

# Studies of $W + \text{jets}$ and Prompt Diphoton Production with the CDF Detector

## Tests of enhanced leading order QCD in $W$ boson plus jets events from 1.96 TeV $p\bar{p}$ collisions

Soushi Tsuno<sup>1</sup> for the CDF Collaboration

Institute of Physics, University of Tsukuba, Tsukuba, Ibaraki, Japan, e-mail: [tsuno@fnal.gov](mailto:tsuno@fnal.gov)

Received: October 9, 2003 / Revised version: date

**Abstract.** We have studied the  $W + \geq n$  jets process in Tevatron Run II experiment. This is the first result for the CDF Run II experiment. The data used corresponds to a total integrated luminosity of  $72 \text{ pb}^{-1}$  taken from Mar.2002 through Jan.2003. The lowest order QCD predictions have been tested with a new prescription of the parton-jet matching, which allows to construct the enhanced LO phase space. We found a good agreement between data and theory in the typical kinematics distributions. Number of events for each inclusive samples up to 3 jets are compared with Monte Carlo calculations. The comparison with Run I results is also presented.

**PACS.** 13.87.Ce Jets Production in large- $Q^2$  scattering

## 1 Introduction

Understanding the jet production mechanism is remarkable topics, because most of the search channels like Higgs boson or SUSY particles contain one jet or two more jets in that signature of the final state. In the top quark physics which is the main physics program at Tevatron Run II, it is the most crucial part of the analysis to measure the mass or production cross section. The  $W$  boson plus jets process thus has been a practical analysis process in hadron colliders not only for the dominant background of the most of precision measurements but also for the probing sample for new physics, due to an ease triggering of a high transverse momentum of the leptons from their bosons.

In this study, we present a comparison with data and theory calculated by the lowest order perturbative calculation. The fixed cone jet algorithm is used to identify a jet. To avoid a theoretical ambiguity of the collinear/infrared enhancement at the lowest order calculation, we apply a parton-jet matching procedure by requiring the clear definition of the parton separation, where each parton distributes within the particular cone size of a jet with an assumption that the doubly counted phase space will happen presumably in the collinear region, as well as the merging/splitting procedure of the cone jet algorithm. This is the first result for the CDF Run II experiment.

## 2 Data set

The CDF is successfully taking the colliding data since 2002. The data used in this analysis corresponds to a to-

tal integrated luminosity of  $72.0 \text{ pb}^{-1}$  taken from Mar.2002 through Jan.2003. The high- $p_T$  electron triggered samples are used. After good qualitative cuts on an isolated high- $p_T$  electron and a requirement of an imbalance of calorimeter energy due to the undetected neutrino (missing  $E_T$ ), the JetClu algorithm [1] is used to collect/count a jet. The transverse energy and pseudo-rapidity ( $\eta$ ) coverage of jets are required as

$$E_T \geq 15 \text{ GeV} \quad , \quad |\eta| \leq 2.4 \quad (1)$$

The clustering cone size is 0.4. The merging/splitting criteria is followed by the Jet Separation Method [2] which requires the iterative separation cone between two jets with 95% separation probability estimated by the two partons at the lowest order calculation. We collect the jets samples inclusively, that is, group the  $W + \geq n$  jets event samples, where, for instance, an event which has 2 jets is a member of the  $W + \geq 2$  jets event sample but at the same time it can be a member of the  $W + \geq 1$  jets event sample.

## 3 Comparisons of Theory to Data

### 3.1 Jet $E_T$ distribution

The jet transverse energy  $E_T$  is presented in Figure 3.1 for each jet process. From the upper-most side, the distributions are the highest  $E_T$  in  $W + \geq 1$  jets events, the second highest  $E_T$  in  $W + \geq 2$  jets events, and so forth. The data points are presented as a circle dot. The statistical error is only included in this data point. The shade

band among the data point is estimated by the fluctuation of the 10% jet energy scale uncertainty. The solid and dashed lines are the LO QCD predictions, except in  $W \geq 4$  jets events, produced by GR@PPA [3] event generator with the energy scale of the squared mass of a  $W$  boson ( $M_W^2$  (GeV $^2$ )) and the square of the average value of the parton  $p_T$  ( $\langle p_T \rangle^2$  (GeV $^2$ )), respectively, where the renormalization and factorization scales are equivalent denoted as the energy scale. The LO QCD prediction in  $W \geq 4$  jets events is produced by Alpgen [4] event generator with the energy scale of  $M_W^2 + p_{TW}^2$ . Those Matrix Element-based event generators are embedded into HERWIG [5] showering Monte Carlo simulation, and then the generated events are passed through the full detector simulation. The MC predictions are normalized by the total number of events in each  $W + \geq n$  jets data sample.

For the MC prediction, the energy scale of  $\langle p_T \rangle^2$  varies with the parton  $p_T$ 's in event by event. The lower energy scale is enhanced by the larger strong coupling  $\alpha_s$  since the size of a strong coupling constant increases with the lower energy scale. The shape of the jet  $E_T$  distribution thus depends on an order of magnitude of the  $\alpha_s$  by event basis. At least, the requirement of the parton-jet matching, one parton in one cone in the LO calculation, gives a clear logic that the size of the  $\alpha_s$  directly translates the jet  $E_T$  shape. Hence, we can expect that the jet  $E_T$  distribution arises a sensitivity to the choice of the energy scale. We can see the steeper decline of the  $E_T$  distribution in the case of  $\langle p_T \rangle^2$  than that of  $M_W^2$ . The choice of the energy scale  $M_W^2$  is useful as a good bench mark point to compare not only to the different energy scale but also to the higher order calculation because the running strong coupling constant (scale running) is less meaningful in the NLO calculation. The both MC predictions is a good agreement with the data. The choice of  $\langle p_T \rangle^2$  seems to be better to describe the data well, but is not clear due to the large jet energy uncertainty.

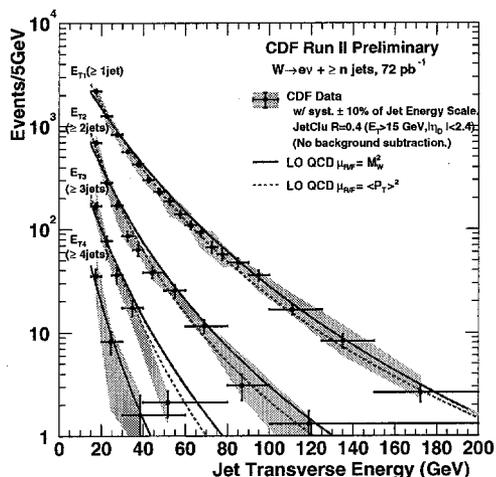


Fig. 1. Jet transverse energy. From the up-most side, the distributions are the highest  $E_T$  in  $W + \geq 1$  jets events, the second highest  $E_T$  in  $W + \geq 2$  jets events, and so forth.

### 3.2 Angular and mass distributions

The invariant mass and angular distribution ( $\Delta R_{jj}$ ) between two jets is a sensitive variable to the collinear/infrared singularity. Some differences may be an indicator to the higher order perturbative calculation. In Figure 2, we present the dijet mass distribution and angular distribution between the highest  $E_T$  jet and the second highest  $E_T$  jet in the  $W + \geq 2$  jets events and the  $W + \geq 3$  jets events, respectively.

A discrepancy in both mass distributions of  $W + \geq 2$  jets and 3 jets events in the data and MC predictions can be seen in this plot. The mass distributions of MC predictions are harder than those of the data. The distribution is better reproduced by the energy scale of  $\langle p_T \rangle^2$ . On the other hand, the  $\Delta R_{jj}$  distributions are insensitive to the energy scale. These features could be seen in Run I measurement [6]. We see that the theory predictions for the  $\Delta R_{jj}$  distribution remain valid to the resolution limit of jet-jet separation for our analysis.

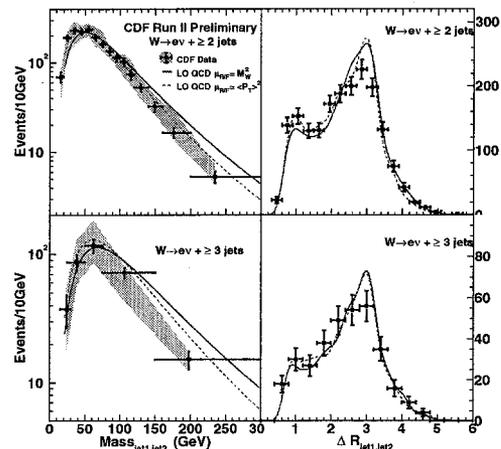


Fig. 2. Dijet mass distribution and jet separation angle between the highest  $E_T$  jet and the second highest  $E_T$  jet in  $W + \geq 2$  jets events and  $W + \geq 3$  jets events, respectively.

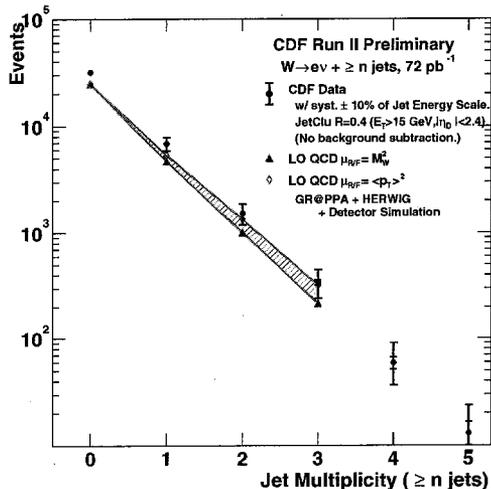
### 3.3 Jet multiplicity

Using the cross section of the MC, we can compare the number of jets distribution with the data. We present the jet multiplicity distribution in Figure 3. The errors on the data points are the sum of the statistical and systematic uncertainty by the jet  $E_T$  scale. The lower and upper band on the LO QCD predictions correspond to the energy scale of  $M_W^2$  and  $\langle p_T \rangle^2$ , respectively. All the acceptance or detection efficiencies are already included into the number of detected events because the MC events are also passed through the detector simulation. The lower energy scale of  $\langle p_T \rangle^2$  yields higher cross sections since it correlates with a larger value of  $\alpha_s$ . We have also plotted the leading order theory prediction for the inclusive  $W$  production

cross section ( $W + \geq 0$  jets) by HERWIG built-in process. On this plot, we did not consider any background contributions. However, those background contaminations are almost negligible in the  $W + 0,1,2,3$  jets events. Indeed, those fractions are  $\sim 2.8\%$ ,  $\sim 4.4\%$ ,  $\sim 4.7\%$ , and  $\sim 10.1\%$  in the  $W + 0,1,2,3$  jets events, respectively.

The ambiguity for the unphysical parameter like the kinematic cuts on the generator level has been already rejected by the requirement of the parton-jet matching. Since there is only one parton from the ME calculation in the jet cone, the number of jets is proportional to the number of partons, that is, an order of the strong coupling constant. We can see almost linear relation of the jet multiplicity in both the data and MC's. This shows our analysis method well describes the enhance lowest order phase space. The difference of the absolute cross section will be addressed as a lack of the higher order calculations.

It is useful to see the fraction of the jet multiplicity presented in Figure 4. The number of events in each  $n$  jet bin is normalized by the number of events in  $W + \geq 0$  jets events. The relative size of the jet multiplicity on the data is well reproduced by the LO calculations.



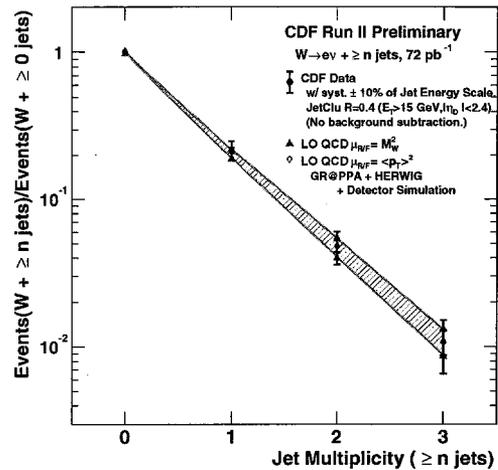
**Fig. 3.** Jet multiplicity distribution. The errors on the data points are the sum of the statistical and systematic uncertainty by the jet  $E_T$  scale. The lower and upper band on the LO QCD predictions correspond to the energy scale of  $M_W^2$  and  $\langle p_T \rangle^2$ , respectively.

### 3.4 Ratio of the jet multiplicity

We show various ratio plots to each jet bin in  $W + \geq n$  jets events in Figure 5. From the top, the ratio of theory to data, the ratio of  $n$  jets events to  $n-1$  jets events, and the ratio to the ratio of  $n$  jets events to  $n-1$  jets events,

$$R_{n/(n-1)} = \frac{\sigma_n}{\sigma_{n-1}}, \quad (2)$$

are presented. Taking the ratio of the physics variable is to cancel out the uncertainties from the absolute source like



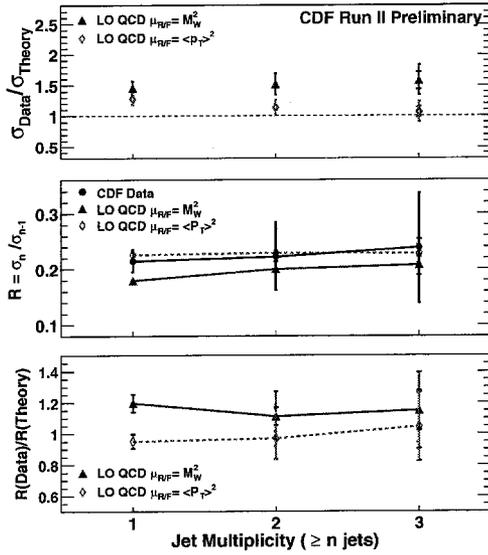
**Fig. 4.** Fraction of the jet multiplicity. The number of events in each  $n$  jet bin is normalized by the number of events in the  $W + \geq 0$  jets events.

the luminosity. The identification efficiency or acceptance etc. may also cancel somehow out.

We see that the absolute cross section predictions agree with the data less than factor 2. Those factors are  $\sim 1.2$  for the energy scale of  $\langle p_T \rangle^2$  and  $\sim 1.5$  for  $M_W^2$  over the range to the  $W + \geq 3$  jets events, respectively. Remarkable feature is that the MC predictions are showing the almost constant behavior on this ratio plot. That means that our analysis method and MC prediction well describe the data. It is interesting to see the ratio  $R_{n/(n-1)}$  (middle). The jet counting uncertainties will be reduced except for  $R_{10}$ . The  $R_{n/(n-1)}$  comparison is valid if higher order QCD corrections to the LO cross sections are not strongly dependent on the number of final state partons. The ratio  $R_{n/(n-1)}$  measures the decrease in cross section with the addition of 1 jet. The value of  $R_{n/(n-1)}$  thus is clearly dictated by the magnitude of the strong coupling constant since adding an extra jet adds a factor of  $\alpha_s$ . We can see the energy scale  $\langle p_T \rangle^2$  is a better agreement than the  $M_W^2$ . In the  $R_{n/(n-1)}$  plot, the particular value of  $R_{n/(n-1)}$  will vary as a function of the specific jet  $E_T$  requirement that define a jet. To remove this dependence to some degree, the ratio (bottom) of data and theory for  $R_{n/(n-1)}$  give a sensitivity to an independent comparison of the jet definition and its systematics. With accurate theory predictions and accurate data measurements the value of this ratio is 1.0. If the QCD predictions reproduce the jet kinematics accurately, the ratio of data to theory is independent of the choice of the jet  $E_T$  requirement so that the quantity may be of more general interest.

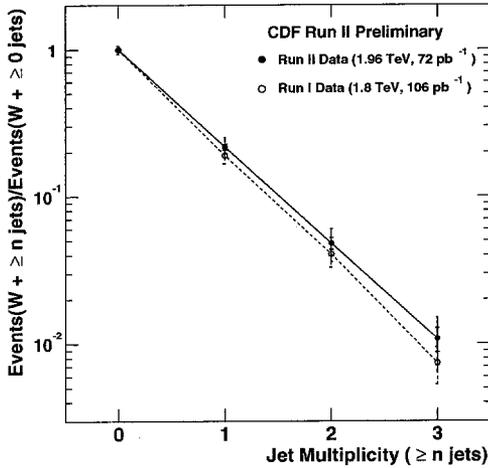
### 3.5 Comparison with Run I measurement

It is practical to compare with the Run I measurement. We present the fraction of jets in Figure 6. The fraction of the Run II data is slightly larger than the Run I results. This is not obvious feature. The upgraded collision energy of 1.96 TeV and higher instantaneous luminosity will make



**Fig. 5.** Ratio of jet multiplicity. From the top, the ratio of theory to data, the ratio of  $n$  jets events to  $n-1$  jets events, and the ratio to the ratio of  $n$  jets events to  $n-1$  jets events are shown.

larger size of the production cross section and fake backgrounds. The fraction of the jet multiplicity normalized by  $W + \geq 0$  jets events however should be independent of the collision energy. One of the doubtful source is the different definition of a jet and a size of backgrounds. In our analysis, no background subtraction is applied. The detailed study will be needed for more accurate measurement.



**Fig. 6.** Fraction of jets with a comparison to Run I measurement.

## 4 Conclusion

Data have been compared to the theory predictions at the lowest order perturbative calculation level. Jet Separation

procedure based on the parton-jet matching requirement is used for the data and theoretical predictions. This requirement is to construct the enhanced LO phase space. For the theory prediction, two choices of the energy scale,  $\langle p_T \rangle^2$  and  $M_W^2$ , where the renormalization and factorization scales are equivalent, has been tested. The jet transverse energy, mass and jet-jet separation distribution are compared with data and theory predictions. All the distributions are good agreements. The choice of the energy scale of  $\langle p_T \rangle^2$  is preferred to describe data well. Jet multiplicity distribution is also compared up to  $W + \geq 3$  jets events. Non-background condition is assumed. But those background effects are almost negligible in the less than 3 jets events. The constant (flat) behavior can be seen in the various ratio plots. This is very important feature to certify our rightness of the MC generation and analysis scheme, which will be crucial for the measurement of the strong coupling constant. In the comparison with Run I results, Run II results were slightly larger than Run I results. We'd also like to mention that the NLO event generator is a key point for this study.

## References

1. F. Abe *et al.*, *Phys. Rev. Lett.* **74** (2626) 1995.
2. S. Tsuno, in preparation.
3. S. Tsuno *et al.*, *Comput. Phys. Commun.* **151** (2003) 216.
4. M.L. Mangano *et al.*, *J. High Energy Phys.* **07** (2003) 001.
5. G. Marchesini and B. R. Webber, *Nucl. Phys.* **B 310** (1988) 461;  
G. Marchesini *et al.*, *Comput. Phys. Commun.* **67** (1992) 465.
6. T. Affolder *et al.*, *Phys. Rev.* **D 63** (2001) 072003.