Computational tools for the Future Colliders

Myeonghun Park

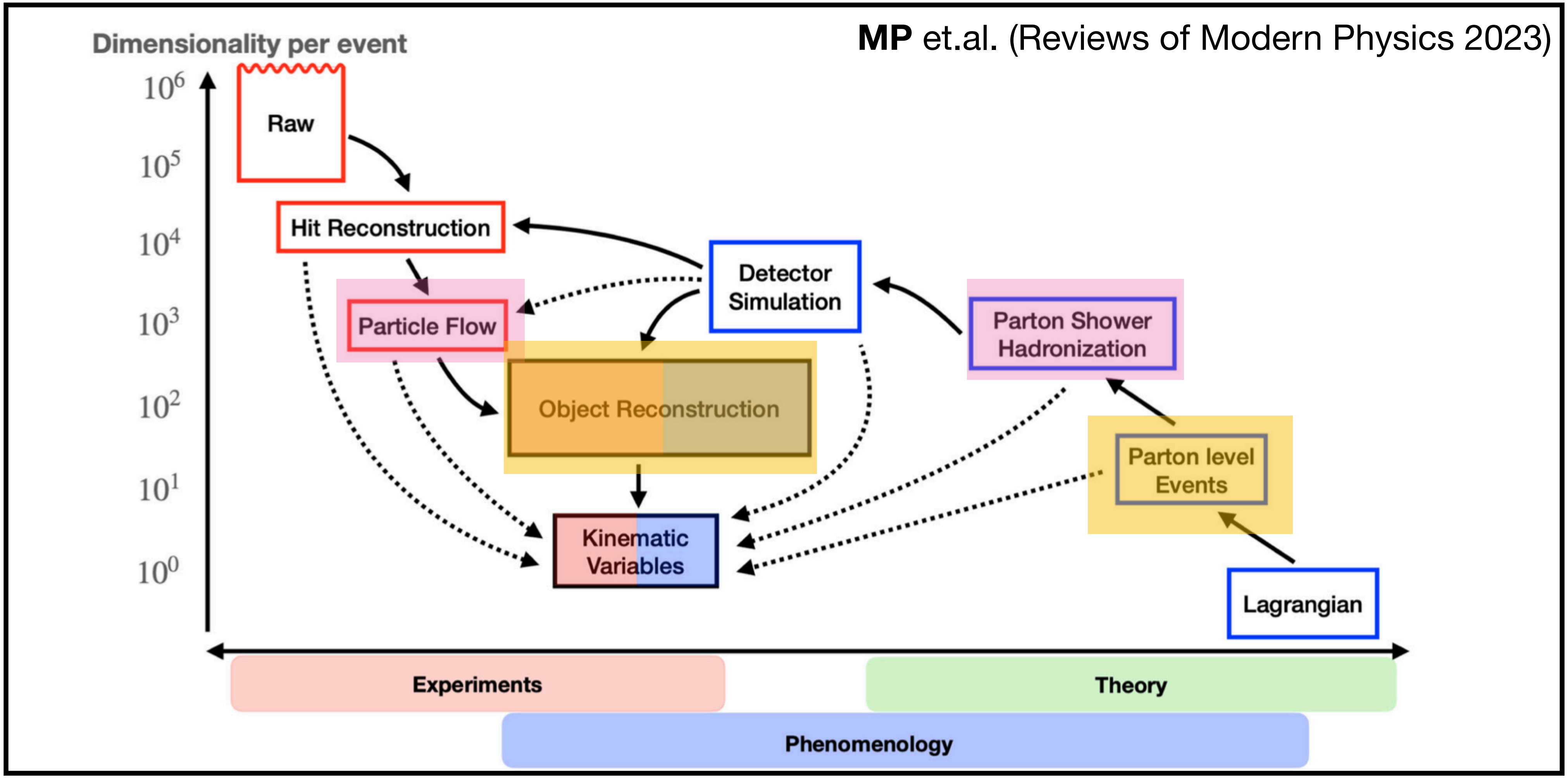
Seoul National University of Science and Technology (SeoulTech)

The 2025 International Workshop on the High Energy Circular Electron Positron Collider

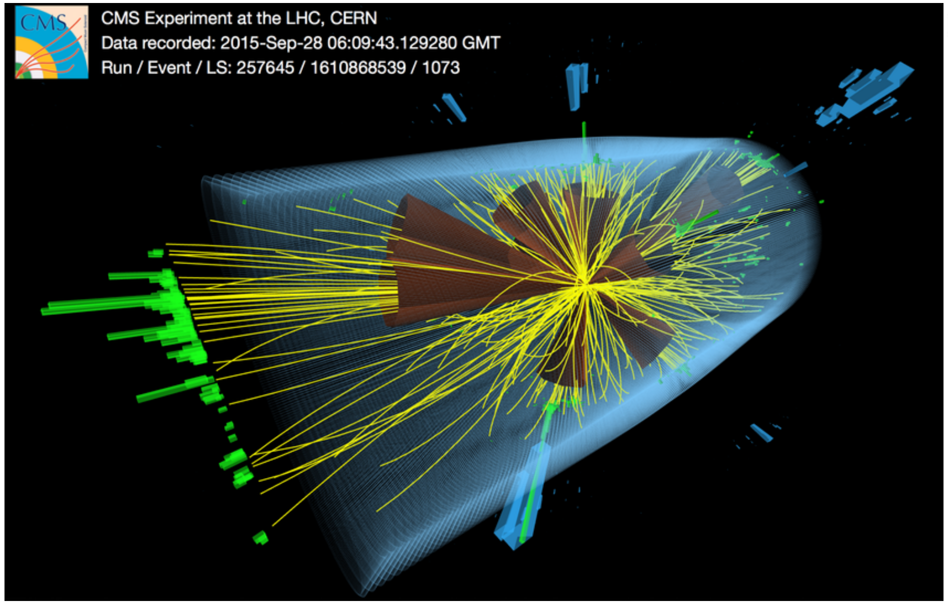
Nov 08, 2025. Guangzhou CHINA

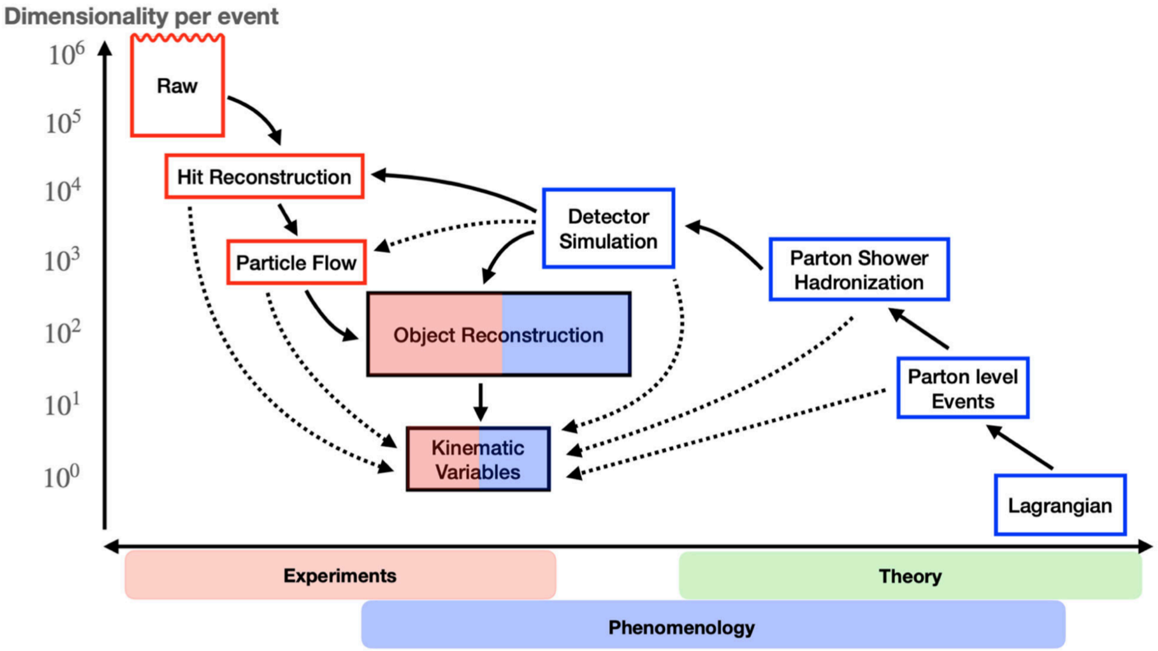
Understanding Physics using Colliders

- Recent advancements in Machine Learning (**ML**) allow us to fully exploit **all available information**, including raw-level data such as detector responses.



High multiplicity in High Energy collisions



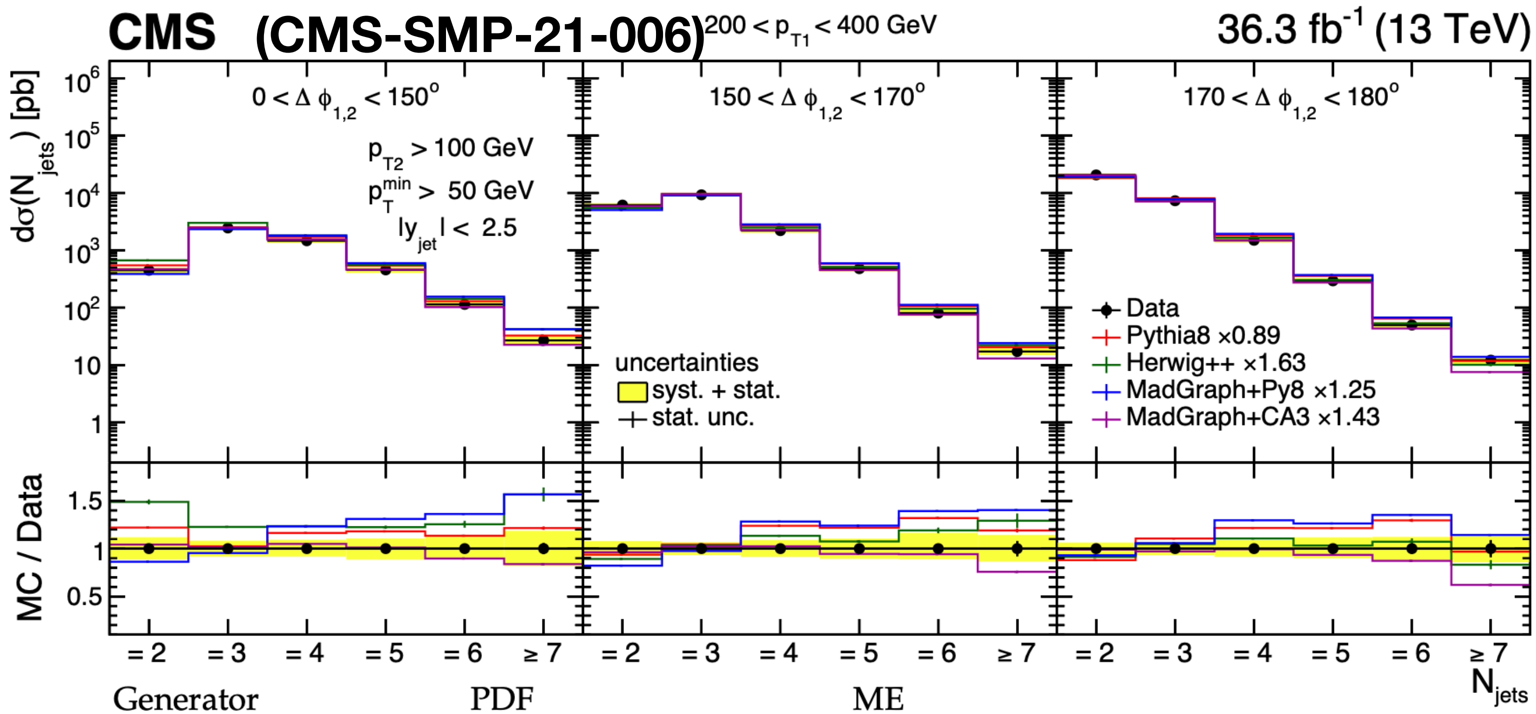


- Precise expectations from a simulation is required
  (1) Monte Carlo integration using ML

- Data-driven approaches without theoretical bias to fully exploit all possibilities.
  (2) Event reconstruction using Quantum Algorithm

Importance of Precision

- For Machine Learning which requires "**training**", **the big amount of well-understood data is necessary.**



Generator	PDF	ME
PYTHIA8 [23]	NNPDF 2.3 (LO) [25]	LO $2 \to 2$
MADGRAPH+PY8 [4]	NNPDF 2.3 (LO) [25]	LO $2 \to 2,3,4$
MADGRAPH+CA3 [4]	PB-TMD set 2 (NLO) [1]	LO $2 \to 2,3,4$
HERWIG++ [26]	CTEQ6L1 (LO) [27]	LO $2 \to 2$

- At large jet multiplicities, all generators tend to deviate from the data, indicating that our current simulations still lack sufficient precision in this complex region

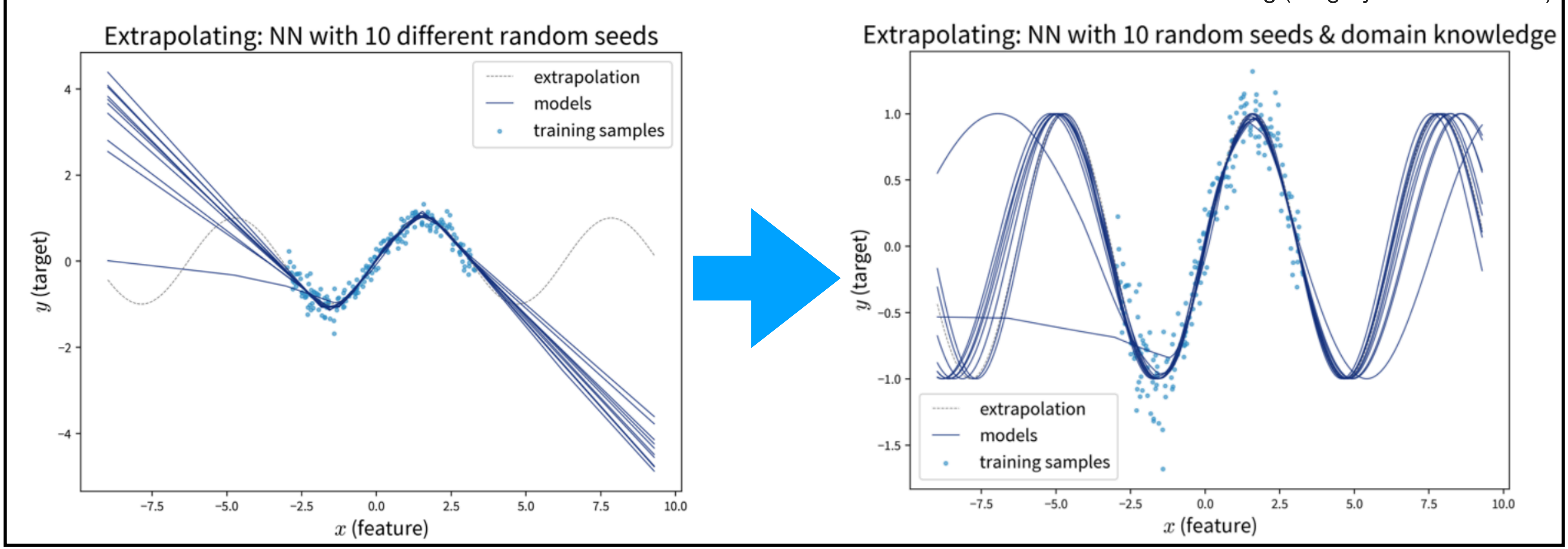
Importance of Theory

- We need **HUGE "training data"** to feed the **"data hungry" Neural Net**.

- One can dream of "data-driven" machine learning.

  - We cannot guarantee the estimation out of Controlled samples.
    : **NO magic can do "Exploration".**
    : Domain knowledge is strongly required.

A Practitioner's Guide to Machine Learning (Blog by Franziska Horn)



(1) Monte Carlo Integration

- $$\sigma = \frac{1}{2s} \int \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \delta\left(k_1 + k_2 - \sum_{i=1}^{n} p_i \right) |M_{fi}|^2 \theta_{\text{cut}}(p1, \cdots, p_n)$$

- For an observable $O(p_1, \cdots, p_n)$, we need to calculate the differential distribution of

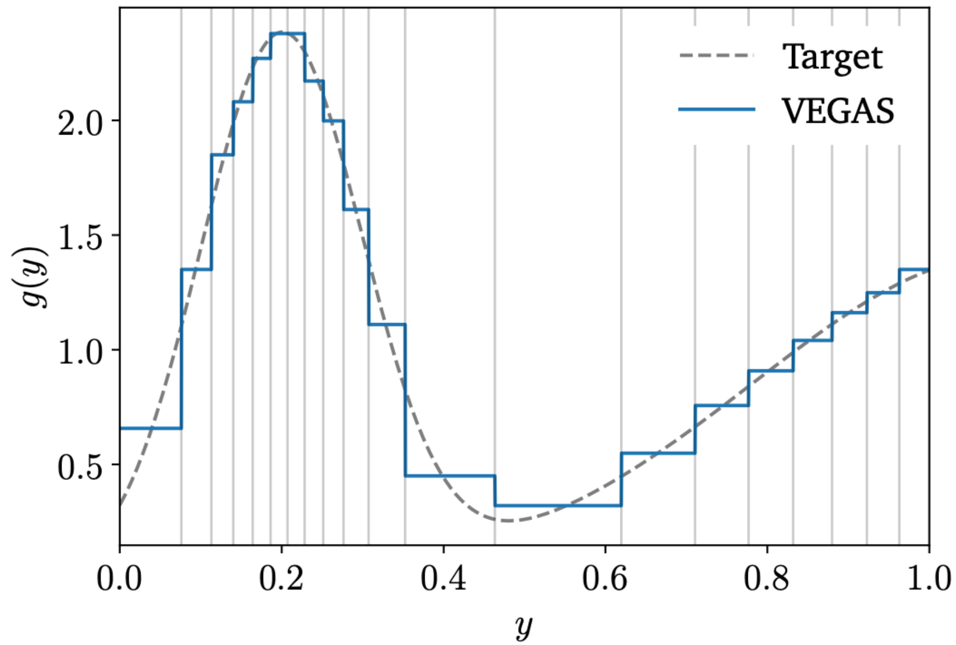
$$\frac{d\sigma}{dO} = \frac{1}{2s} \int \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \delta\left(k_1 + k_2 - \sum_{i=1}^{n} p_i \right) |M_{fi}|^2 \theta_{\text{cut}}(p1, \cdots, p_n) \delta(O - O(p1, \cdots, p_n))$$

- **Precise numerical Integration in high dimensional phase space that may contain nontrivial singularities**

Monte Carlo Integration: Domain vs Value Partitioning

- **Riemann-style: VEGAS**

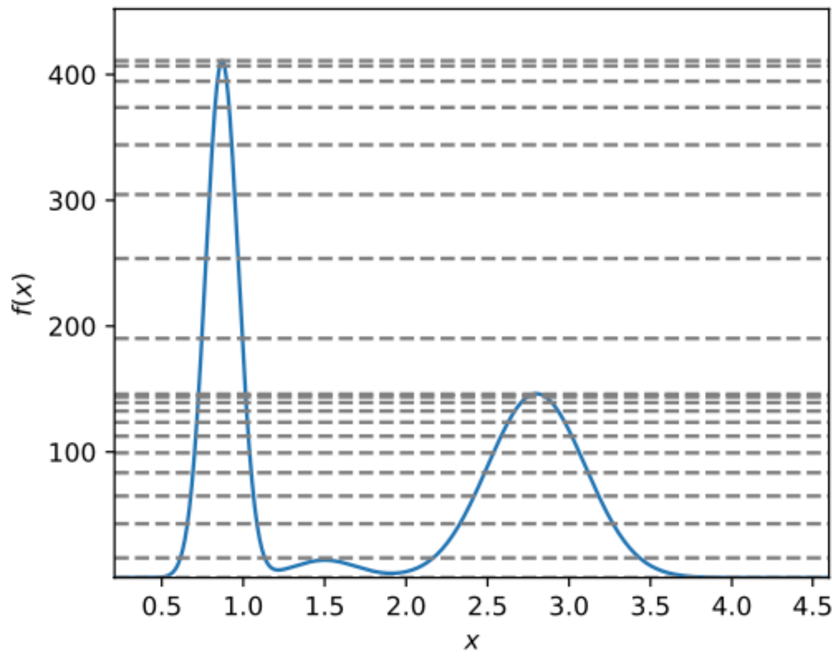
from MadNIS (Theo Heimel et.al. arXiv:2311.01548)



- **Stratified Sampling**: Divide domain **into sub-domains.**
  For example, if we divide the domain into N divisions, $\sigma \propto \frac{1}{N}$ instead of $\sigma \propto \frac{1}{\sqrt{N}}$

- **VEGAS**: Adaptive importance sampling
  - divides the domain along coordinates.

- **Lebesgue**-style

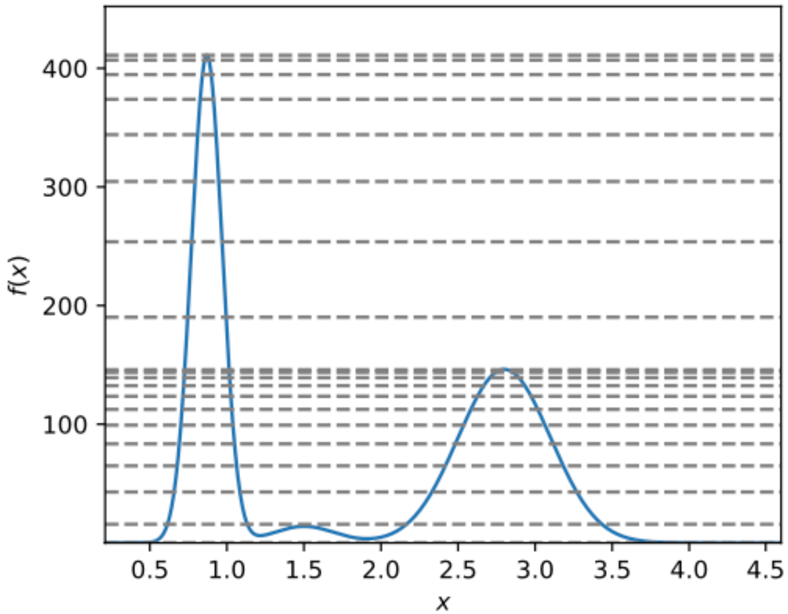


- **Alternative approach**: Integrate by function values instead of coordinates.

- **Lebesgue integration**: group by $f(x)$ rather than x.
  - More efficient when the integrand has a **singular or sharply peaked** structure.

  - A classical example: $f(x) = \begin{cases} 1 & \text{if } x \in P \\ 0 & \text{if } x \in Q \end{cases}$

Our approach: Lebesgue



- **Divide the space of integrand (classes)**

$$\Phi_j = \{\vec{x} \,|\, l_j < f(\vec{x}) \leq l_{j+1}\}$$

- The integral : $I_\Phi\left[f(\vec{x})\right] = \int_\Phi \mathrm{d}^d x\, f(\vec{x}) = \sum_{j=1}^{n} \int_{\Phi_j} \mathrm{d}^d x\, f(\vec{x}) = \sum_{j=1}^{n} V_{\Phi_j} \langle f \rangle_{\Phi_j}$

V_{Φ_j} : Volume of Φ_j .

- We recast the **problem of integration** → **classification problem**

MP. et al, Computational Physics Communication 319 (2026) 109907

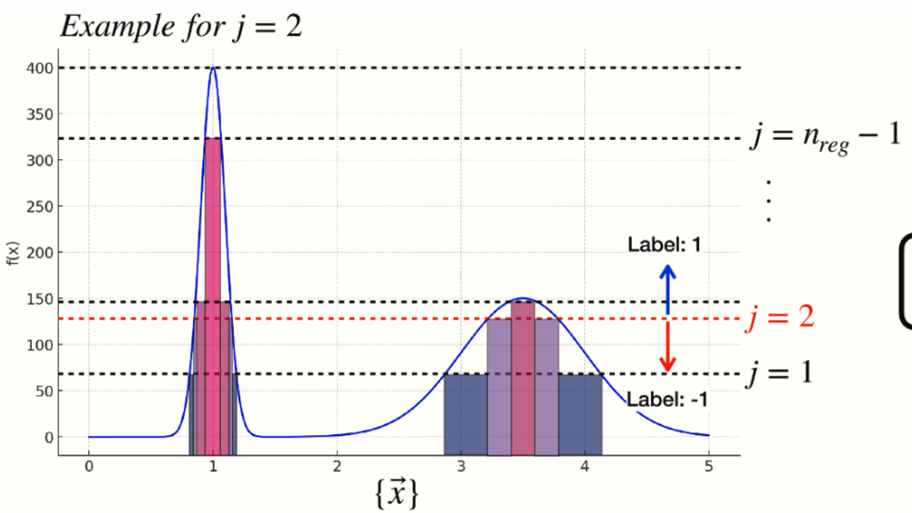
Monte Carlo with ML

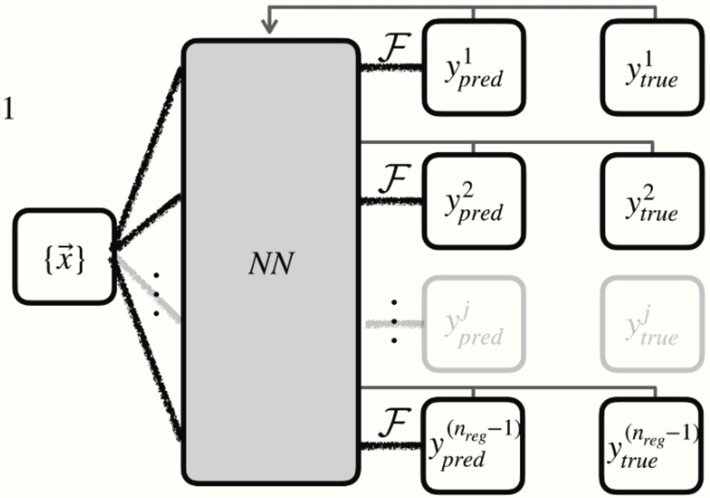
$$I_\Phi[f(x)] = \int_\Phi \mathrm{d}^d x\, f(x) = \sum_{j=1}^{n} \int_{\Phi_j} \mathrm{d}^d x\, f(x) = \sum_{j=1}^{n} V_{\Phi_j} \langle f \rangle_{\Phi_j}$$

- here, if we can "correctly" decide $x \in \Phi_j$, we can calculate

$$V_{\Phi_j} \simeq \frac{N_j}{N_{\text{total}}} V_{\text{total}}\,, \quad \langle f \rangle_{\Phi_j} \simeq \frac{1}{N_j} \sum_{i=1}^{N_j} f(x_i)$$ with large sample N_{total}

- **It is crucial to estimate V_{Φ_i}.** With iterative ML algorithm,





$[ex]: \ \{\ \vec{x}_i\,,\ y_i = (\ y^1, y^2, \cdots, y^{(n_{reg}-1)}\ ) = (\ -1,\ -1,\ 1,\ \cdots,\ 1\ )\ \}$

example

- $\infty - \infty = \text{finite}$: We are testing "fine-tuning" function of

with gaussians, $g(x;\mu,\sigma) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[\dfrac{(x-\mu)^2}{2\sigma^2}\right]$ and $G(\vec{x};\vec{\mu},\sigma) = \prod_{j=1}^{7} g(x_j;\mu_j,\sigma)$

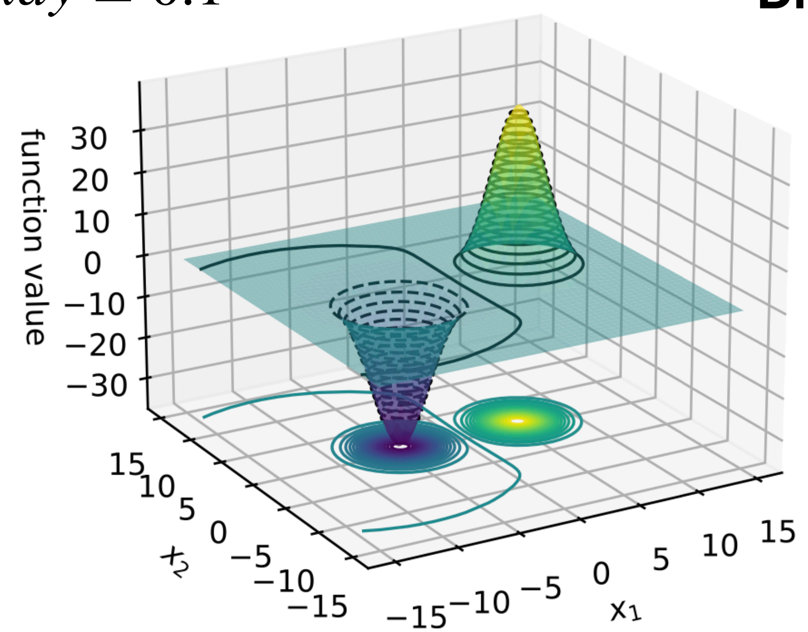
Two large peaks which are canceled each other, and one small contribution from broad distribution

$$f_{7D}(\vec{x}) = 100 \times \left[G(\vec{x};\vec{\mu}_+,\sigma_+) - G(\vec{x};\vec{\mu}_-,\sigma_-)\right] + 0.1 \times G(\vec{x};\vec{0},\sigma_0)$$

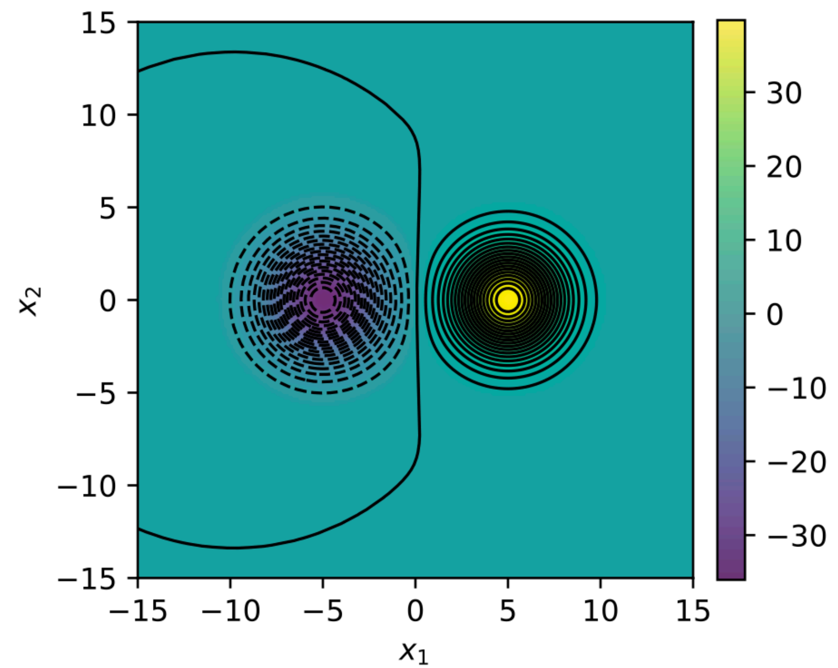
$$\sigma_+ = \sigma_- = 0.3$$

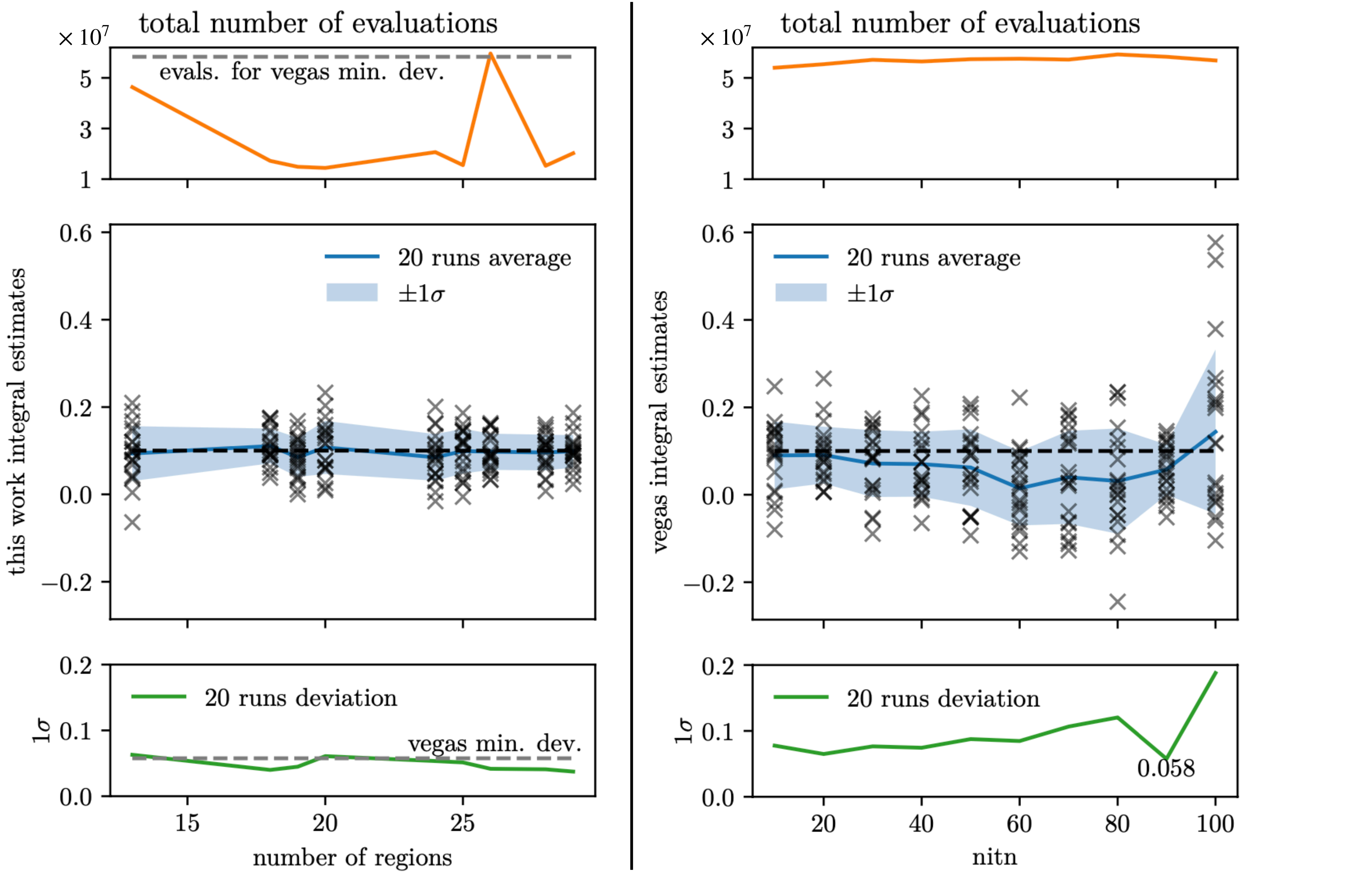
$$\sigma_0 = 1.0$$

$$\int f(\vec{x})dxdy \simeq 0.1$$



Divisions for equal "absolute" contributions

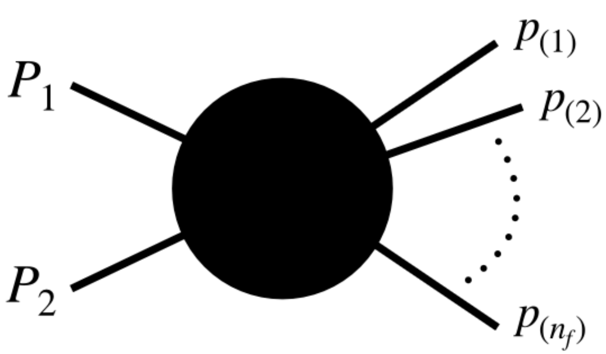


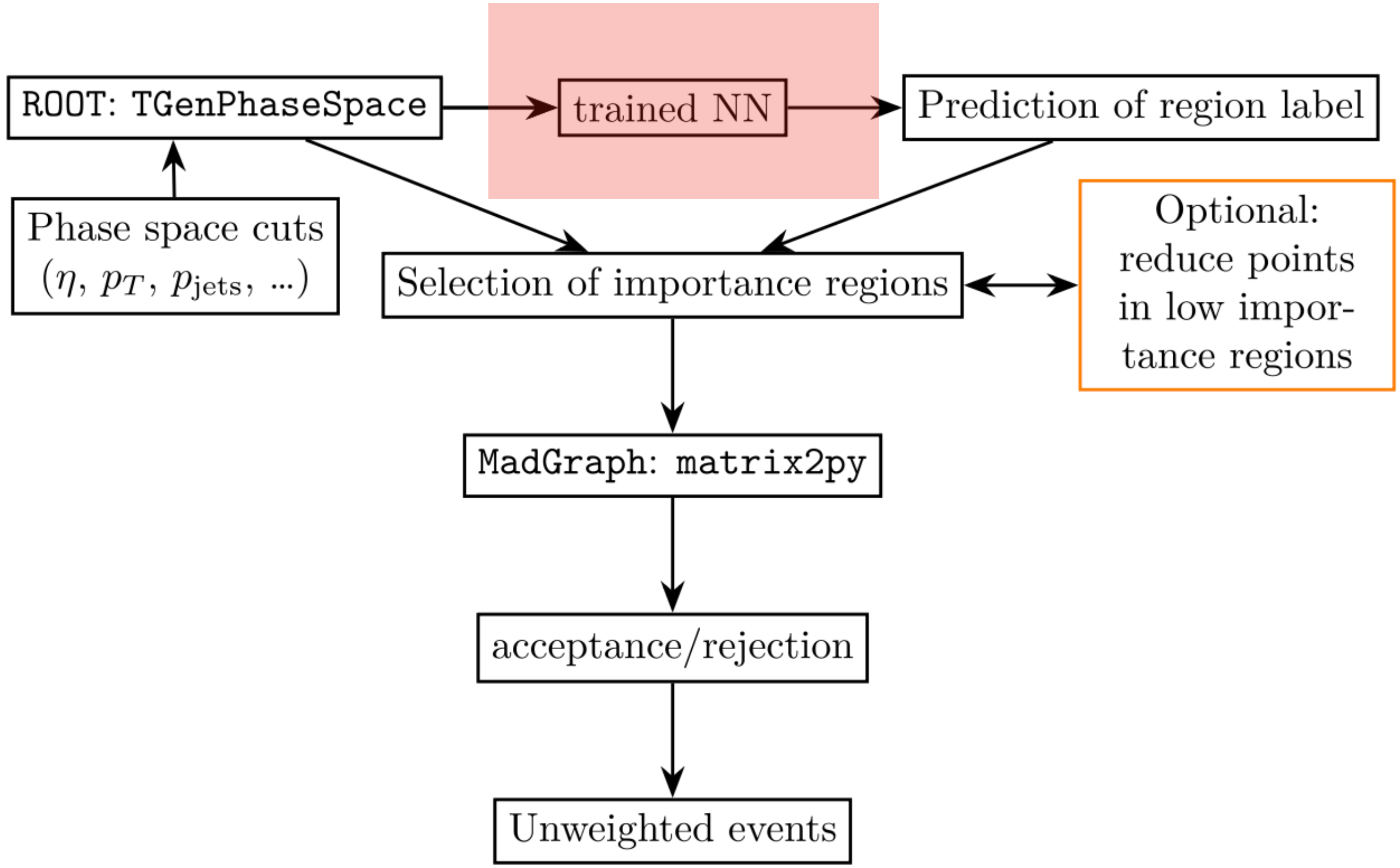


LeStrat-Net **VEGAS**

- LeStrat-Net achieves comparable precision with significantly fewer function evaluations and exhibits improved statistical stability compared to VEGAS.

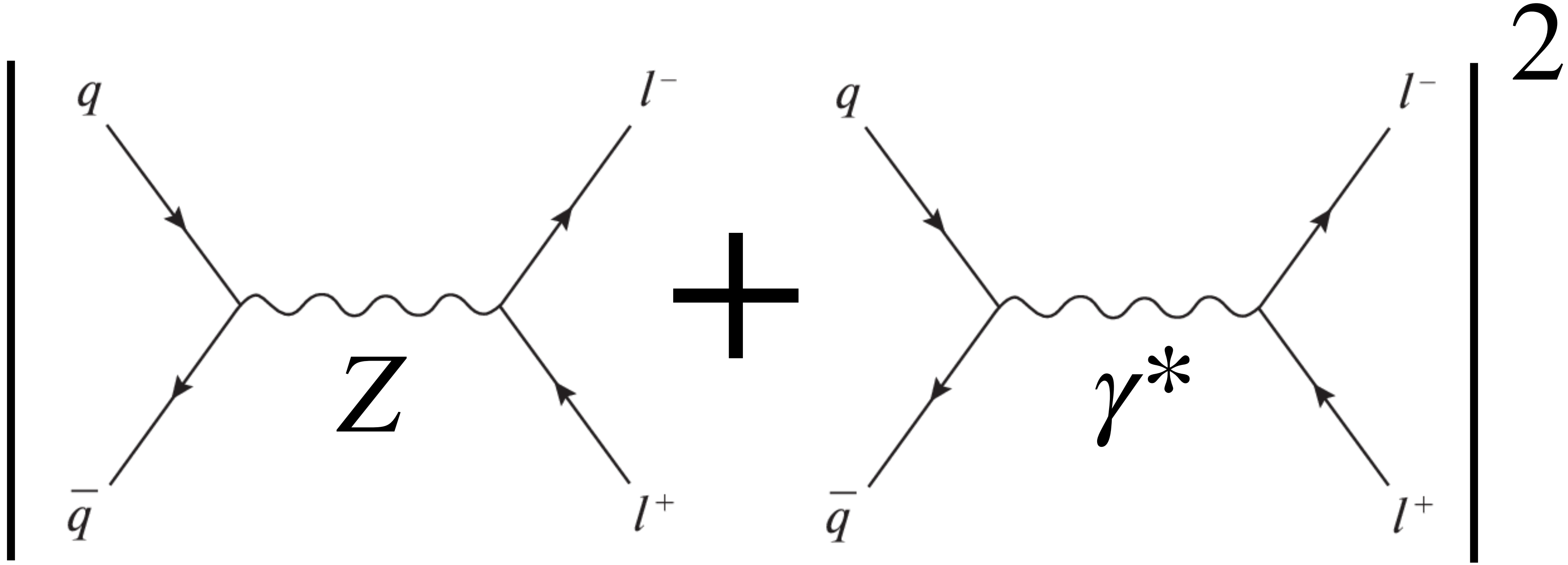
Generating MC samples with NN

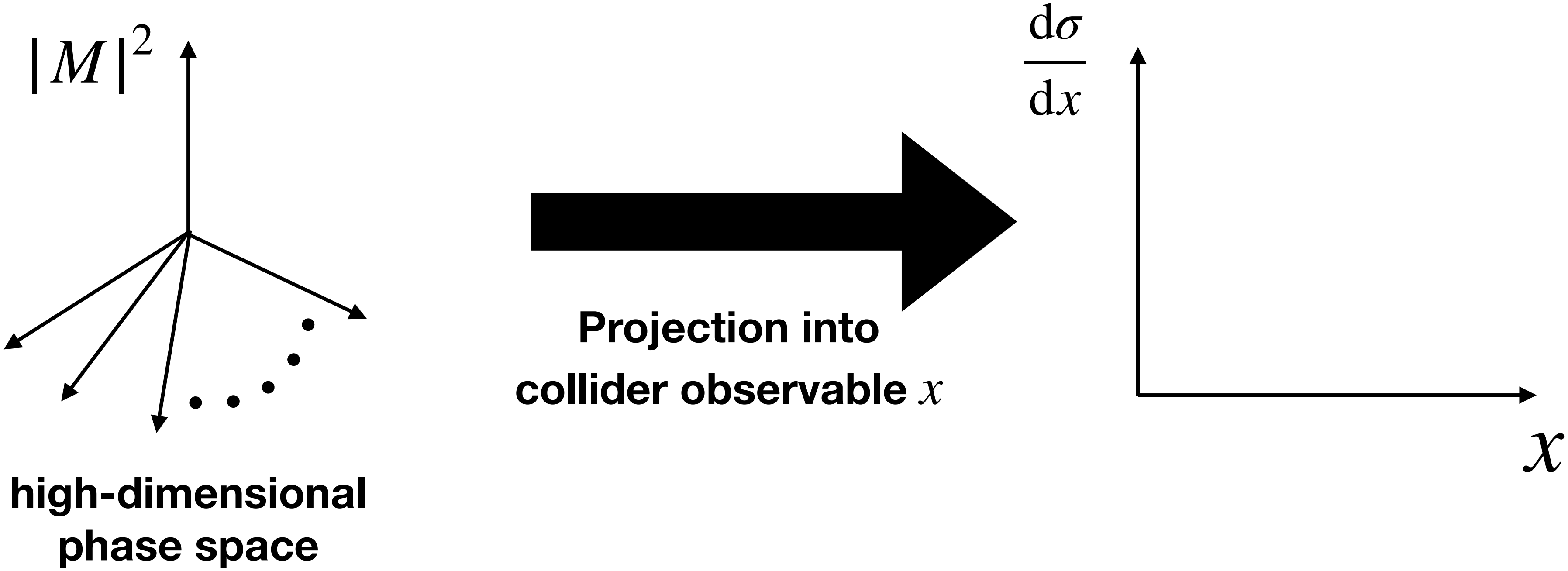




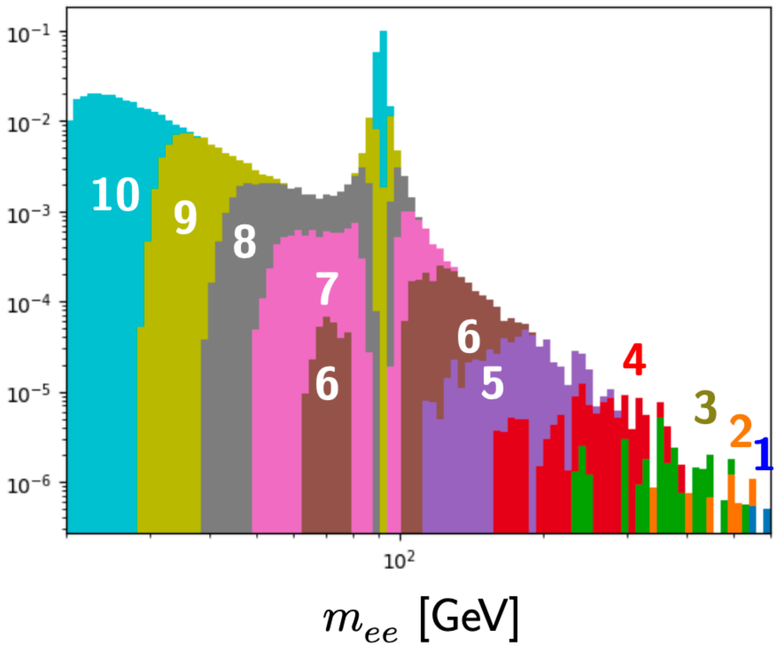
- Weight of a phase space point p is $\omega(p) = \Pi_j f(x_j, \mu) \, |M_{n_i \to n_f}(p)|^2 \, |J(p)|$
- Apply acceptance/rejection to unweight the events.

$2 \to 2$ process with interference



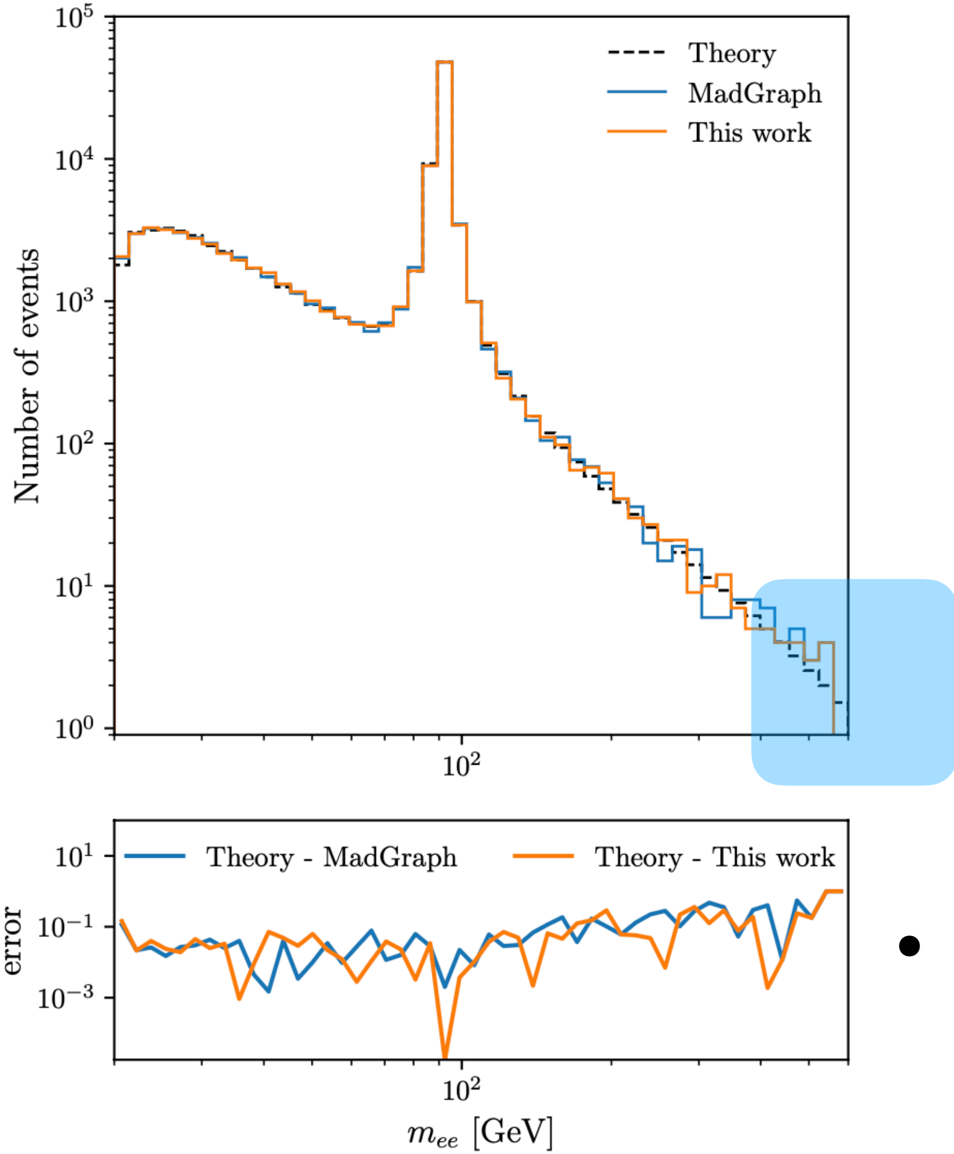


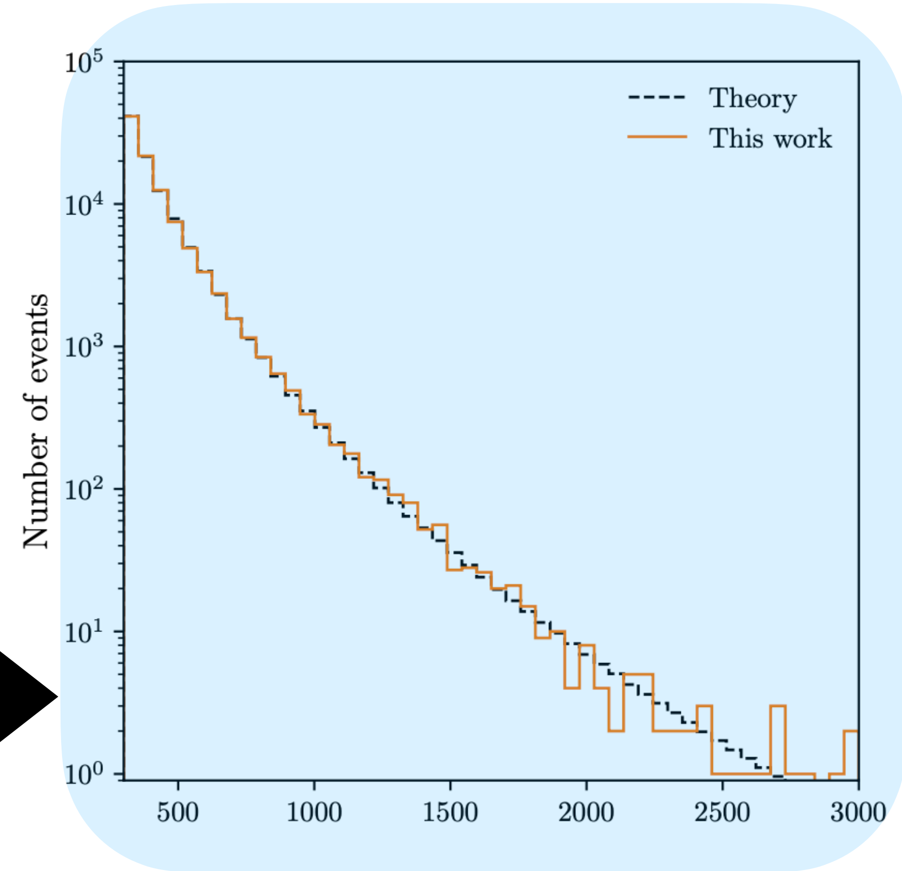
e^-e^+ invariant mass projection



- Sample each region until enough events are accumulated. **NN can tell which regions points belong to**.
- Select points using correct result.

Sample as long as we want





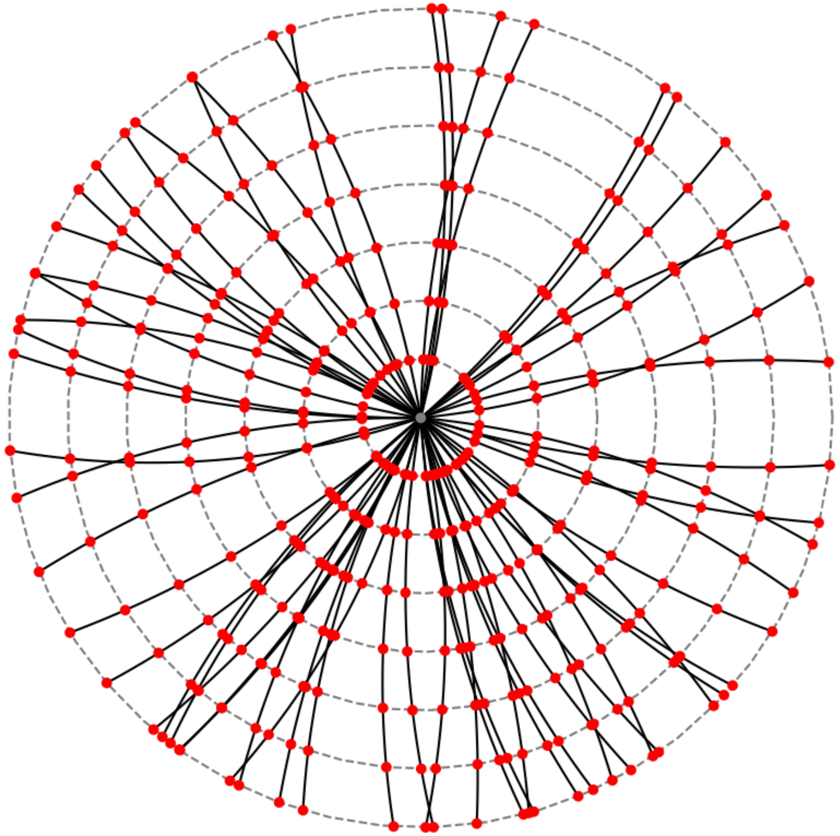
- As we divide phase-space region according to importance, we can **zoom into "rare events"**

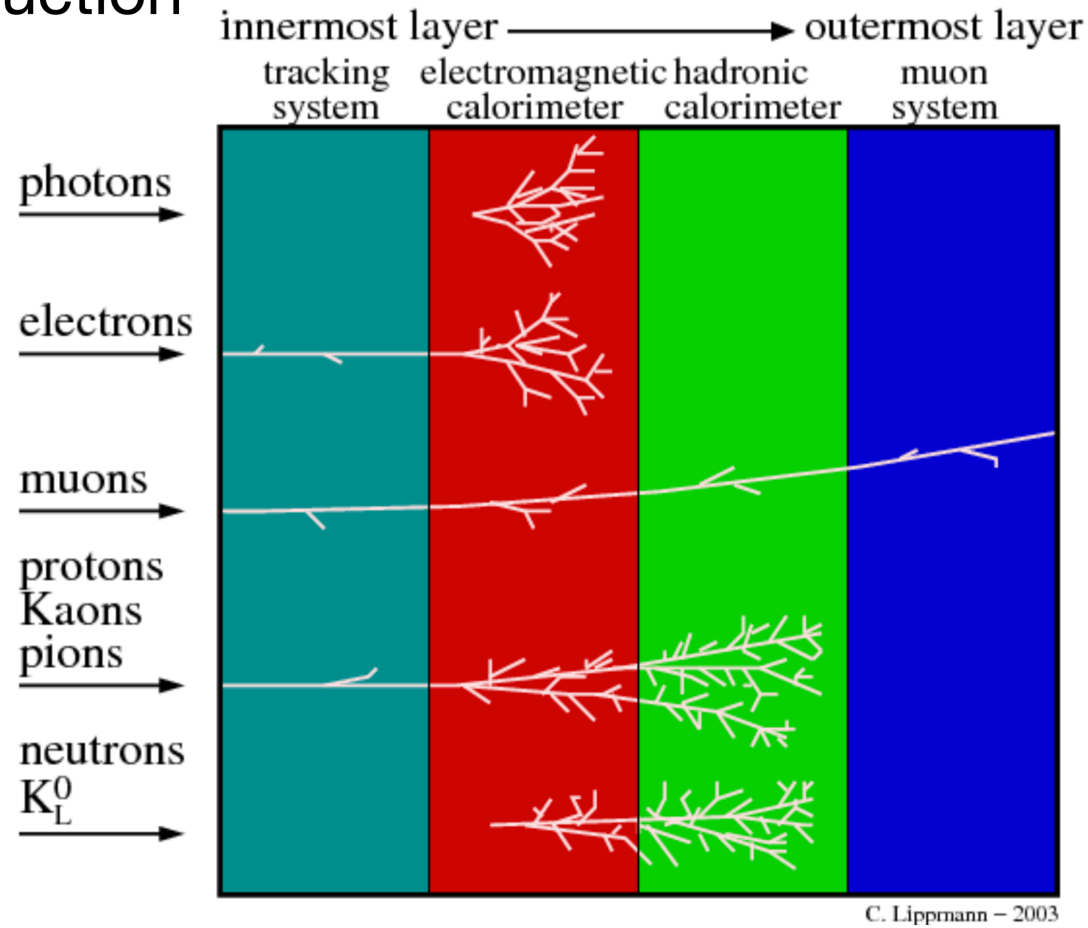
(2) **Event reconstruction on an event-by-event basis using Quantum Computations**

- to identify a physics behind complicated data

Combinatorial problem

- At the level of particle reconstruction

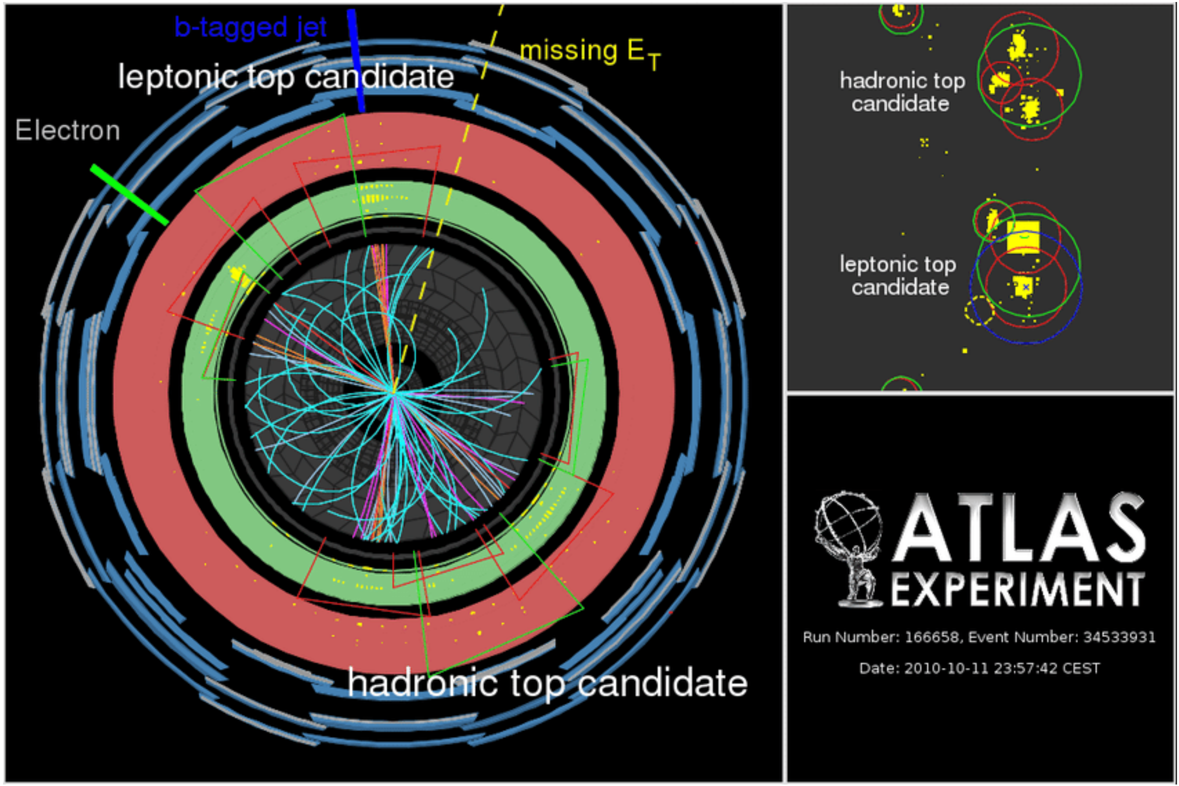


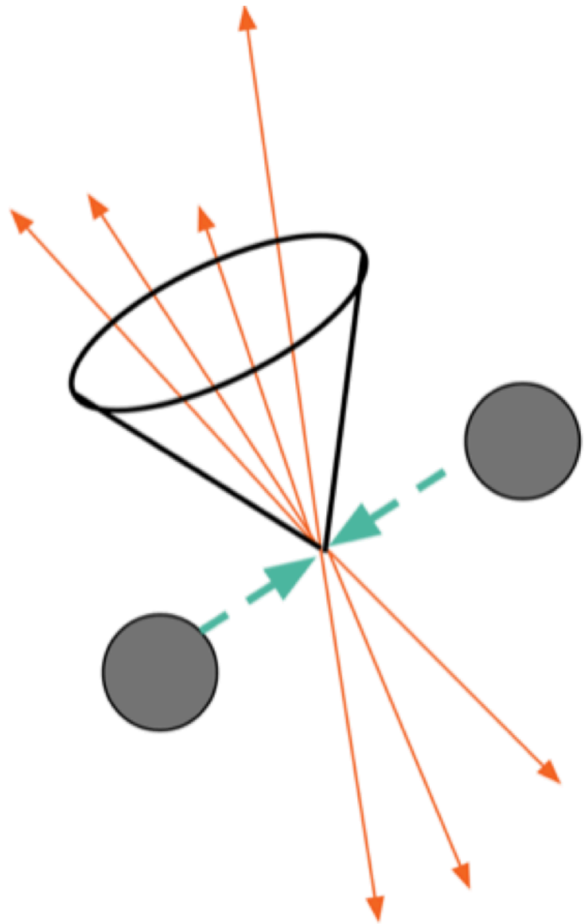


- Utilizing sub-detectors,
  Muons, electrons/photons, π , K, p ...

Combinatorial problem

- At the level of object reconstruction

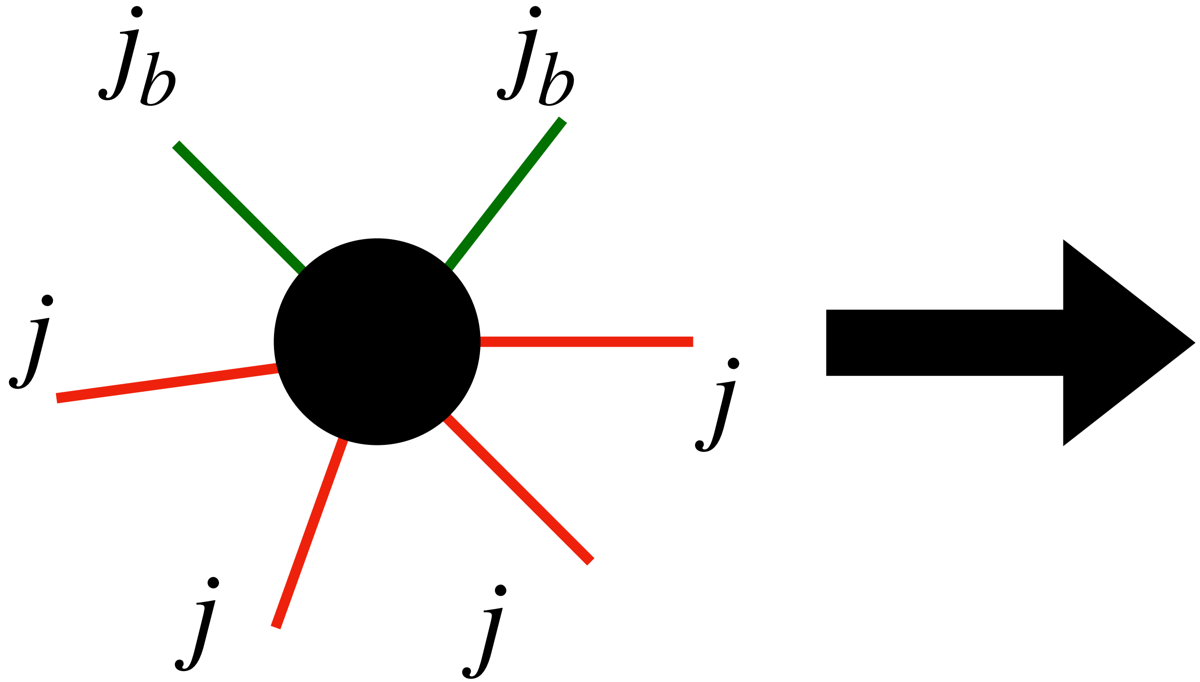


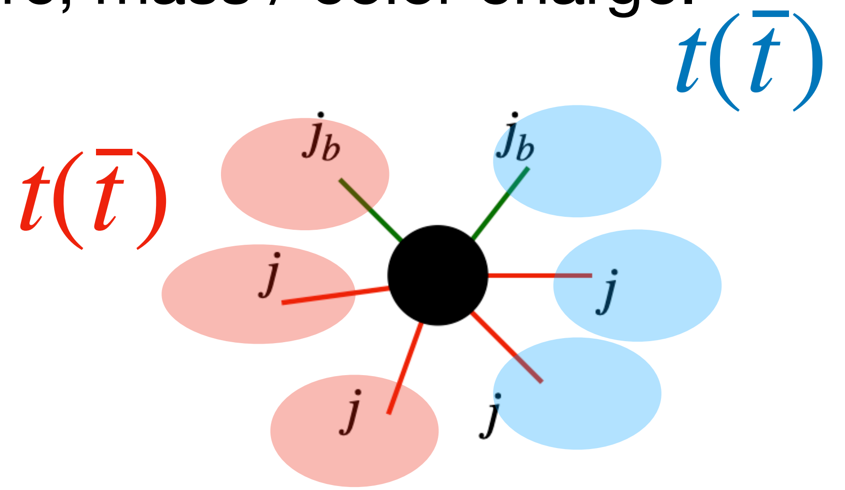


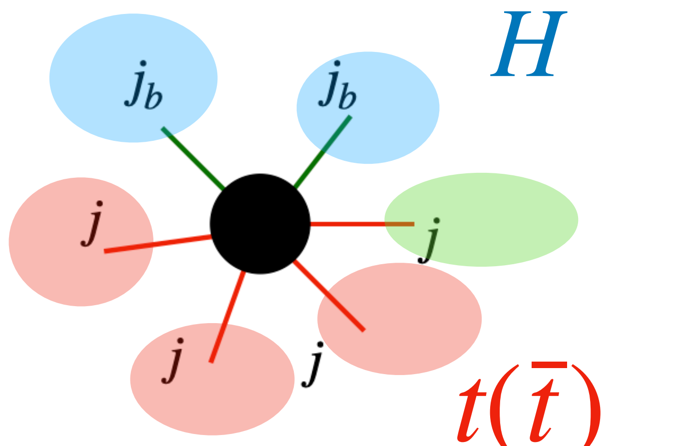
- Utilizing particle information,
  (1) isolated (good) leptons/photons,
  (2) jet objects
  (→ jet charge, particle ID can be used for jet properties)

Combinatorial problem

- At the level of "**mother particle**" reconstruction
- If **we know (assume)** the decay structure, mass / color charge.







- Guided by kinematic/QCD information,

For example, see arXiv:2311.116674
(Exploring the Synergy of Kinematics and Dynamics for Collider Physics)

But, we **need (should) not assume** anything from High Energy Collider.

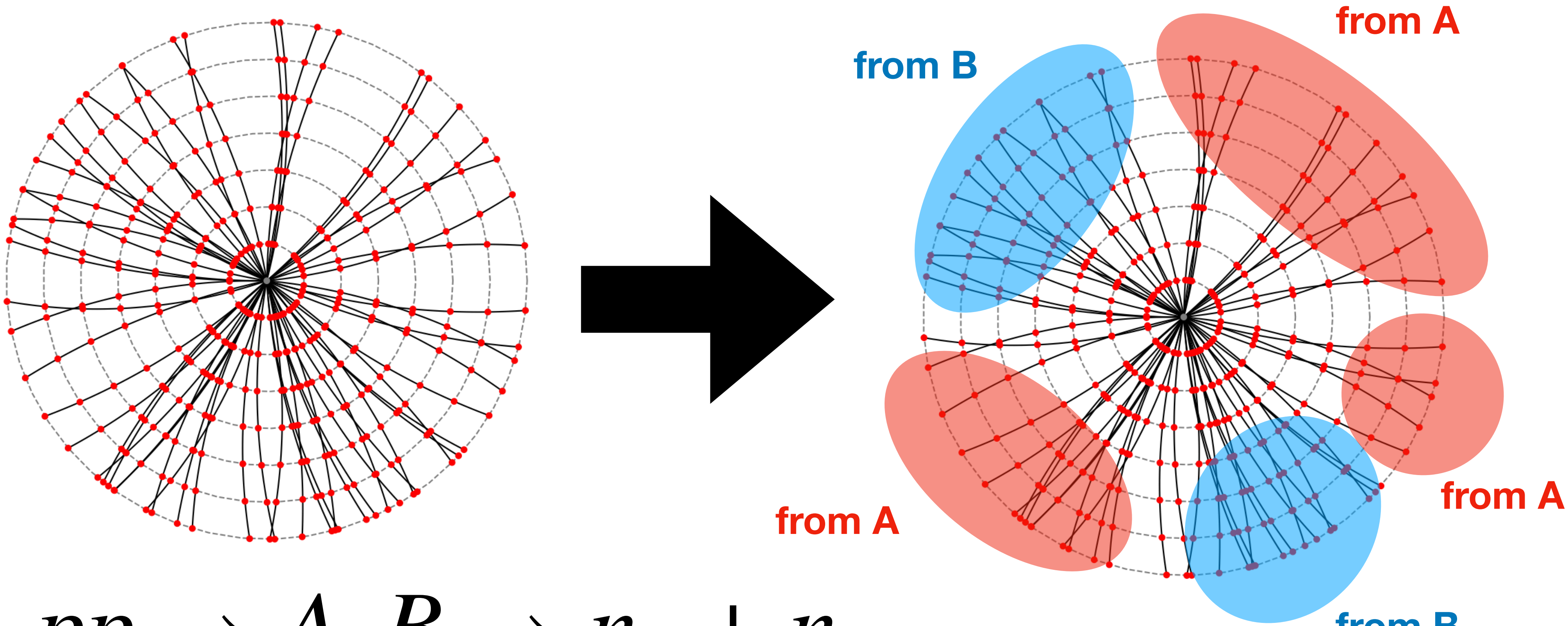
- Of course, nobody stops you to check what you want to see.

Event Reconstruction without prejudices

- As the **Colliders** (including the current LHC) **have killed various favorable new physics models**, providing us a precise details of the Standard Model (SM).

- If we see something beyond the expectation from the SM,
  - We need to reconstruct the "event-topology", to identify a physics behind the signal.
  - **To minimize assumptions, we start from unsupervised (or weakly supervised) methods**

Combinatorial problem



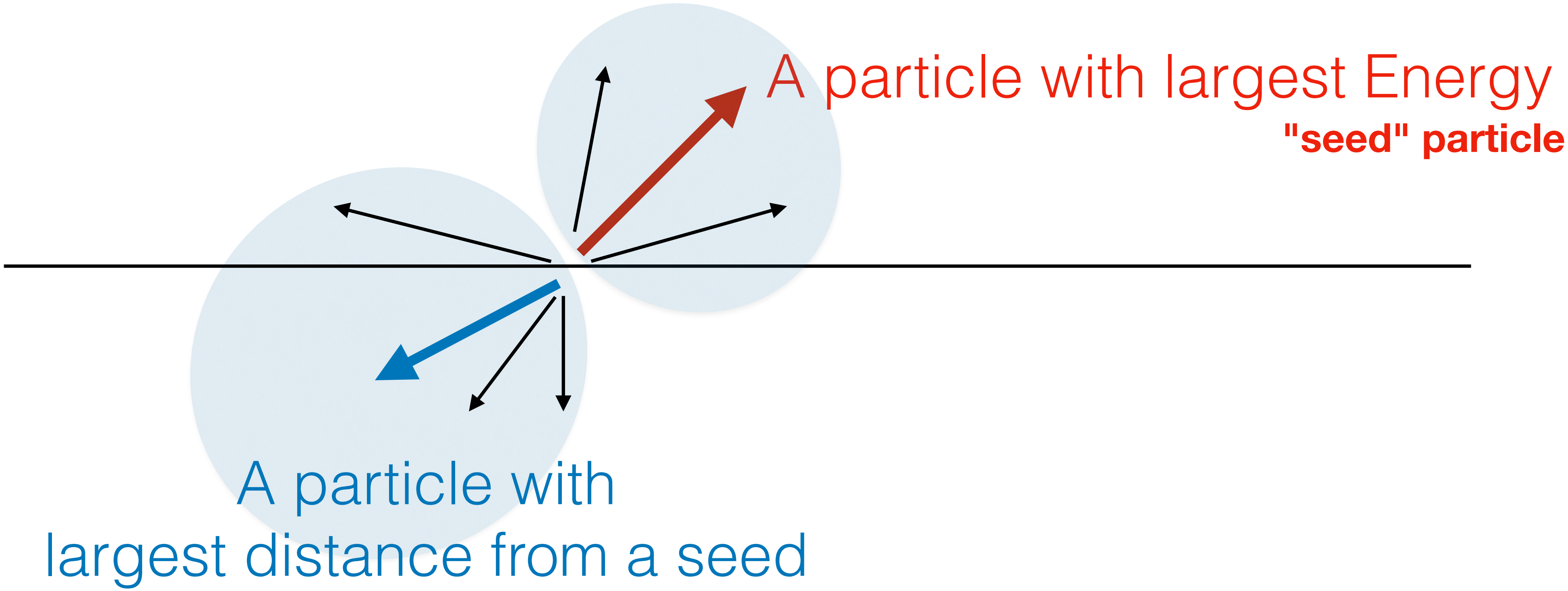
$$pp \rightarrow A, B \rightarrow n_A + n_B$$

(number of particles from the decay of A or B)

- It would be natural to start with two particles
  - Particle–antiparticle creation ensured by **CPT** and **gauge symmetries**
  - Z_2 symmetry to ensure the **stability** of (possible) dark matter

A Classic algorithm

- **Hemisphere method**: a seed-based method (iterative and converge)
  - Take two particles and clustering other particles based on these two.



With a proper metric d, one decides which hemisphere a particle belongs to.

Combinatorial problems at the LHC

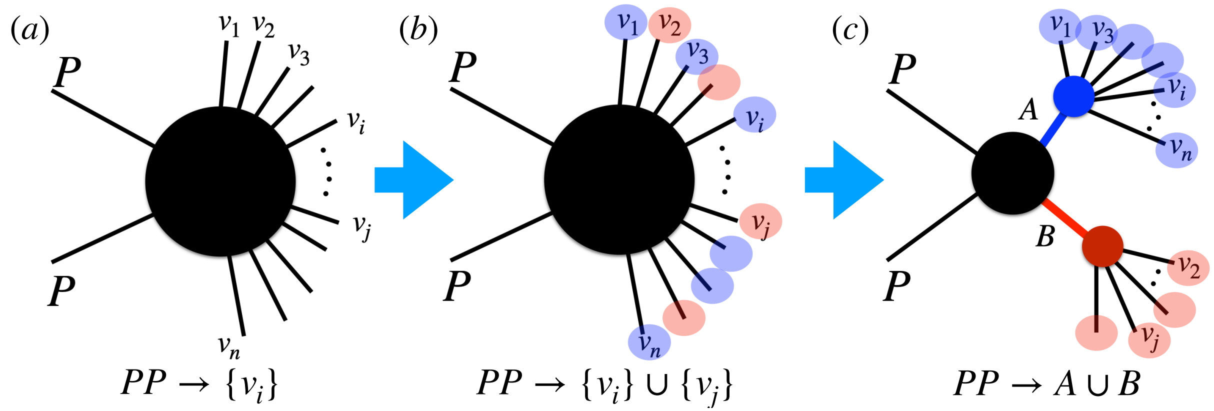


FIG. 1. (a) n-observed particles (b) Dividing n particles into two groups for $2 \to 2$ process (c) Identified event-topology with A and B.

- Assuming $2 \to 2$ production with subsequent decays, identification of an event-topology becomes a binary classification, with 2^{n-1} possibilities.
- Combinatorial problem: What would be an efficient way of assigning all observed particles in two decay chains?

p_i is the momentum of constituent of A if $x_i = 1$ $\qquad P_1 = \sum_i p_i \, x_i$

p_i is the momentum of constituent of B if $x_i = 0$ $\qquad P_2 = \sum_i p_i \, (1 - x_i)$

Minimize the mass difference: $H = \left(P_1^2 - P_2^2\right)^2$ for all possible combinations of x_i

Hybrid quantum-classical approach for combinatorial problems at hadron colliders, 2410.22417

Leveraging Quantum Annealer to identify an Event-topology at High Energy Colliders, 2111.07806

Minimization using Ising model

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

$$H_P = \left(P_A^2 - P_B^2\right)^2 \to H_P + \lambda\left(P_A^2 + P_B^2\right)$$

$$= \sum_{ij} J_{ij} s_i s_j + \frac{\lambda}{2} \sum_{ij} P_{ij} s_i s_j \text{ with } P_{ij} = p_i \cdot p_j \text{ and } J_{ij} = \sum_{kl} P_{ik} P_{jl}$$

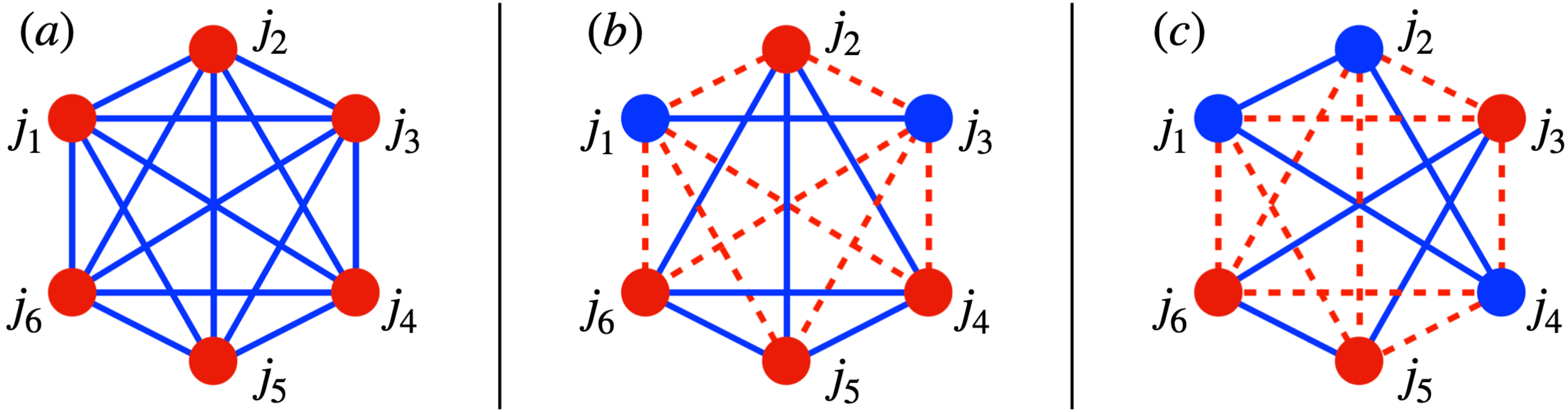
- We take $\lambda = \frac{\min(J_{ij})}{\max(P_{ij})}$

- QUBO : Quadratic Unconstrained Binary Optimization

  A generic QUBO problem: $H_{\text{QUBO}} \equiv \sum_{ij} J_{ij} s_i s_j + \sum_k h_k s_k$

- Turn into a combinatorial problem into a graph problem.

  For example, $pp \to AB \to j_1, j_2, \cdots, j_6$



(a)
- All jets are from **A**

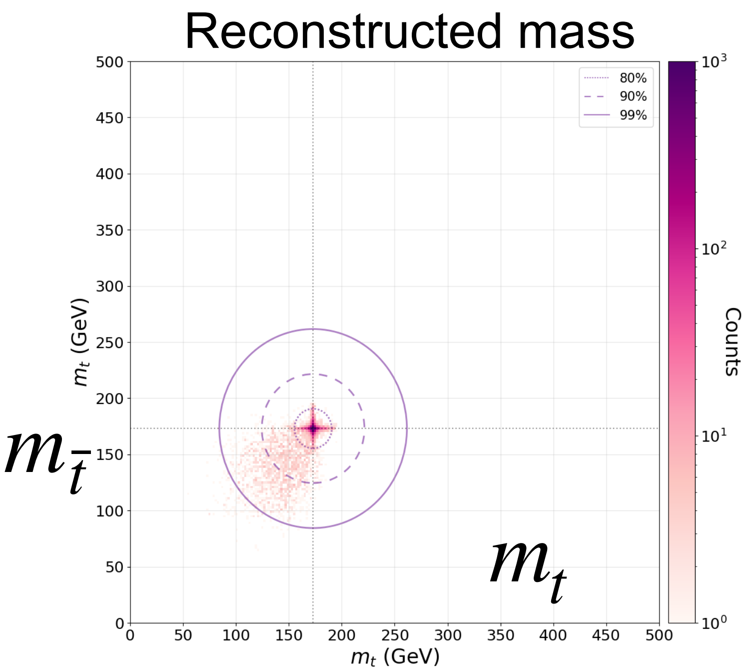
(b)
- $\{j_2, j_4, j_5, j_6\}$ from **A**
- $\{j_1, j_3\}$ from **B**

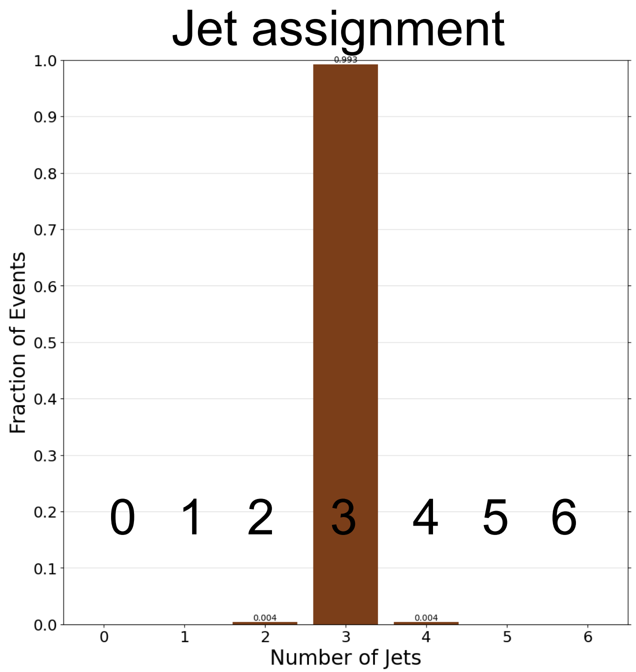
(c)
- $\{j_3, j_5, j_6\}$ from **A**
- $\{j_1, j_2, j_4\}$ from **B**

$$H_p = \sum_{ij} J_{ij} s_i s_j + \frac{\lambda}{2} \sum_{ij} P_{ij} s_i s_j$$

- **Edges connecting nodes from the same parent particle** ($s_i s_j = +1$) contribute **positively** to H_P and are shown as solid (blue) lines
- **Edges connecting nodes from different parent particles** ($s_i s_j = -1$) contribute **negatively** to H_P
- The task of finding the **correct combinatorics** is equivalent to identifying the graph configuration that **minimizes** H_P

Parton-level truth and hemisphere method



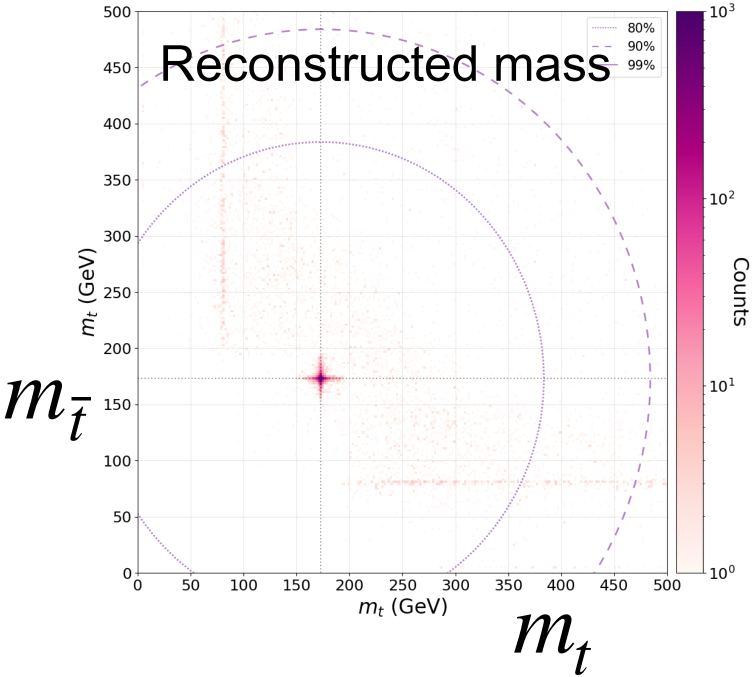


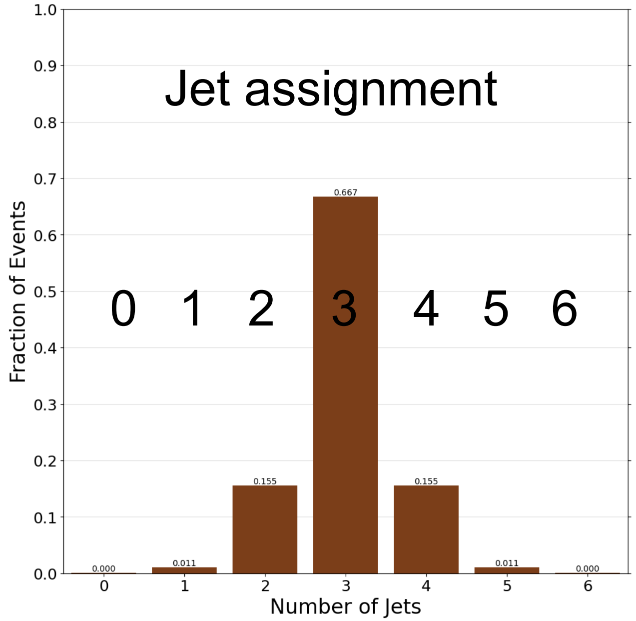
parton-level truth

79% overal efficiency

No mass information used.

Hybrid quantum-classical approach for combinatorial problems at hadron colliders: 2410.22417





Hemisphere method

50% overal efficiency

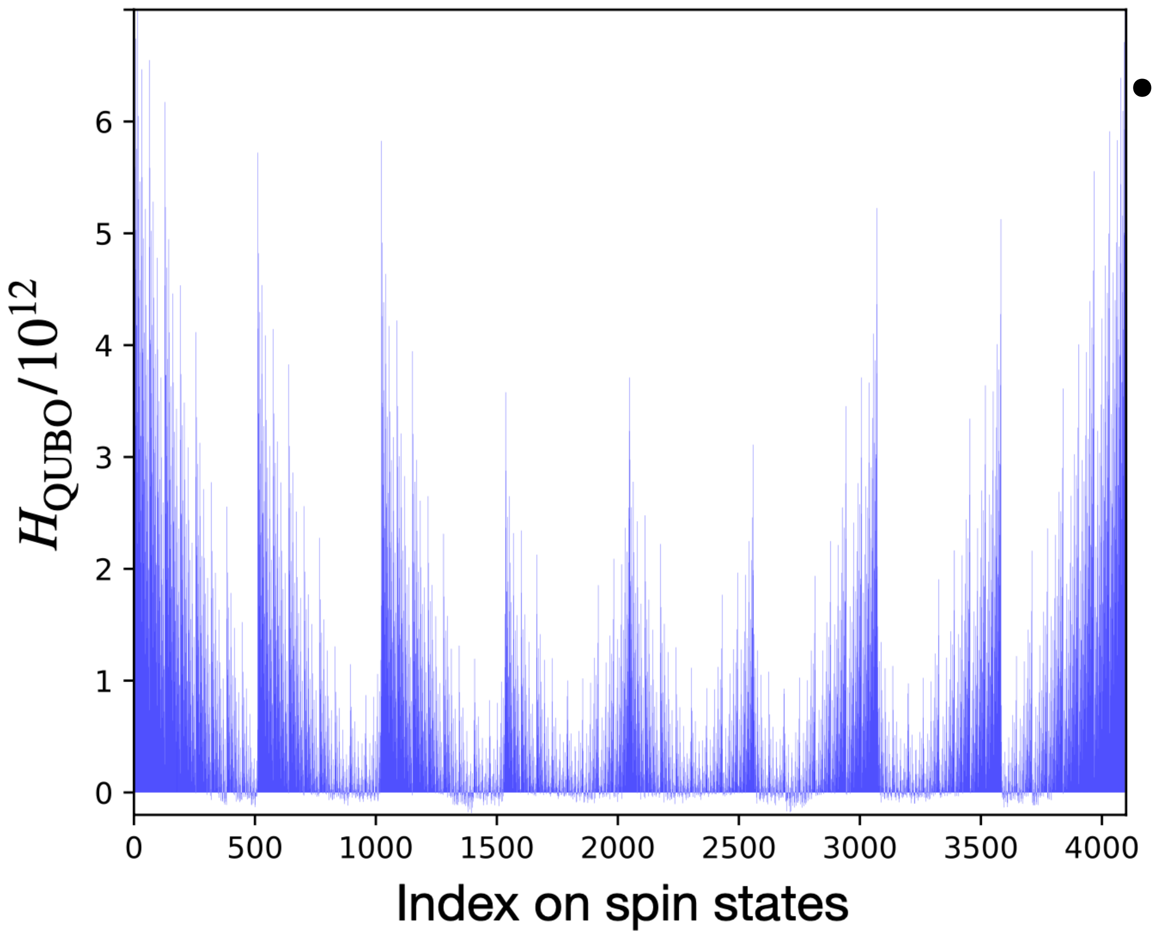
No mass information used.

Issue

Difficulty in finding the ground state of an Ising Hamiltonian

Combinatorial complexity

Landscape of energy distribution



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

- Due to the rapidly changing shape of the potential, **any classical algorithm would fail.** (except the brute-force scanning)

- This example is from an event of 12 particles from a collision

Minimization with an adiabatic process

$$H(t) = \left(1 - \frac{t}{T}\right) H_0 + \left(\frac{t}{T}\right) H_P \text{ from } t = 0 \text{ to } t = T$$

- H_0 : The initial Hamiltonian whose ground state is easy to prepare.

- H_P : The problem Hamiltonian whose ground state is a solution to the optimization problem.

: Thus at $t = T$, we will have a solution for a hamiltonian $H(T) = H_P$

Utilizing Gate Quantum Circuit

- Starting from a well-known QAOA (arXiv:1411.4028) (Quantum Approximation Optimization Algorithm)

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

- with $H_M = \sum \sigma_i^X$ and $H_P = H_{\text{QUBO}} = \sum J_{ij} \sigma_i^z \sigma_j^z + \lambda \sum P_{ij} \sigma_j^z \sigma_j^z$

- Starting from the ground state of H_M : $|\psi_0\rangle = |-\rangle$, evolve to the ground state of H_P : $|\psi\rangle = U(T,0)\,|\psi_0\rangle$

$$U(T,0) = U(T, T - \Delta t) U(T - \Delta t, T - 2\Delta t) \cdots U(\Delta t, 0)$$
$$= \prod_{j=1}^{p} U(j\Delta t, (j-1)\Delta t)$$

- Evolution operator $U(T,0)$

$$U(T,0) \approx \prod_{j=1}^{p} e^{-i\Delta t H(j\Delta t)} = \prod_{j=1}^{p} \exp\Big[-i\Delta t\big[(1-a(j\Delta t))H_M + a(j\Delta t)H_P\big]\Big]$$

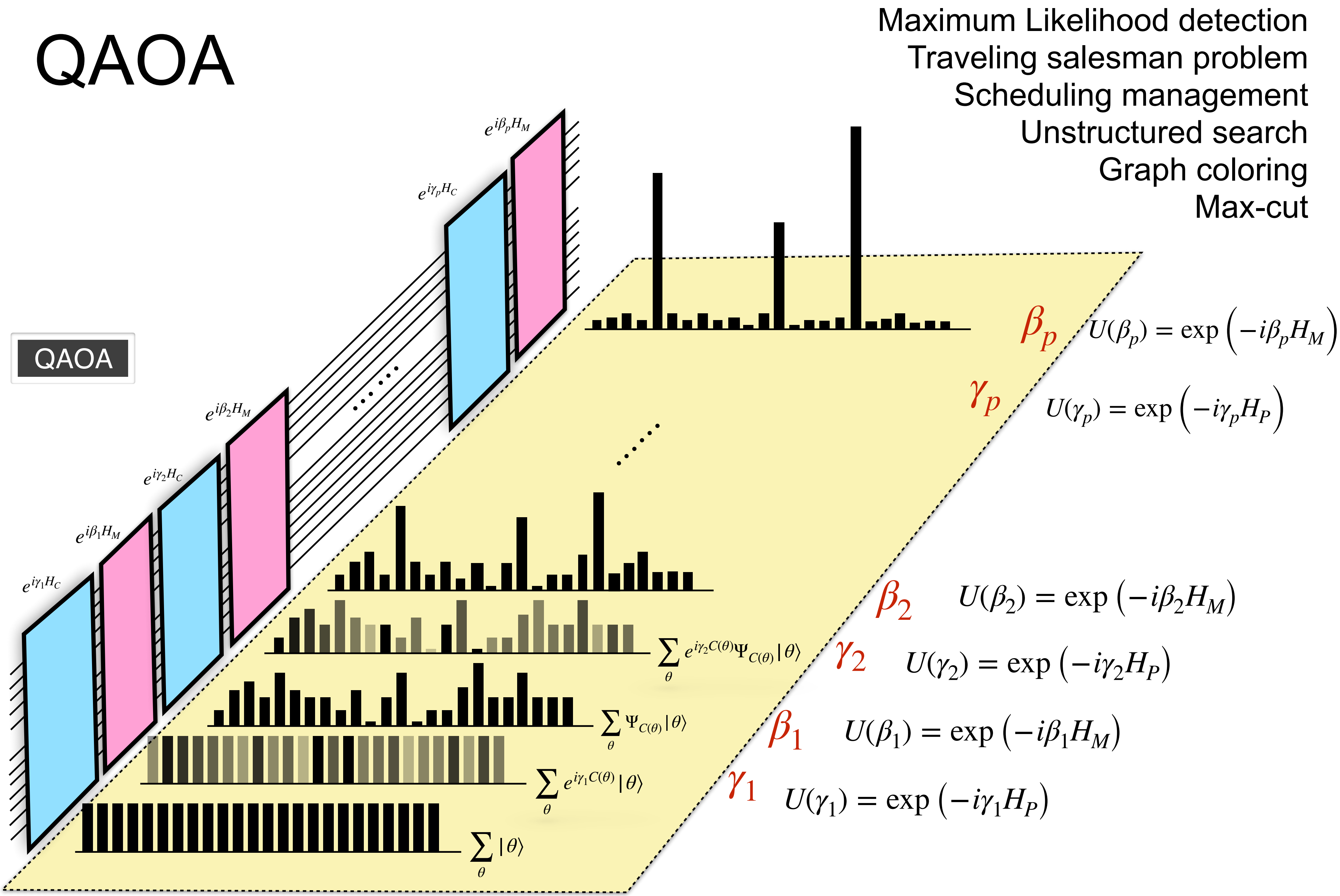
- With Suzuki-Trotter decomposition, $\lim_{p\to\infty} (e^{iAt/p}e^{iBt/p})^p = e^{i(A+B)t}$

$$U(T,0) \approx \prod_{j=1}^{p} \exp\Big[-i\Delta t(1-a(j\Delta t))H_M\Big]\exp\Big[-i\Delta t a(j\Delta t)H_P\Big]$$

$$= \prod_{j=1}^{p} \exp\Big[-i\beta_j H_M\Big]\exp\Big[-i\gamma_j H_P\Big]$$

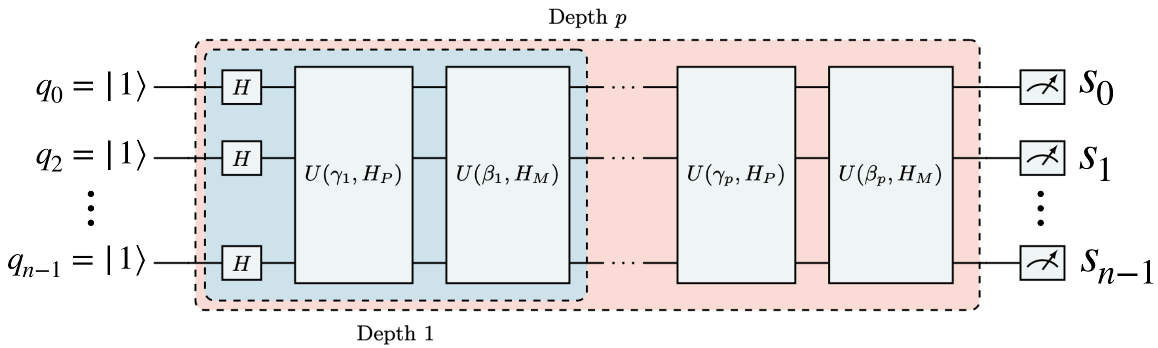
$$|\psi\rangle \equiv |\boldsymbol{\beta},\boldsymbol{\gamma}\rangle = \sum_{j=1}^{p} U(\beta_j, H_M)U(\gamma_j, H_P)|-\rangle^{\otimes n}$$

Works in the adiabatic limit or $p \to \infty$



QAOA
Maximum Likelihood detection
Traveling salesman problem
Scheduling management
Unstructured search
Graph coloring
Max-cut
QAOA
$e^{i\gamma_1 H_C}$
$e^{i\beta_1 H_M}$
$e^{i\gamma_2 H_C}$
$e^{i\beta_2 H_M}$
$e^{i\gamma_p H_C}$
$e^{i\beta_p H_M}$
$\sum_\theta |\theta\rangle$
$\sum_\theta e^{i\gamma_1 C(\theta)} |\theta\rangle$
$\sum_\theta \Psi_{C(\theta)} |\theta\rangle$
$\sum_\theta e^{i\gamma_2 C(\theta)} \Psi_{C(\theta)} |\theta\rangle$
β_p $U(\beta_p) = \exp\left(-i\beta_p H_M\right)$
γ_p $U(\gamma_p) = \exp\left(-i\gamma_p H_P\right)$
β_2 $U(\beta_2) = \exp\left(-i\beta_2 H_M\right)$
γ_2 $U(\gamma_2) = \exp\left(-i\gamma_2 H_P\right)$
β_1 $U(\beta_1) = \exp\left(-i\beta_1 H_M\right)$
γ_1 $U(\gamma_1) = \exp\left(-i\gamma_1 H_P\right)$

Machine Learning (to make QAOA more efficient)



Measurement:

$$L(\beta,\gamma) \equiv \langle \psi(\beta,\gamma) | H_P | \psi(\beta,\gamma) \rangle$$

Using "classical optimizer"

update

$$(\beta,\gamma) \to (\beta^*,\gamma^*)$$

- Find the minimum of a given hamiltonian H.

  - The state $|\psi(\beta,\gamma)\rangle = U(\beta,\gamma)|0\rangle$

  - $L(\beta,\gamma) \equiv \langle \psi(\beta,\gamma) | H_P | \psi(\beta,\gamma) \rangle \simeq \frac{1}{N_s} \sum H_P(s_0, \cdots, s_{n-1})$

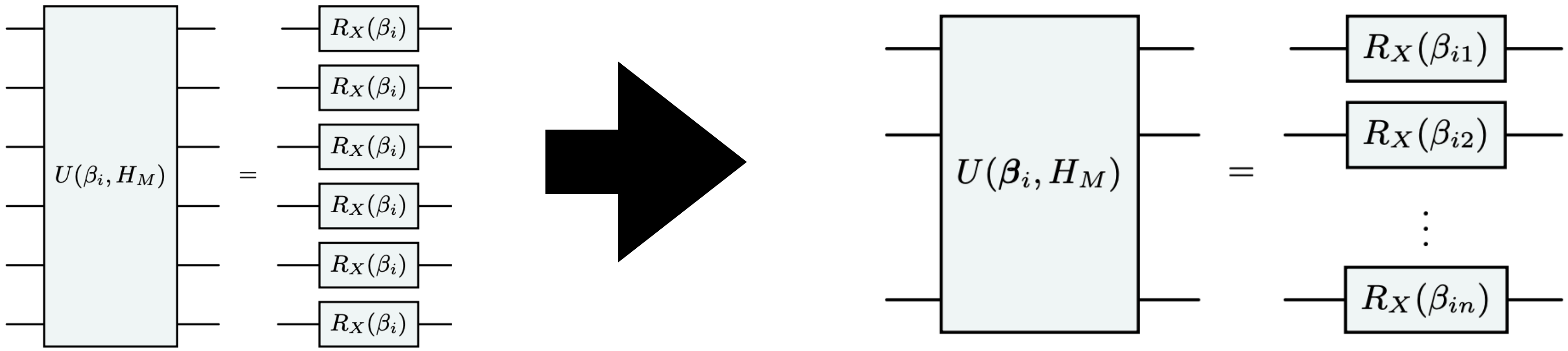
  will have a minimum when $|\psi(\beta^*,\gamma^*)\rangle = U(\beta^*,\gamma^*)|0\rangle \simeq |\psi_{\min}\rangle$

- Finding the $(\beta^*,\gamma^*) = \operatorname{argmin} L(\beta,\gamma)$ with a classical optimizer method. (For example, Gradient Descent Method)

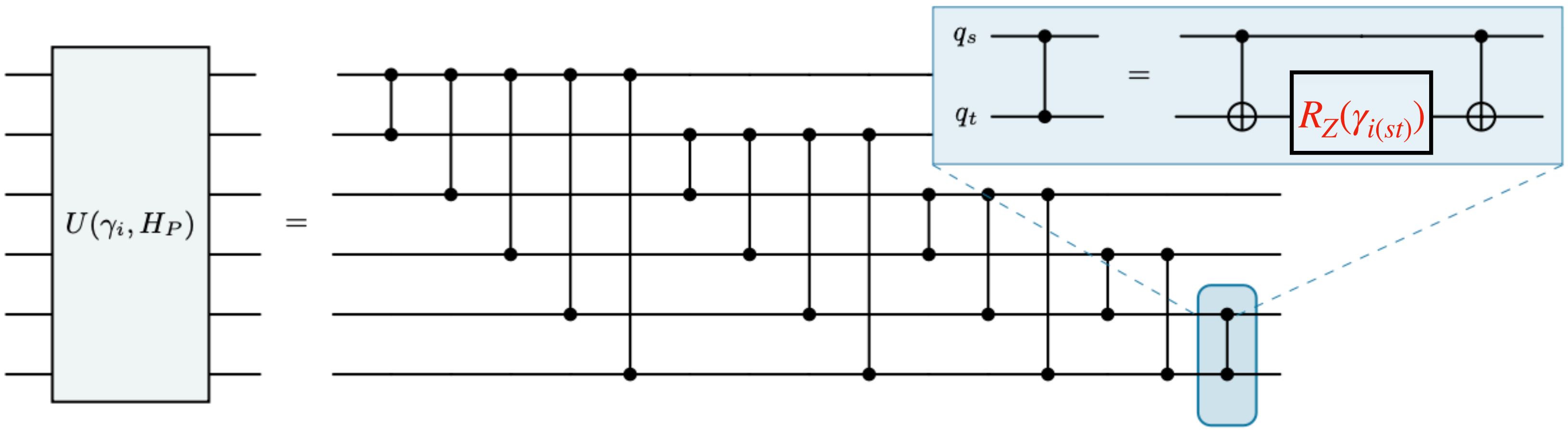
Variations of QAOA

- Multi Angle (MA)-QAOA: What if each qubit has its own free parameter? (**increase degree of freedom**, shifting burdens from QC to Optimizer)

Different parameters for each qubit in H_M

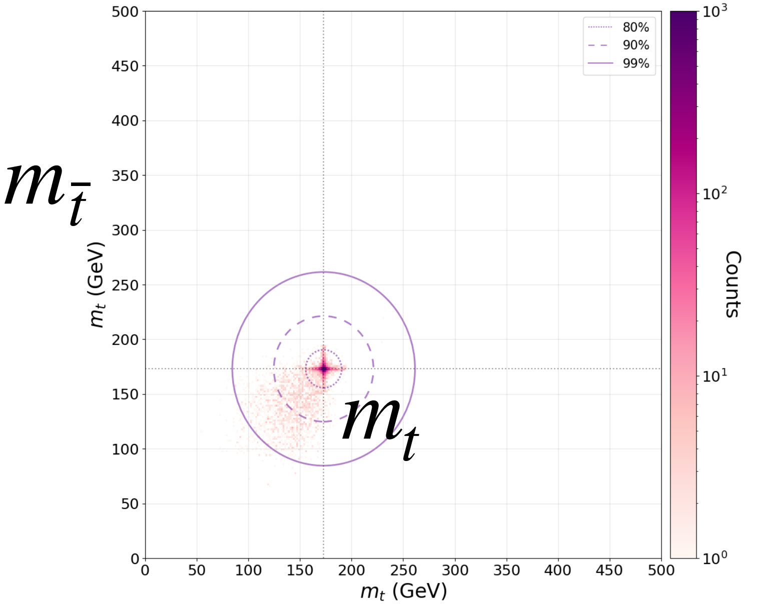


Different parameters for each pair of two qubits in H_P

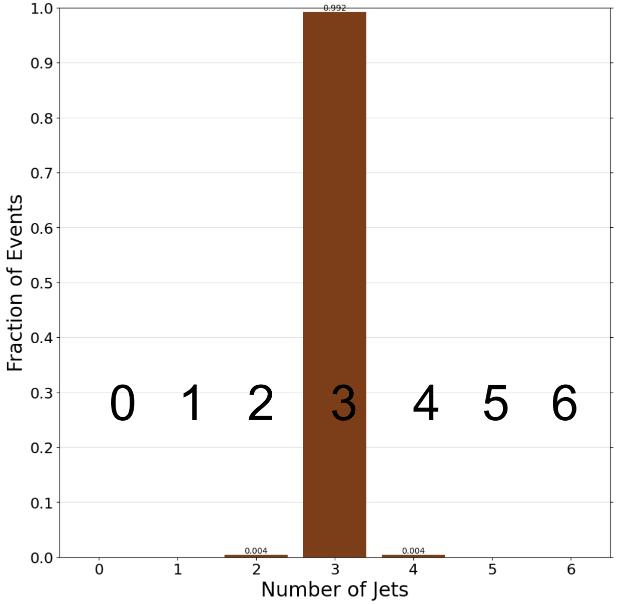


Results from Quantum Optimization Algorithm

Reconstructed mass



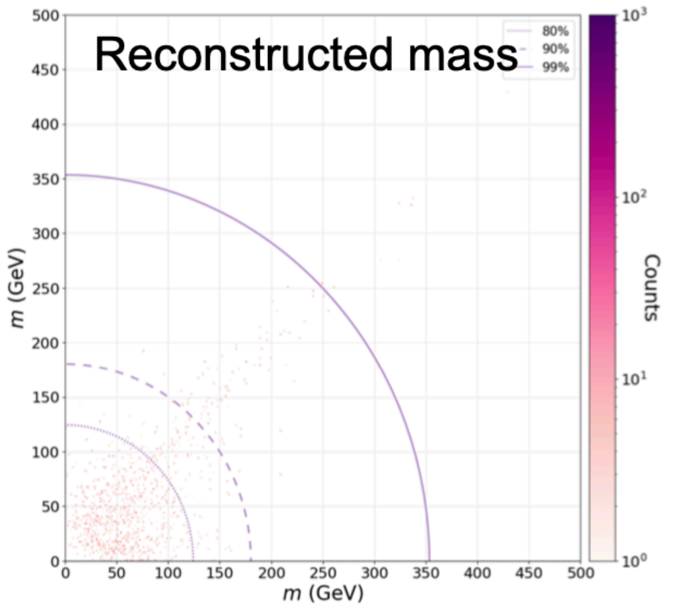
Jet assignment

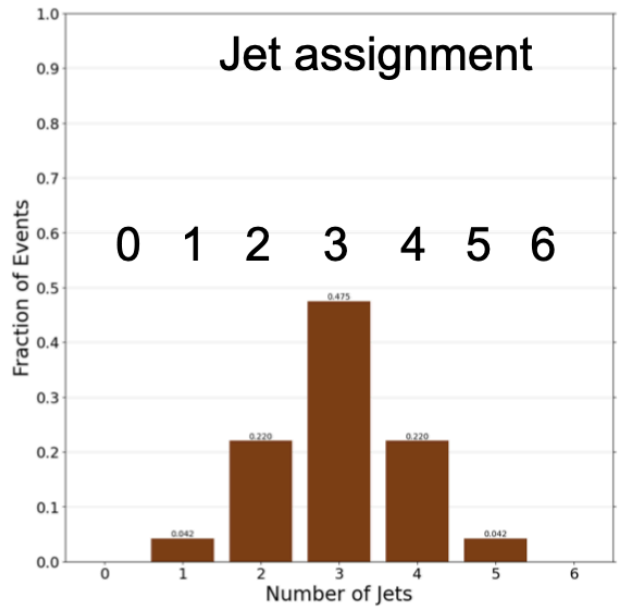


$pp \to t\bar{t} \to 6j$

79% overal efficiency

No mass information used.





$pp \to 6j$ (QCD backgrounds)

multi-jet (6 jets)

Comparison

Methods	matching accuracy (efficiency)		n_{train}	$n_{\text{parameters}}$	depth (p)	n_{CNOT}	n_{R_Z}	n_{R_X}
	parton-level	smeared events						
Hemisphere	50%	48%	N/A					
QAOA	55%	53%	N/A	16	8	240	120	48
ma-QAOA	75%	73%		168				
FALQON	72%	69%		2	250	7,500	3,750	1,500
SPANet	91%	70%	5×10^5	10^6	N/A			
	81%	62%	2×10^4	1.9×10^3				

- Hybrid Quantum Methods are strong with smearing by detector. (**SPANet is a supervised learning algorithm**, “SPANet: Generalized permutation-less set assignment for particle physics using symmetry preserving attention,” SciPost Phys. 12, 178 (2022), arXiv:2106.03898"

Conclusion

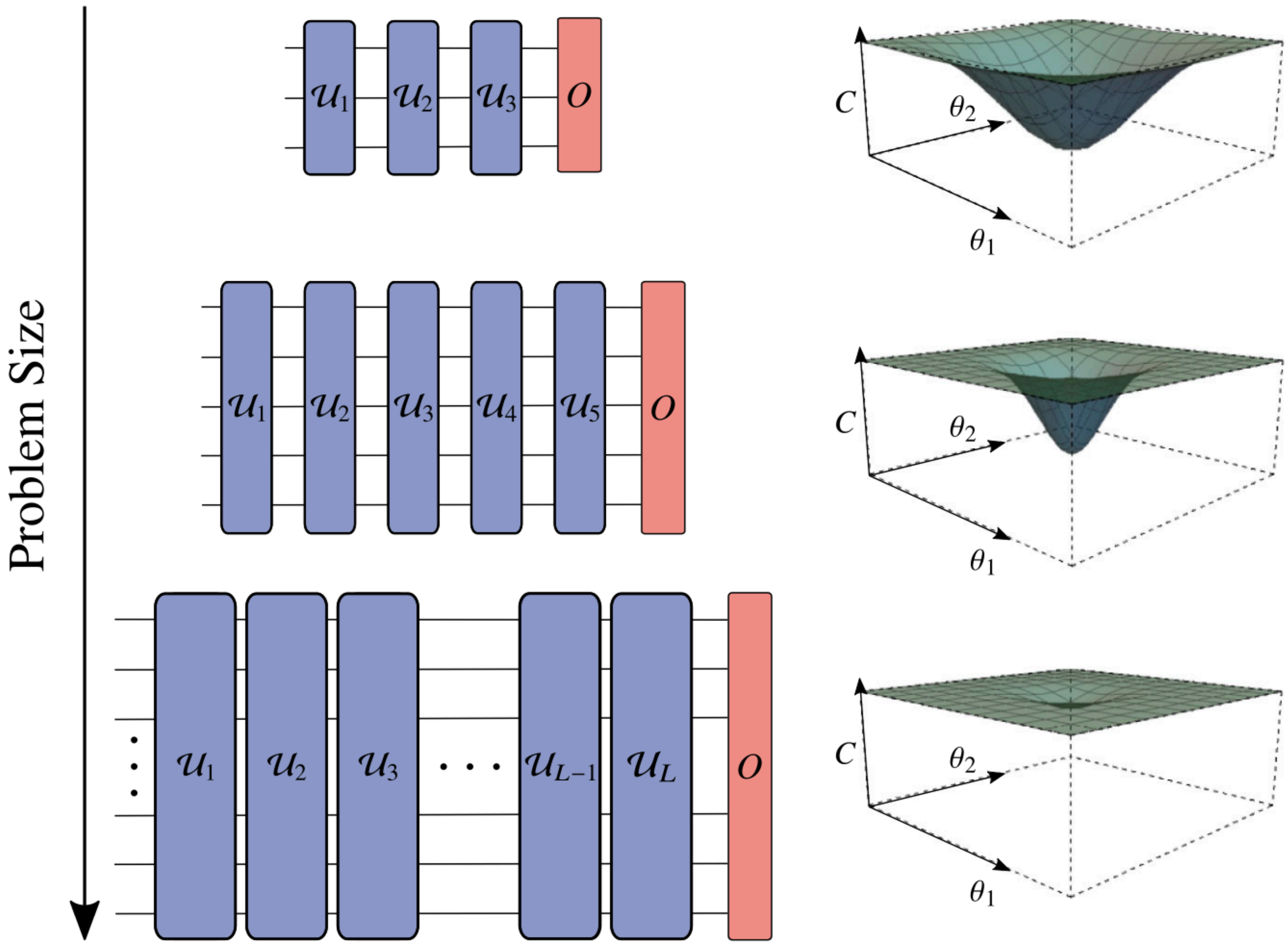
- **Lebesgue-style Monte Carlo with LeStrat-Net**:
  Comparable precision with **far fewer evaluations** and **stable uncertainties**, enabling targeted sampling of rare regions in high-dimensional phase space.

- **Model-agnostic event-topology reconstruction**:
  Formulating the combinatorial assignment as an Ising/QUBO optimization problem, solved by QAOA and its variants, enables robust, unsupervised reconstruction of event topologies **on an event-by-event basis**, without relying on model assumptions.

- **Outlook**: integrate with CEPC simulation chain (generator → detector), benchmark on CEPC-like processes ($e^+e^-\to ZH/WW/ZZ$), and explore hardware-aware quantum algorithms (FALQON/ma-QAOA) for near-term quantum devices.

Back-up

- With increasing size of a circuit, the gradient vanishes exponentially.



Nature Communications volume **12**, Article number: 6961 (2021) by Samson Wang et.al.

- FALQON (Feedback-Based Quantum Optimization): Purely QC

Alicia B. Magann et.al., PRL 129, 250502 (2022)

$$H(t) = H_P + \beta(t) H_M$$

$$i \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = (H_P + \beta(t) H_M) |\psi(t)\rangle \longrightarrow \quad E_\mathrm{p}(|\psi(t)\rangle) = \langle \psi(t) | H_\mathrm{p} | \psi(t) \rangle$$

$$\frac{\mathrm{d}E_P}{\mathrm{d}t} = \langle \psi(t) | i \left[H_P + \beta(t) H_M, H_P \right] | \psi(t) \rangle = \beta(t) \langle \psi(t) | i \left[H_M, H_P \right] | \psi(t) \rangle$$

$$\equiv A(t)\beta(t) < 0$$ **if we take** $\beta(t) = -A(t-\tau)$