PSICHE* in Time and Space Applied in Particle Physics Jet Reconstruction

Margaret Lazarovits, Chris Rogan
University of Kansas

ML4Jets, University of Vienna
September 15, 2026

This material is based upon work supported by the National Science Foundation under awards 1945038 and 2310030.

*Probabilistic, Structure-Intrinsic Clustering with an Hierarchical Embedding, arxiv:2609.16276

Outline

Presenting generally-applicable clustering algorithm with a domain-specific realization

Jet Reconstruction

clustering context

- Cluster size and multiplicity
- Substructure
- Incorporation of additional dimensions

PSICHE

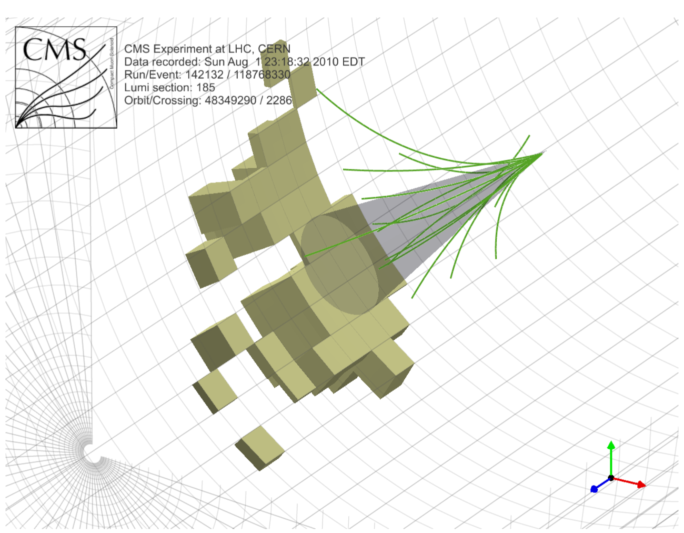
algorithmic highlights

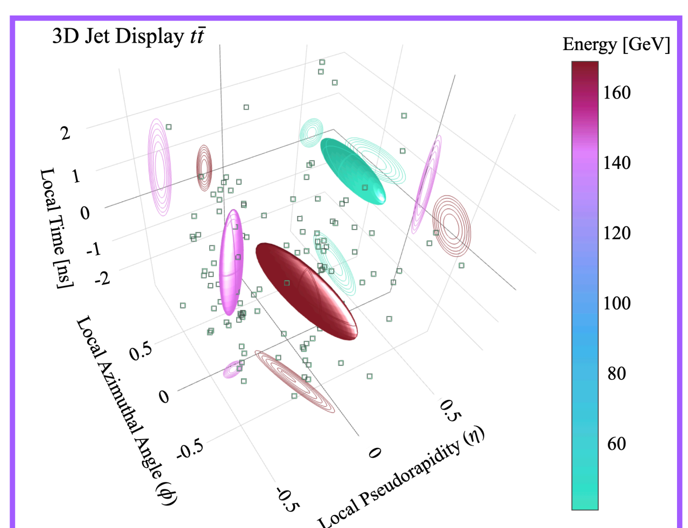
- Hierarchically-embedded models
- Structure-intrinsic design
- Timing-aware capabilities

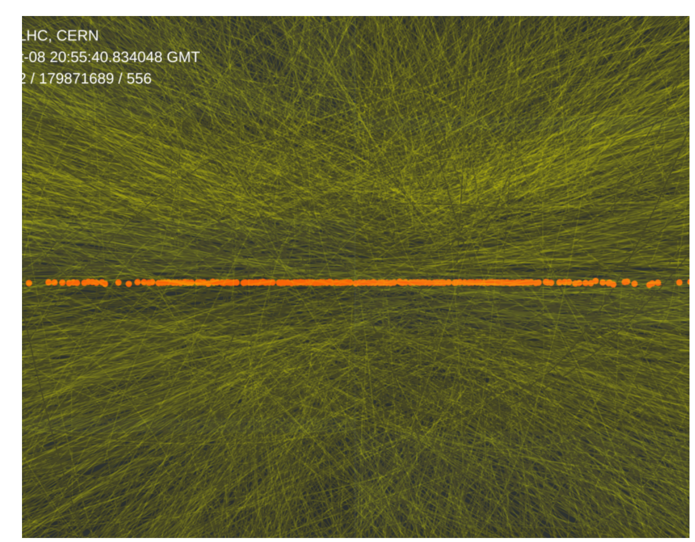
Pileup Mitigation

physics application

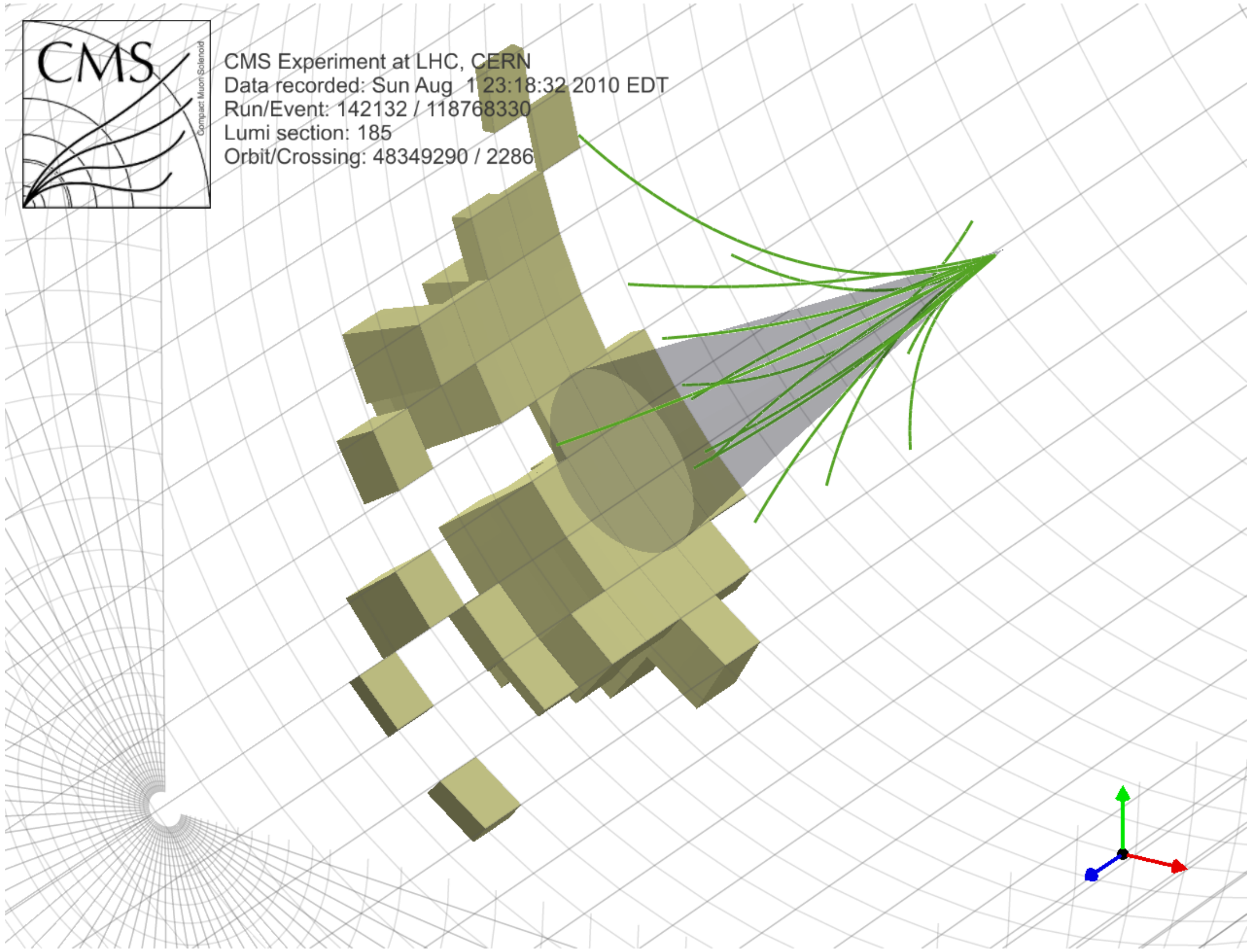
- Dynamic scale
- Simultaneous substructure learning
- Kinematic and temporal characterization



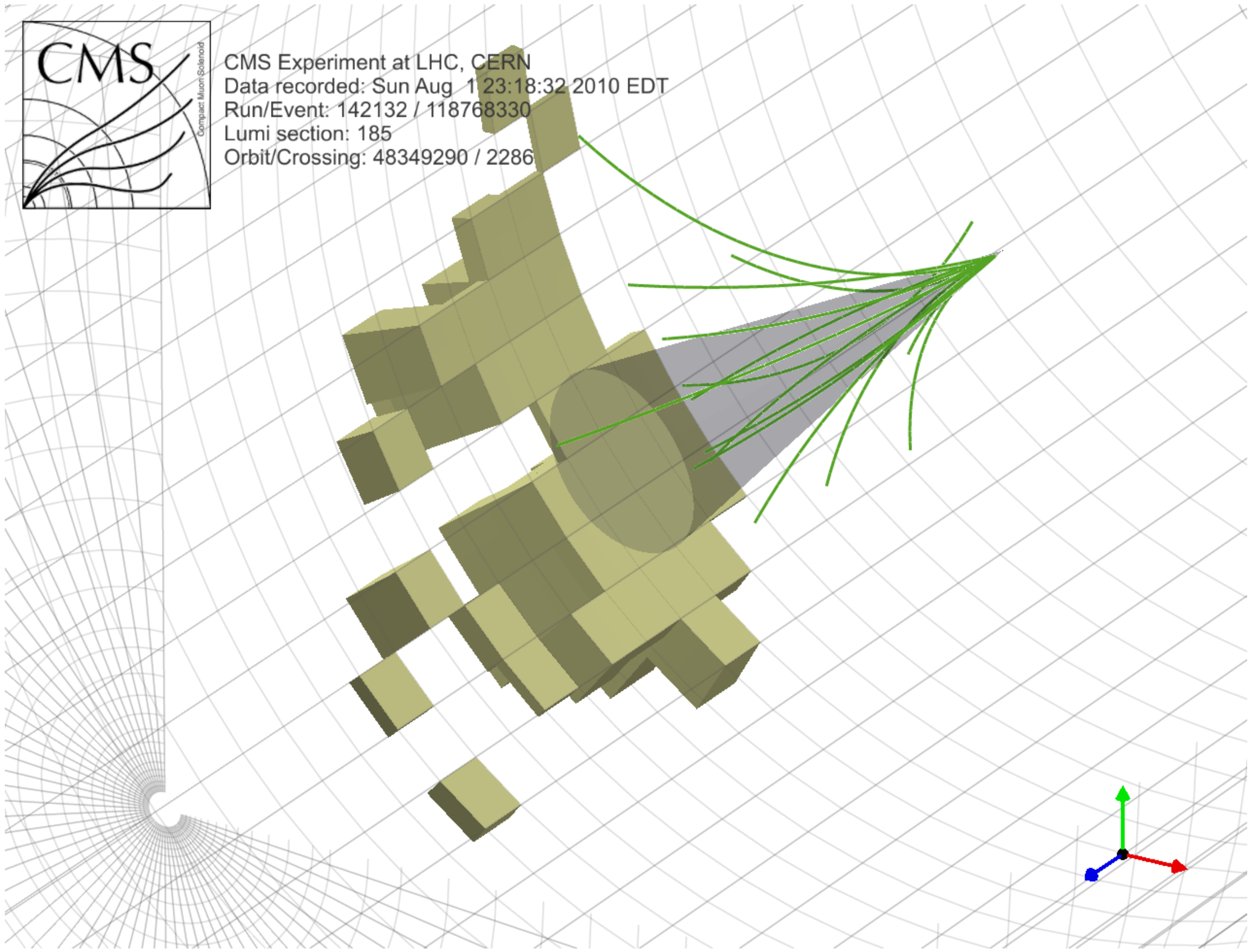


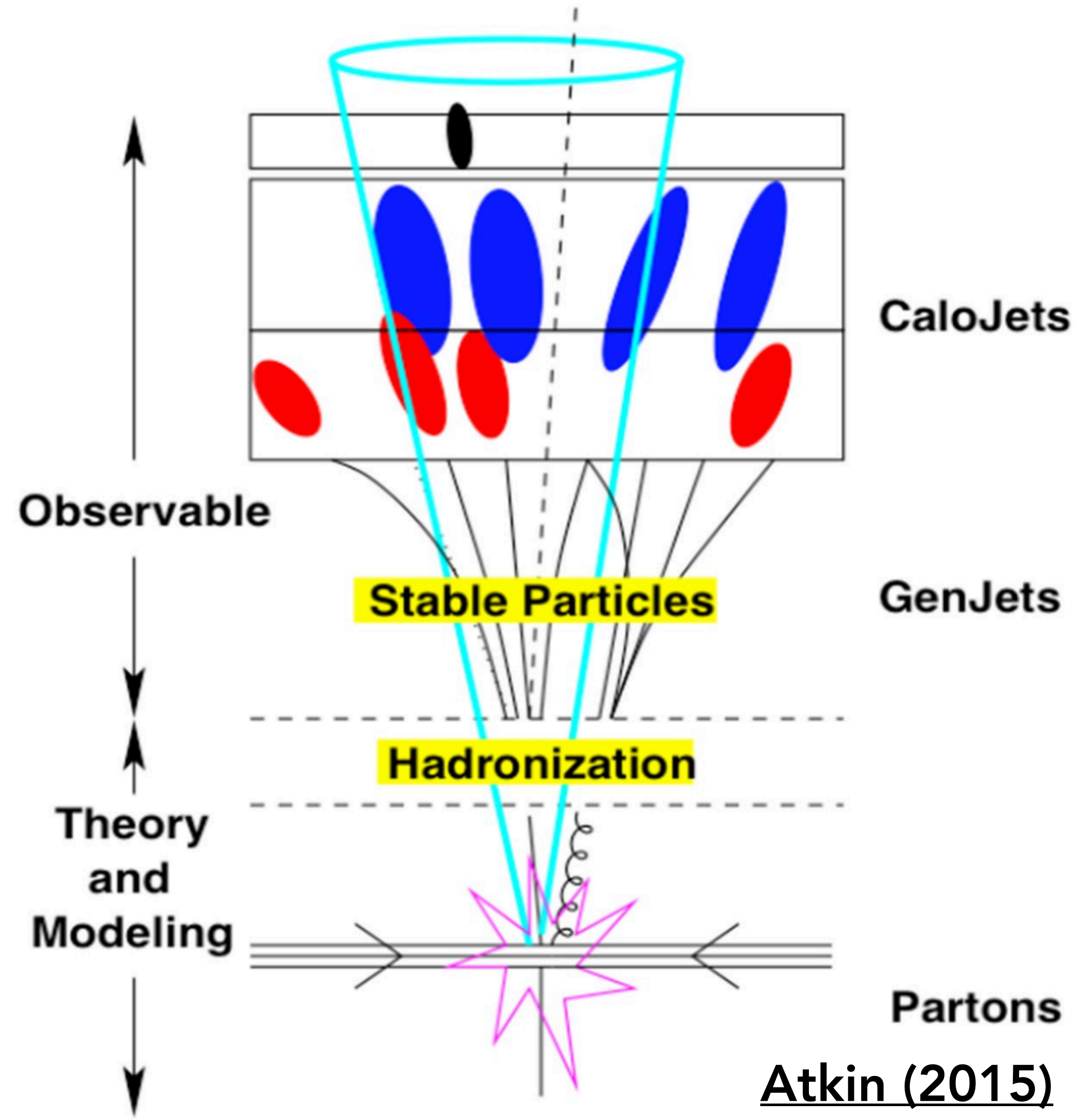


What is a jet in particle physics?



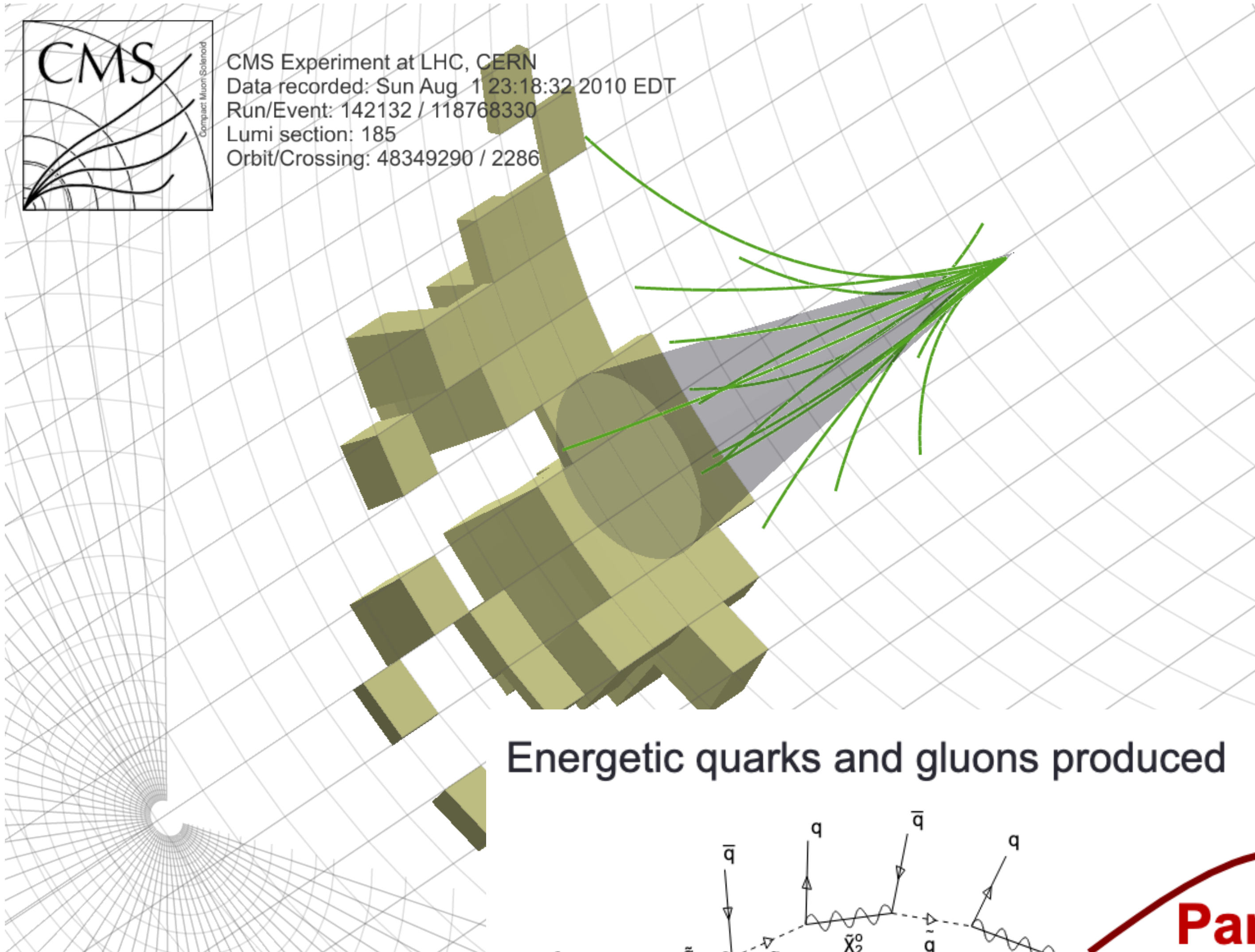
What is a jet in particle physics?

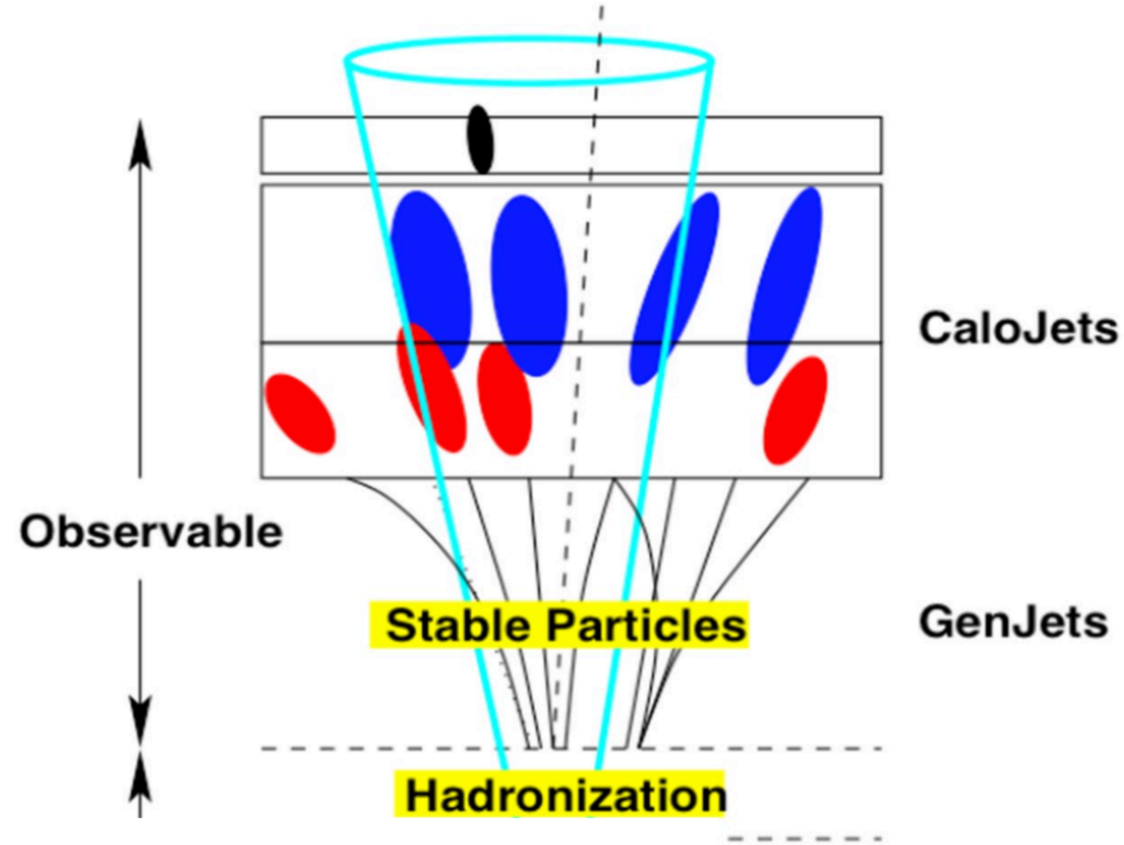




Atkin (2015)

What is a jet in particle physics?



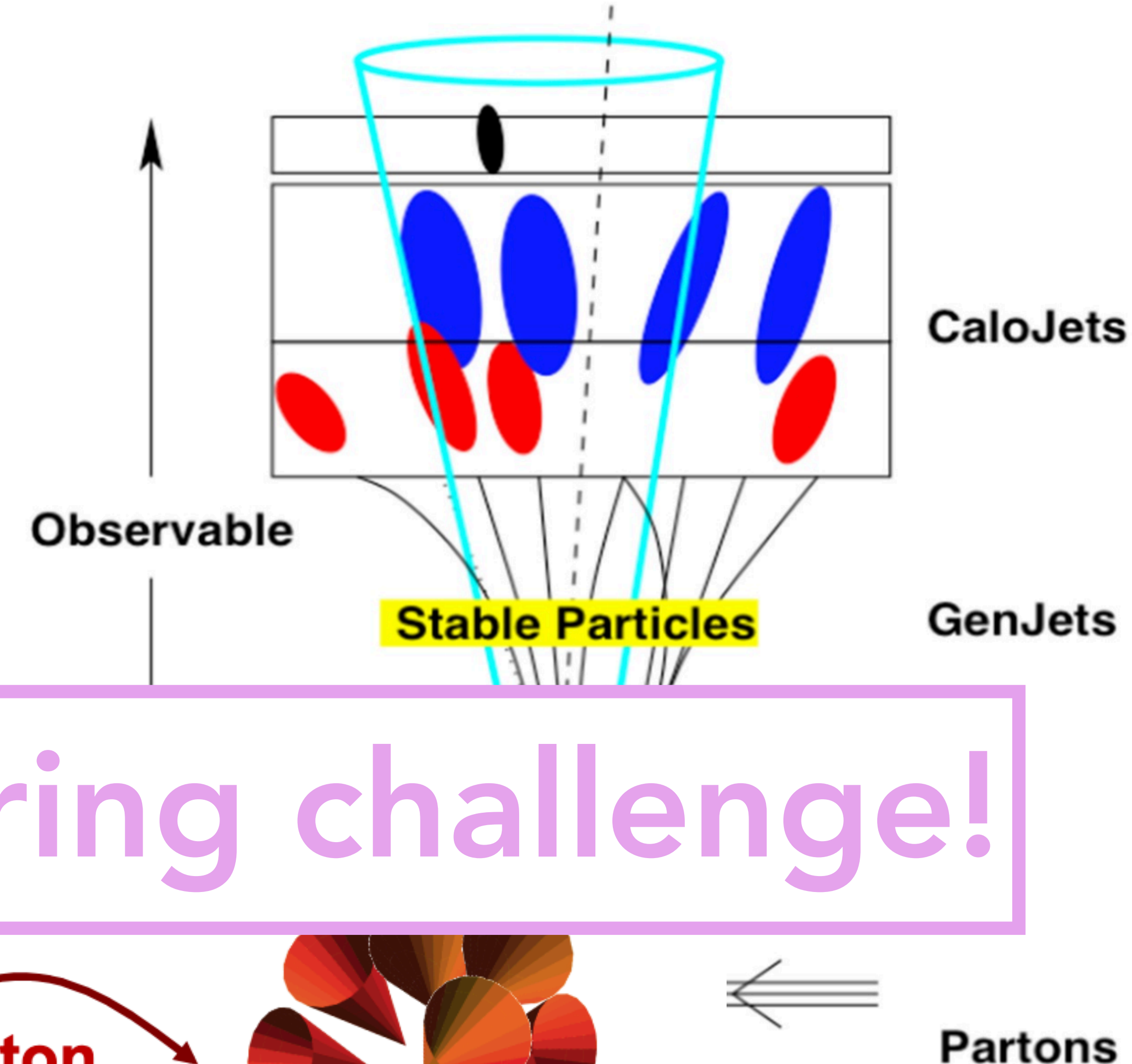


Energetic quarks and gluons produced

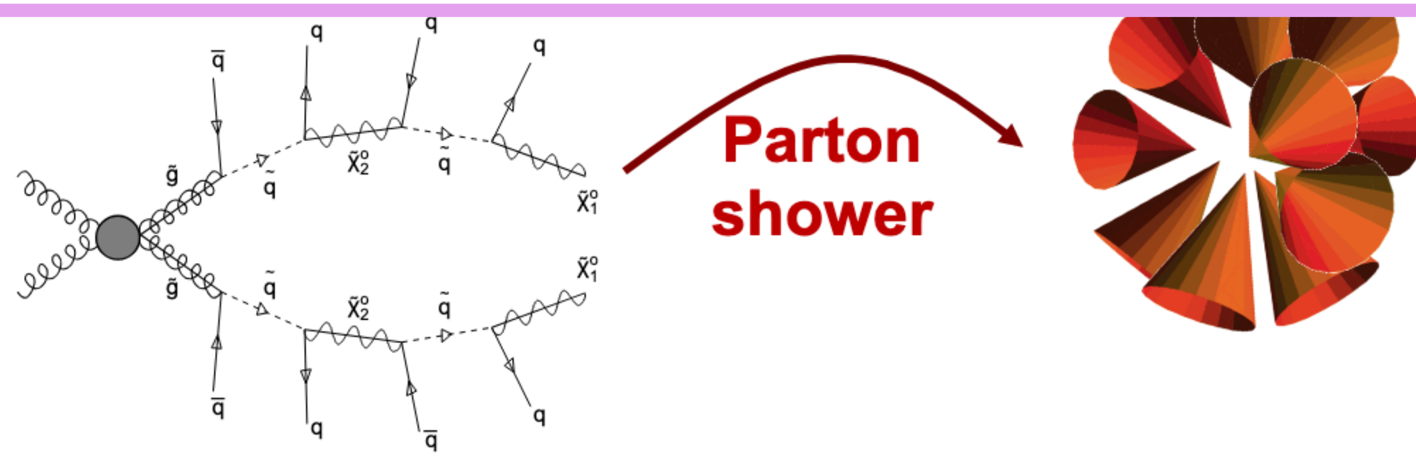
Quarks and gluons “shower” to form jets

Schwartz (2013)

What is a jet in particle physics?



Jets are a clustering challenge!



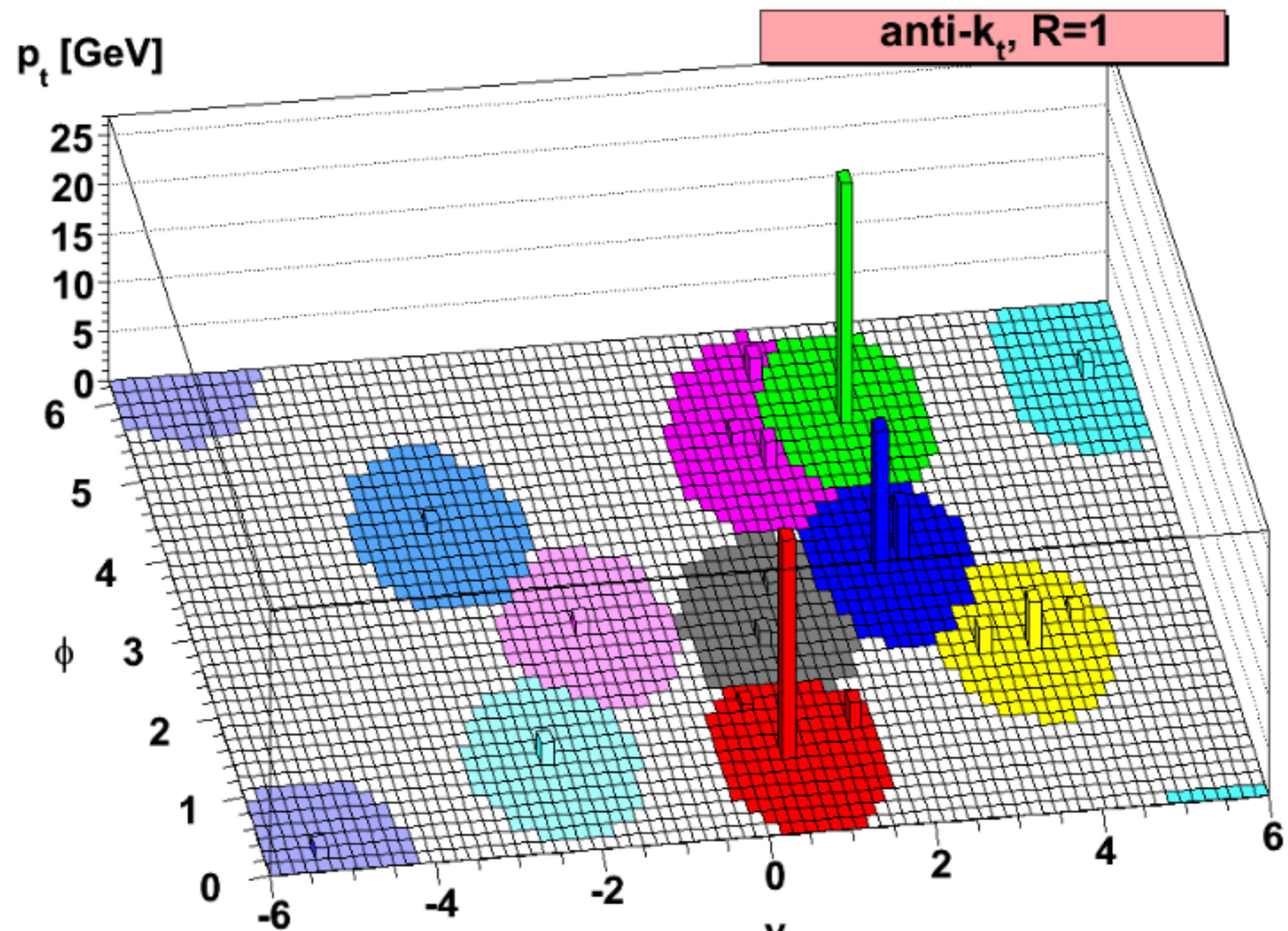
Quarks and gluons “shower” to form jets

Commonly-Used Clustering Methods

- Cone *size parameter* and/or *multiplicity* typically *specified in advance*

- *Different algorithms* for *clustering* and *substructure analysis*

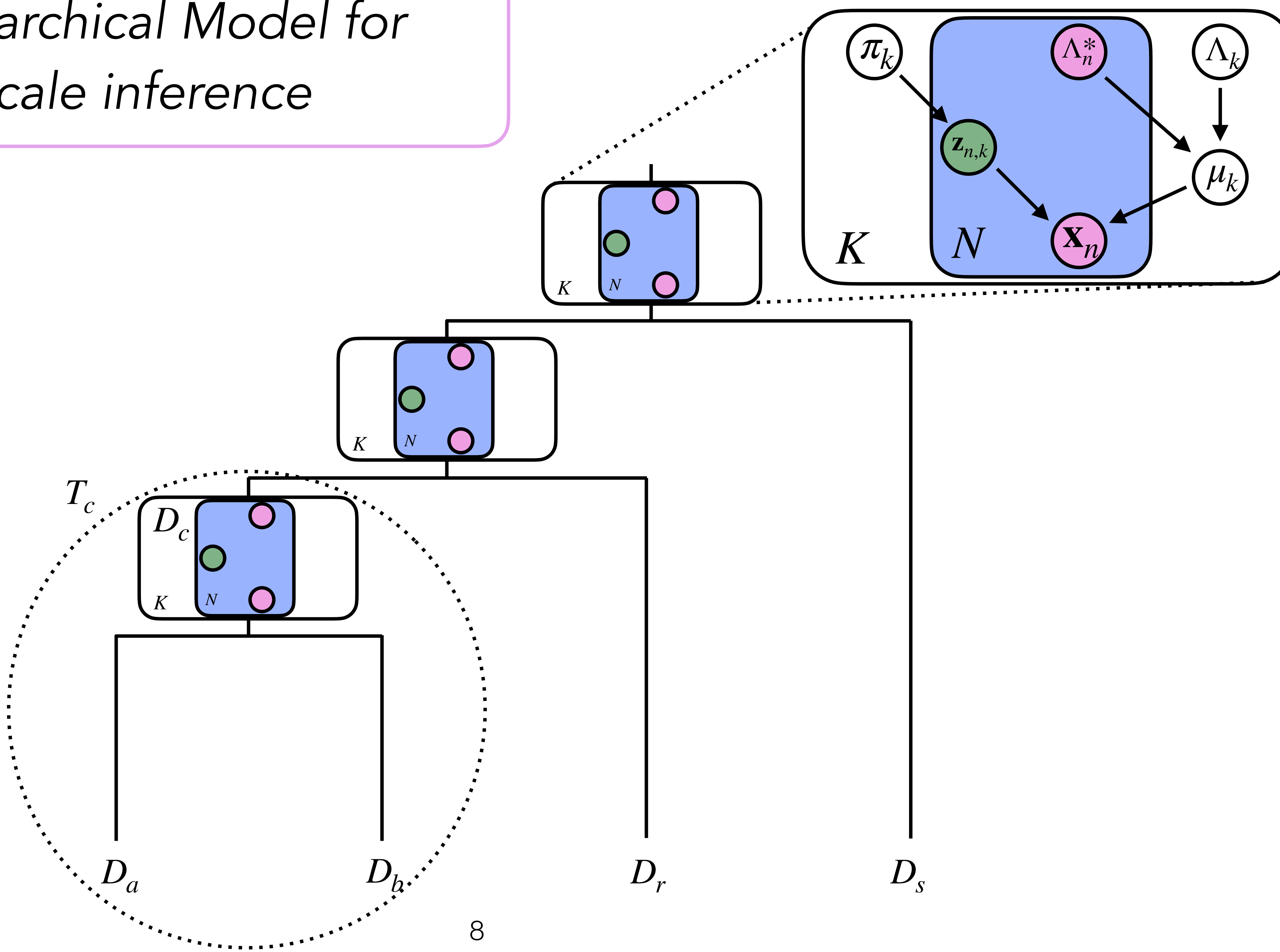
- *Distance-based metrics* for clustering in *space only*, *not straightforward to incorporate time*



Cacciari, et. al. (2008)

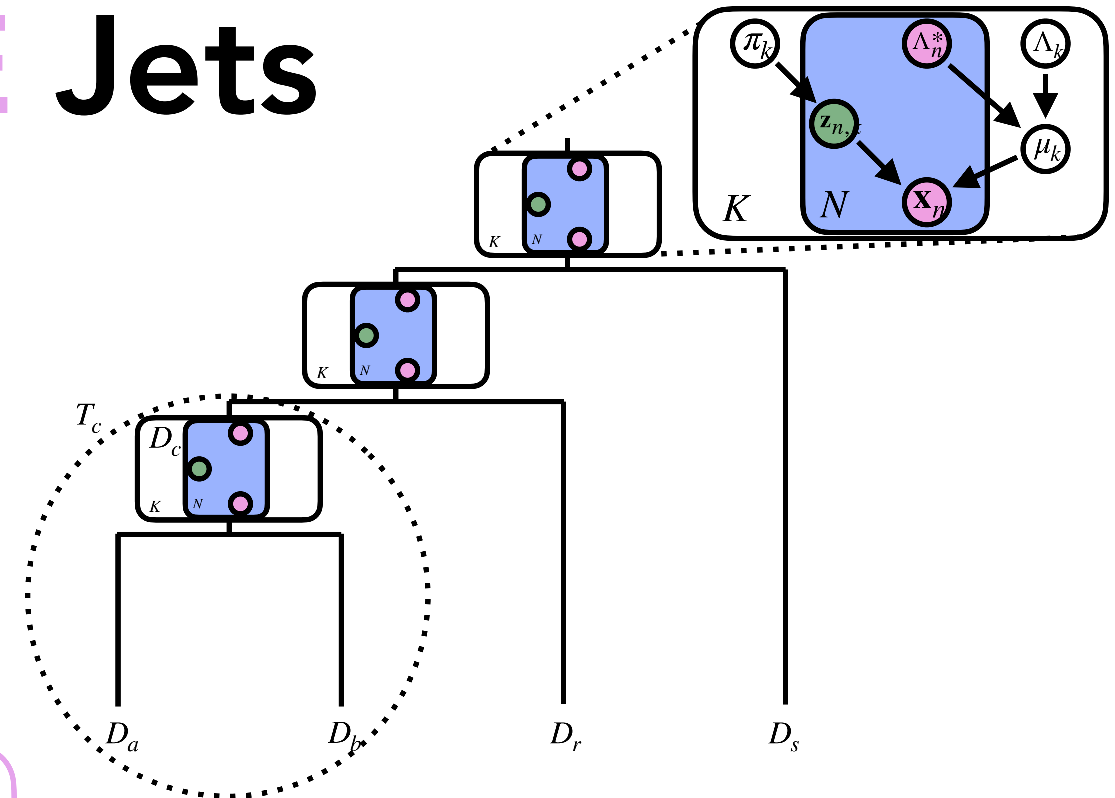
PSICHE Jets

Hierarchical, Probabilistic models:
*mixture of Gaussians **embedded** within Hierarchical Model for multi-scale inference*



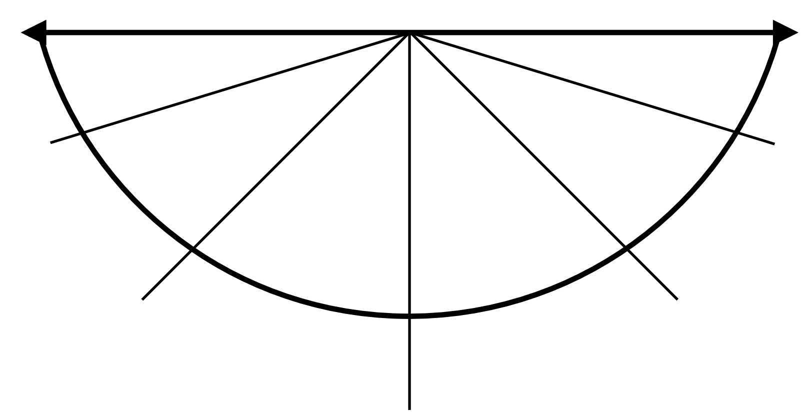
PSICHE Jets

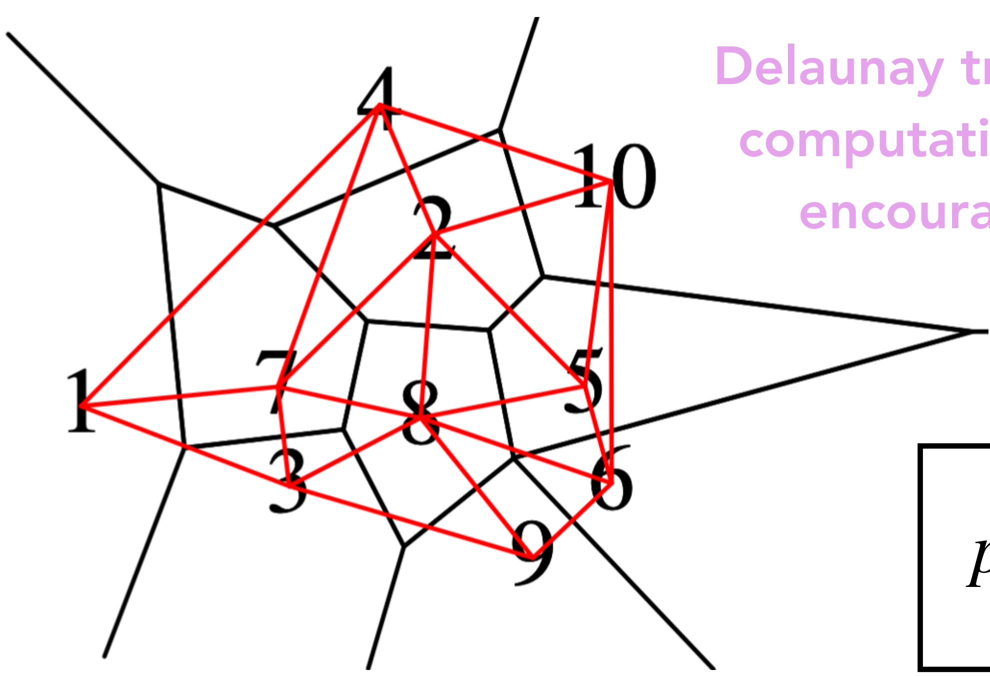
Hierarchical, Probabilistic models: *mixture of Gaussians embedded within Hierarchical Model for multi-scale inference*



Structure-Intrinsic: ***domain-specific inductive biases*** *guide unsupervised learning of physically-meaningful scales*

Angular spatial projection incentivizes jet locality





Delaunay triangulation supports computational tractability and encourages jet continuity

Cacciari, Salam (2006)

Bayesian priors reinforce detector geometries and effects

$$p(\boldsymbol{\mu}\,|\,\boldsymbol{\Lambda})p(\boldsymbol{\Lambda}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{\mu}_k\,|\,\mathbf{m}_0, (\beta_0\boldsymbol{\Lambda}_k)^{-1}, \boldsymbol{\Lambda}_n^*)\mathcal{W}(\boldsymbol{\Lambda}_k\,|\,\mathbf{W}_0, \nu_0)$$

PSICHE Jets

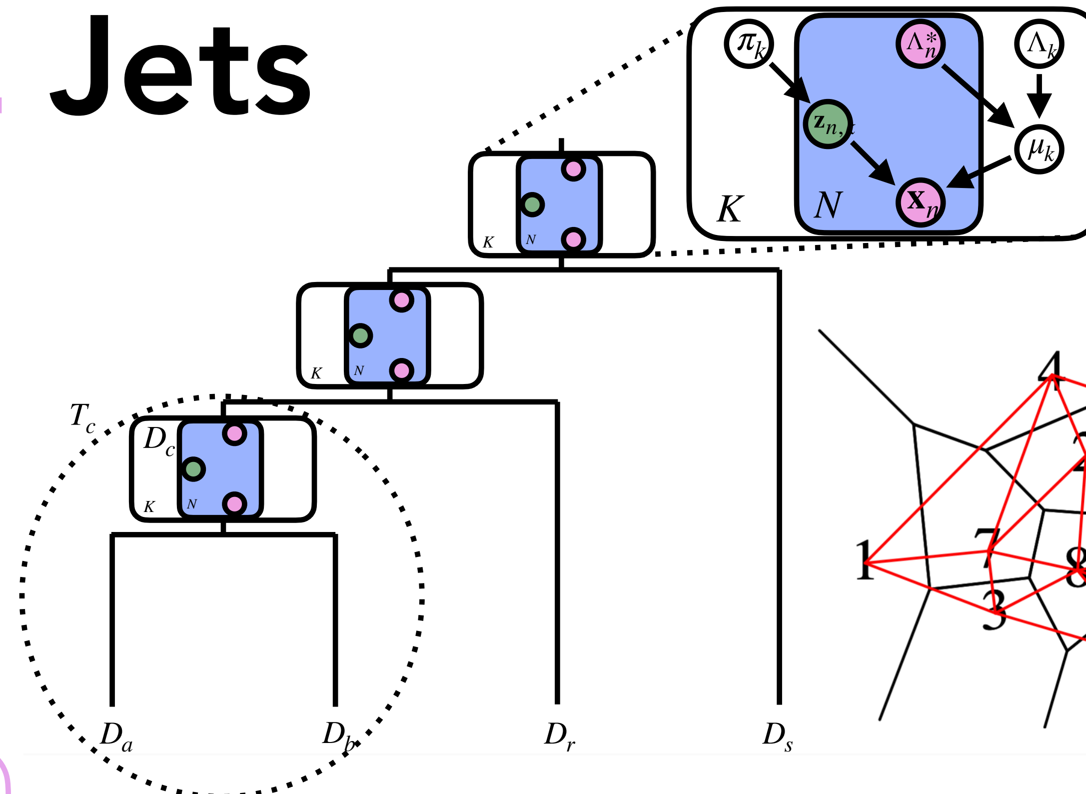
Hierarchical, Probabilistic models: *mixture of Gaussians* ***embedded*** *within Hierarchical Model for multi-scale inference*

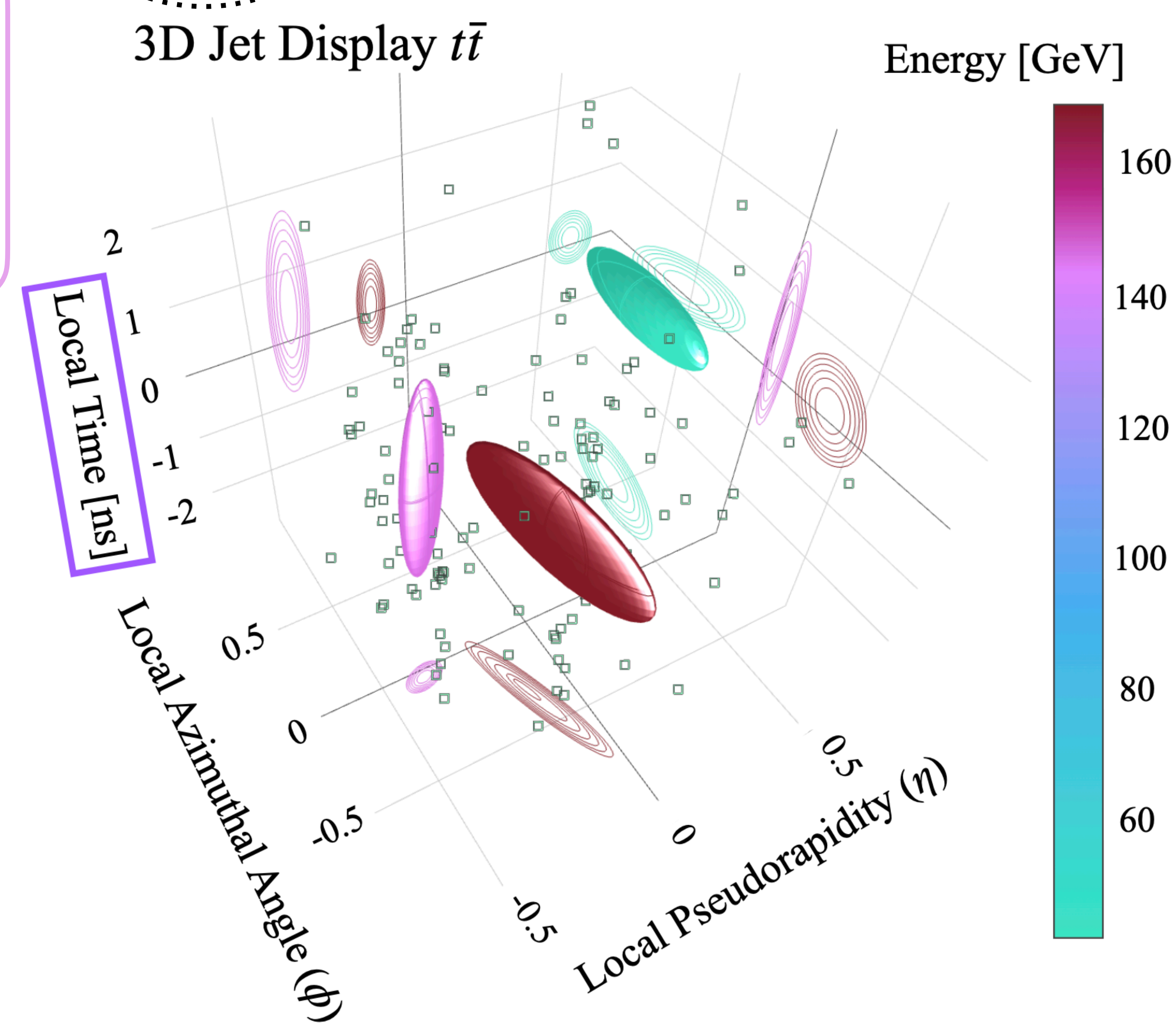
Structure-Intrinsic: ***domain-specific inductive biases*** *guide unsupervised learning of physically-meaningful scales*

Timing-aware: *time integrated appropriately alongside space into model likelihood*

Λ_n^*

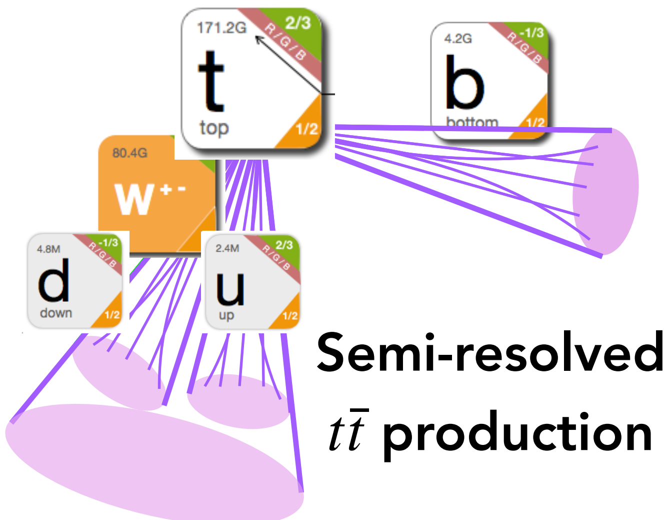
Physically-motivated uncertainties to incorporate nonuniform, experimental resolutions





KU THE UNIVERSITY OF KANSAS

PSICHE Jets: Emergent Features

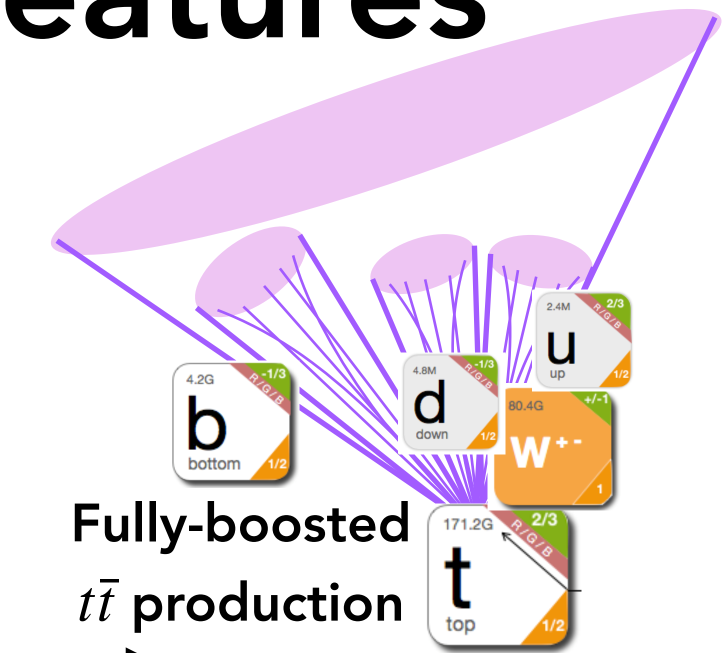


Semi-resolved
$t\bar{t}$ production

Dynamic scale

$$\text{jet size} \propto \frac{m}{p_T}$$

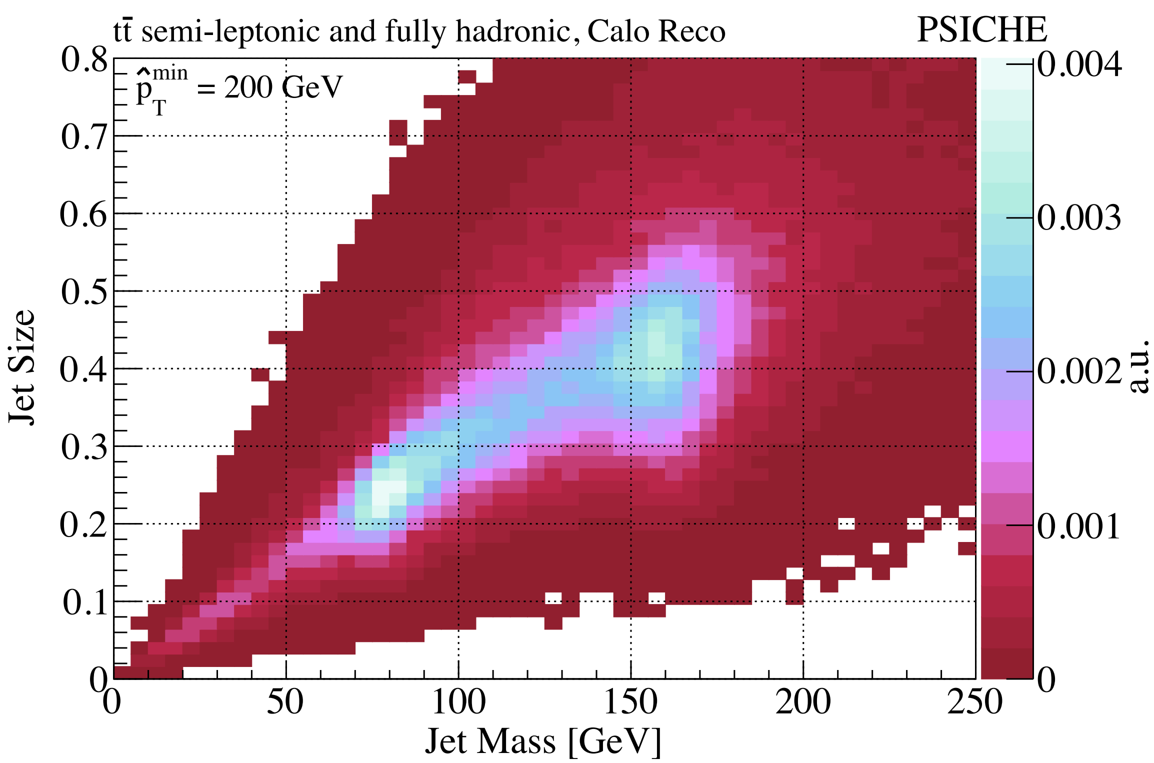
Higher momentum systems

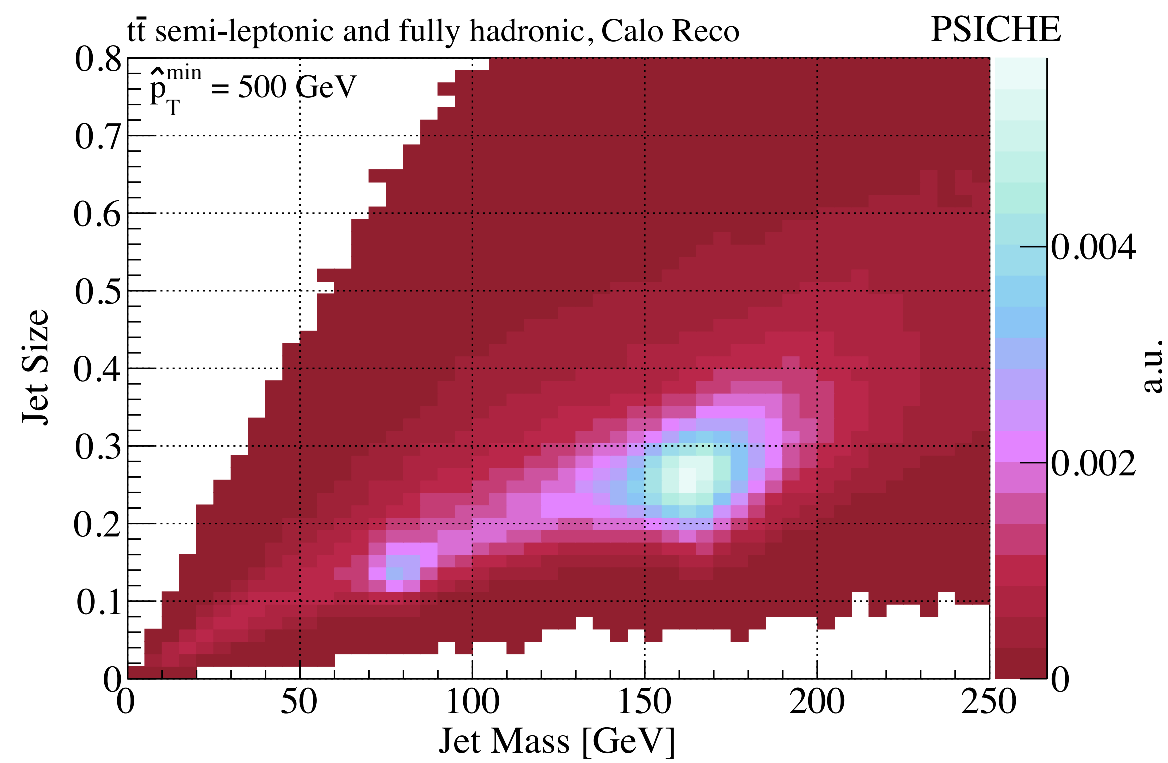


Fully-boosted
$t\bar{t}$ production

Accommodates less collimated boosts with larger size

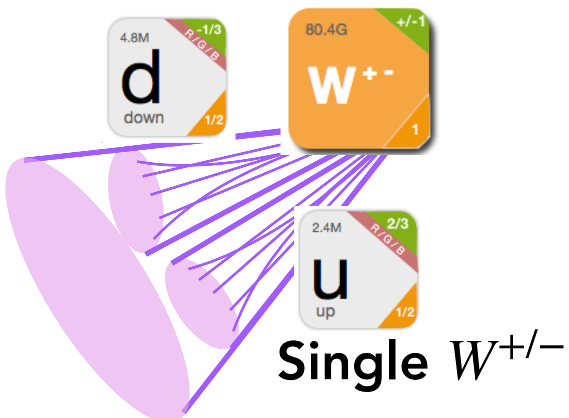
Accommodates more collimated boosts with smaller size



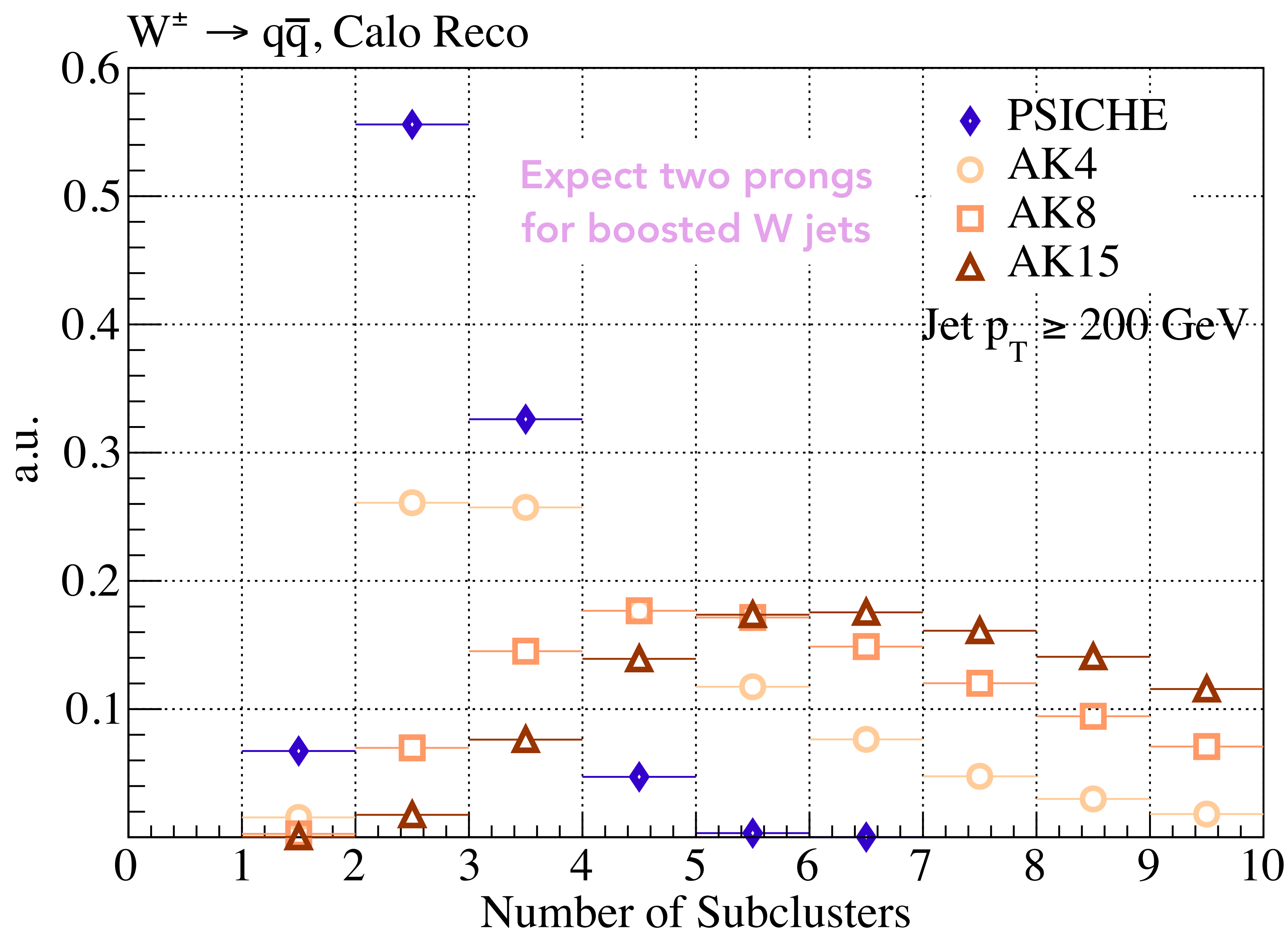


Dynamically reconstructs resonances at *different boost regimes*, without a hard cutoff scale or fine-tuning probabilistic hyperparameters

PSICHE Jets: Emergent Features



Multi-scale learning

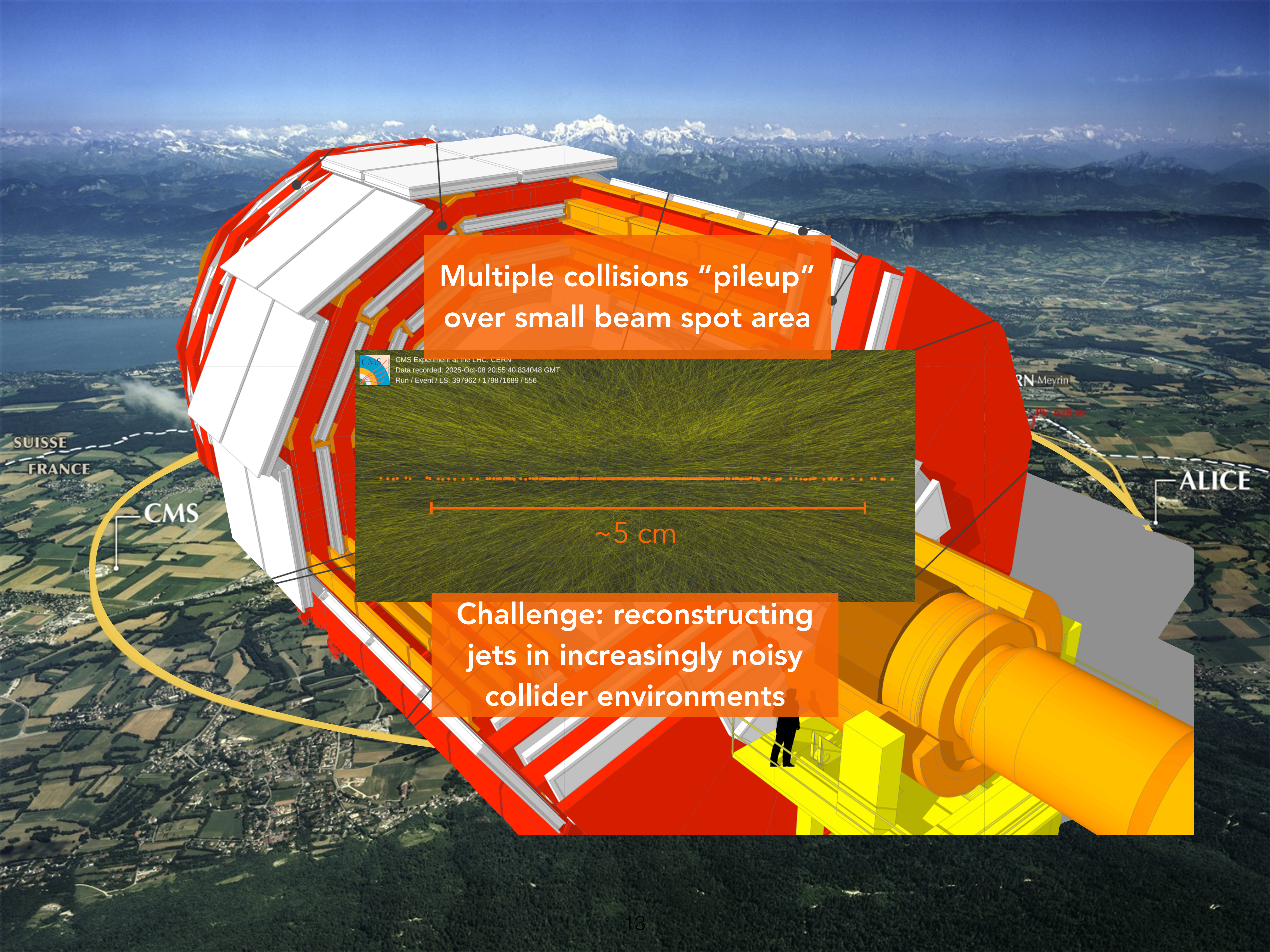


Co-learning of scales facilitated by **nested algorithm design**

Can deploy mixture model on clustered jets, **need pre-determined amount,** scale of learning is not well-defined

PSICHE is learns **parsimonious jet substructure** by design with **regularization mechanics**

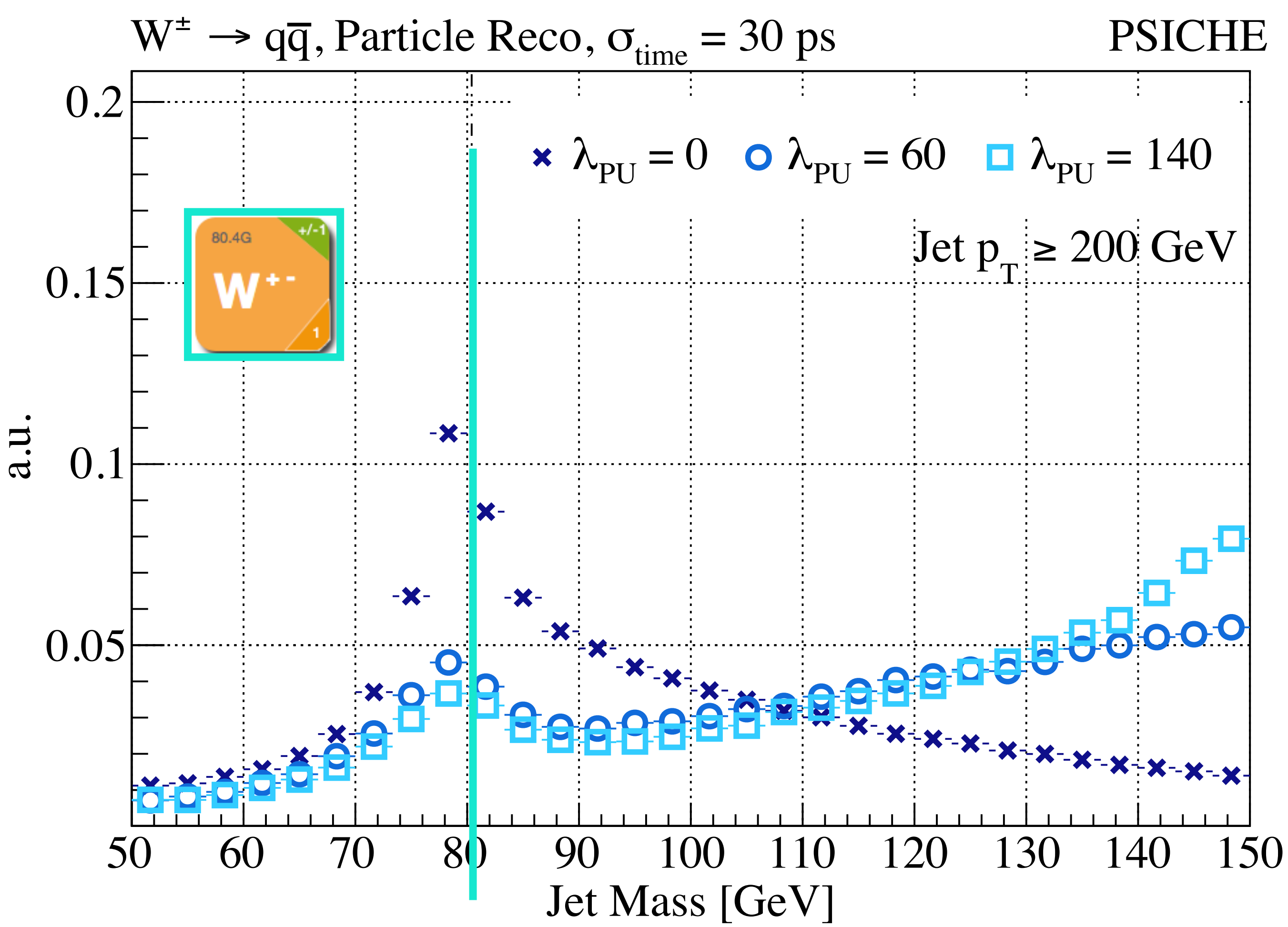
Simultaneously **learns *physically-meaningful,* probabilistically-interpretable jet *substructure***

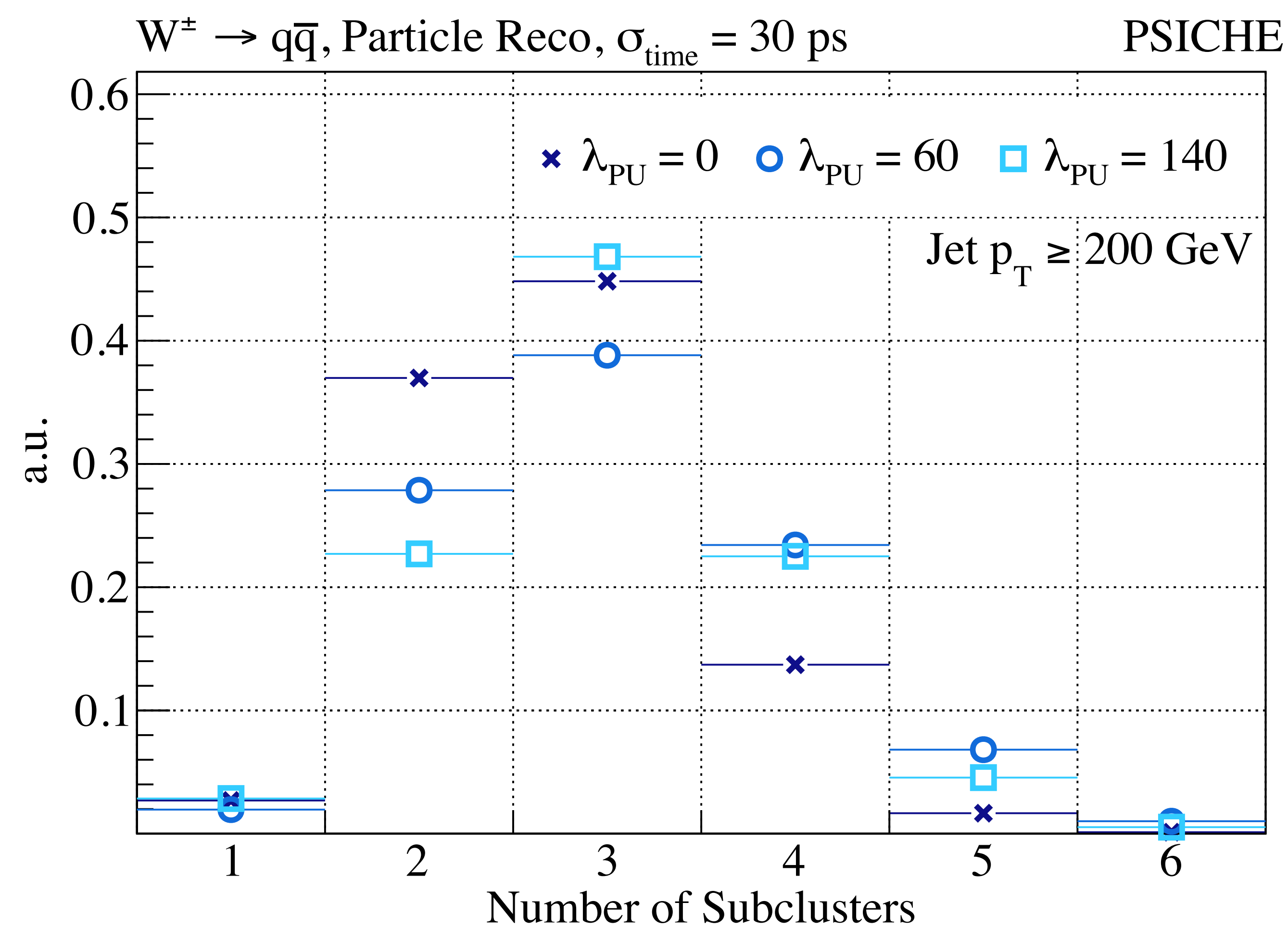


Multiple collisions "pileup" over small beam spot area

Challenge: reconstructing jets in increasingly noisy collider environments

PSICHE Jets:
In the Presence of Pileup





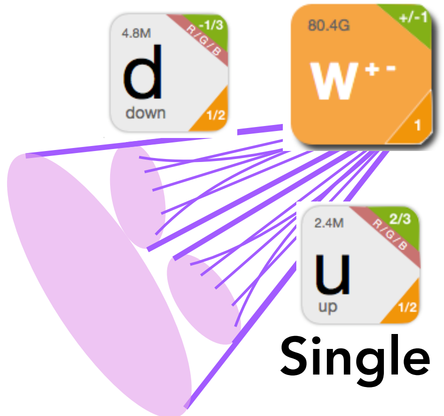
Identifying jet features become less clear at higher levels of pileup (PU)

× Ideal ○ Current □ Projected

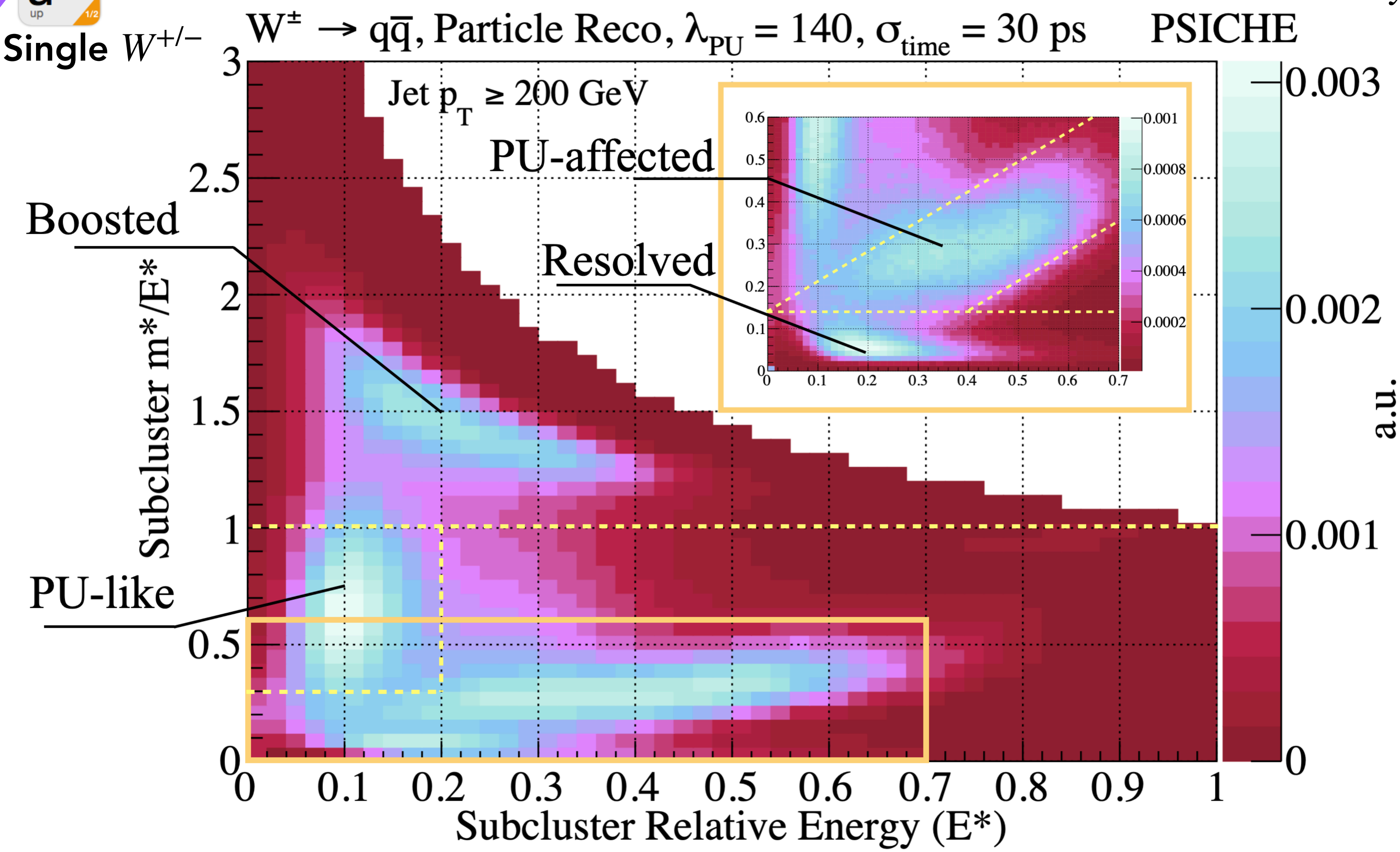
PSICHE Jets: In the Presence of Pileup

Multi-scale learning

$$x^* = \frac{x^{\text{subcluster}}}{x^{\text{jet}}}$$



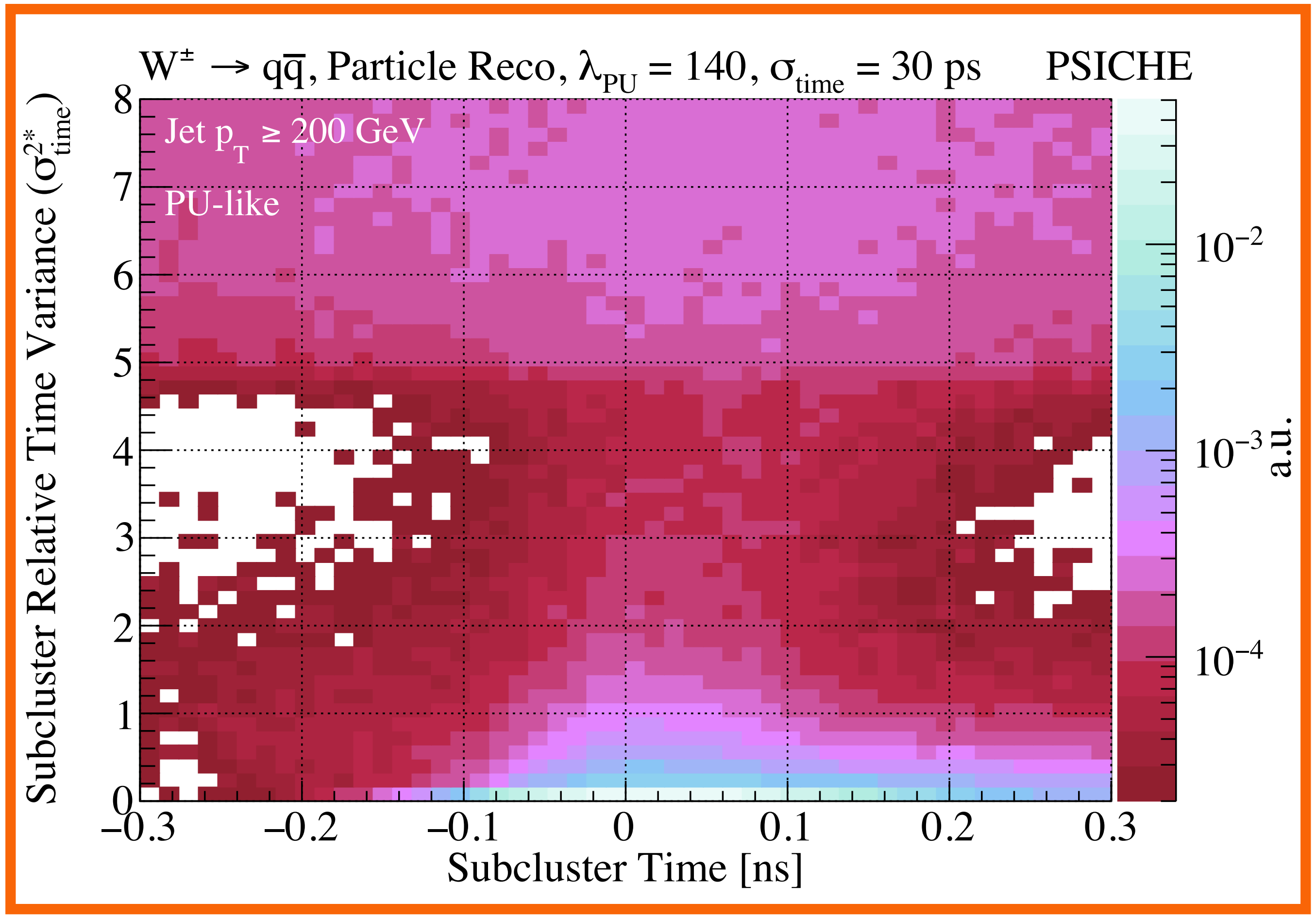
Single $W^{+/-}$

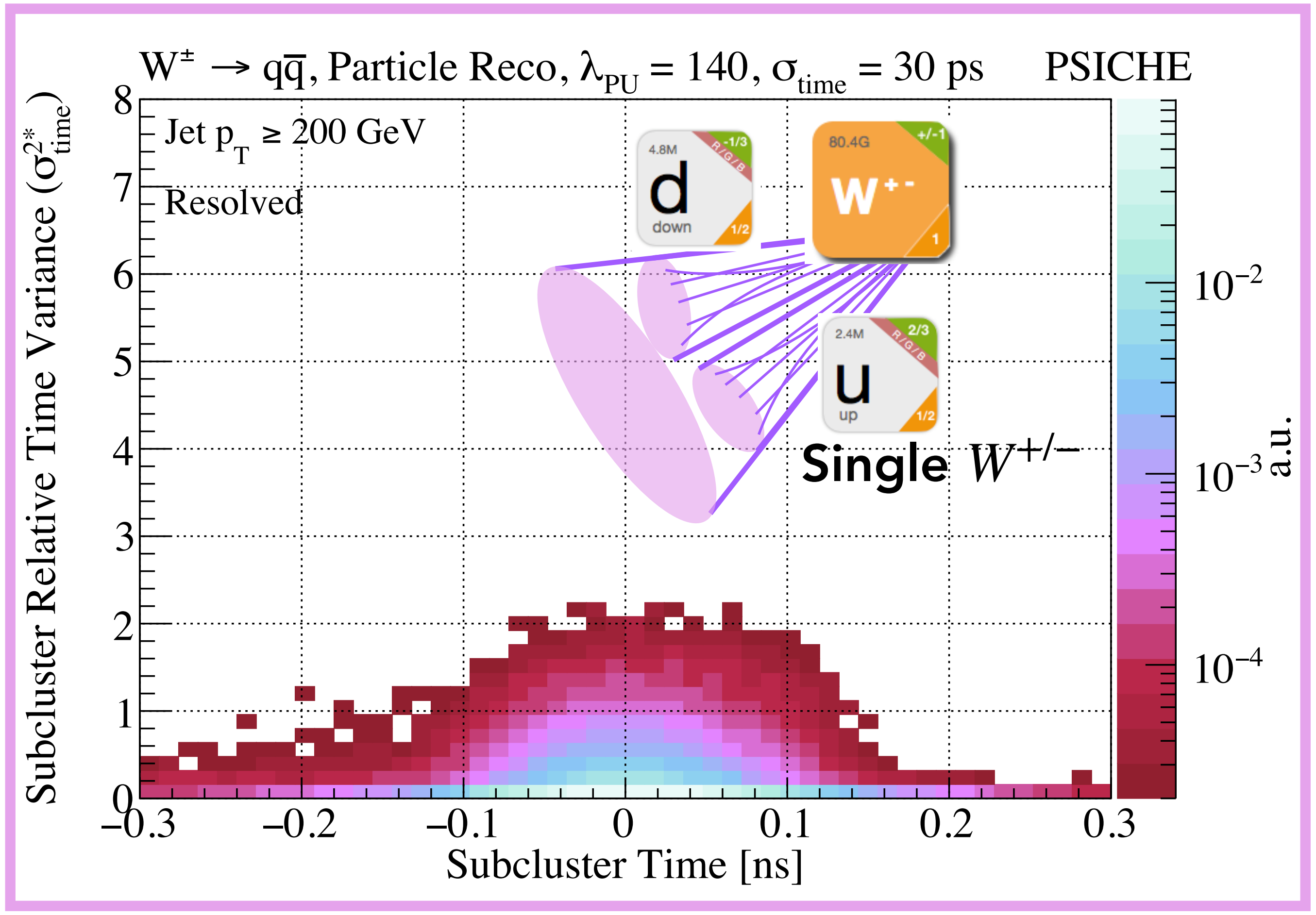


Jet substructure topologies can be *characterized* based on subclusters' *relative kinematic contributions* to respective jets

PSICHE Jets:
In the Presence of Pileup

Multi-scale learning & timing-aware

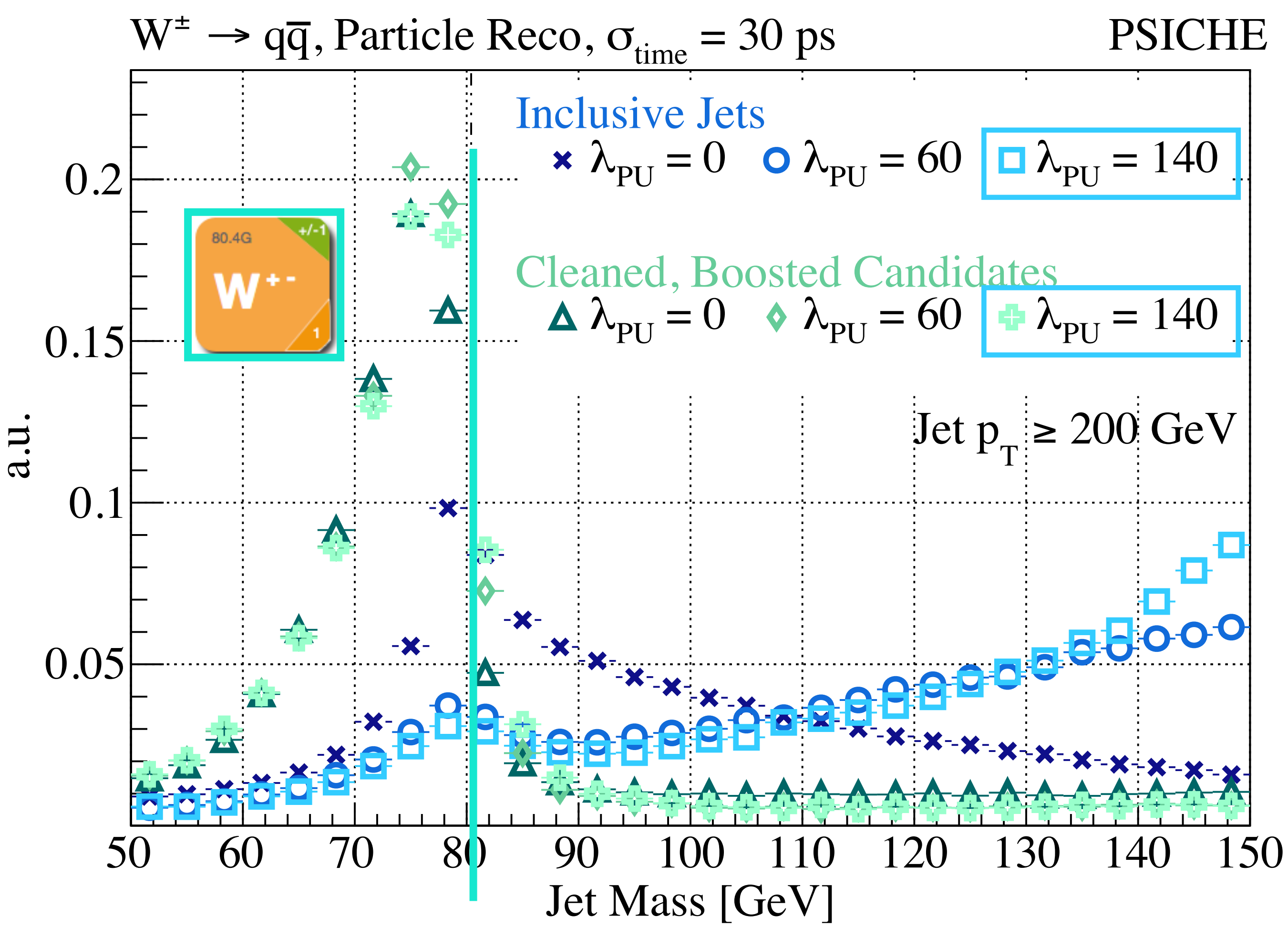


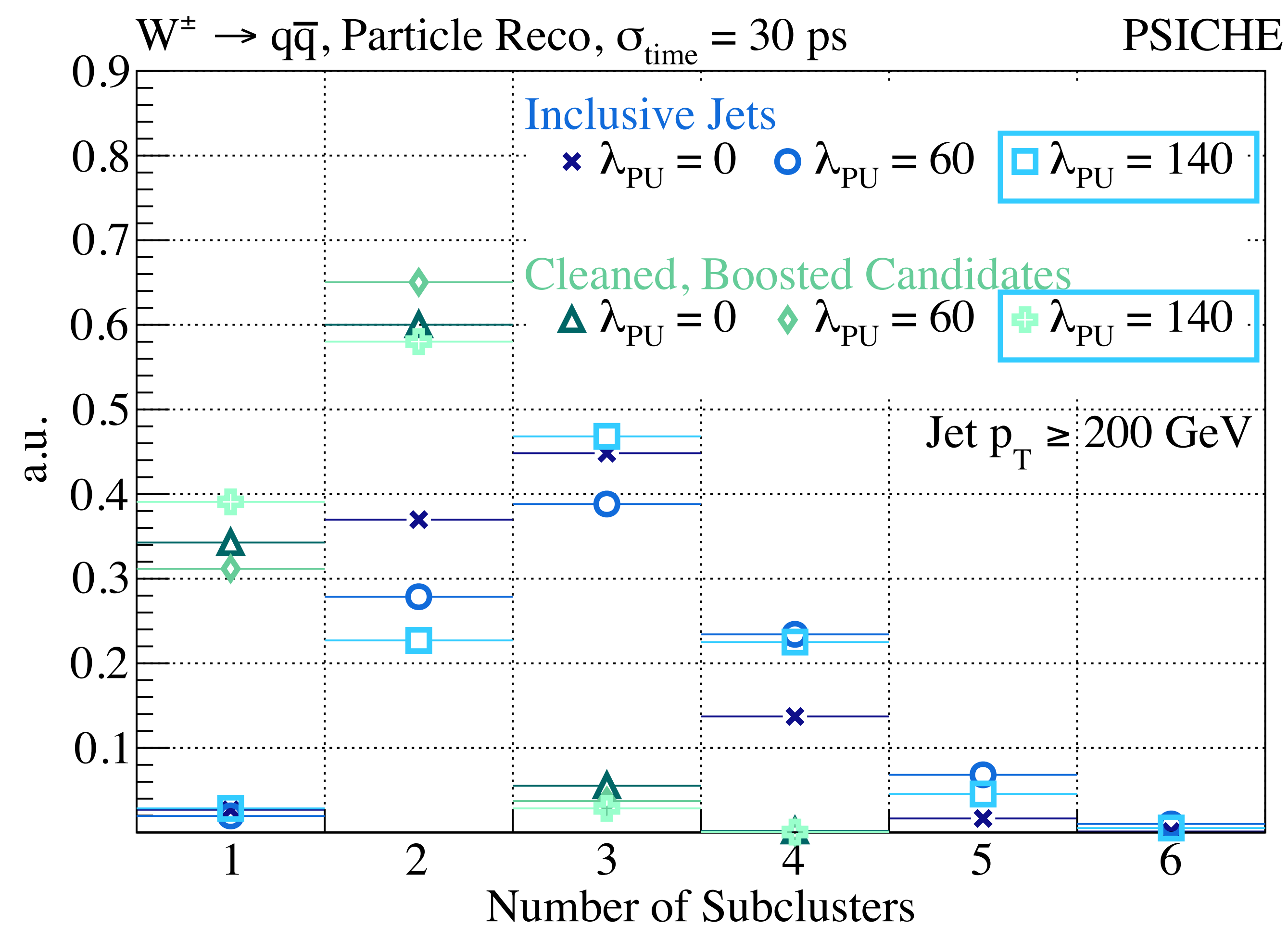


Different subcluster types exhibit **different, characteristic spreads** in **space** and **time**

Addition of timing allows **separation of noise** from **physically-meaningful jet substructure**

PSICHE Jets:
In the Presence of Pileup





*Can leverage **time-** and kinematic-based subcluster selection to clean and recover jet features at **extreme PU levels (HL-LHC)***

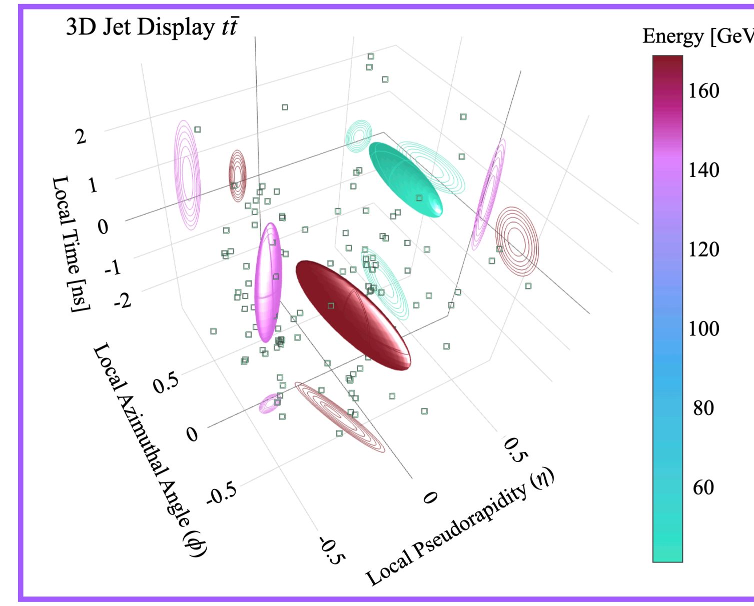
PSICHE: Looking to the Future

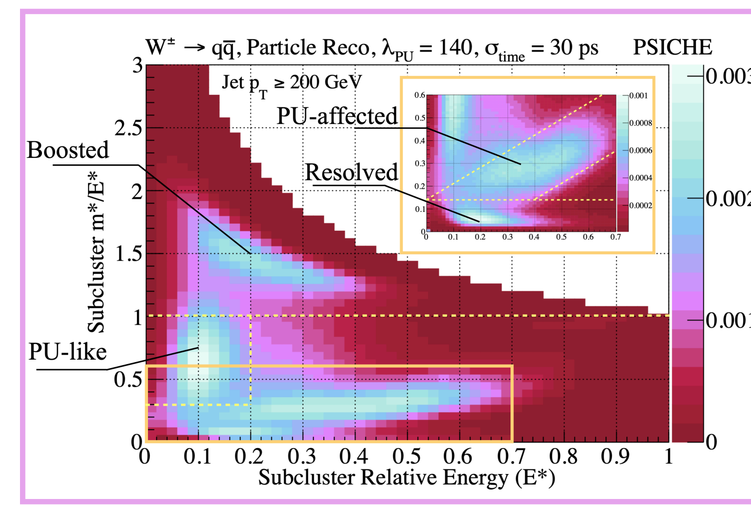
Algorithmic Future Work

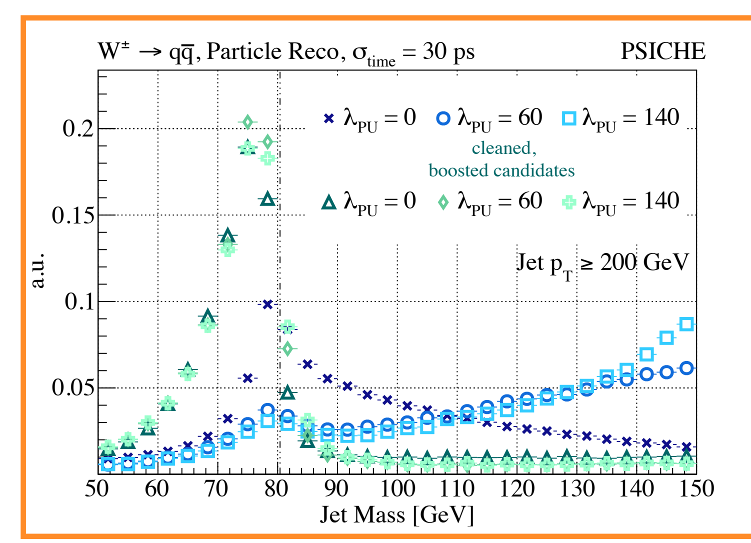
- Experiment with incorporating **additional scales** as nested hierarchies
- Apply algorithm design in **different domains** with appropriate inductive biases

Physics Future Work

- Using dynamically-learned substructure for **jet identification**
- Exploring jet reconstruction with PSICHE in searches for **LLPs/hidden valley** scenarios



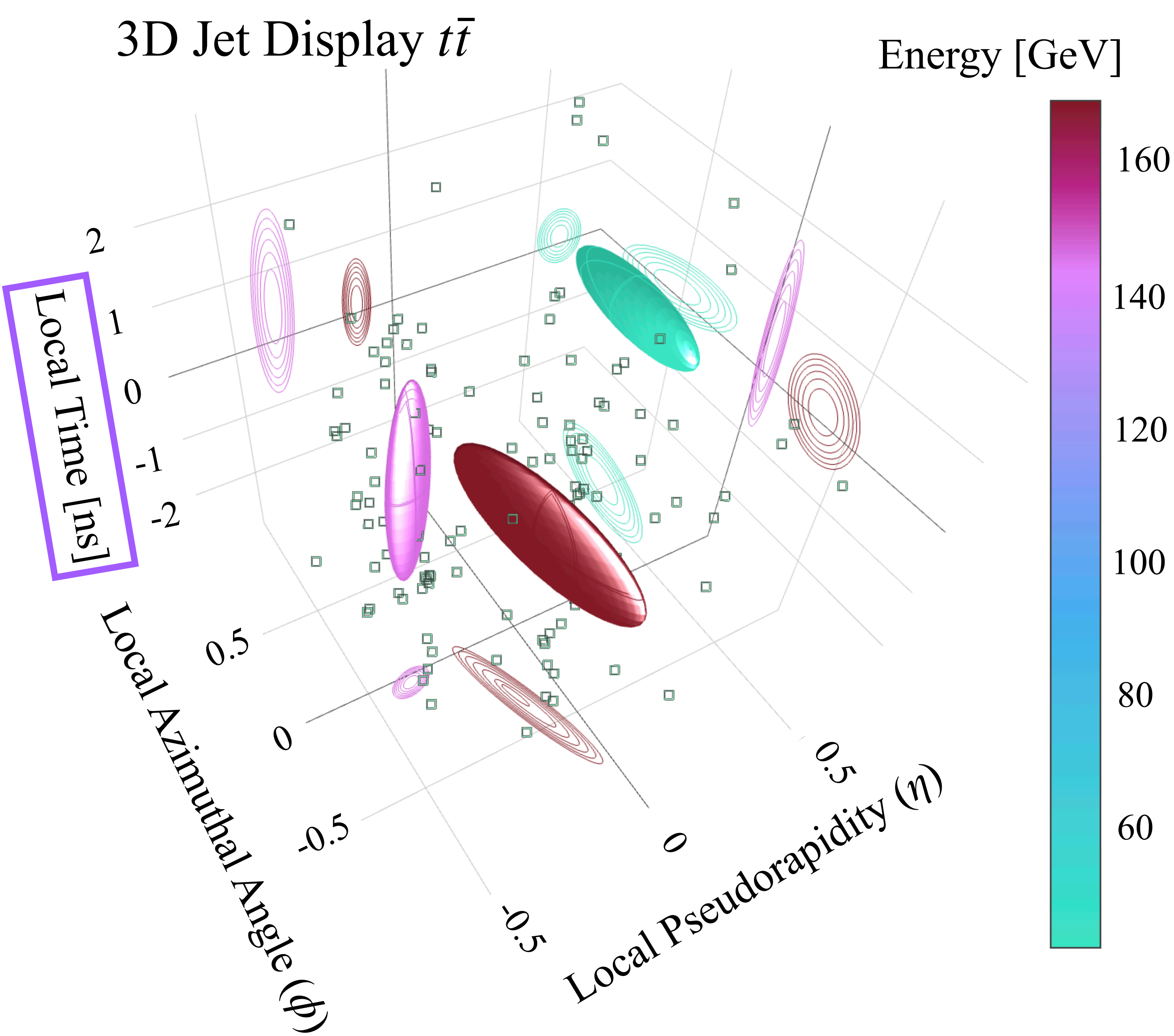




*paper submitted to arXiv, listed in ~day!

PSICHE Jets

Probabilistic, Structure-Intrinsic Clustering with an Hierarchical Embedding

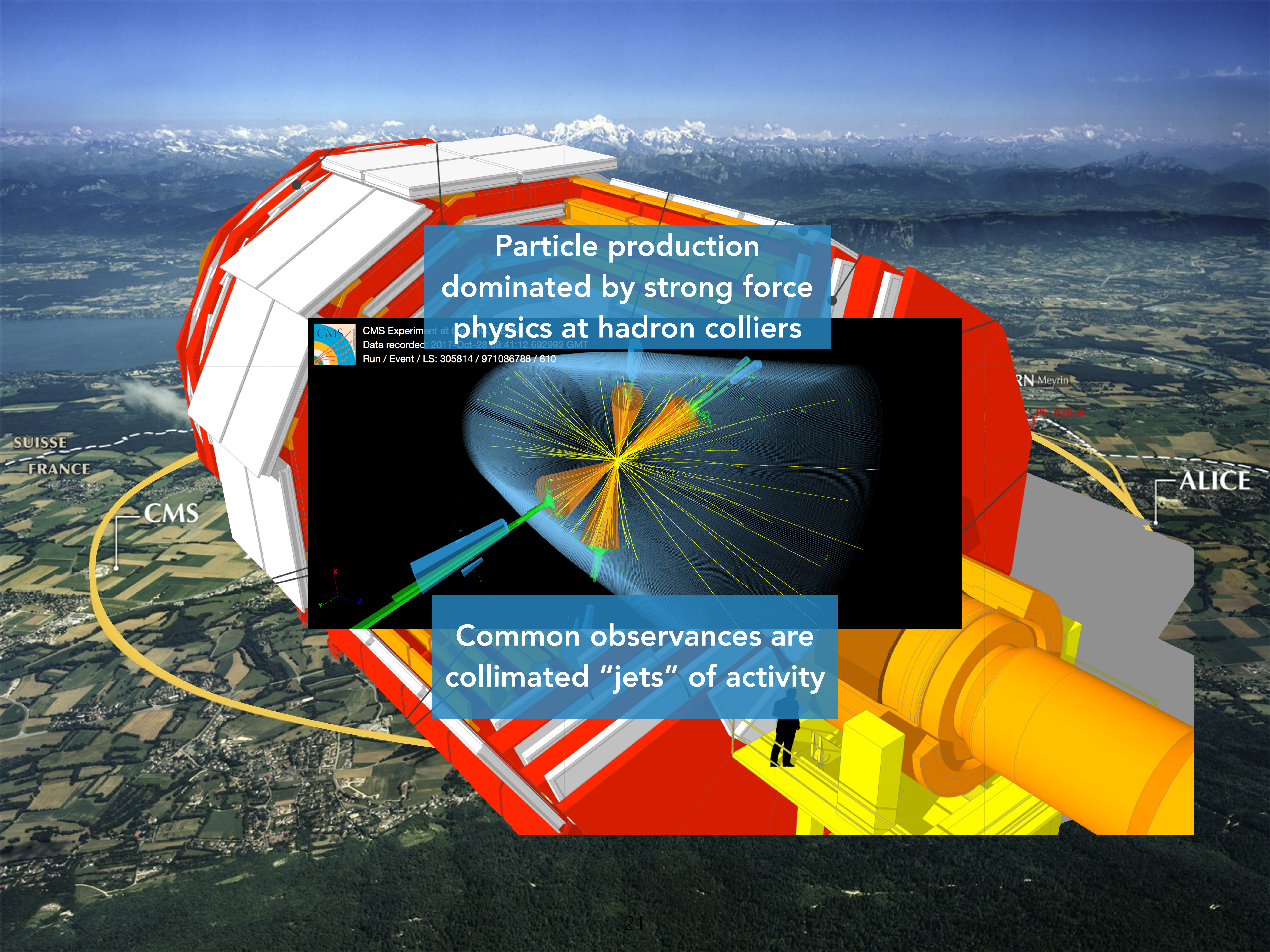


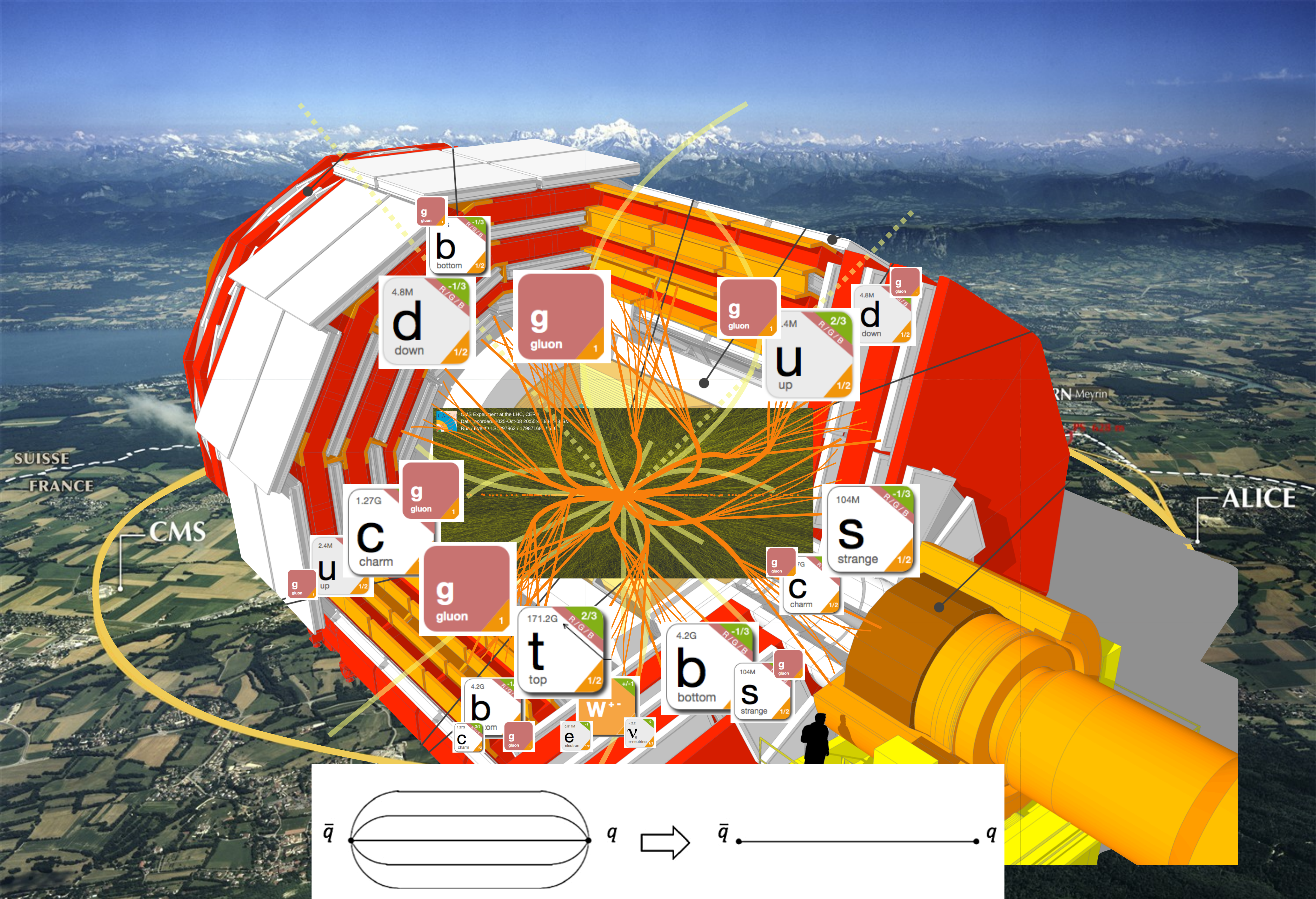
Example jet display of three-prong top quark decay, with jet and substructure scales learned by PSICHE

- Presented hierarchically-embedded, probabilistic algorithm infused with **domain-specific inductive biases**
- **Self-consistent framework** with various emergent features: dynamic jet sizes, multi-scale learning, timing-aware clustering
- Timing and kinematic information learned can be leveraged for **collider challenge** of pileup mitigation

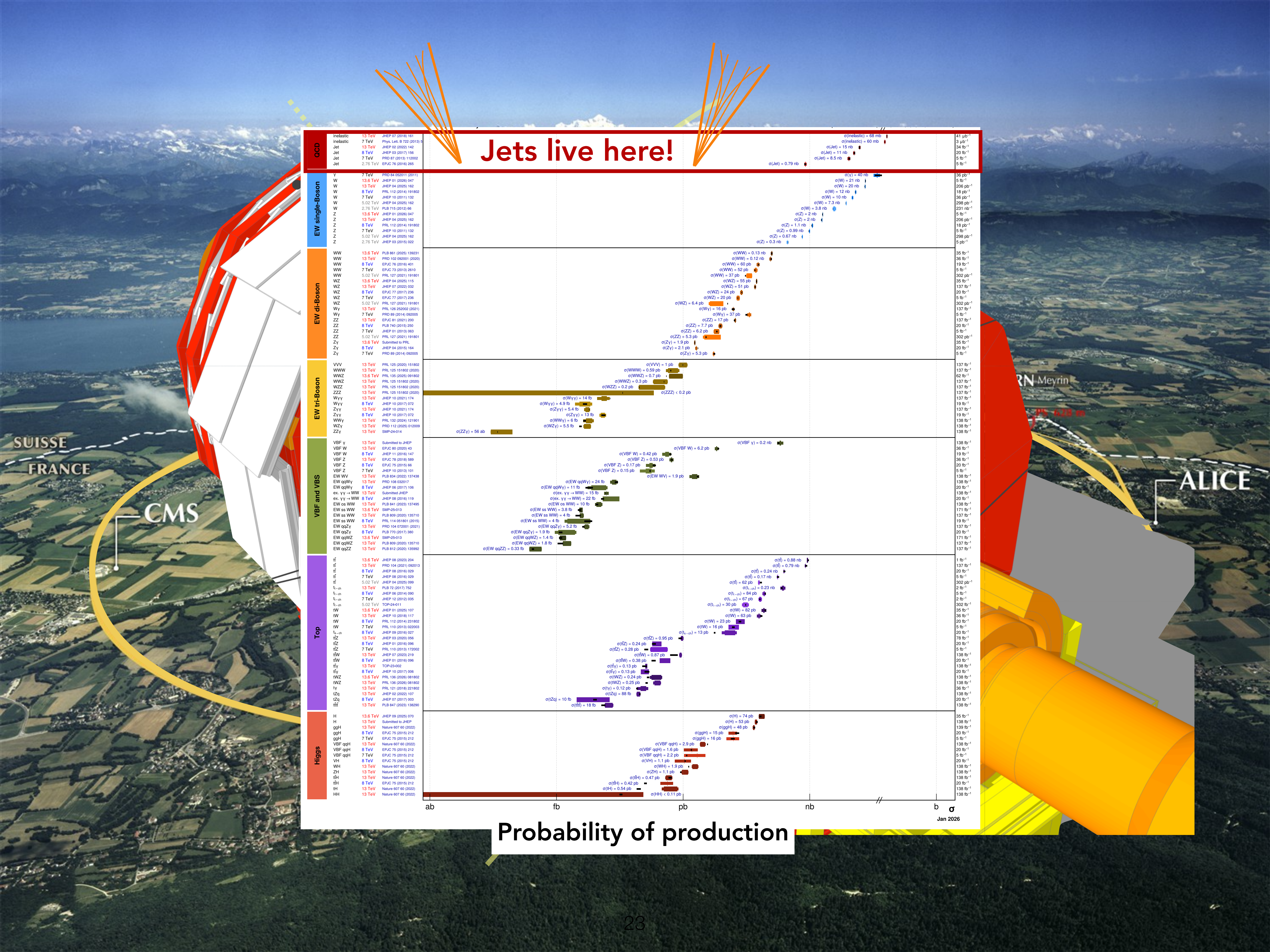
THE UNIVERSITY OF KANSAS

Backup


Particle production dominated by strong force physics at hadron colliers
CMS Experiment at
Data recorded: 2017-Oct-28 09:41:12.692992 GMT
Run / Event / LS: 305814 / 971086788 / 610
Common observances are collimated “jets” of activity
SUISSE
FRANCE
CMS
ALICE
Meyrin

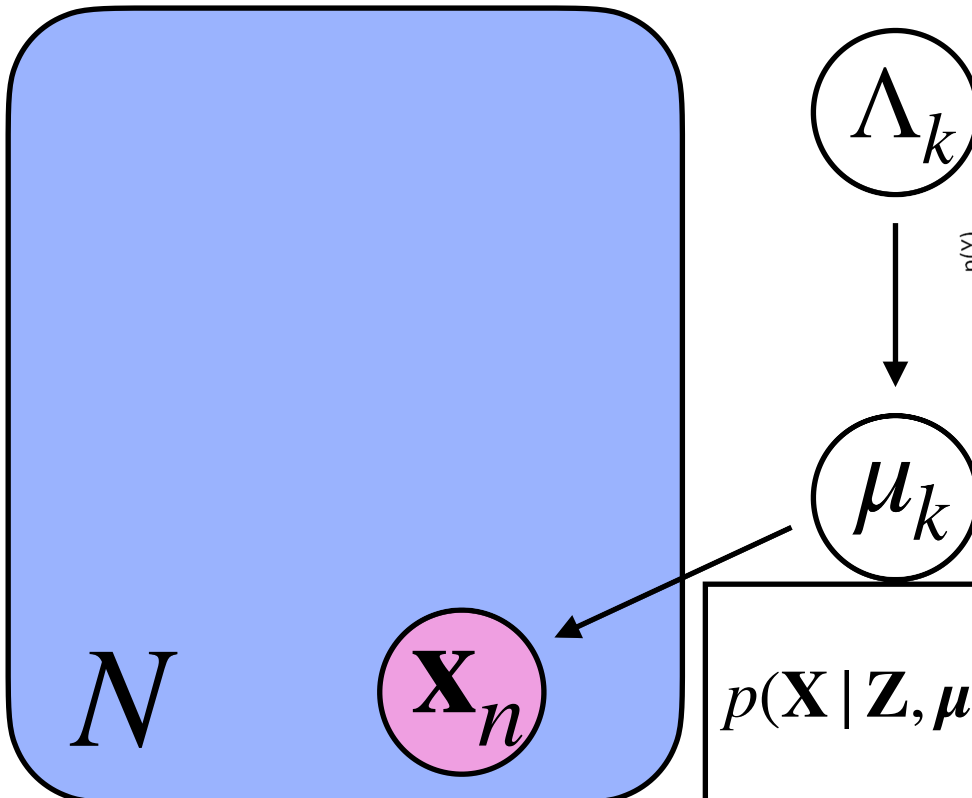


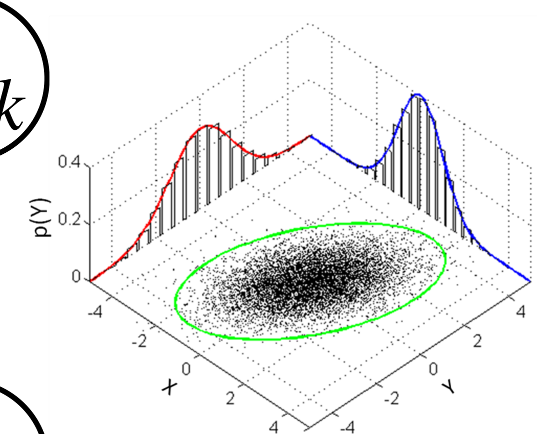
$$V_{\mathrm{QCD}}(r) \approx -\frac{4}{3}\,\frac{\alpha_s}{r} + \kappa\, r$$

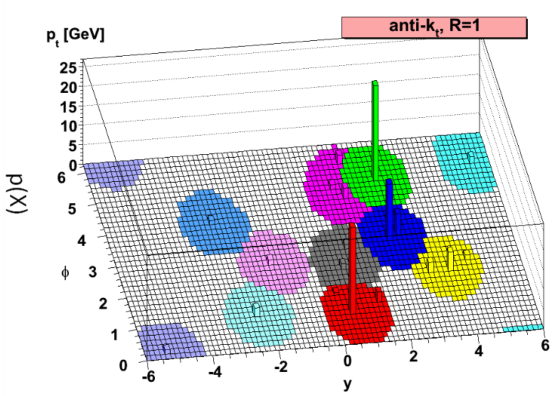

Jets live here!
SUISSE
FRANCE
CMS
ALICE
QCD
EW single-Boson
EW di-Boson
EW tri-Boson
VBF and VBS
Top
Higgs
ab
fb
pb
nb
b
σ
Jan 2026
Probability of production

PSICHE Jets

Probabilistic**: Bayesian statistics forms basis for **scale-dynamic learning







$$p(\mathbf{X}\,|\,\mathbf{Z},\boldsymbol{\mu},\boldsymbol{\Lambda}) = \prod_{n=1}^{N}\prod_{k=1}^{K}\mathcal{N}(\mathbf{x}_n\,|\,\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k^{-1})^{z_{nk}}$$

Input four vector

Gaussian likelihood: location and spread of subclusters

Inputs: energy-weighted position four-vector

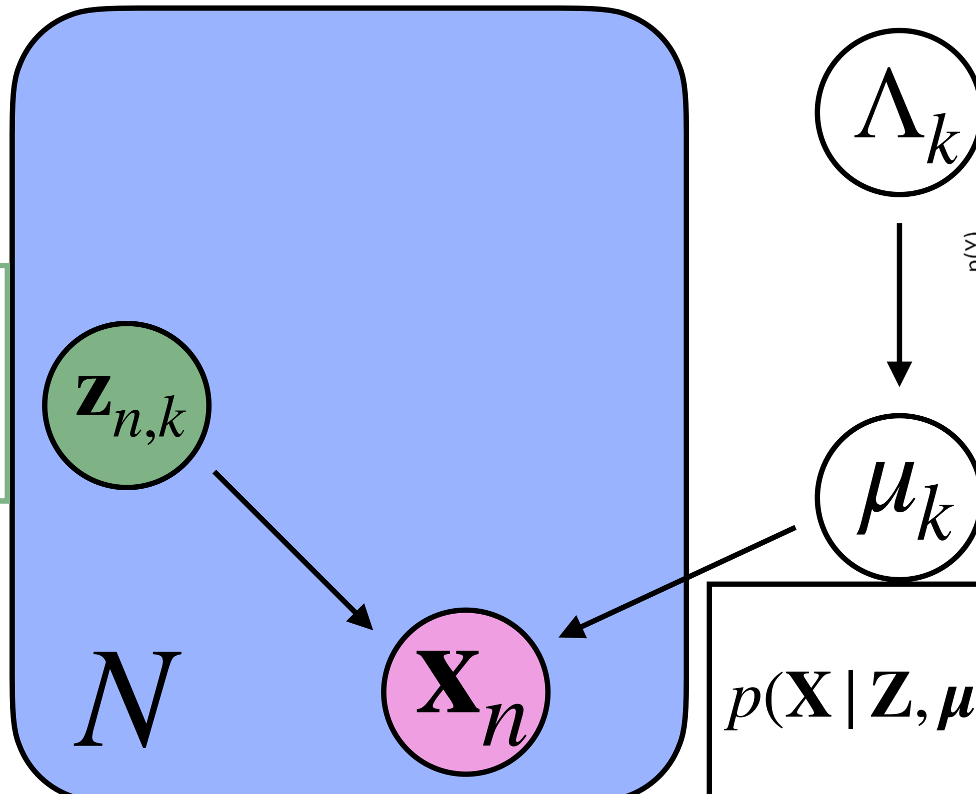
Output: mixture model clusterings

PSICHE Jets

Probabilistic**: Bayesian statistics forms basis for **scale-dynamic learning

$$p(\mathbf{Z} \mid \boldsymbol{\pi}) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{nk}}$$

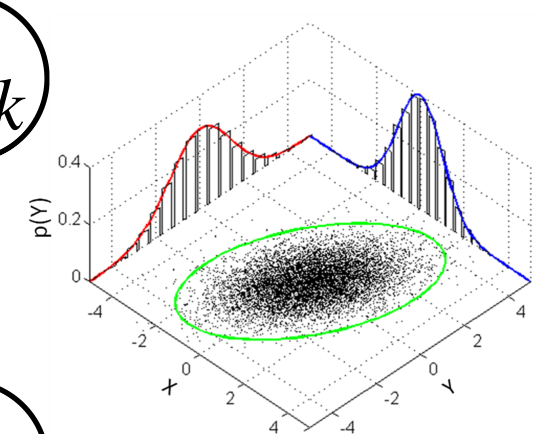
Latent assignment variable

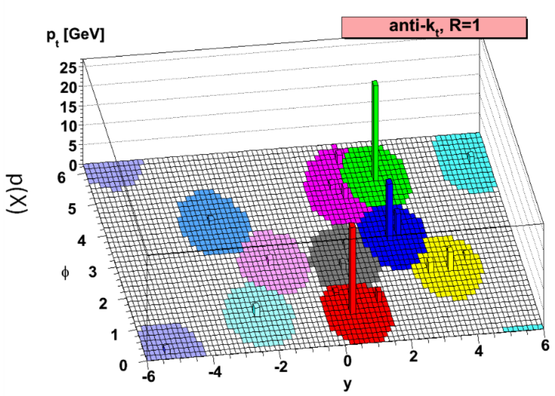


Input four vector

$$p(\mathbf{X} \mid \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \prod_{n=1}^{N} \prod_{k=1}^{K} \mathcal{N}(\mathbf{x}_n \mid \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1})^{z_{nk}}$$

Gaussian likelihood: location and spread of subclusters





Inputs: energy-weighted position four-vector

Output: mixture model clusterings

PSICHE Jets

Probabilistic**: Bayesian statistics forms basis for **scale-dynamic learning

Prior: uncertainty on location and spread

$$p(\boldsymbol{\mu}\,|\,\boldsymbol{\Lambda})p(\boldsymbol{\Lambda}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{\mu}_k\,|\,\mathbf{m}_0, (\beta_0\boldsymbol{\Lambda}_k)^{-1} \quad)\mathcal{W}(\boldsymbol{\Lambda}_k\,|\,\mathbf{W}_0, \nu_0)$$

$$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}\,|\,\boldsymbol{\alpha}_0)$$

Prior: uncertainty in number of subclusters

π_k

$$p(\mathbf{Z}\,|\,\boldsymbol{\pi}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \pi_k^{z_{nk}}$$

Latent assignment variable

$\mathbf{z}_{n,k}$

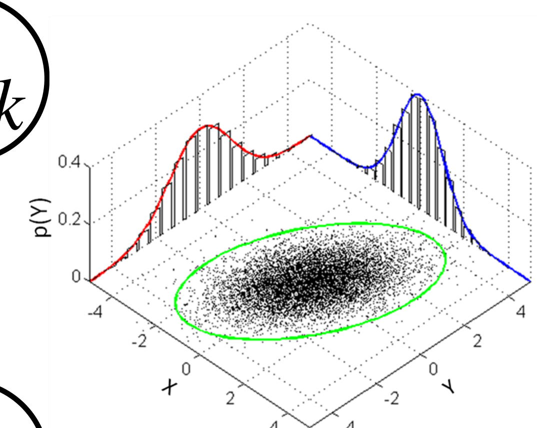
N

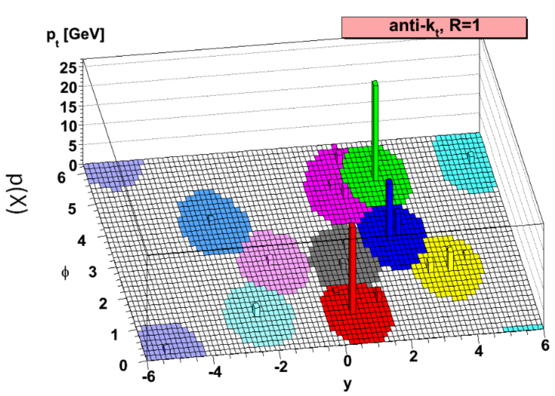
$\mathbf{x}_n$

Input four vector

Λ_k

μ_k





$$p(\mathbf{X}\,|\,\mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \mathcal{N}(\mathbf{x}_n\,|\,\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1})^{z_{nk}}$$

Gaussian likelihood: location and spread of subclusters

Inputs: energy-weighted position four-vector and uncertainty

Output: mixture model clusterings

PSICHE Jets

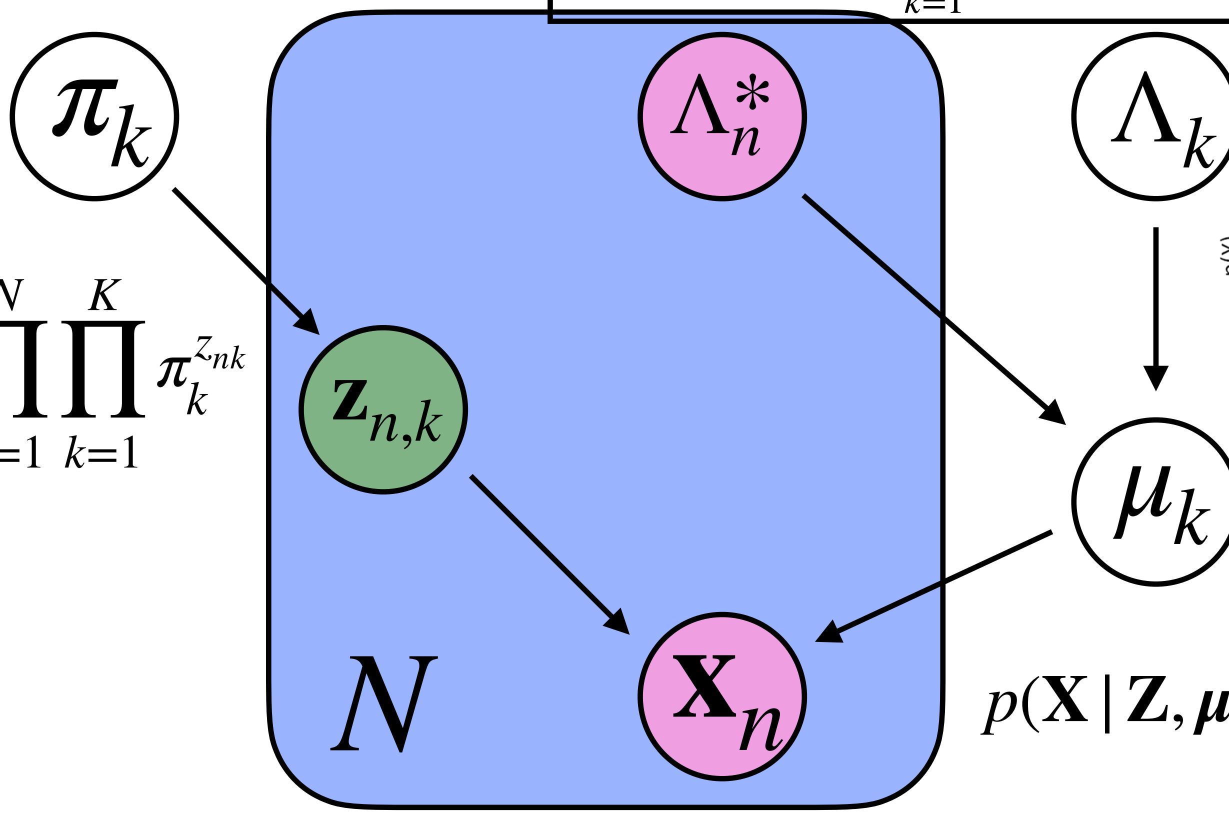
Probabilistic**: Bayesian statistics forms basis for **scale-dynamic learning

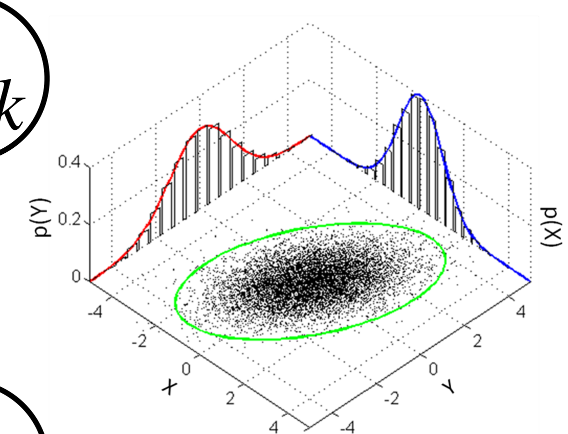
Prior: uncertainty on location and spread, measurement uncertainty

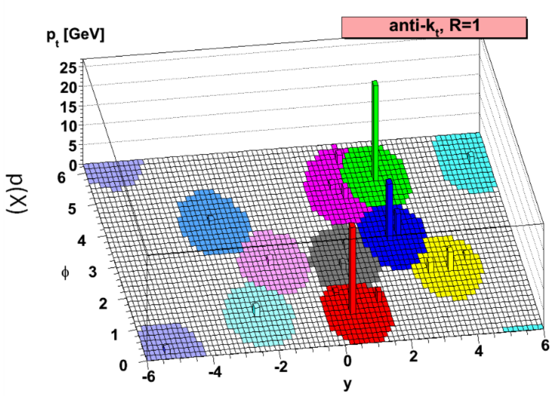
$$p(\boldsymbol{\mu}\,|\,\boldsymbol{\Lambda})p(\boldsymbol{\Lambda}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{\mu}_k\,|\,\mathbf{m}_0, (\beta_0\boldsymbol{\Lambda}_k)^{-1}, \boldsymbol{\Lambda}_n^*)\mathcal{W}(\boldsymbol{\Lambda}_k\,|\,\mathbf{W}_0, \nu_0)$$

$$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}\,|\,\boldsymbol{\alpha}_0)$$

$$p(\mathbf{Z}\,|\,\boldsymbol{\pi}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \pi_k^{z_{nk}}$$







$$p(\mathbf{X}\,|\,\mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \mathcal{N}(\mathbf{x}_n\,|\,\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1})^{z_{nk}}$$

Inputs: energy-weighted position four-vector and uncertainty

Output: mixture model clusterings

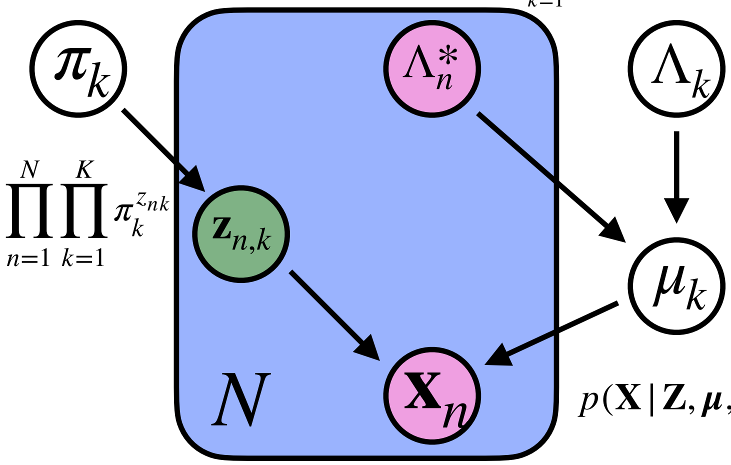
PSICHE Jets

*Hierarchical Embedding: multi-tier inference facilitates **multi-scale learning***

$p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}\,|\,\boldsymbol{\alpha}_0)$

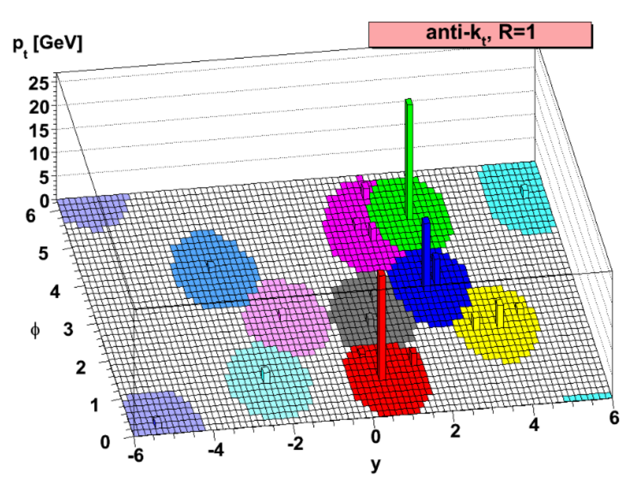
$p(\boldsymbol{\mu}\,|\,\boldsymbol{\Lambda})p(\boldsymbol{\Lambda}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{\mu}_k\,|\,\mathbf{m}_0, (\beta_0\boldsymbol{\Lambda}_k)^{-1}, \boldsymbol{\Lambda}_n^{*})\mathcal{W}(\boldsymbol{\Lambda}_k\,|\,\mathbf{W}_0, \nu_0)$

$p(\mathbf{Z}\,|\,\boldsymbol{\pi}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \pi_k^{z_{nk}}$

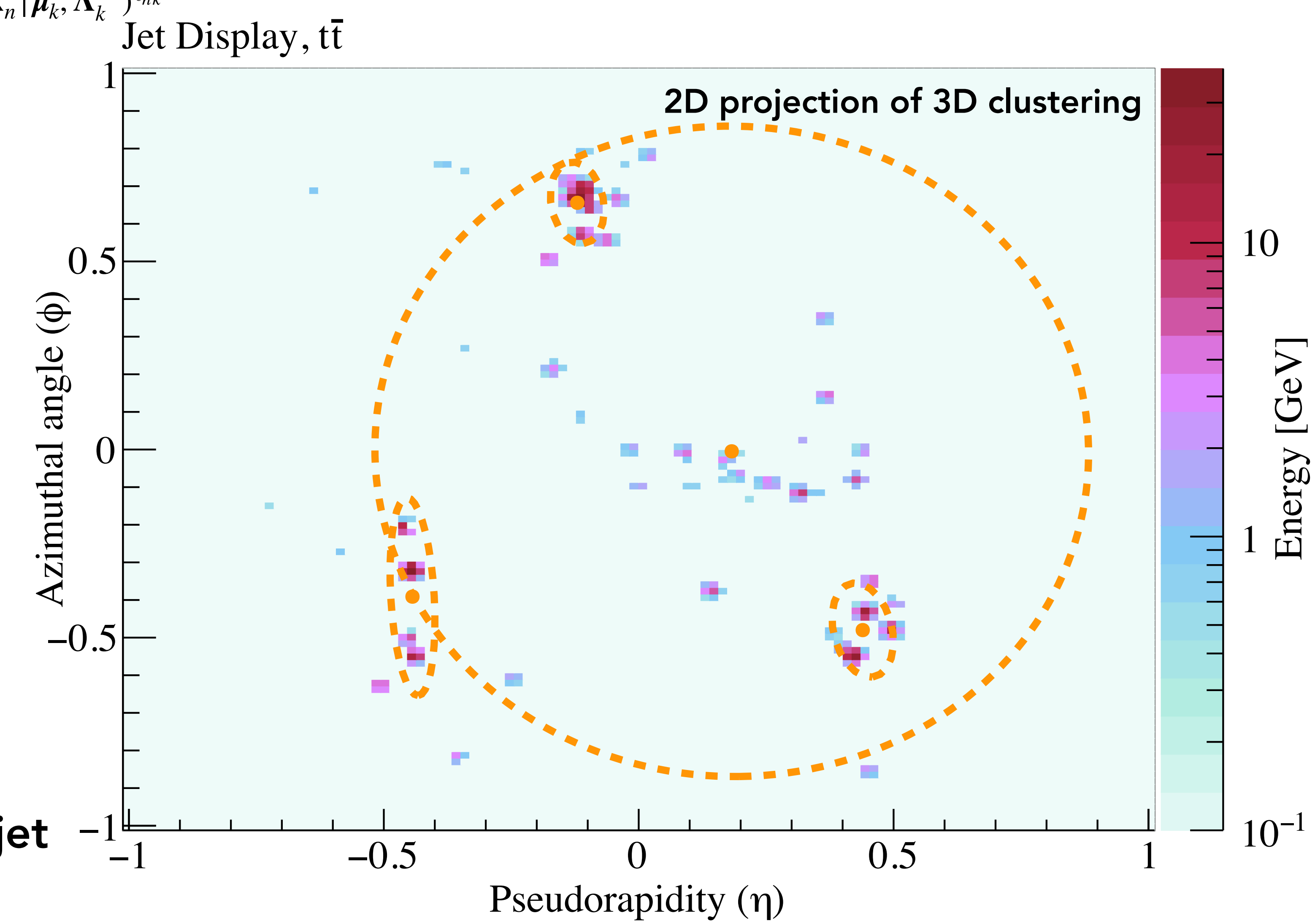


$p(\mathbf{X}\,|\,\mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \mathcal{N}(\mathbf{x}_n\,|\,\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k^{-1})^{z_{nk}}$

Variational mixture model
learns substructure

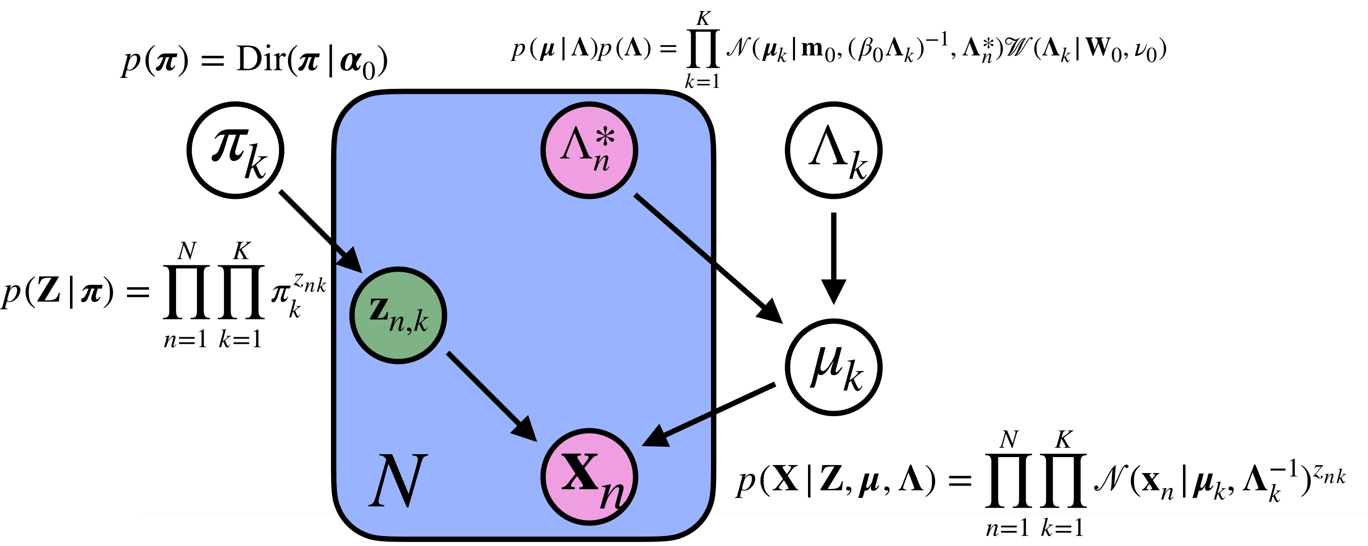


Subclusters within a jet

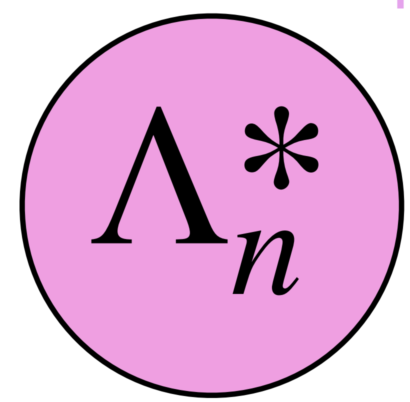


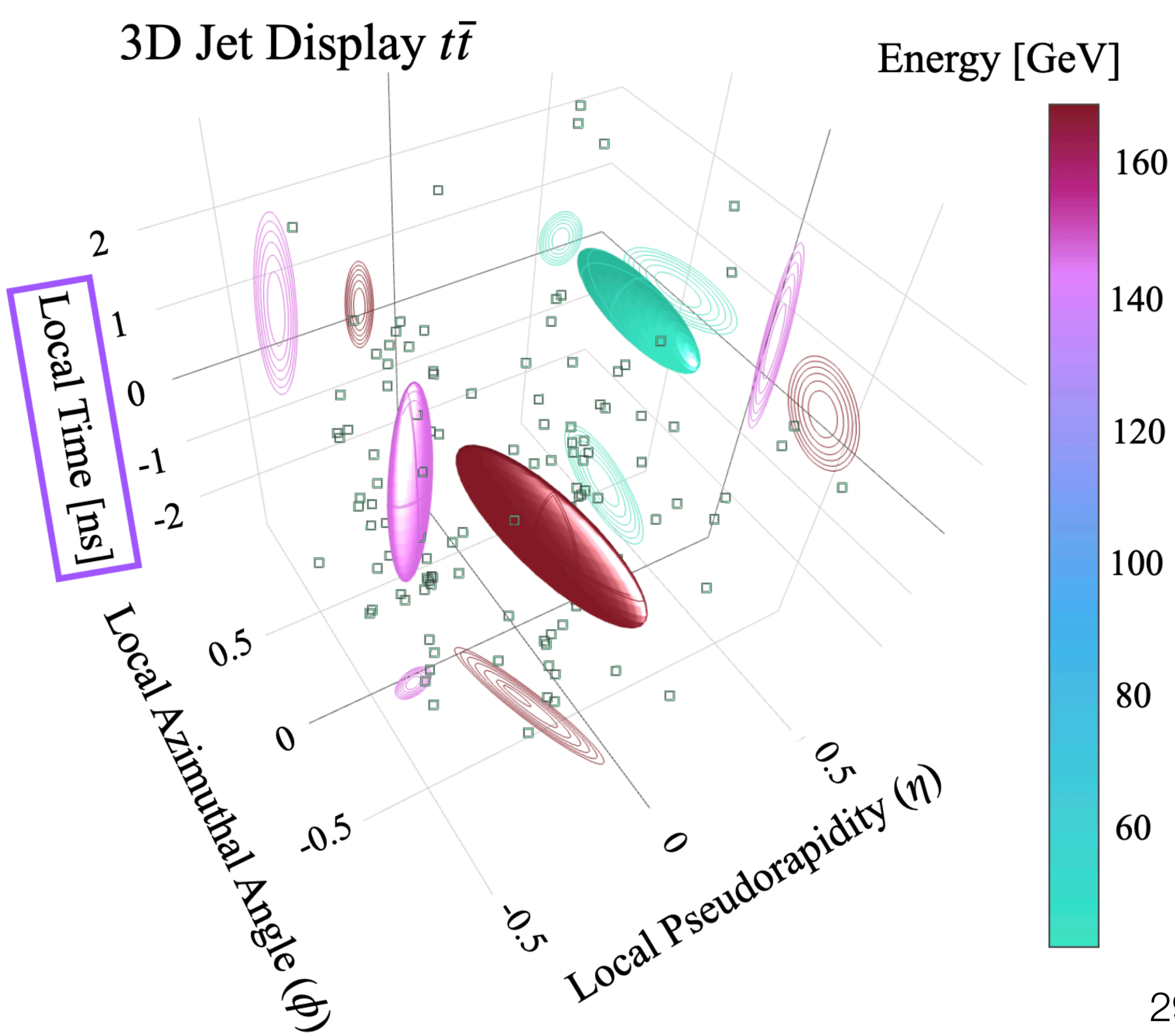
PSICHE Jets

Structure-Intrinsic: domain-specificity incorporated in algorithm design

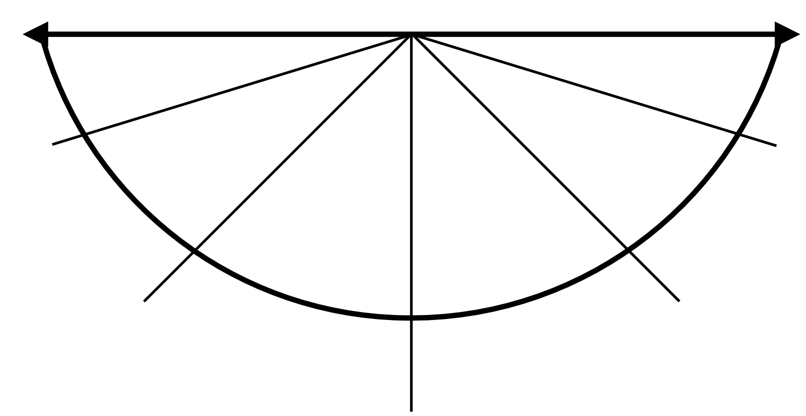


Physically-motivated uncertainties to incorporate experimental resolutions





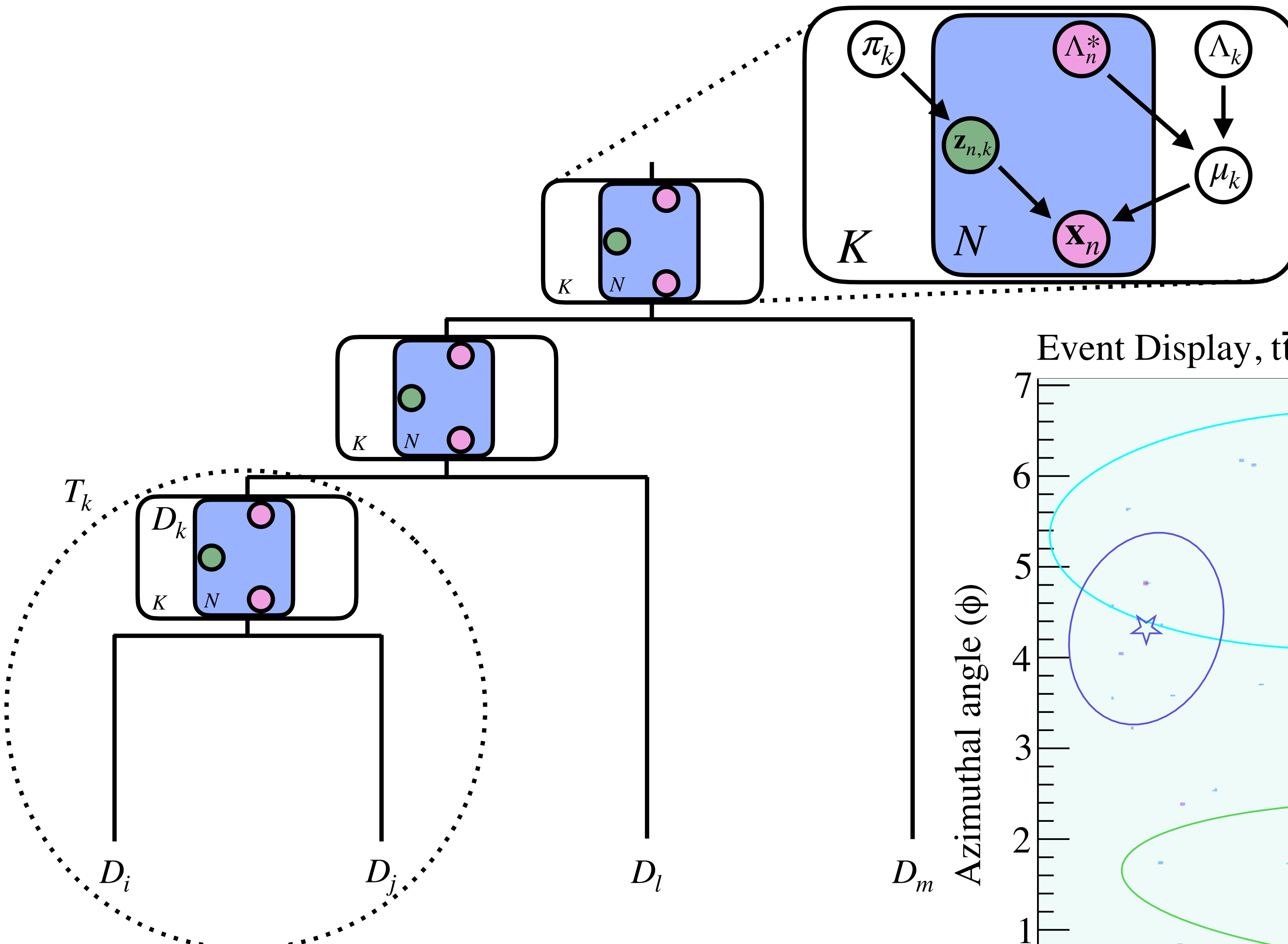
Spatial telescoping to incentivize jet locality



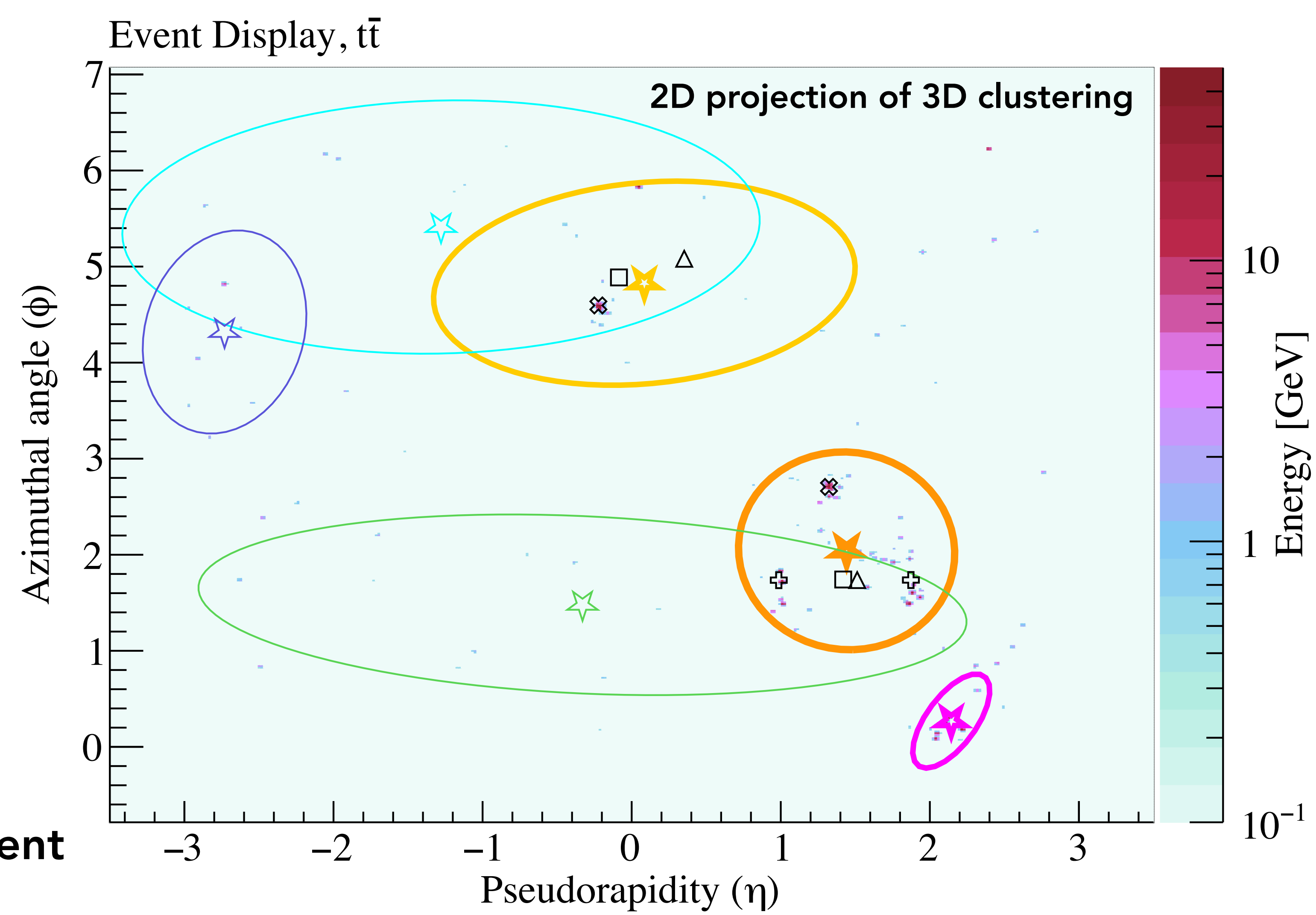
Input data include spatial + temporal position

PSICHE Jets

Hierarchical Embedding: multi-tier inference facilitates ***multi-scale learning***



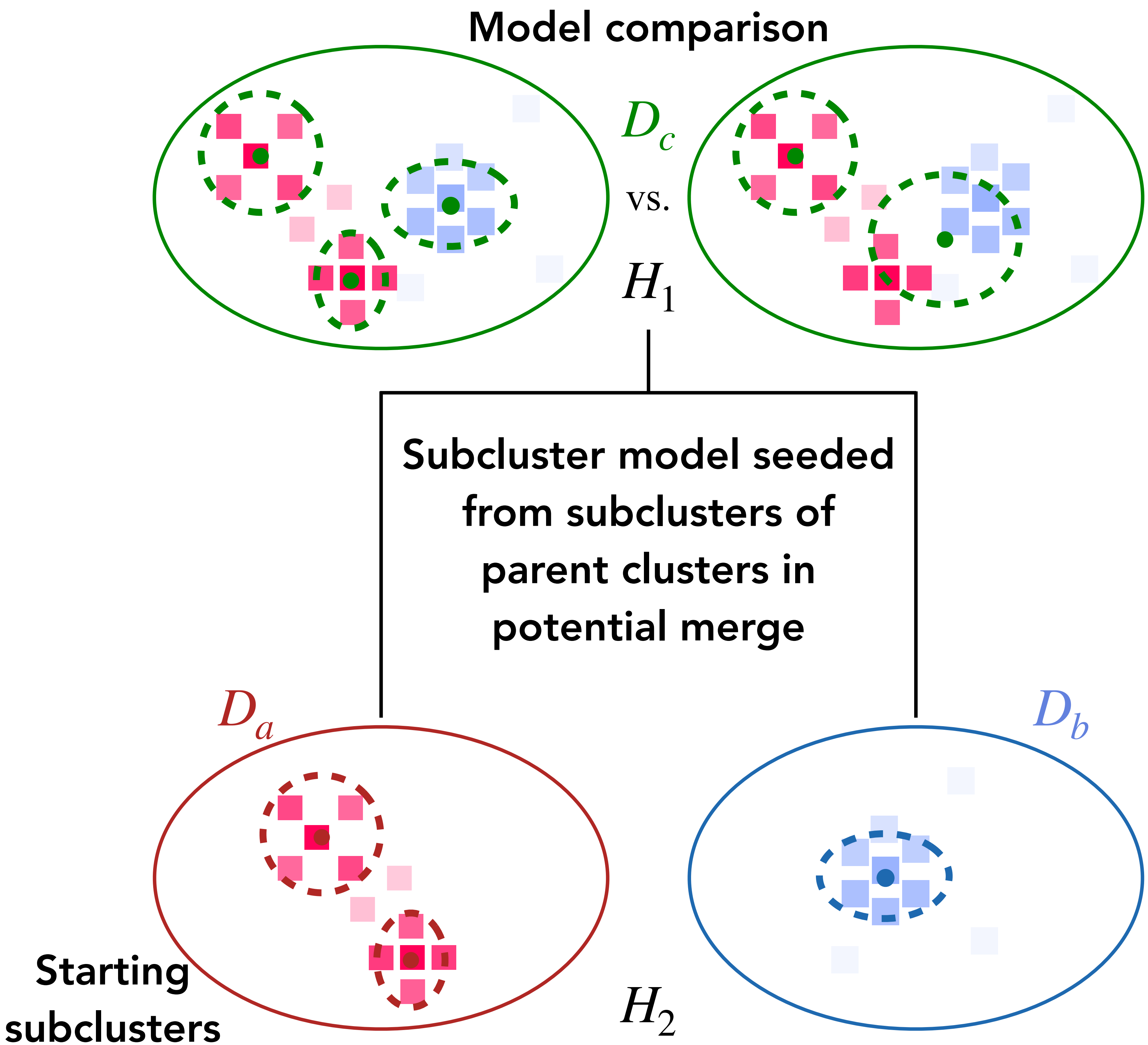
Embedded within Bayesian Hierarchical model [Heller, Ghahramani 2005] **to learn global jet scale**



Jets within a collision event

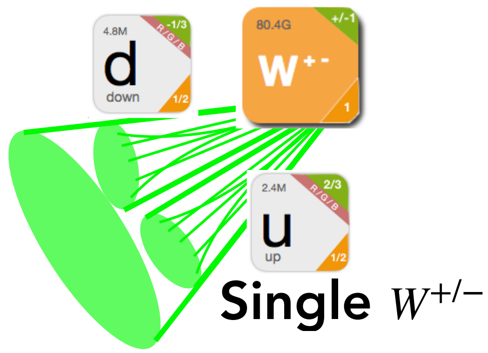
Subcluster merging scheme

Subcluster model regulated via probabilistic comparison

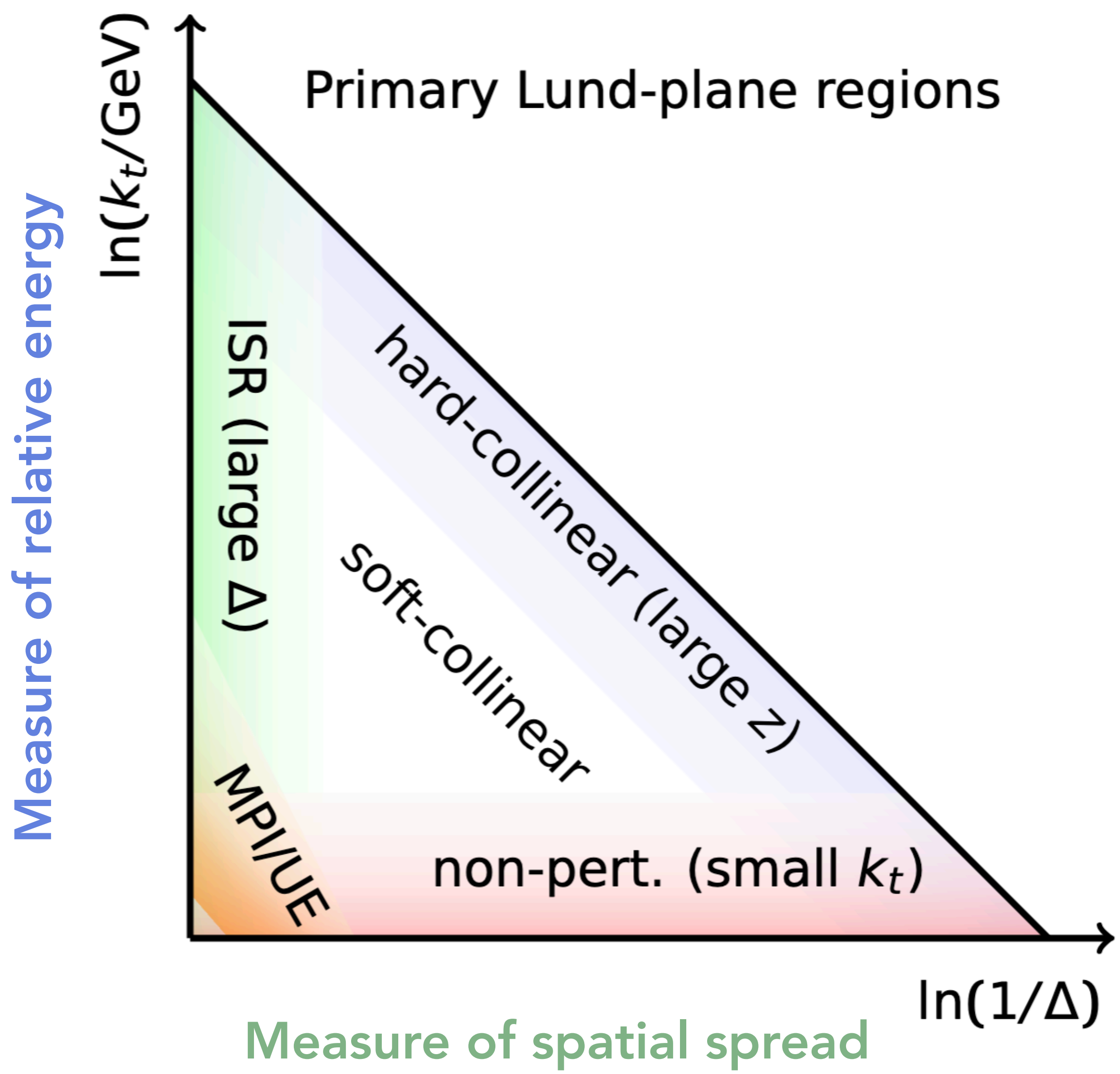


1. Create subcluster pairs, ranked by PDF inner product
2. For each pair, project data into “local” model space
3. From projected data, fit two, local models: one from original starting subclusters, the other from a merged subcluster (average location, covariance of pair)
4. Compare likelihoods of merged and original local models, choose higher likelihood pair to seed merged model
5. Repeat for all subcluster pairs
6. Fit merged model
7. Compare merged model to original model, choose higher likelihood model

PSICHE Jets: Novel Features



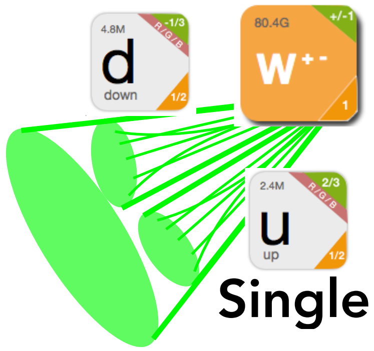
Multi-scale learning



Dreyer, et. al. (2018)

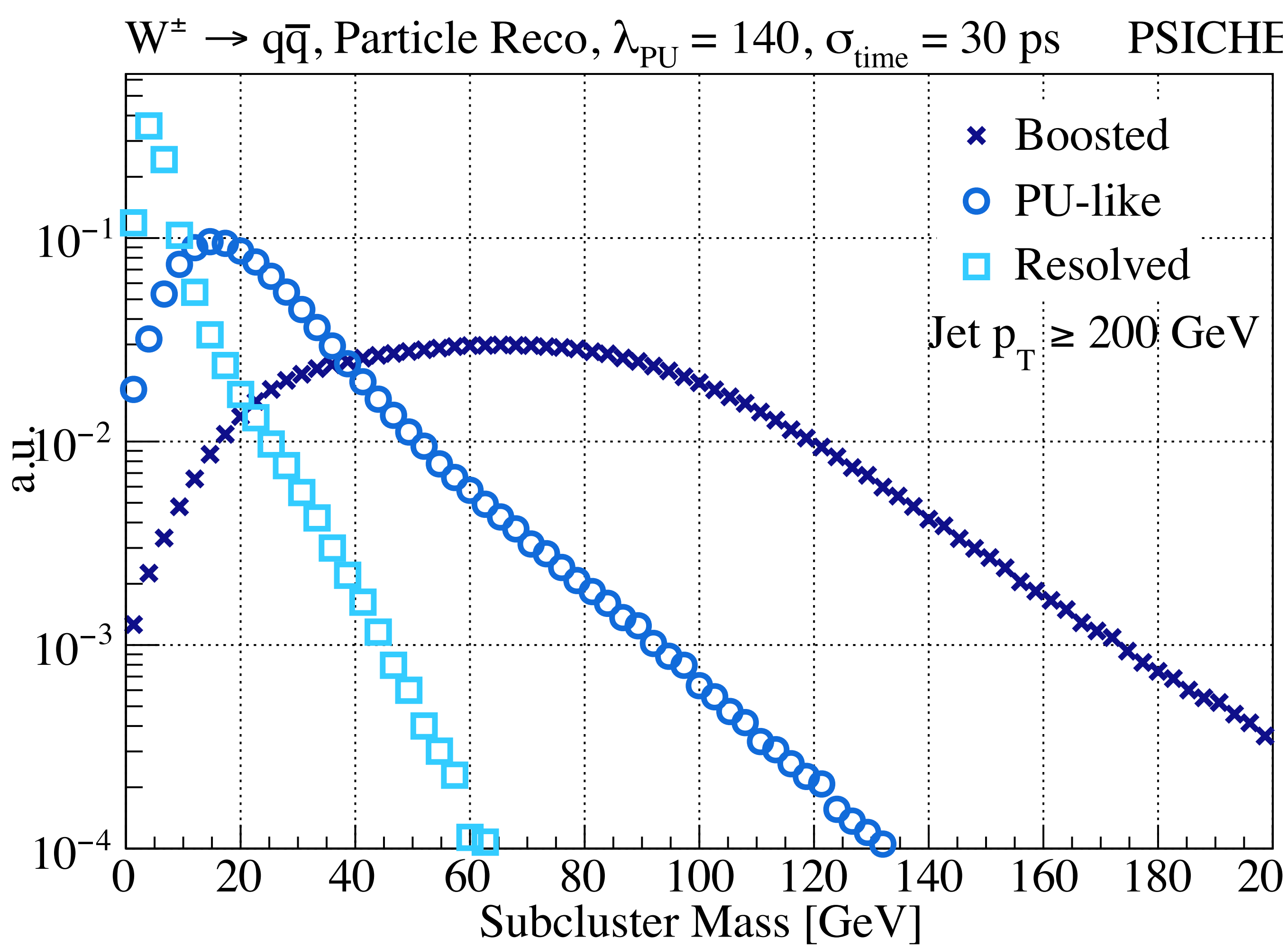
- Lund jet plane: tool to analyze jet substructure
- Defined by spatial spread axis and relative energy axis
- **Different types of subjets** populate **different areas** of the plane

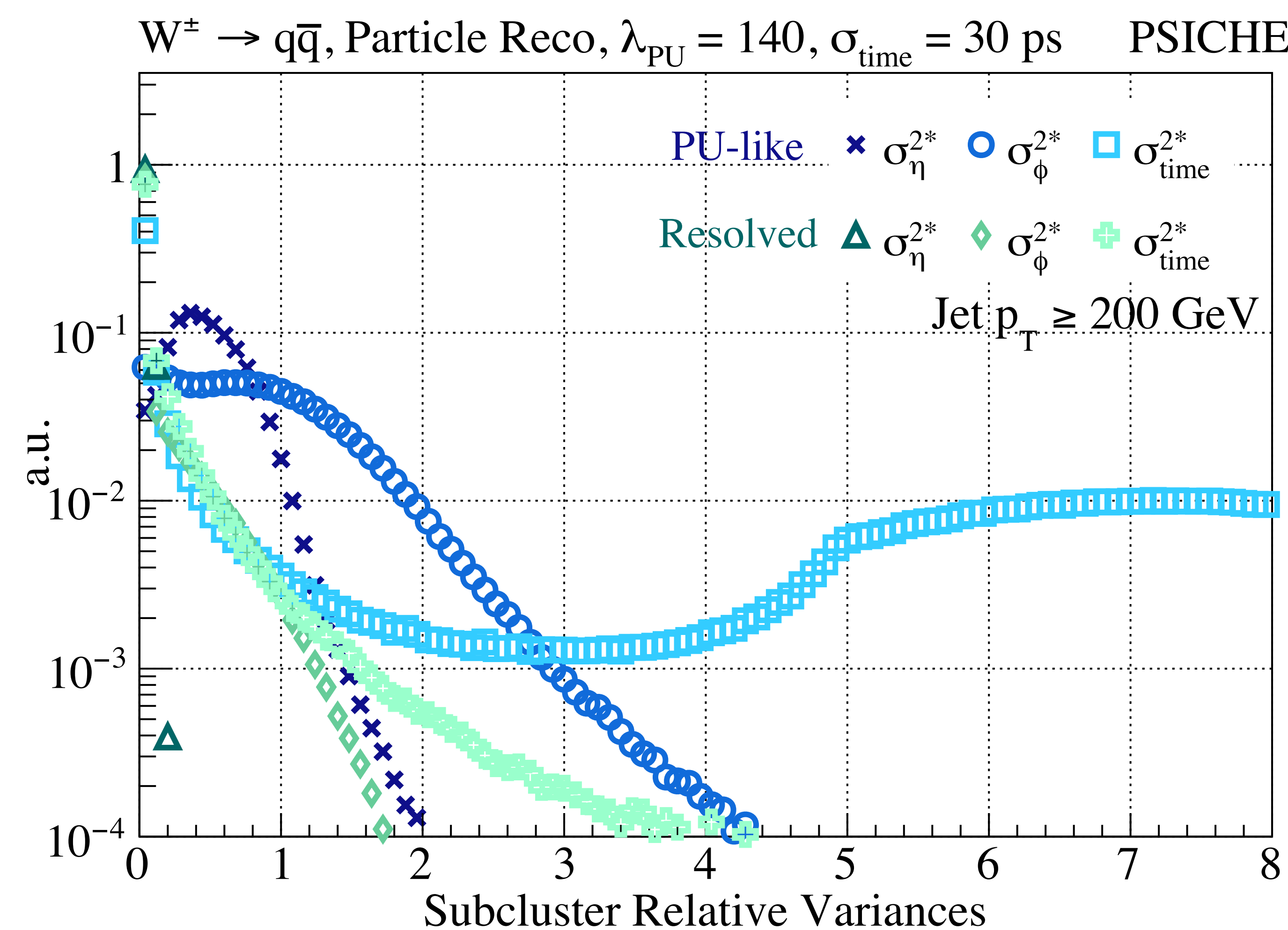
PSICHE Jets: Novel Features



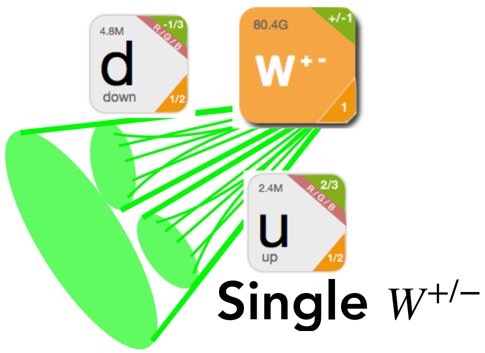
Single $W^{+/-}$

Multi-scale learning

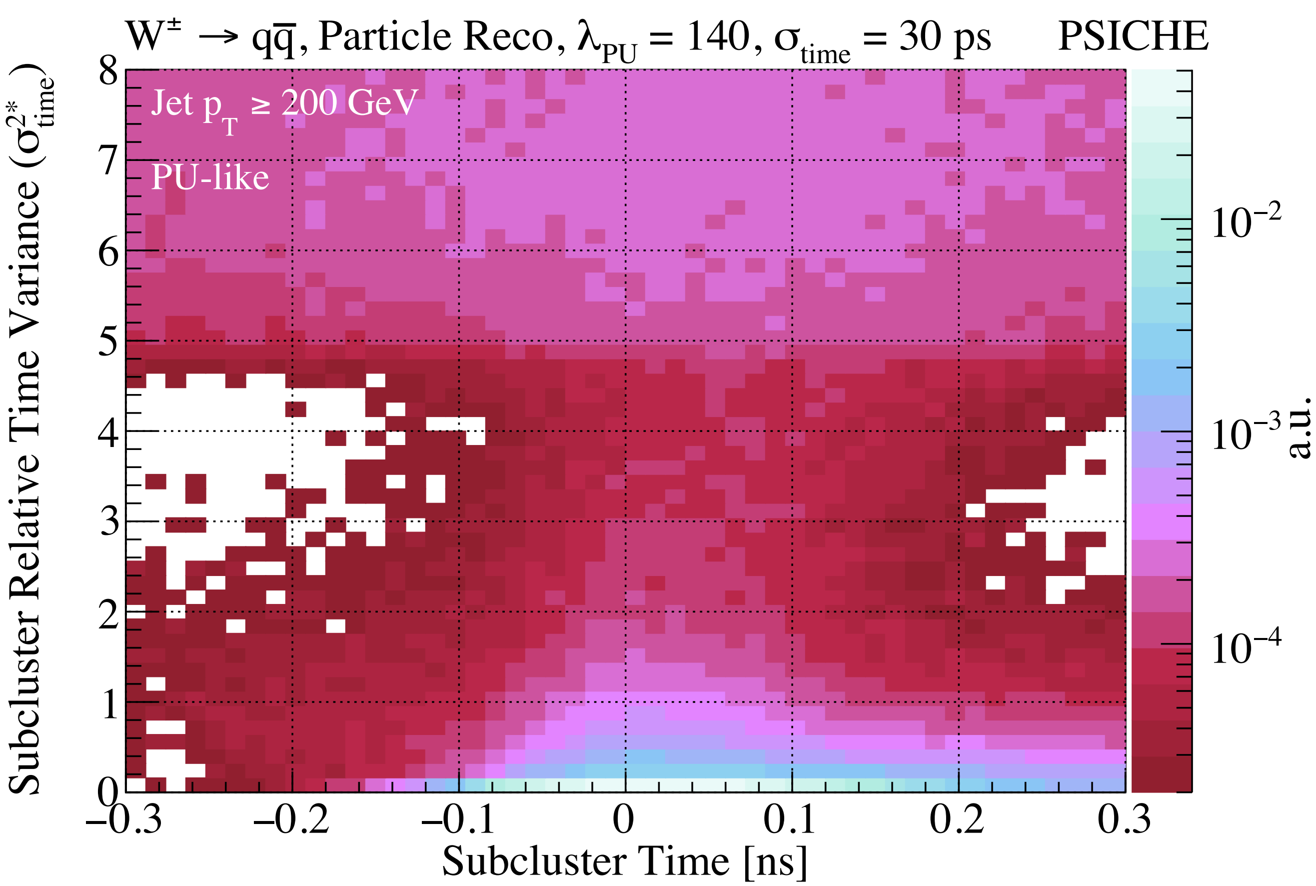


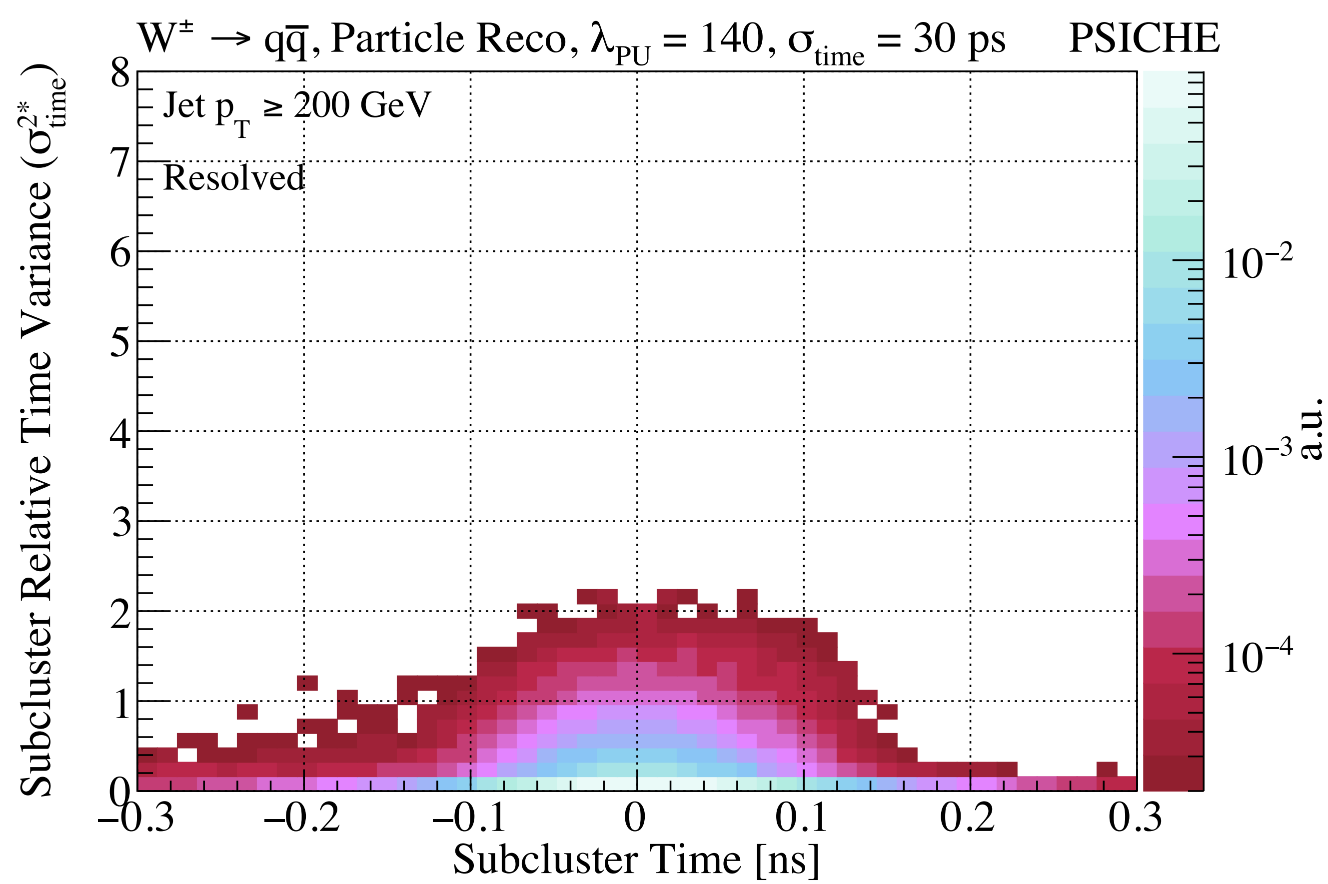


PSICHE Jets: Novel Features



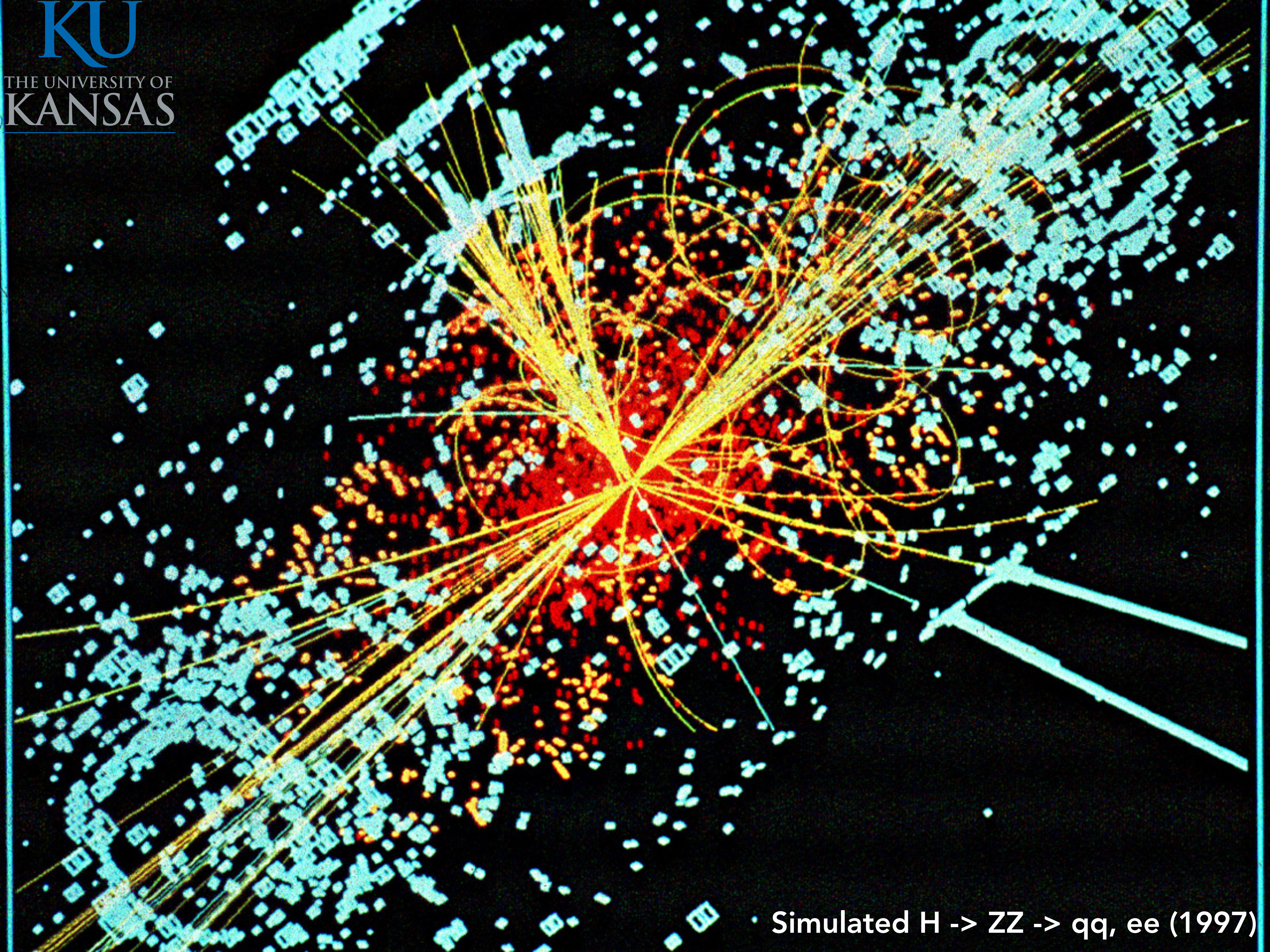
Multi-scale learning &
timing aware





Different subcluster types exhibit **different, characteristic location** in **time**

Leverage **timing** and **subcluster shape** information for **pileup cleaning**



Simulated H -> ZZ -> qq, ee (1997)