A skeptic's guide to jets Part 1: Jet spectra



Christine Nattrass University of Tennessee, Knoxville

Acknowledgements

The following people contributed ideas and/or slides, but of course I take full responsibility for anything you don't like: Rosi Reed, Megan Connors, Sevil Salur Abhijit Majumder, Raghav Kunnawalkam Elayavalli Marta Verweij, Laura Havener Austin Schmier, Charles Hughes, Will Witt

Questions an experimentalist should ask

- What do I want to learn?
- What am I measuring?
- What assumptions am I making?
- What are the dominant uncertainties?
- How do I compare to models?

The answers for jets are highly non-trivial!

What do I want to learn? The cartoon picture



Probing the Quark Gluon Plasma

Medium Probe Detector Model Medium

Want a probe which traveled through the collision QGP is very short-lived (~1-10 fm/c) → cannot use an external probe

Want a probe which traveled through the medium QGP is short lived \rightarrow need a probe created in the collision



Want a probe which traveled through the medium QGP is short lived \rightarrow need a probe created in the collision We expect the medium to be dense \rightarrow absorb/modify probe















"Simple" example: Single hadrons

Nuclear modification factor

- Measure spectra of probe (jets) and compare to those in p+p collisions or peripheral A+A collisions
- If high- p_T probes (jets) are suppressed, this is evidence of jet quenching



Nuclear modification factor рРb • h^{\pm} , p-Pb $s_{NN} = 5.02 \text{ TeV}$, NSD (ALICE) h[±]. Pb-Pb (ALICE) Ľ 1.8 h[±] Pb-Pb (CMS) ★ γ , Pb-Pb s_{NN} = 2.76 TeV, 0-10% (CMS) ⊕ W[±], Pb-Pb s_{NN} = 2.76 TeV, 0-10% (CMS) R_{PbPb} s_{NN} = 2.76 TeV, 0-5% 1.6 ▼ Z⁰, Pb-Pb S_{NN} = 2.76 TeV, 0-10% (CMS) 1.4 1.2 Y Control Control 0.8 0.6 0.4 Probe 0.2 20 30 50 90 100 p_{τ} (GeV/c) or mass (GeV/c²) ALI-DER-95222

- Charged hadrons (colored probes) suppressed in Pb—Pb
- Charged hadrons not suppressed in p—Pb at midrapidity
- Electroweak probes not suppressed in Pb—Pb

Nuclear modification factor RAA



Electromagnetic probes – consistent with no modification – medium is transparent to them

Strong probes – significant suppression – medium is opaque to them - even heavy quarks!

What am I measuring? Definition of a jet

Theoretical calculations

Factorization theorem

- Assumption: Parton distribution functions, perturbative cross section, fragmentation function factorize
- What people really mean by "perturbatively calculable"
 - D and f are explicitly nonperturbative!
 - D is for parton c → hadron h Not what is experimentally measured
- Most theories for jet quenching modify fragmentation function D



Jet finders

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

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

p+p dijet











in pp collisions

- Jet finder: groups final state particles into jet candidates
 - Anti-k_T algorithm
 JHEP 0804 (2008) 063 [arXiv:0802.118
 9]
- 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.



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







Cambridge/Aachen jet finding algorithm



A jet is what a jet finder finds.

Jet cross-section in pp $\sqrt{s} = 2.76$ TeV, R = 0.2 Inclusive

arXiv:1301.3475

PLB: 10.1016/j.physletb.2013.04.026



•Green and magenta bands: NLO on Parton level

•Blue band: NLO + hadronization

•Hadronization calculations necessary to describe data

Jet ratios in pp $\sqrt{s} = 2.76$ TeV, R = 0.2, 0.4 Inclusive





Mini-summary



- Jets are not partons
- Good jet finders:
 - Infrared and colinear safe
 - k_{T} , anti- k_{T} , Cambridge/Aachen, SISCone
- Jet is defined by jet finder, its parameters
- PDFs, fragmentation functions non-perturbative
 → all jet measurements sensitive to somewhat non-perturbative effects
- Good agreement between theory and experiment

Jets in A+A collisions What assumptions am I making?

p+p vs A+A










Jet finding



in AA collisions

- Jet finder: groups final state particles into jet candidates
 - Anti-k_T algorithm
 JHEP 0804 (2008) 063 [arXiv:0802.1189]
- Combinatorial jet candidates
- Energy smearing from background
- Sensitive to methods to suppress combinatorial jets and correct energy
- Focus on narrow/high energy jets

Analysis steps



Understanding the background

TennGen background generator



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 \rightarrow fluctuating σ





Area-based background subtraction



Background density p



Christine Nattrass (UTK), uBNL 2022

Random cones



Random cones in ALICE

- Estimate ρ
 - $k_{_{T}} jet \ finder \rightarrow jet \ candidates$
 - $\rho = Median(p_T/A)$
- Draw Random cone

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



Random cones



Christine Nattrass (UTK), uBNL 2022

Shape of width of the distribution

Single particle spectra

$$f_{\Gamma}(p_{T}, p, b) = \frac{b}{\Gamma(p)}(b p_{T})^{p-1}e^{-bp_{T}}$$

$$\frac{dN}{dy} \propto f_{\Gamma}(p_{T}, 2, b) = b^{2} p_{T}e^{-bp_{T}}$$

$$\mu_{p_{T}} = \frac{p}{b}, \sigma_{p_{T}} = \frac{\sqrt{p}}{b}$$
Tannenbaum, PLB(498),1-2,Pg:29-34(201)
Add non-Poissonian fluctuations in N due to flow

$$\sigma_{total} = \sqrt{N \sigma_{p_{T}}^{2} + (N+2N^{2}\sum_{n} v_{n}^{2}) \mu_{p_{T}}^{2}}$$

Width vs multiplicity



Christine Nattrass (UTK), uBNL 2022

Width vs multiplicity



Christine Nattrass (UTK), uBNL 2022

Shape of width of the distribution



Combinatorial jets are mostly random!

$$\sigma_{total} = \sqrt{N} \sigma_{p_{T}}^{2} + (N + 2N^{2} \sum_{n} v_{n}^{2}) \mu_{p_{T}}^{2}} \approx 8.5 \, GeV/c$$
Fluctuations in p_T
 $\sigma_{p_{T}fluctuations} = \sqrt{N} \sigma_{p_{T}} \approx 2.5 \, GeV/c$
Fluctuations in N
 $\sigma_{N} = \sqrt{N} \mu_{p_{T}} \approx 4.4 \, GeV/c$
Note that this is for **random cones**!
Jet finders are peak finders.
Fluctuations in N

Mini-summary

- Jet finders put all input clusters, tracks in a jet candidate
- Background is *dominated* by random particles
- Models have background too!
 - And it doesn't agree with data!
 - Sensitive to multiplicity, shape of spectrum



Jets in A+A collisions: Dealing with background

Problems caused by background

- Combinatorial jets: unknown fraction of jets → estimate fraction, suppress
- Background in "real" jets: fluctuations in background distort reconstructed spectrum make unfolding unstable

Background subtraction methods

- Area-based (STAR, ALICE)
- Mixed events (STAR)
- Neural net (ALICE)
- Iterative (ATLAS, CMS)
- Reflection about η=0 (CMS)



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 (hjet correlations)
- Did not see such agreement at the LHC

ATLAS

Background subtraction method:

- Iterative procedure
 - Calorimeter jets: Reconstruct jets with R=0.2.
 v₂ modulated <Bkgd> estimated by energy in calorimeters excluding jets with at least one tower with Etower > <Etower>

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_{jet} = 20 GeV
 Phys. Lett. B 719 (2013) 220-241



Neural net Background Subtraction

- Deep Neural Network
 - N,100,100,50,1
- Trained on
- Trained with back propagation



Phys. Rev. C 99, 064904 (2019)

Alternative to area-based subtraction

- Estimate <p_T>
 - k_{τ} jet finder → jet candidates
 - $< p_T > = Median(p_T/N)$

$$p_T = p_T^{tot} - \langle p_T \rangle N$$



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!



Area-based subtraction

•ALICE/STAR

- •Require leading track $p_T > 5 \text{ GeV/c}$
 - Suppresses combinatorial "jets"
 - Biases fragmentation
- •No threshold on constituents
- •Limited to small R

Combinatorial "jets"



Survivor bias



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





What you see depends on what you're looking for

http://walkthewilderness.net/animals-of-india-72-asiatic-elephant/

Bias & background

- 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

http://walkthewilderness.net/animals-of-india-72-asiafic-elephant/

Iterative procedure

- Used by ATLAS & CMS
- ATLAS
 - Calorimeter jets: Reconstruct jets with R=0.2.
 v₂ modulated <Bkgd> estimated by energy in calorimeters excluding jets with at least one tower with

E_{tower} > <E_{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_{jet} = 20 GeV



Jet RAA



Jet RAA



Tension between ATLAS & ALICE/CMS
ATLAS

Background subtraction method:

- Iterative procedure
 - Calorimeter jets: Reconstruct jets with R=0.2.
 v₂ modulated <Bkgd> estimated by energy in calorimeters excluding jets with at least one tower with Etower > <Etower>

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_{jet} = 20 GeV
 Phys. Lett. B 719 (2013) 220-241



What you see depends on where you look



Mini-summary

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

What are the dominant uncertainties?

Analysis steps



Unfolding

• $\vec{\mu}$: the "true" histogram

$$\vec{v} = R\vec{\mu} + \vec{\beta}$$

- \vec{v} : the actual data we measure
- • $\vec{\beta}$: background
- R : the response matrix

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

Simple Solution (Inversion)

- Rearrange $\vec{v} = R\vec{\mu} + \vec{\beta}$ to get $\vec{\mu} = R^{-1}(\vec{v} \vec{\beta})$
- Problem: we don't have \vec{v} , we have \vec{n} , the measured data, which is subject to statistical fluxuations.
- We assume n_i is the maximum likelihood estimator for v_i , then solve for the estimator

$$\hat{\mu} = R^{-1}(\vec{n} - \vec{\beta}).$$

• R^{-1} is obtained from R through simple matrix inversion

Iterative Bayesian Method

• Using prior knowledge, start with an initial guess for the distribution of true histograms $P^{0}(\hat{\mu})$

• Use Bayes' Theorem to invert the response matrix
$$P(\hat{\mu}_i | v_j^{sig}) = \frac{P(v_j^{sig} | \hat{\mu}_i) P^0(\hat{\mu}_i)}{\sum_{l=1}^M P(v_j^{sig} | \hat{\mu}_l) P^0(\hat{\mu}_l)}$$

•
$$\hat{\mu}_i = \frac{1}{\epsilon_i} \sum_{j=1}^N v_j^{sig} P(\hat{\mu}_i | v_j^{sig})$$
 where ϵ_i is the detector efficiency

• Plug in the newly obtained $P(\hat{\mu}_i | v_j^{sig})$ and $\hat{\mu}_i$ as new priors, then repeat

•Terminate before the wildly oscillating true inverse is reached (usually ~4 iterations) to preserve some smoothness

RooUnfold-Bayes

- RooUnfoldTest.cxx
- method = Bayes
- Exponential training and testing



About unfolding...

- d'Agostini (author of Bayesian unfolding algorithm) says you should avoid it if you can
- Necessary when experimental resolution is poor
 - Ex: Single particle spectra $\frac{\sigma_p}{p} \ll w_{bin} \rightarrow$ unfolding unnecessary Ex: Jet spectra $\frac{\sigma_p}{p} \approx w_{bin} \rightarrow$ unfolding necessary
- Algorithm assumes response matrix is correct
 - Matching reconstructed and simulated jets is non-trivial!
- Corrects for multiple experimental effects simultaneously
 - Difficult to disentangle different effects
 - Leads to non-trivial uncertainty correlations between data points due to algorithm
 - May not handle systematic correlations between effects correctly







Response matrix includes assumptions about

- Detector response
 - Including particle composition of jets!
- Fragmentation and hadronization
 - How does hadronization influence the width of your jet?
- Background and/or background fluctuations
- How you match reconstructed ("detector level") and true ("particle level") jets

Jet Momentum Resolution



•Jet resolution

- Dominated by background fluctuations at low momentum
- Dominated by detector effects at high momentum

Mini-summary

- Jet energy resolution is fundamentally large
 - Measuring multiple correlated particles!
 - Be skeptical of jet measurements with <10% uncertainties
- Unfolding is complicated, often unstable, and hard
- Construction of response matrix includes several assumptions

Jets in A+A collisions: 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?



Christine Nattrass (UTK), uBNL 2022

Why use Rivet?

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

Rivetizing Heavy Ion Collisions at RHIC 2020

November 30, 2020 to December 4, 2020 Online US/Eastern timezone





Analysis steps: Full Monte Carlo





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
- Unclear how these influence the measurement
- Safest to do the same analysis on data and model
 - But unfolding is necessary in a full Monte Carlo model!

A skeptic's guide to jets Part 2: Where we are going



Christine Nattrass University of Tennessee, Knoxville





There is no partionic energy loss.



There is only partionic energy redistribution.



What is jet (sub)structure?

- A Whatever I am measuring!
- B Any new jet observable
- C Any observable which measures the structure of jets.
- D A cool buzzword
- E I don't know but it sounds cool and gets me talks/grants

Types of observables



Christine Nattrass (UTK), uBNL 2022

Type I: Energy loss



Christine Nattrass (UTK), uBNL 2022

Type II: Fragmentation


Type III: Distribution of properties

Jet mass



- Quenching models (JEWEL, Q-PYTHIA) show a larger mass than pp-like PYTHIA jets
- Pb-Pb measurement can discriminate among these predictions



Type IV: Declustering

Note: These slides are from Laura Havener

*A selection. Don't be offended if I skip your favorite.





Wider jets form earlier and narrower jets form later

Jet splittings: in medium

Vacuum splittings in/out of the medium

$$t_f^{\rm vac} = \frac{1}{\theta^2 \omega}$$



Medium-induced splittings from gluon radiation

$$t_g^{\text{med}} = \sqrt{\frac{\omega}{\hat{q}}}$$

Jet splittings: in medium

Coherence: subjets unresolved and jet loses energy as a whole

Decoherence: medium *resolves* the subjets resulting in a stronger e-loss

Medium-induced splittings

Vacuum splittings inside medium, resolved Vacuum splittings inside medium, unresolved

Vacuum splittings outside medium





Exploring the Lund Plane: in vacuum

- Lund Diagram*: phase space of jet splitting <u>*Z. Phys. C43 (1989)</u> JHEP 12 (2018)
- log(k_T) > 0 separates perturbative from non-perturbative regime
- Formation time: how long until the splitting occurred 1

Y. L. Dokshitzer, et.al.

Christine Nattrass (UTK), uBNL 2022

 $(1-z)k_{\rm T}\Delta R$



Soft drop grooming

 Reconstruct anti-k_T R=0.4 charged jets with jet-by-jet constituent background subtraction*
<u>*JHEP 06 (2014) 092</u>



Soft drop grooming

 Reconstruct anti-k_T R=0.4 charged jets with *jet-by-jet constituent* background subtraction*
*JHEP 06 (2014) 092

Remove from each constituent inside the jet instead of from the whole jet

Jet-by-jet:

$$p_{\rm T}^{\rm jet, corr} = p_{\rm T}^{\rm jet} - \rho A$$

Track-by-track (i) in jet: $p_{\rm T}^{\rm i, corr} = p_{\rm T}^{\rm i} - \rho A \label{eq:pt_track}$



Groomed variables



zg: jet splitting



Christine Nattrass (UTK), uBNL 2022

z_g: opening angle



Christine Nattrass (UTK), uBNL 2022

Background

Unfolding: jet splitting

Uncorrelated background leads to subjets being picked up as incorrect or "fake" splittings



Non-diagonal response prohibits unfolding

Mini-summary

- "Jet substructure" is used inconsistently
- Search for new observables
 - Haven't really used most of the "old" ones!
- So far it's a mixed bag
 - Many are insensitive
 - Some may have some promise
 - Background tricky





Christine Nattrass (UTK), uBNL 2022

Bayesian Statistical Analysis Models and Data Analysis Initiative



JETSCAPE Event generator

Jet Energy-loss Tomography with a Statistically and Computationally Advanced Program Envelope http://jetscape.wayne.edu/



Christine Nattrass (UTK), uBNL 2022

Event Generator + Bayesian Statistical analysis



Christine Nattrass (UTK), uBNL 2022



Conclusions

- Jets are complicated and hard to measure to high precision
- Much of the physics we want does not require them
- Extra insight from studying them anyways
- Be skeptical, especially of background subtraction
- Make sure the measurement is comparable to model





Exploring the Lund Plane: in medium



Exploring the Lund Plane: in medium

 Jet splittings in heavy-ion (HI) collisions

1: Vacuum splitting outside of medium

2: Vacuum splitting inmedium, resolved (decoherence)

3: Vacuum splittings inmedium, unresolved (coherence)

4: Medium-induced splittings



Jets in ALICE: Response matrix RM_{det}

- "Particle" level jets defined by jet finder on MC particles
- Pythia with Pb-Pb tracking efficiency
- "Detector" level jets defined by jet finder after event reconstruction through GEANT
- Particle level jets are geometrically matched to detector level jets
- Matrix has a dependence on spectral shape and fragmentation

It-finding efficiency is probability of a matched particle level jet



Christine Nattrass (UTK), uBNL 2022

Modified fragmentation



- Enhancement at low z
- No modification/enhancement at high z?

nsd: iterative declustering



Modification: enhancement at small n_{SD} and suppression at intermediate n_{SD}



Christine Nattrass (UTK), uBNL 2022



Au+Au di-jets more imbalanced than p+p for prcut>2 GeV/cChristing Nathras (UTK), wBNL 2022 Christing Nathras (UTK), wBNL 2022 Chri

Width vs multiplicity



Discrepancy not from an excess of jets!

Jet-hadron correlations



- Jets are broader, constituents are softer
- Also seen in:
 - Di-hadron correlations [Lots of papers]
 - Jet shapes [arXiv:1708.09429, arXiv:1512.07882, arXiv:1704.03046]
 - Dijet asymmetry with soft constituents [PRL119 (2017) 62301]