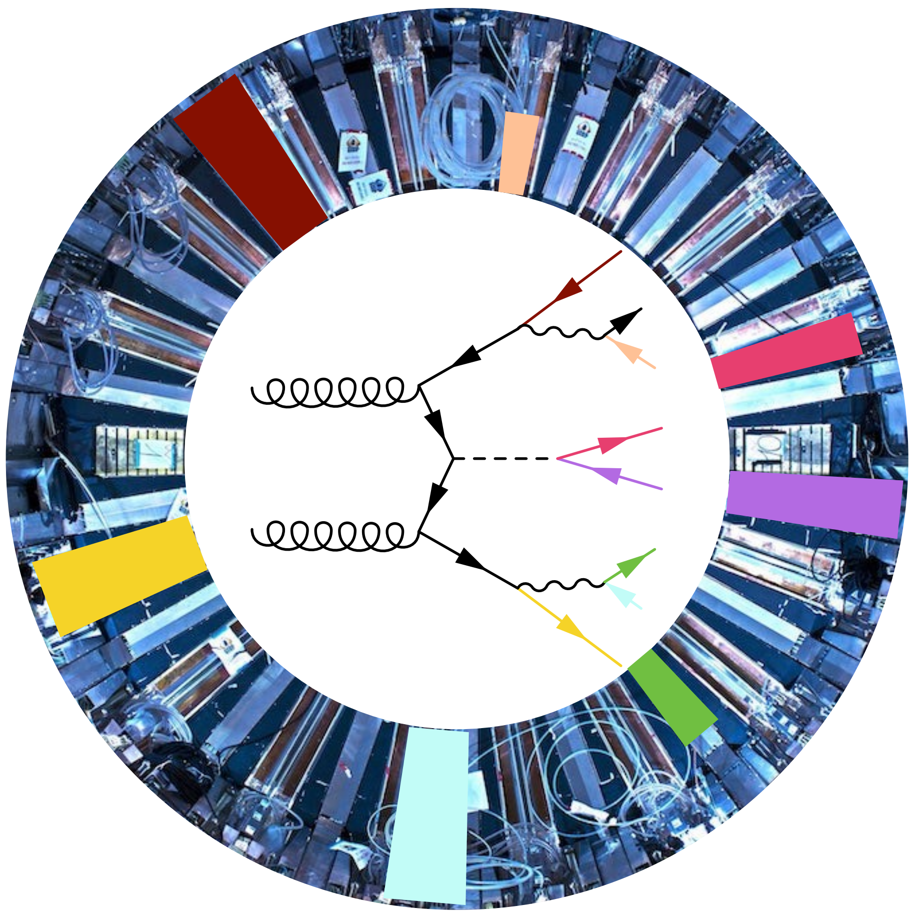
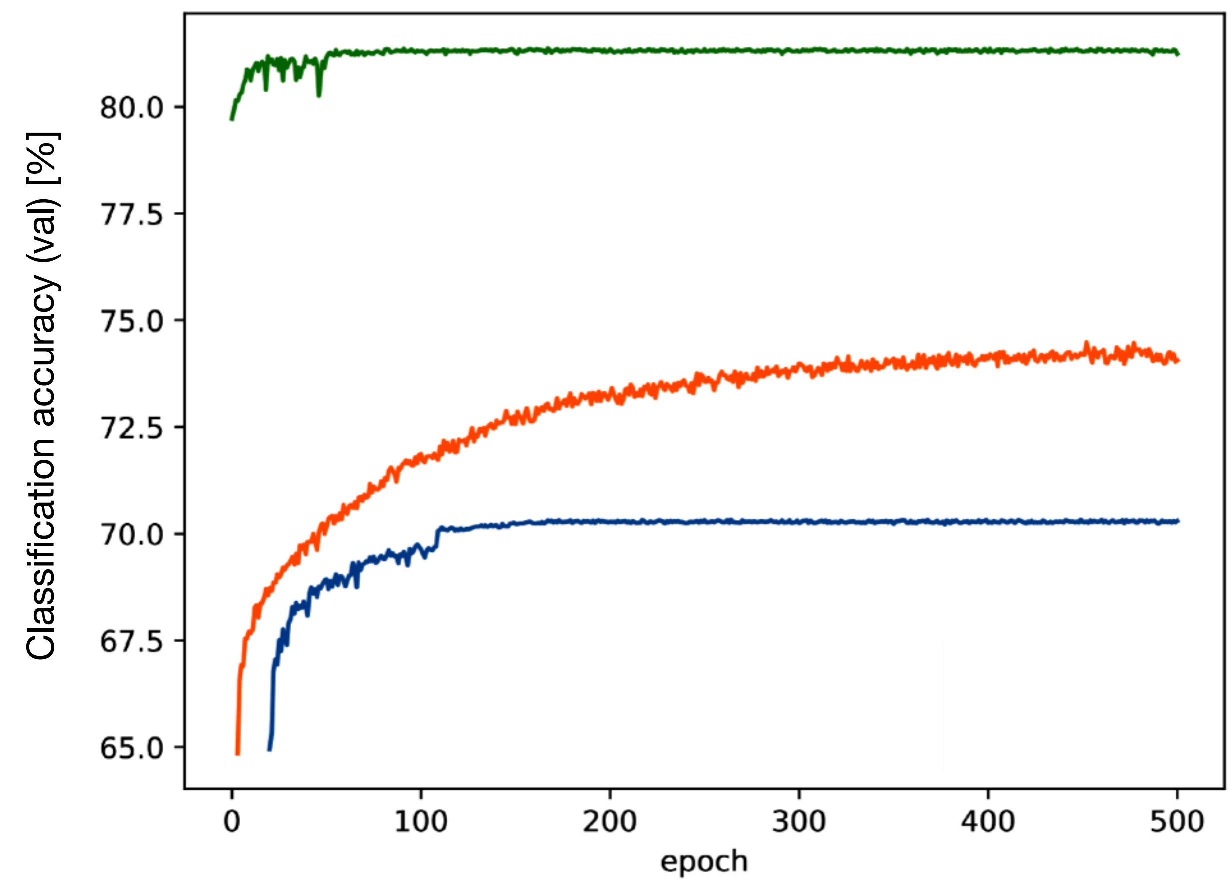




Supervised and Reinforced Jet-Parton Assignment for Particle Physics Analyses

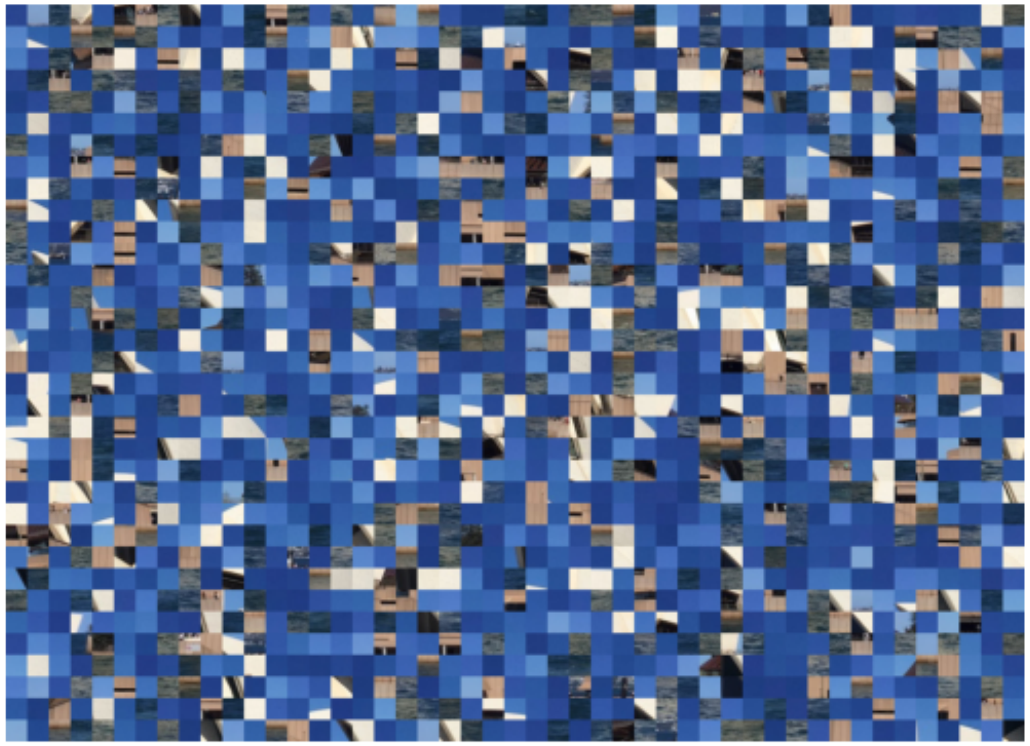
Martin Erdmann, Benjamin Fischer, Dennis Noll

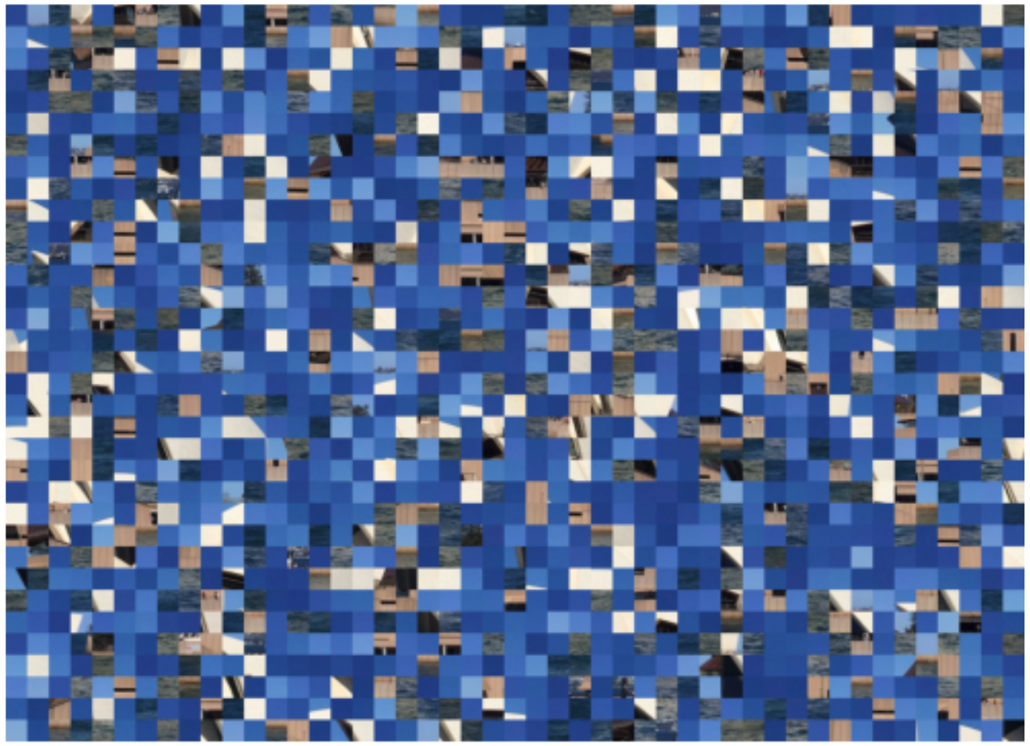
Conceptual Advances Workshop

13.09.22

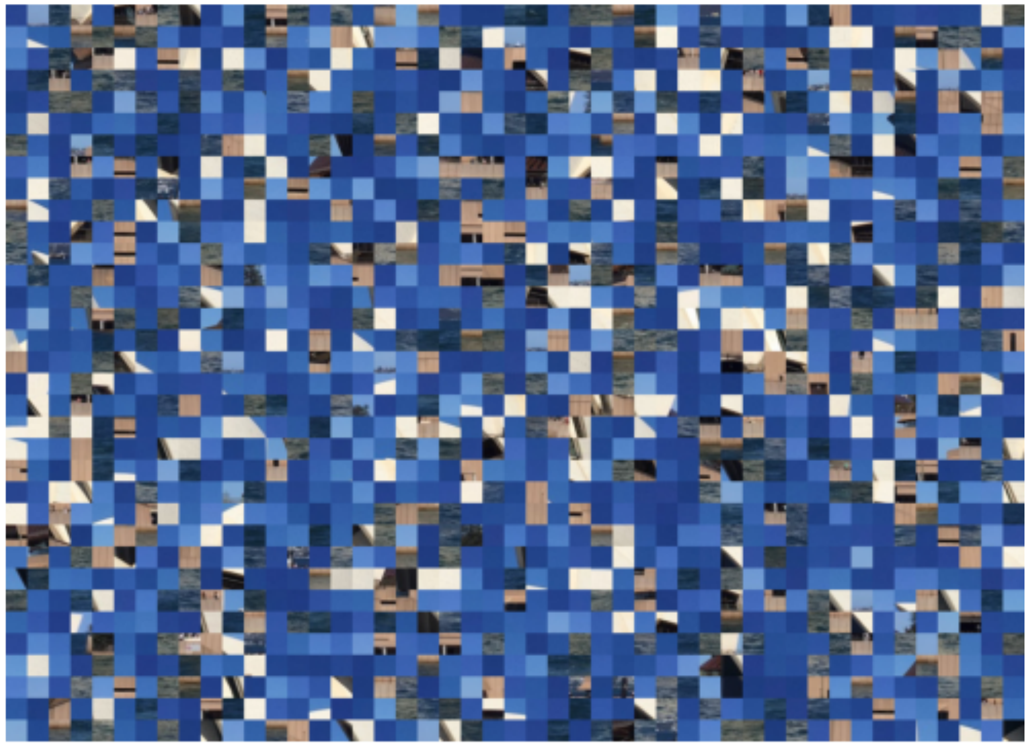
RWTH AACHEN UNIVERSITY

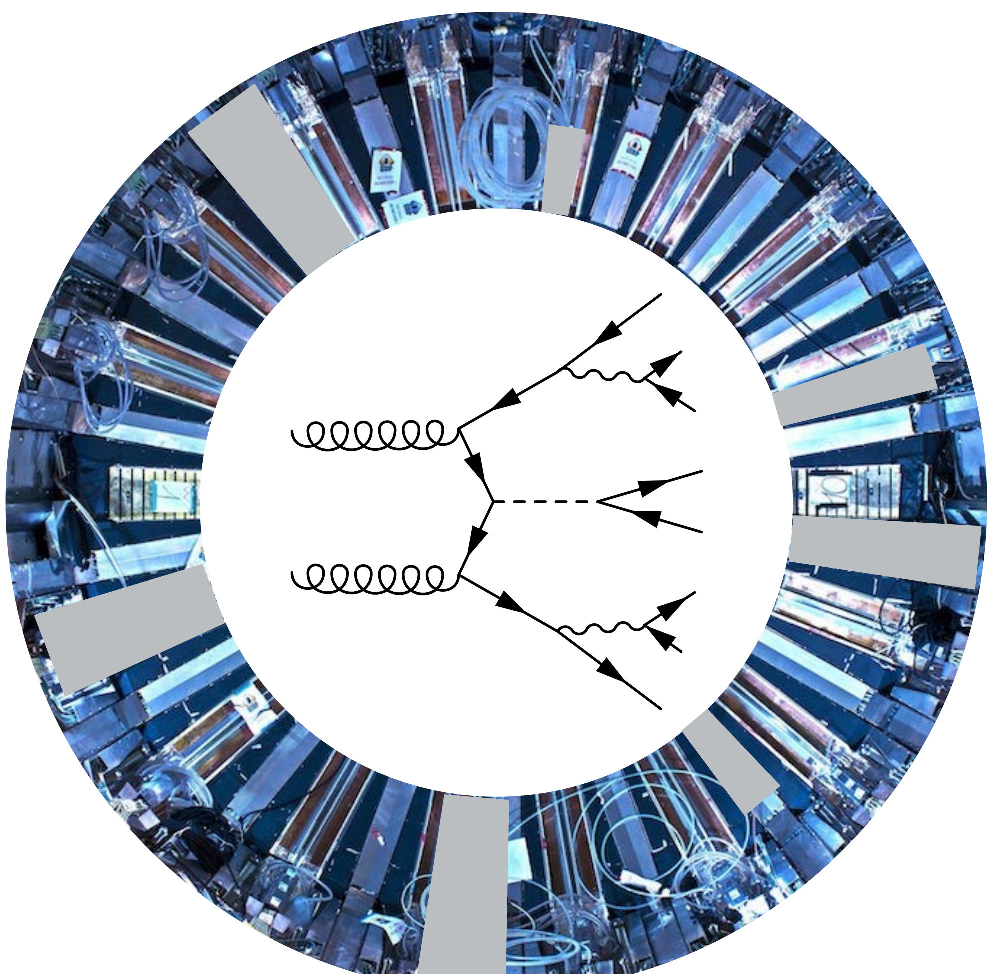
Motivation



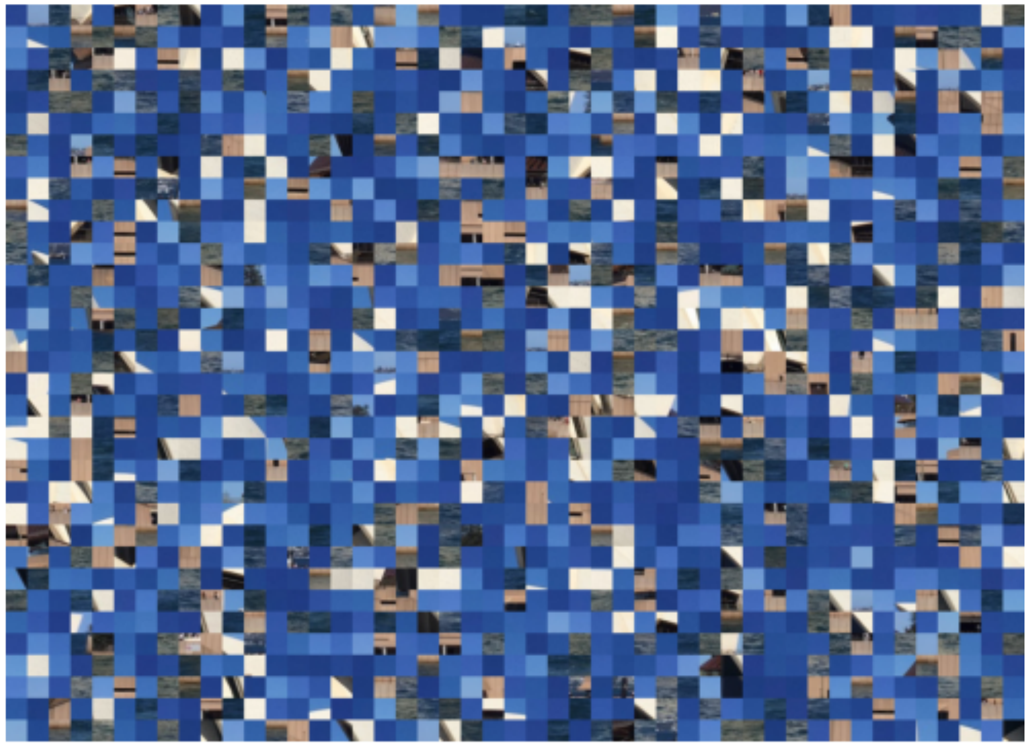


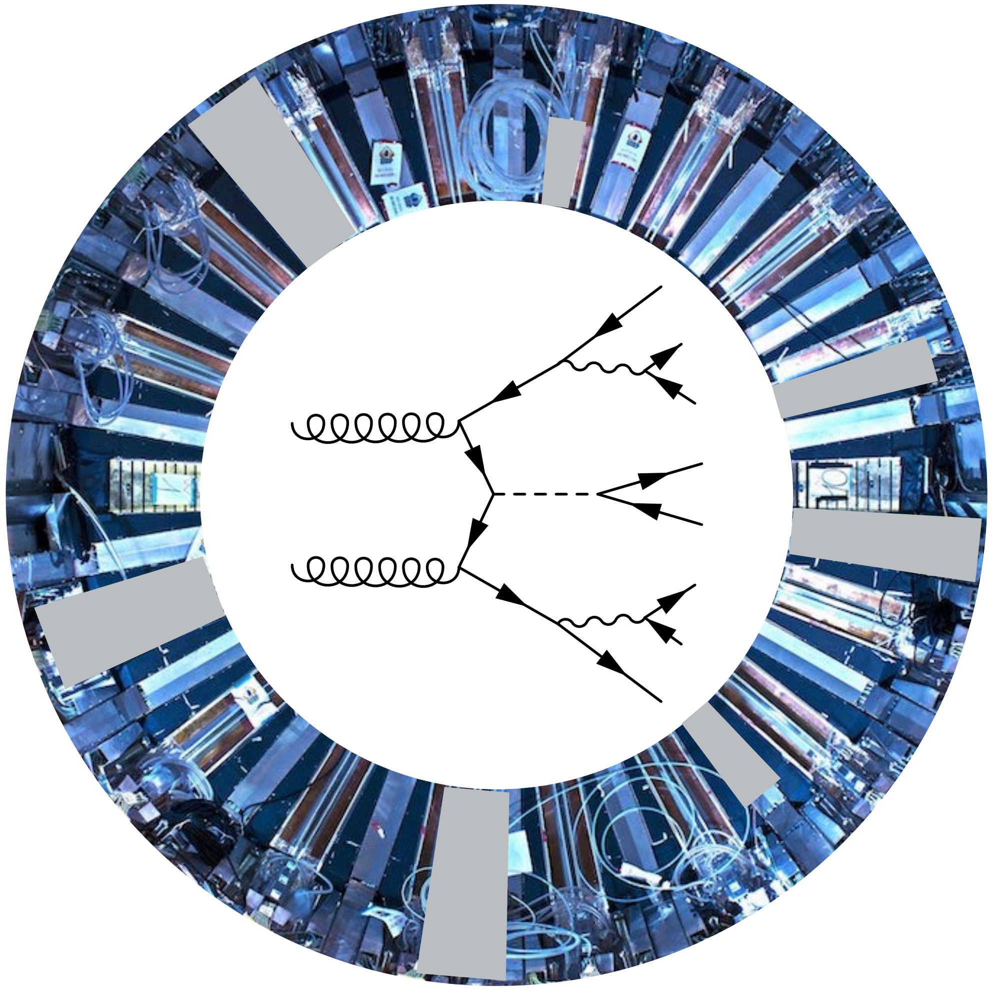
Motivation

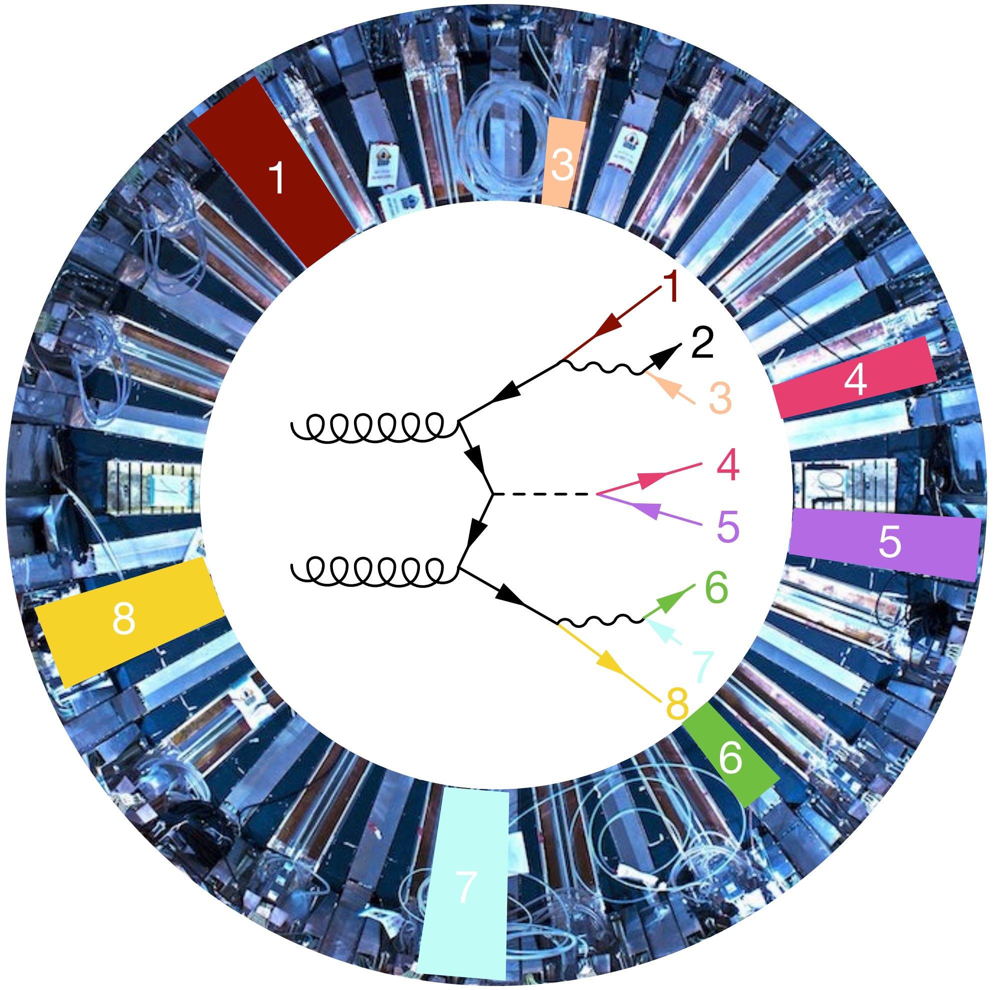




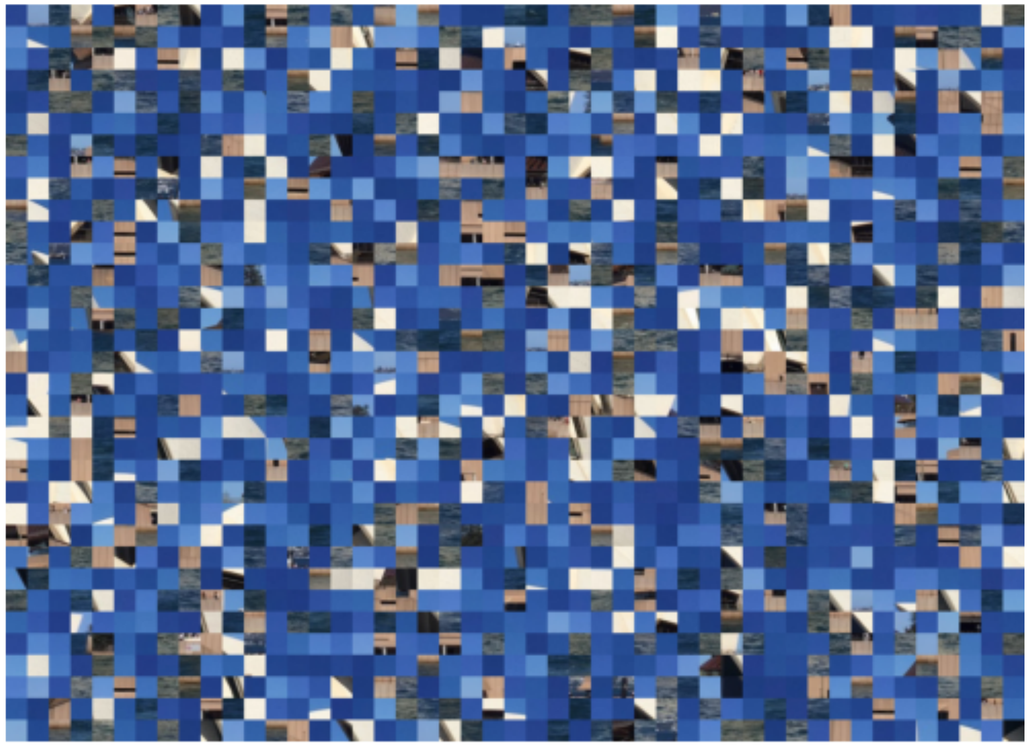
Motivation

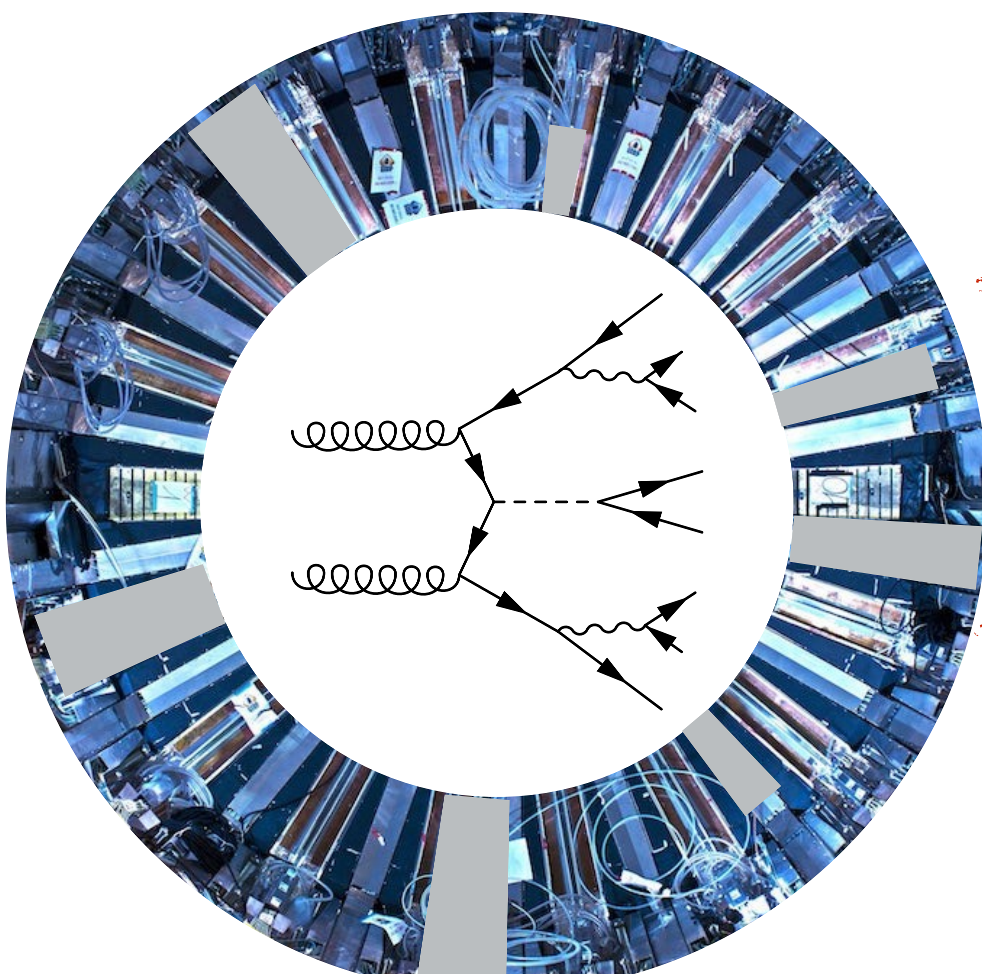






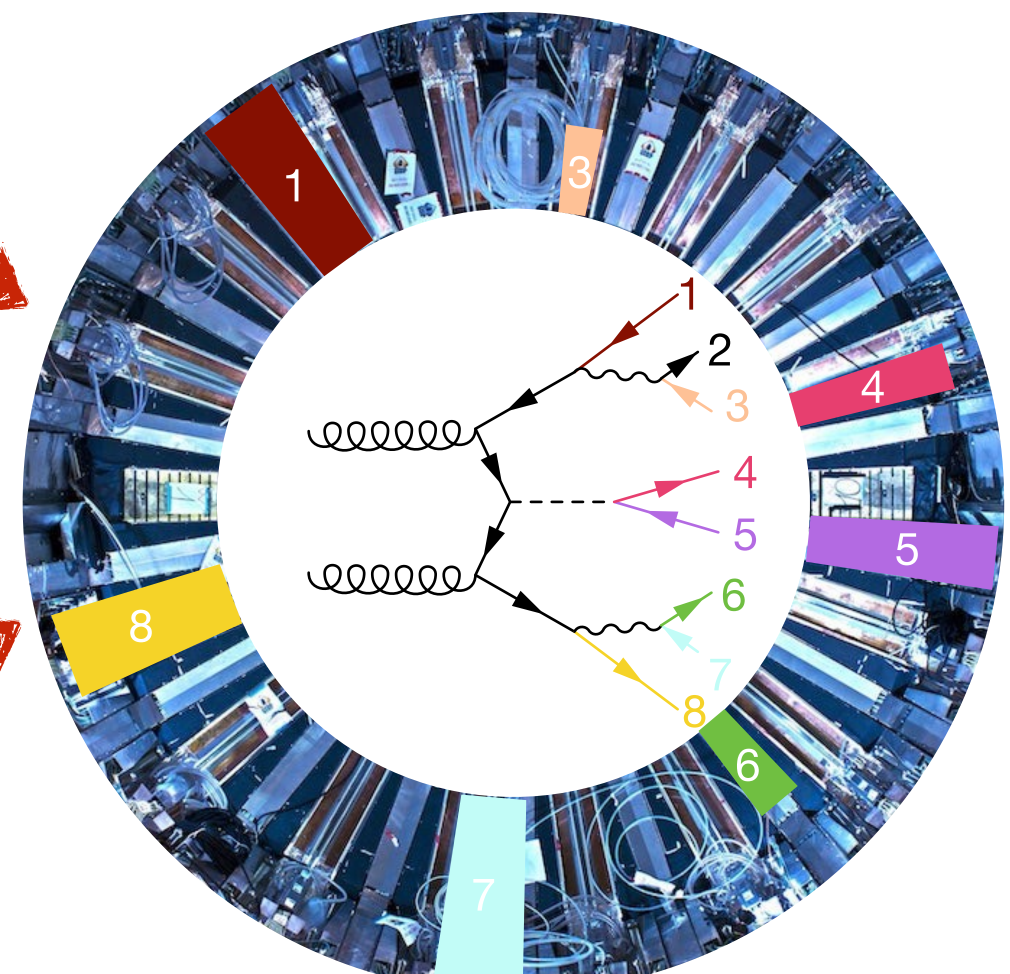
Motivation





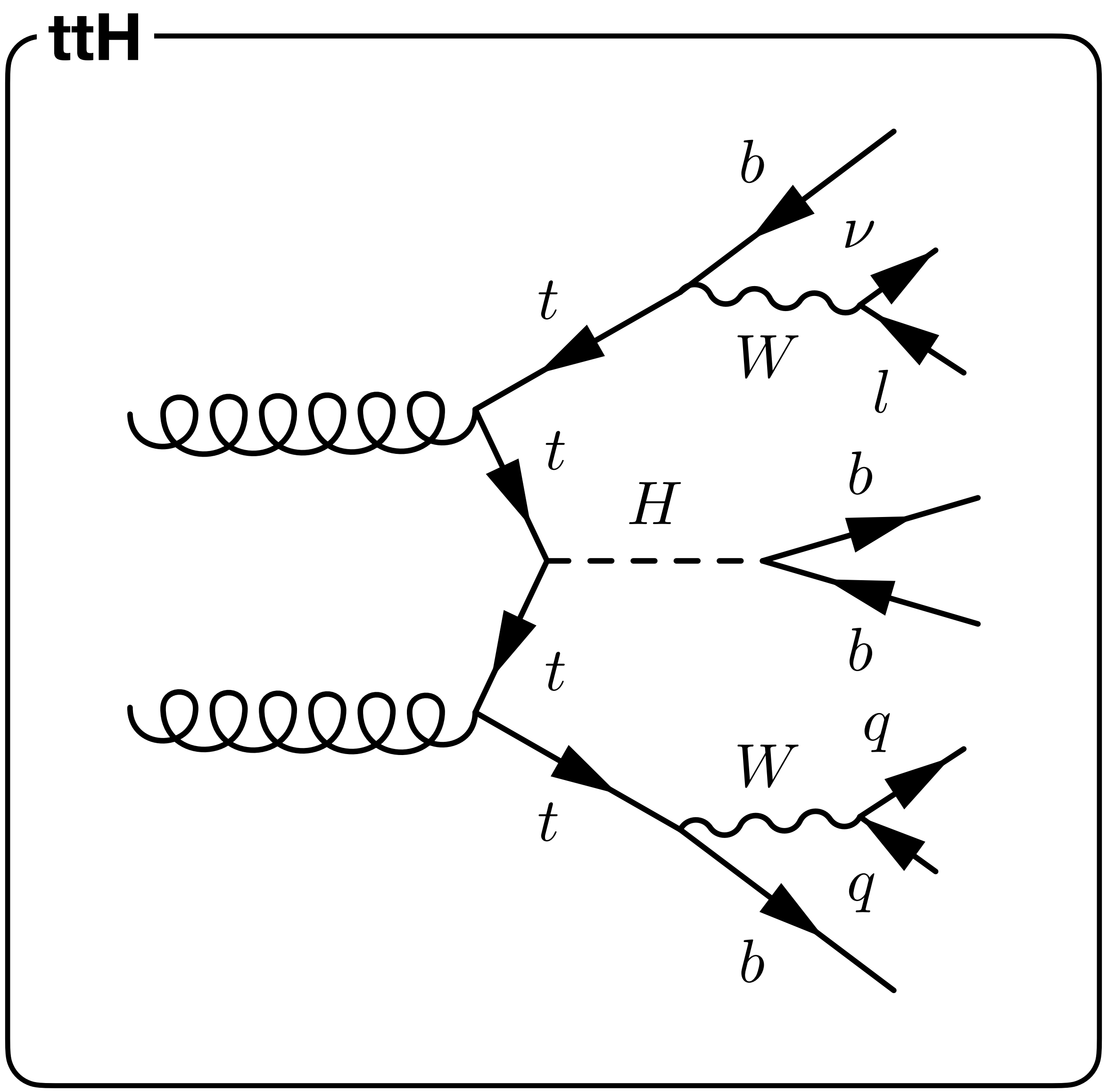
Supervised
(Slides 5-7)

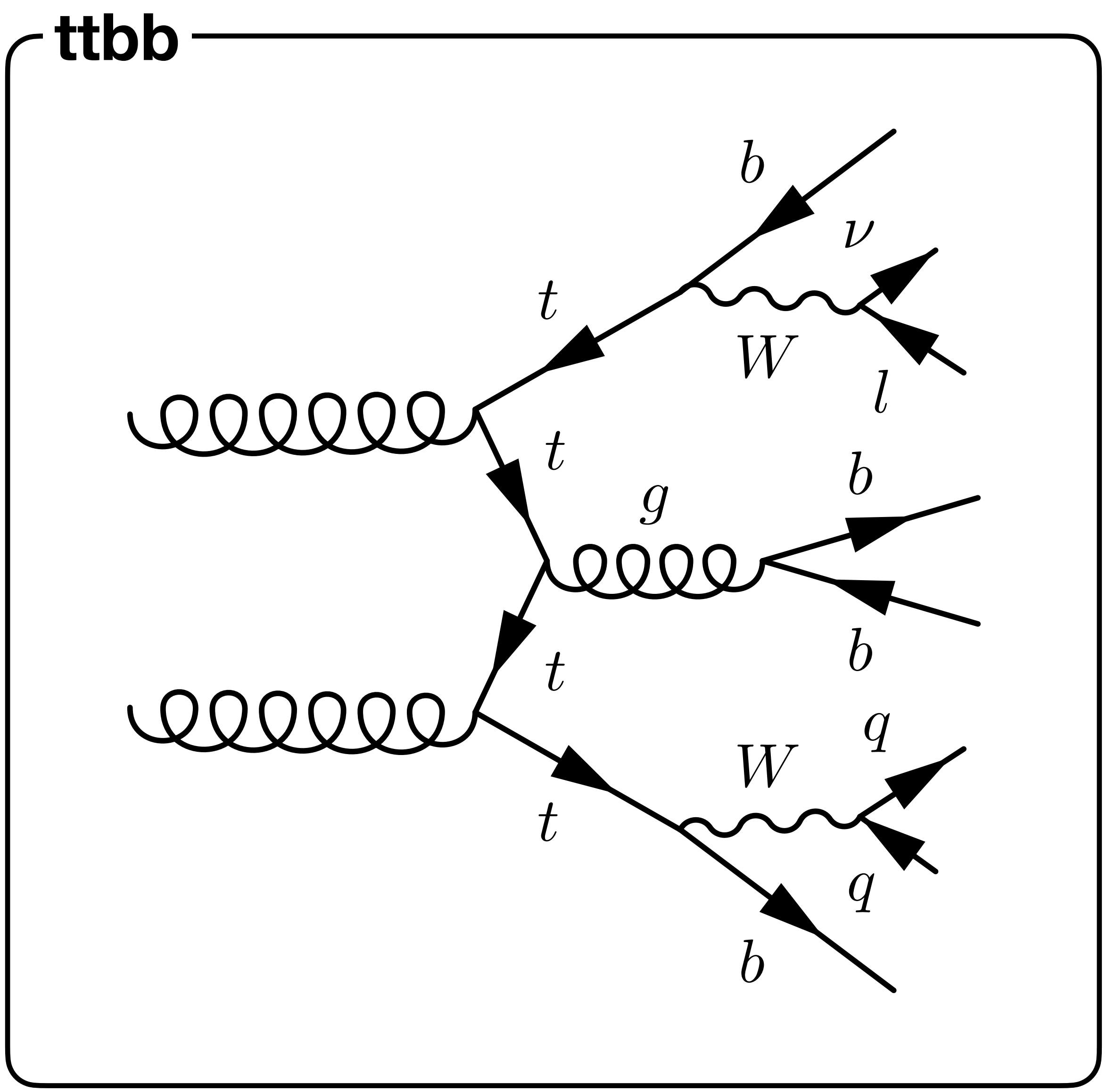
Reinforced
(Slides 8-12)



RWTH AACHEN UNIVERSITY
Dennis Noll - 13.09.22

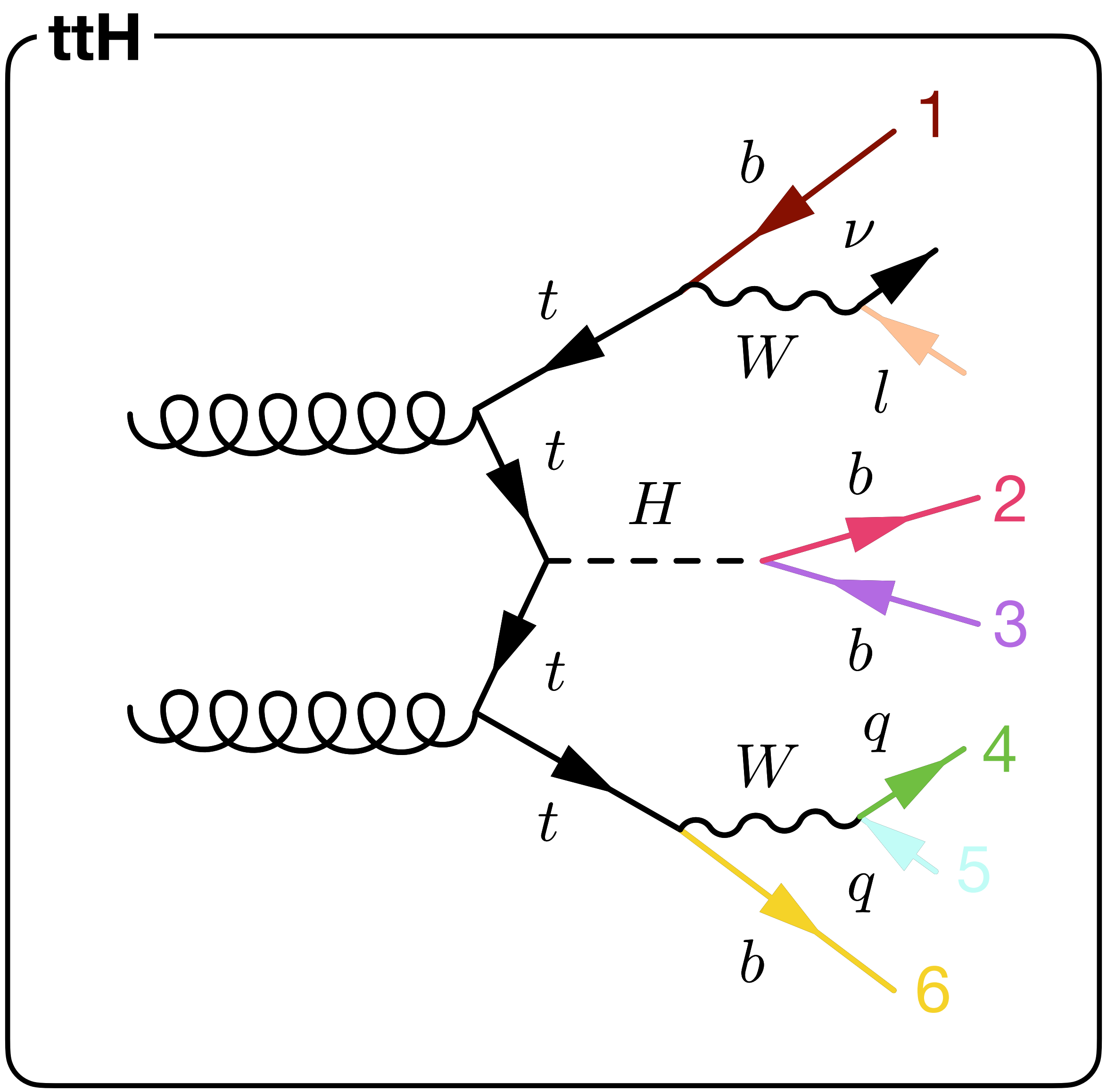
- Classification of ttH vs ttbb (1812.09722):
  - Two processes with same final state
  - Jet Parton Assignment (JPA) crucial:
    - Without JPA: **Complex**
    - With JPA: **Easy** (e.g. m_{bb} from slot 2 and 3)

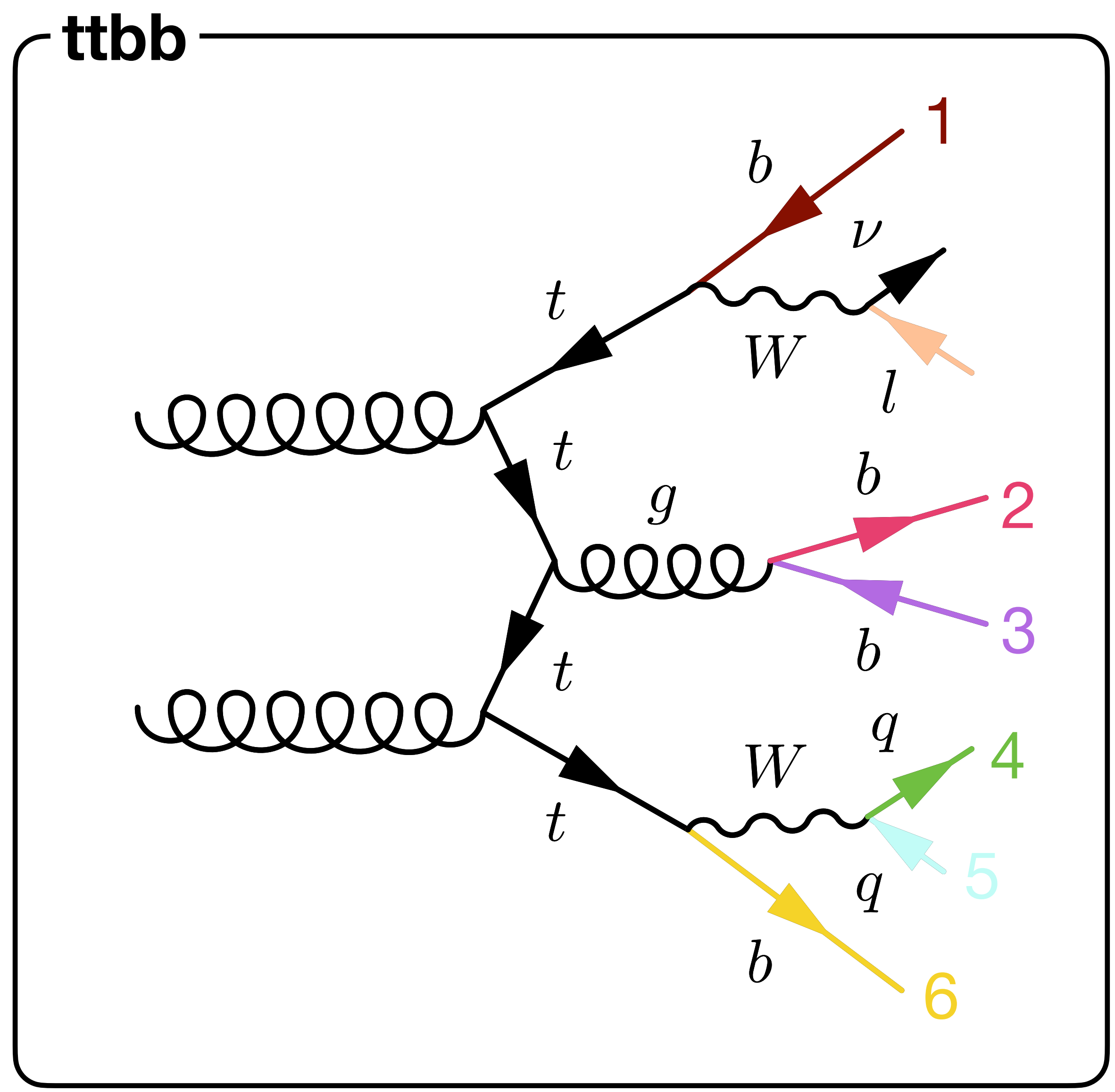




Physics Scenario

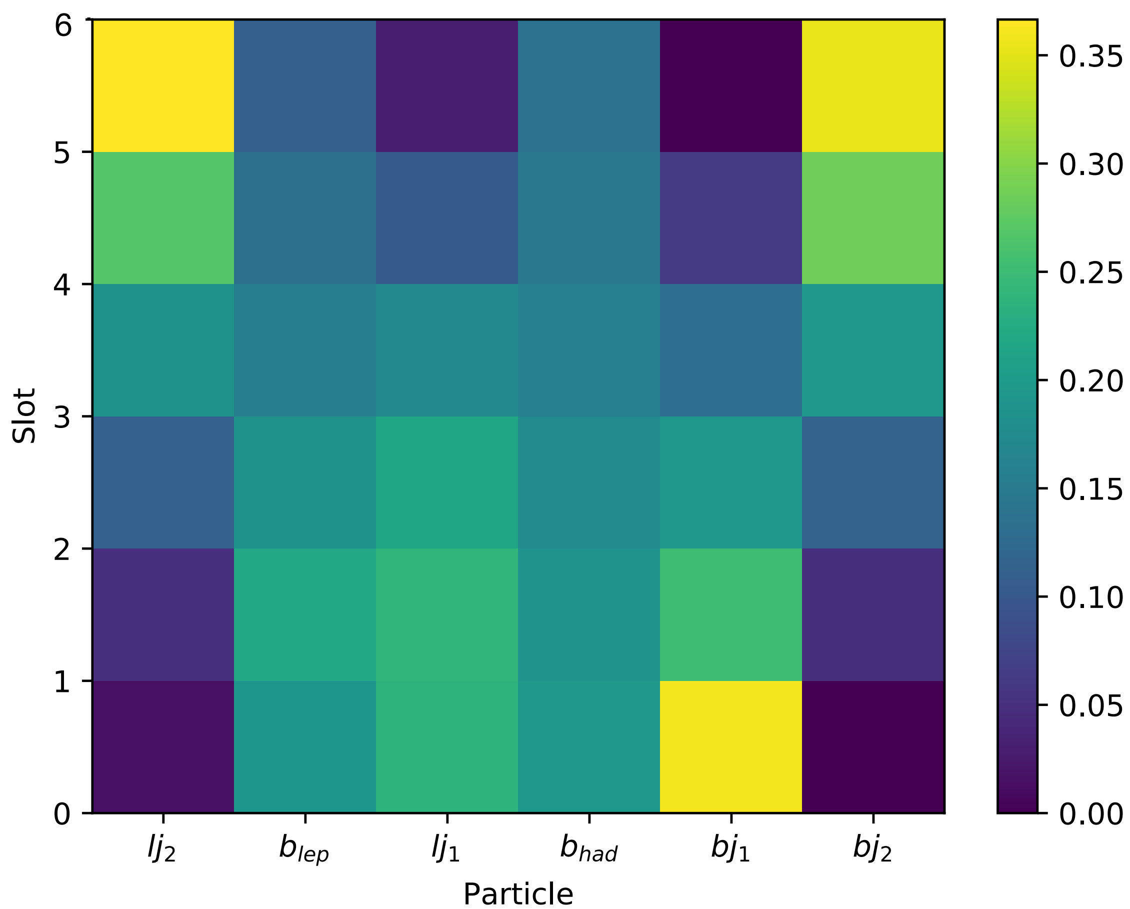
- Classification of ttH vs ttbb (1812.09722):
  - Two processes with same final state
  - Jet Parton Assignment (JPA) crucial:
    - Without JPA: **Complex**
    - With JPA: **Easy** (e.g. m_{bb} from slot 2 and 3)





Particle Based

- Assign with one variable (pT)
- Pro: Fast & works with many part.
- Con: High ambiguity



Permutation Methods

- Evaluate all different assignments
- Pro: Accurate
- Con: Scaling of permutations
  - (6! = 720; 10! > 3Mio)

Method	Probability of assigning all particles correctly
χ^2	37 %
DNN	52 %

Existing Approaches

Particle Based

- Assign with one variable (pT)
- Pro: Fast & works with many part.
- Con: High ambiguity

Permutation Methods

- Evaluate all different assignments
- Pro: Accurate
- Con: Scaling of permutations
  - (6! = 720; 10! > 3Mio)

Method	Probability of assigning all particles correctly
χ^2	37 %
DNN	52 %

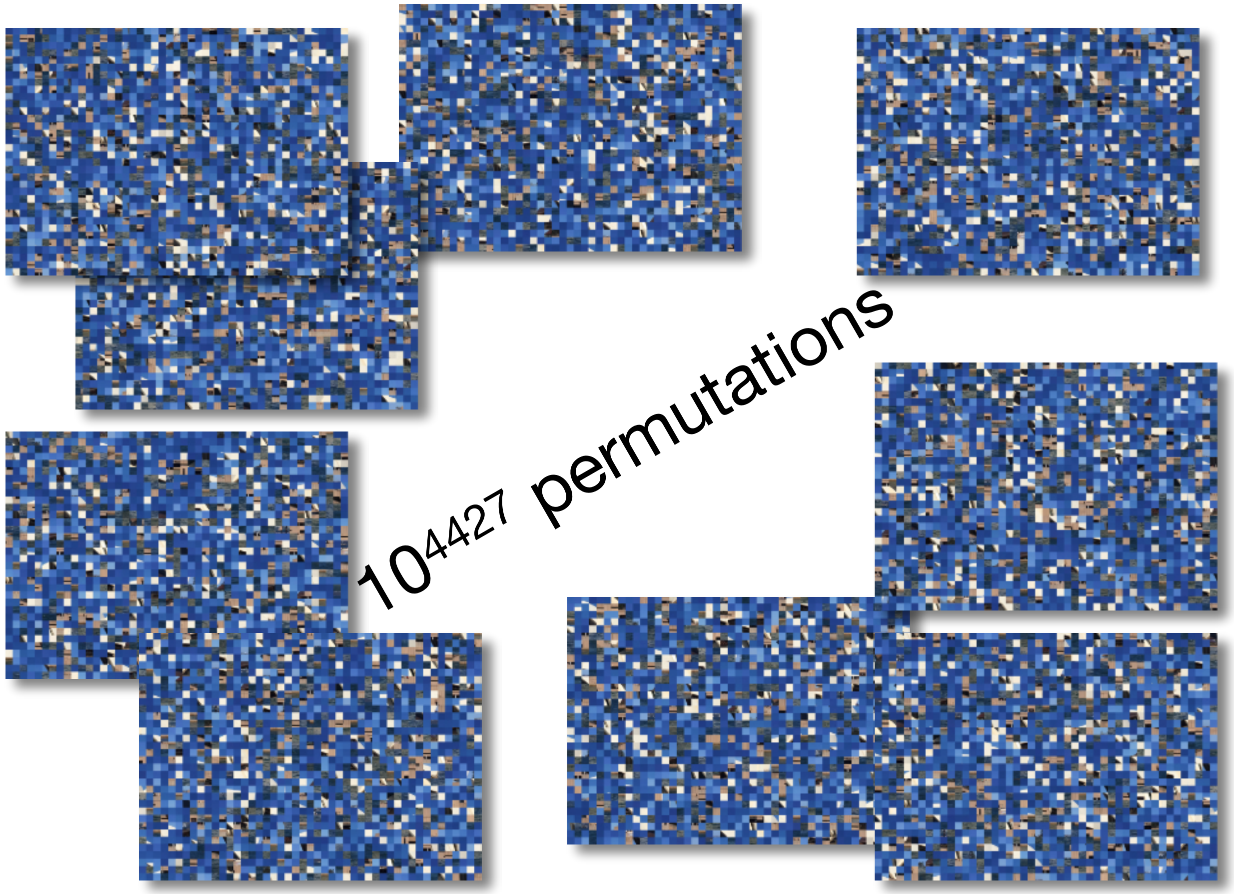
Existing Approaches

Particle Based

- Assign with one variable (pT)
- Pro: Fast & works with many part.
- Con: High ambiguity

Permutation Methods

- Evaluate all different assignments
- Pro: Accurate
- Con: Scaling of permutations

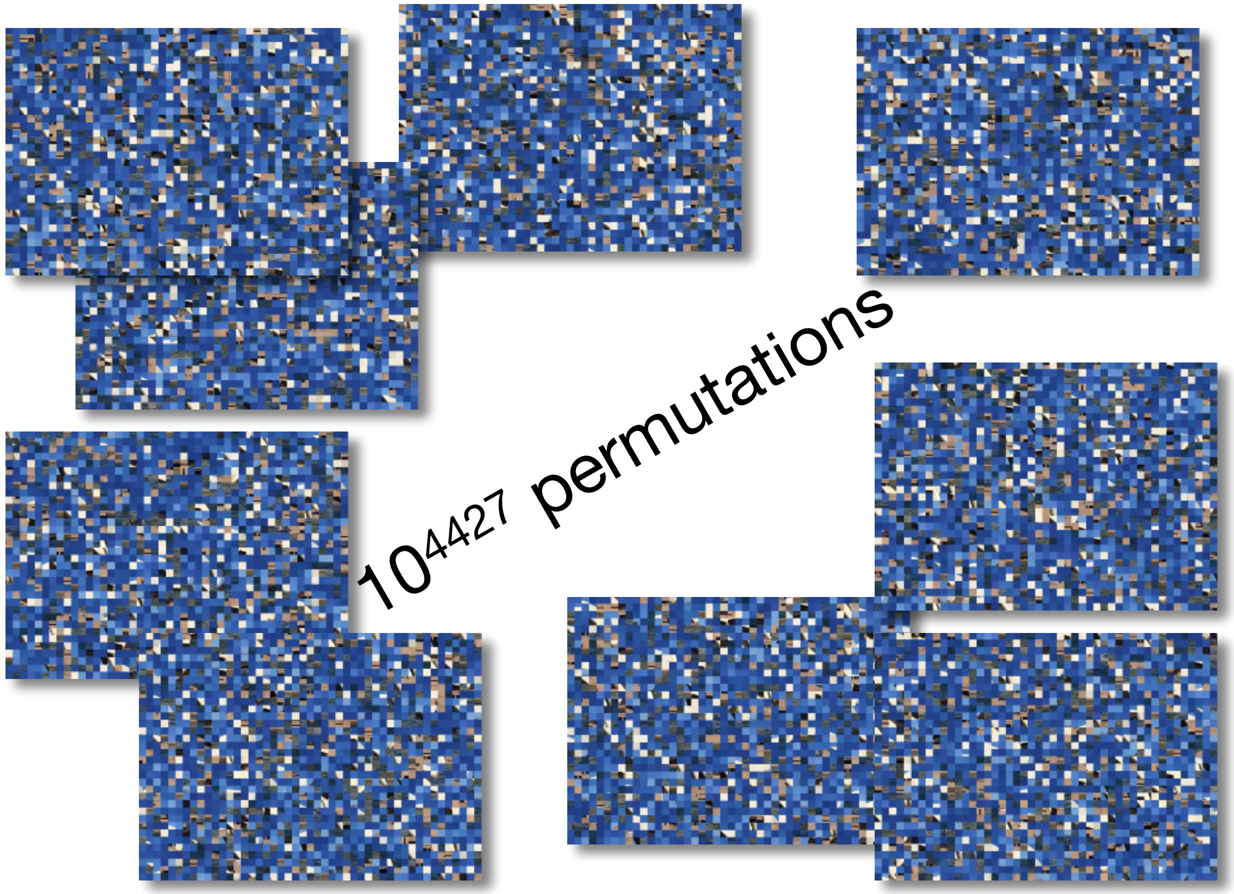


Particle Based

- Assign with one variable (pT)
- Pro: Fast & works with many part.
- Con: High ambiguity

Permutation Methods

- Evaluate all different assignments
- Pro: Accurate
- Con: Scaling of permutations

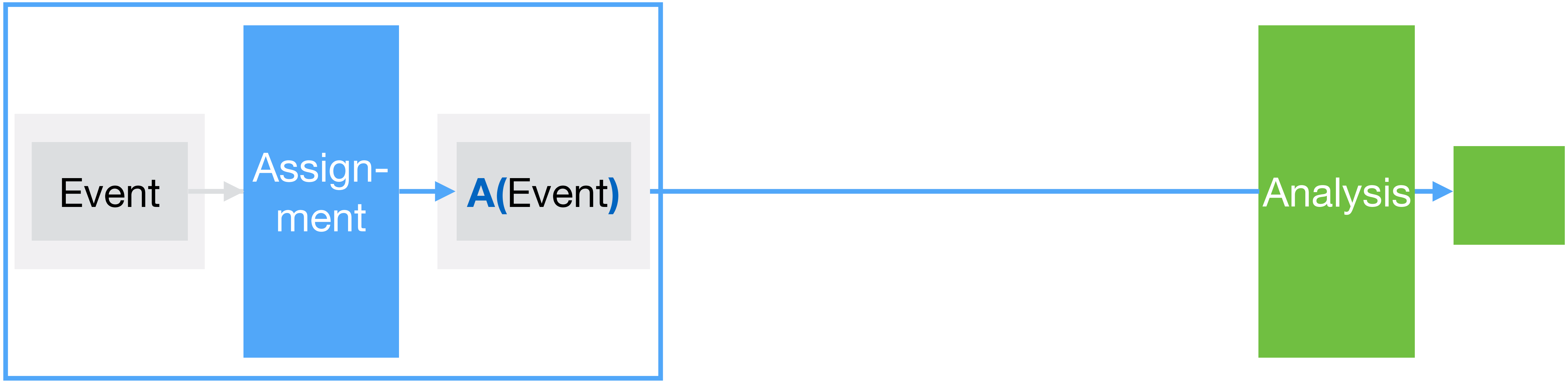


- Take the best out of the existing approaches:
  - Fast as particle based
  - Accurate as permutation method

A new approach to Jet-Parton Assignment

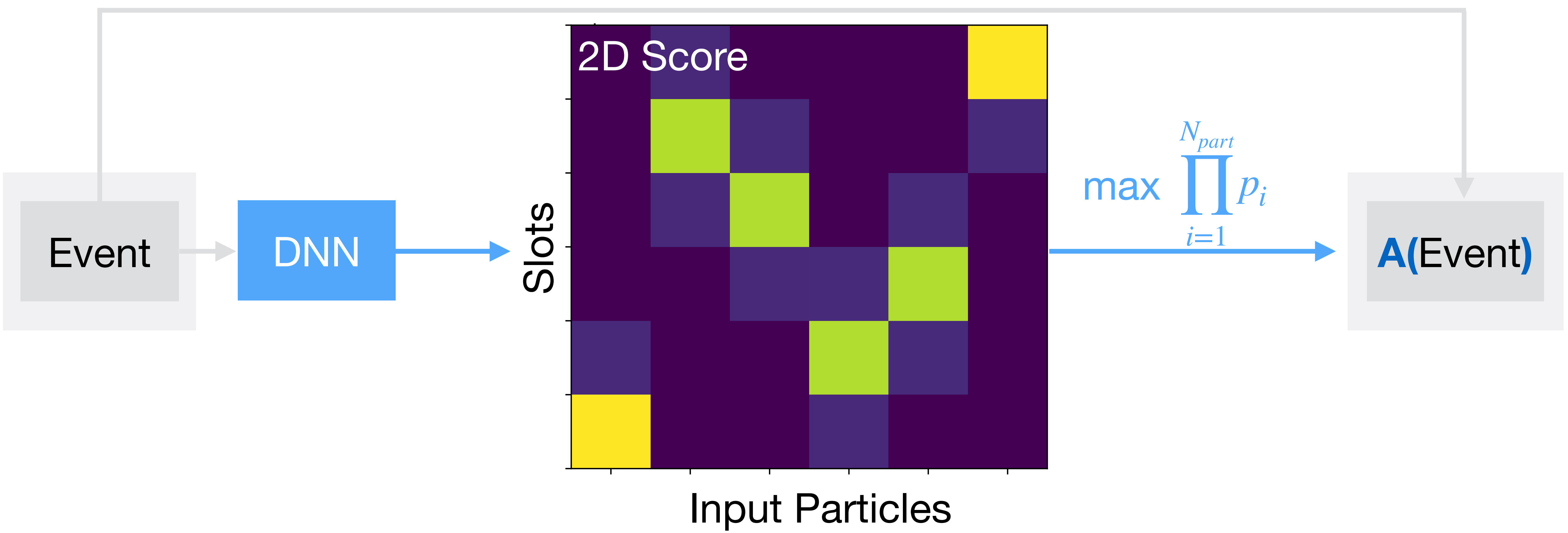
- Evaluation:
  - Input p_T sorted Event
  - Assignment network assigns particles A(Event) (trainable)
  - Analysis evaluates assigned events S(A(Event)) (trainable)

A new approach to Jet-Parton Assignment



- Evaluation:
  - Input p_T sorted Event
  - Assignment network assigns particles A(Event) (trainable)
  - Analysis evaluates assigned events S(A(Event)) (trainable)

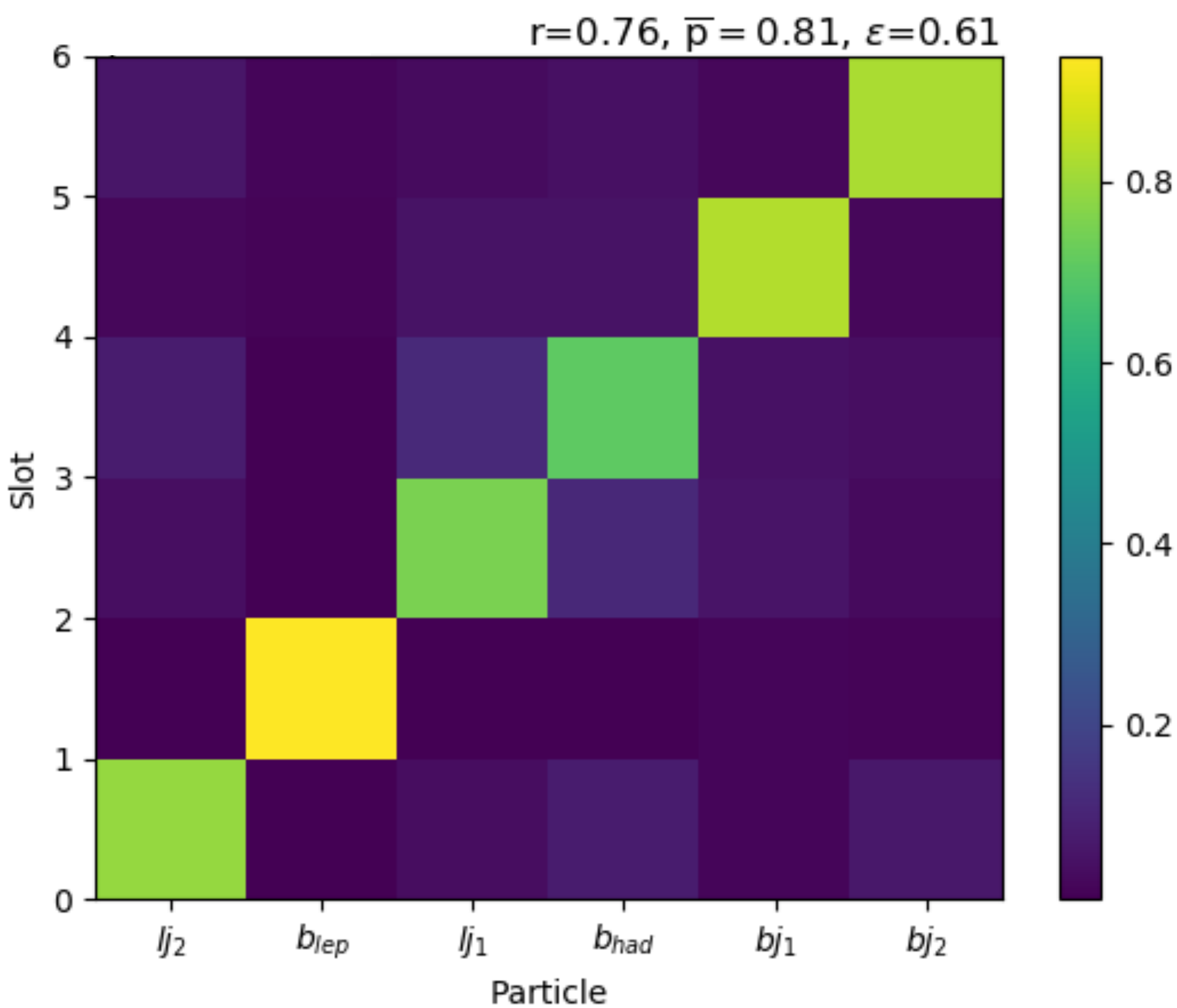
Assignment - Representation



- Multi-classification DNN:
  - Produces 2D Score
  - Bipartite graph:
    - Particle ↔ Slot
- Assignment:
  - Maximise joint score
  - Uses Munkres algorithm

Assignment - Training with generator information

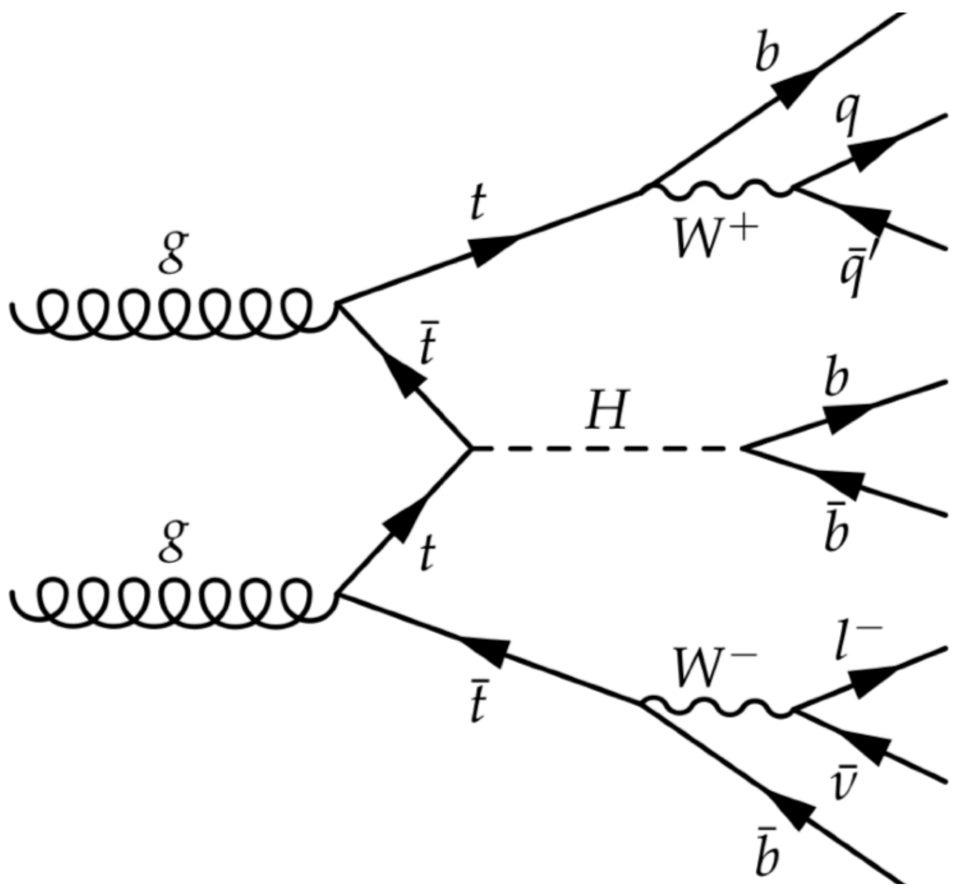
- Supervised training with fixed Particle ↔ Slot
- Correct assignment can be learned
- Probability to correctly identify complete ttH final state is $\epsilon = 61\,\%$
- Better than state-of-the-art parton assignment methods ($\epsilon = 52\,\%$, 1706.01117)



Reinforcement Training - Motivation

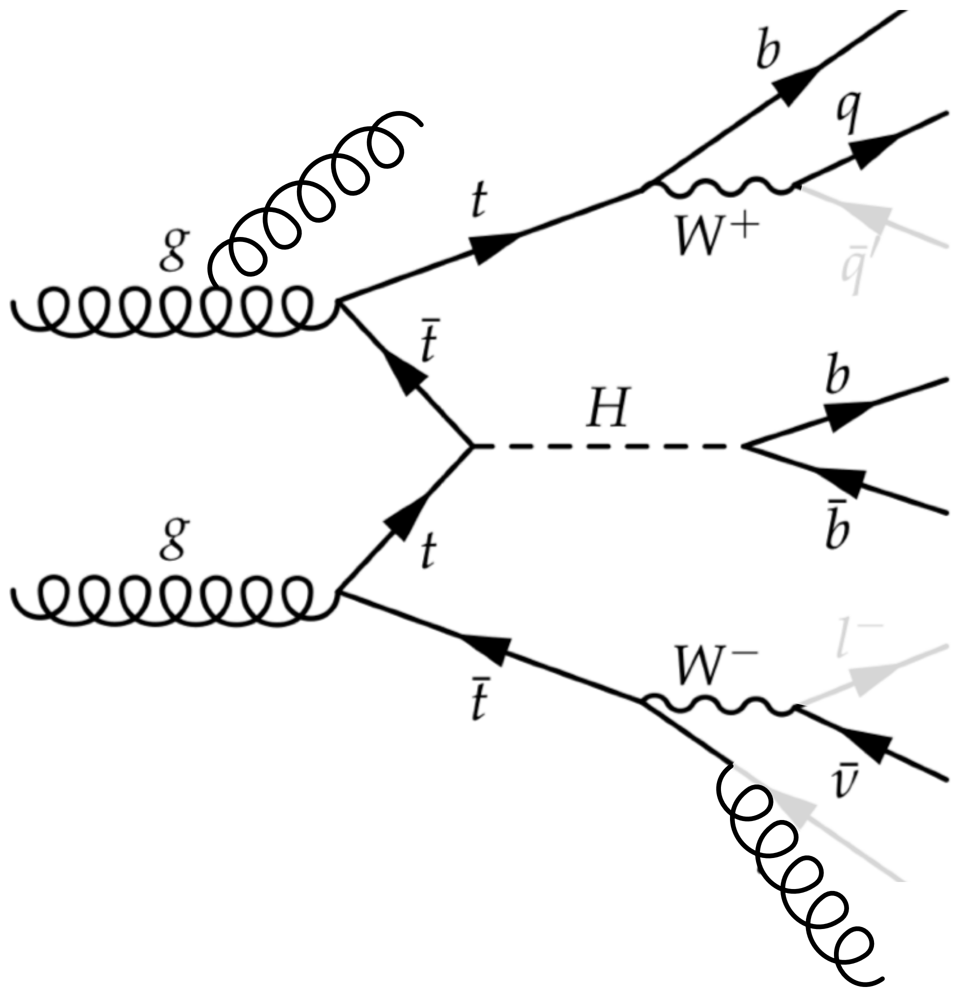
Textbook example

- Particle nature fully known
- One slot for one particle
- <u>Supervised training works</u>



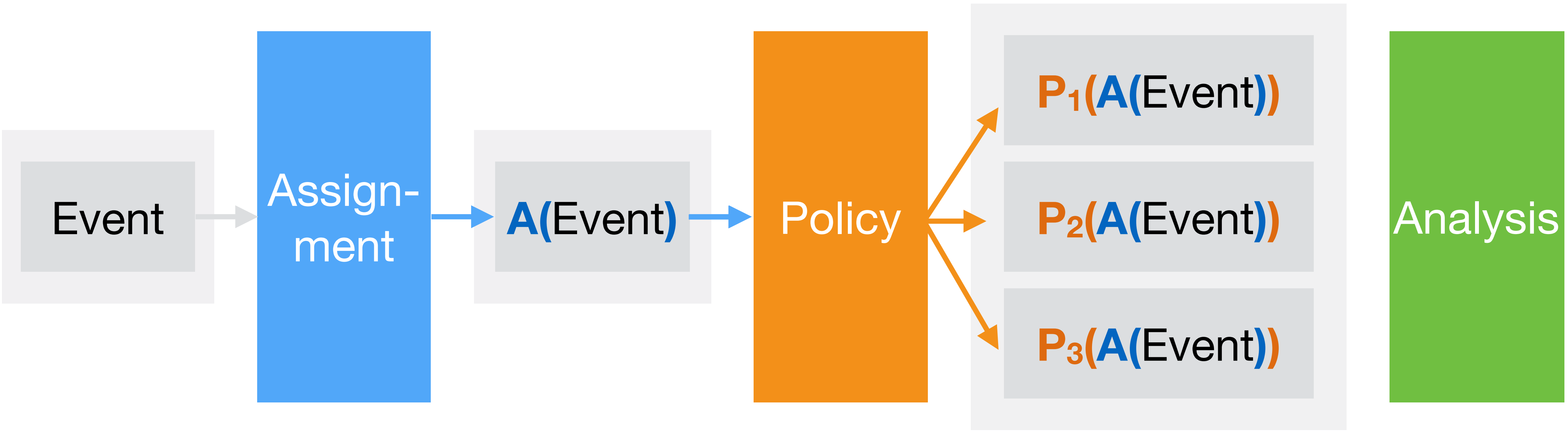
Real World

- Particle nature often not known
- Different particles in every event
- <u>Need autonomy to adapt</u>



Reinforcement Training - Procedure

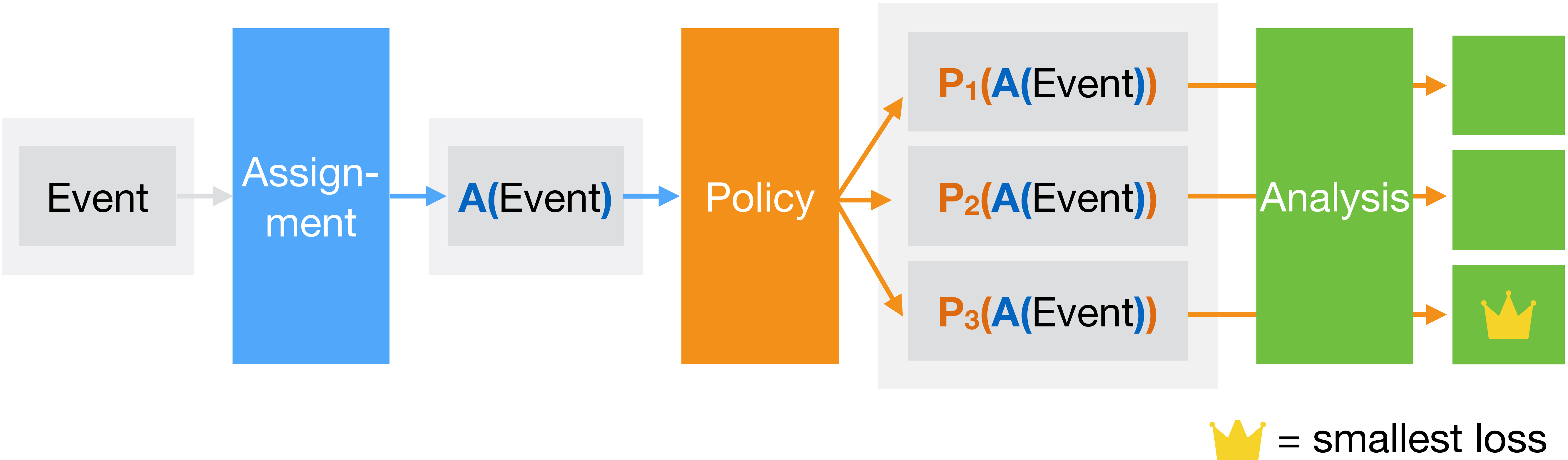
- Training objective: Find order of inputs, which is best suited for analysis
  - Assignment network assigns particles A(Event)
  - Policy suggests new orderings P_1(A(Event)), P_2(A(Event)), P_3(A(Event)), …
  - Analysis evaluates assigned (and permuted) events



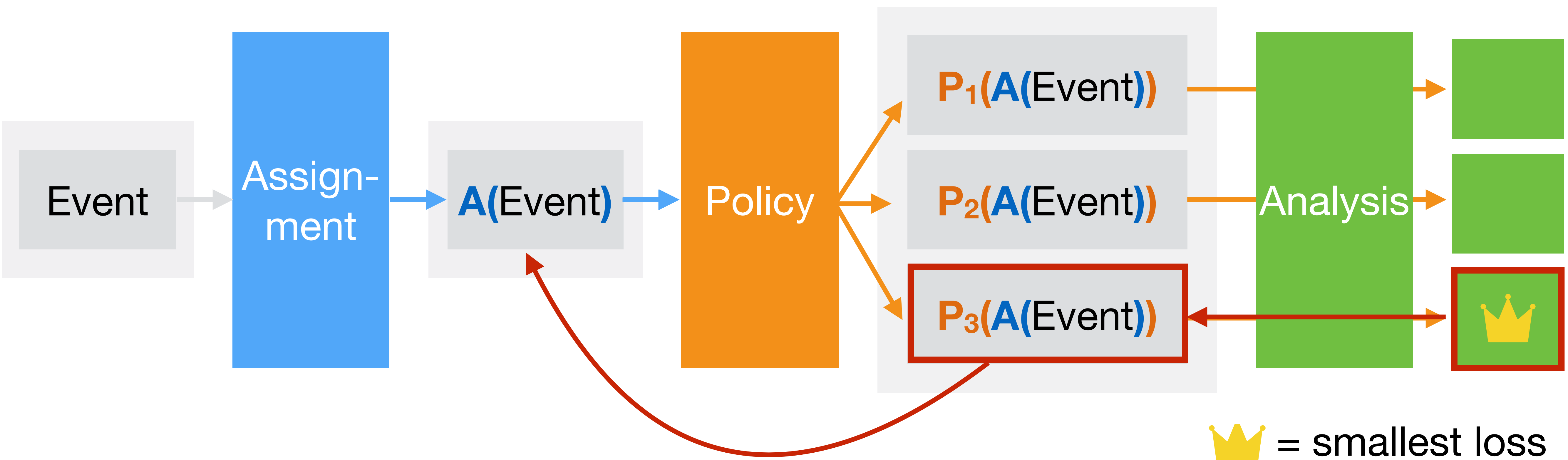
Reinforcement Training - Procedure

Dennis Noll - 13.09.22

- Training objective: Find order of inputs, which is best suited for analysis
  - Assignment network assigns particles A(Event)
  - Policy suggests new orderings $P_1(A(\text{Event}))$, $P_2(A(\text{Event}))$, $P_3(A(\text{Event}))$, …
  - Analysis evaluates assigned (and permuted) events



- Training objective: Find order of inputs, which is best suited for analysis
  - Assignment network assigns particles A(Event)
  - Policy suggests new orderings P_1(A(Event)), P_2(A(Event)), P_3(A(Event)), …
  - Analysis evaluates assigned (and permuted) events

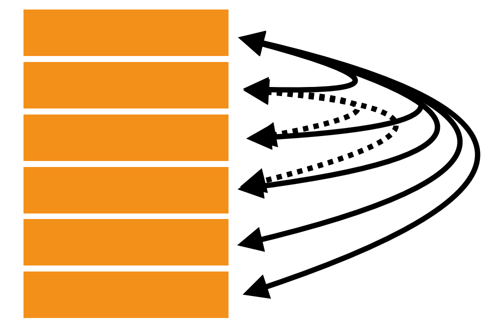


- Step 1: Train assignment A_{new} ← P_3(A) (Policy Gradients)
- Step 2: Train analysis with A_{new}(Event)

Reinforcement Training - Details

Permutation Policy

- Ensures that all permutations can be probed along the training
- All pair-wise swaps [$\mathcal{O}(n^2)$]:
  - $N_{pool}(6) = 15$

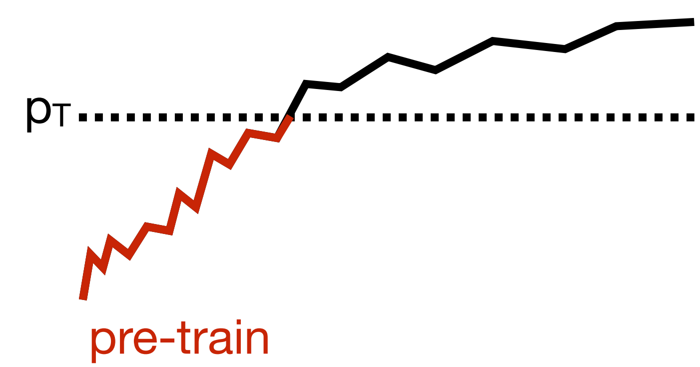


Training

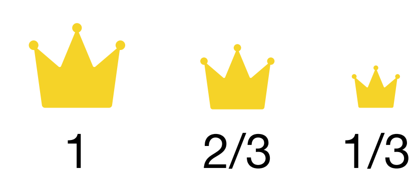
- Epoch wise Schedule:

1 Epoch Assignment
1 Epoch Analysis

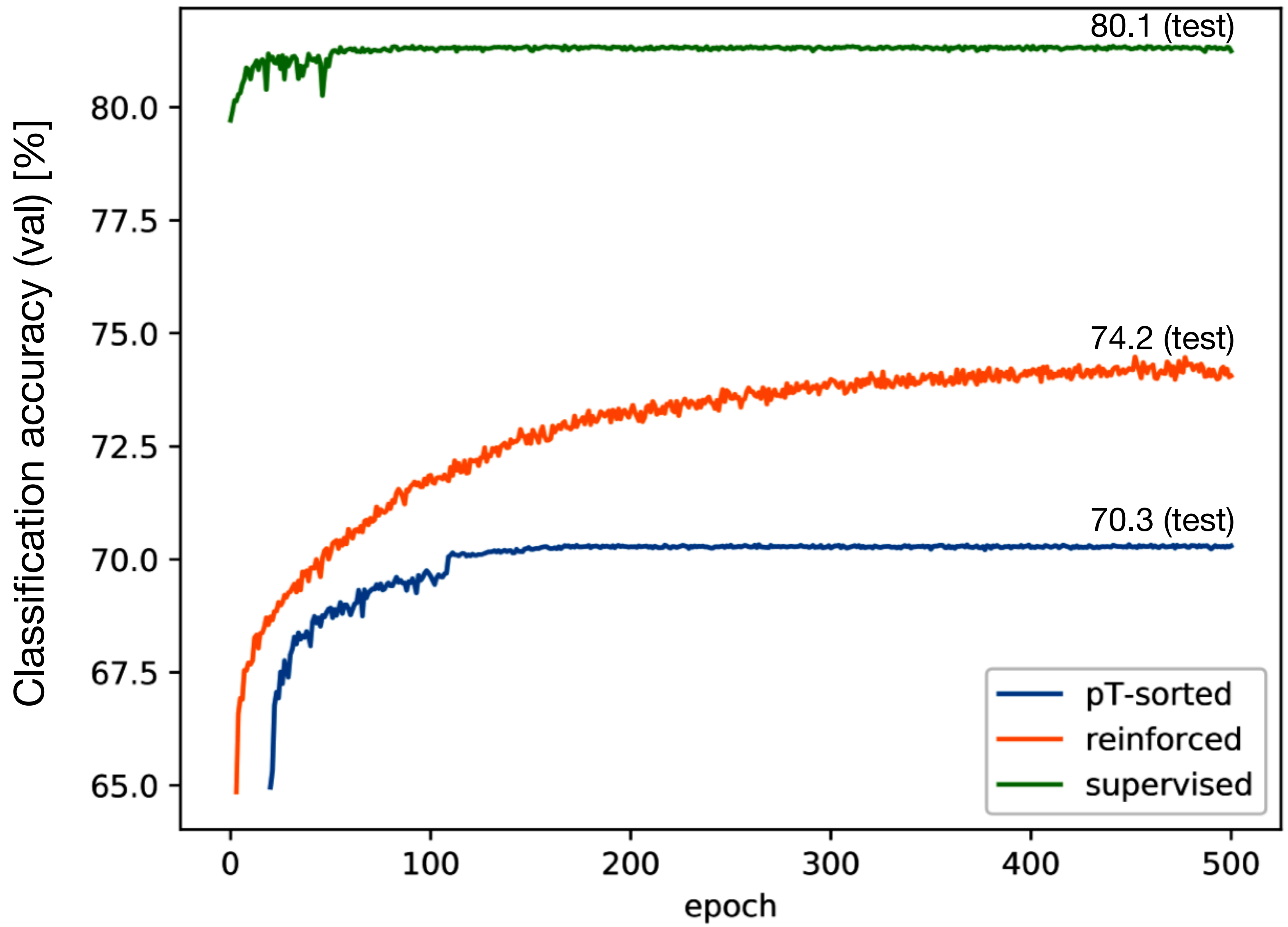
- Pre-training (p_T as baseline):
  - Done for assignment and analysis



- Use three best permutations (weighted)

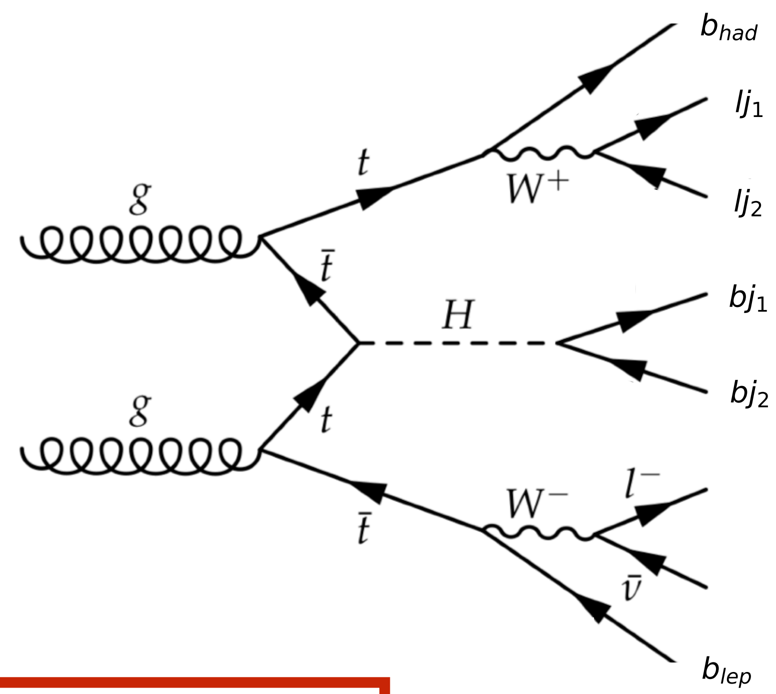


- Evaluate three methods for jet-parton assignment with same analysis network:
  - Standard: p_T - sorted
  - New (supervised): pre-trained **with** generator information
  - New (reinforced): autonomously trained **without** generator information

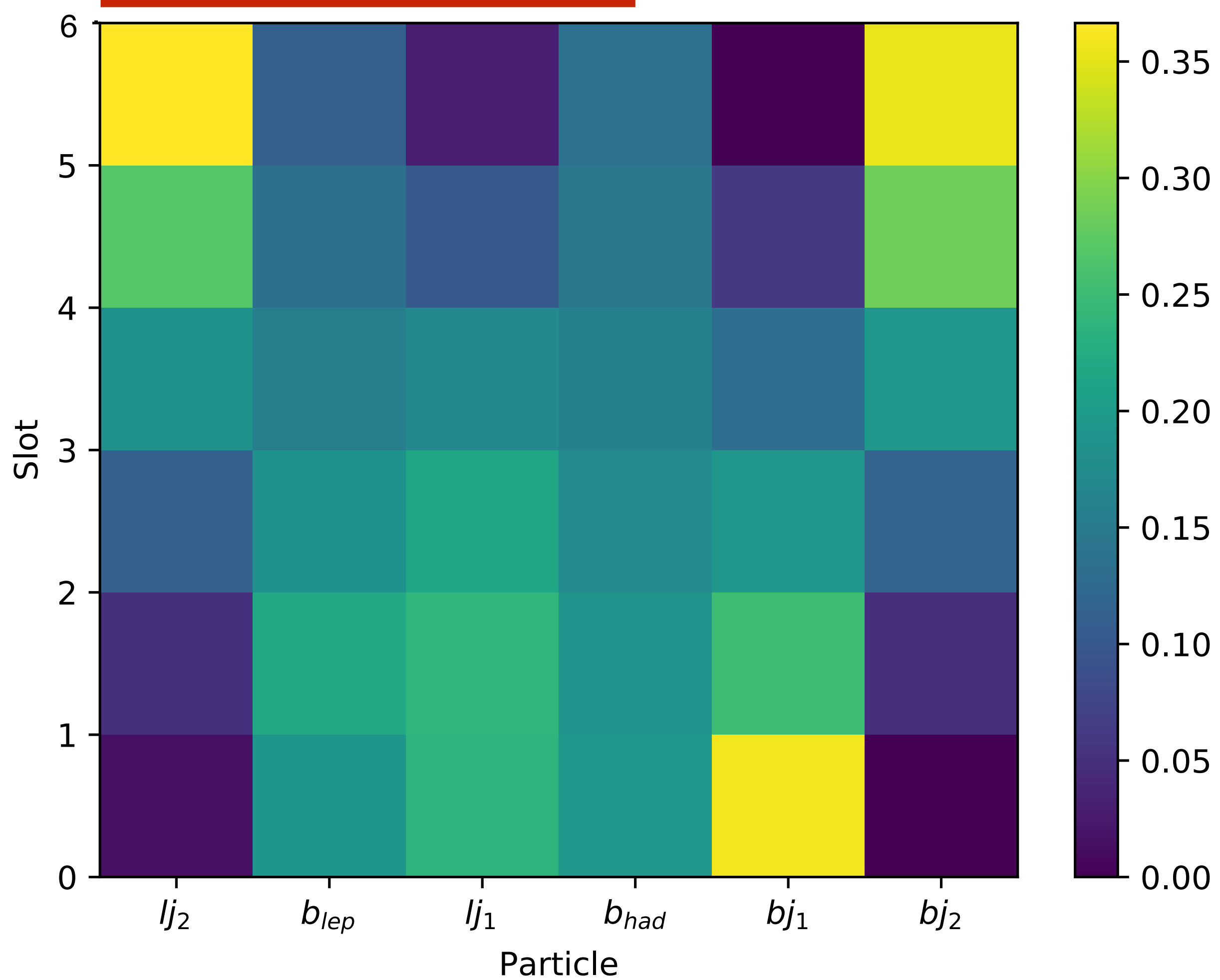


Looking into the assignment

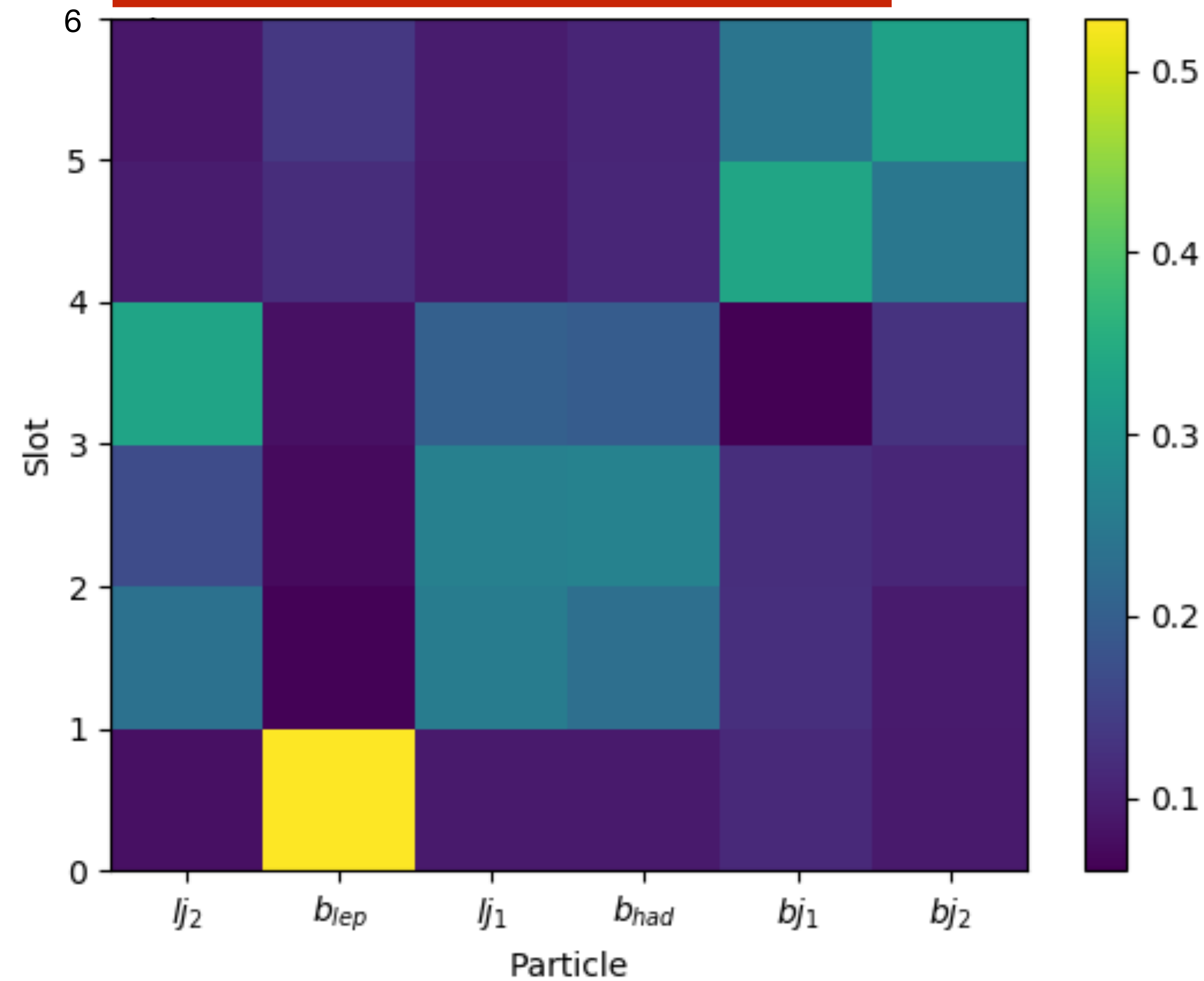
- Assignment network autonomously learns structure:
  - Direct neighbours grouped (e.g. bj_1, bj_2)
  - Do not confuse distinguishable particles (e.g. lj_2, bj_2)
  - Identifies branches of Feynman graph (e.g. b_{lep})



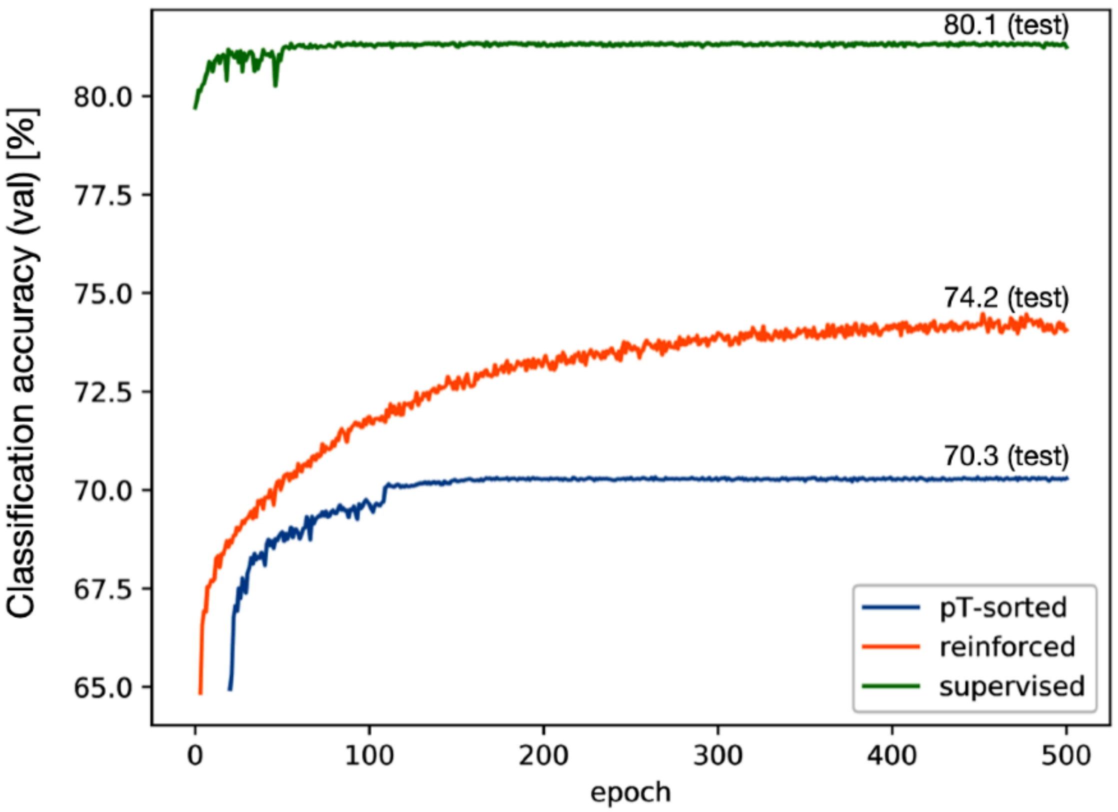
p_T - assignment

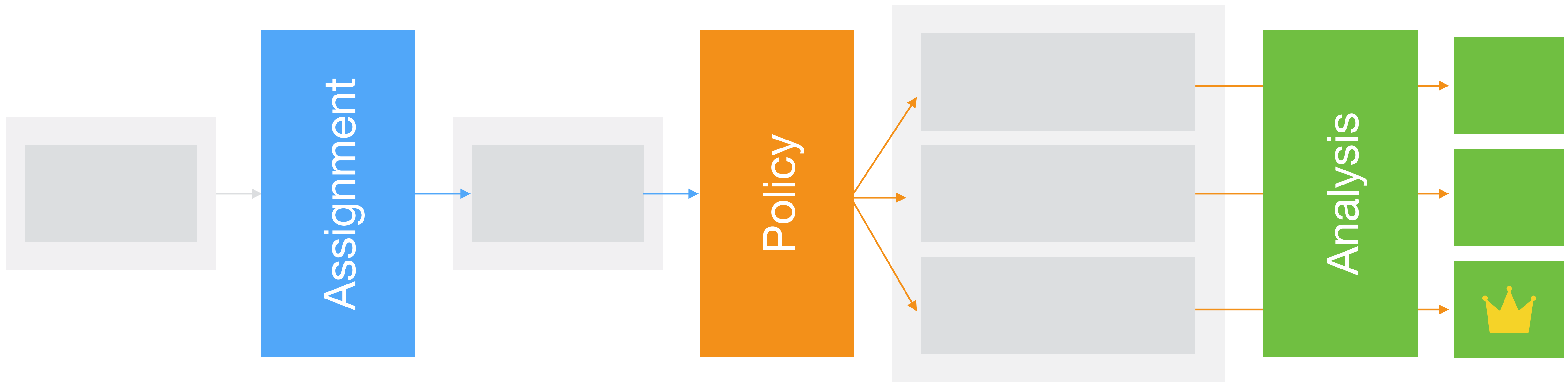


reinforced - assignment



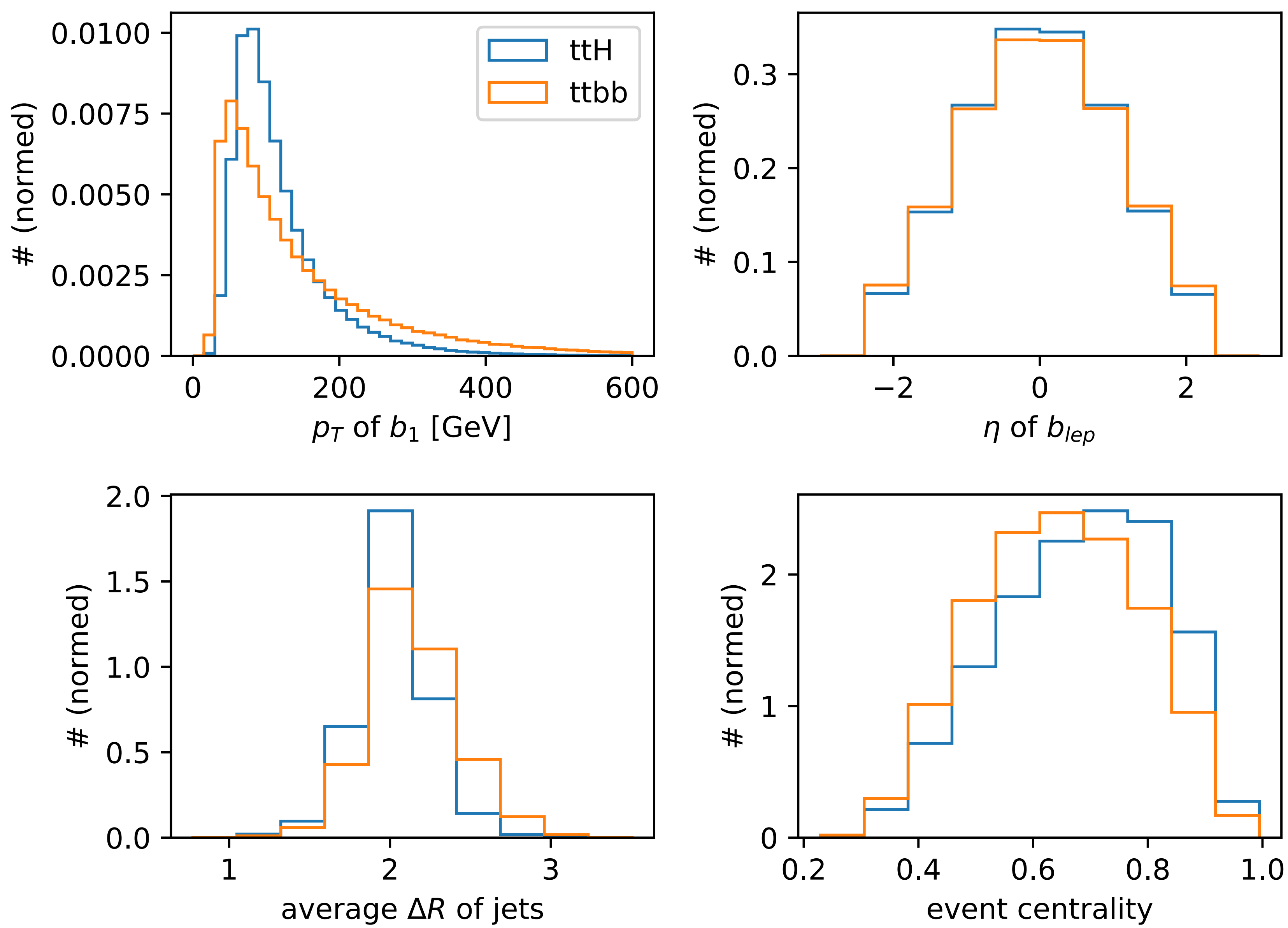
- A new method for jet-parton assignment
  - Supervised training if generator information is known
  - Reinforced training autonomously learns assignment based on the analysis





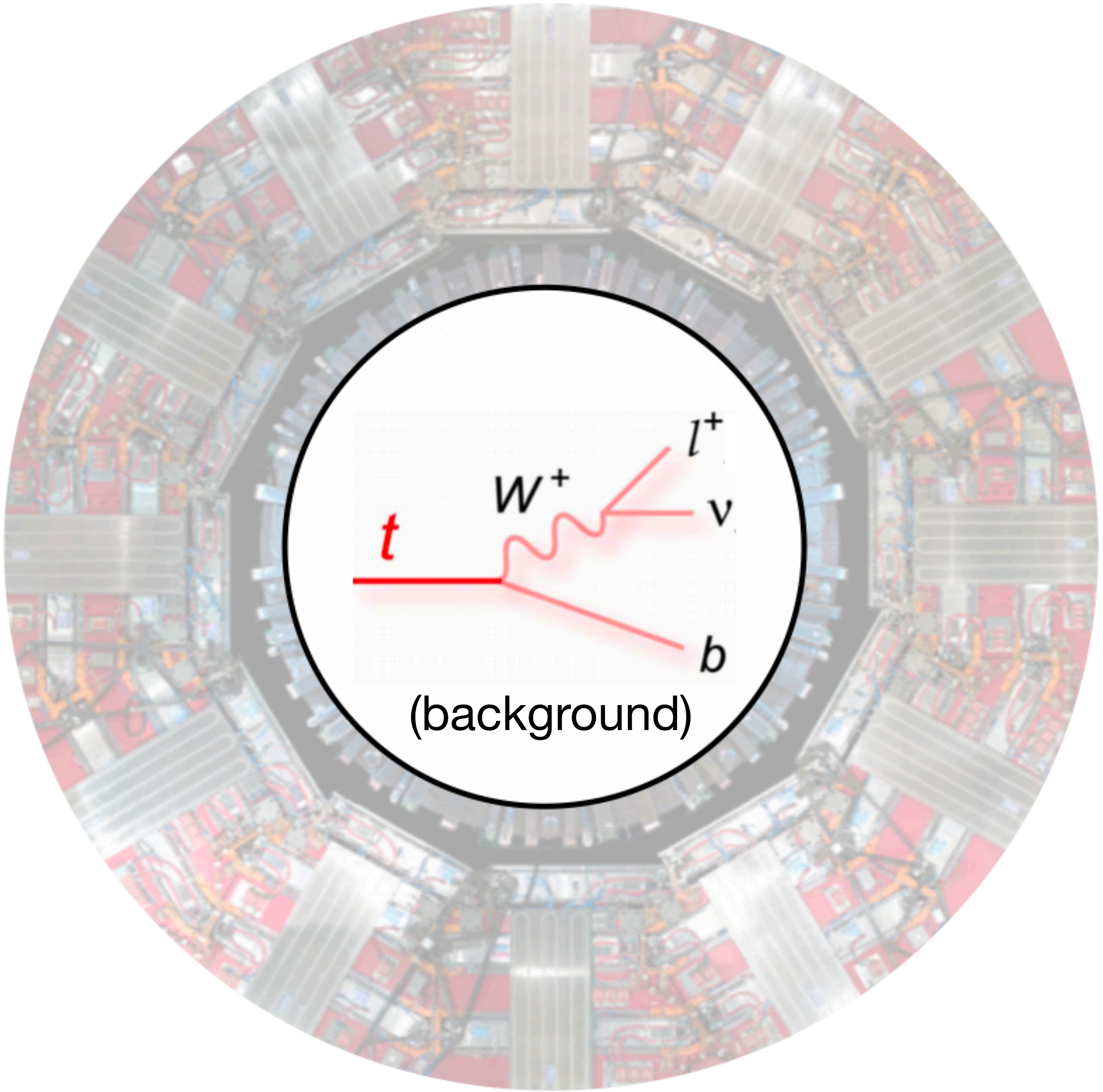
Backup

- Simulations Pythia + Delphes (10^6 events - 50% ttH, 50% ttbb)
- Input variables:
  - Low-level: 4-vector of 8 particles
  - High-level: 26 variables (see backup)



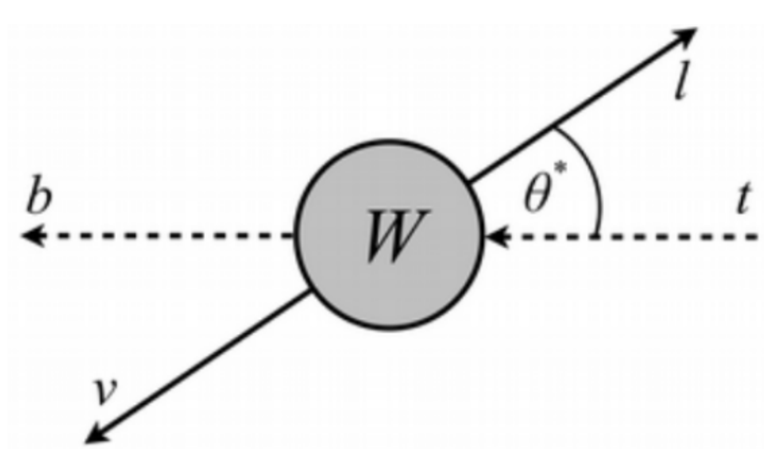
- Assignment Network:
  - Feed forward
    - 5 ELu Layers
      - 500 Nodes
    - Batch Normalization
- Analysis Network:
  - Lorentz-Boost Network (1812.09722)
  - Together with Feed forward:
    - 2 Elu Layers
      - 128 Nodes

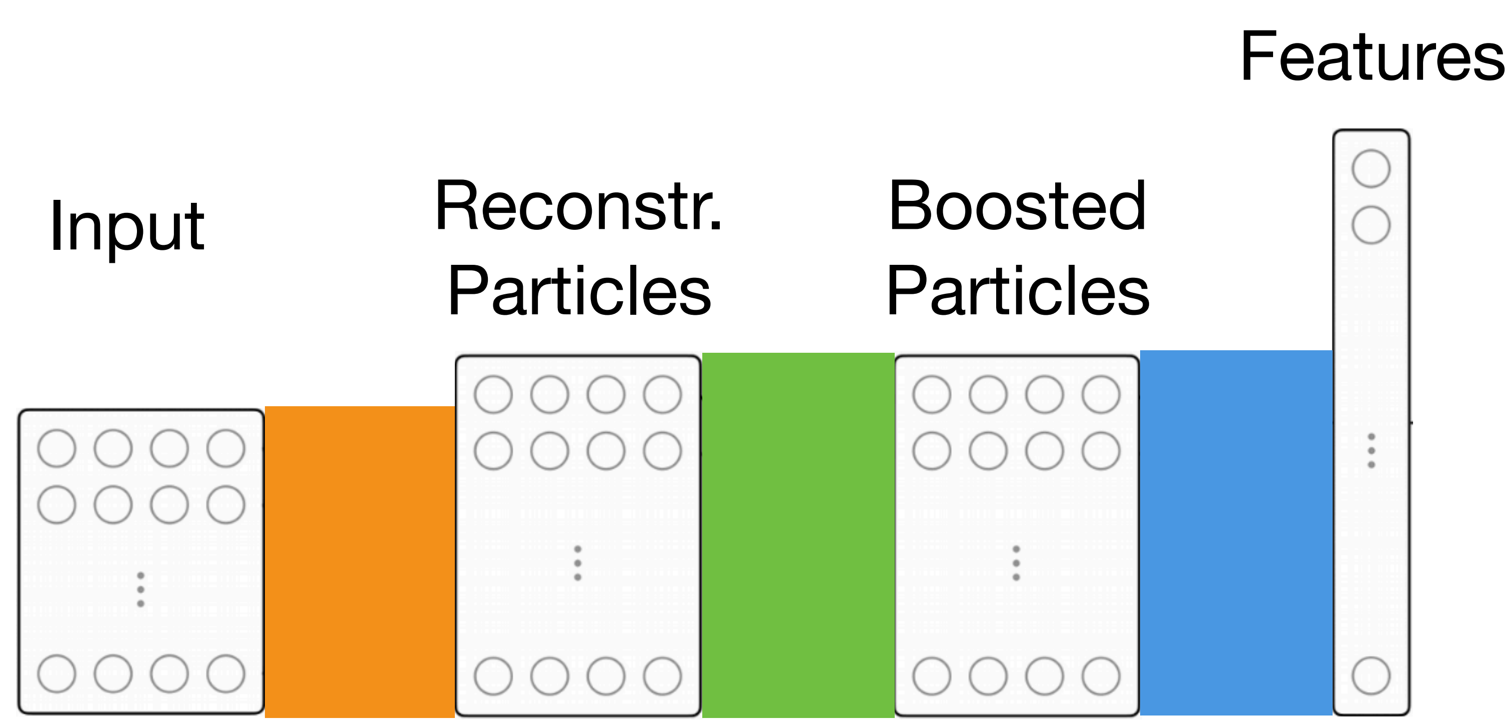
Lorentz Boost Network (1/2) - 1812.09722



Calculate θ^*:

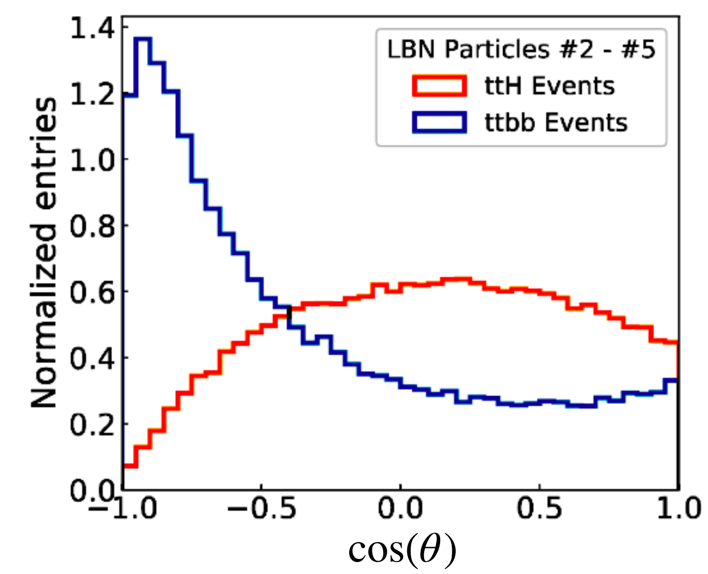
- Reconstruct W
- Reconstruct t
- Boost lepton in W System = l_W
- Boost W in t system = W_t
- Calculate angle $\sphericalangle(l_W, W_t)$

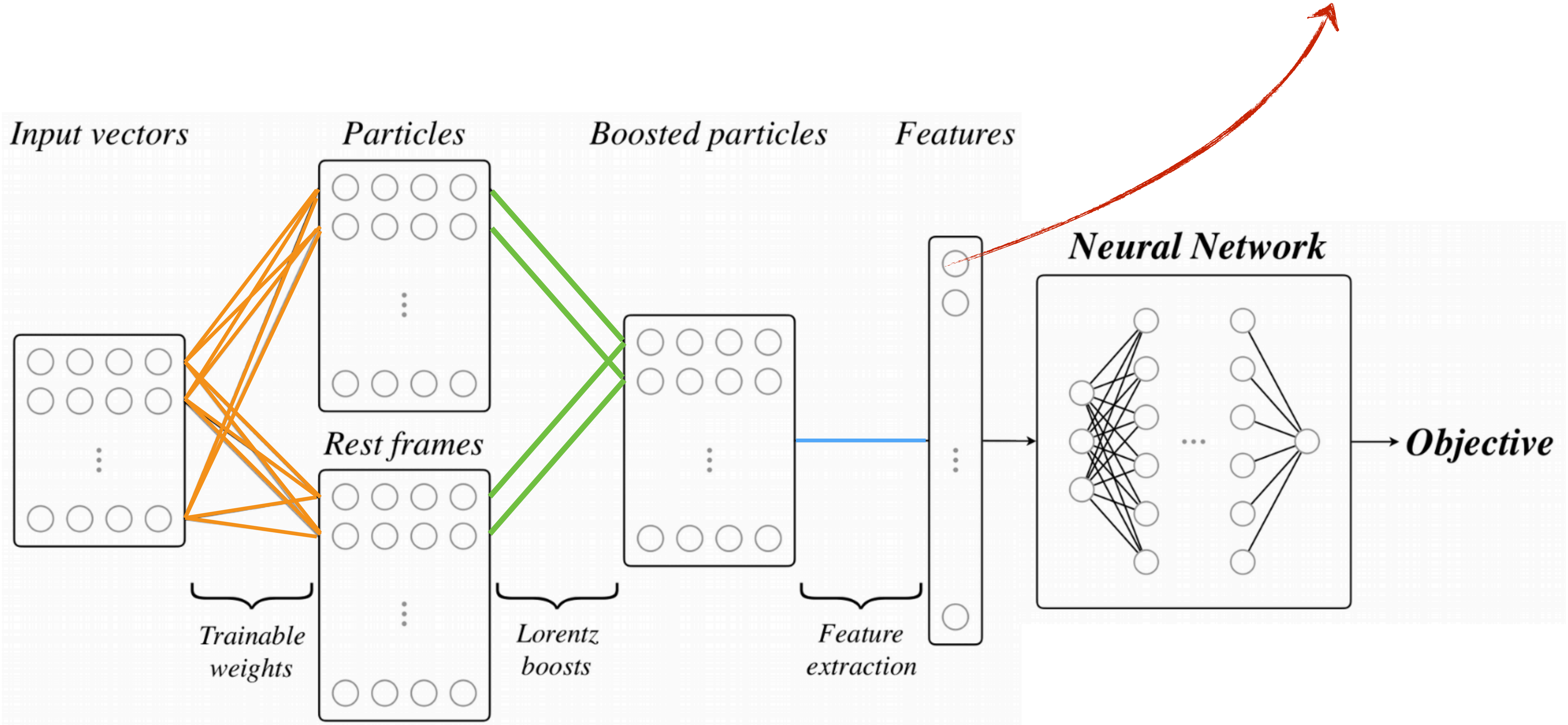




Lorentz Boost Network (2/2) - 1812.09722

- Autonomously build high level variables:
  - Reconstruction: learnable weights
  - Boost: Gamma Matrix
  - Features: Single particle + Pairwise
- Lorentz symmetry within neural network





Munkres algorithm

Dennis Noll - 13.09.22

- Optimally solving the assignment problem Particle ↔ Slot
- Example:

Jobs (slots)

40	60	15
25	30	45
55	30	25

Workers (particle)

Assignment cost

Optimal choice

- Achieves a complexity of $O(n^3)$ compared to all different permutations $O(n!)$
- Involves 5 steps - different complexity:
  1. Row reduction - $O(n^2)$
  2. Column reduction - $O(n^2)$
  3. Test for optimal assignment - $O(n^3)$
  4. If needed: Shift zeros - $O(n^3)$
  5. Making the final assignment - $O(n)$

Algorithm 1 Reinforced Training

Input: Events (e), Permutations ($\mathbb{P}$), Analysis (A_0), Sorting model (S_0)

1: **for** $i \leftarrow 0$ to epochs **do**
2: $\quad \hat{P} \leftarrow \operatorname*{argmax}_{P_n \in \mathbb{P}} A_i(P_n(S_i(\mathbf{e})))$
3: $\quad T \leftarrow \hat{P}(S_i)$
4: $\quad$ train S_{i+1} to approximate T
5: $\quad$ train A_{i+1} with $S_{i+1}(\mathbf{e})$

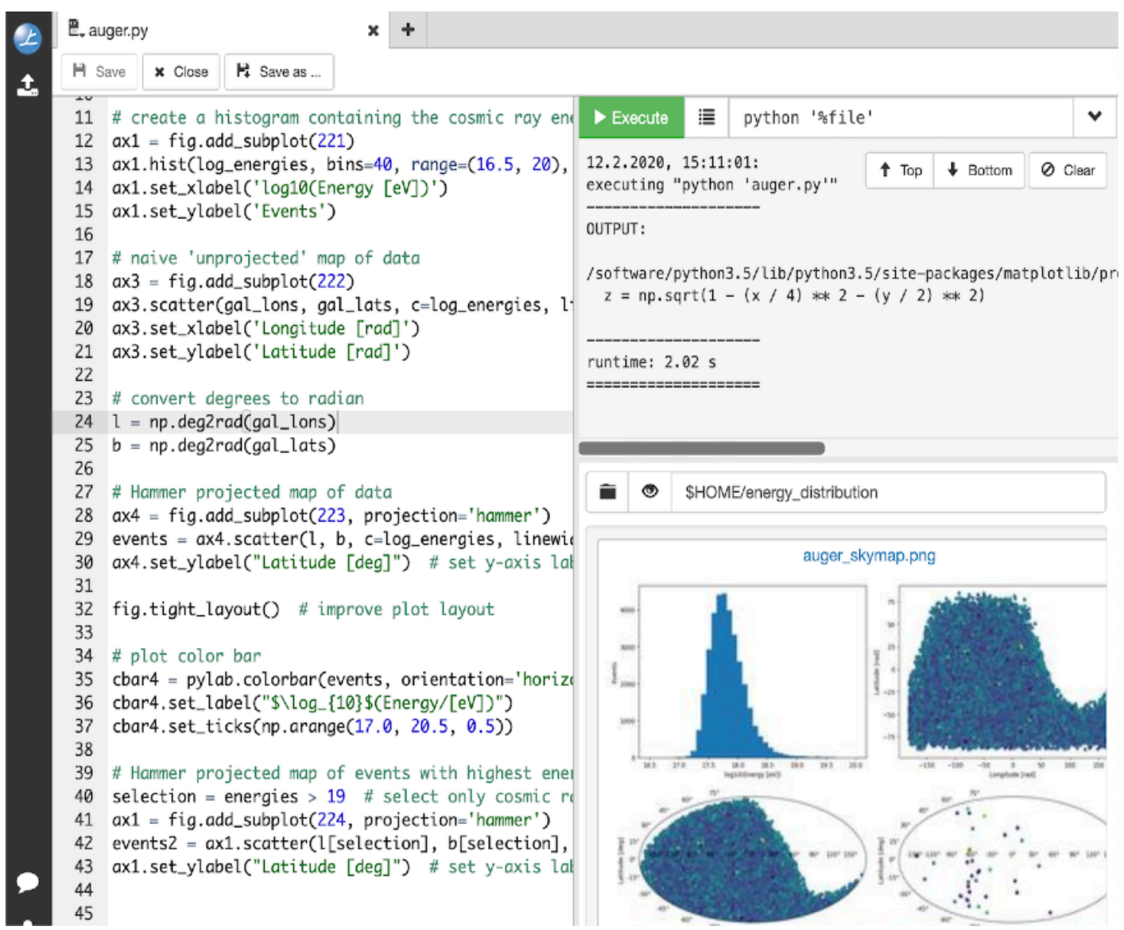
VISual Physics Analysis (VISPA)

RWTH AACHEN UNIVERSITY
Dennis Noll - 13.09.22

VISPA

Software

- Full development environment (editor, file browser, …)
- Runs in your web browser



Hardware

- 200 CPU Cores
- 30 GPU Cards (300 TFlops)
- (ITC RWTH: 1500 TFlops)

- Accessible via https://vispa.physik.rwth-aachen.de/

High-level variables

- Event shape variables: sphericity, transverse sphericity, aplanarity, centrality
- First five Fox-Wolfram moments
- Cosine of spatial angular difference θ^* between the charged lepton in the W boson restframe and the W boson direction when boosted into the rest frame of its corresponding top quark. In the hadronic branch, the down-type quark is used owing to its increased spin analyzing power
- minimum, maximum and average of ΔR of jet pairs
- minimum, maximum and average $|\Delta\eta|$ of jet pairs.
- minimum and maximum of the distance in $\Delta R$of jet-lepton pairs
- minimum, maximum and average $|\Delta\eta|$ of jet-lepton pairs
- sum of the transverse momenta of all jets
- transverse momentum and the mass of the jet pair with the smallestΔR
- transverse momentum and the mass of the jet pair whose combined mass is closest to the Higgs boson mass m_H=125GeV

Inspired by 1804.03682