DESY 10-026
March 2010

ISSN 0418-9833

arXiv:1003.1854v2 [hep-ph] 19 May 2010

Factorization breaking in high-transverse-momentum charged-hadron production at the Tevatron?

S. Albino, B. A. Kniehl, G. Kramer

II. Institut für Theoretische Physik, Universität Hamburg,
Luruper Chaussee 149, 22761 Hamburg, Germany.

Abstract

We compare the transverse momentum (p_T) distribution of inclusive light-charged-particle production measured by the CDF Collaboration at the Fermilab Tevatron with the theoretical prediction evaluated at next-to-leading order in quantum chromodynamics (QCD) using fragmentation functions recently determined through a global data fit. While, in the lower p_T range, the data agree with the prediction within the theoretical error or slightly undershoot it, they significantly exceed it in the upper p_T range, by several orders of magnitude at the largest values of p_T, where perturbation theory should be most reliable. This disagreement is too large to be remedied by introducing additional produced particles into the calculation, and potentially challenges the validity of the factorization theorem on which the parton model of QCD relies. Clearly, a breakdown of the factorization theorem, being a fundamental property of QCD, would be extremely difficult to understand.

PACS: 12.38.Cy, 12.39.St, 13.66.Bc, 13.87.Fh

Inclusive single-hadron production processes in e^+e^-, ep, and $pp(\bar{p})$ reactions can be well described [1, 2] by cross section calculations that depend on universal fragmentation functions (FFs), consistent with the validity of the factorization theorem at CERN LEP, BNL RHIC, and DESY HERA energies (for a recent review, see Ref. [3]). However, in order to prepare for incoming data at much higher energy from the CERN LHC, from the late 2009 run and/or from future runs, this validity needs to be tested also at the Tevatron. Such tests were performed in Ref. [1], using CDF data for which p_T reached 10 GeV [4] and also UA1 and UA2 data from CERN S$p\bar{p}$S runs at sub-TeV energies. Very recently, new data have been obtained by the CDF Collaboration [5] for which p_T reaches 150 GeV. Such data provide a unique opportunity to test the factorization theorem applied to the final state in a high-p_T regime of inclusive hadron production that has never been accessible before, where we expect perturbation theory in the QCD-improved parton model to work perfectly, and these data are the focus of this Letter. We note that agreement between theory and recent CDF measurements of inclusive jet production [6] has been obtained at much higher p_T values, which confirms the inital-state factorization theorem for these data. Such measurements originate from the same physical processes as the lower-p_T hadron production measurements do, since the produced hadron carries just a fraction of the jet's energy.

We present here the first rigorous interpretation of the CDF data of Ref. [5] in the context of perturbative QCD. The experimental analysis resorted to the Monte Carlo event generator PYTHIA, which failed to give useful results for $p_T > 50$ GeV. We use the latest FF sets [9, 8, 7] extracted from global data fits that included accurate primary-quark-tagged measurements from e^+e^- reactions. Unfortunately, e^+e^- reaction data do not sufficiently constrain the gluon FF, because it only enters the calculation at next-to-leading order (NLO), while it enters the calculation of $p\bar{p}$ reactions at the same order as the quark FFs. The two most recent FF sets, AKK08 [7] and DSS [8], also used pp reaction data from RHIC, but the gluon FFs in both cases differ substantially, suggesting that the RHIC data used in those fits did not significantly improve the constraints on the gluon FFs. The FF errors in both sets were not determined. Therefore, the best estimate of the error on the predictions for CDF data at present is given by the general spread in the calculation using different sets. Besides the AKK08 and DSS FF sets, we also include the HKNS one [9], which was obtained from a fit only to e^+e^- data. We assume that the unidentified charged particle in these CDF measurements is a light charged hadron. That is, we take the sum of the $\pi^\pm$, $K^\pm$ and $p/\bar{p}$ production cross sections, and assume the possible contamination with other charged particles to be negligible.

Our calculations are performed to NLO in the modified minimal-subtraction ($\overline{\rm MS}$) renormalization and factorization schemes. We set the renormalization scale to be $\mu = kp_T$ and the factorization scale to be $\mu_f = k_f p_T$, where $k = k_f = 1$ unless otherwise stated. For the initial protons, we use the parton density function (PDF) set CTEQ6.6M [10]. Since the produced-hadron mass was accounted for only in the AKK08 analysis, we account for it in the same way here for calculations using the AKK08 FF sets, but set it to zero when using the DSS and HKNS sets. In any case, neglecting the hadron masses makes very little difference, as we shall see later. We refer to the above calculation as the AKK08

default one (labeled "AKK08(default)" in our figures).

The comparison with the recent data [5] collected in Run II (with center-of-mass energy $\sqrt{s} = 1.96$ TeV) in the central region of the CDF detector (with pseudorapidity $|\eta| < 1$) and older CDF data [4] from Run I ($\sqrt{s} = 1.8$ TeV, $|\eta| < 1$) is shown in Fig. 1. The most striking observation to be made is the disagreement at $p_T > 50$ GeV (see left plot). Looking closely at the experimental line shape, there appears to be a transition to a softer slope, corresponding to a fall-off with a lower power in $1/p_T$, at about $p_T = 20$ GeV, which eventually leads to a systematic departure from the theoretical prediction. In fact, fitting the function A/p_T^n to the CDF data yields $A = 71.50$ and $n = 6.52$ with $\chi^2/\text{d.o.f.} = 0.886$ for 5 GeV $< p_T <$ 20 GeV and $A = 0.030$ and $n = 3.97$ with $\chi^2/\text{d.o.f.} = 2.18$ for 20 GeV $< p_T <$ 150 GeV. We evaluate the theoretical error, indicated by the shaded region labeled "AKK08(Min/Max)" in Fig. 1 and later plots, by plotting, for each p_T value, the maximum and minimum cross sections in the range $1/2 < k, k_f < 2$. To achieve this, we use a 17×17 grid of points equidistant in $\ln k_{(f)}$ ranging from $-\ln 2$ to $\ln 2$. As Fig. 1 shows, the theoretical error calculated in this way is not nearly sufficient to accommodate the recent CDF data at large p_T. Within this error, the calculation in the region $p_T \lesssim 40$ GeV is consistent with most of the data. In the region $p_T \lesssim 20$ GeV, the data lie in the lower range of the theoretical error, and in fact favour a large value of $k = k_f$, as can be seen by the dashed curve in Fig. 1, which corresponds to $k = k_f = 4$. The error is largest at small $p_T = \mathcal{O}(1)$, where both $\alpha_s(\mu)$ and the partonic cross sections become very large, and for this reason we limit our discussion to those data for which $p_T > 2$ GeV.

In Fig. 2, the same calculation is performed with the DSS [11] and HKNS [9] FF sets. We note that FFs for unidentified charged hadrons were also obtained by the authors of Ref. [11]. However, the difference between the calculation using those FFs and the one using the sum of the DSS FF sets for $\pi^\pm$, $K^\pm$, and $p/\bar{p}$ production (which respectively make up about 75, 15, and 10% of the sample) turns out to be negligible over the whole p_T range considered. Assuming that the difference between the calculations using the AKK08, DSS, and HKNS FF sets gives an estimate of the error due to the uncertainties in the FFs, we conclude that the latter are not nearly sufficient to account for any discrepancy between theory and the new CDF data [5]. We stress that all these FF sets are tightly constrained in the very region of the (z, μ_f) parameter space that is probed by our predictions, in particular where the discrepancy occurs, so that no extrapolation error arises. In fact, the average value of z lies in the range $0.44 \lesssim \langle z \rangle \lesssim 0.65$ for 2 GeV $< \mu_f = p_T <$ 150 GeV. The wealth of data from TRISTAN, LEP1, SLC, and LEP2, which populate the range 58 GeV $< \mu_f = \sqrt{s} <$ 191 GeV, are quite confining in this intermediate z range. We conclude that the error due to the uncertainties in the FFs is small against the overall one.

In Fig. 3, we investigate the effect of the produced hadron's mass m_h, by comparing our default prediction with the one obtained by setting all particles' masses to zero. The difference also gives an estimate for the size of any neglected small-p_T effect of a similar order of magnitude, namely $O(m_h^2/p_T^2)$, such as higher twist. Clearly, such effects are negligible relative to the theoretical uncertainty shown in Fig. 1.

As mentioned above, the gluon FF is somewhat less well constrained than the quark FFs. In order to determine whether the discrepancy with the recent CDF data [5] could be due to large errors on the gluon FF, we also show in Fig. 3 the contribution coming from quark fragmentation only (labeled "no gluon"). We conclude that a reasonable modification to the gluon FF may improve the description of the low-p_T data, but certainly not at the largest p_T values where gluon fragmentation becomes negligible.

The only other non-perturbative quantities in our analysis besides FFs are PDFs. Also these are tightly constrained by global analyses in the range of x and μ_f probed by the CDF data. In order to illustrate this, we study in Fig. 4 the variation due to the choice of PDF set, by comparing our default result based on the CTEQ 6.6M PDF set [10] with the calculations using the MSTW2008 [12] and HERAPDF0.1 [13] ones. All these PDF sets have been constrained in very different ways. Specifically, the HERAPDF0.1 PDF set [13] was obtained from a combined fit to HERA data, which puts strong constraints on the gluon PDF at low x. Clearly this type of error is the least significant of all.

Finally, we compare with the recent accurate data from STAR [14] ($\sqrt{s} = 200$ GeV, $|y| < 0.5$) in Fig. 5, and find perfect agreement with our default prediction, well within the theoretical error. By contrast, the CDF data from Runs I [4] and II [5] tend to undershoot the theoretical predictions in the same p_T range (see Fig. 1). A possible explanation for the worse agreement at the larger $\sqrt{s}$ values is that perturbation theory eventually fails if the fraction of available momentum $z = 2p_T \cosh y/\sqrt{s}$ taken away by the produced hadron becomes too low.

In conclusion, the p_T distribution of inclusive charged-hadron hadroproduction recently measured by the CDF Collaboration [5] at the Tevatron was confronted here for the first time with rigorous NLO predictions relying on the factorization theorem of QCD, which, unlike MC event generators, have no left-over adjustable parameters. Surprisingly, a spectacular disagreement was discovered, by up to 3 orders of magnitude, at the largest p_T values, where perturbation theory is expected to work most reliably. In fact, the χ^2 per degree of freedom turned out to be as large as 7.3 for the last 4 data points. Perfoming a careful error analysis, we estimated the overall uncertainty at the largest p_T values to be about $\pm 30\%$, being mainly due to unknown corrections beyond NLO manifesting themselves in residual scale variations. In fact, the theoretical uncertainty actually reduces with increasing value of p_T simply because of asymptotic freedom causing the strong-coupling constant to fade, so that the high-p_T region is where we would expect the agreement with the data to be the best. Being at most of $O(\Lambda_{\rm QCD}/p_T)$, higher-twist effects are bound to be negligibly small there. The only large logarithms that need to be resummed are those in $\alpha_s(\mu)$, the PDFs, and the FFs, which are duly taken care of by the renormalization group and evolution equations, respectively. The partonic cross sections are devoid of large logarithms requiring resummation, which is ensured by the scale choice $\mu = \mu_f = O(p_T)$. Because $\langle z \rangle \approx 0.5$, no large logarithms of the types $\ln z$ or $\ln(1-z)$ need to be resummed, and resummation may not even be appropriate. Because even the highest p_T bin is far away from the edge of phase space, no high-p_T resummation is required either. A different philosophy of error analysis may lead to a somewhat different result, but certainly not to an error of $O(10^5\%)$, required to remove the disagreement. The latter is even far too

large to be remedied by introducing new heavy virtual particles with resonating propagators or charged particles in the final state. Some more drastic modification either to the theoretical preduction or to the experimental data is required. We do not claim that the discrepancy discovered by us actually challenges the factorization theorem, nor that the data is wrong, nor that new physics is at work. We do not even wish to speculate about possible problems in the experimental analysis, such as the suppression of the cosmic-ray background, etc. We only state that there is a discrepancy, whatever its reason might be. We recall that no such discrepancy was observed in the inclusive single-jet production data of the same experiment [6] at considerably higher p_T values. High-p_T data from the LHC is needed to verify whether this apparent final-state factorization breaking is a genuine phenomenon. We note that the recent inclusive charged-hadron production data from the ATLAS and CMS Collaborations [15] at the LHC are limited to the region $p_T < 20$ GeV. Unfortunately, they are presented as yields differential in p_T and cannot be converted to cross section distributions because the conversion factors are not presently known.

This work was supported in part by BMBF Grant No. 05H09GUE, by DFG Grant No. KN 365/5–3, and by HGF Grant No. HA 101.

Note added: We were informed by the CDF Collaboration that they are currently reanalyzing their published data [5]. In the meantime, four more papers [16] on the same subject have appeared.

References

[1] B. A. Kniehl, G. Kramer, and B. Pötter, Nucl. Phys. **B597**, 337 (2001).

[2] S. Albino, B. A. Kniehl, G. Kramer, and C. Sandoval, Phys. Rev. D **75**, 034018 (2007).

[3] S. Albino, arXiv:0810.4255 [hep-ph], to be published in Rev. Mod. Phys.

[4] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **61**, 1819 (1988).

[5] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. D **79**, 112005 (2009).

[6] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. D **78**, 052006 (2008); **79**, 119902(E) (2009); D0 Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. **101**, 062001 (2008).

[7] S. Albino, B. A. Kniehl, and G. Kramer, Nucl. Phys. **B803**, 42 (2008).

[8] D. de Florian, R. Sassot, and M. Stratmann, Phys. Rev. D **75**, 114010 (2007).

[9] M. Hirai, S. Kumano, T. H. Nagai, and K. Sudoh, Phys. Rev. D **75**, 094009 (2007).

[10] CTEQ Collaboration, P. M. Nadolsky *et al.*, Phys. Rev. D **78**, 013004 (2008).

[11] D. de Florian, R. Sassot, and M. Stratmann, Phys. Rev. D **76**, 074033 (2007).

[12] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009).

[13] H1 and ZEUS Collaborations, DESY Report Nos. H1prelim–08–045, ZEUS–prel–08–003.

[14] Y. Xu, Eur. Phys. J. C **62**, 187 (2009); private communication.

[15] CMS Collaboration, M. Abbrescia *et al.*, JHEP 1002, 041 (2010); ATLAS Collaboration, G. Aad *et al.*, arXiv:1003.3124 [hep-ex].

[16] F. Arleo, D. d'Enterria, and A. S. Yoon, arXiv:1003.2963 [hep-ph]; M. Cacciari, G. P. Salam, and M. J. Strassler, arXiv:1003.3433 [hep-ph]; A. S. Yoon, E. Wenger, and G. Roland, arXiv:1003.5928 [hep-ph]; B.L. Ioffe, arXiv:1005.1078 [hep-ph].

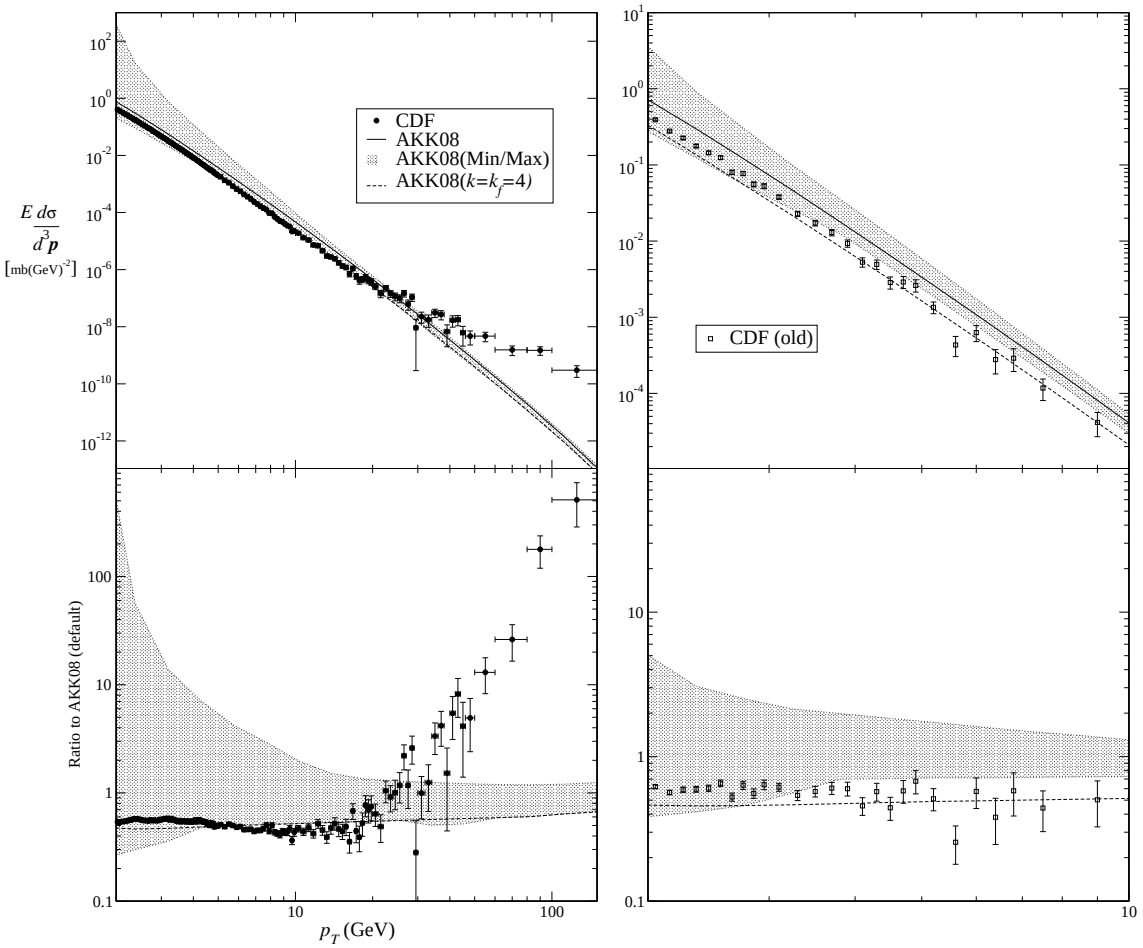

Figure 1: **Top left**: The CDF data of Ref. [5] for inclusive unidentified-charged-particle production is compared with NLO QCD predictions evaluated using the AKK08 FF sets [7]. The shaded region labeled "AKK08(Min/Max)" indicates, for each p_T value, the maximum and minimum values of the cross section within the ranges $1/2 < k < 2$ and $1/2 < k_f < 2$. Also shown is the result for $k = k_f = 4$. **Top right**: As in the top left plot, except for the older CDF data of Ref. [4]. Each lower plot is identical to the one above it, except that all results are divided by the central NLO predictions for clarity. (The latter hence correspond to unity.)

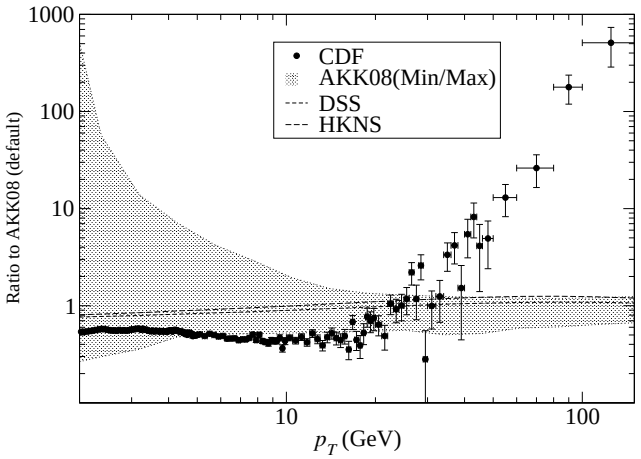

Figure 2: As in Fig. 1 (bottom left). Also shown is the calculation performed using the DSS [11] and HKNS [9] FF sets.

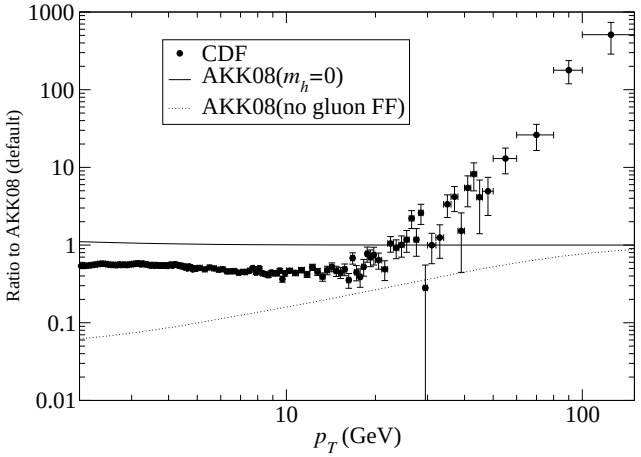

Figure 3: As in Fig. 1 (bottom left) for the data. The AKK08 calculation is performed with all hadrons' masses set to zero (labeled "$m_h = 0$") and again without the contribution from the evolved gluon FF ("no gluon").

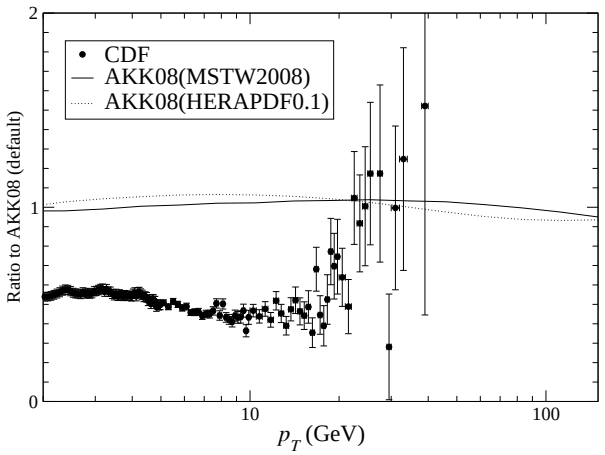

Figure 4: As in Fig. 1 (bottom left) for the data. The AKK08 calculation is performed using the PDF sets of Ref. [12] (labeled "MSTW") and Ref. [13] (labeled "HERAPDF0.1"). The data for which $p_T > 50$ GeV lie far above the calculation.

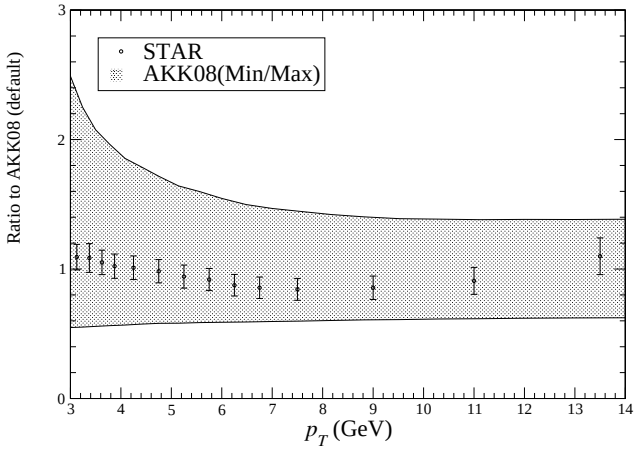

Figure 5: The STAR data of Ref. [14] for inclusive charged-hadron production is compared with NLO QCD predictions evaluated using the AKK08 FF sets [7].