Expanding the HEP frontier with boosted b tags and the QCD power spectrum

Keith Pedersen
(kpeders1@hawk.iit.edu)

ILLINOIS INSTITUTE
OF TECHNOLOGY

Doctoral defense (10 Apr 2018)

Advisor: **Zack Sullivan**

Slides at `www.hepguy.com` (thesis coming soon)

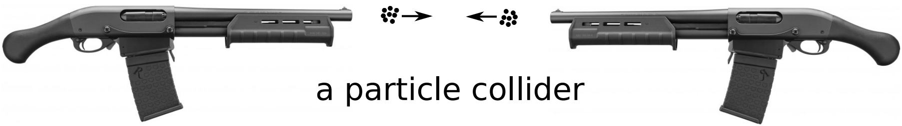
Outline

1. **Challenges of the high energy physics (HEP) frontier**
   - The search for new physics
   - Collider basics
   - High energy and high luminosity

2. **The μ_x boosted-bottom jet tag**
   - Boosted bottom hadron signatures
   - A scale-free observable
   - Extending the LHC's reach with the μ_x tag

3. **The QCD power spectrum**
   - The potential of correlated information
   - The power spectrum H_l
   - Discrete sampling noise
   - Power jets and pileup

What is matter, and how does it work?



a particle collider

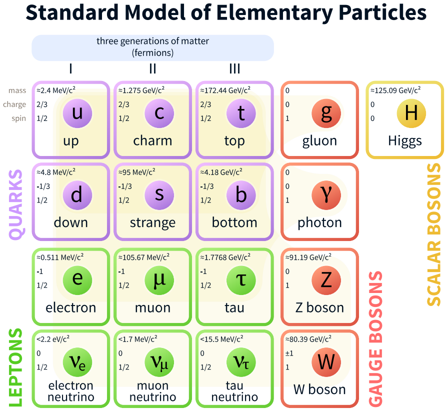
Scattering jargon

σ = scattering cross section

L = collider luminosity

$$L_{\text{int}} = \int L \, \mathrm{d}t$$

$$\mathrm{Ex}\,(\text{collisions}) = \sigma(\text{cm}^2)\; L_{\text{int}}(\text{cm}^{-2})$$

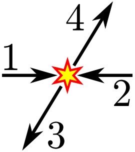


Standard Model of Elementary Particles

three generations of matter (fermions)

	I	II	III		
mass charge spin	≈2.4 MeV/c² 2/3 1/2 u up	≈1.275 GeV/c² 2/3 1/2 c charm	≈172.44 GeV/c² 2/3 1/2 t top	0 0 1 g gluon	≈125.09 GeV/c² 0 0 H Higgs
QUARKS	≈4.8 MeV/c² -1/3 1/2 d down	≈95 MeV/c² -1/3 1/2 s strange	≈4.18 GeV/c² -1/3 1/2 b bottom	0 0 1 γ photon	SCALAR BOSONS
	≈0.511 MeV/c² -1 1/2 e electron	≈105.67 MeV/c² -1 1/2 μ muon	≈1.7768 GeV/c² -1 1/2 τ tau	≈91.19 GeV/c² 0 1 Z Z boson	
LEPTONS	<2.2 eV/c² 0 1/2 ν_e electron neutrino	<1.7 MeV/c² 0 1/2 ν_μ muon neutrino	<15.5 MeV/c² 0 1/2 ν_τ tau neutrino	≈80.39 GeV/c² ±1 1 W W boson	
				GAUGE BOSONS	

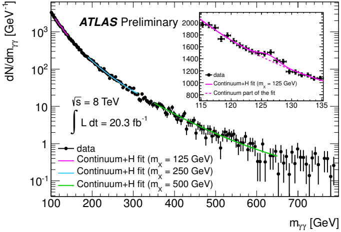
Rise of the 750 GeV excess



$$p_1 p_2 \to X \to p_3 p_4$$

$$m_X = \sqrt{(p_3 + p_4)^2}$$

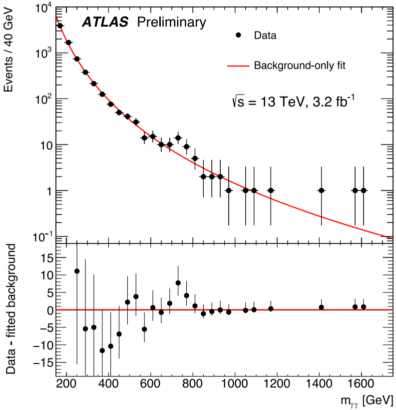
At $\sqrt{S} = 8\,\mathbf{TeV}$, the LHC saw the Higgs boson at $m_{\gamma\gamma} = 125$ GeV



Nucl.Part.Phys.Proc. **273–275** (2016) 2460-2462.

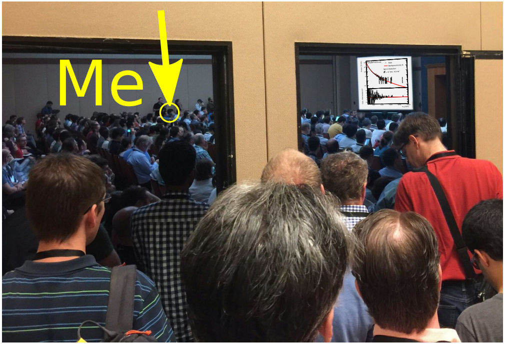
In 2015, first data at $\sqrt{S} = 13\,\mathbf{TeV}$ saw excess in $m_{\gamma\gamma} \approx 750$ GeV.

A new particle?

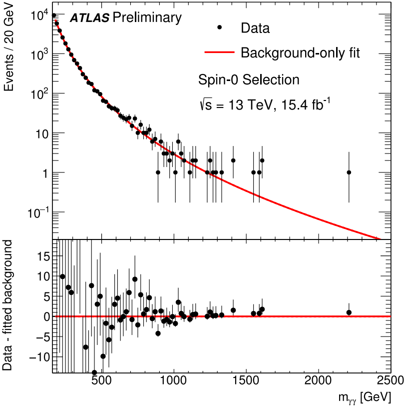


JHEP **1609** (2016) 001

Fall of the 750 GeV excess



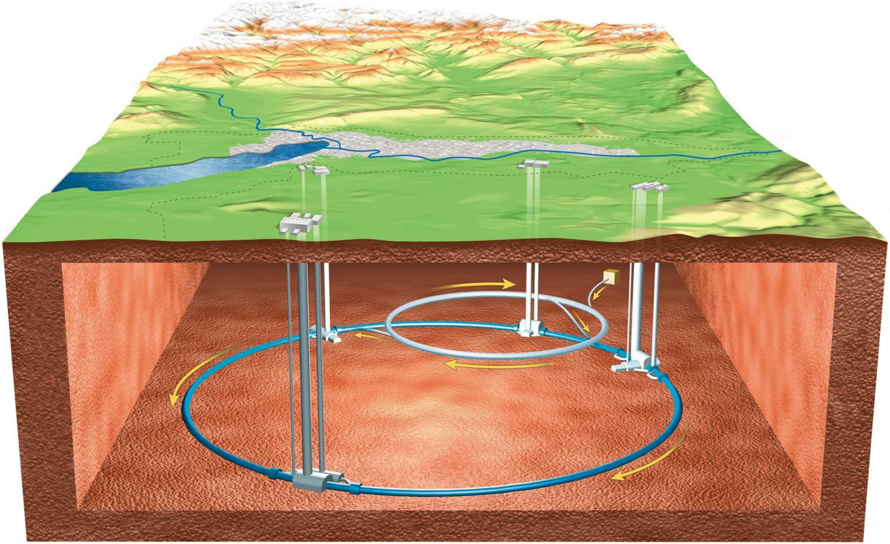
On Aug 5, 2016, with 5 × **more** data at $\sqrt{S} = 13$ TeV, the bump was gone.



The high-energy and high-luminosity frontier will face harder problems than statistical anomalies!.

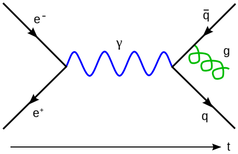
The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC)



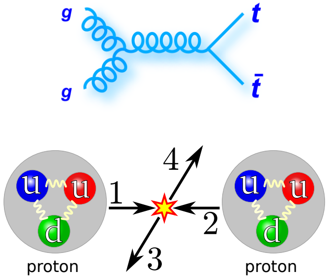
Collider energy and invariant mass

electron-positron collider



- $e^+e^- \to q\bar{q}g$
- $\boldsymbol{P}_1 = E_{\text{beam}}[1,\, +\hat{z}]$
- $\sqrt{S} = \sqrt{(\boldsymbol{P}_1 + \boldsymbol{P}_2)^2} = 2E_{\text{beam}}$

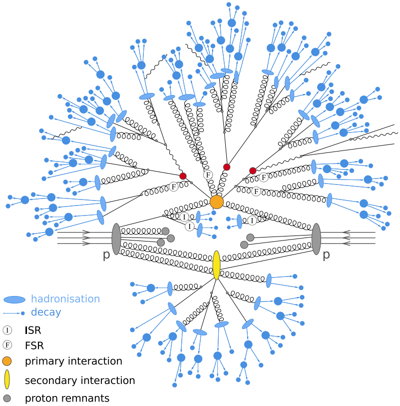
A proton collider is really a parton collider

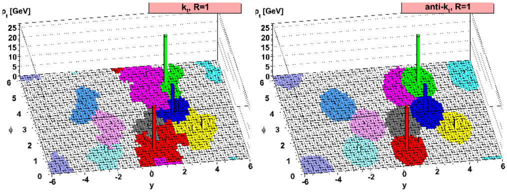


- not $pp \to q\bar{q}$ but:

  $q\bar{q} \to q'\bar{q}'$, $qg \to qg$, $gg \to q\bar{q}$
- $\boldsymbol{p}_1 = x_1 \boldsymbol{P}_1$
- $\sqrt{s} = \sqrt{(\boldsymbol{p}_1 + \boldsymbol{p}_2)^2} = 2\sqrt{x_1 x_2}\, E_{\text{beam}}$

New physics is wrapped in QCD





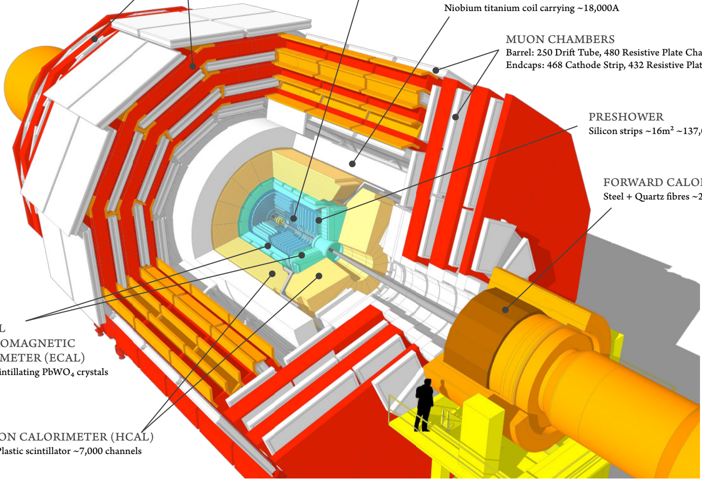
QCD has **asymptotic freedom**; hard scatter ↦ busy final state.

- Initial-state radiation.
- Final-state radiation from quarks/gluons creates **jets**;
  - **jet-parton duality**
- Confinement . . . colored particles must hadronize.

Reconstructing quark/gluon jets requires a **jet definition**.

- k_T jets *rewind* QCD shower.
- anti-k_T less sensitive to soft physics.

An LHC detector



CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m^2 ~66M channels
Microstrips (80x180 μm) ~200m^2 ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m^2 ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

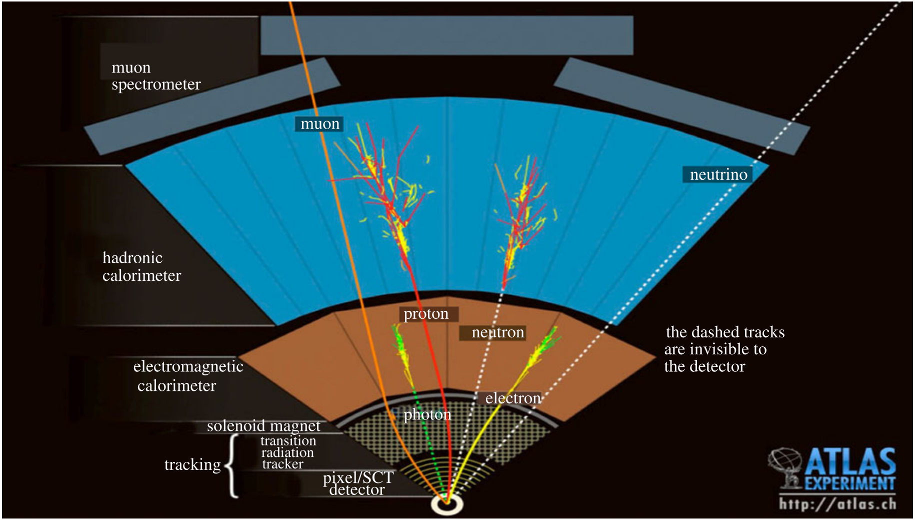
CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76,000 scintillating $PbWO_4$ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

What a detector sees

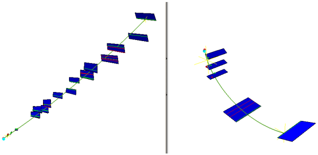
charged ↦ **tracks**

neutral ↦ **towers**

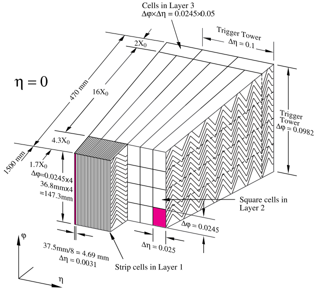


Physics objects: tracks and towers

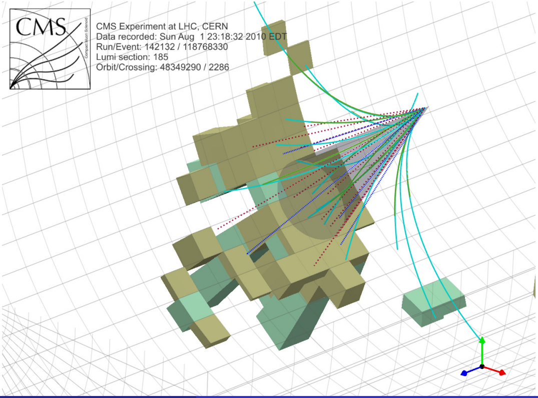
tracks



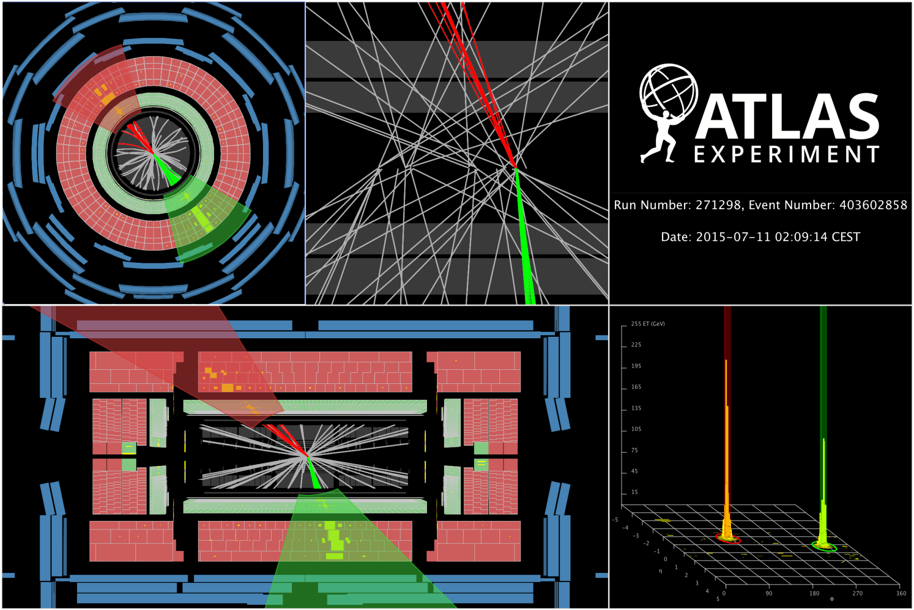
towers



- **Neutral tower**: track energy subtracted from tower that was struck.
- **massless** tracks and neutral towers are clustered into **massive** jets.



A two-jet event with $\sqrt{s} = 3.25$ TeV



Run Number: 271298, Event Number: 403602858

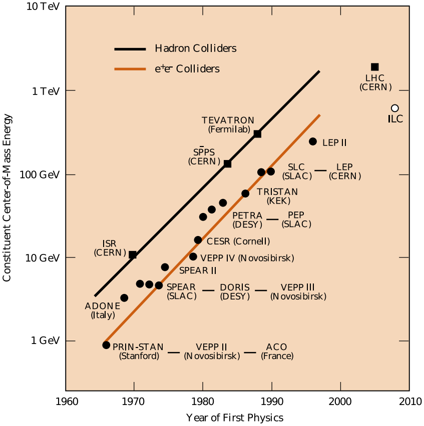
Date: 2015-07-11 02:09:14 CEST

New physics is rare physics

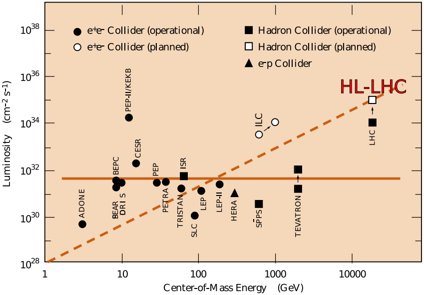
How to find new physics:

① Increase collider energy $\sqrt{S}$.

② Increase luminosity L_{int}.



- Higher $\sqrt{S}$ unlocks new physics.
- Higher L creates more **events** (better stats, more precision).

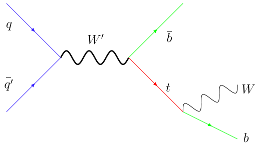


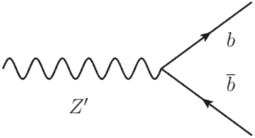
Caveats:

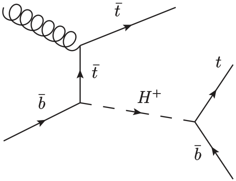
- High-E jets are harder to tag.
- More collisions ↦ more pileup.

The need for a boosted-bottom jet definition

A bottom quark jet (*b* jet) is useful for finding new physics

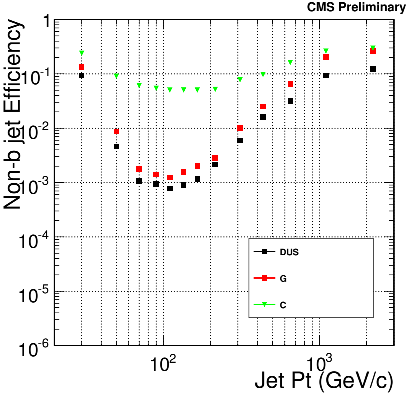






Maintaining 50% *b* jet efficiency

CMS-PAS-BTV-09-001

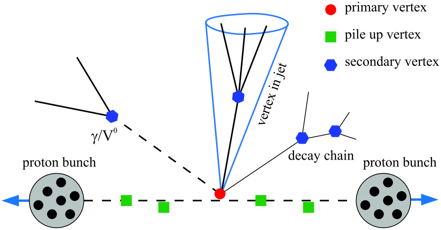


We need a more robust tag for boosted *b* jets!

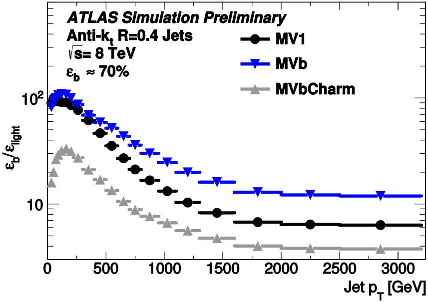
Outline

1. Challenges of the high energy physics (HEP) frontier
    - The search for new physics
    - Collider basics
    - High energy and high luminosity
2. The μ_x boosted-bottom jet tag
    - Boosted bottom hadron signatures
    - A scale-free observable
    - Extending the LHC's reach with the μ_x tag
3. The QCD power spectrum
    - The potential of correlated information
    - The power spectrum H_l
    - Discrete sampling noise
    - Power jets and pileup

Light jets dominate boosted track-vertex tags



- b jets contain a b hadron:
  $$b \to (B^0/B^\pm/\Lambda_b) \to X$$
- Tracks point to b hadron decay.
- Faked by light jets ($u\,d\,s\,g$).

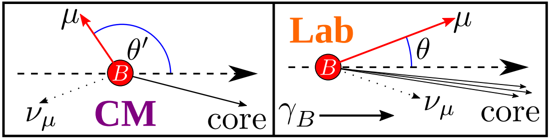


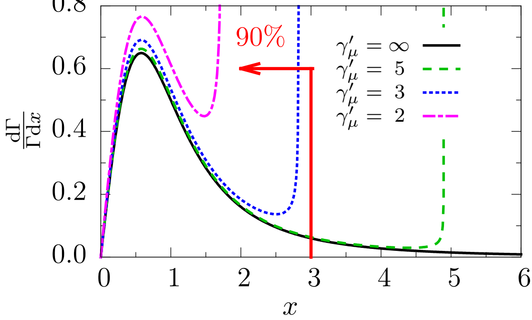
ATL-PHYS-PUB-2014-014

The μ_x boosted-b tag

$$\boxed{B^0 \to \mu^+ \nu_\mu X \approx 10\%}$$

Phys. Rev. D **93**, 014014 (2016) [KP/ZS]





$$p_{\text{subjet}} = p_\mu + p_\nu + p_{\text{core}}$$

1. Muon angle

$$x \equiv \gamma_{\text{B}} \tan\theta = \frac{\sin\theta'}{\kappa + \cos\theta'}$$

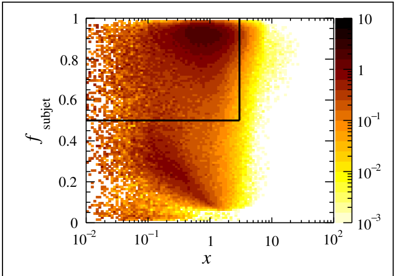
$$\kappa = \beta'_\mu / \beta_{\text{B}} \to 1$$

2. Subjet hardness

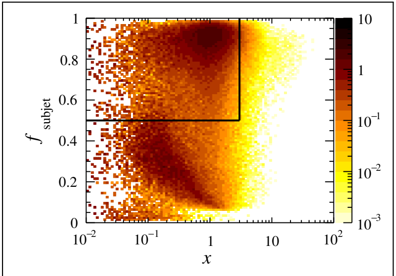
$$f_{\text{subjet}} \equiv \frac{p_T^{\text{subjet}}}{p_T^{\text{jet}}}$$

$$\boxed{x \leq 3 \quad \text{and} \quad f_{\text{subjet}} \geq 0.5}$$

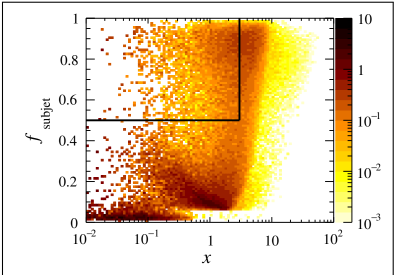
μ_x tag for different flavors



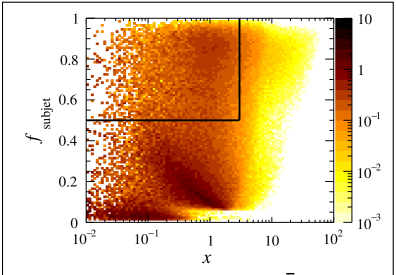
bottom



charm



light-**light**



light-**heavy** ($g \to b\bar{b}$)

Predicted μ_x boosted-b tag efficiencies at the LHC

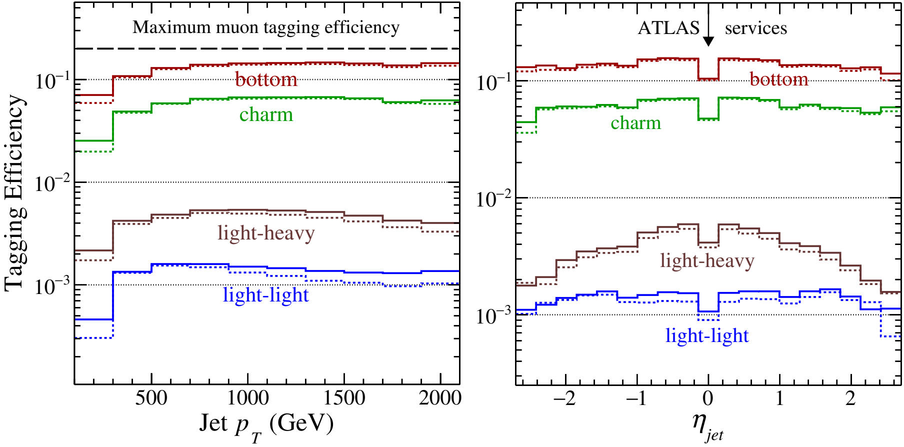
Signal efficiencies

- $\sim$14% of b-jets
- $\sim$6.5% of c-jets

Light jet **fake rate**

- Light-light $\mathcal{O}(0.1\%)$
- All light $\mathcal{O}(0.6\%)$

Robust to pileup!



μ_x tag applied to a Z' model

Many extensions of the Standard Model (SM) predict heavy, narrow particles which couple via a **vector current** ... the W' and Z'

- Sequential Standard Model
- broken $SU(2)_L \times SU(2)_R$
- Kaluza-Klein from extra dimensions
- GUT models
- non-commuting extended technicolor
- and more ...

$Z' \to l^+ l^-$ ... clean reconstruction, low BG.

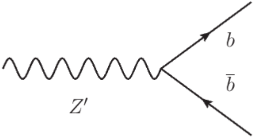
- Excluded below 2.9 TeV with SM-like coupling (ATLAS/CMS).

Is new physics *leptophobic* (no coupling to leptons)?

$$\boxed{\mathcal{L} \supset \frac{g_B}{6} Z'_{B,\mu} \bar{q} \gamma^\mu q}$$

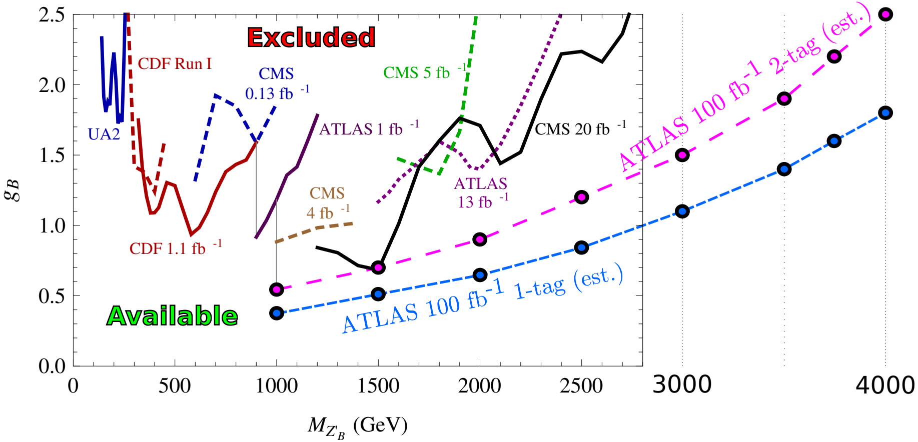
$U(1)'$ from Dobrescu & Yu

(Phys. Rev. D **88** (2013) 035021, 1506.04435)



Predicted reach at the LHC

MadGraph5 → PYTHIA 8 → Delphes 3 (w/ `FastJet` 3) at $\sqrt{S} = 13$ TeV

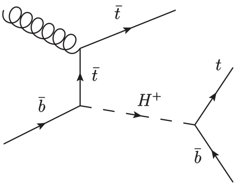


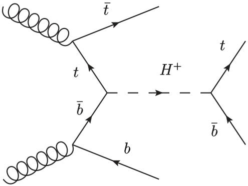
The μ_x tag greatly improves the LHC's reach for heavy, leptophobic Z'.

Searches for a charged Higgs boson

Supersymmetry (SUSY) requires extra Higgs doublets. We examined a type-II two Higgs doublet model (2HDM): Φ_1 and Φ_2

- $\tan\beta \equiv v_2/v_1$
- 5 bosons: $(h,\ H,\ \text{and}\ H^{\pm})$ and (A)
- Measurements restrict $h_{SM} \to h$.
- Degenerate heavy Higgs' mass ensures $H^{\pm}$ only couples to t and b quarks!

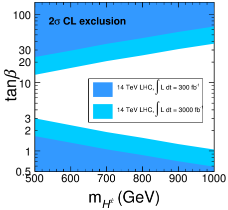




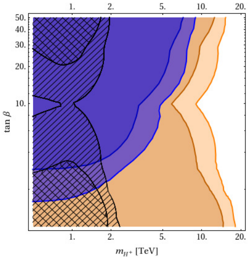
μ_x boosted b and boosted hadronic top

Prior predictions disagree

JHEP **1506** (2015) 137 **(14 TeV)**



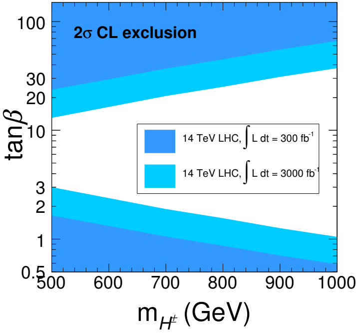
JHEP **1511** (2015) 124 **(14 & 100 TeV)**

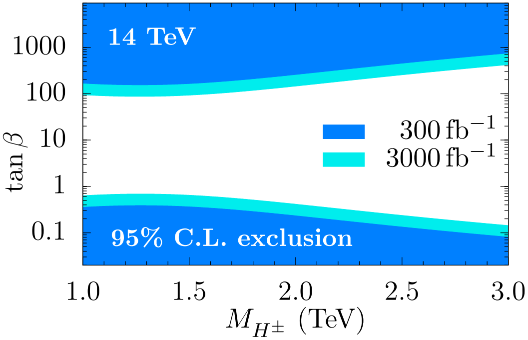


Extending the mass reach at 14 TeV

Phys. Rev. D **95** (2016) 035037 [KP/ZS]

- Our predictions **extend** Craig et al. above $m_{H^\pm} = 1\ TeV$.
- The 14 TeV cross section is too small to plug “the wedge”!



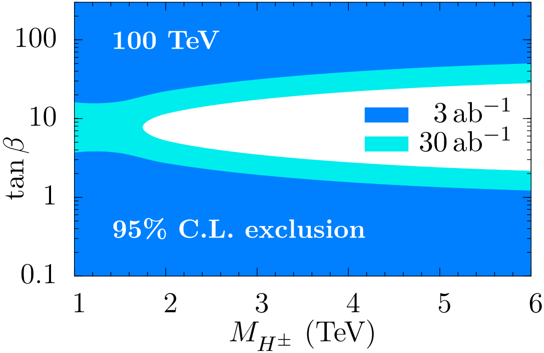


Do we need a 100 **TeV collider?**

A case for 100 TeV

Nearly independent of model, 100 TeV collider can see $H^{\pm}$.

- The existence of $H^{\pm}$ maps to the existence of 2HDM.
- The existence of 2HDM maps to the existence of SUSY.
- We may need a 100 TeV collider for a final ruling on SUSY.



A new tool for experiments

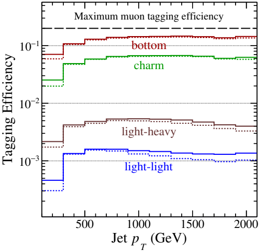
μ_x boosted bottom jets:

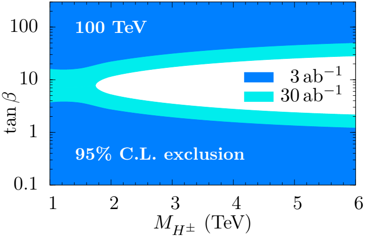
- Well motivated by simple kinematics.
- Maintains its signal/noise ratio in the TeV regime.
- Exhibits robust performance in Monte Carlo.

We developed a custom module (`HighPtBTagger`) to implement μ_x tagging in a popular detector simulator (`Delphes`), available on GitHub:

https://github.com/keith-pedersen/delphes/tree/HighPtBTagger_devel

---





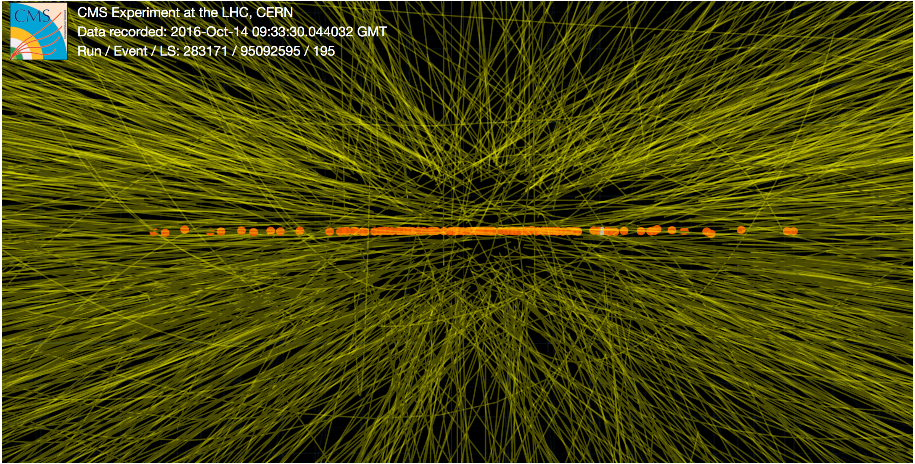
Outline

1. **Challenges of the high energy physics (HEP) frontier**
   - The search for new physics
   - Collider basics
   - High energy and high luminosity

2. **The μ_x boosted-bottom jet tag**
   - Boosted bottom hadron signatures
   - A scale-free observable
   - Extending the LHC's reach with the μ_x tag

3. **The QCD power spectrum**
   - The potential of correlated information
   - The power spectrum H_l
   - Discrete sampling noise
   - Power jets and pileup

Pileup is here to stay



- Most pileup from other vertex — **charged** pileup is largely reducible.
- The LHC is currently averaging 40 pileup events per hard scatter!
- The HL-LHC is expected to average over 100!

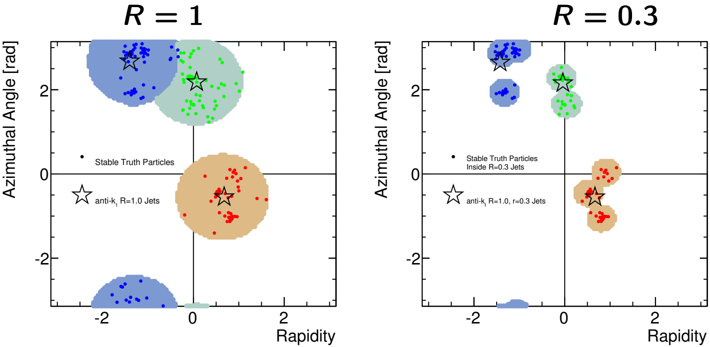
Pileup in anti-k_T jets

anti-k_T jets use one correlation at a time; find the smallest “distance”

$$d_i^2 = p_{T,i}^{-2}$$ cluster becomes a jet

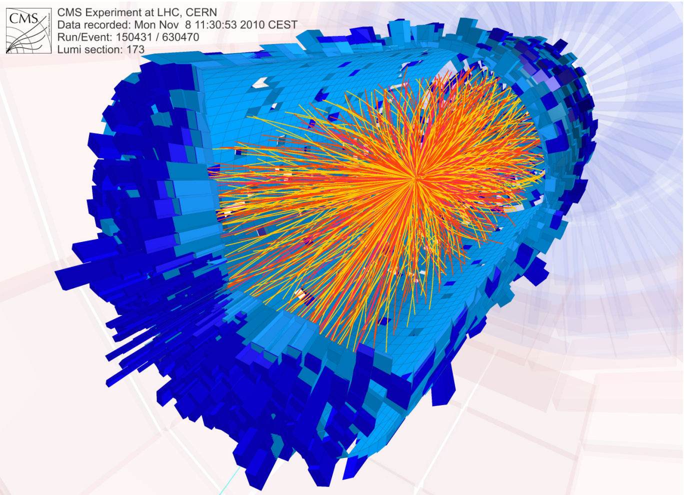
$$d_{ij}^2 = \min(p_{T,i}^{-2}, p_{T,j}^{-2})\frac{\Delta y_{ij}^2 + \Delta\phi_{ij}^2}{R^2}$$ merge two clusters

---

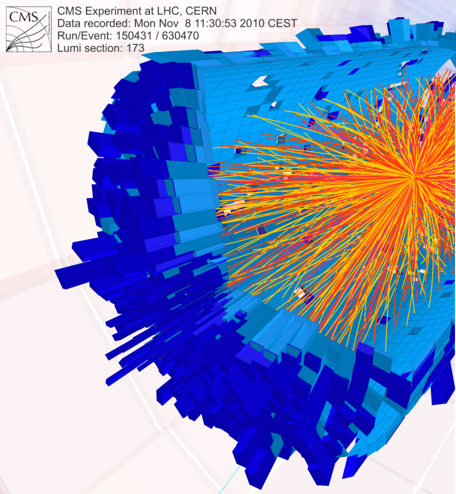


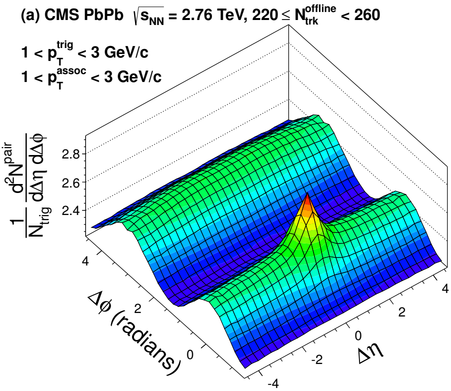
Eur.Phys.J. **C76** (2016) 581

Learning from heavy-ion collisions



Learning from heavy-ion collisions



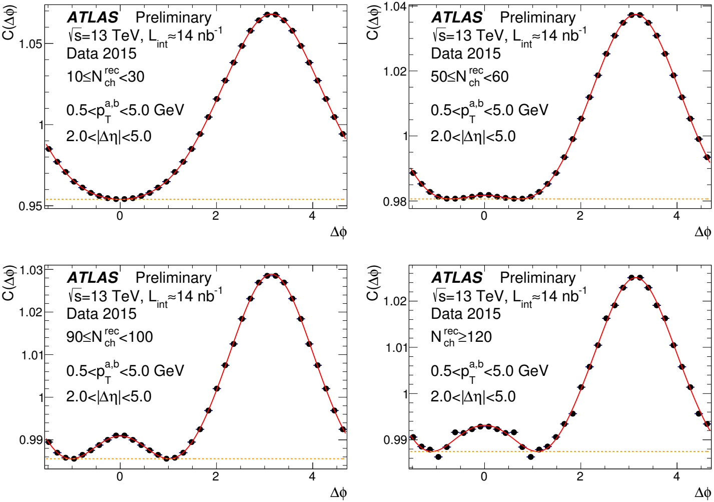


Phys.Lett. **B724** (2013) 213–240

The same-side ridge is attributed to quark-gluon plasma

Connecting lead-lead to proton-proton

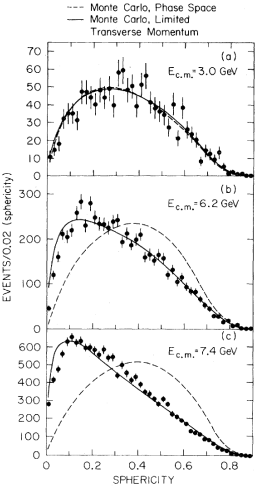
ATLAS-CONF-2015-027



Why is the same-side correlation seen in high-multiplicity *pp* collisions?

Event shape variables

Sphericity



Phys.Rev.Lett. **35** (1975) 1609–1612

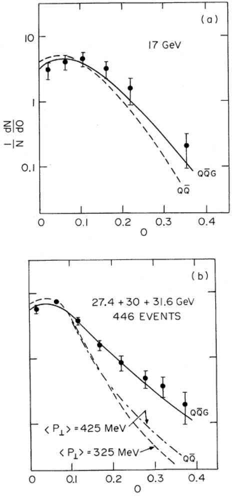
$\Longleftarrow$ 2-jet structure
$(e^+e^- \to q\bar{q})$;
first seen with Sphericity.

3-jet structure $\Longrightarrow$
$(e^+e^- \to q\bar{q}g)$;
first seen with Oblateness.

Event shape variables:

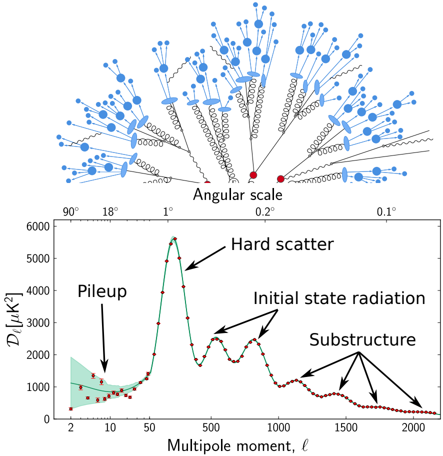
- Condense each event to a single number.
- Shape curves from many events.

Oblateness

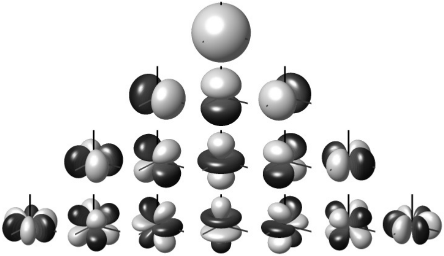


Phys.Rev.Lett. **43** (1979) 830

The power spectrum of QCD radiation



$$S_l \equiv \sum_{m=-l}^{l} |E_l^m|^2$$



Spherical harmonics $Y_l^m(\theta,\phi)$

- degree l — number of lobes.
- order m — lobe orientation.

$$E(\hat{r}) = \sum_i E_i\,\delta(\hat{r}-\hat{p}_i)$$

$$E_l^m = \int_\Omega \mathrm{d}\Omega\, Y_l^{m*}(\hat{r})\,E(\hat{r})\,.$$

Rotational invariance and energy scale

$$S_l \equiv \sum_{m=-l}^{l} |E_l^m|^2 = \frac{2l+1}{4\pi}\int_{\Omega,\,\Omega'} d\Omega\, d\Omega' E(\hat{r})E(\hat{r}')P_l(\cos\xi)$$

A dimensionless power spectrum scales out $E_{\text{tot}} = \int_\Omega d\Omega E(\hat{r}) = \sum_i E_i$

$$H_l \equiv \frac{1}{2l+1}\frac{S_l}{S_0} = \frac{1}{4\pi}\int_{\Omega,\,\Omega'} d\Omega\, d\Omega' \rho(\hat{r})\rho(\hat{r}')P_l(\cos\xi)$$

$\rho(\hat{r}) = \sum_i f_i\, \delta(\hat{r} - \hat{p}_i)$	$f_i \equiv \frac{E_i}{E_{\text{tot}}}$	$\hat{p}_i \equiv \frac{\vec{p}_i}{\lvert\vec{p}_i\rvert}$
event shape	energy fraction	direction of travel

★ $$H_l = \langle f|\, P_l\Big(|\hat{p}\rangle \cdot \langle\hat{p}|\Big)\, |f\rangle = f_i\, P_l(\cos\xi_{ij}) f_j$$

$H_0 = 1$

$0 \le H_l \le 1$

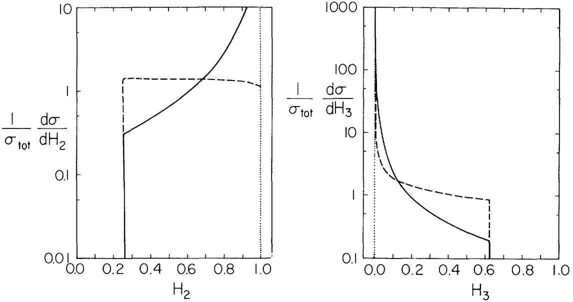
$\xi_{\text{res}} = \frac{2\pi}{l}$

The Fox-Wolfram distributions $f(H_l)$

Fox and Wolfram defined H_l to differentiate two final states:

- $e^+e^- \to \gamma \to q\bar{q}g$ . . . . generic QCD.
- $e^+e^- \to X \to ggg$ . . . . a new, heavy resonance.

Phys. Rev. Lett 41 (1978) 1581



$q\bar{q}$ (dotted), $q\bar{q}g$ (solid), $X \to ggg$ (dashed)

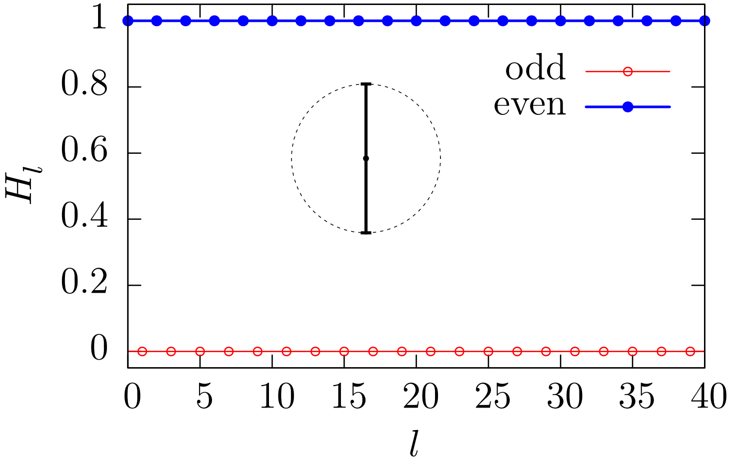
Unsuccessful in 1978:

- Too few particles at a lepton collider.
- Variable conditions at a hadron collider:
  - Boost y_{cm}
  - Scale $Q = \sqrt{s}$
- Fox-Wolfram $f(H_l)$ are not independent.

Skip the middle man!

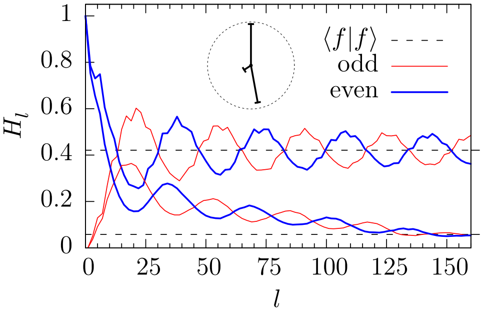
Every 2-particle event

A two-particle event in the CM frame is not just H_2.



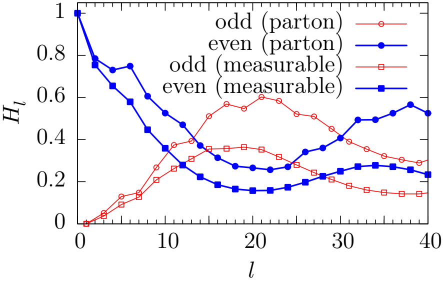
$$(\delta(\hat{r}+\hat{z})+\delta(\hat{r}-\hat{z})) \neq Y_2^0(\hat{r})$$

A 2-jet-like event

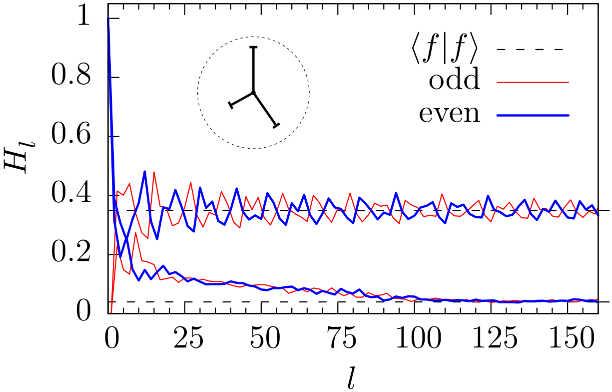


- No broad CMB-like shapes!
- oscillation ↦ correlation; Fox-Wolfram $f(H_l)$ are not independent, can't focus on just a few.
- $H_l \sim \langle f|f\rangle$

- measureable ↦ partons at low l, but stabilize to lower power.
- H_2 is large.
- H_3 is small.



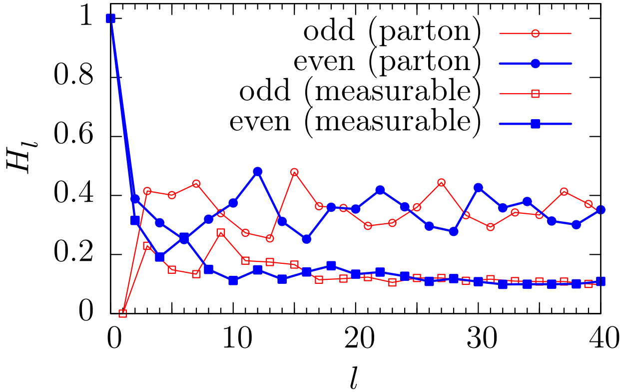
A 3-jet-like event



Important features

- ***H_l* rapidly oscillates**: ~~CMB~~
- ***H_l* is unending**: $H_l \sim \langle f|f\rangle$
- ***$N \neq n$***: N measurable particles don't match n original partons.

A 3-jet-like event



Important features

- **H_l rapidly oscillates**: ~~CMB~~
- **H_l is unending**: $H_l \sim \langle f | f \rangle$
- **$N \neq n$**: N measurable particles don't match n original partons.

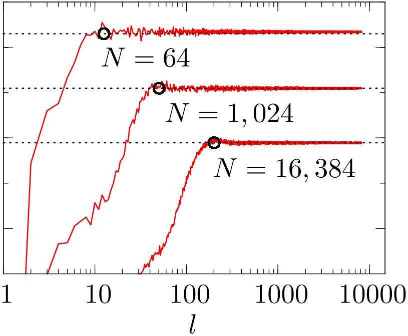
The multiplicity plateau and sampling noise

$$H_l = \langle f|\, P_l\Big(\, |\hat{p}\rangle \cdot \langle \hat{p}| \,\Big)\, |f\rangle$$
$$= \langle f|f\rangle + (\text{inter-particle})$$

$$\xi_{\text{res}} = \frac{2\pi}{l}$$

$$\langle f|f\rangle \propto \frac{1}{N}\,; \quad \langle f|f\rangle \geq \frac{1}{N}$$

Random isotropic particles ($\rho(\hat{r}) = \frac{1}{4\pi}$)



Sample multiplicity limits useful information — discrete sampling noise!

$$f(k) = \int_{-\infty}^{\infty} \delta(x)\, e^{-2\pi k\, x} \mathrm{d}x = 1$$

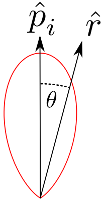
We need a low-pass filter for angular correlations; make particles extensive.

The extensive particle

Event shape $\rho(\hat{r})$ from particle shape functions:

$$\xcancel{\rho(\hat{r}) = \sum_i f_i \, \delta(\hat{r} - \hat{p}_i)}$$

$$\rho(\hat{r}) = \sum_i f_i \, h_i(\hat{r})$$

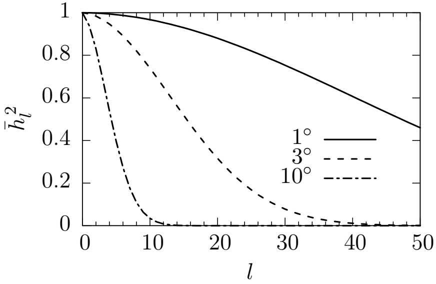


pseudo-Gaussian
in polar angle θ:

$$h(\theta) \approx C \, \exp\left(-\frac{\theta^2}{2\lambda^2}\right)$$

$$h_i(\hat{r}) = C \, \exp\left(-\frac{1 - \hat{r} \cdot \hat{p}_i}{\lambda^2}\right)$$

A low-pass filter



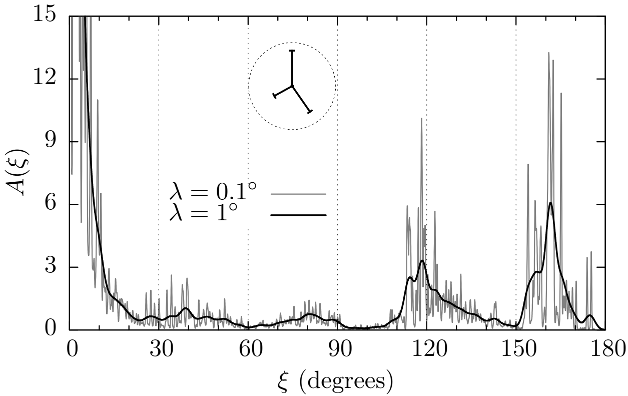
$$H_l = \bar{h}_l^2 \, H_l^{\delta\text{-particle}}$$

Larger λ discards *more* small-angle (high-l) information.

Infrared and collinear safety

$$\boxed{A(\xi) = \sum_l (2l+1)\, H_l\, P_l(\xi)}$$

$$h(\theta) \approx C \exp\left(-\frac{\theta^2}{2\lambda^2}\right)$$



The power spectrum at low-l is stable to soft or collinear radiation.

The power jets fit

observable power spectrum

$$\rho(\hat{r})_{\text{obs}} = \sum_{i=1}^{N} f_i\, h_i(\hat{r})$$

$$\Downarrow$$

$$H_l^{\text{obs}}$$

$n \ll N$

$$\chi_l = H_l^{\text{fit}} - H_l^{\text{obs}}$$

n-prong power spectrum

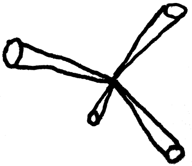
$$\rho(\hat{r})_{\text{fit}} = \sum_{j=1}^{n} f_j\, h_{(j)}(\hat{r})$$

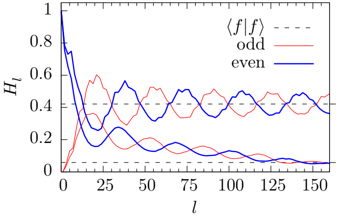
$$\Downarrow$$

$$H_l^{\text{fit}}$$

To see QCD radiation and **extract jet-like structures**. . .

- Reproduce H_l^{obs} of N tracks and towers by fitting an n-“prong” model



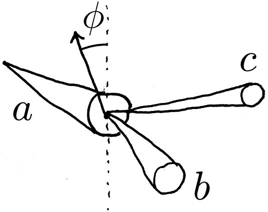


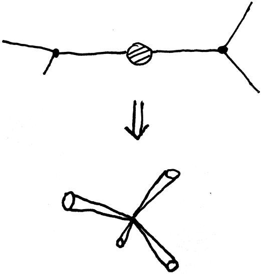
The power jets model

Describe hard QCD radiation with a binary splitting tree ($a \to b\,c$).

$$\boldsymbol{p}_a = \boldsymbol{p}_b + \boldsymbol{p}_c$$

Four degrees of freedom per splitting node ($\boldsymbol{p}_b = [E_b,\, \vec{p}_b\,]$).

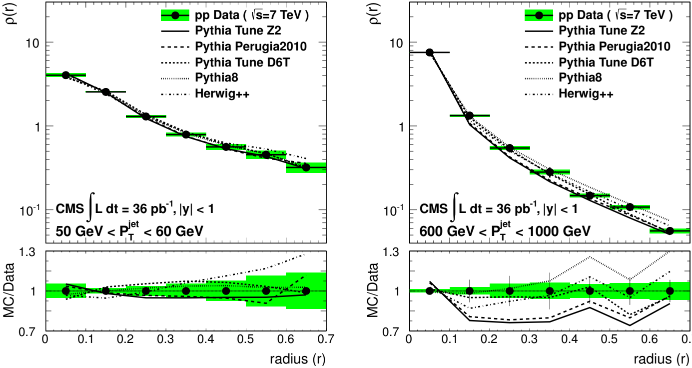


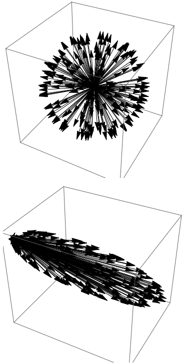


Prong shape $h_{(j)}(\hat{r})$ needs physical basis (not pseudo-Gaussian a priori).

Jets are shaped by their boost

JHEP **1206** (2012) 160



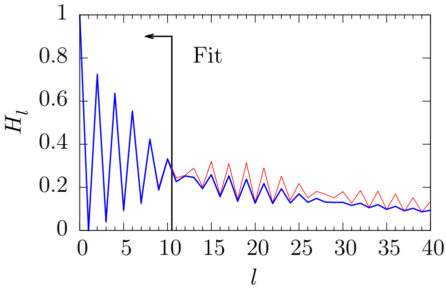


An energetic jet is boosted centrally; the most basic jet shape:

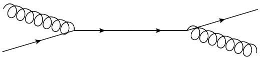
- Assume scalar decay ($J = 0$) in the rest frame of the QCD shower.
- Boost the scalar shape into the lab frame (mass ↦ shape).

Fitting a 2-jet-like event

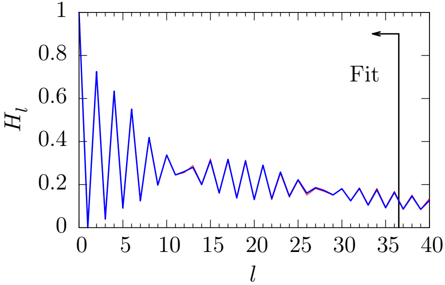
3-prong

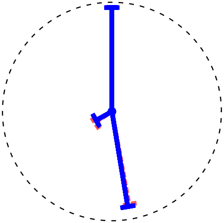


The 3-prong fit fails for $l > 10$;
another prong is needed.

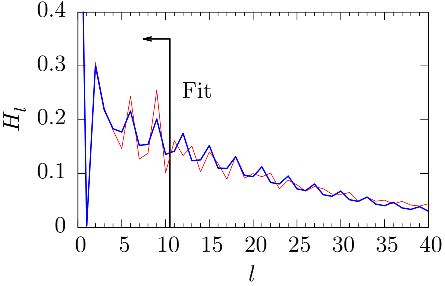


4-prong

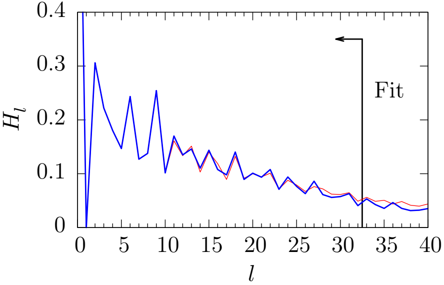




Fitting a 3-jet-like event

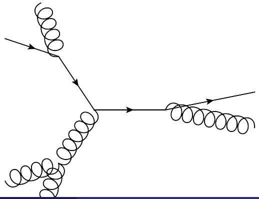


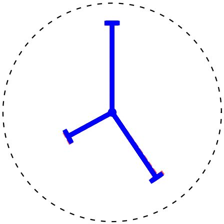
3-prong



6-prong

For a 3-jet-like event — 6 prongs:





Power jets provide superb reconstruction

Table : Reconstructed 3-jet kinematics for the 2-jet-like event.

(GeV)	E_1	E_2	E_3
parton	190.1	172.8	37.00
power jets	190.4(0)	174.2(1)	35.52(8)
error	0.1%	0.7%	-4%

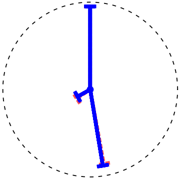
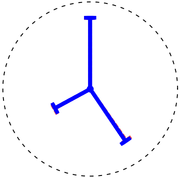


Table : Reconstructed 3-jet kinematics for the 3-jet-like event.

(GeV)	E_1	E_2	E_3
parton	163.0	143.5	93.56
power jets	162.0(1)	146.3(4)	91.68(4)
error	-0.6%	2.0%	-2.0%

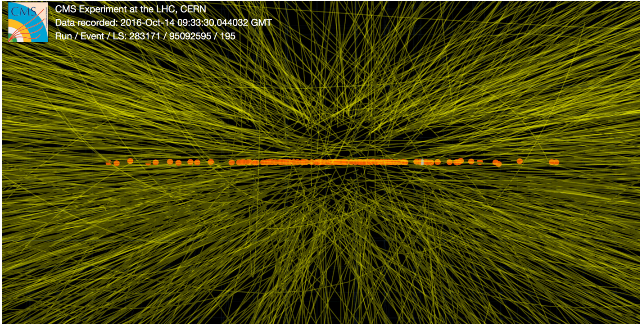


Pileup is just another shape

Add pileup to the event shape

$$\begin{aligned}\rho(\hat{r}) &= \rho(\hat{r})_{\text{hard}} + \rho(\hat{r})_{\text{PU}} \\ &= (1 - f_{\text{PU}}) \sum_j f_j\, h_{(j)}(\hat{r}) + f_{\text{PU}}\, h_{\text{PU}}(\hat{r})\end{aligned}$$

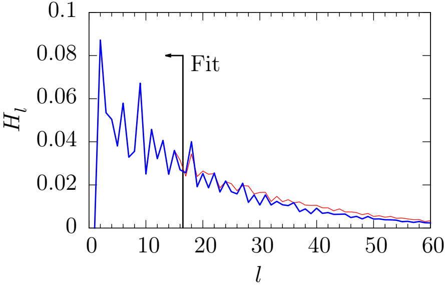
$h_{\text{PU}}(\hat{r})$ can **measured** in high-pileup events lacking hard scatter!

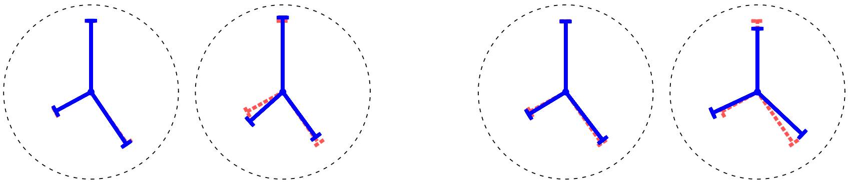


Using noise-noise correlations to see the signal

$S/N = 1 \quad (f_{\text{PU}} = 0.5)$

$S/N = 1/5 \quad (f_{\text{PU}} = 0.8)$





power jets

anti-k_T

power jets

anti-k_T

The promise of the QCD power spectrum

Power jets:

- Simultaneous fit to all detector information.
- Continuous shape function:
  - For tracks and towers: accounts for discrete sampling noise.
  - For n-prong model: permits $n \ll N$.
- From fully correlated QCD radiation spectrum:
  - Reconstruct jet energies and dijet invariant mass.
  - No boundaries (no *fixed* radius R) and no constituents.
  - Extremely robust to pileup because pileup shape can be **fit**!

The promise of the QCD power spectrum

Power jets:

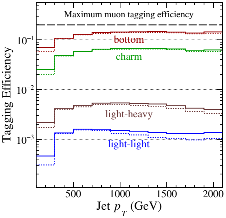
- Simultaneous fit to all detector information.
- Continuous shape function:
  - For tracks and towers: accounts for discrete sampling noise.
  - For n-prong model: permits $n \ll N$.
- From fully correlated QCD radiation spectrum:
  - Reconstruct jet energies and dijet invariant mass.
  - No boundaries (no *fixed* radius R) and no constituents.
  - Extremely robust to pileup because pileup shape can be **fit**!

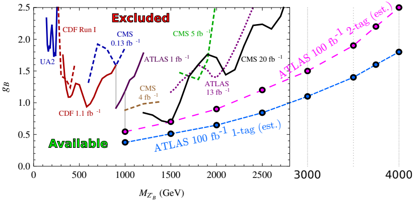
Next steps:

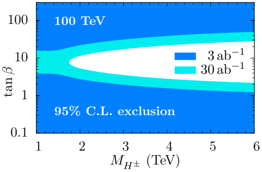
- Recover absolute orientation during fit.
- Incorporate longitudinal boost of proton collider (i.e., fit y_{cm}).
- Apply power jets to interesting physics:
  - Jet substructure (QCD shower).
  - Direct measurement of m_t from all hadronic final state.
  - How do we see long-distance (same-side) correlation?

Necessary tools for today and the future

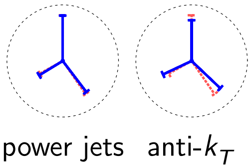
Use **more robust** information to tag b jets (muon angle)

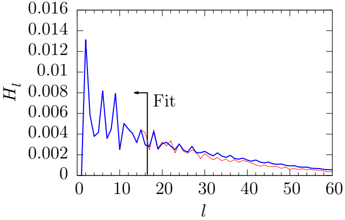






Use **all** information to reconstruct and study QCD





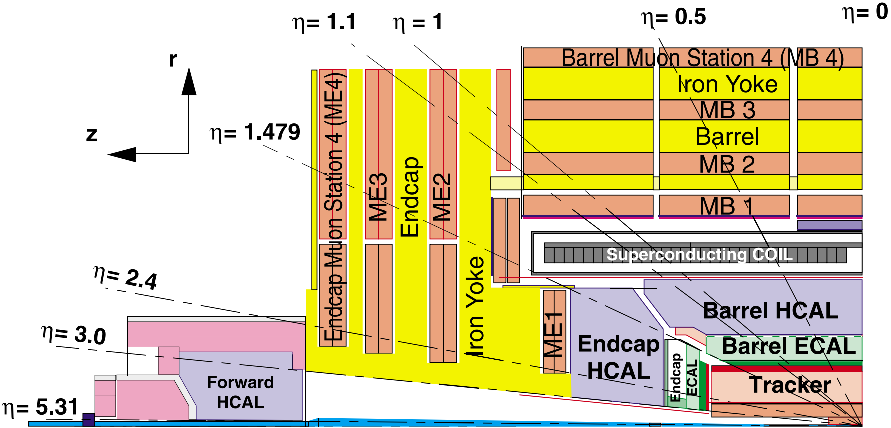
Thank you

Thank you for your attention!

The longitudinal boost of the CM frame

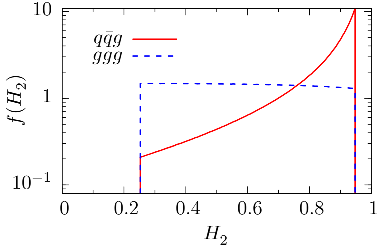
$y = \operatorname{arctanh}\frac{p_L}{E}$ $\quad$ $y_{\text{cm}} = \frac{x_1 - x_2}{x_1 + x_2}$ $\quad$ Δy is invariant to y_{cm}

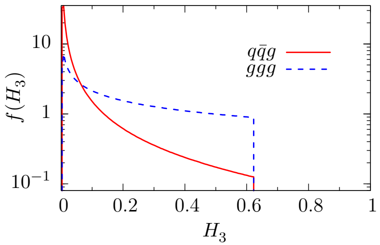
$\boxed{\eta = y \text{ for massless particles}}$

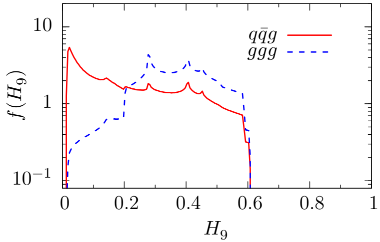


Fox-Wolfram distributions $f(H_l)$ don't work

$H_2 < 0.95$







$0.45 < H_2 < 0.5$