Parameter Setting for Quantum Annealing

D-Wave Users Group
28 September 2016

Kristen Pudenz, Ph.D.
Senior Quantum Applications Engineer
Limitations on the problem type define quantum annealing as a subset of adiabatic quantum computation

\[ H(t) = A(t)H_X + B(t)H_{Ising} \]

\[ H_X = \sum_i \sigma_i^x \]

begin in a uniform superposition of all computational states

\[ H_{Ising} = \sum_i h_i \sigma_i^z + \sum_{i,j \in \text{edges}} J_{ij} \sigma_i^z \sigma_j^z \]

finish in the ground state of an Ising spin glass over some graph
Mixed SAT

Boolean Satisfiability: Is there an assignment of variables that makes a given Boolean formula evaluate to TRUE? If so, what is it?

Multiple variants based on the clause structure of the formula.

Mixed SAT has clauses of multiple lengths.

\[(b_1 \overline{b}_2) \land (\overline{b}_5 b_4 b_2 \overline{b}_7 b_6) \land (b_3 b_6 b_1)\]

Most versions of SAT with more than 3 variables per clause are hard

Mixed SAT problems can be a product of a software verification approach we are pursuing, and therefore become an application of interest.

We’re interested in finding multiple solutions to these problems.

Problems were generated randomly, then restricted to those with less than a million solutions.
For the purposes of this study, we specify a logical Hamiltonian and a physical Hamiltonian using spin variables $s_i$

$$H_{\text{logical}} = \sum_{i=1}^{n} h_i s_i + \sum_{(i,j) \in E_l} J_{i,j} s_i s_j$$

$$s_i \in \{-1, 1\}$$

$$H_{\text{physical}} = \sum_{i=1}^{n} \sum_{k=1}^{K_i} h_{i,k} s_{i,k} + \sum_{(i,j) \in E_l} \sum_{(i,k;j,m) \in E_p} J_{i,k;j,m} s_{i,k} s_{j,m} + H_{\text{chain}}$$

$$H_{\text{chain}} = \sum_{i=1}^{n} \sum_{(i,k;i,m) \in E_p} c S_{i,k} s_{i,m}$$
Parameter Setting Strategies (2/2)

logical Hamiltonian

logical coupler

logical bias

chain couplings

physical Hamiltonian

physical coupler

physical bias
Choose one physical device to represent each logical problem term

**logical Hamiltonian**

\[ h_{i,k} = \begin{cases} h_i, & \text{if } k \text{ selected} \\ 0, & \text{otherwise} \end{cases} \]

\[ J_{i,k;j,m} = \begin{cases} J_{ij}, & \text{if } (i, k; j, m) \text{ selected} \\ 0, & \text{otherwise} \end{cases} \]

**physical Hamiltonian**

physical qubit selected due to adjacent physical coupler(s)
Distribute logical problem terms evenly over the physical devices available to represent them.

**logical Hamiltonian**

$$h_{i,k} = \frac{h_i}{K_i}$$

$$J_{i,k;j,m} = \frac{J_{ij}}{|(i,k;j,m) \in E_p|}$$

**physical Hamiltonian**

$$h_{1/4}$$

$$h_{1/4}$$

$$h_{1/4}$$

$$h_{2/3}$$

$$h_{2/3}$$

$$h_{1/4}$$

$$J_{12}/2$$

$$J_{12}/2$$

$$c$$
Assign physical bias to qubits via a weight calculated by the number of physical couplers assigned to that qubit.

**logical Hamiltonian**

\[ w_{i,k} = \frac{d_{i,k}}{D_i} \]

number of couplers on qubit \( i,k \)

\[ h_{i,k} = h_i w_{i,k} \]

**physical Hamiltonian**

\[ J_{i,k;j,m} = \frac{J_{i,j}}{|(i,k;j,m) \in E_p|} \]
Assign physical bias to qubits via a weight calculated by the number of physical couplers assigned to that qubit, but put some bias on each physical qubit.

**logical Hamiltonian**

\[
h_{i,k} = h_{\text{min}} \text{sign}(h_i) + h_{\text{remainder}} w_{i,k}
\]

\[
h_{\text{remainder}} = h_i - K_i h_{\text{min}} \text{sign}(h_i)
\]

\[
J_{i,k;j,m} = \frac{J_{ij}}{|(i, k; j, m) \in E_p|}
\]

**physical Hamiltonian**

\[
h'_{1/2} = h_1 - 2 h_{\text{min}}
\]

\[
h'_{2/2} = h_2 - h_{\text{min}}
\]
Decoding Strategies

**Single qubit:** use the state of the qubit with the highest weight

\[ w_{i,k} = \frac{d_{i,k}}{D_i} \]

- number of couplers on qubit \( i,k \)
- total couplers on logical qubit \( i \)

**Majority vote:** take the state of the majority of the physical qubits

\[ \text{vote result } \rightarrow v_i = \text{sign} \left( \sum_{k=1}^{K_i} s_{i,k} \right) \]

**Weighted majority vote:** give the higher weight qubits more influence

\[ v_i = \text{sign} \left( \sum_{k=1}^{K_i} w_{i,k} s_{i,k} \right) \]

We use all strategies together because they are computationally trivial
We see the best performance at chain coupling $c = 1.6$ across all parameter setting strategies.

Results over DW2 problem set; 927 instances of sizes $n = \{10, 20\}$
On the newer, larger chip, we also see the best performance at chain coupling $c = 1.6$ across all parameter setting strategies.

Results over DW2X problem set; 123 instances of sizes $n = 30$
Median Success by Strategy

The single strategy’s performance falls off particularly sharply for high chain couplings.

DW2 problems
\[ n = \{10, 20\} \]

DW2X problems
\[ n = 30 \]
It’s useful to characterize the performance of the different parameter setting strategies by considering the minimum parameter distance (MPD) they produce.

MPD is the smallest distance between two physical terms in the final Hamiltonian.

\[ MPD = \min_{ijkl} |A_{ij} - A'_{kl}| \]

Small MPD can lead to parameters crossing in implementation, i.e.

\[ A_{ij} < A_{kl} \quad A'_{ij} > A'_{kl} \]

where \( A_{ij} \) and \( A_{kl} \) are the terms specified, and \( A'_{ij}, A'_{kl} \) are what was actually implemented in hardware.
MPD for the single strategy falls off very sharply as chain coupling magnitude increases. The other strategies are more consistent.

Results over DW2 problem set; 927 instances of sizes $n = \{10, 20\}$
MPD for the DW2X problem set behaves similarly to that for the DW2 problem set, because the same parameter setting strategies are used.

Results over DW2X problem set; 123 instances of sizes $n = 30$
Effect of Spin Reversal Transformations

Most problem instances yield more unique answers when subjected to a SRT than when re-parameterized.

Results over DW2X problem set; 123 instances of sizes $n = 30$
For an overwhelming majority of instances, all SRTs studied yield an enhancement in success probability over the identity SRT.

Results over DW2X problem set; 123 instances of sizes $n = 30$
Conclusions

We have identified parameter setting strategies with differentiated performance in distinct chain coupling regimes.

Specifically, the single device parameter setting strategy is more effective for lower chain couplings.

Minimum parameter distance (MPD), a quantity that is easy to calculate, is a strong indicator of the performance of any parameter setting strategy.

Using multiple spin reversal transformations (SRTs) is a better way to boost success probability than using multiple parameter settings.
Any SAT clause of size $k$ can be represented in Ising form using a “cascading-OR” construction involving $2(k-1)$ qubits.

For two variables, the penalty is simple:

$$H_2(b_1, b_2) = -\sigma_{b_1}^z - \sigma_{b_2}^z + \sigma_{b_1}^z \sigma_{b_2}^z$$

Another important building block is the penalty function for an OR gate:

$$H_{OR}(b_1, b_2, a) = \sigma_{b_1}^z + \sigma_{b_2}^z - 2\sigma_a^z + \sigma_{b_1}^z \sigma_a^z - 2\sigma_{b_2}^z \sigma_a^z$$
We construct a three variable clause by substituting the output bit $a$ from $H_{OR}$ for one of the variables in the two variable penalty:

$$H_3(b_1, b_2, b_3) = H_{OR}(b_1, b_2, a) + H_2(a, b_3)$$
A four variable clause can be represented by attaching an OR gate to any of the active variables from $H_3$:

Clauses of any size can be built in this way, and addition of multiple clause penalty functions serves as the AND between the disjunctive clauses.
Adiabatic Quantum Computation

Solve a hard problem by slowly changing the system Hamiltonian

\[ H(t) = A(t) H_{\text{begin}} + B(t) H_{\text{problem}} \]

begin in an easy-to-construct ground state

finish in a hard-to-find ground state

\[ A(0) \gg B(0) \]
\[ A(T) \ll B(T) \]
\[ 0 < t < T \]
\[ T \sim \frac{1}{g_{\text{min}}^2} \]

computation time depends on energy spectrum

Philipp Hauke et. al. arXiv:1411.7933
AQC Advantages and Disadvantages

Advantages:
- Computationally universal
- No sharp transitions; thermodynamically stable
- Inherent robustness to noise
  - Adiabatic evolution is stable
  - System can relax from excited states
- Compatible with open-loop error suppression
- Error correction under active development

Disadvantages:
- No algorithms designed for AQC with exponential speedup
  - Factoring can be translated with polynomial overhead
- Threshold theorem from circuit model QC does not apply
  - We don’t have a winning strategy for error correction in AQC