Contact Interactions Search Using Inclusive Jet Spectrum PDF Uncertainties

Harrison B. Prosper

Florida State University

PDF@CMS Meeting

20 November 2014

Outline

- **Introduction**
- Analysis Overview
- PDF Uncertainties in CI Search
- Summary

Introduction

Goal

Use the p_T spectrum of inclusive jet events

$$p + p \rightarrow \text{jet} + X$$

to search for new QCD-like interactions whose details cannot be resolved at current energies

Search Region

$$|y| < 0.5$$

$$507 \leq p_T \leq 2500 \text{ GeV}$$

In the process of finalizing the 8 TeV search

Introduction

Model

Use the effective Lagrangian

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

where $\lambda = 1/\Lambda^2$, κ_i are free parameters, and each O_i is a sum of dim-6 operators:

$$O_{1,2} \sim \bar{u}_L \gamma_\mu u_L \; \bar{u}_L \gamma^\mu u_L$$
$$O_{3,4} \sim \bar{u}_L \gamma_\mu u_L \; \bar{u}_R \gamma^\mu u_R$$
$$O_{5,6} \sim \bar{u}_R \gamma_\mu u_R \; \bar{u}_R \gamma^\mu u_R$$

Introduction

The QCD+CI cross section, @NLO, can be written as

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

$$+ \lambda \sum_{i=1}^{6} \kappa_i [b_i + a_i g + a_i f]$$

$$+ \lambda^2 \sum_{i=1}^{6} \kappa_i^2 [b_{ii} + a_{ii} g + a_{ii} f]$$

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

$$+ \lambda^2 \sum_{i=1,2,5,6} \kappa_i \kappa_4 [b_{i4} + a_{i4} g + a_{i4} f]$$

where $g = -\ln(\mu_0 \sqrt{k})$ and $f = \ln(\sqrt{(k / \lambda)})$

The CI term contains 57 coefficients

Introduction

Theoretical Input

- QCD @ NLO

  fastnlo_toolkit-2.3.1pre-1871 + fnl3332y0.tab

- CI @ NLO

  CIJET-1.1

- Non-perturbative corrections (Sanmay Ganguly)

- PDFs (LHAPDF-6.1.4)

  CT10nlo, MSTW2008nlo68cl, NNPDF23_nlo

Summary of PDF Sets Used

PDF sets used in CI search @ 8 TeV

	CT10nlo	MSTW2008	NNPDF2.3
No. of PDFs	6	7	7
No. of Parameters	26	28	259
HERA DIS	✓	✓	✓
Fixed-target DIS	✓	✓	✓
Fixed-target DY	✓	✓	✓
Tevatron W, Z+jets	✓	✓	✓
LHC W, Z+jets			✓

S. Forte and G.Watt, Ann.Rev.Nucl.Part.Sci. 63 (2013) 291-328

Outline

- Introduction
- **Analysis Overview**
- PDF Uncertainties in CI Search
- Summary

Analysis Overview

Analysis Steps

1. For CT10, MSTW2008, generate 100 *randomly* sampled PDF sets (using hessian2replicas in LHAPDF6), plus the 100 random sets from NNPDF.
2. For each of the randomly sampled PDF sets and 7 combinations of the renormalization and factorization scales, compute the QCD inclusive jet p_T spectrum and the CI contribution for given values of κ and Λ.
3. Convolve each spectrum with the jet response function taking into account JES and JER uncertainties.
4. Average likelihood over pooled ensemble of spectra.
5. Compute limit on Λ for specific values of κ.

Analysis Overview

Bayesian Justification for Pooling PDF ensembles

Each PDF fitting group M provides (in effect) a prior $p(\omega_M \mid M)$

Using PDF model M, we approximate the marginal likelihood

$$p(D \mid \Lambda, M) = \int p(D \mid \Lambda, \omega_M, M)\, p(\omega_M \mid M)\, d\omega_M$$

$$\approx \frac{1}{K} \sum_i^K p(D \mid \Lambda, \omega_{M,i}, M)$$

by averaging the likelihood with respect to the randomly generated PDFs.

Analysis Overview

Bayesian Justification for Pooling PDF ensembles

Given a discrete prior $\pi(M)$, we can *average* over models

$$p(D \mid \Lambda) = \sum_M p(D \mid \Lambda, M)\,\pi(M)$$

to obtain an overall likelihood $p(D \mid \Lambda)$ for computing limits.

Moreover, for a specific physics model, say QCD ($\Lambda = \infty$), we can calculate the probability of the PDF model M

$$p(M \mid D) = p(D \mid M)\,\pi(M) \,/\, \sum_M p(D \mid M)\,\pi(M)$$

and thereby rank PDF models according to which best fits the data.

Outline

- Introduction
- Analysis Overview
- **PDF Uncertainties in CI Search**
- Summary

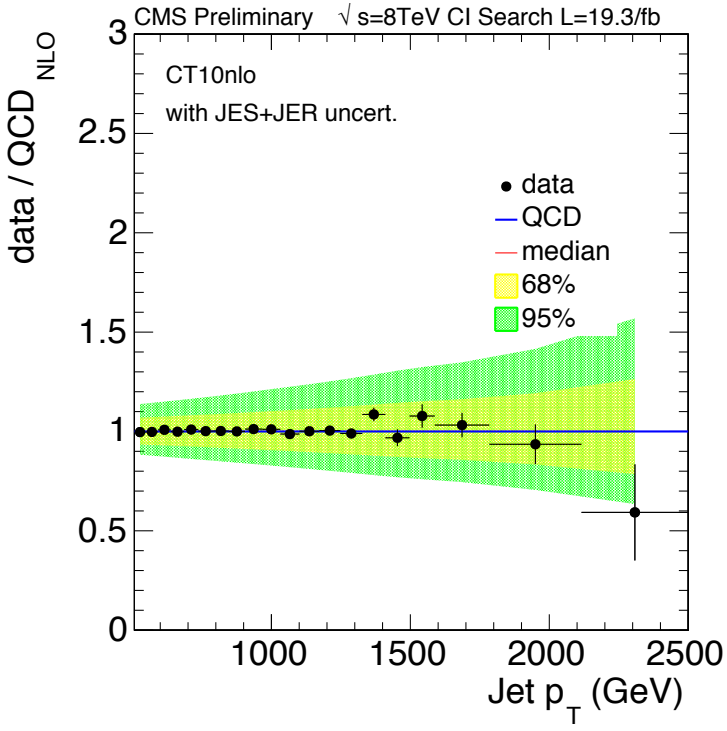
PDF Uncertainties

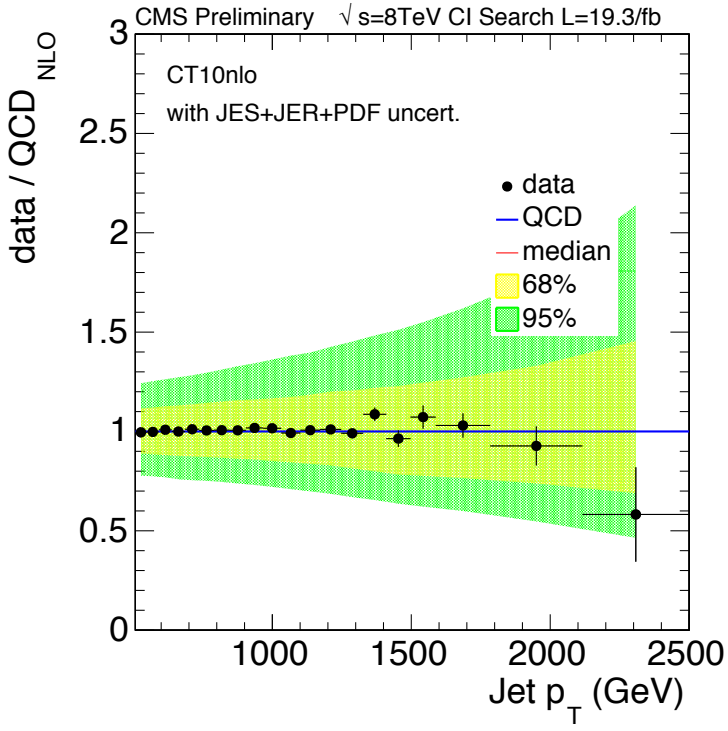
- Contact interaction effects are most prominent at $y = 0$ and at high jet p_T .

- But this is also where the uncertainties in the jet energy scale (JES) and the PDFs are largest.

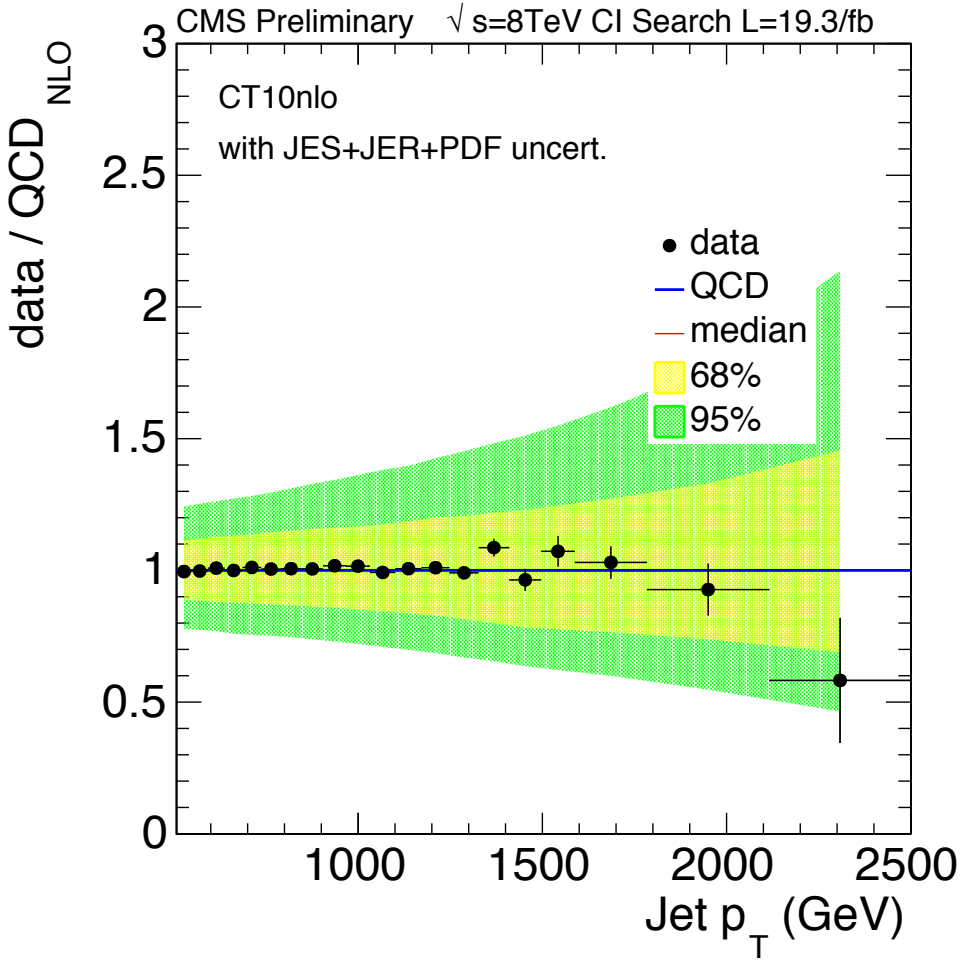
- Moreover, at present, the PDF uncertainties are comparable to the JES uncertainties and will become dominant in the early part of Run II.

CT10nlo PDF Uncertainties

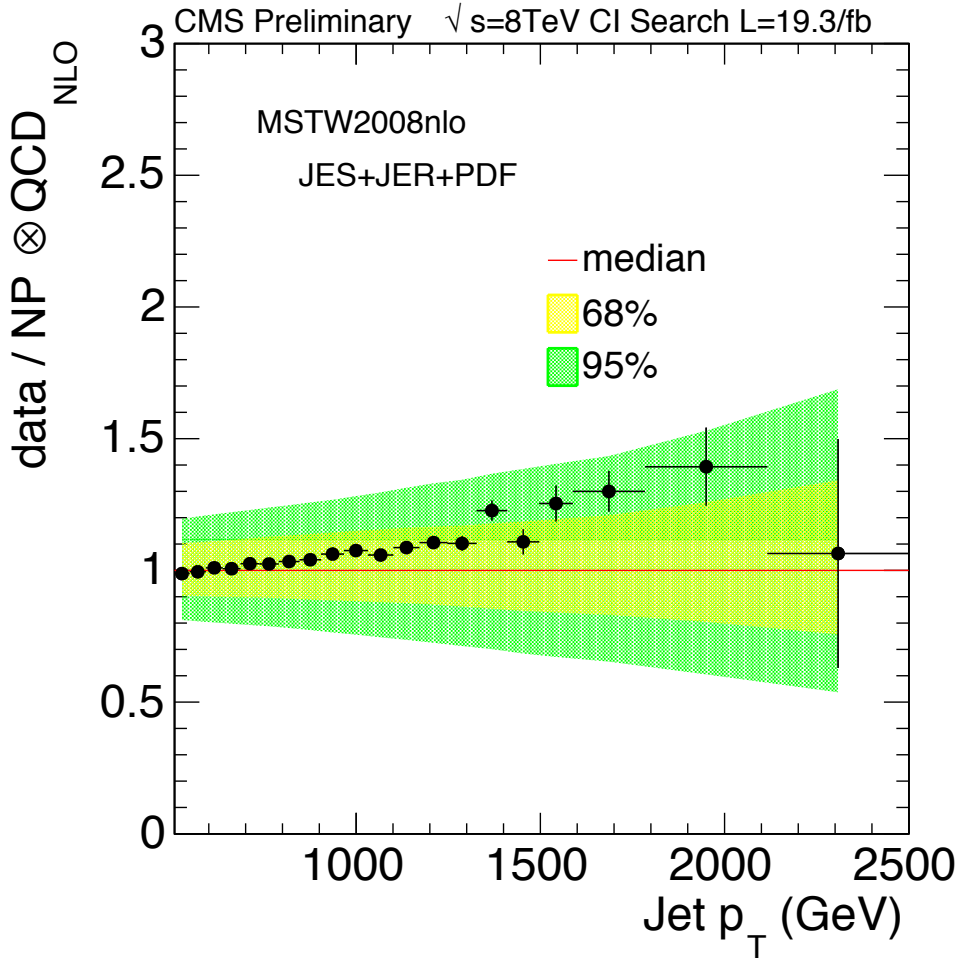




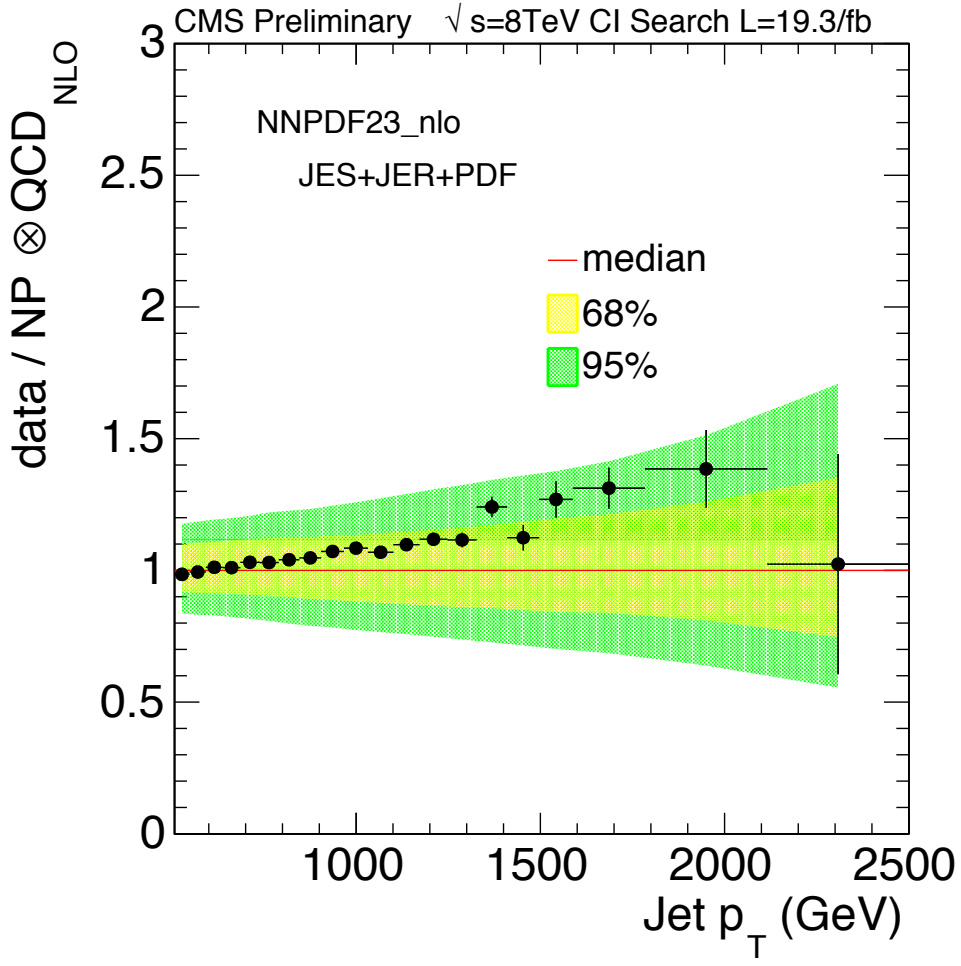
CT10nlo PDF Uncertainties



MSTW2008nlo PDF Uncertainties



NNPDF23 PDF Uncertainties



Issues with PDF Fits

Currently, all PDF fitting groups extract PDF parameters by minimizing a quadratic form (typically, referred to as χ^2)

$$Y(\theta) = [D - T(\theta)]^T \Sigma^{-1} [D - T(\theta)]$$

where θ are the parameters of the PDFs, and Σ is the covariance matrix of the data D, which are either real or replicas, X, generated from real data by sampling from

$$X \sim \text{Gauss}\{-[X - D]^T \Sigma^{-1} [X - D] / 2\}$$

The criterion $\Delta Y = Y(\theta) - \min Y(\theta) = T$ with $T = 1$ yields unreasonably small 68% confidence sets. Therefore, CTEQ and MSTW use a value of $T \sim 50 - 100$.

Issues with PDF Fits

- There is no consensus about the value of T nor about which of the following
  - discrepant data
  - incomplete theory
  - PDF parameterization

  is the principal reason why T must be ~ 50 – 100.

- It is not clear (at least to me) why NNPDF and MSTW arrive at similar uncertainties. Nor is it clear that NNPDF has solved the problem of PDF uncertainties in a statistically sound way.

Issues with PDF Fits

The basic question that no one has answered is this:

How do we know that the PDF uncertainties are "correct"?

And this cannot be answered sensibly without first answering:

What does "correct" mean?

Summary

- Owing to the advances in LHAPDF6, the CI search is now able to use multiple PDF models in a well-founded manner. The old PDF4LHC recommendation should be retired.

- Unfortunately, PDF uncertainties are now, or soon will, the dominant uncertainty in analyses that use jets of high p_{T}.

- However, in spite of impressive advances in PDF theory and methodology, more work is urgently needed to put PDF uncertainties on a statistically solid foundation.