Variational Quantum Algorithms

for Combinatorial Problesm at Colliders

Jacob Scott

In collaboration with

KC Kong, Cosmos Dong, Taejoon Kim, Myeonghun Park

Particle Physics on the Plains

October 14, 2023

Structure of talk

1. Problem at hand
   - And current alternatives
2. Quantum computing basics
   - Qubits & quantum gates
3. What are variational quantum algorithms (VQAs)?
   - QAOA + its one of its derivatives
   - & FALQON
4. Preliminary results

Interesting picture

The problem

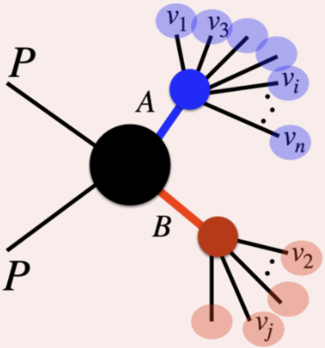


FIGURE 1: $2 \to 2 \to n$ collision, from Kim et al. 2021 [2111.07806]

- Binary classification: did particle ν_i come from A or B?
- A QUBO[1] problem: Quadratic Unconstrained Binary Optimization problem:

For an n-bit string x, find x^* such that it minimizes

$$f_w(x) = \sum_{i,j=1}^{n} w_{ij} x_i x_j$$

[1] A rare occurance of a 'Q' *not* standing for 'quantum'

Specific case: $pp \to t\bar{t}$

Two dominant decay modes for t:

- Leptonic: $t \to W^+ b \to \ell \nu_\ell b$
  - Cleaner – only one jet
  - But there's missing momentum
- Hadronic: $t \to W^+ b \to q\bar{q}b$
  - No missing momentum
  - But messier – 3 quarks creating 3 jets

Our focus has been on the latter case:

$$pp \to t\bar{t} \to q\bar{q}q'\bar{q}'b\bar{b}$$

Alternate methods

- Kinematic methods like the hemisphere method
  - Make assertions and assumptions of the kinematics of the system
- Machine learning
  - Great at finding and learning patterns
  - Lot of research that can be applied in HEP [6, 4]
  - Classical, so still limited by exponential growth of complexity
- Quantum annealing
  - Can find the global minimum
  - Smaller energy gaps between eigenvalues requires larger relaxation times
  - Limited with degenerate eigenvalues

Qubits

- A 2-dimensional state that exists somewhere on the Bloch sphere:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \quad \text{where} \quad |\alpha|^2 + |\beta|^2 = 1$$

- Measurements done in "computational basis":

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- Also important, the Hadamard basis:

$$|\pm\rangle = \frac{1}{\sqrt{2}}\big(\,|0\rangle \pm |1\rangle\,\big)$$

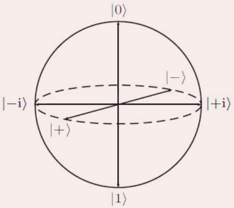


FIGURE 2: Bloch sphere, from wikipedia

Quantum Gates

One (qu)bit

- Classically, only the NOT gate
- Quantum mechanically, any 2x2 unitary that moves a qubit around the Bloch sphere:

$$U_{2\times 2}\,|\psi\rangle = |\psi'\rangle \Longrightarrow \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix}$$

- As a circuit:

$$|\psi\rangle \; -\!\boxed{U_{2\times 2}}\!- \; |\psi'\rangle$$

Quantum Gates

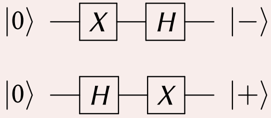
One (qu)bit

- Useful in this talk are Pauli rotation gates generated by Pauli matrices:

$$R_X(\theta) = e^{-i\theta\sigma_x}, \quad R_Y(\theta) = e^{-i\theta\sigma_y}, \quad R_Z(\theta) = e^{-i\theta\sigma_z}$$

- When $\theta = \pi/2$, $R_X \propto \sigma_x$, etc. These are the X, Y, and Z gates: π rotations on the Bloch sphere about their respective axes.
  - e.g. $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$ so often called the NOT gate
- Also Hadamard gates: $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$

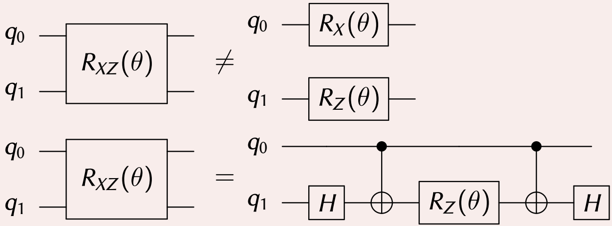
Examples of using gates:



Quantum Gates

Two (qu)bits

- Classically, AND, NOR, XOR, etc. Not unitary, i.e. can't work backwards from result to initial conditions
- Quantum mechanically, can start to exploit entanglement
- Of import this talk are tensor products of Pauli matrices
  - e.g. $\exp\left[-i\theta\sigma_x^0\sigma_z^1\right]$ is an R_{XZ} gate on qubits 0 and 1.



other common or important gates: CNOT [controlled NOT], Tiffoli [controlled AND], SWAP [interchanges qubits], Bill [short for BILLONAIRE]

Variational Quantum Algorithms

If you know neural networks keep that in mind

- We have some operator C whose expectation value, $\langle \boldsymbol{\theta} | C | \boldsymbol{\theta} \rangle$ we want to extremize
  - The state $|\boldsymbol{\theta}\rangle$ is parameterized by parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$
  - e.g. for this talk, think C is a Hamiltonian and we wish to find the ground state energy
- $|\boldsymbol{\theta}\rangle$ can be created with a quantum circuit: $U(\boldsymbol{\theta})\,|0\rangle = |\boldsymbol{\theta}\rangle$
- Use a classical optimizer to update the parameters $\boldsymbol{\theta}_n \to \boldsymbol{\theta}_{n+1}$ and repeat

Quantum Approximation Optimization Algorithm (QAOA)

- Start with Hamiltonian

e.g. $a(t) = t/T$

$$H(t) = \big(1 - a(t)\big) H_M + a(t) H_P$$

such that $a(0) = 0$ and $a(T) = 1$

  - Means that we start in system H_M and, if T is large, slowly evolve into state H_P
  - **Mixer Hamiltonian**: H_M, an easily solvable system with easily initializable eigenstates, usually $H_M = \sum \sigma_x^k$
  - **Problem Hamiltonian**: H_P, Hamiltonian whose minimum energy state is the answer we wish to find

- Exploit the adiabatic theorem: start in ground state of H_M, evolve slowly into ground state of H_P

Original paper: Farhi et al. 2014 [1411.4028]

Quantum Approximation Optimization Algorithm (QAOA)

Remember from our quantum mechanics courses:

$$H|\psi\rangle = i\frac{\partial}{\partial t}|\psi\rangle \implies |\psi\rangle = e^{-iHt}|\psi_0\rangle$$

In our case, $|\psi\rangle = |\boldsymbol{\theta}\rangle$. What do we choose for $|\psi_0\rangle$? How do we do time-evolution on a circuit?

Result:

Ground state of H_M

- Our state is: $|\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle = \prod_{j=1}^{p} U(\beta_j, H_M) U(\gamma_j, H_P) |+\rangle^{\otimes n}$

$$\text{where} \quad U(\beta_j, H_M) = \exp\left[-i\beta_j H_M\right],$$
$$U(\gamma_j, H_P) = \exp\left[-i\gamma_j H_P\right]$$

and p is the **depth** of the circuit

Original paper: Farhi et al. 2014 [1411.4028]

Quantum Approximation Optimization Algorithm (QAOA)

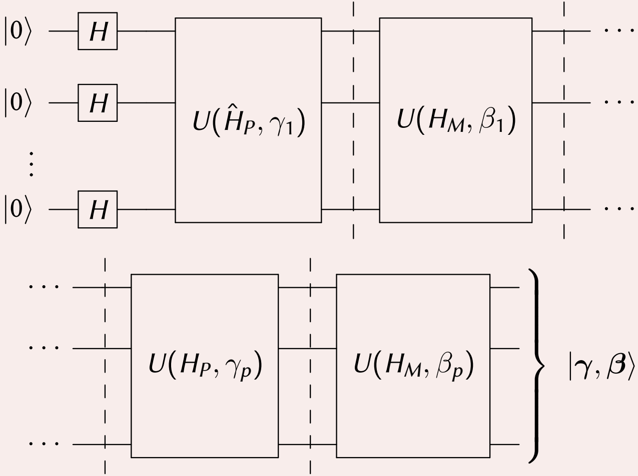
Continuous time-evolution approximated by small discrete steps alternating applications of the Hamiltonians

- **Circuit**: $|\boldsymbol{\beta}, \boldsymbol{\gamma}\rangle = \prod_{j=1}^{p} U(\beta_j, H_M) U(\gamma_j, H_P) \, |+\rangle^{\otimes n}$
  - This approximation is exact when $p \to \infty$
- **Mixer Hamiltonian**: $H_M = \sum_{k=1}^{n} \sigma_x^k$

  Just X rotations on every qubit
  - As quantum gates:
    $$\exp\left[-i\beta_j \sum_{k=1}^{n} \sigma_x^k\right] = \prod_{k=1}^{n} e^{-i\beta_j \sigma_x^k} = \prod_{k=1}^{n} R_X(\beta_j)$$
- **Problem Hamiltonian**: H_P depends on problem
  - Can we write it as Pauli matrices?
- **Goal**: minimize $\langle \boldsymbol{\beta}, \boldsymbol{\gamma} | H_P | \boldsymbol{\beta}, \boldsymbol{\gamma} \rangle$

Original paper: Farhi et al. 2014 [1411.4028]

Quantum Approximation Optimization Algorithm (QAOA)



Original paper: Farhi et al. 2014 [1411.4028]

Quantum Approximation Optimization Algorithm (QAOA)

But there are problems!

- How quickly does it converge?
  - We don't have the technology for a circuit of $\mathcal{O}(10)$ depth, let alone ∞ depth
- How easy is it to navigate the parameter space?
  - Are there many local minima to get stuck in?
  - Barren plateaus?
- "...short-depth QAOA is not really the digitized version of the adiabatic problem, but rather an ad hoc ansatz, and as a result should not be expected to perform optimally, or even well."
  - From Zhu et al. 2022 [2005.10258]

So we say good riddance to the justification of this approximate adiabaticity and consider other ansatzes

---

Original paper: Farhi et al. 2014 [1411.4028]

Multi-Angle QAOA (ma-QAOA)

- **Expressibility** is a circuit's ability to explore its Hilbert space
  - Or, for a single qubit, to traverse the Bloch sphere
- Idea: give *every* gate its own free parameter:

$$U(\beta_j, H_M) = \exp\left[-i\beta_j \sum_{k=1}^{n} \sigma_x^k\right]$$

$$\Downarrow$$

$$U(\boldsymbol{\beta}_j, H_M) = \exp\left[-i \sum_{k=1}^{n} \beta_{jk}\sigma_x^k\right]$$

- This allows for lower depth circuits that are more accessible with current quantum computers, i.e. NISQ era
- Tradeoff between the quantum and classical computers

Original paper: Herrman et al. 2021 [2109.11455]

Feedback-based ALgorithm for Quantum OptimizatioN (FALQON)

- A purely quantum algorithm – no optimization
- Considers the Hamiltonian: $H(t) = H_P + \beta(t) H_M$
- Want

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \psi(t) | H_P | \psi(t) \rangle \leq 0 \Longrightarrow A(t)\beta(t) \leq 0$$

  where $A(t) = \langle \psi(t) | i[H_M, H_P] | \psi(t) \rangle$
- Choose $\beta(t) = -A(t - 2\Delta t)$ and descritize: $\beta_{k+1} = -A_k$
- Our state is: $|\boldsymbol{\beta}\rangle = \prod_{j=1}^{p} U(\beta_j, H_M) U(H_P)$
  - where $U(\beta_j, H_M) = e^{-i\beta_j H_M \Delta t}$ and $U(H_P) = e^{-iH_P \Delta t}$

Setup

- With QAOA on our mind, is there a Hamiltonian that we can use?

$$H = (p_A^2 - p_B^2)^2 \quad \text{where} \quad \begin{cases} p_A = \sum x_i p_i \\ p_B = \sum (1 - x_i) p_i \end{cases}$$

  and $x_i = 1$ if particle i is associated with A, otherwise 0
- Setting $x_i = (1 + s_i)/2$,

$$H = \sum_{ij} J_{ij} s_i s_j \quad \text{where} \quad J_{ij} = \sum_{k\ell} (p_i \cdot p_k)\,(p_j \cdot p_\ell)$$

  Since $s_i = \pm 1$, we can write our Hamiltonian operator as

$$H_P = \sum_{ij} J_{ij} \sigma_z^i \sigma_z^j.$$

  Eigenvalues are ± 1 and eigenvectors are computation basis vectors

Setup

- One qubit for each final state particle
  - $n = 6$ qubits in the $t\bar{t}$ decay
- The mixer layer consists of 6 1-qubit R_X gates with rotation β_k
  - For ma-QAOA, there are rotations β_{ki} for the ith qubit
- The problem layer consists of 6 choose $2 = 15$ R_{ZZ} gates with rotation $J_{ij}\gamma_k$ between the i and j qubit
  - Or rotations $J_{ij}\gamma_{k,ij}$ for ma-QAOA

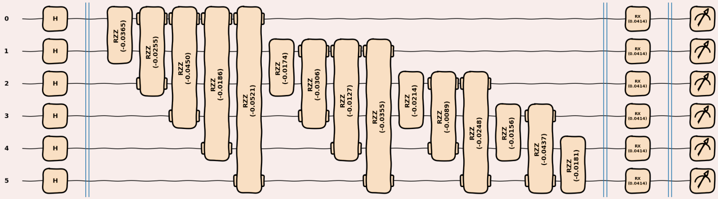
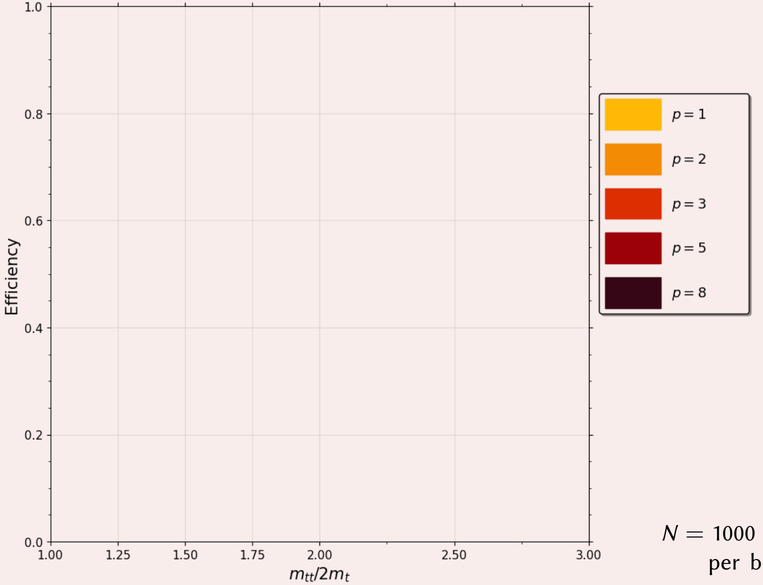


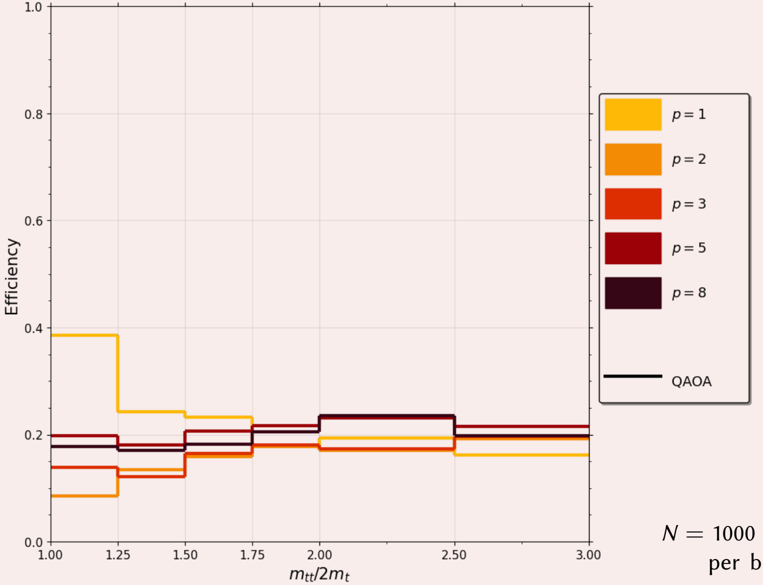
FIGURE 3: One-layer QAOA circuit

Parton-Level Results



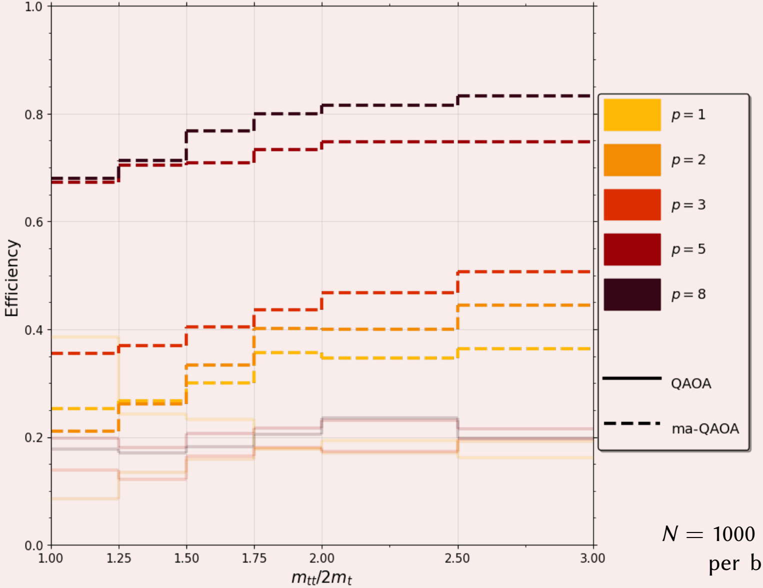
$N = 1000$ events per bin

Parton-Level Results



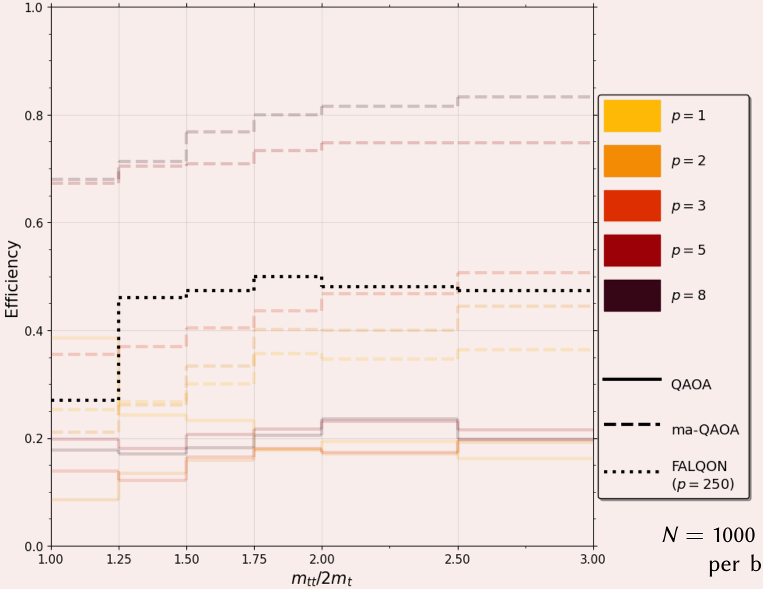
$N = 1000$ events per bin

Parton-Level Results



$N = 1000$ events per bin

Parton-Level Results



$N = 1000$ events per bin

Conclusion

Same interesting picture

1. VQAs can be used
2. Show promising results comparable or better than QA and hemisphere method
3. Linear growth in number of qubits

Further...

1. Added complications:
   - Missing momentum, e.g. dilepton channel
   - More jets, e.g. ISR, FSR
2. Better ansatzes/algorithms out there/to be found?
3. More appropriate H_P

References

[1] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. *A Quantum Approximate Optimization Algorithm.* 2014. arXiv: 1411.4028 [quant-ph].

[2] Rebekah Herrman et al. *Multi-angle Quantum Approximate Optimization Algorithm.* 2021. arXiv: 2109.11455 [quant-ph].

[3] Minho Kim et al. *Leveraging Quantum Annealer to identify an Event-topology at High Energy Colliders.* 2021. arXiv: 2111.07806 [hep-ph].

[4] Jason Sang Hun Lee et al. *Zero-Permutation Jet-Parton Assignment using a Self-Attention Network.* 2023. arXiv: 2012.03542 [hep-ex].

[5] Alicia B. Magann et al. "Feedback-Based Quantum Optimization". In: *Physical Review Letters* 129.25 (2022). DOI: 10.1103/physrevlett.129.250502. eprint: 2103.08619. URL: https://doi.org/10.1103%2Fphysrevlett.129.250502.

[6] Alexander Shmakov et al. "SPANet: Generalized permutationless set assignment for particle physics using symmetry preserving attention". In: *SciPost Physics* 12.5 (May 2022). DOI: 10.21468/scipostphys.12.5.178. eprint: 2106.03898. URL: https://doi.org/10.21468%2Fscipostphys.12.5.178.

[7] V. Vijendran et al. *An Expressive Ansatz for Low-Depth Quantum Optimisation.* 2023. arXiv: 2302.04479 [quant-ph].

[8] Linghua Zhu et al. *An adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer.* 2022. arXiv: 2005.10258 [quant-ph].

eXpressive QAOA (XQAOA)

Sacrifice away adiabaticity for even more expressibility

- Add another mixer Hamiltonian to ma-QAOA: $H_X = \sum_{k=1}^{n} \sigma_y^k$
- Our quantum state is now

$$|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}\rangle = \prod_{j=1}^{p} U(\boldsymbol{\alpha}_j, H_X) U(\boldsymbol{\beta}_j, H_M) U(\boldsymbol{\gamma}_j, H_P) \left|+\right\rangle^{\otimes n}$$

- ma-QAOA is just XQAOA when $\boldsymbol{\alpha} = \mathbf{0}$

Original paper: Vijendran et al. 2023 [2302.04479]

Adaptive Derivative Assembled Problem Tailored QAOA (ADAPT-QAOA)

The kitchen sink emporium

- The choice of mixer Hamiltonian is not fixed but rather chosen from a *pool*, $\mathcal{A}$, in an iterative fashion
- Choice made by whichever maximizes energy gradient. Calculate

$$\Delta E_k(A_j) = \frac{\partial}{\partial \beta_k} \langle \psi_k | H_P | \psi_k \rangle \Big|_{\beta_k = 0}$$

  for each A_j then choose $A_k = \underset{A_j \in \mathcal{A}}{\text{argmax}}\, \Delta E_k(A_j)$ as the mixer Hamiltonian for layer k.
- Optimize circuit as with normal QAOA and repeat for next layer

---

Original paper: Zhu et al. 2022 [2005.10258]