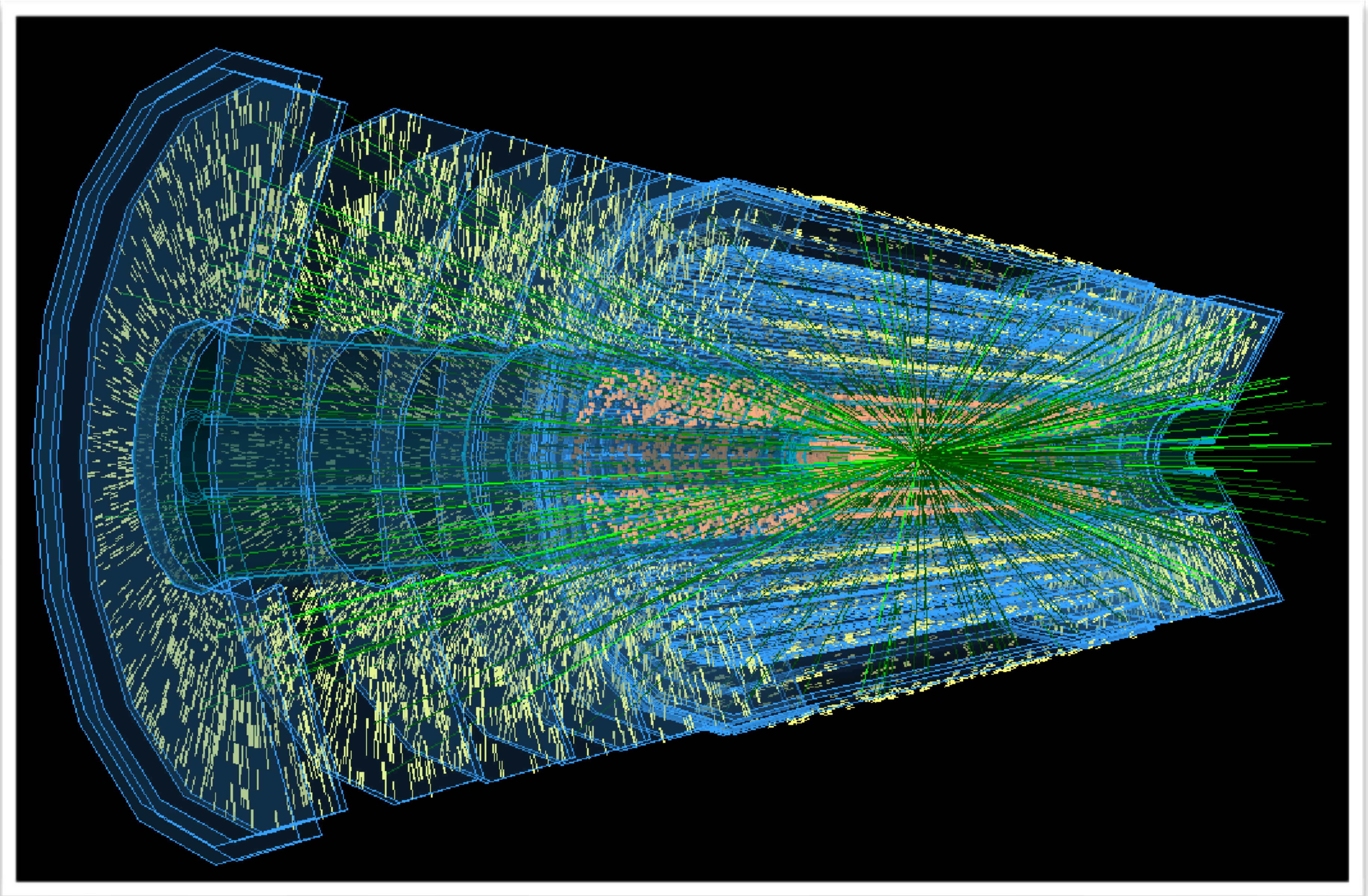
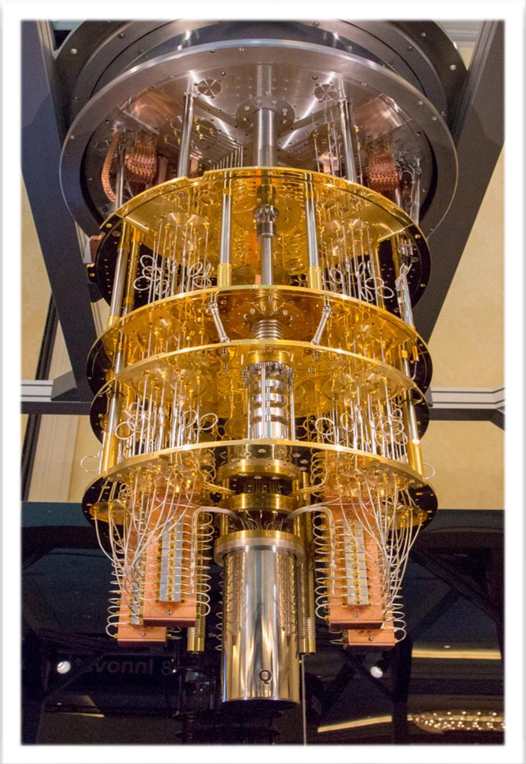
Recent developments in quantum computing and its application to high-energy physics





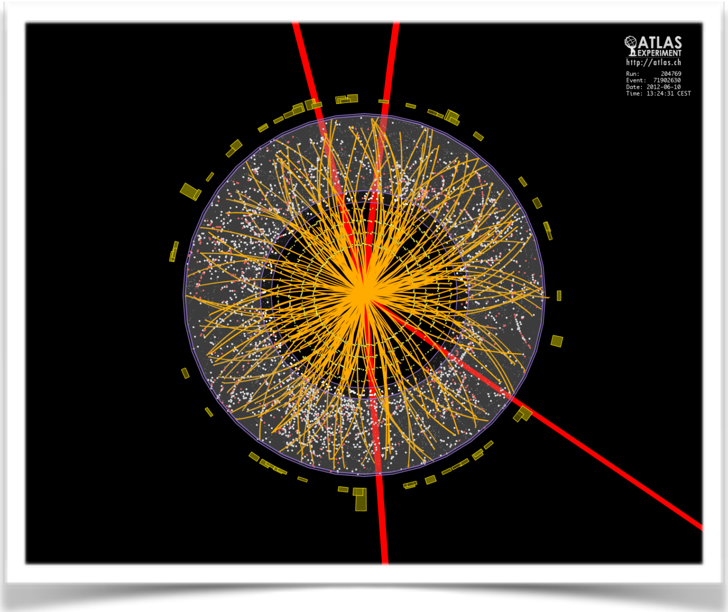
Heather M. Gray, UC Berkeley/LBNL

Talk Outline

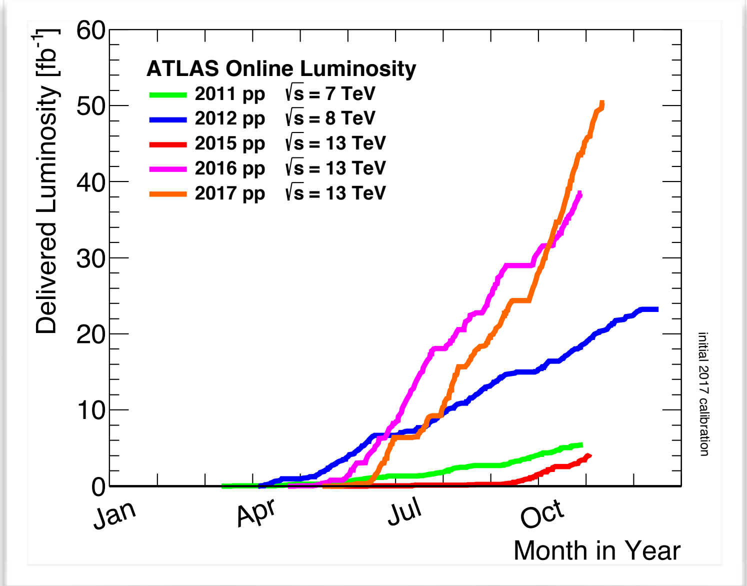
- Computing challenge of the HL-LHC
- Recent developments in quantum computing
  - IBM-Q
  - D-Wave
- Ideas for areas in HEP using quantum computing
  - Monte Carlo simulation
  - Track reconstruction
  - Analysis via machine learning

Coming soon: HL-LHC

First beam in ATLAS
(2009)



Higgs discovery
(2012)

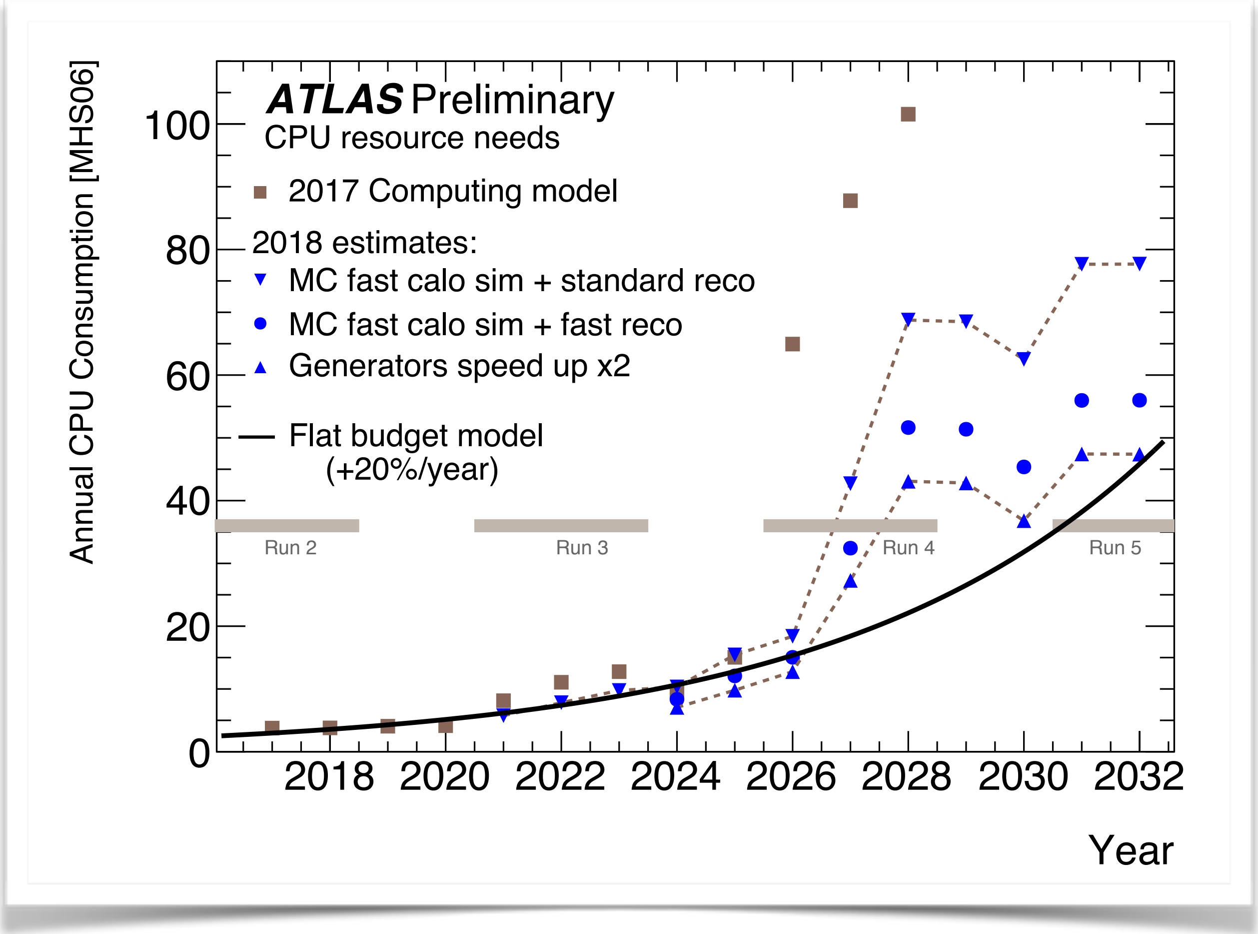


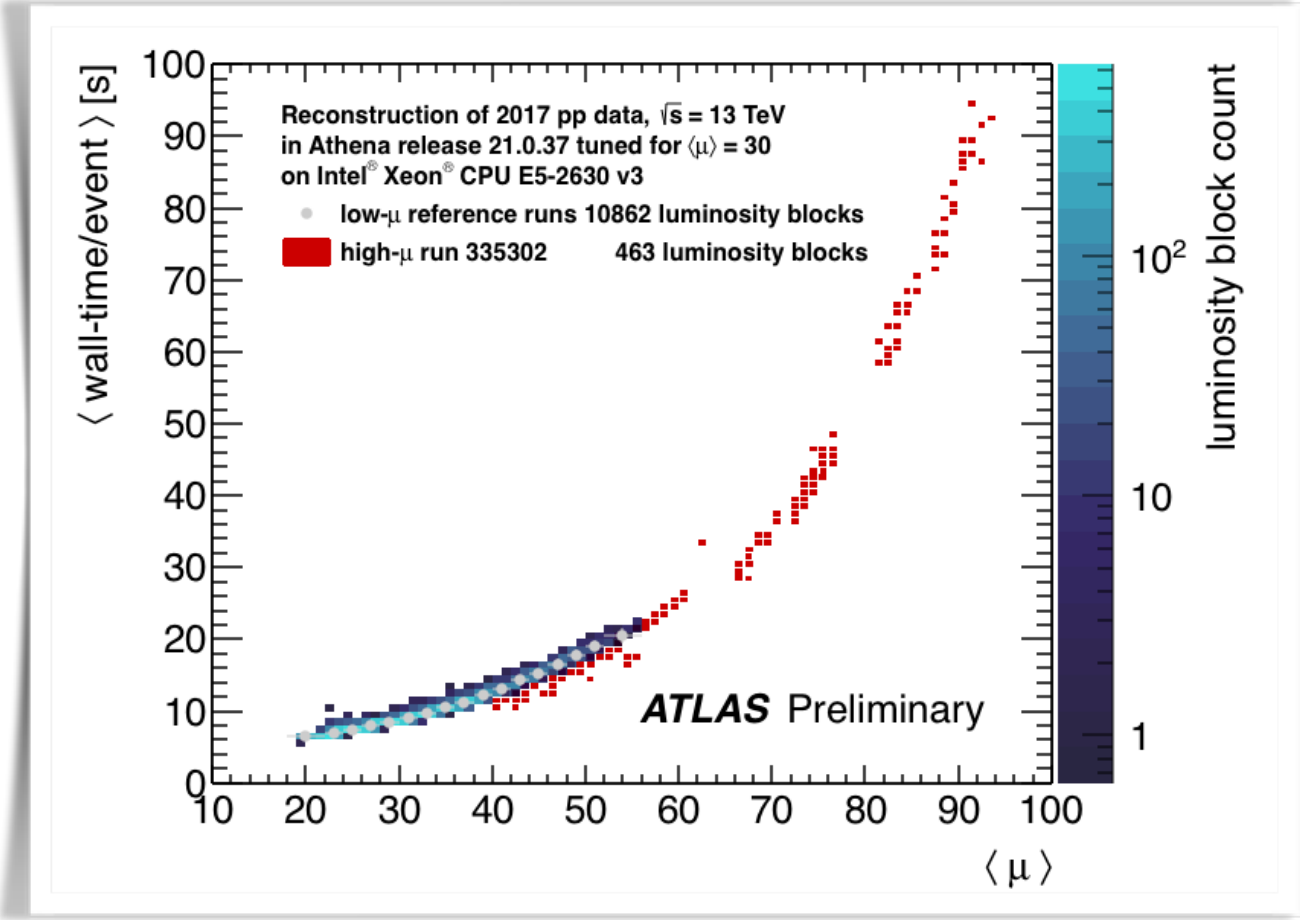
Only ~5% of total expected data

Computing for HL-LHC

- The HL-LHC environment is expected to pose a **challenge** for computing
  - Increased **luminosity**
  - Increased **read-out rates** (trigger+detector upgrades)
  - Increased **pile up**
- Currently project needing more CPU time than will be available
  - Dominated by **track reconstruction**
- Also expect to need 10x more disk storage

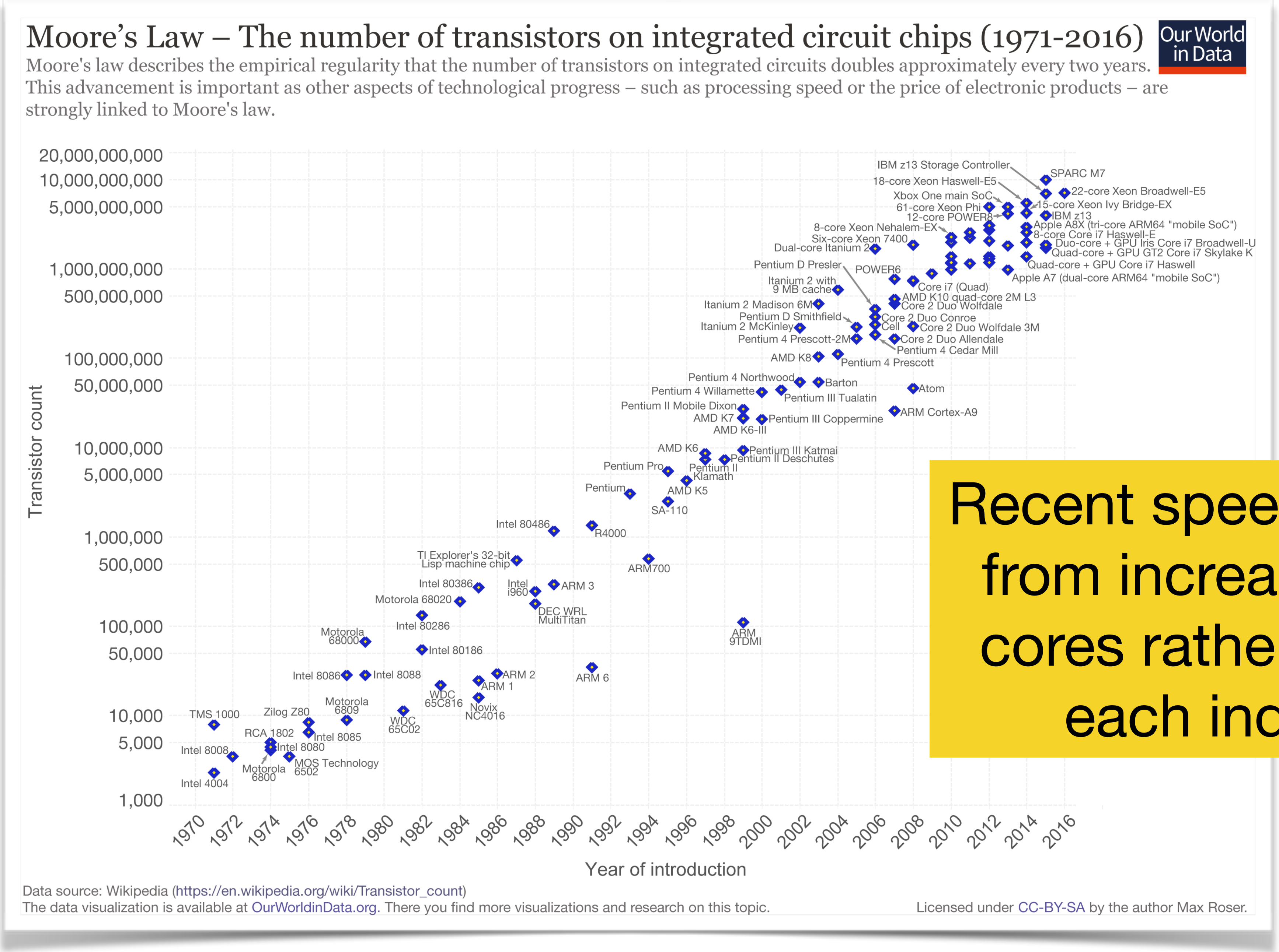
ref





Developments in Computing

Moore's law: number of transistors on integrated circuits doubles every two years



Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.

Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)
The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

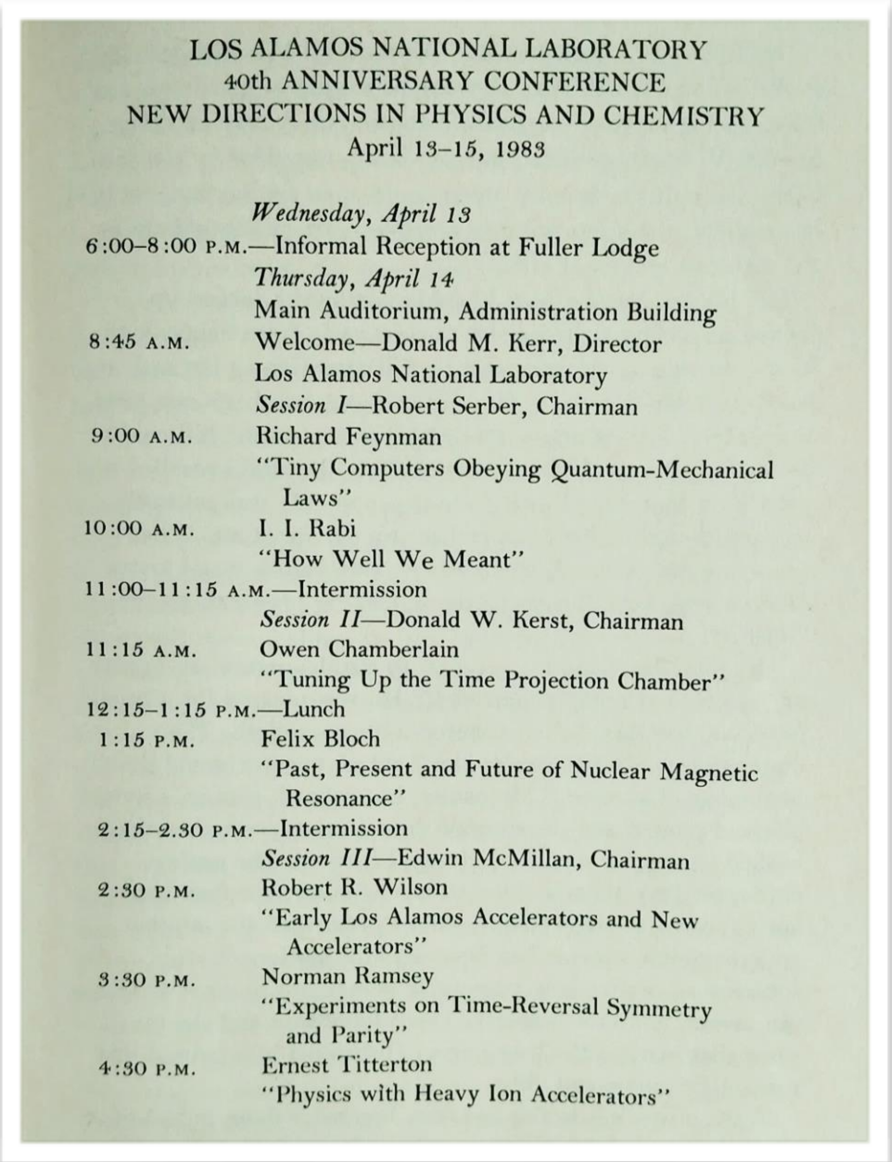
Licensed under CC-BY-SA by the author Max Roser.

Recent speed improvements from increasing number of cores rather than power of each individual core

Not expected to continue indefinitely: approaching the size of atoms

Initial ideas of quantum computing

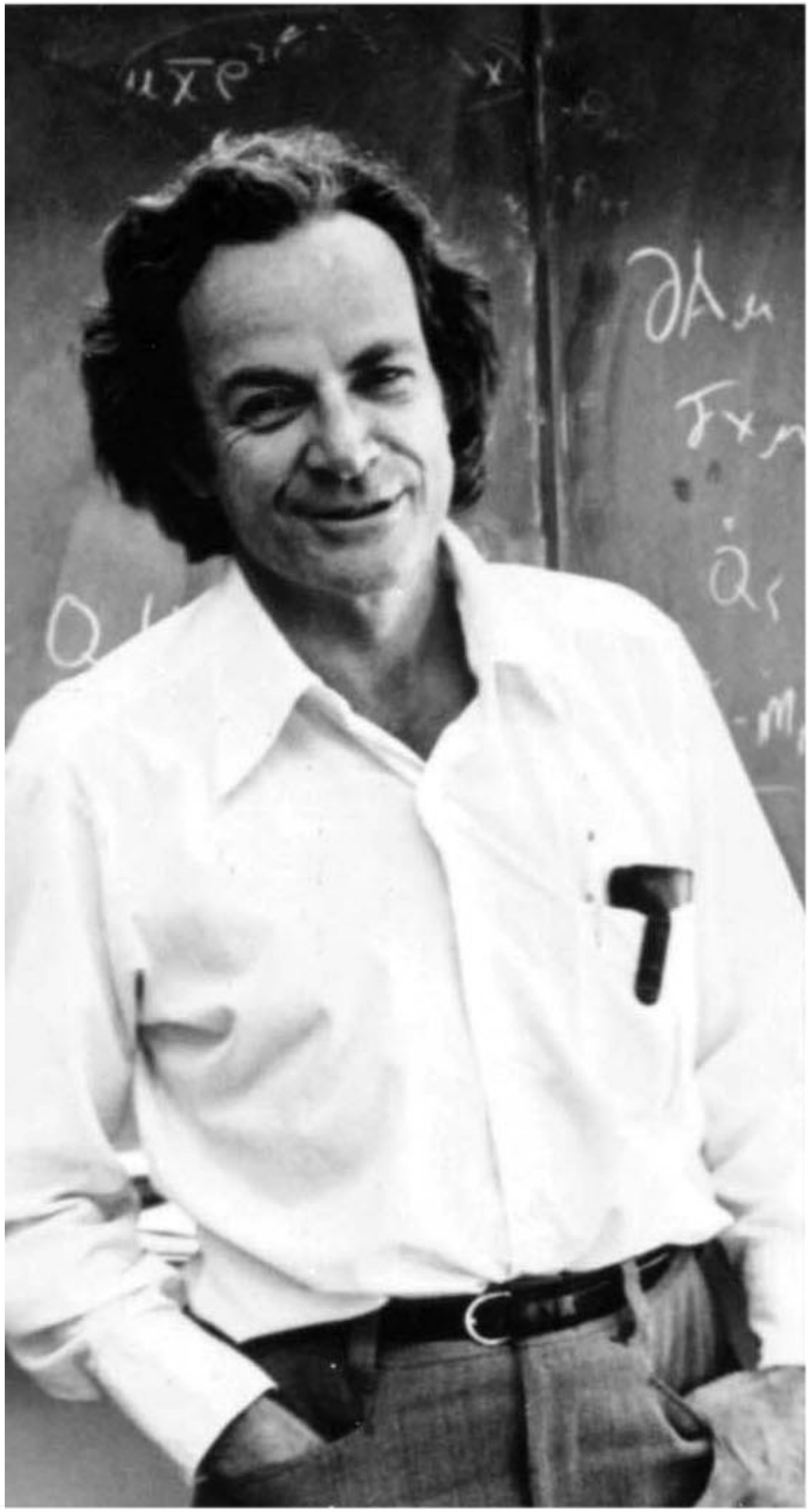
> *"Let the computer itself be built of quantum mechanical elements which obey quantum mechanical laws."*



LOS ALAMOS NATIONAL LABORATORY
40th ANNIVERSARY CONFERENCE
NEW DIRECTIONS IN PHYSICS AND CHEMISTRY
April 13–15, 1983

Wednesday, April 13
6:00–8:00 P.M.—Informal Reception at Fuller Lodge

Thursday, April 14
Main Auditorium, Administration Building
8:45 A.M. Welcome—Donald M. Kerr, Director
Los Alamos National Laboratory
Session I—Robert Serber, Chairman
9:00 A.M. Richard Feynman
"Tiny Computers Obeying Quantum-Mechanical Laws"
10:00 A.M. I. I. Rabi
"How Well We Meant"
11:00–11:15 A.M.—Intermission
Session II—Donald W. Kerst, Chairman
11:15 A.M. Owen Chamberlain
"Tuning Up the Time Projection Chamber"
12:15–1:15 P.M.—Lunch
1:15 P.M. Felix Bloch
"Past, Present and Future of Nuclear Magnetic Resonance"
2:15–2.30 P.M.—Intermission
Session III—Edwin McMillan, Chairman
2:30 P.M. Robert R. Wilson
"Early Los Alamos Accelerators and New Accelerators"
3:30 P.M. Norman Ramsey
"Experiments on Time-Reversal Symmetry and Parity"
4:30 P.M. Ernest Titterton
"Physics with Heavy Ion Accelerators"

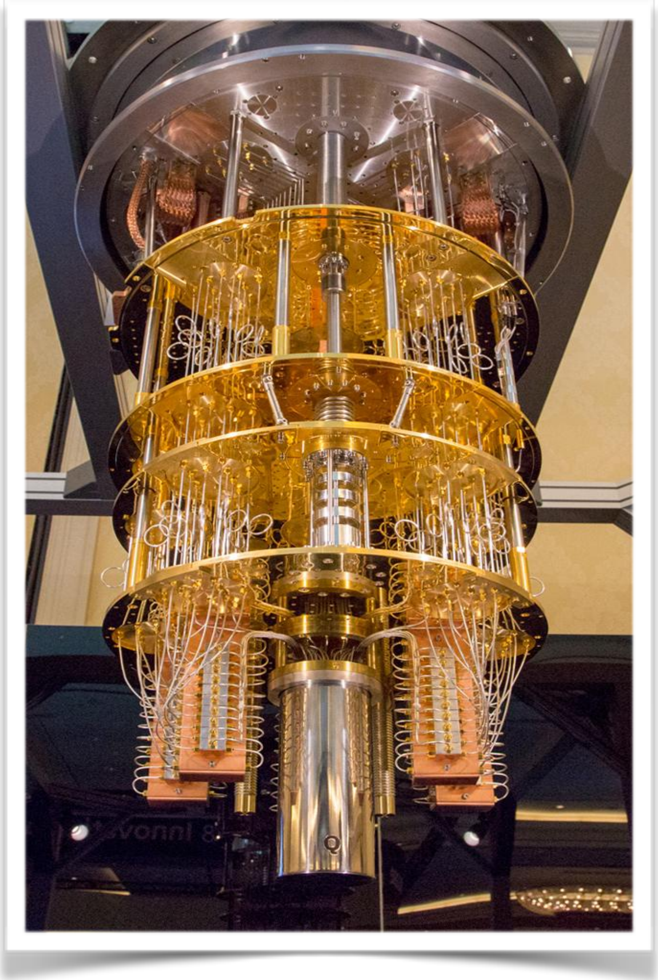


RICHARD FEYNMAN (1982)

Noisy Intermediate Scale Quantum (NISQ)

- Current state of the art quantum computers fall into two main categories
  - Quantum annealers, e.g. D-Wave (2000 qubits)
  - Universal quantum computers, e.g. IBM Q (20 quits)
- All quantum computers are not equal: challenges include connectivity and noise

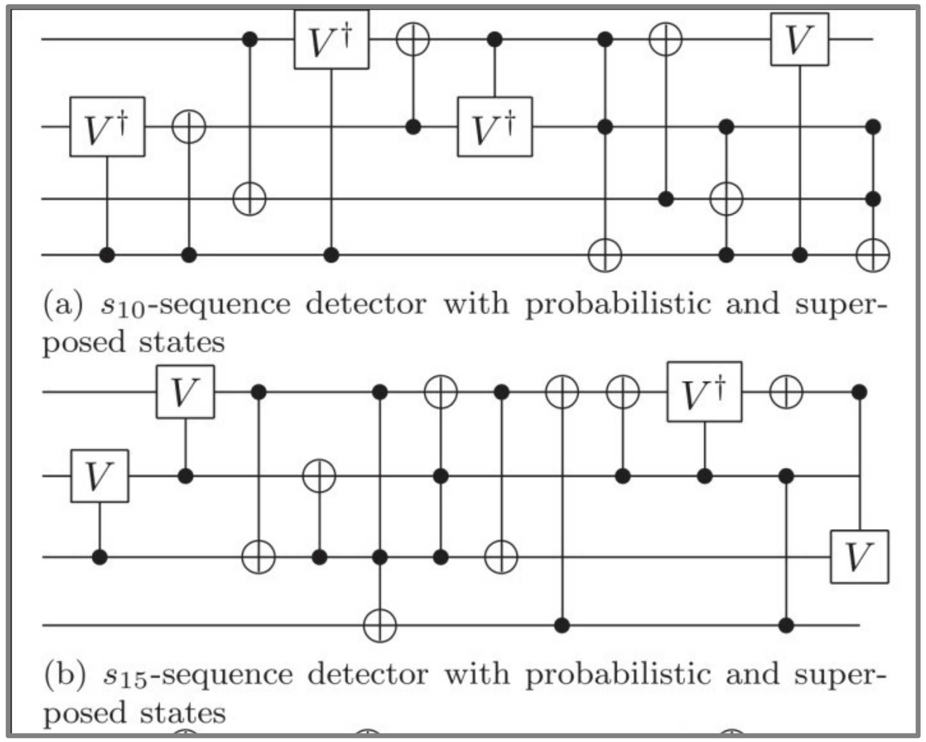
IBM 20Q Tokyo chip



D Wave

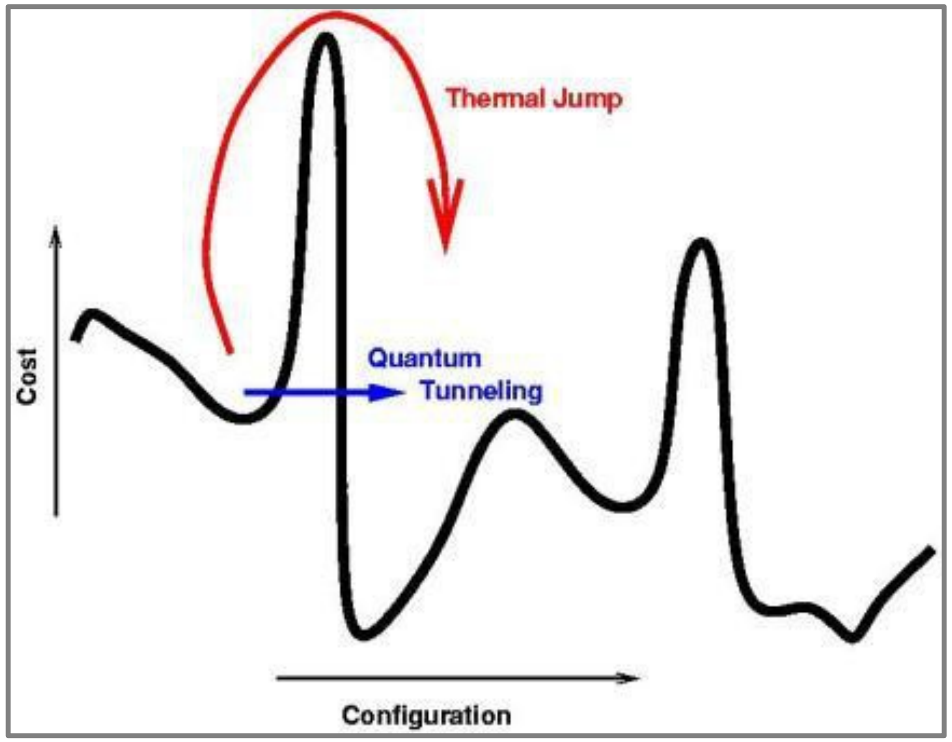
arXiv:1801.00862.pdf

Qubit and qunit



(a) s_{10}-sequence detector with probabilistic and superposed states

(b) s_{15}-sequence detector with probabilistic and superposed states

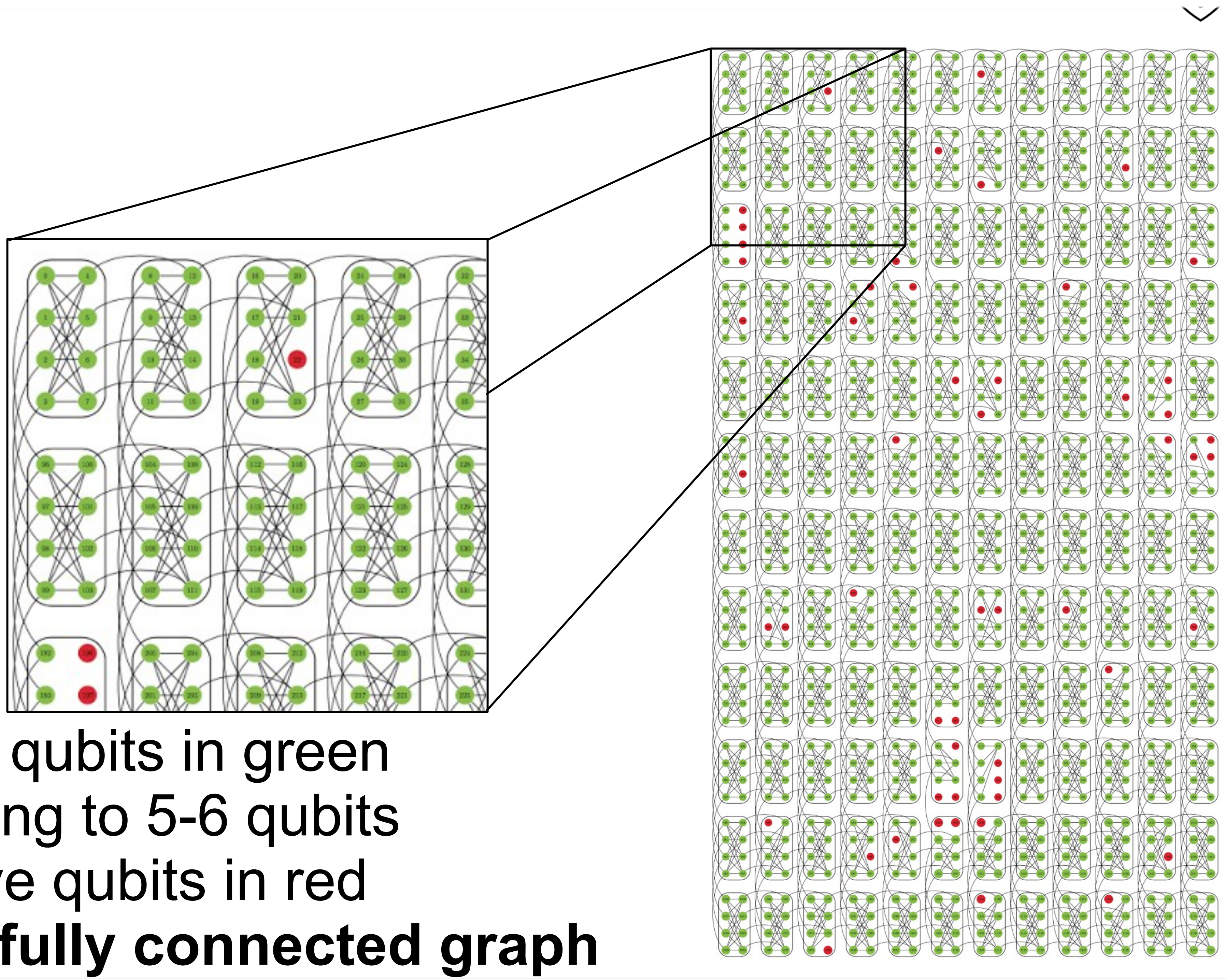


<u>Quantum Circuits</u>
Series of quantum gates operating on a set of quantum states.

<u>Quantum Annealing</u>
Evolution of a quantum system to a low T Gibbs state
That's D-Wave !

Slide courtesy of J.R. Vlimant

D-Wave Connectivity

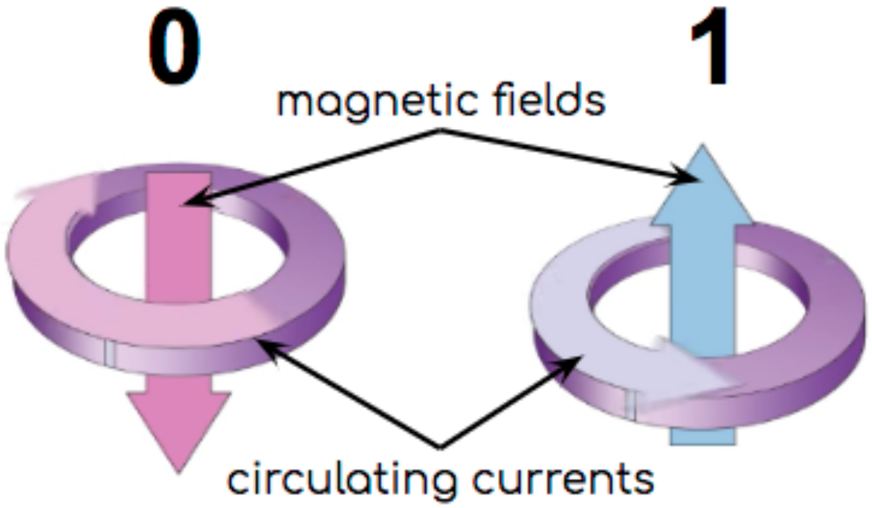


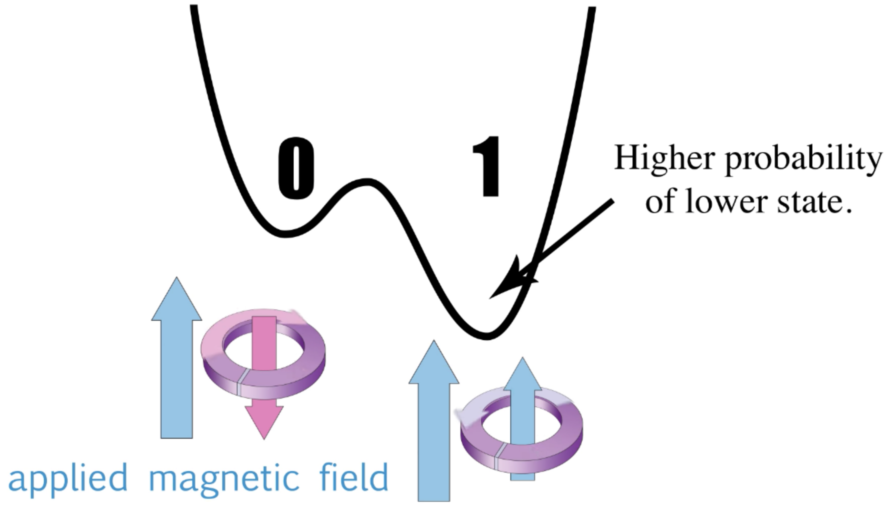
Active qubits in green
Coupling to 5-6 qubits
Inactive qubits in red
Not a fully connected graph

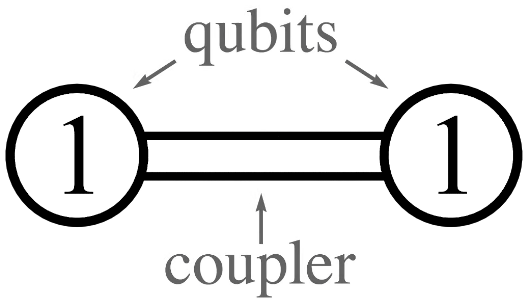
Slide courtesy of J.R. Vlimant

Quantum Annealing on D-Wave computers

source: dwavesys on YouTube







qubits ⇒ q_i

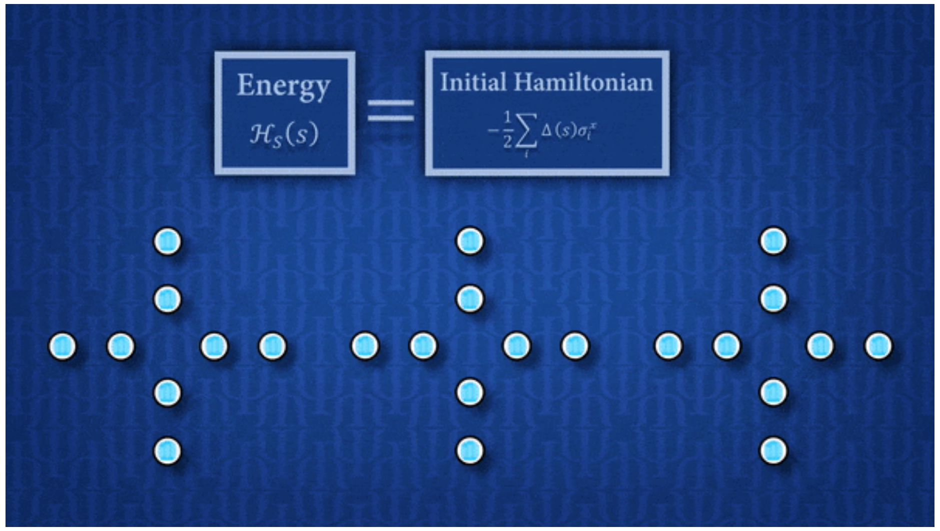
bias weights ⇒ a_i

coupling strength ⇒ b_{ij}

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

QUBO

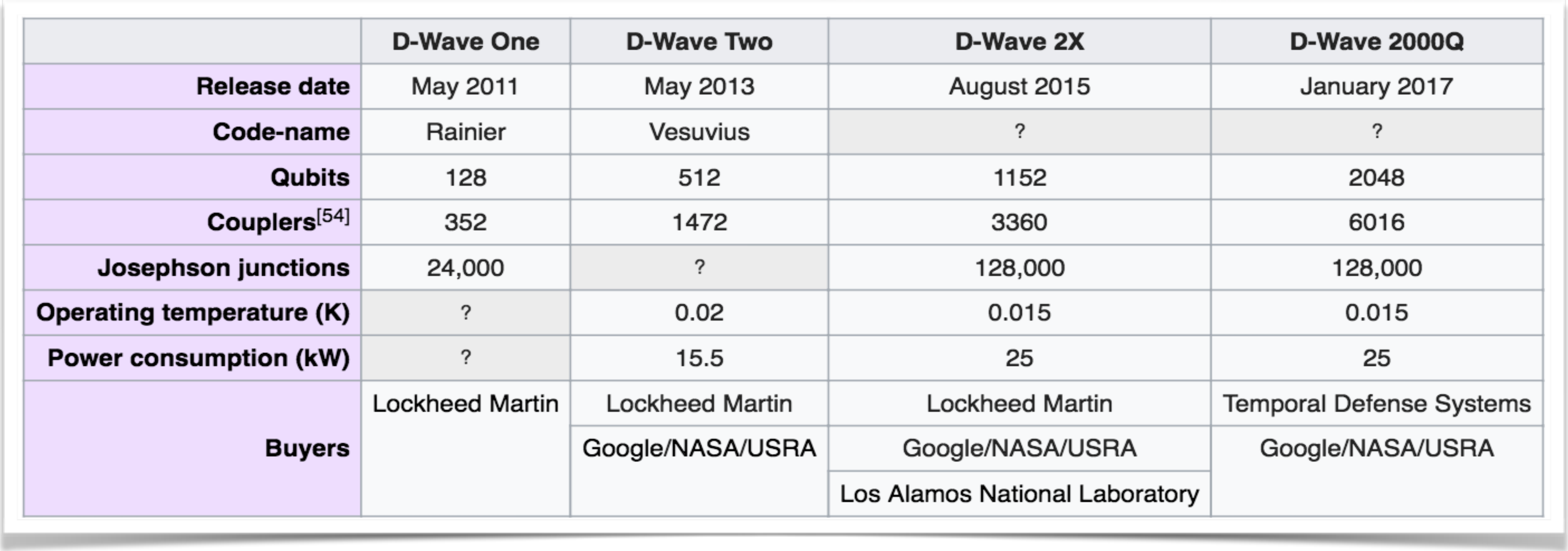
Quadratic
Unconstrained
Binary
Optimisation



anneal time
~20μs

Slide credit: L. Linder

Evolution of D-Wave Hardware



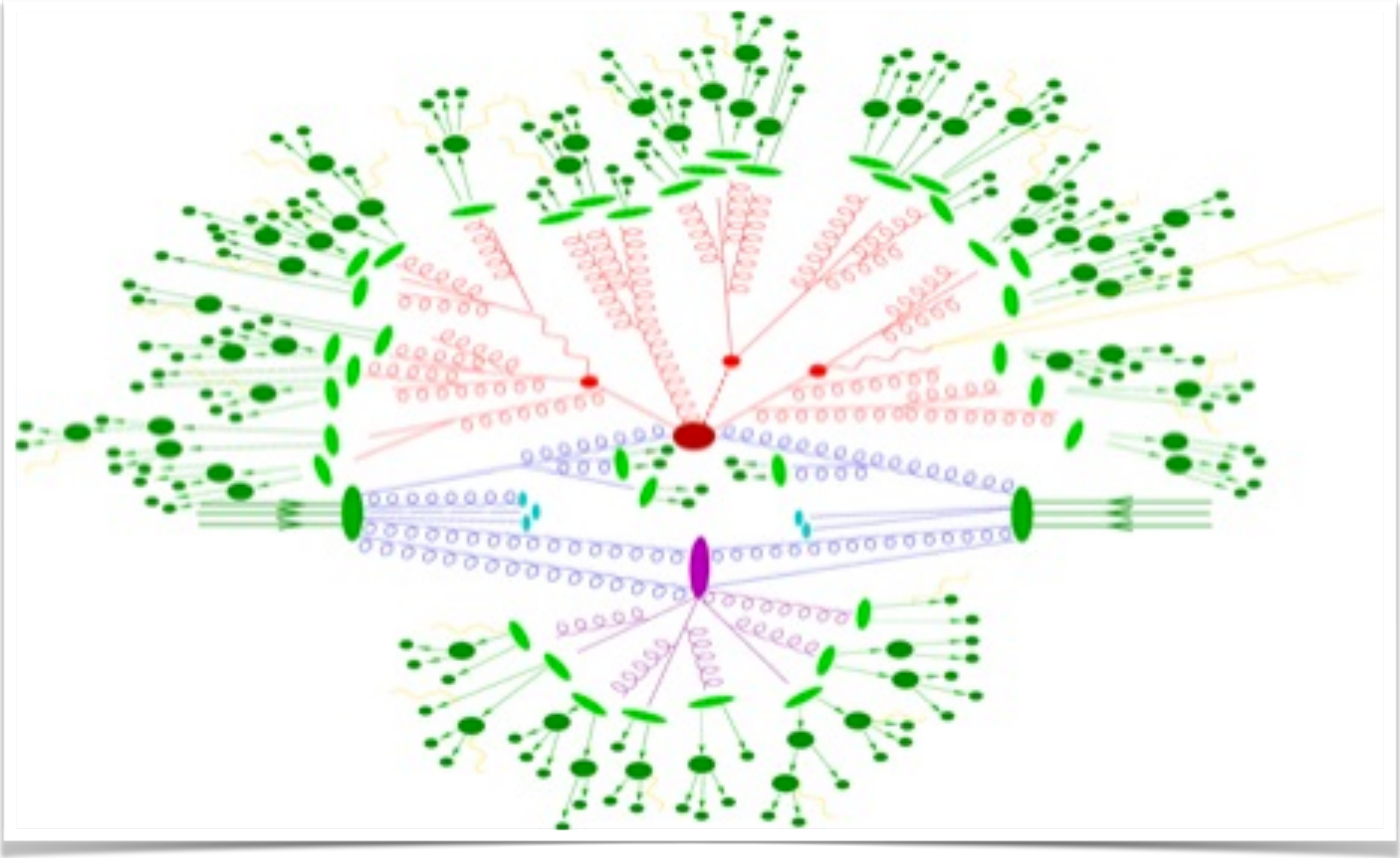
	D-Wave One	D-Wave Two	D-Wave 2X	D-Wave 2000Q
Release date	May 2011	May 2013	August 2015	January 2017
Code-name	Rainier	Vesuvius	?	?
Qubits	128	512	1152	2048
Couplers[54]	352	1472	3360	6016
Josephson junctions	24,000	?	128,000	128,000
Operating temperature (K)	?	0.02	0.015	0.015
Power consumption (kW)	?	15.5	25	25
Buyers	Lockheed Martin	Lockheed Martin Google/NASA/USRA	Lockheed Martin Google/NASA/USRA Los Alamos National Laboratory	Temporal Defense Systems Google/NASA/USRA

Slide credit: L. Linder

Ideas for Quantum Computing in HEP

Event Generation

A typical pp event at the LHC



In current MC generators, we largely neglect the correlations between particles in the parton shower

Entanglement

- This isn’t the full picture
  - Particles are quantum mechanical objects
  - Correlations exist between them
- Idea: exploit entanglement between qubits on a quantum computer to improve the description of the parton shower

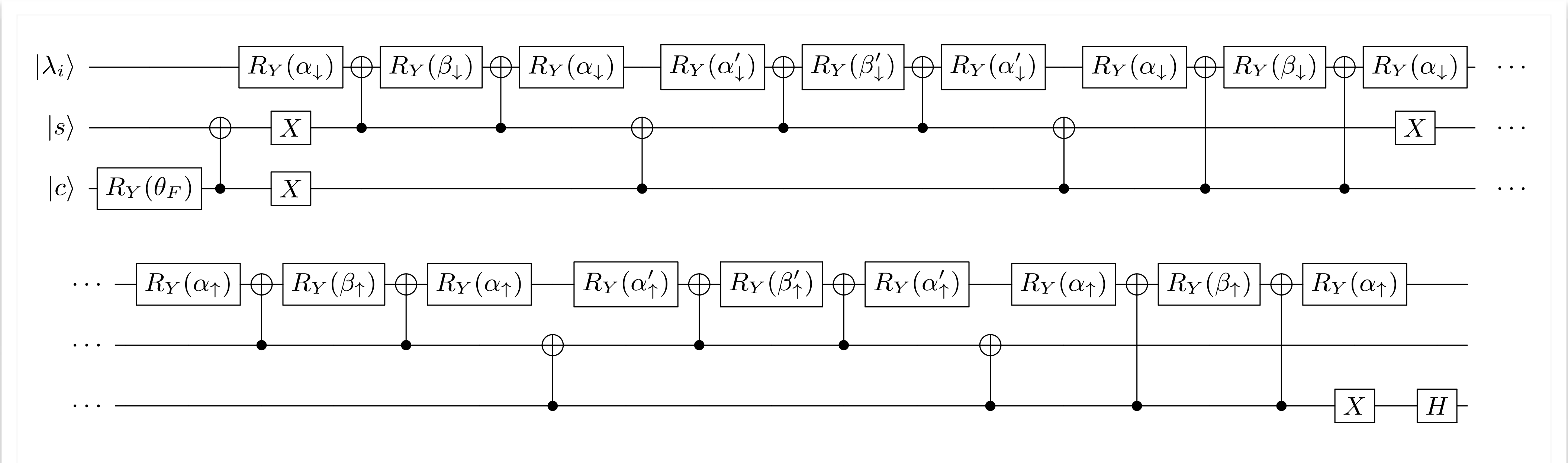


FIG. 8. The complete quantum circuit for a single step after the decomposition into only single qubit and CNOT gates. The angles labeled with ↓ and ↑ correspond to the decomposition of $U_{A,\downarrow}$ and $U_{A,\uparrow}$ respectively. Note that $\alpha'_{\downarrow,\uparrow} = -\alpha_{\downarrow,\uparrow}$ and $\beta'_{\downarrow,\uparrow} = -\beta_{\downarrow,\uparrow}$

arXiv:1901.08148

Early Results

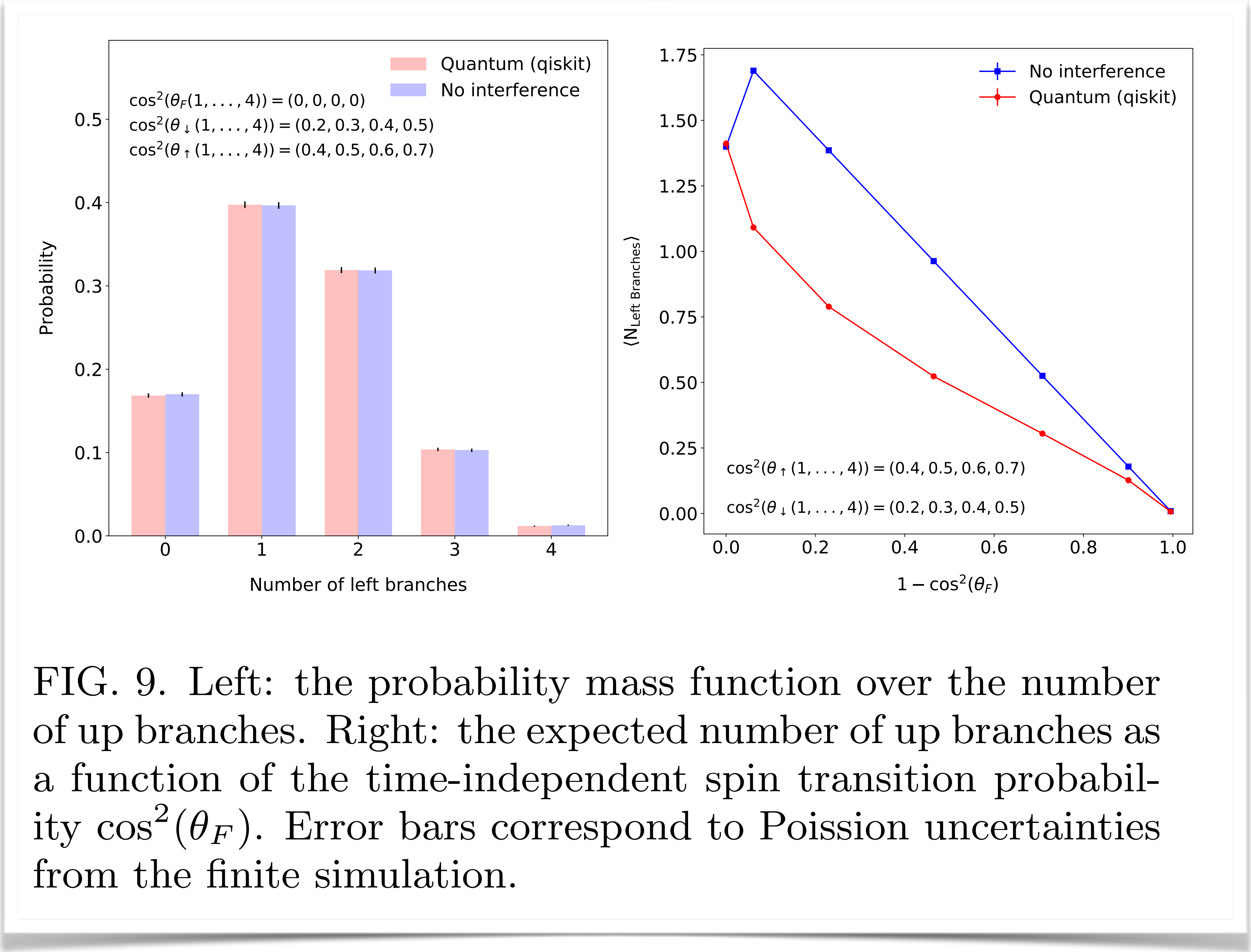


FIG. 9. Left: the probability mass function over the number of up branches. Right: the expected number of up branches as a function of the time-independent spin transition probability $\cos^2(\theta_F)$. Error bars correspond to Poission uncertainties from the finite simulation.

arXiv:1901.08148

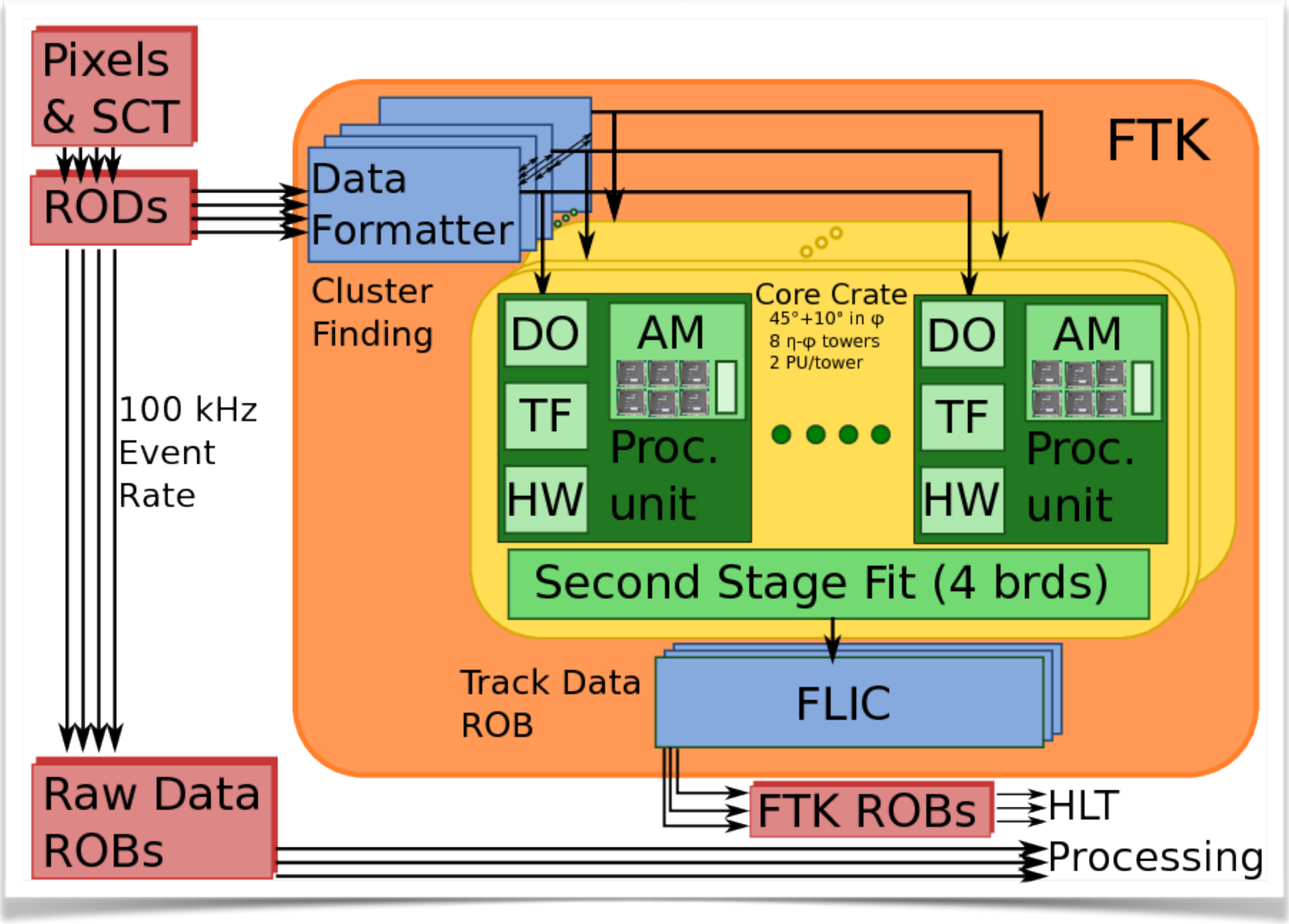
Track Reconstruction

https://sites.google.com/lbl.gov/hep-qpr

Track Reconstruction
1. Associative Memory

Tracking with Associative Memory

- Store possible tracks patterns directly in hardware
  - Direct mapping from hit patterns to tracks
  - Avoids scaling with combinatorics
  - Can be sensitive to changes in detector conditions
- Currently being installed within ATLAS as the Fast Track Trigger (FTK)



Quantum Associative Memory

- Theoretically proven asymptotic advantages of circuit-based QC
  - Optimal recall of unstructured memories
  - Optimal memory capacity

Strategy

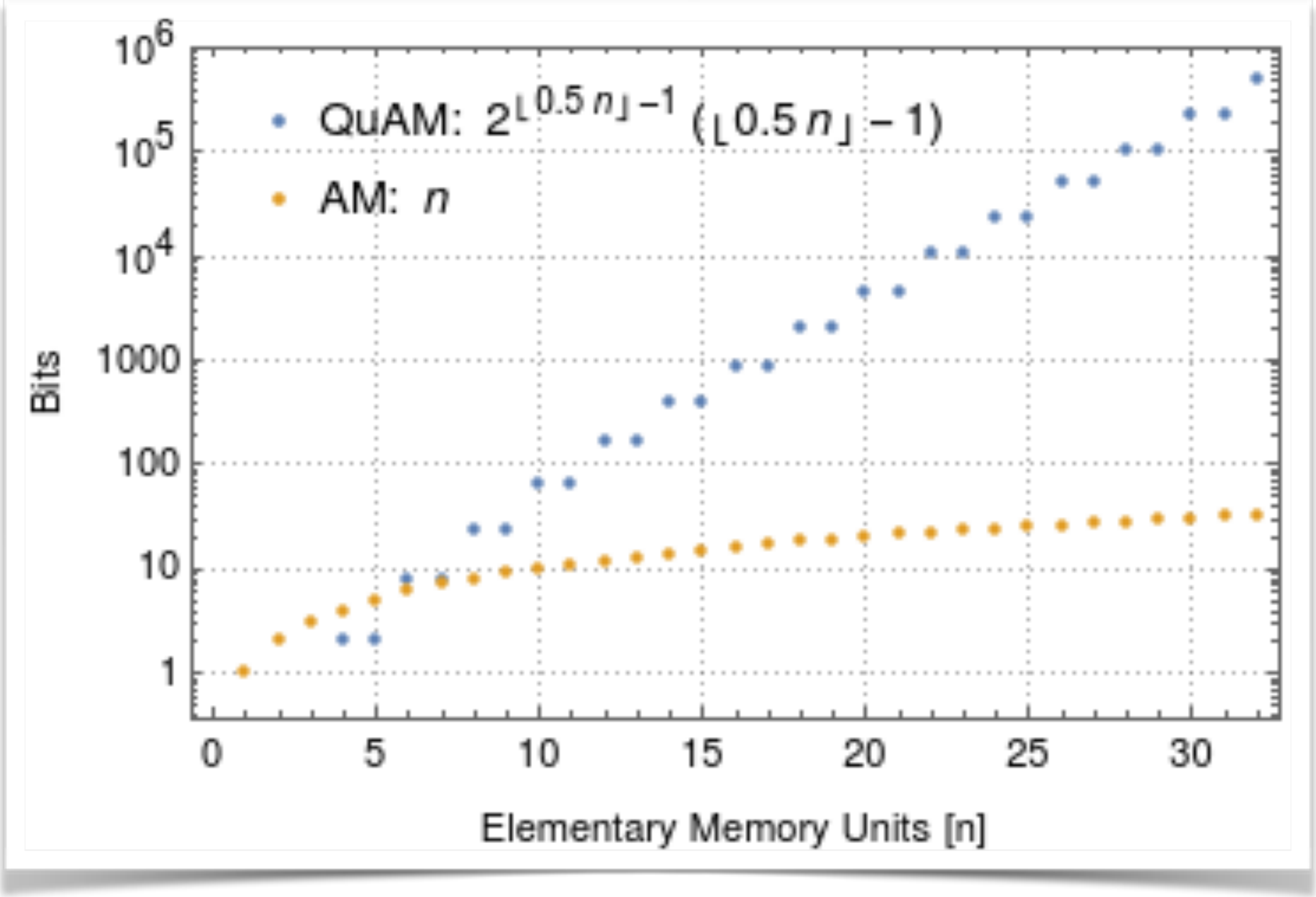
- Memorize N patterns by assembling a quantum superposition:

$$|\Xi\rangle = \sum^{N} \alpha_i |\xi^i\rangle, \qquad \alpha_i \in \mathbb{C} \;\wedge\; N \leqslant 2^n \;\wedge\; \sum^{N} |\alpha_i|^2 = 1$$

- Apply generalized Grover's algorithm to amplify the amplitude of a pattern being recalled.
- Measure memory.

Storage Capacity

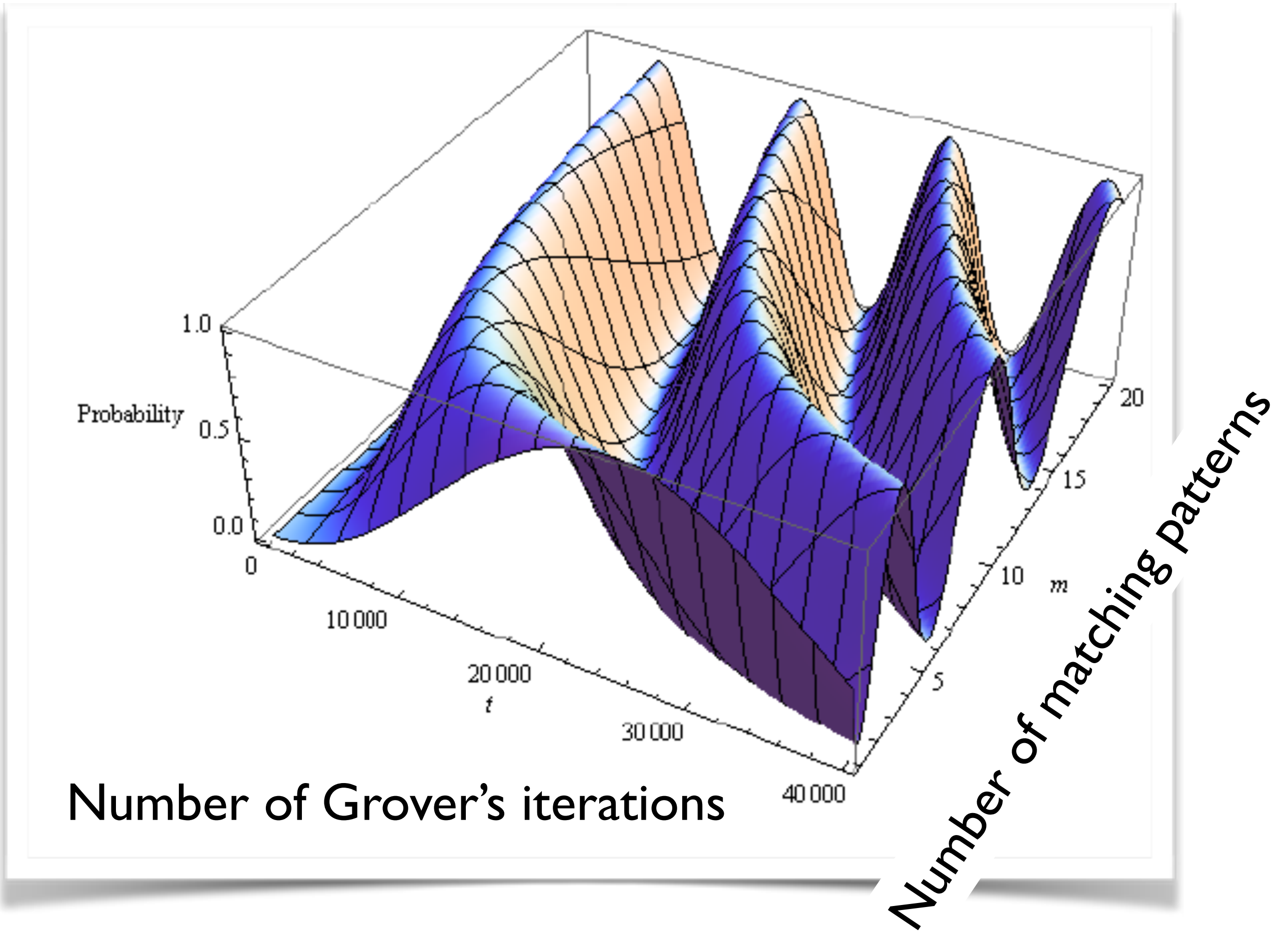
Detector hit identifier (bits)	8	16	32
Binary track pattern (bits)	64	128	256
QuAM register (qubits)	130	258	514
QuAM capacity (patterns)	**~10^{19}**	**~10^{38}**	**~10^{77}**

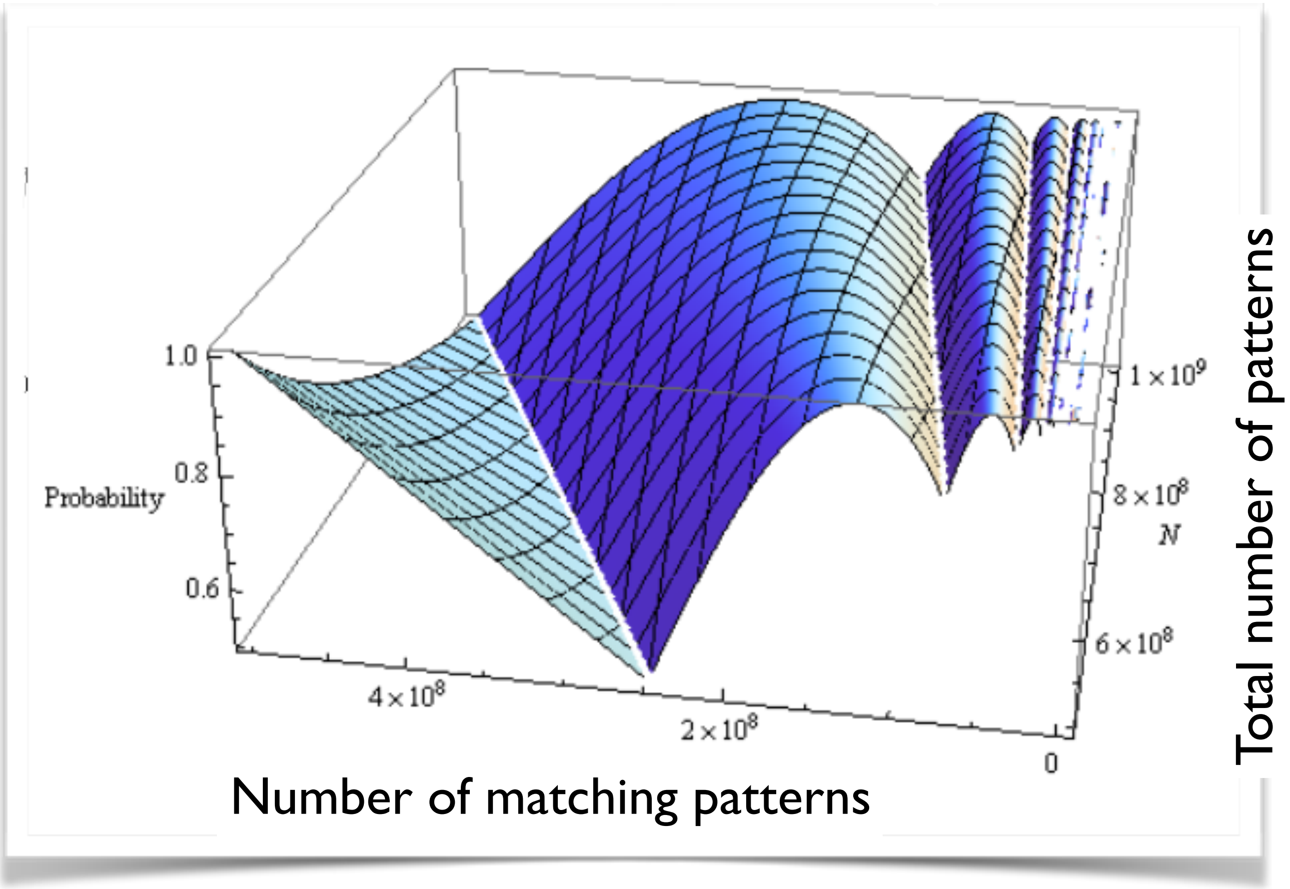


cf: 10^{78}-10^{82} atoms in the known universe

Recall Efficiency

- Theoretical probability of measuring a solution as a function of the number of Grover’s iterations and matching patterns (for N = 10^9)
- Peak probability for measuring a solution as a function of the number of matching and total number of patterns stored in QuAM.
- Both estimates assume the special case of uniform initial superposition and errorless quantum dynamics.

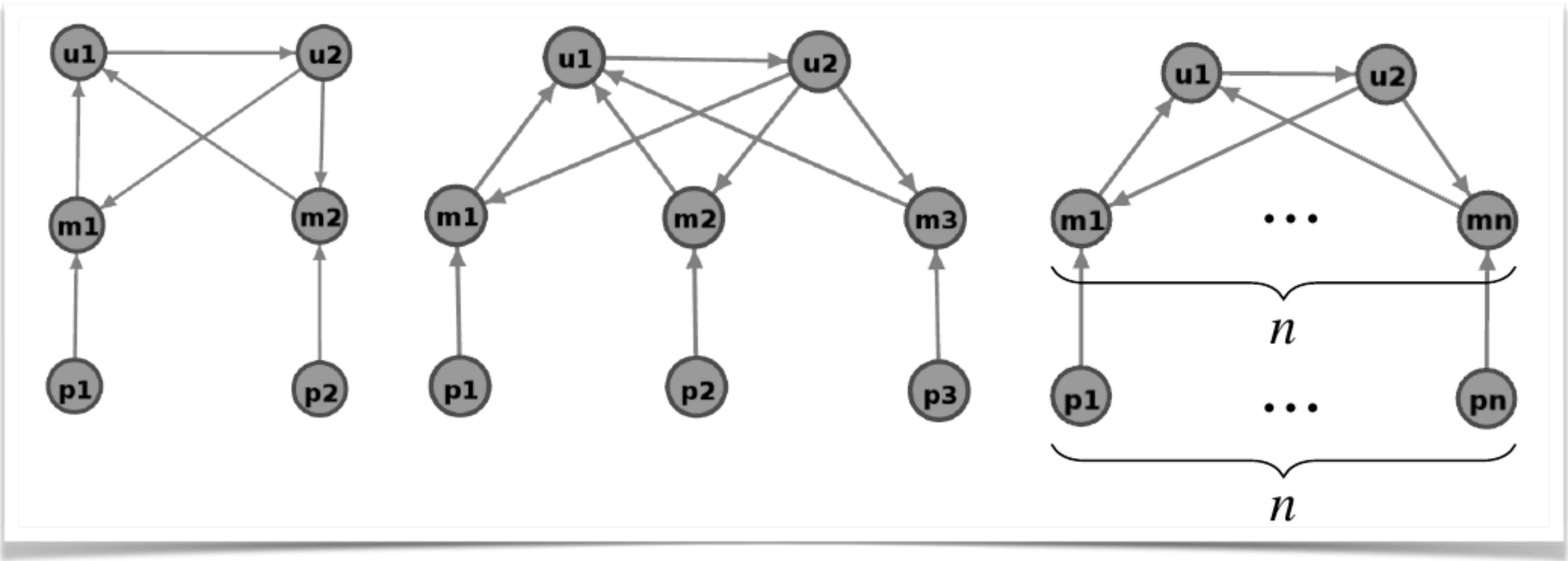




Topological Constraints

- Integral (storage and recall) topological requirements for patterns containing different numbers of bits
  - u: control register
  - p: temporary storage register
  - m: “permanent” storage register

2-bit 3-bit n-bit

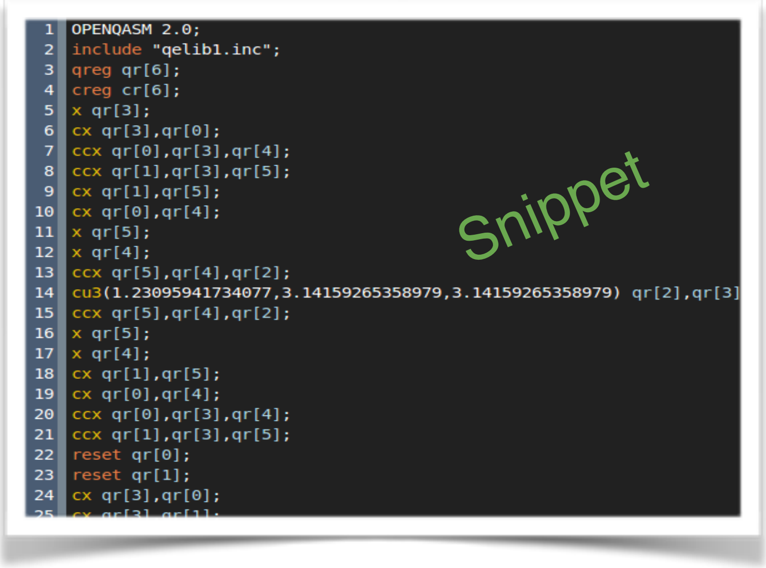


Implementation (I)

arXiv:1902.00498

- We developed QuAM circuit generators implementing the Trugenberger's initialization and generalized Grover's algorithms.
  - use open-source quantum computing platform, Qiskit
- Supported backends
  - IBM QE cloud-based quantum chips [5Q Yorktown/Tenerife, 14Q Melbourne, 20Q Tokyo]
  - Local/remote noisy simulators

Storage QuAM



```
OPENQASM 2.0;
include "qelib1.inc";
qreg qr[6];
creg cr[6];
x qr[3];
cx qr[3],qr[0];
ccx qr[0],qr[3],qr[4];
ccx qr[1],qr[3],qr[5];
cx qr[1],qr[5];
cx qr[0],qr[4];
x qr[5];
x qr[4];
ccx qr[5],qr[4],qr[2];
cu3(1.23095941734077,3.14159265358979,3.14159265358979) qr[2],qr[3]
ccx qr[5],qr[4],qr[2];
x qr[5];
x qr[4];
cx qr[1],qr[5];
cx qr[0],qr[4];
ccx qr[0],qr[3],qr[4];
ccx qr[1],qr[3],qr[5];
reset qr[0];
reset qr[1];
cx qr[3],qr[0];
```

Retrieval QuAM

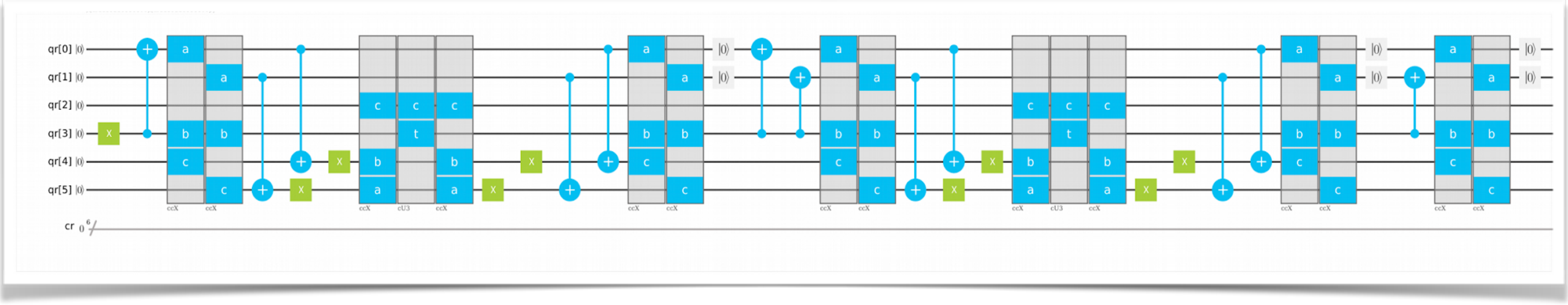
```
s qr[5];
h qr[5];
cx qr[4],qr[5];
h qr[5];
s qr[5];
h qr[4];
h qr[5];
x qr[4];
x qr[5];
h qr[5];
cx qr[4],qr[5];
h qr[5];
x qr[4];
x qr[5];
h qr[4];
h qr[5];
h qr[5];
cx qr[4],qr[5];
h qr[5];
```

Implementation (2)

arXiv:1902.00498

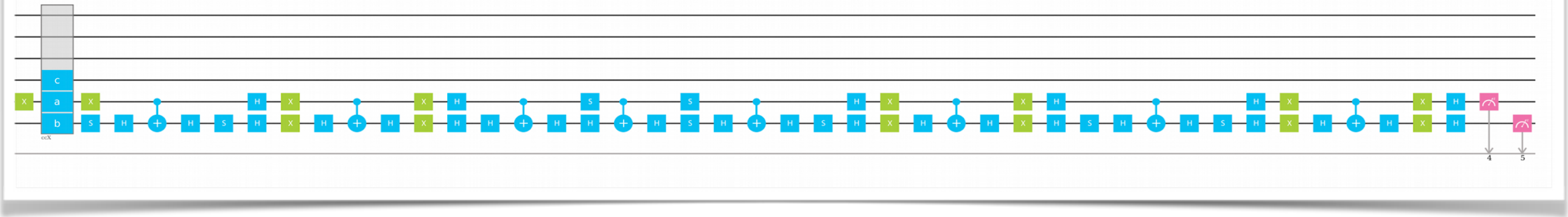
QuAM storage circuit generator

Ex.: complete circuit for retrieving one 2-bit pattern



QuAM retrieval circuit generator

Ex.: complete circuit for retrieving one 2-bit pattern

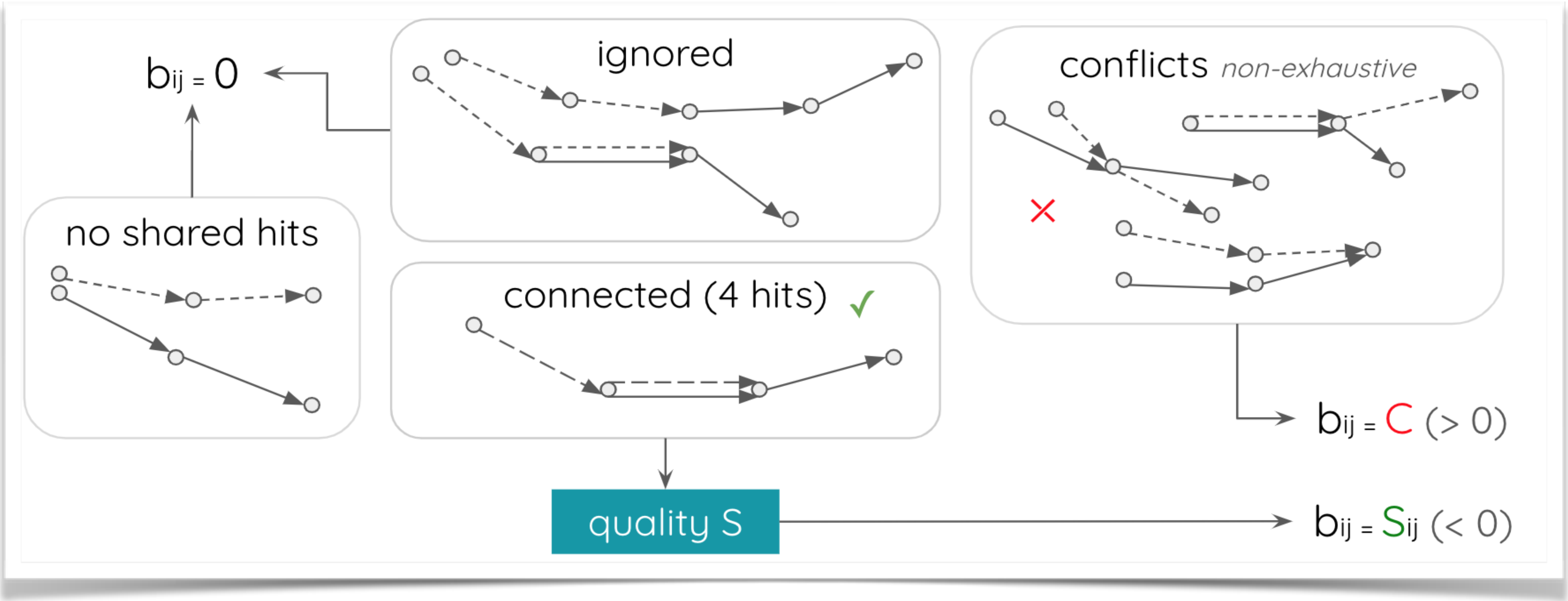


Track Reconstruction
2. Quantum Annealing

https://github.com/derlin/hepqpr-qallse

Quantum Annealing

- Explore how Quantum Annealing can bring speed improvements to pattern recognition
  - Implement QUBO minimisation on D-Wave and study scaling with track multiplicity
- Inspired from [1], but use triplets as binary variables
- Encode the quality of the triplets based on physics properties. The pair-wise connections b act as constraints (>0) or incentives (<0):

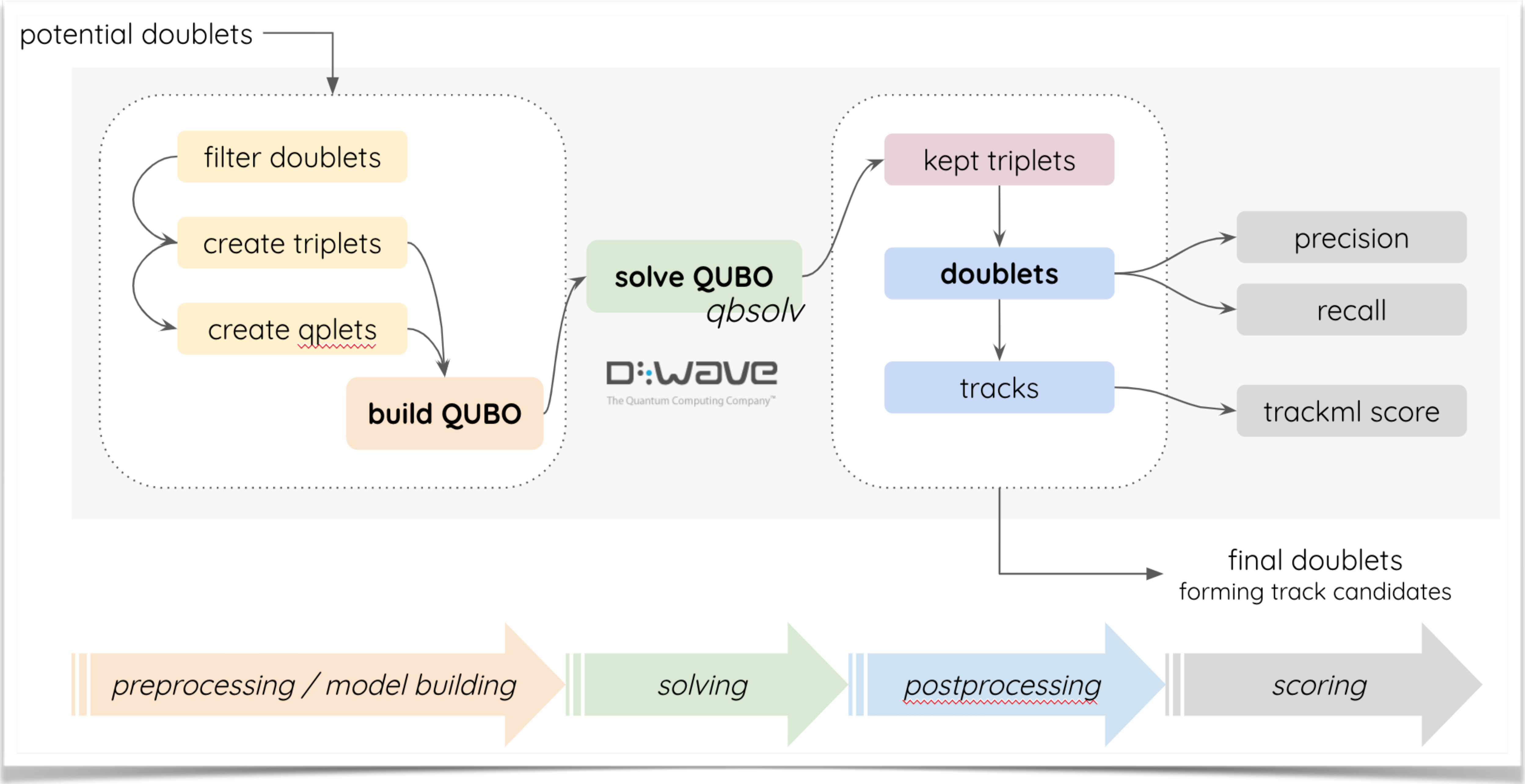


- Minimizing O means selecting the best triplets to form track candidates.

[1] Stimpfl-Abele, G., & Garrido, L. (1991). Fast track finding with neural networks. *Computer Physics Communications*, *64*(1), 46-56.

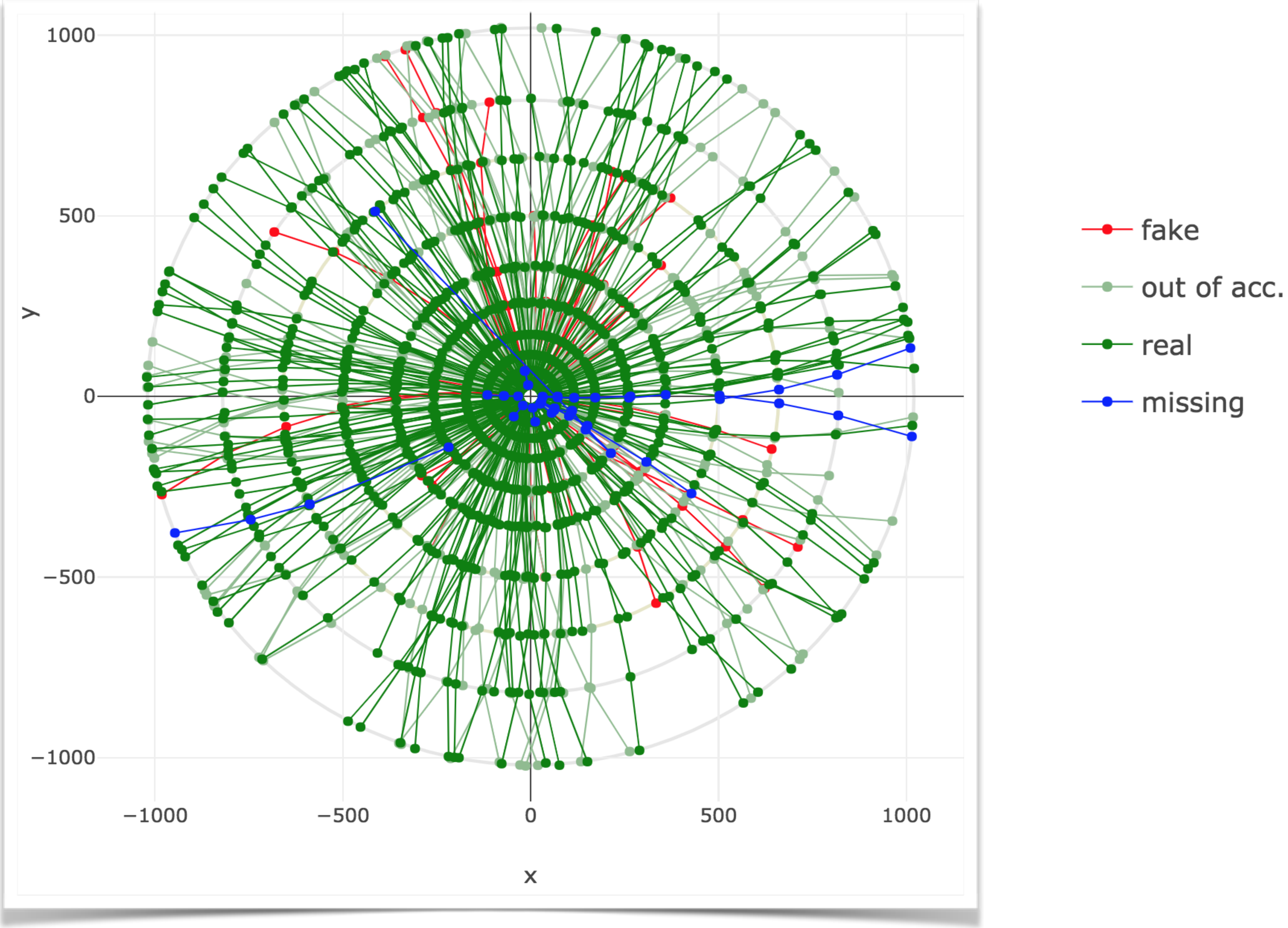
Implementation

- Dataset: simplified TrackML dataset, focus on barrel, 1+ GeV, 5+ hits
  - Toy dataset, but representative of expected conditions at the HL-LHC
- QUBO solvers: qbsolv (D-Wave + simulation), neal (simulation)
- D-Wave 2X (1152 qubits), D-Wave 2000Q (2048 qubits)



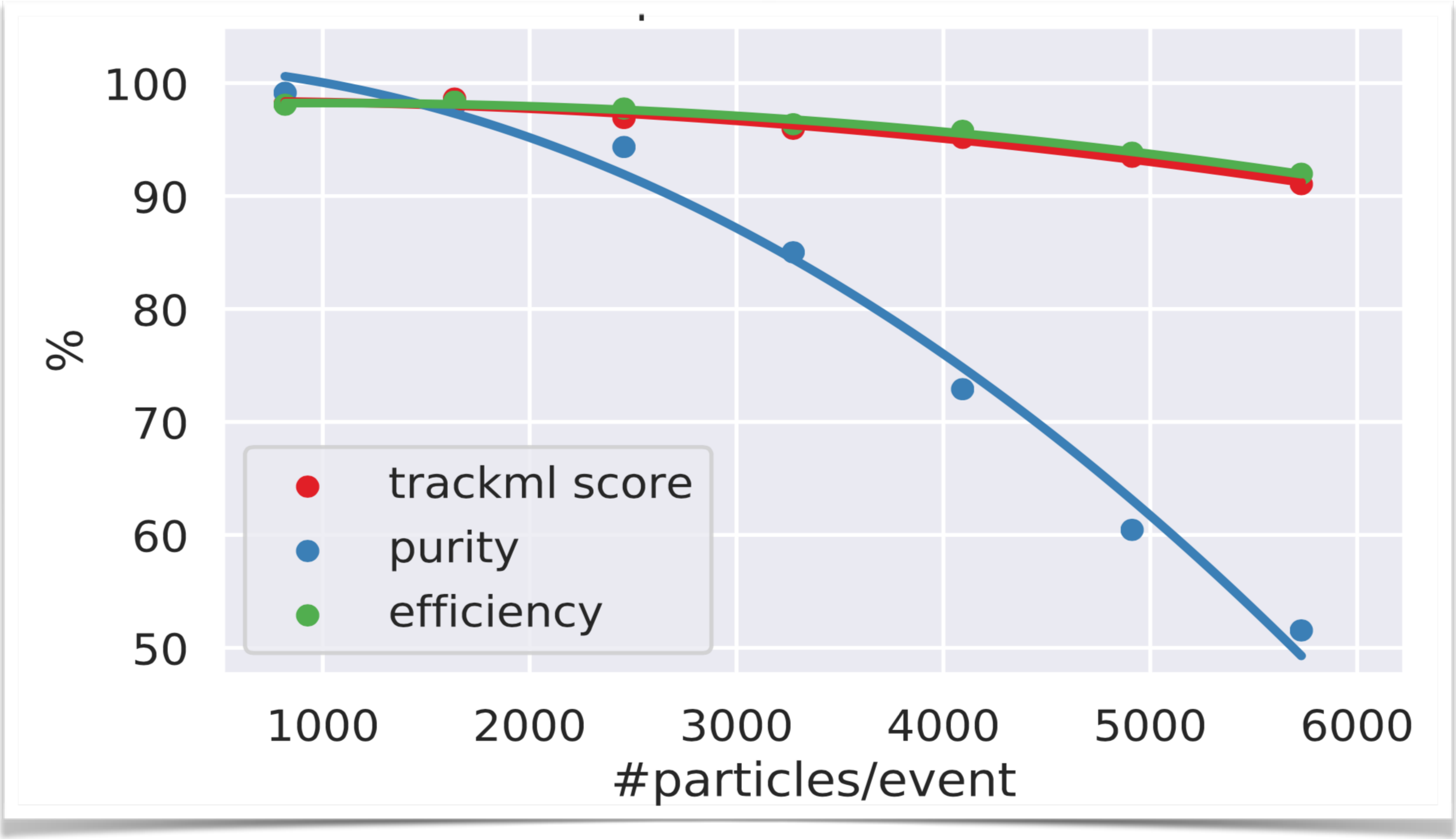
Performance

Doublets for a dataset of 2456 particles and 16855 hits



Performance (2)

Physics performance as a function of occupancy using a D-Wave 2X (qbsolv).

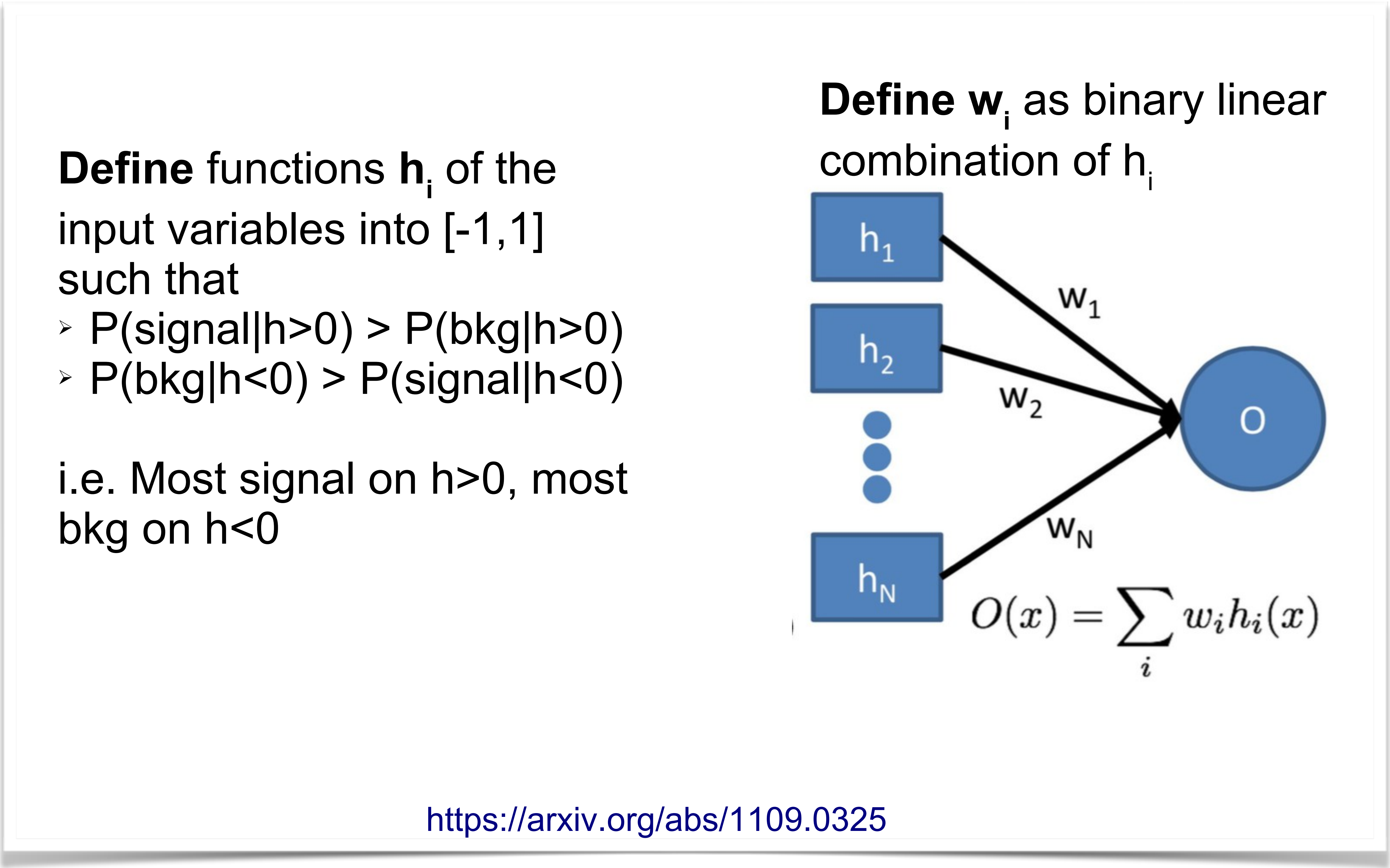


Timing building: 0-20 min | solving: 0-12s (sim), 0-56 min (D-Wave)

D-Wave | sim. Same physics, important time overhead with D-Wave

Analysis: Quantum Assisted Machine Learning (QAML)

QAML Classifiers



Define functions **h$_i$** of the input variables into [-1,1] such that

- P(signal|h>0) > P(bkg|h>0)
- P(bkg|h<0) > P(signal|h<0)

i.e. Most signal on h>0, most bkg on h<0

Define w$_i$ as binary linear combination of h$_i$

$$O(x) = \sum_i w_i h_i(x)$$

https://arxiv.org/abs/1109.0325

Slide courtesy of J. R. Vlimant

QAML Objective Function

Define as a “target” function

$$y(x) = \begin{cases} +1, & \text{if } \in S, \\ -1, & \text{if } \in B \end{cases}$$

Per event error

$$E(x) = y(x) - \sum_{i=1}^{N} w_i h_i(x)$$

Full error

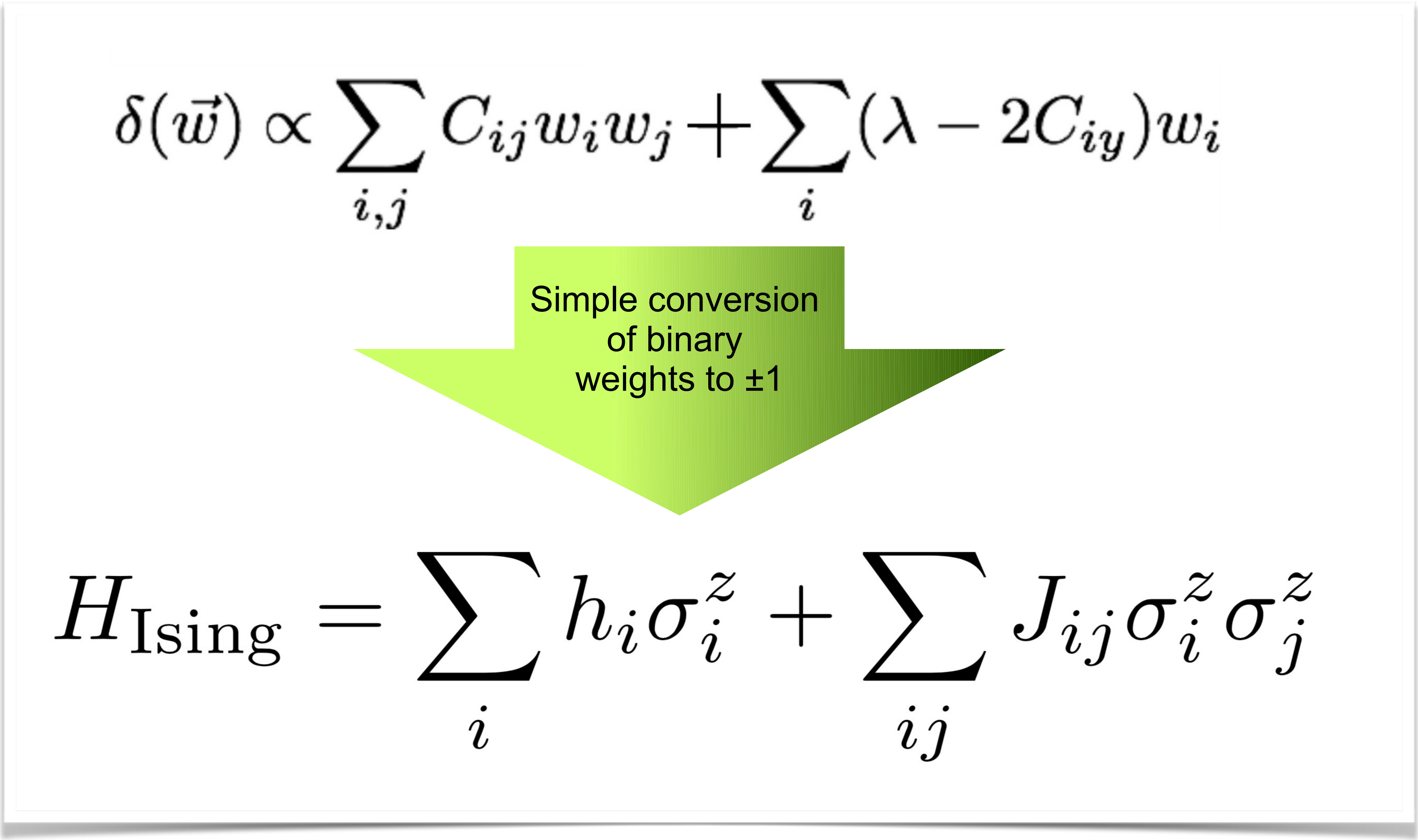
$$\delta(\vec{w}) \propto \sum_{i,j} C_{ij} w_i w_j + \sum_{i} (\lambda - 2C_{iy}) w_i$$

- ➔ C_{ij} and C_{iy} are summations over the values of h_i over the training set
- ➔ λ is a parameter penalizing the number of non-zero w_i

https://arxiv.org/abs/1109.0325

Slide courtesy of J. R. Vlimant

Implementation with QUBO



$$\delta(\vec{w}) \propto \sum_{i,j} C_{ij} w_i w_j + \sum_{i} (\lambda - 2C_{iy}) w_i$$

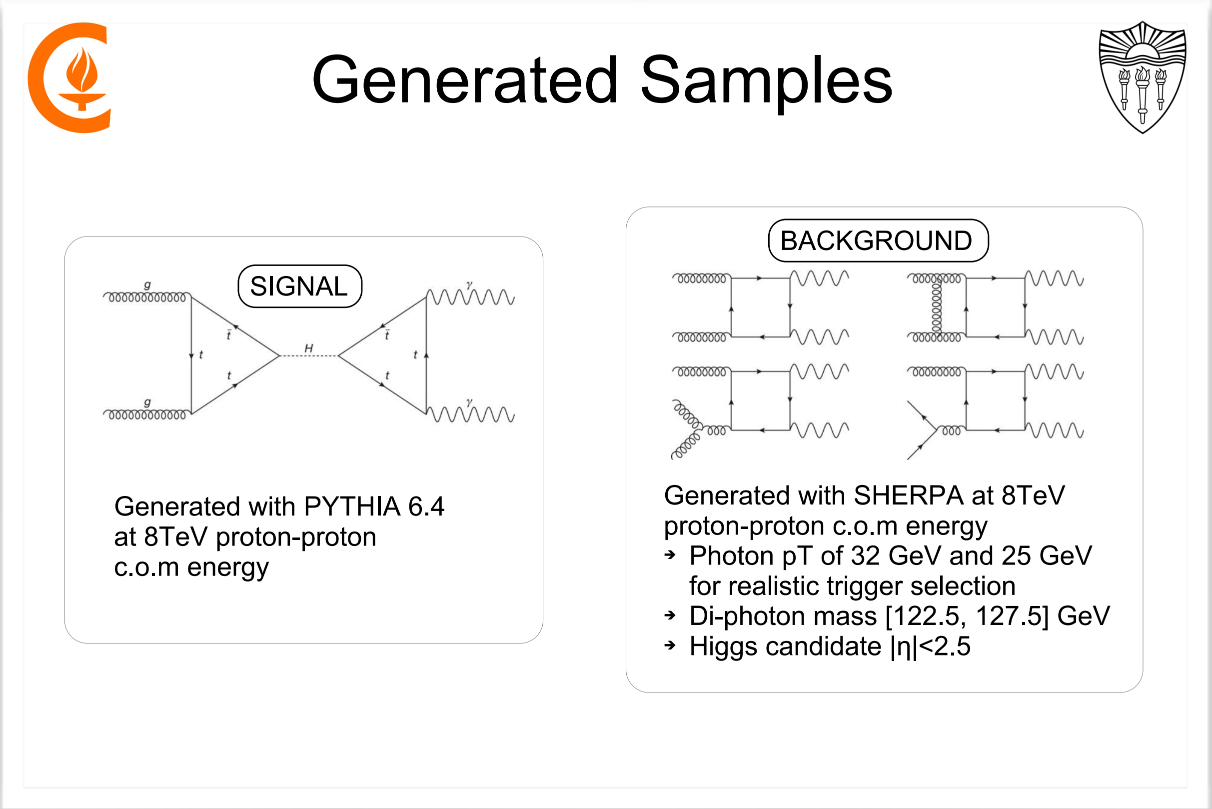
Simple conversion of binary weights to ±1

$$H_{\text{Ising}} = \sum_{i} h_i \sigma_i^z + \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z$$

Slide courtesy of J. R. Vlimant

Example: H→γγ

Slides courtesy of J. R. Vlimant



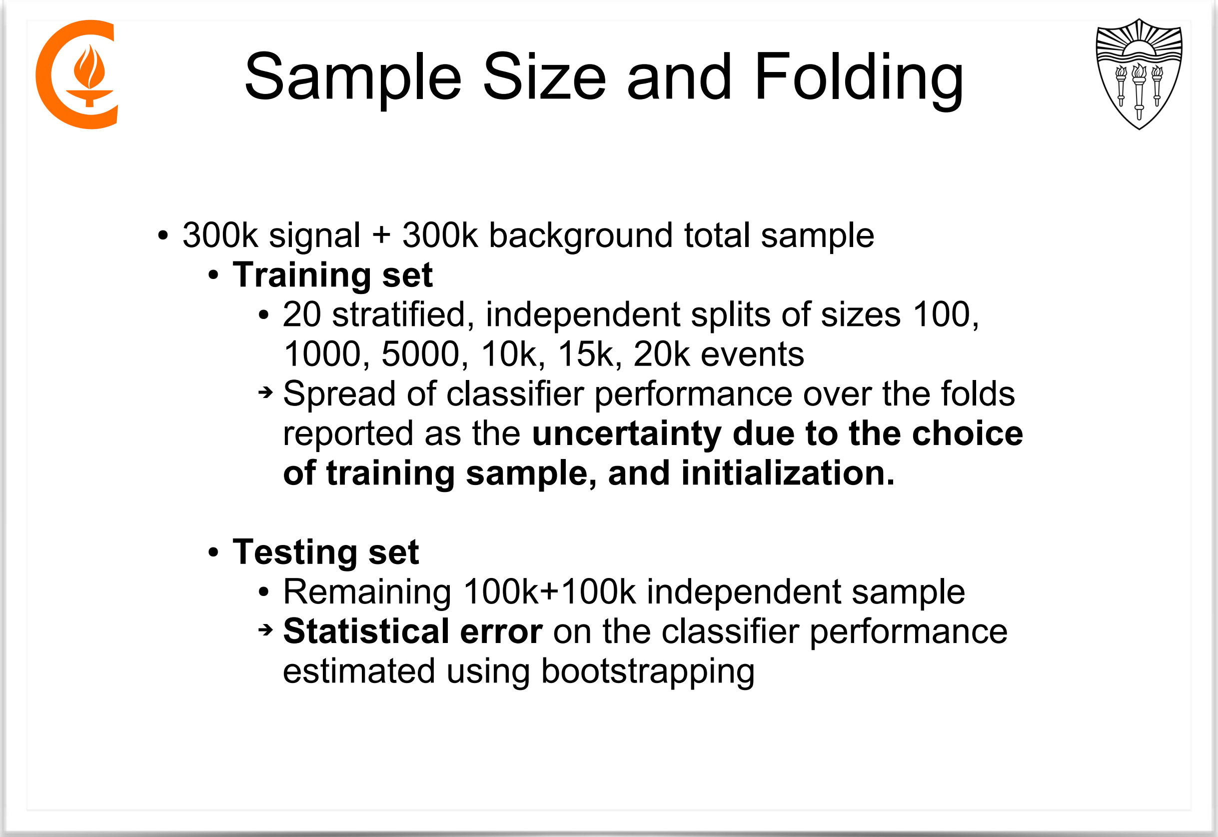
Generated Samples

SIGNAL

Generated with PYTHIA 6.4 at 8TeV proton-proton c.o.m energy

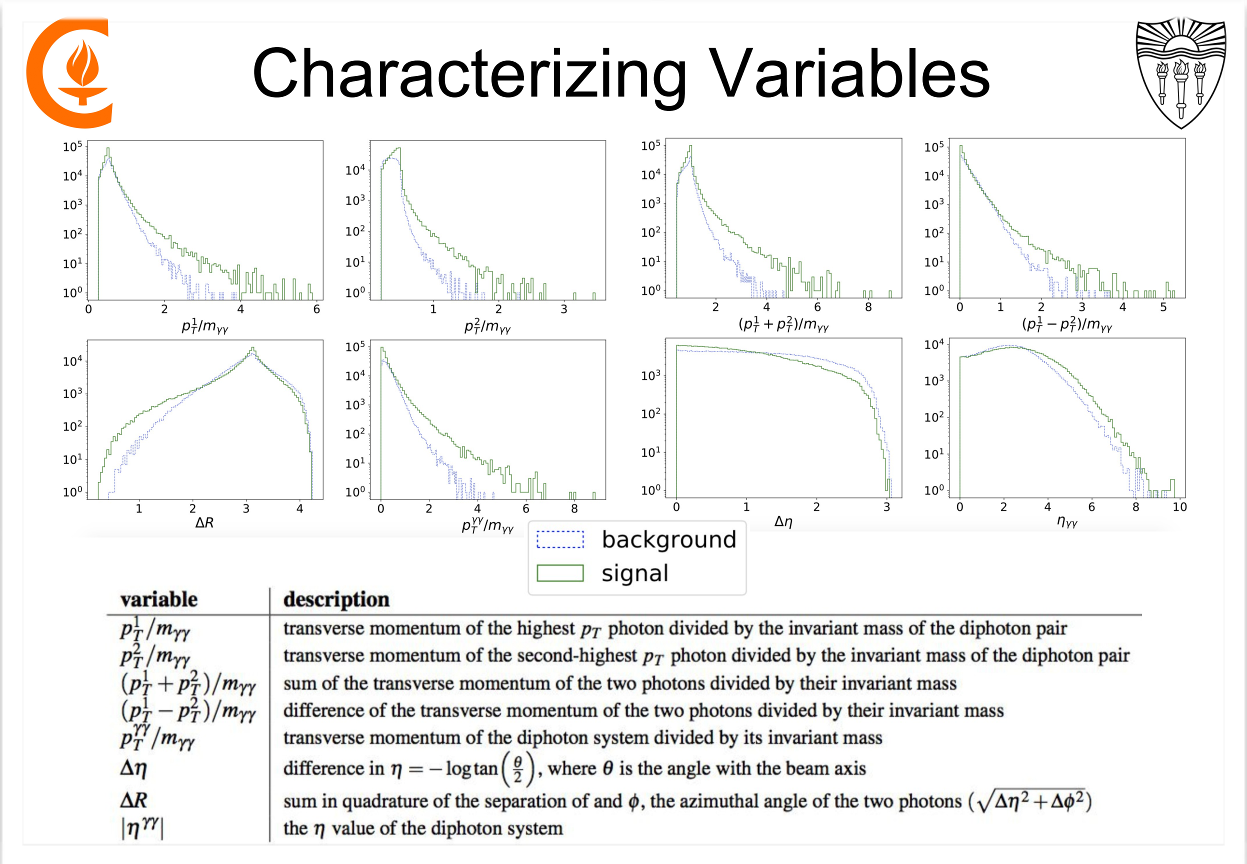
BACKGROUND

Generated with SHERPA at 8TeV proton-proton c.o.m energy
- → Photon pT of 32 GeV and 25 GeV for realistic trigger selection
- → Di-photon mass [122.5, 127.5] GeV
- → Higgs candidate |η|<2.5



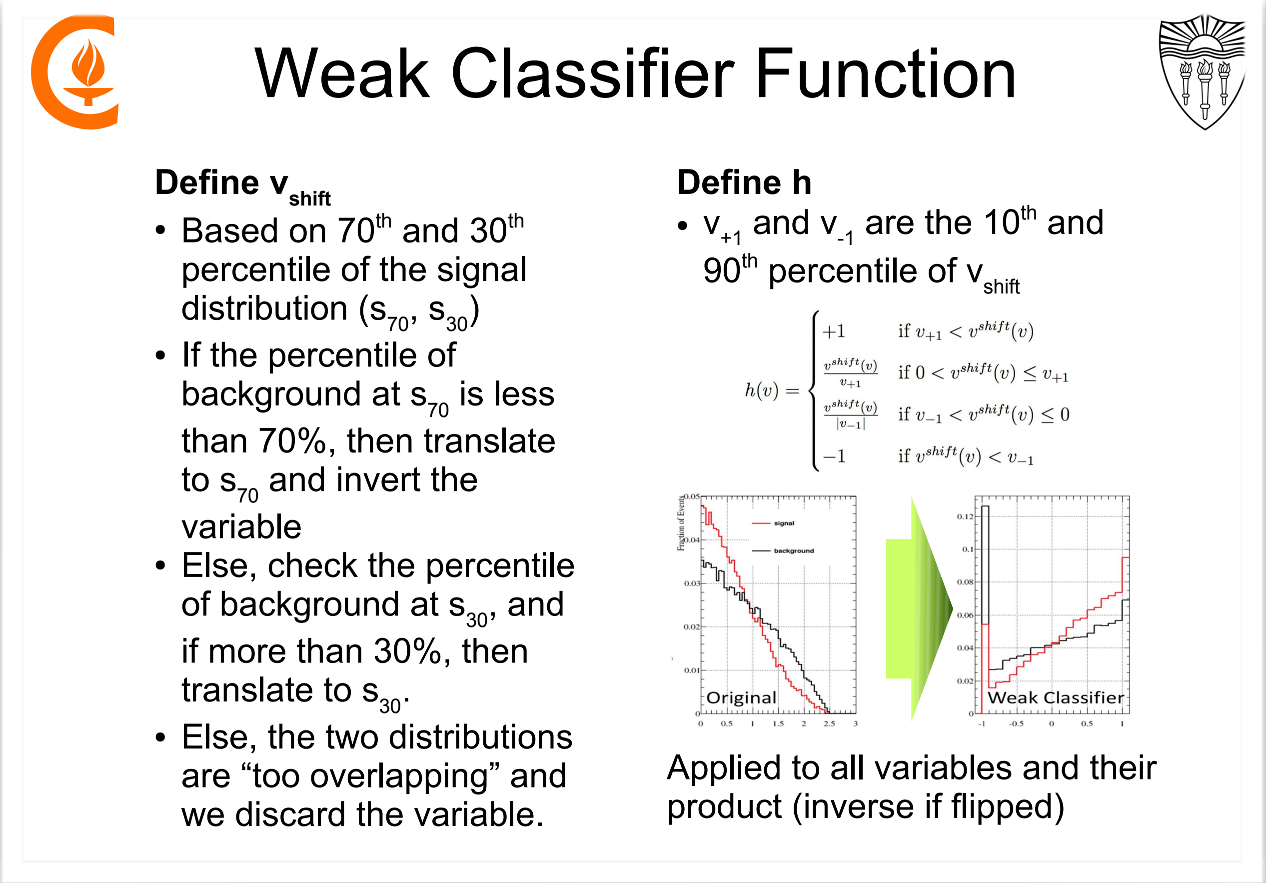
Sample Size and Folding

- 300k signal + 300k background total sample
  - **Training set**
    - 20 stratified, independent splits of sizes 100, 1000, 5000, 10k, 15k, 20k events
    - → Spread of classifier performance over the folds reported as the **uncertainty due to the choice of training sample, and initialization.**
  - **Testing set**
    - Remaining 100k+100k independent sample
    - → **Statistical error** on the classifier performance estimated using bootstrapping



Characterizing Variables

variable	description
$p_T^1/m_{\gamma\gamma}$	transverse momentum of the highest p_T photon divided by the invariant mass of the diphoton pair
$p_T^2/m_{\gamma\gamma}$	transverse momentum of the second-highest p_T photon divided by the invariant mass of the diphoton pair
$(p_T^1+p_T^2)/m_{\gamma\gamma}$	sum of the transverse momentum of the two photons divided by their invariant mass
$(p_T^1-p_T^2)/m_{\gamma\gamma}$	difference of the transverse momentum of the two photons divided by their invariant mass
$p_T^{\gamma\gamma}/m_{\gamma\gamma}$	transverse momentum of the diphoton system divided by its invariant mass
$\Delta\eta$	difference in $\eta = -\log\tan\left(\frac{\theta}{2}\right)$, where θ is the angle with the beam axis
ΔR	sum in quadrature of the separation of and ϕ, the azimuthal angle of the two photons ($\sqrt{\Delta\eta^2+\Delta\phi^2}$)
$\lvert\eta^{\gamma\gamma}\rvert$	the η value of the diphoton system



Weak Classifier Function

Define v_{shift}
- Based on 70th and 30th percentile of the signal distribution (s_{70}, s_{30})
- If the percentile of background at s_{70} is less than 70%, then translate to s_{70} and invert the variable
- Else, check the percentile of background at s_{30}, and if more than 30%, then translate to s_{30}.
- Else, the two distributions are "too overlapping" and we discard the variable.

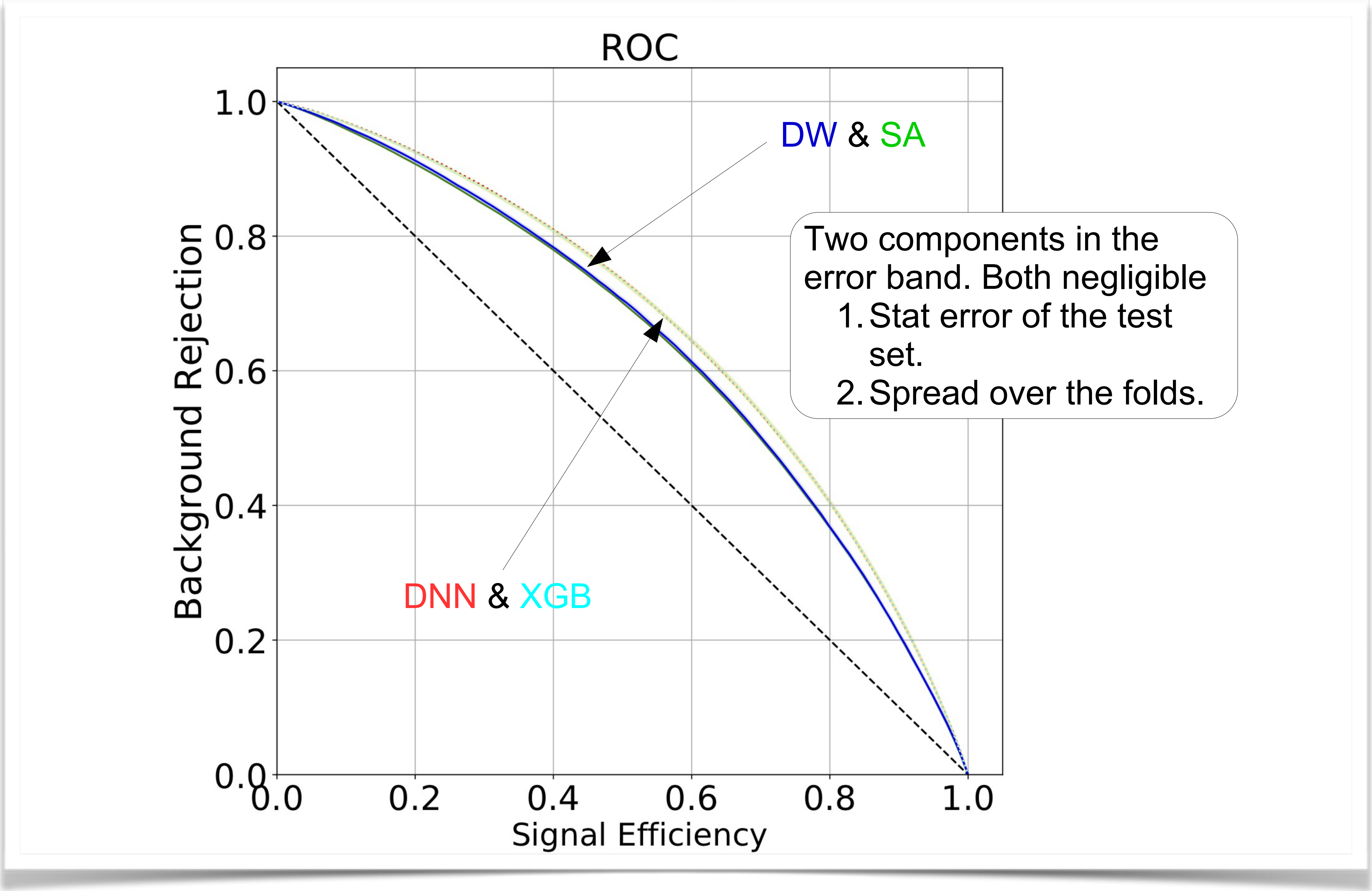
Define h
- v_{+1} and v_{-1} are the 10th and 90th percentile of v_{shift}

$$h(v) = \begin{cases} +1 & \text{if } v_{+1} < v^{shift}(v) \\ \frac{v^{shift}(v)}{v_{+1}} & \text{if } 0 < v^{shift}(v) \le v_{+1} \\ \frac{v^{shift}(v)}{|v_{-1}|} & \text{if } v_{-1} < v^{shift}(v) \le 0 \\ -1 & \text{if } v^{shift}(v) < v_{-1} \end{cases}$$

Applied to all variables and their product (inverse if flipped)

Performance

Nature 550, 375-379 (2017)



Two components in the error band. Both negligible

1. Stat error of the test set.
2. Spread over the folds.

- First application of D-Wave quantum annealing in high-energy physics
  - Via solving an Ising model
- Good performance, but not at the same level of DNN & xgboost

Image courtesy of J. R. Vlimant

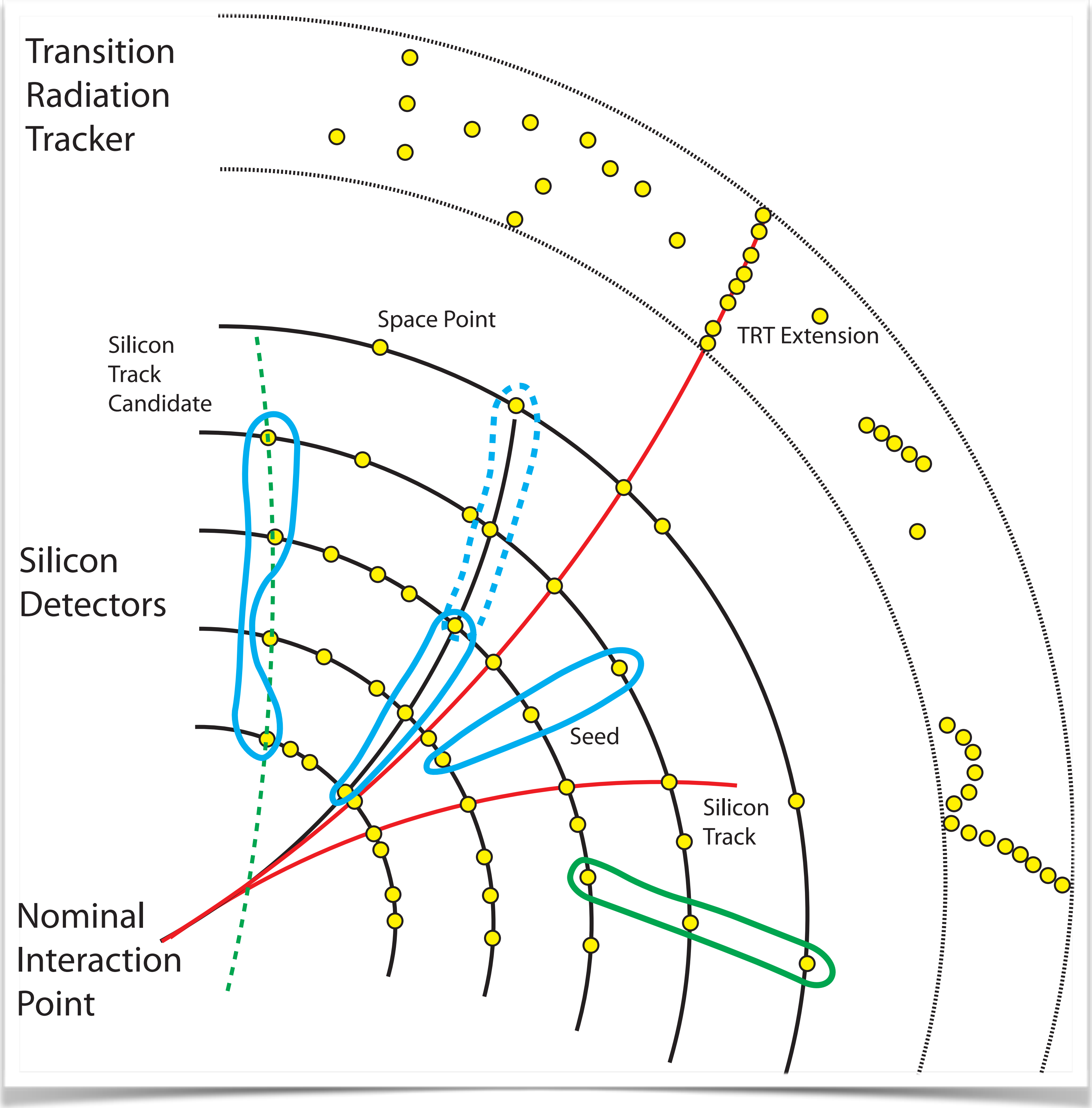
Conclusion

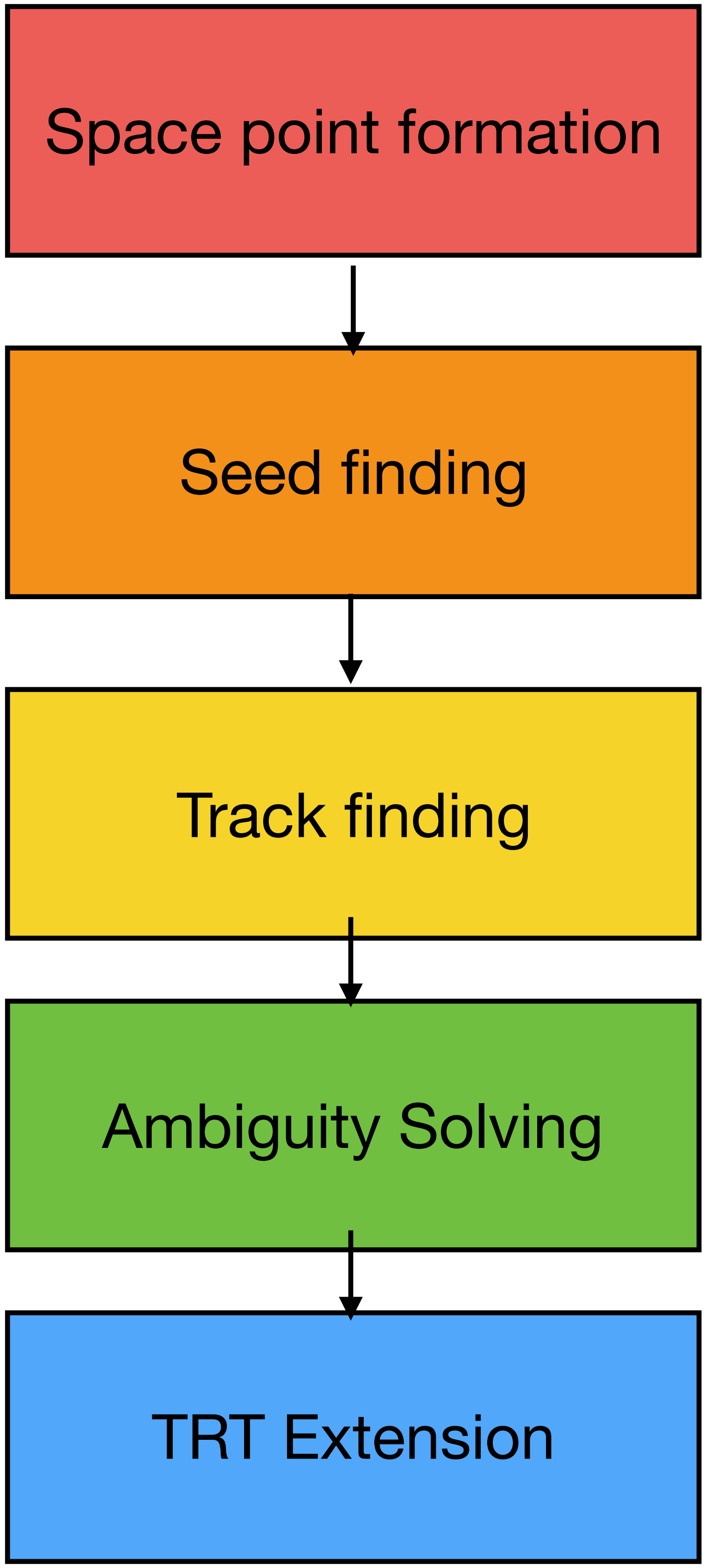
- HL-LHC is coming very soon!
- Major upgrade to accelerator and detectors which will enable an extensive and exciting physics program
- Fully exploiting this data is the topic of an extensive R&D program in software and computing
- Too early to tell, but, perhaps quantum computing can play a role?
    - Showed examples of preliminary research in pattern recognition on quantum computers
    - Ideas currently being pursued include MC simulation, track reconstruction, Higgs analyses

Back up

Track Reconstruction: Current Approaches

Multi-step iterative Kalman filter approach





Space point formation

Seed finding

Track finding

Ambiguity Solving

TRT Extension