CMS Collaboration

Contact Interaction Search with Inclusive Jets

Suman Beri[1] and Harrison B. Prosper[2]

[1]*Department of Physics, Panjab University, India*

[2]*Department of Physics, Florida State University, USA*

(Dated: October 1, 2014)

I. INSTALLATION

This work requires three packages, `lhapdf-5.9.1`, `fastnlo_toolkit-2.3.1pre-1871`, and `CIJET1.1`. The `fastNLO` program, which uses LHAPDF, will be used to calculate the binned inclusive jet p_T spectrum, while `CIJET` will be used to calculate, at next-to-leading order (NLO) accuracy, the spectra arising from contact interactions.

A. Instructions

1. First create in your home directory an area into which the codes will be installed:

```
cd
mkdir external
cd external
mkdir bin
mkdir lib
mkdir -p share/lhapdf/PDFsets
```

2. Unpack and compile `lhadpdf-5.9.1`

```
tar zxvf lhapdf-5.9.1.tar.gz
cd lhadpf-5.9.1
./configure --prefix=$HOME/external
make
make install
```

3. Unpack and compile `fastnlo_toolkit-2.3.1pre-1871`

```
tar zxvf fastnlo_toolkit-2.3.1pre-1871.tar.gz
cd fastnlo_toolkit-2.3.1pre-1871
./configure --prefix=$HOME/external
make
make install
```

4. Unpack and compile `CIJET-1.1`

```
tar zxvf CIJET-1.1.tar.gz
cd CIJET1.1
make
```

II. ANALYSIS DETAILS FOR 8 TEV CMS DATA

We shall use the measured inclusive jet p_{T} spectrum at 8 TeV with the following characteristics:

- Jets defined by anti-k_{T} algorithm with distance parameter $D = 0.7$.
- Jet $|y| < 0.5$
- Jet p_{T} binning, range $507 \leq p_{\mathrm{T}} \leq 2500$ GeV

For the models, we shall use the following:

- Renormalization (μ_r) and factorization (μ_f) scales with nominal value $\mu =$ jet p_{T} and each scale changed independently by the factors 1/2, 1, and 2.
- PDFs: CT10nlo, NNPDF21, and MSTW2008nlo68cl.

Spectra will be computed for random samplings of PDFs from each PDF group, at least 100 PDF sets per group. Ensembles already exist for NNPDF, but for CTEQ and MSTW we shall generate our own samples using the procedure described in JHEP 08 (2012) 052 [arXiv:1205.4024 [hep-ph]] and http://mstwpdf.hepforge.org/random.

Let ΔF represent a random change in the predicted cross section for a given jet p_T bin due to a random sampling from the space of PDFs. This change can be approximated by

$$\Delta F = \frac{1}{2f} \sum_{j=1}^{n} |F_j^+ - F_j^-| \, R_j, \tag{1}$$

where the R_j are random numbers sampled from n independent Gaussians of zero mean and unit variance, $F_j^{\pm}$ are the $\pm$ shifts in predicted cross sections due to the $\pm$ shifts in the n PDF parameters, and $f = 1$ or 1.64 depending on whether the $\pm$ shifts are 68% or 90% C.L. shifts. We have assumed that $|F^+ - F^-|/(2z)$ approximates one standard deviation and the shifts are approximately Gaussian. CTEQ provides 90% shifts, while MSTW provides both 68% and 90% shifts. For CT10 $n = 26$ and for MSTW2008 $n = 20$.

Shifts in the PDF parameters will induce correlations across all jet p_T bins and across all models. In order to account for these correlations correctly, for a given PDF set we need to use the same set of random numbers R_j for all bins and all models. For each PDF set it would probably be safer to generate these 100 sets of n numbers R_j and store them in a file for later use.

III. OBSERVATIONS

The high-p_T end of the observed inclusive jet spectrum (CMS PAS SMP-12-012, CMS Analysis Note AN2012_223_V16, 2013 and its updates) with 19.34fb^{-1} of data is shown in Table I.

TABLE I: Jet yield for each p_T bin and $|y| < 0.5$.

bin	p_T range	jet yield	bin	p_T range	jet yield
1	507-548	701198	11	1032-1101	7112
2	548-592	452728	12	1101-1172	4348
3	592-638	284560	13	1172-1248	2706
4	638-686	178899	14	1248-1327	1581
5	686-737	115347	15	1327-1410	1025
6	737-790	72274	16	1410-1497	529
7	790-846	46057	17	1497-1588	335
8	846-905	29217	18	1588-1784	282
9	905-967	18615	19	1784-2116	88
10	967-1032	11602	20	2116-2500	6

IV. MODELS

For the 8 TeV data set, we shall consider the contact interaction (CI) model defined by the effective Lagrangian [1],

$$L = 2\pi\lambda \sum_{i=1}^{6} \eta_i \, O_i, \tag{2}$$

where $\Lambda = 1/\sqrt{\lambda}$ is the CI mass scale, η_i [1] are constants and O_i are dimension 6 operators. This model is defined by seven parameters: Λ, $\eta_1, \cdots, \eta_6$. In practice, we shall follow the CMS paper arXiv:1202.5535v1 [hep-ex] and consider specific combinations of values for the η_i. Writing, $\eta_{LL} = \eta_1$, $\eta_{RL} = \eta_3/2$, and $\eta_{RR} = \eta_5$, we shall consider the models in Table II. At next-to-leading order (NLO), the inclusive jet cross section for jet p_T bin j can be written

TABLE II: Models to be considered in this analysis

Model	η_{LL}	η_{RL}	η_{RR}
LL	± 1	0	0
RR	0	0	± 1
VV	± 1	± 1	± 1
AA	± 1	∓ 1	± 1
V-A	0	± 1	0

as [2],

$$\sigma_j \; = \; \sigma_{\mathrm{QCD}} + (b - b' \ln \lambda)\, \lambda + (a - a' \ln \lambda)\, \lambda^2, \tag{3}$$

where σ_{QCD} is the QCD cross section computed at NLO using the `fastNLO` program. We can use Gao's program, `CIJET`, to calculate the coefficients $b, b', a,$ and a' for each of the models listed in Table II. This will allow us to construct a family of 1-parameter models with which to analyze the 8 TeV data.

We shall not unfold the spectrum; instead we shall convolve the predicted spectra using the jet response function

$$R(p_T|z, x, y) = \mathrm{Gaussian}(p_T, xz, y\sigma_z), \tag{4}$$

which is assumed to be a Gaussian with mean z — the true jet p_T — and standard deviation given by

$$\sigma_z = zC_{Data}\sqrt{\frac{N^2}{z^2} + \frac{S^2}{z} + C^2}, \tag{5}$$

where $C_{Data} = 1.12$, $N = 6.130$ GeV, $S = 0.949$ GeV$^{1/2}$, and $C = 0.031$. In the simplest case, the scale factors x and y are used to model the uncertainty in jet energy scale (JES)

[1] We use η_i instead of λ_i, which is the notation used in Ref. [1], in order to avoid possible confusion with the parameter λ.

and the jet energy resolution (JER) σ_z. These constants (derived from simulated jets) are for the rapidity bin $|y| < 0.5$. As is clear from Table 6 in AN2012_223_V16, 2013, the jet resolution depends slightly on rapidity[2]. In the simplest case, x and y are Gaussian variates with unit mean and standard deviations of 0.04 and 0.10, respectively, representing the overall 4% uncertainty in the jet energy scale at high p_T and 10% uncertainty in the jet energy resolution. For a more accurate modeling of the JES, we shall use the fact that (currently) x is a linear sum of 33 independent jet (p_T, y)-dependent Gaussian components, each with its own standard deviation. The details of the calculations may be found at

`https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources`

In practice we shall convolve $d\sigma_{\text{QCD}}/dp_T$ and the differential distribution of the coefficients in Eq. (3) with the jet energy response function, Eqs. (4) and (5), using

$$c_j^{\text{obs}}(x, y) = \int_{p_T \text{ bin } j} dp_T \int_0^\infty dz\, R(p_T|z, x, y)\, f(z), \tag{6}$$

where $f(z)$ represents a smooth interpolation of either $d\sigma_{\text{QCD}}/dp_T$ or each of the differential distributions of the coefficients. We shall generate random samples of x, y pairs and calculate the smeared binned spectra using Eq. (6), using the same x, y pair for QCD and the CI models and for all PDF sets in order to maintain the correct correlation across all models induced by the JES and JER uncertainties. Another x, y pair will be sampled and another set of smeared spectra will be calculated. This procedure will be repeated 100 times for each PDF set. The results from the three PDF sets will be pooled into a single 300 member ensemble of smeared spectra.

Note that each PDF set will be sampled independently, but for a given PDF set we shall use the same set of sampled PDFs for the QCD and CI models.

A. Likelihood

We shall follow the 7 TeV analysis and use the multinomial likelihood function

$$p(D|\lambda, \omega) = \binom{N}{N_1, \cdots, N_K} \prod_{i=1}^{K} \left(\frac{\sigma_i}{\sigma}\right)^{N_i}, \tag{7}$$

[2] There is a typo in Table 6; the numbers for N and C are switched.

where $\sigma \equiv \sum_{i=1}^{K} \sigma_i$, $N \equiv \sum_{i=1}^{K} N_i$ is the total observed count, N_i the count in jet p_T bin i, and ω denotes the nuisance parameters, $\omega = \sigma_\mathrm{QCD}, b, b', a, a'$.

---

[1] J. Gao et al., "Next-to-leading QCD effect to the quark compositeness search at the LHC", *Phys. Rev. Lett.* **106** (2011) 142001, `arXiv:1101.4611`.

[2] J. Gao, CIJET, `arXiv:1301.7263`.