Combinatorial Optimization with "Quantum" machine

Myeonghun Park

(Seoultech / KIAS)

based on arXiv:2111.07806

with Minho Kim, Pyungwon Ko, Jae-hyeon Park

2023 Ai and Qi Application in Fundamental Physics

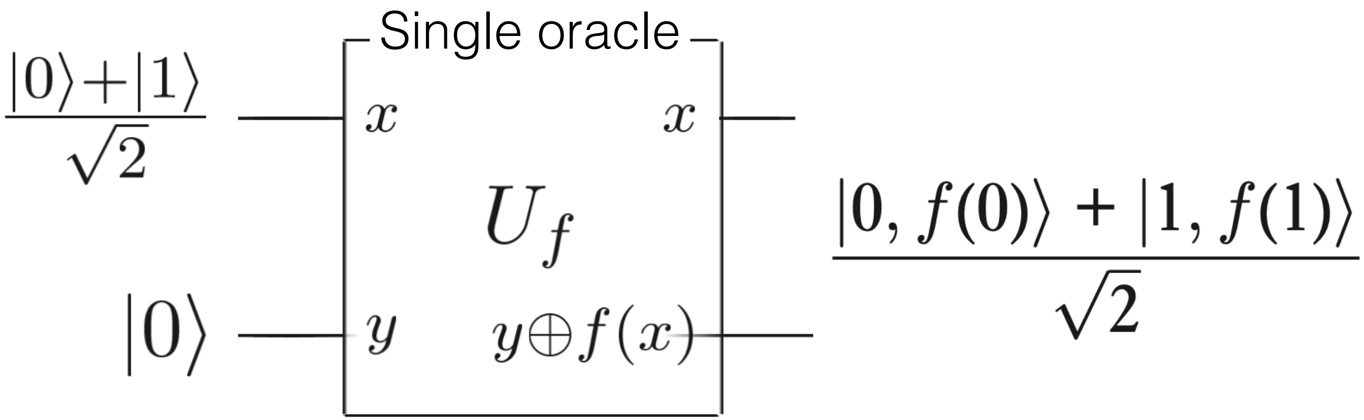
Quantum Computung

- There have been great works and focusing on QC.
  - Daniel's talk

- But at least it is **very interesting** to think about **computing algorithms being operated with the "quantum nature"** as a "physicist".

My "naive" understanding on quantum advantage

Quantum superposition

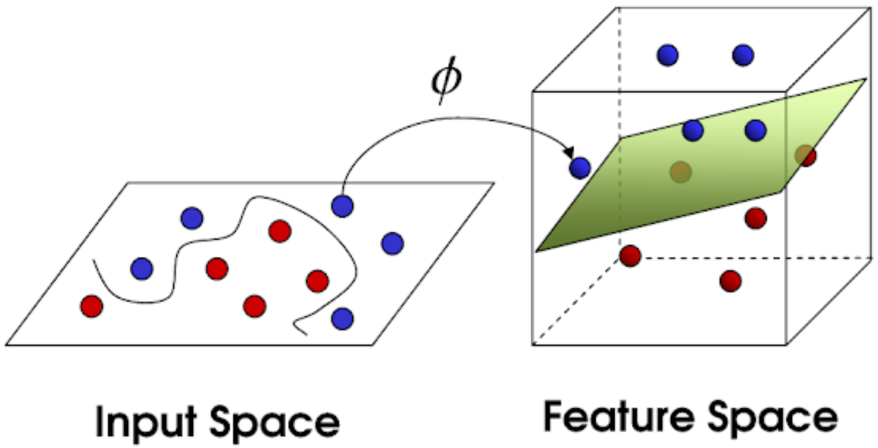


copied from N&C CH1

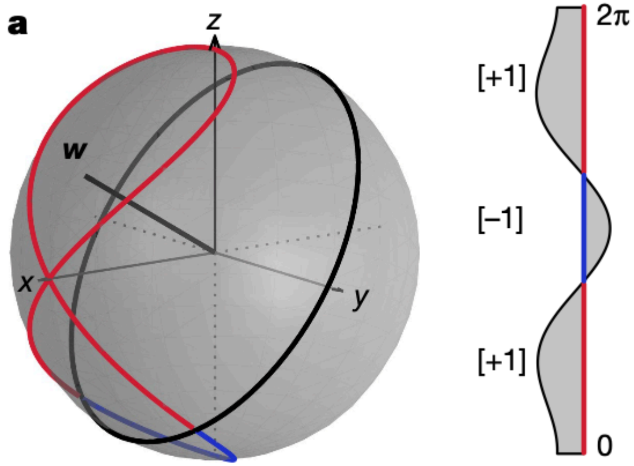
- leads to "**quantum parallel**" computing

Quantum Hilbert space

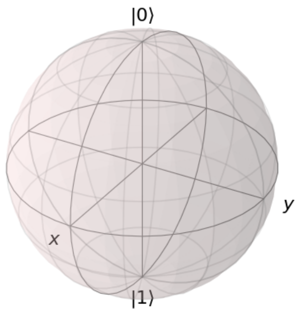
classical ML Kernel Method



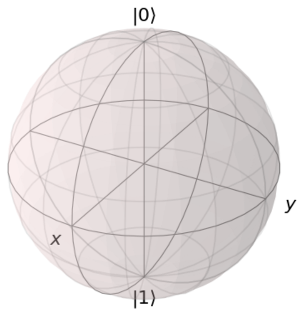
Quantum kernel function



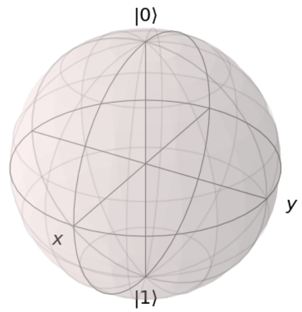
$$|\Psi\rangle = \sum a_{i_1 i_2 \cdots i_n}$$



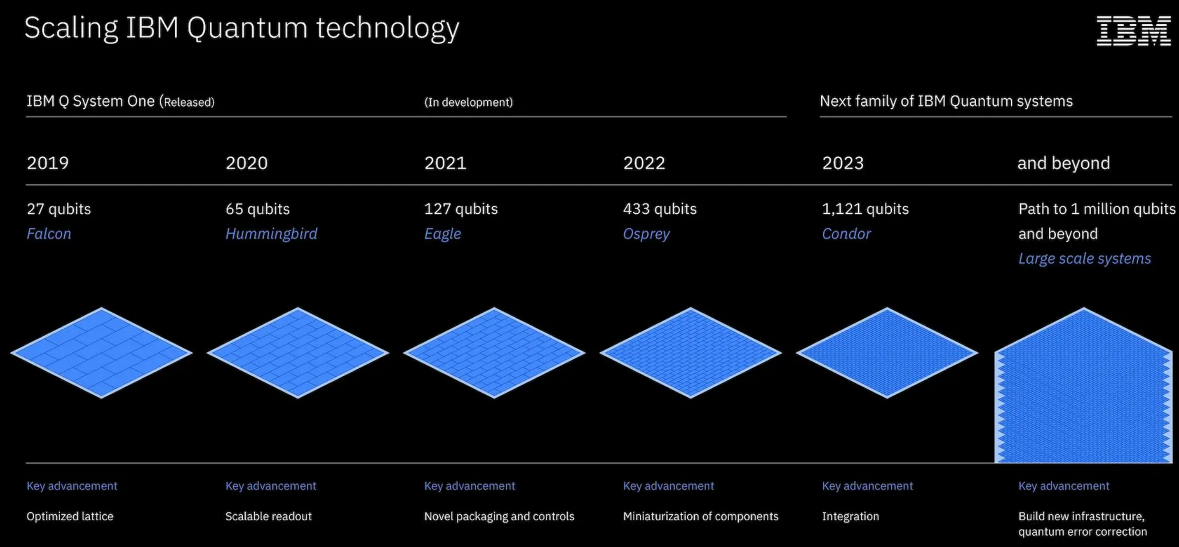
$\otimes$



$\cdots \otimes$



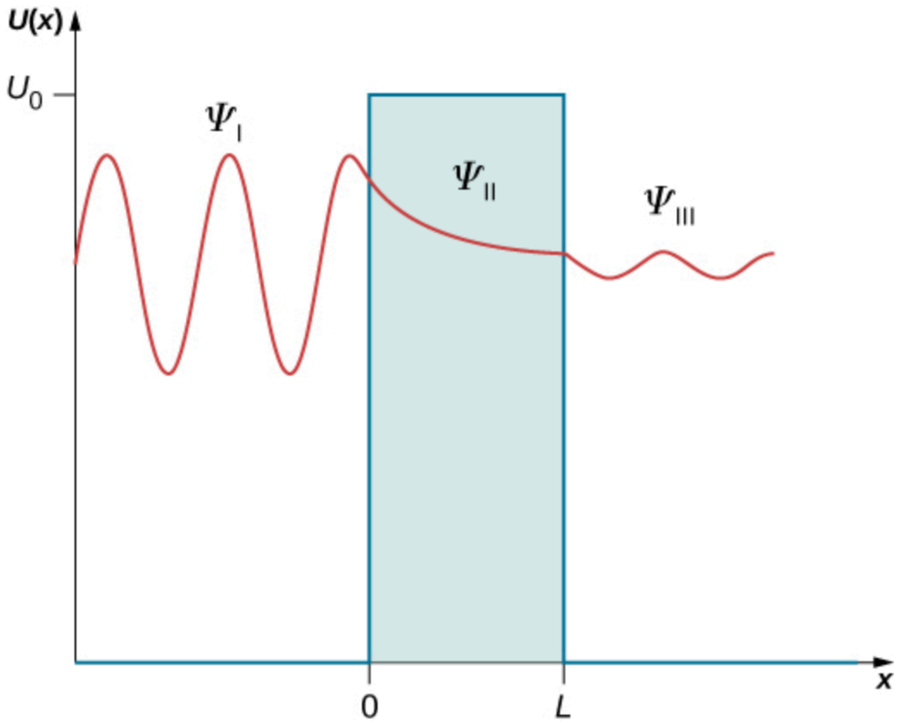
- Expressing and manipulating input data in an **exponentially large** **(2^n) and "compact" Quantum Hilbert Space**.

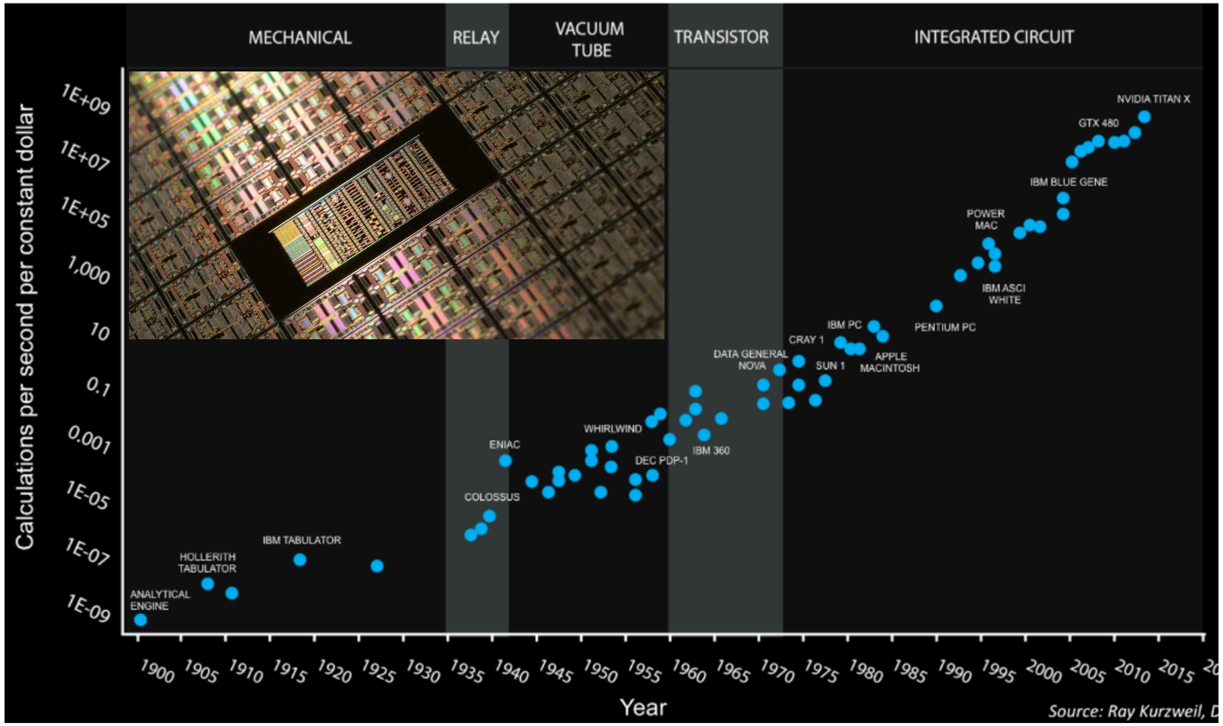


Current 433 Qubits (IBM Ospery)

What else ?

Quantum tunneling



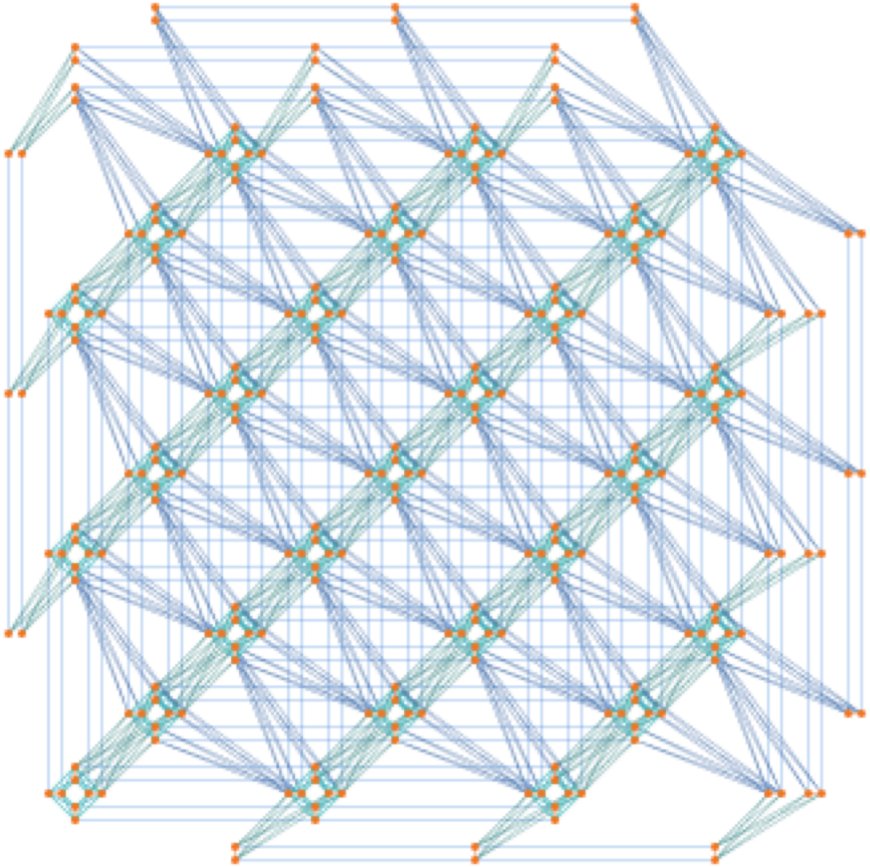


- which is the big "barrier" to make very tiny chip,

  There would be the end of Moore's law
  - uncontrolled leakage from Quantum tunneling gives the errors in computing

"Quantum" annealer

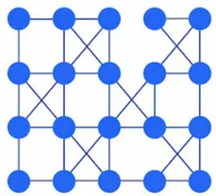
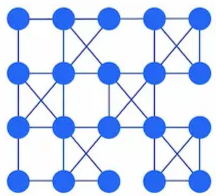
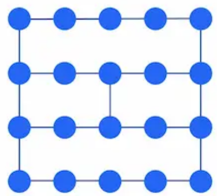
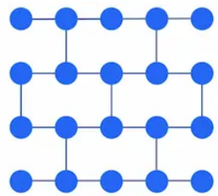
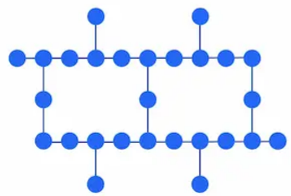
	2000Q	Advantage
Graph topology	Chimera	Pegasus
Graph size	C16	P16
Number of qubits	> 2000	> 5000
Number of couplers	> 6000	$> 35,000$
Couplers per qubit	6	15



- Current "Advantage" machine has **5000+ qubits** (though limited couplers ~ 35,000 $\ll {}_{5000}C_2 \simeq 10^7$

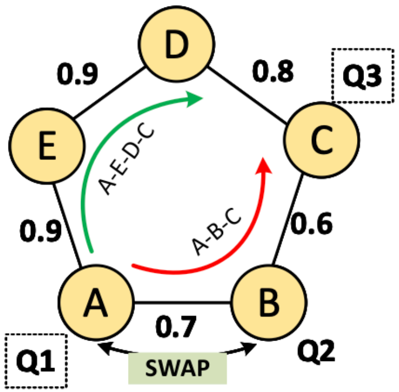
For Gated-QC

- We need a "connection" to operate between arbitrary two qubits (e.g. controlled-gate)

Processor	Penguin v1	Penguin v2	Penguin v3	Penguin v4	Falcon r4
					
Avg. qubit connectivity	3.9	3.7	2.3	2.3	2.1

- Due to "error-propagation", Gated-QC reduces the connectivity

= The number of required qubits to program > number of qubits in your circuit



(a)

Path	#SWAP	P(Success)
A-B-C	1	0.42
A-E-D-C	2	0.567

(b)

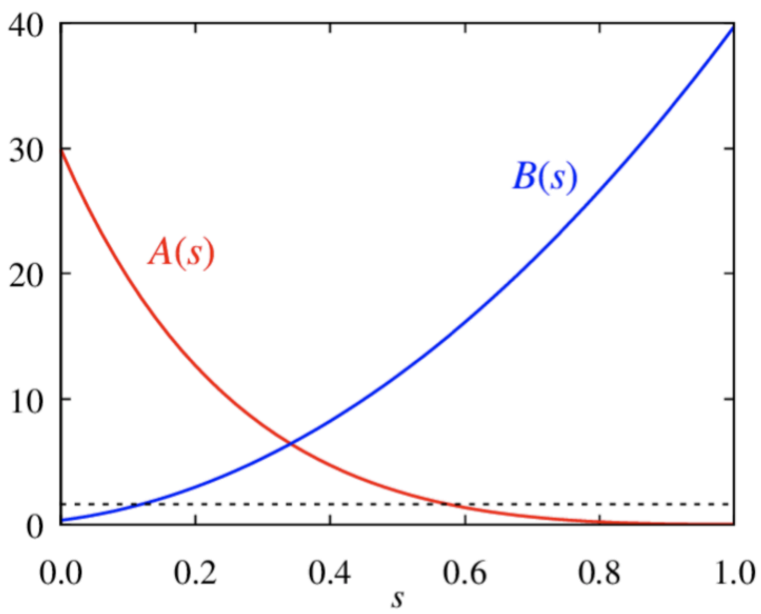
- Shortest path does not guarantee the lowest error rate (arXiv:1805.10224)

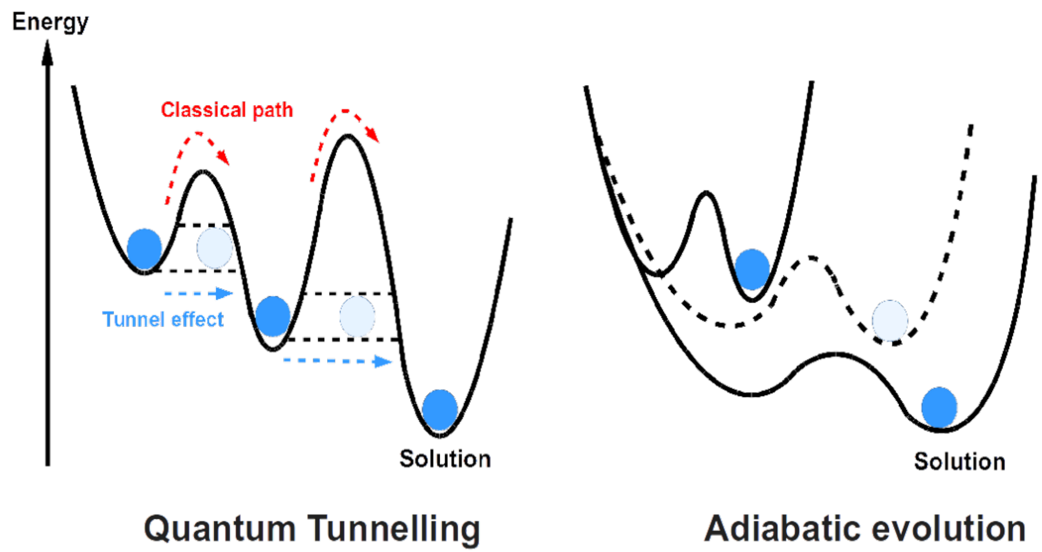
Quantum Annealing

- With adiabatic theorem, we can find the ground state of a complicate hamiltonian H_{QUBO} starting from simple H_0.

$$H_{\text{QA}} = A(s)\,H_0 + B(s)\,H_{\text{QUBO}} \text{ with } H_0 = \sum \sigma_i^x \text{ and } H_{\text{QUBO}} = \sum J_{ij}\sigma_i^z\sigma_j^z + \sum h_i\sigma_i^z$$

(T. Kadowaki and H. Nishimori, Quantum annealing in the transverse ising model, 1998)





- Annealing time $< 2000\mu s$, (mostly) $\mathcal{O}(10)\mu s$

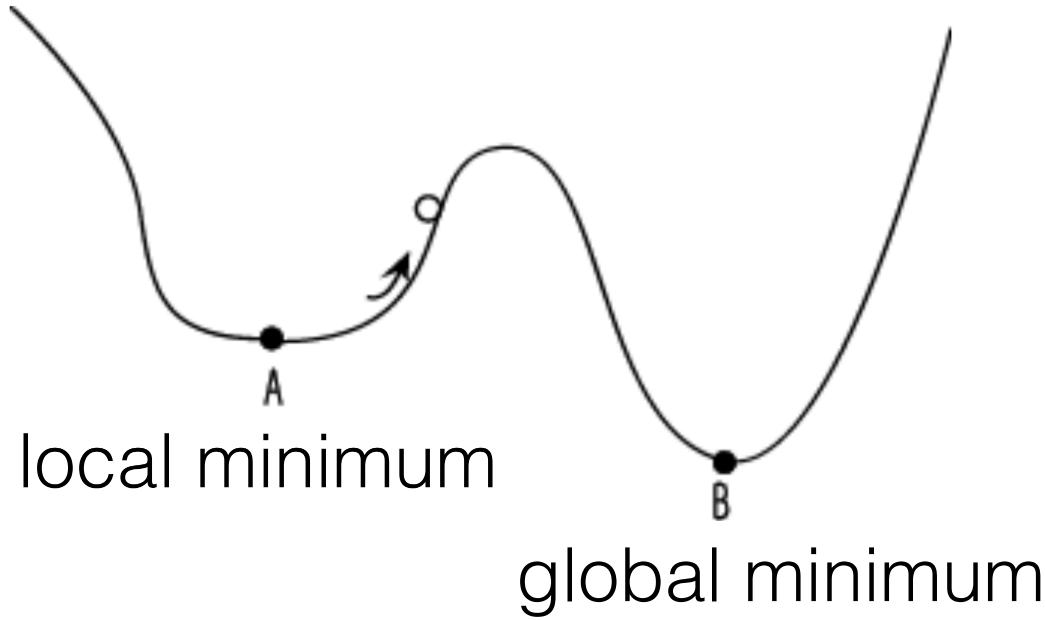
"Quantum annealing"

- claims to utilize "quantum tunneling" to find the minimum of the hamiltonian

$$H_{\text{QUBO}} = \sum J_{ij}\sigma_i\sigma_j + \sum h_i\sigma_i$$

for **Q**uadratic **U**nconstrained **B**inary **O**ptimization problems.

"Classic" minimization method (for Ising hamiltonian)



Simulated annealing

- Go to the next spin state $s_n \rightarrow s_{n+1}$

  1) If $E_n > E_{n+1}$: go to the lower energy

  2) If $E_n < E_{n+1}$, go with a probability of $e^{-\frac{E_{n+1} - E_n}{k_B T}}$ to **jump out**

(A "temperate $T \rightarrow 0$. With large T, SA can jump out local minimum)

Any good example

in High Energy Physics to demonstrate

the **advantage** from "Quantum tunneling",

which **cannot** be solved with "classical" optimization method ?

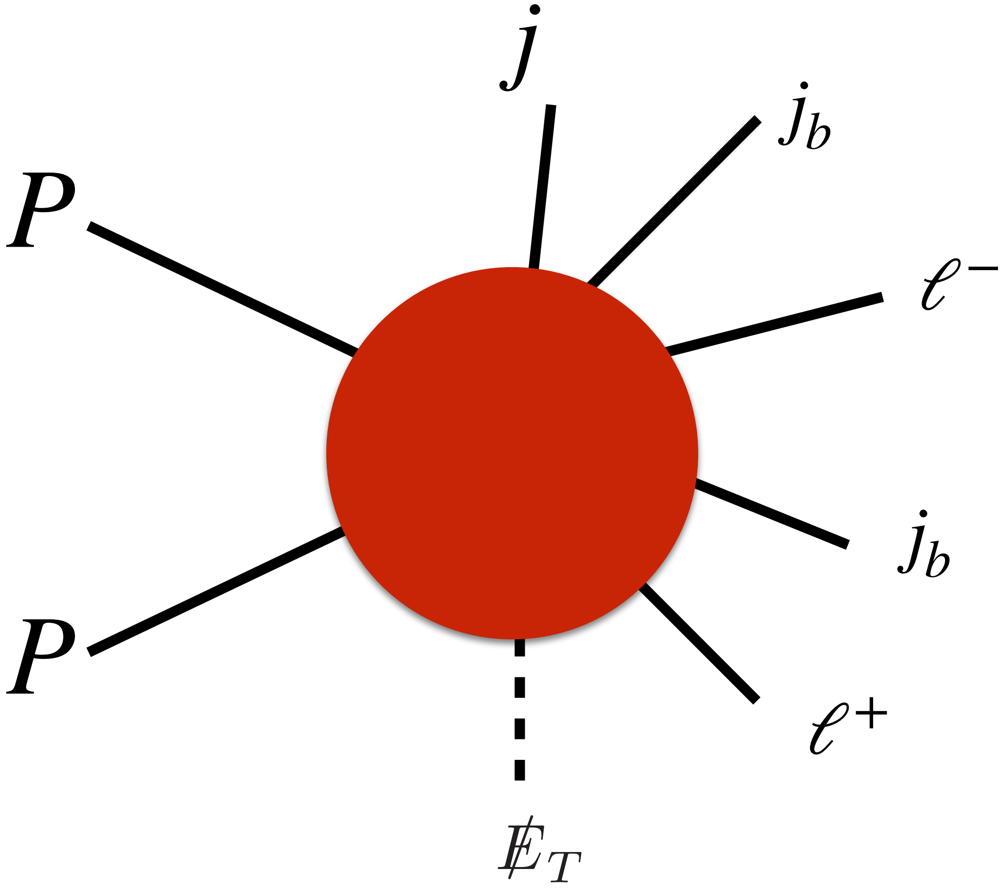
Hunt for new physics afterwards

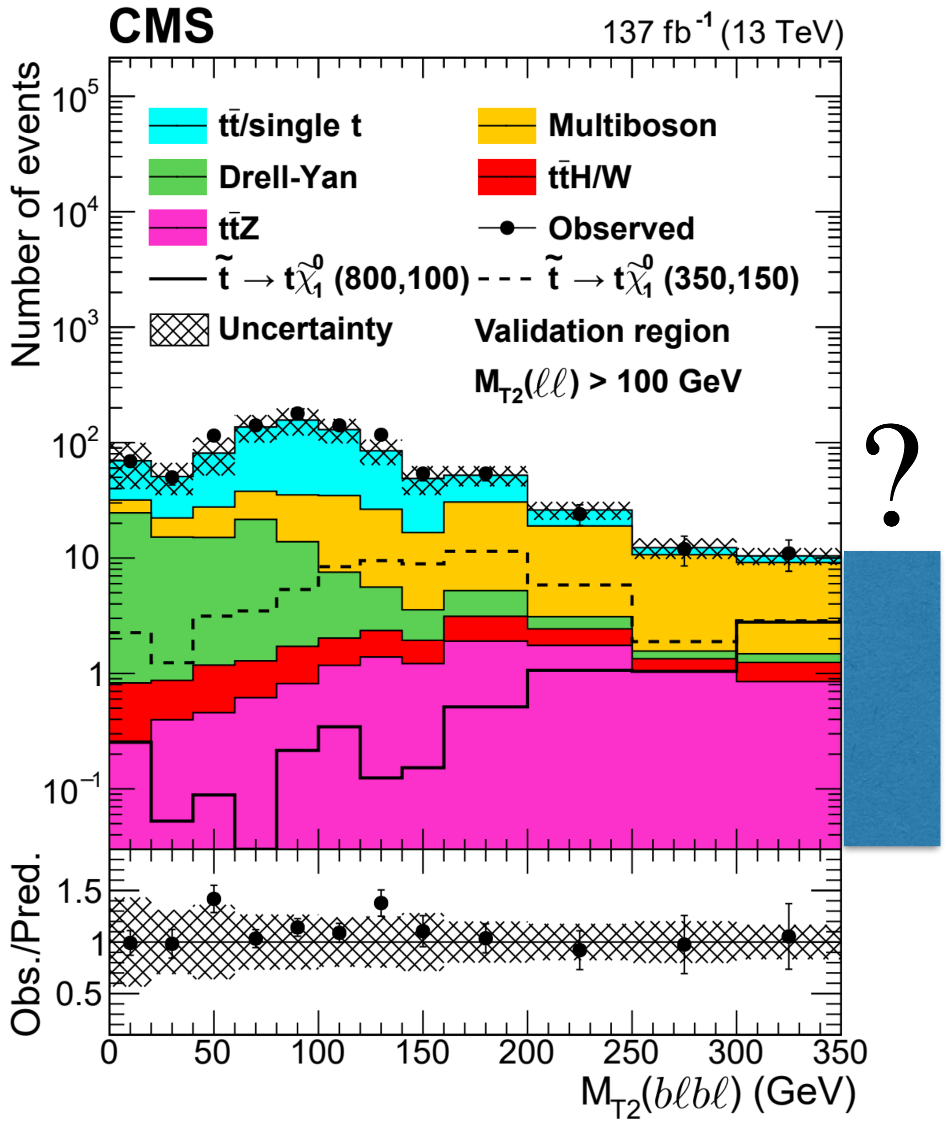
1. **Anomaly detection** (different from SM expectations)
   - Need to have precise tools (importance of MC)

2. Try to interpret a new signal with **various** model assumptions or **Model-independent way** so called simplified model
   - For each model, we start with **specific** "feynman-diagram"
     (**event-topology**, without specific spin assignment.)

   - Determine parameters (spin, mass) with various methods

Example: anomaly





a simple kinematic variable

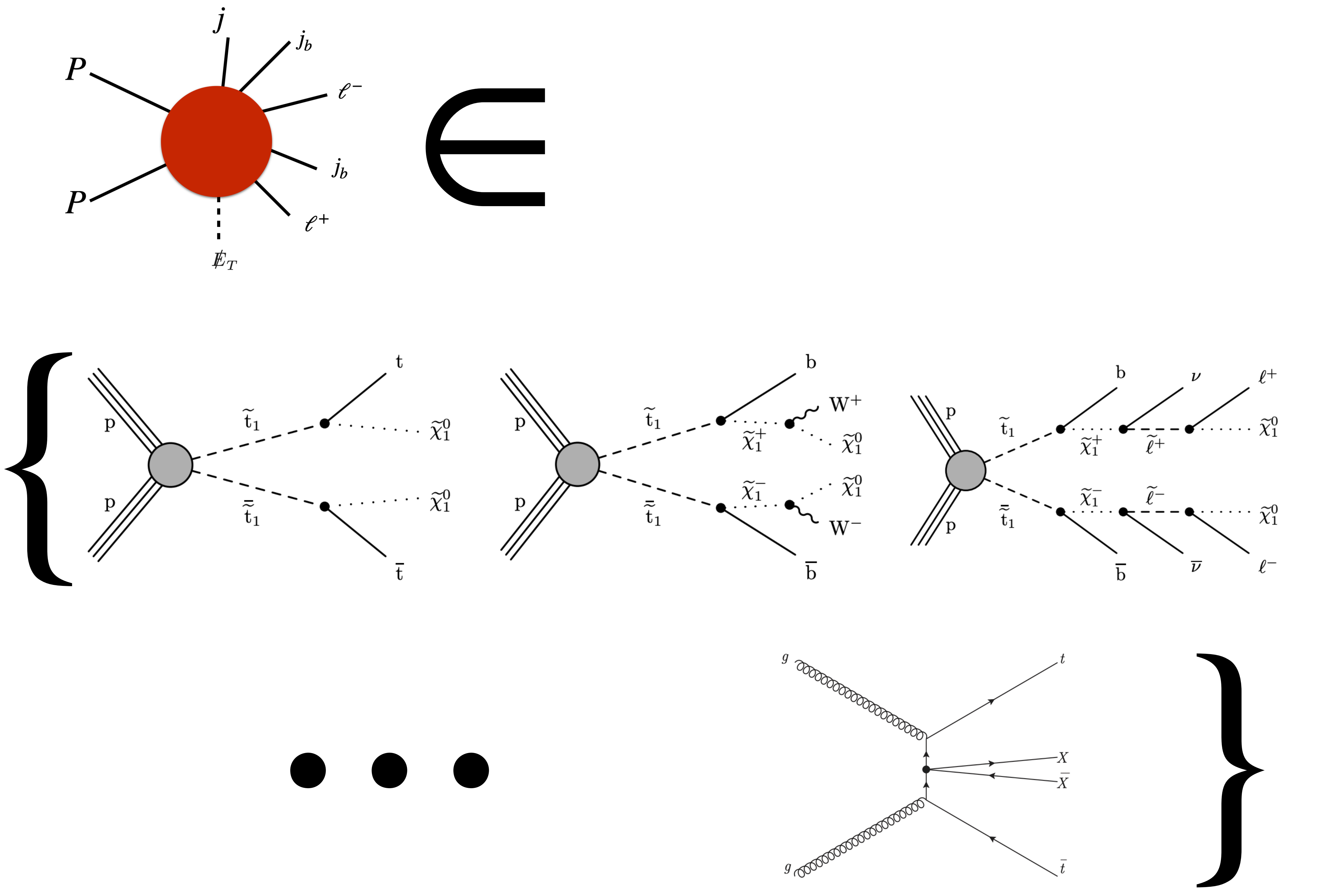


diagram from Lian-Tao Wang et.al. arxiv:1303.6638

Purely bottom-up approach

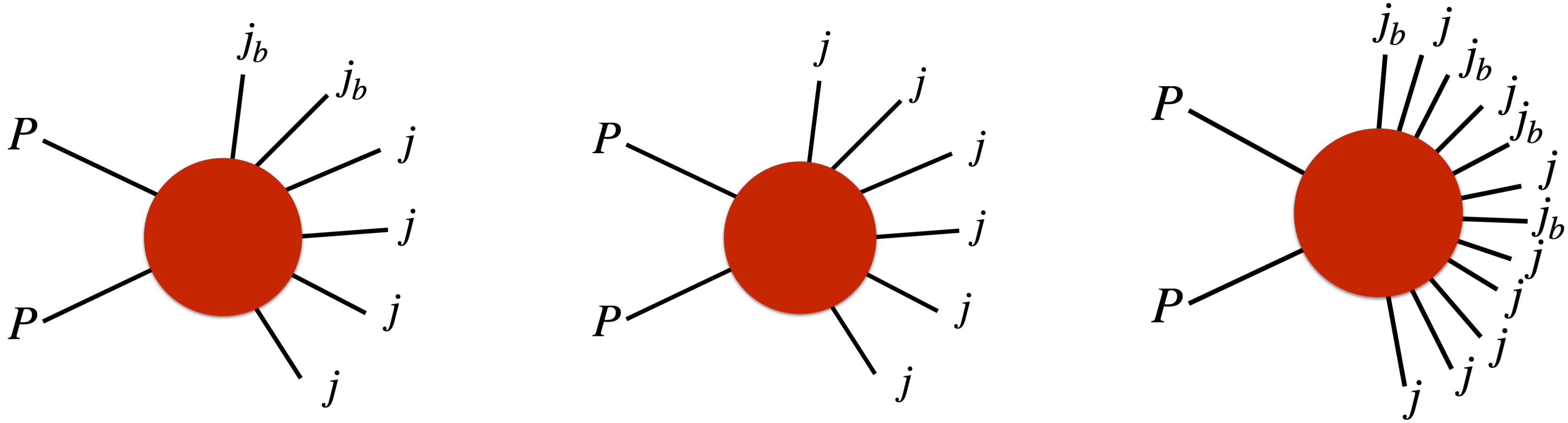
1. **Figure out what is the relevant event-topology behind anomalous (deviation from SM) events.**

2. Check the mass spectrum.

3. Check spin configuration.

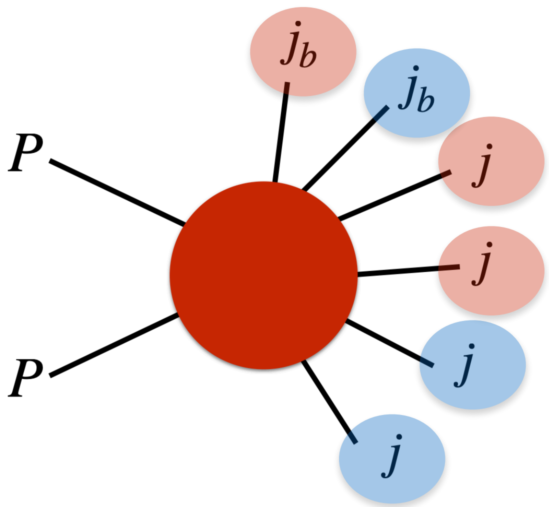
- So far, there are very few literatures for ***#1.***
  Here I will introduce how one can **identify the event-topology**

Our examples (multi-jets)



1. Under the a simple assumption: $pp \to X, Y \to \{j_x\} \cup \{j_y\}$
   (No prejudices on X and Y)

2. Find a right **combination** to reconstruct X and Y particles.
   $\to$ Read off information on **Mass** and **Spin** from event reconstruction.



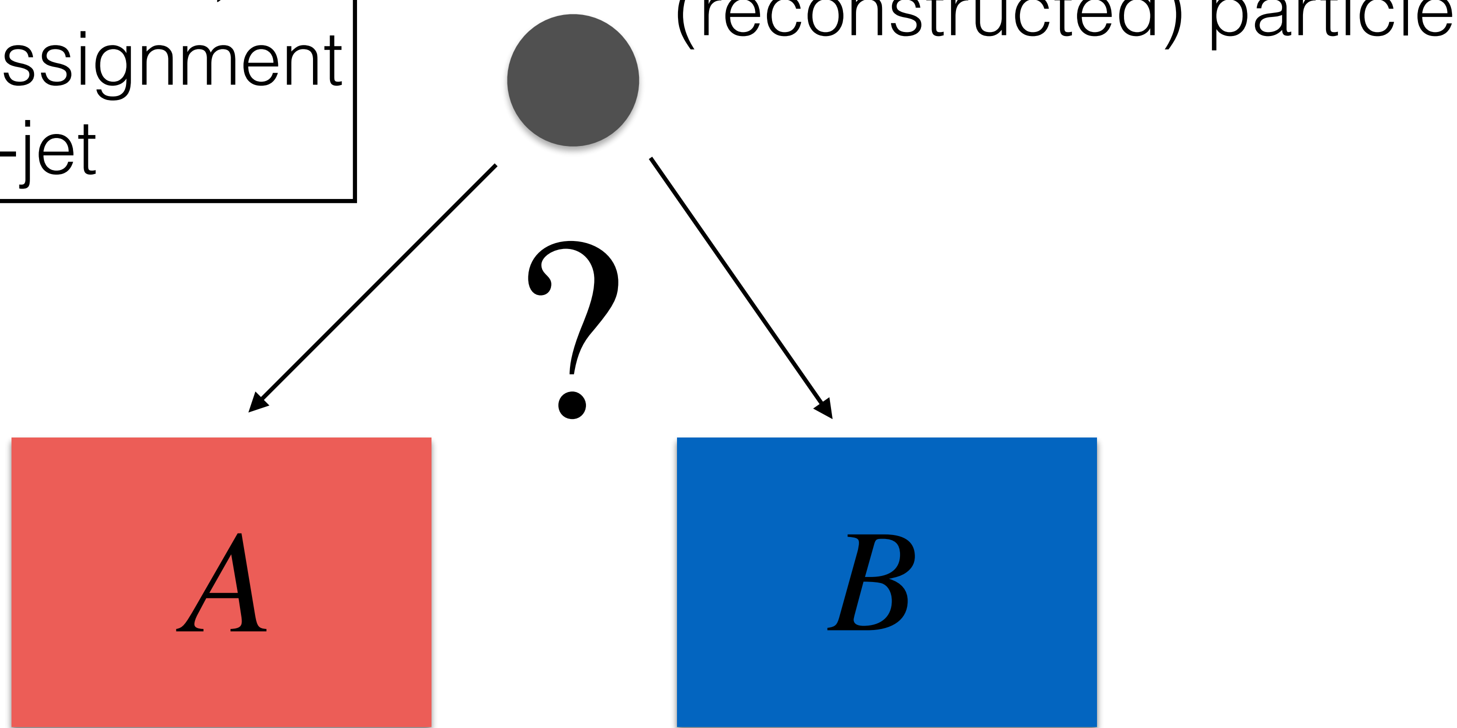
- Standard example of six jets

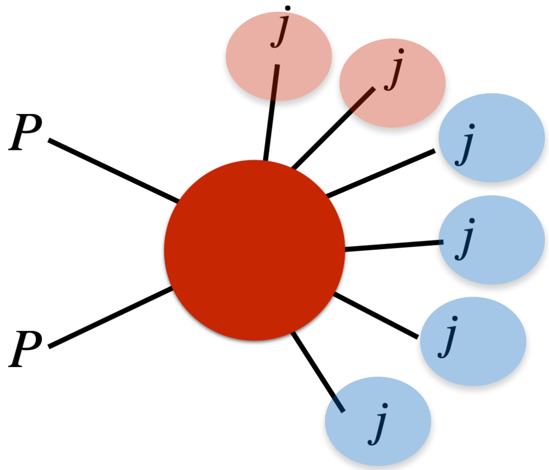
$$pp \to t\bar{t} \to \boxed{\{j_b, (W \to jj)\}} \cup \boxed{\{j_b, (W \to jj)\}}$$

(when A and B have same mass)

- Right answer is $(n_A, n_B) = (3,3)$

$2^6 = 64$ cases,
no special assignment
for b-jet





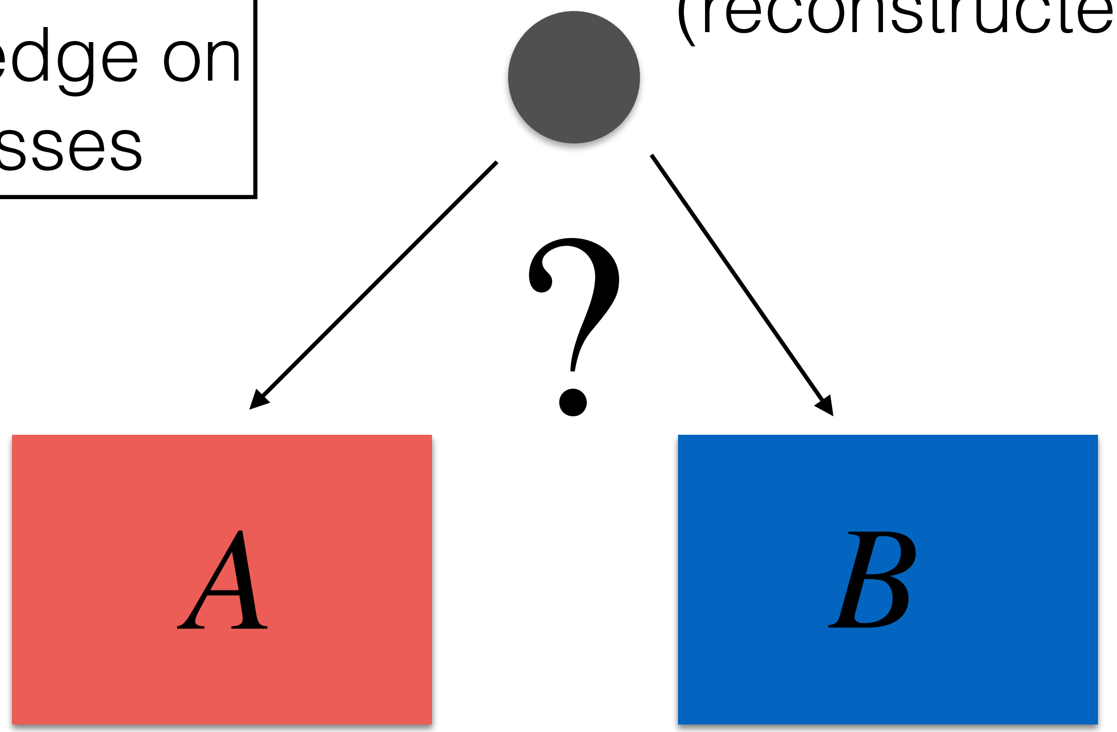
- Different mother particles

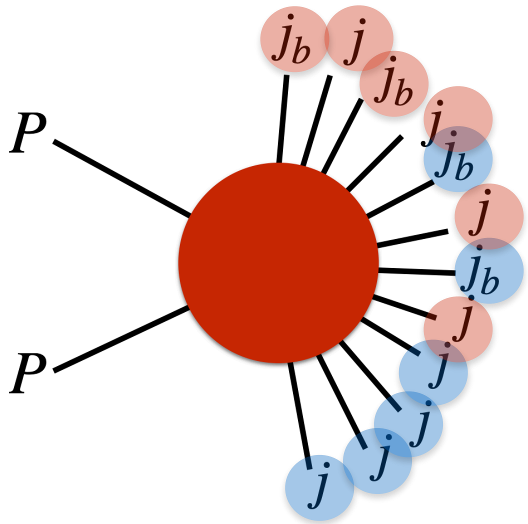
$$pp \to ZH \to \{j, j\} \cup \{(W \to jj), (W^* \to jj)\}$$

- Right answer is $(n_A, n_B) = (2,4)$

$2^6 = 64$ cases,
no prior knowledge on
A and B masses

(reconstructed) particle





- Complicate situation (12 jets)

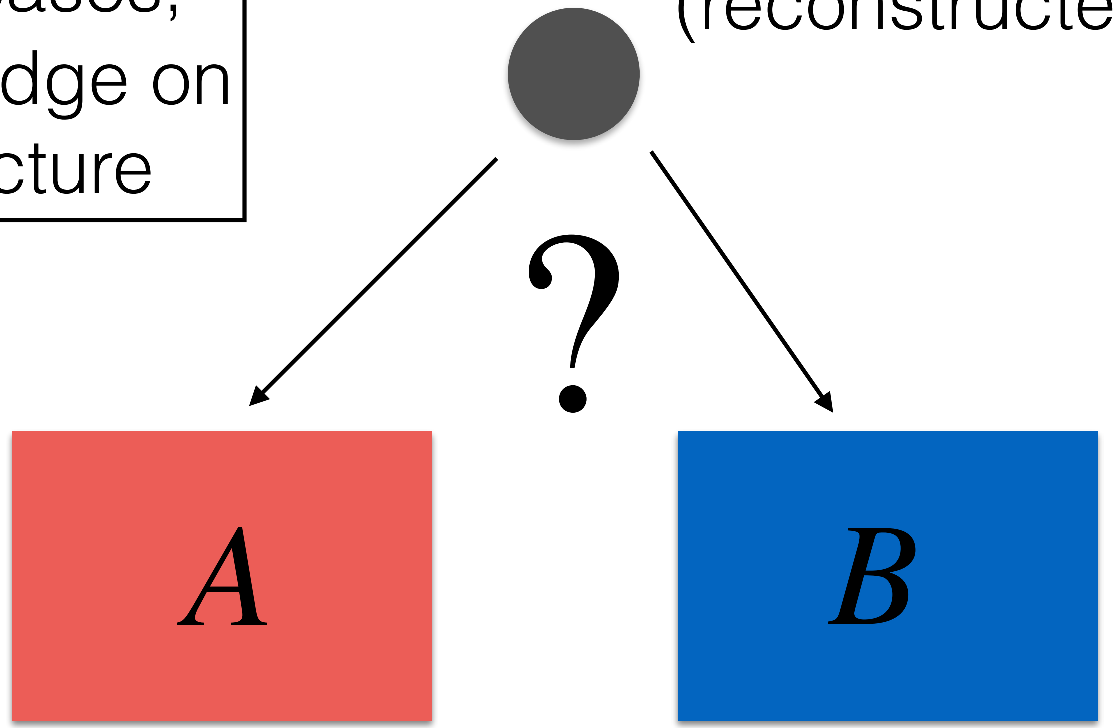
$$pp \to o\tilde{o} \to \{t, \bar{t}\} \cup \{t, \bar{t}\}$$

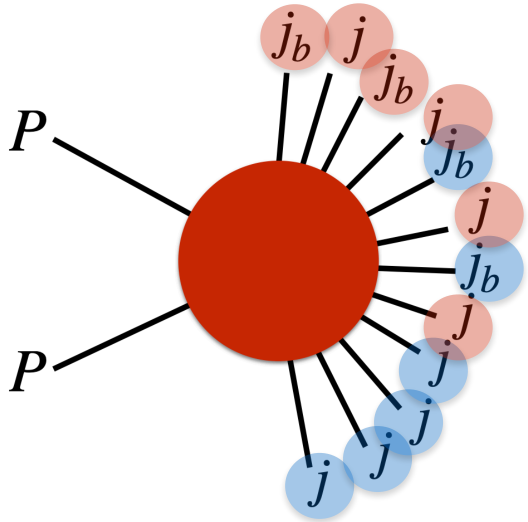
$$o \to t\bar{t} \to \{j_b, (W \to jj)\} \cup \{j_b, (W \to jj)\}$$

$$\tilde{o} \to t\bar{t} \to \{j_b, (W \to jj)\} \cup \{j_b, (W \to jj)\}$$

$\boxed{\text{$2^{12} = 4096$ cases, no prior knowledge on a decay-structure}}$

(reconstructed) particle





- Complicate situation (12 jets)

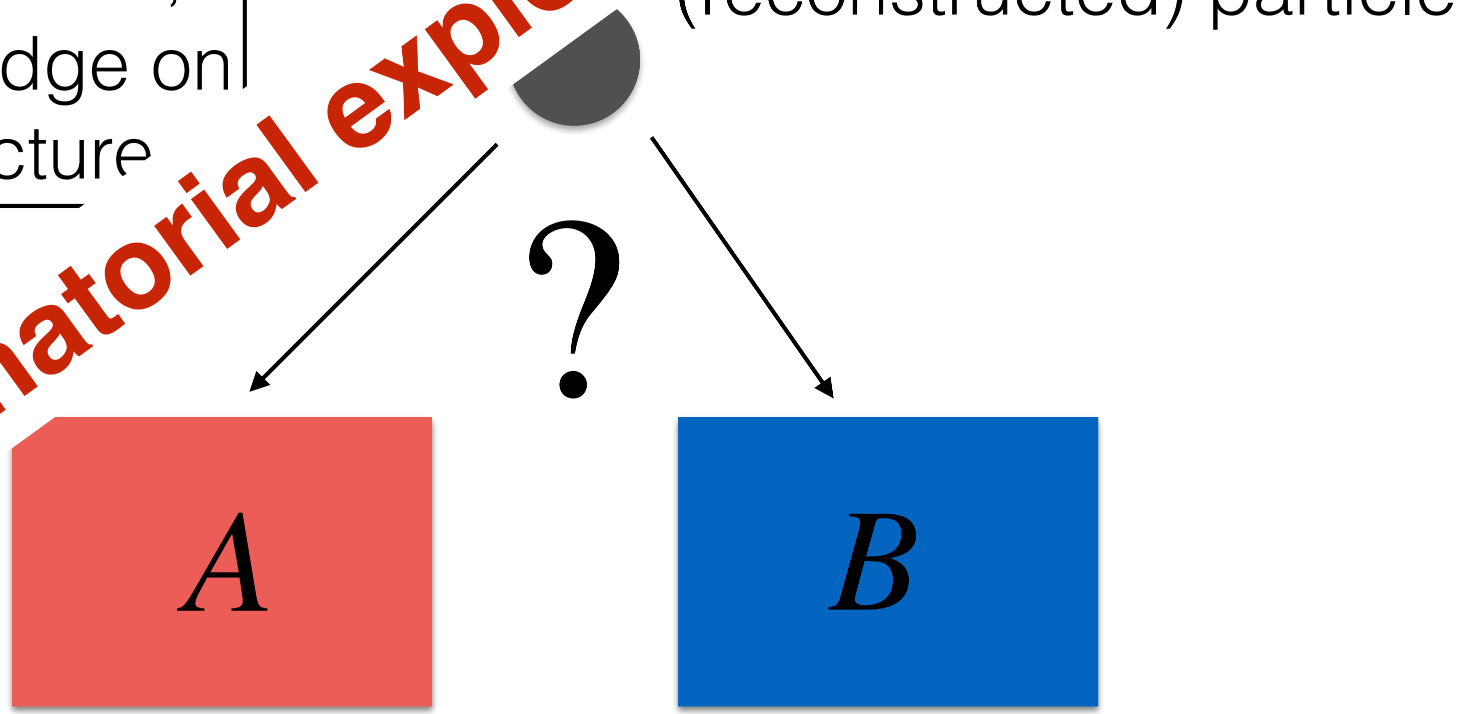
$pp \to o\tilde{o} \to \{t, \bar{t}\} \cup \{t, \bar{t}\}$

$o \to t\bar{t} \to \{j_b, (W \to jj)\}$ … $W \to jj)\}$

$\tilde{o} \to t\bar{t} \to \{j_b, (W$ … $\{j_b, (W \to jj)\}$

Combinatorial explosion will occur !

$2^{12} = 4096$ cases,
no prior knowledge on
a decay-structure



An algorithm ?

- With the only assumption of $2 \to (2 \to n)$ process

  - No special treatment on any flavor-tagged particle

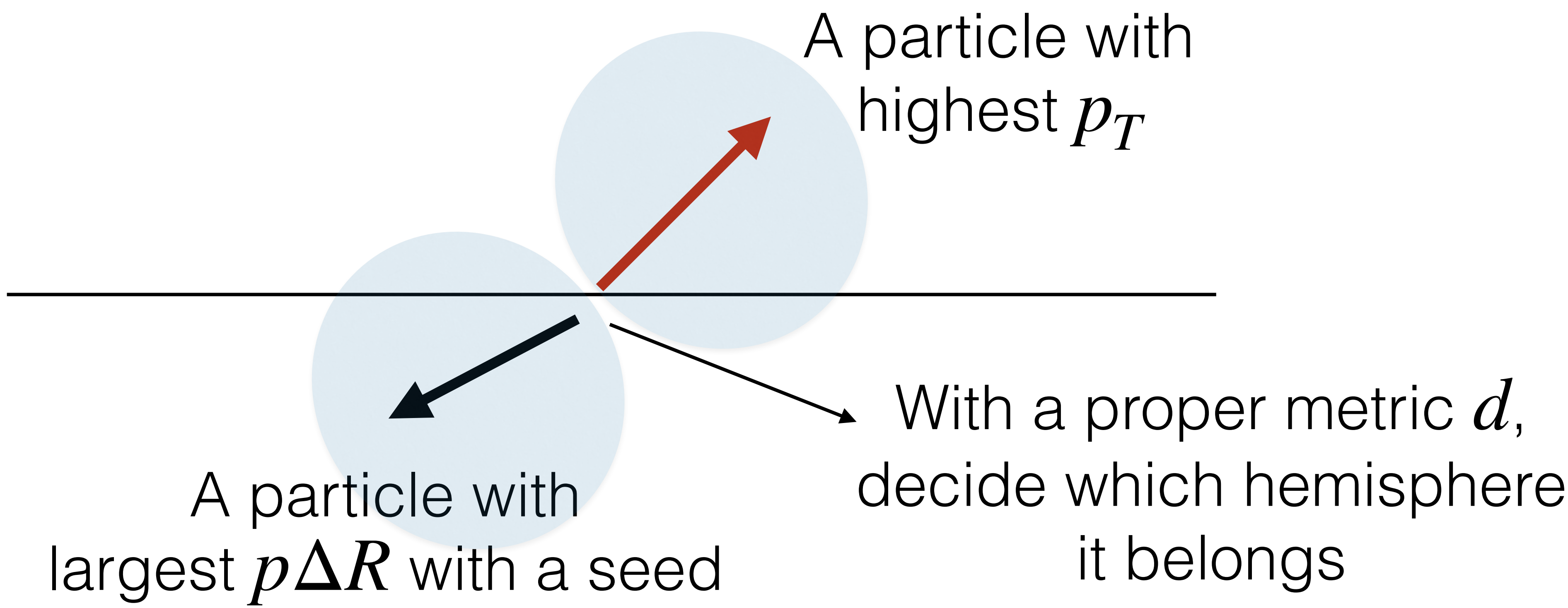
  - No assumption on M_A and M_B

  - No assumption on any decaying structure

- What could be a good guide line ?

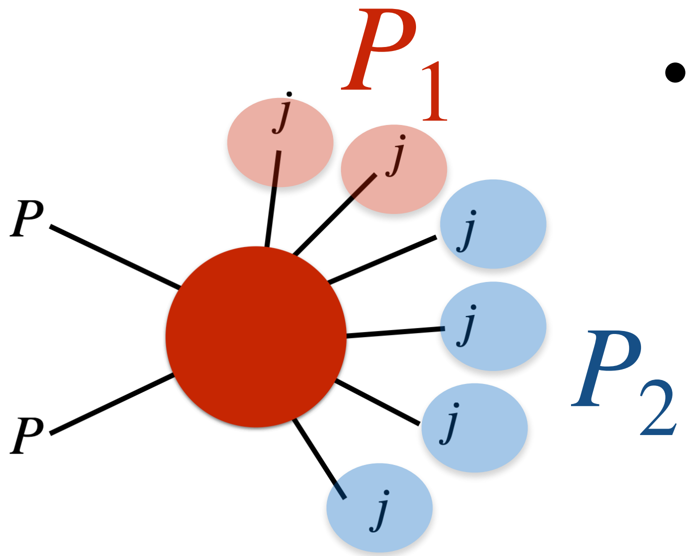
A Classic algorithm

- Hemisphere method: a seed-based method (iterative and converge)



CMS hemisphere TDR,
Shigeki Matsumoto, Mihoko M Nojiri, and Daisuke Nomura (2006)

Non-geometric algorithm

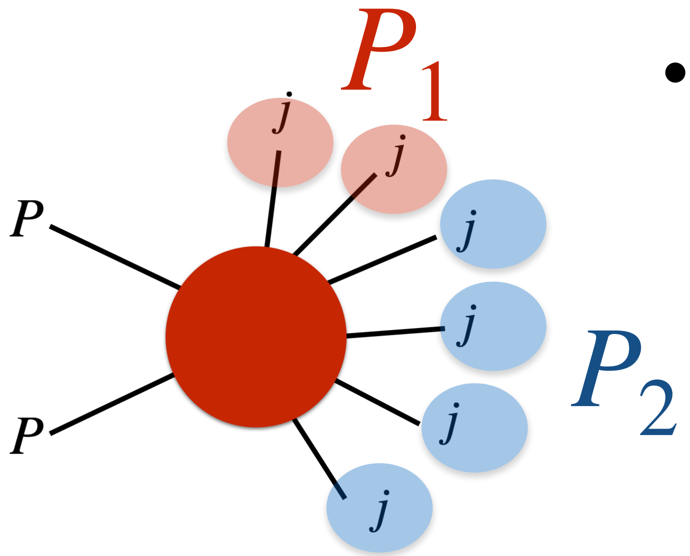


- For each assignment, calculate invariant mass

$$(M_A^2, M_B^2) = (P_1^2, P_2^2)$$

- Try to **minimize** the mass difference $H = (M_A^2 - M_B^2)^2$

Non-geometric algorithm

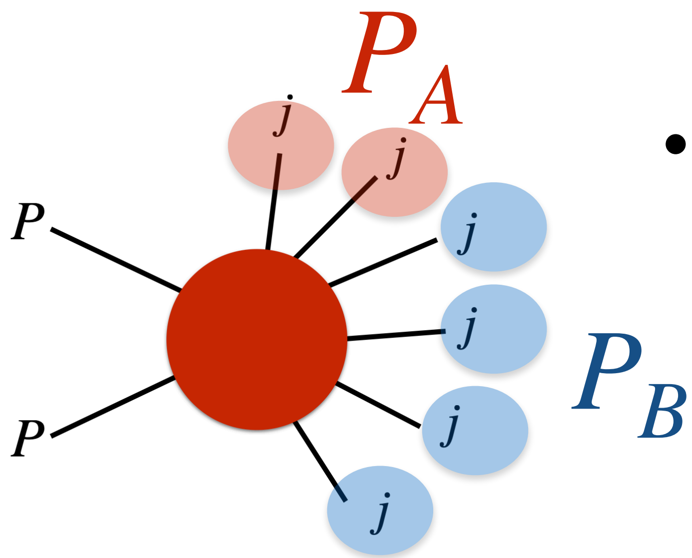


- For each assignment, calculate invariant mass

$$(M_A^2, M_B^2) = (P_1^2, P_2^2)$$

- Try to **minimize** the mass difference $H = (M_A^2 - M_B^2)^2$

- **How can we deal with the case of $M_A \neq M_B$?**



- Try to **minimize** the mass difference

$$H = (M_A^2 - M_B^2)^2$$

- **How can we deal with the case of $M_A \neq M_B$?**

  + (even with $M_A = M_B$) we need to handle
  1) off-shell mass due to the width of A and B
  2) from smearing effects due to imperfect detectors

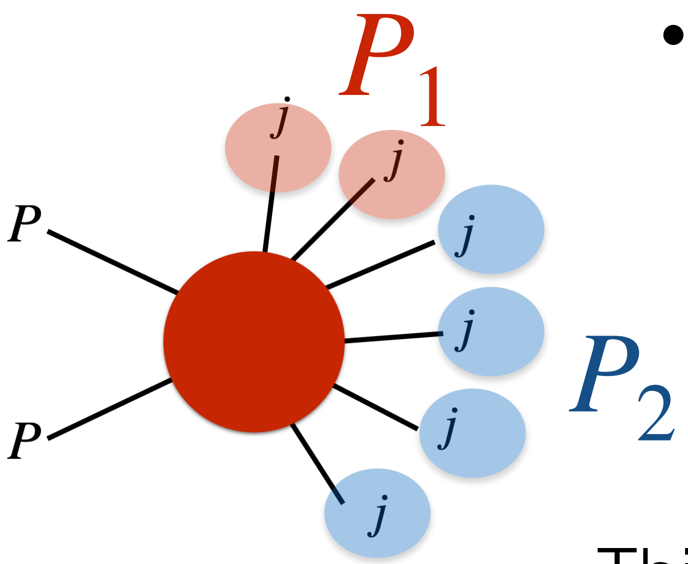
- One suggestion: Add a regularization term of $\lambda(P_1^2 + P_2^2)$

  (λ is a dimension full "**hyper-parameter**")

- $2 \rightarrow 2$ process: $\{p_i\} \rightarrow P_1 \cup P_2$
  Using a binary operation $x_i \in \{0,1\}$

  For p_i to be either in P_1 $(x_i = 1)$ or in P_2 $(x_i = 0)$

$$P_1 = \sum_i p_i x_i \,, \quad P_2 = \sum_i p_i \left(1 - x_i\right)$$



- Try to **minimize**

$$H = (P_1^2 - P_2^2)^2 + \lambda (P_1^2 + P_2^2)$$

for each "assignment" ?!

- This problem now becomes well-known...

Minimization using Ising model

- If we replace $x_i \to \dfrac{1+s_i}{2}$ with $s_i \in \{+1, -1\}$

$$H = \left(P_1^2 - P_2^2\right)^2 \to H + \lambda\left(P_1^2 + P_2^2\right)$$

$$= \sum_{i,j}\left(C_{ij} + 2\lambda S_{ij}\right)s_i s_j + \sum_i\left(J_i - 2\lambda\sum_j S_{ij}\right)s_i$$

- To maintain the importance of original H,

we take $\lambda = \dfrac{\min\left(C_{ij}\right)}{\max\left(S_{ij}\right)}$

But our "mindless"

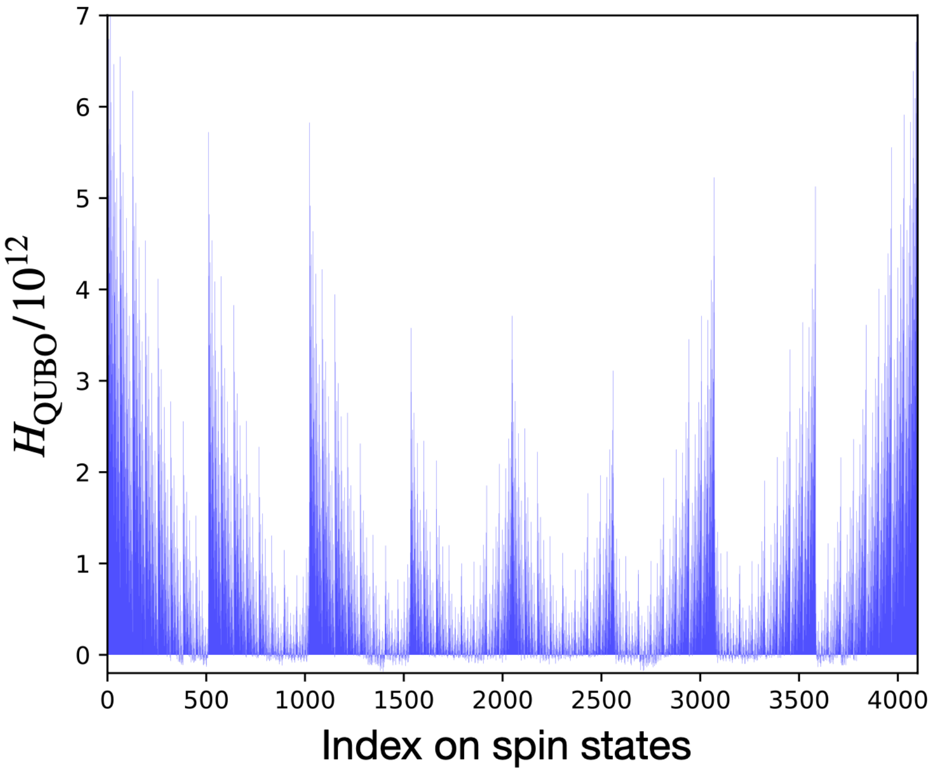
=**minimally assumed** Collider example

is **not so easy**

for a classical SM

Combinatorial complexity arises (for a random Ising model)

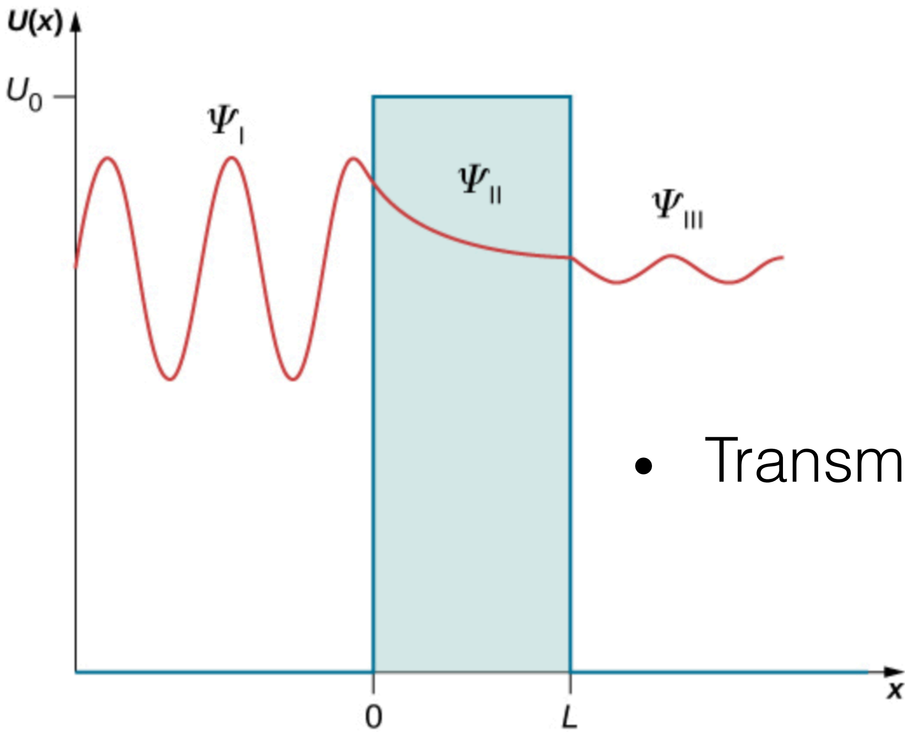
Landscape of energy distribution



$$\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \to \uparrow\uparrow\uparrow\uparrow\uparrow\downarrow \to \cdots \to \downarrow\downarrow\downarrow\downarrow\downarrow\downarrow \quad (n_{\mathrm{spin}} = 2^{12} = 4096)$$

SA cannot jump this random potential!

Quantum tunneling



- Transmission coefficient $T \propto e^{-2L\sqrt{2m(U_0 - E)}}$

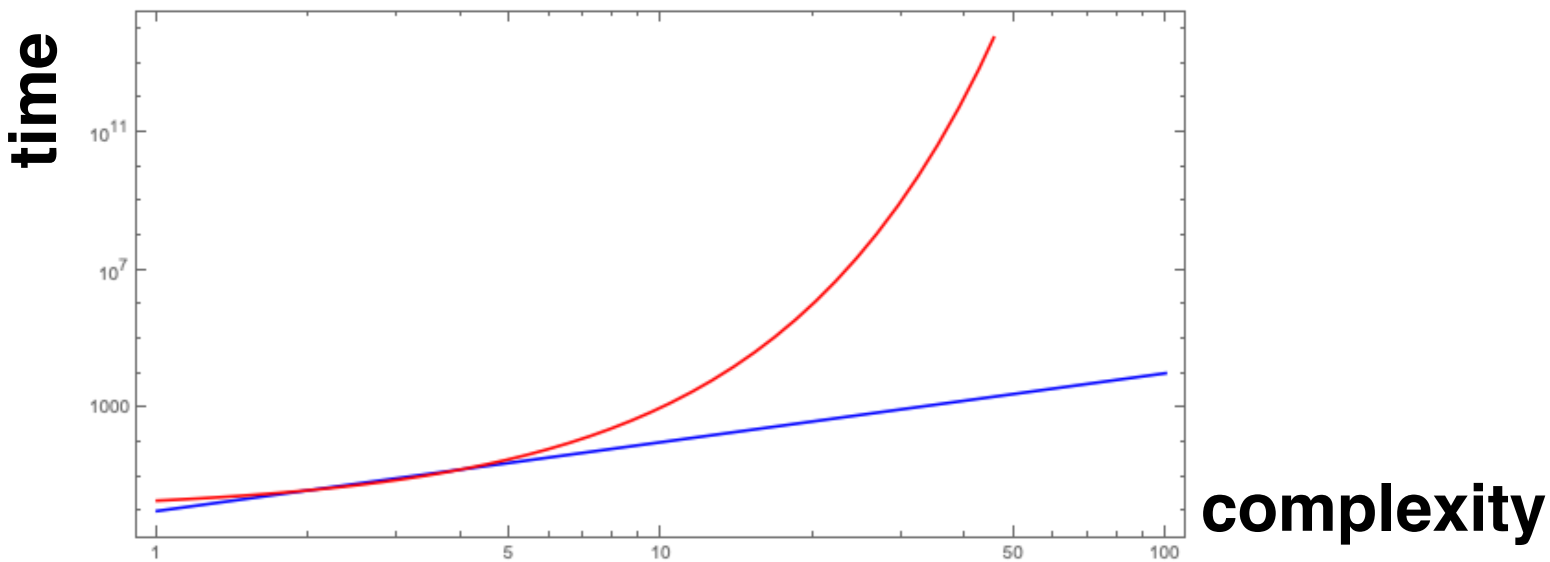
1) The effect of energy difference becomes mild

2) Effective for **shallow barrier** !

This would be
a Good example for D-Wave!

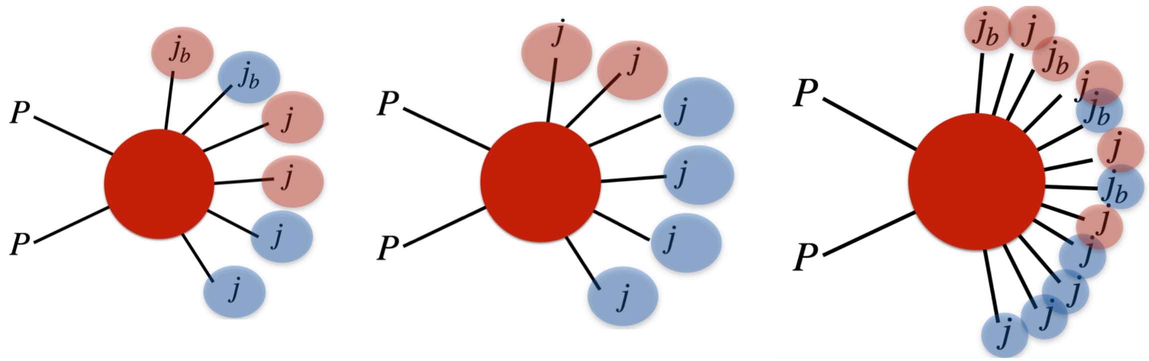
(small) Quantum advantage

- QA v.s. Brute-force scanning:
  The required time (mostly preparation time T_{QUBO})
  of QA machine: $T_{\mathrm{QUBO}} = \mathcal{O}(n^2)$
  The complete scanning with n input takes $\mathcal{O}(2^n)$



(big) Quantum advantage

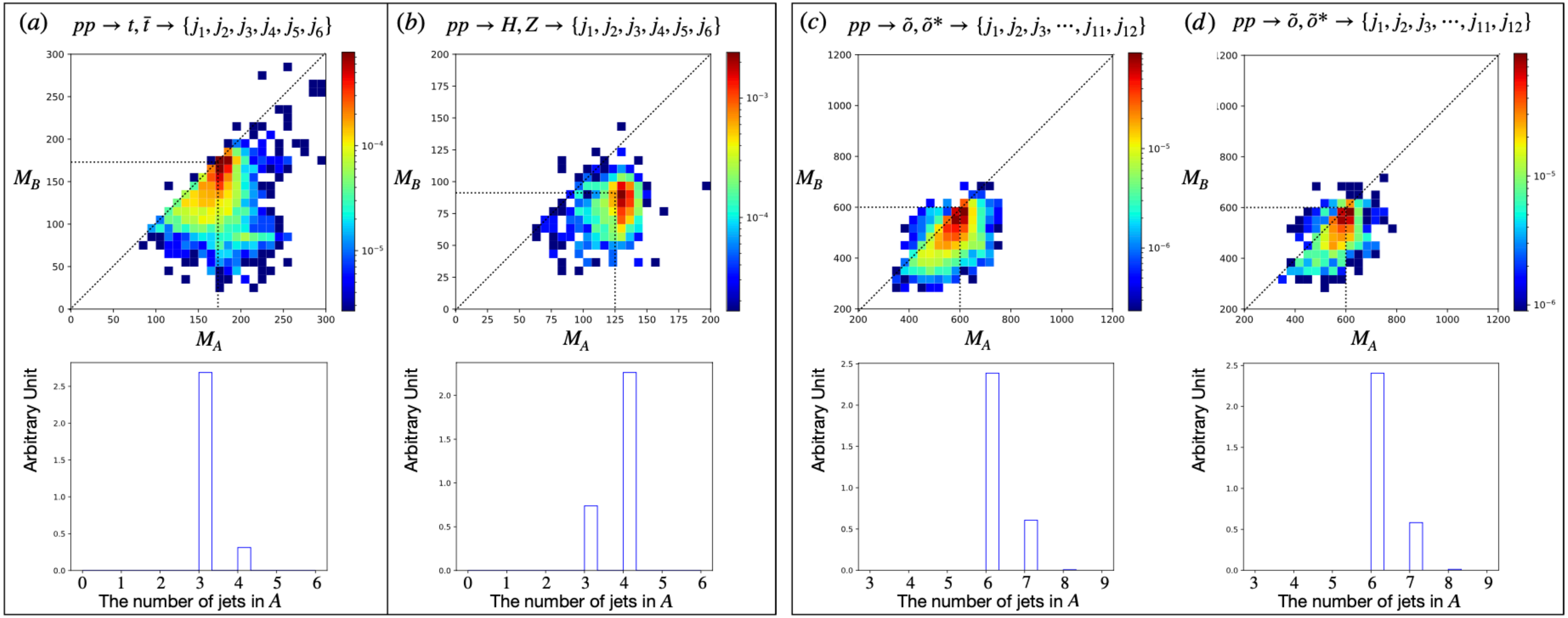
- QA v.s. SA



Process	$pp \to t\bar{t}$ $(2 \to 6)$	$pp \to HZ$ $(2 \to 6)$	$pp \to \tilde{o}\tilde{o}^*$ $(2 \to 12)$
Quantum annealing	100%	100%	74.3%
Simulated annealing	36.7%	45.7%	1%

Percentage to get a **global minimum energy state** (**does not guarantee** a true combinatorial assignment)

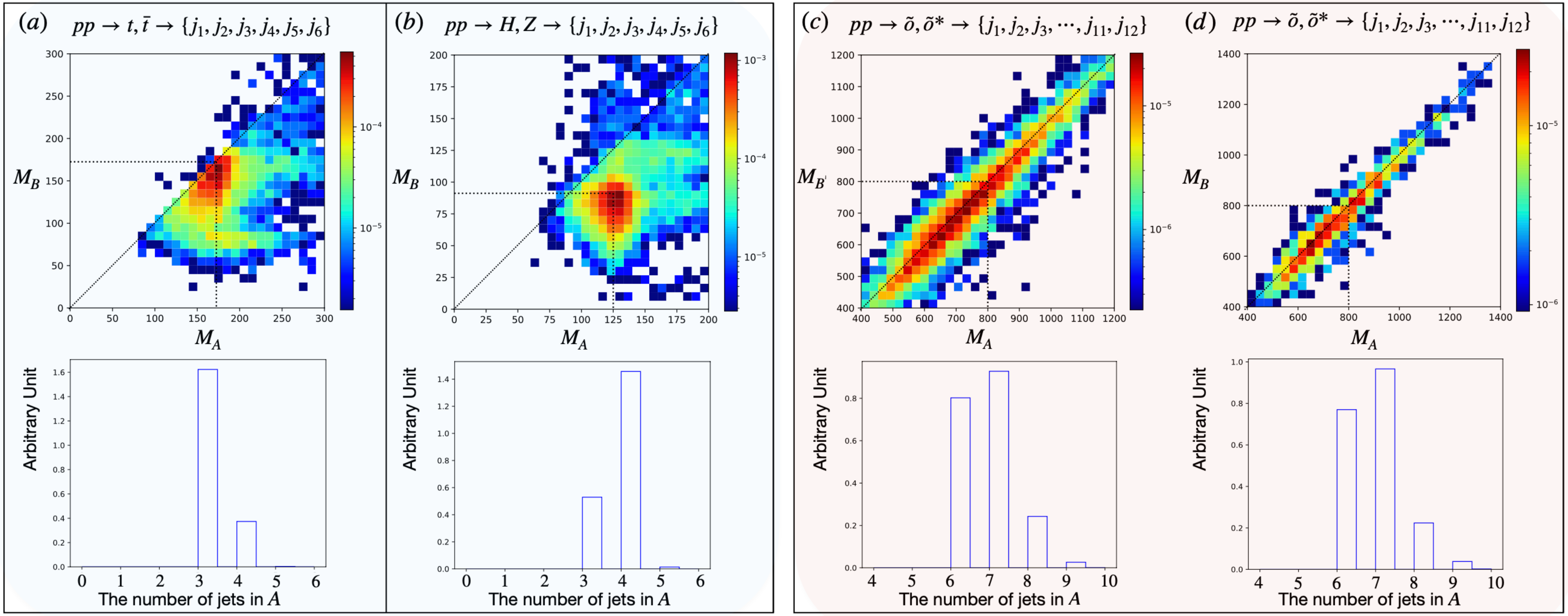
results



- Madgraph → Pythia (ISR/FSR/MPI turned off) → Delphes

*** a to c: brute force scanning for H_{QUBO} to check the fidelity of our algorithm**
d is from D-Wave computer (expensive...)

results



- Madgraph → Pythia (**ISR/FSR/MPI turned ON**) → Delphes

(As we give a priority to hardest jets,
effect of hard ISR is emerging for hard scale,
here $2m_{\tilde{o}} = 1.2\text{TeV}$)

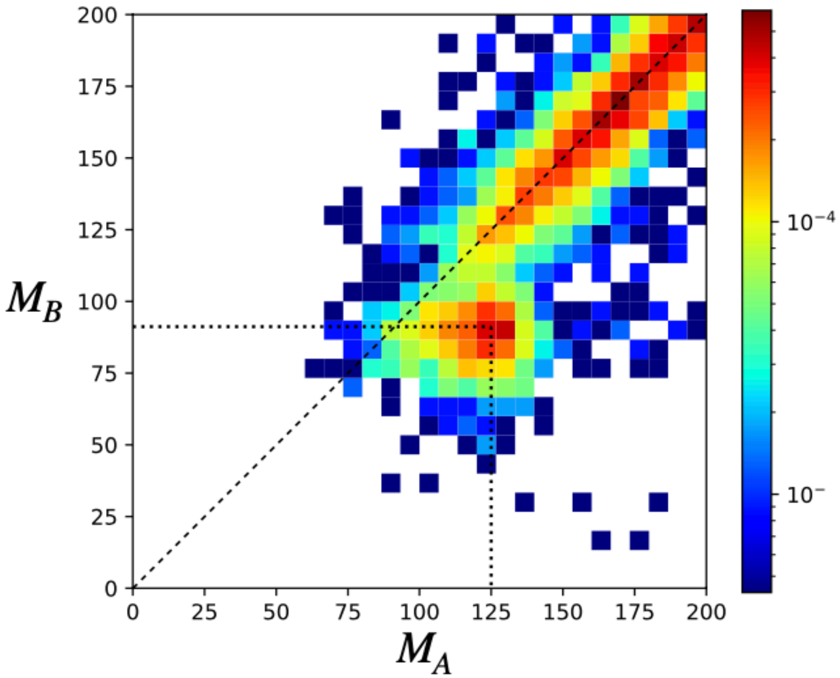
Effect of additional constraints

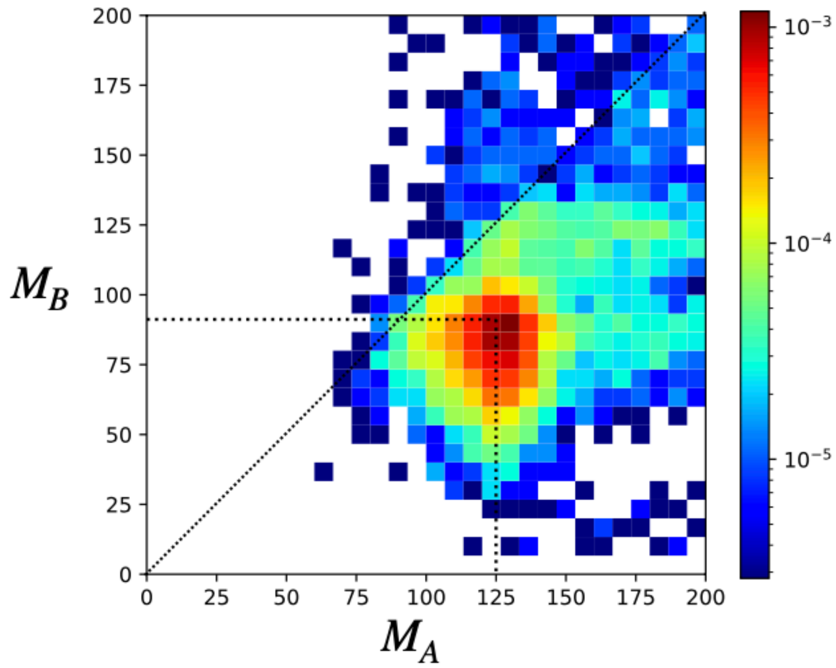
$$H = \left(P_1^2 - P_2^2\right)^2 \to H + \lambda\left(P_1^2 + P_2^2\right)$$

- For different mother particle cases: $pp \to HZ$

$$H = \left(P_1^2 - P_2^2\right)^2$$

$$H \to H + \lambda\left(P_1^2 + P_2^2\right)$$



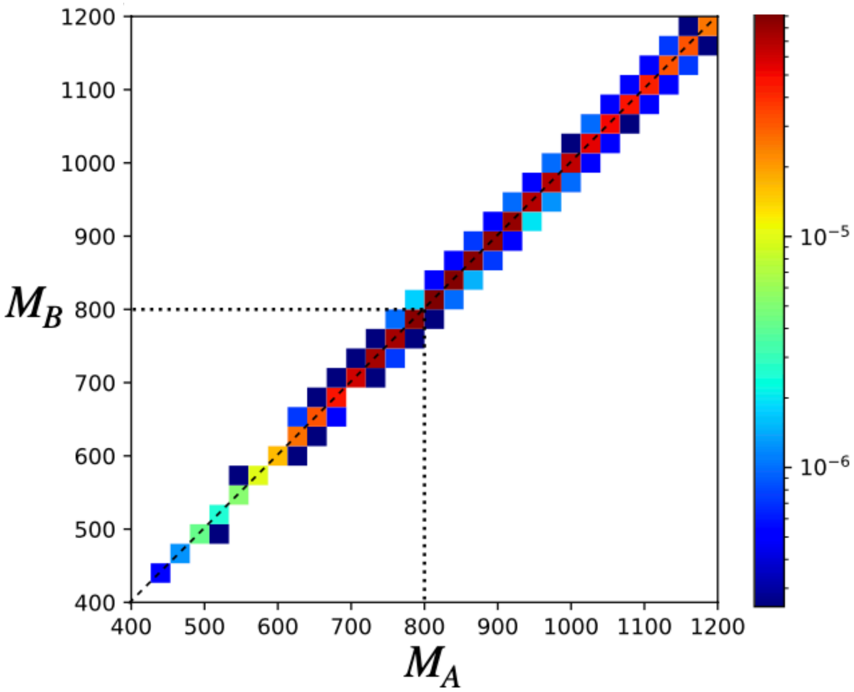


Effect of additional constraints

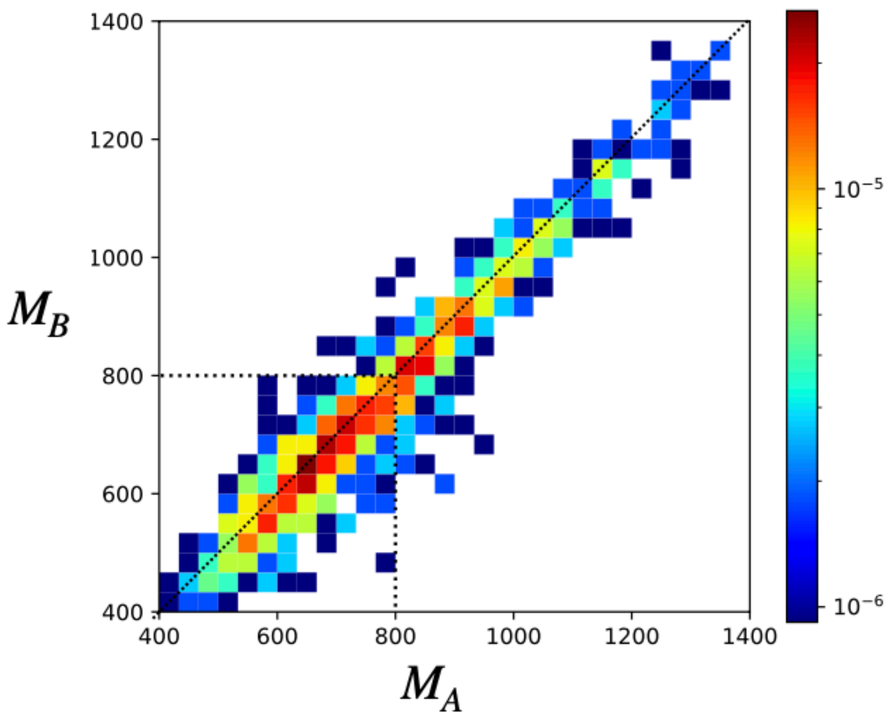
$$H = \left(P_1^2 - P_2^2\right)^2 \to H + \lambda\left(P_1^2 + P_2^2\right)$$

- For smearing effects : $pp \to \tilde{o}\tilde{o} \to t\bar{t}t\bar{t}$

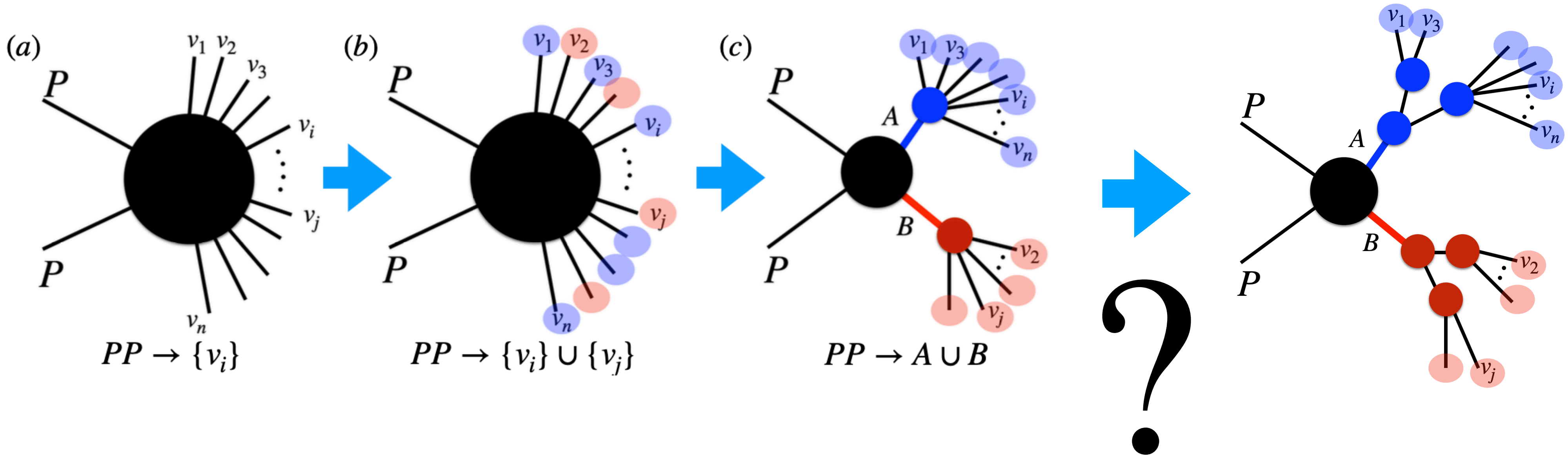
$$H = \left(P_1^2 - P_2^2\right)^2$$



$$H \to H + \lambda\left(P_1^2 + P_2^2\right)$$



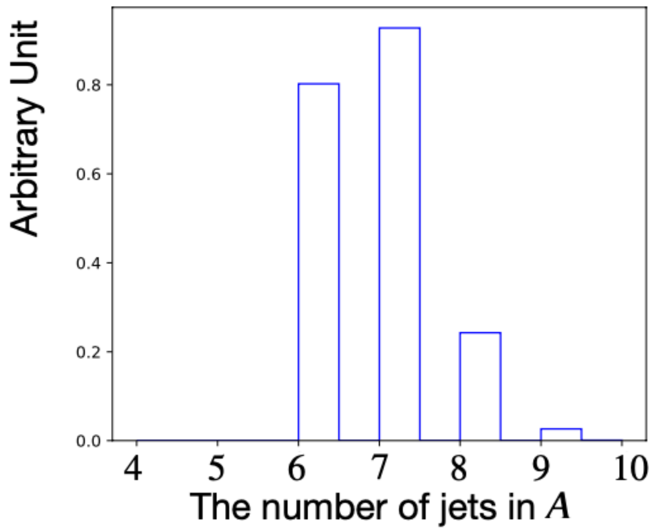
Sequential algorithm



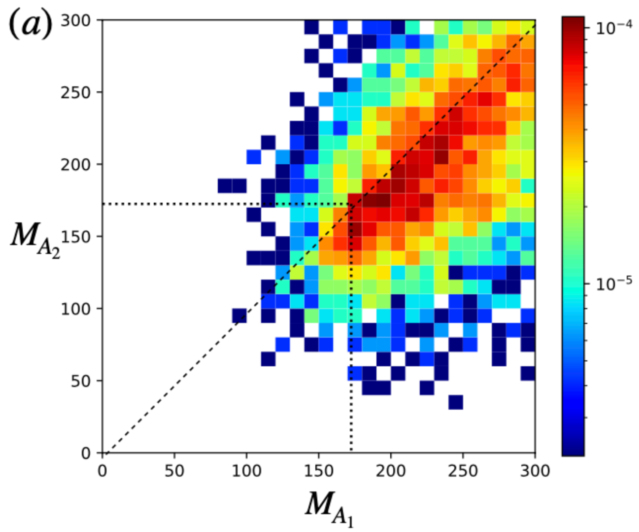
$$H^{(A)}_{\rm QUBO} = \sum_{ij=1}^{\ell} J'^{\alpha}_{ij} s^{\alpha}_i s^{\alpha}_j + \sum_{i=1}^{\ell} h'^{\alpha}_i s^{\alpha}_i,$$

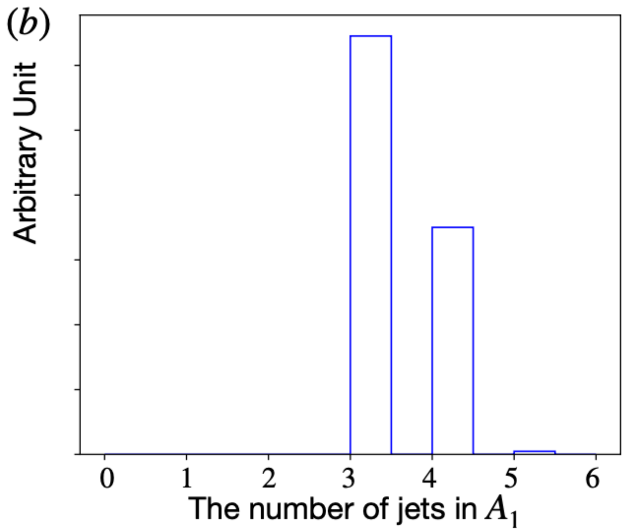
$$H^{(B)}_{\rm QUBO} = \sum_{ij=1}^{m} J'^{\beta}_{ij} s^{\beta}_i s^{\beta}_j + \sum_{i=1}^{m} h'^{\beta}_i s^{\beta}_i,$$

- For 12 hard-jets production, it would be worthy if we can check whether this is four-tops events or not !



$$pp \to A, B$$





$$A \to A_1, A_2$$

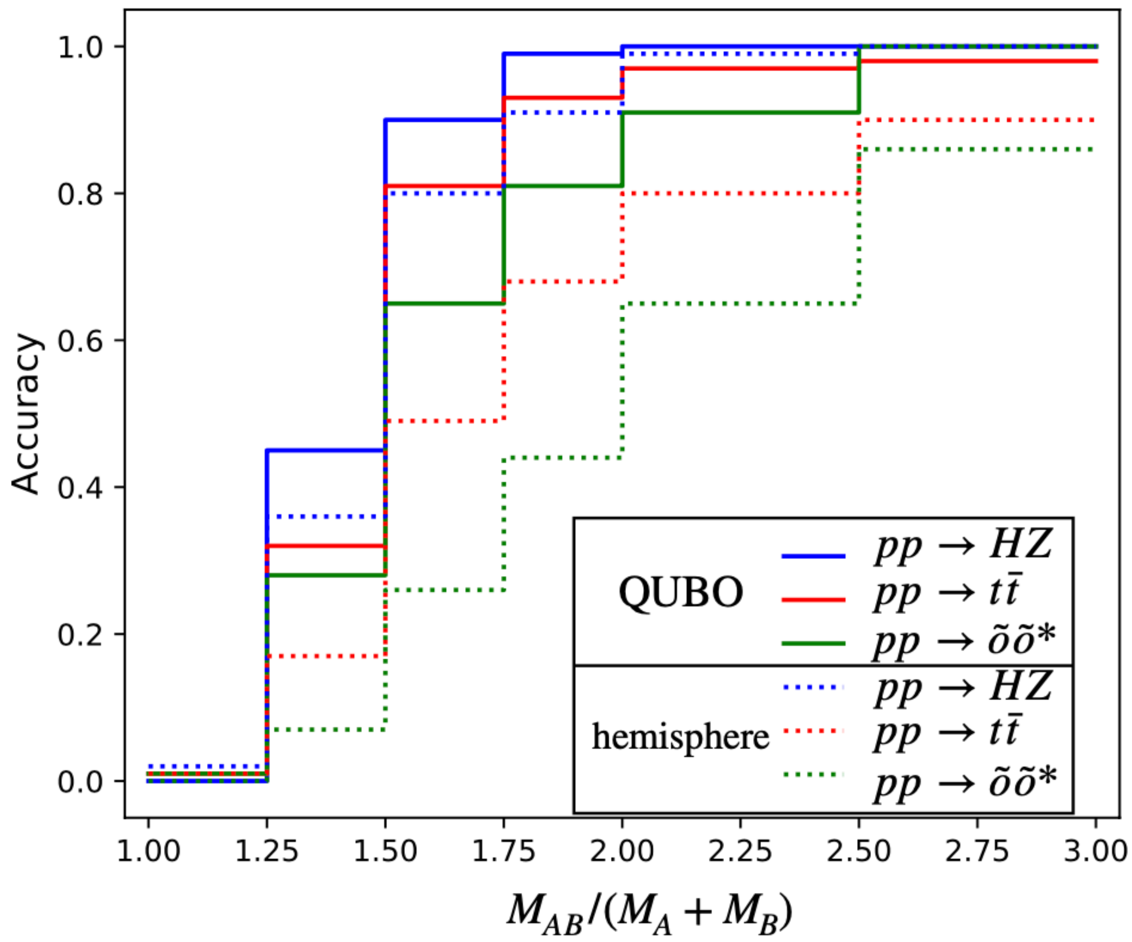
- We can "guess" that $A_i = t(\bar{t})$ as their **mass** and **number of children** are identical to the case of a top-quark.

Bench mark?

- There are not many studies on identifying event-topology. (as far as I have searched... if I missed, plz let me know)
- Hemisphere method: **seed-based** algorithm (**our algorithm is seedless** one)

Process		$pp \to t\bar{t}$ Eq. (7a)	$pp \to HZ$ Eq. (7b)	$pp \to \tilde{o}\tilde{o}^*$ Eq. (7c)
Algorithm	QUBO	47.3%	89.5%	15.1%
	Hemisphere	33.6%	86.2%	5.84%

(Parton-level analysis with detector cuts)



- Performance of an algorithm based on "**seed**" becomes weak when particles are **not boosted enough to develop structures.**

- Lorentz boost factor $\gamma_A = \dfrac{E_A}{M_A} = \dfrac{M_{AB}}{2M_A}$ (for A=B case)

Current limits for QA

- Number of couplers is limited
  - **spin-chain** method to encode a hamiltonian (connections)

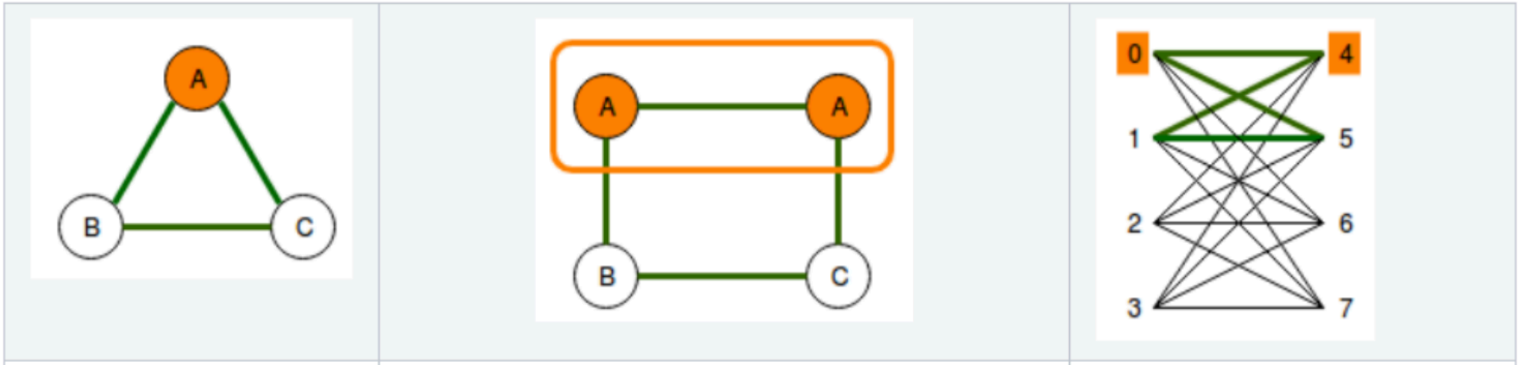
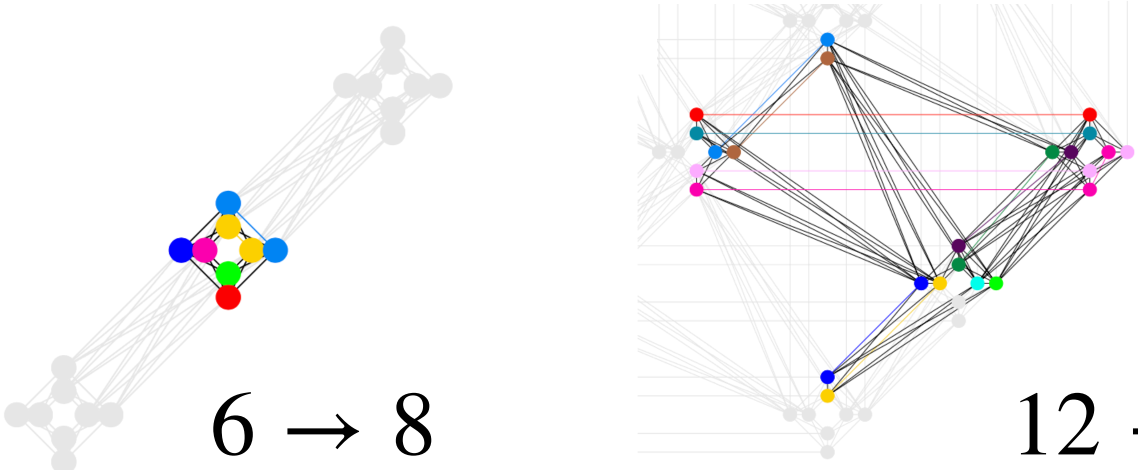


Fig 1: Abstract graph | Fig 2: Intermediate representation of the graph | Fig 3: Embed graph onto QPU

- Number of required qubits for our problem



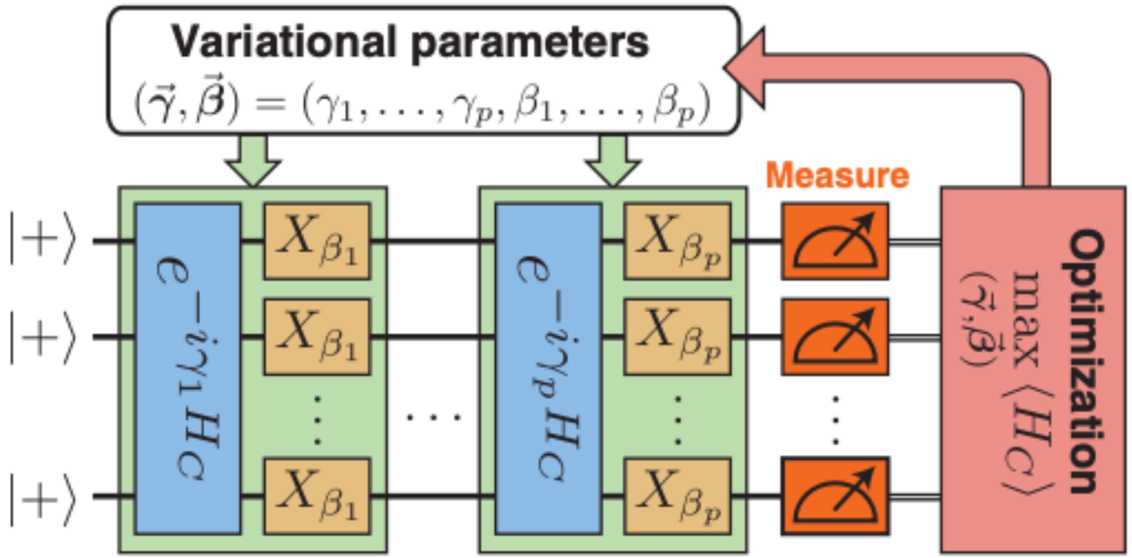
$6 \rightarrow 8$ $12 \rightarrow 24$

Conclusion

- I presented a **simple quantum annealing method for clustering** reconstructed particles.
  - We are interested in expanding this work including Missing particles. (KC and me)

- Gate-based QC can be used via a variational algorithm.

- We can use Gate-based QC for QUBO, which KC is working on. Details about this, plz check Dr. Bae's talk afternoon



- As a **desperate** seeker, we have tried to take advantages of new computing methods, ML, QC, QML.

- In this talk, I presented a **bottom-up** collider algorithm to identify a new physics from a signal (if we can have)

- There could be many good examples to demonstrate **Quantum Advantage** in the field of HEP:

  - check Jae-hyun's talk afternoon

- QC can be the next "Galieo's telescope"

- At least with QC and QI, we can "teach" QM-1 to students in a very interesting and "modern" way!

Thank you for

- Coming to this far away,
- Sharing your ideas and visions,
- "Enjoying" night discussions and drinks.