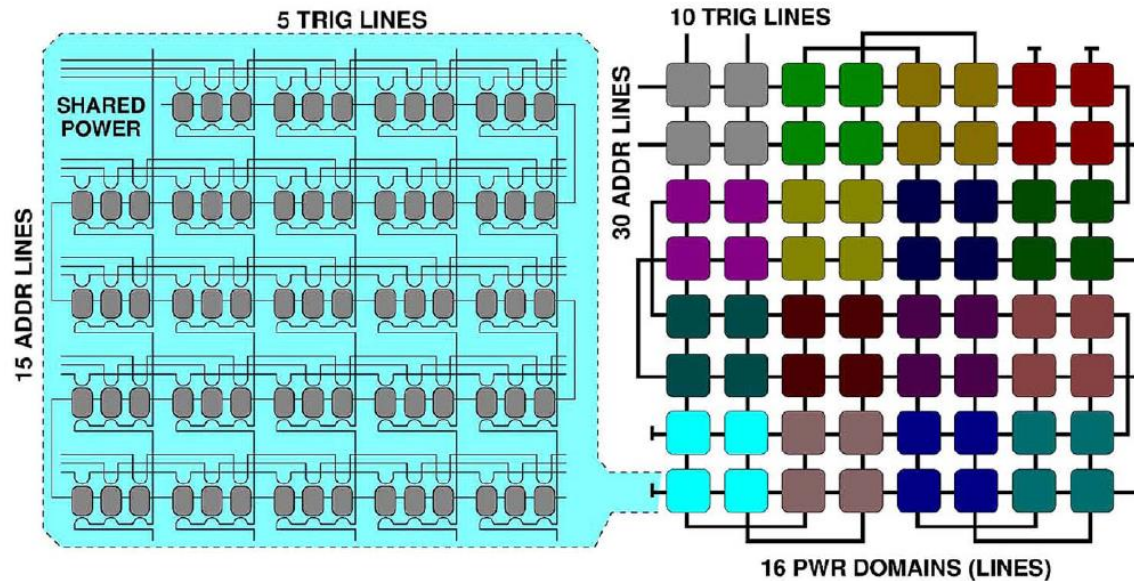
Reconstructed Cell



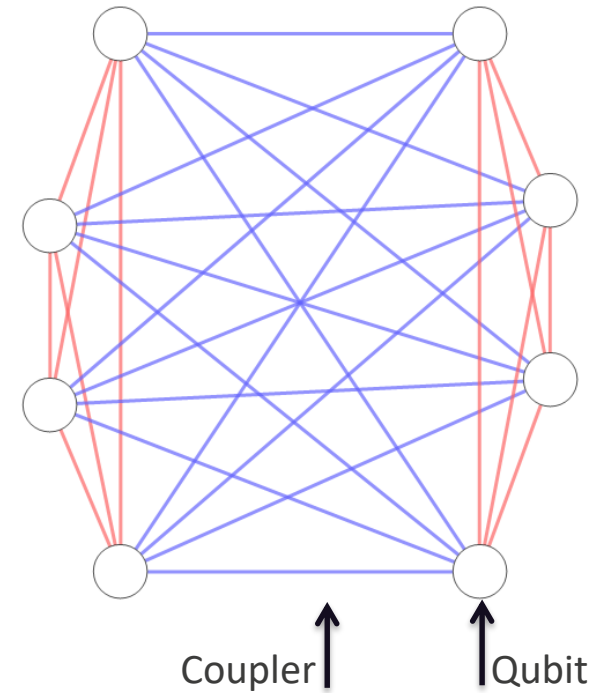
Reference & figure credit: [Bunyik et al., *IEEE Trans. Appl. Supercond.* (2014)]

Spurious Interactions & Echo of the Chip Architecture

Power Lines for Controllers



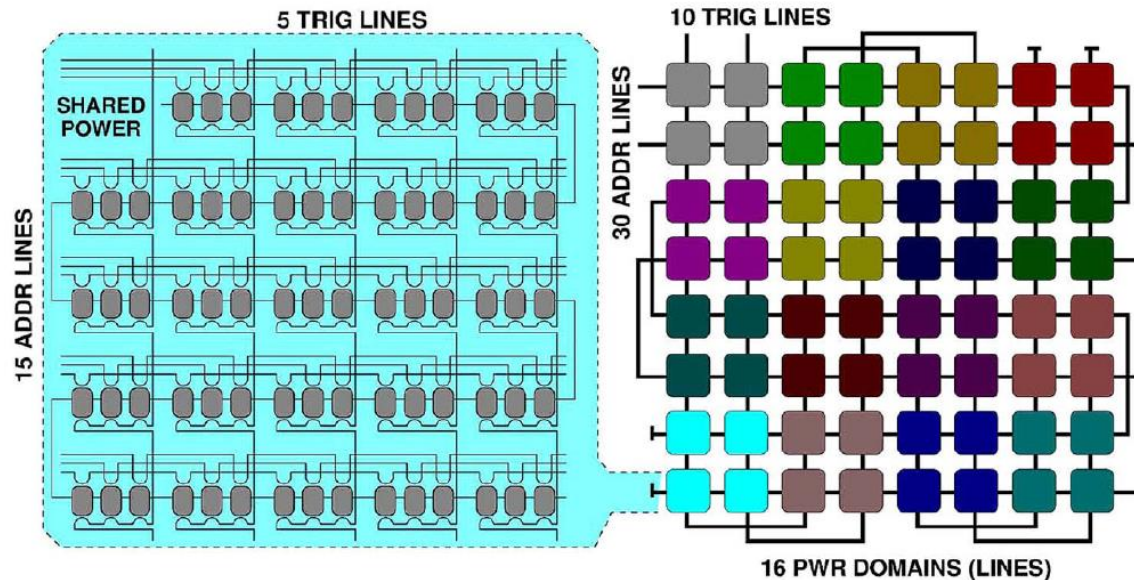
Reconstructed Cell



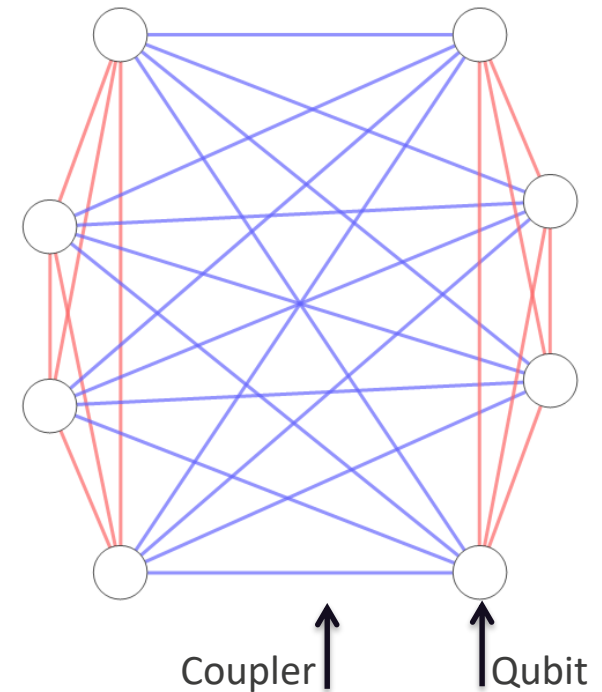
Reference & figure credit: [Bunyik et al., *IEEE Trans. Appl. Supercond.* (2014)]

Spurious Interactions & Echo of the Chip Architecture

Hypothesis: **Spurious** Interactions Reflects **Shared Control Lines**



Reconstructed Cell



Reference & figure credit: [Bunyik et al., *IEEE Trans. Appl. Supercond.* (2014)]

Coupling Input-Output Response

What is the Input-Output Response?



What is the Input-Output Response?



Single Scaling (temperature) response:

$$h_{out} = \beta_{eff} h_{in}$$

$$J_{out} = \beta_{eff} J_{in}$$

Benedetti et al. *Phys. Rev. A*, (2016)

Raymond et al. *Front. In ICT*, (2017)

Marshall et al. *arXiv*, (2017)

What is the Input-Output Response?



Single Scaling (temperature) response:

$$h_{out} = \beta_{eff} h_{in}$$

$$J_{out} = \beta_{eff} J_{in}$$

Benedetti et al. *Phys. Rev. A*, (2016)

Raymond et al. *Front. In ICT*, (2017)

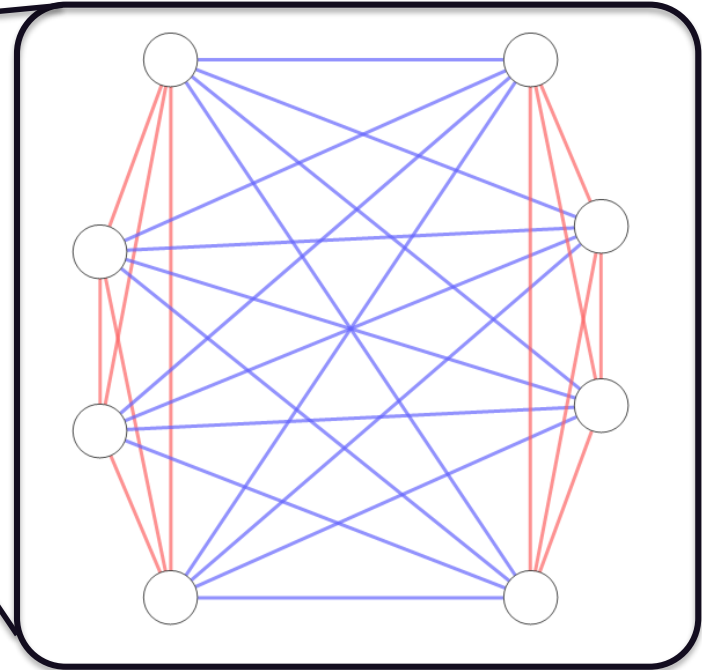
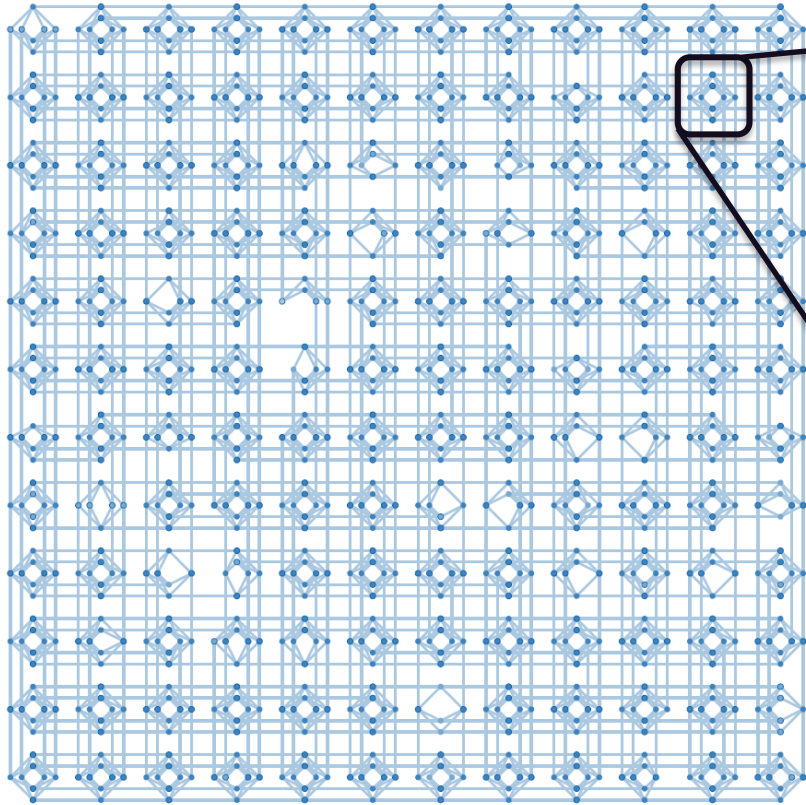
Marshall et al. *arXiv*, (2017)

General Linear response:

$$\begin{bmatrix} h_{out} \\ J_{out} \end{bmatrix} = \begin{bmatrix} \beta_{hh} & \beta_{hJ} \\ \beta_{Jh} & \beta_{JJ} \end{bmatrix} \begin{bmatrix} h_{in} \\ J_{in} \end{bmatrix} + \begin{bmatrix} c_h \\ c_J \end{bmatrix}$$

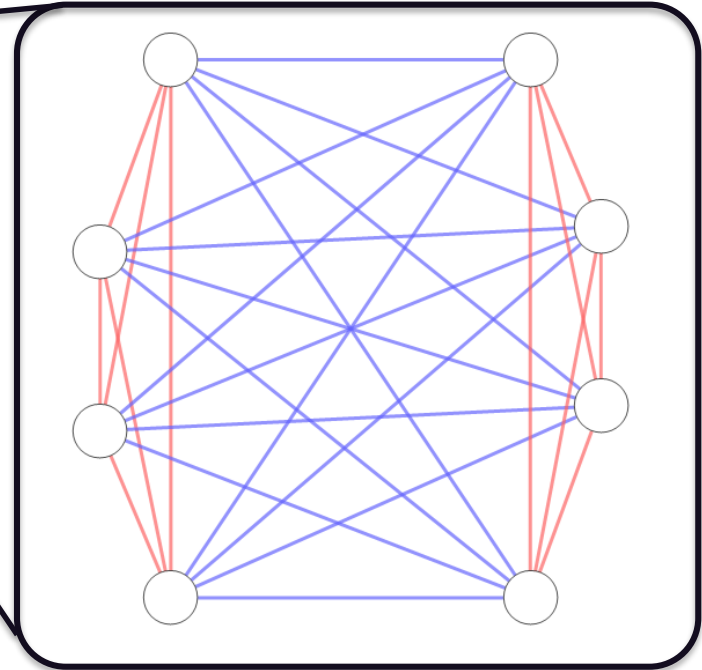
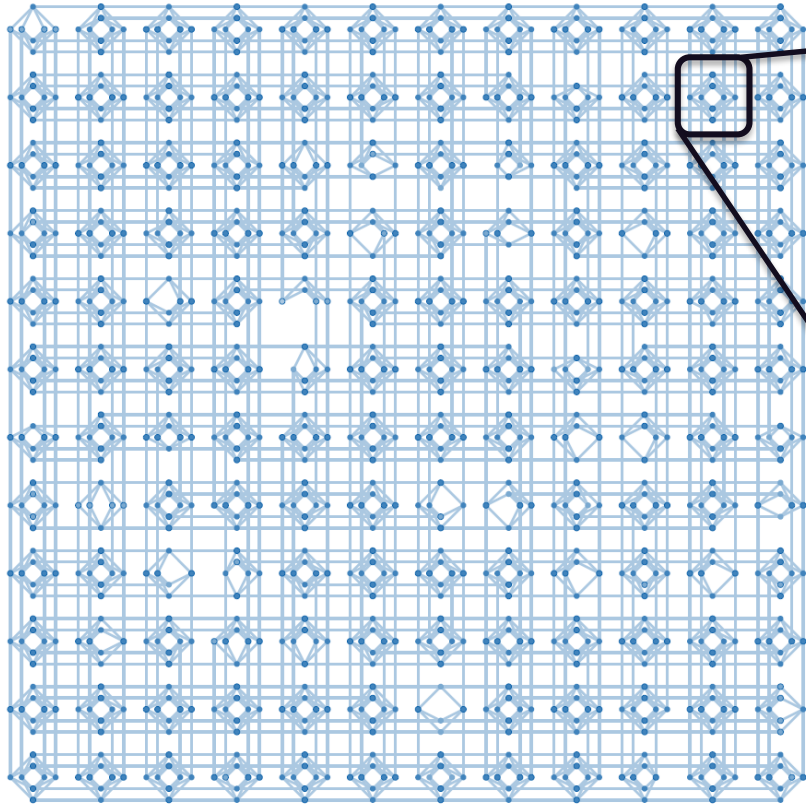
Testing effective temperature hypothesis

Input-Output Response: Results on 1 Chimera Cells



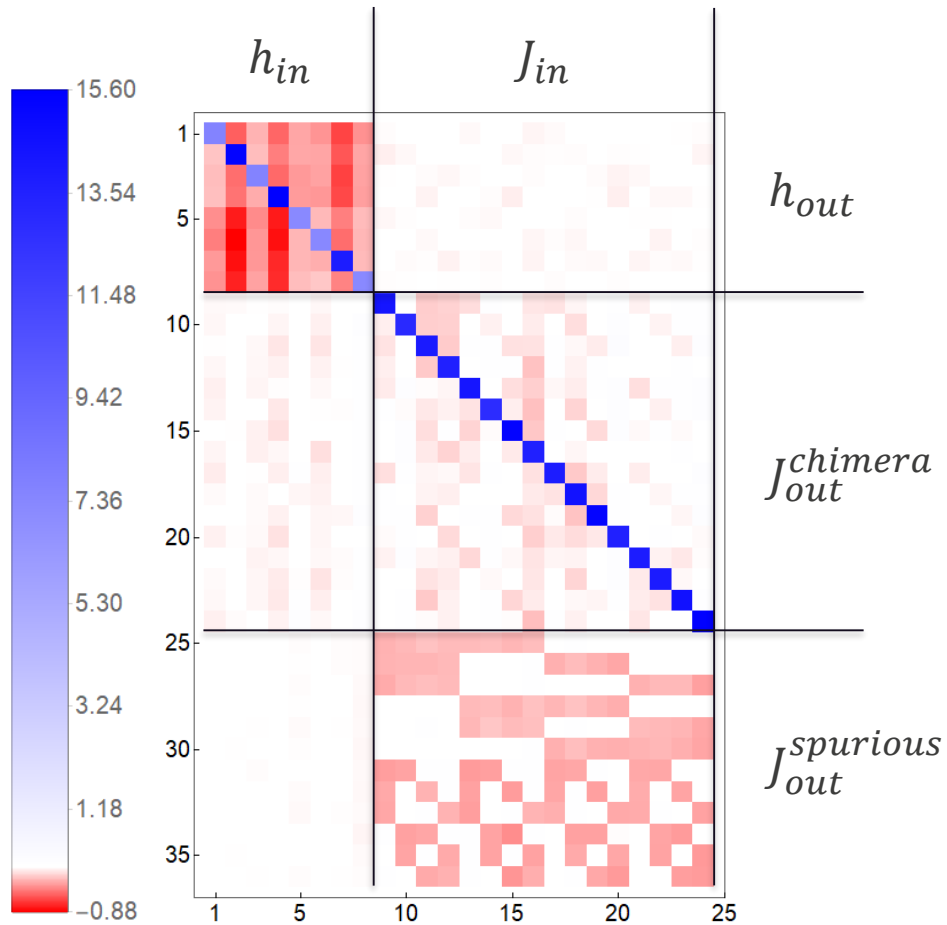
- 16 **Input** Couplers at 0.025
- 8 **Input** Fields at 0.0

Input-Output Response: Results on 1 Chimera Cells

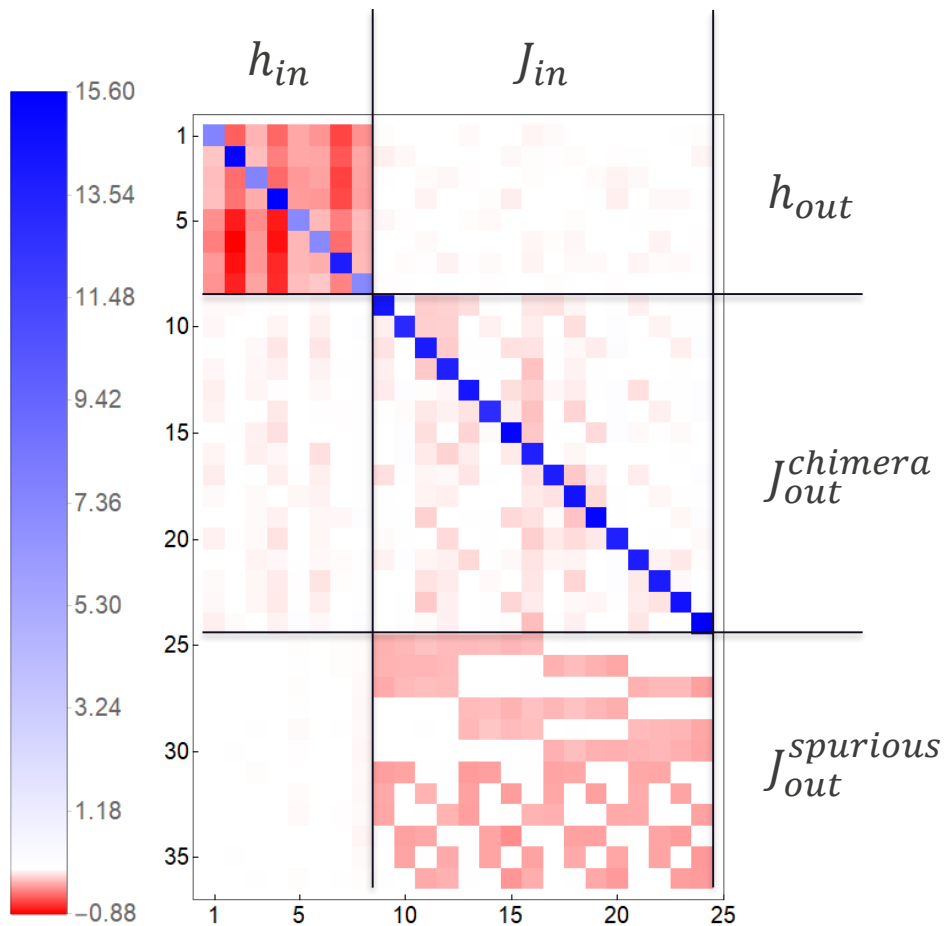


- 16 **Input** Couplers at 0.025
- 8 **Input** Fields at 0.0
- 16 **Output** Chimera Couplings
- 8 **Output** Spurious Couplings
- 8 **Output** Fields

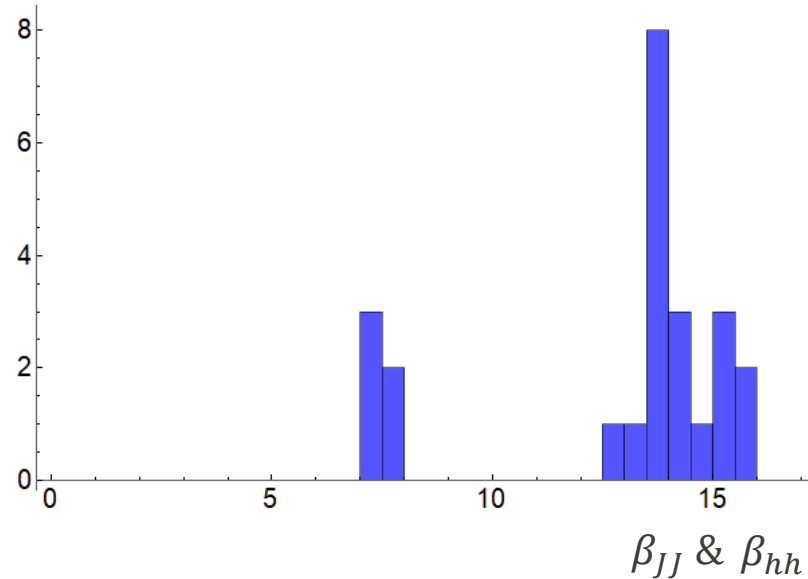
Input-Output Response Matrix



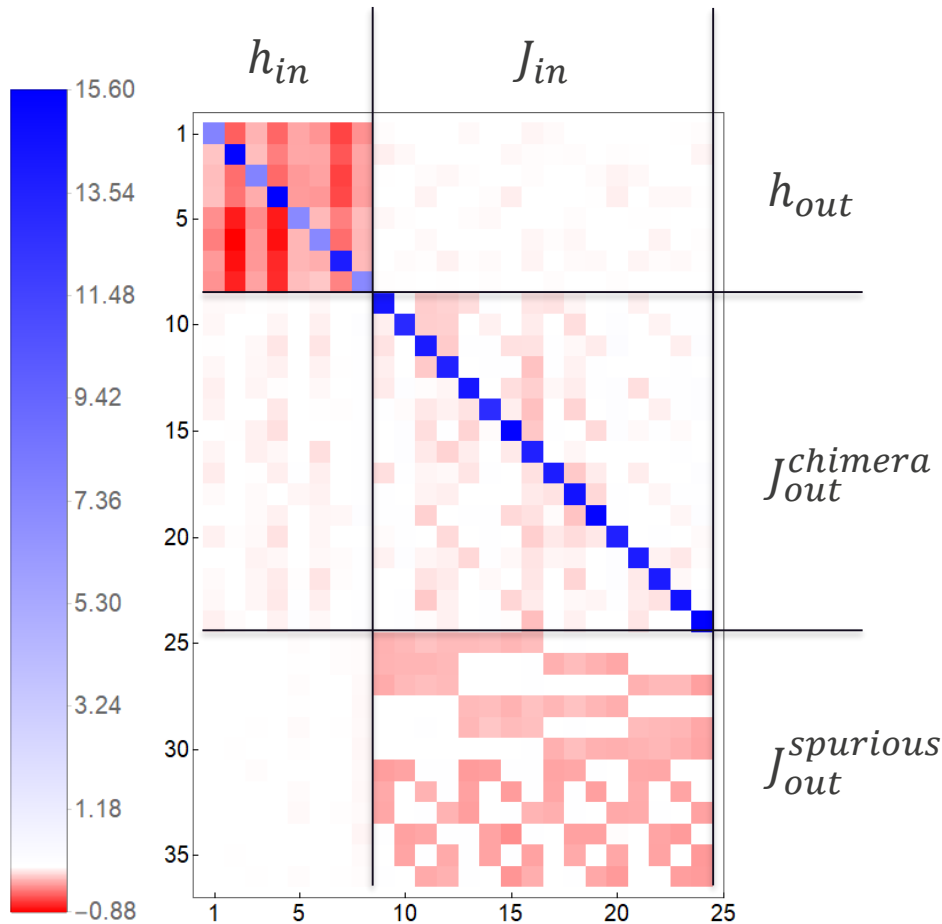
Input-Output Response Matrix



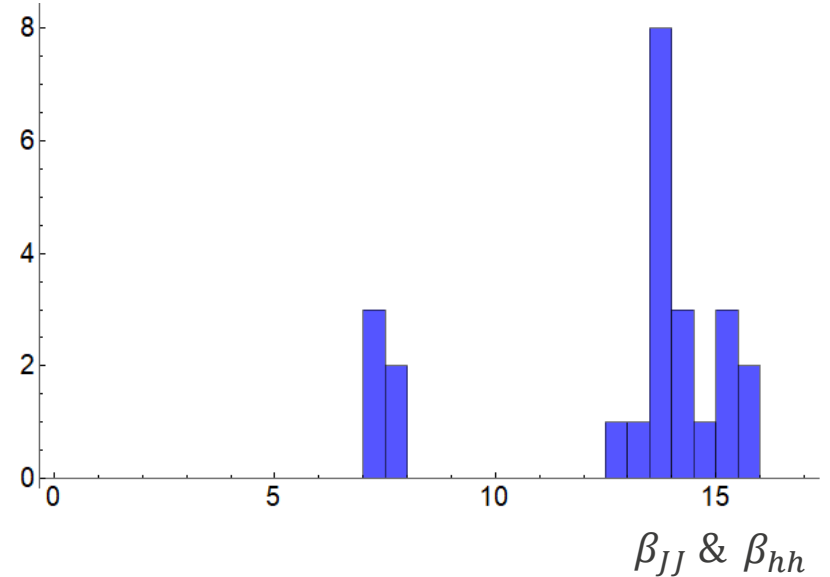
Diagonal Input-Output Response



Input-Output Response Matrix

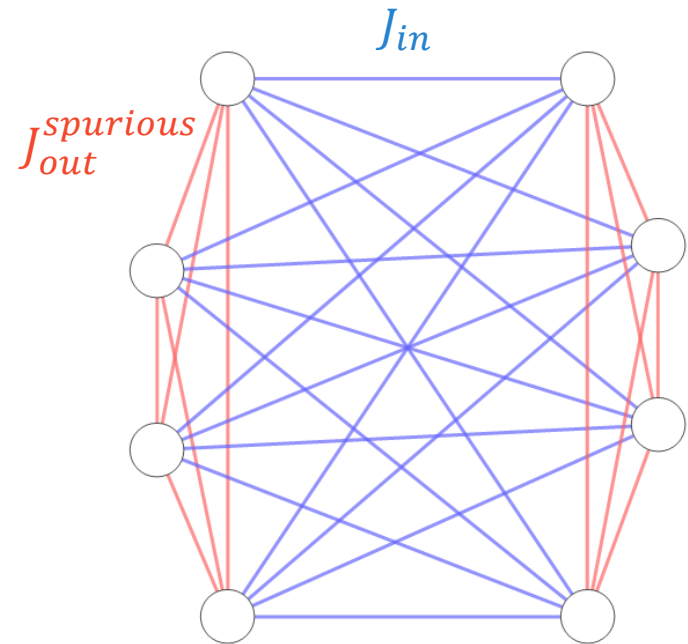
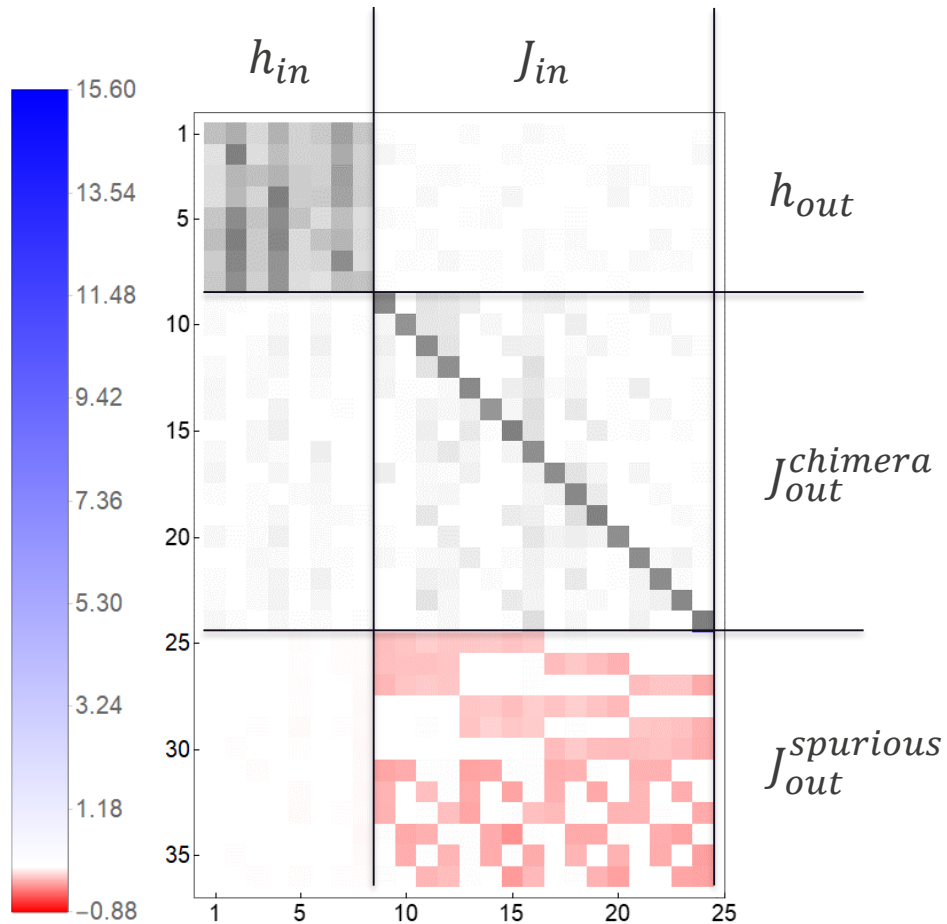


Diagonal Input-Output Response



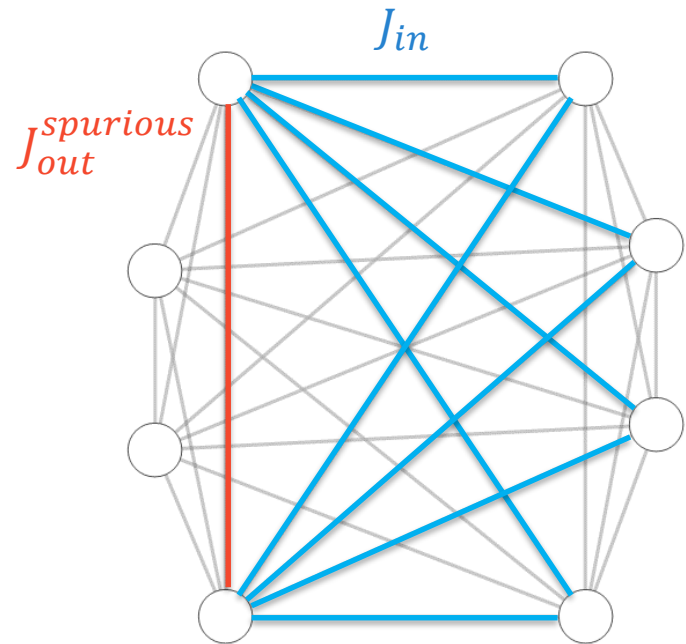
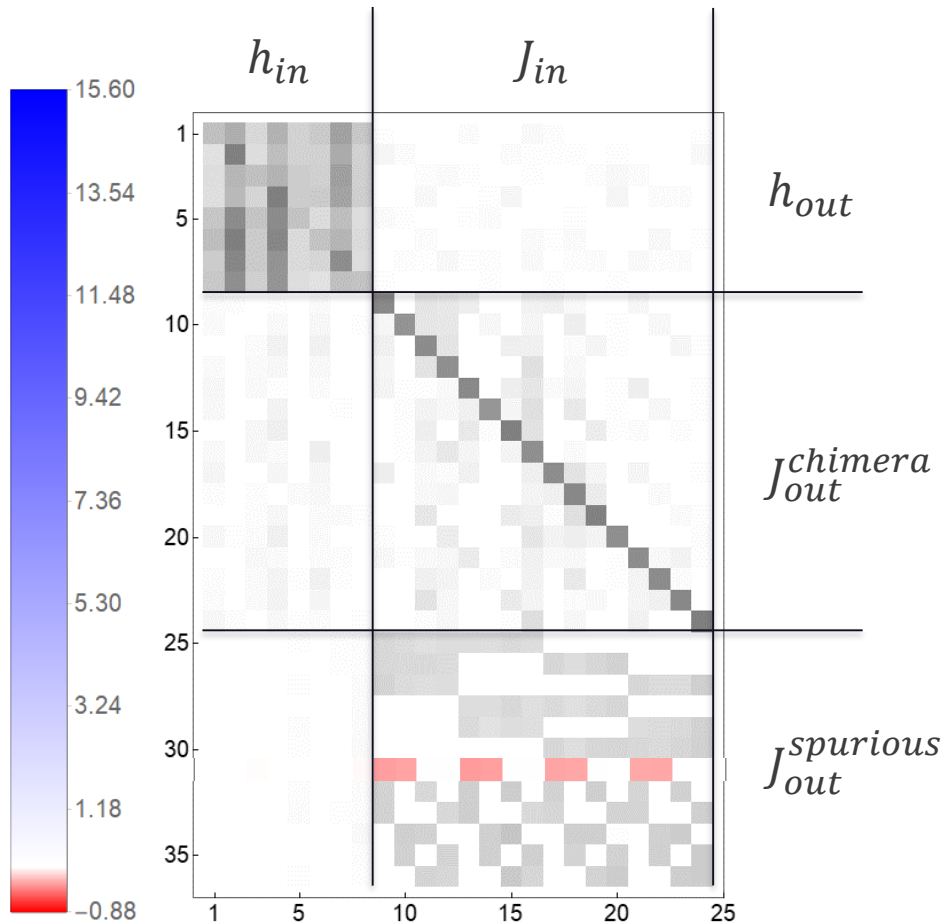
Different Scaling
↓
No single effective temperature

Input-Output Response Matrix



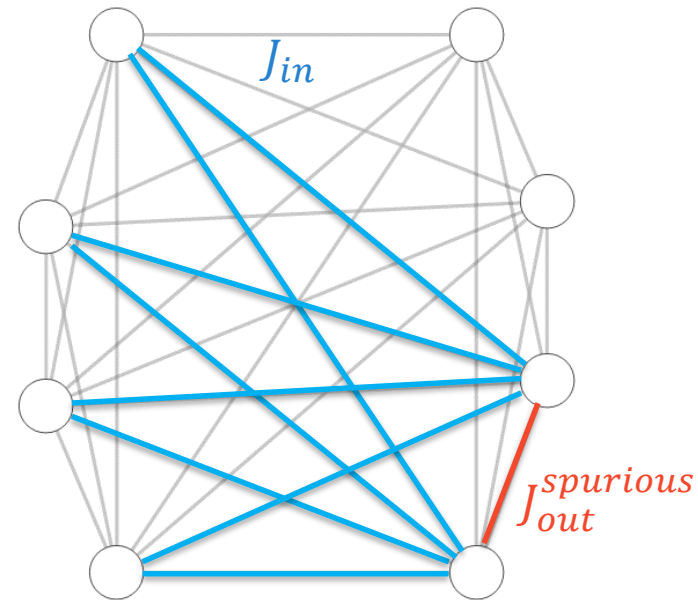
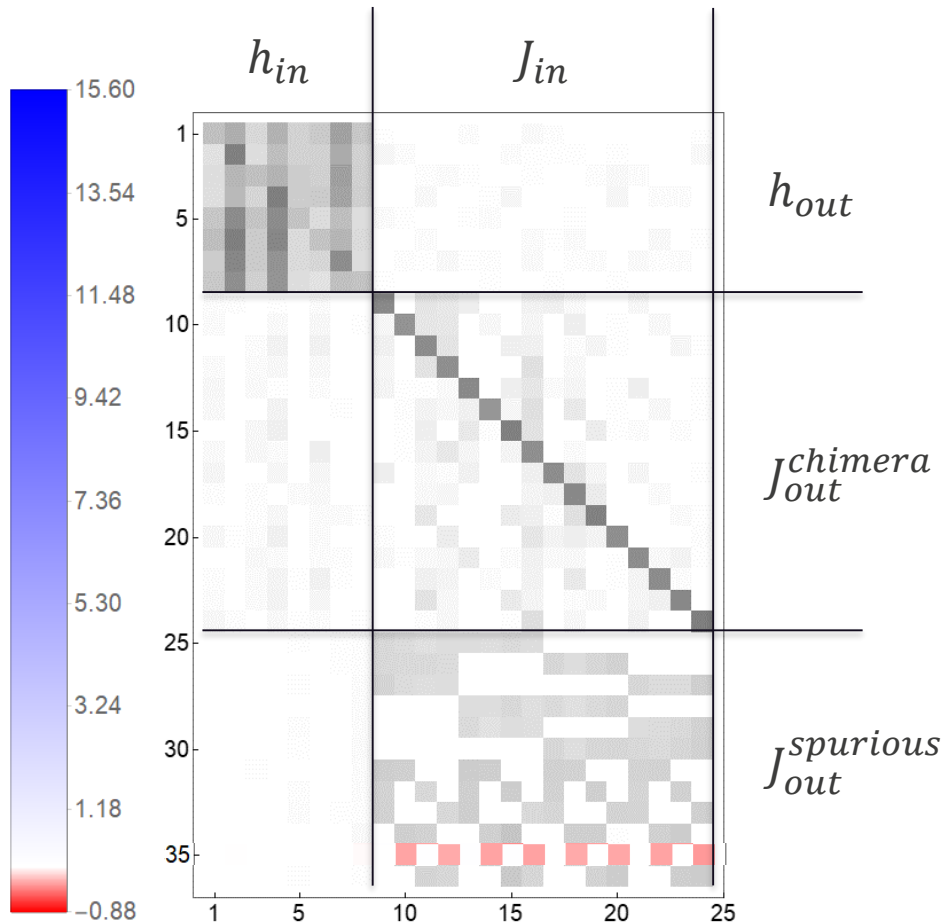
Spurious couplings response from input couplers

Input-Output Response Matrix



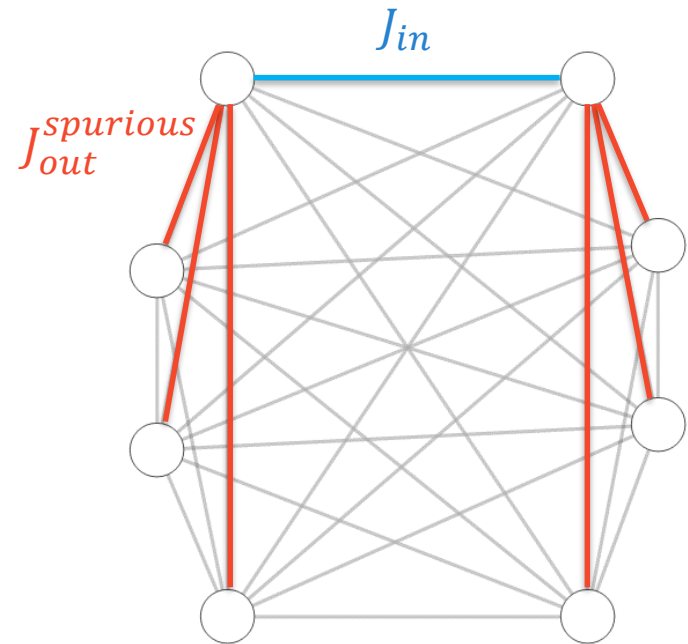
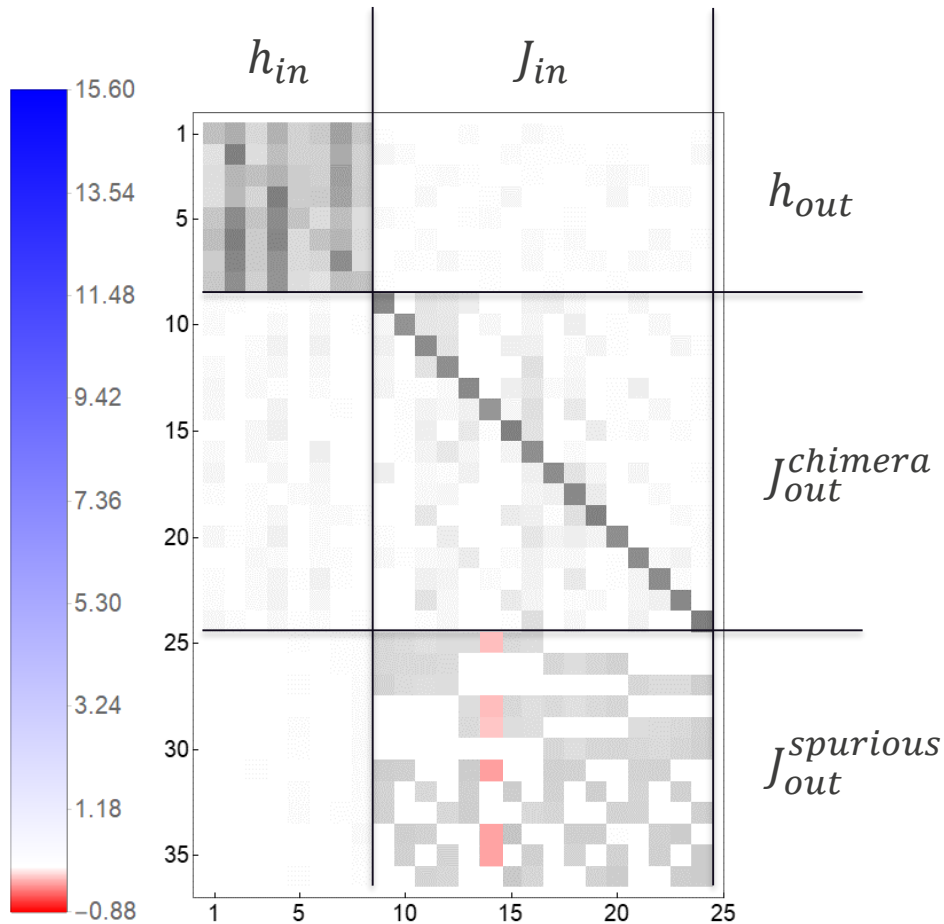
Spurious couplings only respond to neighboring input couplers

Input-Output Response Matrix



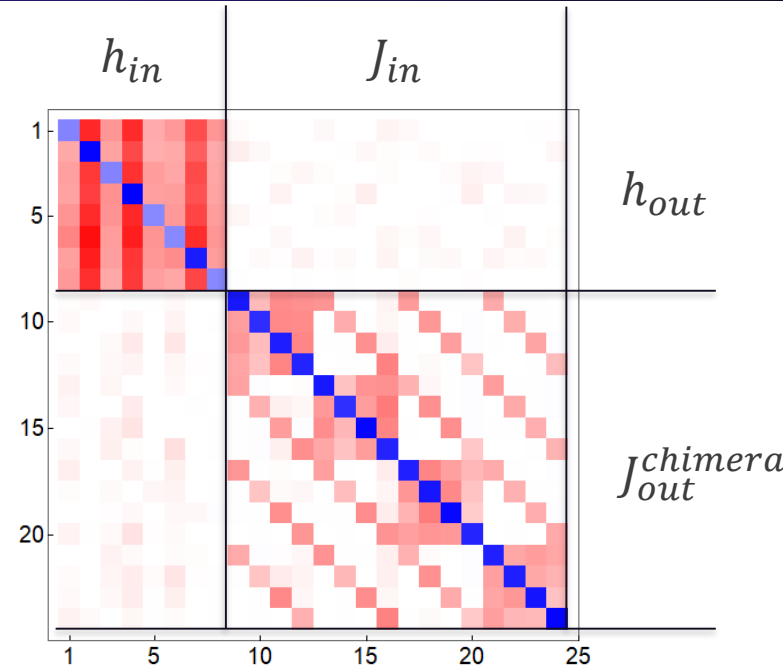
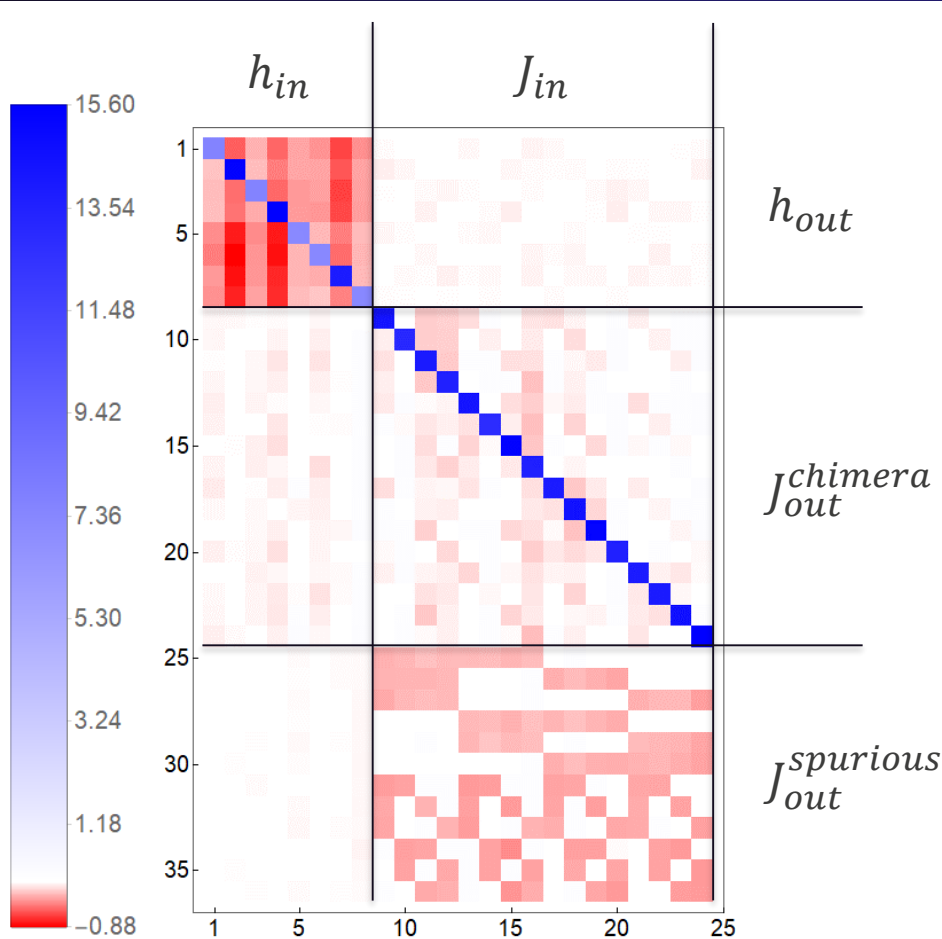
Spurious couplings only respond to neighboring input couplers

Input-Output Response Matrix



Input couplers induces response on neighboring spurious links

Input-Output Response Matrix



Best response matrix ignoring spurious links

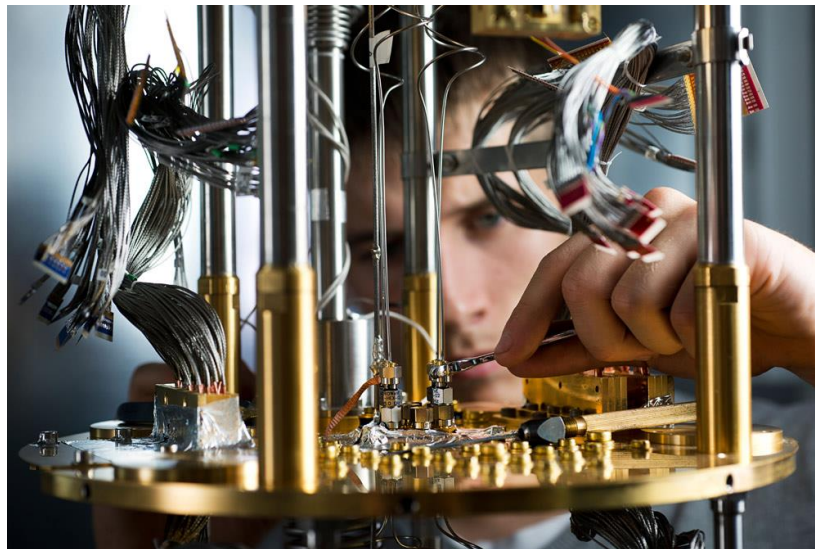
Conclusions

Summary

- ✓ D-Wave **samples** are well described by **pairwise Ising Models**
- ✓ **Multi-body** interactions **exists** but are **small**
- ✓ **Pairwise** structure is **approximately chimera** but with **spurious interactions**
- ✓ **Output-Input** response is mainly **diagonal**
- ✓ **No Single** effective **temperature**

Path forward

- Deeper study of **deviations** from **pairwise** Ising models → Enhance chip **architecture**
- Use **Input-Output** response for **calibration** → Enhance **optimization**/sampling



Acknowledgements



advanced network
science initiative
(ansi)