Machine-Learning assisted reconstruction of hadron-collider events using mini-jets

IMPRS recruiting workshop, Max Planck Institute for Physics

Josef Murnauer, 25.11.2024

Background - Autogenerated with gencraft from title

Education & Research

- Bachelor of Science in Physics at LMU Munich, 10.2018 - 09.2021
- Master of Science in Physics at LMU Munich, 10.2021 - 05.2024
- Master Thesis at Max-Planck-Institute for Physics, 04.2023 - 05.2024
- Currently working at Infineon Technologies in Munich, since 07.2024

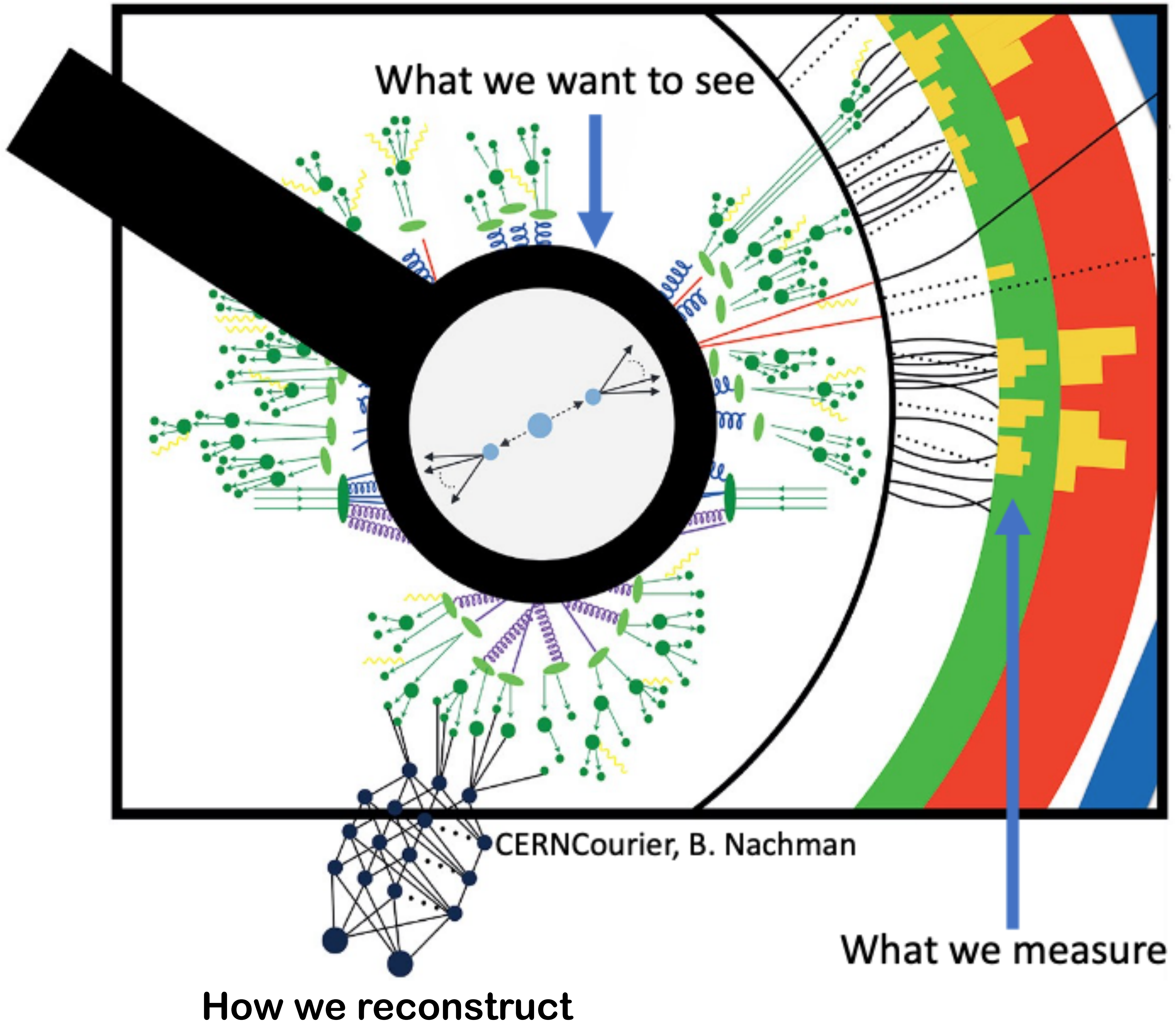
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MAX-PLANCK-INSTITUT FÜR PHYSIK

infineon

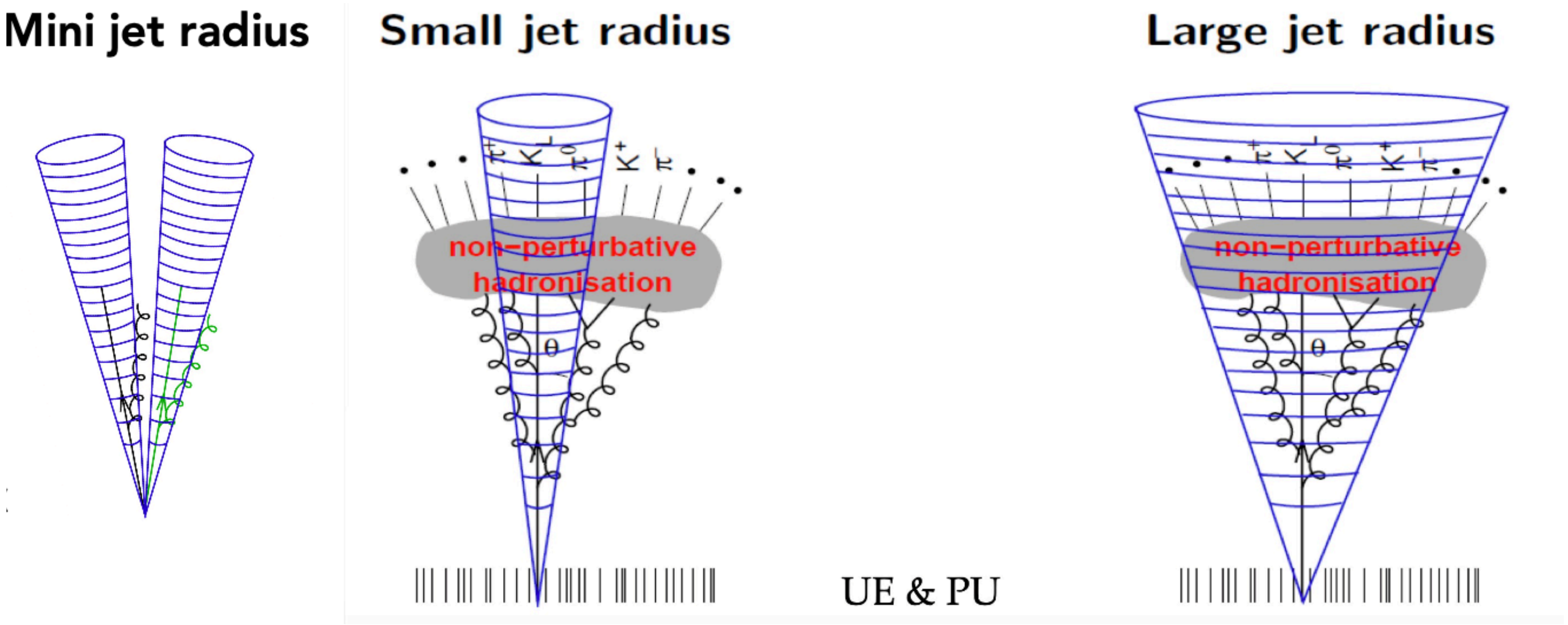
Introduction

- At hadron collider, like the LHC, a critical task is the reconstruction of intermediate resonances from Higgs-, W- or Z-bosons, or from a top quark decay
- Decays into partons (hadronic decays) are particularly difficult (but also leptonic decays are challenging when neutrinos are involved)
- Due to color confinement, partons cannot be observed directly in experiment
- Instead, they form jets of particles
- Jet algorithms are applied as a proxy for a single quark, or to reconstruct an intermediate resonance directly (H, W, Z, t)



Jet Reconstruction

- Many different approaches for jet-analysis
- Most commonly used: Recursive clustering algorithms
- Quantity to describe the size of the jets
  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$
- Typical values at the LHC:
  - Large-R jets - R=0.8
  - Small-R jets - R=0.4
- New approach:
  - **Mini-Jets - R=0.1**
  - Idea: parton-jet duality for sufficiently hard partons from matrix element or hard radiation in Parton shower, but less sensitive to soft radiation
  - Challenge: reconstruct H, W, Z-bosons or tops quarks from mini-jets

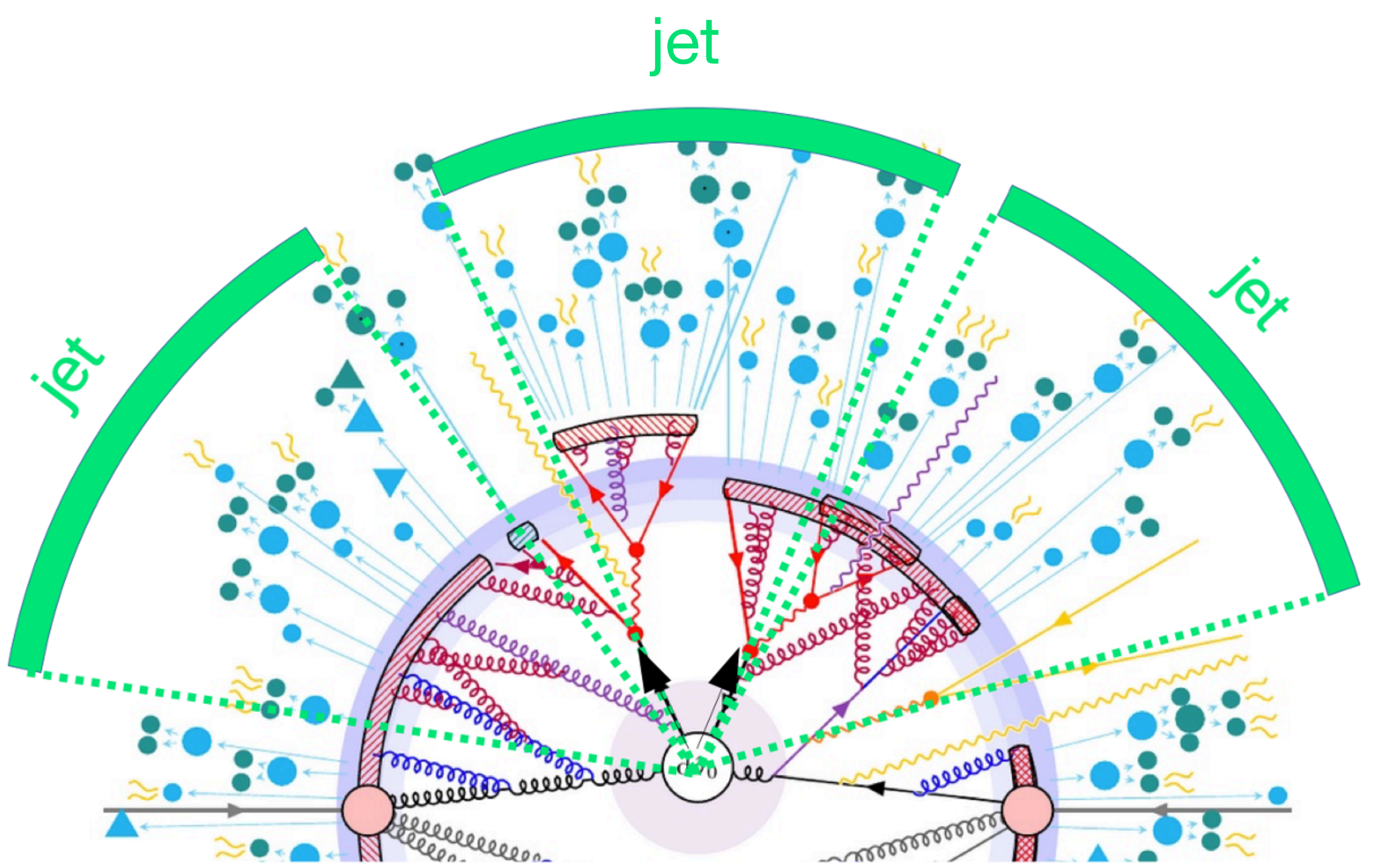


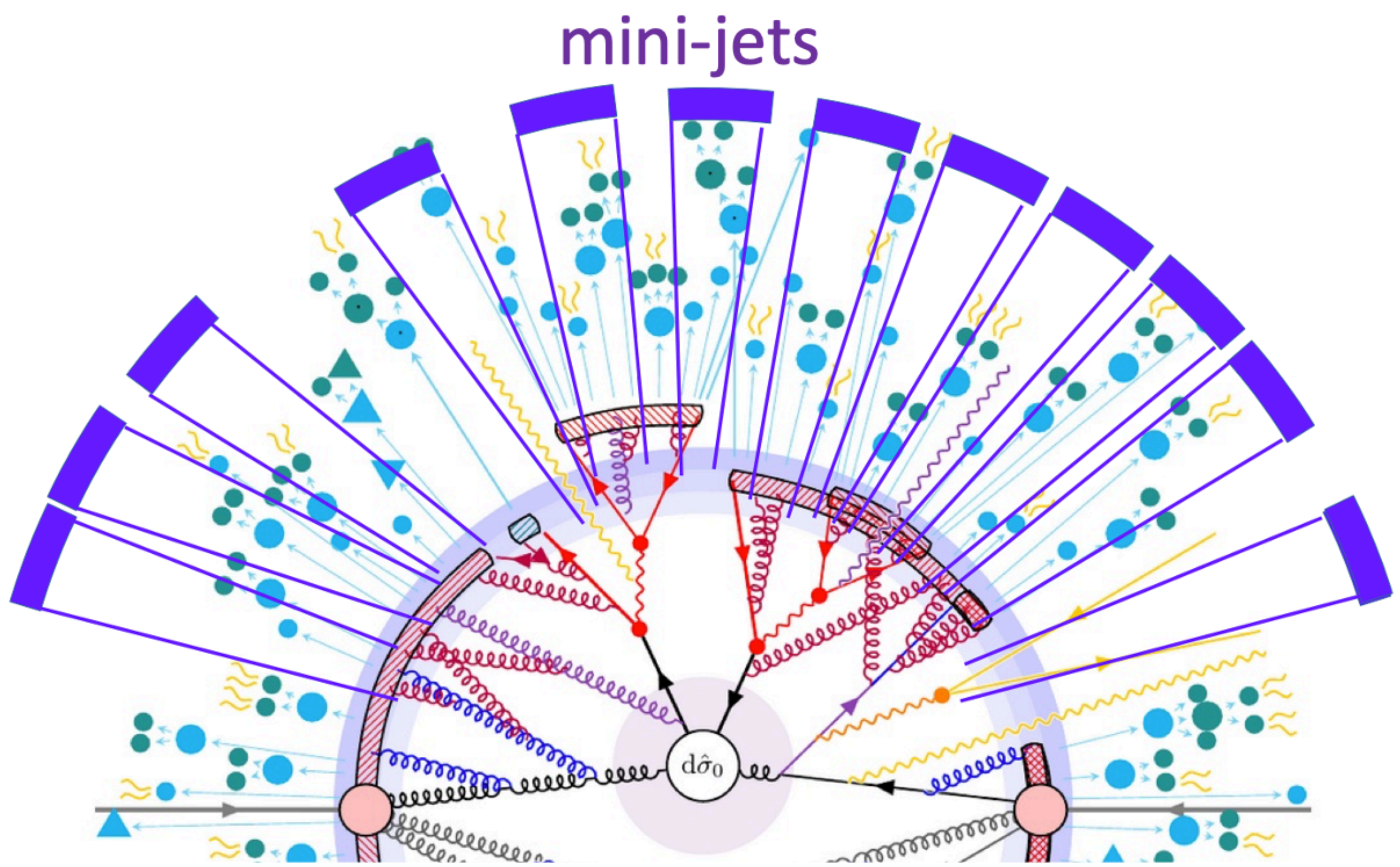
Jets at hadron colliders (2); Gavin Salam

Mini-Jets

- Idea of jet algorithms in common LHC analyses:
  - Define a suitable jet algorithm and size (R=0.4, 0.8, etc.)
  - Assumption: All information of the hard process is contained in the selected jets and (isolated) leptons
  - Limitations through different event topologies and hard radiation
- Idea of mini-jets:
  - Use many very small jets and collect all information
  - Capture details of hard radiation
  - Sensible jet multiplicity of <n> ~15
  - But: Difficult to handle combinatorics to reconstruct underlying hard physics

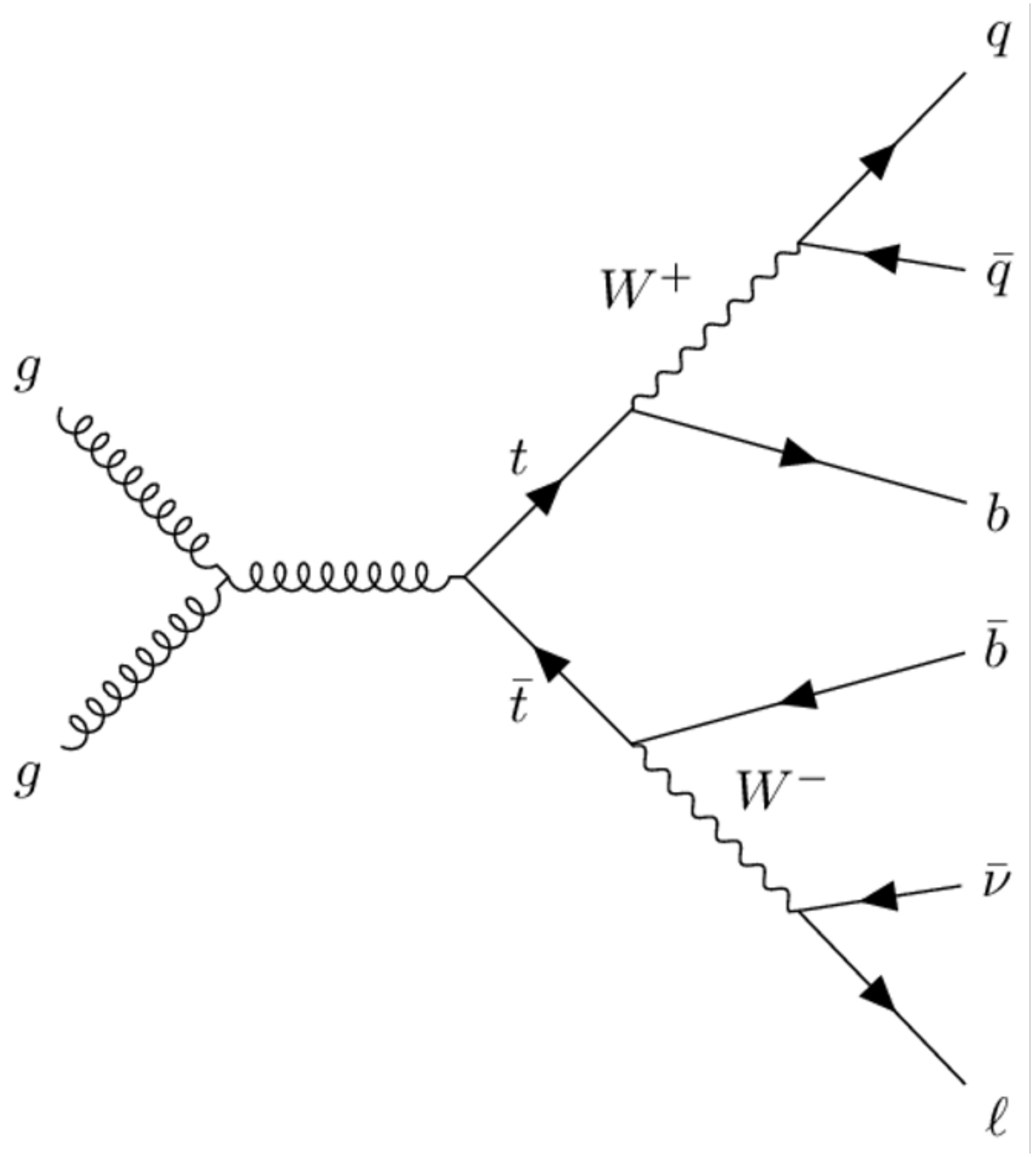
    → **Solution**: Let a neural network handle the excessive information and high combinatorics





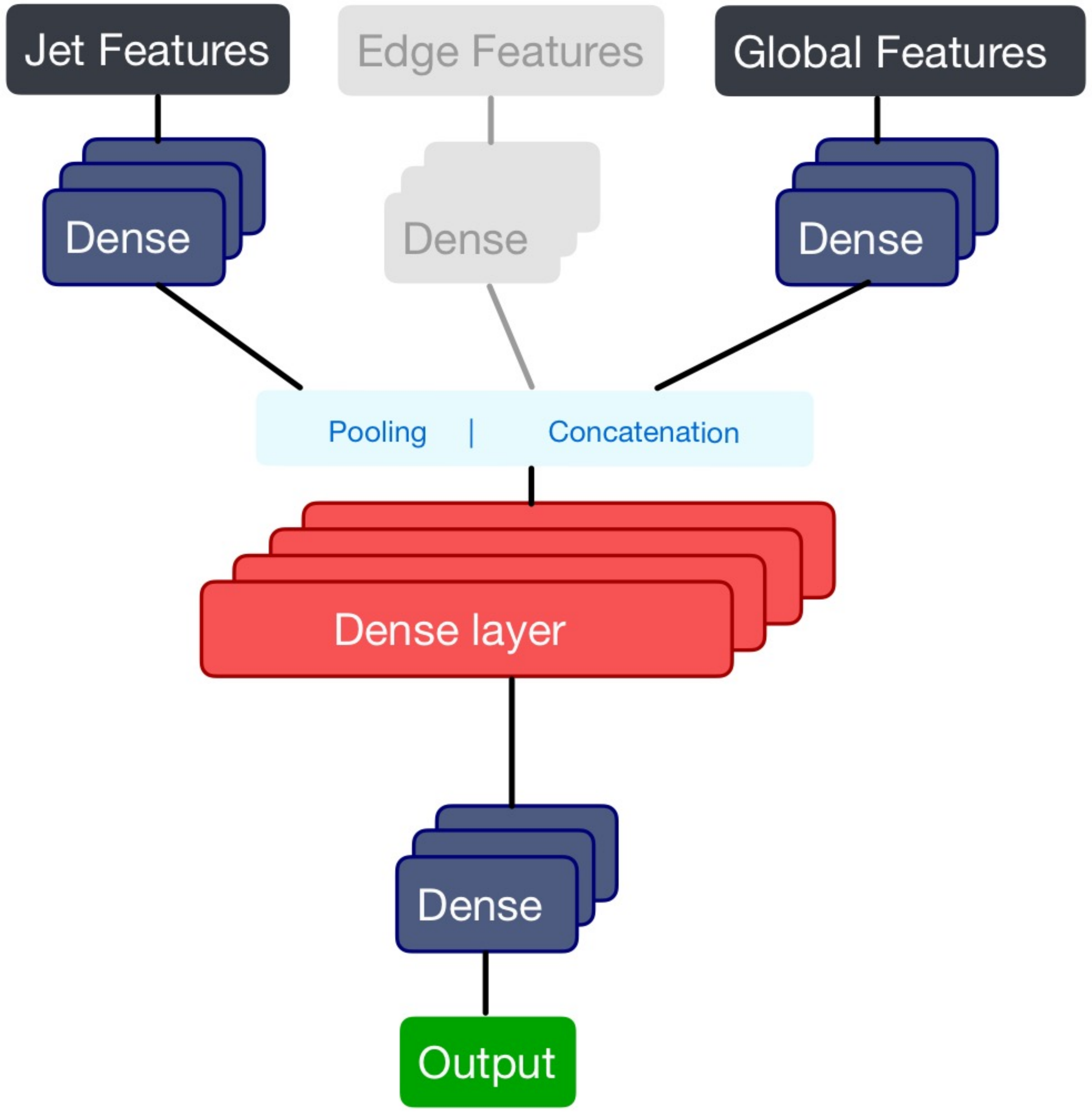
Underlying Physics and Data

- Pair of top quarks, produced in proton-proton collisions, decays into two W-bosons and two bottom quarks
- Centre of mass energy: 13 TeV
- W-bosons decay semi leptonically
  - Leptonic W-boson → lepton and neutrino
  - Hadronic W-boson → quark and anti-quark
- Monte Carlo Simulation:
  - Particle-level analysis
  - Anti-kt algorithm (different R-values)
  - ME, parton-shower and hadronization from Pythia 8.3



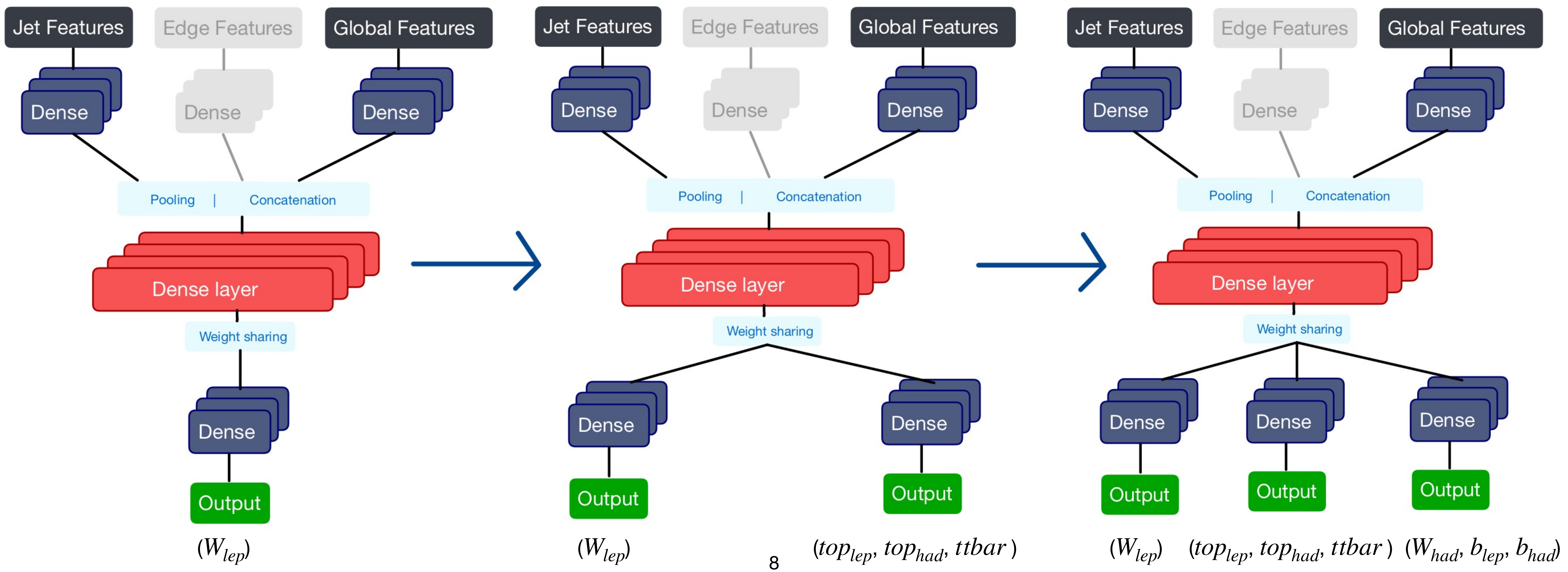
Neural Network Architecture

- Jet Features:
  - Mini-jets with 9 features ($p_T, \eta, \phi, p_x, p_y, p_z, E, m,$ *b-tag*)
- Global Features (Lepton and missing energy):
  - 3 final-state lepton features (p_T, η, ϕ)
  - 2 Missing features (p_x, p_y)
- Global pooling for jets & concatenation of jets and global features
- Core layer:
  - Multiple dense layers with all relevant information of the event
- Output:
  - Extracting information for a specific resonance
- Size of network $\mathcal{O}(10^6)$ parameters



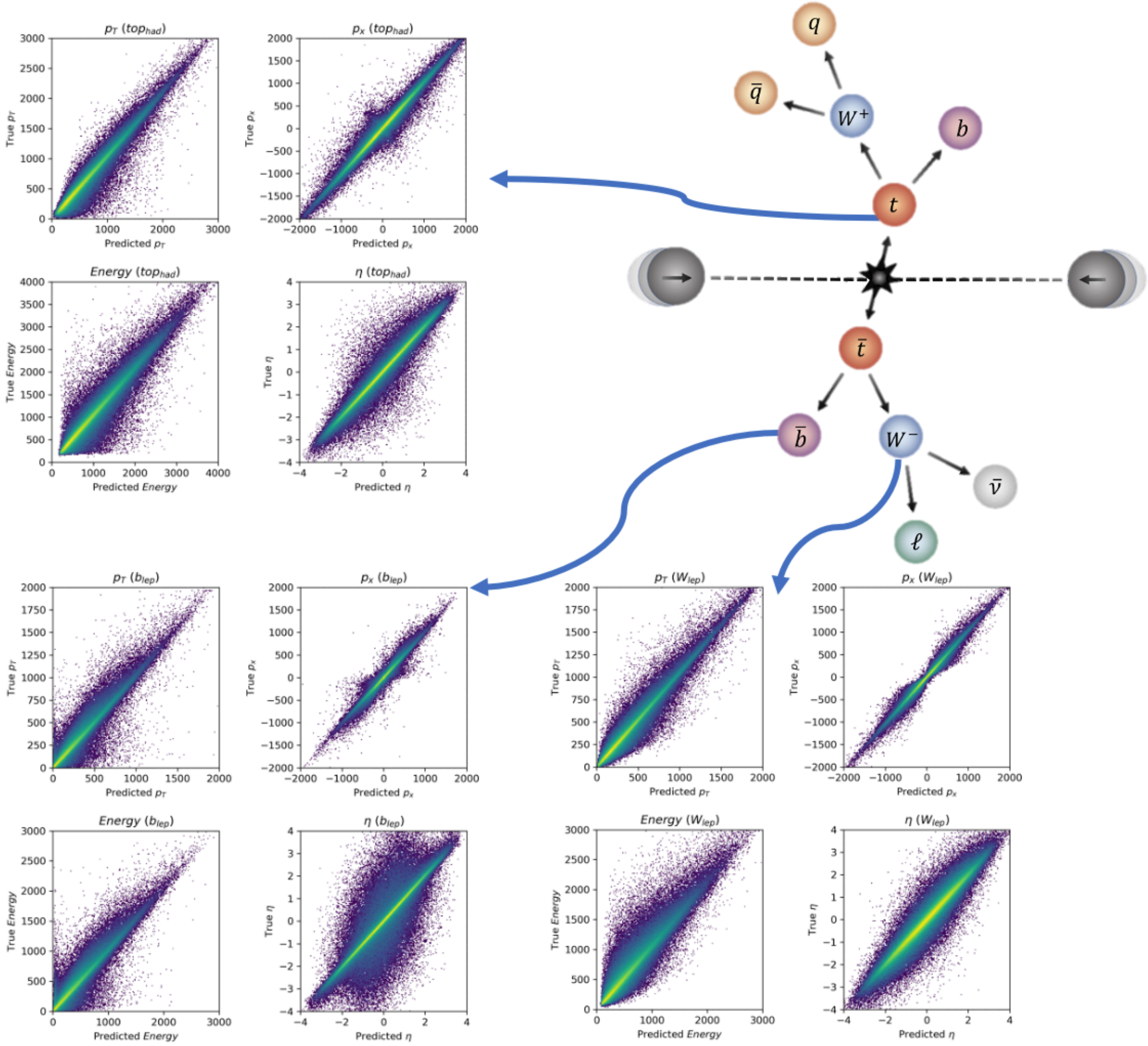
Training the Model

- Start with training one output particle
  - Easiest particle for the model in this case: W-boson decaying leptonically
- Get to more difficult particles step by step
  - Use weights from leptonic W-boson to train top quarks and the invariant mass of top quark pair
  - Use weights from these two models to train bottom quarks and hadronic W-boson



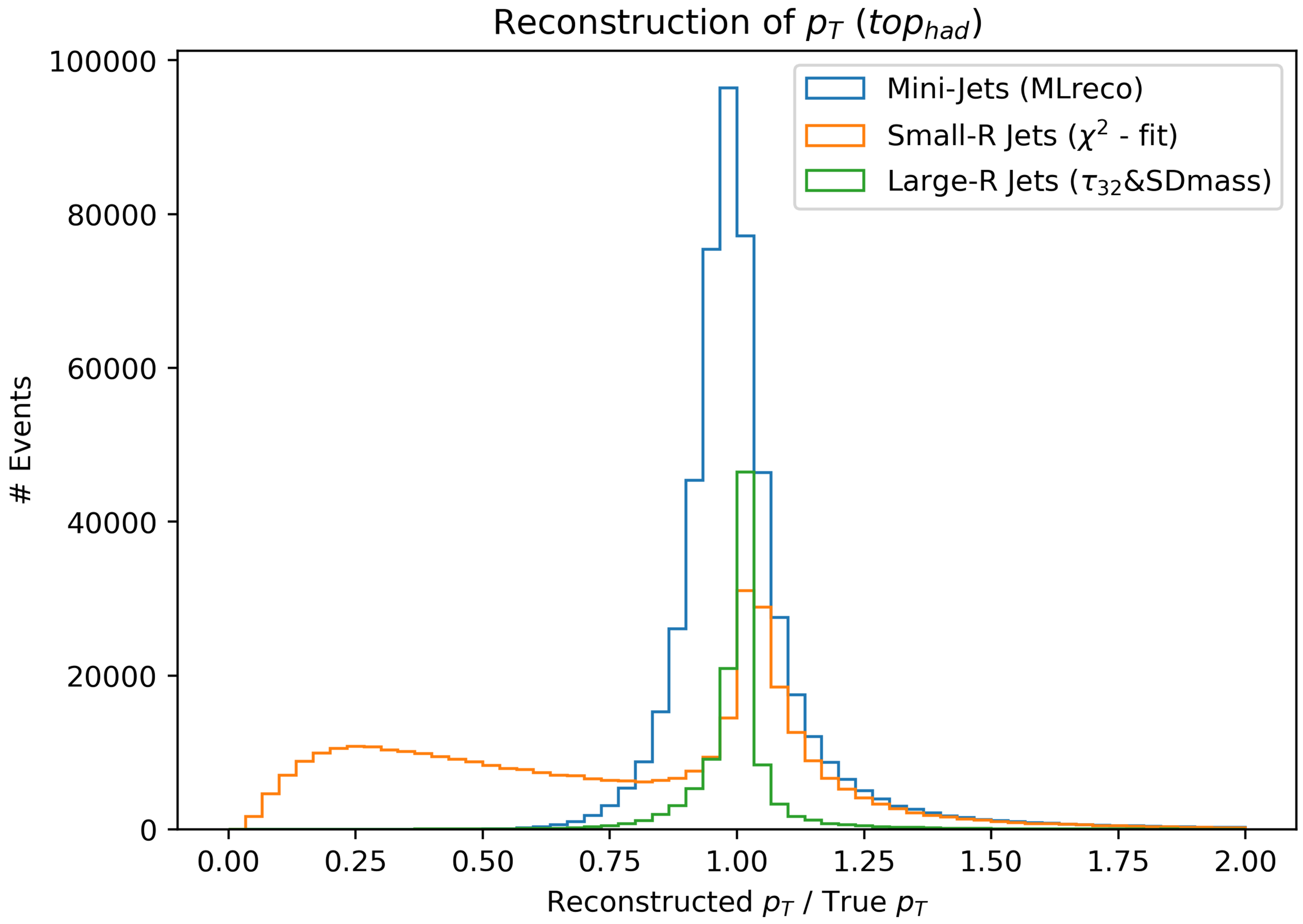
Reconstruction of Particle Properties

- X-axis: predicted values of network
- Y-axis: true values of target observables
- Properties of hadronic top quark, leptonic b-quark and leptonic W-boson are shown
- B-tagging information included



Performance of the p_T - Reconstruction

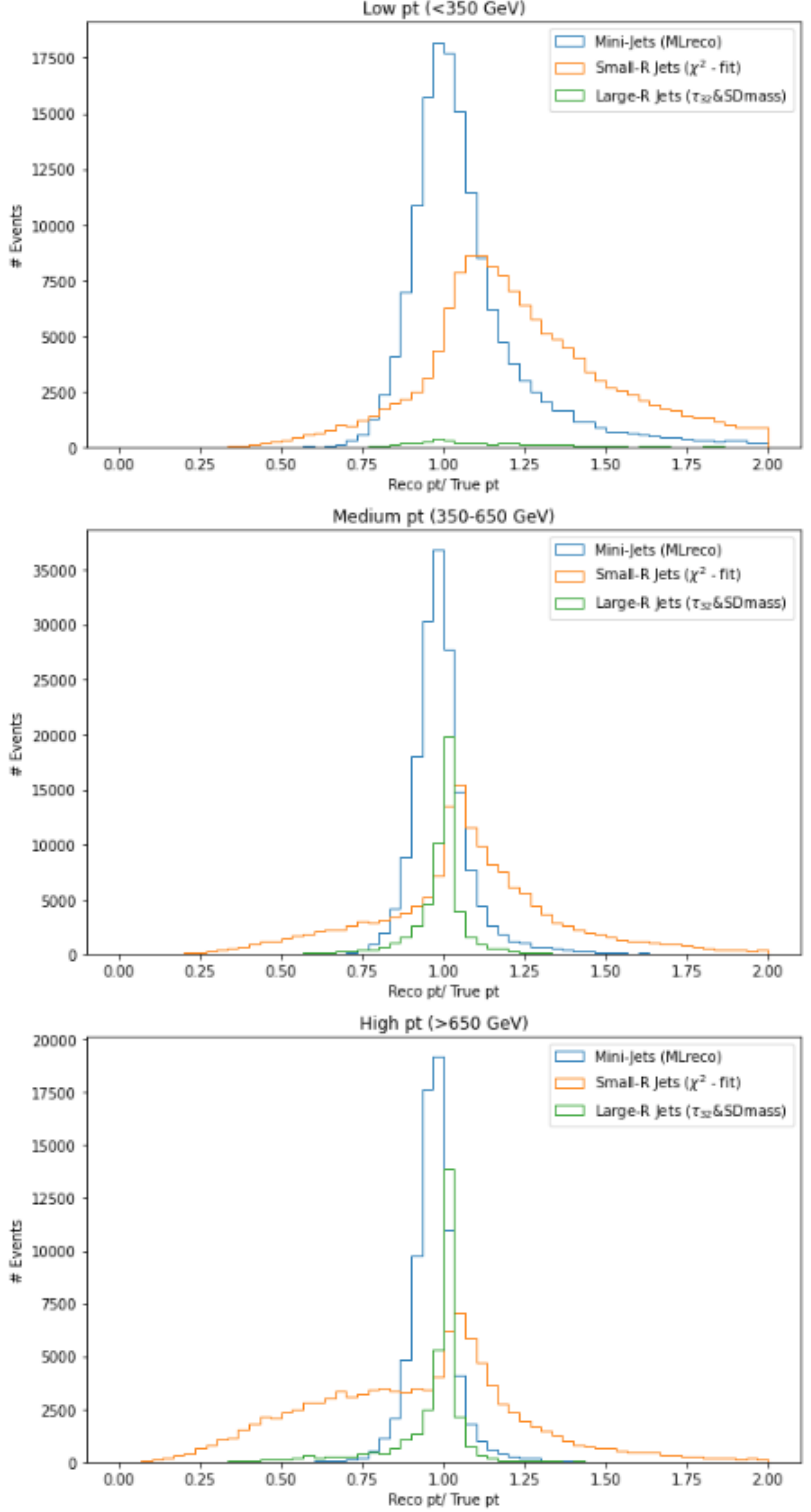
- Ratio of reconstructed and true transverse momentum of hadronically decaying top quark
- Classical reconstruction for small-R jets:
    - Select 6 leading jets and combine 3 of them. Combination with closest mass to the actual top quark mass is selected
- Classical reconstruction for large-R jets:
    - Select „2-jet-like“ events with $\tau_{32} < 0.54$
    - Softdropmass: 110 GeV < SDM < 210 GeV
    - Efficiency ~30%
- Small-R and large-R reconstruction algorithms suffer from efficiency losses due to additional cuts
- Machine-Learning model efficiency outperforms classical reconstruction with small-R and large-R jets



Performance of the p_T - Reconstruction

For Different Event Topologies

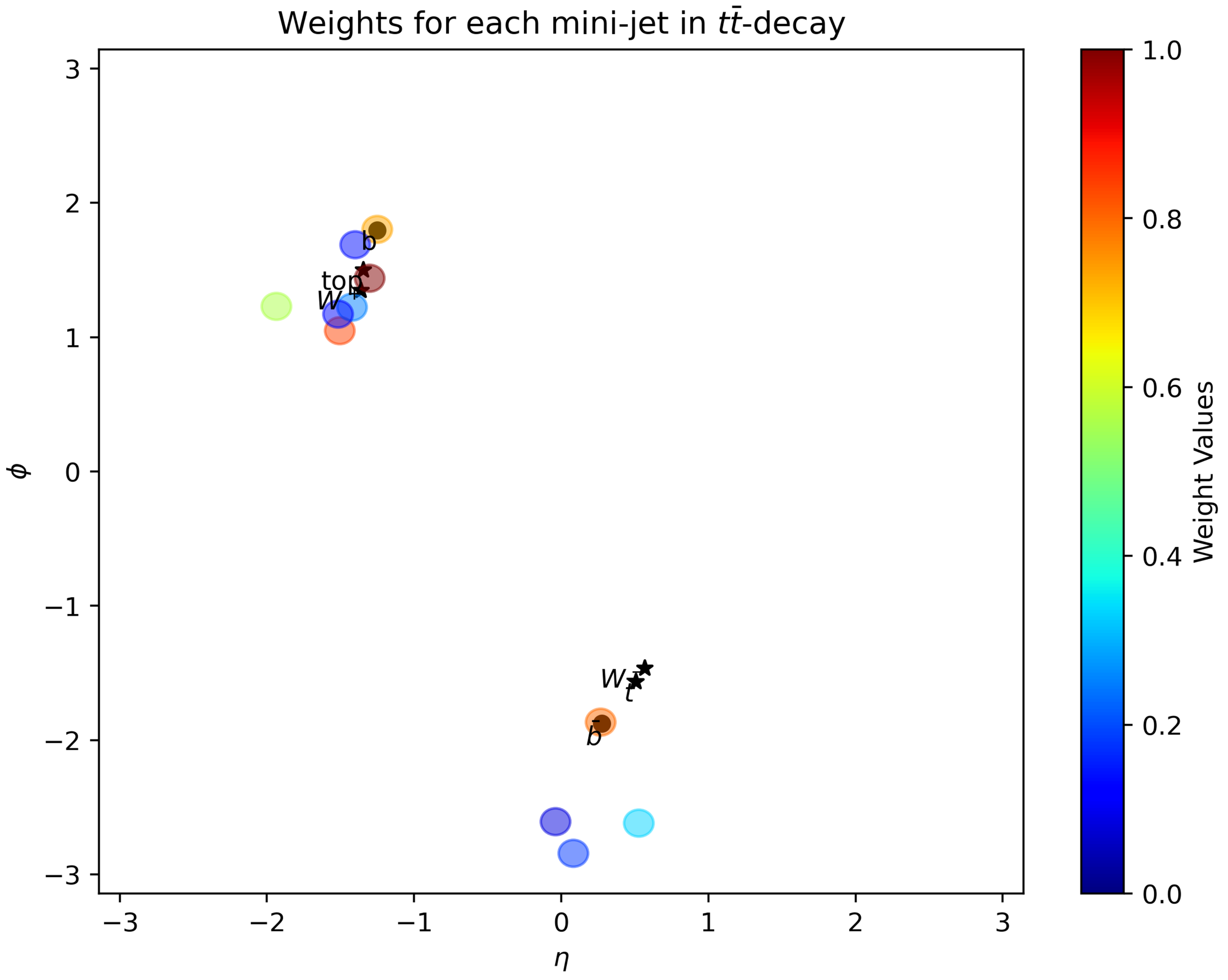
- Low p_T (< 350 GeV):
  - Classical reconstruction methods have problems due to non-boosted decay topology
- Medium p_T (350 GeV - 650 GeV):
  - Better description of physics with classical methods due to more boosted topology
- High p_T (> 650 GeV):
  - Large-R reconstruction very good due to highly boosted topology
  - Large-R jets with efficiency of (only) ~40%



Alternative Network Architecture - Interpretability

- **Goal**: Try to understand the reconstruction process of the model
- Idea: Assign a weight to each mini-jet (importance for reconstruction)
    - Input: 4-vector of each mini-jet
    - Penultimate layer: Dense layer with same number of nodes as number of mini-jets
    - Every mini-jet gets a weight which corresponds to its importance of the reconstruction
    - Last layer: Custom layer with multiplication of weights and input
    - Output: 4-vector of target particle (top-quark)
    - Custom layer:

$$\begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \sum_{i}^{25} w_i \begin{pmatrix} E^{input} \\ p_x^{input} \\ p_y^{input} \\ p_z^{input} \end{pmatrix}_i$$



Summary & Outlook

- We have studied a new reconstruction methodology for collider events using mini-jets and a deep neural network
- Mini-Jets
  - Anti-kt jets with radius of R=0.1
  - Goal: Preserve physics of hard interaction, as well as hard emissions in the parton shower
  - Reasonable multiplicity of <n> = 15 in semi-leptonic ttbar events at $\sqrt{s} = 13$ TeV
- Mini-Jets are input to a ML-based reconstruction of target observables
  - W/Z-boson, top-quark, Higgs-boson quantities
- ML-based reconstruction outperforms classical reconstruction algorithms over a large kinematic range
  - Top-quark, W-boson and b-quark properties were studied
  - Algorithm can handle different event topologies at different scales in a single algorithm: e.g., resolved and boosted top-quarks
- My work opens the path to state-of-the-art and automatized reconstruction of collider events, and can similarly also be applied in other research fields
- Paper in preparation (Britzger, Kluth, Kogler, Murnauer)

Backup

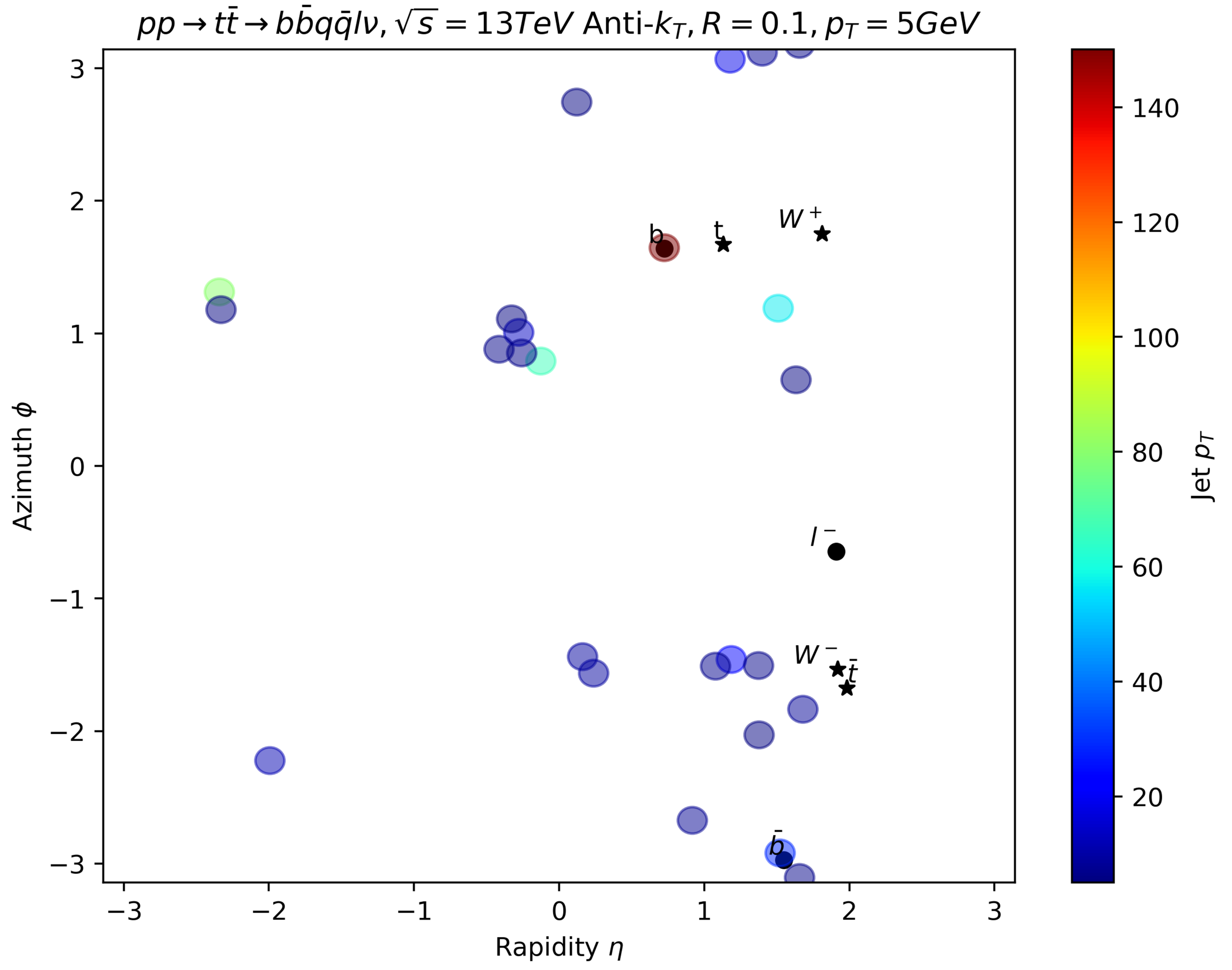
Results

Performances of different jet algorithms			
Jet Sizes	Mean Value $\bar{x}$	Standard Deviation σ	Efficiency ϵ
ML-Reco	1.0340	0.4112	100 %
χ^2	0.7709	0.4819	74.14 %
Jet-tagging	1.0065	0.2247	21.42 %

Performance of different jet sizes for various topologies									
Jets	Low p_T			Medium p_T			High p_T		
	$\bar{x}$	σ	ϵ	$\bar{x}$	σ	ϵ	$\bar{x}$	σ	ϵ
ML-Reco	1.1138	0.6604	100%	0.9960	0.1175	100%	0.9753	0.0792	100%
χ^2	1.0035	0.5461	77.61%	0.7408	0.3440	76.01%	0.4016	0.3187	65.48%
Jet-tagging	1.2642	0.9690	2.47%	1.0012	0.0941	28.72%	0.9873	0.0904	38.28%

Challenging Events

- Peak position of ML reco slightly shifted towards < 1
- Few events with very bad reconstruction
  - Low momentum decaying top quark is predicted as a high momentum top quark:
  - $\frac{predicted}{true} = \mathcal{O}(100) \gg 1$
- Other mini-jets with high momentum are present, coming potentially from leptonic top quark or final state radiation



B-tagging

- Small and large R jets are b-tagged via ghost-matching of B-meson or B-baryon
- Distance of ΔR to the next b-tagged classical R = 0.4 jet is given as an extra feature of the mini-jets
- DNN learns from distance which and how many mini-jets are used for reconstructing a b-quark

Architecture of the Neural Network

- Current research is dominated by GNNs (Graph Neural Networks)
  - We have tried, but didn’t see improvements
  - Possible explanation: enough parameters for DNN to learn graph structure of data
- We wanted a proof of concept: flexibel model which is capable of regressing many target observables at once ~ 30
- Want to test other architectures in the future
- PELICAN, Lorentz-Equivariant Transformer, …

Implying Physical Constraints

- **Goal**: Reduce number of trainable parameters of the neural network
  - Model could learn features that are misleading
  - Model with huge number of weights is hard to interpret
  - Reduce computational effort
- Benefits of physical constraints:
  - Guide the model while learning (respecting fundamental laws of physics)
  - Simplify the complexity of the model
- Results:
  - Used mass of top quark as a constraint
  - Reduced number of parameters by a factor of 10 with comparable results to the initial model