A tale of two jets
Christine Nattrass, University of Tennessee, Knoxville
Acknowledgements

Antonio Da Silva

Patrick Steffanic

Charles Hughes
What a theorist needs to know about background

• You have background too!
• The distinction between signal and background is somewhat arbitrary
• Experimental background subtraction techniques may lead to non-trivial bias
• The gold standard is treating the model exactly like the data
Background is not just an experimental problem

**TennGen** background generator

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

### Track properties
- **Momentum spectra**
  - Random $p_T$
  - Blast Wave Fit
  - $K^+$ 
  - ALICE PLB720 (2013) 52-62

- **Polynomial Fit**
  - $v_n$
  - ALICE JHEP 1609 (2016) 164

- No jets! No resonances
- Emulates hydro correlations
PYTHIA Angantyr

Based on PYTHIA 8
Sjöstrand, Mrenna & Skands,
JHEP05 (2006) 026

Based on Fritiof & wounded nucleons

N-N collisions w/fluctuating radii → fluctuating $\sigma$

Lots of jets! And resonances!
No hydrodynamics, no jet quenching
Area-based background subtraction


**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:
  \[ d_{ij} = \min \left( p_{T,i}^2, p_{T,j}^2 \right) \Delta R_{ij}^2 \]
  \[ d_{iB} = p_{T,i} \]
- Combine smallest \(d_{ij}\):
  If \(d_{iB}\) smallest, \(d_{iB} \rightarrow \text{jet}\)
- Repeat until no particles left

Jet candidates

Median \(\rho = p_T / A\)

\[ p_T^{\text{jet}} = p_T^{\text{reco}} - \rho_{\text{median}} A^{\text{jet}} \]
Background density $\rho$

FastJet $k_t$ ($p_T^{\text{min}} = 0.15$ GeV/c)
Fit: $(-3.3\pm0.3)$ GeV/c $+$ $(0.0623\pm0.0002)$ GeV/c $\times N_{\text{raw}}^{\text{input}}$

Pb-Pb $\sqrt{s} = 2.76$ TeV

ARXIV:2005.02320
Random cones

- Real jets
- Excluded

$R=0.4$

$\phi$

$\eta$
Random cones in ALICE

- Estimate $\rho$
  - $k_T$ jet finder $\rightarrow$ jet candidates
  - $\rho = \text{Median}(p_T/A)$
- Draw Random cone

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

JHEP 03 (2012) 053

\( \delta p_T = p_{T,\text{cone}} - \rho A_{\text{cone}} \) (GeV/c)

\( \delta p_T \) (0-10\%) Pb-Pb
- Data
  \( \mu = -0.50 \pm 0.01 \) (GeV/c)
  \( \sigma = 9.72 \pm 0.01 \) (GeV/c)
- Angantyr
  \( \mu = -1.74 \pm 0.30 \) (GeV/c)
  \( \sigma = 9.97 \pm 0.18 \) (GeV/c)
- Background Generator
  \( \mu = -0.48 \pm 0.12 \) (GeV/c)
  \( \sigma = 7.36 \pm 0.05 \) (GeV/c)

Shape of width of the distribution

Single particle spectra

\[ f_{\Gamma}(p_T, p, b) = \frac{b}{\Gamma(p)}(bp_T)^{p-1} e^{-bx} \]
\[ \frac{dN}{dy} \propto f_{\Gamma}(p_T, 2, b) = b^2 p_T e^{-kp_T} \]
\[ \mu_{p_T} = \frac{p}{b}, \sigma_{p_T} = \frac{\sqrt{p}}{b} \]

Σp_T of N particles→N-fold convolution:

\[ f_N(p_T, p, b) = f_{\Gamma}(p_T, Np, b) \]
\[ \frac{dp_T^{\text{total}}}{dy} \propto f_N(p_T, Np, b) \]
\[ N = \frac{N_{\text{total}}}{A_{\text{total}}} \pi R^2 \]
\[ \mu_{\text{total}} = \frac{Np}{b} = N \mu_{p_T}, \sigma_{\text{total}} = \sqrt{Np} = \sqrt{N} \sigma_{p_T} \]

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

Add non-Poissonian fluctuations in N due to flow

\[ \sigma_{\text{total}} = \sqrt{N} \sigma_{p_T}^2 + (N + 2 \sum \nu_n^2) \mu_{p_T}^2 \]

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

\[ \sigma(p_T) \text{ (GeV/c)} \]

- random cones
- RC (w/o lead. jet)
- RC randomized \( \gamma \psi \)
- Poissonian limit
- Poissonian limit + \( v_2 \) (\( \langle \gamma \psi \text{pp} \rangle = 2 N^2 \langle v_2 \rangle \))
- Poissonian limit + \( v_2, v_3 \) (\( \langle \gamma \psi \text{pp} \rangle = 2 N^2 (v_2 + v_3) \))

\[ N_{\text{input}} \]

\[ N_{\text{raw}} \]

\[ P_{-} \text{ width (GeV/c)} \]

- Background Generator, no \( v_n \)
- Background Generator, with \( v_n \)
- Equation 3
- Equation 4

TennGen

Small deviations

Christine Nattrass (UTK), INT, 28 July 2021
Width vs multiplicity

\[ \sigma(\delta p_T) \text{ (GeV/c)} \]

- random cones
- RC (w/o lead. jet)
- RC randomized \( \eta \phi \)
- Poissonian limit
- Poissonian limit + \( v_2 \) (\( \sigma_{N^2}^{\text{pp}} = 2N^2_2 v_2^2 \))
- Poissonian limit + \( v_2 \),\( v_3 \) (\( \sigma_{N^2}^{\text{pp}} = 2N^2_3 (v_2^3 + v_3^3) \))

ALICE

\[ \text{Pb-Pb } \sqrt{s_{\text{NN}}} = 2.76 \text{ TeV} \]
\[ R = 0.4, p_T^{\text{min}} = 0.15 \text{ GeV/c} \]

Data / Prediction

Christine Nattrass (UTK), INT, 28 July 2021
Shape of width of the distribution

Single particle spectra

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

\[ \frac{dN}{dy} \propto f_{\Gamma}(p_T, 2, b) = b^2 p_T e^{-k_pT} \]

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

\[ \Sigma p_T \text{ of } N \text{ particles} \rightarrow N\text{-fold convolution:} \]

\[ f_N(p_T, p, b) = f_{\Gamma}(p_T, Np, b) \]

\[ \frac{dp_{T\text{total}}}{dy} \propto f_N(p_T, Np, b) \]

\[ N = \frac{N_{\text{total}}}{A_{\text{total}}} \pi R^2 \]

\[ \mu_{\text{total}} = \frac{Np}{b} = N \mu_{p_T}, \sigma_{\text{total}} = \sqrt{Np} = \sqrt{N} \sigma_{p_T} \]

Add Poissonian fluctuations in N:

\[ \sigma_{\text{total}} = \sqrt{N \sigma_{p_T}^2 + N \mu_{p_T}^2} \]

Add non-Poissonian fluctuations in N due to flow:

\[ \sigma_{\text{total}} = \sqrt{N \sigma_{p_T}^2 + (N - \sum v_n^2) \mu_{p_T}^2} \]

Assumes shape

Assumes uncorrelated number fluctuations

Tannenbaum, PLB(498), 1–2, Pg. 29-34 (2001)
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
Signal and background overlap

- Real jets
- Combinatorial jets
Signal vs Background:
The standard paradigm

Background

Signal
Signal vs Background:
The standard paradigm

Background

Combinatorial jets

Signal
Signal vs Background:
The standard paradigm

Background

Combinatorial jets = “fake” jets

Signal
Signal vs Background:
The standard paradigm

Background

Combinatorial jets

Signal

*Some gray areas
Technique

• Anti-\(k_T\) jet finder, \(|\eta_{\text{jet}}|<0.5\)
• **Combinatorial jets**: Only contain TennGen particles
• **Real jets**: Add a PYTHIA pp event. Real jets contain >80% of \(p_{\text{Thard}}^{\text{min}}\)
Real jets

Area vs. Jet $p_T$

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

Mean $p_T$ vs. Jet $p_T$

Leading Hadron $p_T$ vs. Jet $p_T$

Log z scale

Christine Nattrass (UTK), INT, 28 July 2021
$p_{\text{Thard}} > 40 \text{ GeV/c, } R=0.2$

Log z scale
$p_{\text{Thard}} > 40 \text{ GeV}/c, \ R=0.6$

**Log z scale**
Silhouette Values

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

\[ d_{i,j} = \sqrt{\left( \frac{A_i - A_j}{A_{\text{max}} - A_{\text{min}}} \right)^2 + \left( \frac{\alpha_i - \alpha_j}{\alpha_{\text{max}} - \alpha_{\text{min}}} \right)^2 + \left( \frac{\langle p_T \rangle_i - \langle p_T \rangle_j}{\langle p_T \rangle_{\text{max}} - \langle p_T \rangle_{\text{min}}} \right)^2 + \left( \frac{p_{T,i} - p_{T,j}}{p_{T,\text{max}} - p_{T,\text{min}}} \right)^2} \]

Area

\[ \alpha = \frac{1}{p_{T_{\text{jet}}}} \sum z_k(\vec{R}_k) \]

Angularity

Leading \( p_T \)

Average \( p_T \)
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 \( s_i = \frac{b_i - a_i}{\max[b_i, a_i]} \)

-1

Looks more like another cluster

0

Indistinguishable from other clusters

1

Looks more like its own cluster
Silhouette values

Example from Wikipedia

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

R = 0.2, p_T hard min = 40

s<0: look more like background

Real jets look more real if PYTHIA p_T is higher

s~0: look similar to signal
Silhouette values – decreasing $p_T$

$R = 0.2$, $p_T$ hard min = 40

Real Jets vs. Jet $p_T$

Real Jets vs. pythia $p_T$ in jet

Combinatorial Jets vs. Jet $p_T$
Silhouette values – decreasing $p_T$

$R = 0.2$, $p_T$ hard min = 30

Real Jets vs. Jet pT

Real Jets vs. pythia pT in jet

Combinatorial Jets vs. Jet pT
Silhouette values – decreasing $p_T$

$R = 0.2$, $p_T$ hard min = 20

Real Jets vs. Jet $p_T$

Real Jets vs. pythia $p_T$ in jet

Combinatorial Jets vs. Jet $p_T$
Silhouette values – decreasing $p_T$

Real jets look more like combinatorial jets

Combinatorial jets look more like real jets

These aren’t random jets!

$R = 0.2$, $p_T$ hard min = 10
Silhouette values – increasing R

R = 0.2, p_T hard min = 30

Real Jets vs. Jet pT

Real Jets vs. pythia pT in jet

Combinatorial Jets vs. Jet pT
Silhouette values – increasing R

R = 0.3, p_T hard min = 30
Silhouette values – increasing R

R = 0.4, p_T hard min = 30

Real Jets vs. Jet pT

Real Jets vs. pythia pT in jet

Combinatorial Jets vs. Jet pT
Silhouette values – increasing $R$

$R = 0.5$, $p_T$ hard min = 30
Silhouette values – increasing R

Real jets look more like real jets

Tail in distribution of real jets gets smaller

Combinatorial jets look more like real jets

These aren’t random jets!

R = 0.6, p_T hard min = 30

Log z scale
Mini-summary

• “Signal” and “background” have different properties, but...
• Always overlap somewhat
• Any procedure to remove “background” will also cut signal
How to compare to models
Iterative procedure

- Used by ATLAS & CMS
- ATLAS
  - **Calorimeter jets**: Reconstruct jets with $R=0.2$. $v_2$ modulated $<\text{Bkgd}>$ estimated by energy in calorimeters excluding jets with at least one tower with $E_{\text{tower}} > <E_{\text{tower}}>$.  
  - **Track jets**: Use tracks with $p_T > 4$ GeV/c

- Calorimeter jets from above with $E > 25$ GeV and track jets with $p_T > 10$ GeV/c used to estimate background again.
- Calorimeter tracks matching one track with $p_T > 7$ GeV/c or containing a high energy cluster $E > 7$ GeV are used for analysis down to $E_{\text{jet}} = 20$ GeV

Constituent biases don't matter that much up here!

But they do matter down here!
Survivor bias

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

- **Experimental background subtraction methods**: complex, make assumptions, apply biases
- **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
**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?
Monte Carlo Model

HepMC

HEPData

Rivet

Comparison to data
Why use Rivet?

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

Tracks → Jet finding algorithm → Jet candidates → Background subtraction

Jet spectrum smeared by energy resolution, background fluctuations

**Unfolding** – corrects for single track reco $\epsilon$, $E$ resolution, background fluctuations

Corrected spectra
Jets in ALICE: Response Matrix Construction

\[ \text{RM}_{\text{det}} \times \text{RM}_{\text{bkg}} = \text{RM}_{\text{tot}} \]

\( \text{RM}_{\text{det}} \) and \( \text{RM}_{\text{bkg}} \) are approximately factorizable

- (a) \( \text{RM}_{\text{det}} \) Detector response matrix
- (b) \( \text{RM}_{\text{bkg}} \) Background fluctuation matrix
- (c) \( \text{RM}_{\text{tot}} = \text{RM}_{\text{bkg}} \times \text{RM}_{\text{det}} \)

Pb-Pb \( s_{_{NN}} = 2.76 \) TeV
0-10\% Centrality

**ALICE PERFORMANCE**
19/06/2013

**Christine Nattrass (UTK), INT, 28 July 2021**
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
Closure

- **Methods**
  - Use $\delta p_T$ method to measure width of fluctuations with varying numbers of leading jets (LJ) discarded
  - Embed PYTHIA pp event into PYTHIA heavy ion event
  - The PYTHIA pp event is “true”

- Only embedding leads to full closure
Comparison to data

Pb-Pb at $\sqrt{s_{NN}} = 2.76$ TeV

Unfold to correct for fluctuations and other effects

Christine Nattrass (UTK), INT, 28 July 2021
Conclusions

- “Background” is not just an experimental problem!
- “Signal” and “background” jets overlap → impossible to suppress background without biasing jets
- Gold standard is to use Rivet
  - But it requires treating the model exactly like data
  - A number of issues specific to jets need to be discussed in the field
  
  Recorded tutorials from Rivetizing Heavy Ion Collisions at RHIC
Few heavy ion analyses in Rivet


Theorists don’t use Rivet
Undergraduates!*

*And one beginning graduate student with no programming experience

Left to right: Ricardo Santos (Berea), James Neuhaus, Jerrica Wilson, Mariah McCreary, Christine Nattrass, Austin Schmier (UTK)
Course-based undergraduate research experience

Ask me if you want more info!

Early Engagement in Course-Based Research Increases Graduation Rates and Completion of Science, Engineering, and Mathematics Degrees

Stacia E. Rodenbusch, Paul R. Hernandez, Sarah L. Simmons, and Erin L. Dolan
Jennifer Knight, Monitoring Editor:
Published Online: 13 Oct 2017 [https://doi.org/10.1187/cbe.16-03-0117

Abstract

National efforts to transform undergraduate biology education call for research experiences to be an integral component learning for all students. Course-based undergraduate research experiences, or CURES, have been championed for engaging students in research at a scale that is not possible through apprenticeships in faculty research laboratories. Yet there are if any studies that examine the long-term effects of participating in CURES on desired student outcomes, such as graduat from college and completing a science, technology, engineering, and mathematics (STEM) major. One CURE program, the Freshman Research Initiative (FRI), has engaged thousands of first-year undergraduates over the past decade. Using propensity score-matching to control for student-level differences, we tested the effect of participating in FRI on students' probability of graduating with a STEM degree, probability of graduating within 6 yr, and grade point average (GPA) at graduation. Students who completed all three semesters of FRI were significantly more likely than their non-FRI peers to earn a STEM degree and graduate within 6 yr. FRI had no significant effect on students' GPAs at graduation. The effects were similar for diverse students. These results provide the most robust and best-controlled evidence to date to support calls for early involvement of undergraduates in research.

Phys 494 – Course-based Undergraduate Research Experience in Relativistic Heavy Ion Physics

Instructor:
Dr. Christine Nattrass
Office: SERF 609
Phone: 974-6211
Email: christine.nattrass@utk.edu
Office hours: TBA

Teaching assistant: N/A

Class time & Location: TR 12:40-1:55 SERF 210

Course Description:
This course will incorporate undergraduates into a research project in high energy nuclear physics in a course setting. Each student will be responsible for implementing a heavy ion analysis in the program RIVET so that it can be used by the JETSCAPE collaboration to make comparisons between Monte Carlo models and data. Each student’s project will be incorporated into a public software repository so that it is available to the field and, if possible, it will be validated by the relevant experiment and incorporated into the official RIVET software.

3 semesters
15 students
8 women
3 minorities
3 non-traditional

All Rivet students
22 students
11 women
7 minorities
4 non-traditional

Christine Nattrass (UTK), INT, 28 July 2021
Learn Rivet yourself!
Or send your students & postdocs!

https://indico.bnl.gov/event/8843/
https://indico.bnl.gov/event/8840
Backup: jet properties
$p_{\text{Thard}}>40 \text{ GeV/c, } R=0.2$
$p_{\text{Thard}} > 30 \text{ GeV}/c, R=0.2$

Log $z$ scale
$p_{\text{Thard}} > 20 \text{ GeV}/c, \ R=0.2$

Log z scale

Christine Nattrass (UTK), INT, 28 July 2021
$p_{\text{Thard}} > 10 \text{ GeV/c, } R=0.2$

Log z scale

Christine Nattrass (UTK), INT, 28 July 2021
p_{\text{Thard}} >40 \text{ GeV}/c, R=0.2

Log z scale

Christine Nattrass (UTK), INT, 28 July 2021
$p_{\text{Thard}} > 40 \text{ GeV/c}, R=0.3$

Log $z$ scale

Christine Nattrass (UTK), INT, 28 July 2021
$p_{\text{Thard}} > 40 \text{ GeV/c}, R=0.4$

Log $z$ scale
$p_{\text{Thard}} > 40 \text{ GeV}/c, R=0.5$

Log z scale
$p_{\text{Thard}} > 40 \text{ GeV}/c$, $R=0.6$

Log z scale
Backup: silhouette scores
Silhouette values – decreasing $p_T$

Real jets look more like combinatorial jets

Combinatorial jets look more like real jets

$R = 0.2$, $p_T$ hard min = 10

Log $z$ scale
Silhouette values – increasing R

R = 0.6, p_T hard min = 30

Real jets look more like real jets

Tail in distribution of real jets gets smaller

Combinatorial jets look more like real jets
Backup: jet definition
Jets in principle

- Jet measures partons
- Hadronic degrees of freedom are integrated out
- Algorithms are infrared and collinear safe

- OK
- BAD: 2 jets are merged in one
**k_\text{T} jet finding algorithm**

- **Particles, clusters**

  **k_\text{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:
    \[ d_{ij} = \min\left(\frac{\Delta R_{ij}^2}{R^2}, p_{T,i}^2, p_{T,j}^2\right) \]
    \[ d_{iB} = p_{T,i}^2 \]

  - Combine smallest \( d_{ij} \):
    - If \( d_{iB} \) smallest, \( d_{iB} \rightarrow \text{jet} \)
    - Repeat until no particles left

- **Jet candidates**

---

Christine Nattrass (UTK), INT, 28 July 2021
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:
  \[ d_{ij} = \min \left( \frac{\Delta R_{ij}^2}{R_{ij}^2}, \frac{\Delta R_{ij}^2}{R_{ij}^2} \right) \]
  \[ d_{ij} = p_{T,i}^2, \Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \]
- Combine smallest $d_{ij}$
  If $d_{iB}$ smallest, $d_{iB} \rightarrow$ jet
  Repeat until no particles left

Jet candidates
Cambridge/Aachen jet finding algorithm

**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:
  \[ d_{ij} = \Delta R_{ij}^2 \]

- Combine smallest \( d_{ij} \):
  If \( d_{ib} \) smallest, \( d_{ib} \rightarrow \) jet

Repeat until no particles left

Particles, clusters

Jet candidates
Backup: misc
Unfolding

\[ \hat{\nu} = R\hat{\mu} + \hat{\beta} \]

- $\hat{\mu}$: the “true” histogram
- $\hat{\nu}$: the actual data we measure
- $\hat{\beta}$: background
- $R$: the response matrix

$\nu_i = \sum_{j=1}^{M} (R_{ij}\mu_j) + \beta_i$

May correct for “missing” jets!
Mixed events

- Gets background up to a normalization factor
- Good agreement with the data… but 20% discrepancies still within uncertainties
- In measurement with background suppressed (h-jet correlations)
- Did not see such agreement at the LHC for jet spectra
Mini-summary

- Experimental techniques can bias measurement in subtle ways
  - Background subtraction
  - Kinematic cuts
  - Choice of jet finder, R
  - Centrality determination
  - Technique for finding reaction plane
- Larger influence at low momentum
- Safest to do the same analysis on data and model
  - But unfolding is necessary in a full Monte Carlo model!
Experimental techniques for background
Focus on smaller angles

- **Pros**
  - Background is smaller
  - Background fluctuations smaller

- **Cons:**
  - Modifications expected at higher R
  - Biases sample towards quarks

Aside: “quark” and “gluon” jet only defined at leading order.
Focus on high $p_T$

- **Pros:**
  - Reduces combinatorial background

- **Cons:**
  - Cuts signal where we expect modifications
  - Could bias towards partons which have not interacted
  - Biases sample towards quark jets

"Quark" and "gluon" jets only defined at leading order!

PoS High-pTphysics09:023,2009
Area-based subtraction

- ALICE/STAR
- Require leading track $p_T > 5$ GeV/c
  - Suppresses combinatorial “jets”
  - Biases fragmentation
- No threshold on constituents
- Limited to small R – unstable unfolding

Combinatorial jets
Jet $R_{AA}$

LHC Run1 Data; PbPb (0-10%) $\sqrt{s_{NN}} = 2.76$ TeV

CMS 1609.05383

ALICE PLB 746(2015) 1-14

ATLAS PRL 114(2015) no.7

arXiv:1705.01974
Tension between ATLAS & ALICE/CMS
Mini-summary

• Most studies do one or more of the following:
  - Explicitly apply a (non-perturbative) bias
  - Implicitly apply a (non-perturbative) bias
  - Focus on small R
  - Focus on high pT
• May also → survivor bias
• Background subtraction should be part of definition of algorithm