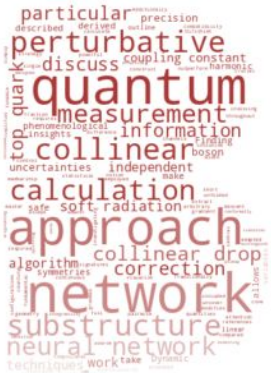
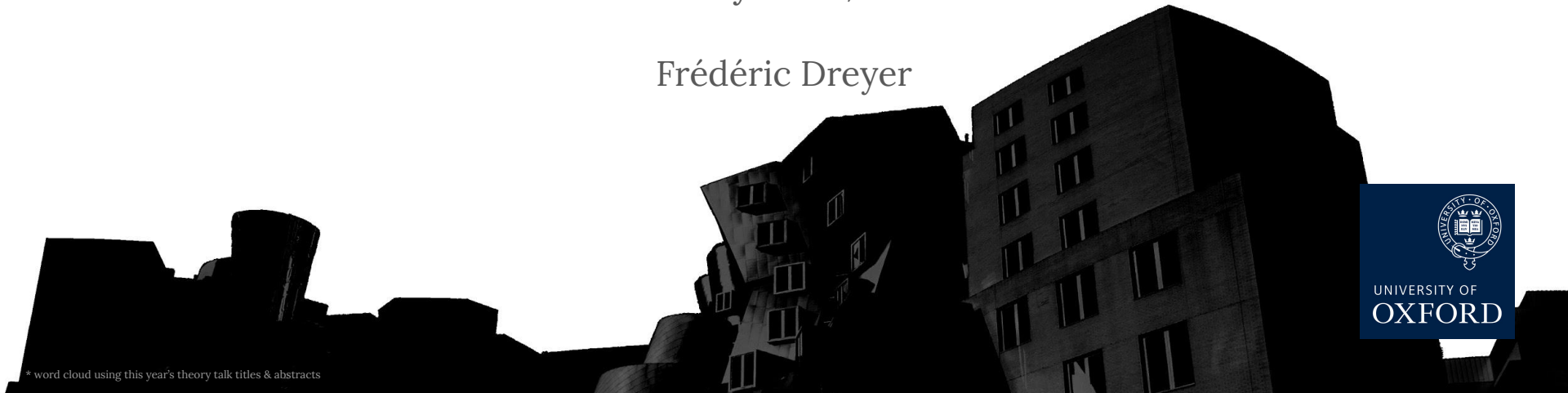


Theory Summary

26 July 2019, MIT

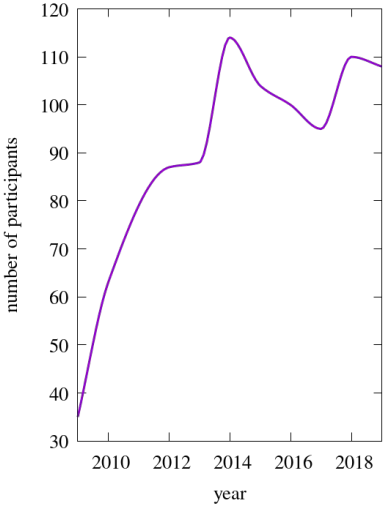
Frédéric Dreyer

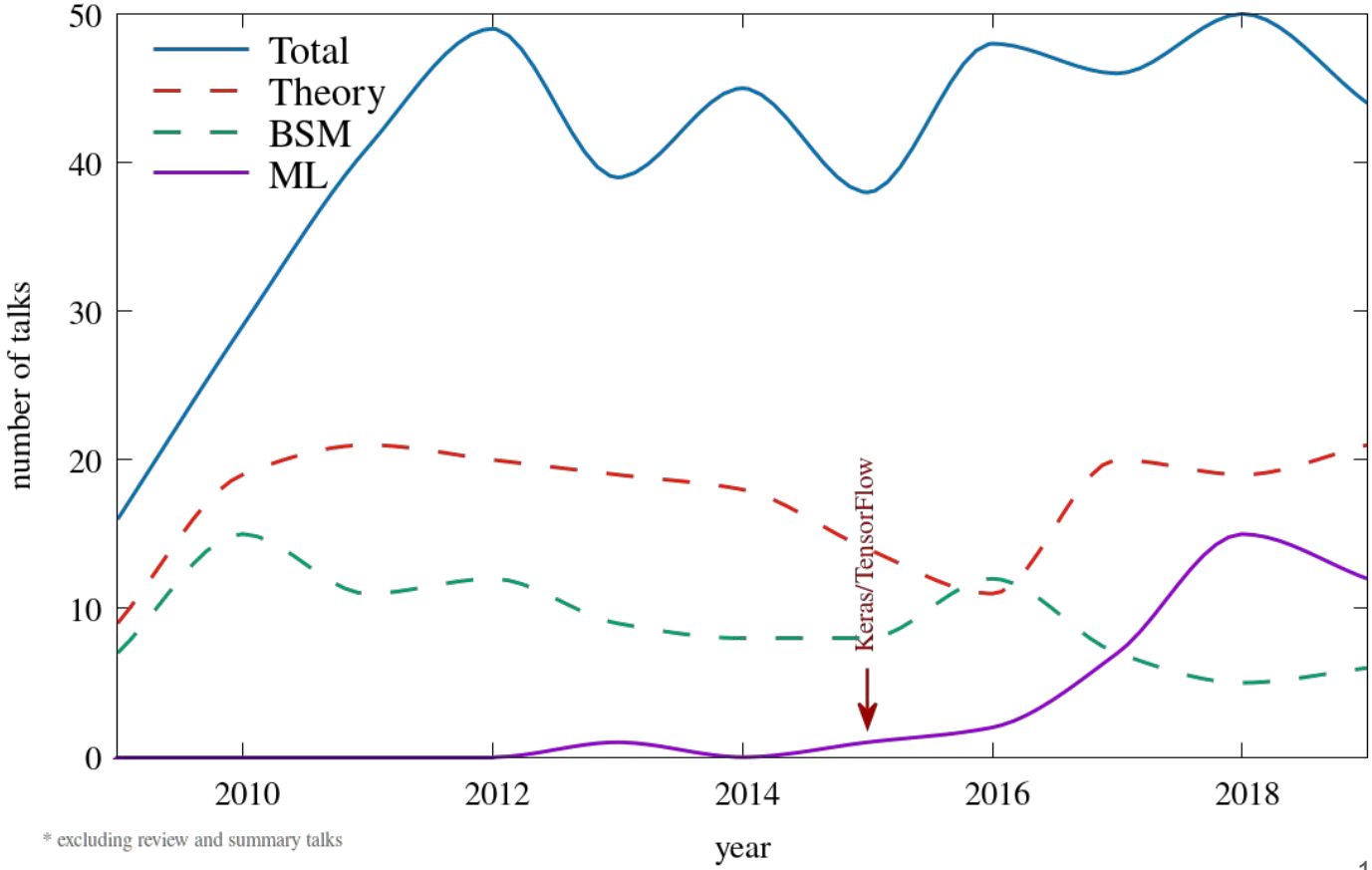


UNIVERSITY OF OXFORD

* word cloud using this year's theory talk titles & abstracts

BOOST over the years





* excluding review and summary talks

From ideas...

Modified Mass Drop Tagger

[Marzani '13 '14]

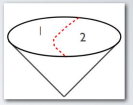
1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2 ($m_1 > m_2$)
2. If $m_1 < \mu m$ (mass drop) and the splitting was not too asymmetric ($y_{ij} > y_{cut}$), tag the jet.
3. Otherwise redefine j to be the subjet with highest transverse mass and iterate.

Soft Drop

Larkoski, SM, Soyez and Thaler (2014)

1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2.
2. If $\frac{\min(p_{T1}, p_{T2})}{p_{T1}+p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$

   then deem j to be the soft-dropped jet.



3. Otherwise redefine j to be the harder subjet and iterate.

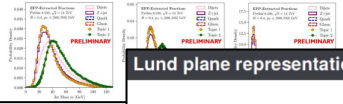
ets can be either kept (grooming mode) or discarded (tagging mode)

- An operational definition of quark and gluon jets defined directly in terms of hadronic cross sections:

$$p_{\text{quark}}(x) \equiv \frac{p_A(x) - \kappa_{AB}\, p_B(x)}{1-\kappa_{AB}} \qquad p_{\text{gluon}}(x) \equiv \frac{p_B(x) - \kappa_{BA}\, p_A(x)}{1-\kappa_{BA}}$$

[Metodiev '18]

- Allows quark and gluon jet distributions to be measured separately without fraction or template inputs:



Lund plane representation

To create a Lund plane representation of a jet, recluster a jet j with the Cambridge/Aachen algorithm then decluster the jet following the hardest branch.

1. Undo the last clustering step, defining two subjets j_1, j_2 ordered in p_t.
2. Save the kinematics of the current declustering

$$\Delta \equiv (y_1 - y_2)^2 + (\phi_1 - \phi_2)^2, \quad k_t \equiv p_{t2}\Delta,$$

$$m^2 \equiv (p_1 + p_2)^2, \quad z \equiv \frac{p_{t2}}{p_{t1}+p_{t2}}, \quad \psi \equiv \tan^{-1}\frac{y_2 - y_1}{\phi_2 - \phi_1}.$$

3. Define $j = j_1$ and iterate until j is a single particle.

[FD '18]

From ideas... ...to measurements

Modified Mass Drop Tagger

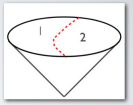
1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2 ($m_1 > m_2$)
2. If $m_1 < \mu m$ (mass drop) and the splitting was not too asymmetric ($y_{ij} > y_{cut}$), tag the jet.
3. Otherwise redefine j to be the subjet with highest transverse mass and iterate.

[Marzani '13 '14]

Soft Drop

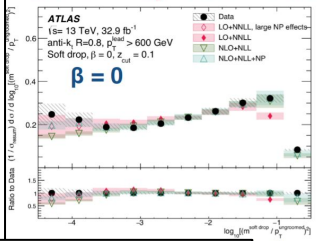
Larkoski, SM, Soyez and Thaler (2014)

1. Undo the last stage of the C/A clustering. Label the two subjets j_1 and j_2.
2. If $\frac{\min(p_{T1}, p_{T2})}{p_{T1}+p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$ then deem j to be the soft-dropped jet.
3. Otherwise redefine j to be the harder subjet and iterate.



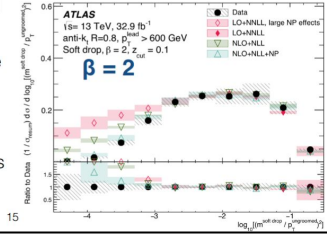
[Roloff '18]

Soft Drop Jet Mass



- Three main regimes in the mass distribution:
  - $\rho < -3$: Non-perturbative regime
  - $-3 < \rho < -1$: Resummation regime
  - $\rho > -1$: Fixed order regime
- See this paper on NLO+NLL and this paper on LO+NNLL for more information about the predictions

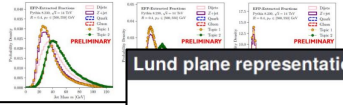
[Mozer '18]



- An operational definition of quark and gluon jets defined directly in terms of hadronic cross sections:

$$p_{\text{quark}}(x) \equiv \frac{p_A(x) - \kappa_{AB}\, p_B(x)}{1-\kappa_{AB}} \qquad p_{\text{gluon}}(x) \equiv \frac{p_B(x) - \kappa_{BA}\, p_A(x)}{1-\kappa_{BA}}$$

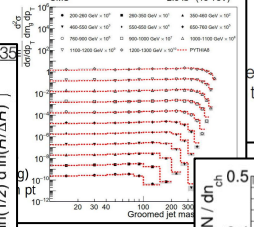
- Allows quark and gluon jet distributions to be measured separately without fraction or template inputs:



[Metodiev '18]

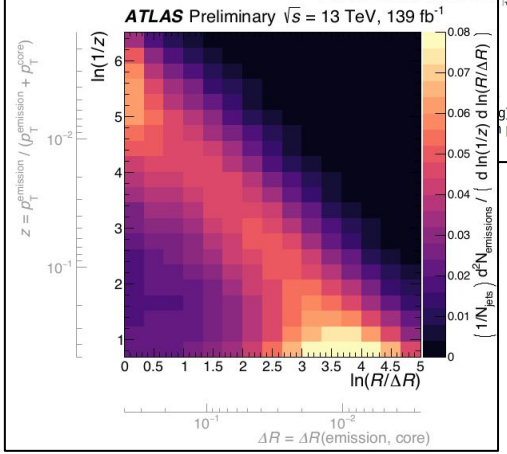
Measureing Jet structure

- SMP-16-010
- Studied with and without



[Roloff '19]

ATLAS-CONF-2019-035

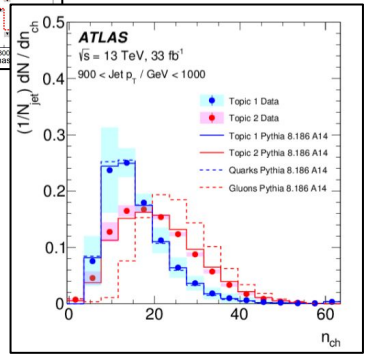


Lund plane representation

To create a Lund plane representation of a jet, recluster a jet j with the Cambridge/Aachen algorithm then decluster the jet following the hardest branch.

1. Undo the last clustering step, defining two subjets j_1, j_2 ordered in p_t.
2. Save the kinematics of the current declustering
$$\Delta \equiv (y_1 - y_2)^2 + (\phi_1 - \phi_2)^2, \quad k_t \equiv p_{t2}\Delta,$$
$$m^2 \equiv (p_1 + p_2)^2, \quad z \equiv \frac{p_{t2}}{p_{t1}+p_{t2}}, \quad \psi \equiv \tan^{-1}\frac{y_2 - y_1}{\phi_2 - \phi_1}.$$
3. Define $j = j_1$ and iterate until j is a single particle.

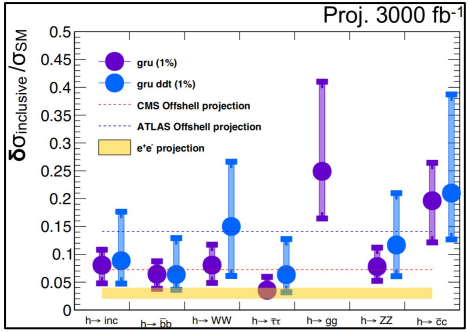
[FD '18]



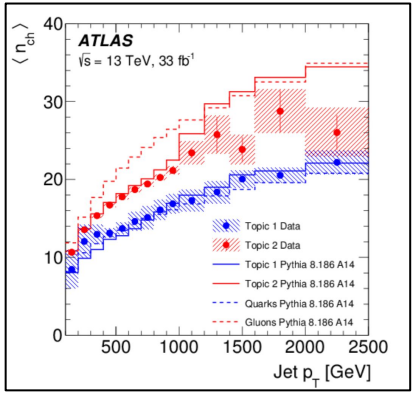
[Kar '19]

BOSTON 2019

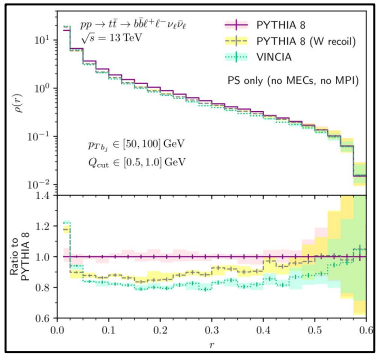
Many interesting talks on a broad range of topics!



[Mantilla Suarez]



[Kar]

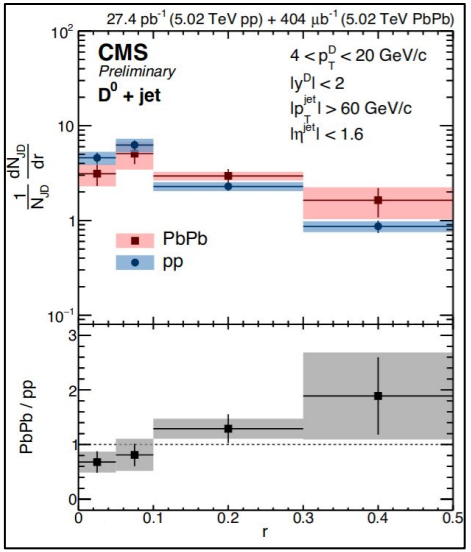


[Brooks]

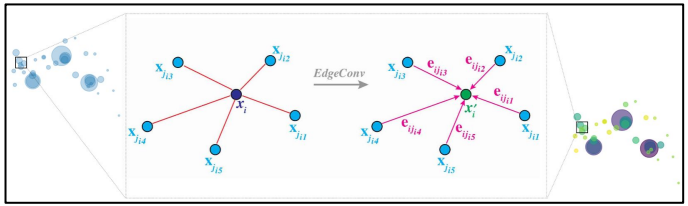
$$
\begin{aligned}
z\Sigma(z) = \frac{\alpha_s}{4\pi}\frac{3C_F}{2} &\\
&+ \left(\frac{\alpha_s}{4\pi}\right)^2 \left[C_A C_F \left(-\frac{50\zeta_2}{3} + 4\zeta_3 - \frac{107\log(z)}{15} + \frac{35366}{675} \right) + C_F n_f \left(\frac{53\log(z)}{60} - \frac{4913}{900} \right) \right.\\
&\left. + C_F^2 \left(\frac{86\zeta_2}{3} - 8\zeta_3 + \frac{25\log(z)}{4} - \frac{8263}{216} \right) \right]\\
&+ \left(\frac{\alpha_s}{4\pi}\right)^3 \left[C_A C_F n_f \left(\frac{379579\zeta_2}{5400} + \frac{3679\zeta_3}{30} - \frac{118\zeta_4}{3} + \left(-\frac{108\zeta_2}{5} + \frac{16\zeta_3}{3} + \frac{6644267}{54000} \right) \log(z) \right.\right.\\
&\left. - \frac{16259\log^2(z)}{1800} - \frac{1025118113}{2160000} \right)\\
&+ C_A C_F^2 \left(-\frac{400\zeta_2^2}{3} + \frac{137305\zeta_2}{216} - 72\zeta_2\zeta_3 + \frac{10604\zeta_3}{15} + \frac{4541\zeta_4}{6} - 216\zeta_5 \right.\\
&\left. + \left(-\frac{1100\zeta_2}{3} + \frac{262\zeta_3}{3} + \frac{105425}{144} \right) \log(z) - \frac{340}{9}\log^2(z) - \frac{105395741}{51840} \right)\\
&+ C_A^2 C_F \left(-\frac{906257\zeta_2}{2700} + 24\zeta_2\zeta_3 - \frac{47483\zeta_3}{90} - \frac{481\zeta_4}{6} + 56\zeta_5 + \left(\frac{503\zeta_2}{5} - \frac{74\zeta_3}{3} - \frac{2916859}{6750} \right) \log(z) \right.\\
&\left. + \frac{8059\log^2(z)}{300} + \frac{964892417}{540000} \right)\\
&+ C_F^2 n_f \left(-\frac{15161\zeta_2}{120} - \frac{7994\zeta_3}{45} + \frac{236\zeta_4}{3} + \left(\frac{416\zeta_2}{9} - \frac{32\zeta_3}{3} - \frac{6760183}{64800} \right) \log(z) + \frac{4619\log^2(z)}{720} \right.\\
&\left. + \frac{164829499}{486000} \right) + C_F n_f^2 \left(\frac{6\zeta_2}{5} + \frac{23\log^2(z)}{45} - \frac{8867\log(z)}{1350} + \frac{88031}{4500} \right)\\
&+ C_F^3 \left(\frac{688\zeta_2^2}{3} - \frac{18805\zeta_2}{216} + 48\zeta_2\zeta_3 + 52\zeta_3 - 1130\zeta_4 + 208\zeta_5 + \left(\frac{1849\zeta_2}{9} - \frac{172\zeta_3}{3} - \frac{723533}{2592} \right) \log(z) \right.\\
&\left. + \frac{625\log^2(z)}{48} + \frac{742433}{1944} \right) \Bigg]\\
&+ \mathcal{O}(\alpha_s^4)
\end{aligned}
$$

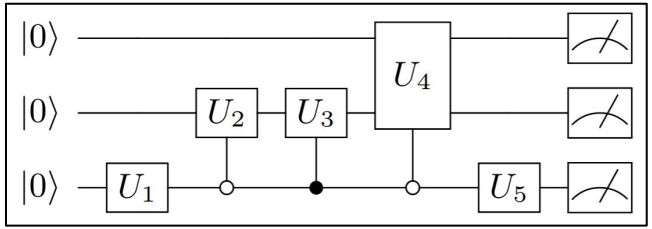
[Moult]

[Wang]



[Qu]





[Nachman]

Many interesting talks on a broad range of topics!

Can we classify them with ML?

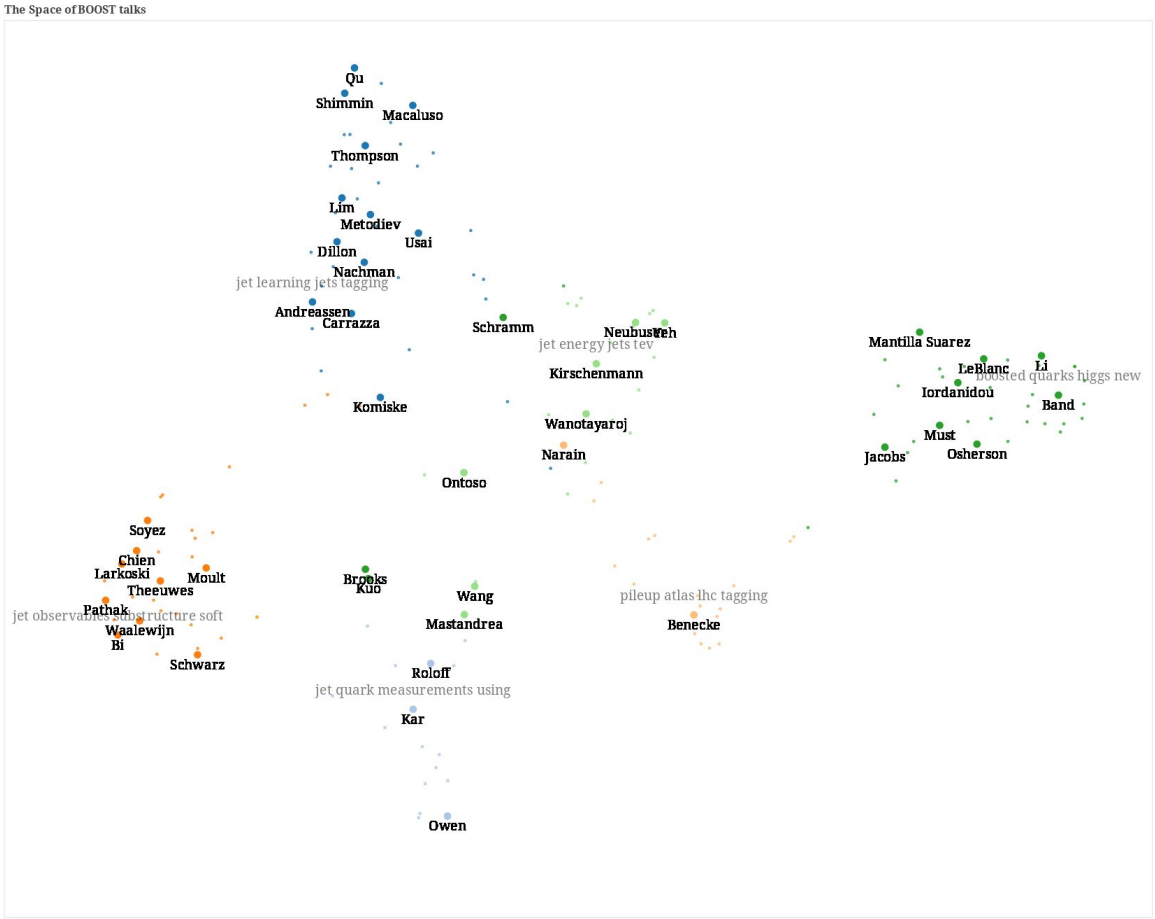
BOOSTON 2019

Many interesting talks on a broad range of topics!

Can we classify them with ML?

Topic modelling with LDA on titles and abstracts

⇒ not so bad!



(projected to two dimensions with t-SNE)

[inspired by Metodiev's "Space of Jets" plot]

Progress this year

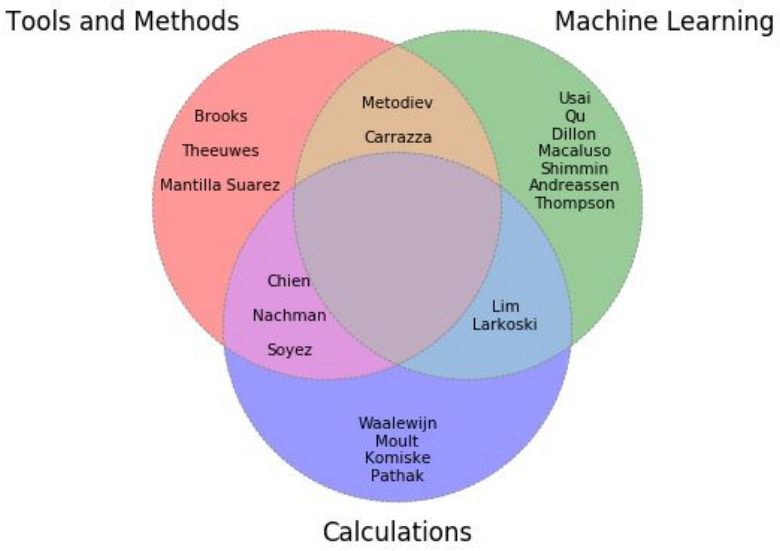
I will split the talks in three broad categories

- New Calculations
- Advances in Tools and Methods
- Machine Learning

Progress this year

I will split the talks in three broad categories

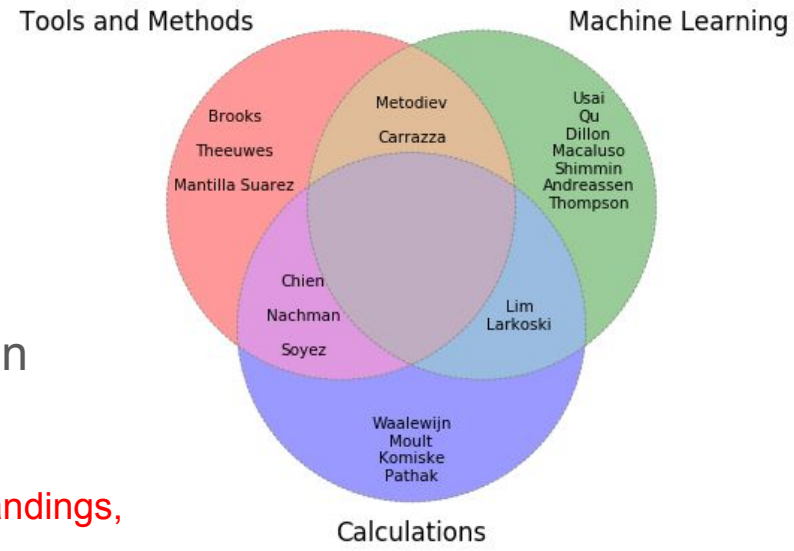
- New Calculations
- Advances in Tools and Methods
- Machine Learning



Progress this year

I will split the talks in three broad categories

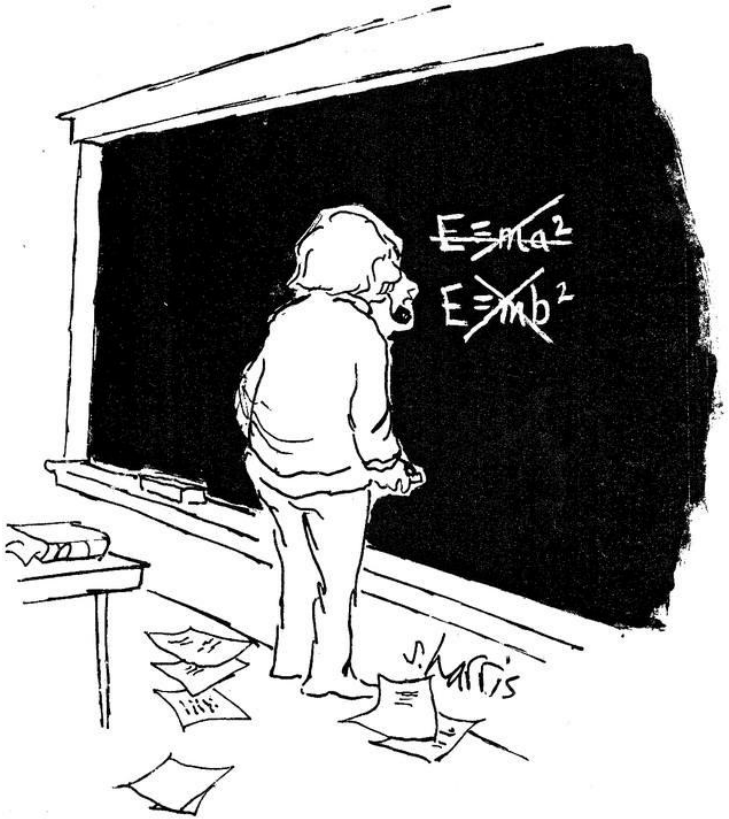
- New Calculations
- Advances in Tools and Methods
- Machine Learning



Good for the field, but bad for the summary speaker: many talks at the interface between different topics.

Disclaimer: Apologies for any omission or misunderstandings, please refer to the original talks !

Calculations



Credit: Sydney Harris

New Calculations

This wouldn’t be BOOST without a range of new calculations!

Several new results pushing the precision boundaries:

- NNLL calculation of substructure observable
- NLL’ calculation of jet shape
- Calculation of power corrections for groomed jet mass

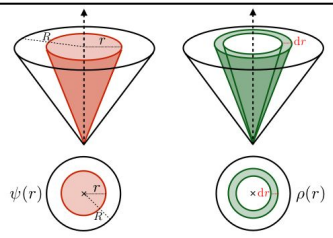
In parallel to precision effort we also saw:

- Calculation of new observables
- Algorithmic improvements for correlators
- Theory view on quark and gluon discrimination

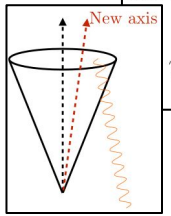
[Waalewijn]

Precision calculation of the jet shape

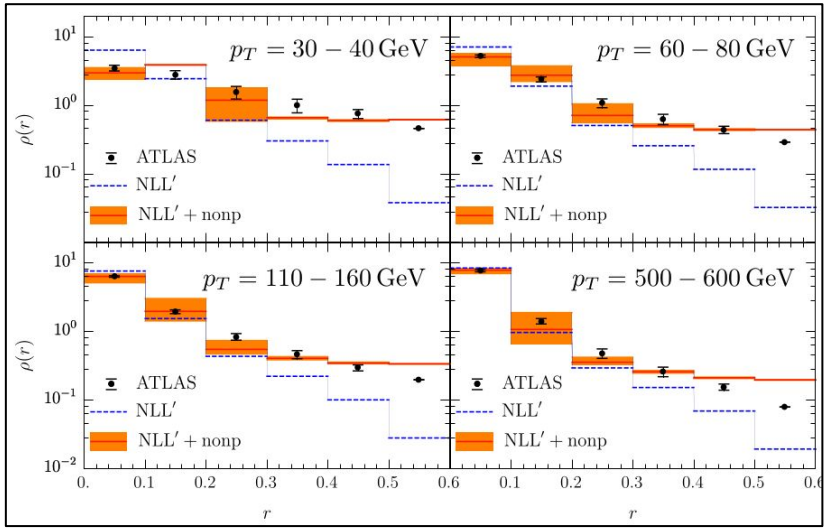
- New state-of-the-art NLL' calculation of jet shape, resumming R and r/R logs
- Recoil of the jet axis due to soft radiation taken into account
- Model non-pert. effects by adding localized energy, with displacement of jet axis



Jet shape is average $z_r = p_T^{\text{subjet}}/p_T$



$$\psi(r) = \int_0^1 \mathrm{d}z_r\, z_r \frac{\mathrm{d}\sigma}{\mathrm{d}p_T\,\mathrm{d}\eta\,\mathrm{d}z_r} \Big/ \frac{\mathrm{d}\sigma}{\mathrm{d}p_T\,\mathrm{d}\eta} \qquad \rho(r) = \frac{\mathrm{d}\psi}{\mathrm{d}r}$$



- Good agreement with ATLAS & CMS data
- Non-perturbative model critical for description of distribution tail

[Pathak]

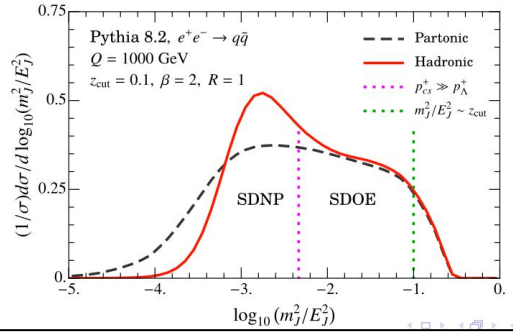
Power Corrections for Soft Drop Jet Mass

Distinguish regions of the groomed jet mass spectrum:

a) soft drop operator expansion (SDOE) region, $p_{cs}^{+} \gg p_{\Lambda}^{+}$: $\frac{Q\Lambda_{\rm QCD}}{m_J^2}\left(\frac{m_J^2}{QQ_{\rm cut}}\right)^{\frac{1}{2+\beta}} \ll 1$

b) soft drop nonperturbative (SDNP) region, $p_{cs}^{+} \sim p_{\Lambda}^{+}$: $m_J^2 \lesssim Q\Lambda_{\rm QCD}\left(\frac{\Lambda_{\rm QCD}}{Q_{\rm cut}}\right)^{\frac{1}{1+\beta}}$

c) ungroomed resummation region: $m_J^2 \gtrsim z_{\rm cut}\frac{Q^2}{4}$.

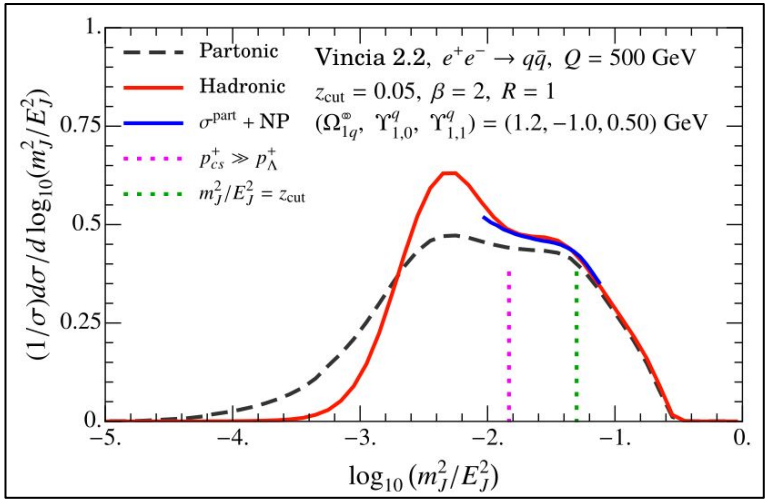


Leading power corrections for the full cross section can be parameterized as

$\Upsilon_1^{\circledcirc}(\beta) = \Upsilon_{1,0}^{\circledcirc} + \beta\,\Upsilon_{1,1}^{\circledcirc}$

$$\frac{d\sigma_\kappa^{\rm had}}{dm_J^2} = \frac{d\hat{\sigma}_\kappa}{dm_J^2} - Q\,\Omega_1^{\infty}\frac{d}{dm_J^2}\left(C_1(m_J^2, Q, z_{\rm cut}, \beta)\frac{d\hat{\sigma}_\kappa}{dm_J^2}\right) + \frac{Q\Upsilon_1^{\circledcirc}(\beta)}{m_J^2}\,C_2(m_J^2, Q;\, z_{\rm cut}, \beta)\frac{d\hat{\sigma}_\kappa}{dm_J^2}$$

- Groomed jet mass robust to hadronization corrections, interesting candidate for α_s extraction
- Requires accurate theoretical predictions over wide mass range
- Identify Soft Drop Operator Expansion region where leading power corrections can be included

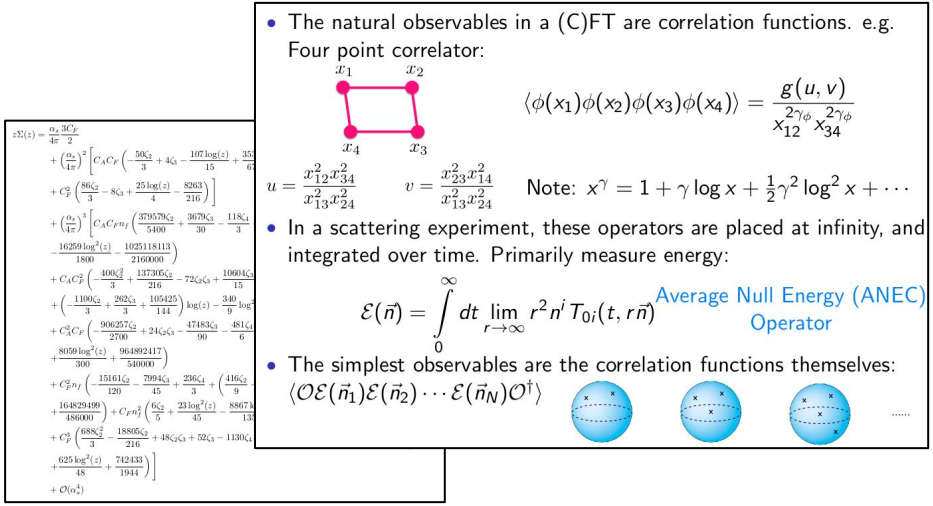


- Fit universal coefficients using parton and hadron-level MC
- Good description of hadronization corrections down to the fully non-perturbative region

[Moult]

The Collinear Limit of the Energy-Energy Correlator: From CFTs to Jet Substructure

- Can we understand all-order structure of substructure observables in a simpler theory and apply lessons to QCD?
- Consider collinear limit of N-point EEC in CFT, interesting because: no double logs, no non-global logs, unaffected by grooming
- Achieve NNLL accuracy by computing twist-two anomalous dimensions, mapping CFT jet function to QCD equivalent



- The natural observables in a (C)FT are correlation functions. e.g. Four point correlator:

$$\langle \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle = \frac{g(u,v)}{x_{12}^{2\gamma_\phi} x_{34}^{2\gamma_\phi}}$$

$$u = \frac{x_{12}^2 x_{34}^2}{x_{13}^2 x_{24}^2} \qquad v = \frac{x_{23}^2 x_{14}^2}{x_{13}^2 x_{24}^2} \qquad \text{Note: } x^\gamma = 1 + \gamma \log x + \tfrac{1}{2}\gamma^2 \log^2 x + \cdots$$

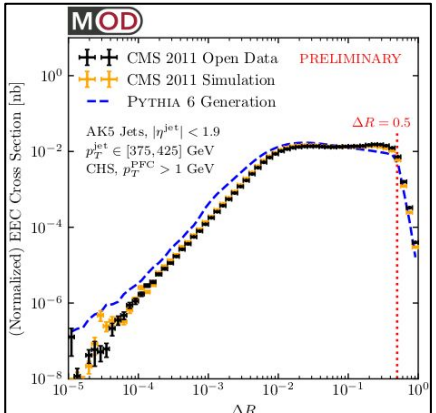
- In a scattering experiment, these operators are placed at infinity, and integrated over time. Primarily measure energy:

$$\mathcal{E}(\vec{n}) = \int_0^\infty dt \lim_{r\to\infty} r^2 n^i T_{0i}(t, r\vec{n})$$

Average Null Energy (ANEC) Operator

- The simplest observables are the correlation functions themselves:

$$\langle \mathcal{O}\mathcal{E}(\vec{n}_1)\mathcal{E}(\vec{n}_2)\cdots\mathcal{E}(\vec{n}_N)\mathcal{O}^\dagger\rangle$$



EEC is sensitive probe of initiating parton and can be computed at high perturbative accuracy

[Komiske]

Cutting Multiparticle Correlators Down to Size

Energy Flow Moments (EFMs)

[PTK, Metodiev, Thaler, *to appear soon*]

$\theta_{ij} = \sqrt{2 n_i^\mu n_{j\mu}}$ $\beta = 2$ removes square root

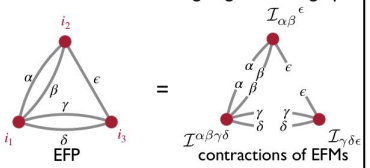
Factors of n_i^μ can be organized in optimal way

EFM$_v$ is a totally symmetric little group tensor

$$\mathcal{I}^{\mu_1 \cdots \mu_v} = \sum_{i=1}^{M} z_i n_i^{\mu_1} \cdots n_i^{\mu_v}$$

v	0	1	2	3	4	5	6
$n^{(d=4)}_{\text{components}}$	1	4	10	20	35	56	84

EFMs result from cutting edges of EFP graph



All $\beta = 2$ EFPs are $\mathcal{O}(M)$

$\text{ECF}_N^{(\beta=2)}$ are all $\mathcal{O}(M)$

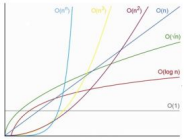
$D_2^{(\beta=2)}, C_2^{(\beta=2)}$ are $\mathcal{O}(M)$

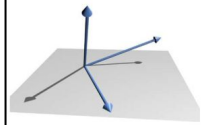
- Multiparticle correlators are ubiquitous observables in collider physics
- Naive computation complexity is M^N but tree graphs can be computed in with M^2 Variable Elim

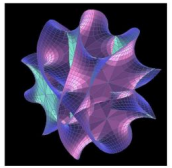
- Reorganize sums with Energy Flow Moments
- β=2 EFPs become O(M)
- Can be used to find Tensor identities

$$T^{a_1 \cdots a_k}_{b_1 \cdots b_\ell [c_1 \cdots c_m]} = 0$$

and to count basis of amplitudes







Experiment

Computational Complexity

Multiparticle correlators are $\mathcal{O}(M^N)$ to compute in general

$\beta = 2$ EFPs can be computed in $\mathcal{O}(M)$

Why not use $D_2^{(\beta=2)}$? Performance in backup

Linear Tensor Identities

Multiparticle correlators exhibit mysterious linear redundancies

All redundancies understood via cutting graphs and applying master antisymmetrization identity

Counting Superstring Amplitudes

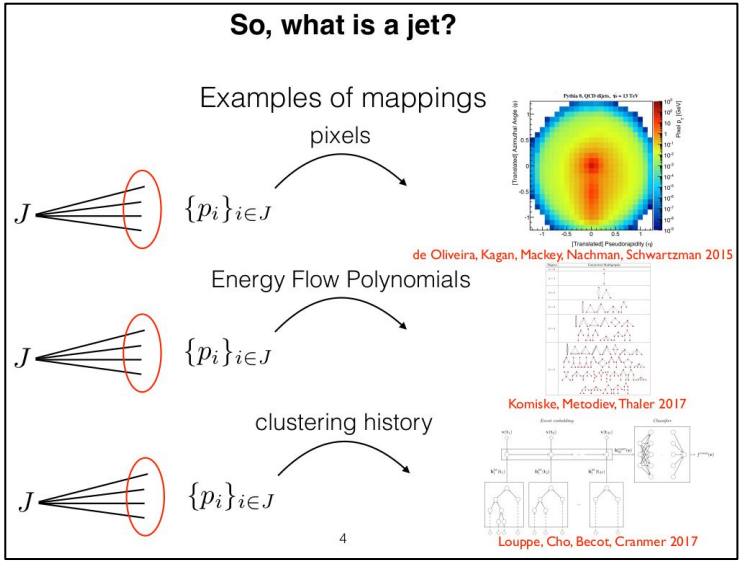
Counting independent kinematic polynomials difficult

Immediate enumeration through multigraphs and new OEIS sequences!

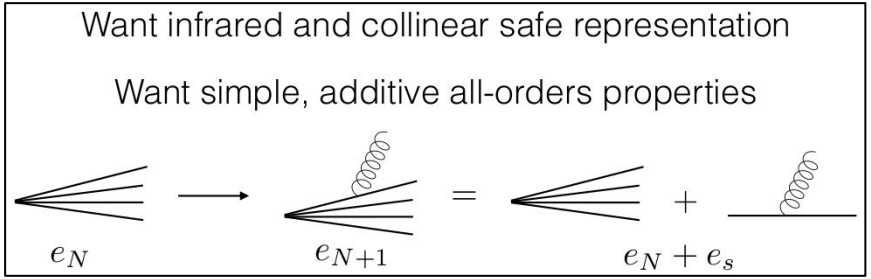
Theory

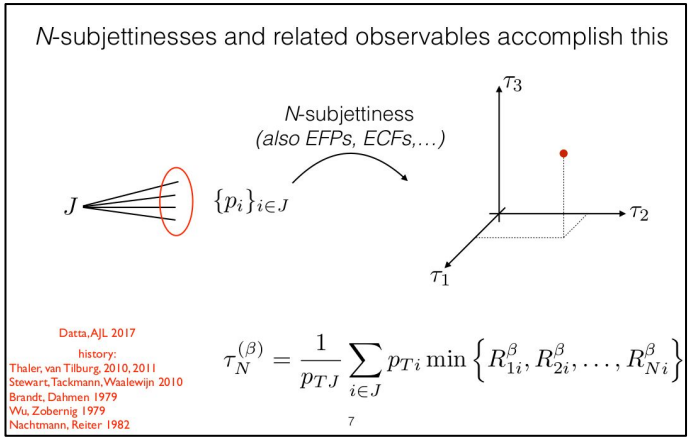
[Larkoski]

A Theory of Quark vs. Gluon Discrimination



- Different representations of jets should be roughly equivalent, but some key properties should be observed
- N-subjettiness, ECFs, EFPs, well suited for theoretical study
- What kinematic properties are driving separation power between quark and gluon initiated jets?



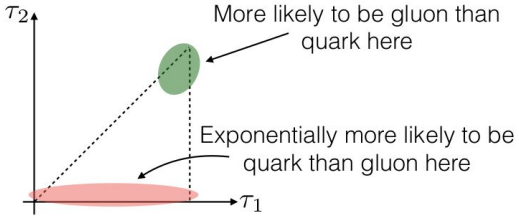


[Larkoski]

A Theory of Quark vs. Gluon Discrimination

Where do jets live?

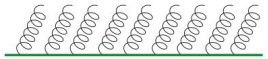
For visualization simplicity, just consider (τ_1,τ_2)



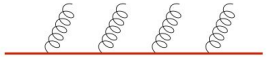
Gluon jets are always contaminated by some quark jets

In practice small because $C_F/C_A \sim 0.44$

Resolving only 6 emissions: $\left(\frac{C_F}{C_A}\right)^{\geq 6} < 1\%$



gluon jet $\propto C_A$



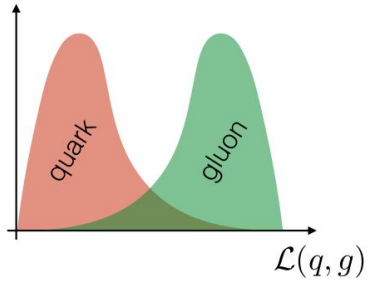
quark jet $\propto C_F < C_A$

Additivity then implies $\tau_N^{\text{gluon}} > \tau_N^{\text{quark}}$

- Consider (τ_1,τ_2) space, gluon jets contaminated by quarks

Quark likelihood = Gluon "reducibility factor" = $\kappa_g \sim \left(\frac{C_F}{C_A}\right)^N$

- But quark sample can be made arbitrarily pure for $\tau_N \to 0$



Practical

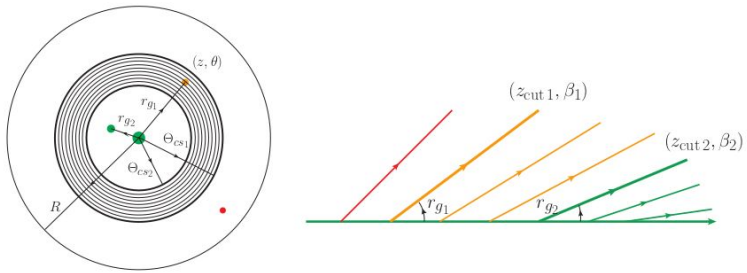
IRC safe observables are good q/g discriminants out of the box

good τ_N discrimination well-known

Gallicchio, Schwartz 2012

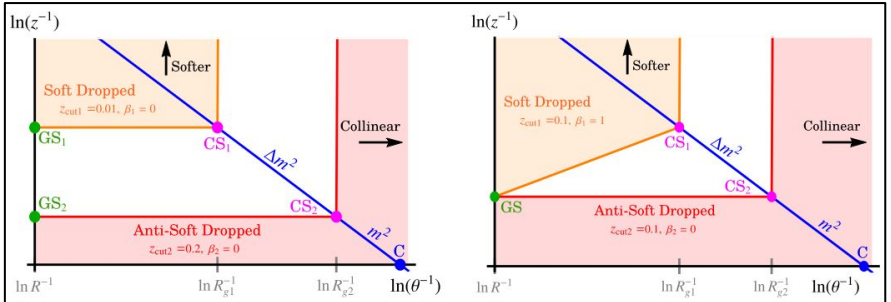
[Chien]

Collinear Drop

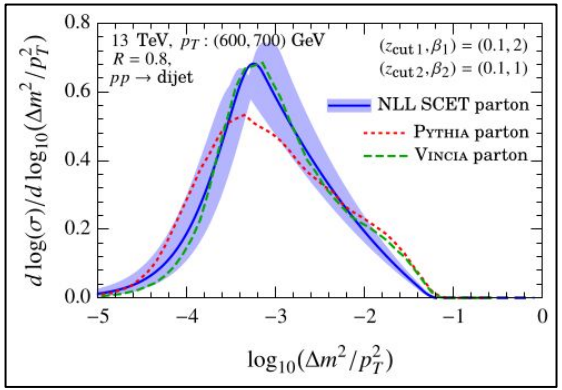


$\Delta m^2 = m^2_{\mathrm{SD}_1} - m^2_{\mathrm{SD}_2}$ probes the soft radiation within the ring

- New observable probing radiation sandwiched between two Soft Drop configurations
- Can probe soft physics while removing collinear contributions
- Calculated up to NLL with SCET and compared to MC, good agreement with Vincia
- Sensitivity to underlying event can be tuned through first SD parameters



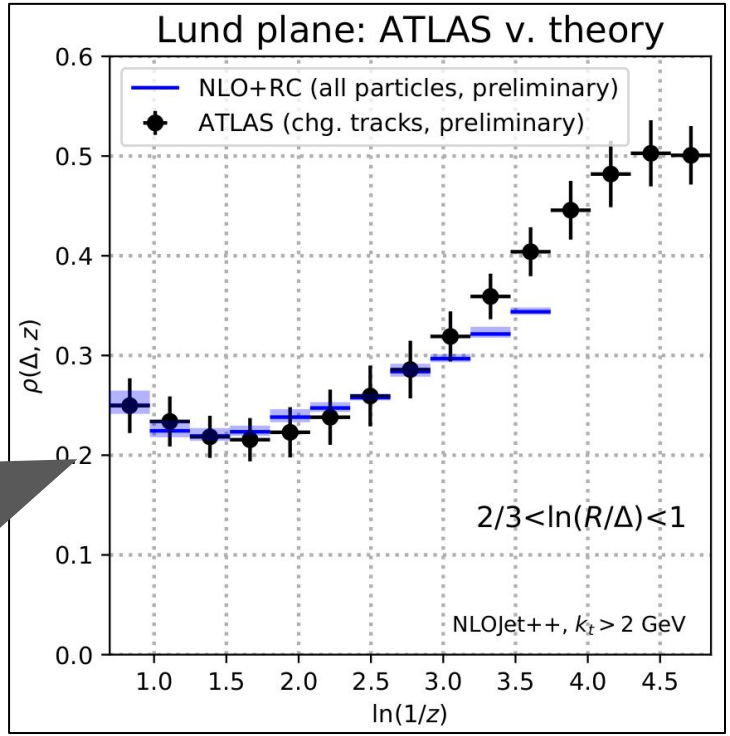
- Two ways to create hierarchical soft-drop conditions $\mathrm{SD}_1 < \mathrm{SD}_2$
  - fixing β and varying z_{cut}
  - fixing z_{cut} and varying β (ATLAS 13 TeV data: $z_{\mathrm{cut}} = 0.1$, $\beta = 0, 1, 2$. 1711.08341)

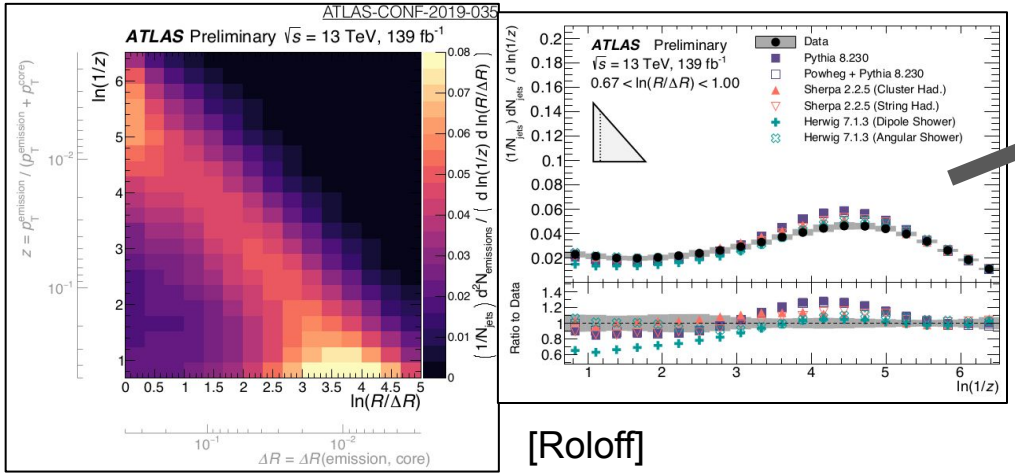


[Thanks to Gregory for the plot]

Bonus: Lund plane comparison

- We saw a first measurement of the Lund plane using charged tracks on Thursday
- Preliminary calculation using NLO + running coupl. shows reasonably good agreement with data!





[Roloff]

Tools and Methods

Box of dental tools, c. 1830-1850, New England
Museum of Fine Arts, Boston

Advances in Tools and Methods

After more than 10 years of BOOST, we are still finding new approaches to extract the most out of the data, and improving modeling of LHC collisions.

- Improvements of parton showers for heavy ion and coloured resonance decays
- New approach for parton showers with quantum computers
- Novel methods to extract fundamental Higgs and QCD constants
- ML-inspired metric for collider events

[Soyez]

Parton shower and jets in the quark-gluon plasma

Seeking driving physics mechanisms

"jet ↔ parton shower"
⇒ need to understand parton showers

Main idea
Start from the most fundamental (simple) point of view in pQCD

Assumptions:

1. target a "leading-logarithmic" (often double logs) accuracy
2. simple, fixed, medium of size L and transport coefficient $\hat{q}$
   ($\hat{q}$ is the averaged $k_\perp^2$ per unit length via kicks from the medium)

Factorised evolution in 3 stages:

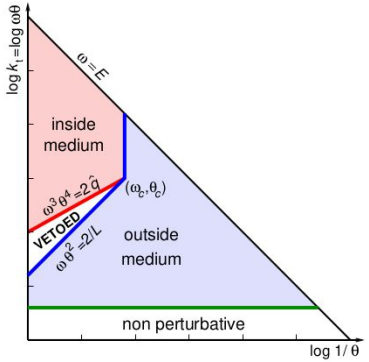
1. in-medium angular-ordered VLEs
2. each VLE sources MIEs propagating through the medium
3. out-medium VLEs with first emission at any angle

- Can we build a leading-log shower for heavy ions?
- Consider both **Vacuum-like** and **medium-induced emissions**.
- VLEs: angular-ordered emissions derived from soft-collinear limit of splitting function
- MIEs: additional emissions ordered in formation time $t_f = 2/\omega\theta^2$, only soft divergence

Factorised physical picture

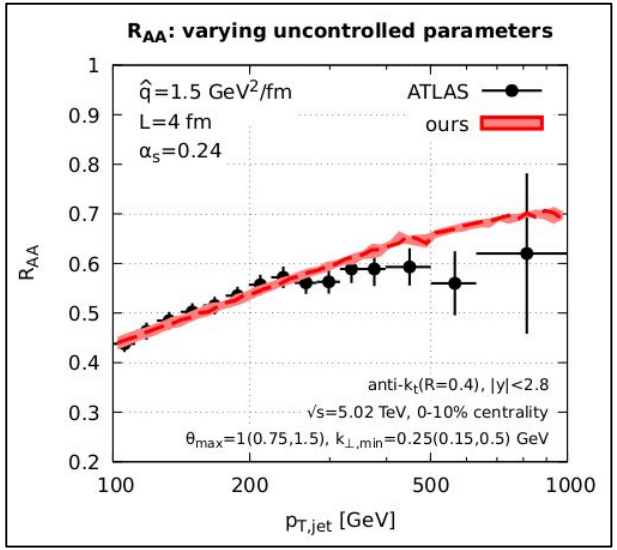
Double-log accuracy:

- in-medium VLEs
- medium length ($t_f = L$)
- VLEs vetoed in between
- colour (de)coherence
  - ✓ in-medium has $\theta > \theta_c$
  - ✓ in-medium: angular-ordered
  - ✓ in→out jump: no orgering

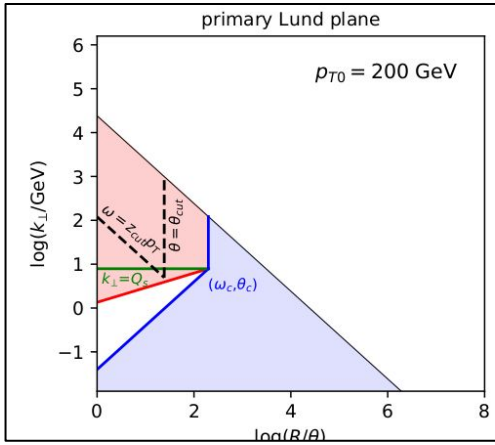


[Soyez]

Parton shower and jets in the quark-gluon plasma



- This method provides reasonable description of data
- VLE shower provides multiple sources for the MIEs, critical in increasing energy loss at high pt



Note: (i) use mMDT (SD, $\beta = 0$)
(ii) impose condition $\theta_g > \theta_{cut}$

High-p_T ($p_T z_{cut} \theta_{cut} \gg \omega_c \theta_c$)
only VLEs can contribute

Low-p_T ($p_T z_{cut} \theta_{cut} \lesssim \omega_c \theta_c$)
both VLEs and MIEs
MIEs add peak at small z_g

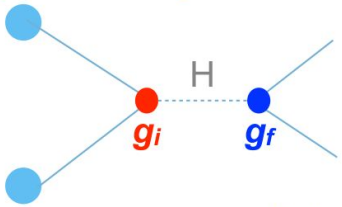
Two basic phenomena:

1. SoftDrop condition triggered either by a VLE or by a MIE
2. both subjets lose energy by MIE emissions

[Mantilla Suarez]

Constraining the Higgs width at the HL-LHC

Total cross section depends on coupling strengths in production g_i and decay g_f stages, and width Γ_H



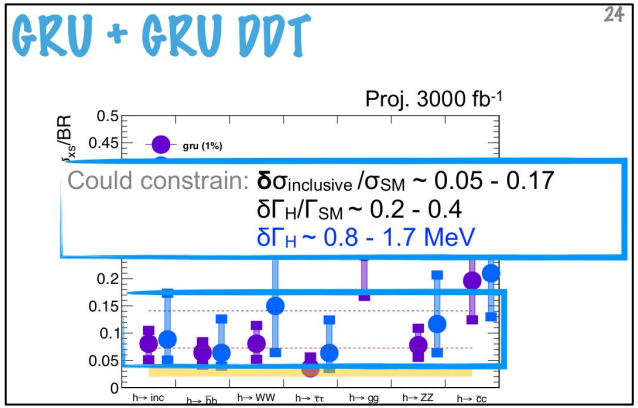
$$\sigma_{i\to H\to f} \sim \frac{g_i^2\, g_f^2}{\Gamma_H} \times m_H \text{ (if on-shell)}$$

How to extract Γ_H from an inclusive cross section measurement?

- Measuring or constraining the Higgs width is critical to pinning down the Higgs sector
- Current strategies rely on off-shell measurements
- Could we place constraints from inclusive cross section and decays, similar to approach for lepton colliders?

1. Measure inclusive cross section from reconstructed m_h
2. Use existing measurements to constrain Γ_H:
   1. boosted h → bb
   2. W+h → bb
   3. W+h → WW

$$\Gamma_h \propto \frac{1}{\sigma(W+h\to WW)} \times \left(\sigma(gg\to h) \times \frac{\sigma(W+h\to \bar{b}b)}{\sigma(ggh\to \bar{b}b)}\right)^2$$

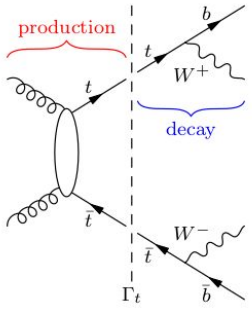


- Proof of concept analysis: could potentially constrain Γ_H to ~0.8-2 MeV, competitive with off-shell method

  ⇒ Interesting avenue for further study

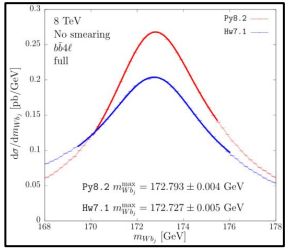
[Brooks]

Coherent Showers in Decays of Coloured Resonances



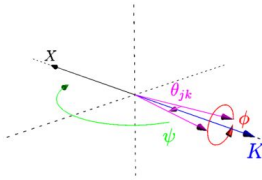
This talk: goal is to improve modelling of radiation produced in decay of (coloured) resonances by including coherence.
→ Potential implications for e.g. b-taggers

- Large uncertainties associated to specific shower model
- New approach to coherent parton showers in the decays of coloured resonances, based on resonance-final QCD antennae, implemented in VINCIA
- Same singularity structure as IF antennae but different phase-space factorisations and kinematic mapping
- Noticeable differences in b-jet profile and top mass spectrum

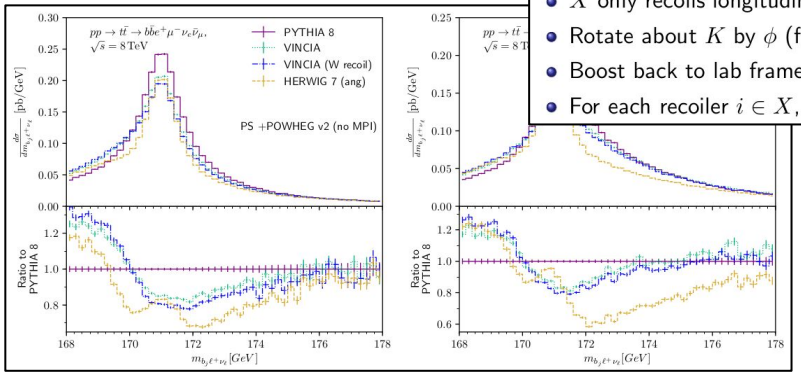


Kinematic Map (recoil strategy)

- Construct in A rest frame.
- X only recoils longitudinally.
- Rotate about K by ϕ (flatly sampled).
- Boost back to lab frame.
- For each recoiler $i \in X$, boost p_i by $p_{X'} - p_X$



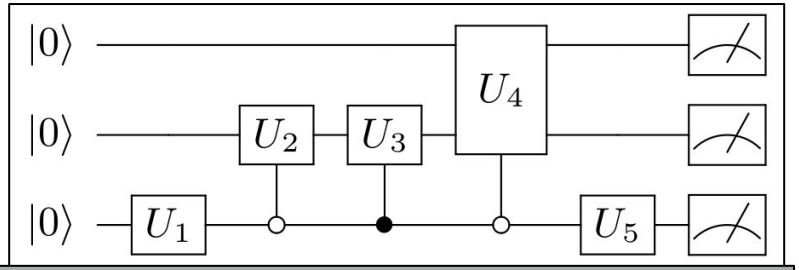
Parton showers are widely used, important to study their uncertainties and accuracy



[Nachman]

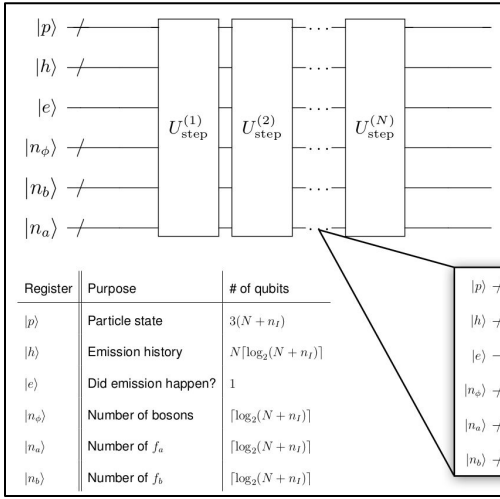
Investigating the use of Quantum Computers for Final State Radiation

- Quantum computers provide a natural framework to perform computations of quantum phenomena
- Classical Markov Chain MC parton showers forced to neglect some interference/subleading color effects

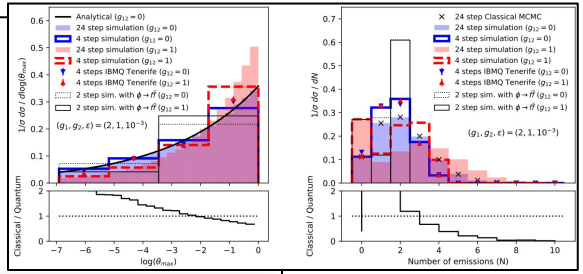


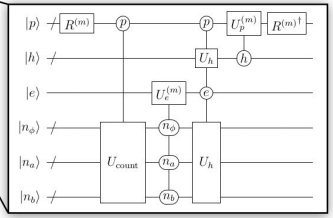
Can a piece of the calculation that is hard/impossible with classical computers & accelerate it on a QPU?

⇒ Develop quantum circuit describing quantum properties of PS



Register	Purpose	# of qubits
$\lvert p\rangle$	Particle state	$3(N+n_I)$
$\lvert h\rangle$	Emission history	$N\lceil\log_2(N+n_I)\rceil$
$\lvert e\rangle$	Did emission happen?	1
$\lvert n_\phi\rangle$	Number of bosons	$\lceil\log_2(N+n_I)\rceil$
$\lvert n_a\rangle$	Number of f_a	$\lceil\log_2(N+n_I)\rceil$
$\lvert n_b\rangle$	Number of f_b	$\lceil\log_2(N+n_I)\rceil$





Toy QED

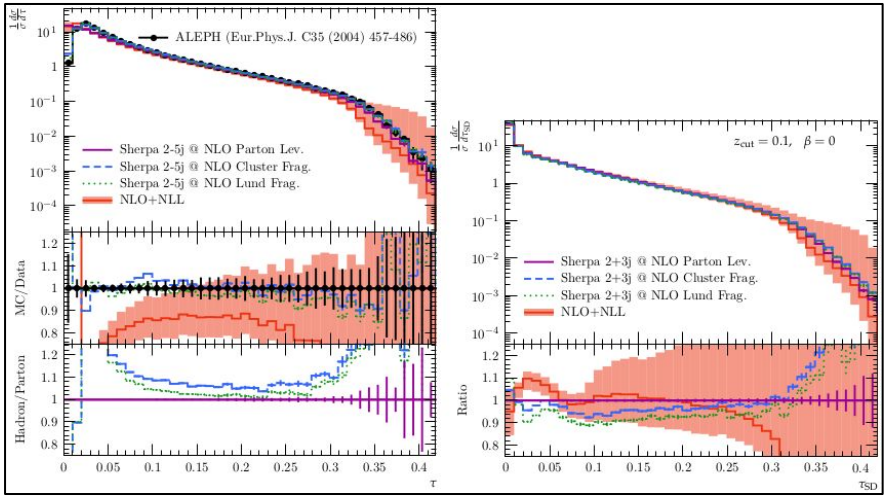
- Quantum circuit can capture interferences, scaling is polynomial

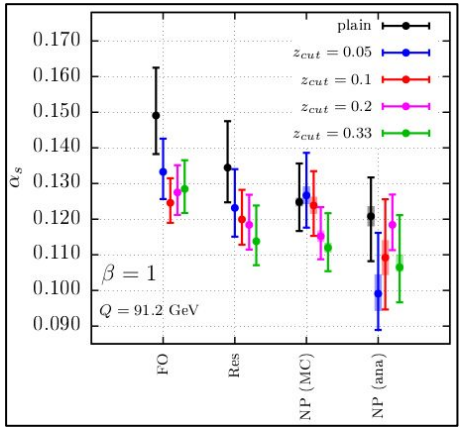
⇒ promising avenue for future applications of quantum computers in HEP!

[Theeuwes]

Fitting the strong coupling constant for soft-drop thrust

- Precise measurement of α_s critical, uncertainties feed back into all other theoretical predictions
- Can jet physics provide a competitive measurement?
- Fit from Thrust with Soft-Drop grooming:
  - reduced hadronization effects
  - known at NLO+NLL
  - can be measured to high accuracy

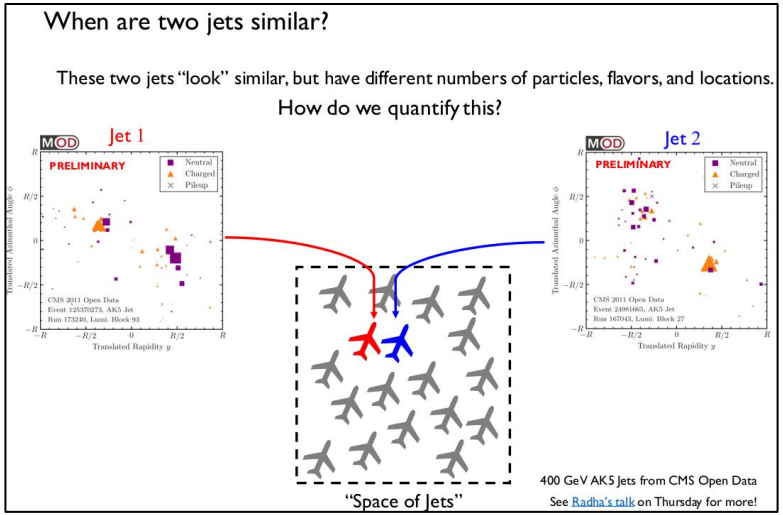




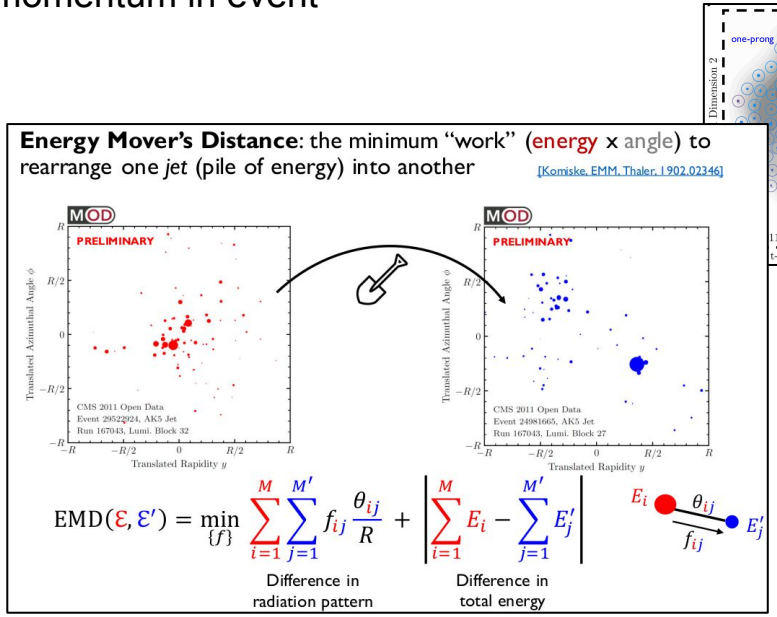
- Monte Carlo study indicates potential to extract α_s to ~5% at LEP
- Will it be competitive with real data and uncertainties?
- Can it be extended to the LHC?

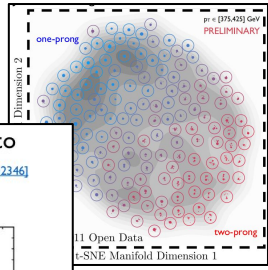
[Metodiev]

The Space of Collider Events



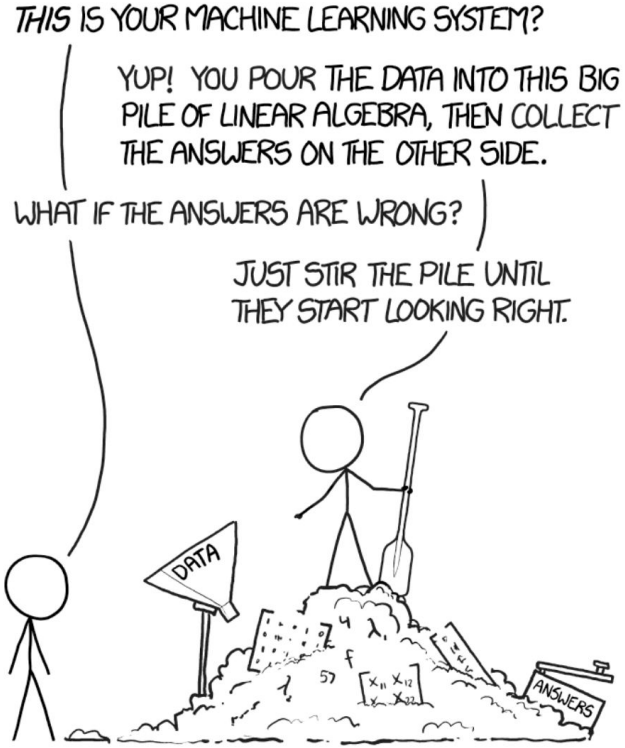
- Can we systematically quantify the difference between two jets or events?
- Energy Mover's distance: metric between two probabilities, quantifies work required to rearrange momentum in event





- Provides an interesting tool to classify or visualize the space of jets
- Can this be used to improve pileup subtraction or to train GANs?

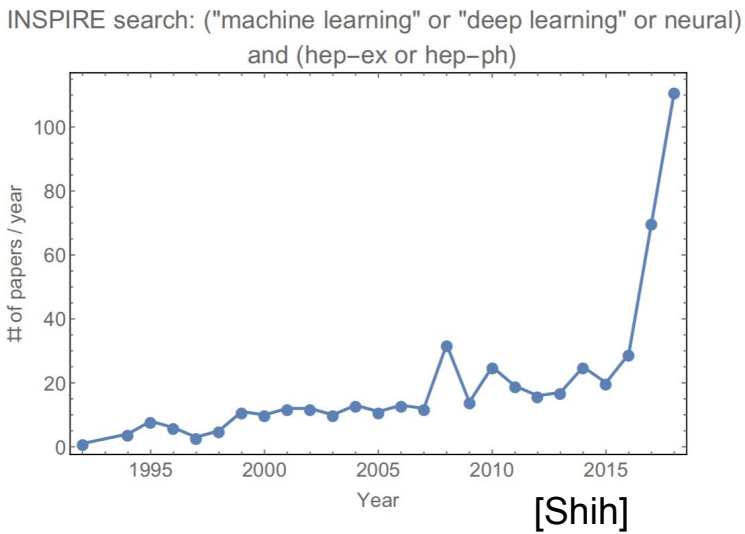
Machine Learning



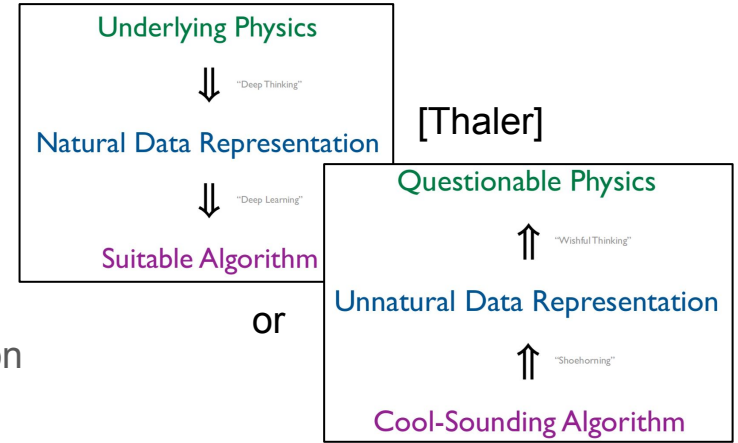
Credit: xkcd.com

ML @ BOOST

- Explosion of applications of machine learning in jet substructure over the past ~2 years
- Impact of ML in HEP has spun off into a number of separate venues: ML4Jets, ICML workshop, Aspen conference, ...
- Many interesting talks this year on new ML-driven approaches to solve a range of existing problems
  - Tagging
  - Grooming
  - Reweighting
  - NN uncertainties
  - Interpretability
  - Event-wide identification



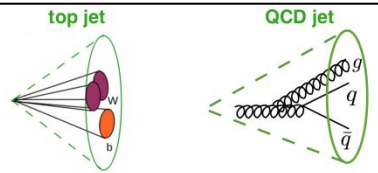
[Shih]



[Thaler]

[Macaluso]

The Machine Learning Landscape of Top Taggers



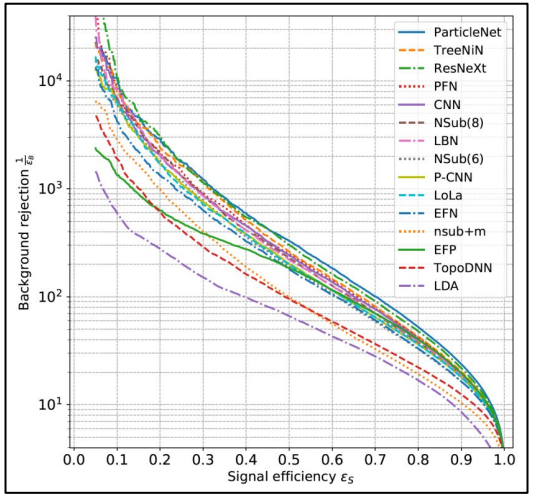
- Typical classifiers search for:
  - Kinematic features induced by the top and W boson mass.
  - Number of prongs: N-subjettiness.
  - Flavor: b-tagging
- Goal of this study:

Compare ML based setups to classify top vs QCD jets.

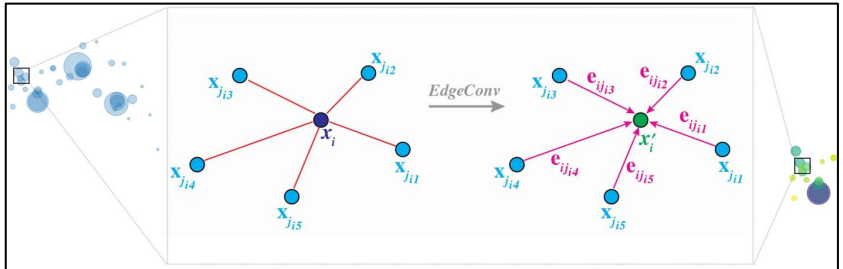
- Comparison of wide range of ML Top taggers using common data set
- Provides a useful benchmark of existing tools, and performance goal for new approaches
- Many taggers provide similar performance

⇒ are there other metrics to consider?
interpretability, calculability, robustness, speed?

- **Image-based taggers**
  - CNN
  - ResNeXt
- **4-Vector-based taggers**
  - TopoDNN
  - Multi-Body N-Subjettiness
  - TreeNiN
  - P-CNN
  - ParticleNet
- **Theory-inspired taggers**
  - Lorentz Boost Network
  - Lorentz Layer
  - Latent Dirichlet Allocation (LDA)
  - Energy Flow Polynomials
  - Energy Flow Networks
  - Particle Flow Networks



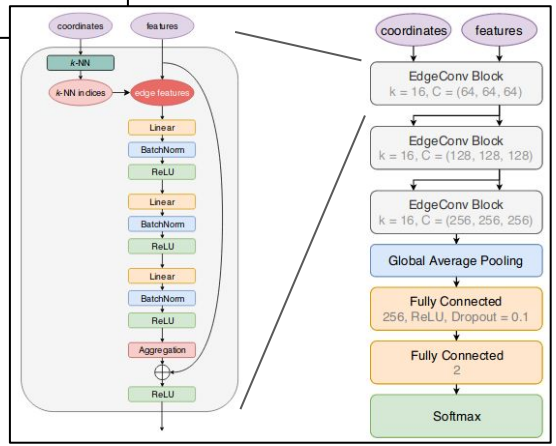
ParticleNet: Jet Tagging via Particle Clouds

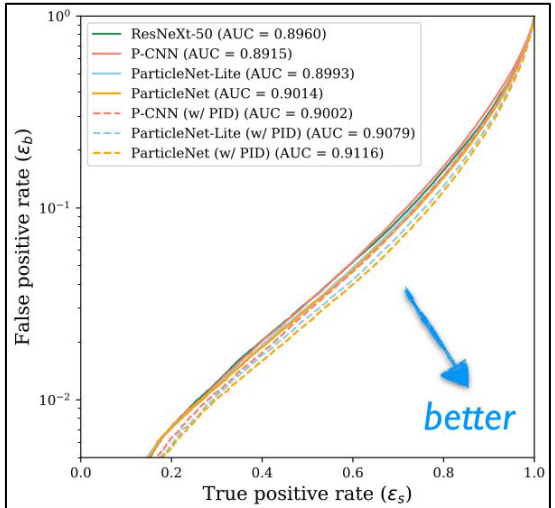


- Convolution on point clouds: *EdgeConv* [arXiv:1801.07829]
  - treating a point cloud as a graph: each point is a vertex
  - for each point, a local patch is defined by finding its k-nearest neighbors
  - designing a symmetric "convolution" function

- Consider jets as "particle clouds", i.e. unordered set of its constituent particles
- Map particle cloud to a graph, with each vertex representing a particle
- Use this representation with customized neural network architecture based on Dynamic Graph CNN
- Outperforms all other current jet tagging techniques

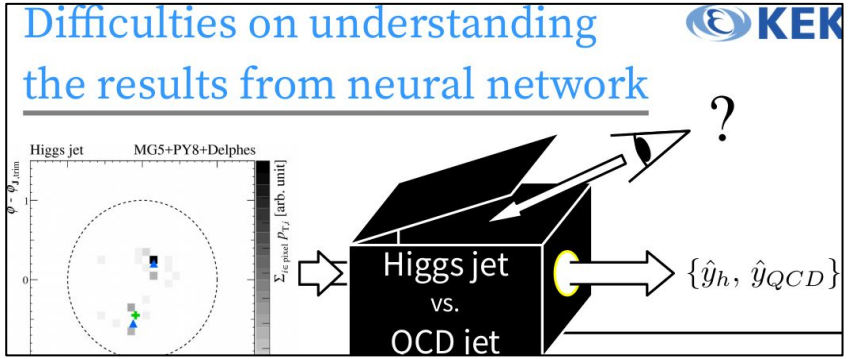
- How robust is this method to non-perturbative effects/soft-collinear splittings?
- Can some IRC safety be implemented in a way that retains performance?





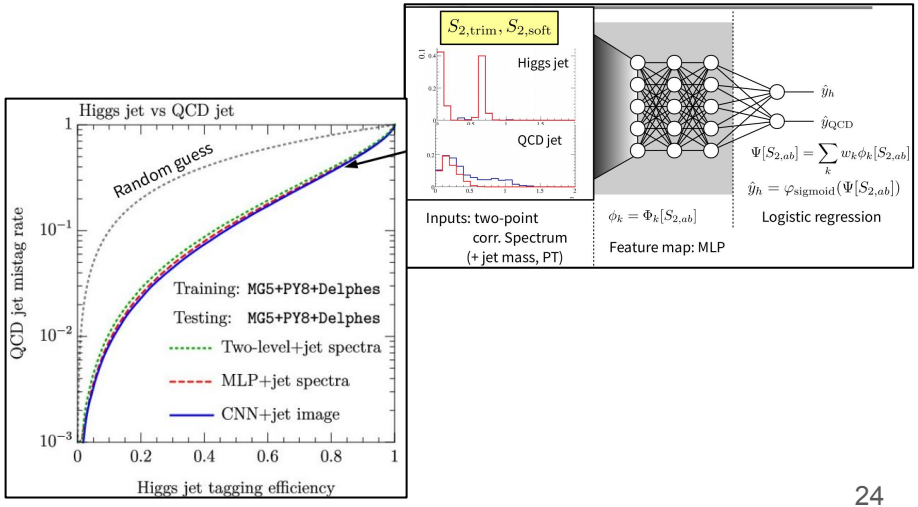
[Lim]

Interpretable Deep Learning for Two-Prong Jet Classification with Jet Spectra



- Can we open the box and gain some first principle understanding of what a NN is learning?
- Use two-point correlation spectrum as input to NN

$$S_{2,\mathrm{trim}}(R) = \int d\vec{R}_1\, d\vec{R}_2\; P_{T,\mathbf{J}_{\mathrm{trim}}}(\vec{R}_1) P_{T,\mathbf{J}_{\mathrm{trim}}}(\vec{R}_2)\delta(R - R_{12})$$



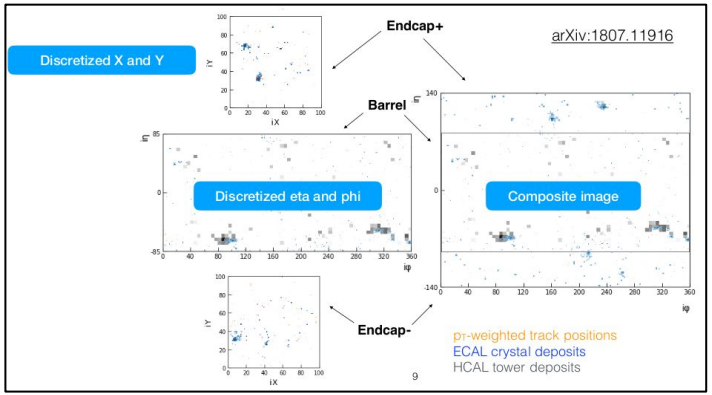
- Can achieve good performance with truncated series, equivalent to logistic regression
- Method could be extended to three-pronged jets

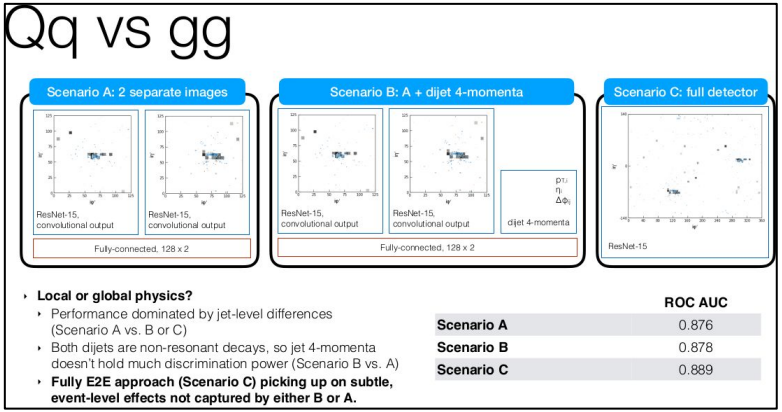
Can this interpretability be used for analytic calculations or to study non-pert. effects?

[Usai]

End-to-end particle and event identification at the Large Hadron Collider with CMS Open Data

- Is it possible to do end-to-end identification using low-level detector data?
- Jet tagging competitive with alternative approaches
- Performs well on full events, can separate qq v. gg dijets better with a fully end-to-end approach





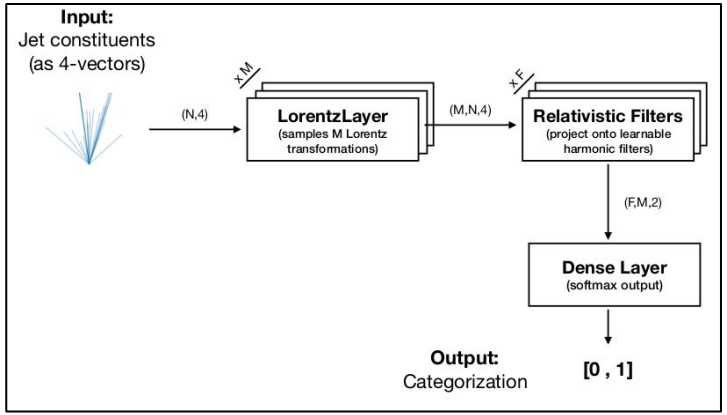
	ROC AUC
Scenario A	0.876
Scenario B	0.878
Scenario C	0.889

Is it possible to understand what additional correlations the event-level classification is learning?
Does it depend on the jet radius, ISR?

[Shimmin]

Relativistic Harmonic Networks

- Can we construct NN architecture that builds in symmetries to rotations and boosts?
- Group equivariant networks tend to learn faster and with less parameters when applied to problems with appropriate symmetries



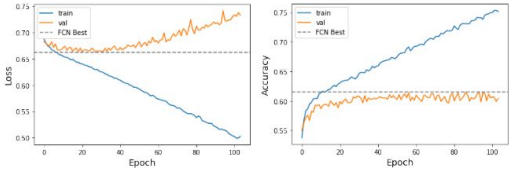
Relativistic Filters

As before, we can construct arbitrary, **learnable** filters:

$$\psi(\xi,\theta,\phi) = \int_0^{N_{\max}} dN \sum_{\ell,m}^{\ell_{\max}} a_{\ell m}(N)\Psi_{N\ell m}(\xi,\theta,\phi)$$

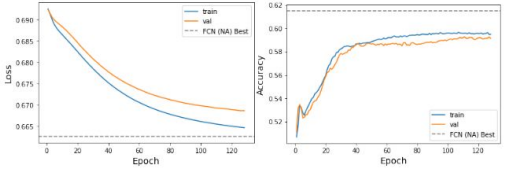
Network 1: Fully Connected
4 layers, 90k parameters

Accuracy: ~62%



Network 2: Relativistic
Harmonic Net (l_{max}=0), 596 parameters

Accuracy: ~59%

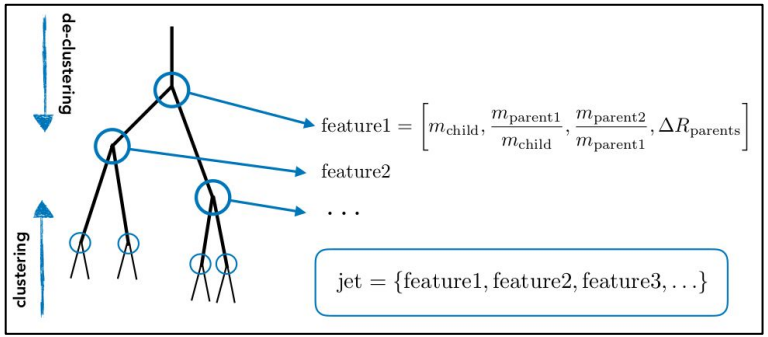


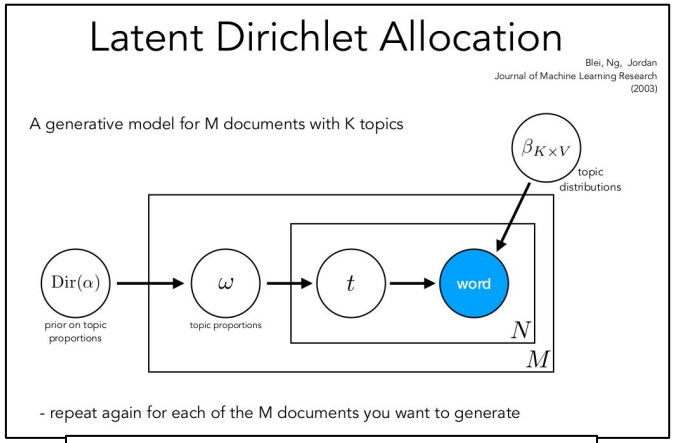
- Proof-of-concept study with single-layer network seems to indicate no gain in performance, but smaller network sizes

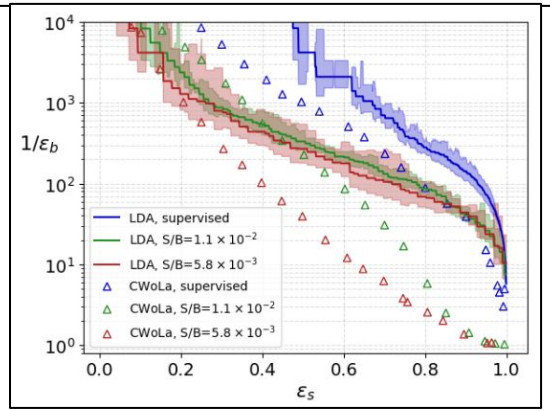
[Dillon]

Uncovering latent jet substructure

- Can we use Latent Dirichlet Allocation to extract jet topics differentiating signal and background?
- Use clustering history to extract features





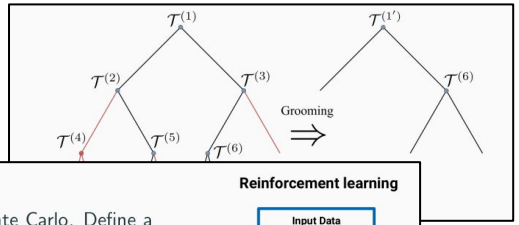


- Application to classification of new physics signal
- Efficient background rejection even with small signal sample

[Carrazza]

Jet grooming through reinforcement learning

- Jet grooming critical tool in jet physics
- Can performance of grooming algorithms be improved?
- Use reinforcement learning to train a NN-based groomer with optimal mass resolution
- Desired behaviour is enforced through reward fct



Input data:

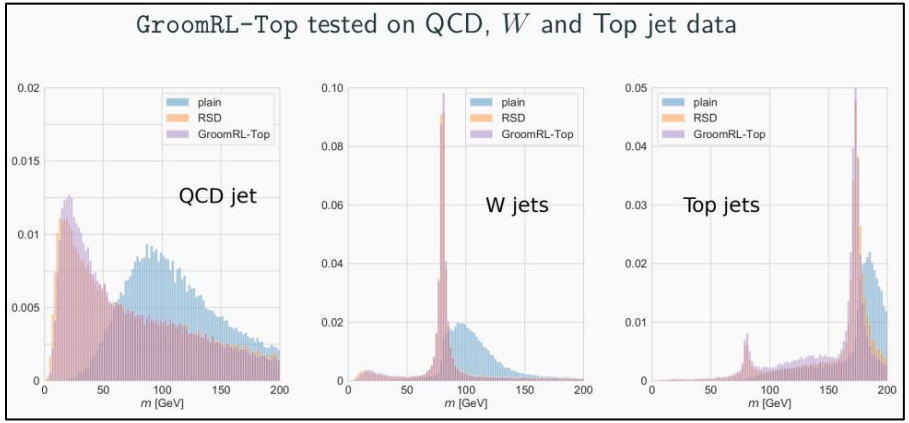
Generate jet events with Monte Carlo. Define a set of possible **states** in a five dimensional box:

$$s_t = \{z, \Delta_{ab}, \phi, m, k_t\}$$

Methodology:

Jet grooming is characterized by a policy and a sequential set of actions/cuts, so:

- Train a reinforcement learning agent which learns how to decide which action to take.
- Define an environment reward which motivates the agent to groom efficiently.



- ML grooming leads to improved mass resolution
- Algorithm retains calculability and robustness of heuristic methods

[Thompson]

Deep-Learning Jets with Uncertainties and More

- Is it possible to use Bayesian NN as a handle on uncertainties in ML?

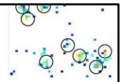
$$p(c^*|C) \approx \int d\omega\, p(c^*|\omega)\, q_{\mu,\sigma}(\omega) \approx \frac{1}{N}\sum_j^N p(c^*|\omega_j(\mu,\sigma)) \equiv \mu_{\text{pred}}$$

- Same performance as deterministic NN

Bayesian networks

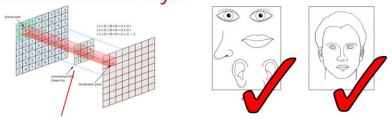
- Consider each weight as a Gaussian distribution
  - Parameterise as mean and standard deviation
- Output distribution has:
  - Mean prediction (best model prediction, μ)
  - Standard deviation (uncertainty, σ)

Whole event tagging



- We want to include large- and small-scale correlations
- Convolutions focus on small scale structure: e.g. for a face:

Convolutional layer



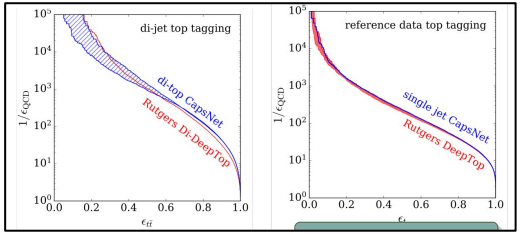
Network learns:
- 2 eyes
- 1 nose
- 1 mouth
- 2 ears
- No relative positioning

Filter of ~size of facial features

- **Solution:** Convolutional capsule networks
  - Capsules describe features (e.g. eyes, nose…)
  - Correlations between capsules gives a face
  - Work well in dense environments

8

- Event wide tagging with Capsule Networks
- Model is able to use correlations from full event



[Andreassen]

Reweighting via classification for MC tuning

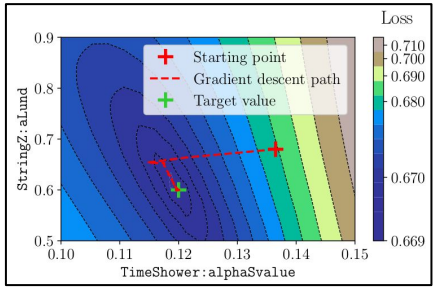
DCTR Summary:

A deep neural network classifier to estimate the likelihood ratio

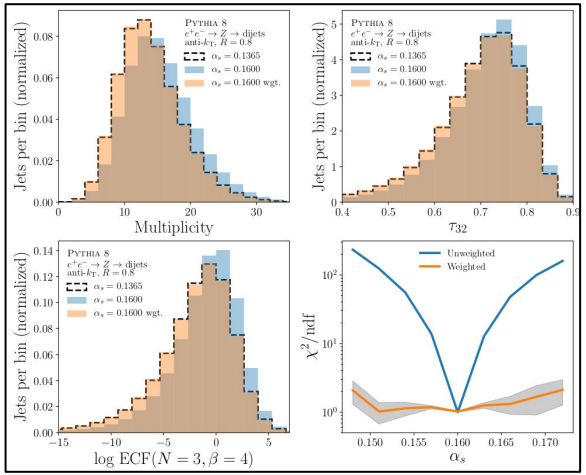
$$w(x,\theta) = p_1(x,\theta)/p_0(x,\theta)$$

as a continuous function of any MC parameter θ using the full phase space $x \in \Omega$

Use DCTR to :
- reweight one simulation to another
- tune MC to match an unknown sample, e.g. real data



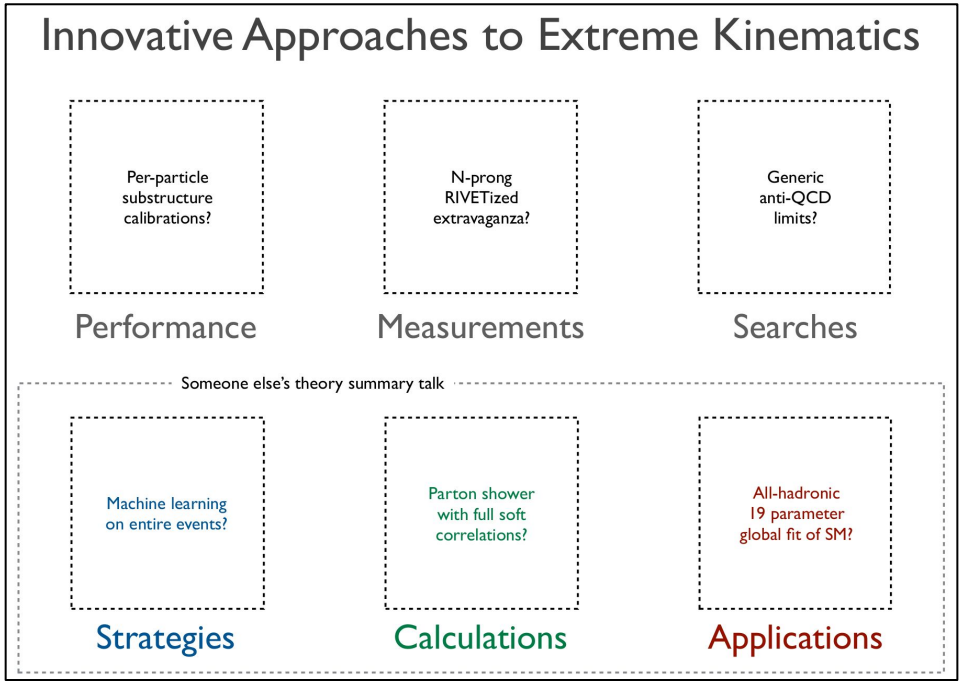
- Is it possible to reweight an event through a NN
- Use high-dimensional classifiers to reweight phase space and identify the best parameters for describing data
- Both continuous or discrete reweighting of Monte Carlo



Precise full phase-space reweighting has a wide number of applications for simulations, parameter tuning, and modeling of systematic uncertainties

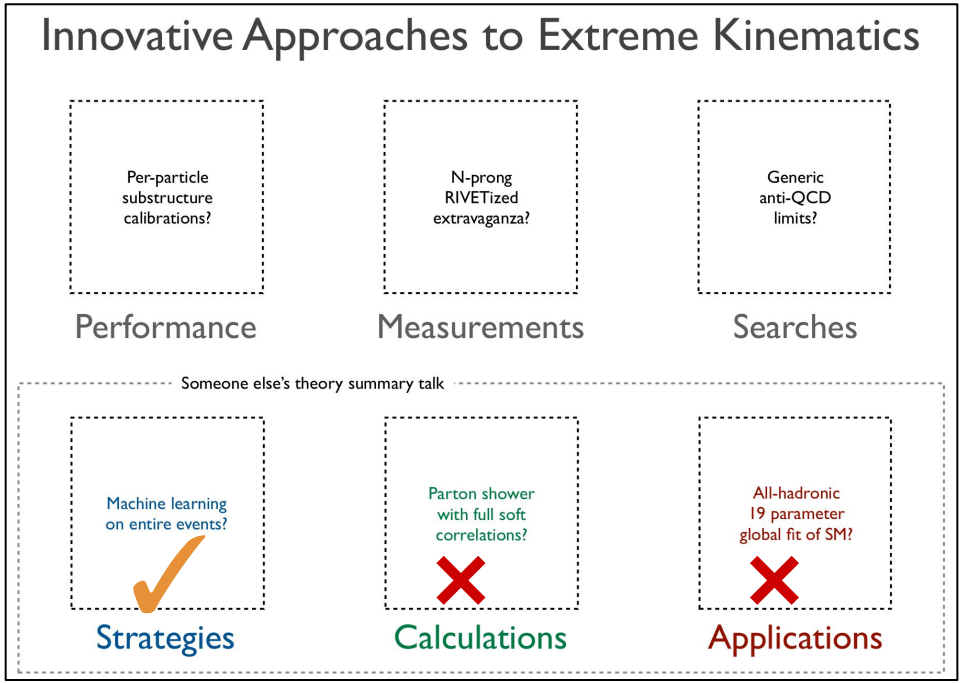
Looking back to ‘18

Jesse Thaler’s hopes for this year:



Looking back to ‘18

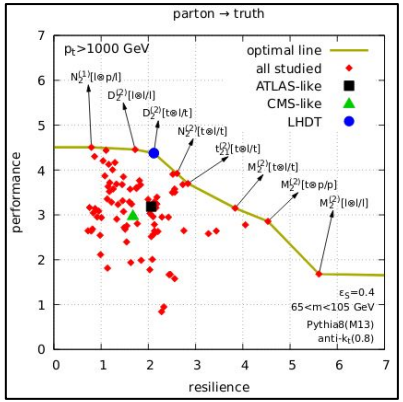
Jesse Thaler’s hopes for this year:



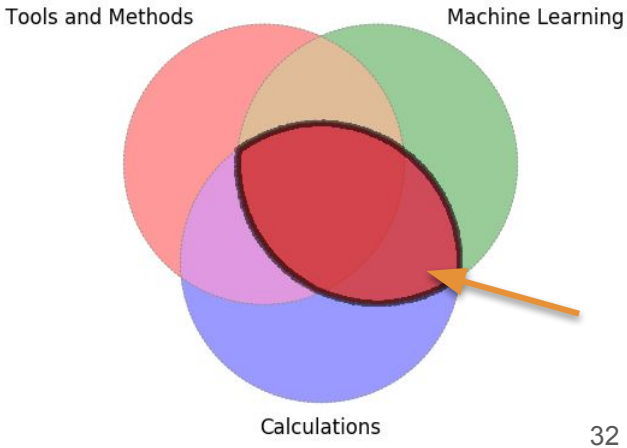
There is a lot of opportunities for progress left for BOOST 2020!

My hopes for 2020

- Can we come up with quantifiable metrics beyond performance for comparisons of different ML algorithms? Different metrics for different applications?
- Is it possible to find ways to leverage performance gains from ML methods in calculable and robust frameworks?
- Can we perform precision calculations for other key jet substructure observables? And compare these calculations with measurements?
- Is it possible to improve non-perturbative modelling, e.g. through improvements of perturbative component of parton showers?



[arXiv:1803.07977]



Conclusions

- Machine learning is here to stay.
- If a problem can be framed in the “right” way, ML can lead to real insights - but it should be primarily viewed as tool!
- Precision calculations of jet substructure observables will be critical for future measurements, notably for α_s extractions
- See you in Hamburg!