D-Wave

The Quantum Computing Company™

# D-Wave Quantum Computing

## Quantum Computing for the Real World Today


Dr. Colin P. Williams

Vice President of Strategy & Corporate Development

D-Wave Systems Inc.

# D-Wave's Mission & Activities

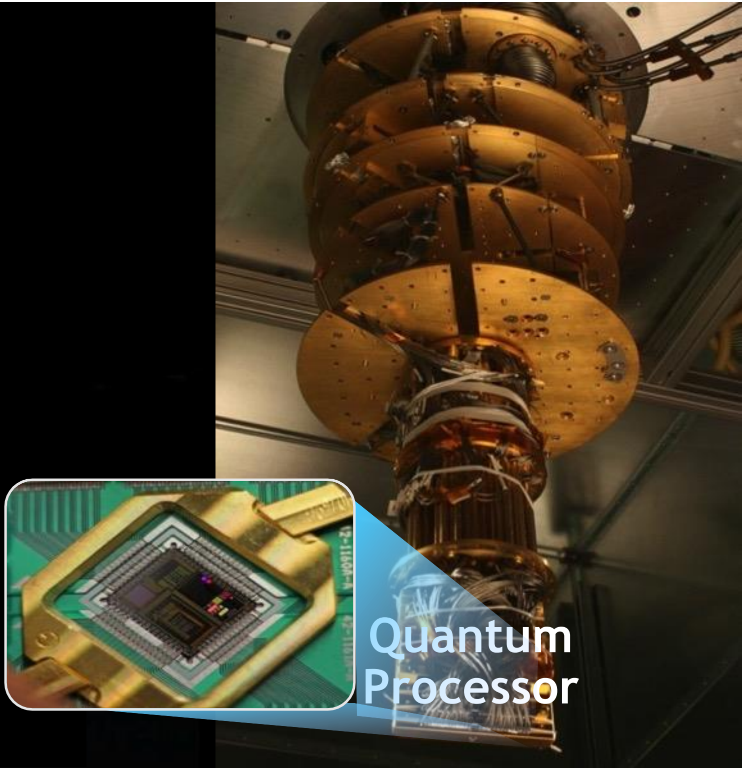
- **Mission**
  - To solve the world's hardest problems especially in the areas of artificial intelligence and machine learning


- **Core technologies**
  - Superconducting annealing-based quantum computers
  - Hybrid quantum/classical architectures
- **Business model**
  - Quantum computer system sales
  - Quantum computer cloud services
  - Quantum machine learning services

# What are Quantum Computers?

# What are Quantum Computers?

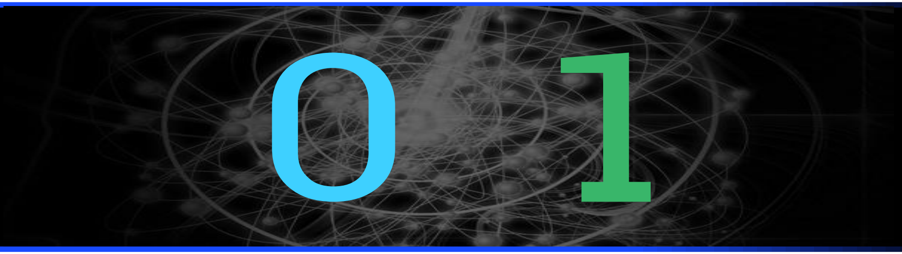
Computers that harness quantum physical effects not available to conventional computers



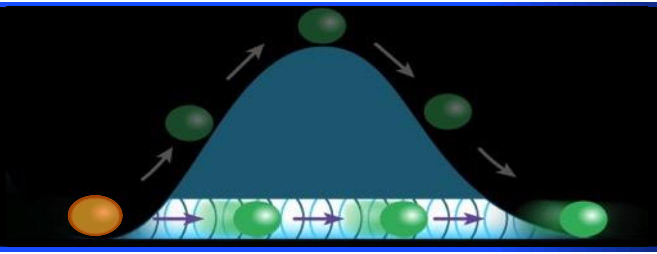
# Which Quantum Effects are Used?

Superposition




Entanglement

Quantum Tunneling



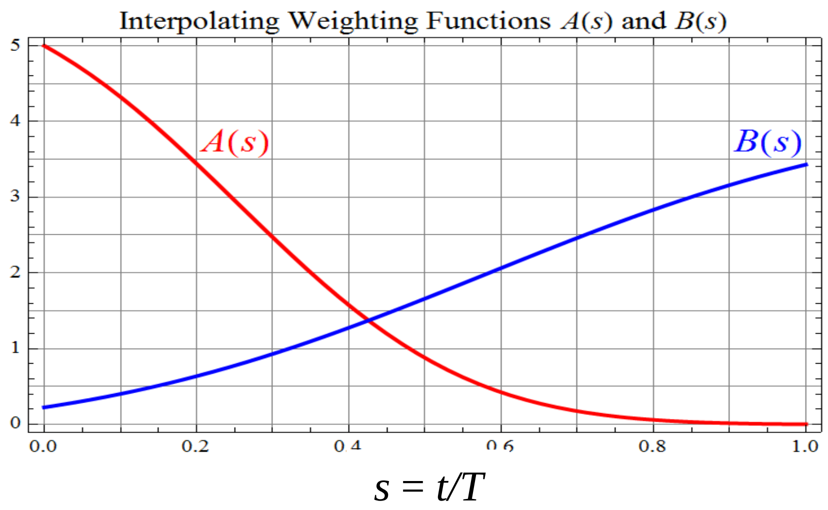
# Our Approach in Context

- **Gate Model** (Google, IBM, Intel, Alibaba, Rigetti)
  - Most common approach / based on analogy with Boolean logic circuits
  - Very difficult to scale; requires **massive** qubit overhead for error correction
- **Topological** (Microsoft)
  - Like gate model but without need for error correction (in theory)
  - Needs exotic quasi-particle whose robustness is now in dispute
    (Phys. Rev. Lett., 118, 046801, 26 Jan 2017)
- **Annealing** (D-Wave, Google, IARPA)
  - Harnesses Nature's ability to find low energy configurations via quantum tunneling
  - Resilient to noise / does not require long coherence times / MIT pedigree
  - Handles a wide range of important problems
  - Currently non-universal but could be made universal

# Q. Annealing is Inspired by Adiabatic Theorem of QM …

- **If …**
  - Start in ground state of $H_{\text{initial}}$
  - Change $H_{\text{initial}}$ to $H_{\text{final}}$ in total time $t = T$

    $H(s) = A(s)\, H_{\text{initial}} + B(s)\, H_{\text{final}}$
  - where $0 \leq s = t/T \leq 1$

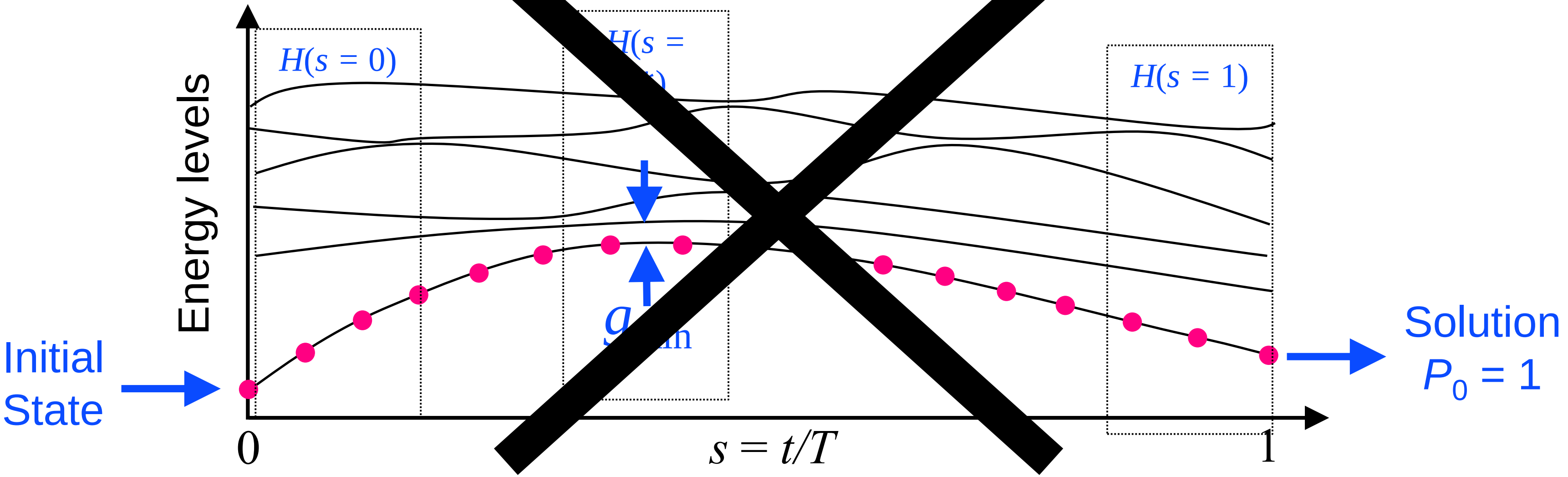



- **then …**
  - System will remain in ground state of all the instantaneous Hamiltonians passed through
  - Provided change is made sufficiently slowly

# How Long “Ought” the Computation Take?

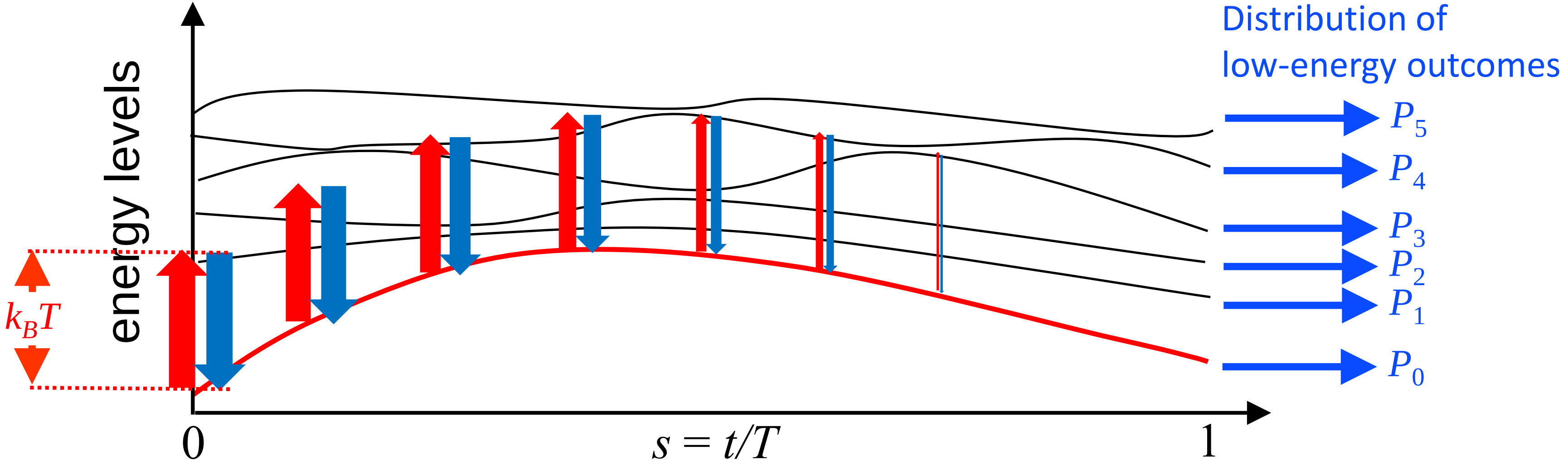
$$H(s) = A(s)\,H_{\text{initial}} + B(s)\,H_{\text{final}} \quad \text{where } s = t/T$$

Theorem says running time $T \sim O(1/g_{\min}^2)$



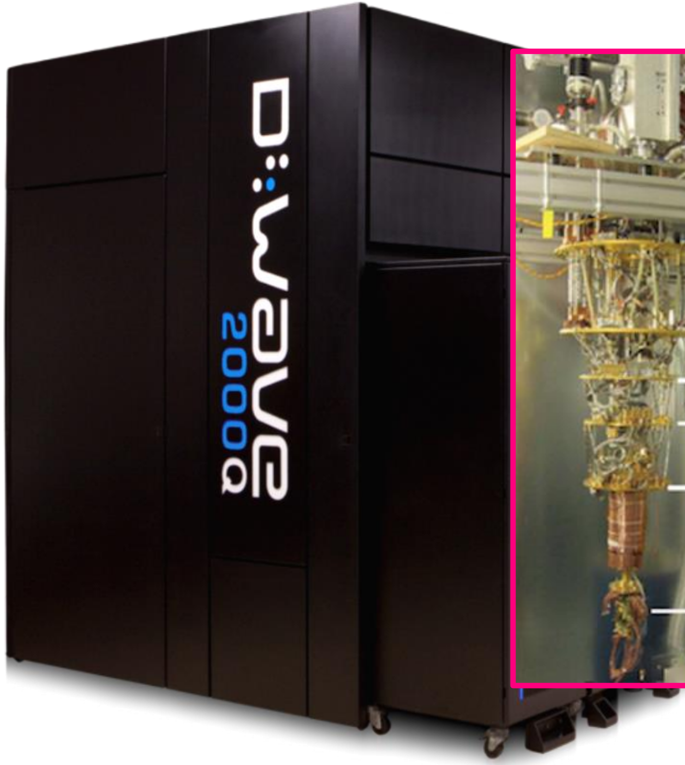
# Quantum Annealing

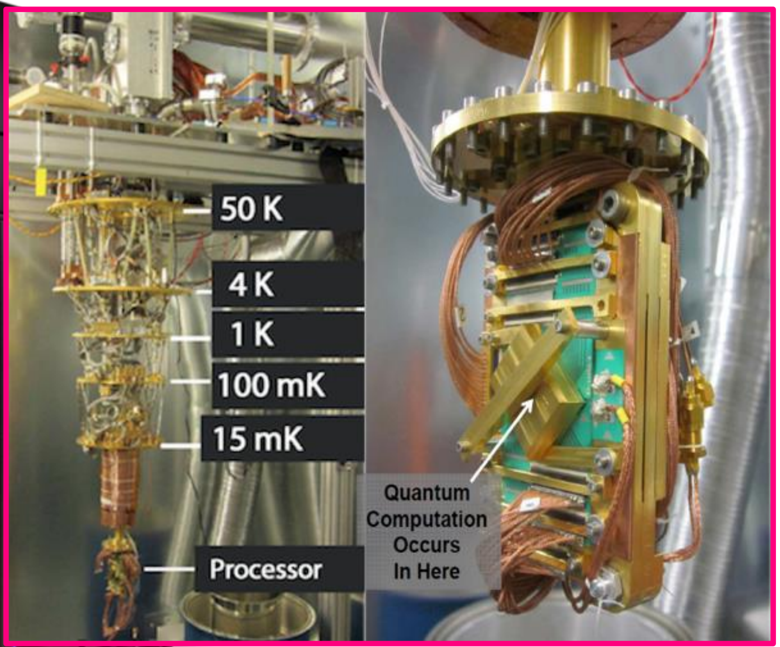
- AQC: assumes isolation from env. & 0 Kelvin; run slowly once
- QA: assumes coupling to env. & > 0 Kelvin; run quickly and repeat



# D-Wave's
# Quantum Computer

# New Product: D-Wave 2000Q™ (January 2017)

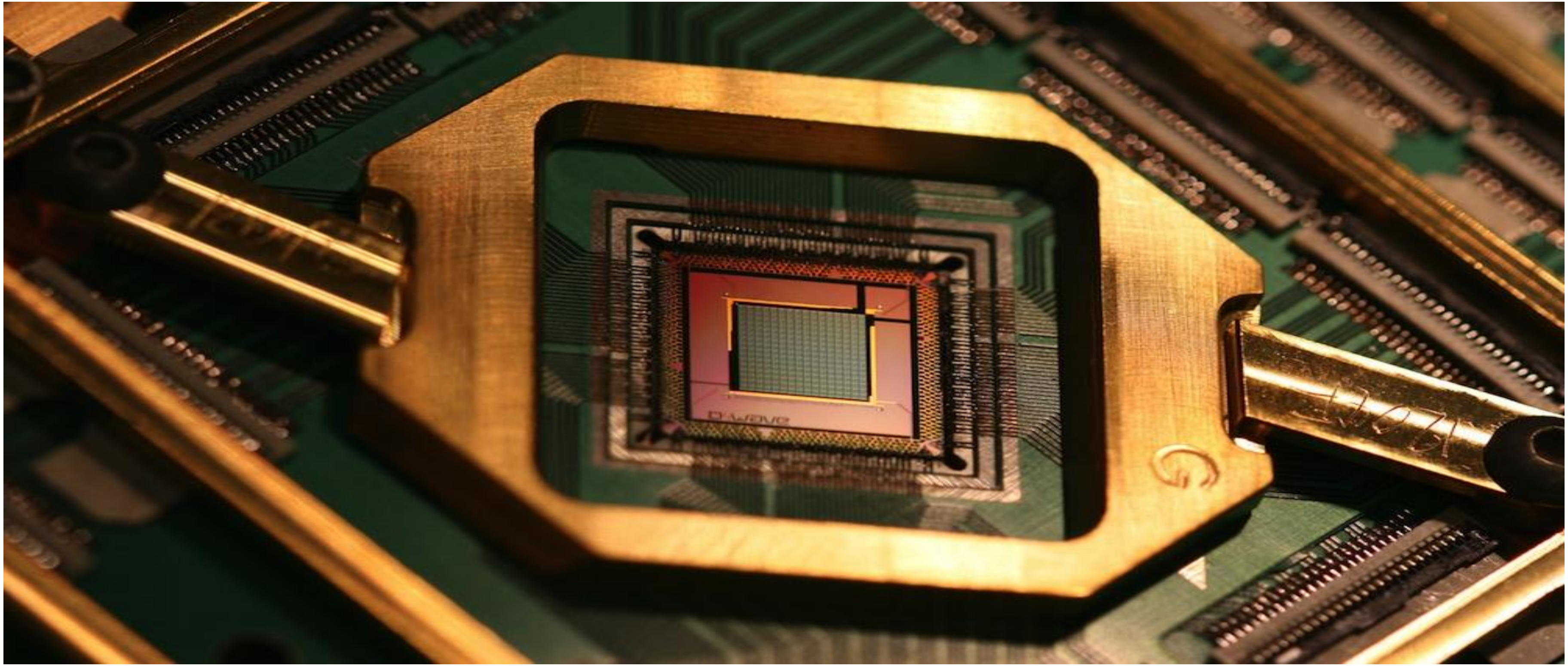


D:Wave
The Quantum Computing Company™

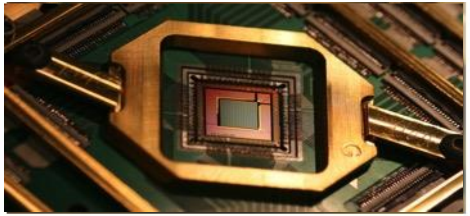
# World’s Most Advanced Quantum Processor



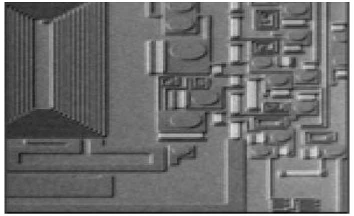
 

# Superconducting yet made in a CMOS Foundry

$1B World Class Production Facility



2000-qubit Circuits at 300,000 JJs



.25µm design rules

ASML 193nm lithography 65nm

# Problem Machine Solves Natively

- Find a vector of spin values, $\mathbf{s}$, that corresponds to a low value of an energy function, $E(\mathbf{s})$

$$\operatorname*{argmin}_{s_1,\cdots,s_n} E(\mathbf{s}) \qquad E(\mathbf{s}) = \sum_i \underset{\text{Local Biases}}{h_i} s_i + \sum_{i,j\in\mathcal{E}} \underset{\text{Couplings}}{J_{i,j}} s_i s_j, \quad s_i = \pm 1$$

easily mapped to a 0/1 variables via $s_i = 2\,x_i - 1$

- Energy values close to Boltzmann distributed (end of anneal) or close to quantum Boltzmann distributed (mid-anneal)

# Functional Quantum Computation Established

- Papers show superposition, entanglement & co-tunneling
  - Johnson et al., "Q. Annealing with Manufactured Spins," Nature 473, 194-198, 12th May (2011).
  - T. Lanting et al., "Cotunneling in pairs of coupled flux qubits," Phys. Rev. B 82, 060512(R) (2010).
  - T. Lanting et al., "Entanglement in a Q. Annealing Processor," Phys. Rev. X 4, 021041 (2014).
- These quantum effects play a functional role in the computations
  - Boixo, et al., "Computational multiqubit tunneling in programmable quantum annealers," Nature Communications 7, Article number: 10327, Published 07 January (2016).
- UCL/USC showed that none of the classical models so far proposed as explanations for the D-Wave machine are correct
  - Albash et al., "Consistency Tests of Classical and Quantum Models for a Quantum Annealer," Phys. Rev. A 91, 042314, Published 13 April (2015).
- USC & D-Wave showed q. annealing can occur successfully on timescales orders of magnitude longer than the coherence time
  - Albash et al., "Decoherence in adiabatic quantum computation," Phys. Rev. A 91, 062320 (2015).
  - N G Dickson et al. "Thermally assisted quantum annealing of a 16-qubit problem", Nature Communications 4, Article number: 1903, 21 May (2013).

# Why is Quantum Computing Exciting?

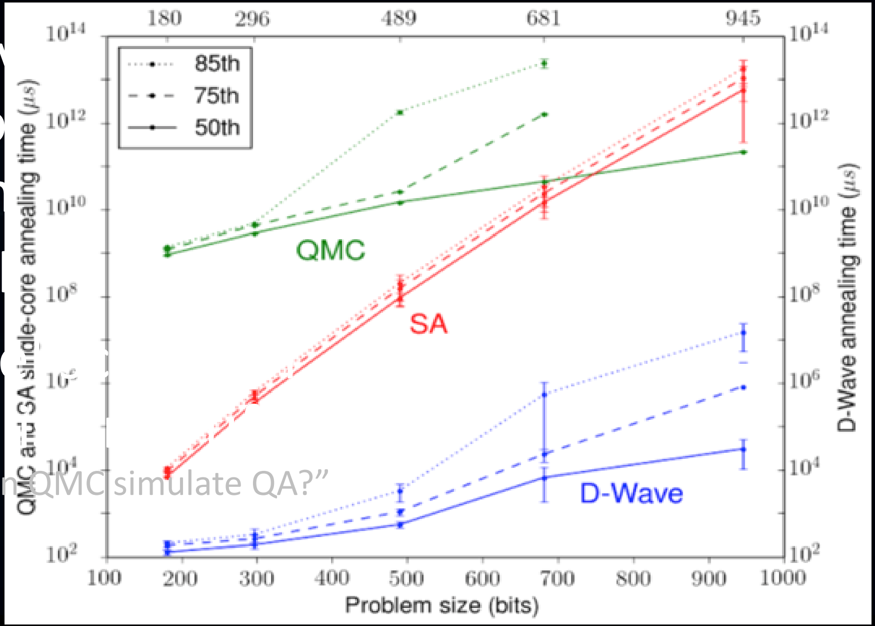
# Potential for Massive Speedups

See “What is the Computational Value of Finite Range Tunneling?” arXiv:1512.02206v3

- Google found D-Wave 2X was 100,000,000x faster than QMC and SA on a particular problem (their “quantumess” test)

- More competiti algorithms beco as we move to h connectivity chi

- Apparent paralle QMC & D-Wave (see E. Andriyash et al. “Can QMC simulate QA?” arXiv:1703.09277)

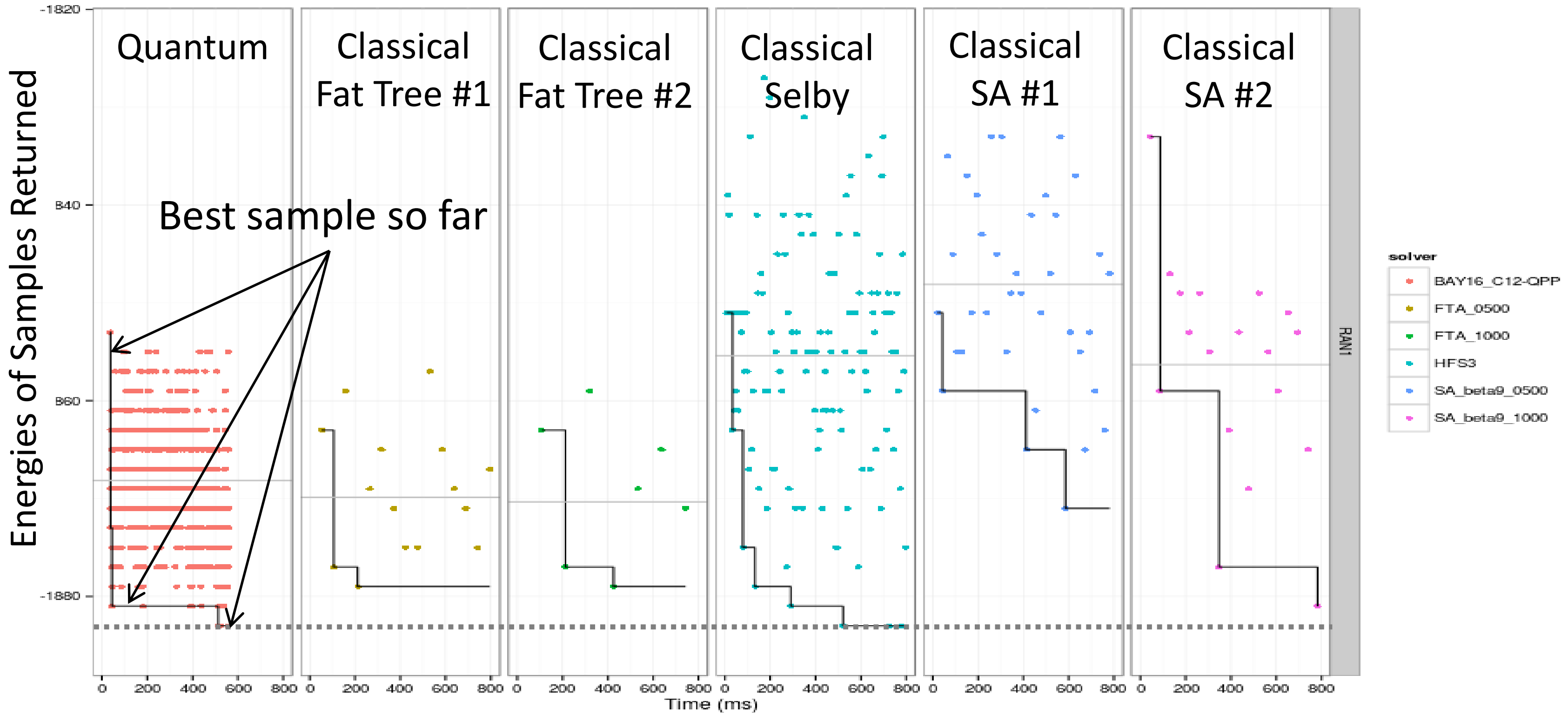






D-Wave The Quantum Computing Company

# Potential for Faster & Better (Lower Energy) Sampling







D::WAVE The Quantum Computing Company™

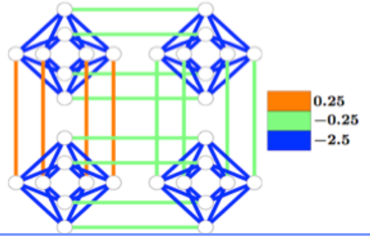
# Quantum Sampling Accelerates Probabilistic ML

D. Korenkevych et al., "Benchmarking Quantum Hardware for Training of Fully Visible Boltzmann Machines," arXiv:1611.04528

## Goal

- Compare rate of learning of a fully visible probabilistic graphical model classically vs. quantumly

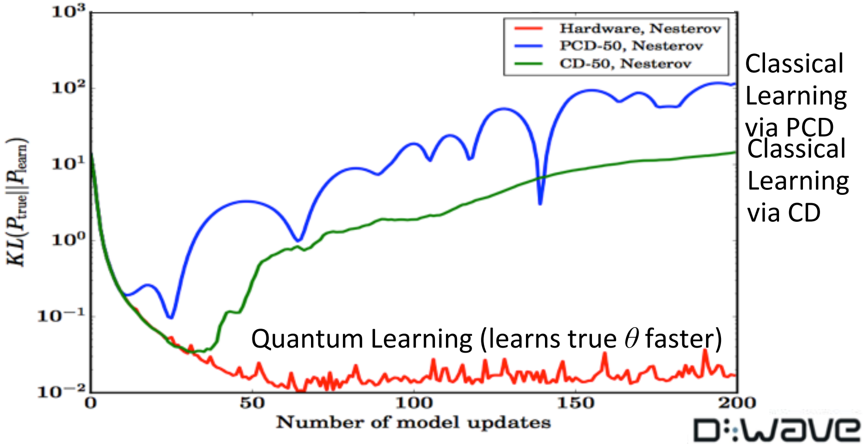
## Model to Learn




## Procedure

- Specify model parameters $\theta_{\text{true}}$, draw exact Boltzmann samples from $\theta_{\text{true}}$, and estimate $\theta$ from samples
- Compare efficacy of CD, PCD, and QA-seeded MCMC chains at estimating the true distribution

## Result: Quantum Learns Faster



# Why Probabilistic Machine Learning Matters

D:Wave
The Quantum Computing Company™

# What Current A.I. Doesn't Do Well

*"Unsupervised learning had a catalytic effect in reviving interest in deep learning, but has since 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."*

Yann LeCun, Yoshua Bengio & Geoffrey Hinton,
"Deep Learning," Nature, Vol. 521, 28th May (2015)

- **A.I. is not yet sufficiently good at unsupervised learning**

D-Wave The Quantum Computing Company™

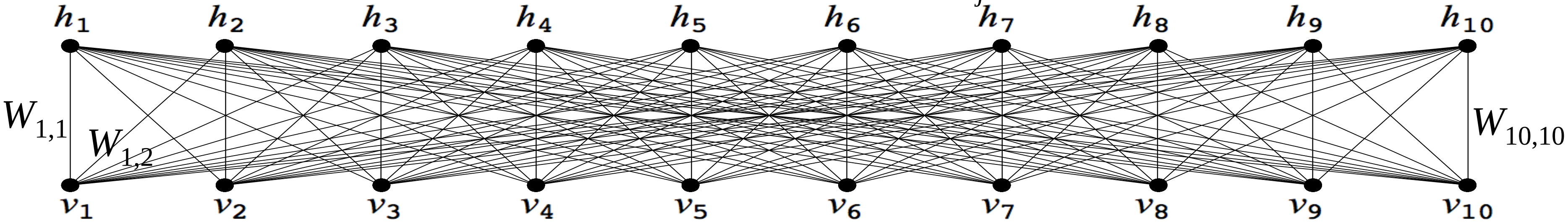
# Quantum Computers Can Help

- Unsupervised learning can use **probabilistic models**
- These rely on **sampling**
- Quantum computers have potential to revolutionize A.I. by making unsupervised learning models feasible to train efficiently
- Because quantum computers are **fast native samplers**

# How does Sampling Arise in Probabilistic Models?

- Consider a Restricted Boltzmann Machine (RBM)
- RBMs can be components of more complex neural networks

Hidden units with biases $c_j$




Visible units with biases $b_i$

Model parameters $q$° $\{b_i\}$È$\{c_j\}$È$\{W_{ij}\}$

# What does Training Entail? Discrete Sampling!

- Given training data (visible vectors) $\boldsymbol{v}_t$ s.t. $1 \leq t \leq T$
- Adjust model parameters s.t. model most likely reproduces the training data
- Done by maximizing the log-likelihood of the observed data distribution w.r.t. the model parameters, $\theta_i$

$$\frac{\partial}{\partial \theta_i}\left(\sum_{t=1}^{T} \log(p(\boldsymbol{v}_t))\right) = -\sum_{t=1}^{T}\left\langle \frac{\partial}{\partial \theta_i} E(\boldsymbol{v}_t, \boldsymbol{h})\right\rangle_{p(\boldsymbol{h}|\boldsymbol{v}_t)} + T\left\langle \frac{\partial}{\partial \theta_i} E(\boldsymbol{v}, \boldsymbol{h})\right\rangle_{p(\boldsymbol{v},\boldsymbol{h})}$$

where $\boldsymbol{\theta} \equiv \{b_i\} \cup \{c_j\} \cup \{W_{ij}\}$

- Positive Phase
- Expectation over $p(\boldsymbol{h}|\boldsymbol{v}_t)$ in “clamped” condition
- Requires *sampling* over the (given) data distribution
- **Simple!**

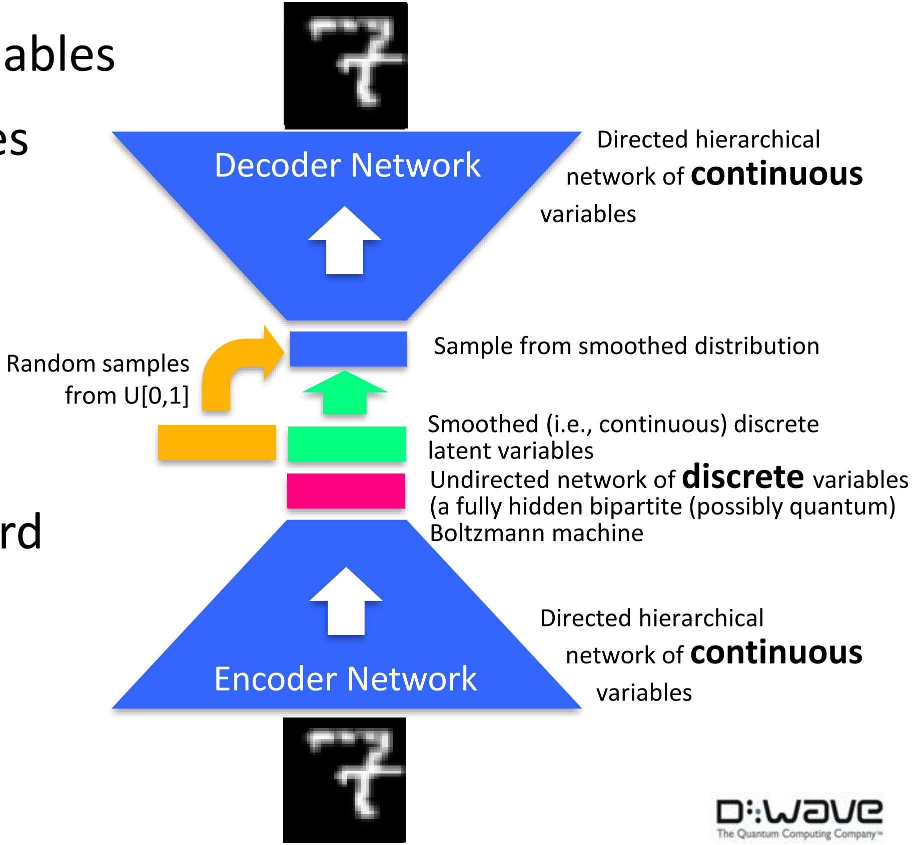
- Negative Phase
- Expectation over $p(\boldsymbol{v}, \boldsymbol{h})$ in “unclamped” condition
- Requires *sampling* over the (predicted) model distribution
- **Intractable!**



D-Wave The Quantum Computing Company™

# Discrete Sampling in Complex Architectures (DVAE/QVAE)

- Real data has discrete & continuous variables
- Natural to want discrete hidden variables
- Can’t backpropagate through discrete variables
- DVAE solves this problem
  - See J. Rolfe, “Discrete Variational Autoencoders”, arXiv:1609.02200
- Exceeds state of the art on three standard machine learning datasets
- DVAE (classical) / QVAE (quantum)



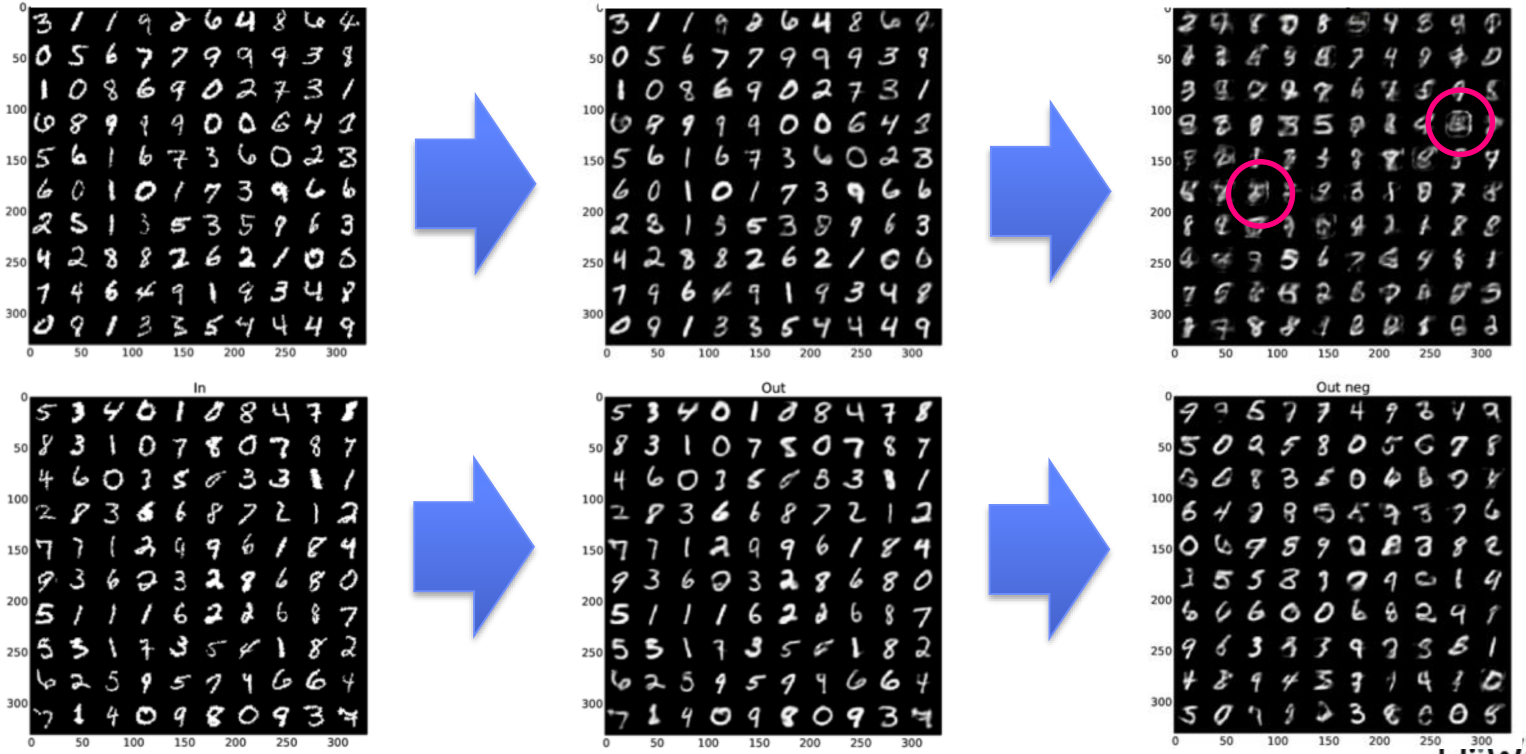
# DVAE Exceeds State-of-the-Art on a Generative Task

J. Rolfe, “Discrete Variational Autoencoders”, arXiv:1609.02200 [stat.ML]

Training Digits | Reconstructed Digits | Machine-Generated Novel Digits

**Previous State of the Art**

**DVAE**

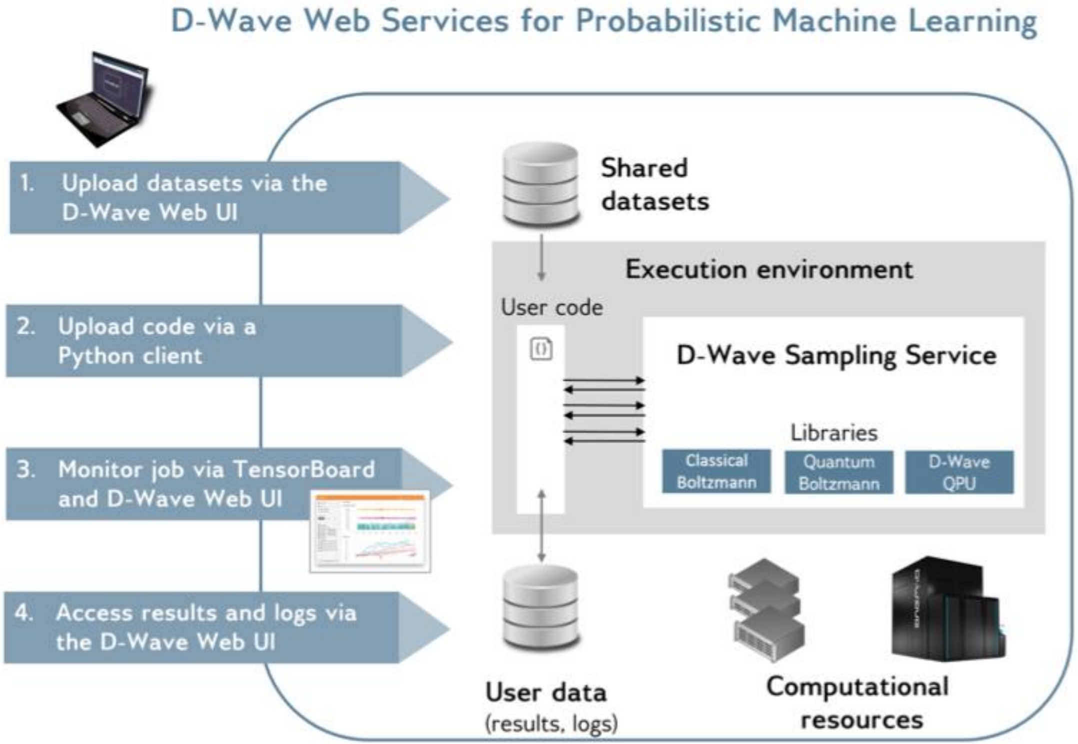






D-Wave The Quantum Computing Company™

# NEW! D-Wave Web Services for Probabilistic Machine Learning




- **D-Wave web services are designed to make it easier to train PML models**
- **Capabilities**
  - Learns from noisy / incomplete data
  - Quantifies confidence in predictions
  - Balances knowns versus unknowns
  - Reveals hidden correlations in data
  - Infers missing data
- **Functionality (Web Services for PML):**
  - Classical Boltzmann sampling (GPU)
  - Quantum Boltzmann sampling (CPU)
  - Raw QPU sampling (QPU)
- **Supports both ML/QML models**
- **Called from TensorFlow or Python**




D-Wave The Quantum Computing Company™

# Conclusions

Contact: cpwilliams@dwavesys.com

- Quantum computing will turbo-charge unsupervised learning
- Quantum and hybrid machine learning models already running
- New (classical) sampling services will be released in 2017
  - Reinvigorate probabilistic machine learning and prepare ground for future quantum & hybrid ML services
  - Available today (classically) / faster tomorrow (quantumly)
- DVAE already surpassing state of the art / QVAE coming in 2018
- **Seeking users for sampling services & probabilistic machine learning models that use them via the cloud**

# Thank you!

Email: cpwilliams@dwavesys.com

D-Wave
The Quantum Computing Company™