Adiabatic Quantum Optimization Processor Performance

Richard Harris

D-Wave Systems Inc.
Burnaby, BC Canada

September 2009

Outline

Outline

Optimization Problem Specification

Let there be an optimization problem of the form

$$E_c(\vec{s}) = -\sum_i h_i s_i + \sum_{i,j>i} K_{ij} s_i s_j \tag{1}$$

where $-1 \leq h_i \leq +1$ and $-1 \leq K_{ij} \leq +1$, for which there exists an optimal solution $\vec{s}$ that minimizes the energy $E_c(\vec{s})$, where the individual elements $s_i = \pm 1$. This problem can be mapped onto a dimensionless Ising spin Hamiltonian of the form

$$\mathcal{H}(t) = -\sum_i h_i \sigma_z^{(i)} + \sum_{\langle i,j \rangle} K_{i,j} \sigma_z^{(i)} \sigma_z^{(j)} \tag{2}$$

Here, the groundstate configuration of the spins $\vec{s}$ represents the lowest energy solution to the optimization problem. The objective is to implement an algorithm in hardware that reliably finds this groundstate.

Adiabatic Quantum Optimization (AQO)

Introduce quantum mechanical degrees of freedom to create a quantum Ising spin Hamiltonian of the form

$$\mathcal{H}(t) = -\sum_i h_i \sigma_z^{(i)} + \sum_{\langle i,j \rangle} K_{i,j} \sigma_z^{(i)} \sigma_z^{(j)} - \sum_i f(t) \sigma_x^{(i)} \quad (3)$$

Use physical system to find the groundstate by evolving $f(t)$ such that

$$\begin{aligned} f(0) &\gg h_i, K_{ij} \\ f(t \to \infty) &\ll h_i, K_{i,j} \end{aligned}$$

See Farhi *et al*, Science **292**, 472 (2001).

Implementation

CJJ rf-SQUID Flux Qubits

Low energy rf-SQUID Hamiltonian can be approximated as a qubit:

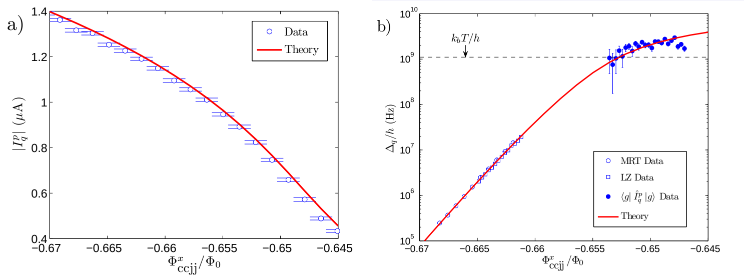
$$\mathcal{H}_q = -\frac{1}{2}\left[\epsilon\sigma_z + \Delta_q\sigma_x\right] \tag{4}$$

where $\epsilon = 2\left|I_q^p\right|\Phi_q^x$ is the bias energy and Δ_q is the tunneling energy. As such, $\left|I_q^p\right|$ and Δ_q are the defining properties of a flux qubit.

We use compound-compound Josephson junction (CCJJ) rf-SQUID flux qubits. Here, an external flux bias Φ_{ccjj}^x controls the tunnel barrier height, thus allowing us to initialize qubits with $\Delta_q \gg k_B T$ and then *annealing* the qubit by slowly raising the tunnel barrier such that $\Delta_q \to 0$. A full disclosure of the CCJJ rf-SQUID flux qubit has been posted as arXiv:0909.4321.

Implementation

Example $|I_q^p|$ and Δ

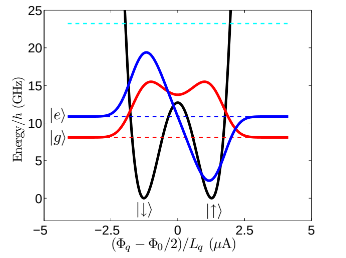


Theory from CCJJ rf-SQUID Hamiltonian with *independently* calibrated parameters. MRT (incoherent regime), see Harris *et al*, PRL **101**, 117003 (2008). LZ (incoherent regime), see Johannson *et al*, PRB **80**, 012507 (2009). $\langle g|\,\hat{I}_q^p\,|g\rangle$ (coherent regime), see Harris *et al*, arXiv:0909.4321.

Implementation

Flux Qubits are Used in the Flux Basis in AQO

Qubits are used in the flux basis in AQO, not the energy eigenbasis as in many (but not all) implementations of gate model quantum computation.



State	Energy Basis	Flux Basis
$\|0\rangle$	$\|g\rangle$	$\|\downarrow\rangle$
$\|1\rangle$	$\|e\rangle$	$\|\uparrow\rangle$
Groundstate	Eigenstate, $\|0\rangle$	$(\|0\rangle + \|1\rangle)/\sqrt{2}$
Excited State	Eigenstate, $\|1\rangle$	$(\|0\rangle - \|1\rangle)/\sqrt{2}$

Implementation

A Network of Inductively Coupled rf-SQUID Flux Qubits

Use rf-SQUID flux qubits as quantum Ising spins. Provide tunable mutual inductances between qubits to implement pairwise interactions. Low energy Hamiltonian for general bias conditions can be written as

$$\begin{aligned}\mathcal{H}(t) &= -\frac{1}{2}\sum_{i}\left[\Phi_i^x(t)2|I_i^{\mathrm{p}}\left(\Phi_{\mathrm{ccjj}}^x(t)\right)|\sigma_z^{(i)} + \Delta_q\left(\Phi_{\mathrm{ccjj}}^x(t)\right)\sigma_x^{(i)}\right] \\ &\quad + \sum_{\langle i,j\rangle} M_{i,j}\left(\Phi_{i,j}^{\mathrm{cpl}}\right)\left|I_q^{\mathrm{p}}\left(\Phi_{\mathrm{ccjj}}^x(t)\right)\right|^2\sigma_z^{(i)}\sigma_z^{(j)} \qquad (5)\end{aligned}$$

$$\begin{aligned}\Phi_i^x, \Phi_{\mathrm{ccjj}}^x &\equiv \text{qubit external flux biases} \\ M_{i,j}(\Phi_{i,j}^{\mathrm{cpl}}) &\equiv \text{tunable mutual inductance between qubits} \\ \Delta_q &\equiv \text{qubit tunnel energy} \\ |I_q^{\mathrm{p}}| &\equiv \text{qubit persistent current}\end{aligned}$$

Implementation

Mapping Problems onto Hardware

Assume that all qubits are identical: $|I_i^p(t)| \rightarrow |I_q^p(t)|$ and $\Delta_i(t) \rightarrow \Delta(t)$. Map problem parameters h_i and K_{ij} in the dimensionless problem Hamiltonian onto the hardware Hamiltonian. Scale problem energies to the antiferromagnetic (AFM) coupling energy.

$$\mathcal{H}(t) = J_{\mathrm{AFM}}(t)\left\{-\sum_i h_i \sigma_z^{(i)} + \sum_{\langle i,j\rangle} K_{i,j}\sigma_z^{(i)}\sigma_z^{(j)}\right\} - \frac{1}{2}\sum_i \Delta(t)\sigma_x^{(i)} \qquad (6)$$

$$\begin{aligned} h_i \times J_{\mathrm{AFM}}(t) &= |I_q^p(t)|\Phi_i^x(t) \\ K_{ij} \times J_{\mathrm{AFM}}(t) &= M_{ij}|I_q^p(t)|^2 \\ J_{\mathrm{AFM}}(t) &= M_{\mathrm{AFM}}|I_q^p(t)|^2 \end{aligned}$$

Implementation

Removing Time-Dependence In Problem Mapping

The problem parameters h_i and K_{ij} are supposed to be *time-independent*. However, the mapping onto physical hardware yielded

$$h_i = \frac{|I_q^p(t)|\Phi_i^q(t)}{M_{\mathrm{AFM}}|I_q^p(t)|^2} = \frac{\Phi_i^x(t)}{M_{\mathrm{AFM}}|I_q^p(t)|} \quad (7)$$

$$K_{ij} = \frac{M_{ij}|I_q^p(t)|^2}{M_{\mathrm{AFM}}|I_q^p(t)|^2} = \frac{M_{ij}}{M_{\mathrm{AFM}}} \quad (8)$$

Clearly K_{ij} are naturally time-independent. However, in order to render h_i time-independent one must apply *time-dependent* qubit flux biases

$$\Phi_i^x(t) = h_i \times M_{\mathrm{AFM}}|I_q^p(t)| \quad (9)$$

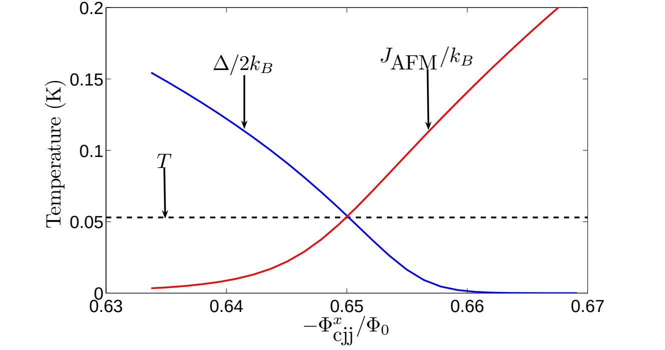
See Harris *et al*, arXiv:0903.3906.

Implementation

Running the AQO Algorithm

To run AQO on this hardware, use given mappings for h_i and K_{ij} and sweep qubit tuning parameters from regime where $\Delta \gg J_{\mathrm{AFM}}$ to $\Delta \ll J_{\mathrm{AFM}}$.

$$\mathcal{H}(t) = J_{\mathrm{AFM}}(t)\left\{-\sum_i h_i \sigma_z^{(i)} + \sum_{\langle i,j\rangle} K_{i,j}\sigma_z^{(i)}\sigma_z^{(j)}\right\} - \frac{1}{2}\sum_i \Delta(t)\sigma_x^{(i)} \quad (10)$$



C1 Processor Architecture

Processor Unit Cell

Hardware has been detailed in several publications. Key references:

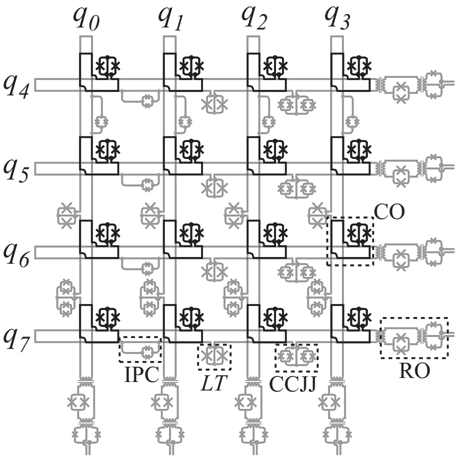
- Programmable Control Circuitry: Johnson *et al*, arXiv:0907.3757
- High Fidelity Readout: Berkley *et al*, arXiv:0905.0891
- Inter-qubit Coupler: Harris *et al*, PRB **80**, 052506 (2009)
- CCJJ Qubit: Harris *et al*, `arXiv:0909.4321`

A new publication describing how this hardware is operated is currently being drafted and will be posted on arXiv soon.

C1 Processor Architecture

Processor Unit Cell

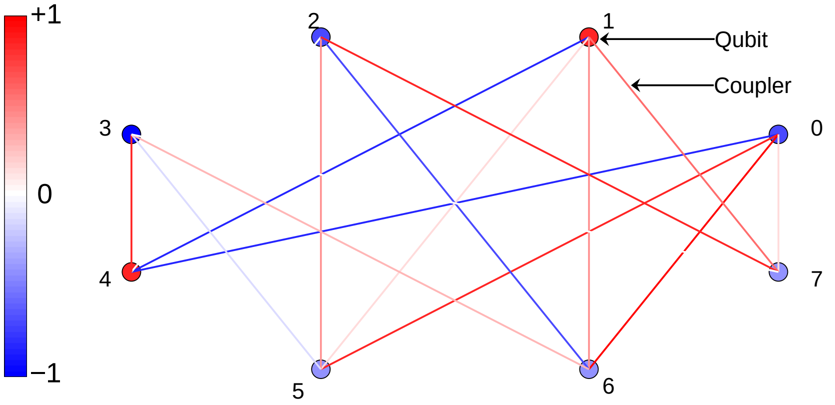
8 flux qubits (shaded) connected by 16 inter-qubit couplers (dark).



Embedding Problems on C1 Processor

Optimization Problems As Graphs

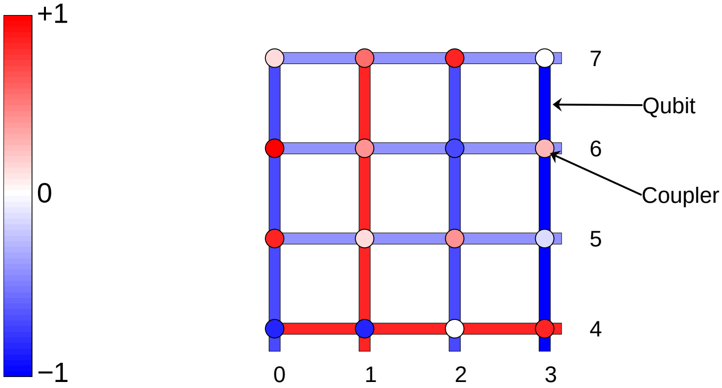
Problems posed as connected graphs. One can assign $-1 \leq h_i \leq +1$ to any vertex and $-1 \leq K_{ij} \leq +1$ to any edge.



Embedding Problems on C1 Processor

Optimization Problems on Hardware

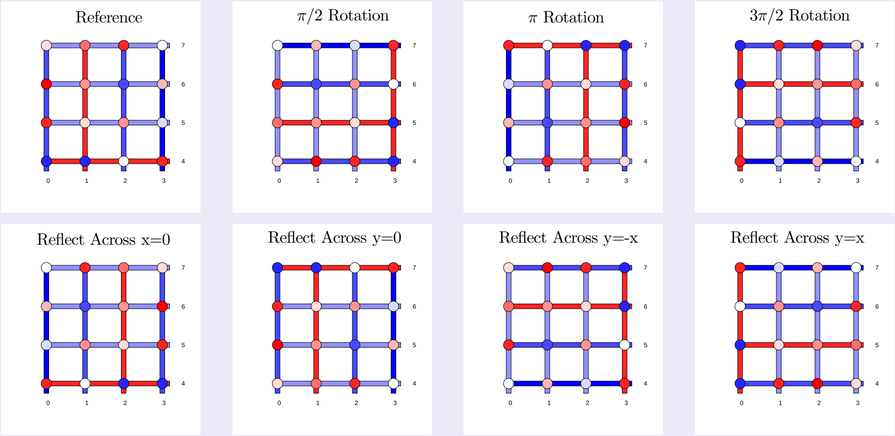
Assign vertices and edges of graphs (problems) to qubits and couplers, respectively. Qubits are extended objects in the hardware and couplers are localized at intersections of qubits.



Embedding Problems on C1 Processor

Symmetry Leads to Multiple Embeddings

There are many ways to embed the same graph. Some examples . . .



Outline

Benchmarking The Hardware

Definition of Benchmark Problem Set

A set of 8 permutations of 800 unique graphs (total 6400 problems) were generated. The graphs were created by choosing sets of 8 values of h_i and 16 values of K_{ij} from a uniform random distribution of values within the set

$$h_i, K_{ij} \in [\ -1 \quad -6/7 \quad -5/7 \quad \dots \quad +5/7 \quad +6/7 \quad +1\ ].$$

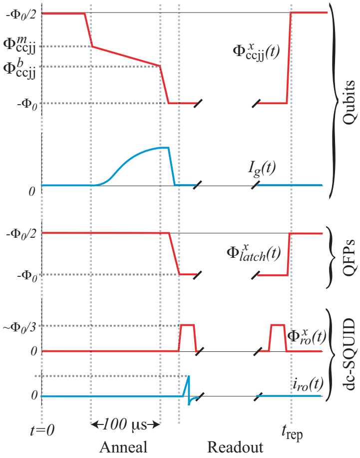
We denote such a problem as being specified to 4 bits of precision (4BOP), where h_i, K_{ij} are sampled from $2^4 - 1$ evenly spaced values between ± 1.

Connections To Complexity and Noise

The precision to which one wants to specify values of h_i and K_{ij} is related to the complexity or difficulty of the optimization problem that is to be solved. The precision to which one can specify values of h_i and K_{ij} in practice is intimately linked to the physics behind dephasing via flux noise in superconducting gate model quantum computing hardware.

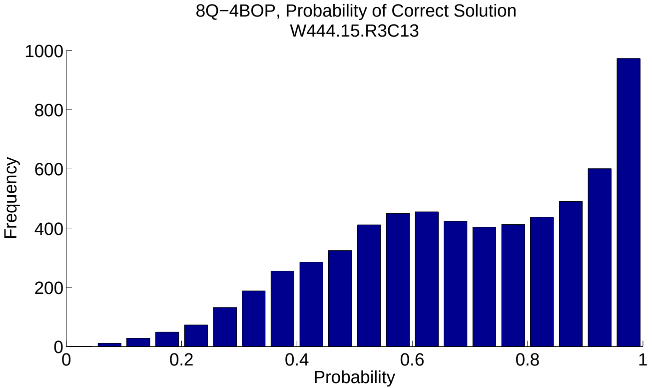
Control Signals

Only 2 time-dependent control signals during annealing: $\Phi^x_{\mathrm{ccjj}}(t)$ and a *single* persistent current compensation signal $I_g(t)$ that simultaneously drives all qubit flux biases $\Phi^x_i(t)$.



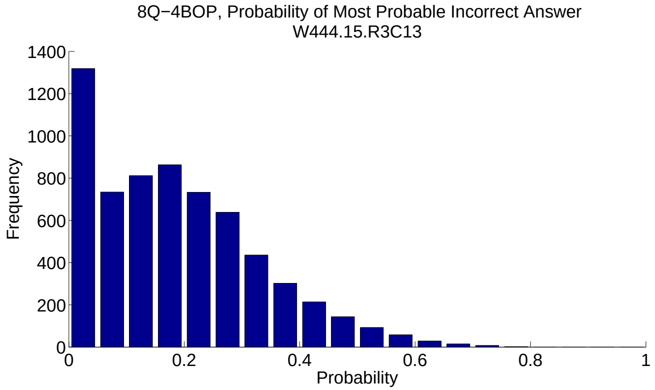
AQC Processor Performance

Adiabatic quantum computation (AQC) implies system should be in groundstate at all times with $P = 1$. Run each of the 6400 problems 1024 times and bin the probability of finding the groundstate. Data show structure around $P \sim 0.6$ and a tail that is rapidly suppressed at lower P.



Failures of the AQC Processor

Track the probability of finding the system in the most prominent incorrect state. If the errors were simply random, P should be very low for all cases. Data indicate otherwise.



Summing It Up Thusfar ...

The Key Experimental Observations

The hardware seemingly only yields a groundstate with $P = 1$ for only a portion of the benchmarking problems. However, the failures do not appear to be random.

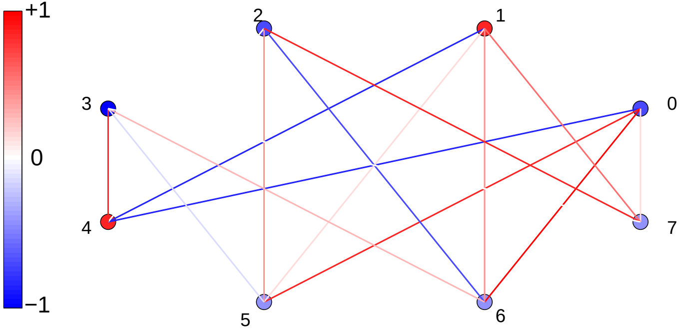
Should one even *expect* this system to return a groundstate with $P = 1$?

The remainder of this lecture will address the question posed above. It will be demonstrated that acknowledging that the AQO processor is operated at finite temperature is essential to understanding its performance.

Outline

A Typical Problem

Consider Problem 793 (permutation 0 pictured below). Graph contains a seemingly typical mix of h_i and K_{ij}, nothing special in its structure. This was a problem that returned the groundstate with $P \approx 0.6 \pm 0.1$ for all permutations.



States Above the Single Count Level

Permutation	State ($\sum_i \frac{s_i+1}{2} 2^{8-i}$)	State ($\vec{s}$)	Probability
0	200	(1 1 -1 -1 1 -1 -1 -1)	563/1024 *
0	201	(1 1 -1 -1 1 -1 -1 1)	91/1024
0	232	(1 1 1 -1 1 -1 -1 -1)	362/1024
.	.	.	.
.	.	.	.
.	.	.	.
7	140	(1 -1 -1 -1 1 1 -1 -1)	656/1024 *
7	142	(1 -1 -1 -1 1 1 1 -1)	212/1024
7	156	(1 -1 -1 1 1 1 -1 -1)	152/1024

Key Observation

Only 3 states showed up for all 8 embeddings of the same problem. The error states are always related to the groundstate (*) by a single bit flip.

A Pattern In the Error States

Permutation	Bits That Fail	Mapping	Transformed Bits
0	2,7	Reference	2,7
1	0,6	Rotated +pi/2	2,7
2	1,4	Rotated pi	2,7
3	3,5	Rotated -pi/2	2,7
4	1,7	Reflect across V	2,7
5	2,4	Reflect across H	2,7
6	0,5	Reflect across x=-y	2,7
7	3,6	Reflect across x=+y	2,7

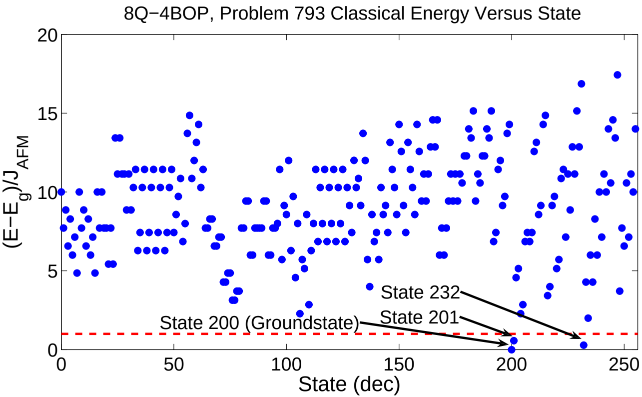
Key Conclusion

The failures are independent of permutation. The loss of probability in the groundstate is due to something intrinsic to the problem structure.

The Failure States are Low Lying States

Choose permutation 0 and calculate the energy for all classical states $\vec{s}$:

$$E_c(\vec{s}) = J_{\mathrm{AFM}} \left[-\sum_{i=0}^{7} h_i s_i + \sum_{i,j>i}^{7} K_{ij} s_i s_j \right] \quad (11)$$



Temperature Matters

Thermalization of the Processor State

If our hardware really does implement an internally adiabatic evolution coupled to a bath at temperature T, then one really ought to observe a thermal distribution of processor states $\vec{s}_i$:

$$P_i = \frac{e^{-E_c(\vec{s}_i)/k_B T}}{\mathcal{Z}} \tag{12}$$

$$\mathcal{Z} = \sum_i e^{-E_c(\vec{s}_i)/k_B T}$$

For the C1 chip, we know $T \approx 50\,\mathrm{mK}$ and $M_{\mathrm{AFM}} = 1.563\,\mathrm{pH}$. Since J_{AFM} grows during annealing, one can *speculate* that the device will successfully thermalize until it enters the incoherent regime where the dynamics will be frozen: the dividing line is roughly located at $\Phi^x_{\mathrm{cjj}} \sim -0.66\,\Phi_0$, at which point $\left|I^p_q\right| \approx 1.1\,\mu\mathrm{A}$. This yields $J_{\mathrm{AFM}}/k_B = M_{\mathrm{AFM}}\left|I^p_q\right|^2/k_B \approx 150\,\mathrm{mK}$.

Temperature Matters

A Quick Sanity Check

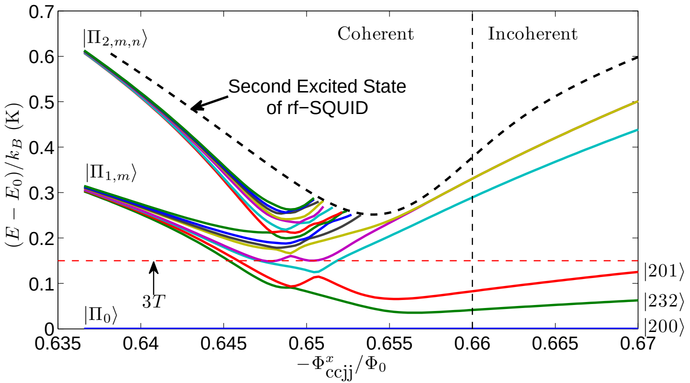
Compare observed (averaged over 8 permutations) and predicted probabilities for problem 793:

State	P Observed	P Predicted
200	0.55	0.58
232	0.35	0.29
201	0.09	0.12
Others	$\ll 0.01$	$\ll 0.01$

The thermalization picture does pick out the right states and gives a decent estimate of their probabilities.

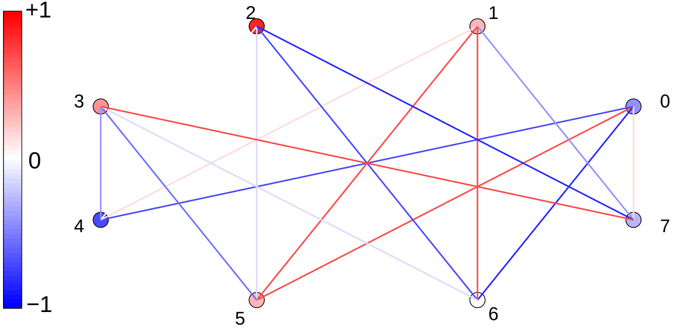
Annealing Trajectories and Thermalization

Consider lowest 16 levels from a simulation of the system running Problem 793: System localizes into classical configurations when $-\Phi^x_{\text{ccjj}} \gtrsim 0.65\,\Phi_0$ (where $\Delta_q \approx 2J_{\text{AFM}}$). We see evidence of the device thermalizing up to $-\Phi^x_{\text{ccjj}} \sim 0.66\,\Phi_0$, which is well beyond the phase transition.

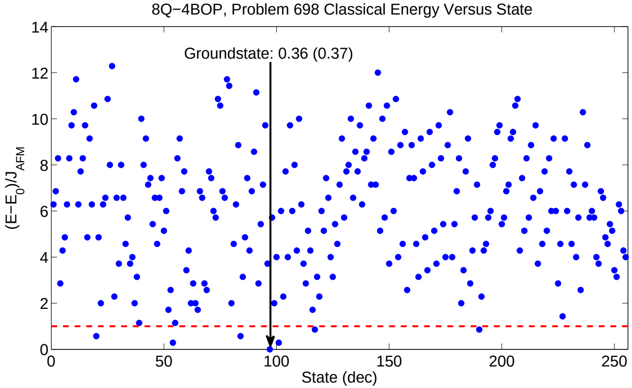


Is This a 1-Qubit Fluke?

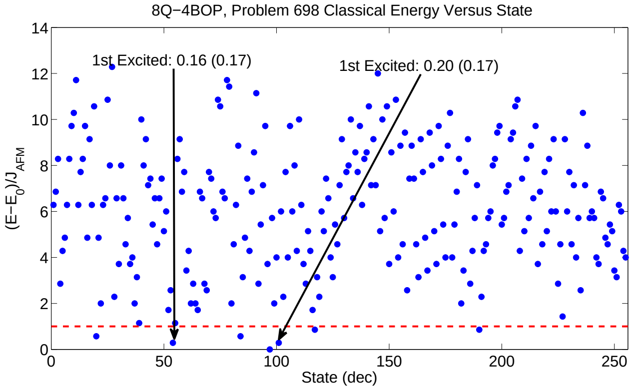
The low lying states for problem 793 were both 1 bit flip from the groundstate. Is there evidence of low lying states that are multiple bit flips from the groundstate being thermally occupied? Consider Problem 698:



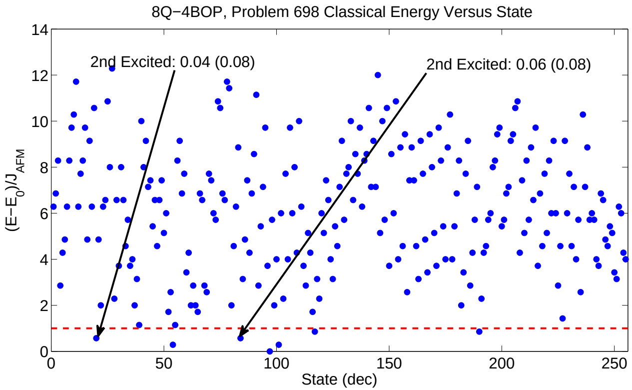
Experimental (Theoretical) Occupation of Ground State



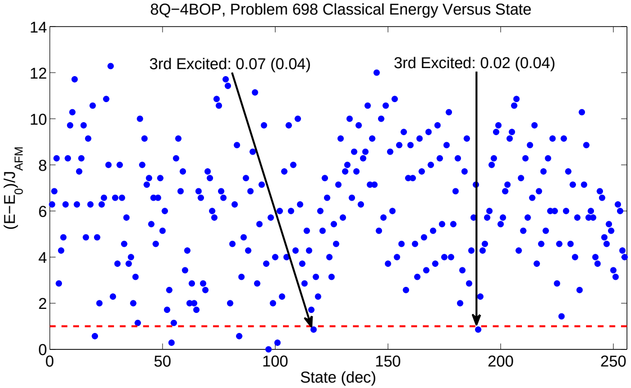
Experimental (Theoretical) Occupation of Excited States



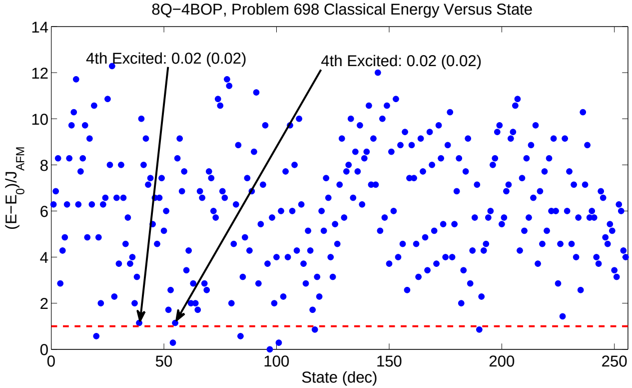
Experimental (Theoretical) Occupation of Excited States



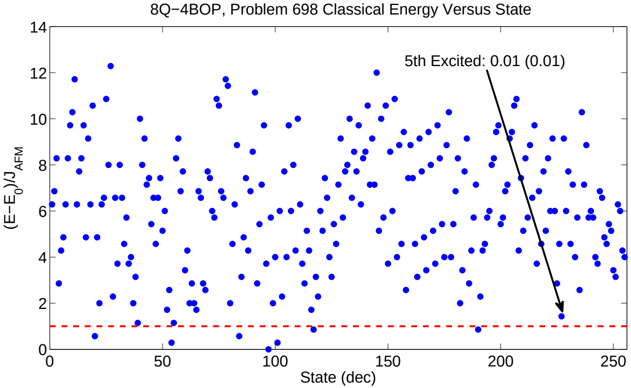
Experimental (Theoretical) Occupation of Excited States



Experimental (Theoretical) Occupation of Excited States



Experimental (Theoretical) Occupation of Excited States



Observation of Multiple Bit Flips

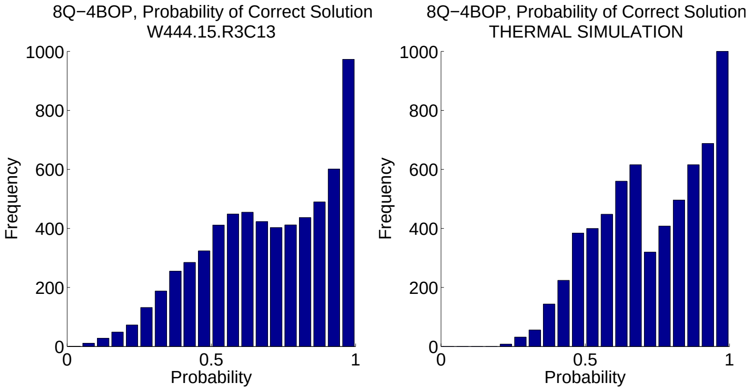
State	Hamming Distance	P Observed	P Predicted
97	-	0.36	0.37
101	1	0.20	0.17
54	5	0.16	0.17
84	4	0.06	0.08
20	5	0.04	0.08
117	2	0.07	0.04
190	7	0.02	0.04
55	4	0.02	0.02
39	4	0.02	0.02
227	2	0.01	0.01
Others	-	$\ll 0.01$	$\ll 0.01$

Multi-bit differences are observed with appreciable probability. The discrepancies between experiment and theory may be due to the target Hamiltonian being slightly miscalibrated.

Simulated Processor Performance

Nature Won't Allow a Perfect AQC ...

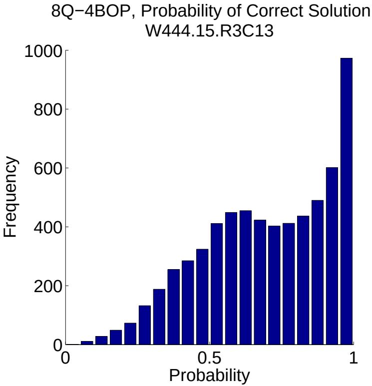
The humped structure appears to be due to a conspiracy between thermalization and problem set structure - it will never go away at finite T.

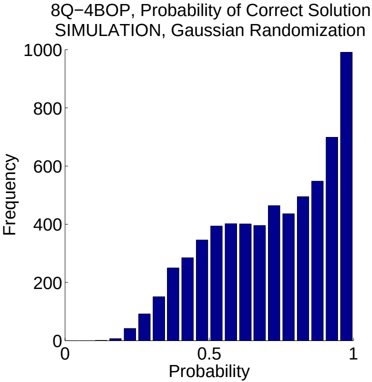


Simulated Processor Performance

Nature Won't Allow a Perfect AQC ...

Adding a small amount of Gaussian distributed random noise on h_i to the simulation reproduces the salient features.





Outline

Choosing a New Performance Metric

AQO $\neq$ AQC when $T > 0$

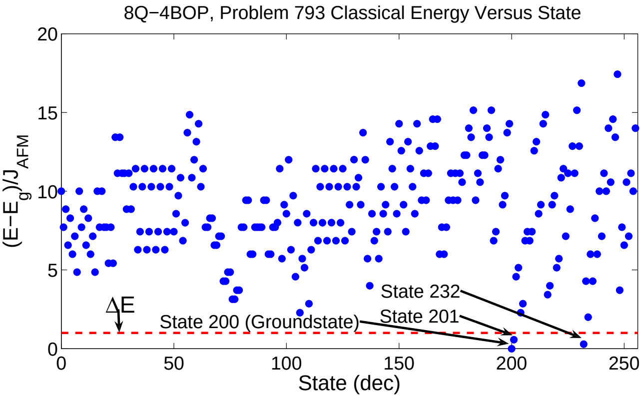
Simply counting the probability of achieving the groundstate alone can distort one's perception of processor performance. When characterizing AQO, one really ought to specify the probability of obtaining a low energy state within a thermal energy window above the groundstate.

So what is a fair width ΔE for this window? Consider a system that has two low lying levels that are well separated from all others. Let us define the success metric as being a probability of 0.95 of being within the groundstate. Solving for the excited state probability then yields:

$$\begin{aligned} 0.05 &= \frac{e^{-\Delta E/k_B T}}{1 + e^{-\Delta E/k_B T}} \\ \Delta E &\approx 3 k_B T \end{aligned}$$

A Dividing Line Between Low and High Energy States

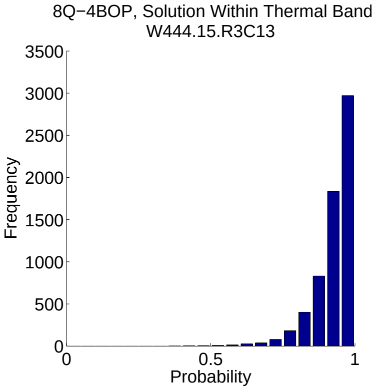
Revisit the classical energy plot for problem 793, permutation 0: The suggested definition of $\Delta E = 3k_B T$ makes sense.

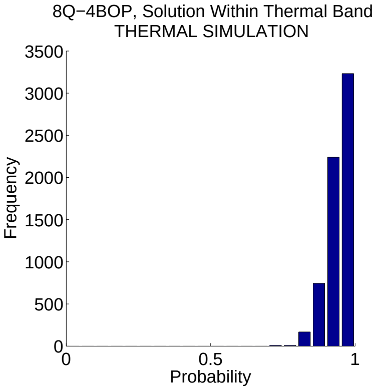


AQO Performance

Low Energy Solutions

The hardware returns low energy solutions with very high probability. This is the expected behaviour of a AQO processor operated at finite T.

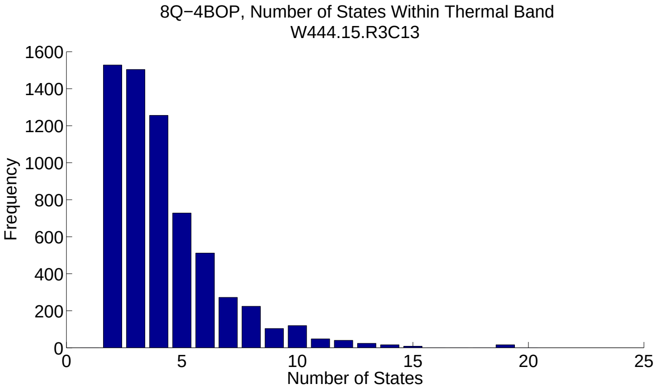




Number of Low Lying States

How many states lie within the thermal window?

The number is $<< 256$, but certainly > 1, which seems to be a feature of the choice of test problems. This means that finding the true groundstate out of relatively few occupied states is easy.



A Scalar Success Metric

Is the groundstate the most probable state?

Thermodynamics dictates that the groundstate ought to be the *most probable* state. We ran each problem 1024 times, identified the state $\vec{s}$ that showed up with the highest probability and then asked whether this state was a groundstate: This was confirmed to be the case for 6073 out of 6400 problems (success rate 95.89%).

Even for those cases where a groundstate was not the most probable state, a groundstate was always observed with finite probability. Since only a small number of states ($\ll 2^8$) showed up with finite probability for all problems in the benchmarking set, then it was very easy to identify an exact solution to a given optimization problem.

Implications for Hardware Operation

Differing Modes of Hardware Operation

Suggested modes for operation of the AQO processor:

- Run problem a statistically large number of times, calculate

$$E_c(\vec{s}) = -\sum_i h_i s_i + \sum_{\langle i,j \rangle} K_{i,j} s_i s_j$$

  for all states $\vec{s}$ that appear above the readout noise level. Take state that yields the lowest E_c as the solution.
- Run hardware once, calculate E_c using measured state $\vec{s}$. If E_c is less than some threshold that is dictated by the user, then answer can be deemed an acceptable solution. If $E_c >$ threshold, iterate.

Number of Repetitions

How many times must one run the hardware to find an exact solution?

Assume for some particular problem that one obtains the groundstate with probability 0.05, so 95% of the time one gets another state. In order to observe the groundstate *at least* once in M measurements with probability P_{gs},

$$1 - P_{\mathrm{gs}} = (0.95)^M \tag{13}$$

P_{gs}	M
0.95	59
0.99	90
0.999	135
...	...

Even in the worst case scenario for this processor, one could find an exact solution to an optimization problem in a small number of iterations.

Outline

Is This a Quantum Processor?

Point 1:

The data shown herein have simply demonstrated that we have built hardware that solves optimization problems in a manner that is consistent with thermodynamics. As such, they cannot be used to argue whether this is a classical or quantum mechanical system as *all* physical systems must obey thermodynamics. Rather, the details must be in *how* the system thermalizes.

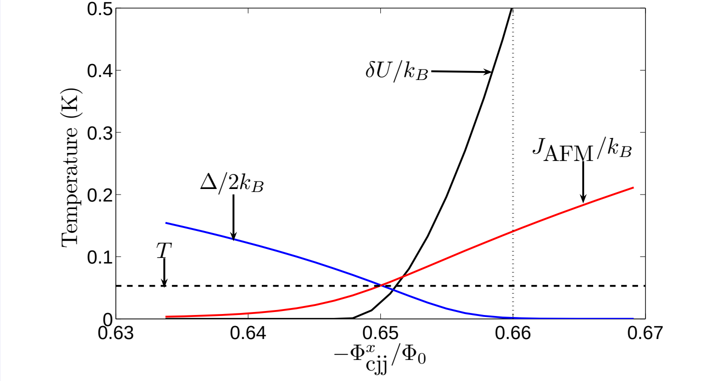
Is This a Quantum Processor?

Point 2:

The fact that measured single qubit parameters $\left|\hat{I}_q^p\right|$ and Δ agree with the results from an independently calibrated quantum mechanical rf-SQUID Hamiltonian indicates that the constituent elements of this processor are quantum mechanical. However, this does not immediately imply quantum mechanical behavior of the collective system. We need to perform specific, targeted experiments to characterize the dynamics of the collective system and compare the results to *quantitative* models, be they quantum or classical, of the system evolution.

... But, System Thermalizes in the Presence of Large Tunnel Barriers

The system thermalizes when the single qubit tunnel barrier height $\delta U \gg k_B T$, but *only* in the presence of finite tunneling energy Δ. We see signatures of thermalization up to $-\Phi^x_{\text{cjj}} \sim 0.66\,\Phi_0$, where $\delta U / k_B T \gtrsim 10$ but $\Delta / h \sim 50\,\text{MHz}$. Thermal activation over such a barrier seems unlikely, but tunneling is still relatively fast.



Outline

Conclusions

- It is critical to include the effects of finite T when discussing the performance of a AQO processor.
- The dominant failure mode for the AQO processor can be described by thermalization processes. These are arguably soft failures to low energy solutions. The width of the band of probable low lying states is given roughly by $\sim 3k_B T$.

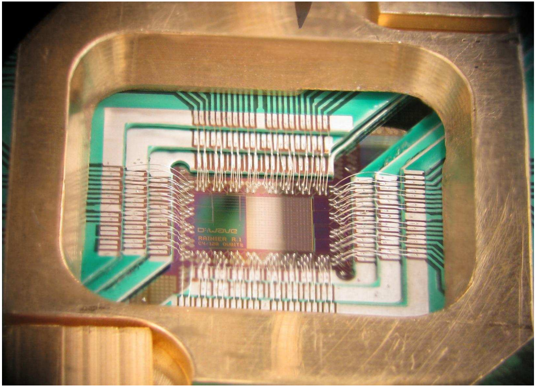
What's Next - The Key Issues

Two Key Issues That Need to Be Addressed:

- **Scaling:** 8-qubit system is too small to state anything meaningful regarding the scaling of AQO. We need to experimentally characterize larger processors at a size that at least pushes the performance of classical optimization systems (hardware and software).
- **Mechanism:** We need to perform carefully crafted experiments to confirm or refute that quantum mechanics is essential to the processor operation. This boils down to determining whether the off-diagonal elements of the single qubit Hamiltonians (Δ_q) affect the dynamics of the collective processor.

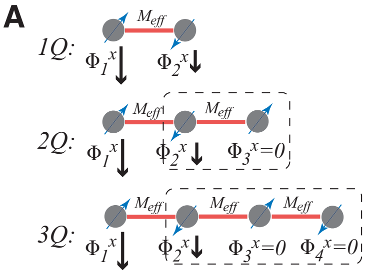
What's Next - Scaling

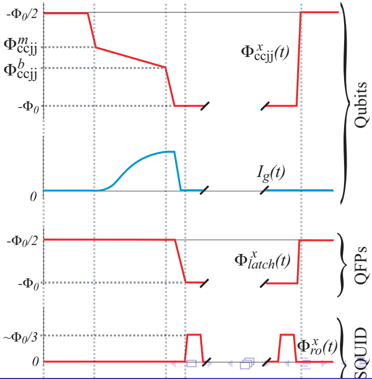
We currently have three 128-qubit chips cold and being calibrated at the moment. Our intention is to benchmark these devices against large sets of random Ising spin glass problems.



What's Next - Mechanism

We are looking into using antiferromagentic chains of varying length to probe the dynamics of collective systems. Such chains have 2 low lying states that are well separated from all others, thus simplifying modelling.





What's Next - Mechanism

By carefully choosing *static* flux biases applied to the antiferromagnetic chain one can steer the magnitude of the gap at the anticrossing $0 < g < \Delta_q$. This provides a means for slowing down the dynamics of the collective system such that it can be probed with our low bandwidth bias lines.