Harnessing the QCD power spectrum for high energy physics

Zack Sullivan[1,*] and *Keith* Pedersen[1]

[1]Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616-3793, USA

Abstract. We introduce a new framework to encode information in the QCD power spectrum on an event-by-event basis from which jets and their substructure can be extracted. Our new "power jets" scheme requires no jet radius parameter, and allows pileup to be extracted in situ. The result is an accurate kinematic reconstruction of underlying hard physics which is robust to extremely high pileup.

1 Introduction

High multiplicity events at the CERN Large Hadron Collider (LHC) present challenges to the observation of low-energy phenomena, while hinting at important long-range QCD phenomena that have been largely ignored by high energy physics. The high-luminosity LHC will see an average of 200 proton collisions per crossing, leading to extremely poor energy resolution for "low-energy" jets ($\sim$ 40 GeV) [1]. Interesting tails of QCD jet distributions are discarded by sequential jet algorithms, like anti-k_T [2], while pileup subtraction removes even more correlated low-energy information. Jets are helpful for high-statistics searches for new particles, but we want to be able to understand events on an event-by-event basis.

Jets are useful, but there is more to QCD at LHC than jet physics. While heavy-ion collisions expect to find long-range same-side particle correlations due to thermodynamic effects, recent high-multiplicity events in proton-proton collisions at LHC also find a same-side ridge [3]. This raises the intriguing question of just how much of the ion collision ridge is due to collisions forming a quark-gluon plasma, and how much is intrinsic to the underlying hadronic collisions. One of the goals of this research is to provide a fully consistent framework in which to address the question.

In a recent thesis [4] and forthcoming paper [5], we present a new framework to consistently define the energy-weighted QCD angular power spectrum on an event-by-event basis, and introduce a method called "power jets" with which we can extract jet-like phenomena with high precision in events with large pileup. In Sec. 2 we describe improvements to Fox-Wolfram moments needed to construct the QCD power spectrum, and discuss important limitations on the extraction of information from the spectrum. In Sec. 3 we describe the "power jet" method of extracting hard QCD phenomena from the power spectrum, and demonstrate that it is both accurate and stable under large backgrounds due to pileup. We conclude in Sec. 4 with a look toward future applications of our framework.

---

*e-mail: Zack.Sullivan@IIT.edu

2 The QCD power spectrum

There is an old history of studying event shape variables to describe gross features of the QCD radiation spectrum. Traditional variables, such as sphericity or oblateness, average over a large number of events, and can be used to extract information like the ratio of the three-jet to two-jet cross section in e^-e^+ data. We wish to use the full QCD power spectrum on an event-by-event basis to understand QCD at all scales.

In order to construct the QCD power spectrum out of energy-weighted angular correlations we start by revisiting the dimensionless power spectrum introduced by Fox and Wolfram [6]. The moments H_l were calculated from

$$H_l \equiv \frac{4\pi}{2l+1} \frac{\sum_m \left|E_l^m\right|^2}{E_{\text{tot}}^2} = \int_\Omega d\Omega \int_{\Omega'} d\Omega' \rho(\hat{r})\rho(\hat{r}')P_l(\hat{r}\cdot\hat{r}') \,, \tag{1}$$

where $E_l^m = \int_\Omega d\Omega \, Y_l^{m*}(\hat{r}) \, E(\hat{r})$ are the energy-weighted coefficients of the spherical harmonics, P_l are the Legendre polynomials, and $\rho(\hat{r}) = \sum_i f_i \delta(\hat{r} - \hat{p}_i)$ is a delta-function particle shape with energy fraction $f_i = E_i/E_{\text{tot}}$.

Originally it was proposed that the distributions of different moments (e.g., H_2 and H_3) be used to separate generic QCD from searches for new heavy resonances. This turned out to be impractical for several reasons including: multiplicity is low at the lepton colliders, hadron colliders introduce an additional boost that smears the moments, and most importantly the Fox-Wolfram distributions are not independent. In Fig. 1 we see that information in a *single* 3-jet event ($e^-e^+ \to q\bar{q}g \to$ hadrons) is spread out across all moments (not just $l = 3$).
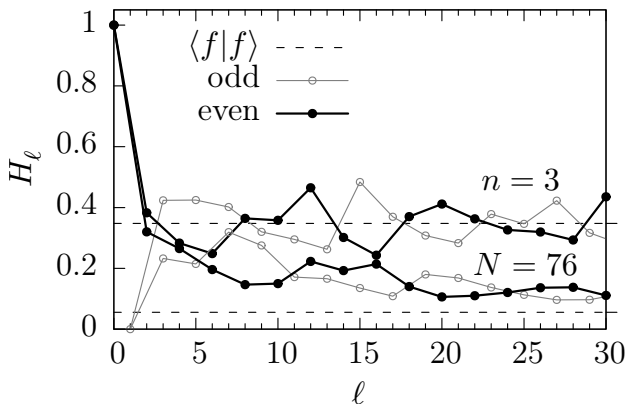

Figure 1. Power spectrum for a 3-jet event (data from Fig. 1b of Ref. [5]) produced by $n = 3$ partons (upper) or $N = 76$ resulting hadrons (lower).

Figure 1 also demonstrates that the exact power spectrum will vary uniquely between events. Given the spread due to sampling, it is not obvious how to relate the power spectrum of the measured hadrons to the underlying hard process, whose information we want to extract. In addition, we see there will be a multiplicity dependent plateau (denoted by the dashed lines) that limits the number of moments that contain information. In Refs. [4, 5] we provide a proof that this plateau is proportional to $1/N$, where N is the multiplicity of the event. This is an important information-theory limit that controls how deeply we can probe

into jet-like structure (high moments correspond to small angles). It also points out a formal problem with the Fox-Wolfram construction — the integral over the power spectrum is infinite. We will solve this shortly.

First, in order to understand what structures and artifacts are visible in the angular power spectrum, we examine in Fig. 2 the measured power spectrum obtained from an isotropic sample of neutral particles with increasing multiplicities. Again we see a plateau due to finite sampling arises at $1.278/N$ [4], setting an effective upper limit on the moments, and hence angular resolution, that can be probed. When multiplicity is large enough we can probe the beam holes at small $l \sim 2$, and a ringing due to the 6° resolution of the towers used at large $l \sim 60$. For neutral particles the angular resolution of the calorimeter towers will generally dominate over the intrinsic sampling resolution. In contrast, charged tracks have typical angular resolutions of $\sim 2°/p_T[\text{GeV}]$, and exhibit fewer artifacts in the power spectrum. Fortunately, the geometric artifacts that do arise are limited to a low absolute contribution to the power H_l.
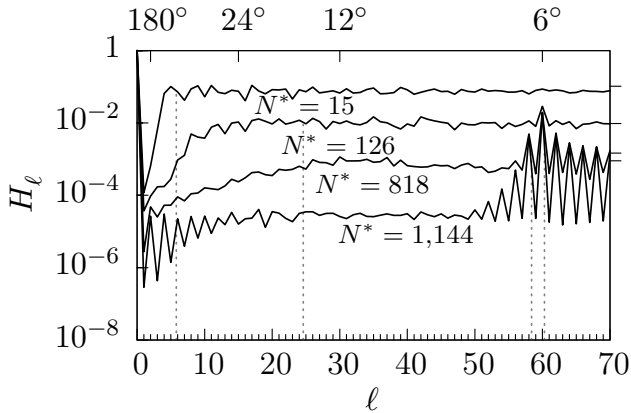

Figure 2. Power spectrum of an isotropic sample reconstructed by 6° towers for various multiplicities. Vertical dashes denote maximum angular resolution of the sample (Fig. 2b of Ref. [5]).

In addition to multiplicity effects, there is a fundamental sampling limit on QCD information in an event set by the smallest interparticle angle that can be probed. This resolution can be calculated by finding the first n angles whose weight in H_l is larger than the $1/N$ plateau, and taking the weighted geometric mean

$$\xi_{\min} = \exp\left(\frac{\sum_{k=1}^{n} w_k \ln \xi_k}{\sum_{k=1}^{n} w_k}\right), \tag{2}$$

where $\xi = \arccos(\hat{p}_i \cdot \hat{p}_j)$ and $w_k = f_i f_j$.

Given the artifacts and limits on information inherent in the construction of the QCD angular power spectrum, we introduce a modification of the Fox-Wolfram moment construction to naturally regulate the high-moment noise (and its contribution to the measurement of the power spectrum). The key is to replace the event shape $\rho(\hat{r}) = \sum_i f_i \delta(\hat{r} - \hat{p}_i)$ in Eq. 1 with a particle shape function

$$\rho(\hat{r}) = \sum_i f_i h_i(\hat{r}), \tag{3}$$

$$h_i(\hat{r}) = C \exp\left(-\frac{1 - \hat{r} \cdot \hat{p}}{\lambda^2}\right), \tag{4}$$

where C is a normalization constant, and λ is related to $\xi_{\min}$ above (for explicit formulae see Ref. [5]). This shape acts as a band pass filter to suppress the multiplicity-induced plateau.

Now, we can approximate h_i for small angles as

$$h_i(\hat{r}) \approx C \exp\left(-\frac{\theta^2}{2\lambda^2}\right) . \tag{5}$$

This allows us to include the effects of experimental resolution in a given event by solving for the λs associated with tracks (t) or towers (T) subject to the constraint that the angle between them $\theta \geq \xi_{\min}$. One subtlety here is that the minimum angle should be determined as a weighted sum over the constituent shapes

$$\xi_{ij}^* = \int d\Omega \int d\Omega' \, h_{(i)}(\hat{r}) \, h_{(j)}(\hat{r}') \arccos(\hat{r} \cdot \hat{r}') , \tag{6}$$

where ξ_{ij}^* is an effective $\xi_{\min}$ that accounts for the additional smearing of the angular resolution implied by the shape functions.

At this point we have a new framework in which to calculate the energy-weighted QCD angular power spectrum. This framework can be shown to be infrared and collinear safe [4, 5], and incorporates all QCD energy and angular information contained in an event in a finite and consistent manner. The next step is to use this framework to extract useful physics.

3 Power jets

In this section we address how to extract jet-like phenomena from our framework for evaluating the energy-weighted QCD angular power spectrum. Our "power jets" model uses "prongs" with a shape function $\rho(\hat{r})_{\text{fit}}$ based on the boost of isotropic radiation to calculate a predicted H_l^{fit}. The energies and angles of the prongs are then modified to minimize the χ^2 between the observed power spectrum H_l^{obs} and H_l^{fit}.

The number of prongs in a fit is related to the number of uniquely resolvable directions of collimated hard radiation, while the prong mass produces a shape that is related to the accompanying soft radiation. We have examined fits to two jet-like and three jet-like events, and find that we generally minimize the χ^2 with two prongs per hard jet-like object. It is important to realize we are not calculating "jets" in the standard sense — asymptotic objects with finite radius. Instead, every *power jet* includes a contribution from the full energy and angle-weighted correlation information in the event. We are extracting all of the QCD radiation that resulted from the hard underlying interaction, not cutting out a small region of phase space. This has several useful consequences.

In Fig. 3 we see the η-ϕ distribution of a reconstructed three jet-like event. This event was produced by showering and hadronizing $e^- e^+ \to q\bar{q}g$. The power jets reconstruction identifies and clusters three jet-like objects — two of which split into isolated subjets. No artificial cone-like shape is imposed on the jet (like you find with anti-k_T jets). The energy-weighted angular correlations pick out the correct combinations. This is apparent in Table 1, where we see the input partonic energies in the event compared to the power jets reconstruction. The reconstructed power jet angles map to the initial parton angles to a fraction of a percent.

One goal of our framework is to use the natural separation of correlation lengths between QCD radiation and pileup to provide a better fit to jet-like phenomena. In Fig. 4 we take the same three jet-like event above, but add *five* times the total energy in pileup events to the detector. We then plot the azimuthal reconstruction of the power jets (solid lines) over underlying hard partons (dashed lines) and compare to the anti-k_T reconstruction. Numerically the power jet fits are within 1% of the fits in Table 1, but anti-k_T has difficulty with both
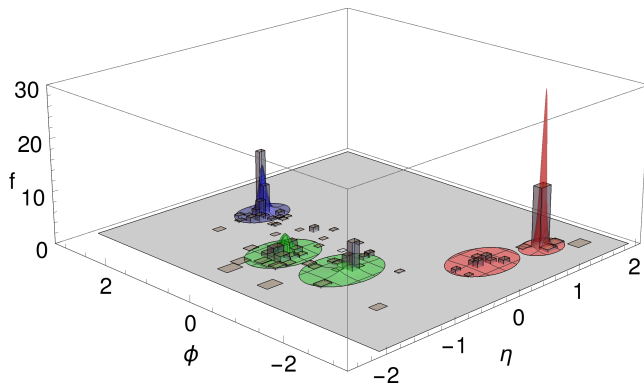
Figure 3. Energy distribution in the η-ϕ plane for a three jet-like event in an e^-e^+ collision, with an overlay of the prong shapes that fit it. A cutoff is applied to the shapes for visualization purposes.

Table 1. Reconstructed three-jet energies (with fit error) for a three jet-like event compared to the input parton energies.

	E_1 (GeV)	E_2 (GeV)	E_3 (GeV)
parton	163.0	143.5	93.56
power jet	162.0(1)	146.3(4)	91.68(4)

energy reconstruction and angle. The reason power jets do well is we fit the influence pileup shape across our entire power spectrum, while anti-k_T must perform a subtraction procedure that removes some of the hard QCD radiation and shifts the centroid of what remains. While several procedures to improve anti-k_T exist, none of them are required for power jets — the correlation structure is different.
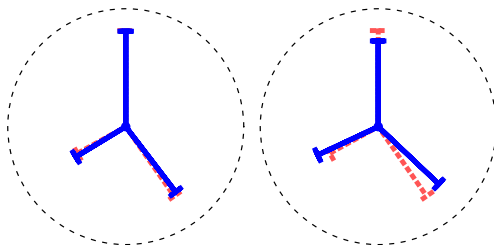

Figure 4. Azimuthal projection of the power jet fit of a three jet-like event compared to an anti-k_T fit in very high pileup. The length of the lines are the energies of the fit (solid) and parton (dashed) from Ref. [4].

4 Conclusions

In this paper we explore the deficiencies of the energy-weighted QCD angular power spectrum and its multiplicity-dependent artifacts due to finite sampling. We correct these deficiencies by introducing a shape function into the fit of the moments that acts as a band pass filter to suppress these artifacts. In the process we provide a way to include the effects

of experimental resolution that differs for tracks and towers into the extraction of the power spectrum that maintains the maximal information available in an event.

With the new framework in place, we explore how to extract jet-like phenomena from an event using "power jets." These power jets very accurately reconstruct the energies associated with the underlying hard processes — even in the presence of large pileup. Their success lies in the fact that all measured particles contribute their correlated information to the jet (not just those in a small nearby cone), and different QCD phenomena separate due to their different correlation lengths.

With the framework in place, there are now several directions to explore — many of which were requested by conference attendees. The initial analytic framework was developed for e^-e^+ events. It is a trivial extension to compute the power spectrum in hadronic collisions, but we can only perform it numerically. We can explore jet substructure by looking at the distribution of reconstructed prongs, and provide first-order QCD shower corrections to the prong shape we have adopted. We can directly calculate higher-order perturbative corrections to the QCD power spectrum to match onto a real QCD calculation. We can explore the power spectrum in heavy-ion collisions, and attempt to extract both short- and long-range correlations on an event-by-event basis. We can extract long-range phenomena in proton collisions and attempt to improve the Monte Carlo modeling of LHC physics. This is just the beginning.

Acknowledgments

References

[1] R. Calkins *et al.*, "Reconstructing top quarks at the upgraded LHC and at future accelerators," in *Planning the Future of U.S. Particle Physics: Report of the 2013 Community Summer Study of the APS Division of Particles and Fields*, edited by Norman A. Graf, Michael E. Peskin, and Jonathan L. Rosner, eConf C1307292, SNOW13-00075 (2014) [arXiv:1307.6908 [hep-ph]].

[2] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012) [arXiv:1111.6097 [hep-ph]].

[3] ATLAS collaboration, "Measurement of two–particle correlations in $\sqrt{s}$ = 13 TeV proton–proton collisions at the LHC with the ATLAS detector," ATLAS-CONF-2015-027.

[4] Keith Pedersen, "Expanding the HEP frontier with boosted *B*-tags and the QCD power spectrum," Ph.D. thesis, Illinois Institute of Technology (2018).

[5] Keith Pedersen and Zack Sullivan, "Harnessing the gobal correlations of the QCD power spectrum," in production.

[6] G. C. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e+ e-Annihilation and Other Processes," Phys. Rev. Lett. **41**, 1581 (1978).