Abstract of my presentation:
Quantum computing promises an unprecedented ability to solve intractable problems by harnessing quantum mechanical effects such as tunneling, superposition, and entanglement. The Quantum Artificial Intelligence Laboratory (QuAIL) at NASA Ames Research Center is the space agency's primary facility for conducting research and development in quantum information sciences. QuAIL conducts fundamental research in quantum physics but also explores how best to exploit and apply this disruptive technology to enable NASA missions in aeronautics, Earth and space sciences, and space exploration. At the same time, machine learning has become a major focus in computer science and captured the imagination of the public as a panacea to myriad big data problems. In this talk, we will discuss how classical machine learning can take advantage of quantum computing to significantly improve its effectiveness. Although we illustrate this concept on a quantum annealer, other quantum platforms could be used as well. If explored fully and implemented efficiently, quantum machine learning could greatly accelerate a wide range of tasks leading to new technologies and discoveries that will significantly change the way we solve real-world problems.

Quantum Machine Learning

Dr. Rupak Biswas

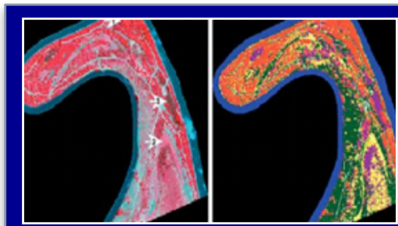
Director, Exploration Technology Directorate
Manager, NASA High End Computing Capability (HECC) Project

NASA Ames Research Center, Moffett Field, Calif., USA

SIAM Conference on Parallel Processing for Scientific Computing
Waseda University, Tokyo, Japan, 7–10 March 2018

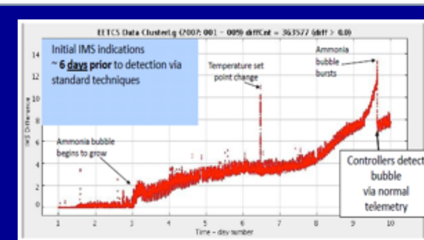
Why Quantum Computing at NASA

Key: Potential quantum speedup

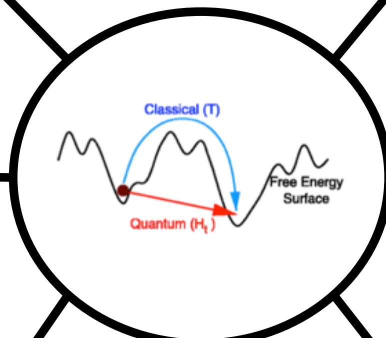


Data Analysis and Data Fusion

Anomaly Detection and Decision Making

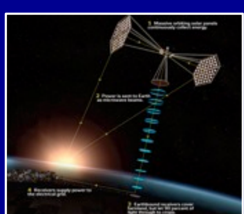


Air Traffic Management



V&V and optimal sensor placement

Mission Planning and Scheduling, and Coordination



Topologically aware Parallel Computing

Common Feature: Intractable (NP-hard / NP-complete) problems!

QuAIL: Quantum Artificial Intelligence Laboratory

Brief Development Timeline

2000–2011: Occasional NASA research on quantum computing, including seminal papers on adiabatic quantum computing & quantum annealing

Jan 2012: NASA organizes the First Quantum Future Technologies Conference attracting eminent researchers worldwide and participation from companies such as Google and D-Wave Systems

Nov 2012: NASA signs innovative 3-way Non-Reimbursable Space Act Agreement (NRSAA) with Google and USRA

Jan 2013: Site preparations begin at NASA Ames using Center investment funds for installation of D-Wave quantum annealer

Sept 2013: 512-qubit D-Wave 2 system comes on-line at Ames

June 2014: AFRL funding for research in quantum annealing

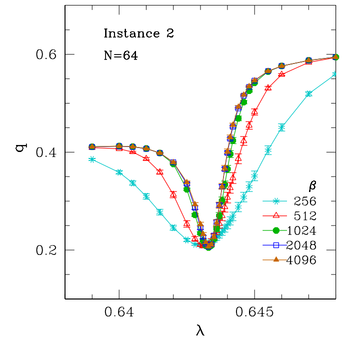
Aug 2014: IARPA funding for MIT-LL led QEO collaboration among NASA, TAMU, ETH-Z, UC Berkeley, and MIT

July 2015: Upgraded D-Wave 2X quantum annealer comes on-line with over 1000 qubits

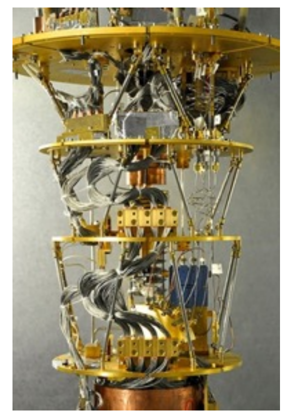
Feb 2017: NASA signs NRSAA with Rigetti Computing for collaborative work on their prototype universal quantum processor

April 2017: Latest upgrade underway for D-Wave system with over 2000 qubits

May 2017: NASA to lead T&E effort for IARPA QEO program



Phys. Rev. Lett. 104, 020502 (2010)




QuAIL team, Feb 2015

QuAIL team has published 40+ papers since 2012

NASA

Office of the Director of National Intelligence
IARPA
BE THE FUTURE

QUANTUM ENHANCED OPTIMIZATION

TARDEC

AIR FORCE RESEARCH LABORATORY

NASA QuAIL Team Focus

Long Term

- Determine the breadth and range of quantum computing applications
- Explore potential quantum algorithms and applications of relevance to NASA
- Evaluate, influence, and utilize emerging quantum hardware
  - Develop programming principles, compilation strategies, etc.
  - Characterize the hardware capabilities, noise, etc.
  - Evaluate and implement the most promising NASA applications
- Projections based on fundamental understanding of quantum physics

Ongoing Efforts

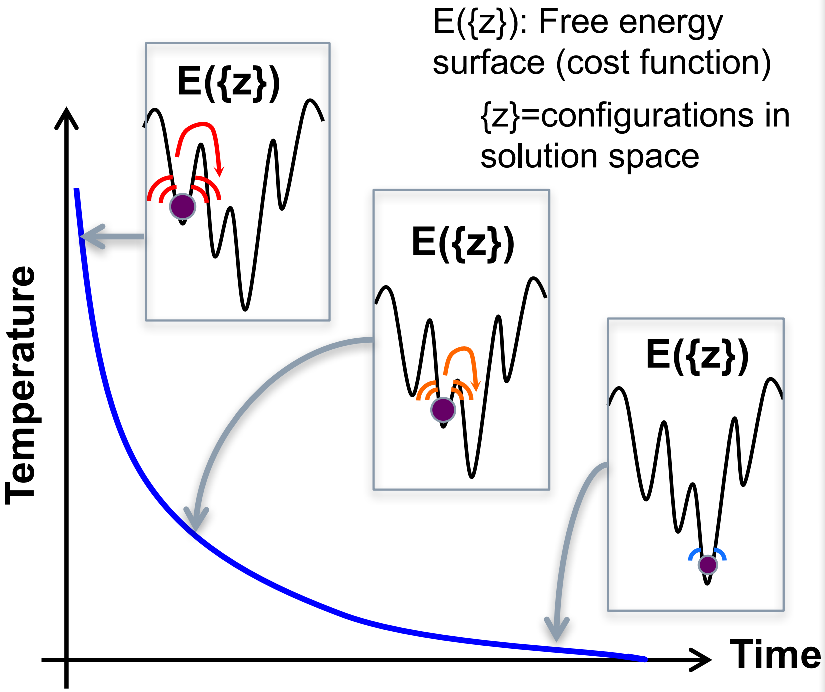
- Initial target: Quantum Annealing
  - Only significant quantum hardware available are quantum annealers from D-Wave Systems
  - Currently the most prominent quantum heuristic
  - Widely applicable to optimization problems, and sampling for ML
  - Early hardware used to develop intuitions and identify potential
- Near-term target: Emerging quantum computing hardware
  - Small universal quantum systems
  - Advanced quantum annealers
  - Alternative approaches to optimization, sampling for ML, etc.

Foundational Theory of Quantum Annealing

Simulated Annealing

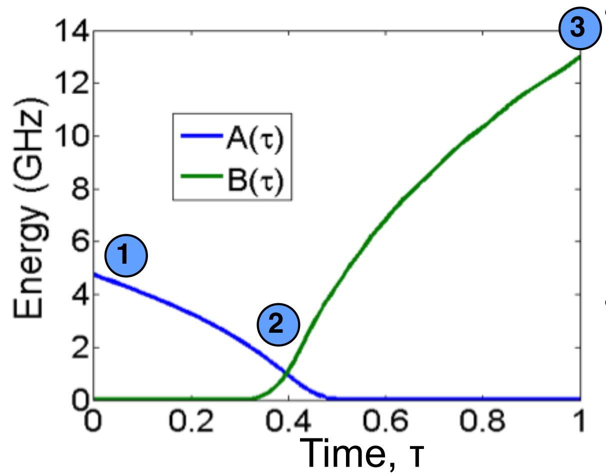
(Kirkpatrick et al., 1983)

- **Algorithm:** Start with high temperature; then, gradually reduce intensity of thermal fluctuations to obtain optimal configuration
- Transitions between states via jumping over barriers due to thermal fluctuations

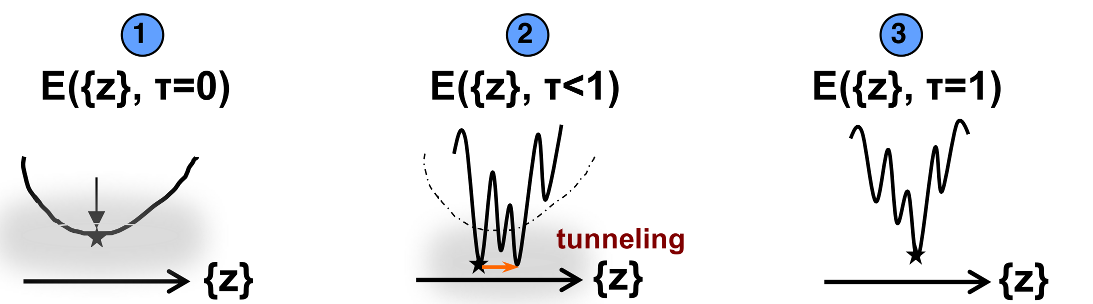


Quantum Annealing

(Finnila et al., 1994, Kadawaki & Nishimori, 1998, Farhi et.al., 2001)



- **Algorithm:** Start with large amplitude A(τ) responsible for quantum fluctuations; then, gradually turn it off while turning on the cost function of interest B(τ)
- Transitions between states via tunneling through barriers due to quantum fluctuations

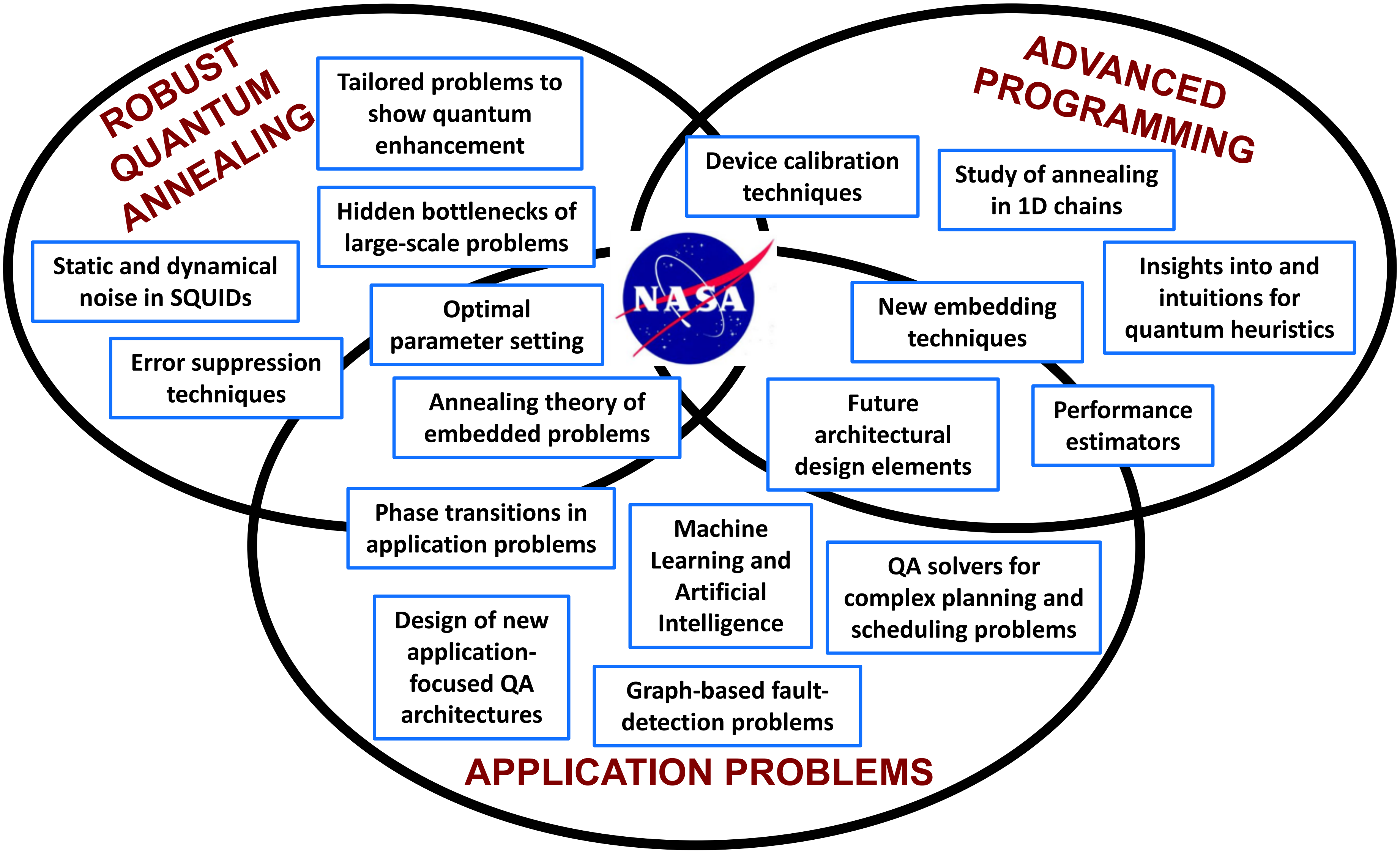


Initialize in an easy to prepare full quantum superposition

Quantum states explored by quantum tunneling

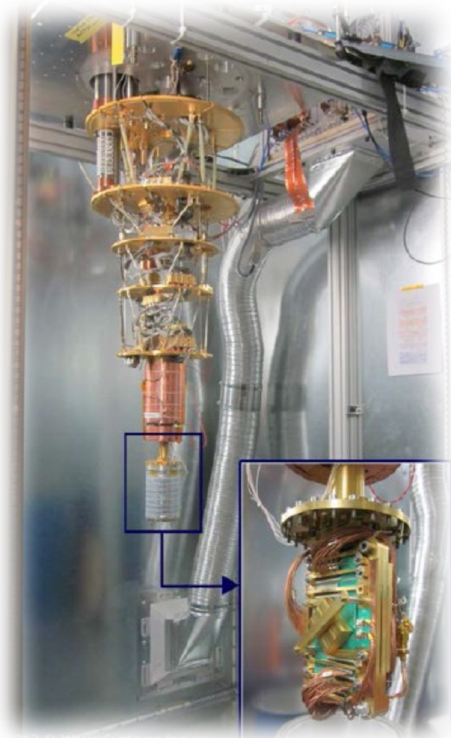
Final state a bit string encoding the solution with probability

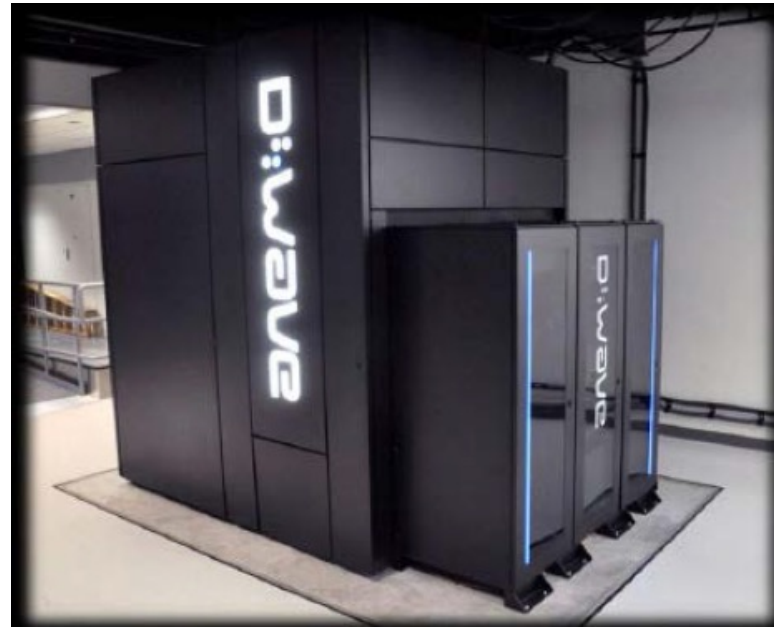
NASA Quantum Research Approach

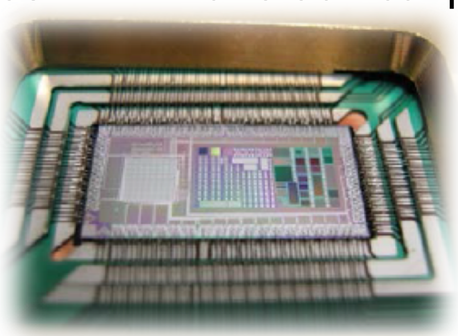


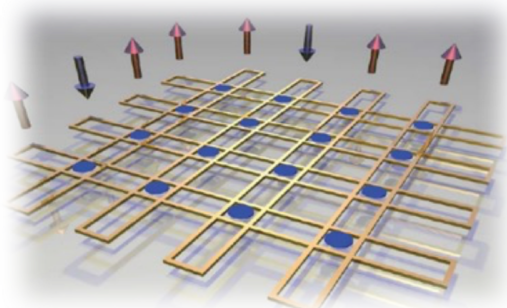
D-Wave System Hardware

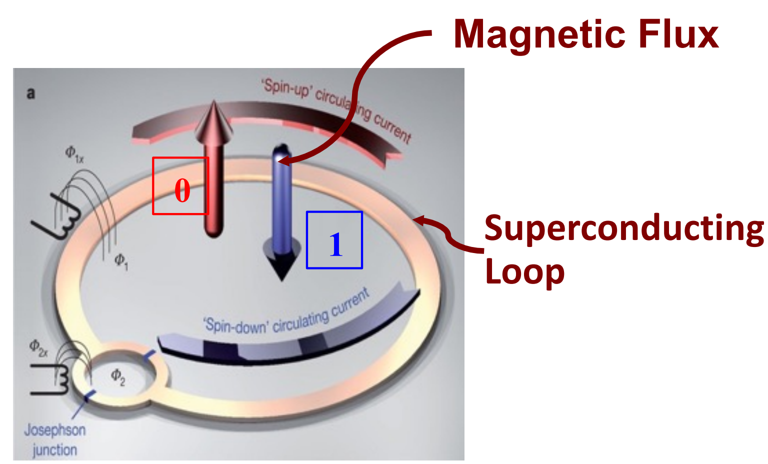
- Collaboration with Google and USRA led to installation of system at NASA Ames in 2013
- Started with 512-qubit Vesuvius processor (currently upgrading to 2000-qubit Whistler)
- 10 kg metal in vacuum at ~15 mK
- Magnetic shielding to 1 nanoTesla
- Protected from transient vibrations
- Single annealing takes 20 μs
- Typical run of 10,000 anneals (incl. reset & readout takes ~4 sec)
- Uses 12 kW of electrical power











Focused on solving discrete optimization problems using quantum annealing

D-Wave System Capability

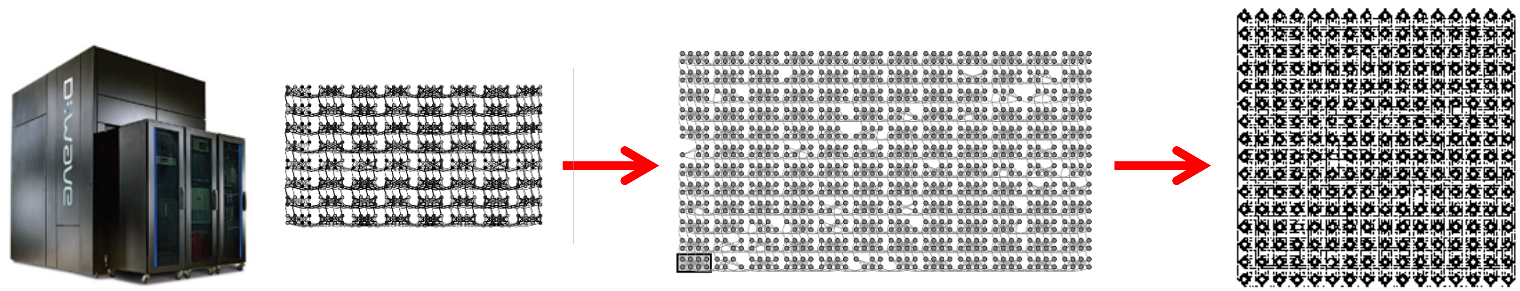
The system solves only one binary optimization problem:

Given $\{ h_j , J_{ij} \}$**, find** $\{ s_k = \pm 1 \}$

that minimizes

$$\xi(s_1, \ldots, s_N) = \sum_{j=1}^{N} h_j s_j + \sum_{i,j \in E}^{N} J_{ij} s_i s_j$$

Vesuvius to *Washington* to *Whistler*



D-Wave Two	D-Wave 2X	D-Wave 2000Q
512 (8x8x8) qubit "Vesuvius" processor	**1152** (8x12x12) qubit "Washington" processor	**2048** (8x16x16) qubit "Whistler" processor
509 qubits working – 95% yield	**1097** qubits working – 95% yield	**2038** qubits working – 97% yield
1472 *J* programmable couplers	**3360** *J* programmable couplers	**6016** *J* programmable couplers
20 mK max operating temperature (**18 mK** nominal)	**15 mK** max operating temperature (**13 mK** nominal)	**15 mK** max operating temperature (*nominal to be measured*)
5% and **3.5%** precision level for *h* and *J*	**3.5%** and **2%** precision level for *h* and *J*	*To be measured*
20 us annealing time **12 ms** programming time	**5 us** annealing time (4X better) **12 ms** programming time	**5 us** annealing time **9 ms** programming time (25% better) **New:** anneal offset, pause, quench
6 graph connectivity per qubit	**6** graph connectivity per qubit	**6** graph connectivity per qubit

Programming the D-Wave System

1 Map the target combinatorial optimization problem into QUBO

No general algorithms but smart mathematical tricks (penalty functions, locality reduction, etc.)

$$\alpha_{ijk} z_i z_j z_k$$

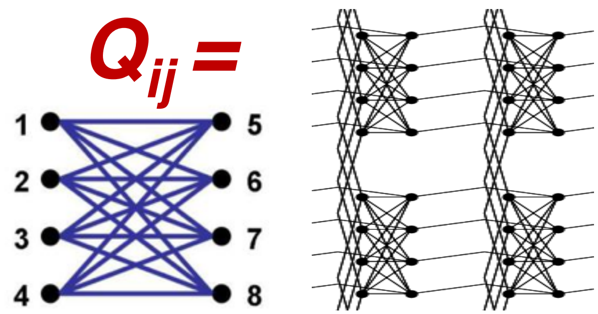
$$\alpha_{ijk}\ y_{ij} z_k + \beta_{ijk}(3y_{ij} - 2z_i y_{ij} - 2z_j y_{ij} + z_i z_j)$$

$$\sum_{ij} Q_{ij} z_i z_j \to$$

$$\sum_i h_i s_i + \sum_{i,j} J_{ij} s_i s_j$$

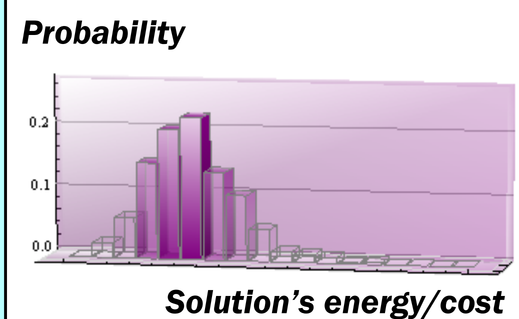
2 Embed the QUBO coupling matrix in the hardware graph of interacting qubits

D-Wave qubit hardware connectivity is a Chimera graph, so embedding methods mostly based on heuristics



3 Run the problem several times and collect statistics

Use symmetries, permutations, and error correction to eliminate the systemic hardware errors and check the solutions



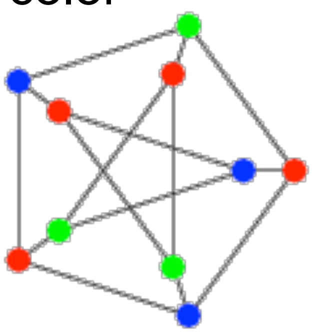
Mapping not needed for random spin-glass models

Embedding not needed for native Chimera problems

Performance can be improved dramatically with smart pre-/post-processing

Mapping to QUBO: Graph Coloring Example

Graph Coloring Problem:
Assign one of *k* colors to each vertex so that no two vertices sharing an edge have the same color



Binary variable:

$$x_{v,c} = \begin{cases} 1 & \text{vertex } v \text{ with color } c \\ 0 & \text{vertex } v \text{ not with color } c \end{cases}$$

Violation of requirements encoded as cost:

Costing cases

(1) No color or Multi-colored

- (1) unique assignment: Each vertex v must be assigned exactly one color:

$$H_v^{(unique)} = \left(\sum_{c \in C} x_{v,c} - 1\right)^2 \Leftarrow \sum_{c \in C} x_{v,c} = 1$$

- (2) Connected vertices cannot use the same color

(2) Same color for connected vertices

$$H_{v,v',c}^{(exclude)} = x_{v,c} x_{v',c} \text{ if } vv' \in E$$

Final QUBO form:

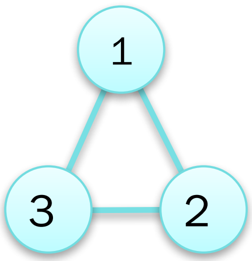
$$H = \sum_v H_v^{(unique)} + \sum_{v,v' \in E} \sum_c H_{v,v',c}^{(exclude)}$$

H = 0 corresponds to a valid coloring

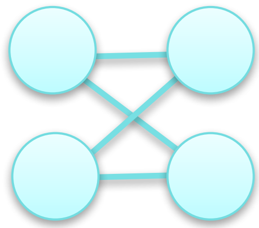
Embedding the QUBO

Embed a triangle onto a bipartite graph

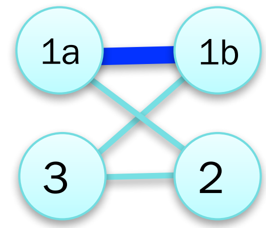
original QUBO



hardware connectivity



QUBO embedded



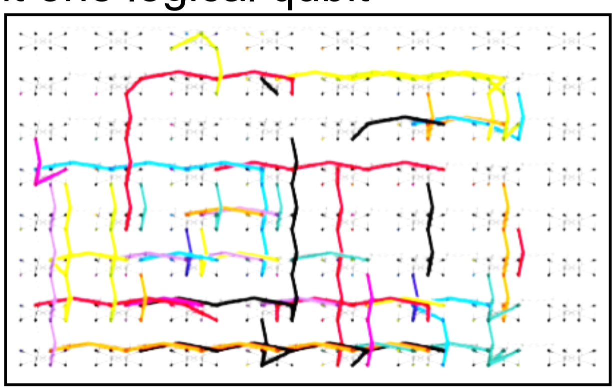
$$H_0 = J_{12}x_1x_2 + J_{23}x_2x_3 + J_{13}x_1x_3$$

$$H_1 = J_{12}x_{1a}x_2 + J_{23}x_2x_3 + J_{13}x_{1b}x_3 + \underline{J_{\mathtt{Ferro}}}x_{1a}x_{1b}$$

Strong, but not too strong, ferromagnetic coupling between physical qubits x_{1a} and x_{1b} encourages them to take the same value, thus acting as a single logical qubit x_1

Embedding a realistic problem instance:

Physical qubits on each colored path represent one logical qubit

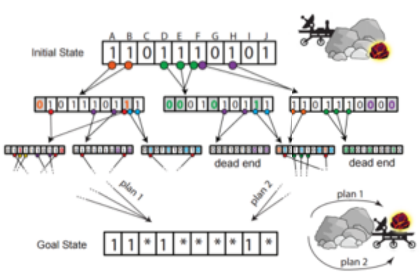


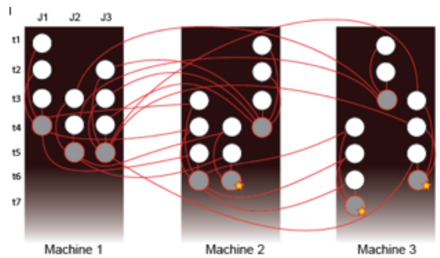
H_0 and H_1 have the same ground state but the energy landscape of the search space differs

Current research investigation: How best to set the magnitude of these "strong" couplings to maximize probability of success

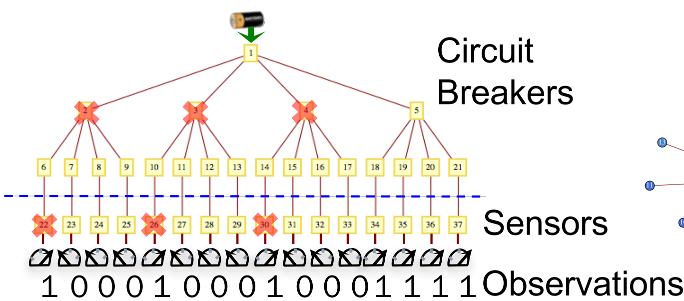
Current NASA Research in Applications

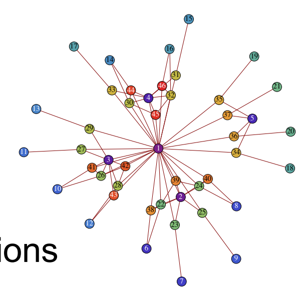
Complex Planning and Scheduling



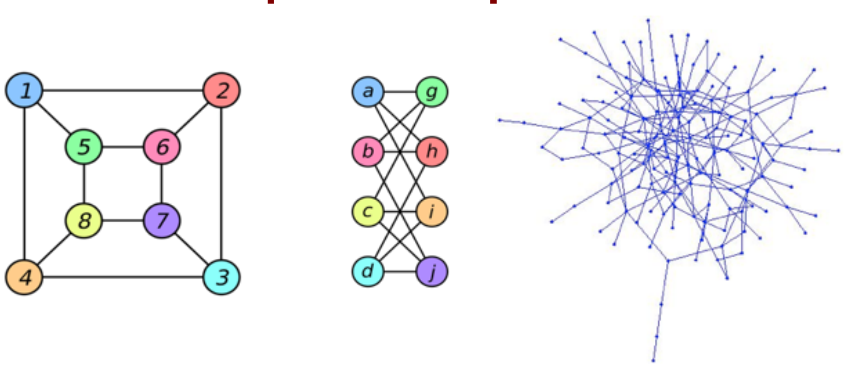


Graph-based Fault Detection





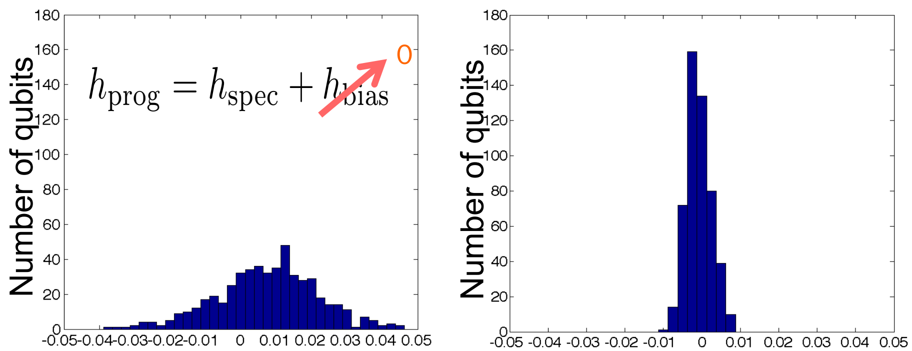
Graph Isomorphism



- General **Planning Problems** (e.g., navigation, scheduling, asset allocation) can be solved on a quantum annealer
- Developed a quantum solver for **Job Shop Scheduling** that pre-characterizes instance ensembles to design optimal embedding and run strategy – tested at small scale (6x6) but potentially could solve intractable problems (15x15) with 10x more qubits
- Analyzed simple graphs of **Electrical Power Networks** to find the most probable cause of multiple faults – easy and scalable QUBO mapping, but good parameter setting (e.g., gauge selection) key to finding optimal solution – now exploring digital circuit **Fault Diagnostics and V&V**
- **Subgraph Matching Problems** are common in applications of interest to the intelligence community – similarly, finding **Longest Matching Sequences** important in genomics and bioinformatics

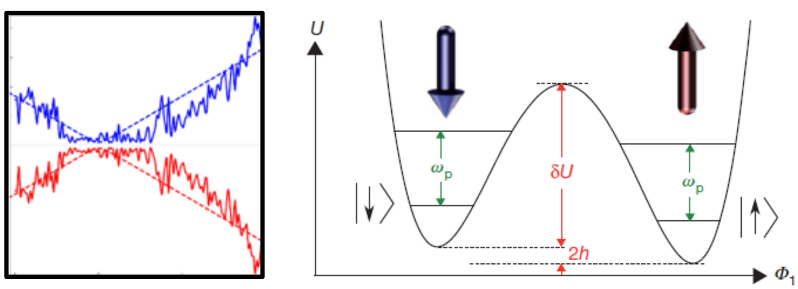
Current NASA Research in Quantum Physics

Calibration of Quantum Annealers

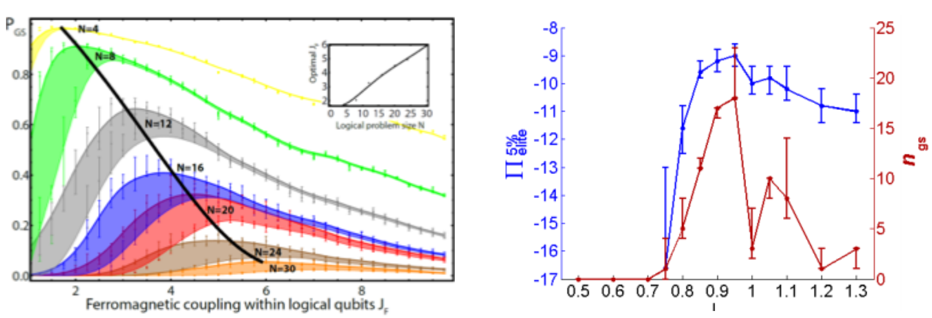


h biases (before) correction   *h* biases after correction

Effect of Noise on Quantum Annealing



Optimal Embedding & Parameter Setting



- Developed technique to determine and correct **residual persistent biases** in the programmable parameters of quantum annealers (*h* and *J*) – correction significantly improves performance and reliability (reduction in variability)
- First **realistic noise analyses** show how low-frequency noise dramatically affects the performance of quantum annealers – results being used to design hardware improvements
- Limited hardware connectivity makes embedding challenging – good runtime parameters determined by considering the **nature and dynamics of chains** – quick scans can be used to predict performance of extensive scans
- Small instances of **hard problems at phase transitions** in combinatorial optimization are intractable – they can be designed by looking at solvability phase transitions
- Predict **tractability of application problems** by studying the scaling of energy gaps and density of bottlenecks in spin glass phase

Quantum annealing capabilities

1) As a discrete optimization solver:

Given $\{h_j, J_{ij}\}$, find $\{s_k = \pm 1\}$ that minimizes

NP-hard problem

$$\xi(s_1, \ldots, s_N) = \sum_{j=1}^{N} h_j s_j + \sum_{i,j \in E}^{N} J_{ij} s_i s_j$$

Potential NASA applications:

- *planning*
- *scheduling*
- *fault diagnosis*
- *graph analysis*
- *communication networks, etc.*

QUBO: Quadratic Unconstrained Binary Optimization (Ising model in physics jargon).

2) As a physical device to sample from Boltzmann-like distributions:

$$P_{Boltzman} \propto exp[-\xi(s_1, \ldots, s_N)/T_{eff}] \longrightarrow \langle v_i h_j \rangle_{p(\mathbf{h},\mathbf{v})}$$

Computationally bottleneck

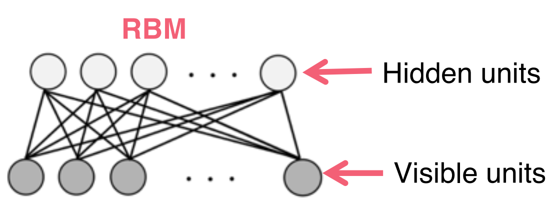
Early work:
Bian et al. 2010. The Ising model: teaching an old problem new tricks.

Follow-up work:
Raymond et al. Global warming: Temperature estimation in annealers. Frontiers in ICT, 3, 23 (2016).

Our work: Benedetti et al. PRA 94, 022308 (2016)

- We provide a robust algorithm to estimate the effective temperature of problem instances in quantum annealers.
- Algorithm uses the same samples that will be used for the estimation of the gradient

Widely used in **generative unsupervised** learning



Potential NASA applications:

- *machine leaning (e.g., training of deep-learning networks)*

Unsupervised learning relies on sampling

Lesson 1: Move to intractable problems of interest to ML experts (e.g., **generative models in unsupervised learning**). **Quantum advantage in near term.**

"Unsupervised learning [... has] been overshadowed by the successes of purely supervised learning. [... We] expect **unsupervised learning to become far more important in the longer term**. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object."

LeCun, **Bengio**, **Hinton**, *Deep Learning*, **Nature 2015**

"In the context of the deep learning approach to undirected modeling, it is rare to use any approach other than Gibbs sampling. **Improved sampling techniques are one possible research frontier.**"

Goodfellow, Bengio, Courville, *Deep Learning*, book in preparation for MIT Press, **2016**

"Most of the previous work in **generative models** has focused on variants of **Boltzmann Machines** [...] While these models **are very powerful**, each iteration of **training requires a computationally costly step of MCMC** to approximate derivatives of an intractable partition function (normalization constant), making it **difficult to scale** them **to large datasets.**"

Mansimov, Parisotto, Ba, **Salakhutdinov**, ICLR **2016**

Unsupervised learning (generative models)

Learn the "best" model distribution that can generate the same kind of data

MODEL

P **(Image)**

Learning algorithm

NO LABELS

DATASET

Example application:

Image reconstruction

Reconstructed image

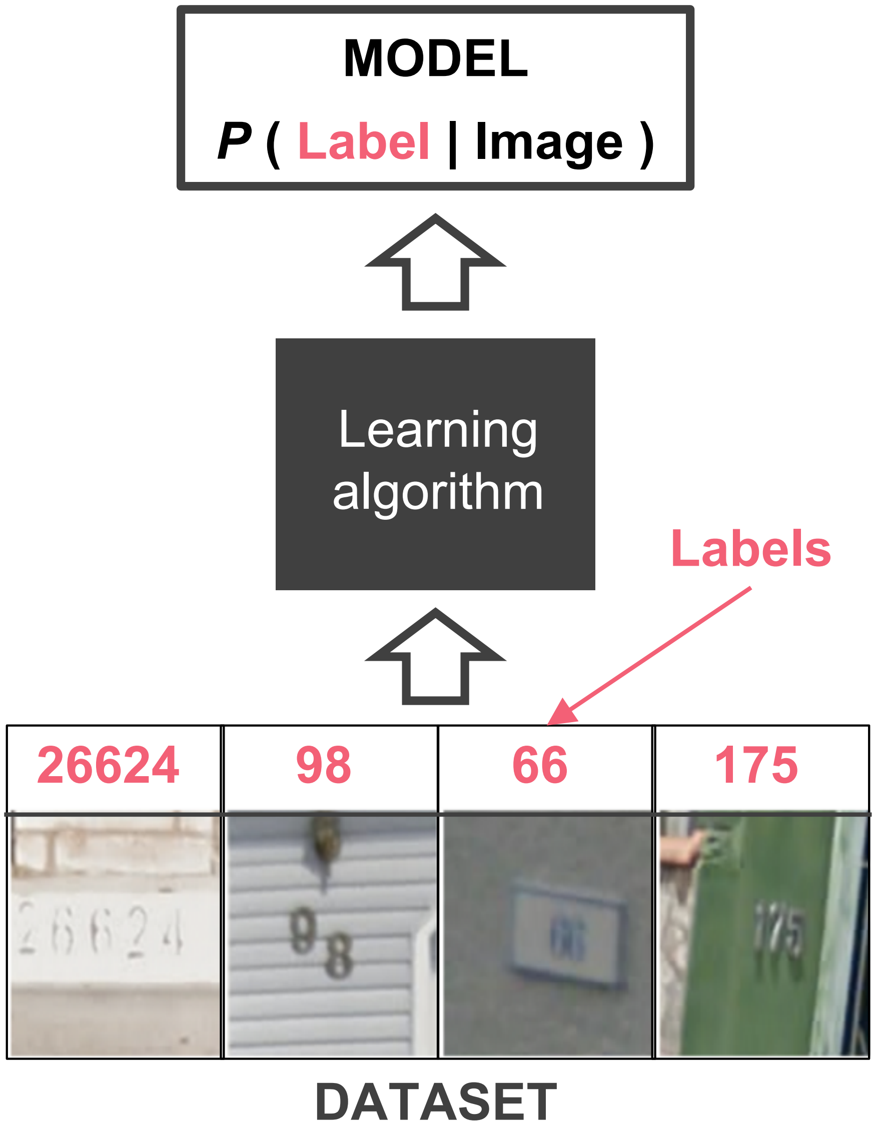
LEARNED MODEL

P **(Image)**

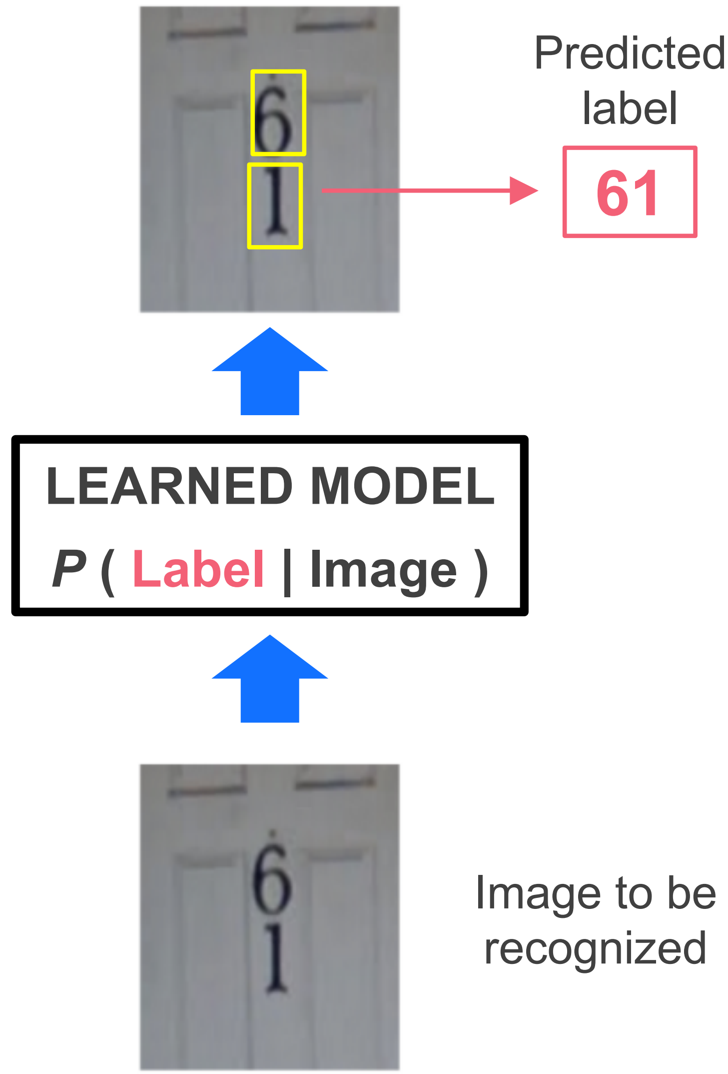
Damaged image

Supervised learning (discriminative models)

Learn the “best” model that can perform a specific task

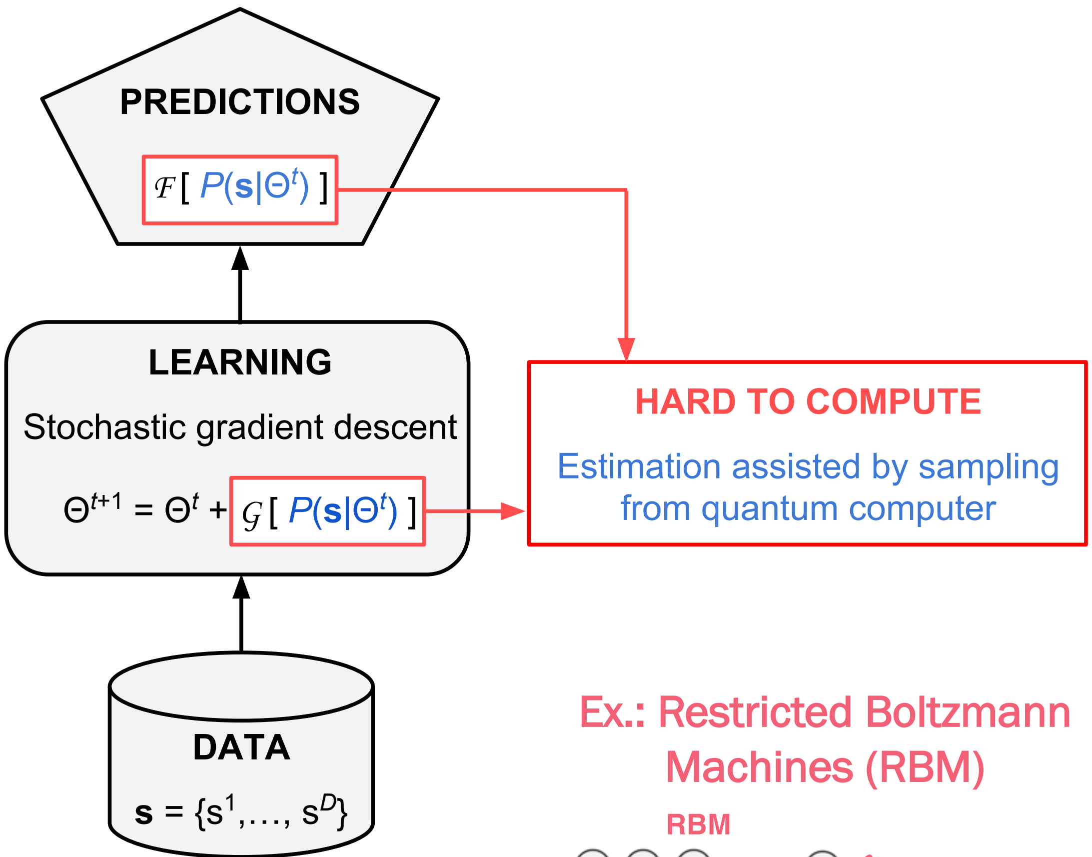


Example application:
Image recognition

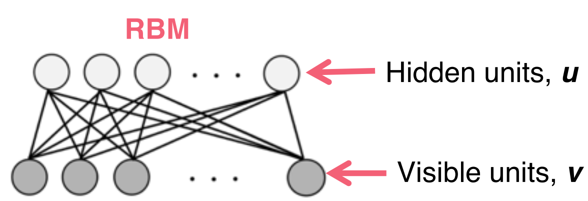


A near-term approach for quantum-enhanced machine learning

Lesson 2: Hybrid approaches for generative modeling in unsupervised machine learning.



Ex.: Restricted Boltzmann Machines (RBM)



Computationally bottleneck

$$\langle v_i u_j \rangle_{p(\mathbf{v},\mathbf{u})}$$

Where, $p(\mathbf{v},\mathbf{u}) = \dfrac{e^{-E(\mathbf{v},\mathbf{u}|\theta)/T_{\mathrm{eff}}}}{Z(\theta)}$

Widely used in **unsupervised** learning

Challenges solved:

Benedetti, et al. Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning. **PRA 94, 022308** (2016).

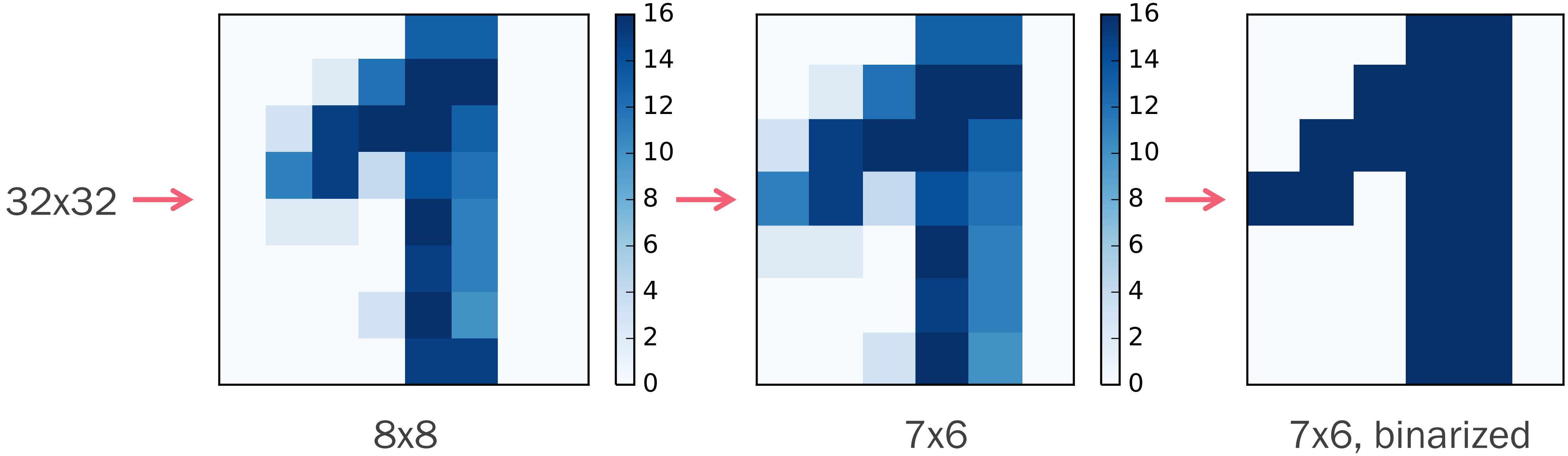
Benedetti, et al. Quantum-assisted learning of hardware-embedded probabilistic graphical models. **arXiv:1609.02542** (2016). Accepted in PRX.

Perdomo-Ortiz, et al. Opportunities and Challenges in Quantum-Assisted Machine Learning in Near-term Quantum Computer. **arXiv:1708.09757**. (2017). Invited article to special QST issue.

Benedetti, et al. Quantum-assisted Helmholtz machines: A quantum-classical deep learning framework for industrial datasets in near-term devices. **arXiv:1708.09784** (2017).

Quantum-assisted unsupervised learning on digits

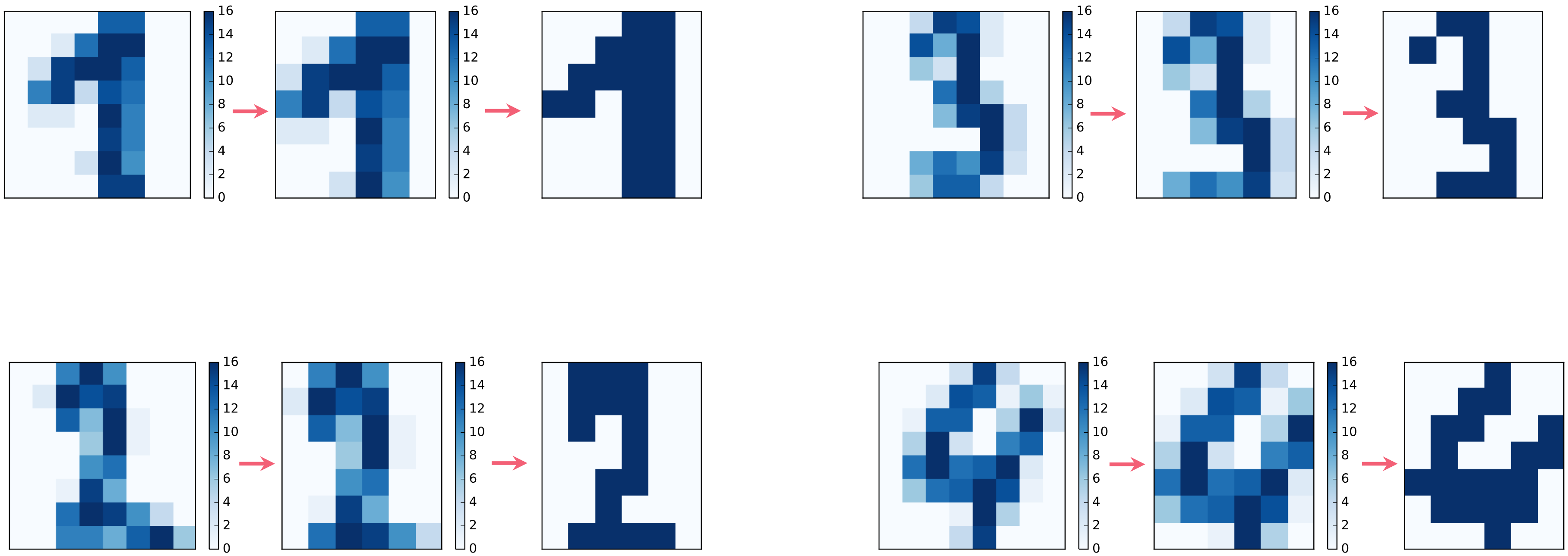
OptDigits Datasets



Dataset: Optical Recognition of Handwritten Digits (OptDigits)

Quantum-assisted unsupervised learning on digits

OptDigits Datasets

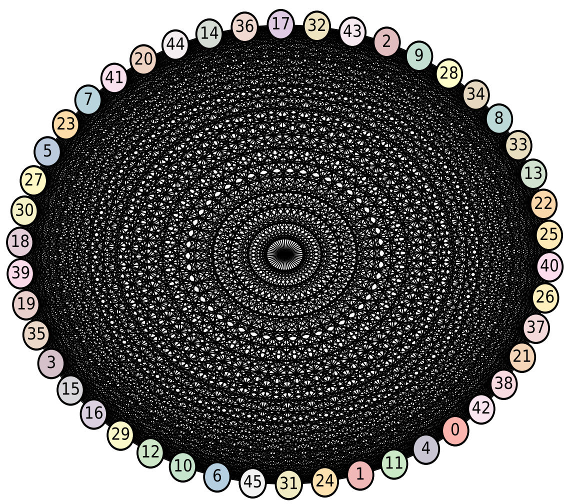


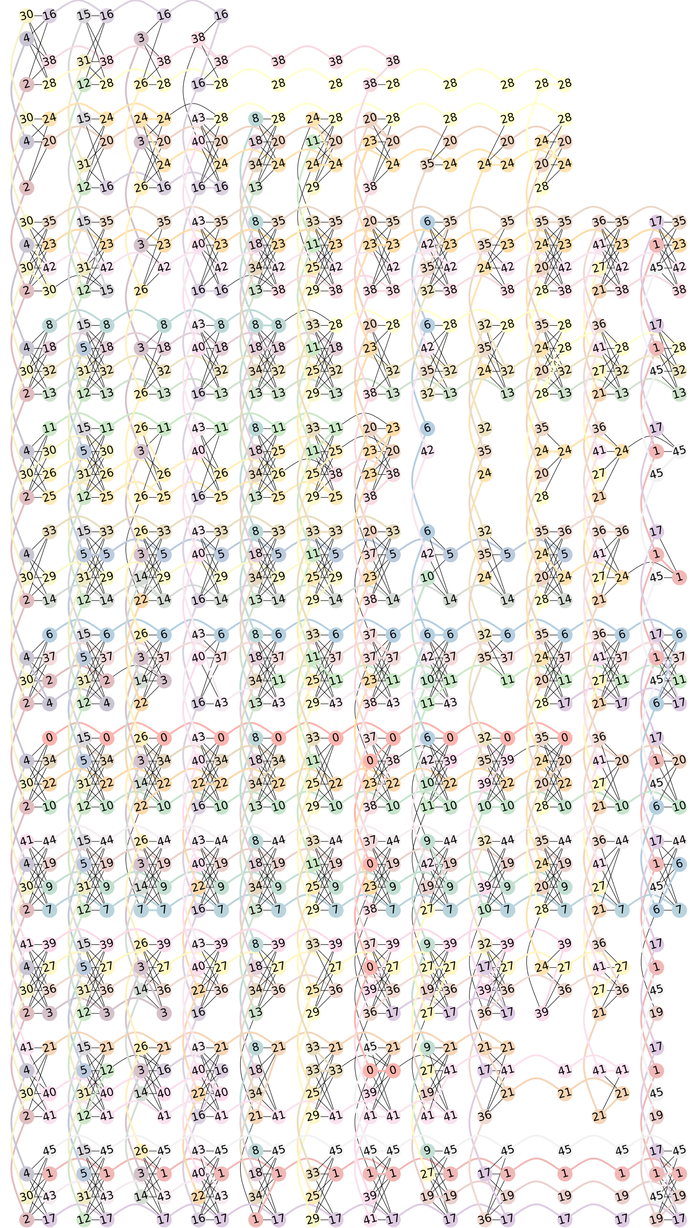
Dataset: Optical Recognition of Handwritten Digits (OptDigits)

Quantum-assisted unsupervised learning on digits

Overcoming the curse of limited connectivity in hardware.

46 fully-connected logical (visible) variables





940 physical qubits

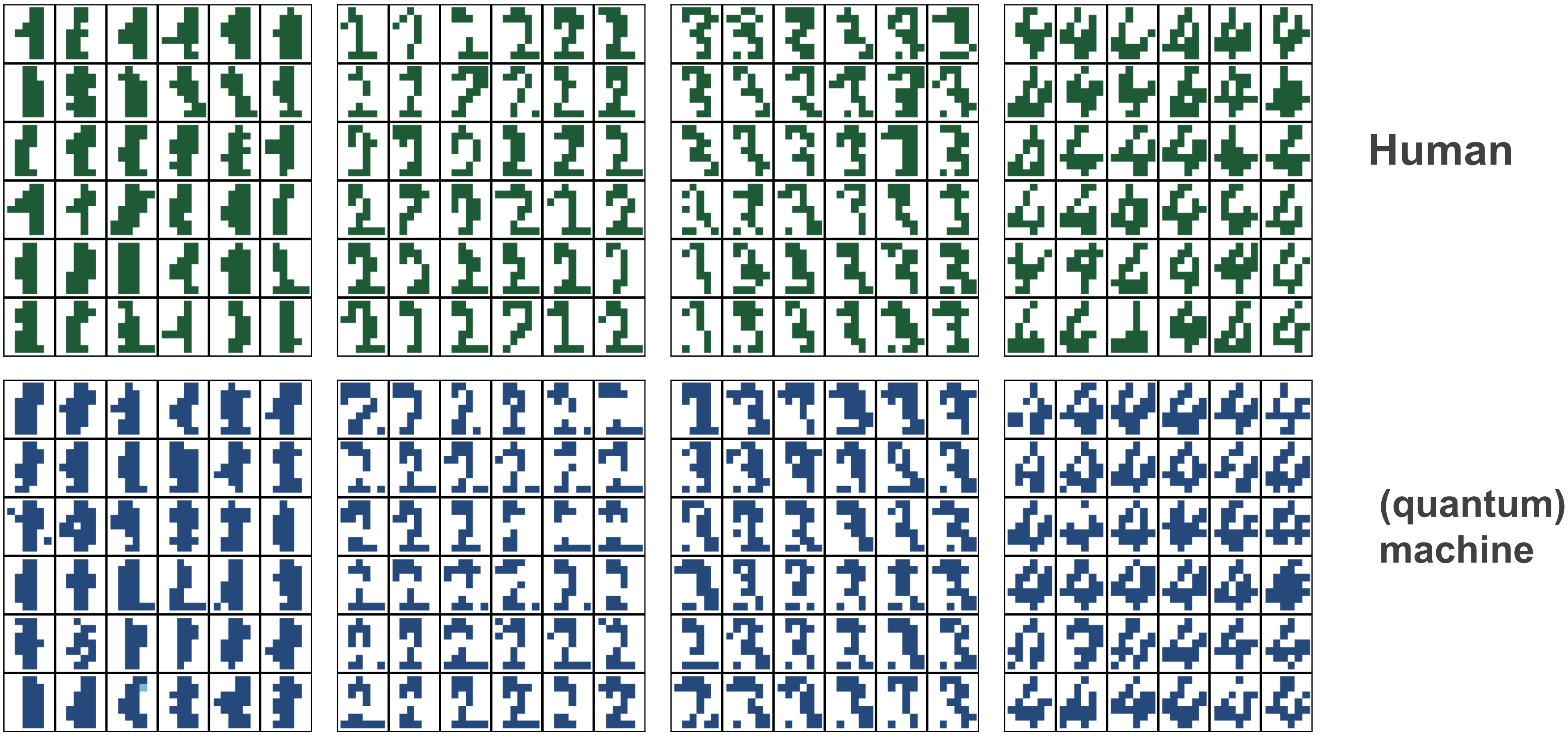
42 for pixels + 4 to one-hot encode the class (only digits 1-4)

- Are the results from this training on 940 qubit experiment meaningful?
- Is the model capable of generating digits?

Benedetti, et al. Quantum-assisted learning of **hardware-embedded** probabilistic graphical models. **arXiv:1609.02542** (2016). Accepted in PRX.

Quantum-assisted unsupervised learning on digits

Human or (quantum) machine? (Turing test)



Dataset: Optical Recognition of Handwritten Digits (OptDigits)

Results from experiments using 940 qubits, without post-processing.
The hardware-embedded model represents a 46 node fully connected graph.

Benedetti, et al. **arXiv:1609.02542** (2016). Accepted in PRX.

A near-term approach for quantum-enhanced machine learning

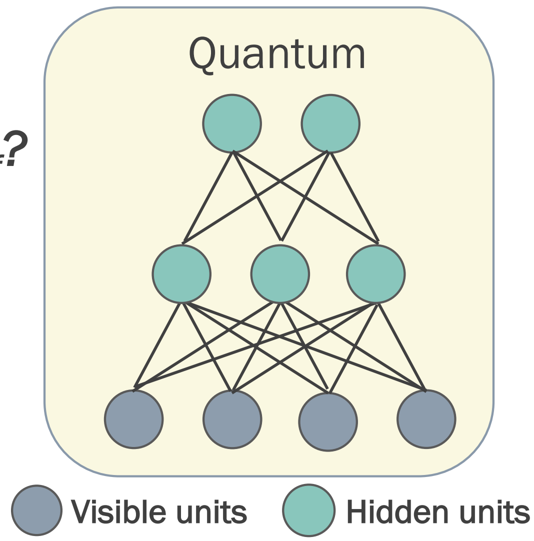
Challenges of the hybrid approach:

- Need to solve classical-quantum model mismatch

Training Method: Stochastic gradient ascent

$$\sum_{\boldsymbol{v} \in S} \frac{\partial \ln \mathcal{L}(\boldsymbol{\theta}|\boldsymbol{v})}{\partial w_{ij}} \propto \boxed{\langle v_i h_j \rangle_{\text{data}}} - \boxed{\langle v_i h_j \rangle_{\text{model}}}$$

Classical - T_{eff}?



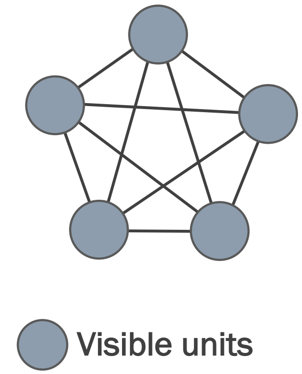
Benedetti et al . Phys. Rev. A 94, 022308 (2016)

No progress in five years since QA sampling was proposed as a promissing appplication.

- Robustness to noise, intrinsic control errors, and to deviations from sampling model (e.g., Boltzmann)

- *Curse of limited connectivity –* parameter setting

Fully visible models

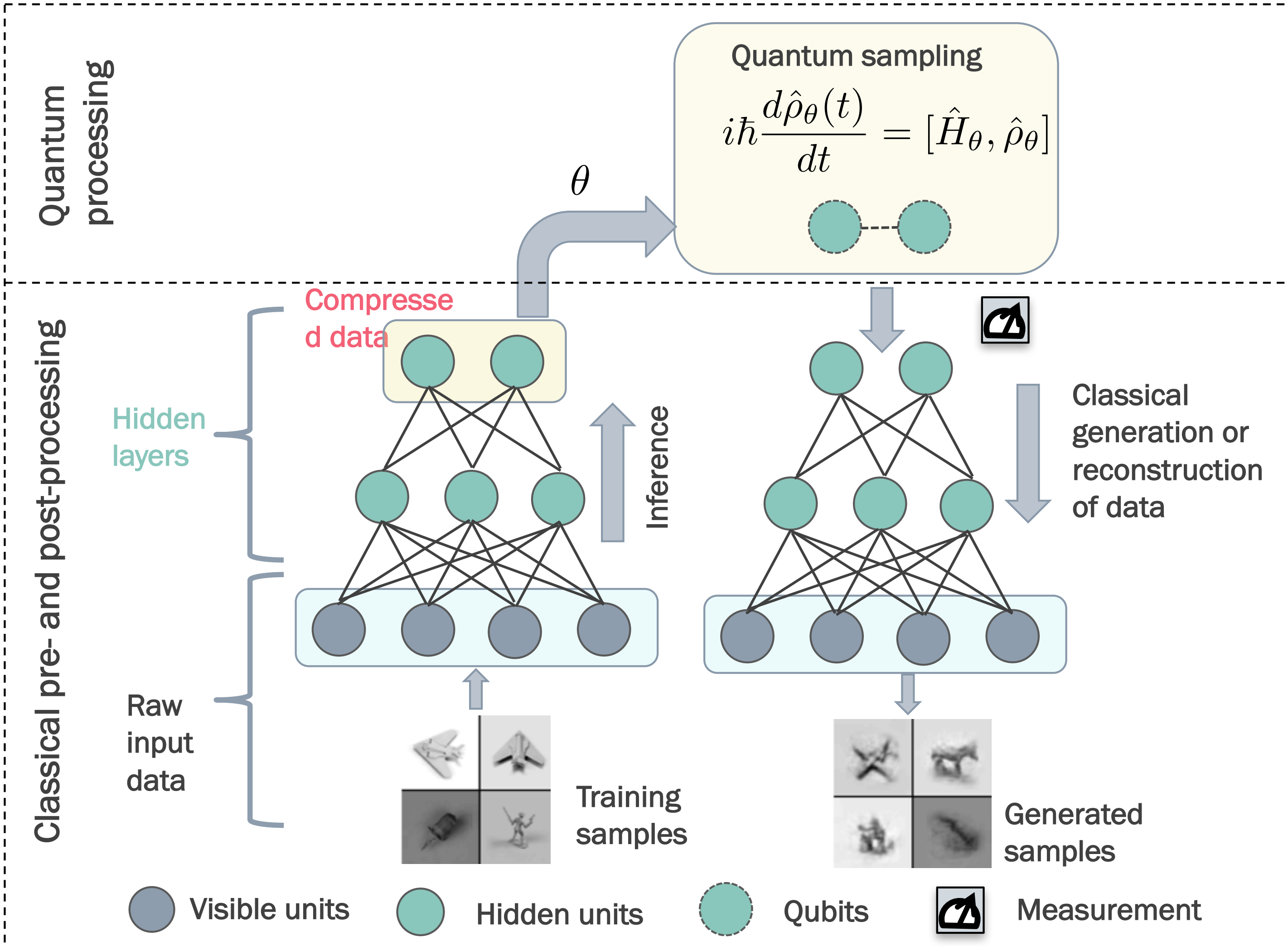


Benedetti et al. arXiv:1609.02542

How about large complex datasets with continuous variables?
All previous fail to do that (fully quantum and hybrid here)

Perspective on quantum-enhanced machine learning

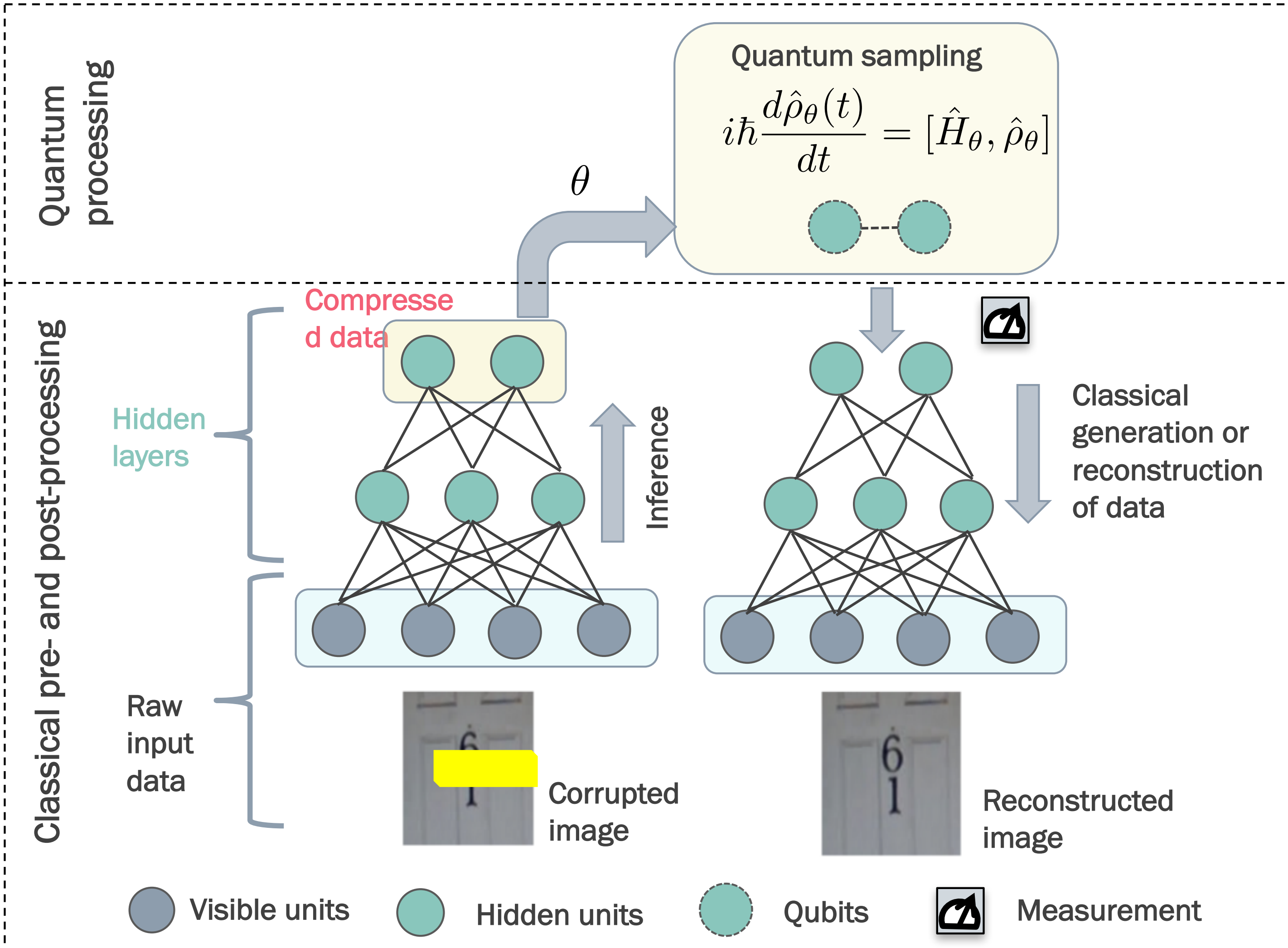
- New hybrid proposal that works directly on a low-dimensional representation of the data.



Benedetti, Realpe-Gomez, and Perdomo-Ortiz. Quantum-assisted Helmholtz machines: A quantum-classical deep learning framework for industrial datasets in near-term devices. **arXiv:1708.09784** (2017).

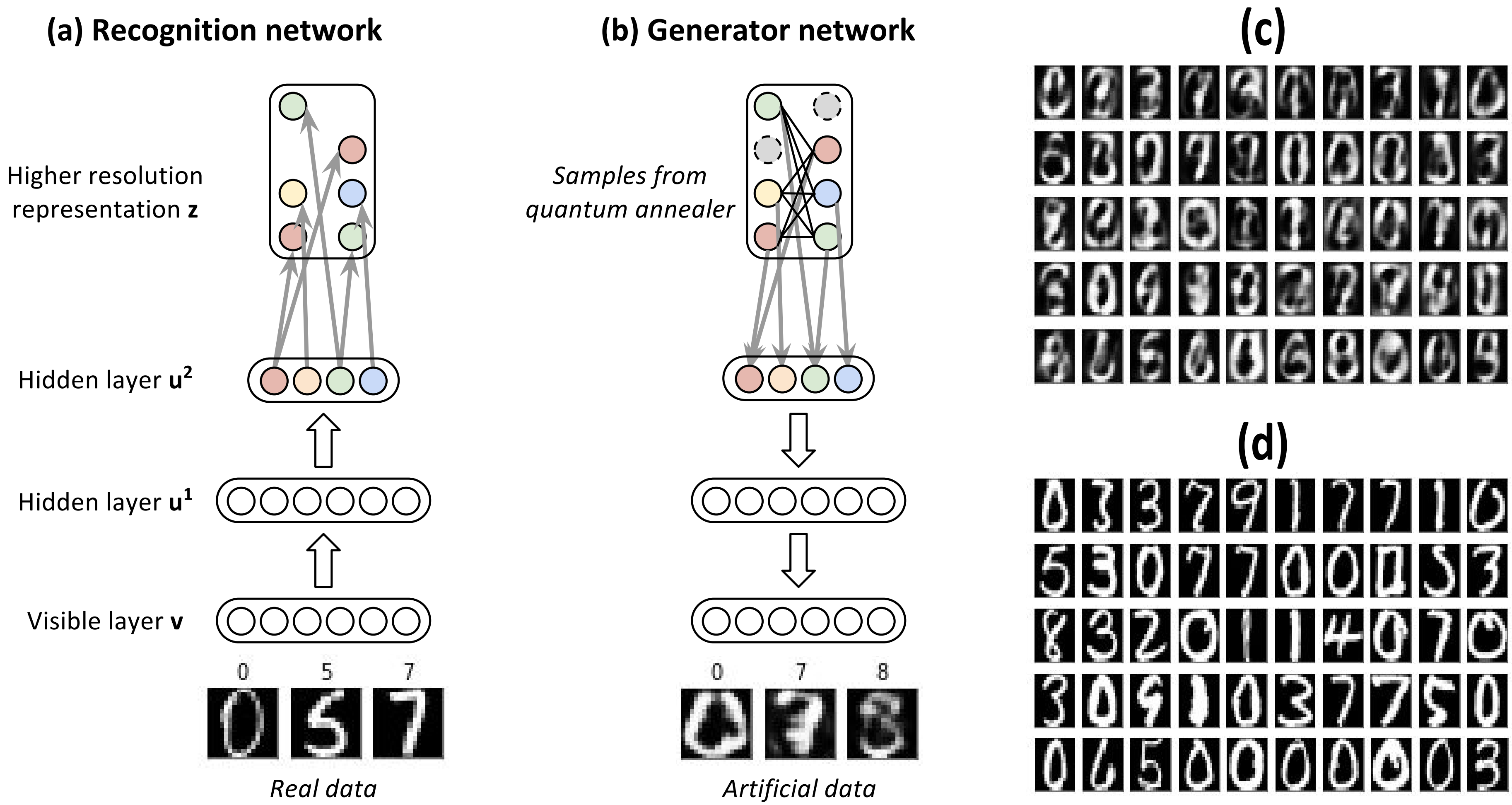
Perspective on quantum-enhanced machine learning

- New hybrid proposal that works directly on a low-dimensional representation of the data.



Benedetti, Realpe-Gomez, and Perdomo-Ortiz. Quantum-assisted Helmholtz machines: A quantum-classical deep learning framework for industrial datasets in near-term devices. **arXiv:1708.09784** (2017).

Experimental implementation of the QAHM



Experiments using 1644 qubits (no further postprocessing!)

Max. CL = 43

Benedetti, Realpe-Gomez, and Perdomo-Ortiz. Quantum-assisted Helmholtz machines: A quantum-classical deep learning framework for industrial datasets in near-term devices. **arXiv:1708.09784** (2017).

Lesson 1: Focus on the **hardest** problems of interest to ML experts (e.g., **generative** models in **unsupervised learning**).
Quickest path to demonstrating quantum advantage in the near-term

Lesson 2: Focus on novel **hybrid** quantum-classical approaches.
Cope with hardware constrains. Exploitation of available quantum resources

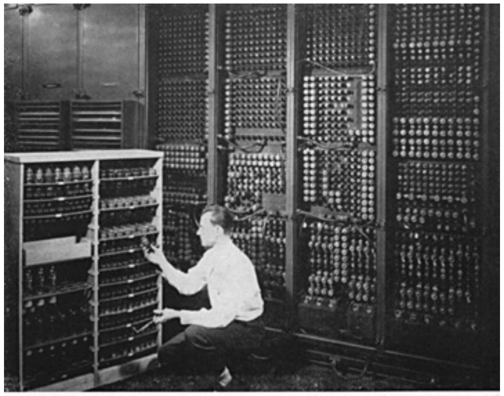
arXiv:1708.09757. (2017). To appear in the Quantum Science and Technology (QST) invited special issue on "What would you do with a 1000 qubit device?"

Conclusions

- Understanding and harnessing the fundamental power of quantum computing is a formidable challenge that requires:
  - New insights in physics and mathematics
  - Innovations in computer and computational science
  - Breakthroughs in engineering design to produce robust, reliable, scalable technologies
- NASA QuAIL team has successfully demonstrated that discrete optimization problems can be run on quantum annealers
  - Effectively using such systems needs judicious mapping, embedding, execution strategies
- Exciting decade in quantum computing ahead of us
  - Compilation and performance capabilities of today's annealers are improving rapidly
  - New and better quantum algorithms, particularly quantum heuristics, are emerging
  - Small-scale universal quantum computers are becoming available

ENIAC (1946), the first "general-purpose" computer

The task of taking a problem and mapping it onto the machine was complex, and usually took weeks. After the program was figured out on paper, the process of getting the program "into" ENIAC by manipulating its switches and cables took additional days. This was followed by a period of verification and debugging […] (source: http://en.wikipedia.org/wiki/ENIAC)



Replacing a bad tube meant checking among ENIAC's 19,000 possibilities.