A tale of two jets

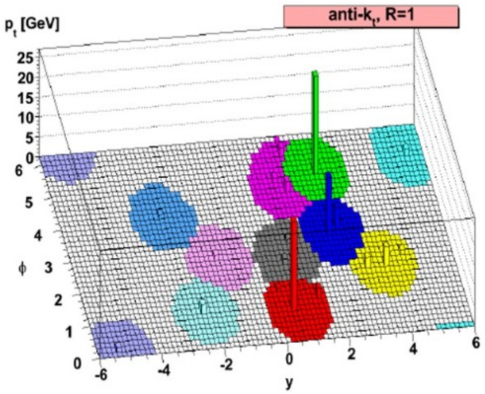
Christine Nattrass, University of Tennessee, Knoxville

Acknowledgements

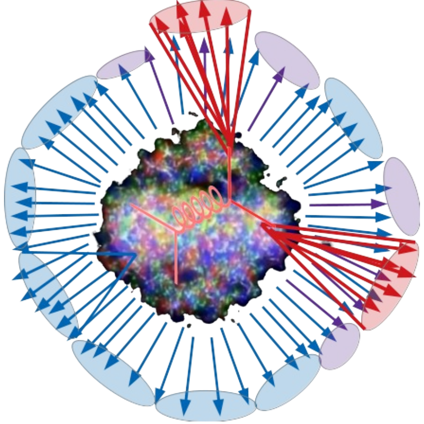

Antonio Da Silva


Patrick Steffanic

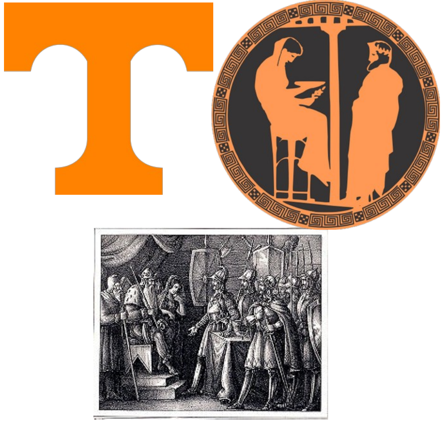

Charles Hughes



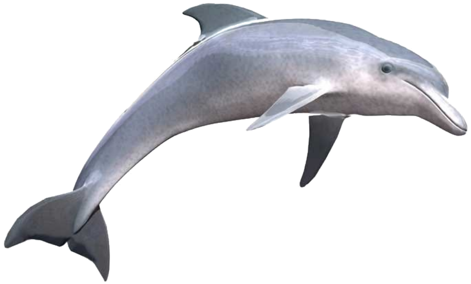
1. What is a jet?



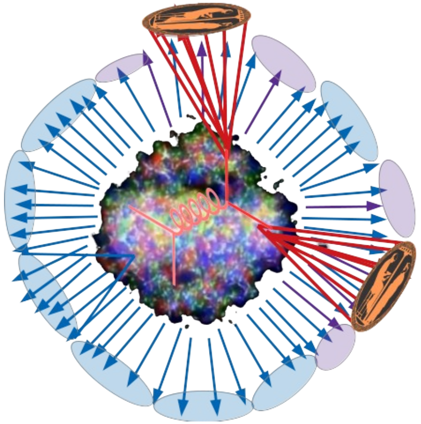
2. Background paradigm



3. Models



4. Suppressing combinatorial jets



5. Correcting for background in MC

6. How to compare to models

1. What is a jet?

What is a jet?

What is a jet?

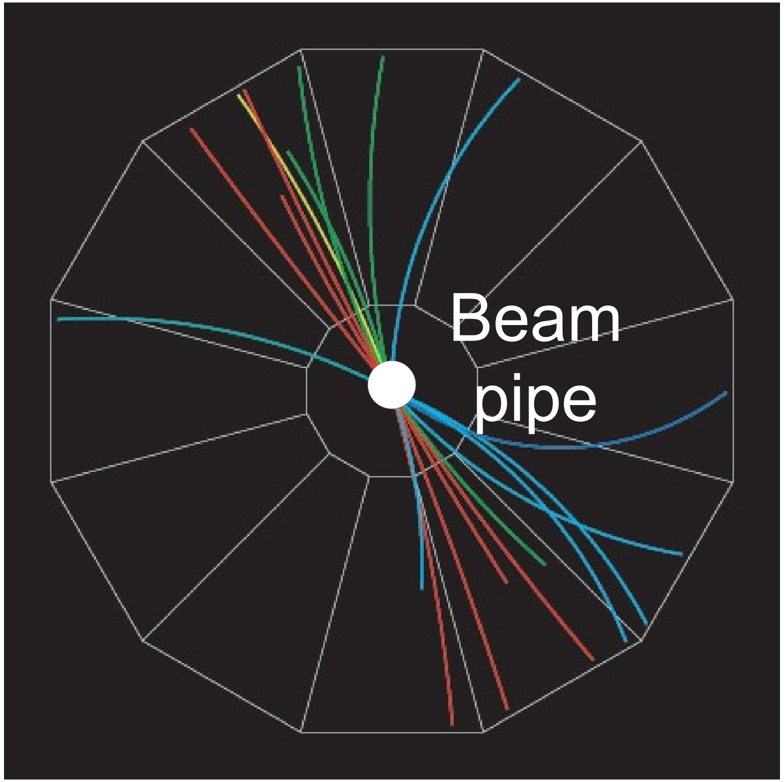
A measurement of a jet is a measurement of a parton.

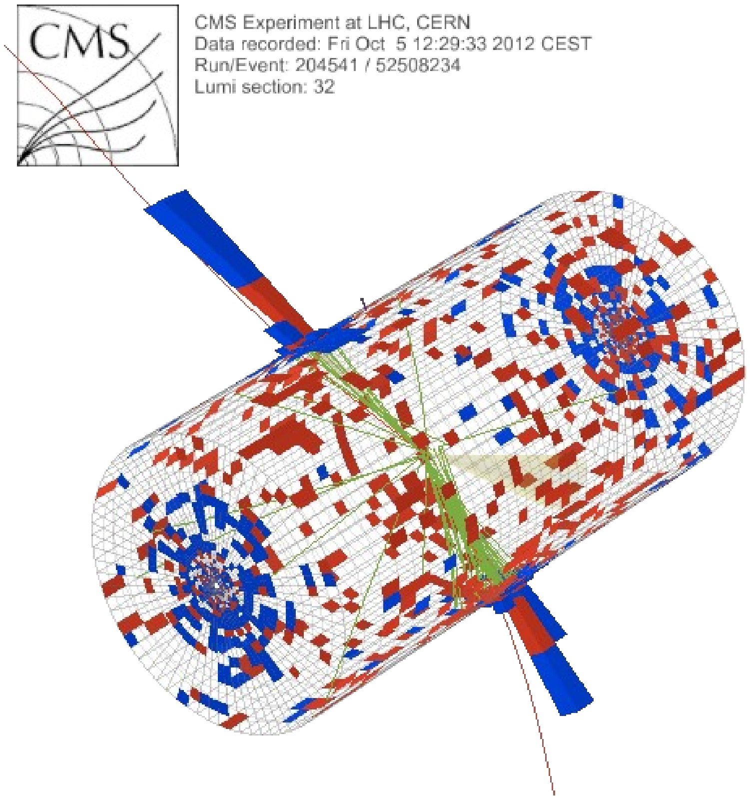
What is a jet?

~~A measurement of a jet is a measurement of a parton.~~

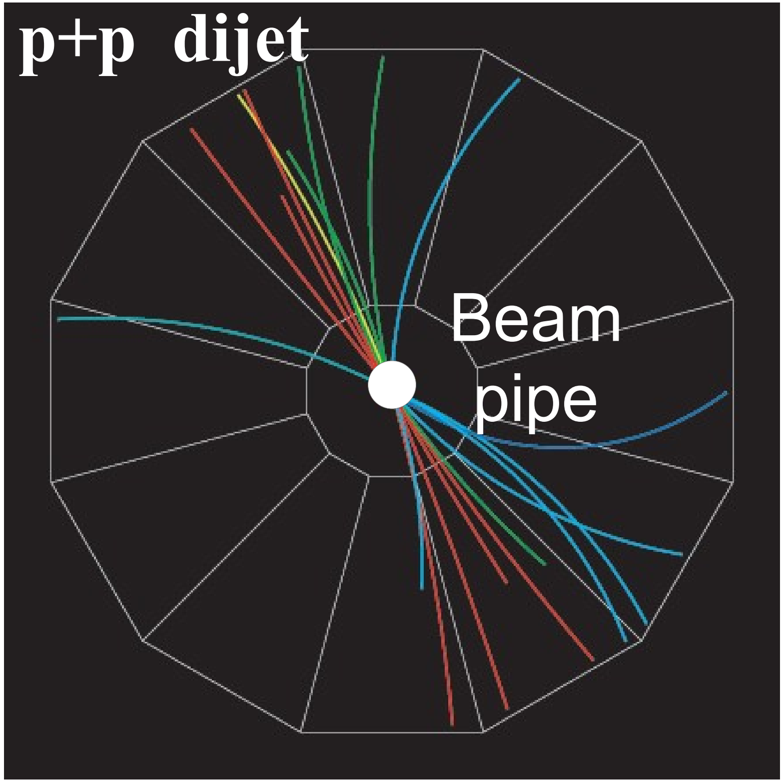
What is a jet?

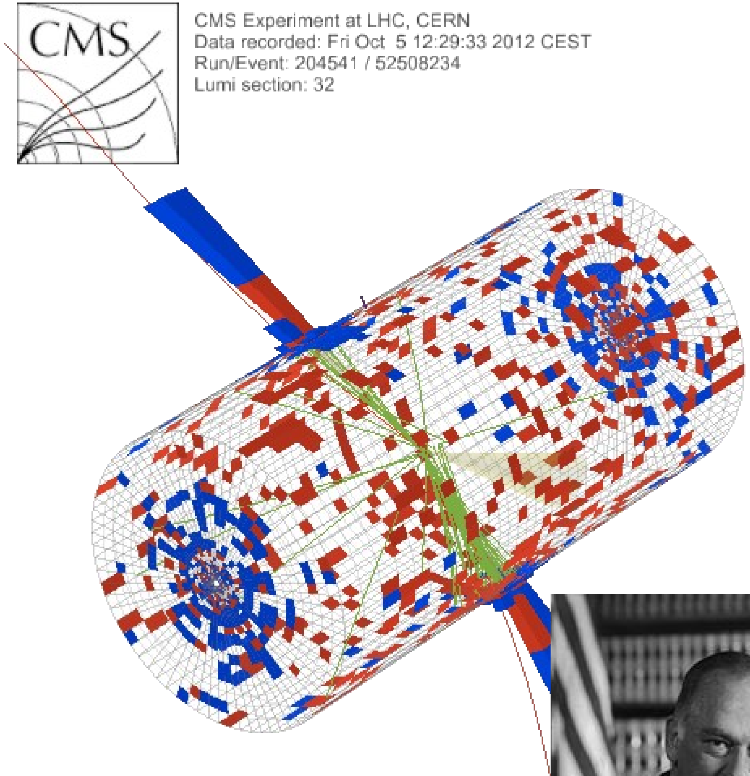
p+p dijet





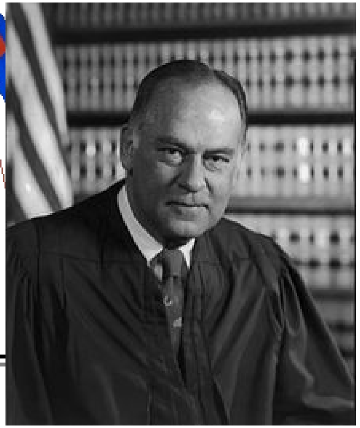
What is a jet?





"I know it when I see it"

US Supreme Court Justice Potter Stewart, Jacobellis v. Ohio



Jet finding *in pp collisions*

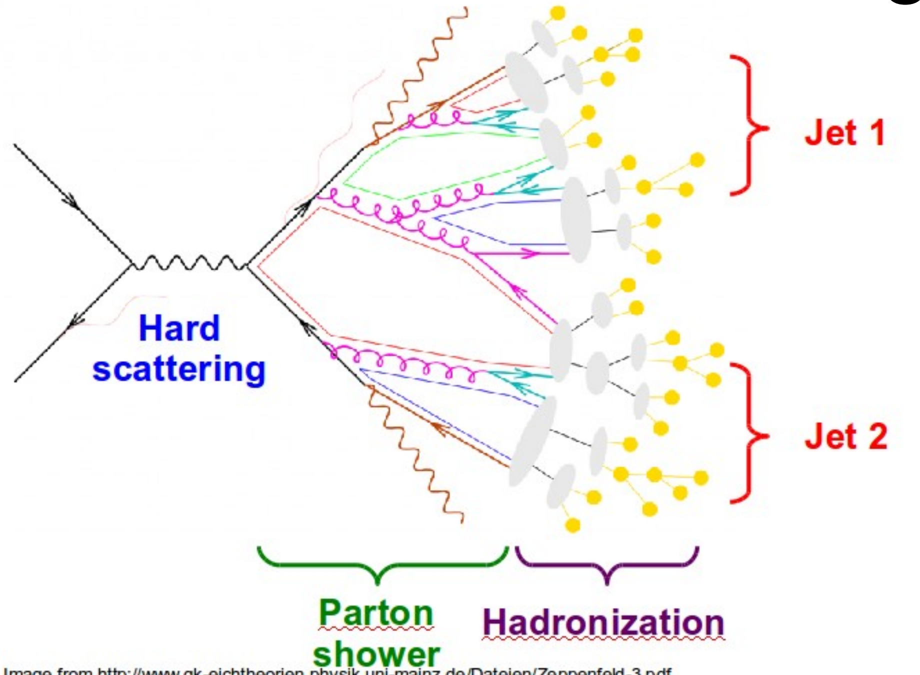


Image from http://www.gk-eichtheorien.physik.uni-mainz.de/Dateien/Zeppenfeld-3.pdf

- Jet finder: groups final state particles into jet candidates
  - Anti-k_T algorithm JHEP 0804 (2008) 063 [arXiv:0802.1189]
- Depends on hadronization
- Ideally
  - Infrared safe
  - Colinear safe

Snowmass Accord: Theoretical calculations and experimental measurements should use the same jet finding algorithm. Otherwise they will not be comparable.

anti-k_T jet finding algorithm

Particles, clusters

↓

k_T algorithm

$$k_T = p_T, \Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$

- For all *i,j* calculate:

$$d_{ij} = min\boxed{\left(p_{T,i}^{-2}, p_{T,j}^{-2}\right)}\frac{\Delta R_{ij}^2}{R^2}$$

$$d_{iB} = p_{T,i}^{-2}$$

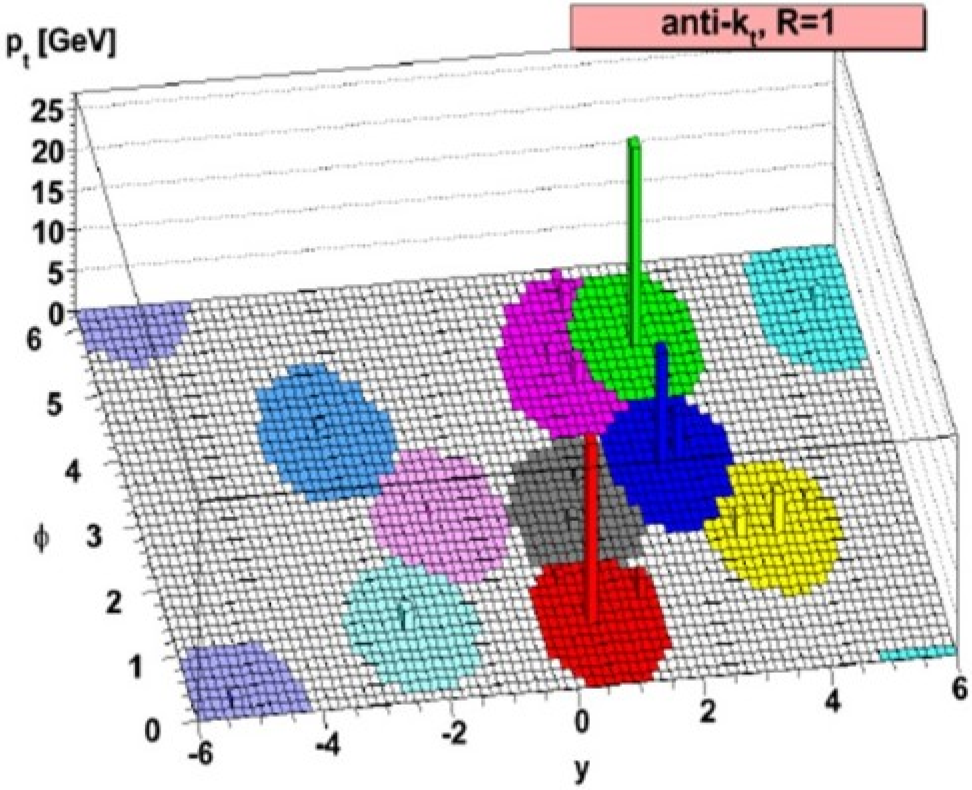
- Combine smallest d_{ij}.

  If d_{iB} smallest, $d_{iB} \rightarrow$ jet

Repeat until no particles left

↓

Jet candidates



Jet finding algorithms

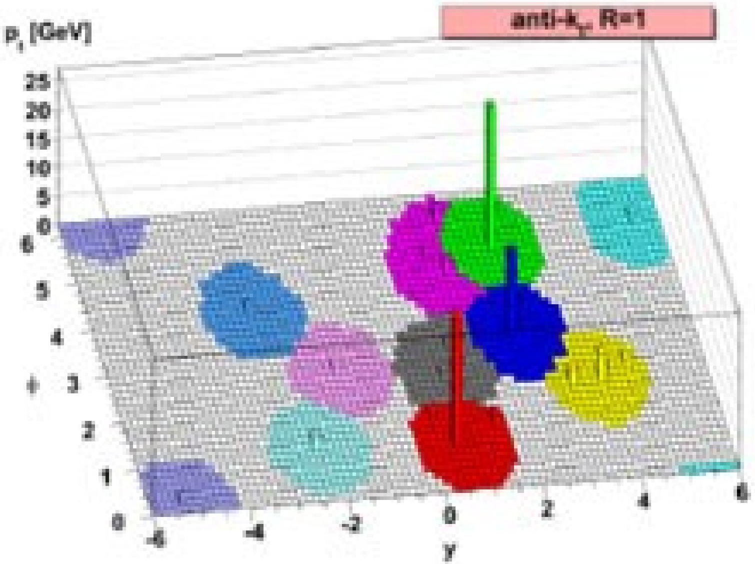
Tracks

Clusters

Particles

Jet finding algorithm

Jet candidates



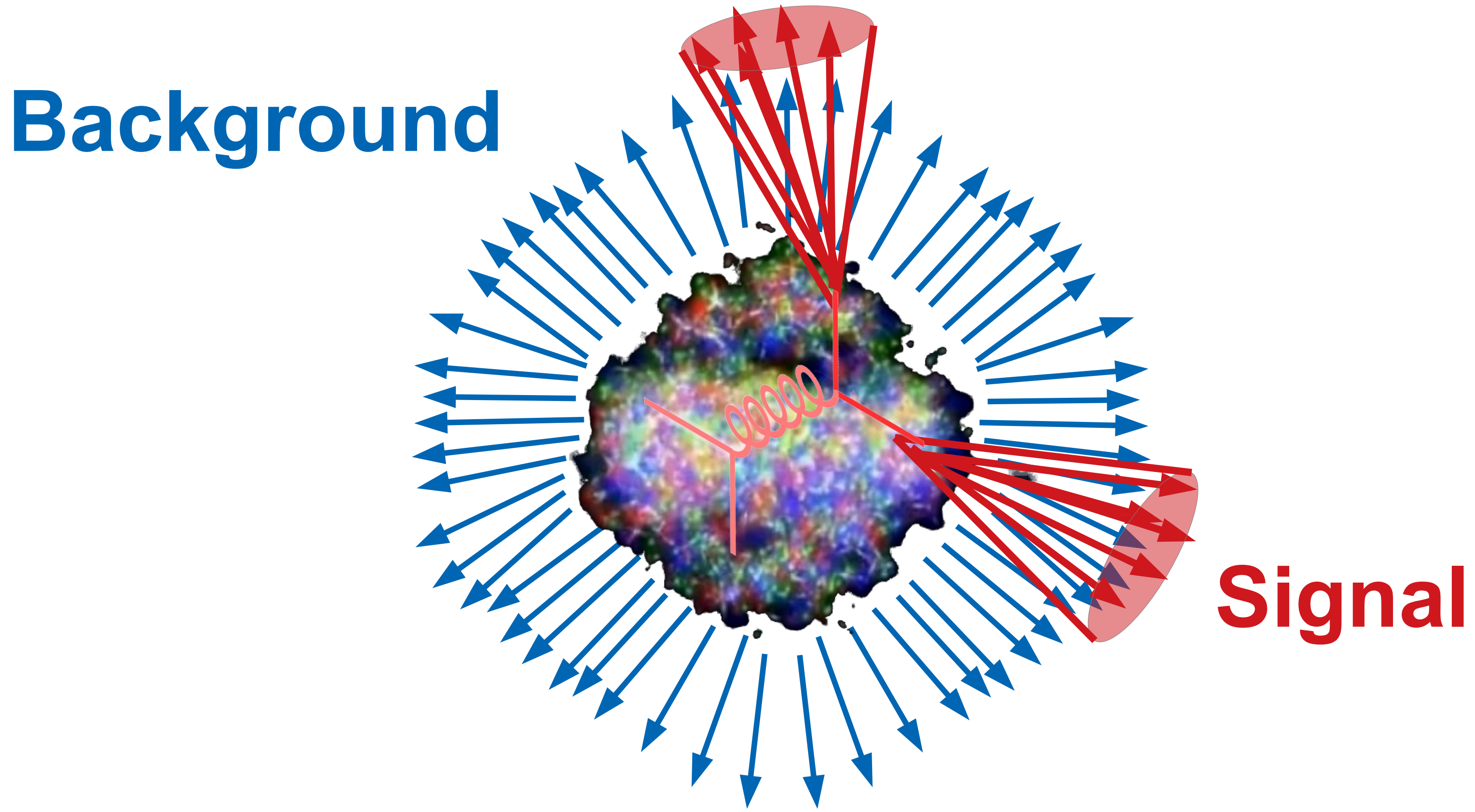
- Any list of objects works as input
- Use the same algorithm on theory & experiment
- Output only as good as input

A jet is what a jet finder finds.

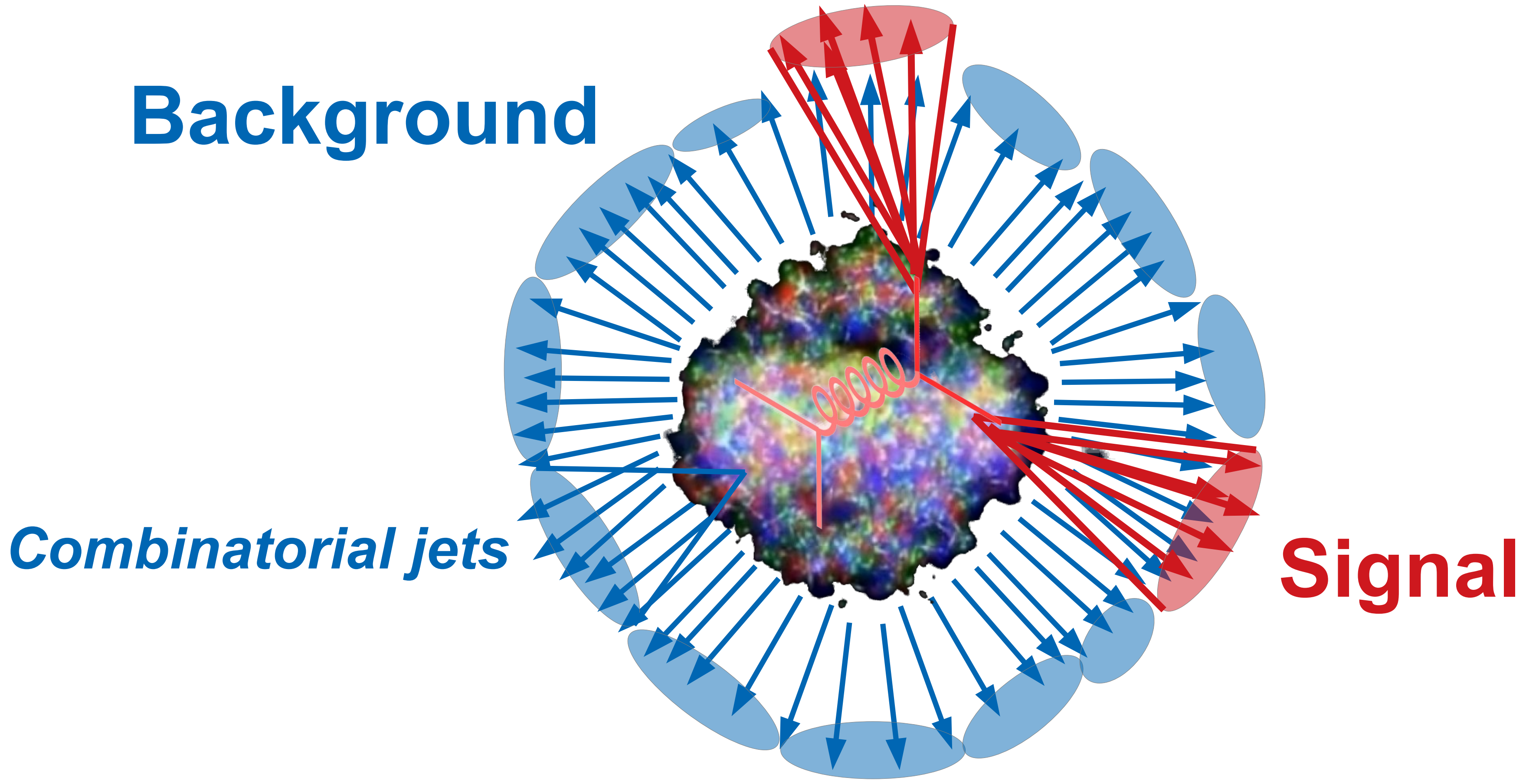
Use the same algorithm on data and the model.
Then the two will be comparable.

2. Standard paradigm of background

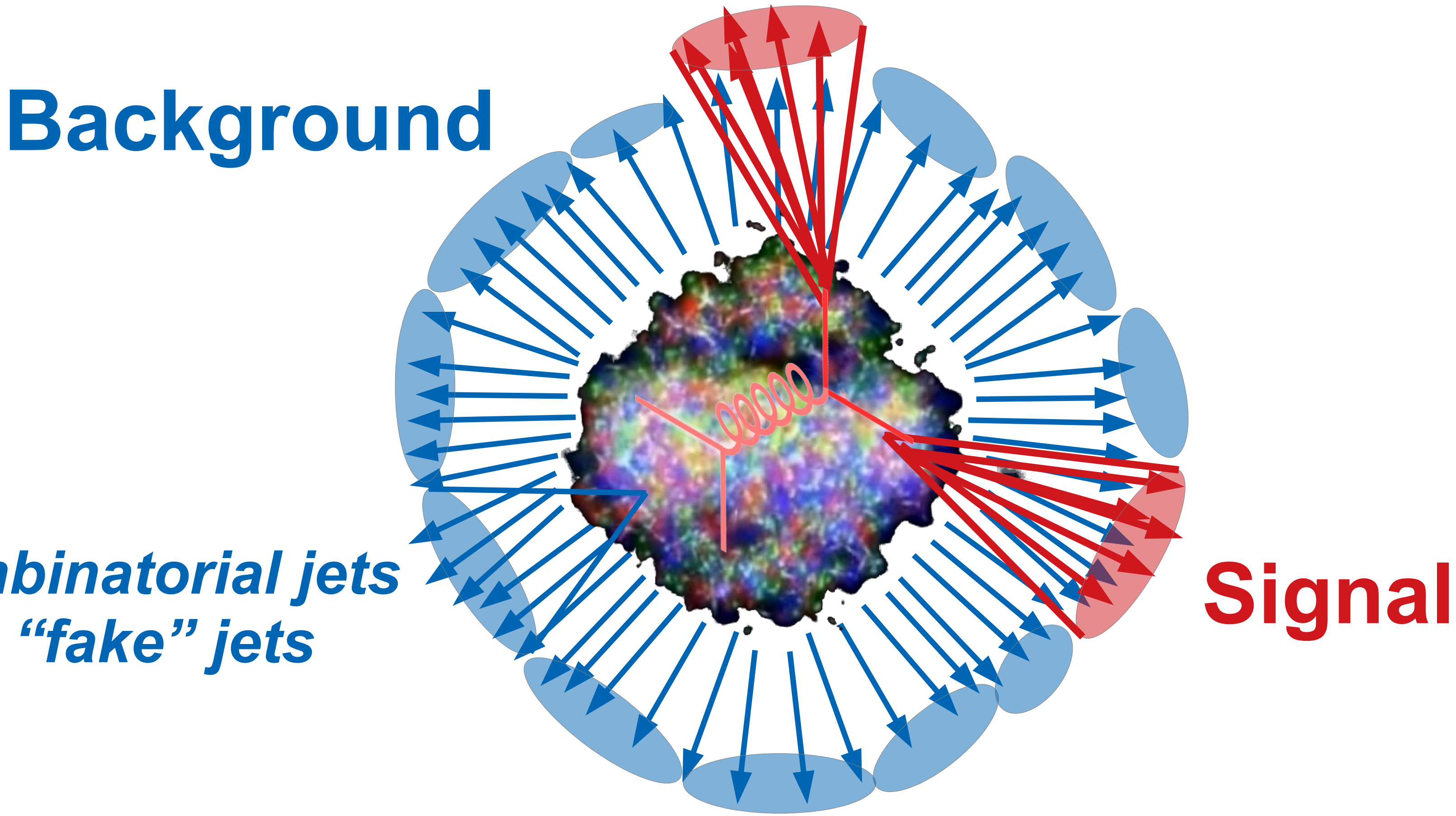
Signal vs Background:
The standard paradigm



Signal vs Background:
The standard paradigm

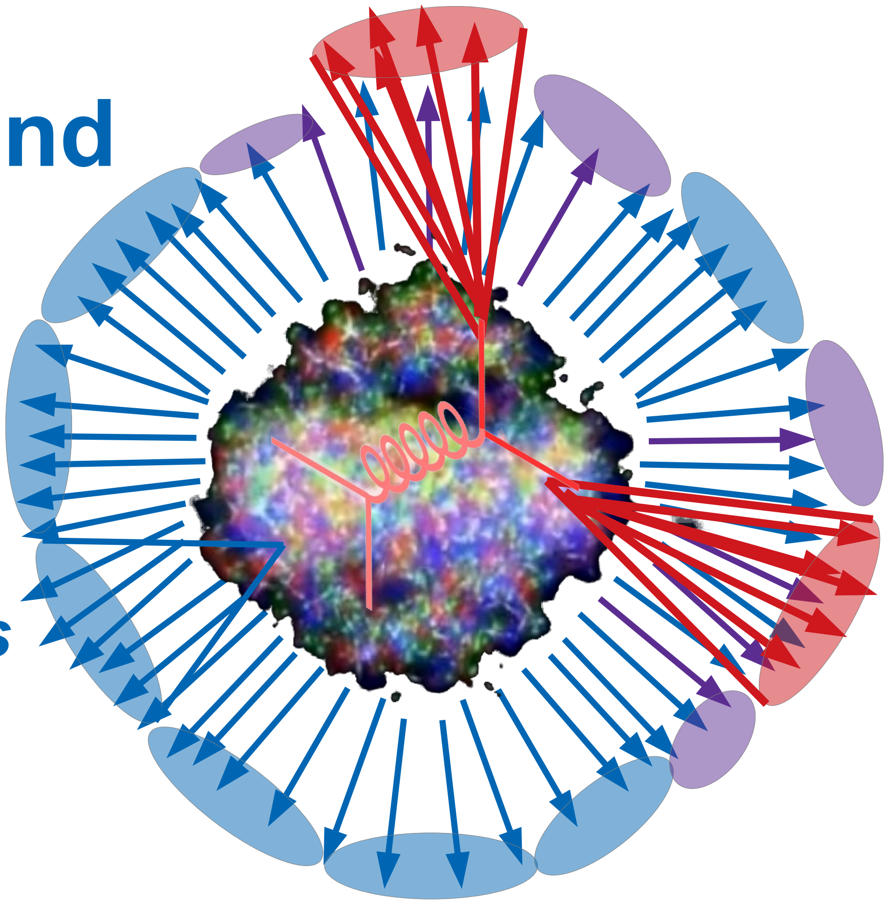


Signal vs Background:
The standard paradigm



Signal vs Background:

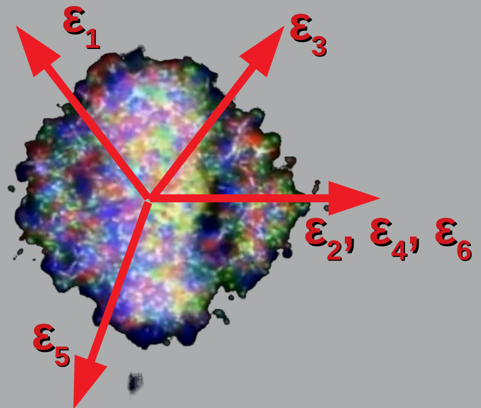
The standard paradigm



3. Models

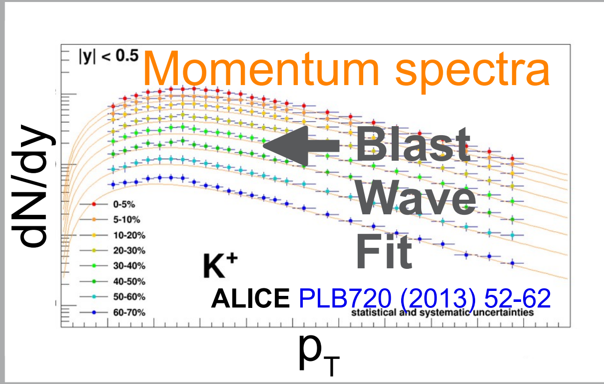
TennGen background generator

Event properties

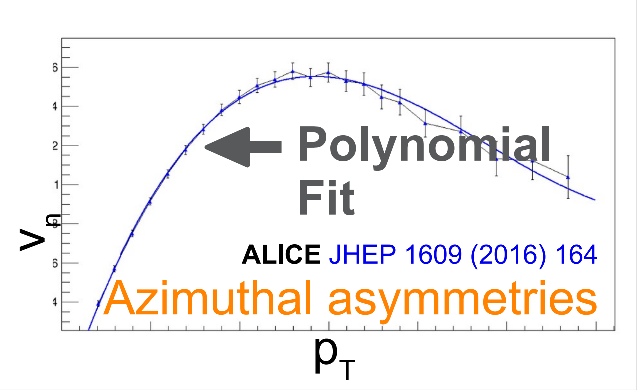


- Even event planes fixed at $\Psi=0$
- Odd planes at random φ
- Multiplies from ALICE PRC88 (2013) 044910

Track properties



→ Random p_T



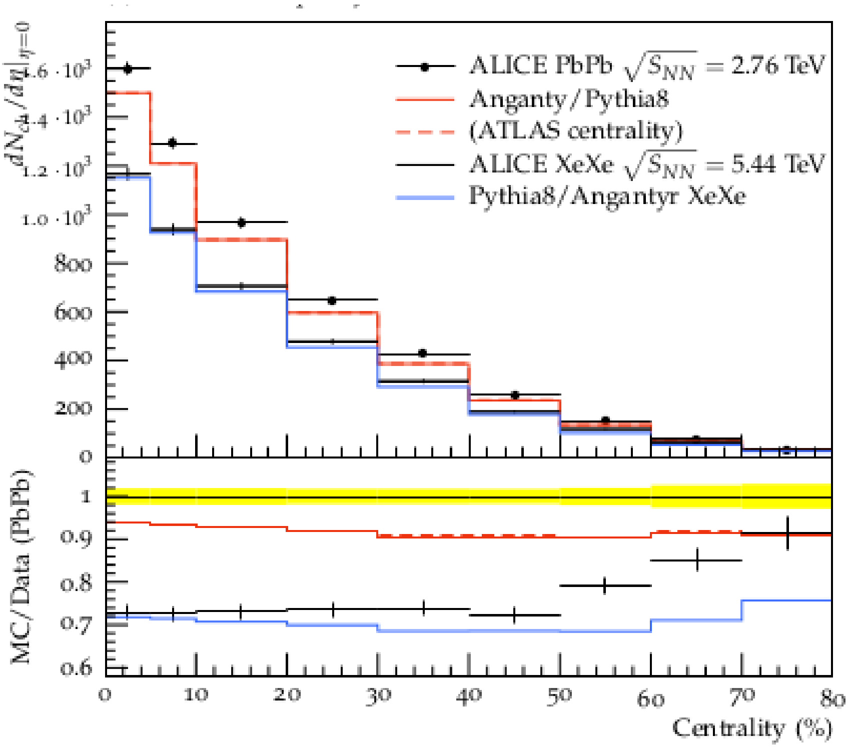
→ v_n → Random φ

No jets! No resonances
Emulates hydro correlations

PYTHIA Angantyr

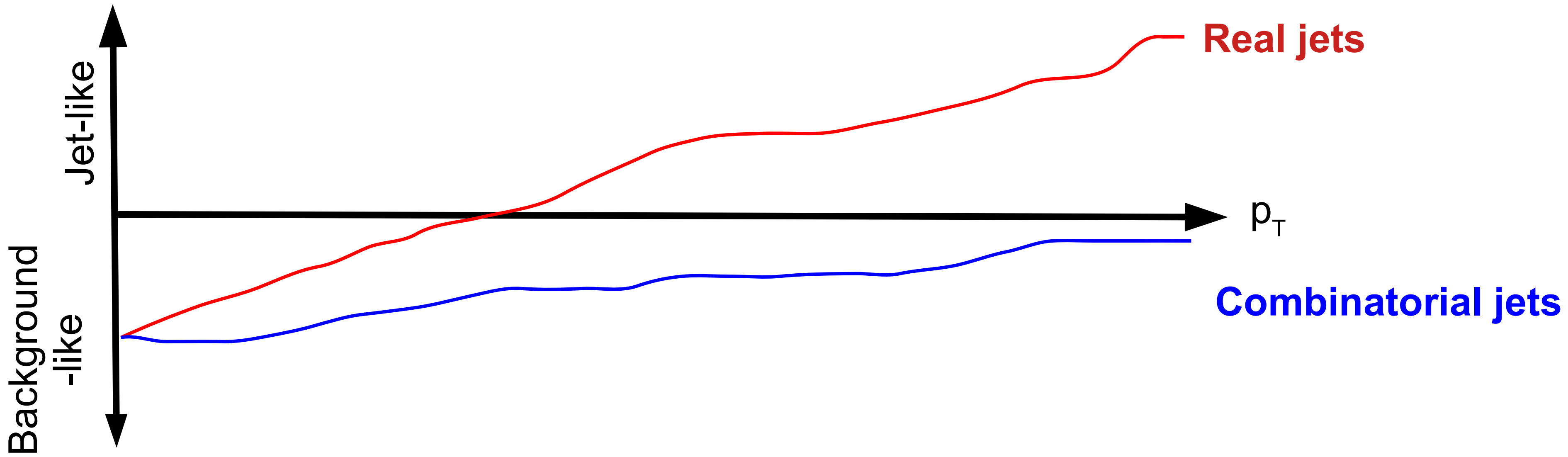
JHEP (2018) 2018: 134

- Based on PYTHIA 8
  Sjöstrand, Mrenna & Skands,
  JHEP05 (2006) 026
  Comput. Phys. Comm. 178 (2008) 852.
- Based on Fritiof & wounded nucleons
- N-N collisions w/fluctuating radii
  → fluctuating σ



Lots of jets! And resonances!
No hydrodynamics, no jet quenching

4. Suppressing combinatorial jets



Technique

- Anti-k_T jet finder, $|\eta_{jet}|<0.5$
- **Combinatorial jets:** Only contain TennGen particles
- **Real jets:** Add a PYTHIA pp event. Real jets contain >80% of p_{Thard}^{min}
- **Squishy jets:** Everything else
- R=[0.2,0.3,0.4,0.5,0.6], p_T=[10,20,40,80]

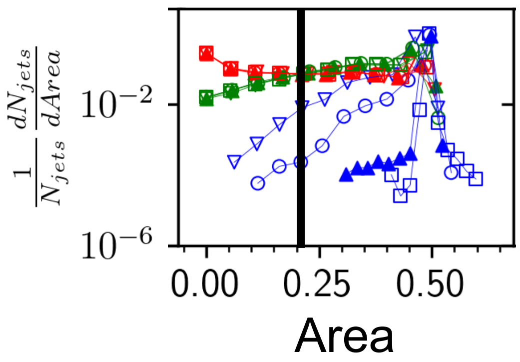
Jet properties – R=0.4

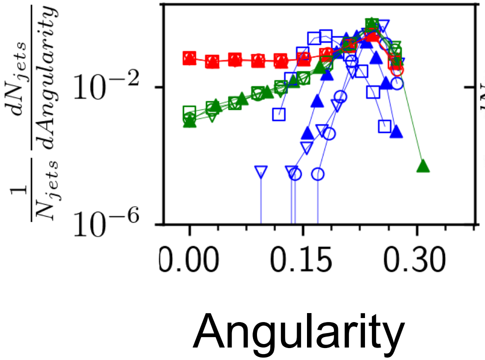
<4% Signal loss
>50% Comb. loss

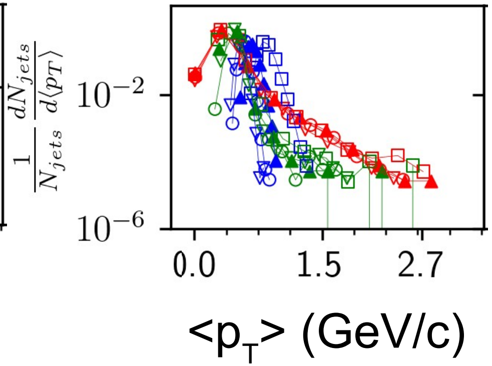
$Area > 0.6\,R$

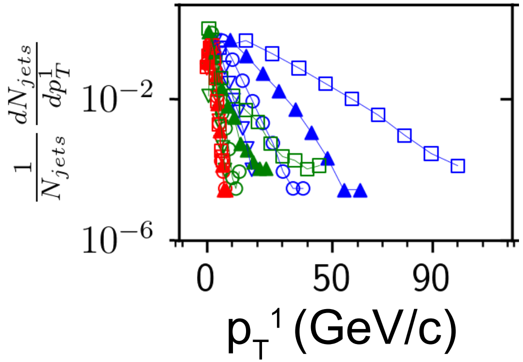
▼-Signal ▼-Combinatorial ▼-Squishy

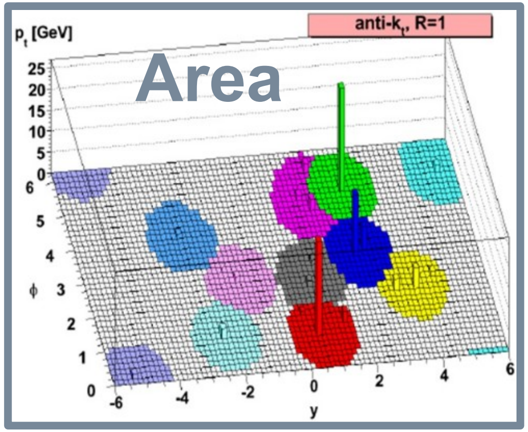
▽ 10 GeV/c
○ 20 GeV/c
▲ 40 GeV/c
□ 80 GeV/c



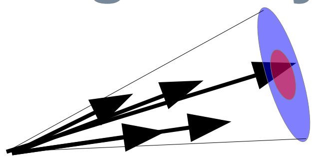






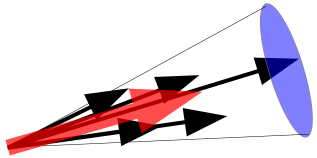


Angularity

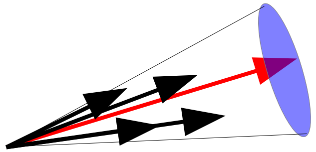


$$\alpha = \frac{1}{p_T^{jet}} \sum_k z_k (\vec{R}_k)$$

Average p_T



Leading p_T

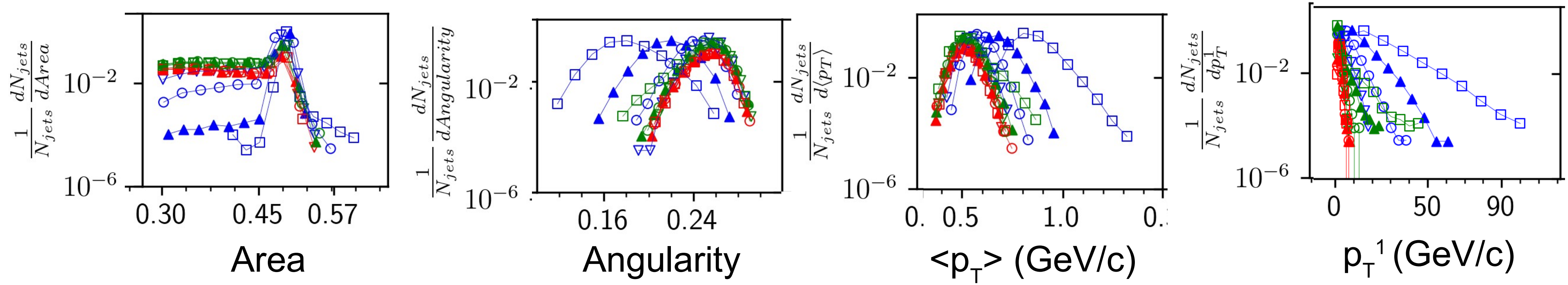


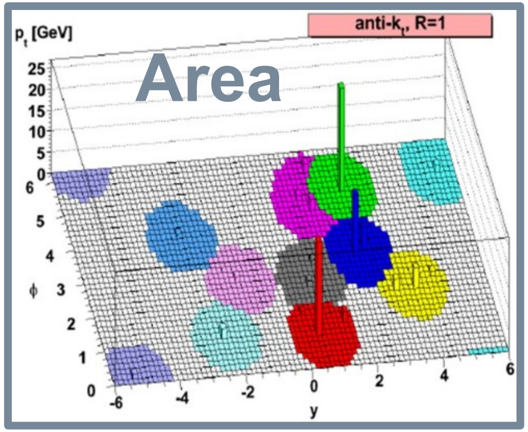
Jet properties – R=0.4

$Area > 0.6\,R$

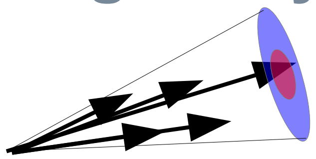
▼-Signal ▼-Combinatorial ▼-Squishy

▽ 10 GeV/c
○ 20 GeV/c
▲ 40 GeV/c
□ 80 GeV/c



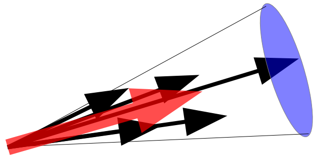


Angularity

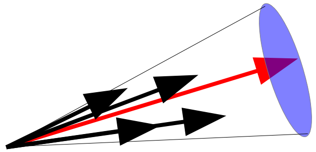


$$\alpha = \frac{1}{p_T^{jet}} \sum_k z_k \left(\vec{R}_k\right)$$

Average p_T



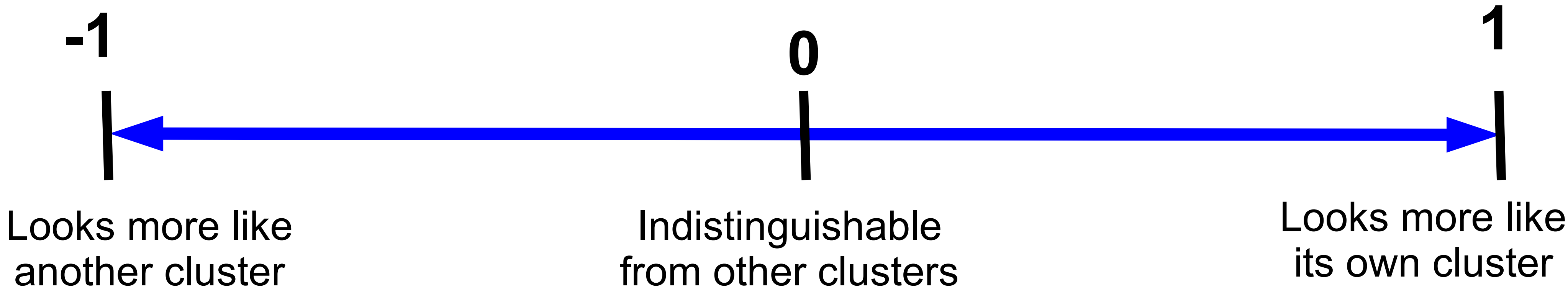
Leading p_T



Silhouette Values

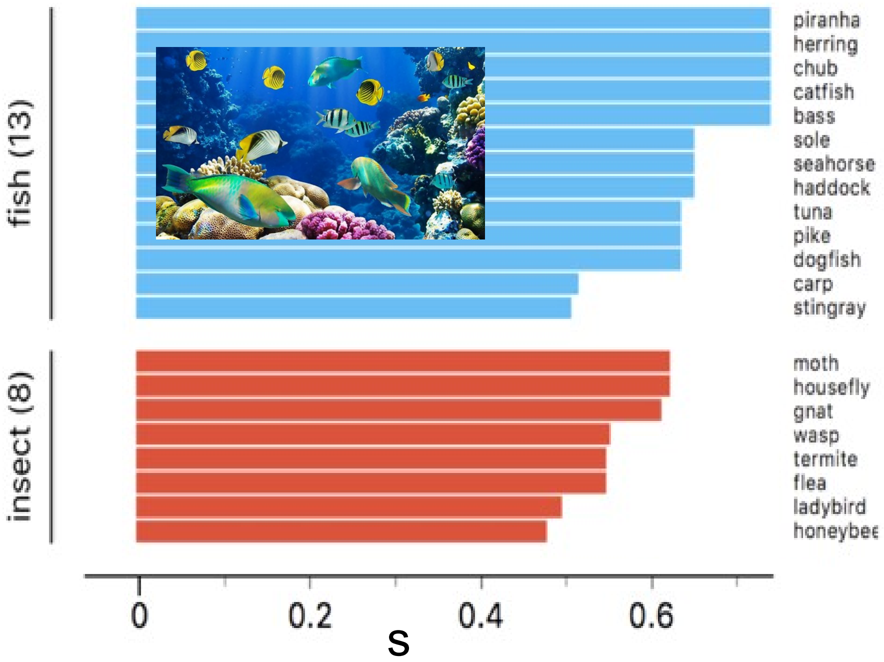
- Average distance between a jet candidate and other jet candidates in its cluster (signal or background) $a_i = \langle d_{i,j} \rangle_{j \neq i}$
- Average distance between jet candidate and jet candidates in the other cluster $b_i = \langle d_{i,j} \rangle$
- Silhouette value

$$s_i = \frac{b_i - a_i}{max[b_i, a_i]}$$

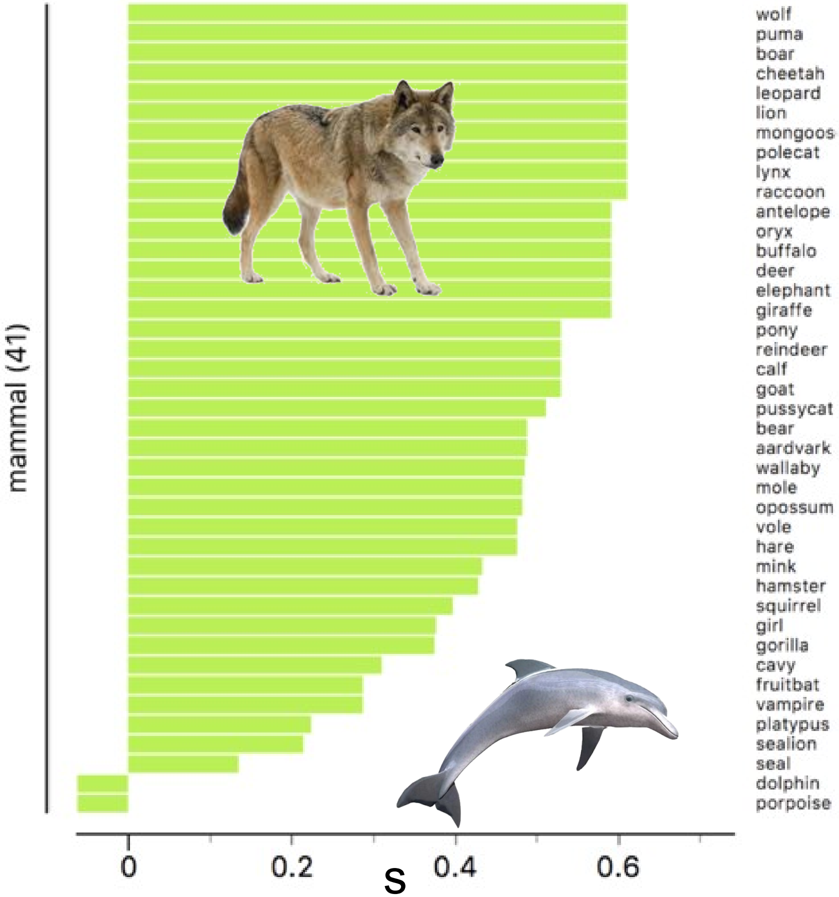


Silhouette values

Example from Wikipedia



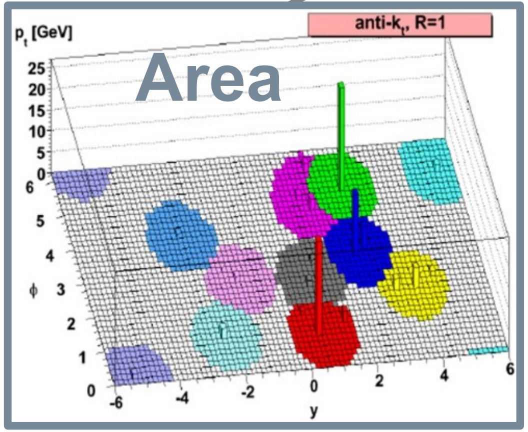
Silhouette scores from three types of animals rendered by Orange data mining suite.



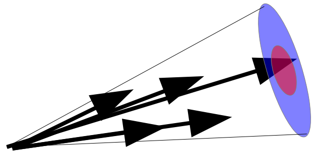
Silhouette Values

- Define a distance between two jet candidates to determine how similar they are

$$d_{i,j}=\sqrt{\left(\frac{A_i-A_j}{A^{max}-A^{min}}\right)^2+\left(\frac{\alpha_i-\alpha_j}{\alpha^{max}-\alpha^{min}}\right)^2+\left(\frac{\langle p_T\rangle_i-\langle p_T\rangle_j}{\langle p_T\rangle^{max}-\langle p_T\rangle^{min}}\right)^2+\left(\frac{p_{T,i}^L-p_{T,j}^L}{p_T^{L,max}-p_T^{L,min}}\right)^2}$$

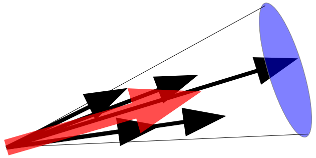


Angularity

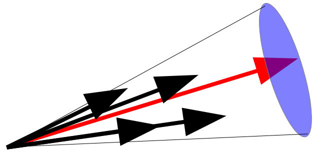


$$\alpha=\frac{1}{p_T^{jet}}\sum_k z_k\left(\vec{R}_k\right)$$

Average p_T



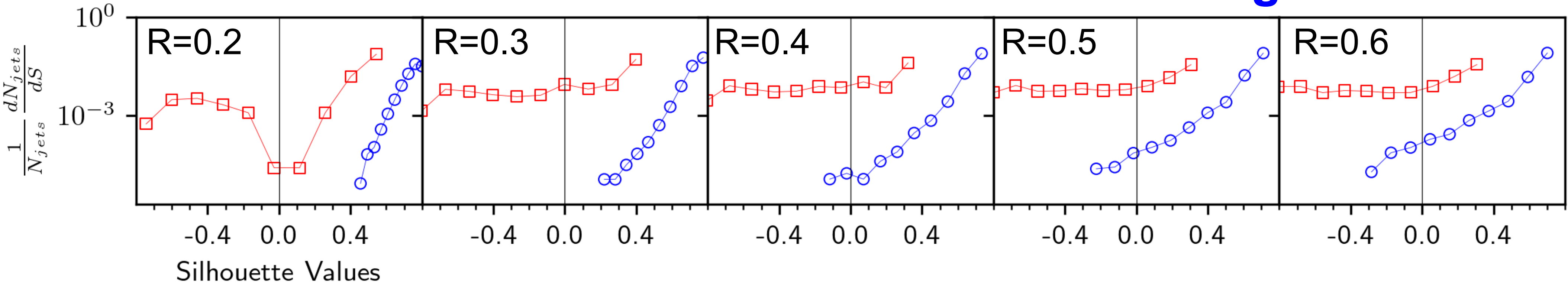
Leading p_T



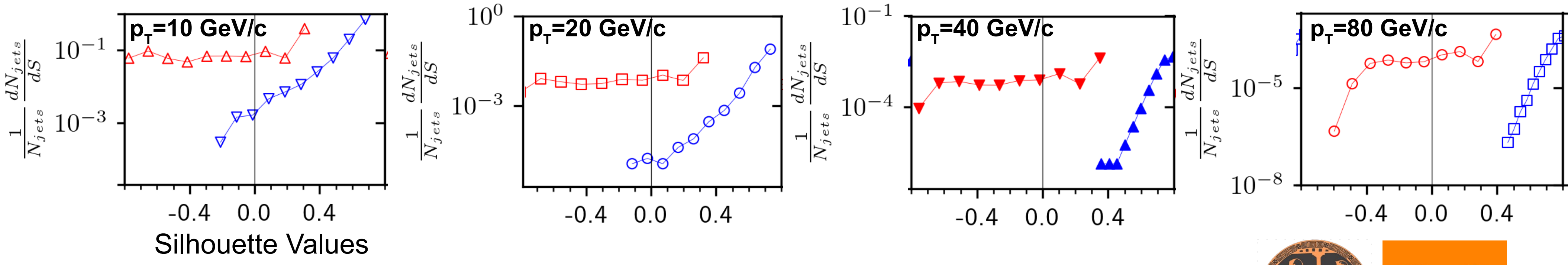
Before area cut

Combinatorial
Signal

p_T=20 GeV/c



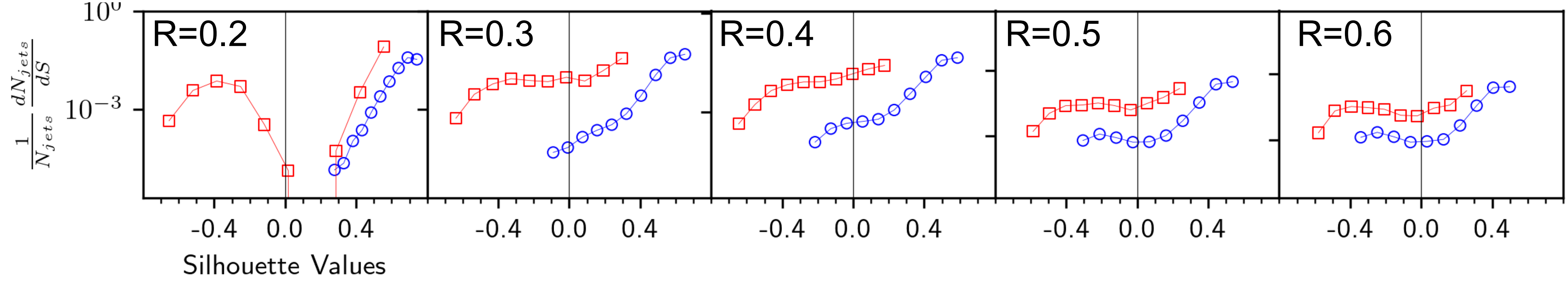
R=0.4



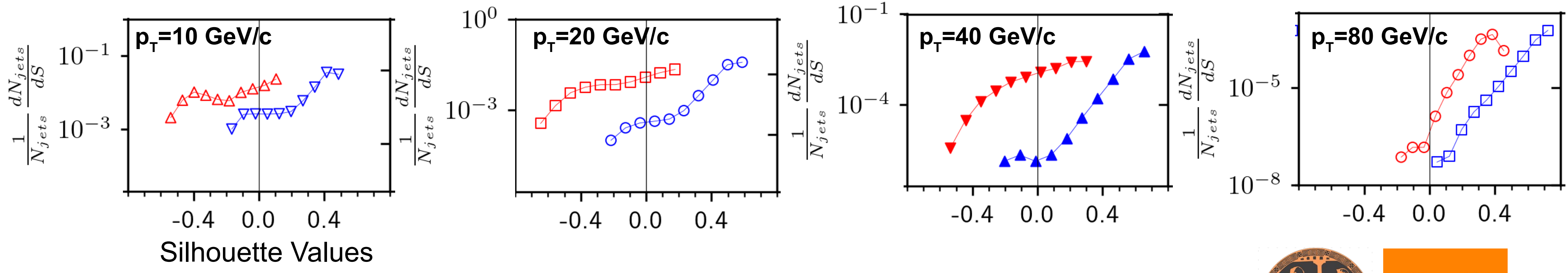
After area cut

Combinatorial
Signal

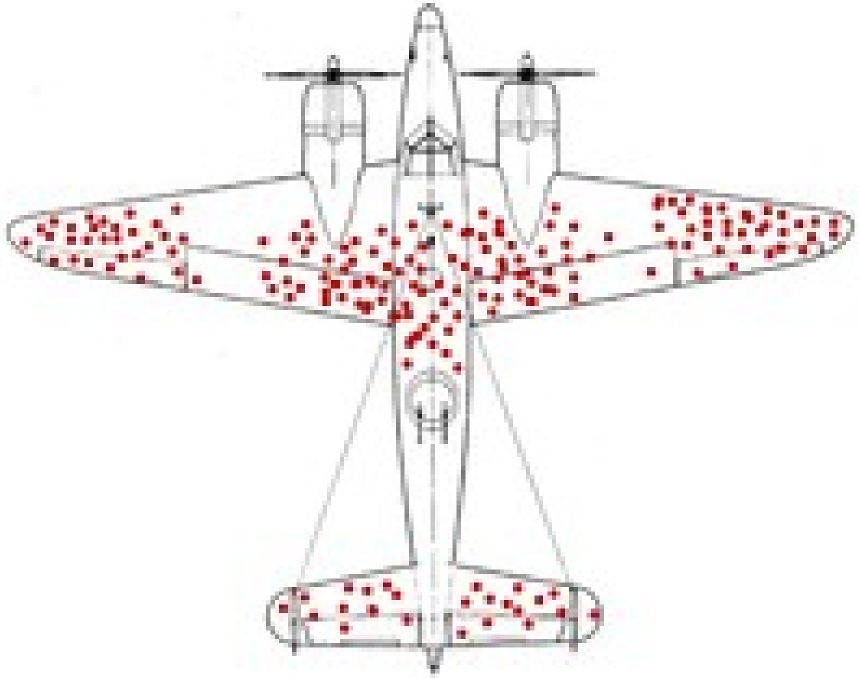
p_T=20 GeV/c



R=0.4



Survivor bias

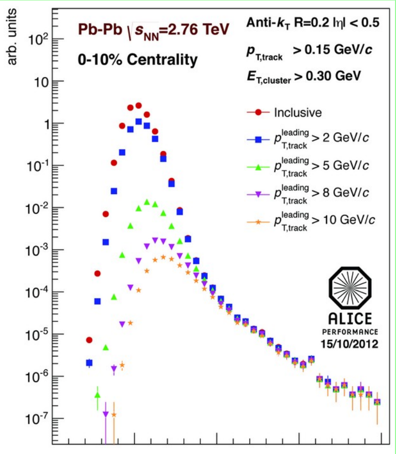


- WWII Example: holes planes returning indicate where it's *safer* to get hit
- We're looking at the real and combinatorial jets which *remain*

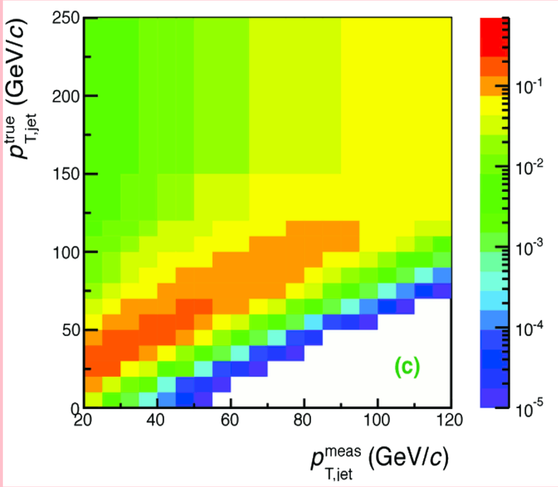
5. Background corrections in Monte Carlo

Analysis steps

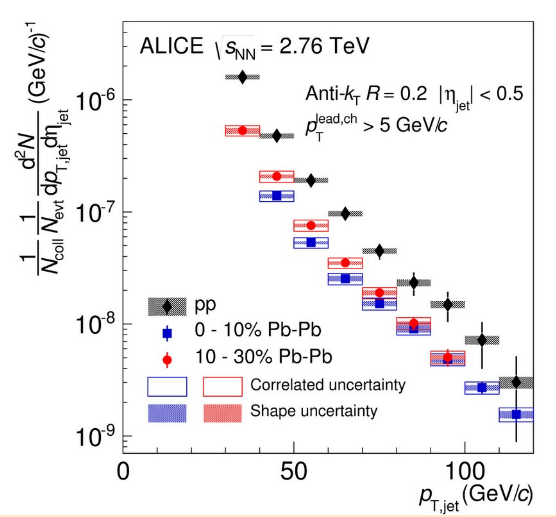
Tracks → Jet finding algorithm
Clusters → Jet finding algorithm
Jet finding algorithm → Jet candidates → Background subtraction



Jet spectrum smeared by energy resolution, background fluctuations

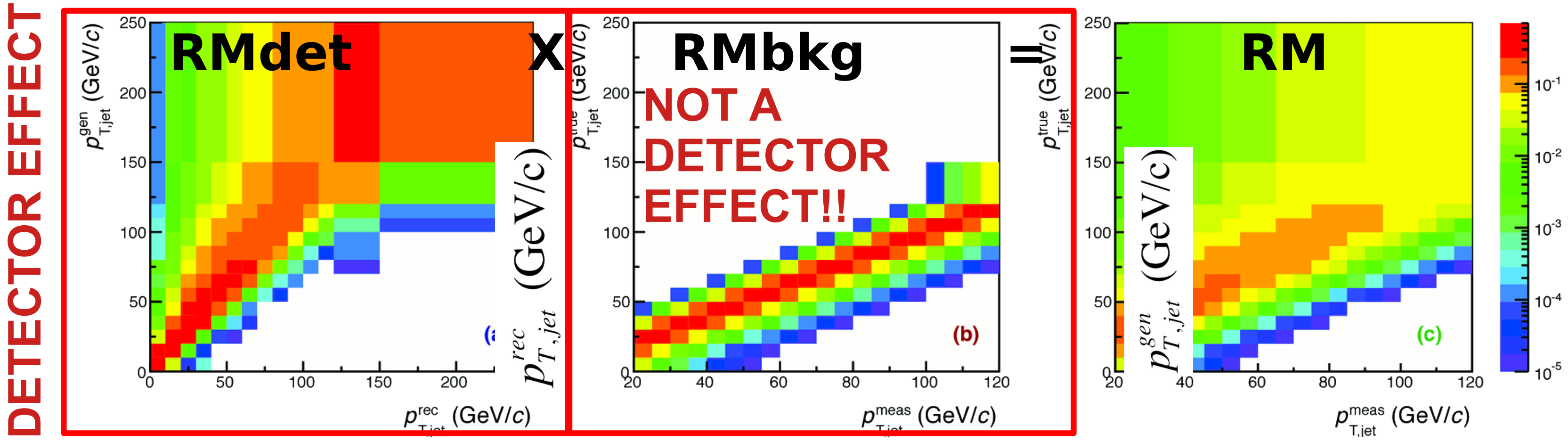


Unfolding – corrects for single track reco ε, E resolution, background fluctuations



Corrected spectra

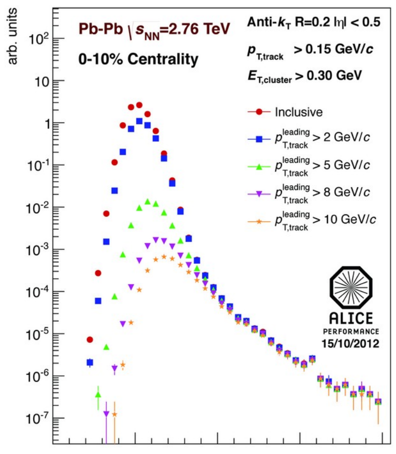
Jets in ALICE: Response Matrix Construction



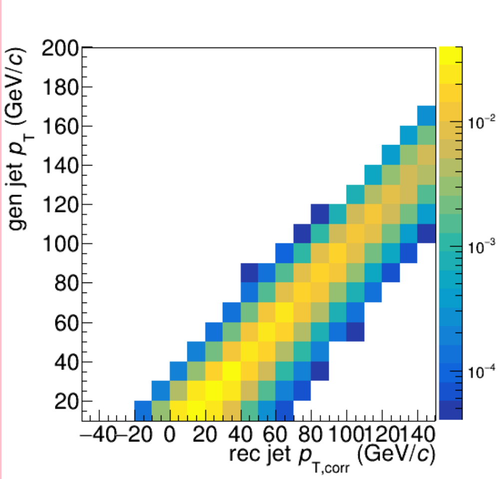
RM$_{bkg}$ and RM$_{det}$ are approximately factorizable

Analysis steps: Full Monte Carlo

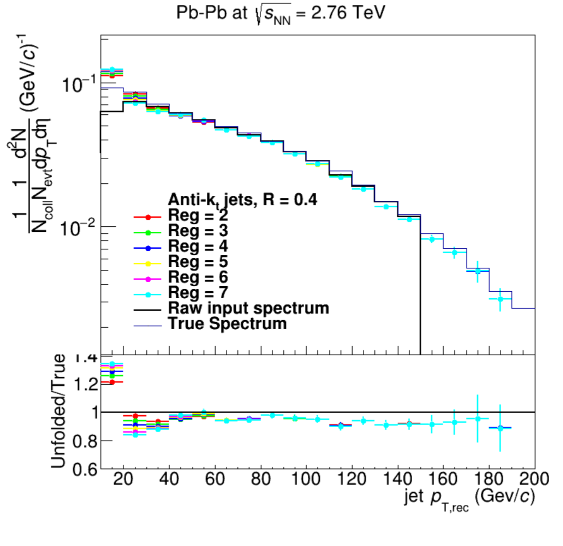
Particles → Jet finding algorithm → Jet candidates → Background subtraction



Jet spectrum smeared by energy resolution

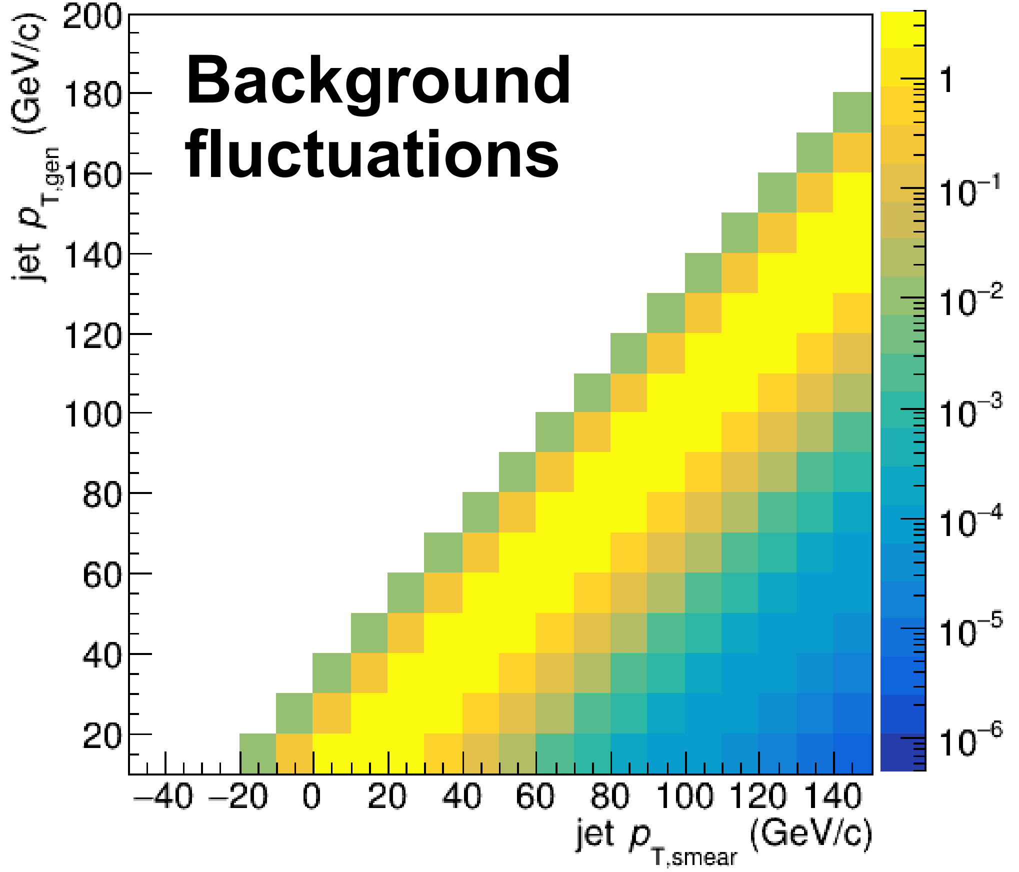


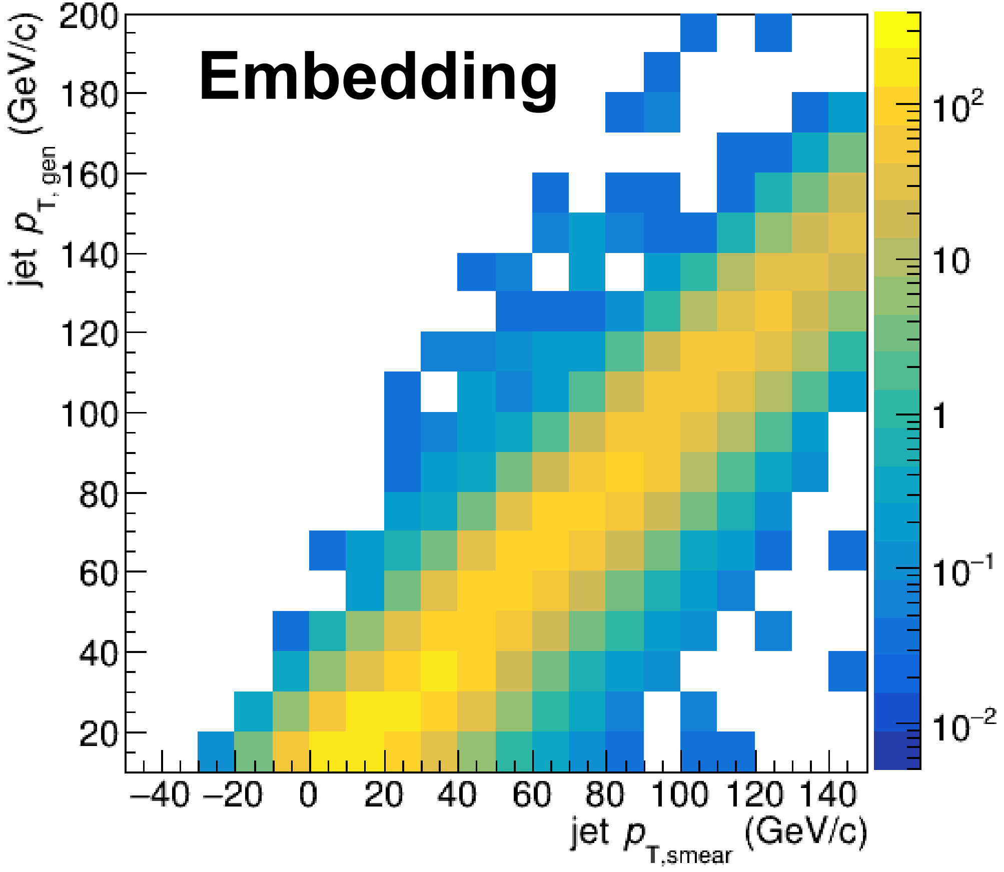
Unfolding – ***corrects for background fluctuations and other effects***

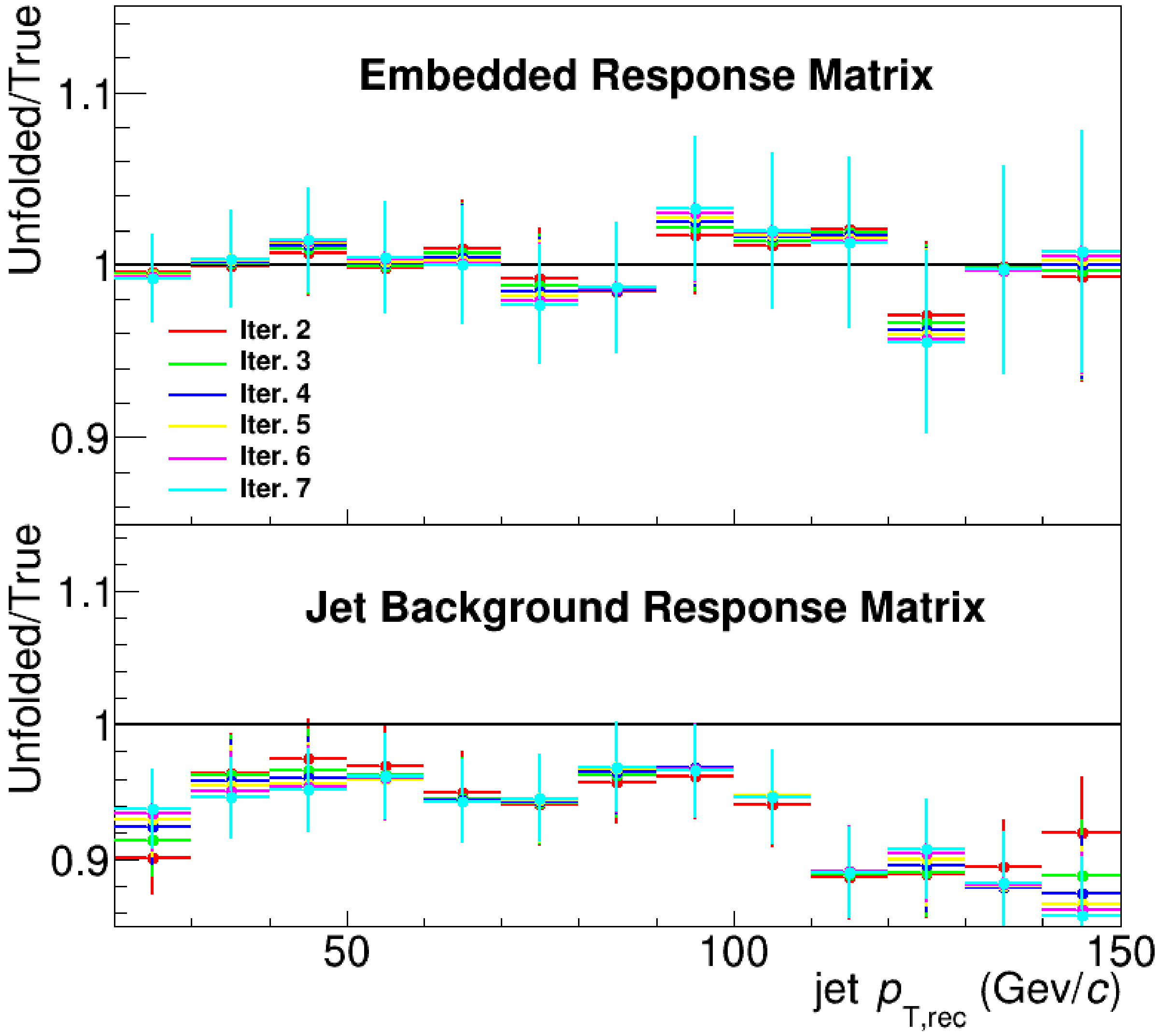


Corrected spectra

Construct a response matrix in Monte Carlo



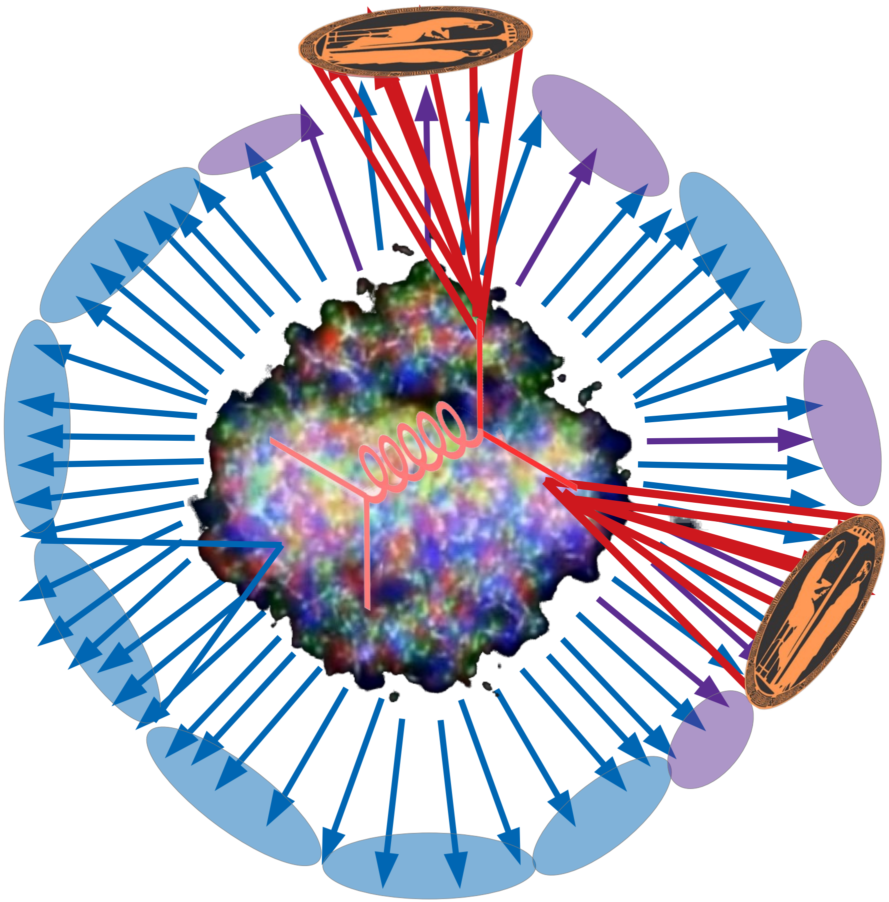




Closure

- Methods
  - Use δp_T method to measure width of fluctuations with varying numbers of leading jets (LJ) discarded
  - Embed PYTHIA event into heavy ion event
- Only embedding leads to full closure
- Not due to jet finder behavior in background → interplay between background and jet finder

Need unfolding and embedding in MC!



6. How to compare to models

Snowmass Accord: Apply the same algorithm to data and your model. Then the measurement and the calculation are the same.

Rivet: Apply the same algorithm to data and your model. Then the measurement and the calculation are the same.

What is Rivet?

AMPT

HIJING

PYTHIA

Monte Carlo Model

JEWEL

EPOS

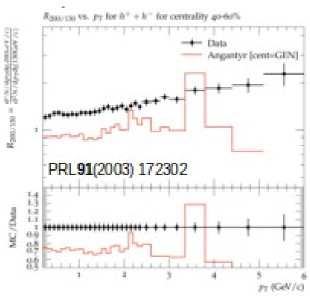
smash

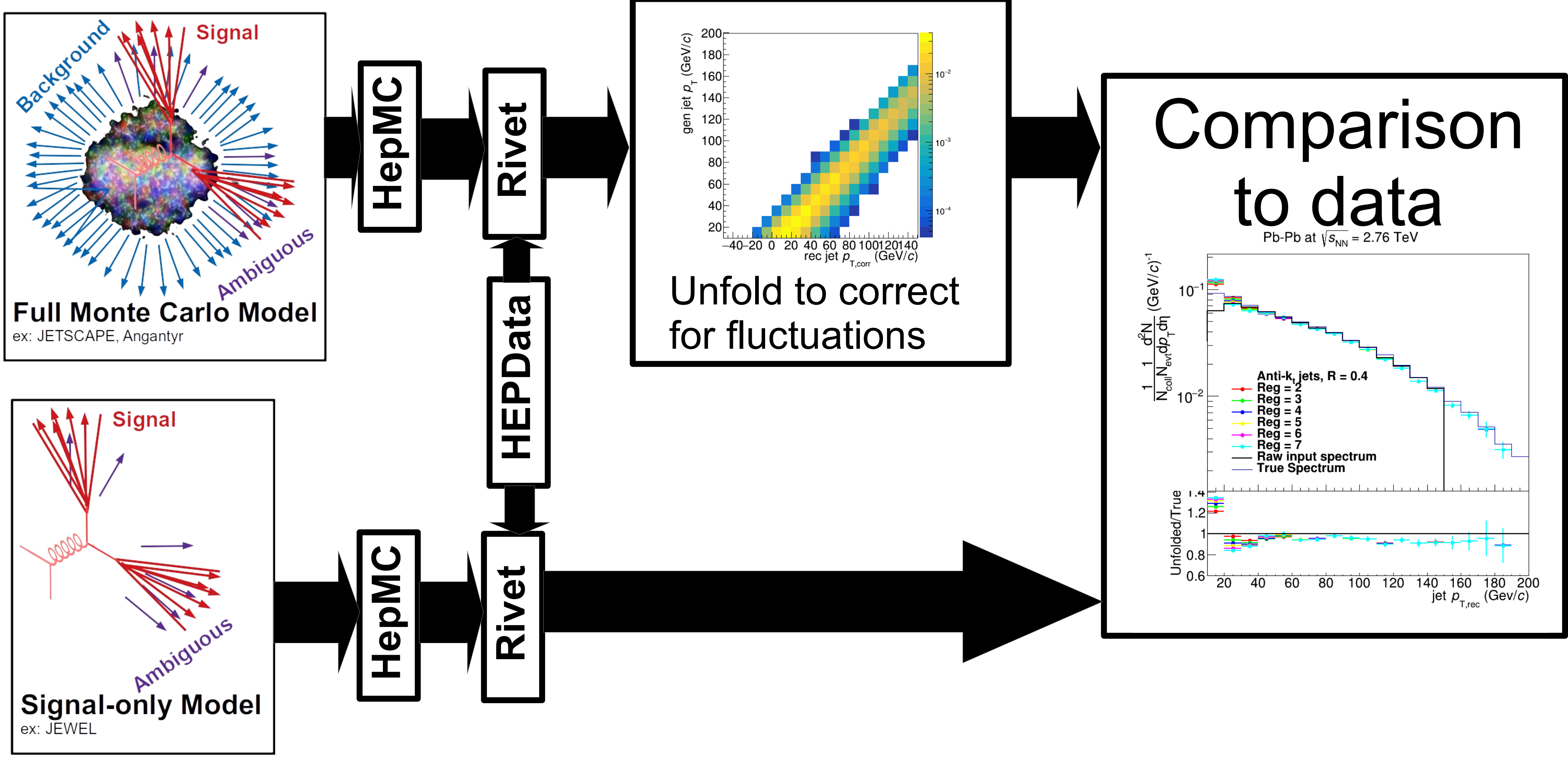
HepMC

HEPData

Rivet

Comparison to data



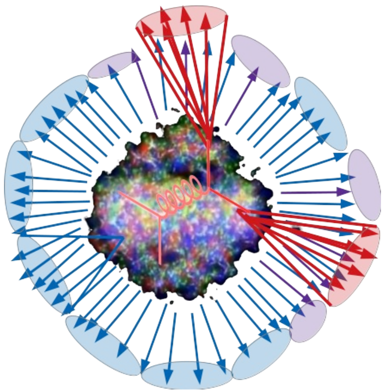


Background
Signal
Ambiguous
Full Monte Carlo Model
ex: JETSCAPE, Angantyr
HepMC
Rivet
HEPData
gen jet p_T (GeV/c)
rec jet $p_{T,corr}$ (GeV/c)
Unfold to correct for fluctuations
Comparison to data
Pb-Pb at $\sqrt{s_{NN}}$ = 2.76 TeV
Anti-k_t jets, R = 0.4
Reg = 2
Reg = 3
Reg = 4
Reg = 5
Reg = 6
Reg = 7
Raw input spectrum
True Spectrum
Unfolded/True
jet $p_{T,rec}$ (Gev/c)
Signal
Ambiguous
Signal-only Model
ex: JEWEL
HepMC
Rivet

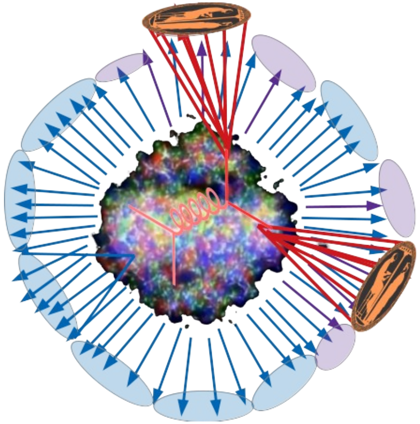
Why use Rivet?

- Facilitates comparisons between Monte Carlos and data
- It's not that hard
- It preserves analysis details

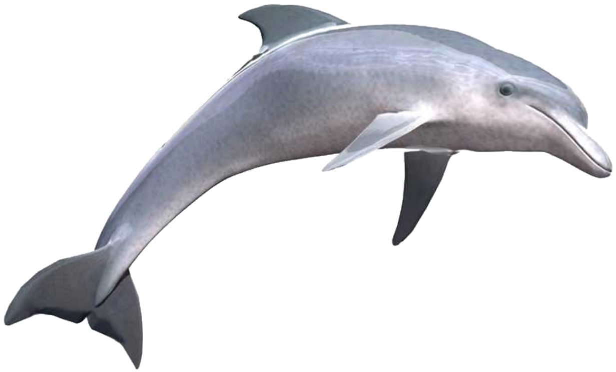
Conclusions



Models have background too!



Correcting for it requires unfolding, embedding



Background suppression → combinatorial jets which look like real jets

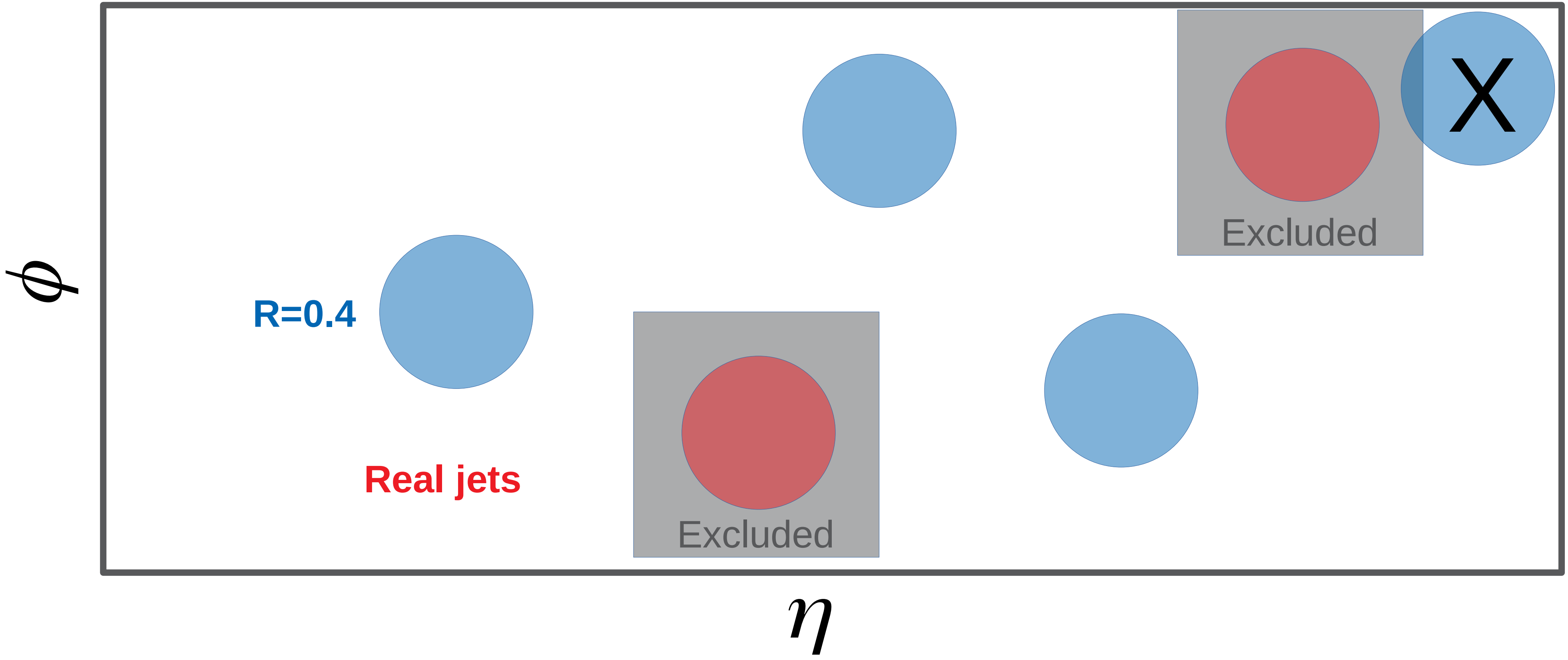
Treat models like data

Recorded tutorials from Rivetizing Heavy Ion Collisions at RHIC

4. Background:
Not just an experimental problem!

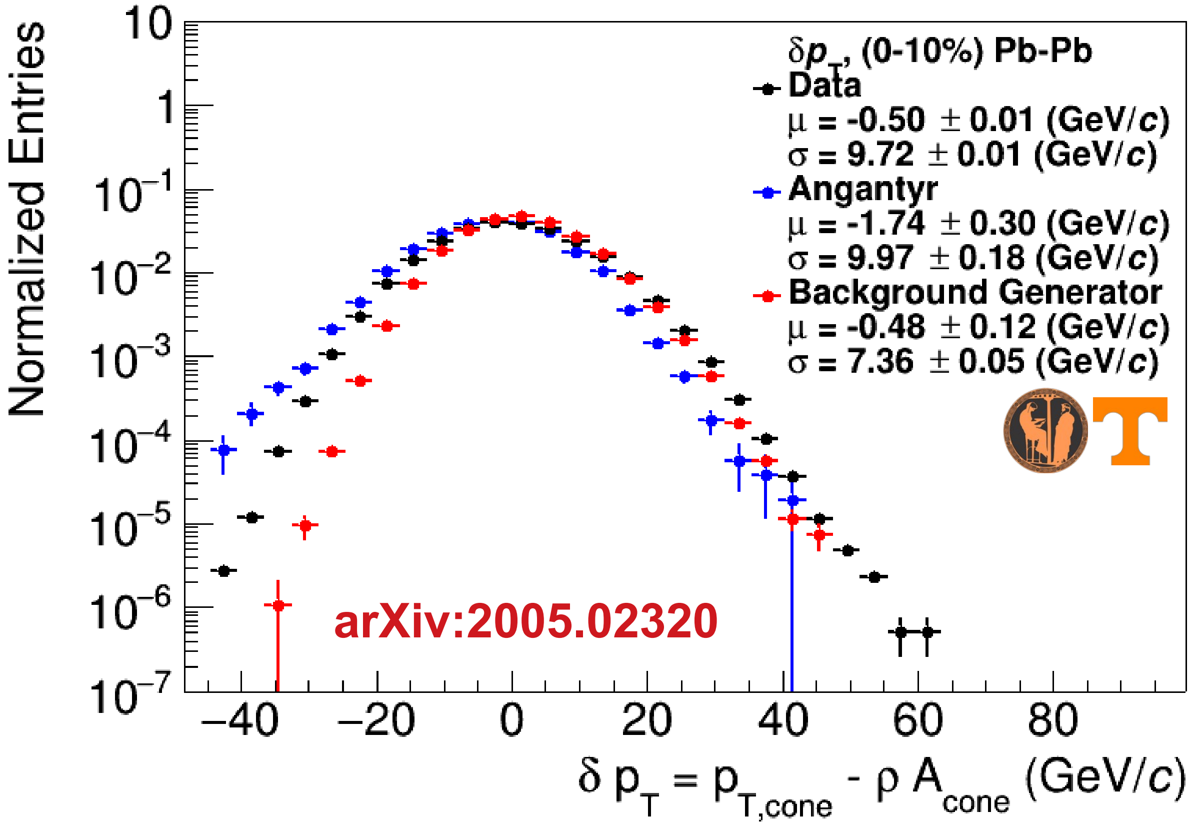
arXiv:2005.02320

Random cones



Random cones

ALICE Data: JHEP 03 (2012) 053



$\rho = \mathrm{Median}(p_T/A)$

Shape of width of the distribution

Single particle spectra

$$f_\Gamma(p_T, p, b) = \frac{b}{\Gamma(p)} (b\, p_T)^{p-1} e^{-bx}$$

$$\boxed{\frac{dN}{dy} \propto f_\Gamma(p_T, 2, b) = b^2 p_T e^{-k p_T}}$$

$$\mu_{p_T} = \frac{p}{b}, \sigma_{p_T} = \frac{\sqrt{p}}{b}$$

Tannenbaum, PLB(498),1–2,Pg.29-34(2001)

Assumes shape

Σp_T of N particles → N-fold convolution:

$$f_N(p_T, p, b) = f_\Gamma(p_T, Np, b) \qquad \frac{dpT^{total}}{dy} \propto f_N(p_T, Np, b)$$

$$N = \frac{N_{total}}{A_{total}} \pi R^2 \qquad \mu_{total} = \frac{Np}{b} = N \mu_{p_T}, \sigma_{total} = \frac{\sqrt{Np}}{b} = \sqrt{N} \sigma_{p_T}$$

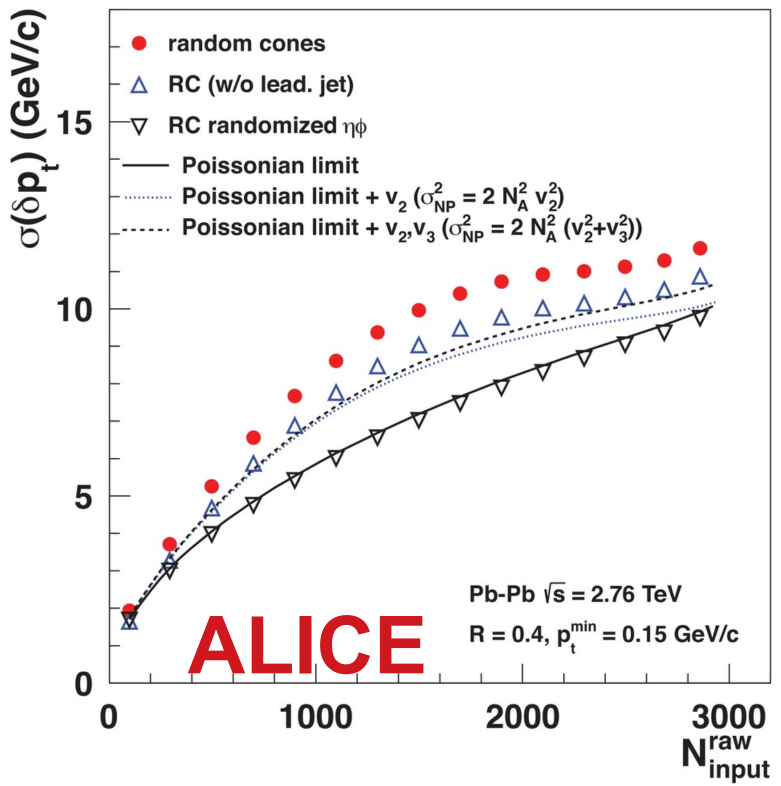
Add Poissonian fluctuations in N: $\sigma_{total} = \sqrt{N \sigma_{p_T}^2 + N \mu_{p_T}^2}$

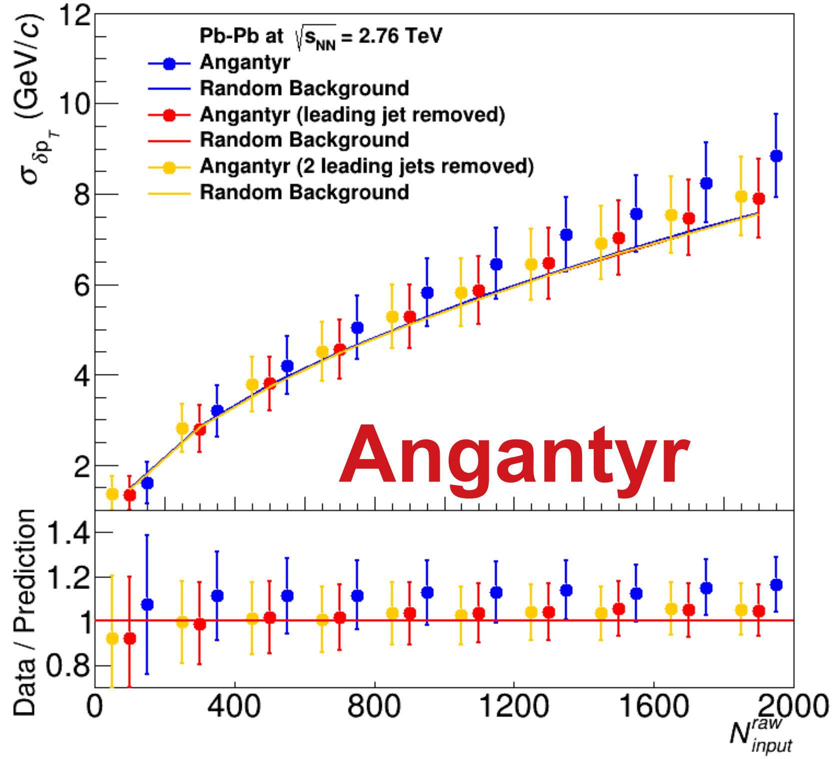
Add non-Poissonian fluctuations in N due to flow

$$\sigma_{total} = \sqrt{N \sigma_{p_T}^2 + \left(N + \boxed{2 \sum_n v_n^2}\right) \mu_{p_T}^2}$$

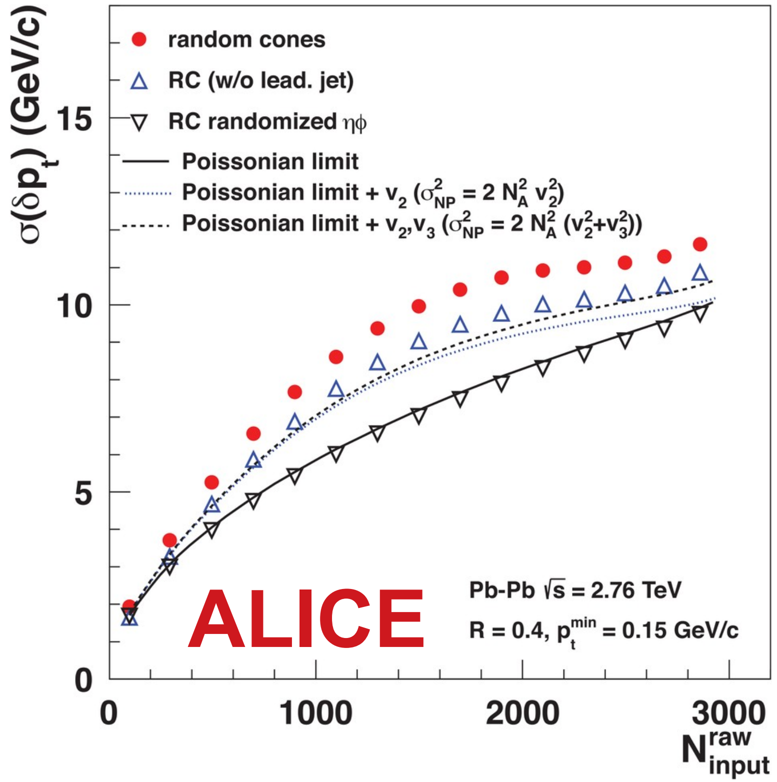
Assumes uncorrelated number fluctuations

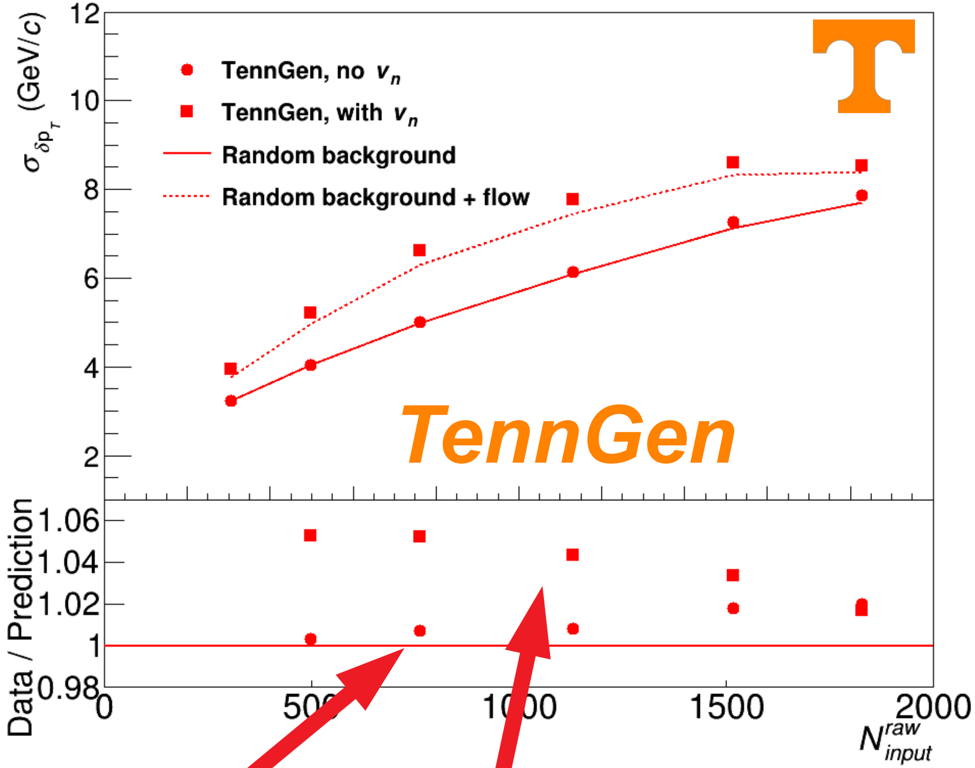
Width vs multiplicity





Width vs multiplicity





Impact of shape of spectrum

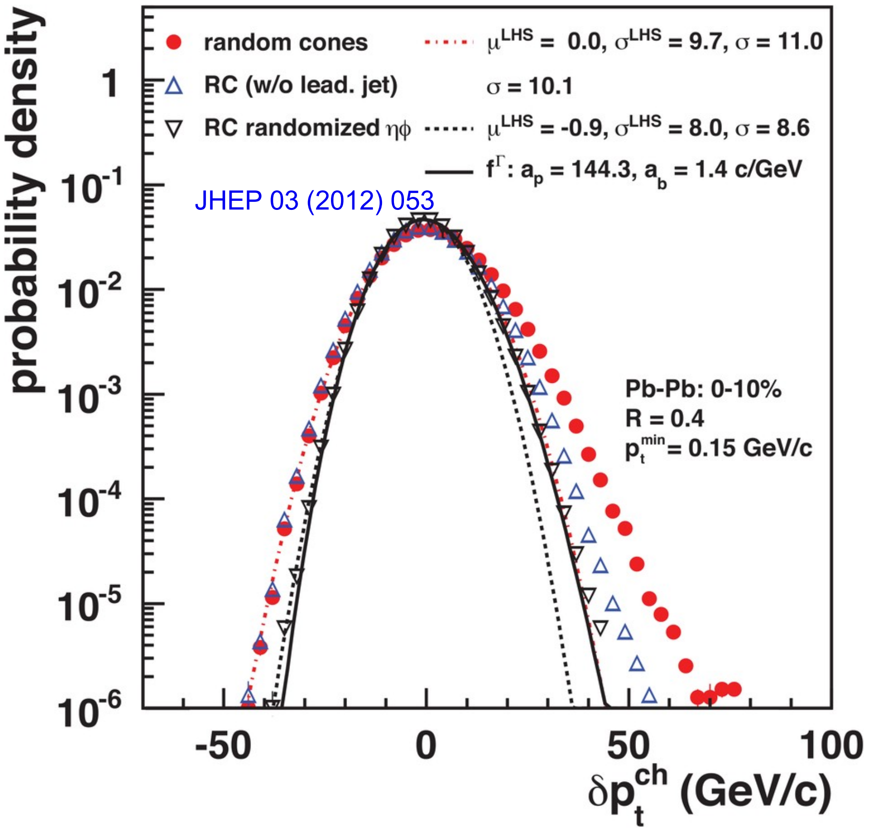
Correlations between event planes

Backup

Random cones in ALICE

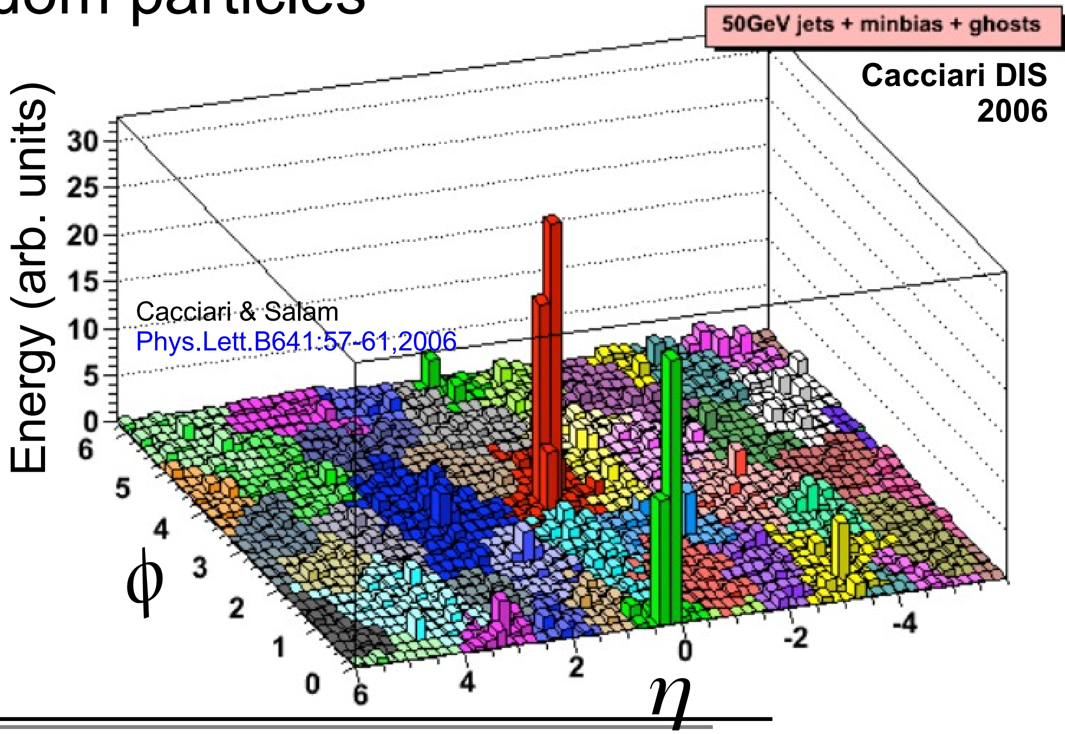
- Estimate ρ
  - k_T jet finder → jet candidates
  - ρ = Median(p_T/A)
- Draw Random cone

$$\delta p_T = p_T^{reco} - \rho A$$



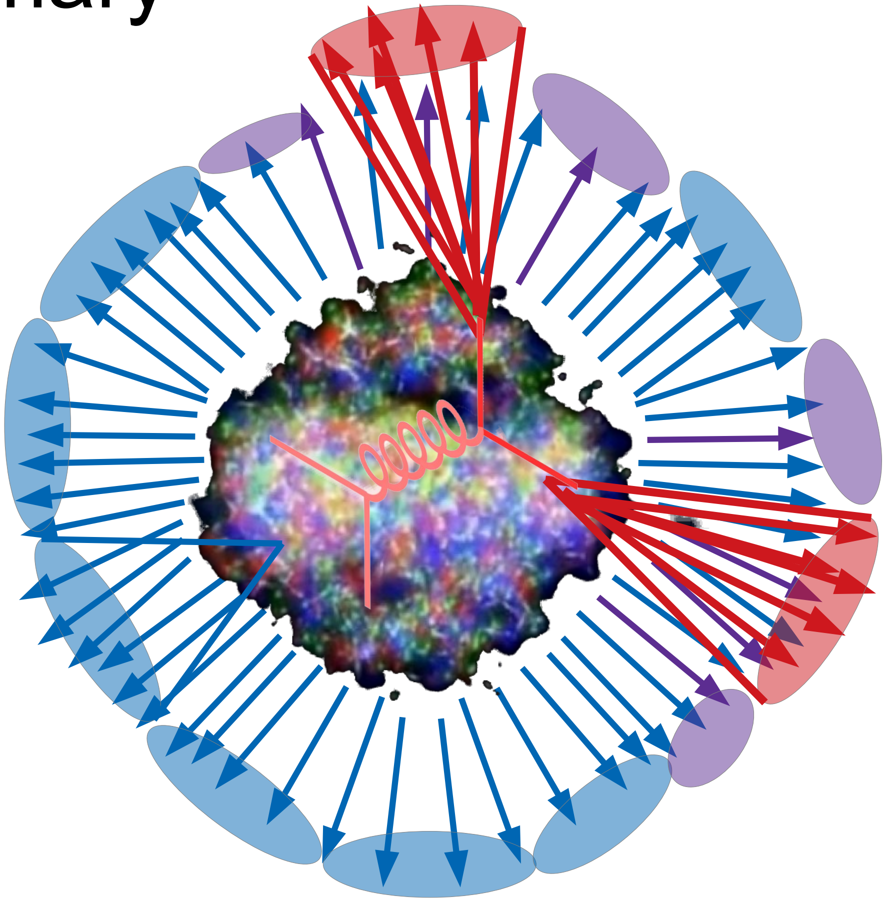
Mini-summary

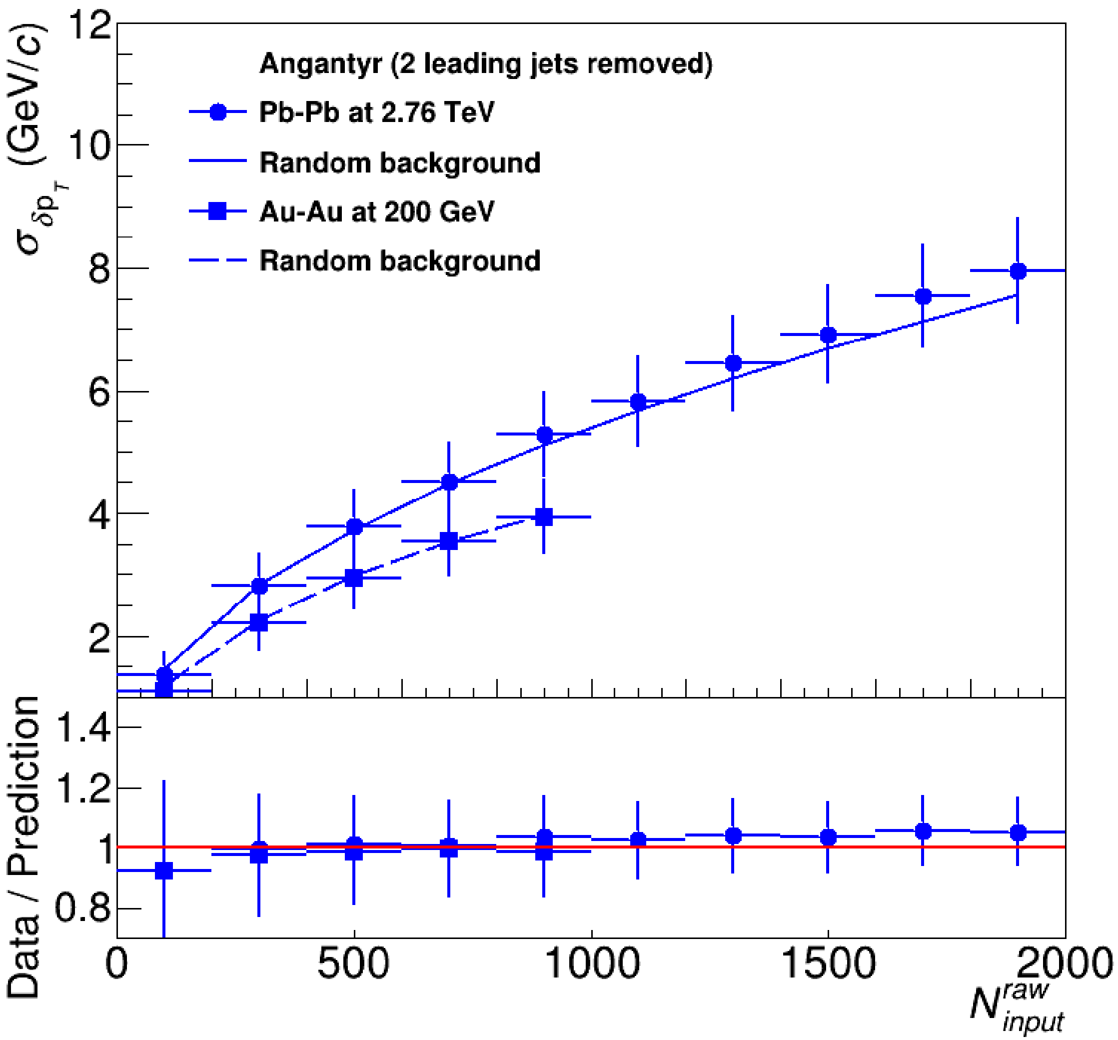
- Jet finders put all input clusters, tracks in a jet candidate
- Background is *dominated* by random particles
  - But ~5% effects from non-Poissonian fluctuations
- Models have background too!
  - Sensitive to multiplicity, implementation of flow

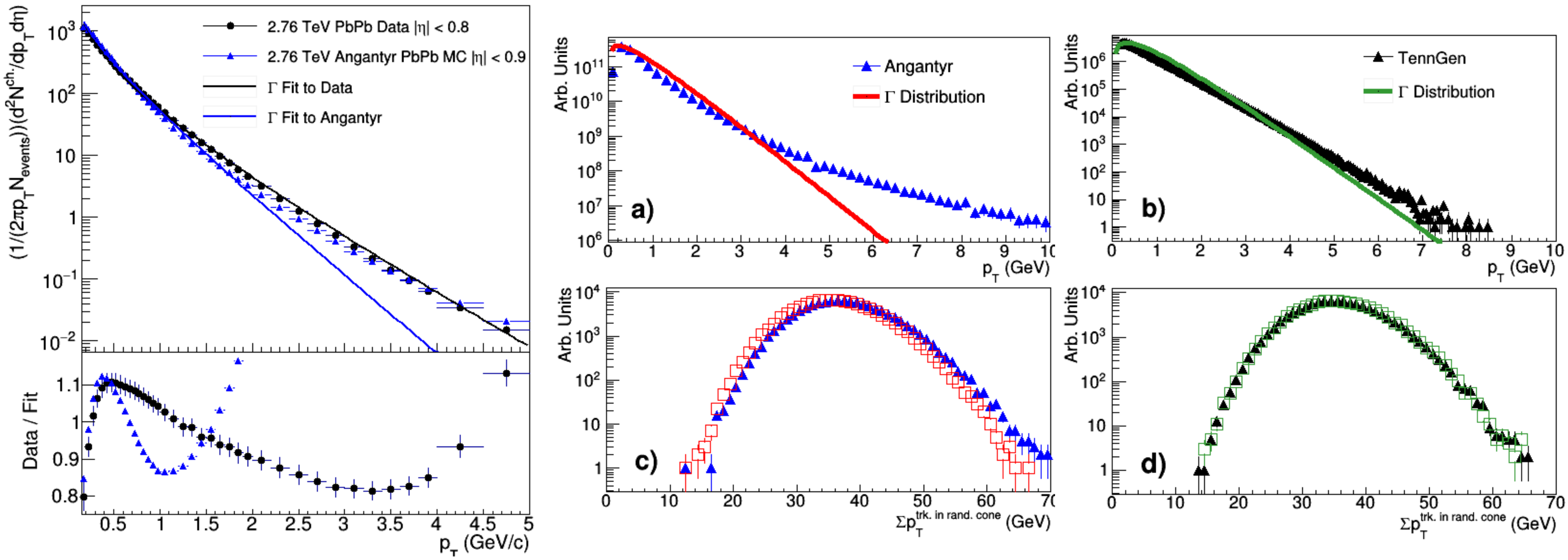


Mini-summary

- “Signal” and “background” have different properties, but...
- Always overlap somewhat
- Any procedure to remove “background” will also cut signal




Angantyr (2 leading jets removed)
Pb-Pb at 2.76 TeV
Random background
Au-Au at 200 GeV
Random background
$\sigma_{\delta p_T}$ (GeV/c)
Data / Prediction
N_{input}^{raw}
0
500
1000
1500
2000
2
4
6
8
10
12
0.8
1
1.2
1.4



2.76 TeV PbPb Data |η| < 0.8
2.76 TeV Angantyr PbPb MC |η| < 0.9
Γ Fit to Data
Γ Fit to Angantyr
(1/(2πp_T N_events))(d²N^ch./dp_T dη)
Data / Fit
p_T (GeV/c)
Angantyr
Γ Distribution
a)
Arb. Units
p_T (GeV)
TennGen
Γ Distribution
b)
c)
d)
Σp_T^trk. in rand. cone (GeV)

Area-based background subtraction

Cacciari & Salam, PLB659:119–126,2008

Particles, clusters

k_T algorithm

$$k_T = p_T, \Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$$

- For all *i,j* calculate:

$$d_{ij} = min(p_{T,i}^2, p_{T,j}^2) \Delta R_{ij}^2$$

$$d_{iB} = p_{T,i}$$

- Combine smallest d_{ij}.

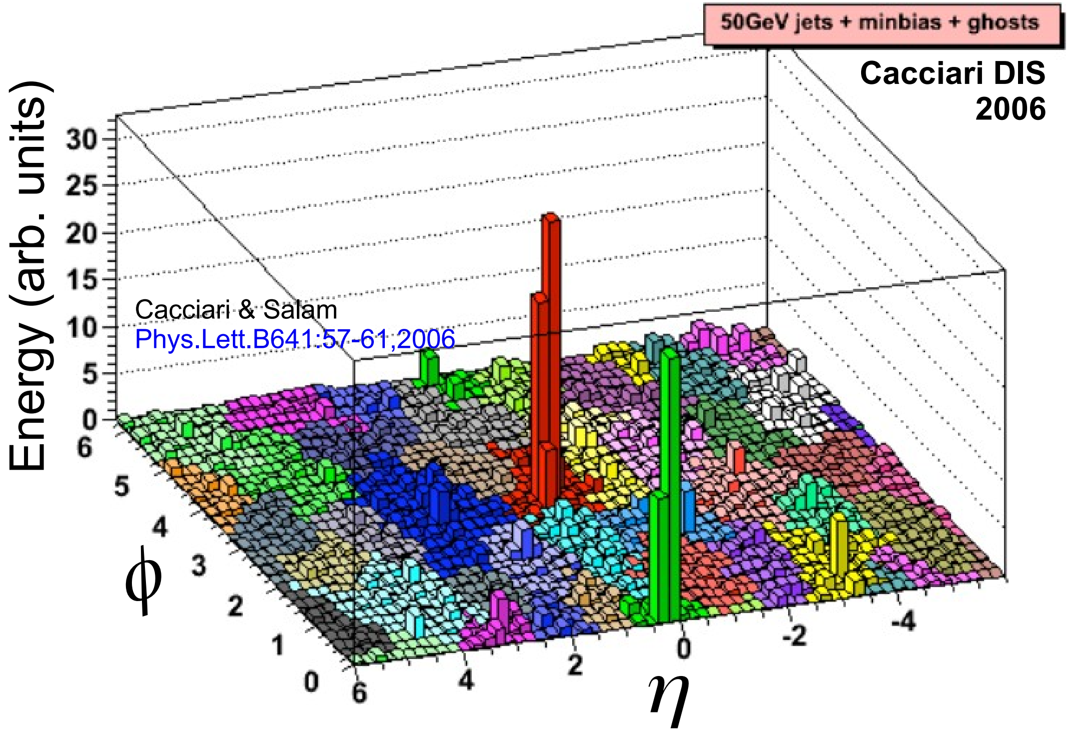
If d_{iB} smallest, $d_{iB} \rightarrow$ jet

Repeat until no particles left

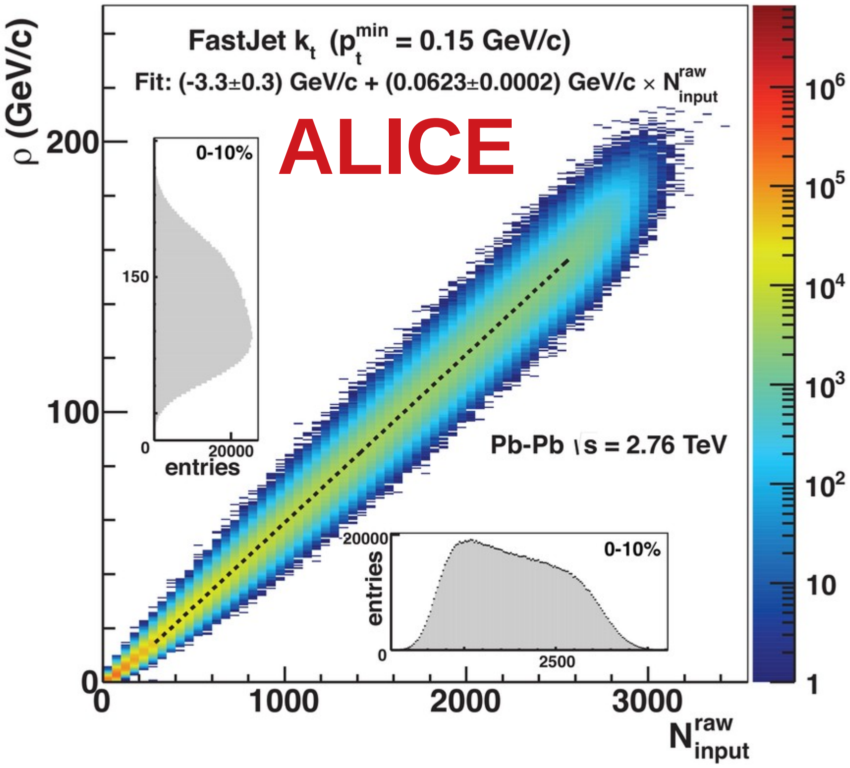
Jet candidates

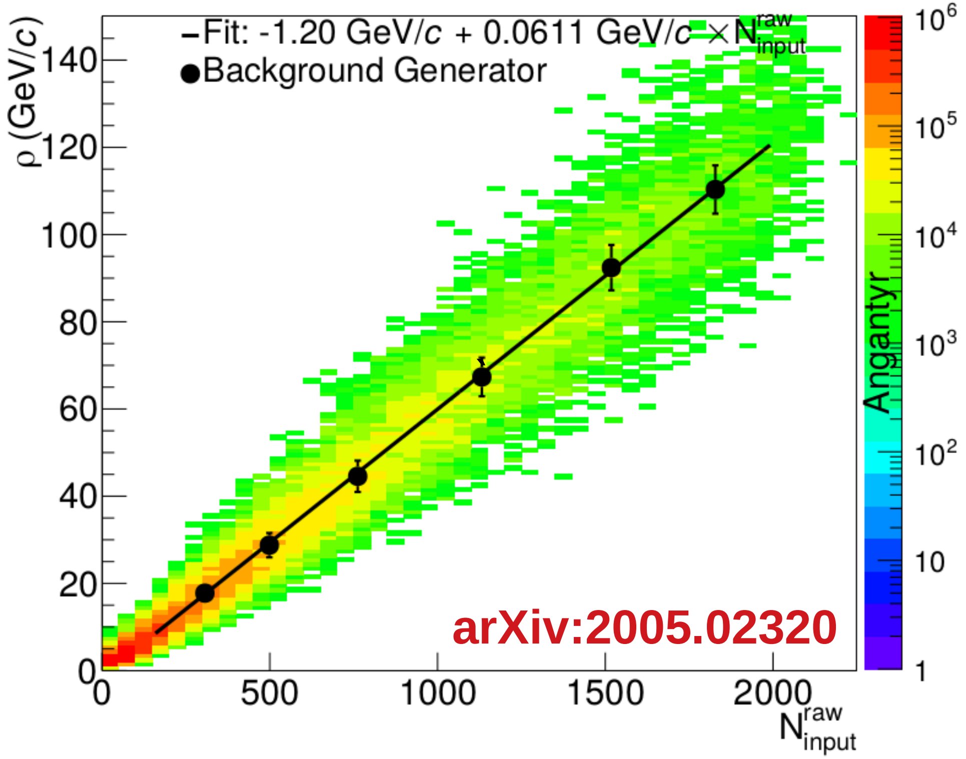
Median $\rho = p_T/A$

$$p_T^{jet} = p_T^{reco} - \rho_{median} A^{jet}$$

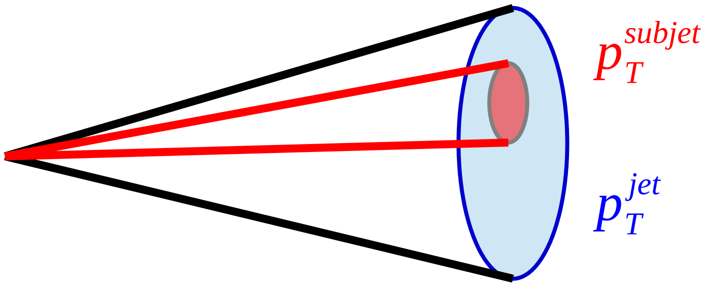


Background density ρ

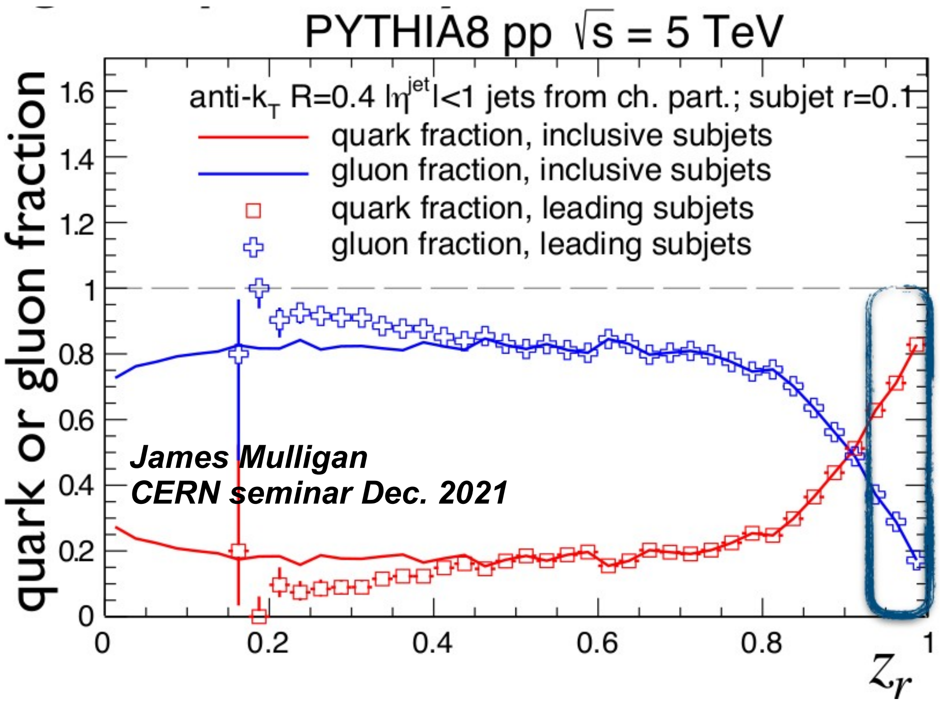




Subjet z



$$z_r = \frac{p_T^{subjet}}{p_T^{jet}}$$



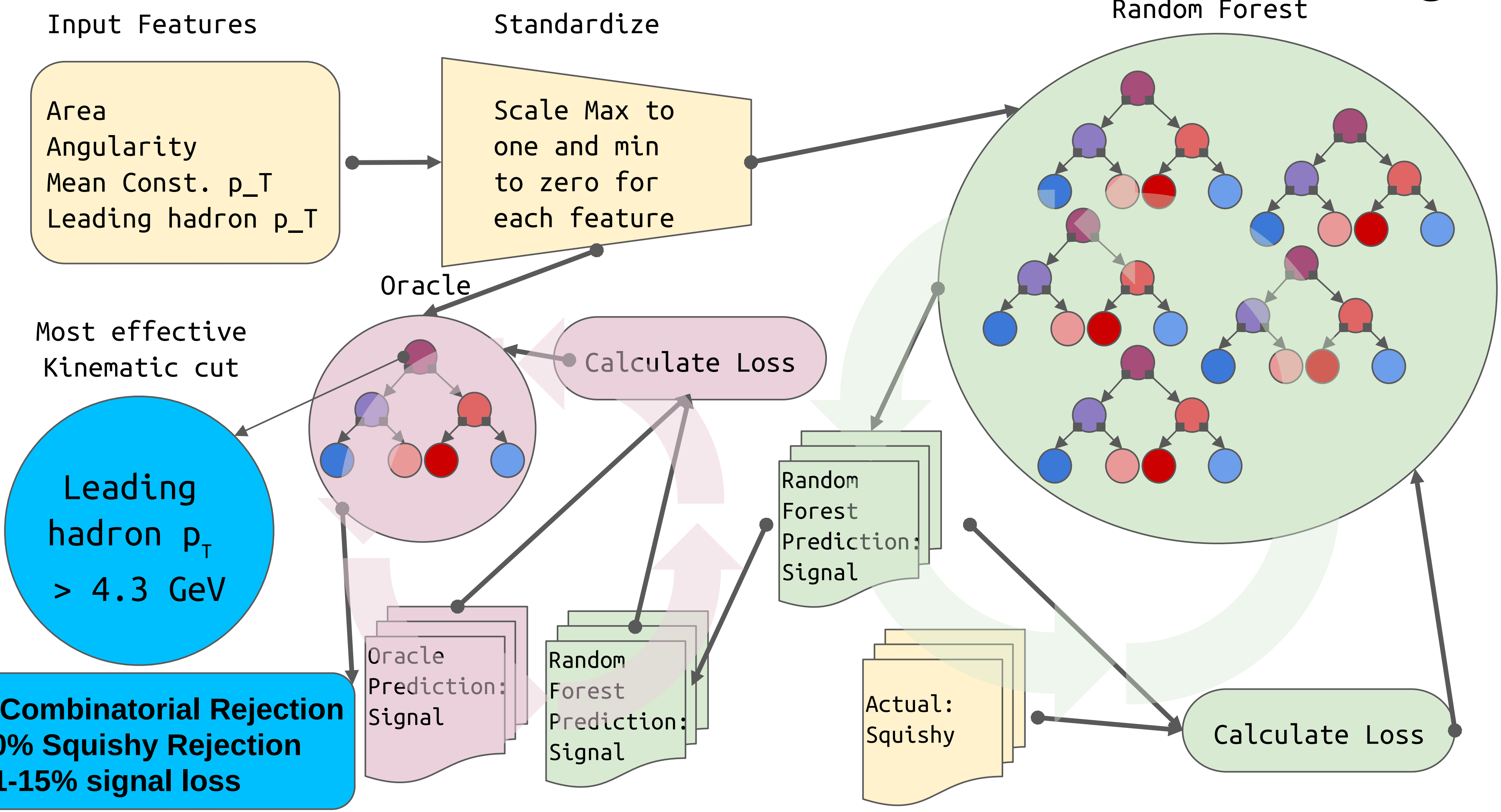
- Cluster jets with anti-k_T with resolution parameter R
- Recluster constituents with anti-k_T with resolution parameter r
- Some discriminating power between quark-like and gluon-like jets
  - Strained at low momentum, small R

Subjet z_r: Area cut

Subjet z_r: Leading hadron p_T^1 cut

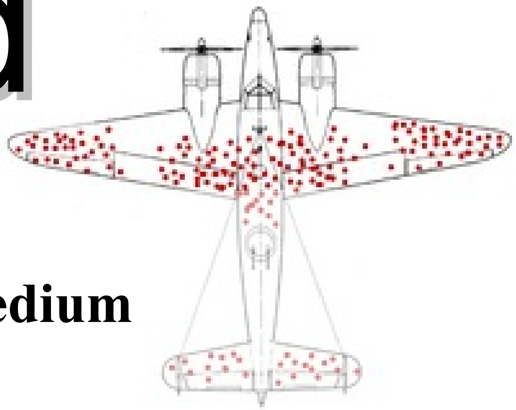
Subjet z_r: Leading hadron p_T^1 cut

Look for a clever solution with Machine Learning



75-90% Combinatorial Rejection
40-90% Squishy Rejection
1-15% signal loss

Bias & background



- **Background suppression → Bias**
- **Survivor bias: Modified jets probably look more like the medium**
- **Quark/Gluon bias:**
  - **Quark jets are narrower, have fewer tracks, fragment harder [Z Phys C 68, 179-201 (1995), Z Phys C 70, 179-196 (1996),]**
  - **Gluon jets reconstructed with k_T algorithm have more particles than jets reconstructed with anti-k_T algorithm [Phys. Rev. D 45, 1448 (1992)]**
  - **Gluon jets fragment into more baryons [EPJC 8, 241-254, 1998]**
- **Fragmentation bias: Experimental measurements explicitly select jets with hard fragments**