The Lund jet plane: organising QCD radiation at colliders

Gavin P. Salam* Rudolf Peierls Centre for Theoretical Physics & All Souls College, Oxford

based on arXiv:1807.04758 with F. Dreyer, G. Soyez (with some of their slides)

* on leave from CERN and CNRS



European Research Council Established by the European Commission



THE ROYAL SOCIETY

Birmingham 4/3/2019

































jets





CMS Experiment at the LHC, CERN Simulated event at 13 TeV centre-of-mass energy









CMS Experiment at the LHC, CERN Simulated event at 13 TeV centre-of-mass energy









jets are about organizing the information from hundreds (or thousands) of particles into a form that we as humans can understand and process









The Quantum-Chromodynamic (QCD) origin of jets



Start off with a qqbar system



. . . .

. . .



a gluon gets emitted at small angles





. . .





it radiates a further gluon

. . . .





and so forth



meanwhile the same happened on the other side



• • • • •



then a non-perturbative transition occurs







giving a pattern of hadrons that "remembers" the gluon branching (hadrons mostly produced at small angles wrt qqbar directions — two "jets")

The tools used by ATLAS & CMS









The tools used by ATLAS & CMS

Jets – 2nd most widely used single tool after Geant

(since 2017)

2019-02-13, excluding self-citations; all papers > 0.2 Powner Sinderoo (2010) Review or Briticle Physics CL(S) technique Puthia 8.2 MNDDE3 DDES CTIO PDFS WSTW2008 PDFS CTEQ6 PDFS PDF4LHC-AunII Sherba 7.7 Herwig++ MC Eurgen (2001) t_{00++}





a pp collision that produces a high p_t top-antitop pair, resulting in two "top-jets", each with subjets

Such events probe point-like nature of top quarks to TeV scale & allow you to search for new ttbar resonances











Parton energy loss leads to jet suppression

Q

from X-N. Wang



Energy deposition leads to medium excitation









jet substructure for pp



what should a jour administration administration and the second s



projection to jets should be resilient to QCD effects



hadron level

Jet substructure for boosted hadronc W/Z/H/t etc. decays

- with $p_t \gg m$, leading to collimated decays.





2 jets

[Figure by G. Soyez]

At LHC energies, EW-scale particles (W/Z/t...) are often produced

Hadronic decay products are thus often reconstructed into single jets.













Convolutional neural networks and jet images

- Project a jet onto a fixed $n \times n$ pixel image in rapidity-azimuth, where each pixel intensity corresponds to the momentum of particles in that cell.
- Can be used as input for classification methods used in computer vision, such as deep convolutional neural networks.





Recurrent neural network on clustering trees

- Techniques inspired from Natural Language Processing with powerful applications in handwriting and speech recognition.
- Train a recurrent neural network on successive declusterings of a jet.





using full event information: jet substructure for W tagging



QCD rejection with use of full jet substructure

5–10x better



jet substructure for HI collisions



fragmentation function

 $D(z) = \left\langle \sum_{i \in jet} \delta(z - p_{ti}/p_{t,jet}) \right\rangle_{jets}$







fragmentation function

differential jet shape

 $D(z) = \left\langle \sum_{i \in jet} \delta(z - z_i) \right\rangle$

 $\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ wit} \\ \Delta R_{kJ} \in [r,]}} \sum_{k \text{ wit} \in [r,]} \sum_{k \in [r,]$

$$-p_{ti}/p_{t,jet})\bigg\rangle_{jets}$$

$$p_{\perp}^{(k)},$$

$$p_{\perp}^{(k)},$$

$$p_{\perp}^{(k)},$$

$$(i \in \mathcal{F})$$





fragmentation function

differential jet shape

 $D(z) = \left\langle \sum_{i \in jet} \delta(z - p_{ti}/p_{t,jet}) \right\rangle_{jet}$

 $\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r+\delta r]}} p_{\perp}^{(k)},$

$$g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta R_{kJ} \, ,$$













fragmentation function

differential jet shape

 $girth \equiv broadening$

jet mass, groomed & ungroomed

 $D(z) = \left\langle \left| \sum_{i \in jet} \delta(z - p_{ti}/p_{t,jet}) \right\rangle_{jet} \right\rangle_{jet}$



 $g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta$

 $m^2 = \left(\sum_{i \in (\text{sub})\text{jet}} p_i^{\mu}\right)$

$$p_{\perp}^{(k)}$$

h
,r+ δr]

$$\Delta R_{kJ}$$











fragmentation function

differential jet shape

 $girth \equiv broadening$

jet mass, groomed & ungroomed

 $z_g, \Delta R_{12}$

 $D(z) = \left\langle \left| \sum_{i \in jet} \delta(z - p_{ti}/p_{t,jet}) \right\rangle_{jet} \right\rangle_{jet}$

 $\rho(r) = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{\substack{k \text{ with} \\ \Delta R_{kJ} \in [r, r+\delta r]}} p_{\perp}^{(k)},$

 $g = \frac{1}{p_{\perp}^{\text{jet}}} \sum_{k \in J} p_{\perp}^{(k)} \Delta$

 $m^2 = \left(\sum_{i \in (\text{sub})\text{jet}} p_i^{\mu}\right)$

 $\min(p_{\perp,1}, p_{\perp})$ $z_g = - p_{\perp,1} + p_{\perp,2}$



$$\Delta R_{kJ}$$

$$\frac{2}{2} > z_{\rm cut} \left(\frac{\Delta R_{1,2}}{R_J}\right)^{\beta}$$
















JEWEL+PYTHIA Pb+Pb (0 – 10 %) $\sqrt{s_{NN}} = 5.02 \text{ TeV}$



JEWEL v. data

<u>arXiv:1707.01539</u>, by Milhano,
 Wiedemann and Zapp with
 medium response



recurrent theme in heavy-ion calculations: 2d phasespace plots





recurrent theme in heavy-ion calculations: 2d phasespace plots





the "Lund plane"

can we construct observables that are

(a) more transparent in terms of the physical info they extract? (b) close to optimal for multivariate techniques & machine-learning?





the Cambridge / Aachen (C/A) jet algorithm

- 1. Identify pair of particles, i & j, with smallest ΔR_{ij}
- 2. If $\Delta R_{ij} < R$ (jet radius parameter)
 - A. recombine i & j into a single particle
 - B. loop back to step 1
- 3. Otherwise, stop the clustering

Dokshitzer, Leder, Moretti & Webber '97 Wobisch & Wengler '98









A sequence of jet substructure tools taggers

- ► 1993: k_t declustering for boosted W's: [Seymour]
- 2002: Y-Splitter (k_t declustering with a cut) [Butterworth.
 Cox, Forshaw]
- 2008: Mass-Drop Tagger (C/A declustering with a k_t/m cut) [Butterworth, Davison, Rubin, GPS]
- ➤ 2013: Soft Drop, β=0 [Dasgupta, Fregoso, Marzani, GPS]
- ▶ 2014: Soft Drop, $\beta \neq 0$ [Larkoski, Marzani, Soyez, Thaler]
 - Undo last clustering of C/A jet into subjets 1, 2
 Stop if z = min(p_{t1}, p_{t2})/(p_{t1} + p_{t2}) (ΔR₁₂/R)^β > z_{cut}
 Else discard softer branch, repeat step 1 with harder branch

[Seymour] a cut) [Butterworth.

oso, Marzani, GPS] ani, Soyez, Thaler]



A sequence of jet substructure tools taggers

- ► 1993: k_t declustering for boosted W's: [Seymour]
- 2002: Y-Splitter (k_t declustering with a cut) [Butterworth.
 Cox, Forshaw]
- 2008: Mass-Drop Tagger (C/A declustering with a k_t/m cut) [Butterworth, Davison, Rubin, GPS]
- ► 2013: Soft Drop, $\beta=0$ [Dasgupta, Fregoso, Marzani, GPS]
- ▶ 2014: Soft Drop, $\beta \neq 0$ [Larkoski, Marzani, Soyez, Thaler]
- > 2017: Iterated Soft Drop [Frye, Larkoski, Thaler, Zhou] count number of iterations until you reach 1 particle
- > 2018/19: ?

[Seymour] a cut) [Butterworth.



Phase space: two key variables (+ azimuth)



 $k_{t} = p_{t} \Delta$



ΔR (or just Δ) opening angle of a splitting

p_t (or p_{\perp}) is transverse momentum wrt beam

 k_t is ~ transverse momentum wrt jet axis



jet with R = 0.4*,* $p_t = 200 \text{ GeV}$



jet with R = 0.4, $p_t = 200 \text{ GeV}$



0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



ΔR_{ij} $k_t = \min(p_{ti}, p_{tj}) \Delta R_{ij}$

0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



ΔR_{ii} $k_t = \min(p_{ti}, p_{tj}) \Delta R_{ij}$

0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



ΔR_{ij} $k_t = \min(p_{ti}, p_{tj}) \Delta R_{ij}$

0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$



ΔR_{ij} $k_t = \min(p_{ti}, p_{tj}) \Delta R_{ij}$

0.01

Introduced for understanding Parton Shower Monte Carlos by B. Andersson, G. Gustafson L. Lonnblad and Pettersson 1989





jet with R = 0.4, $p_t = 200 \text{ GeV}$











decluster a C/A jet: at each step record ΔR,kt as a point in the Lund plane repeatedly follow harder branch

5th heavy-ion workshop @ CERN, 1808.03689 Dreyer, Soyez & GPS, <u>1807.04758</u> (for pp applications)

constructing the Lund plane





decluster a C/A jet: at each step record ∆R,kt as a point in the Lund plane repeatedly follow harder branch

5th heavy-ion workshop @ CERN, <u>1808.03689</u> Dreyer, Soyez & GPS, <u>1807.04758</u> (for pp applications)

constructing the Lund plane







PRIMARY LUND PLANE

LUND DIAGRAM

JET

 $\ln 1/\Delta$





jet with R = 0.4, $p_t = 200 \text{ GeV}$



jet with R = 0.4, $p_t = 200 \text{ GeV}$



jet with R = 0.4, $p_t = 200 \text{ GeV}$



jet with R = 0.4, $p_t = 200 \text{ GeV}$



application to pp QCD studies



average pp Lund density: parton level











average pp Lund density: hadron level (no underlying event / MPI)











average pp Lund density: hadron level (with underlying event / MPI)











average pp Lund density: cross sections





average pp Lund density: cross sections

log(kt [GeV])





analytic perturbative QCD control

To leading order in perturbative QCD and for $\Delta \ll 1,$ one expects for a quark initiated jet

$$\rho \simeq \frac{\alpha_s(k_t)C_F}{\pi} \bar{z} \left(p_{gq}(\bar{z}) + p_{gq}(1-\bar{z}) \right), \quad \bar{z} = \frac{k_t}{p_{t,jet}\Delta}$$



Lund plane can be calculated analytically.

Calculation is systematically improvable.


application to HI collisions



jet with R = 0.4, $p_t = 200 \text{ GeV}$







Figure 4: Lund diagram reconstructed from jets generated by QPYTHIA (left column), JEWEL without recoils (middle column) and JEWEL with recoils (right column). The lower panels correspond to the difference of the radiation pattern with and without jet quenching effects. Note that the scale of the z-axes varies between the panels.



HI MC studies

Andrews et al, <u>1808.03689</u>

 clear potential for distinguishing
 between models,
 with clear physical
 picture of where the differences arise





application to high-pt physics

e.g. new-physics searches and Higgs studies



Comparing quark/gluon v. W-induced jets







Beyond average density: in the (primary) Lund plane





Lund declustering points as inputs to machine-learning

- Simple recurrent networks unable to handle dependencies that are widely separated in the data.
- LSTM networks designed to have memory over longer periods, by adding four layers for each module and including a no-activation function.



[Hochreiter, Schmidhuber (1997)]

 (\mathbf{X}_{t+1})

long-short-term memory networks (LSTMs) gave us the best performance

Figures from http://colah.github.io/posts/2015-08-Understanding-LSTMs/





Lund declustering points as inputs to hand-crafted likelihood calculation

- Identify emission that generates the jet mass (with Soft-Drop)
- Assume all other emissions are independent of each other, i.e. random distribution just set by average density
- Get MC ratio of average densities for W (Signal=S) v.
 QCD (background = B) jets
- Build likelihood discriminator

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_{\ell}(m^{(\ell)}, z^{(\ell)}) + \sum_{i \neq \ell} \mathcal{L}_{n\ell}(\Delta^{(i)}, k_t^{(i)}; \Delta^{(\ell)})$$
$$\mathcal{L}_{n\ell}(\Delta, k_t; \Delta^{(\ell)}) = \ln\left(\rho_S^{(n\ell)} / \rho_B^{(n\ell)}\right)$$
$$\rho_X^{(n\ell)}(\Delta, k_t; \Delta^{(\ell)}) = \frac{dn_{\text{emission}, X}^{(n\ell)}}{d\ln k_t \, d\ln 1 / \Delta \, d\ln \Delta^{(\ell)}}$$







signal efficiency

Performance: background rejection v. signal efficiency

Lund + machine-learning (LSTM) Lund + likelihood (gets to within 70-80% of performance of best machine learning)







resilience to non-perturbatitve effects

$$\zeta = \left(\frac{\Delta \epsilon_W^2}{\langle \epsilon \rangle_W^2} + \frac{\Delta \epsilon_{\rm QCD}^2}{\langle \epsilon \rangle_{\rm QCD}^2}\right)^{-\frac{1}{2}}$$

Performance: S/\sqrt{B} v. resilience to non-perturbative QCD

Lund + likelihood performs better than machine learning when you exclude nonperturbative region ($k_t < 1$ GeV(Lund + machine-learning (LSTM)





Conclusions

The QCD radiation in collider events (pp & HI) is a rich source of information, which we're only just starting to tap into.

The difficulty is that there's a lot of it: how do we condense it down to something we can understand, measure & exploit quantitatively?

The Lund plane "construction" offers an approach that

- maps transparently onto physically meaningful kinematic regions
- ► is amenable to calculations in QCD (work in progress)
- provides a powerful input to machine learning, but also can be used almost as effectively in simpler multivariate frameworks.









Initial—final symmetry





choice of C/A for declustering





why the C/A algorithm?







why the C/A algorithm?



Figure 6: The $\rho(\Delta, k_t)$ results as obtained with k_t (left) and anti- k_t (right) declustering, normalised to the result for C/A declustering.

If you use jet algorithms other than C/A to provide the initial (de)clustering sequence, the jet algorithm itself introduces strong "unphysical" structure





why the C/A algorithm?



Figure 5: Evaluations with Event2 of the second-order contribution to the Lund plane, in a bin of $\ln 1/\Delta$, as a function of κ , for (de)clustering sequences obtained with the k_t , anti- k_t and C/A jet algorithms. In (a) and (b) the dashed line corresponds to the analytic expectations, Eqs. (2.9) and (2.10) for clustering-induced double-logarithms in the k_t and anti- k_t algorithms. In (c), for the C/A algorithm, which is seen here to be free of double logarithms, the dot-dashed line corresponds to the (single-logarithmic) running coupling correction, Eq. (2.11), illustrating that it dominates the second-order correction.

mathematically, the unphysical structure is driven by double logarithms, $(\alpha L^2)^n$ in the Lund-plane density.

C/A only produces at most single logarithms, (αL)ⁿ



choice of original jet alg.



the declustering sequence from C/A v. anti– k_{t} starting points





consequence for Lund plane density







detector effects

.



Detector effects: with Delphes simulation (+ particle flow)

- Detector effects have significant impact on the Lund plane at angular scales below the hadronic calorimeter spacing.
- Two enhanced regions corresponding to resolution scale of HCal and ECal.







subjet-particle rescaling algorithm (SPRA)

Mitigate impact of detector granularity using a subjet particle rescaling algorithm:

- Recluster Delphes particle-flow objects into subjets using C/A with $R_h = 0.12.$
- Taking each subjet in turn, scale each PF charged-particle (h^{\pm}) and photon (γ) candidate that it contains by a factor f_1

$$f_1 = \frac{\sum_{i \in \text{subjet}} p}{\sum_{i \in \text{subjet}(h^{\pm}, \gamma)}}$$

and discard the other neutral hadron candidates.

If subjet doesn't contain photon or charged-particle candidates, retain all of the subjet's particles with their original momenta.

Recluster the full set of resulting particles (from all subjets) into a single large jet and use it to evaluate the mass and Lund plane.

$$p_{t,i}$$

not a new idea!

- [82] A. Katz, M. Son, and B. Tweedie, Jet Substructure and the Search for Neutral Spin-One Resonances in Electroweak Boson Channels, JHEP 03 (2011) 011, [arXiv:1010.5253].
- [83] M. Son, C. Spethmann, and B. Tweedie, Diboson-Jets and the Search for Resonant Zh Production, JHEP 08 (2012) 160, [arXiv:1204.0525].
- [84] S. Schaetzel and M. Spannowsky, Tagging highly boosted top quarks, Phys. Rev. D89 (2014), no. 1 014007, [arXiv:1308.0540].
- [85] A. J. Larkoski, F. Maltoni, and M. Selvaggi, Tracking down hyper-boosted top quarks, JHEP **06** (2015) 032, [arXiv:1503.03347].
- [86] S. Bressler, T. Flacke, Y. Kats, S. J. Lee, and G. Perez, Hadronic Calorimeter Shower Size: Challenges and Opportunities for Jet Substructure in the Superboosted Regime, Phys. Lett. **B756** (2016) 137–141, [arXiv:1506.02656].
- [87] Z. Han, M. Son, and B. Tweedie, Top-Tagging at the Energy Frontier, Phys. Rev. D97 (2018), no. 3 036023, [arXiv:1707.06741].
- [88] CMS Collaboration, C. Collaboration, V Tagging Observables and Correlations,
- [89] **ATLAS** Collaboration, T. A. collaboration, Jet mass reconstruction with the ATLAS Detector in early Run 2 data,





subjet-particle rescaling algorithm (SPRA)





Δ

Pythia v. Sherpa

.



average pp Lund density: parton level











average pp Lund density: hadron level (no underlying event / MPI)











average pp Lund density: hadron level (with underlying event / MPI)









