Quantum Machine Learning on a Quantum Annealer

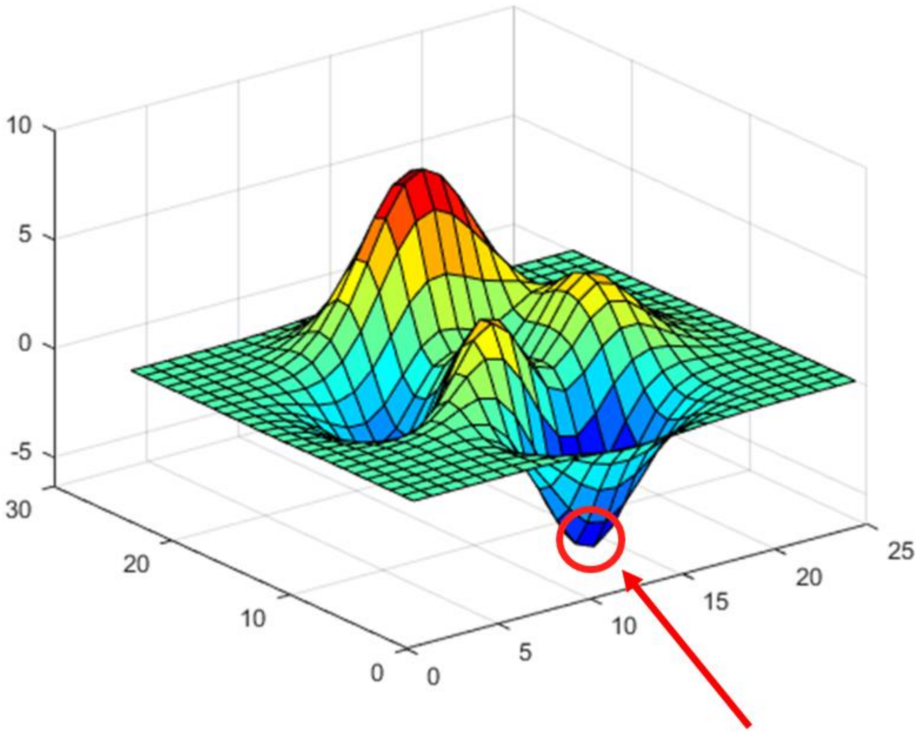
Restricted Boltzmann Machines and Recommendation Systems

Lorenzo Rocutto – PhD @ University of Bologna

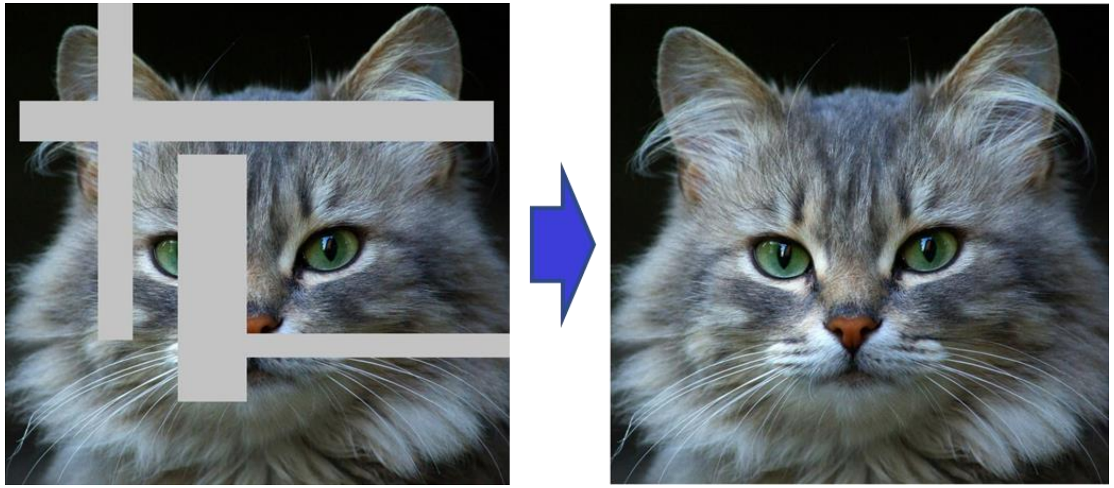
CINECA QUANTUM COMPUTING LAB

A new problem for AQCs

1. Optimization problem



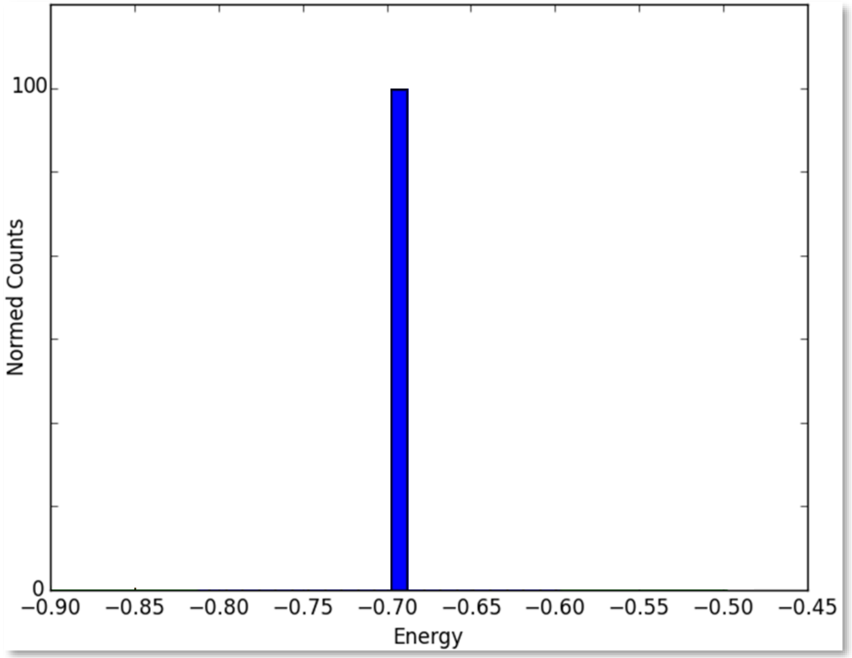
2. Artificial Intelligence: Generative model



A new problem for AQCs

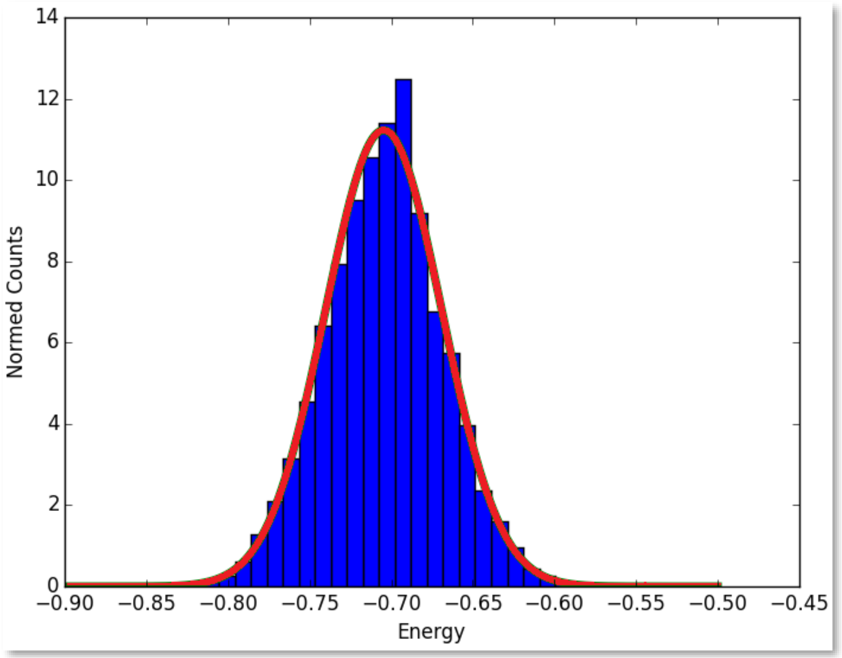
1. Optimization problem

Quantum processor should return only the correct answer



2. Artificial Intelligence: Generative model

Quantum processor should return a probabilistic answer



CINECA QUANTUM COMPUTING LAB

Ideal adiabatic quantum computer

Hypothesis:

1. Complete superposition
2. Slow annealing
3. Absence of environment coupling
4. H_P is implemented exactly
5. No readout errors

Ideal adiabatic quantum computer

Hypothesis:

1. Complete superposition
2. Slow annealing
3. Absence of environment coupling
4. H_P is implemented exactly
5. No readout errors

Which is the least satisfied hypothesis?

Ideal adiabatic quantum computer

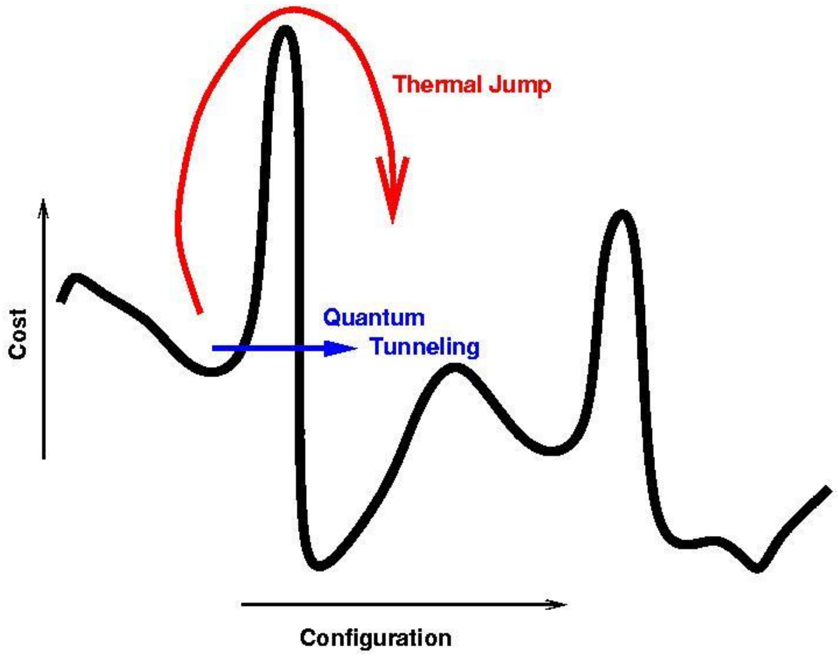
Hypothesis:

1. Complete superposition
2. Slow annealing
3. *Absence of environment coupling*
4. H_P is implemented exactly
5. No readout errors

W h is the least

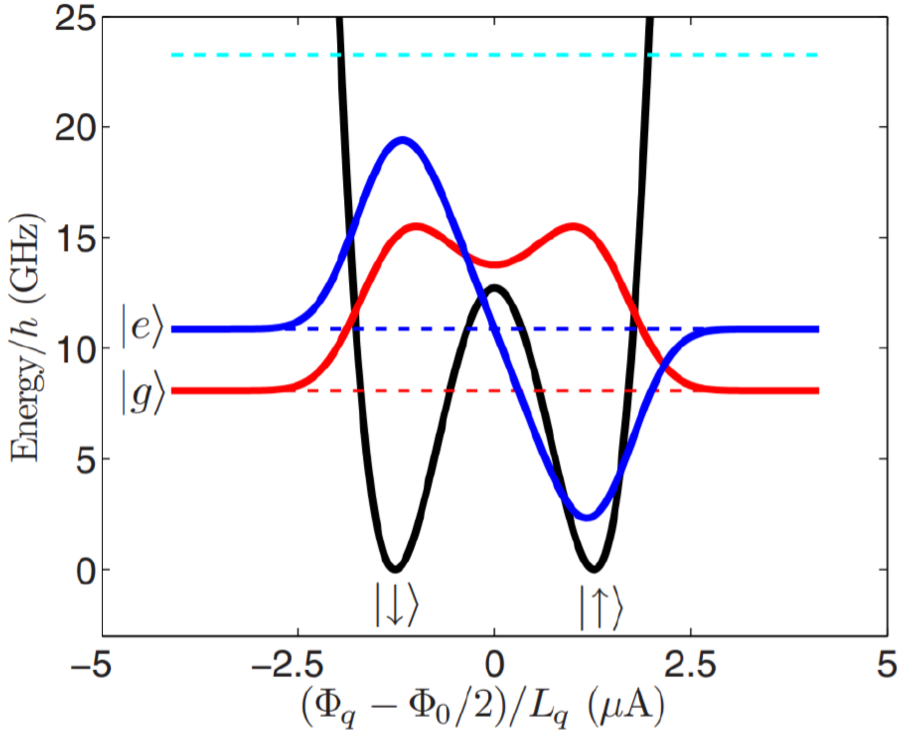
hypothesis?

Exploiting the limits of AQCs



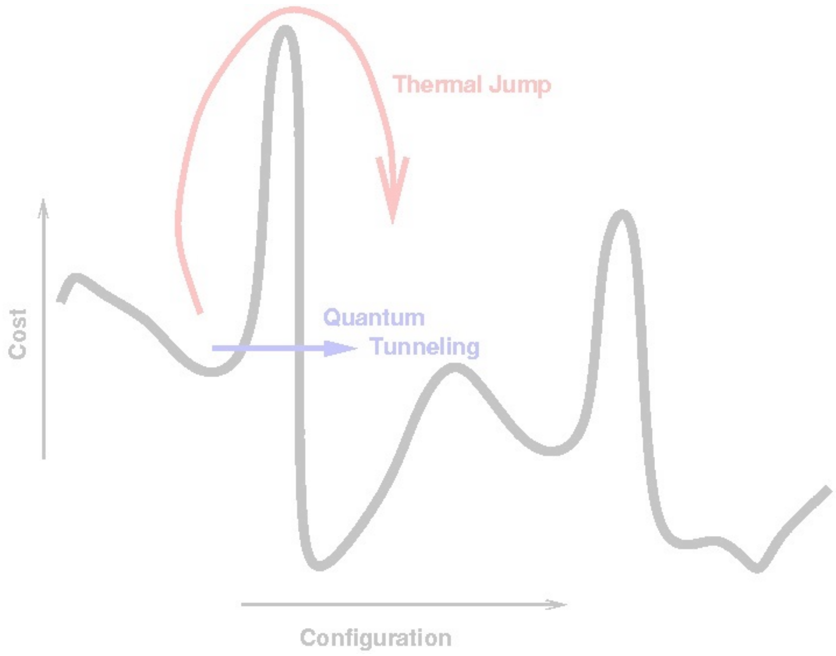
- Quantum annealing tries to tunnell through energy barriers….
- … But also thermal fluctuations can help the process

Single-qubit energy barrier



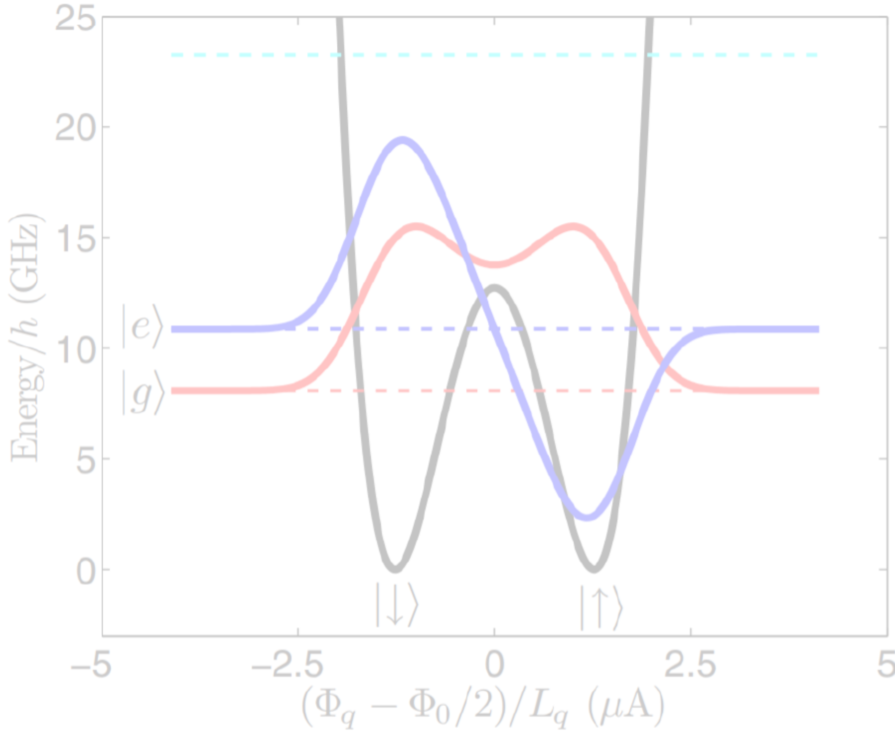
CINECA QUANTUM COMPUTING LAB

Exploiting the limits of AQCs

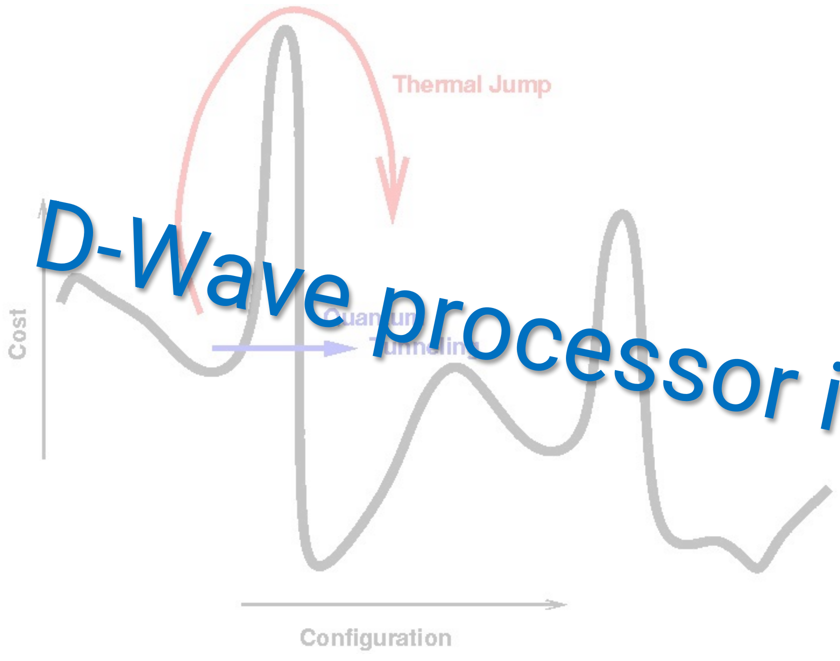


- Quantum annealing tries to tunnell through energy barriers….
- … But also thermal fluctuations can help the process

Single-qubit energy barrier

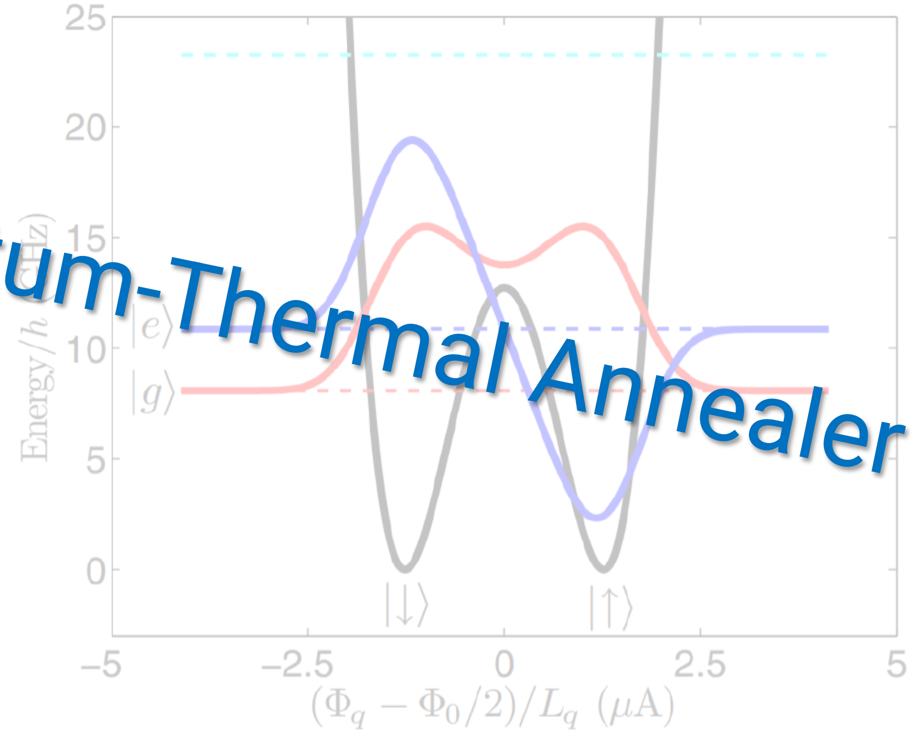


Exploiting the limits of AQCs



- Quantum annealing tries to tunnell through energy barriers….
- … But also thermal fluctuations can help the process

Single-qubit energy barrier



D-Wave processor is a Quantum-Thermal Annealer

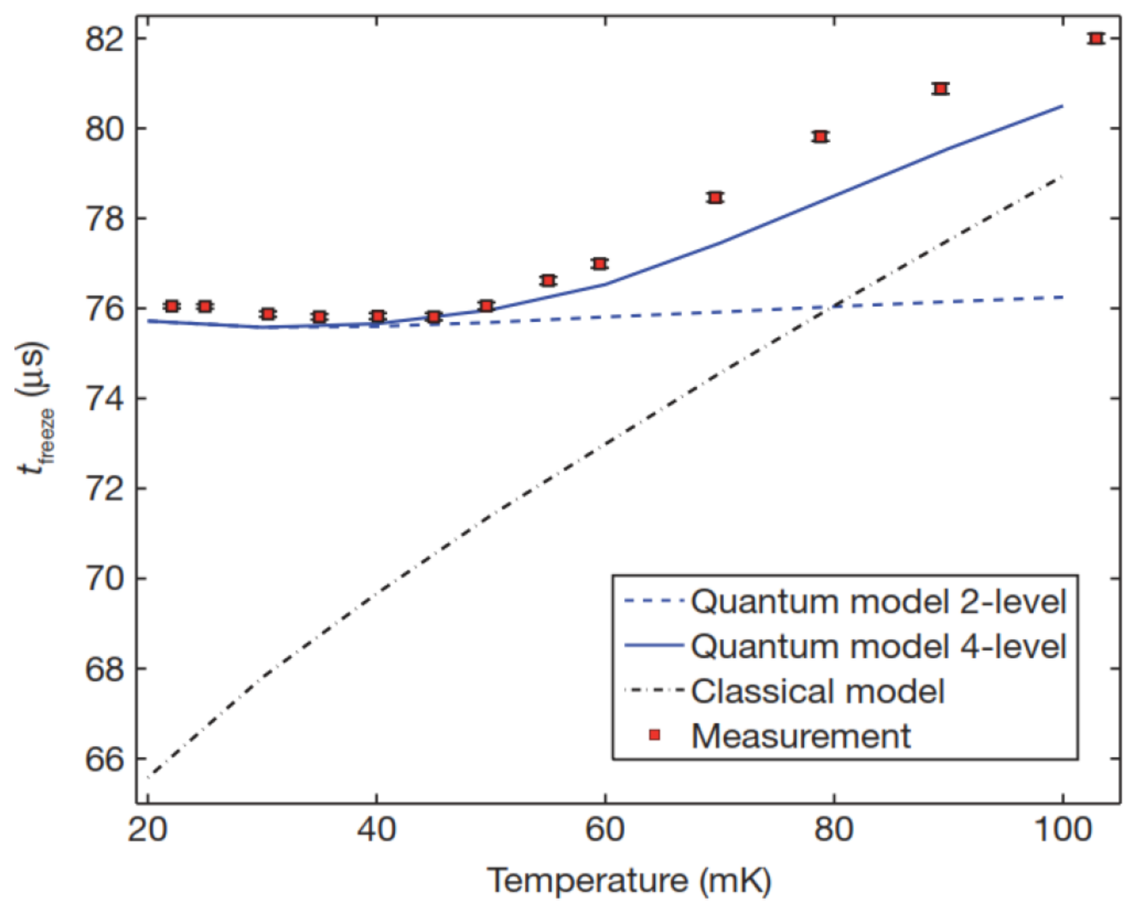
The process freezes

When the thermal fluctuations dominate the process (high T):

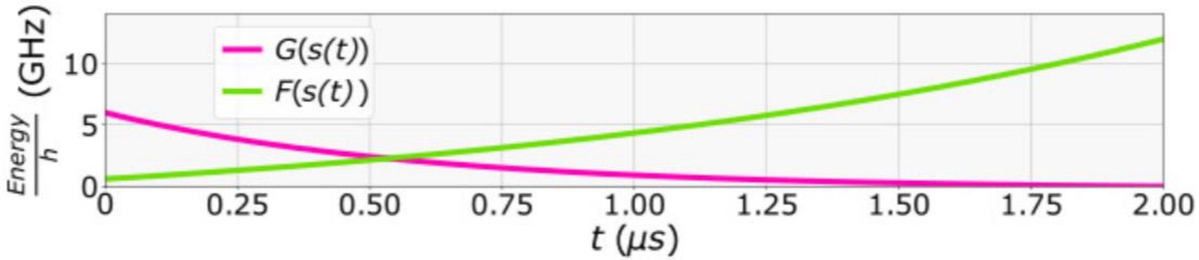
$$P(spin\ flip) \sim e^{-\delta U/k_B T}$$

When $\delta U \gg k_B T$ thermal fluctuations stop.

Quantum tunnelling probability depends on the energy barriers. Not on T.

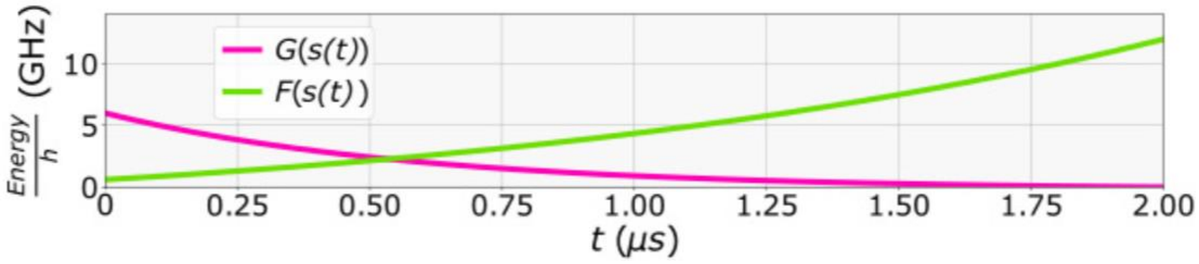


Distribution produced by AQCs



D-Wave quantum annealers can be used as a generator of samples that follows the *Boltzmann distribution* of the classical cost function encoded in H_P.

Distribution produced by AQCs



D-Wave quantum annealers can be used as a generator of samples that follows the *Boltzmann distribution* of the classical cost function encoded in H_P.

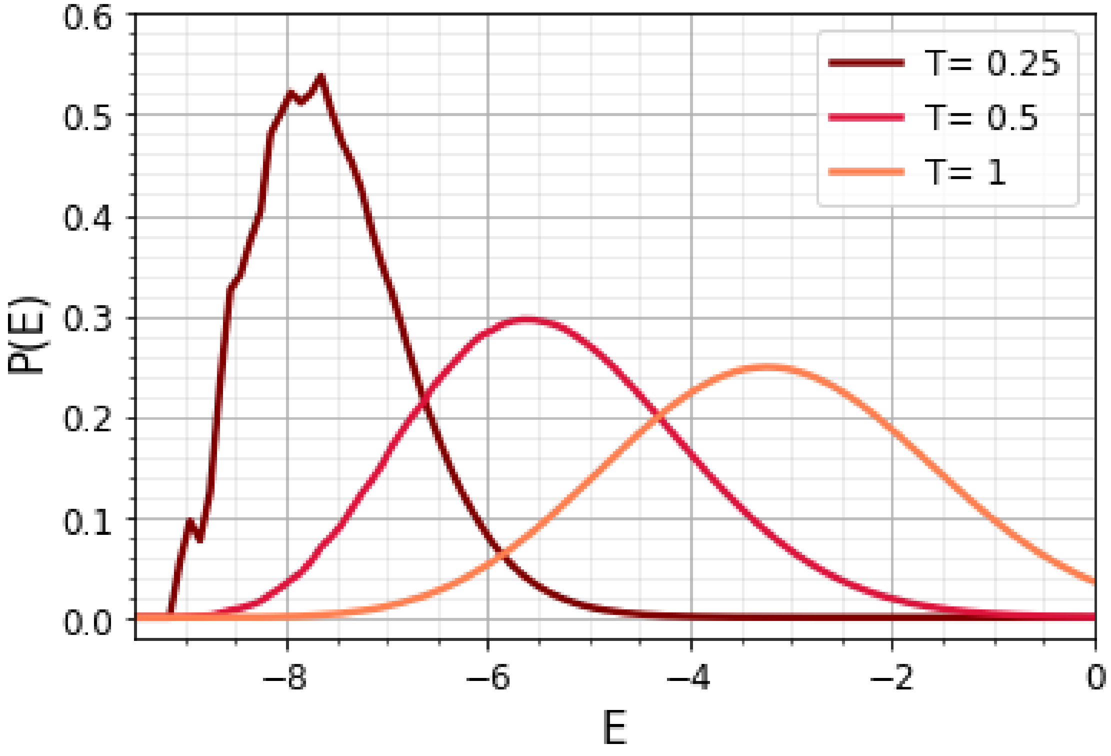
$$J = 7.9\,GHz \rightarrow J = 1.0$$

$$T = 0.38\,K \rightarrow T_{eff} = 1.0$$

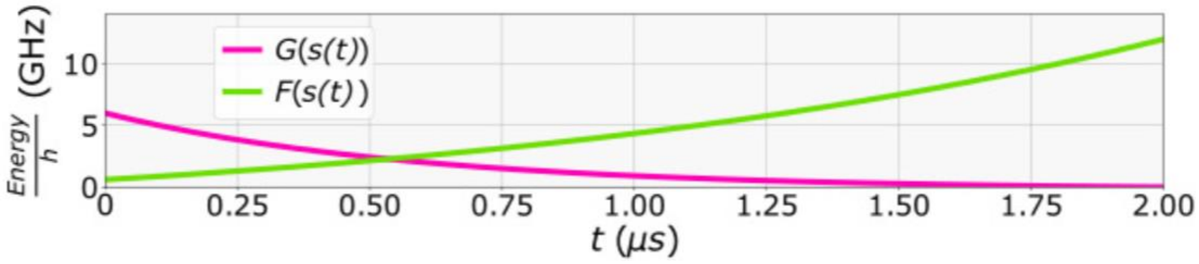
$$T = 12.5\,mK \rightarrow T_{eff} = 0.033$$

Hypothesis:

$$|\phi(\tau > \tau_{\text{freeze}})\rangle \cong \frac{1}{Z^{1/2}} \sum_{j=1}^{2^N} e^{i\theta_j} e^{-\frac{E_j}{2T_{\text{eff}}}} |S_j\rangle$$



Distribution produced by AQCs



D-Wave quantum annealers can be used as a generator of samples that follows the *Boltzmann distribution* of the classical cost function encoded in H_P.

$J \ = \ 7.9\, GHz \ \rightarrow \ J \ = 1.0$

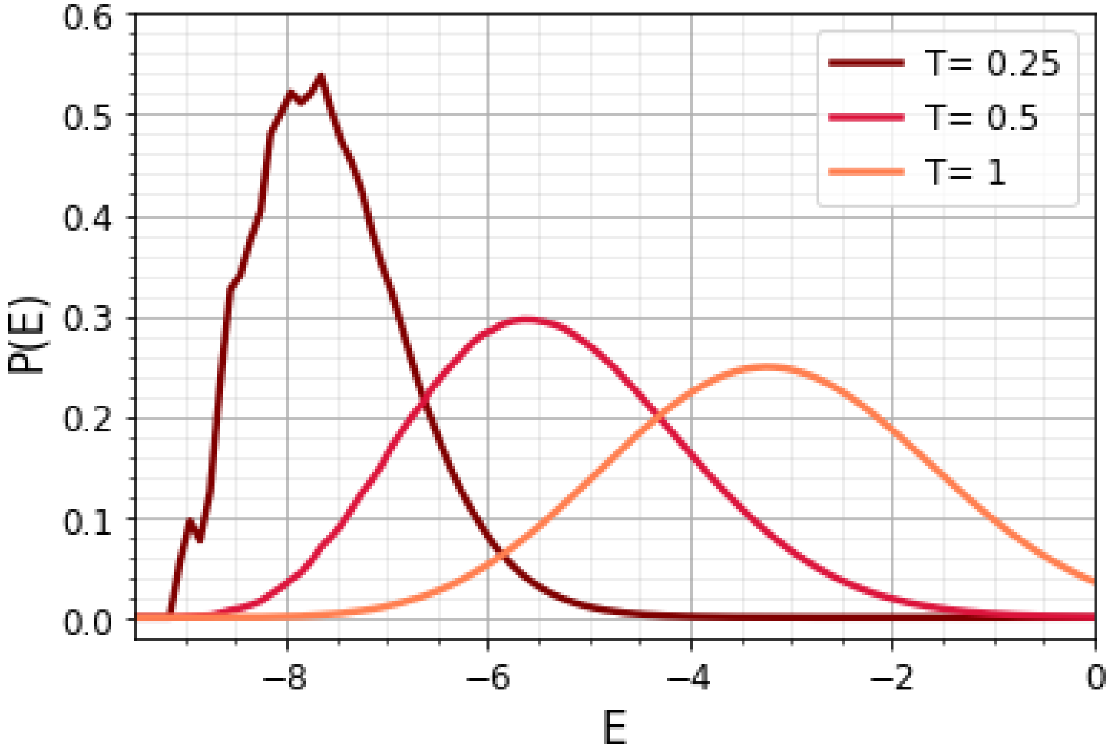
$T \ = \ 0.38\, K \ \rightarrow \ T_{eff} = 1.0$

$T \ = \ 12.5\, mK \ \rightarrow \ T_{eff} = 0.033$

Operating temperature of a D-Wave QPU

Hypothesis:

$$|\phi(\tau > \tau_{\text{freeze}})\rangle \cong \frac{1}{Z^{1/2}} \sum_{j=1}^{2^N} e^{i\theta_j} e^{-\frac{E_j}{2T_{\text{eff}}}} |S_j\rangle$$



CINECA QUANTUM COMPUTING LAB

Boltzmann Machine

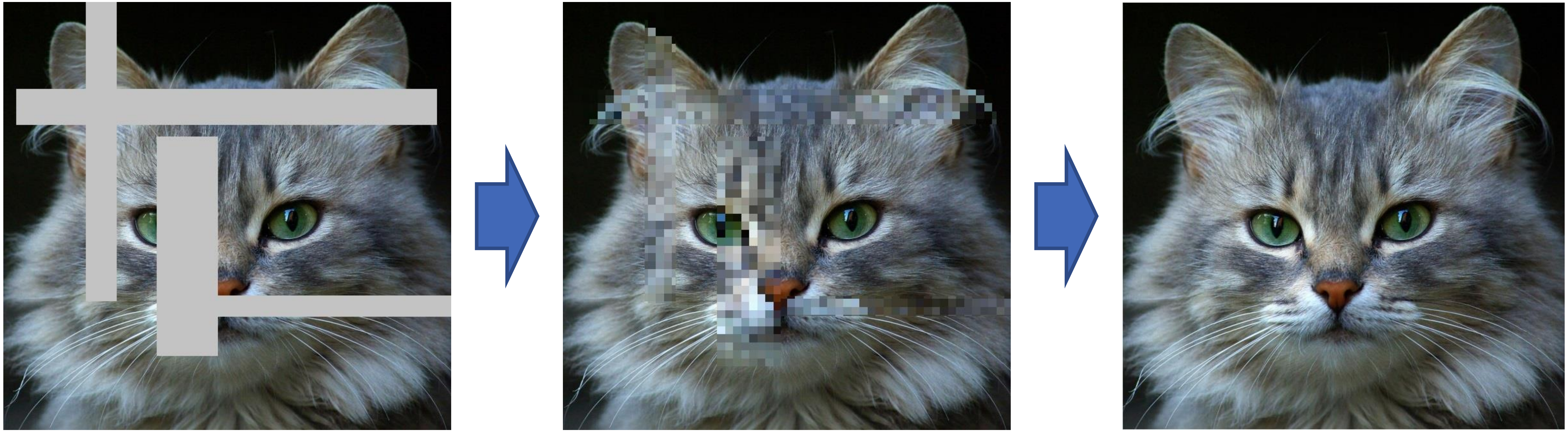
A native neural network for quantum annealers

Lorenzo Rocutto

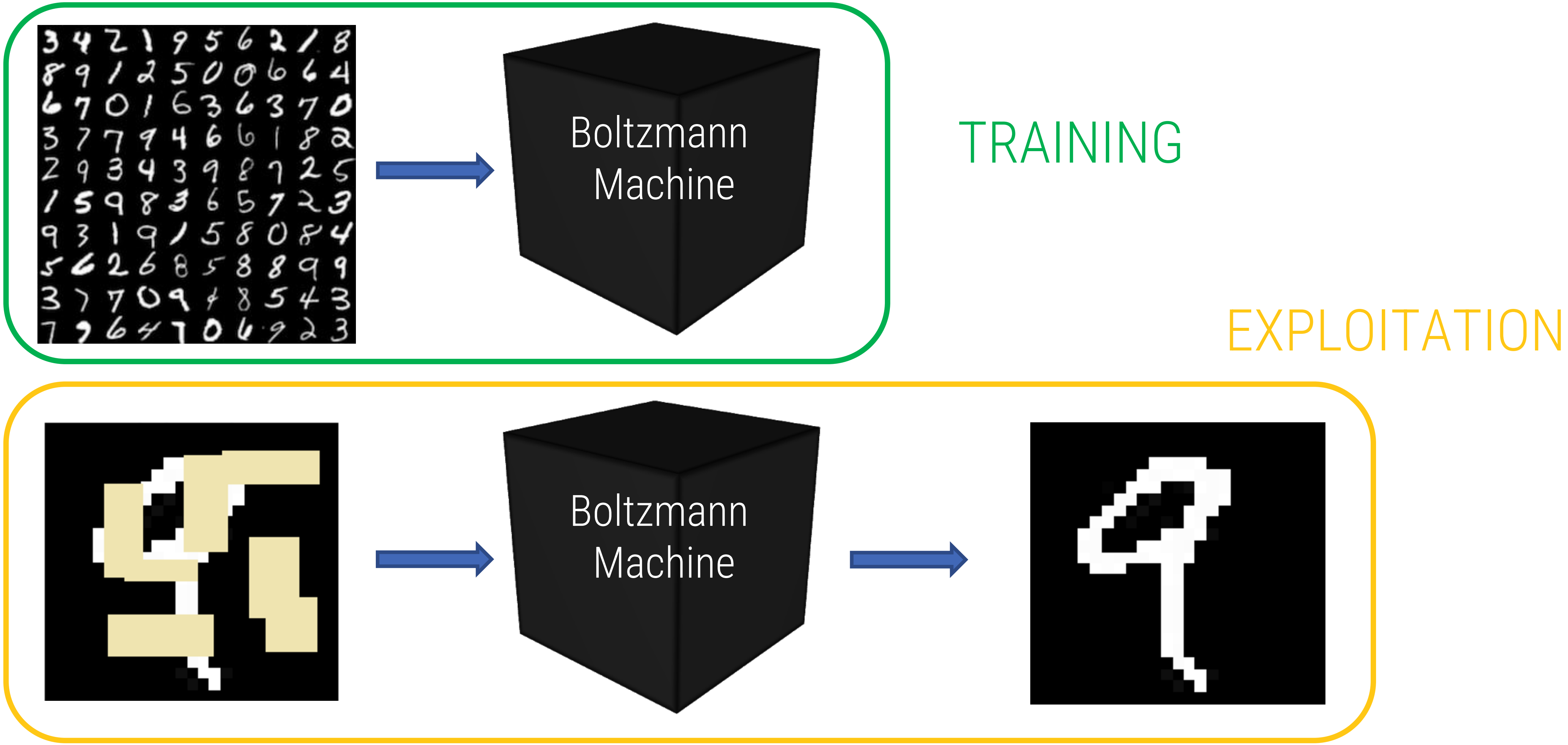
CINECA QUANTUM COMPUTING LAB

Applying a ML model

- We hypothesize the existence of a conditioned probability distribution in the data
- The Boltzmann Machine tries to mimic that distribution to reconstruct missing data



A well known generative model



RBMs can solve many problems

Possible Applications:

- **Recommendation systems,**
- **Network Anomaly Detection,**
- **Fraud detection,**
- **Quantum tomography,**
- **…**

Lessons from the Netflix Prize Challenge
2007
Robert M. Bell and Yehuda Koren

Improved traffic detection with support vector machine based on restricted Boltzmann machine
30 Dec, 2015
Jun Yang[1] · Jiangdong Deng[1] · Shujuan Li[1] · Yongle Hao[1]

Abnormal Traffic Pattern Detection in Real-Time Financial Transactions
13 Mar, 2019
Sean Rastatter, Travis Moe, Amitava Gangopadhyay and Alfred Weaver

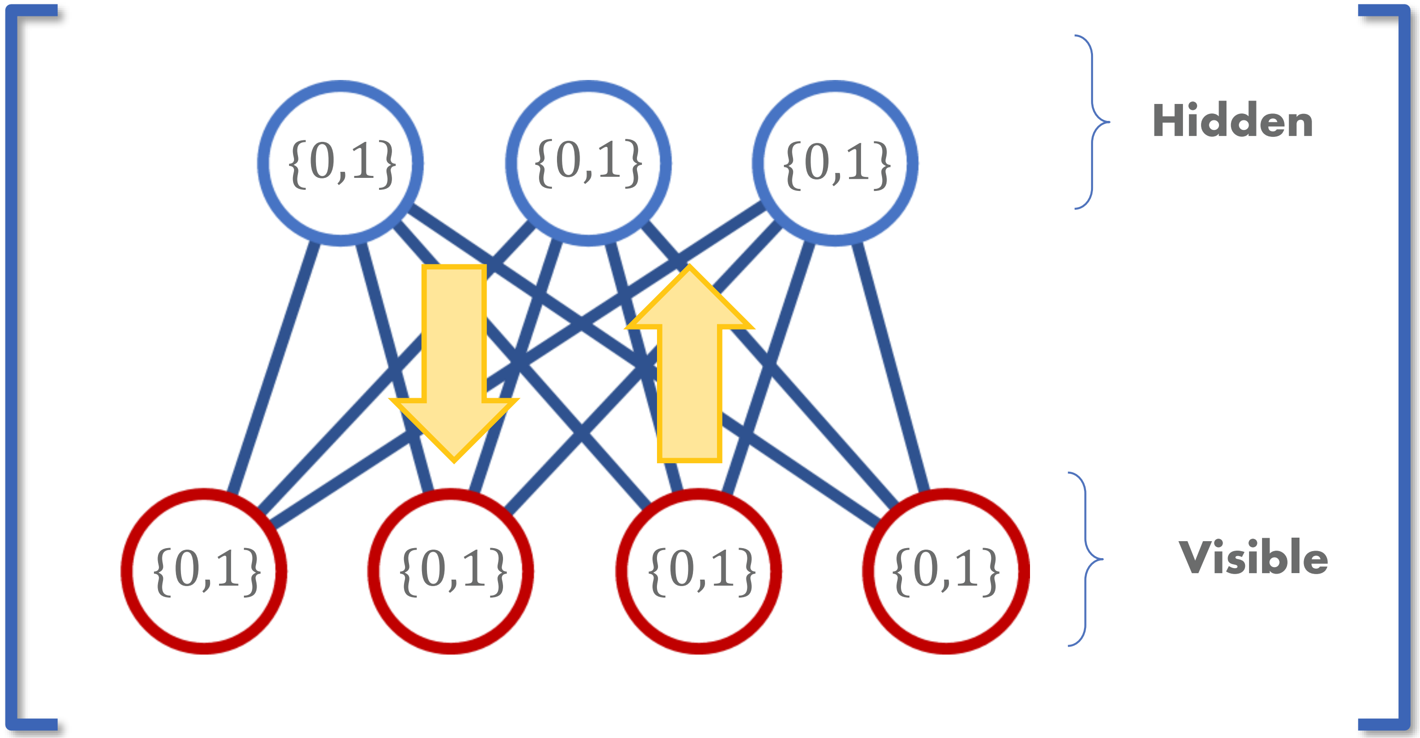
Perspective | Published: 24 June 2019
Restricted Boltzmann machines in quantum physics
Roger G. Melko, Giuseppe Carleo, Juan Carrasquilla & J. Ignacio Cirac
Nature Physics **15**, 887–892(2019) | Cite this article

the computational cost is high…

Classical RBM

A 2 layer Neural Network.

Units assume values in $\{0; 1\}$

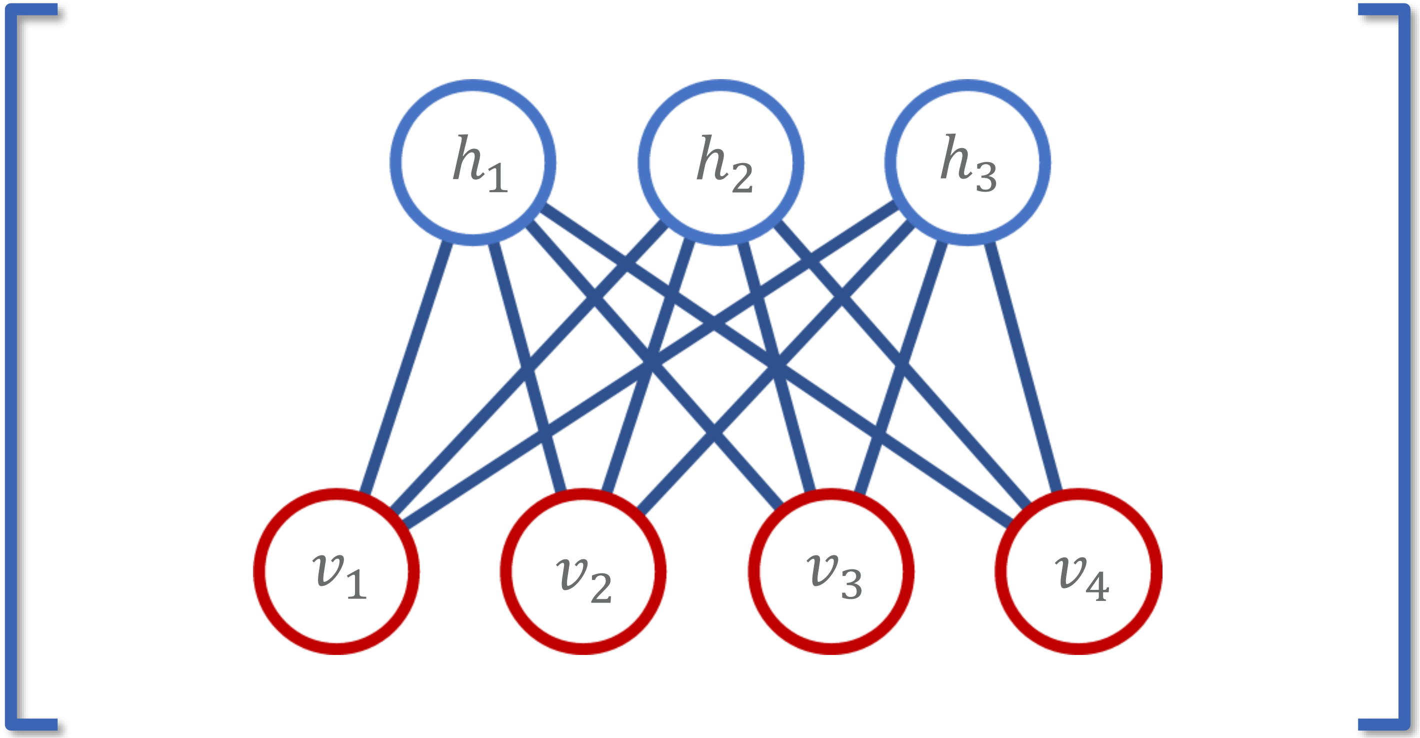


Latent variables

Input and output

Energy model

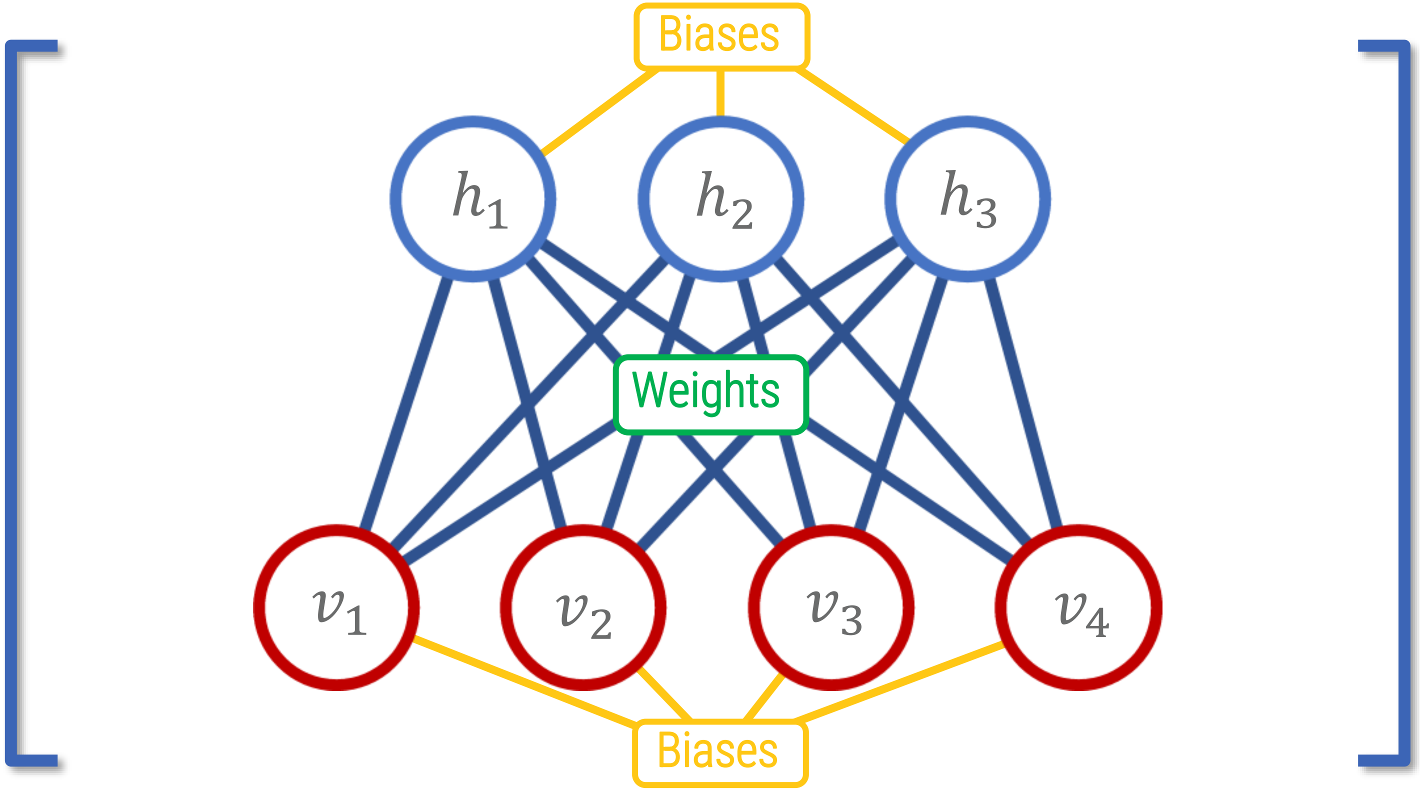
$$E(\mathbf{v},\mathbf{h}) = -\sum_{i\in\text{visible}} a_i v_i - \sum_{j\in\text{hidden}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$



Energy model

Biases — Weights

$$E(\mathbf{v},\mathbf{h}) = -\sum_{i\in\text{visible}} a_i v_i - \sum_{j\in\text{hidden}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$



Energy model

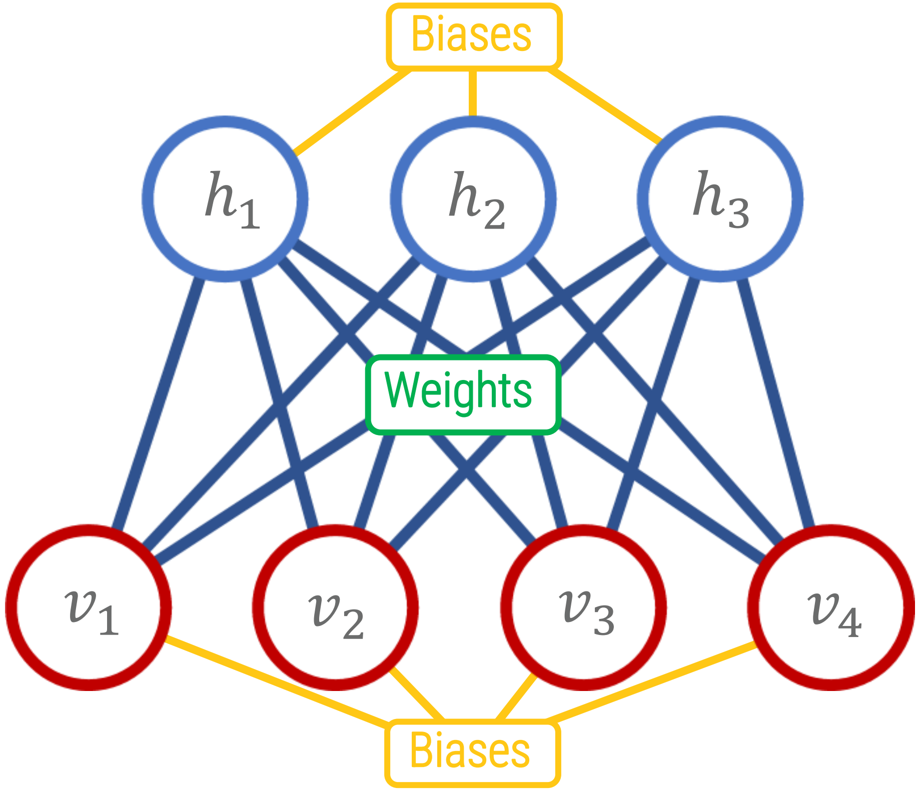
Biases / Weights

$$E(\mathbf{v},\mathbf{h}) = -\sum_{i\in\text{visible}} a_i v_i - \sum_{j\in\text{hidden}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

Boltzmann Distribution

$$p(\mathbf{v},\mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v},\mathbf{h})}$$

G. E. Hinton, **A practical guide to training restricted boltzmann machines**, in Neural networks: Tricks of the trade, Springer, 2012, pp. 599– 619



$$p(\mathbf{v}) = \frac{1}{Z}\sum_{\mathbf{h}} e^{-E(\mathbf{v},\mathbf{h})}$$

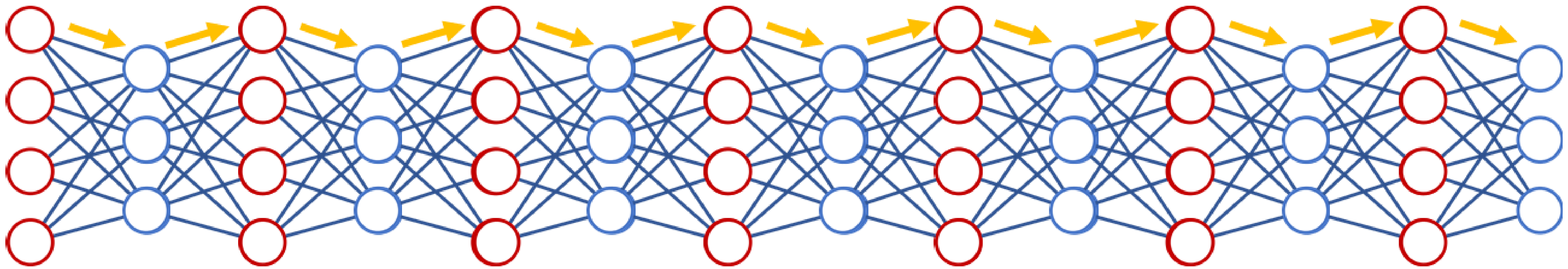
Weights and biases must be trained

Sampling from the RBM

If I have	I can make the update
The state of the **visible** units	$p(h_j = 1 \| \boldsymbol{v}) = \sigma\left(b_j + \sum_i w_{ij} v_i\right)$
The state of the **hidden** units	$p(v_i = 1 \| \boldsymbol{h}) = \sigma\left(a_i + \sum_j w_{ij} h_j\right)$

Reminder:

$$\sigma(x) = \frac{e^x}{1 + e^x}$$



Gibbs Sampling

Ackley, D. H., Hinton, G. E. & Sejnowski, T. J. A learning algorithm for boltzmann machines, Cognitive science 9, 147–169 (1985)

Hard part of training

$$\frac{1}{N_{\mathcal{D}}}\sum_{\mathbf{r}\in\mathcal{D}}\frac{\partial \log p(\mathbf{r})}{\partial w_{ij}} = \frac{1}{N_{\mathcal{D}}}\sum_{\mathbf{r}\in\mathcal{D}}\left(\frac{\sum_{\{\mathbf{h}\}} r_i h_j e^{-E(\{\mathbf{h}\},\mathbf{r})}}{\sum_{\{\mathbf{h}\}} e^{-E(\{\mathbf{h}\},\mathbf{r})}}\right) - \frac{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} v_i h_j e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}}{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}} \equiv$$

$$\equiv \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}$$

Hard part of training

$$\frac{1}{N_{\mathcal{D}}}\sum_{\mathbf{r}\in\mathcal{D}}\frac{\partial \log p(\mathbf{r})}{\partial w_{ij}} = \frac{1}{N_{\mathcal{D}}}\sum_{\mathbf{r}\in\mathcal{D}}\left(\frac{\sum_{\{\mathbf{h}\}} r_i h_j e^{-E(\{\mathbf{h}\},\mathbf{r})}}{\sum_{\{\mathbf{h}\}} e^{-E(\{\mathbf{h}\},\mathbf{r})}}\right) - \frac{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} v_i h_j e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}}{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}} \equiv$$

EASY

$$\equiv \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}$$

Hard part of training

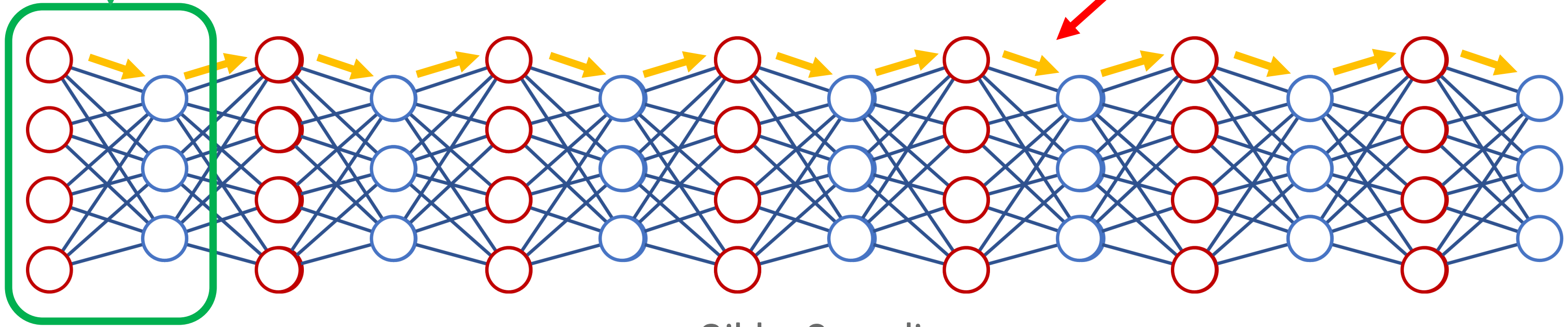
$$\frac{1}{N_\mathcal{D}}\sum_{\mathbf{r}\in\mathcal{D}}\frac{\partial \log p(\mathbf{r})}{\partial w_{ij}} = \frac{1}{N_\mathcal{D}}\sum_{\mathbf{r}\in\mathcal{D}}\left(\frac{\sum_{\{\mathbf{h}\}} r_i h_j e^{-E(\{\mathbf{h}\},\mathbf{r})}}{\sum_{\{\mathbf{h}\}} e^{-E(\{\mathbf{h}\},\mathbf{r})}}\right) - \frac{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} v_i h_j e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}}{\sum_{\{\mathbf{v}\},\{\mathbf{h}\}} e^{-E(\{\mathbf{v}\},\{\mathbf{h}\})}} =$$

$$= \langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{model}$$

EASY

HARD

2^N elements



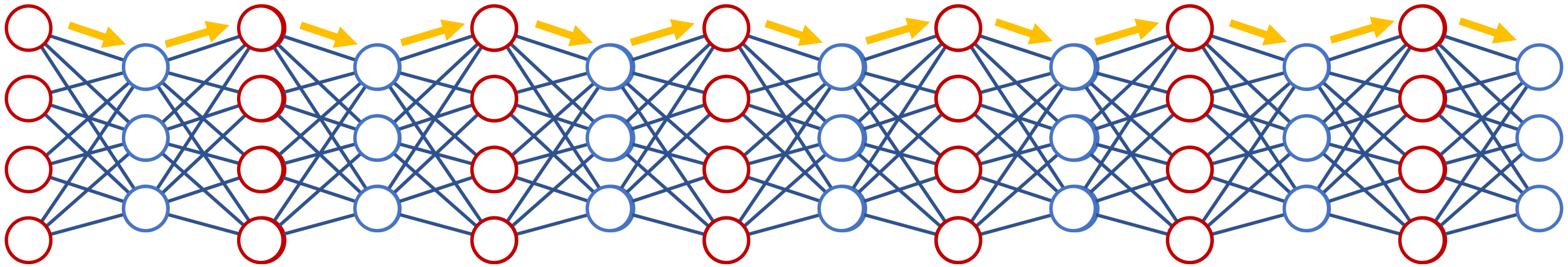
Gibbs Sampling

Ackley, Hinton, Sejnowski, *A learning algorithm for boltzmann machines*

Hard part of training

Problems:
Does not follow the gradient of any function; Works bad for non bipartite graphs

I. Sutskever and T. Tieleman, **On the convergence properties of contrastive divergence**, in Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 789–795.



Gibbs Sampling

CINECA QUANTUM COMPUTING LAB

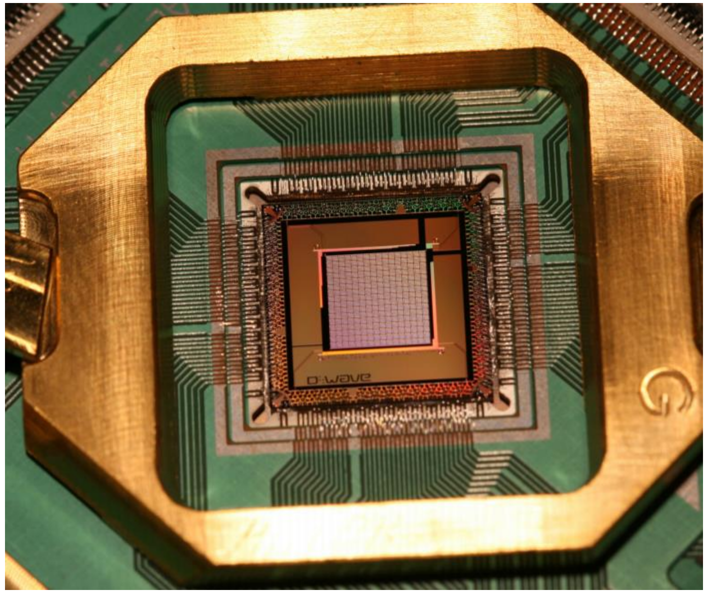
Looking for a better sampling method

On the Challenges of Physical Implementations of RBMs

V. Domoulin et al. – 2014

Proposal of AQCs as physical implementations of RBMs

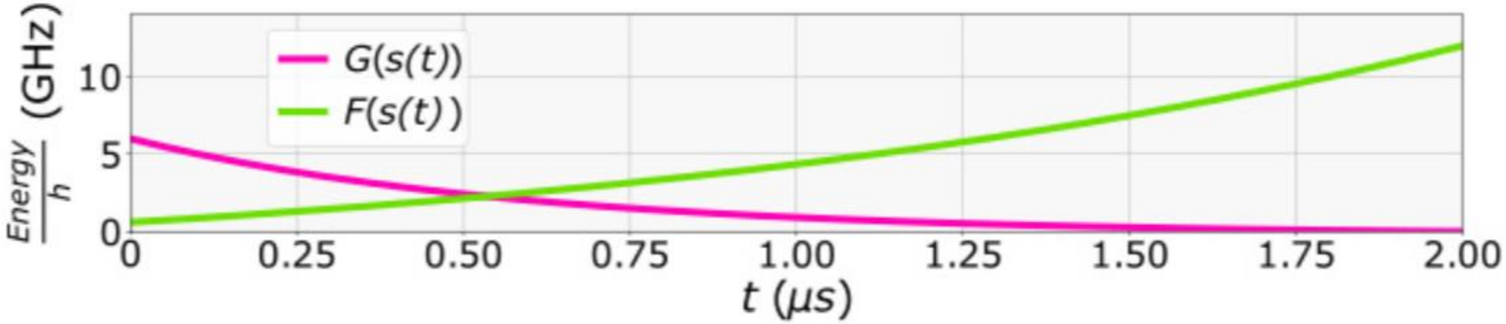
How does an AQC work?

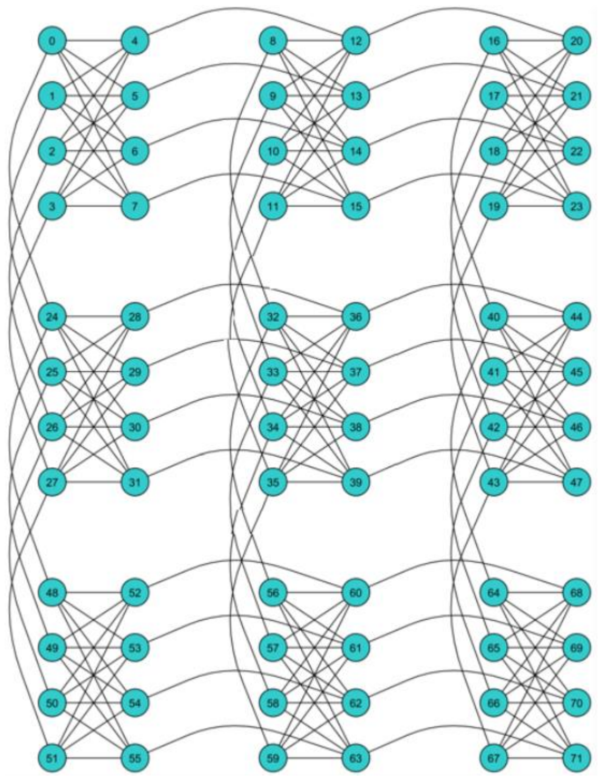


DWave 2000Q^{TM} System chip

AQC and RBM

$$H(t) = -F(s(t))\left(\sum_{i,j} J_{ij}\sigma_i^z\sigma_j^z + \sum_i A_i\sigma_i^z\right) - G(s(t))\sum_i \sigma_i^x$$

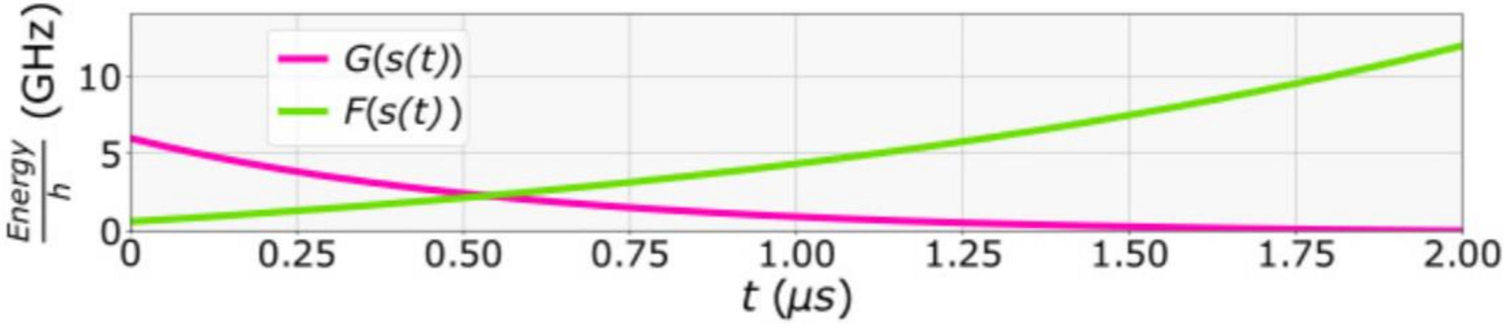


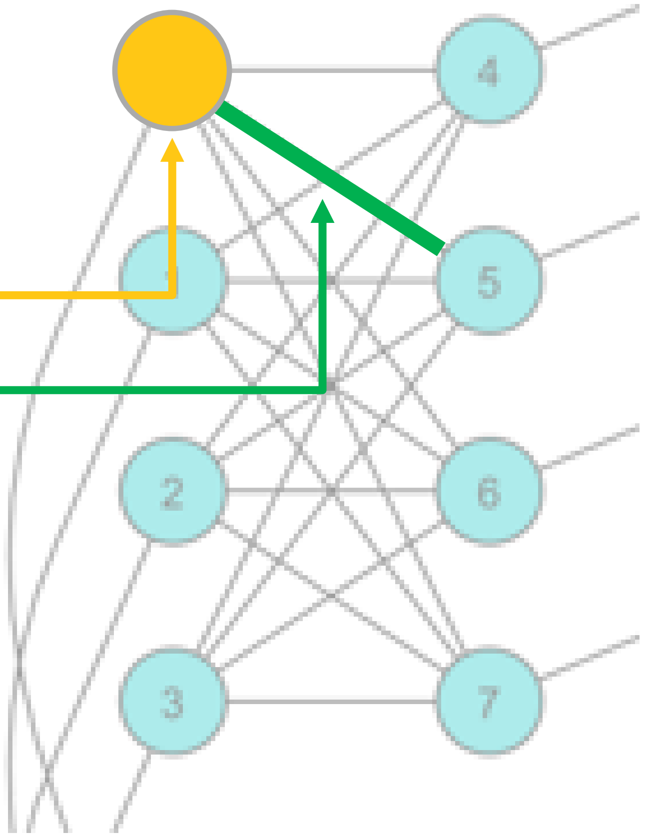


AQC and RBM

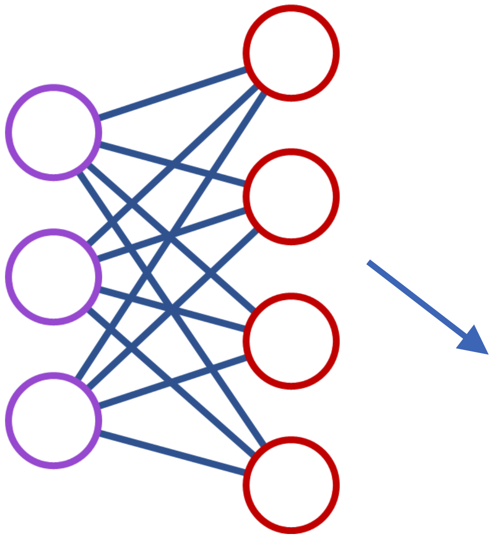
"Weights" "Biases"

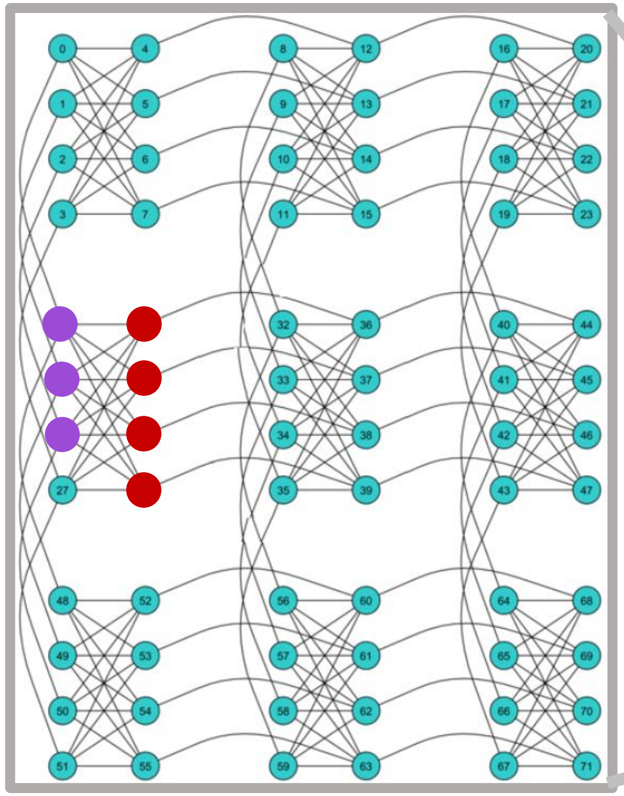
$$H(t) = -F(s(t))\left(\sum_{i,j} J_{ij}\sigma_i^z\sigma_j^z + \sum_i A_i\sigma_i^z\right) - G(s(t))\sum_i \sigma_i^x$$



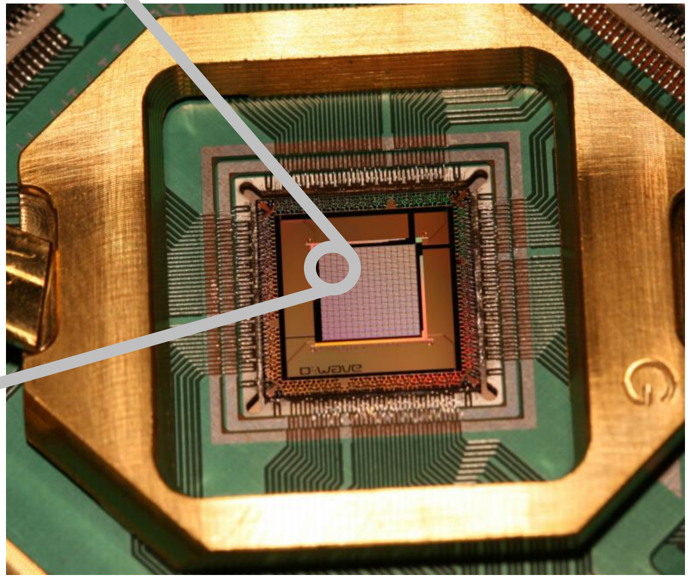


The quantum breakthrough



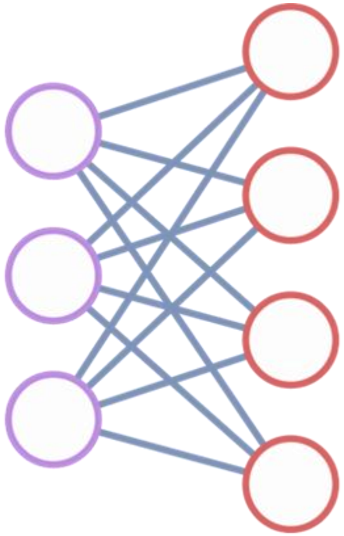


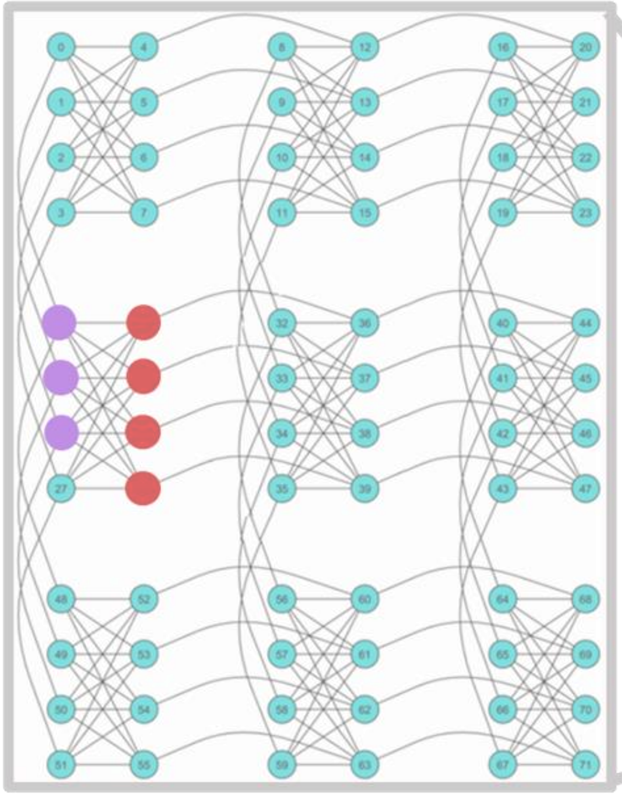
$$P(\mathbf{v},\mathbf{h}) = \frac{e^{-\frac{E_P(\mathbf{v},\mathbf{h})}{T_{\text{eff}}}}}{\sum_{\{\mathbf{v}'\},\{\mathbf{h}'\}} e^{-\frac{E_P(\mathbf{v}',\mathbf{h}')}{T_{\text{eff}}}}}$$



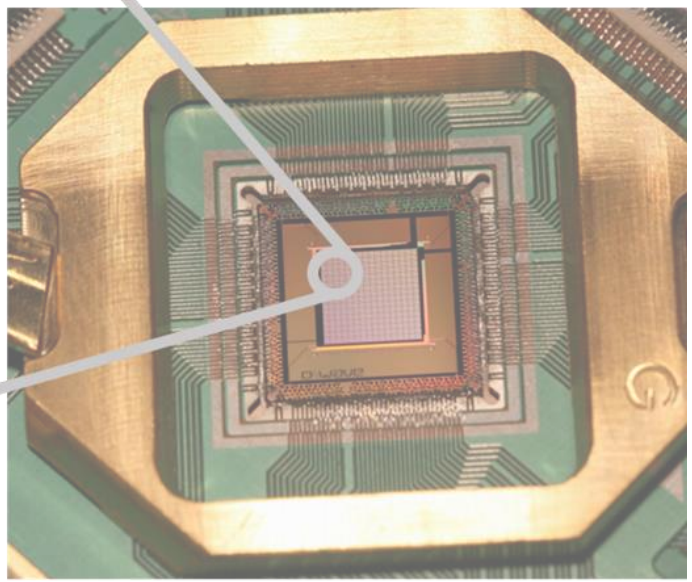
Each sample is produced in a single annealing cycle

The quantum breakthrough





$$P(\mathbf{v},\mathbf{h}) = \frac{e^{-\frac{E_P(\mathbf{v},\mathbf{h})}{T_{\text{eff}}}}}{\sum_{\{\mathbf{v}'\},\{\mathbf{h}'\}} e^{-\frac{E_P(\mathbf{v}',\mathbf{h}')}{T_{\text{eff}}}}}$$



Each sample is produced in a single annealing cycle

Development of Quantum RBMs on AQCs

- *On the Challenges of Physical Implementations of RBMs* – V. Domouli, I. J. Goodfellow, A. Courville, and Y. Bengio– 2014
- **Proposal of AQCs as physical implementations of RBMs**

- *Application of Quantum Annealing to Training of Deep Neural Networks* - H. Adachi, P. Henderson – 2015
- **D-Wave devices produce correctly distributed samples**

LOCKHEED MARTIN

- *Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning* – M. Benedetti,…,Perdomo-Ortiz– 2016
- **Temperature estimation advances RBMs implementation on AQCs**

NASA

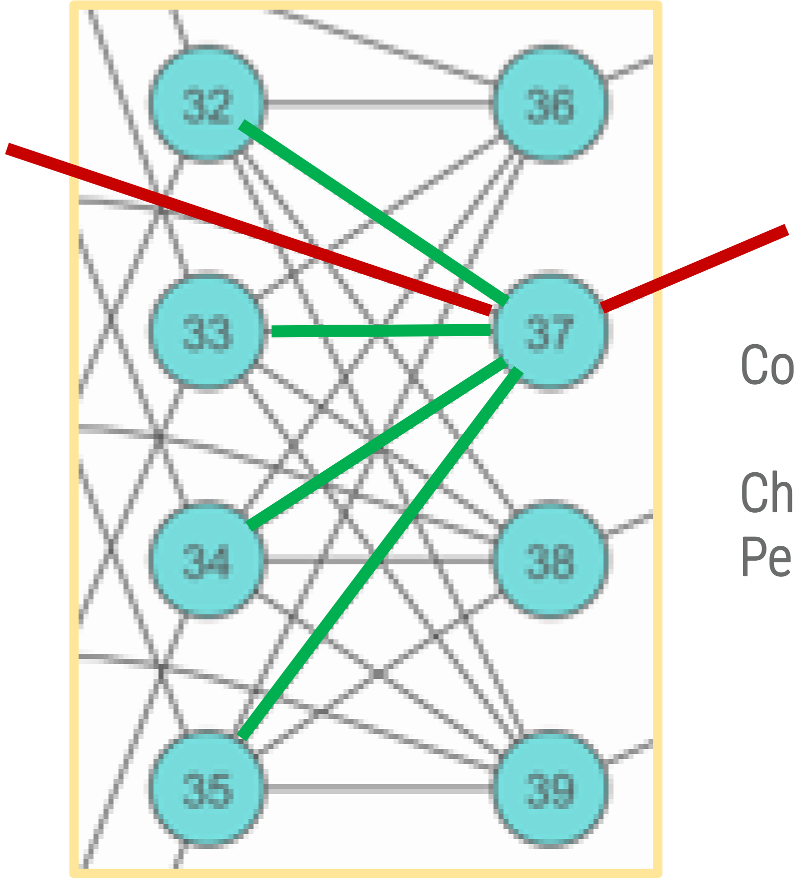
Limitations of the hardware

- Low impact:
  - Parameters noise
  - Constraints on the parameters

Limitations of the hardware

- Low impact:
  - Parameters noise
  - Constraints on the parameters

- High impact:
  - Low connectivity



Couplers:

Chimera 6
Pegasus 15

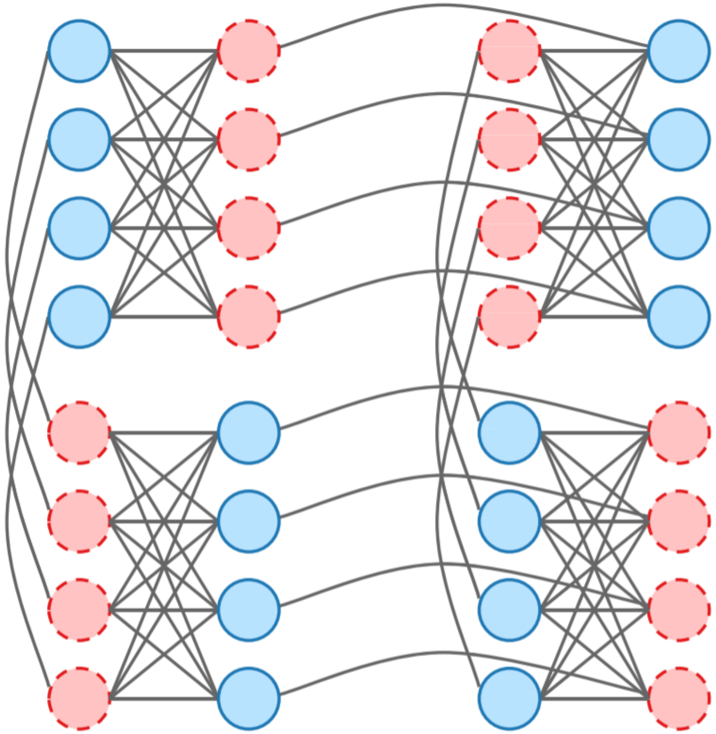
V. Dumoulin, I. J. Goodfellow, A. Courville, and Y. Bengio, On the challenges of physical implementations of rbms, in Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014

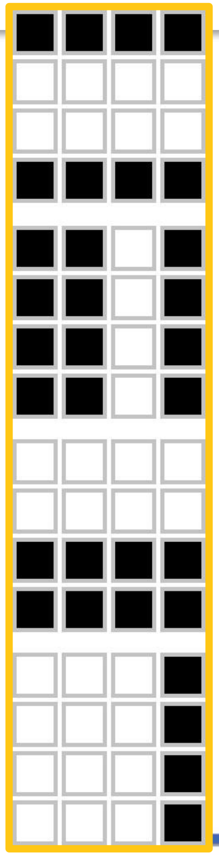
The actual implementation

PHYSICAL REVIEW A **94**, 022308 (2016)

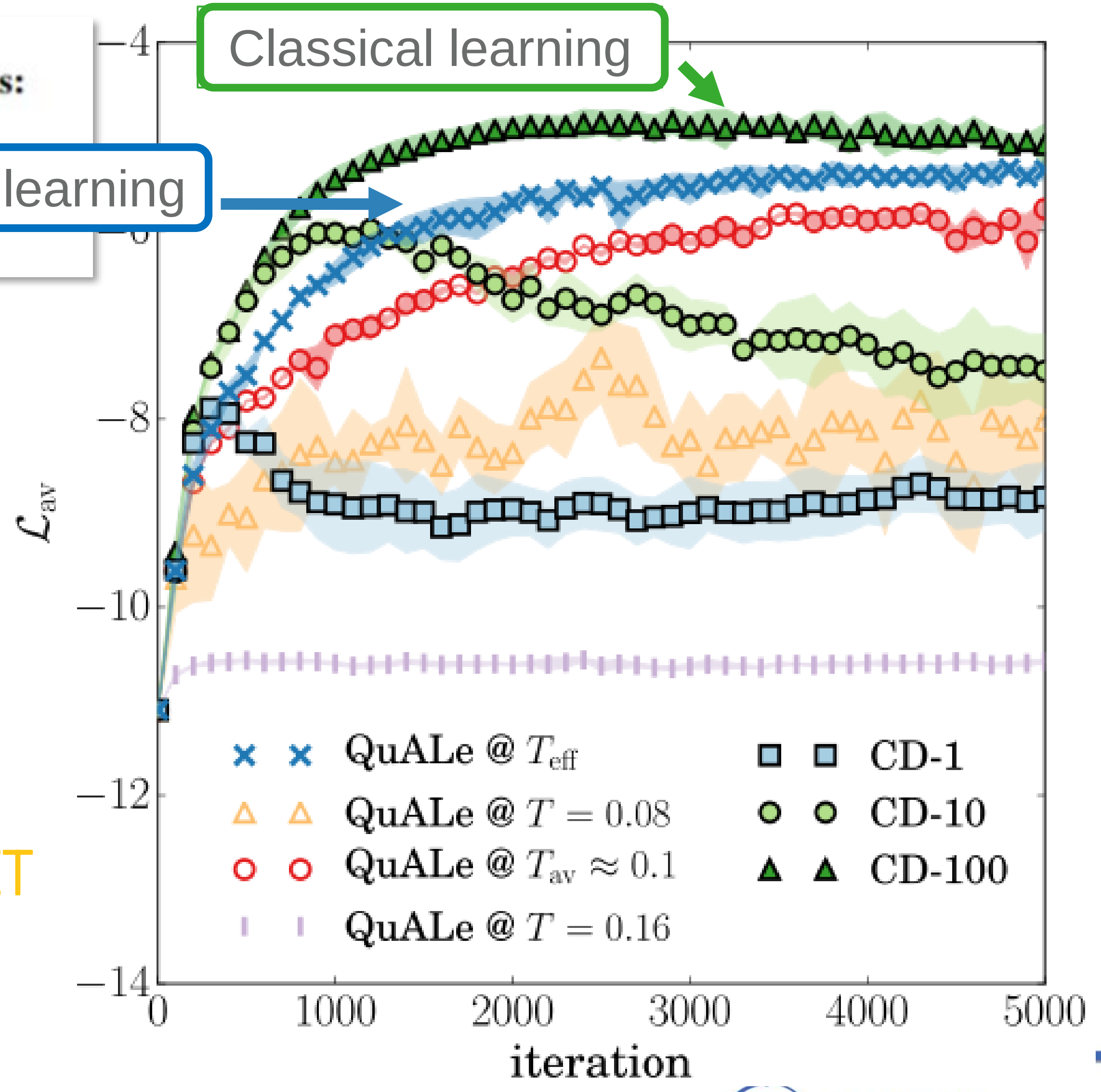
Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning

Marcello Benedetti
John Realpe-Gómez
Rupak Biswas
Alejandro Perdomo-Ortiz





DATASET

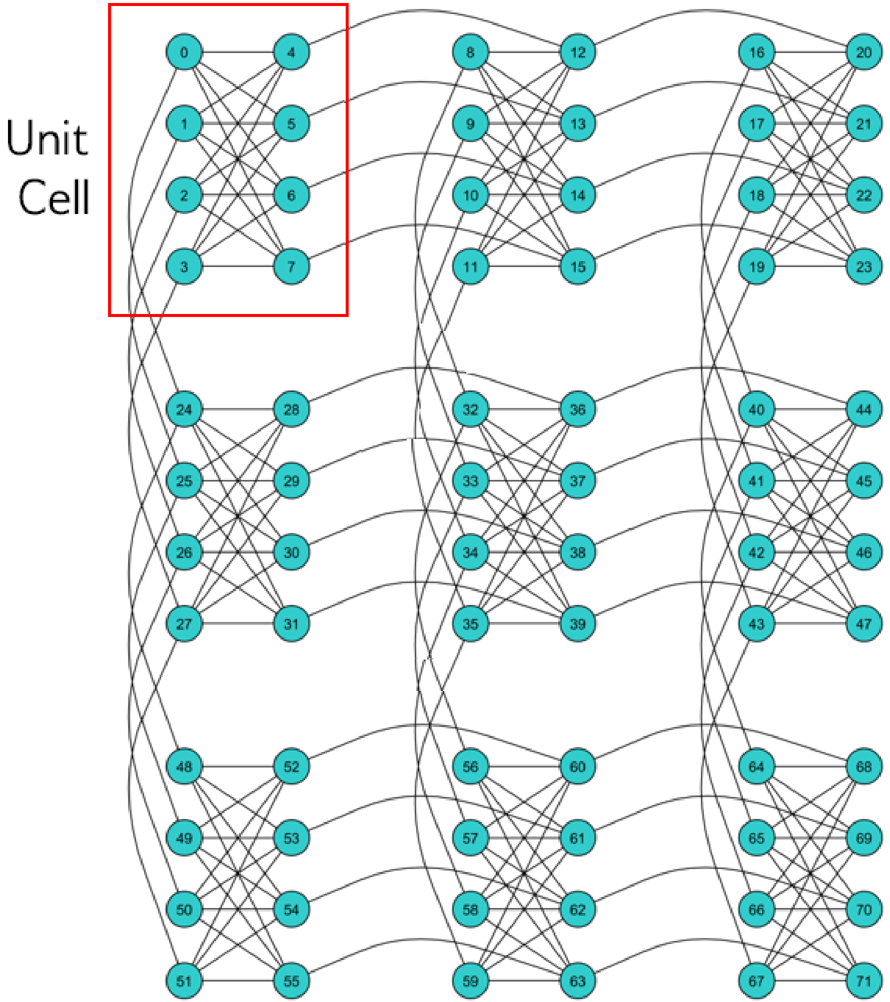


CINECA QUANTUM COMPUTING LAB

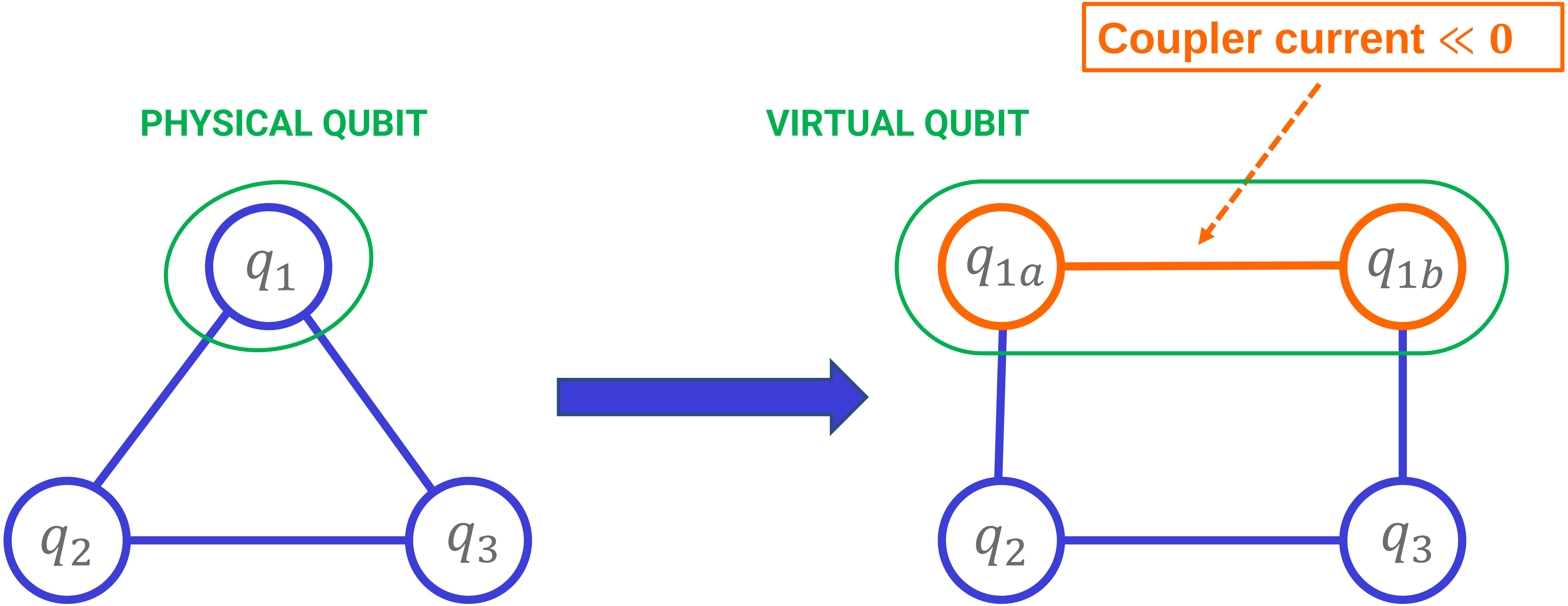
One of the main hardware limits

Many problems can not be implemented on the hardware graph

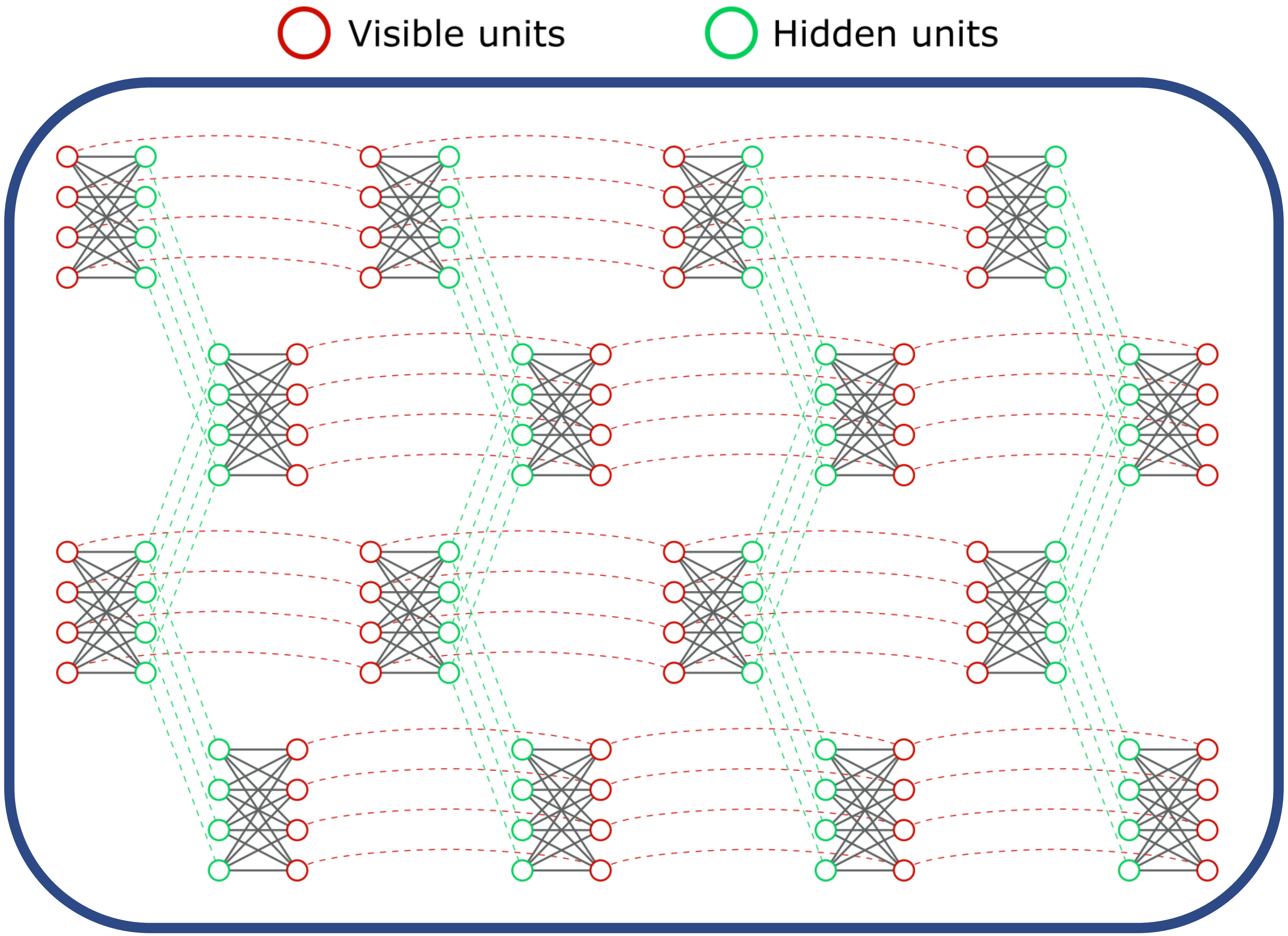
The problem can be overcome using *embbedding* techniques



Embedding techniques

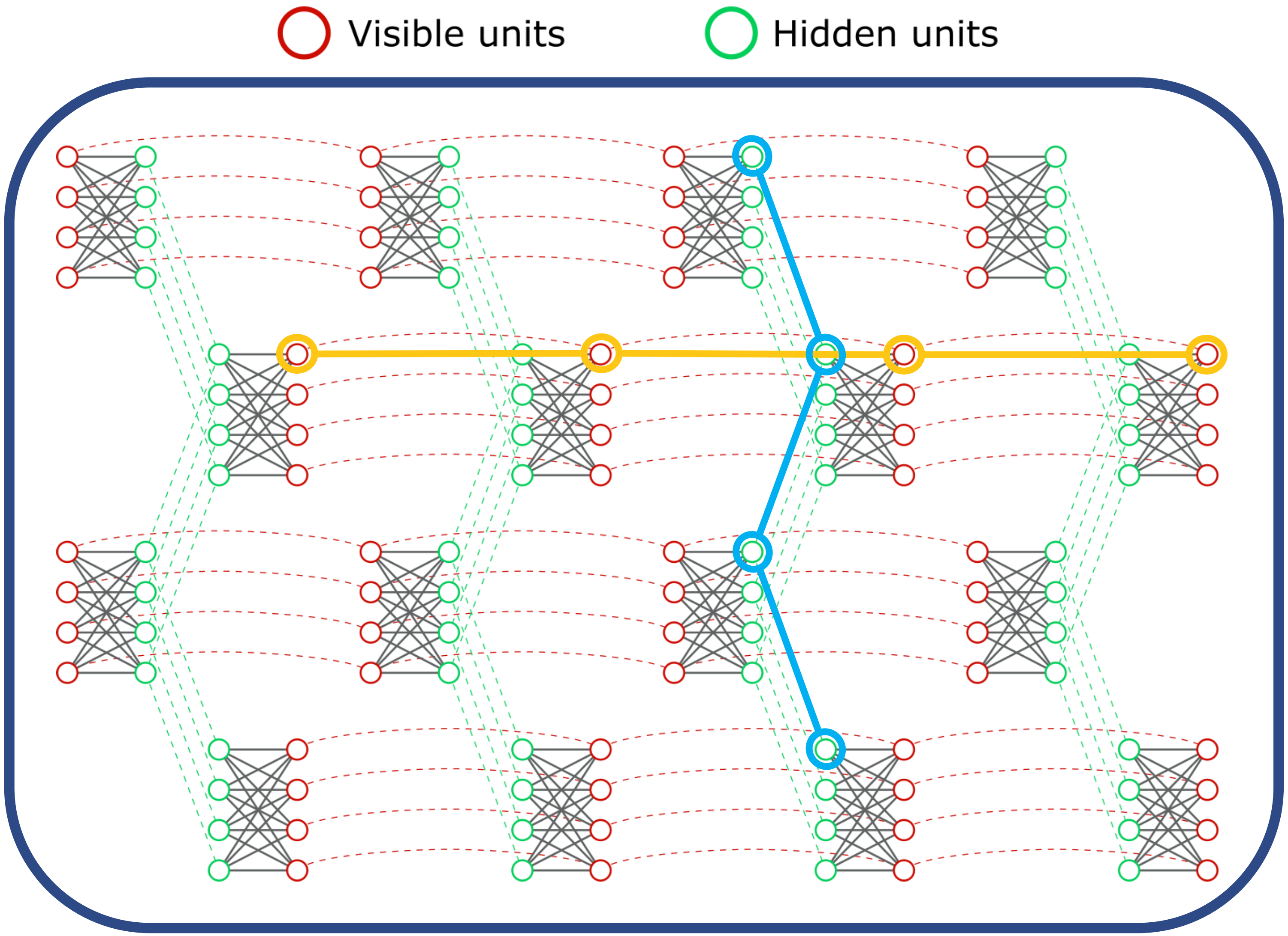


Embedding a complete RBM



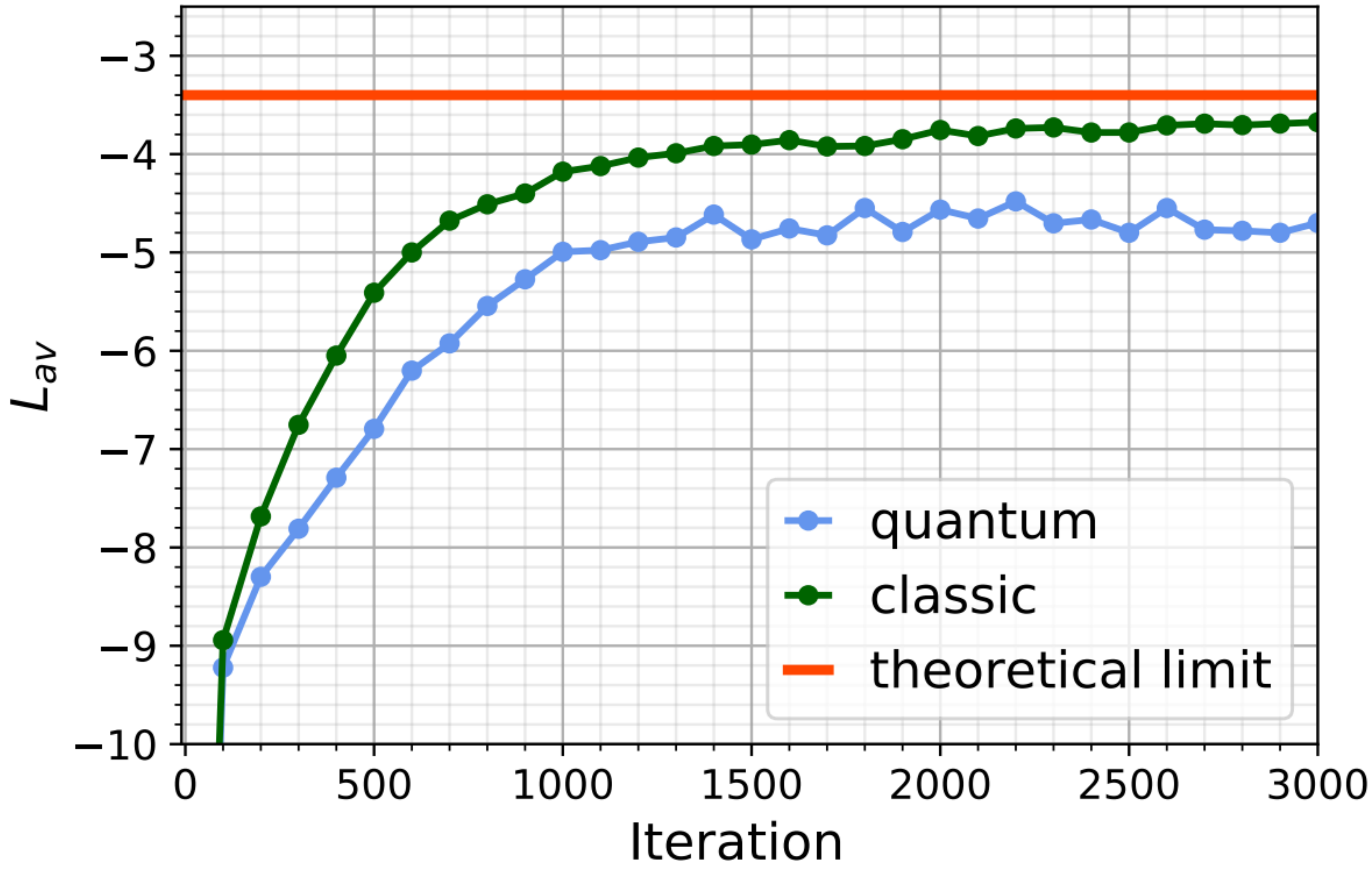
Rocutto, Lorenzo, Claudio Destri, and Enrico Prati. **Quantum Semantic Learning by Reverse Annealing of an Adiabatic Quantum Computer**. Advanced Quantum Technologies

Embedding a complete RBM

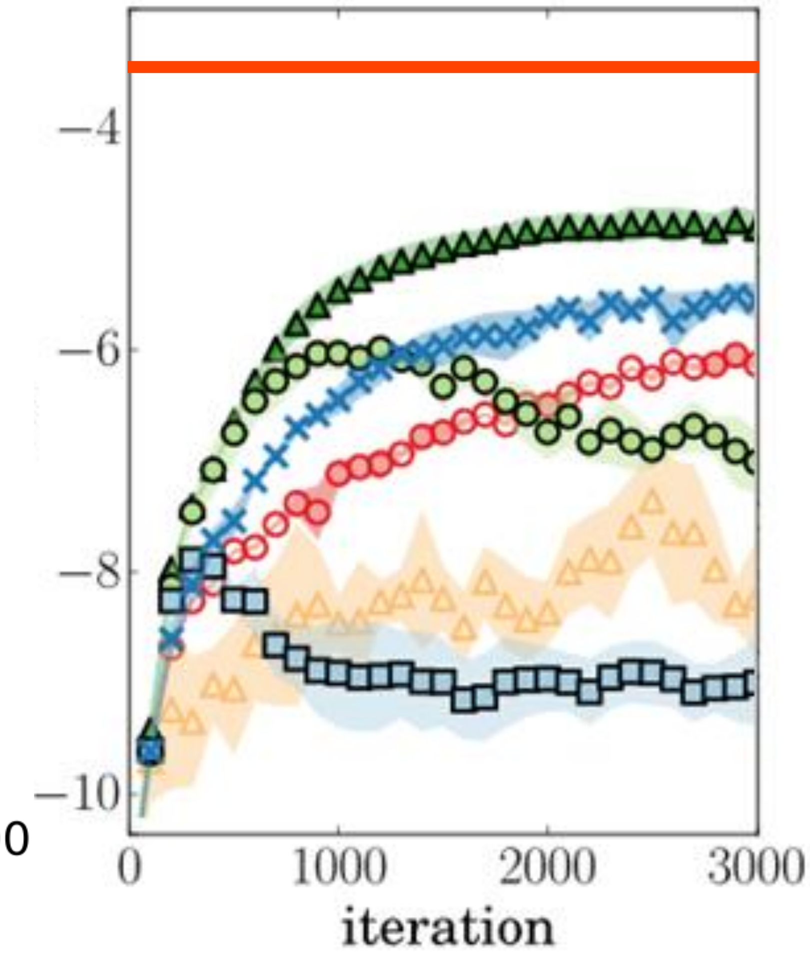


Rocutto, Lorenzo, Claudio Destri, and Enrico Prati. **Quantum Semantic Learning by Reverse Annealing of an Adiabatic Quantum Computer**. Advanced Quantum Technologies

Highly connected RBM

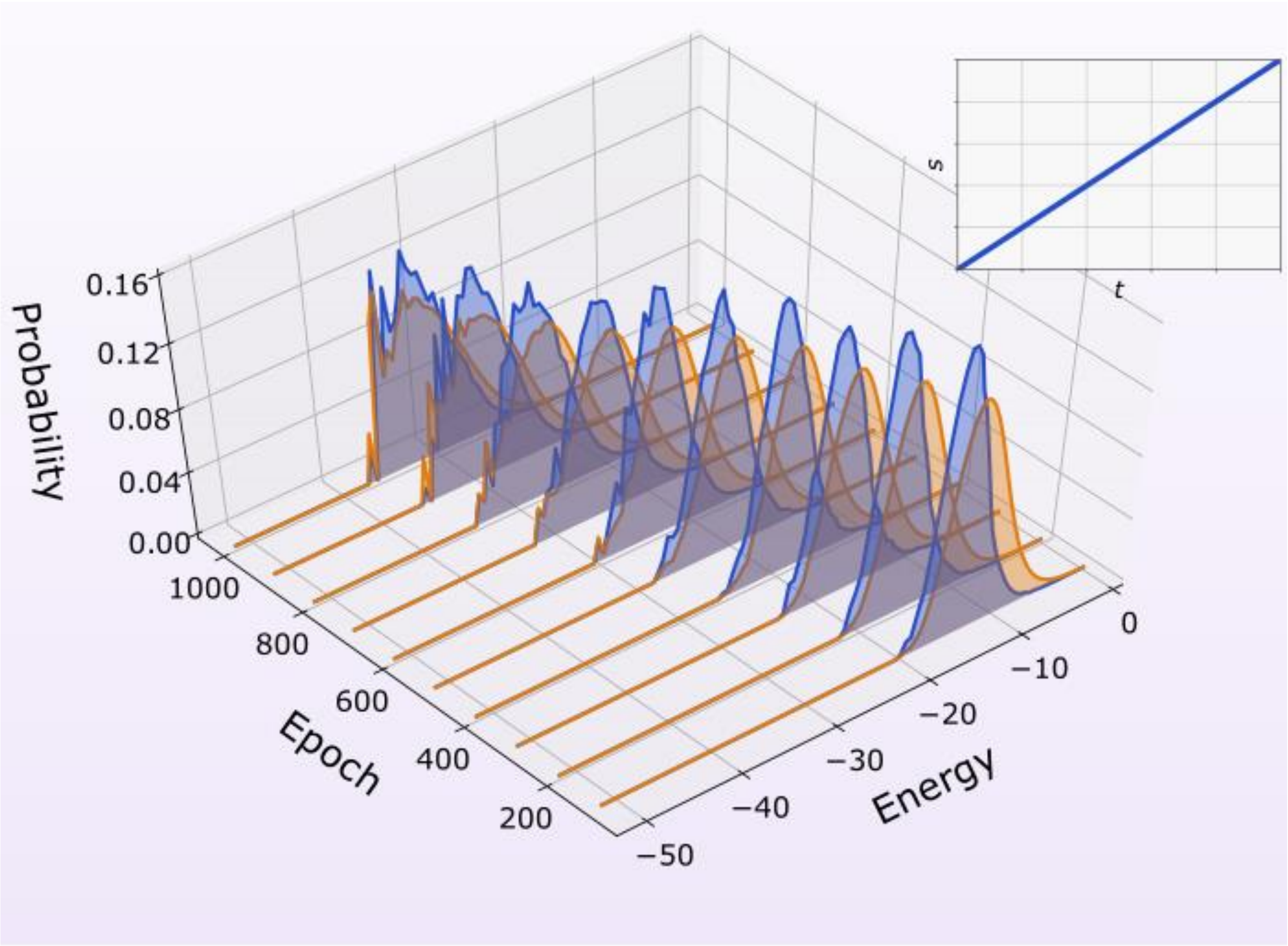


Complete 16x16 RBM



Sparse 16x16 RBM, REF

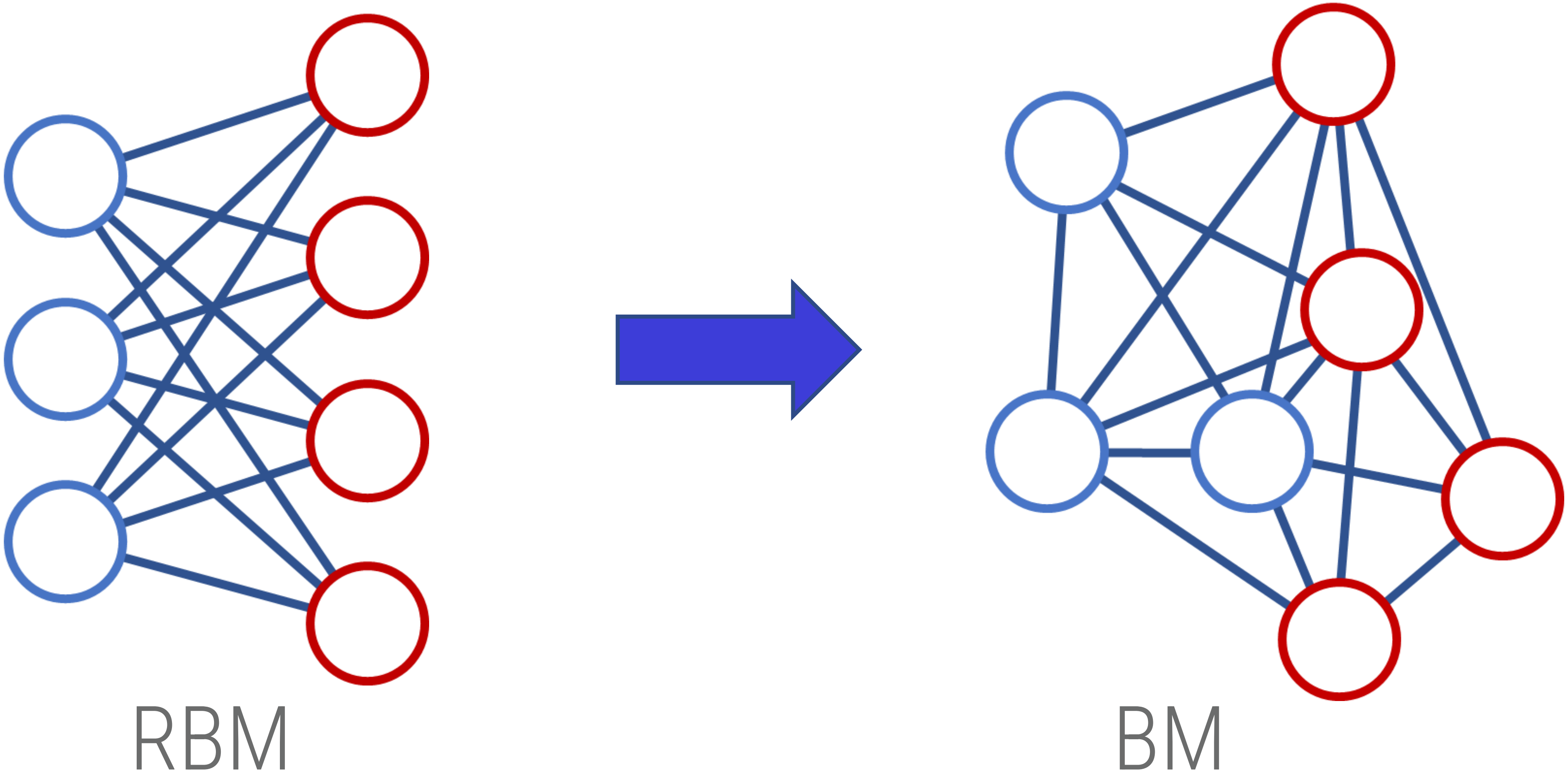
Visualizing Boltzmann Distribution



Orange:
Theoretical Boltzmann distribution of the RBM functional at T=1

Blue:
actual distribution extracted from the DWave annealer

An unfair Benchmark



Matrix factorization

Focus on its application in Machine Learning

Lorenzo Rocutto

CINECA QUANTUM COMPUTING LAB

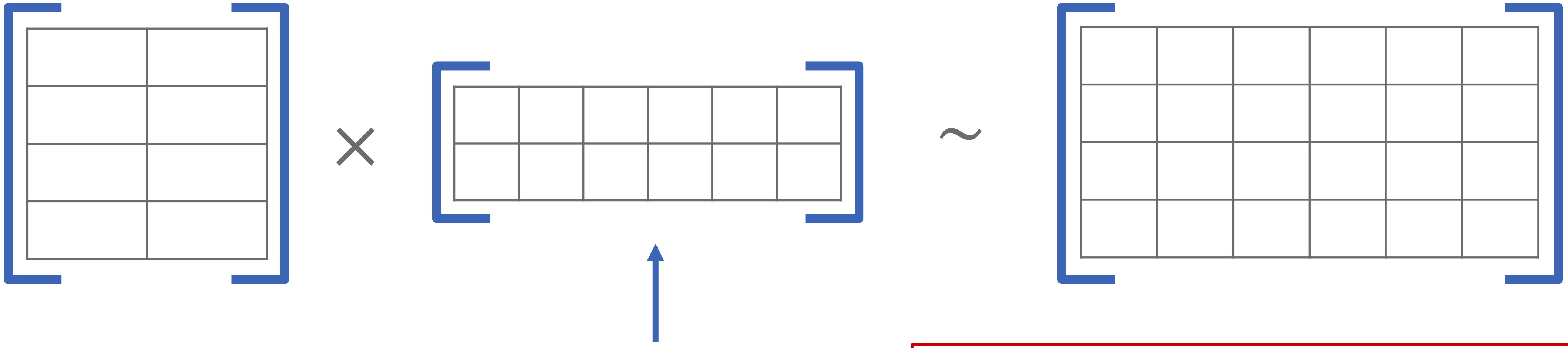
Dimensionality reduction

Learning is easier the smaller are the data in the dataset
We want to compress the data losing the least information possible

W: basis of the decomposition

H: compressed dataset

V: original dataset

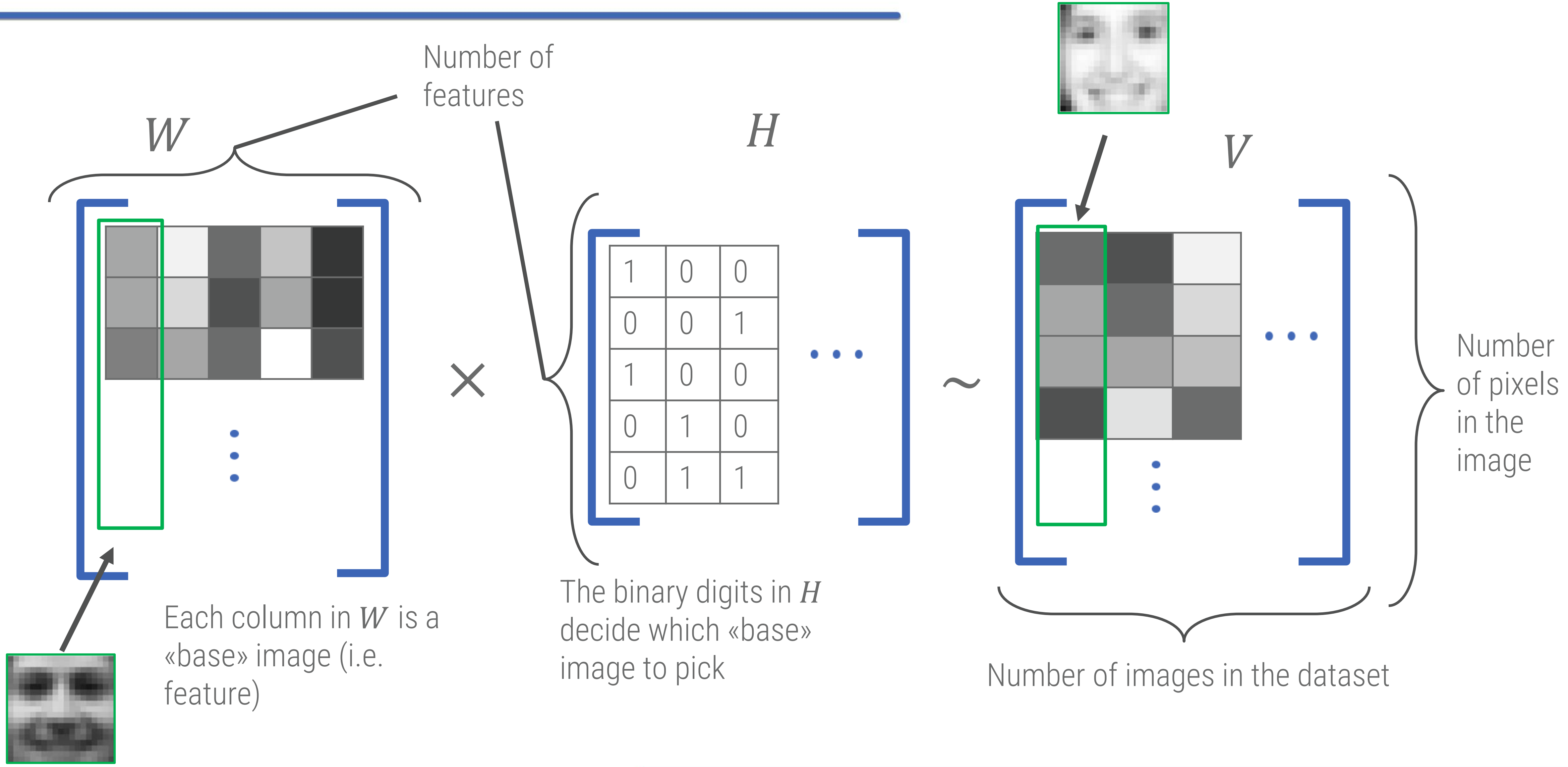


Constrained to binary values

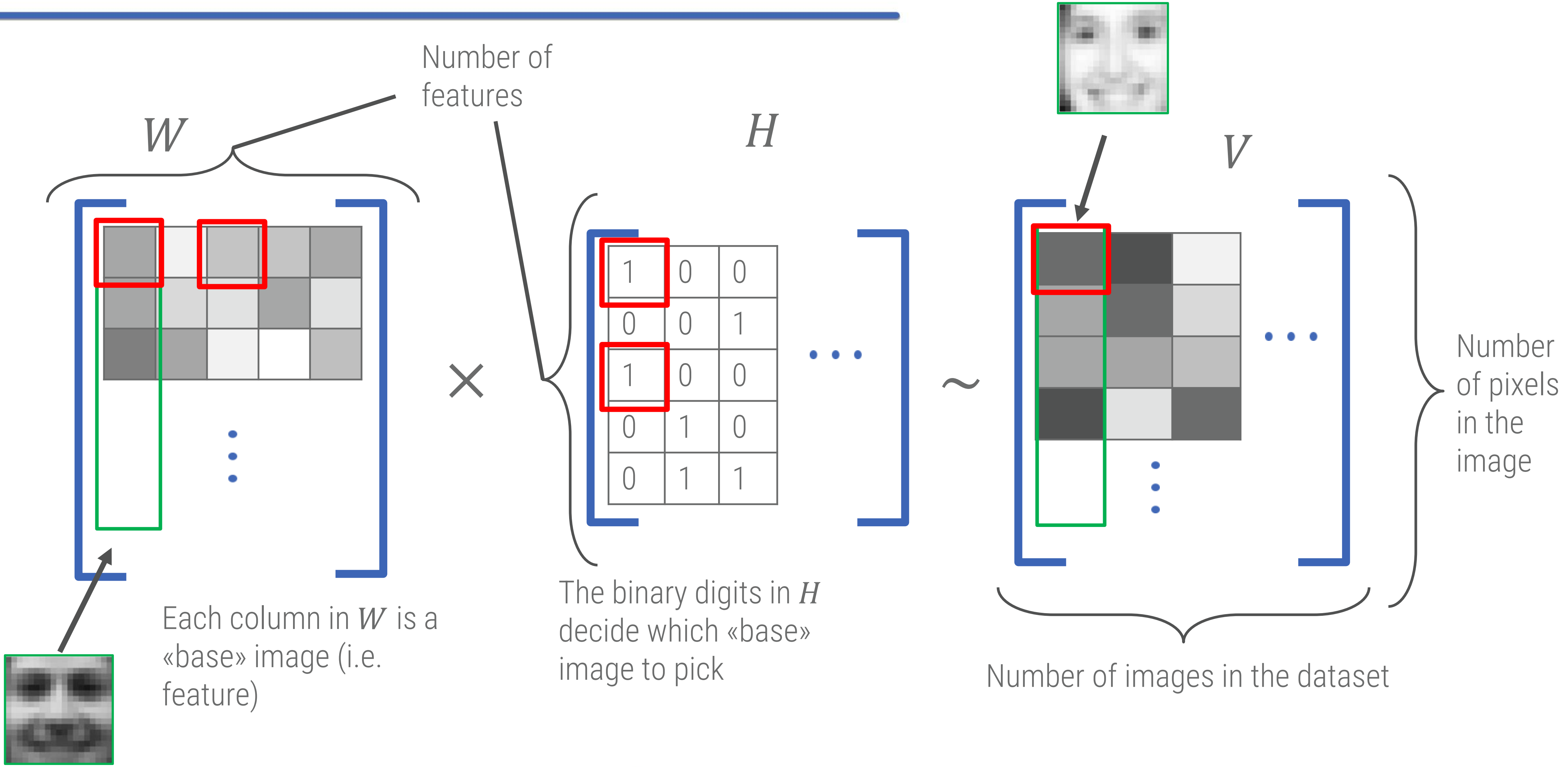
Minimize $error = V - (W \times H)$

Application in face recognition

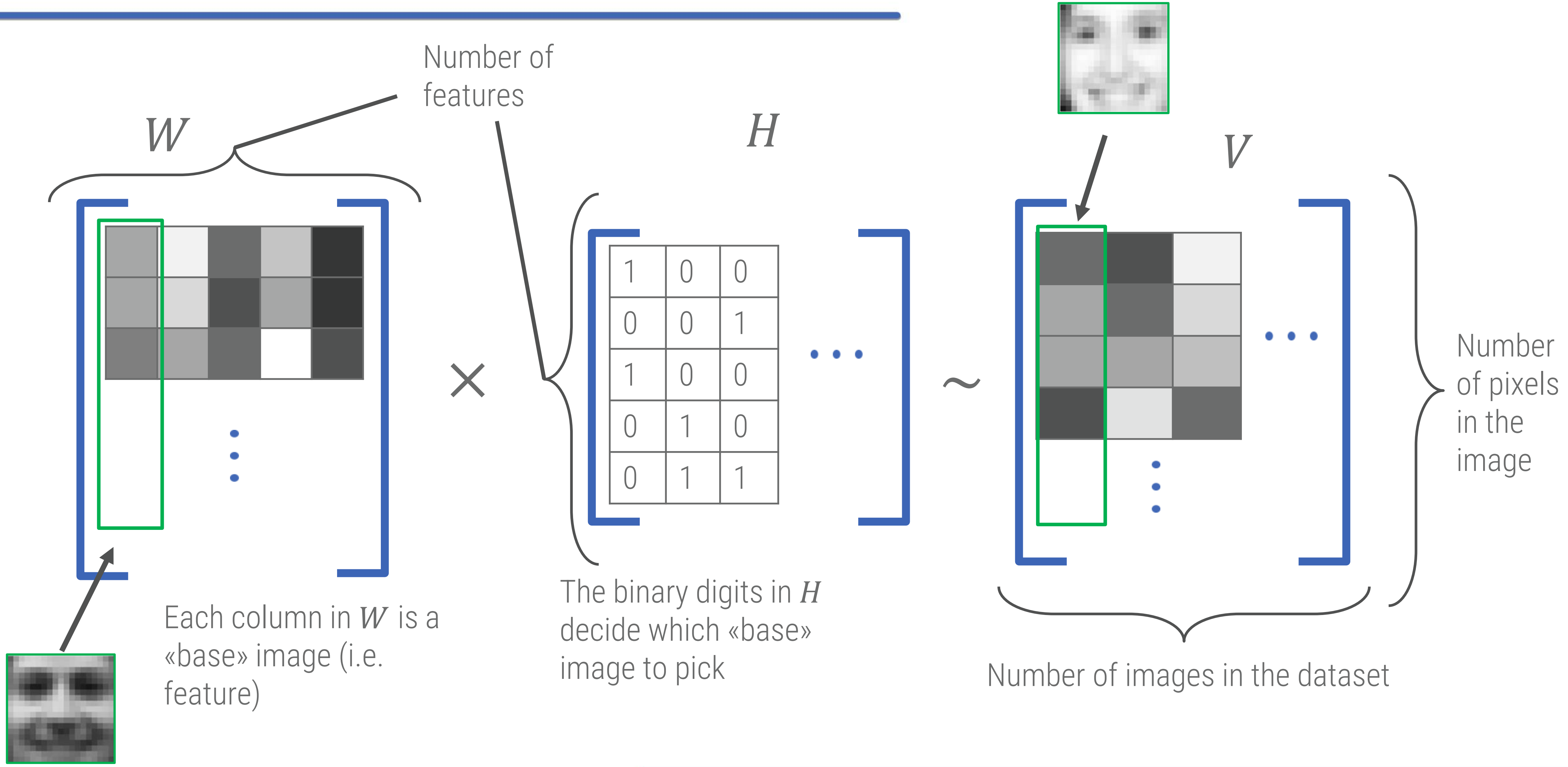
Application in face recognition



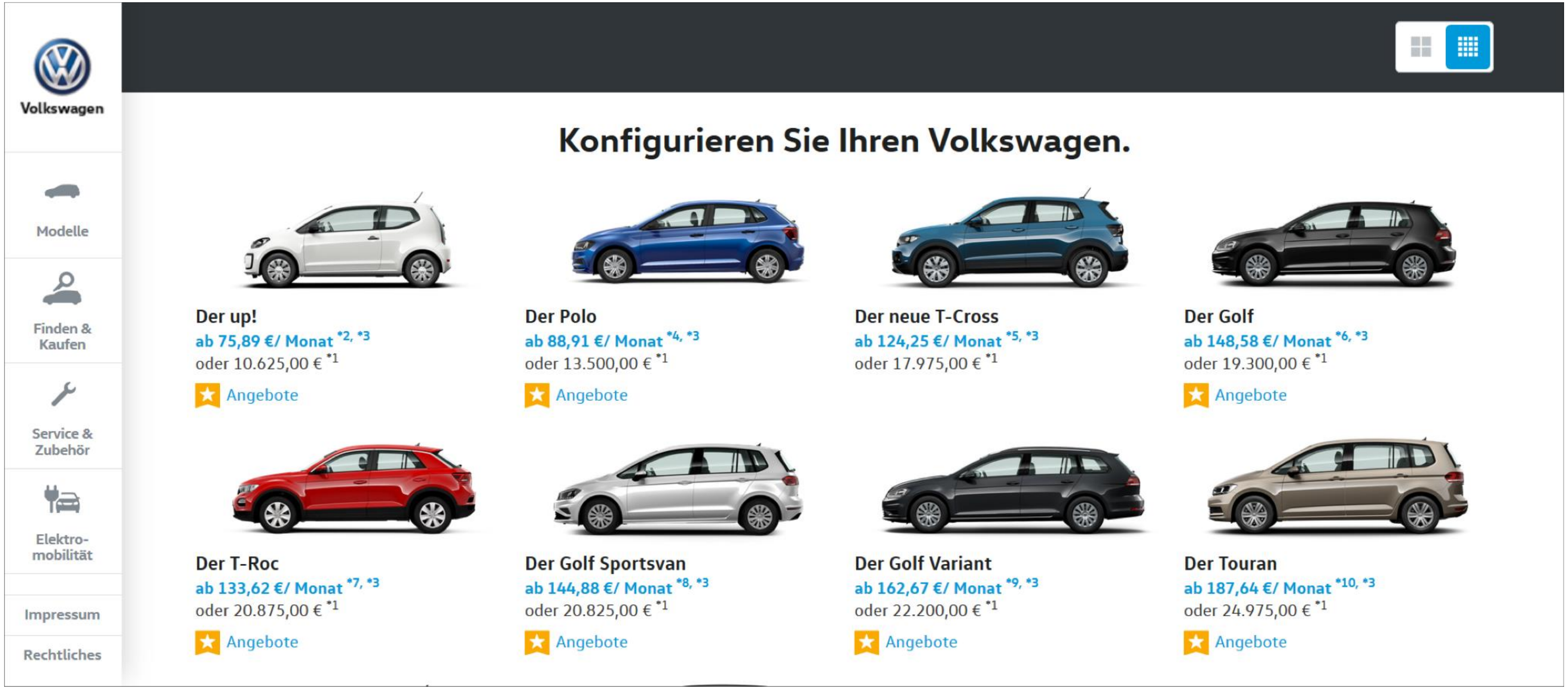
Application in face recognition



Application in face recognition



Application as a recommendation system



Extracted from Andrea Skolik contribution to Qubits 2019, to be found at https://www.dwavesys.com/qubits-europe-2019

- data of 47819 car purchases
- data includes color packages, accessories, …
- one-hot encoded categorical features
- turned them into “user ratings”
- 0: no purchase, 1: purchase

	model_5F11E2	model_5F11MX	...	interior_color_BC	SEAT FULL
0	1	1	...	0	0
1	0	1	...	1	1
2	0	0	...	0	1
3	0	0	...	0	1
4	0	0	...	1	0
5	1	0	...	1	1
6	0	0	...	0	0
7	1	0	...	0	0
8	0	1	...	1	1
9	0	0	...	1	1
10	1	0	...	1	1
11	0	0	...	0	0
12	1	1	...	0	0
13	1	0	...	1	1
14	1	0	...	0	0
15	1	0	...	0	1
16	0	0	...	1	0
17	0	0	...	1	1
18	0	1	...	1	0
19	1	0	...	1	1

Non finita

$$V = W \times H$$

m

n

	model_5F11E2	model_5F11MX	...	interior_color_BC	SEAT FULL
0	1	1	...	0	0
1	0	1	...	1	1
2	0	0	...	0	1
3	0	0	...	0	1
4	0	0	...	1	0
5	1	0	...	1	1
6	0	0	...	0	0
7	1	0	...	0	0
8	0	1	...	1	1
9	0	0	...	1	1
10	1	0	...	1	1
11	0	0	...	0	0
12	1	1	...	0	0
13	1	0	...	1	1
14	1	0	...	0	0
15	1	0	...	0	1
16	0	0	...	1	0
17	0	0	...	1	1
18	0	1	...	1	0
19	1	0	...	1	1

V

=

0.1	0.2	0.3	0.0	0.1
0.0	0.3	0.7	0.2	0.2
0.1	0.5	0.8	0.6	0.4

⋮

×

1	0	0
0	0	1
1	0	0
0	1	0
0	1	1

...

How to actually use a quantum annealer

COMMAND LINE INSTALL

Step 1: D-Wave Python Library

We recommend that you work in a virtual environment. If you are new to Python virtual environments, see our Getting Started guidelines. To set up the required dependencies, run the following command from your terminal:

```
pip install dwave-ocean-sdk && dwave config create
```

COPY

Step 2: Your API token

Follow the prompts and paste the following into your terminal.

••••••••••••••••••••••••••••••

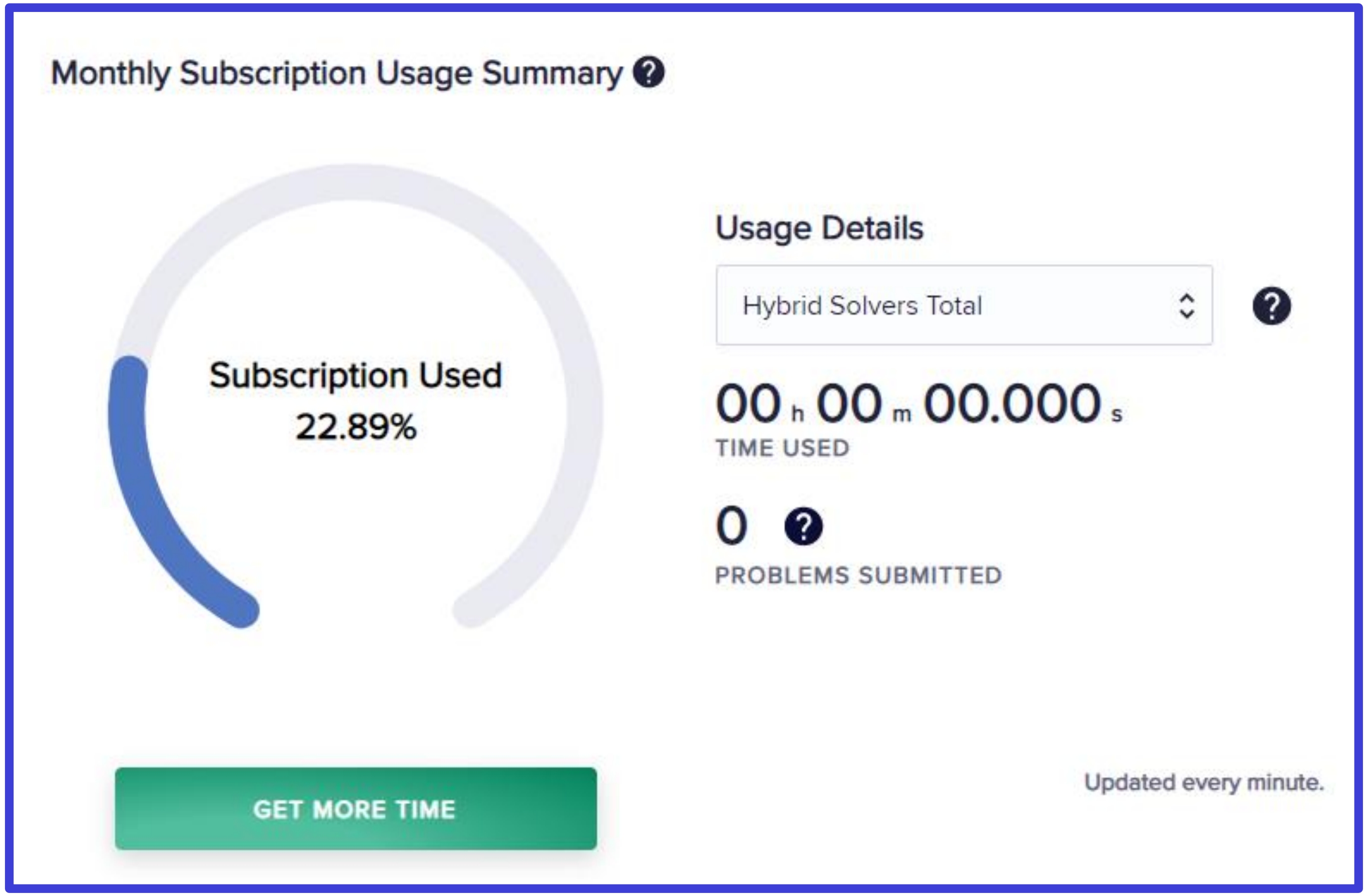
COPY

Step 3: Solve Problems

You should now be able to connect to a D-Wave system. Verify with the **dwave ping** command. Solve your own problems or begin with our end-to-end examples.

Leap plaftorm

Go to: https://cloud.dwavesys.com/leap/login/



Please note:

CINECA is currently distributing computational time to be used on the D-Wave Leap platform!

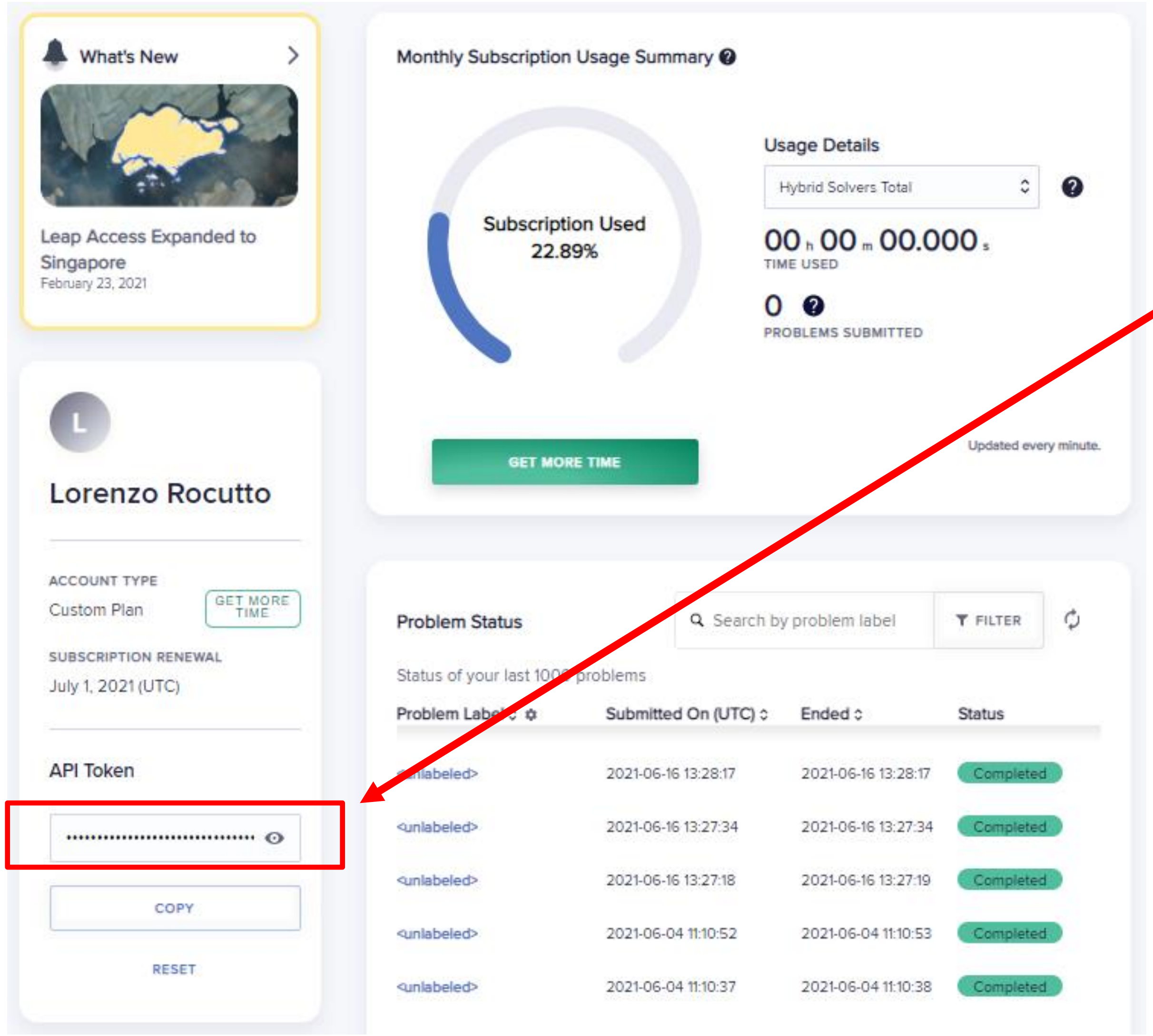
Install ocean sdk

```
(testenv) D:\>pip install dwave-ocean-sdk
```

Insert your token shown on you Leap Home Page

```
(testenv) D:\>dwave config create
Found existing configuration file: C:\Users\[redacted]\AppData\Local\dwavesystem\dwave\dwave.conf
Configuration file path [C:\Users\[redacted]\AppData\Local\dwavesystem\dwave\dwave.conf]:
Profile (create new or choose from: prod):
Input required, please try again.
Profile (create new or choose from: prod): prod
API endpoint URL [https://cloud.dwavesys.com/sapi]:
Authentication token [[redacted]]:
Default client class [qpu]:
Default solver [skip]:
Configuration saved.
```

Install ocean sdk



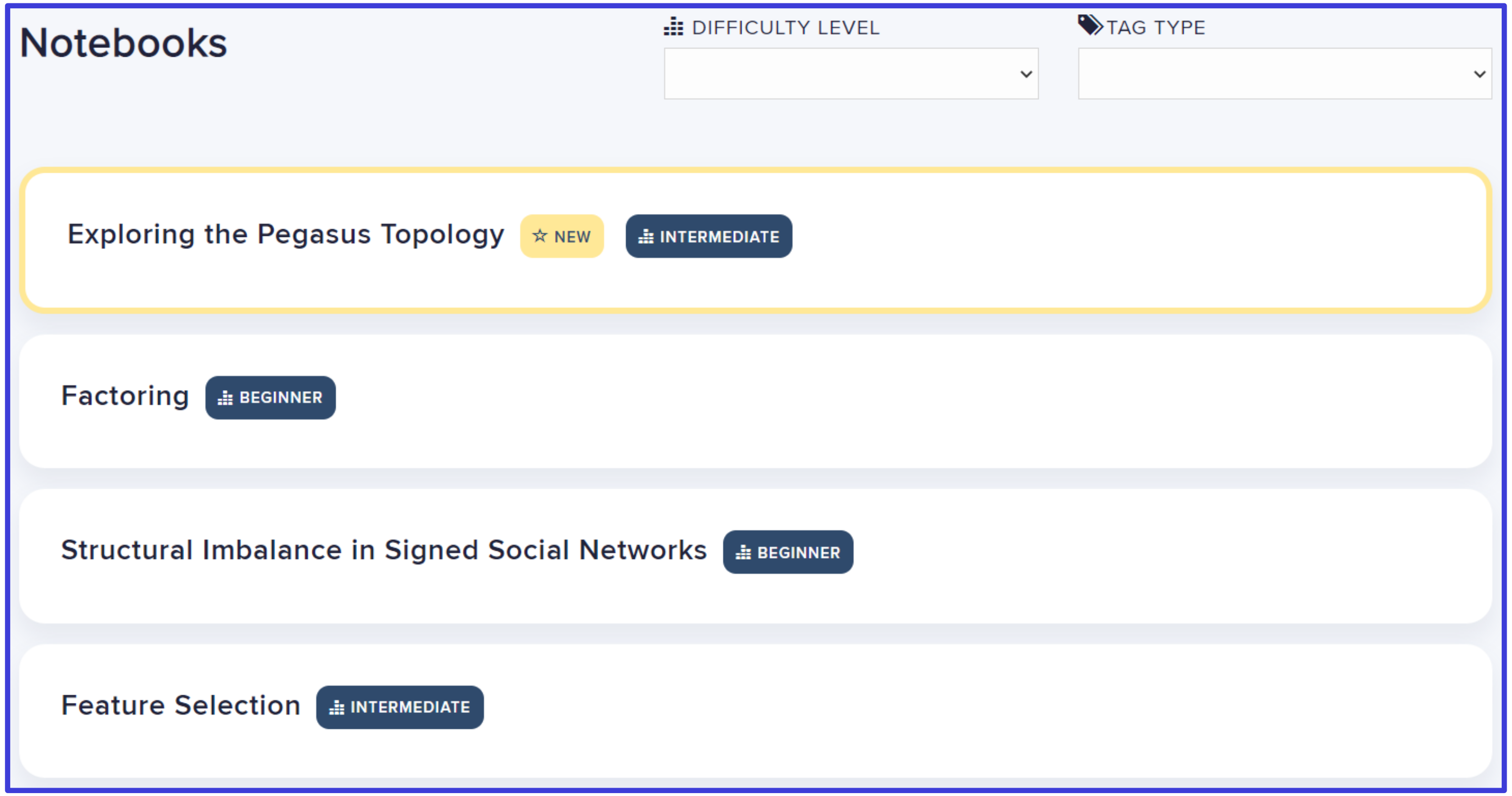
You find it right here

CINECA QUANTUM COMPUTING LAB

Coding

Go to: https://cloud.dwavesys.com/leap/resources#additional-resources

Then click on «jupyter notebooks»



I will show you a simple example that you can later run with your token