A robust approach to QUBO on the chimera topology

Witold Jarnicki

BEIT

KQIS, October 5th, 2021

Introduction

The title

A robust approach to QUBO on the chimera topology

The plan

- Bit graph topologies (“chimera topology”)
- QUBO problem statement (“QUBO”)
- Some applications
- Obstacles for hardware
- Understanding the hardware
- Our reference solution (“robust approach”)
- Scaling up

Bit graphs

Naming

- V — set of $n := |V|$ variables (vertices, qubits).
- K_V — complete graph on V.
- $E \subseteq K_V$ — topology of edges (couplings).
- $S := \{0, 1\}^n$ — bit states.

Bit graphs

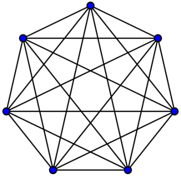
Naming

- V — set of $n := |V|$ variables (vertices, qubits).
- K_V — complete graph on V.
- $E \subseteq K_V$ — topology of edges (couplings).
- $S := \{0, 1\}^n$ — bit states.

Conventions

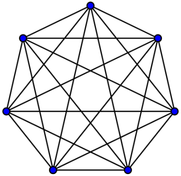
- $V = \{0, 1, \ldots, n-1\}$.
- $E = K_n := K_V$ — complete graph.
- $E = C_k$ — Chimera graph.
- $E = C_{k,\ell}$ — Chimera fragment.
- $E = P_k$ — Pegasus graph.

Bit graphs

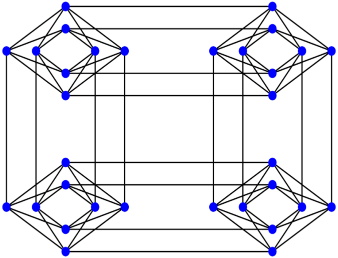


Complete graph K_7

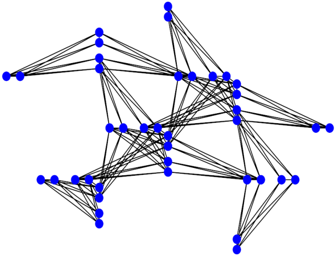
Bit graphs



Complete graph K_7



Chimera graph C_2



Pegasus graph P_2

QUBO

Weights

- $b : V \longrightarrow \mathbb{R}$ — variable biases.
- $c : E \longrightarrow \mathbb{R}$ — coupling weights.
- $\beta \in \mathbb{R}_{>0}$ — Boltzmann distribution parameter.

QUBO

Weights

- $b : V \longrightarrow \mathbb{R}$ — variable biases.
- $c : E \longrightarrow \mathbb{R}$ — coupling weights.
- $\beta \in \mathbb{R}_{>0}$ — Boltzmann distribution parameter.

State functions

- $\mathcal{E} : S \longrightarrow \mathbb{R}$ — state energy,

$$\mathcal{E}(x) := \sum_{\alpha \in V} b(\alpha) x_\alpha + \sum_{\{\alpha,\beta\} \in E} c(\{\alpha, \beta\}) x_\alpha x_\beta.$$

- $W : S \ni x \longmapsto e^{-\beta \mathcal{E}(x)} \in \mathbb{R}_{>0}$ — state weight.
- $Z := \sum_{x \in S} W(x)$ — partition function.
- $p : S \ni x \longmapsto W(x)/Z \in (0, 1]$ — Boltzmann probability distribution.

QUBO

Boltzmann distribution weights

- $W(x) := e^{-\beta \mathcal{E}(x)}$.
- Strongly prefers low energy states when sampling.
- May or may not find a global minimum.
- Used for Quantum Boltzmann Machines.

Some applications

Combinatorial problems

- Potentially lowest value of energy attained $\Longleftrightarrow$ Boolean formula satisfied.
- SAT.
- Factorization.

Some applications

Combinatorial problems

- Potentially lowest value of energy attained $\iff$ Boolean formula satisfied.
- SAT.
- Factorization.

Optimization problems

- Cost optimization.
- Max-variants of combinatorial problems e.g. MAX-SAT.

Some applications

Combinatorial problems

- Potentially lowest value of energy attained $\iff$ Boolean formula satisfied.
- SAT.
- Factorization.

Optimization problems

- Cost optimization.
- Max-variants of combinatorial problems e.g. MAX-SAT.

Boltzmann machines

"Here are some cats. Show me another cat."

Obstacles for hardware

Noisiness

- Low-order weights ignored in practice.
- No official data — needs experimental research.

Obstacles for hardware

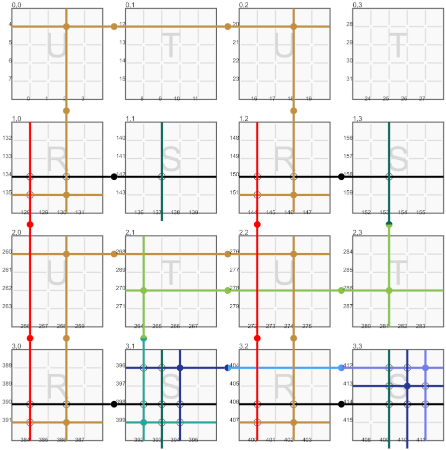
Noisiness

- Low-order weights ignored in practice.
- No official data — needs experimental research.

Topology

- Most problems cannot be directly embedded in target topology.
- Bit chaining required.

Bit chaining



A 2–bit $\times$ 2–bit $\longmapsto$ 4–bit factorization/multiplication circuit.
Embedded into C_4 — colors correspond to logical bits.

Obstacles for hardware

Problems with bit chaining

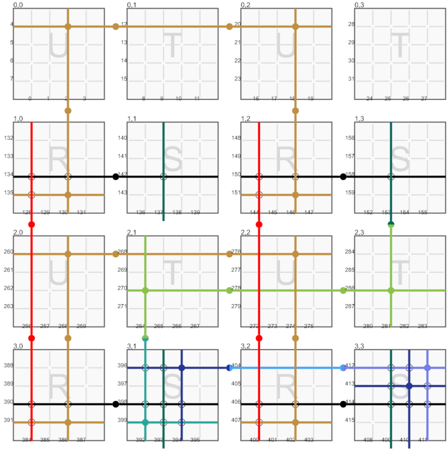
- Low-priority chaining $\Longrightarrow$ solutions cannot be mapped to logical bits.
- High-priority chaining $\Longrightarrow$ non-solutions (noisiness).
- Strongly limits the number of bits e.g. 2048–bit C_{16} can embed at most a 65–bit clique.
- Attempts at hardware-level chaining to counter these issues (Pegasus).

Understanding the hardware

Questions

- How close to the ideal Boltzmann distribution are they?
- How close to the minimum energy they get?
- How noisy are they in general?

Let us prove that 3 is prime



A 2–bit $\times$ 2–bit $\longmapsto$ 4–bit factorization/multiplication circuit.

Hardware sampling depends on more than energy

Factorize 3 by sampling the circuit 1000 times

Attempt	1×3 ($\mathcal{E}_{\text{min}}$)	3×1 ($\mathcal{E}_{\text{min}}$)	failed ($> \mathcal{E}_{\text{min}}$)
1	444	14	542
2	387	61	552
3	307	78	615
4	406	31	563
5	282	96	622
6	240	64	696
7	278	52	670
8	159	117	724
9	190	174	636
10	439	49	512

Hardware sampling depends on more than energy

Use 10 rounds of random bit flips

Attempt	1×3 ($\mathcal{E}_{\text{min}}$)	3×1 ($\mathcal{E}_{\text{min}}$)	failed ($> \mathcal{E}_{\text{min}}$)
1	189	65	746
2	193	67	740
3	230	90	680
4	160	91	749
5	225	54	721
6	232	111	657
7	280	60	660
8	162	76	762
9	166	90	744
10	204	96	700

Attaining minimum energy

Tailored example

- 2001 bits.
- Minimum theoretical energy -1000.
- 10000 samples.
- Sampled energy range $[-690, -538]$.
- Random bit flips do not help, either.

Attaining minimum energy

Tailored example

- 2001 bits.
- Minimum theoretical energy -1000.
- 10000 samples.
- Sampled energy range $[-690, -538]$.
- Random bit flips do not help, either.

However

- Not really a noisiness issue, as the theoretical energy distribution is binomial.
- Problem trivially solved by hybrid approaches involving local search.

Precision of $\beta\mathcal{E}$

Boltzmann distribution

- $W : S \ni x \longmapsto e^{-\beta\mathcal{E}(x)} \in \mathbb{R}_{>0}$ — state weight.
- $Z := \sum_{x \in S} W(x)$ — partition function.
- $p : S \ni x \longmapsto W(x)/Z \in (0, 1]$ — Boltzmann probability distribution.

Precision of $\beta\mathcal{E}$

Boltzmann distribution

- $W : S \ni x \longmapsto e^{-\beta\mathcal{E}(x)} \in \mathbb{R}_{>0}$ — state weight.
- $Z := \sum_{x \in S} W(x)$ — partition function.
- $p : S \ni x \longmapsto W(x)/Z \in (0, 1]$ — Boltzmann probability distribution.

Experiment

- Set variable bias to 0.
- Sample variable 10000 times.
- Repeat $10\times$ for 2000 qubits, resulting in 20000 frequencies.

Precision of $\beta\mathcal{E}$

Boltzmann distribution

- $W : S \ni x \longmapsto e^{-\beta\mathcal{E}(x)} \in \mathbb{R}_{>0}$ — state weight.
- $Z := \sum_{x \in S} W(x)$ — partition function.
- $p : S \ni x \longmapsto W(x)/Z \in (0, 1]$ — Boltzmann probability distribution.

Experiment

- Set variable bias to 0.
- Sample variable 10000 times.
- Repeat $10\times$ for 2000 qubits, resulting in 20000 frequencies.

Result

- Frequency ranging from 30% to 70%.
- Value of $\beta\mathcal{E} = 0$ interpreted as one between $[-0.8, 0.8]$.
- Resolution of $\beta\mathcal{E}$ seems to be even integers, but no finer.

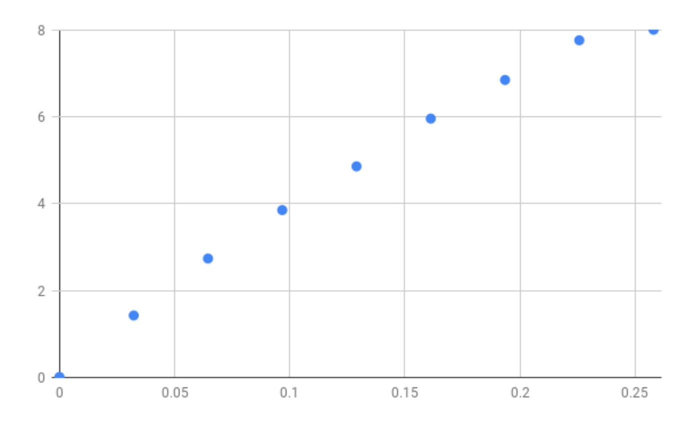
Determining β

Experiment

- Pick some bias values $(b_1, \ldots, b_m)$ spanning $[0, 1]$.
- Determine experimental probability p_j of getting value 1 by sampling multiple qubits with bias p_j, $j = 1, \ldots, m$ (around half a million samples per j).
- Throw away outliers with probabilities too close to 0 or 1.
- As $p_j \simeq 1 - 1/(1 + e^{-\beta b_j})$, determine the experimental value y_j of βb_j by putting $y_j := -\ln(1/(1 - p_j) - 1)$.
- Plot y_j as a function of b_j and use linear regression to determine β.

Determining β

Result



Value of $y_j \simeq \beta b_j$ as a function of b_j. Experimentally, $\beta \simeq 32$.

Random QUBO

Experiment

- Generate a QUBO instance on $C_{6,16}$ choosing biases and couplings uniformly among even numbers in $[-30, 30]$.
- Compute the minimum energy value and its probability using the exact reference solver.
- Sample 100 thousand points from the hardware implementation.

Random QUBO

Experiment

- Generate a QUBO instance on $C_{6,16}$ choosing biases and couplings uniformly among even numbers in $[-30, 30]$.
- Compute the minimum energy value and its probability using the exact reference solver.
- Sample 100 thousand points from the hardware implementation.

Result

- The exact minimum energy of -12000 has probability ~ 0.09 in the Boltzmann distribution, assuming the value of $\beta = 32$ applied to the normalized biases.
- Sampling 100 thousand points failed to produce the minimum energy value. The minimum energy attained was -11992.

Conclusions

Questions

- How close to the ideal Boltzmann distribution are they?
- How close to the minimum energy they get?
- How noisy are they in general?

Conclusions

Questions

- How close to the ideal Boltzmann distribution are they?
- How close to the minimum energy they get?
- How noisy are they in general?

Answers

- Even in random cases, the probabilities seem to be off by many orders of magnitude.
- Sometimes the minimum is not found, even if the theoretical probability guarantees that.
- The noise of $\beta\mathcal{E}$ seems to be on the level of ± 0.8, but rounding βE to even integers still produces results that we cannot explain.

Our QUBO exact reference solver

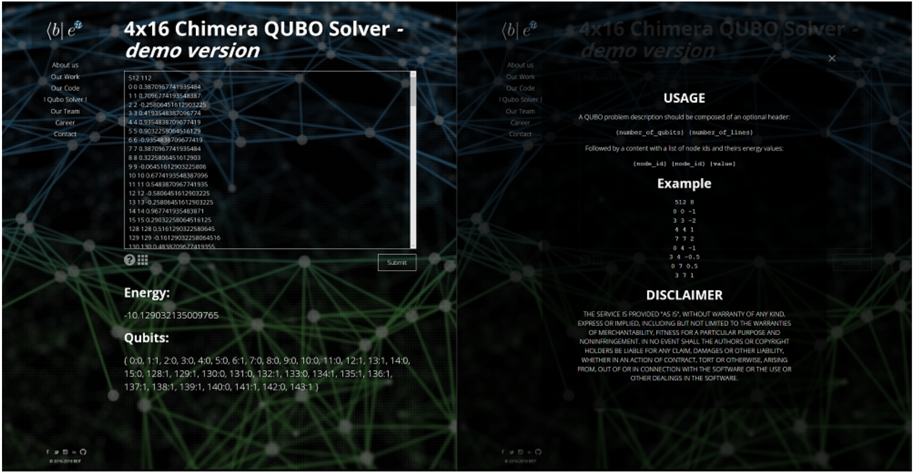
Methods

- Dynamic programming on bounded treewidth graphs.
- Cannot do full C_{16}.
- Uses some ideas by Alex Selby (minimum energy value).
- Computes much more using our new approach.

Our QUBO exact reference solver

Current state for $C_{k,\ell}$ ($k, \ell \in \mathbb{N}$, $k \leq \ell$)

Capability	CPU	GPU
Maximum k	7	8
Typical runtime	$\sim$ 1 minute	$\sim$ 10 seconds
Minimum energy value	$\mathcal{O}(\ell 2^{4k})$	
Minimum energy probability	—	$\mathcal{O}(\ell 2^{4k})$
Minimum energy state	$\mathcal{O}(\ell 2^{4k})$	$\mathcal{O}(\ell^2 2^{4k})$
Boltzmann sample	—	$\mathcal{O}(\ell^2 2^{4k})$
m Boltzmann samples	$\mathcal{O}(k\ell m + \ell^2 2^{4k})$	



Online QUBO solver demo:
finds minimum energy value and state on $C_{4,16}$.

Scaling up

- Our exact solver does not scale that well.
- It can be used to validate hardware platforms or other heuristics.
- Biggest hope lies in hybrid local-search or MCMC solutions.

Thank you for your attention