Proceedings of the DIS'2004, Štrbské Pleso, Slovakia

BOTTOM PRODUCTION

STEFANO FRIXIONE
INFN, Sezione di Genova
Via Dodecaneso 33, 16146 Genova, Italy
E-mail: Stefano.Frixione@cern.ch

I briefly review the theory and phenomenology of bottom production at colliders. When all theoretical uncertainties are taken into proper account, and modern measurements are considered, no significant discrepancy is seen between data and QCD predictions

The physics of bottom quarks is one of the best-studied topics in particle physics. Experimentally, this is due to the abundance with which b quarks are produced at colliders. Theoretically, the reasons of interest are many. The characteristic that sets the b apart is its heaviness,

$$m_b \gg \Lambda_{QCD} \tag{1}$$

which entails peculiar properties. If one is interested in the phenomenology of the decays of the b-flavoured hadrons, eq. (1) suggests to treat the b as infinitely heavy in comparison with its companion light quark(s) in a bound state, paving the way to HQET and its symmetry properties. On the other hand, if one aims at studying the hard production mechanism, eq. (1) implies the possibility of computing the open-b cross section, which is free of collinear and infrared singularities order by order in perturbation theory (as opposed to, say, open-u cross section, whose final-state collinear singularities are cancelled only upon convolution with a non-perturbative fragmentation function). The bottom is also the heaviest quark which hadronizes before decaying, allowing us to test many of the ideas of the factorization theorems in a relatively clear environment.

To be definite, let me consider the b cross section in hadronic collisions, which can be written as follows:

$$d\sigma_{H_1H_2\to b\bar{b}}(S)=\sum_{ij}\int dx_1 dx_2 f_i^{(H_1)}(x_1) f_j^{(H_2)}(x_2) d\hat{\sigma}_{ij\to b\bar{b}}(\hat{s}=x_1x_2S), \tag{2}$$

and can be readily extended to other types of colliding particles. Here, $f_i^{(H)}$ are the parton distribution functions (PDFs), and $d\hat{\sigma}_{ij\to b\bar{b}}$ are the short-distance cross sections, the only pieces in eq. (2) that can be computed in perturbation theory. The LO term (of $\mathcal{O}(\alpha_S^2)$) is trivial to obtain. The NLO one (of $\mathcal{O}(\alpha_S^3)$) was the result of landmark calculations [1,2,3]. Not surprisingly, the NNLO term is not available at the moment; this may be worrisome, since at the NLO the scale dependence is still pretty large, and the corrections are 100% of the Born term. However, there are at least a couple of points that deserve more immediate attention than the lack of the NNLO contribution. The first is that, as in any other cross sections computed in perturbation theory, large logs can appear which spoil the "convergence" of the series. In other words,

$$d\hat{\sigma}=\sum_{i=2}^{\infty}a_i\alpha_S^i\,,\qquad a_i=\sum_{k=0}^{i-2}a_i^{(i-2-k)}\log^{i-2-k}\mathcal{Q}\,,\tag{3}$$

where $\mathcal{Q}$ generically indicates a "large" quantity, in the sense that $\alpha_S\log^2\mathcal{Q}\gtrsim 1$. The second problem is that, although theoretically well defined, the open-b cross section is not physically observable. In order to compare theoretical predictions with data, two things can be done. *a)* Hadron-level experimental data are deconvoluted, and presented in terms of parton-level "measurements" that can be directly compared with open-b results. These deconvolutions are typically performed by means of parton shower Monte Carlos by the experimental collaborations. *b)* The open-b cross section is convoluted with a non-perturbative fragmentation function (NPFF) $D^{b\to H_b}$ according to the following formula for the single-inclusive p_T spectrum

$$\frac{d\sigma(H_b)}{dp_T}=\int\frac{dz}{z}D^{b\to H_b}(z,\epsilon)\frac{d\hat{\sigma}(b)}{d\hat{p}_T},\qquad\qquad p_T=z\hat{p}_T\,.\tag{4}$$

$D^{b\to H_b}$ describes how a b quark transforms ("fragments") into a B hadron; it is not computable in perturbation theory but, being universal in the same sense as PDFs, can be fitted to data in a given type of collision (usually e^+e^-) and used anywhere else.

Common wisdom has it that neither strategy *a)* nor *b)* have been particularly successful, since Tevatron data (and, to some extent, SpS ones) have been shown to be systematically larger than NLO QCD predictions, regardless of whether they were presented in terms of b quarks or of B mesons. In a very recent CDF paper [4] on $B^\pm$ single-inclusive p_T spectrum, the discrepancy was quantified to be $2.9\pm 0.2\pm 0.4$. Taking these comparisons blindly, one is led to conclude that b physics is *the* problem of the SM, and offers the first glance beyond it [5]. Although this remains a viable possibility, it seems premature to buy it without first reassessing carefully all possible sources of mistakes in the past comparisons between theory and data, and considering the uncertainties that so-far uncalculated SM contributions can give. In particular, one should try to answer the following questions:

1) Do large logs spoil the convergence of the series?

2) Is the fragmentation/deconvolution performed appropriately?

But before getting into this, let me point out that, although the discrepancies between data and NLO QCD have been quoted to be large, this is mainly due to the failure to incorporate properly *all* the uncertainties, including the theoretical ones which are very large. Upon doing so one realizes that, on a statistically sound basis, most of the data lie withing 1σ from the default theoretical predictions, and very rarely the discrepancy exceeds the 2σ level. The interested reader can find an informative discussion in ref. [6].

The answer to question *1)* depends on the fact that the logarithms that grow potentially large can be divided into two classes. The first class includes those logs whose arguments don't depend on the observable being measured, such as

$$\mathcal{Q} = 1 - \frac{4m_b^2}{\hat{s}}, \quad \mathcal{Q} = \frac{m_b^2}{\hat{s}}, \tag{5}$$

which are known as threshold logs (relevant when the c.m. energy is not much larger than the quark mass), and small-x logs (relevant when the quark mass is negligible wrt the c.m. energy) respectively. Threshold logs are clearly not a factor at colliders; small-x logs are estimated to give up to 30% effects at the Tevatron [7]. Thus, the overall picture would not change if these logs were properly resummed and matched with NLO QCD predictions. The logs belonging to the second class do have arguments which are directly related to the observables. For example, when measuring the single-inclusive p_T spectrum, the $b\bar{b}$ p_T or their azimuthal distance in the transverse plane, logs of the following arguments are generated

$$\mathcal{Q} = \frac{p_T(b)}{m_b}, \quad \mathcal{Q} = \frac{p_T(b\bar{b})}{m_b}, \quad \mathcal{Q} = 1 - \frac{\Delta\phi(b\bar{b})}{\pi}, \tag{6}$$

and the perturbative series may be badly behaved when $p_T(b) \gg m_b$, $p_T(b\bar{b}) \simeq 0$, and $\Delta\phi(b\bar{b}) \simeq \pi$ respectively. Since single-inclusive p_T spectra are routinely measured, large $p_T(b)/m_b$ logs have a prominent role. The resummation of these logs, i.e. the rearrangement of the perturbative series of eq. (3) in the following form:

$$\frac{d\sigma}{dp_T^2} = \alpha_S^2 \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} r_i^{(j)} \alpha_S^j \left(\alpha_S \log \frac{p_T^2}{m_b^2}\right)^i + \mathrm{PST} \tag{7}$$

can be achieved by using the perturbative fragmentation functions computed in ref. [8]. Each term in the sum which runs over j corresponds to a given logarithmic accuracy ($j = 0$ is LL, $j = 1$ is NLL, and so on), and PST stands for power suppressed terms, i.e. terms which vanish in the limit $m/p_T \to 0$. Typically, eq. (7) is computed up to the NLL; since PST are neglected, the quark behaves as if massless, which implies that predictions based on eq. (7) must strictly be used only when $p_T \gg m_b$. The trouble is that $p_T \gg m_b$ is not a quantitative statement, and often resummed predictions are compared to data even for $p_T \lesssim m_b$. It should be clear that such a comparison is void of sense, and any agreement between theory and measurement must be regarded as accidental. In any event, a matched computation (FONLL) was proposed in ref. [9], which combines the virtues of the fixed order and of the resummed formulae:

$$\frac{d\sigma}{dp_T^2} = a_2\alpha_S^2 + a_3\alpha_S^3 + \alpha_S^2 \sum_{i=2}^{\infty} r_i^{(0)} \left(\alpha_S \log \frac{p_T^2}{m_b^2}\right)^i + \alpha_S^3 \sum_{i=1}^{\infty} r_i^{(1)} \left(\alpha_S \log \frac{p_T^2}{m_b^2}\right)^i \tag{8}$$

and can be used to get sensible predictions in the whole p_T range. Armed with eq. (8), one can answer question *1)* in a quantitative way; it turns out that, in the p_T range probed at the Tevatron, the effects are moderate. As discussed in ref. [10], the NLO cross section used in ref. [4] is only about 20% lower than the FONLL one.

Let me therefore consider question *2)*, and for simplicity discuss the case of fragmentation rather than that of deconvolution. The master equation is (4). The

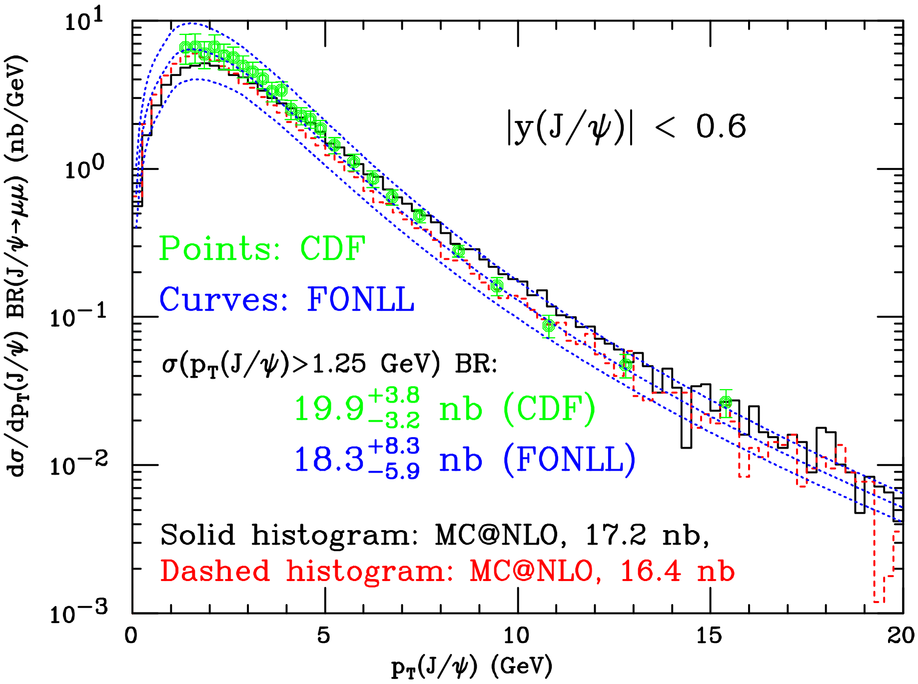

Figure 1. Comparison of CDF Run II data [13] with MC@NLO [14,15] and FONLL [9] predictions.

crucial point is that, while the l.h.s. of this equation is a measurable quantity, neither of the two terms on the r.h.s. is measurable. This is easy to understand if one considers that eq. (4) is used to extract the NPFF from data; the l.h.s. is measured, the short-distance cross section $d\hat{\sigma}$ is computed, and eq. (4) is solved for NPFF: symbolically, NPFF$=d\sigma(H_b)/d\hat{\sigma}(b)$. Thus, if one uses $d\hat{\sigma}$ computed at the LO, the resulting NPFF will clearly differ from the one obtained by computing $d\hat{\sigma}$ to, say, NLO. Notice that this doesn't contradict the universality property of the NPFF; this property merely states that a given NPFF stays the same regardless of the type of hard collisions involved. It follows that sensible predictions can be obtained only if the NPFF has been extracted from e^+e^- data using a cross section computed in the same approximation as that used to predict the $p\bar{p}$ cross section. By reconsidering carefully the fragmentation procedure adopted in ref. [4], the authors of ref. [10] pointed out that the claimed discrepancy of $2.9 \pm 0.2 \pm 0.4$ turns actually out to be $1.7 \pm 0.5 \pm 0.5$ (this also includes the 20% effect mentioned before, for replacing the NLO result with the FONLL one), i.e. data are within 1σ from the default theoretical prediction. This finding is consistent with experimental evidence from D0 [11] that the inclusive rate of jets containing b quarks – a quantity largely insensitive to the details of the perturbative and non-perturbative fragmentation – agrees with NLO QCD predictions [12].

More evidence that things go in the right direction has been achieved thanks to a new CDF measurement [13] of single-inclusive b-hadron p_T spectrum in the central rapidity region. For the first time, these data probe the region of $p_T \simeq 0$, where fragmentation effects play a minor role; thus, an improved agreement wrt the previous comparisons would support the conclusion that what we have to blame is our incomplete understanding of the fragmentation phase. This is in fact what happens. The data show the best-ever agreement (see fig. 1) with FONLL and MC@NLO [14,15]. I stress that FONLL and MC@NLO are both based on the NLO computations of ref. [1], and they differ only in the resummation of logs beyond the leading ones, and in the treatment of the b quark hadronization and decay, which in MC@NLO is performed through the HERWIG [16] cluster model. It is remarkable that, in spite of the differences in the treatment of the resummation of the large-p_T logs, FONLL and MC@NLO can be made to agree perfectly with a proper tuning of the fragmentation and of the clustering parameters (see ref. [17] for a discussion on this point and on the comparison between theory and Run II CDF data).

Let me now turn to b cross section measurements at HERA. A couple of years ago, the situation appeared to be consistent with what was observed at the Tevatron, with NLO QCD predictions systematically undershooting H1 and ZEUS data. This picture has now radically changed, as has been thoroughly documented at this conference [18]. All data lie within 2σ from the theoretical predictions; for the majority of them, the agreement is in fact much better than 2σ, with data basically sitting on top of NLO QCD predictions. It is worth noting that in the case of b production at HERA NLO and FONLL predictions [19] coincide, since the transverse momenta probed are never too large. This also implies that the treatment of the fragmentation mechanism is not as delicate as in the case of the Tevatron measurements.

In the case of HERA, the breakthrough that occurred in the last couple of years has been mainly due to the fact that, thanks to a much larger statistics, cross sections could be presented in the experimentally visible regions (rather than in the form of total rates), and compared to theoretical predictions obtained by applying the same cuts. Former experimental results always involved huge extrapolations from the very narrow visible regions to the whole phase space, performed with standard parton shower Monte Carlo's, which *cannot* give sensible predictions for small p_T's (see ref. [15] for a discussion on this point), a region which gives the dominant contribution to the total rate. MC@NLO is reliable at small p_T's, but it is not yet available for photoproduction and DIS processes, and in any case the publication by the experimental collaborations of the visible cross sections is always the preferred option.

In view of the lesson learned at the Tevatron and HERA, it is unfortunate that the measurements of the $\gamma\gamma \to b\bar{b} + X$ cross sections [20,21] suffer from the drawbacks that prevented a fair comparison between theory and data in $p\bar{p}$ and ep collisions. The three measurements rely on huge extrapolations from the visible regions to the whole phase space, done with standard parton shower Monte Carlo's; the uncertainties associated with the theoretical predictions are too small; the techniques used are identical. For these reasons, I find it difficult to take at face value the discrepancies quoted (data are more than a factor of three larger than the default NLO predictions), since a careful computation of all the uncertainties involved (for example, those relevant to the extrapolation to the whole phase space, which needs to be assessed by using at least two different Monte Carlos, and ideally the NLO computations themselves) would presumably show that data lie within less than 3σ from theory. Clearly, this statement cannot be proved (or disproved) but by the experimental collaborations, which should follow the strategy set by H1 and ZEUS, of quoting results relevant to the visible regions. Let me conclude by mentioning that I'm not aware of any beyond-the-SM mechanism, let alone higher-order QCD corrections, which could explain these huge discrepancies.

In summary, thanks to improvements on both the experimental and the theoretical sides, b data at colliders seem to be in fair agreement with QCD expectations. I stress that the theoretical predictions are essentially based on the NLO computations of the late 80's [1,2,3], and that the changes in the predictions of single-inclusive p_T spectra are due to a better understanding of the fragmentation mechanism and to the use of more precisely determined PDF sets. Newly developed tools, such as MC@NLO, will serve to pin down discrepancies between theory and data in yet unexplored corners of the $b\bar{b}$ phase space.

References

1. P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B **303** (1988) 607.
2. P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B **327** (1989) 49 [Erratum-ibid. B **335** (1990) 260].
3. W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler and J. Smith, Nucl. Phys. B **351** (1991) 507.
4. D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **65** (2002) 052005 [arXiv:hep-ph/0111359].
5. E. L. Berger, B. W. Harris, D. E. Kaplan, Z. Sullivan, T. M. P. Tait and C. E. M. Wagner, Phys. Rev. Lett. **86** (2001) 4231 [arXiv:hep-ph/0012001].
6. M. Mangano, *The 15-year saga of the B cross-section*, Wine and Cheese Seminar, Fermilab, Jan 2004, `http://mlm.home.cern.ch/mlm/`.
7. J. C. Collins and R. K. Ellis, Nucl. Phys. B **360** (1991) 3.
8. B. Mele and P. Nason, Nucl. Phys. B **361** (1991) 626.
9. M. Cacciari, M. Greco and P. Nason, JHEP **9805** (1998) 007 [arXiv:hep-ph/9803400].
10. M. Cacciari and P. Nason, Phys. Rev. Lett. **89** (2002) 122003 [arXiv:hep-ph/0204025].
11. B. Abbott *et al.* [D0 Collaboration], Phys. Rev. Lett. **85** (2000) 5068 [arXiv:hep-ex/0008021].
12. S. Frixione and M. L. Mangano, Nucl. Phys. B **483** (1997) 321 [arXiv:hep-ph/9605270].
13. C. Chen, [CDF], presentation at "Beauty 2003", http://www-cdf.fnal.gov/physics/new/bottom/030904.blessed-bxsec-jpsi/; M. Bishai [CDF], presentation at Fermilab, Dec 5. 2003, http://www-cdf.fnal.gov/~bishai/papers/wandc.pdf
14. S. Frixione and B. R. Webber, JHEP **0206** (2002) 029 [arXiv:hep-ph/0204244].
15. S. Frixione, P. Nason and B. R. Webber, JHEP **0308** (2003) 007 [arXiv:hep-ph/0305252].

16. G. Corcella *et al.*, JHEP **0101** (2001) 010 [arXiv:hep-ph/0011363].
17. M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0407** (2004) 033 [arXiv:hep-ph/0312132].
18. M. Turcato, A. Meyer, and K. Klimek, these proceedings.
19. M. Cacciari, S. Frixione and P. Nason, JHEP **0103** (2001) 006 [arXiv:hep-ph/0102134].
20. M. Acciarri *et al.* [L3 Collaboration], Phys. Lett. B **503** (2001) 10 [arXiv:hep-ex/0011070].
21. OPAL Collaboration, Note PN 455 (2000); DELPHI Collaboration, Note 2003-015-CONF-635.