Using a quantum annealer for particle tracking at the LHC

on behalf of the HEP.QPR project at LBL

Lucy Linder
April 3, 2019

Quantum Annealing & D-Wave
QA in D-Wave computers

qubits ⇒ $q_i$

bias weights ⇒ $a_i$

coupling strength ⇒ $b_{ij}$

quantum machine instruction (QMI)

objective function:

$$O(a; b; q) = \sum_{i=1}^{N} a_i q_i + \sum_{i}^{N} \sum_{j}^{N} b_{ij} q_i q_j \quad q_i \in \{0, 1\}$$

QUBO

Quadratic Unconstrained Binary Optimisation

source: [dwavesys on YouTube](https://www.youtube.com/watch?v=...)
D-Wave hardware & usage

D-Wave 2X (2015)
- 1152 qubits
- 3360 couplers

D-Wave 2000Q (2017)
- 2048 qubits
- 6016 couplers

chimera graph

problem space

problem representation

$$E(x|Q) = \sum_i x_i Q_{ij} x_j$$

BQM (QMI)

SAPI server

API exists

sampling

?post-processing?

minor embedding

un-embedding
QUBO for track reconstruction
"We can regard a track with \( n \) hits as a set of \( n-1 \) consecutive lines [doublets] with a smooth shape and without bifurcation".

1. generate the set of potential doublets (apply early cuts)

2. binary classification task to determine which doublets should be kept in the track candidates, using Hopfield Networks
“We can regard a track with \( n \) hits as a set of \( n-1 \) consecutive lines [doublets] with a smooth shape and without bifurcation”.

1. (generate the set of potential doublets (apply early cuts))
   → use a Python adaptation of the ATLAS online seeding GPU code

2. *binary classification task* to determine which doublets should be kept in the track candidates, using Hopfield Networks Quantum Annealing
Internally, focus on triplets of hits: $T_{a,b,c}$

Triplets can form interesting pairs or be in conflict.

binary classification of triplets
→ any $P_t$
→ any vertex location

unrelated triplets

quality pair

invalid pair

conflicting triplets
The energy function

Select the best triplets, in the form of pairs, that will constitute the track candidates:

\[ E = \alpha \left( \sum_{i}^{N} T_i \right) - \left( \sum_{i,j}^{\text{pairs}} S_{ij}T_iT_j \right) + \zeta \left( \sum_{i,k}^{\text{conflicts}} T_iT_k \right) \quad T \in \{0, 1\} \]

- “bias weight” triplet prior can be set to 0
- “connection strength”, interest of [keeping] a pair of triplets
- avoid “conflicts”, a hit belongs to at most one track
The energy function: QUBO

\[ E = \alpha \left( \sum_{i}^{N} T_i \right) - \left( \sum_{i,j \in \text{qplets}} S_{ij} T_i T_j \right) + \zeta \left( \sum_{i,k \in \text{conflicts}} T_i T_k \right) \quad T \in \{0, 1\} \]

\[ O(a; b; T) = \alpha \left( \sum_{i}^{N} T_i \right) + \sum_{i} \sum_{i < j} b_{ij} T_i T_j \quad T \in \{0, 1\} \]

\[ b_{i,j} = \begin{cases} 
-S_{ij}, & \text{if } (T_i, T_j) \text{ form a pair that should form a track}, \\
\zeta, & \text{if } (T_i, T_j) \text{ are in conflict}, \\
0, & \text{otherwise}. 
\end{cases} \]
A track of \( n \) hits is a set of \( n-2 \) triplets that can be combined into \( n-3 \) pairs.

A set of track candidates is a set of triplets with no conflict.

A triplet \( T_{a,b,c} \) has a higher quality when:

- it has little to no hole: \( H = 0 \)
- the **Menger curvature** \( \text{curv}(a,b,c) \) formed by the three hits in the X-Y plane is small;
- doublets \( ab \) and \( bc \) have similar \( \theta \) angles:
  \[
  \text{drz}(T_{a,b,c}) = |\varsigma(\theta_{ab}, \theta_{bc})| \text{ is small}
  \]

A pair \( (T_i, T_j) \) has a higher quality when:

- it has few holes: \( H_{ij} = 0 \)
- there are similar curvatures:
  \[
  \text{dcurv}_{ij} = |\varsigma(\text{curv}(T_i), \text{curv}(T_j))| \text{ is small}
  \]
- hits are aligned in R-Z:
  \[
  \text{drz}_{ij} = \max(\text{drz}(T_i), \text{drz}(T_j)) \text{ is small}
  \]

The interest of connecting two triplets into a pair can be expressed as:

\[
S_{ij} = \alpha \left( \beta (1 - \text{dcurv}_{ij})^\gamma + (1 - \beta)(1 - \text{drz}_{ij})^\delta \right) / (1 + H_{ij})^\epsilon
\]

\[
S_{ij} = \frac{1 - \frac{1}{2}(\text{dcurv}_{ij} + \text{drz}_{ij})}{(1 + H_{ij})^2}
\]

\begin{align*}
\beta &= 0.5, \\
\epsilon &= 2, \\
\text{others} &= 1
\end{align*}
Experimental setup

Dataset

- **TrackML dataset** (== HL-LHC) with events split into lower multiplicity datasets:
  - select P% of particles
  - select P% of noise

Set weight=0 for hits belonging to particles with:
  - $P_T < 1$ GeV or
  - less than 5 hits

Metrics

- TrackML score
- precision (~purity)
- recall (~efficiency)

Machines

- CORI (1 Haswell node)
- D-Wave 2000Q (leap)
- D-Wave 2X (LANL)

false negative = missings
false positive = fakes

tune the model for that!
iterative hybrid classical/quantum algorithm

QBSOLV

large and/or densely connected QUBOs split into sub-QUBOs fitting the D-Wave hardware.

Tabu search on the recomposed solutions.

The solution is sometimes randomised to escape local minima.
Algorithm overview

potential doublets

filter doublets
create triplets
create pairs
build QUBO

solve QUBO

kept triplets

doublets
trajectories

precision
recall
trackml score

final doublets
forming particle trajectories

preprocessing / model building
solving
postprocessing
scoring
Results

Dataset size: ~20%
1,637 particles, 11,030 hits

Plotting error: too many doublets 392,529

392,529 doublets
p=0.26%, r=99.15%

2,546 doublets
(2,964 triplets)
QUBO size: 14,345

1,512 doublets
p=99.13%, r=97.06%

TrackML score 97.55%

Running on CPU
Performances at low Pt

186 particles in a phi slice of π/3

precision (%): 98.5, recall (%): 98.4, trackml score (%): 98.35

QUBO size 68,043

59,077 ⇒ 752 doublets
Physics performance

Full TrackML event 6,900 - 14,000 particles, Pt $\geq$ 150 MeV, ~15% noise/lower Pt hits

*recall*: no endcaps & no double hits, focus on p_t $\geq$ 1 GeV BUT no hypothesis on vertex location

> 90% recall (efficiency) and trackml score on doublets classification
What about the fakes?

“The biggest difference [with conventional methods] is the number of wrongly associated coordinates [...] soft constraints and very simple geometrical constants is not as good as a sophisticated algorithm based on hard constraints (fits)

Stimpfl-Abele & Garrido, fast track finding with neural networks

→ post-processing should let us filter many fakes
Improvement (TrackML-only)

fixed variable bias weights
penalty based on the triplet’s impact parameters (D0, Z0)

⚠ introduces biases and reintroduces a hypothesis on the vertex location

94%+ precision !

⚠ introduces biases and reintroduces a hypothesis on the vertex location
Future work

improvements & tuning

tests

benchmarks & speedup

• hyperparameter + code optimization
• reduce fake tracks post-processing, other models, ...
• include more properties e.g. magnetic field,

bigger datasets ←
noise level ←
qbsolv alternatives ←

D-Wave
Fujitsu digital annealer
other QUBO solvers

• study physics performance
• study timing performance
Resources

HEP.QPR website
https://hep-qpr.lbl.gov

publication preprint
https://arxiv.org/abs/1902.08324

Github repository (and results)
https://github.com/derlin/hepqpr-qallse

Master Thesis report
https://github.com/derlin/hepqpr-qallse/tree/master/doc

» arXiv:1902.08324 «
Thank you for your attention
Timing performance

Model building

QUBO solving

D-Wave  internet latency, shared QPU, qbsolv .... too early to say !

In the meantime, classical counterparts are surprisingly efficient.

neal (dwavesys)

6'600 particles/event = 12 s.
260,744 QUBO parameters

currently no optimization or parallelization
**Basics of QA**

*adiabatic theorem* of quantum mechanics,

a physical system remains in the ground state if

\[ \text{perturbations are slow} \]
\[ \text{there is a gap} \]

**QA recipe:**

initial Hamiltonian \( H_0 \)

problem Hamiltonian \( H_p \)

\[ H(s) = A(s)H_0 + B(s)H_p \] with

\[ A(s) \downarrow \text{ and } B(s) \uparrow \] given \( s = t/t_f \leq 1 \) (\( t_f = \text{anneal time} \))

If the adiabatic conditions are respected,
the system’s ground state at \( t = t_f \) encodes the solution.
QUBO composition

![Graph showing QUBO composition with parameters, bias weight, connection strength, and conflicts.]
D-Wave timing

results on QallseMp

- total
- dwave
- service
- QPU

QPU access time:

- qpu_sampling_time
- qpu_delay_time_per_sample
- qpu_readout_time_per_sample
- qpu_anneal_time_per_sample
- qpu_programming_time
A slightly different approach

Triplets combined three-by-three.

Model A

In the QUBO, use the same formula for the coupling T1-T2, but a new formula based on the stdev for T1-T3 (+ sum multiple scores): ~ +5% precision on a 60%-dataset.

Model B

Use the stdev formula for all pairs: ~ +15% precision on a 60%-dataset.

Credits: Alex Smith
Should we discard small tracks?

Scores after removing all tracks with 5 hits from the output.
Universal quantum computers
quantum circuits

|ψ⟩

H

|0⟩A

H

|0⟩B

Z

IBM 50Q
Rigetti 19Q
Intel Tangle Lake (49Q)
Google Bristlecone (72Q)
...

» arXiv:1902.00498 «

Quantum annealers

D-Wave (2000Q)

» arXiv:1902.08324 «