Los Alamos
NATIONAL LABORATORY
EST. 1943

SOLVING SPARSE REPRESENTATIONS FOR OBJECT CLASSIFICATION USING QUANTUM D-WAVE 2X MACHINE

Nga Nguyen, Amy Larson, Carleton Coffrin, John Perry, Gary Salazar, *and* Garrett Kenyon

Los Alamos National Laboratory & New Mexico Consortium

*5th **N**euro **I**nspired **C**omputational **E**lements **2017**, San Jose, Mar. 6-8, 2017*

OUTLINE

Quantum D-Wave machine 2X: a quantum annealer

A. METHODOLOGY

Los Alamos NATIONAL LABORATORY EST. 1943

- **Solving a sparse-coding (SC) problem**

Objective function is of the form:

$$E = \min_{\{\vec{a}, \phi\}} \left[\frac{1}{2} |\vec{I} - \phi \vec{a}|^2 + \lambda ||\vec{a}||_p \right].$$

reconstruction error **Lp-sparseness penalty**

Olshausen and Field, Nature 381, 607 (1996)
Rozell, Johnson, Baraniuk, and Olshausen, Neur. Comp. 20, 2526 (2008)

Quantum D-Wave machine 2X: a quantum annealer

A. METHODOLOGY

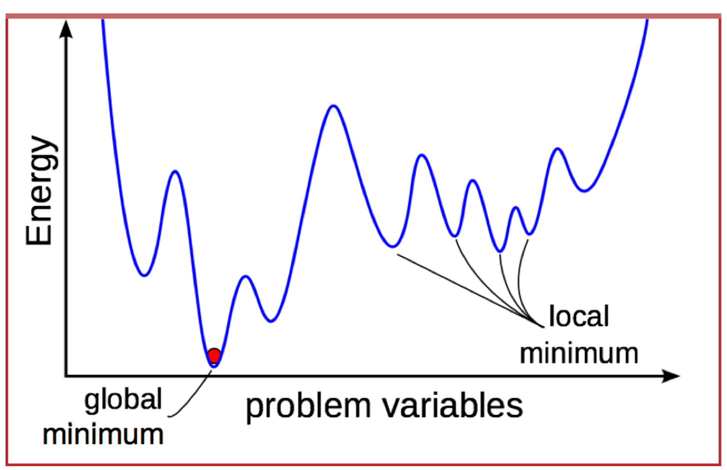
- **Solving a sparse-coding (SC) problem**

Objective function is of the form:

$$E = \min_{\{\vec{a},\phi\}} [\frac{1}{2}|\vec{I} - \phi\vec{a}|^2 + \lambda||\vec{a}||_p].$$

reconstruction error *Lp*-sparseness penalty

Olshausen and Field, Nature 381, 607 (1996)
Rozell, Johnson, Baraniuk, and Olshausen, Neur. Comp. 20, 2526 (2008)



- **non-convex problem**
- **NP-hard class**

Quantum D-Wave machine 2X: a quantum annealer

A. METHODOLOGY

- **Solving a sparse-coding (SC) problem**

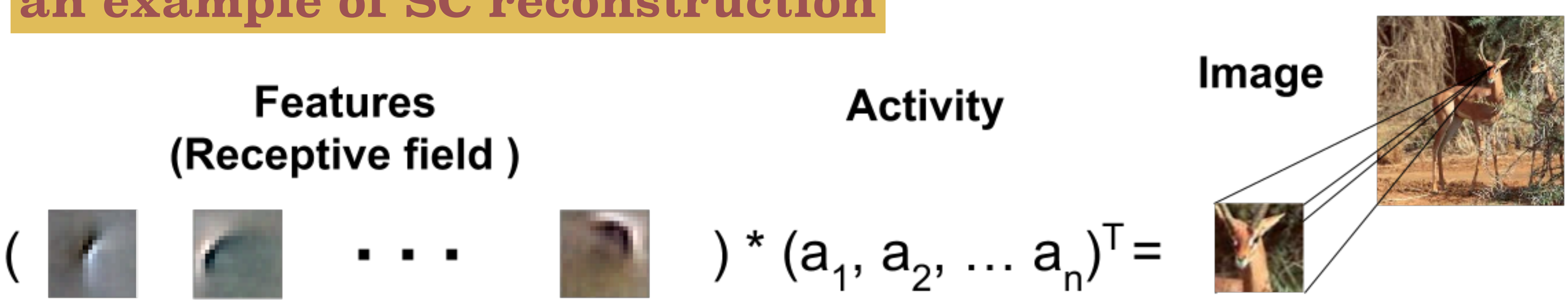
Objective function is of the form:

$$E = \min_{\{\vec{a},\phi\}} \left[\frac{1}{2} |\vec{I} - \phi \vec{a}|^2 + \lambda ||\vec{a}||_p \right].$$

reconstruction error

***Lp*-sparseness penalty**

an example of SC reconstruction



courtesy of Xinhua Zhang

Quantum D-Wave machine 2X: a quantum annealer

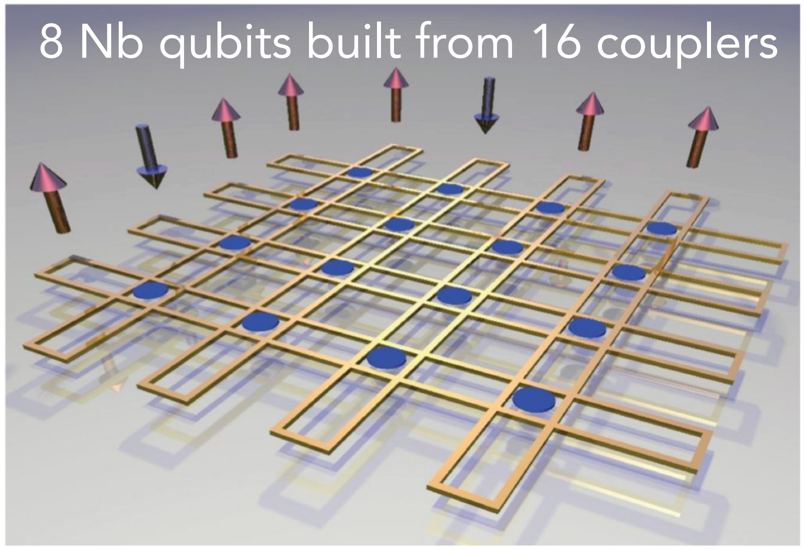
A. METHODOLOGY

- mapping the sparse-coding problem onto a quantum unconstrained binary optimization (QUBO):

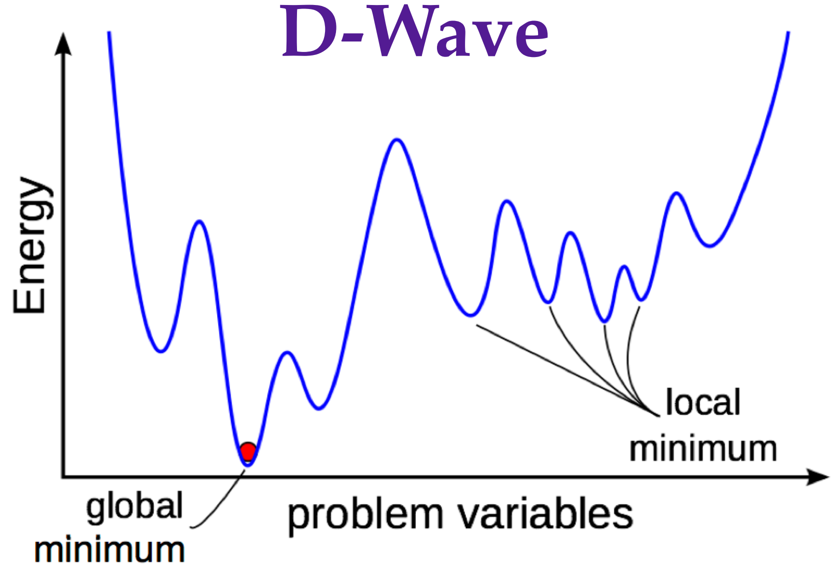
D-Wave Hamiltonian:

$$H(h,Q,a) = \sum_i h_i a_i + \sum_{<i,j>} Q_{ij} a_i a_j$$

where $a_i = \{0,1\} \forall i.$



D-Wave



Quantum D-Wave machine 2X: a quantum annealer

Los Alamos NATIONAL LABORATORY EST. 1943

A. METHODOLOGY

- mapping the sparse-coding problem onto a quantum unconstrained binary optimization (QUBO):

D-Wave Hamiltonian:

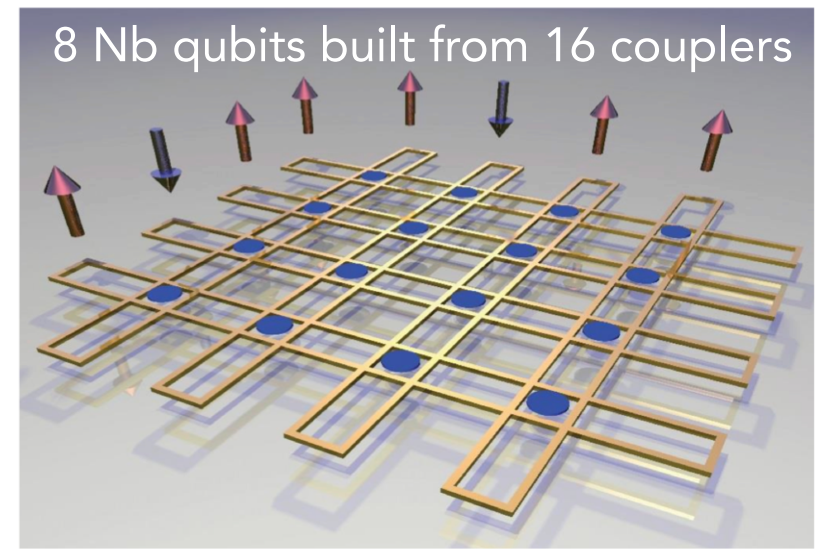
$$H(h,Q,a) = \sum_i h_i a_i + \sum_{<i,j>} Q_{ij} a_i a_j$$

where $a_i = \{0,1\} \forall i.$

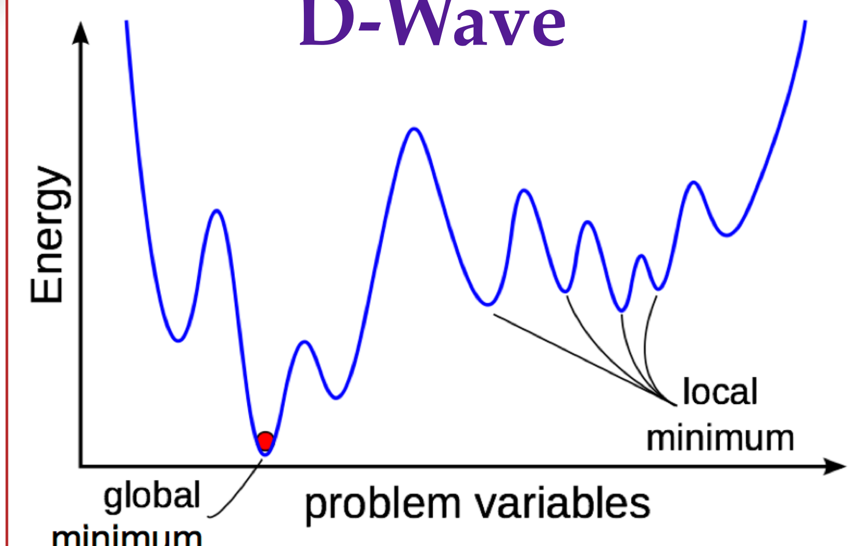
This mapping is achieved by the relations:

$$h = -\phi^T \overrightarrow{I} + (\lambda + \frac{1}{2}),$$

$$Q = \frac{1}{2}\phi^T \phi.$$



D-Wave



- **analogous to L0-sparseness penalty** [Nguyen and Kenyon, PMES-16 (2016)]

A. METHODOLOGY

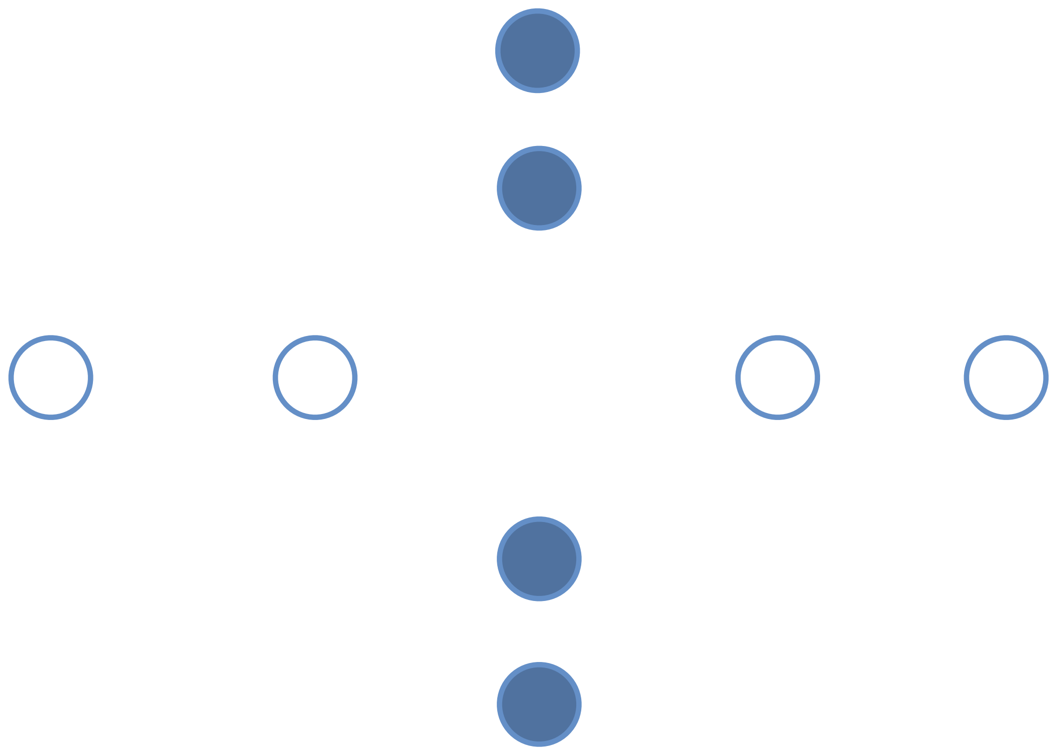
4 “row” qubits

A. METHODOLOGY

4 "column" qubits

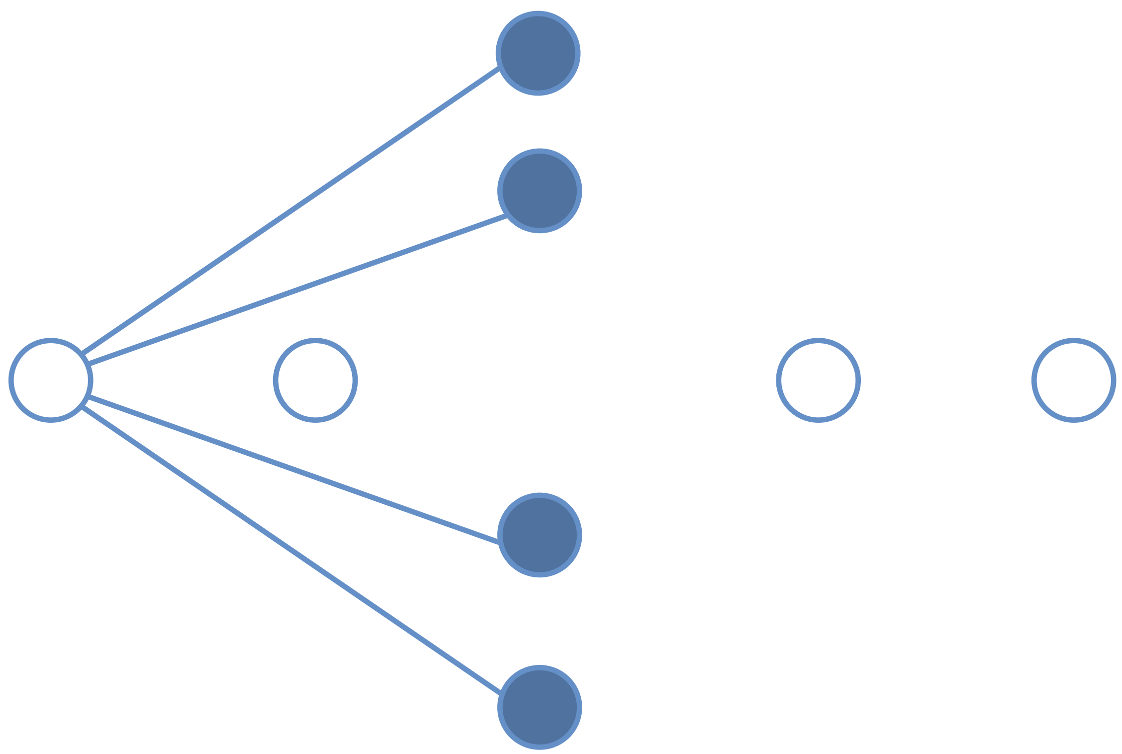
A. METHODOLOGY

unit cell



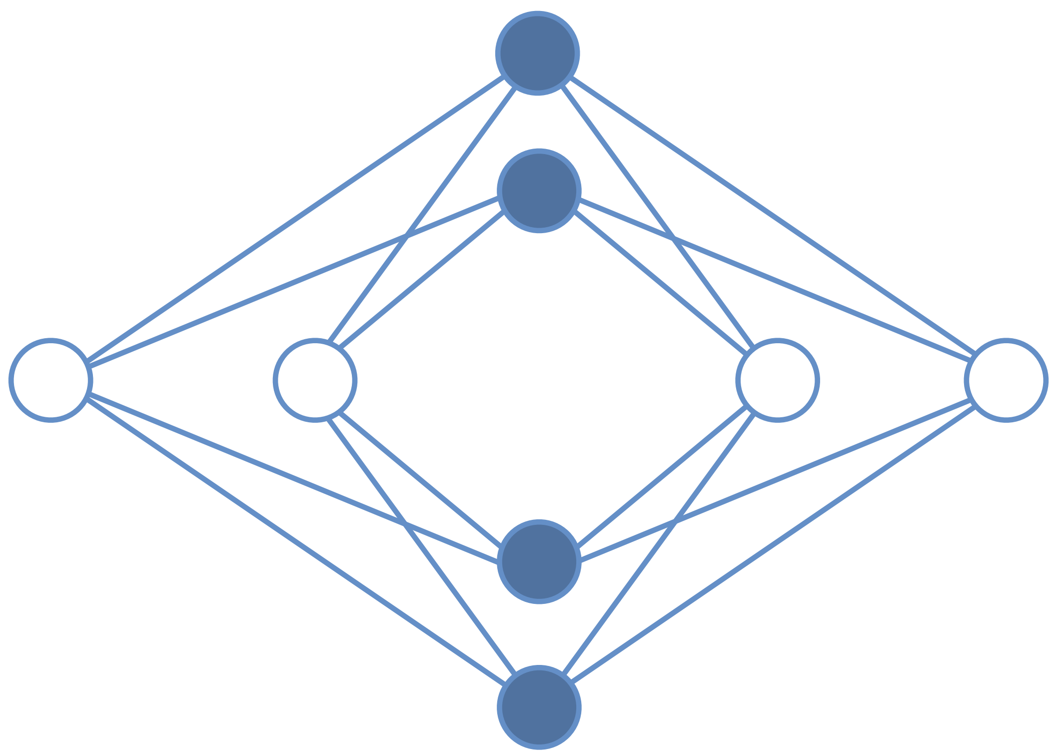
A. METHODOLOGY

intra-cell couplings



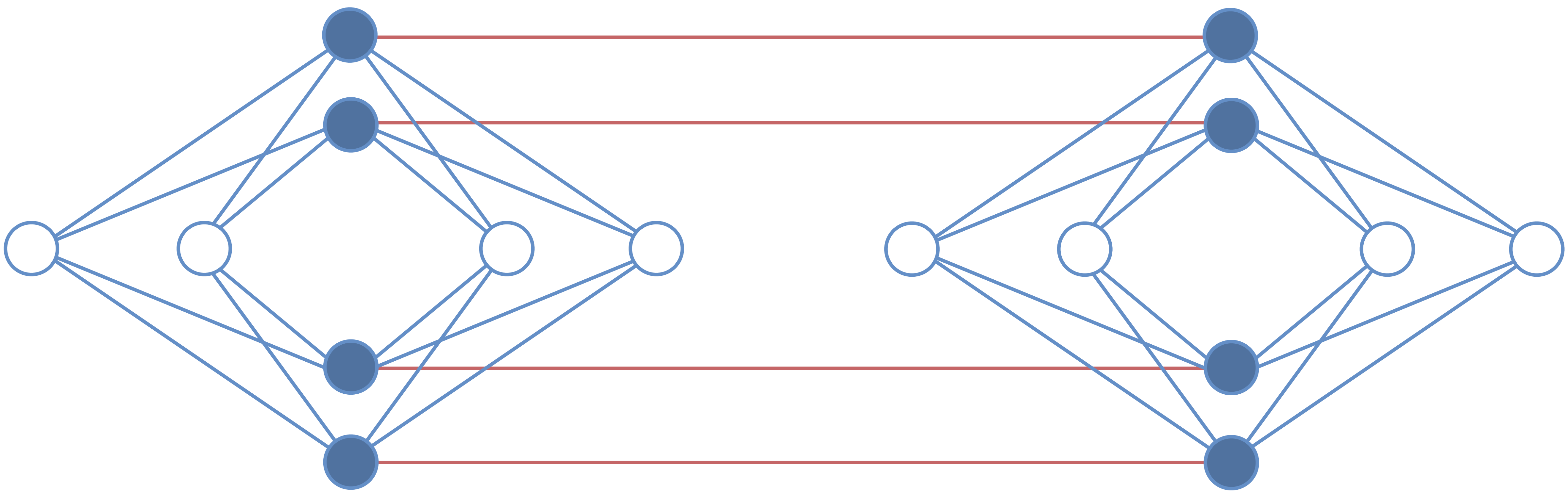
A. METHODOLOGY

intra-cell couplings



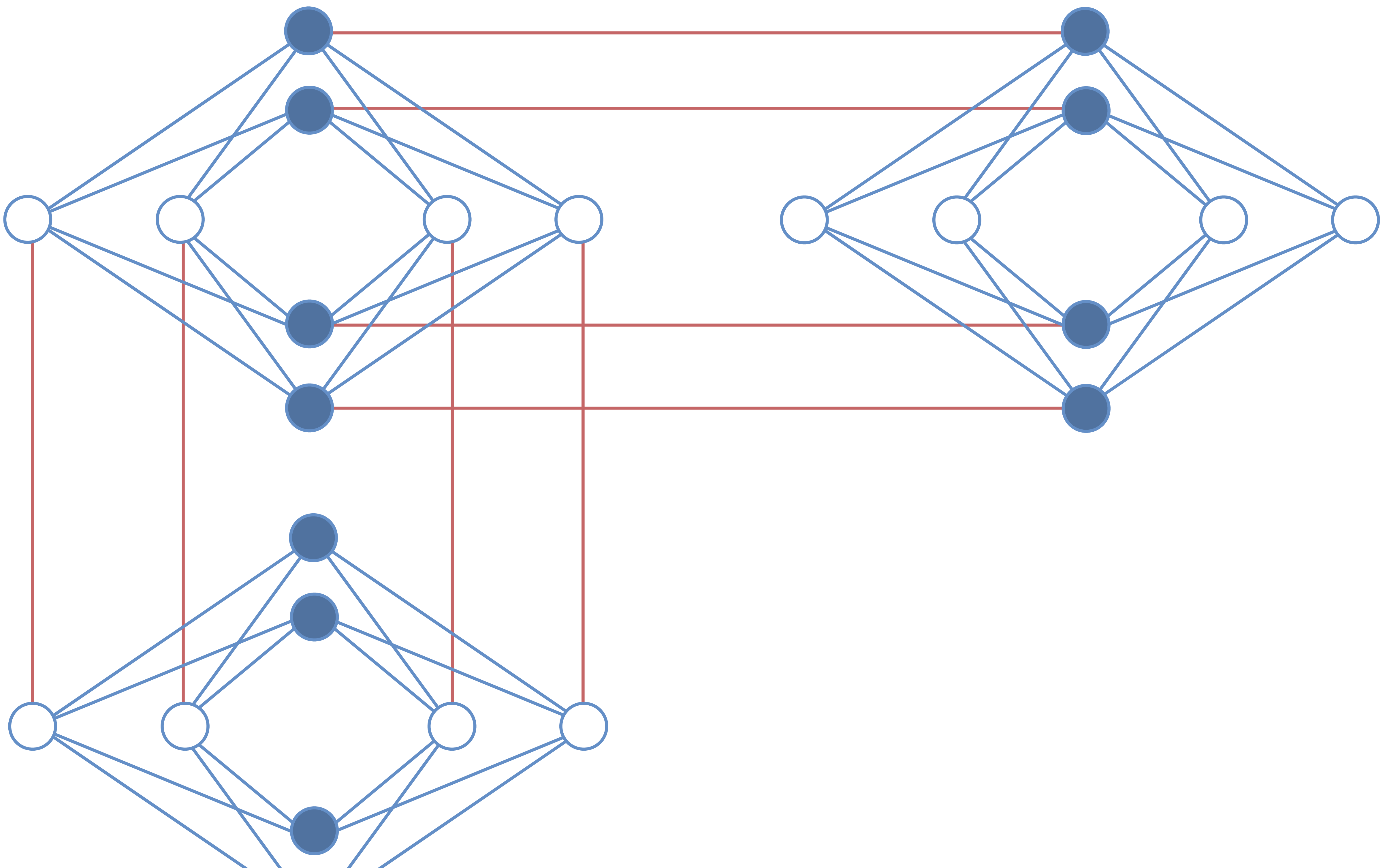
A. METHODOLOGY

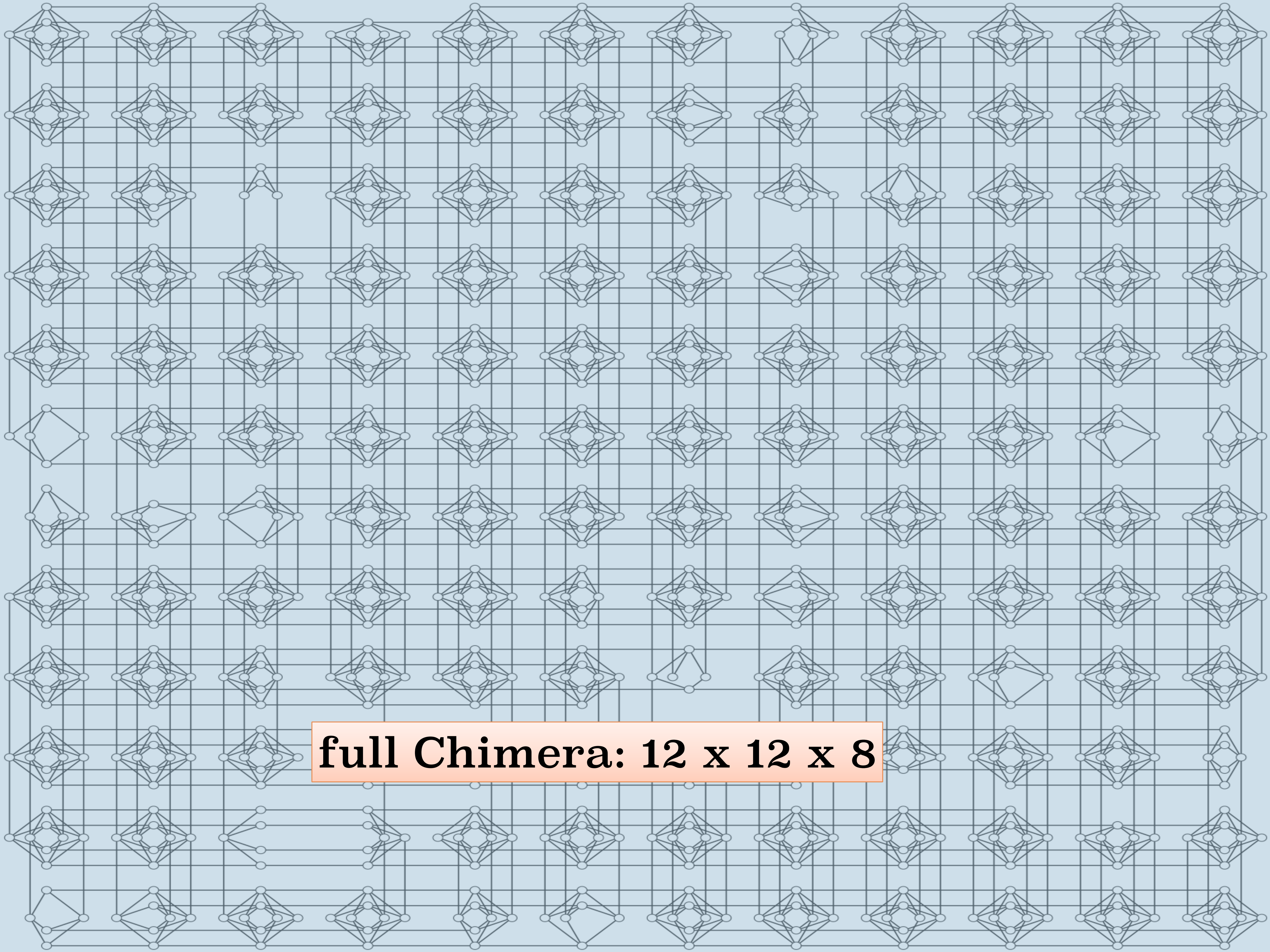
neighboring couplings



A. METHODOLOGY

neighboring couplings



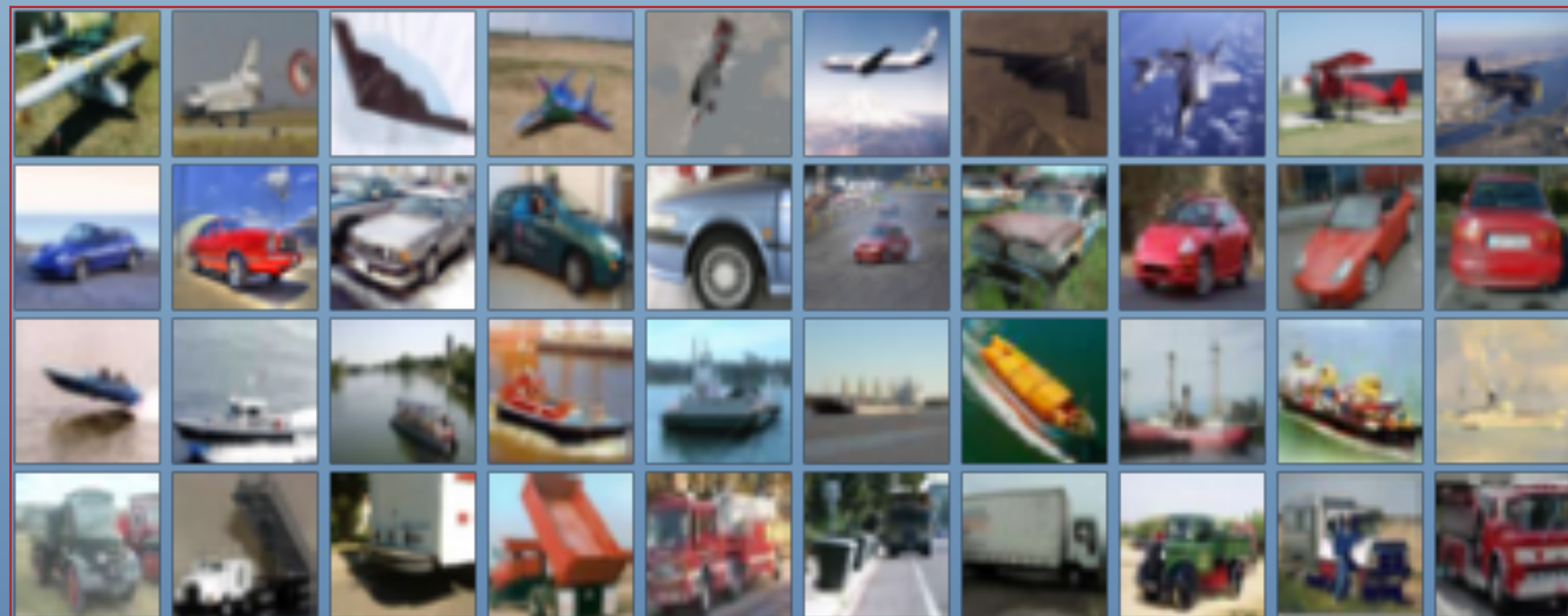

full Chimera: 12 x 12 x 8

B. DATASET

CIFAR-10

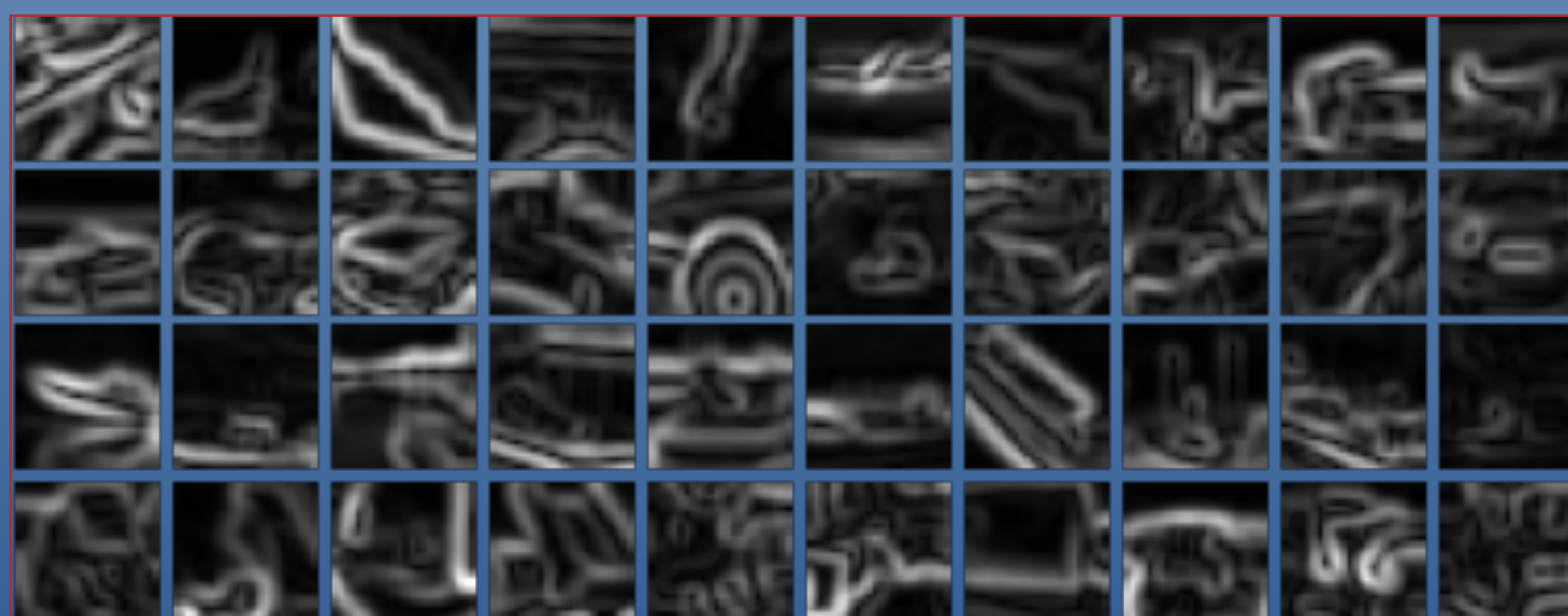
32x32

airplane
automobile
ship
truck



24x24

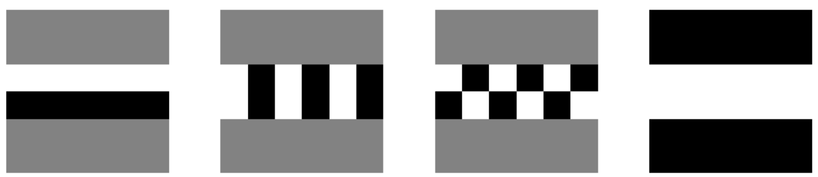
edge detection



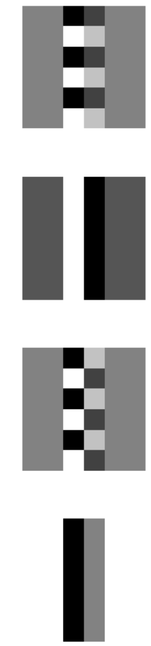
C. IMPLEMENTATION ON D-WAVE MACHINE

8 hand-designed features

"row"



"column"



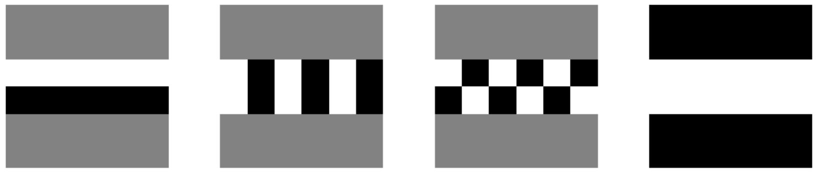
orthogonality!

number of features $N_f = 8$

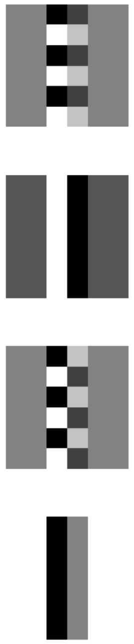
C. IMPLEMENTATION ON D-WAVE MACHINE

8 hand-designed features

“row”



“column”



orthogonality!

number of features $N_f = 8$

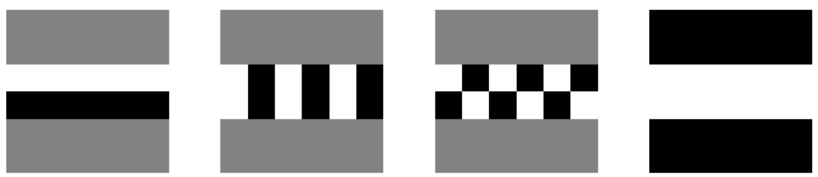
$\{\psi_i\}$

complete set: $\Phi = \Psi * \overrightarrow{I}$

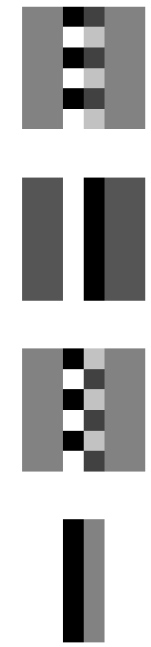
C. IMPLEMENTATION ON D-WAVE MACHINE

8 hand-designed features

"row"



"column"

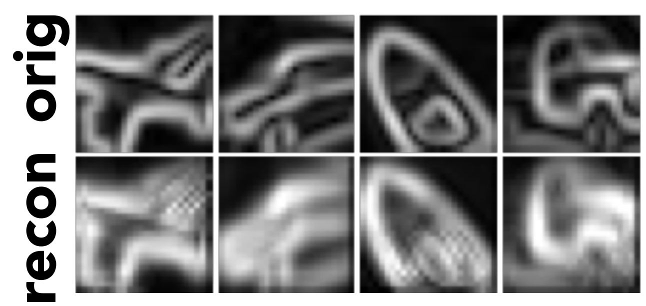


orthogonality!

number of features $N_f = 8$

$\{\psi_i\}$

complete set: $\Phi = \Psi * \vec{I}$

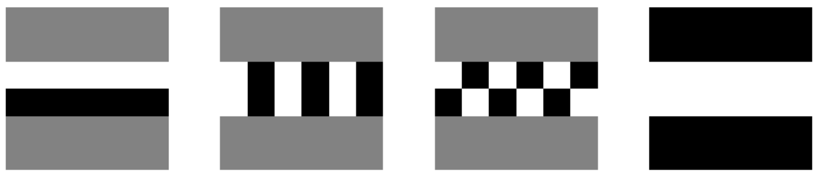


C. IMPLEMENTATION ON D-WAVE MACHINE

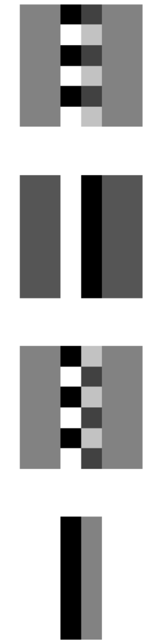
Los Alamos
NATIONAL LABORATORY
EST. 1943

8 hand-designed features

"row"



"column"



orthogonality!

number of features $N_f = 8$

$\{\psi_i\}$

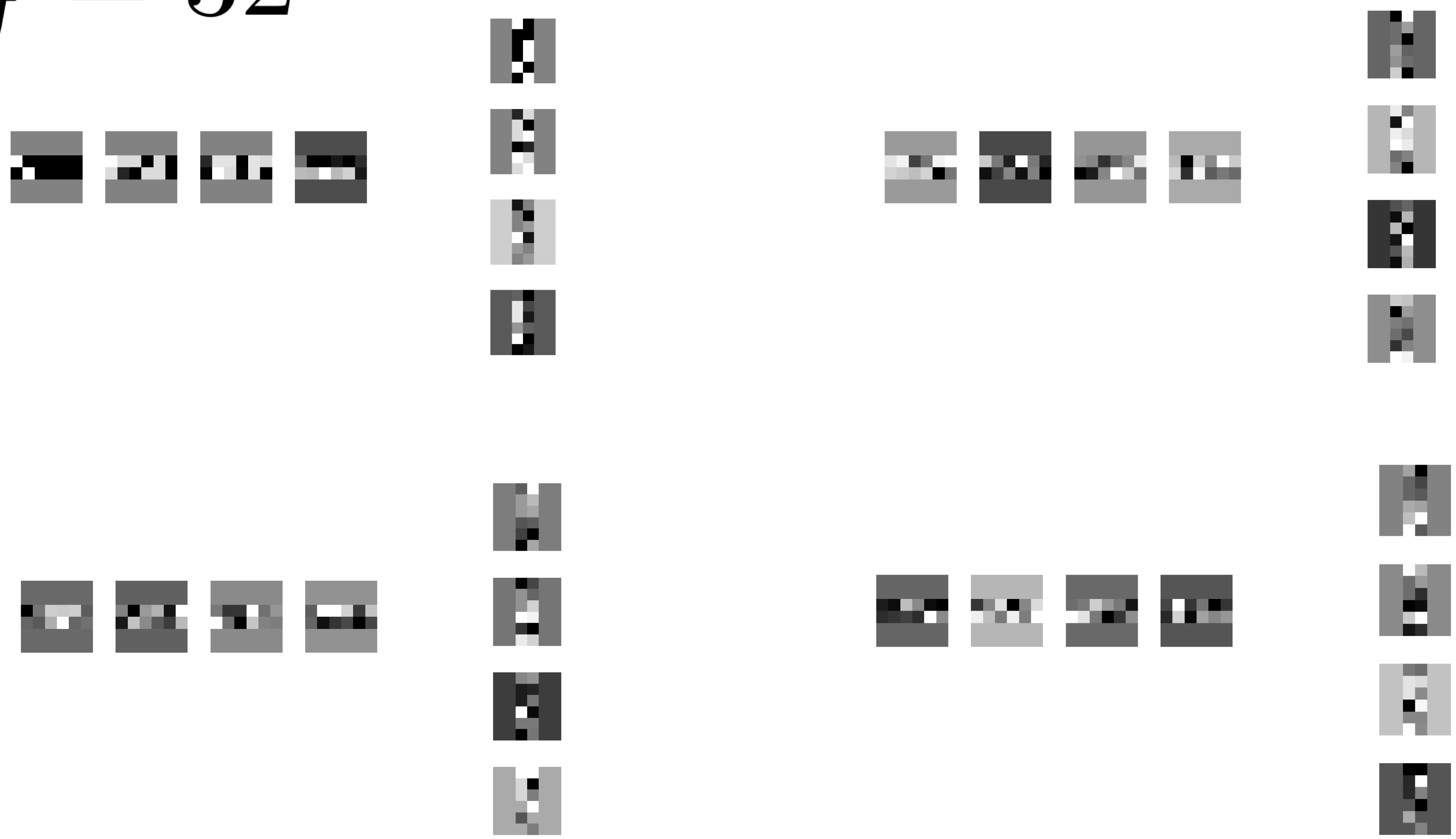
complete set: $\Phi = \Psi * \overrightarrow{I}$

Desire: Randomly generated N_f **:**

$$8 < N_f < 1152$$

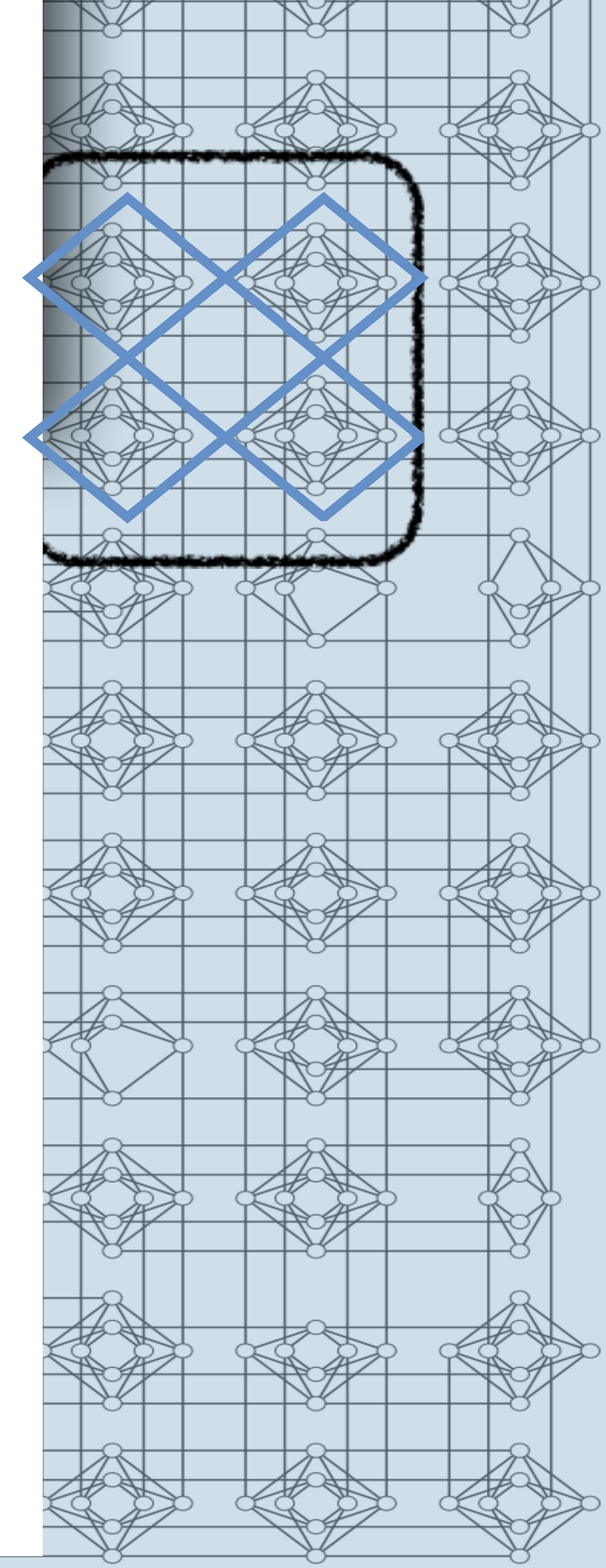
Building features

$N_f = 32$



$\cdots$

$N_f = 1152$



- **24x24 patch images**

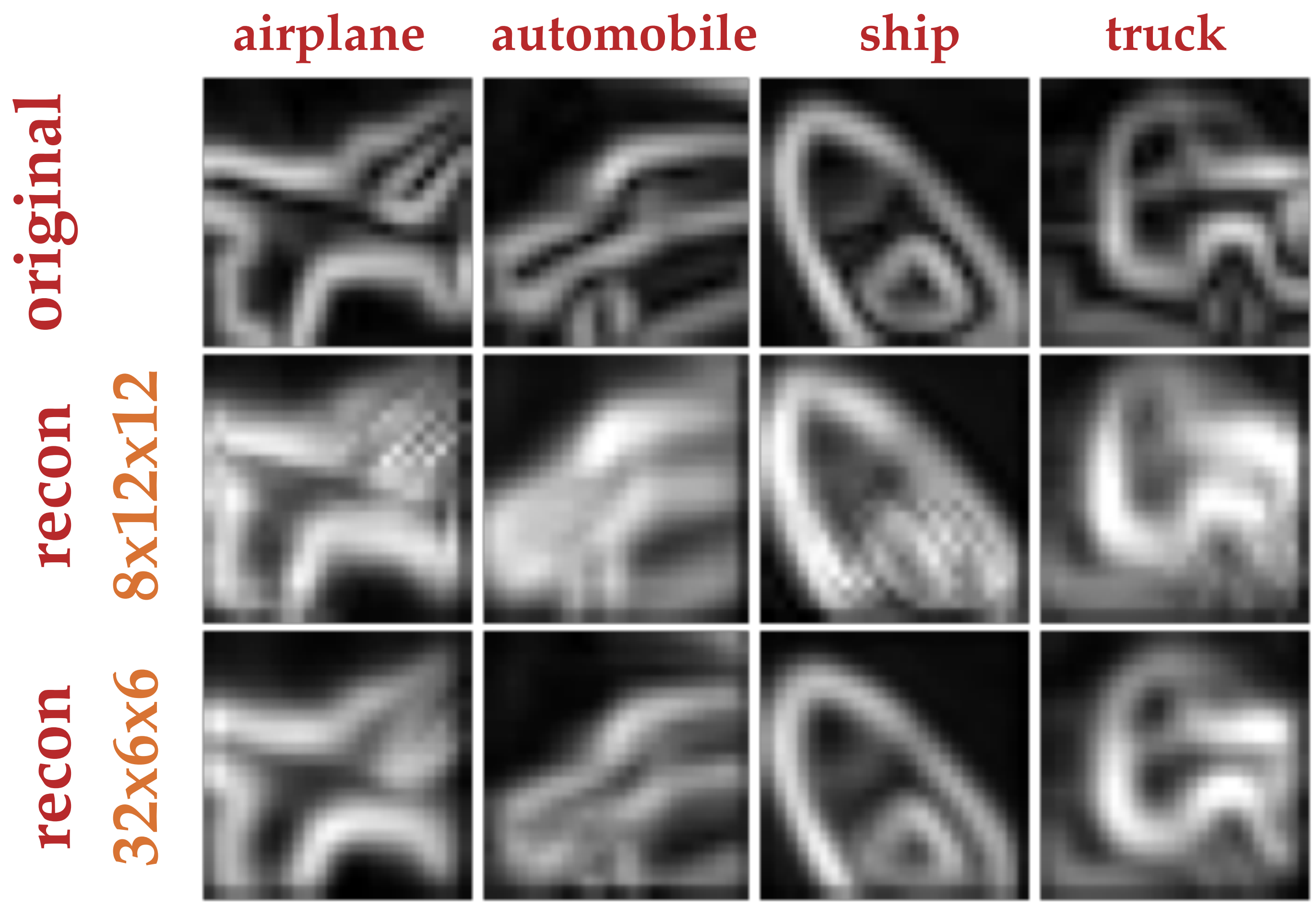
8 and 32 features

1100 *active* qubits
3068 *coupling* strengths

overcomplete order:

$$2 = \frac{12x12x8}{24x24x1}$$

stride: 2, 4



Los Alamos
NATIONAL LABORATORY
EST. 1943

- **24x24 patch images**

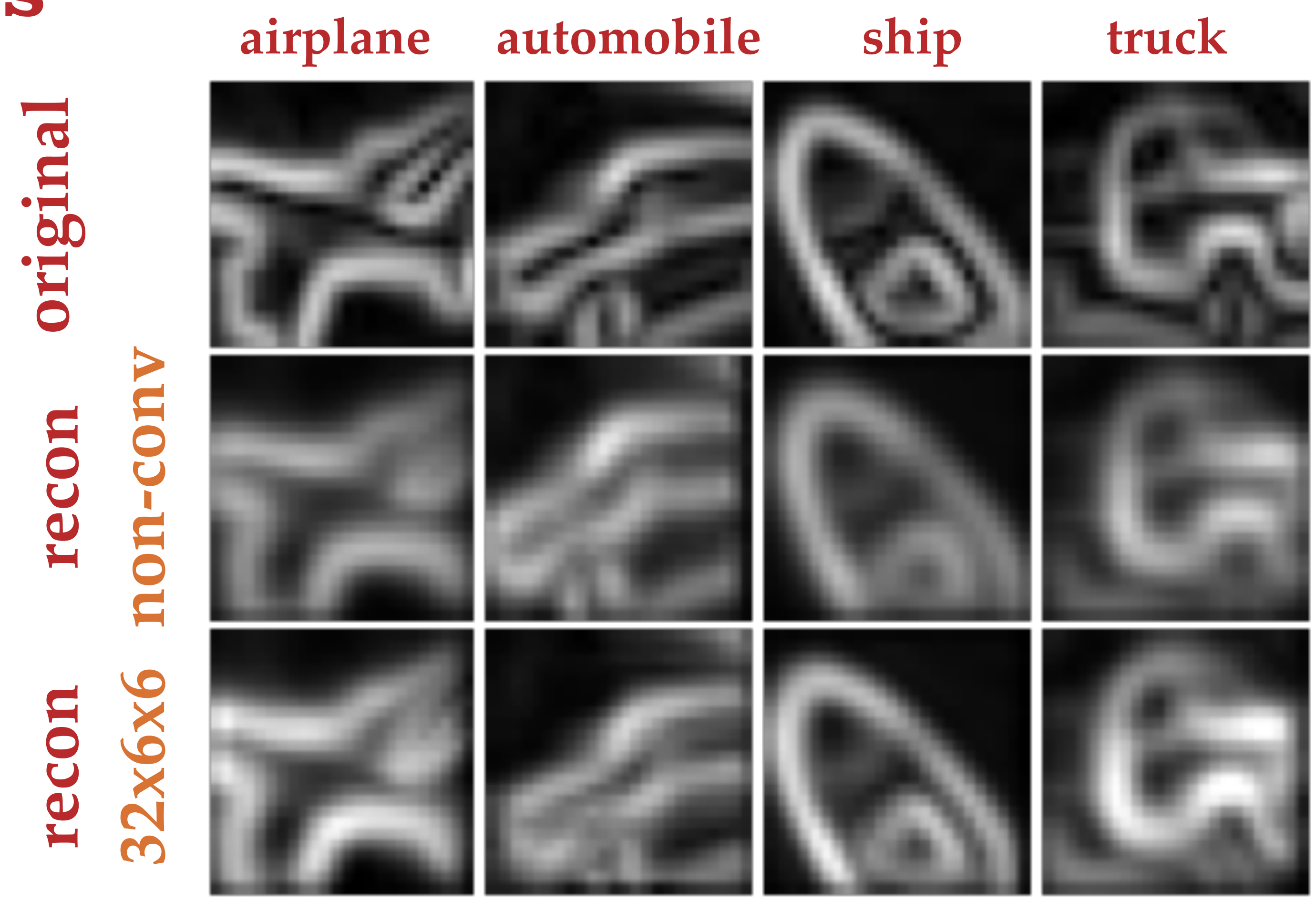
32 and 1152 features

1100 *active* qubits
3068 *coupling* strengths

overcomplete order:

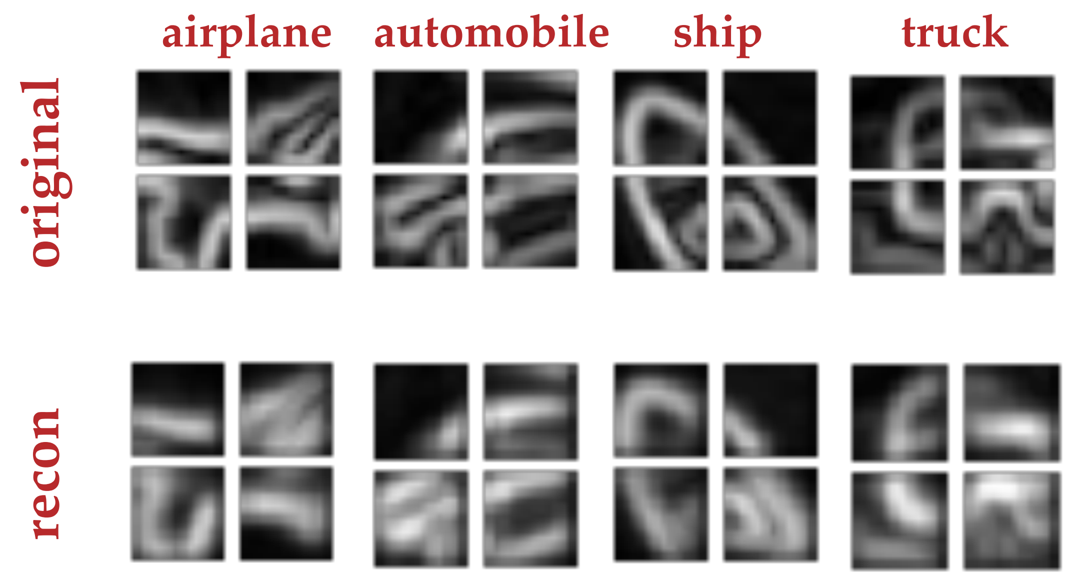
$$2 = \frac{12\text{x}12\text{x}8}{24\text{x}24\text{x}1}$$

stride: 4, 24



RESULTS

- 12x12 patch images



overcomplete order : 8

stride: 4

$$N_f = 32$$

Classification task: SVM (liblinear)

1042 training/208 test images

classes	air	auto	bird	cat	deer	dog	frog	horse	ship	truck
accur. (binary)	89.21%	93.38%	90.87%	89.42%	94.71%	88.94%	87.98%	89.9%	89.9%	85.58%

Nguyen and Kenyon, PMES-16 (2016)

D. COMPARISON WITH A CLASSICAL SOLVER

- So far, quantum computation (D-Wave 2X) has NOT outperformed its classical counterpart (GUROBI). Both are *comparable*.
- We already made the problem hard. We need to make it harder.
- How can we make the SC problem harder for both?

D. COMPARISON WITH A CLASSICAL SOLVER

From SC perspective: more overcomplete, harder to solve...

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity...

D. COMPARISON WITH A CLASSICAL SOLVER

From SC perspective: more overcomplete, harder to solve…

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity…

EMBEDDING technique

(c.f. D-Wave documents)

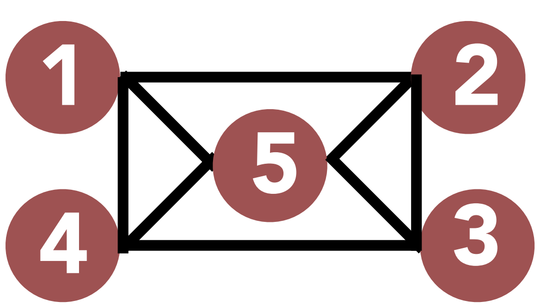
D. COMPARISON WITH A CLASSICAL SOLVER

From SC perspective: more overcomplete, harder to solve…

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity…

EMBEDDING technique

(c.f. D-Wave documents)



D. COMPARISON WITH A CLASSICAL SOLVER

GUROBI OPTIMIZATION

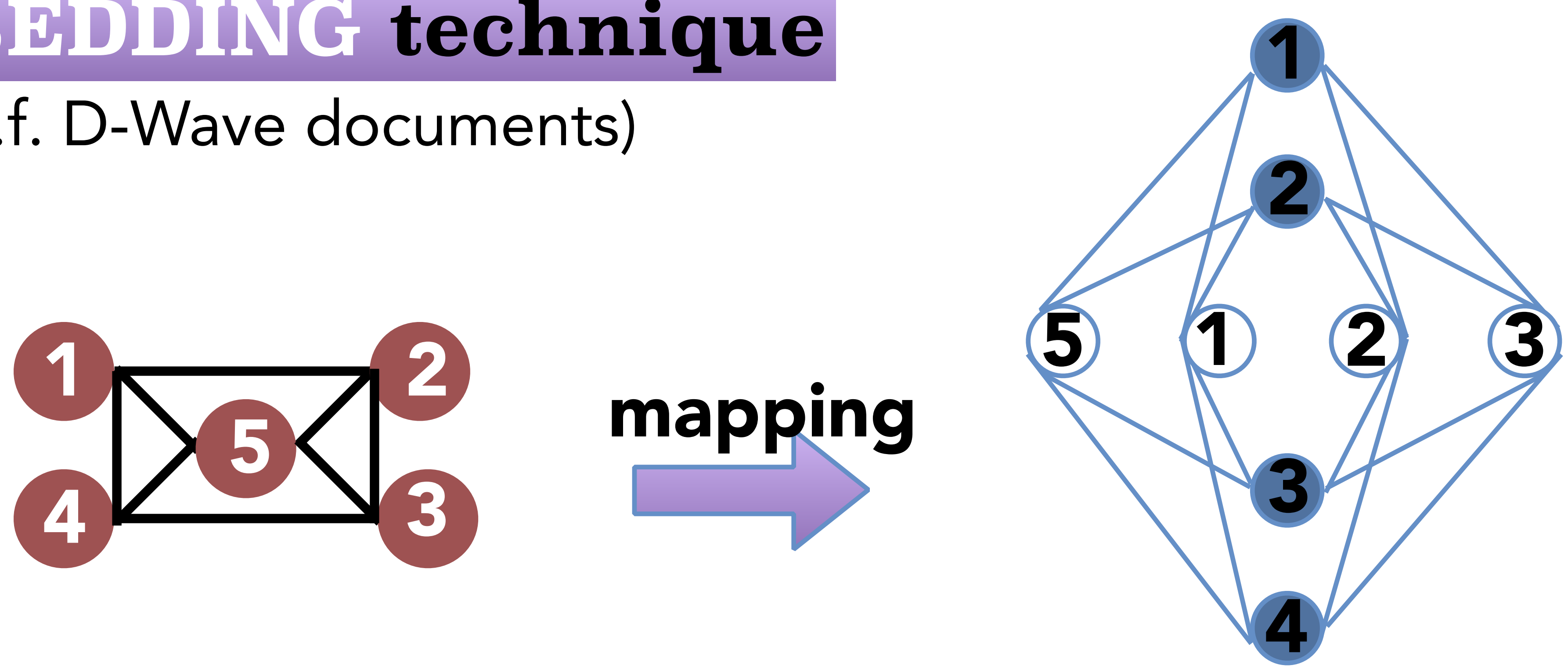
Los Alamos NATIONAL LABORATORY EST. 1943

From SC perspective: more overcomplete, harder to solve…

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity…

EMBEDDING technique

(c.f. D-Wave documents)



D. COMPARISON WITH A CLASSICAL SOLVER

GUROBI
OPTIMIZATION

Los Alamos
NATIONAL LABORATORY
EST. 1943

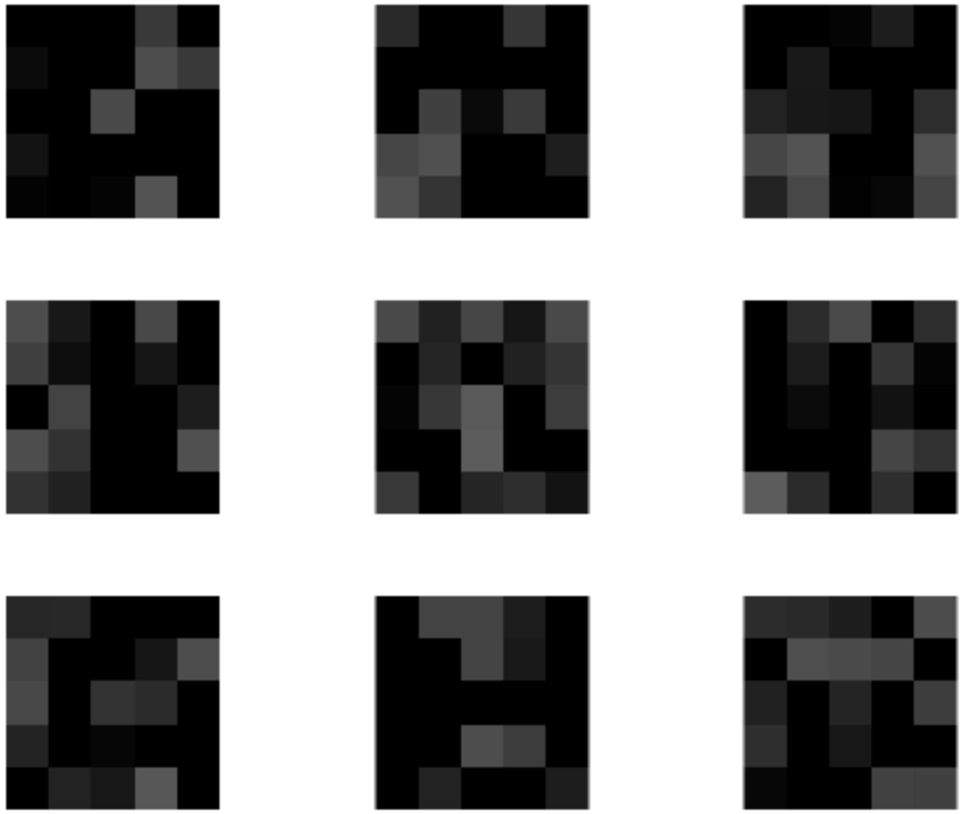
From SC perspective: more overcomplete, harder to solve…

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity…

EMBEDDING technique

- Employ all bipartite couplings
- Small number of nodes (qubits)

5x5



D. COMPARISON WITH A CLASSICAL SOLVER

GUROBI OPTIMIZATION

Los Alamos NATIONAL LABORATORY EST. 1943

From SC perspective: more overcomplete, harder to solve…

Meanwhile: The full Chimera in D-Wave offers a certain set of (*nearest-neighbor*) connectivity…

EMBEDDING technique

- Employ all bipartite couplings
- Small number of nodes (qubits)

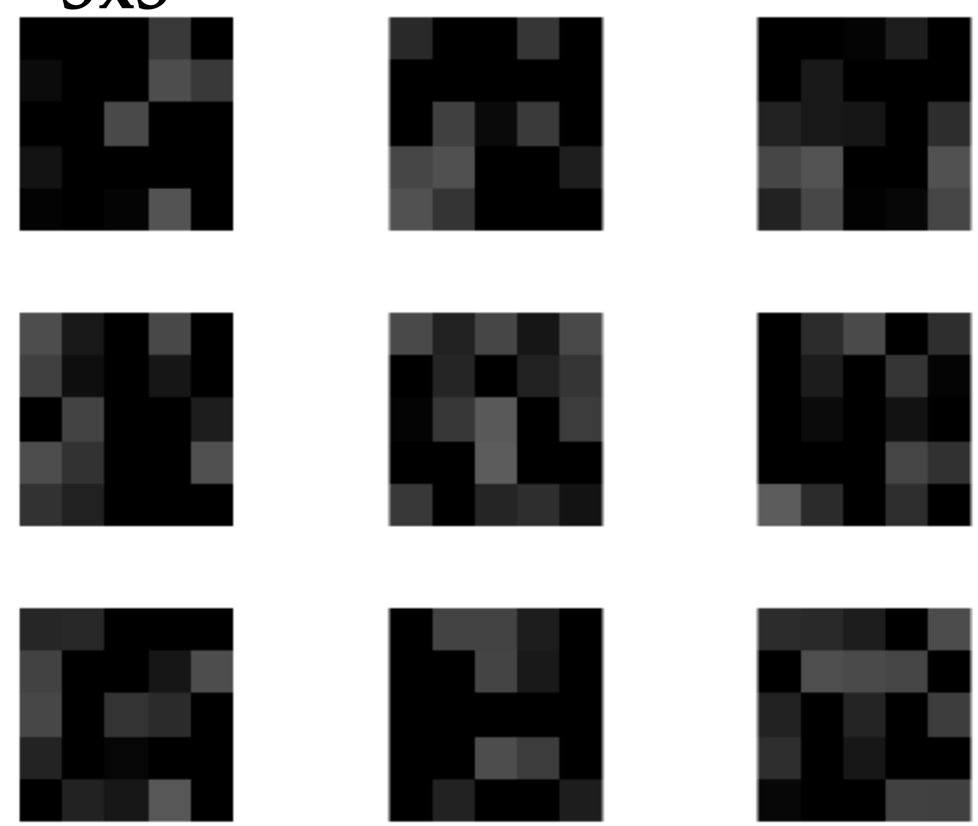
In practice (D-Wave 2X):

Fully connected: 49 nodes

Partially orthogonal: 74 nodes

Feature optimization!

5x5

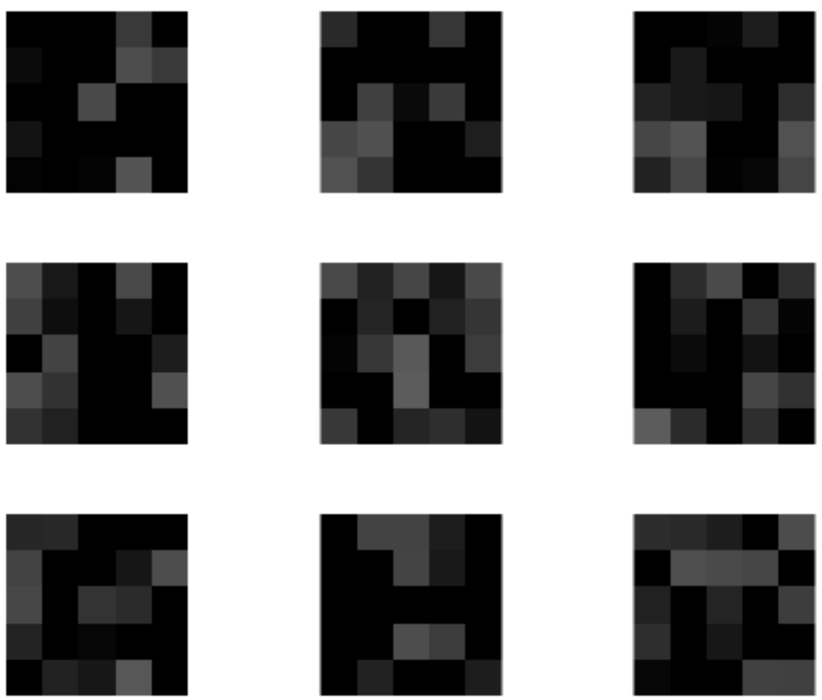


D. COMPARISON WITH A CLASSICAL SOLVER

Feature Learning (in progress)

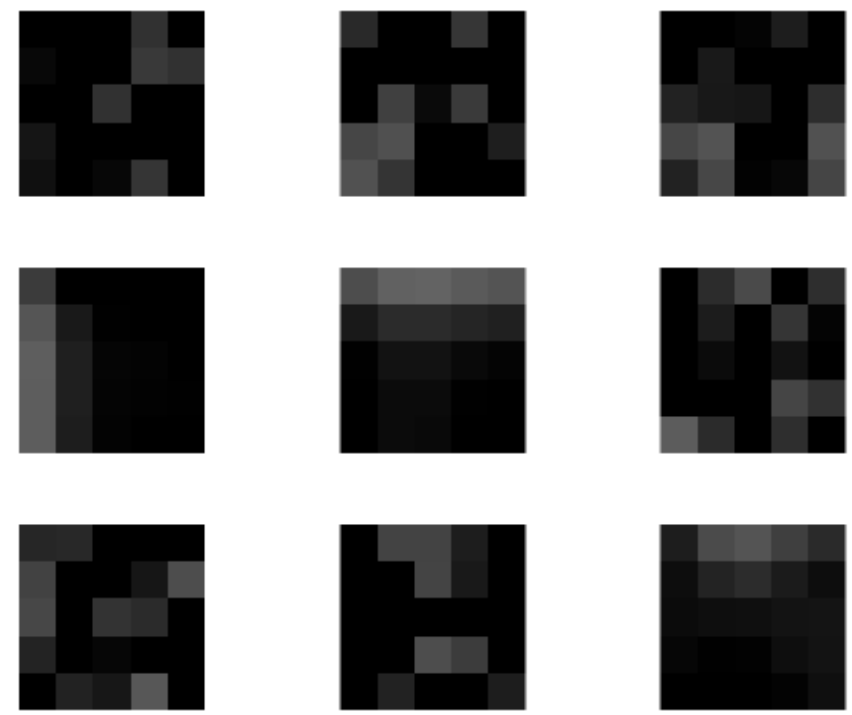
before…

5x5



…after

5x5



D. COMPARISON WITH A CLASSICAL SOLVER

STARTING TO SEE SOMETHING GOOD…

No. of (random) Hamiltonians: 1

solver / problem	GUROBI (best classical solver)		D-Wave 2X (ISING)	
	Energy	Time	Energy	Time
49 nodes: fully connected	-29.99	~ 480 seconds	-29.99	few seconds

D. COMPARISON WITH A CLASSICAL SOLVER

STARTING TO SEE SOMETHING GOOD…

No. of (random) Hamiltonians: 1

problem \ solver	GUROBI (best classical solver)		D-Wave 2X (ISING)	
	Energy	*Time*	*Energy*	*Time*
49 nodes: fully connected	-29.99	~ **480** seconds	-29.99	*few seconds*
	Energy	*Time*	*Energy*	*Time*
72 nodes: partially Chimera-orthogonal	-48.476	~**1800** seconds	-51.295	*few seconds*

COMPRESSIVE SENSING ON A QUANTUM MACHINE

WHY NOT?

Los Alamos NATIONAL LABORATORY EST. 1943

COMPRESSIVE SENSING ON A QUANTUM MACHINE
WHY NOT?

COMPRESSIVE SENSING in a nutshell

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} & \Phi & \end{bmatrix} \begin{bmatrix} x \\ \\ \\ \end{bmatrix}$$

y: measurement, sparse

x: compressible signal

Φ: sensing matrix

COMPRESSIVE SENSING ON A QUANTUM MACHINE
WHY NOT?

COMPRESSIVE SENSING in a nutshell

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} \quad \Phi \quad \end{bmatrix} \begin{bmatrix} x \\ \\ \\ \end{bmatrix}$$

L0: NP-hard
L1: mostly-used

y: measurement, sparse
x: compressible signal
Φ: sensing matrix

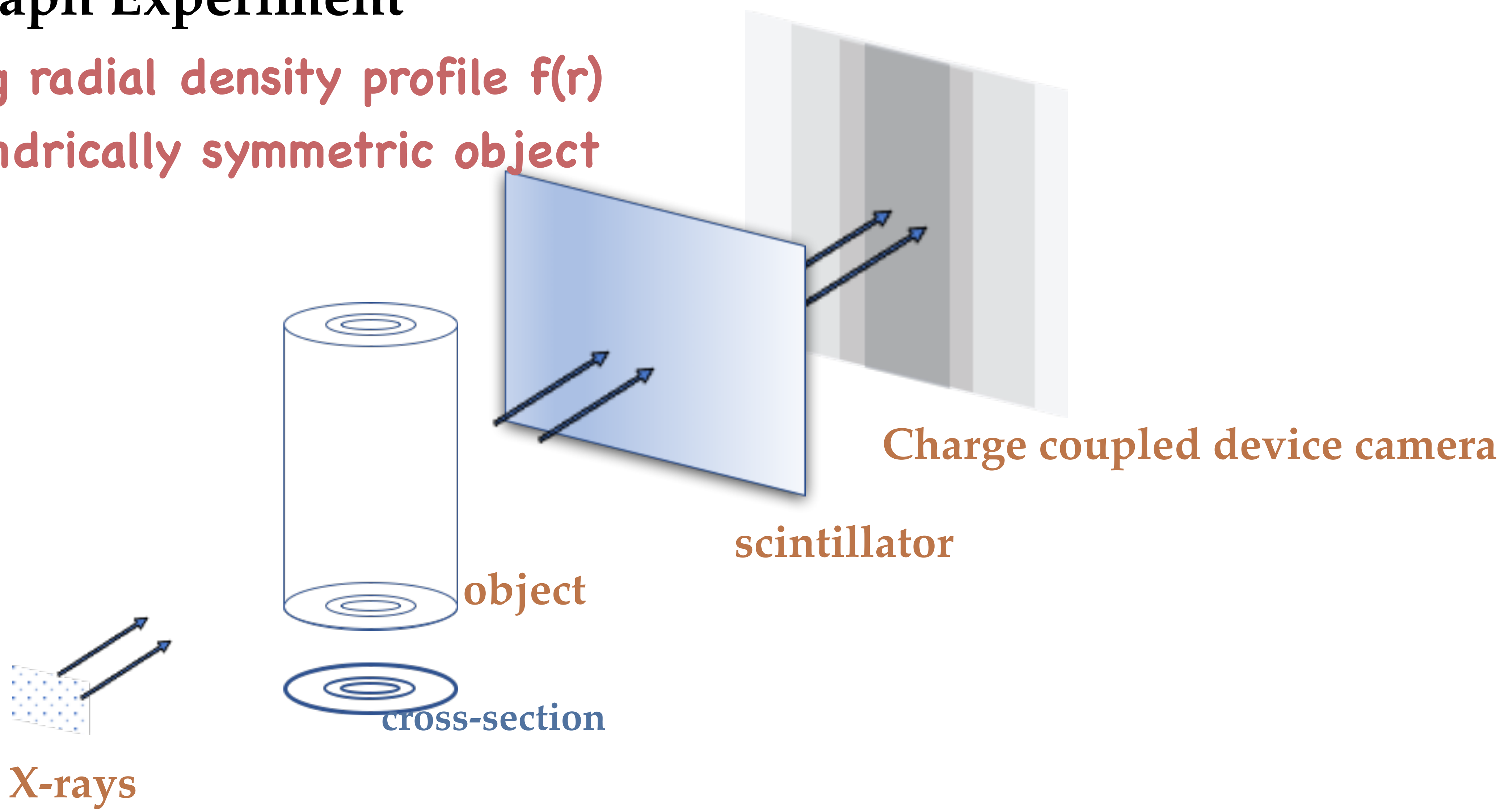
wide applications particularly in image processing (X-ray, CT, …), sampling, etc (c.f. Candes, Baraniuk, *Compressive Sensing*)

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm

Radiograph Experiment

Obtaining radial density profile f(r) of a cylindrically symmetric object



COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm

projected radial density of a cylindrically symmetric object:

$h(y)$ ← projected, measurable

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm

projected radial density of a cylindrically symmetric object:

h(y) ← projected, measurable

Task: Recovering the density profile of the cylindrical object f(r):

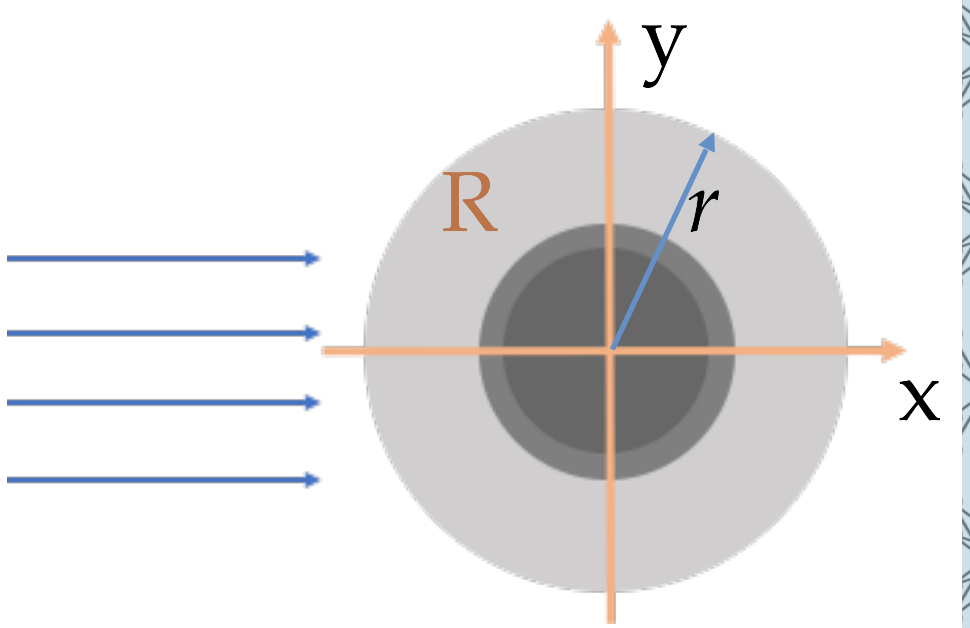
COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm

projected radial density of a cylindrically symmetric object:

h(y) ← projected, measurable

Task: Recovering the density profile of the cylindrical object f(r): $h(y) = 2\int_y^R \frac{f(r)rdr}{\sqrt{(r^2 - y^2)}}$



How?:

one method:

to obtain a slice of the original object ← **Abel inversion**

$$\boxed{h(y) = \mathbf{A} * f(r)}$$

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm + **D-Wave QUBO**

Our method is applicable to e.g. X-ray images as sparse inputs

$$h = \mathbf{A} f$$

We obtain:

$$A_{ij} = 2 \sum_{i|r_i=y_j}^{i|r_i=R} \frac{r_i \Delta_{ij}}{\sqrt{r_i^2 - y_j^2}}$$

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm + **D-Wave QUBO**

Our method is applicable to e.g. X-ray images as sparse inputs

$$h = \mathbf{A} \, f$$

We obtain: $$A_{ij} = 2 \sum_{i|r_i=y_j}^{i|r_i=R} \frac{r_i \Delta_{ij}}{\sqrt{r_i^2 - y_j^2}}$$

Next: to choose (learn) domain basis:

$$f(r) = \sum_n s_n F_n(r)$$

s_n are sparse Fourier coefficients.

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Abel transform algorithm + D-Wave QUBO

Our method is applicable to e.g. X-ray images as sparse inputs

$$h = \mathbf{A} \, f$$

We obtain: $$A_{ij} = 2 \sum_{i|r_i=y_j}^{i|r_i=R} \frac{r_i \Delta_{ij}}{\sqrt{r_i^2 - y_j^2}}$$

Next: to choose (learn) domain basis:

$$f(r) = \sum_n s_n F_n(r)$$

s_n are sparse Fourier coefficients.

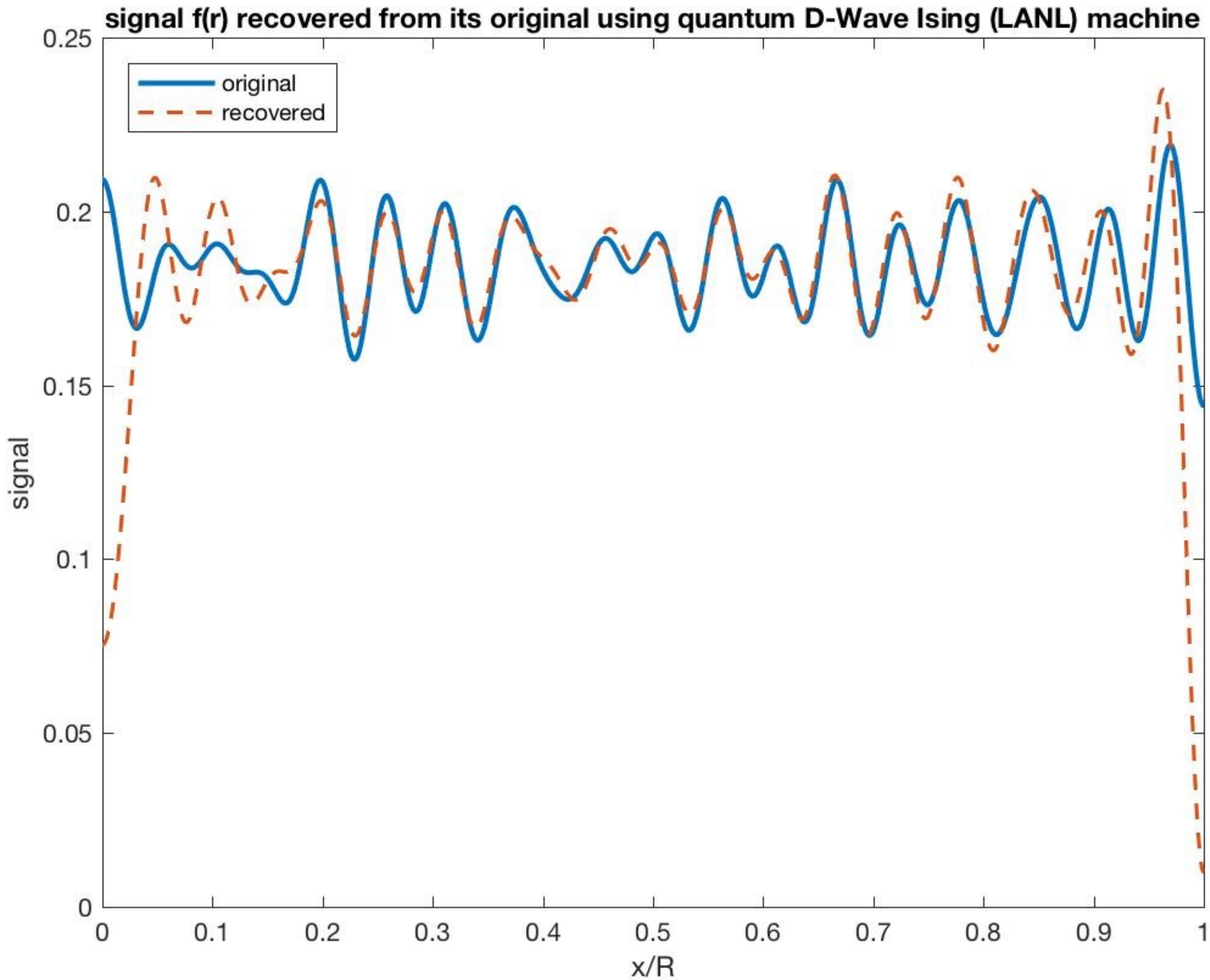
then:

$$h = \mathbf{A} \, \mathbf{F} \, s$$

Ising input

COMPRESSIVE SENSING ON A QUANTUM MACHINE

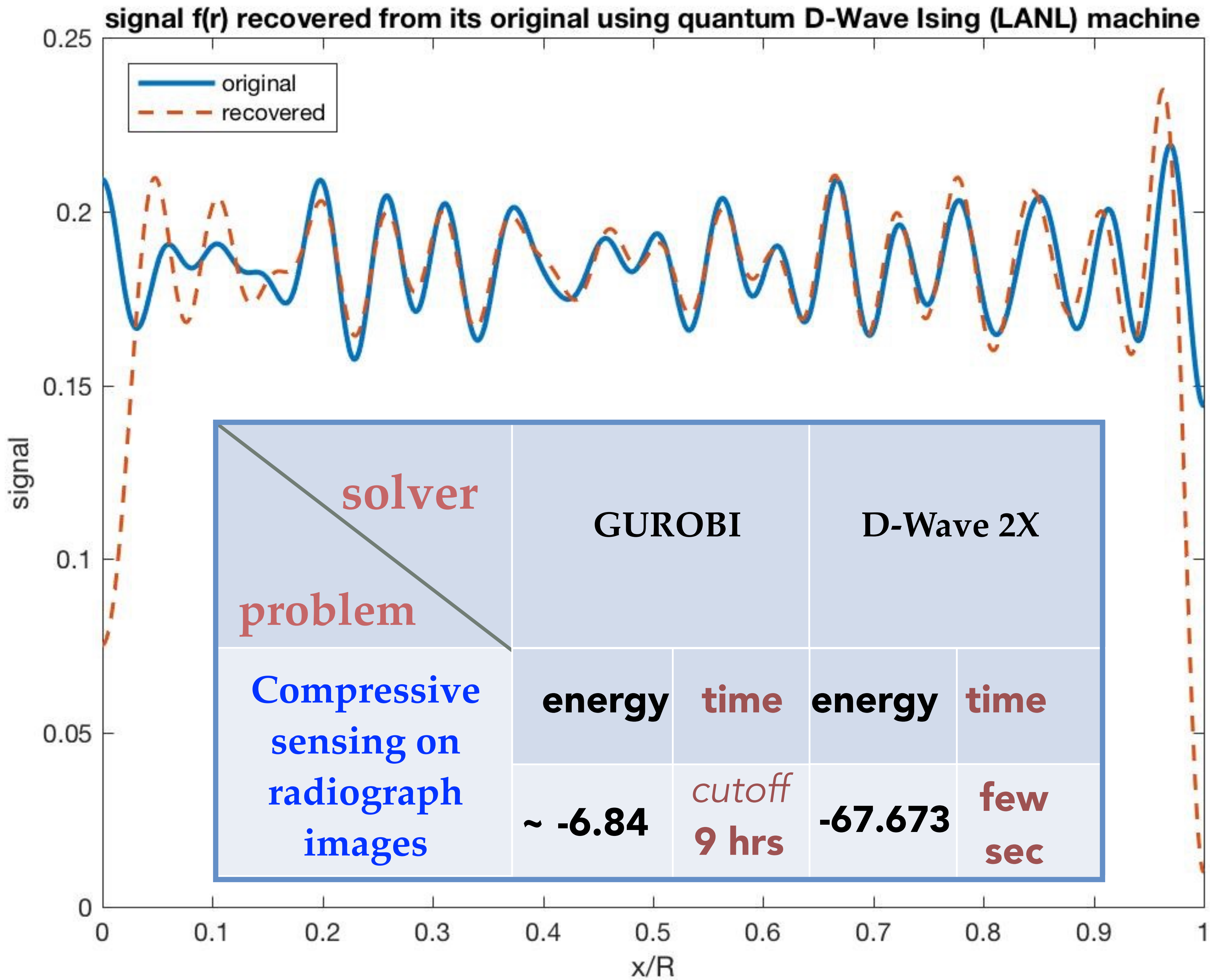
Results



D-Wave Ising pulls out ALL non-zero Fourier coefficients (frequencies)

COMPRESSIVE SENSING ON A QUANTUM MACHINE

Results



solver / problem	GUROBI		D-Wave 2X	
	energy	time	energy	time
Compressive sensing on radiograph images	~ -6.84	*cutoff* 9 hrs	-67.673	few sec

D-Wave Ising pulls out ALL non-zero Fourier coefficients (frequencies)

E. SUMMARY

- first demonstration of sparse coding using quantum computer
- benchmark results on standard image classification task
- mapping of visual features to D-Wave Chimera
- *compare* D-Wave performance with GUROBI
- *compressive sensing* on Ising for density profile detection where D-Wave *significantly* outperforms GUROBI

E. FUTURE WORK

- optimize features
- add colors
- TrueNorth comparison
- hierarchy model
- compressive sensing on real images

THANK YOU FOR YOUR ATTENTION!