Search for Contact Interactions with Inclusive Jets Notes

Suman Beri¹, Suneel Dutt², and Harrison B. Prosper³

¹Department of Physics, Panjab University, India

²Department of Physics, Shoolini University, India

³Department of Physics, Florida State University, USA

(Dated: July 18, 2015)

CMS Collaboration

I. INTRODUCTION

The goal of this work is first to extend the 7 TeV search for contact interactions [?] to 8 TeV and to prepare the groundwork for a search at 13 TeV. For the 8 TeV search, we shall use the measured inclusive jet $p_{\rm T}$ spectrum at 8 TeV, which has the following characteristics:

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

For the QCD and CI models, we shall use the following:

- Renormalization (μ_r) and factorization (μ_f) scales with nominal value $\mu = \text{jet } p_T$ and each scale changed independently by the factors 1/2, 1, and 2, nine pairs in total.
- PDFs: CT10nlo, NNPDF23_nlo_as_0118, and MSTW2008nlo68cl.

Spectra will be computed for random samplings of PDFs from each PDF set, 100 PDF members per set. Ensembles already exist for NNPDF, but for CTEQ and MSTW the ensembles shall be generated using the procedure described in Ref. [?] and released in version 6 of LHAPDF. For each randomly sampled PDF member, we shall compute the (smeared) inclusive jet $p_{\rm T}$ spectrum. In order to maintain the correlations between the QCD and CI models, induced by the PDFs, the same set of sampled PDFs will be used for both.

The data will be interpreted using the contact interaction (CI) model defined by the effective Lagrangian [?],

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

where $\Lambda = \lambda^{-1/2}$ is the CI mass scale, κ_i^{-1} are constants and O_i are dim-6 operators. Clearly, the only unique combinations are the ratios $2\pi\kappa_i/\Lambda^2$; it is therefore only a matter of convention, and convenience, that in Eq. (??) we split these ratios into an overall mass

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

scale Λ and the dimensionless parameters κ_i . This model is defined by seven parameters: Λ , $\kappa_1, \dots, \kappa_6$. In practice, we shall follow Ref. [?] and consider specific combinations of values for the κ_i . Writing, $\eta_{LL} = \kappa_1$, $\eta_{RL} = \kappa_3/2$, $\eta_{RR} = \kappa_5$, and $\kappa_2 = \kappa_4 = \kappa_6 = 0$, we shall consider the models in Table ??.

TABLE I: 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

At next-to-leading order, the inclusive jet $p_{\rm T}$ cross section per bin can be written as [?],

$$\sigma = \sigma_{\text{QCD}}$$

$$+ \lambda \sum_{i=1}^{6} \kappa_{i}(b_{i} + a_{i}g + a_{i}f)$$

$$+ \lambda^{2} \sum_{i=1}^{6} \kappa_{i}^{2}(b_{ii} + a_{ii}g + a_{ii}f)$$

$$+ \lambda^{2} \sum_{i=1,3,5} \kappa_{i}\kappa_{i+1}(b_{ii+1} + a_{ii+1}g + a_{ii+1}f)$$

$$+ \lambda^{2} \sum_{i=1,2,5,6} \kappa_{i}\kappa_{4}(b_{i4} + a_{i4}g + a_{i4}f), \qquad (2)$$

where $\sigma_{\rm QCD}$ is the QCD cross section, $r = \ln(\Lambda/\mu_0)$ is split into two parts denoted by $g = -\ln(\mu_0\sqrt{k})$ and $f(\lambda) = \ln(\sqrt{k/\lambda})$, μ_0 is an arbitrary bin-dependent reference scale (provided by CIJET) and k is an arbitrary bin-independent scale that we are free to choose. At leading order (LO), the a terms vanish. We shall use CIJET to calculate the coefficients b and a for a given PDF set and member. Since the cross section is linear in the coefficients b, $a \times g$, and a, we can smear these coefficients independently of the model parameters λ , $\kappa_1, \dots, \kappa_6$ (see Sec. ??). This will allow us to construct a family of models in which the 7 parameters appear explicitly.

These models are defined using parton-level jets, whereas CMS corrects observed jets so that, on average, their energy scale matches that of jets at the hadron (i.e, particle)

level. Therefore, in principle, the models must be corrected to the hadron level in order to compare data with theory. This non-perturbative (NP) correction, which currently cannot be calculated from first principles, is applied as a multiplicative factor to the parton-level inclusive jet differential cross section. The NP correction

$$C_{\text{NP}}(p_{\text{T}}) = A + B/p_{\text{T}}^{n},$$

$$A = 1.003,$$

$$B = 77.374,$$

$$n = 1.385,$$
(3)

has been derived by the CMS Inclusive Jet measurement group by comparing parton-level jet predictions with predictions at the hadron level, computed using PYTHIA. In addition to the non-perturbative corrections, we also include the far more important electroweak corrections provided by the Inclusive Jet measurement group. These corrections rise to about 12% at jet $p_T \sim 2\text{TeV}$.

II. OBSERVATIONS

The high- $p_{\rm T}$ end of the observed inclusive jet spectrum (CMS PAS SMP-12-012, CMS Analysis Note AN2012_223_V16, 2013 and its updates) with the full 8TeV integrated luminosity of 19.71fb⁻¹ and the Winter 2014, V8, jet energy corrections (JEC) is shown in Table ??.

III. JET RESPONSE

As alluded to above, we shall not unfold the spectrum. Instead, we shall convolve the predicted spectra using the jet response function

$$R(p'_{\rm T}|z) = \text{Gaussian}(p'_{\rm T}, z, \sigma_z), \tag{4}$$

where $p'_{\rm T}$ is the observed jet transverse momentum using the *nominal* jet energy scale, z is the true jet transverse momentum and σ_z , the jet transverse momentum resolution, is given by

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

TABLE II: Jet yield for each $p_{\rm T}$ bin and |y| < 0.5.

bin	$p_{\mathrm{T,min}}$	$p_{\mathrm{T,max}}$	jet yield
1	507	548	722864
2	548	592	468062
3	592	638	296032
4	638	686	186497
5	686	737	120580
6	737	790	76129
7	790	846	48454
8	846	905	31121
9	905	967	19639
10	967	1032	12373
11	1032	1101	7746
12	1101	1172	4670
13	1172	1248	2930
14	1248	1327	1746
15	1327	1410	1111
16	1410	1497	602
17	1497	1588	367
18	1588	1784	311
19	1784	2116	111
20	2116	2500	6

where $C_{Data} = 1.052$, N = 5.7936 GeV, S = 0.984 GeV^{1/2}, and C = 0.029. 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.

Equation (??) is the appropriate response function for the nominal jet energy scale. In order to account for uncertainty in the JES and JER, we introduce two zero mean, unit variance, Gaussian variates x and y, respectively. We account for the 10% uncertainty in the JER, σ_z , by scaling it by the factor $Y = 1 + y\sigma_y$. Likewise, we account

for the p_{T} -dependent uncertainty in the JES, which varies from 1% at $p'_{\mathrm{T}} = 500$ GeV to 1.5% at $p'_{\mathrm{T}} = 2,500$ GeV according to the results encoded in the uncertainty source file Winter14_V5_DATA_UncertaintySources_AK7PDF.txt², by scaling the nominal jet p_{T} , p'_{T} , by the factor $X = 1 + x\sigma_{p'_{\mathrm{T}}}$. With these modifications, the jet response function becomes

$$R(p_{\mathrm{T}}|x,y,z) = \delta(p_{\mathrm{T}} - Xp_{\mathrm{T}}') \operatorname{Gaussian}(p_{\mathrm{T}}',z,Y\sigma_{z}). \tag{6}$$

The details of the JES uncertainty calculations may be found at

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

The bin-by-bin quantities $\sigma_{\rm QCD}$, b, $a \times g$, and a (58 per bin) will be converted to differential quantities by dividing them by the $p_{\rm T}$ bin width and the rapidity bin width |y|=0.5. The differential quantities will then be convolved with the jet energy response function, Eq. (??), for a random pair of values (x, y). This will yield smeared differential distributions,

$$f_{\text{obs}}(p_{\text{T}}|x,y) = \int_{0}^{\infty} dz \, R(p_{\text{T}}|x,y,z) \, f(z),$$
 (7)

for each of the 58 coefficients in Eq. (??), where f(z) represents a smooth interpolation of a differential distribution. The ensemble of functions $\{f_{\text{obs}}(p_{\text{T}}|x,y)\}$ will encode the uncertainty in both the JES and the JER.

As noted, the cross section per bin, Eq. (??), is a linear combination of 58 terms. Therefore, for every PDF set and member within that set we must compute 58 smeared differential distributions, each of which must then be integrated over 20 jet p_T bins. The spectra from the three PDF sets will be pooled into a single ensemble of smeared spectra that will represent the distribution of spectra that incorporates smearing due to PDF uncertainties as well as jet smearing. In order to separate out the effect of jet smearing from that due to PDF uncertainties, we shall repeat the procedure using the nominal PDFs, that is, the PDFs denoted by PDF member zero in each PDF set.

IV. ROADMAP

In order to permit integration over the PDF parameters, we shall generate a sample of PDF set members and place them in the standard LHAPDF area for PDF sets

² We need to migrate to V8.

(share/lhapdf/PDFsets). Note: samples already exists for NNPDF.

We shall make use of two programs from the CIJET package, dijets4ci_mul and ciconv. The first program must be run in a directory at the same level as the directory fastCI, which contains the sub-directory fgrid. Here is a suggested directory structure:

```
work/
data/ run dijets4ci_mul here
output from dijets4ci_mul (*_fitresults.dat files)
fastCI/fgrid/ output from dijets4ci_mul (*_app1.sum files)

fastCI/ run ciconv here
   CT10/ output from ciconv
        000/
   :
      099/
MSTW/
      000/
   :
   099/
NNPDF/
      000/
   :
   099/
```

The calculation of cross sections using CIJET is a two-step process. First, for each true jet $p_{\rm T}$ bin, one calculates a set of coefficients using the program dijets4ci_mul:

```
dijets4ci_mul <tagname>
Example:
dijets4ci_mul ci0507
```

The results of this step, which can take several hours, is a file ci0507_fitresults.dat in the data directory, which are independent of the PDF set, renormalization and factorization scales (μ_r and μ_f , respectively), and of the model parameters Λ , $\kappa_1, \dots, \kappa_6$. Therefore, they need be calculated only once per center-of-mass energy and p_T and rapidity bin.

Next, the program ciconv, which is run from the fastCI directory, is used to compute coefficients for a specific PDF set/member and 9 different pairs of scale choices. Given these coefficients, one can calculate the cross section for any values of the model parameters for the 9 different pairs of scale choices:

ciconv <PDF-set-name> <PDF-number > <fgrid-filename> <output-filename>
Example:

ciconv CT10nlo 0 ci0507_app1.sum CT10/000/ci0507.txt

In the example, we are calculating coefficients for member 0 of CT10nlo using the coefficients stored in the grid file fastCI/fgrid/ci0507_app1.sum, with the results written to file CT1/000/ci0507.txt. This is to be repeated for each of the PDF members.

A. Likelihood

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

$$p(D|\lambda, \kappa_1, \cdots, \kappa_6, \omega) = \binom{N}{N_1, \cdots, N_K} \prod_{i=1}^K \left(\frac{\sigma_i}{\sigma}\right)^{N_i}, \tag{8}$$

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, $\sigma_{\rm QCD}$, b, $a \times g$, and a, over which the likelihood will be marginalized.

The calculations of the QCD cross sections are done in the fastNLO directory.

Appendix A: Installation

This work requires the packages, LHAPDF-6.1.5, fastnlo_toolkit-2.3.1pre-1871, and CIJET-1.1. The fastNLO program is used to calculate the inclusive jet $p_{\rm T}$ spectrum, while CIJET is used to calculate, at next-to-leading order (NLO) accuracy, the spectra arising from contact interactions (CI). The LHAPDF package must be compiled first because it is needed by fastNLO and CIJET.

First create a work area, say CMSCI, and copy the directories CI, corrections, data, fastNLO, and fastCI to CMSCI from the Dropbox/CMSCI.

Below we assume that you are using a bash shell.

1. Create an area into which the codes will be installed:

```
cd
mkdir external
cd external
mkdir bin
mkdir lib
mkdir include
mkdir downloads
mkdir -p share/LHAPDF/PDFsets
```

In your .bash_profile file, which is a hidden file in your home directory, you should update the environment variables PATH and LD_LIBRARY_PATH as follows. (I'm assuming that you have already setup Root.)

```
export PATH=\$HOME/external/bin:$PATH
export LD_LIBRARY_PATH=\$HOME/external/bin:\$LD_LIBRARY_PATH
```

Next time you create a terminal window these variables will be updated. For now, just do

```
cd
source .bash_profile
```

Below we assume that the files to be unpacked are in the directory

\$HOME/external/downloads

2. Install LHAPDF-6.1.5. NB: There is a bug in this version of LHAPDF. It will be fixed in version 6.1.6. In the meantime, to fix the bug do the following. Wherever you see the abs function in the source code LHAPDF-6.1.5/src/PDFSet.cc, simply delete the word abs and nothing else and save the corrected source code.

```
cd
cd external
tar zxvf downloads/LHAPDF-6.1.5.tar.gz
cd LHAPDF-6.1.5
./configure --prefix=$HOME/external
make
make install
```

Next, we need to compile the programs in the directory LHADPDF-6.1.5/examples, in particular, hessian2replicas, which we shall use to generate a random sample of 500 PDF sets from CT10nlo, and later from the MSTW PDF set.

```
cd
cd external/LHAPDF-6.1.5/examples
make
```

cd
cd external/bin
ln -s ../LHAPDF-6.1.5/examples/hessian2replicas .

In the last step, we created a link in **bin** to hessian2replicas.

In order to use LHAPDF, it is necessary to set the environment variable LHAPDF_DATA_PATH to the directory that contains the PDF sets; for example:

```
export LHAPDF_DATA_PATH=$HOME/external/share/LHAPDF/PDFsets
```

Also remember to execute the setup.sh script, which should be copied from the Dropbox and placed in your working directory, assumed to be called CMSCI.

```
cd CMSCI source setup.sh
```

3. Install fastnlo_toolkit-2.3.1pre-1871.

Important: After you have unpacked this file, you should replace the program fnlo-tk-cppread.cc in the fastnlo_toolkit-2.3.1pre-1871/src with the version that is in Dropbox/CMSCI/downloads. The version in the Dropbox contains a version with an extra argument for the PDF member. Also check that you have the most recent version of runfastNLO.py (15-Apr-2015) in your fastNLO/work area.

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

4. Unpack and compile CIJET-1.1.

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

We shall be using the CIJET program ciconv. In order to make it globally accessible create a link to it as follows

```
cd
cd external/bin
ln -s ../CIJET1.1/fastCI/ciconv .
```

5. **Download** PDF Sets

For now, just download the PDF set CT10nlo. Later, download the other two sets.

cd

cd external/share/LHAPDF/PDFsets
wget https://www.hepforge.org/archive/lhapdf/pdfsets/6.1/CT10nlo.tar.gz
tar zxvf CT10nlo.tar.gz

6. Generate randomly sampled PDF sets

cd
cd external/share/LHAPDF/PDFsets
hessian2replicas CT10nlo 5678 500

This will create and new directory in external/share/LHAPDF/PDFsets called CT10nlo_rand5678 with 500 randomly sampled PDFs from CT10nlo. The random number seed is 5678. Each of these 500 PDF members (plus the nominal PDF member identified as member zero) will be used to calculate a QCD and CI spectrum. The spread in these spectra will reflect the uncertainty in the PDFs.

We are now, finally, able to start our calculations!

Appendix B: Computing Spectra

1. Computing unsmeared QCD spectra

We compute QCD spectra for each of the 501 PDF members of CT10nlo using 7 different combinations of renormalization and factorization scales. These calculations are performed with the program fnlo-tk-cppread, which is called by the script fastNLO/work/runfastNLO.py, which in turn is called as follows

cd CMSCI/fastNLO/work
python runfastNLO.py CT10nlo_rand5678

This will create 7 text files per PDF member in the directories labeled

../CT10/NNN

where NNN is the PDF member number (000 to 500). Once this is done, we convert the text files into Root histograms stored in files as follows:

python createQCDhists.py CT10

2. Computing smeared QCD spectra

The next step is to smear the theoretical spectra with the CMS jet response function. First, compile the relevant code in the CI directory under CMSCI:

cd CI

make

Note: If your makefile fails, you can try deleting the code fragments

>& \$*.FAILED

which appear in two places in the Makefile, then try make again.

Now, go to the main work directory under CMSCI, where you will find the script smearSpectra.py. In order to smear the spectra in the directory .../fastNLO/CT10 do

python smearSpectra.py ../fastNLO/CT10

This will create new qcd.root files, containing the smeared QCD spectra, in the directories fastNLO/CT10/NNN/JESJERPDF, where NNN is the PDF member number.

3. Computing unsmeared CI coefficient spectra

As noted above, the calculation of the CI spectra proceeds in two steps. The first step takes several hours. However, it need be done only once for a given center-of-mass energy and $p_{\rm T}$ and η bin. Thereafter, the calculations are fast.

For each of the three PDF sets we proceed as follows:

1. Move to the data directory and run the CIJET program dijets4ci_mul <tag-name>. This will create output in data as well as in fastCI/fgrid. Remember, the directories data and fastCI must be at the same level. In practice, this calculation should be run as a batch job, one for each true jet p_T bin. However, since the bins of true jet p_T are specified in bininput.card, the kinematics in kininput.card, and the process (inclusive jets) in proinput.card (see Appendix), it will be necessary for each batch job to use a different "data" directory for each jet p_T bin. I suggest, therefore, that you create a different data directory for each jet p_T, labeling the data directories dataNNNN

where NNNN is the $p_{\rm T}$ bin (see the example data0049 in the dropbox). Remember to set the bin in bininput.card. In the example in the dropbox, the $p_{\rm T}$ bin is 49 to 56. To compute the coefficients needed for subsequent calculations of the CI cross sections, one does:

cd data0049
dijets4cl_mul ci0049

2. Once the above time-consuming calculations have been performed for every true jet p_T bin, we can use their results to compute for every PDF set and member the coefficients needed to calculate cross sections. Then, these coefficient spectra can be used to calculate differential cross sections for any values of the parameters, Λ and κ₁, · · · , κ₆. We compute CI coefficient spectra for each PDF member of CT10nlo using 9 different combinations of renormalization and factorization scales. The calculations are performed with the program ciconv, which is called by the Python program fastCI/runciconv.py and which must be run in the fastCI directory. The program runciconv.py loops over all PDF members and, for each, over all the grid files fgrid/ci*.sum and calls ciconv for a given PDF set, PDF member, and grid file. The results are written to the directory CT10, which contains a subdirectory for each PDF member.

4. Computing smeared CI coefficient spectra

Just as for QCD, go to the main work directory under CMSCI and run the script smearSpectra.py,

python smearSpectra.py ../fastCI/CT10

to create new *.root files, one for each CI coefficient, containing the smeared coefficient spectra, in the directories fastCI/CT10/NNN/JESJERPDF, where NNN is the PDF member number. This calculation takes about two hours on a laptop.

5. Creating a RooFit workspace

The last step before limit calculations is to create a workspace in which the results of the previous calculations are stored. This is done as follows:

python createWorkspace.py CT10

This creates a RooWorkspac called CT10_JESJERPDF_workspace.root, which is the basis for the limit calculations.

You can check the integrity of the workspace by running the workspace plotting program:

python plotWorkspace.py -m LL CT10_JESJERPDF_workspace.root

This will display the spectra for a Left-Left (LL) CI model with $\Lambda=20\,\mathrm{TeV}$ for both constructive and destructive interference. If this is successful, we shall be ready to calculate limits.

Appendix C: Input Files for CIJET

Example bininput.card

```
massbin [inv. mass or pt bin list in GeV]
507 548
rapbin [angular bin list, chi, ystar, ymax or |y|]
0.0 0.5
```

Example kininput.card

```
jetscheme 1 [jet algorithm, 1 for kT (=CA, anti-kT), 2 for others]
recscheme 2 [jet recombination scheme, 1 for "ET", 2 for "4-vector(LHC)"]
Rcone 0.7 [cone size or distance parameter]
Rsep 1.3 [Rsep used in midpoint algorithm]
ptcut 20 [acceptance condition for jet pT/GeV]
ycut 2.5 [acceptance condition for jet rapidity]
ptadd1 20 [additional cut on pT/GeV of jet1 (harder one) in dijet]
ptadd2 20 [additional cut on pT/GeV of jet2 (softer one) in dijet]
```

yb 1.11 [upper cut on yboost of dijet system]
ys 1.39 [upper cut on ystar of dijet system]

Example proinput.card

```
[pdf name as in LHAPDF]
pdf CT10nlo.LHgrid
pdfmember 0
               [pdf member as in LHAPDF]
                              [fit "fitcs/ll/al"]
mode fital
pseed 2444
                           [seed of the random number generator]
many 1000000
              [typical number of MC points, rec. 3M or adjust accordingly]
           [1 for chi, 2 for ystar, 3 for ymax, 4 for inclusive jet]
                   [1 for LHC, 0 for Tevatron]
ppcollider 1
              [c.m.s. energy of collider, in GeV]
sqrtS 8000
scalescheme 0
                                     [choice of dynamic scales]
```

- [1] CMS Collaboration, Chatrchyan, S. et al., "Search for contact interactions using the inclusive jet pT spectrum in pp collisions at \sqrt{s} =7 TeV", Phys.Rev. **D87** (2013) 5, 052017; CMS-EXO-11-010, CERN-PH-EP-2013-002, DOI: 10.1103/PhysRevD.87.052017, e-Print: arXiv:1301.5023.
- [2] JHEP 08 (2012) 052 [arXiv:1205.4024 [hep-ph]] and http://mstwpdf.hepforge.org/random.
- [3] 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.
- [4] J. Gao, CIJET, arXiv:1301.7263.
- [5] CMS Collaboration, Chatrchyan, S. et al., "Search for quark compositeness in dijet angular distributions from pp collisions at $\sqrt{s} = 7$ TeV", JHEP **1205** (2012) 055 (CMS-EXO-11-017).