Heavy flavour production

Stefano Frixione

INFN, Sezione di Genova
Via Dodecaneso 33, 16146 Genova, Italy

Abstract. I review theoretical and phenomenological results relevant to heavy flavour production at colliders, and discuss the impact of recent measurements at the Tevatron and at HERA

Keywords: Collider Physics, Heavy Flavours

GENERALITIES

Heavy flavour production is one of the most extensively studied topics in high-energy particle physics. An impressive amount of data is available, for basically all kinds of colliding particles, which renders it possible to test QCD predictions to some accuracy. In open heavy flavour production final-state observables must be defined using either the variables of the heavy quarks or of the hadrons containing at most one heavy quark, and must not contain any reference to quarkonium states. By definition, a quark is heavy when

$$m_Q \gg \Lambda_{QCD}\,. \tag{1}$$

According to this equation, up, down, and strange quarks are definitely not heavy; for the remaining flavours, we have

$$m_t/\Lambda_{QCD} \simeq 800, \qquad \Longrightarrow \qquad \alpha_S(m_t) \simeq 0.1, \tag{2}$$
$$m_b/\Lambda_{QCD} \simeq 15, \qquad \Longrightarrow \qquad \alpha_S(m_b) \simeq 0.21, \tag{3}$$
$$m_c/\Lambda_{QCD} \simeq 4, \qquad \Longrightarrow \qquad \alpha_S(m_c) \simeq 0.33, \tag{4}$$

from which one is entitled to consider the top and the bottom to be heavy, while the case of charm is borderline. I shall treat the charm as heavy in what follows, for the simple technical reason that the condition in eq. (1) allows one to define an open-quark cross section without the need to convolute it with fragmentation functions, and this puts the charm formally on the same footing as the bottom and the top. On the other hand, since the quark mass typically sets the hard scale of the process, the values of α_S reported in eqs. (2)–(4) imply that in the case of charm the perturbative results will be affected by the larger uncertainties, and that non-perturbative effects are liable to play a major role.

In perturbative QCD, the production of a pair of heavy quarks $Q\overline{Q}$ in the collision of two hadrons $H_{1,2}$ is written according to the factorization theorem

$$d\sigma_{H_1H_2\to Q\overline{Q}}(S) = \sum_{ij}\int dx_1 dx_2\, f_i^{(H_1)}(x_1)\, f_j^{(H_2)}(x_2)\, d\hat{\sigma}_{ij\to Q\overline{Q}}(\hat{s}=x_1x_2S), \tag{5}$$

and a fairly similar formula holds in the case of γ-hadron collisions (see ref. [1] for a review). As is well known, the parton density functions (PDFs) $f_i^{(H)}$ cannot be computed in perturbation theory, but are universal. On the other hand, the short distance cross sections $d\hat{\sigma}_{ij\to Q\overline{Q}}$ are process-specific, and computable in perturbation theory

$$d\hat{\sigma} = \sum_{i=2}^{\infty} a_i \alpha_S^i = a_2 \alpha_S^2 + a_3 \alpha_S^3 + a_4 \alpha_S^4 + \ldots \tag{6}$$

The coefficients a_2, a_3, and a_4 explicitly indicated in eq. (6) correspond to the LO, NLO, and NNLO contributions respectively. Although the computation of the LO term is almost trivial, that of the NLO term is not, and its achievement represented a breakthrough in heavy flavour physics at the end of the 80's [2, 3]. No exact result beyond NLO is currently available, and it does not seem probable that it will for some time. This is worrisome, since the NLO corrections for c and b production are of the same size as the LO contributions (i.e., the K factor is about 2), and the scale dependence for some observables is very large; NNLO terms may thus be numerically sizable, and quite relevant to the correct predictions of measured quantities.

Even if NNLO (or higher) contributions were available, one must keep in mind that such *fixed-order* results may still be insufficient to obtain phenomenologically sensible predictions. There are two main issues that need be considered.

1) Large logs appear in the perturbative coefficients

$$a_i = \sum_{k=0}^{i-2} a_i^{(i-2-k)} \log^{i-2-k} \mathscr{Q}, \tag{7}$$

where "large" means $\alpha_S \log^2 \mathscr{Q} \gtrsim 1$. $\mathscr{Q}$ *may or may not* depend on the observable. If $\mathscr{Q}$ is large, all terms in the expansion on the r.h.s. of eq. (6) are numerically of the same order, and the convergence of the series is spoiled. The way out is that of keeping only (some of) the leading logs in eq. (7), in such a way that the series of eq. (6) can be summed. This effectively corresponds to rearranging the perturbative expansion; technically, one says that the logs are *resummed*.

2) Bottom and charm quarks, although heavy, cannot be directly observed; therefore, the open heavy flavour cross section in such cases must be supplemented with the description of the quark-to-hadron transition (called fragmentation), which always involves a quantity, the non-perturbative fragmentation function (NPFF), not computable in perturbation theory. For the single-inclusive p_T spectrum, one writes

$$\frac{d\hat{\sigma}(H_Q)}{dp_T} = \int \frac{dz}{z} D^{Q\to H_Q}(z;\varepsilon) \frac{d\hat{\sigma}(Q)}{d\hat{p}_T}, \qquad p_T = z\hat{p}_T, \tag{8}$$

where Q is the heavy quark, H_Q is a given heavy-flavoured hadron, $\hat{p}_T$ (p_T) is the transverse momentum of Q (H_Q), $d\hat{\sigma}(Q)$ is the cross section for open-Q production, and $D^{Q\to H_Q}$ is the NPFF.

I'll start by discussing the case of large perturbative logs. In general, the logs to be resummed can be divided into two broad classes.

Observable-dependent logarithms: these logs depend on the kinematics of the final state (including cuts); a sample of their arguments is given in the equations below

$$\mathscr{Q} = \frac{p_T(Q)}{m_Q}, \qquad p_T(Q) \gg m_Q, \tag{9}$$

$$\mathscr{Q} = \frac{p_T(Q\overline{Q})}{m_Q}, \qquad p_T(Q\overline{Q}) \simeq 0, \tag{10}$$

$$\mathscr{Q} = 1 - \frac{\Delta\phi(Q\overline{Q})}{\pi}, \qquad \Delta\phi(Q\overline{Q}) \simeq \pi. \tag{11}$$

Equation. (9) is relevant to the single-inclusive transverse momentum distributions, whereas eqs. (10) and (11) are relevant to $Q\overline{Q}$ correlations. Analytic resummations for the logs of this class are observable-dependent and technically fairly involved, which renders the resummed cross sections unavailable except for a few simple cases. Even if the resummed observable can be computed, in general it must be matched to the corresponding fixed-order result for the predictions to be physically meaningful. Fortunately, this is the case for the single-inclusive p_T spectrum: the FONLL formalism [4] allows one to consistently combine (i.e., avoiding over-counting) the NLO result with the cross section in which p_T/m logs are resummed to NLL accuracy. Thus, FONLL can describe both the small-p_T ($p_T \sim m$, where resummed results don't make sense) and the large-p_T ($p_T \gg m$, where NLO results are not reliable) regimes.

The technical complications of the analytic resummations can be avoided by letting a parton shower Monte Carlo (PSMC) perform the resummation numerically. This procedure has the advantage that the logs can always be resummed, no matter how complicated the definition of the observable and the final state cuts are. The drawback is that the PSMC resummation is formally less accurate in terms of log accuracy than the analytic resummations, although in practice the difference between the two approaches is almost always negligible. On the other hand, PSMC's are based on a LO computation at the level of matrix elements, which is largely insufficient in the case of heavy flavours. In recent years, however, the problem of the consistent inclusion of NLO matrix elements into a PSMC framework has been successfully solved in QCD (MC@NLO [5, 6]); phenomenological results obtained with this formalism will be presented later.

Observable-independent logs: these logs do not depend on the kinematics of the final state. Those relevant to heavy flavour production are

$$\mathscr{Q} = 1 - 4m_Q^2/\hat{s}, \qquad \hat{s} \simeq 4m_Q^2, \tag{12}$$

$$\mathscr{Q} = m_Q^2/\hat{s}, \qquad \hat{s} \gg m_Q^2, \tag{13}$$

denoted as threshold and small-x logs respectively. Techniques to resum the former logs are rather well established; their effects are rather marginal, however, in c and b physics, except for b production at HERA-B (their role in top production has been mentioned before). On the other hand, small-x logs are theoretically challenging and intriguing. The standard Altarelli-Parisi evolution equations are replaced by those of CCFM; upon doing so, one is forced to introduce the so-called unintegrated PDFs, which have a functional dependence on a transverse momentum in addition to those on x and Q^2 of the standard PDFs.

After all the relevant large logs are properly resummed, one needs to understand whether the description of the fragmentation is physically sensible. I remind the reader that the NPFF is not calculable from first principles, and the free parameters it contains (denoted by ε in eq. (8)) are fitted to data after choosing a functional form in z. This fit is typically performed using eq. (8), identifying the l.h.s. with e^+e^- data. It follows that the value of ε is strictly correlated to the short-distance cross section $d\hat{\sigma}(Q)$ used in the fitting procedure, and thus is *non-physical*. When eq. (8) is used to predict H_Q cross sections, it is therefore mandatory to make consistent choices for ε and $d\hat{\sigma}(Q)$. The failure to do so has been one of the reasons behind the long-standing discrepancy between single-inclusive b measurements at colliders, and QCD predictions.

There is also another peculiar issue relevant to the NPFF: $d\hat{\sigma}(Q)/d\hat{p}_T$ has a power-like form, and one writes

$$\frac{d\hat{\sigma}(Q)}{d\hat{p}_T} \simeq \frac{C}{\hat{p}_T^N} \implies \frac{d\hat{\sigma}(H_Q)}{dp_T} = \frac{C}{p_T^N} D_N^{Q\to H_Q}, \tag{14}$$

$$D_N^{Q\to H_Q} = \int dz z^{N-1} D^{Q\to H_Q}(z;\varepsilon). \tag{15}$$

It turns out that the approximation for the H_Q cross section given in eq. (14) is in excellent agreement with the exact result [7]. Since typically $N = \mathscr{O}(5)$ (for c and b at the Tevatron) it follows that, in order to have an accurate prediction for the p_T spectrum in hadroproduction, it is mandatory that the first few Mellin moments computed with $D(z)$ agreed with those measured. In ref. [8], where b production is considered, it is pointed out that this is not the case, in spite of the fact that the prediction for the inclusive B cross section in e^+e^- collisions, obtained with *the same* $D(z)$, displays an excellent agreement with the data. There may seem to be a contradiction in this statement: if the shape is reproduced well, why this is not true for the Mellin moments? The reason is that, when fitting $D(z)$, one excludes the region of large z, since it is affected by Sudakov logs and by complex non-perturbative effects which are unlikely to be described by the NPFF. On the other hand, the large-z region is important for the computation of D_N (because of the factor z^{N-1} in the integrand of eq. (15)). Therefore, for the purpose of predicting D- and B-meson spectra at colliders, ref. [8] advocates the procedure of fitting the NPFF directly in the N-space. A fit to the second moment (denoted as $N = 2$ fit henceforth) is found to fit well all the D_N's for N up to 10.

HEAVY FLAVOUR PRODUCTION: PHENOMENOLOGY

As discussed above, a typical state-of-the-art theoretical computation matches an NLO prediction with a cross section in which the relevant logs have been resummed. The ideal testing ground for such computation is top production for which, according to eqs. (2)–(4), we expect the perturbative series to be well-behaved, and that has the extra advantage of being unaffected by problems related to the understanding of the fragmentation mechanism. At present, the only $t\bar{t}$ data available are those from the Tevatron, where the c.m. hadronic energy is never much larger than the $t\bar{t}$ invariant mass, and therefore threshold logs may play an important role. It turns out that the resummation of such logs is seen to increase only marginally the NLO result (+3.5%), while it reduces substantially the scale uncertainty (from about 10% to about 5%). The comparison of preliminary Run II data with NLO predictions matched to the NLL-resummed result of ref. [9] is presented in fig. 1; as can be seen, theory and data are in good agreement, which confirms the findings of Run I. It should be stressed that, although the overall theoretical accuracy (about 12%, a 7% of which is due to PDF uncertainty) is small compared to that one finds in other heavy-flavour production processes, its combination with experimental errors will render it unlikely that any significant discrepancies with QCD will be found in Run II, even if top mass measurements will result in a value smaller than the present world average (CDF presented at this conference the most precise result so far obtained, $m_{top} = 173.5^{+4.1}_{-4.0}$ GeV – see ref. [10]).

Predictions used in fig. 1 have been obtained in ref. [11], and have been called NLO+NLL to remind that they are based on an NLO result matched with a resummed one that includes logs up to the next-to-leading accuracy. I stress that this terminology is not universal, and in particular is different from the one used in ref. [13] and subsequent papers. In those papers, one generally finds the notation NNLO+N^{k}LL, with k=2,3,4. It is important to realize that NNLO in this context does *not* mean that the full NNLO result has been computed (in fact, such result is not available, since the corresponding two-loop amplitudes have not been computed); it means that all of the NNLO terms have been retained, which are obtained upon a re-expansion of the resummed result, matched with the full NLO result. As far as the resummation is concerned, while the logs in ref. [9] are counted upstairs (i.e. in the exponent), in ref. [13] they are counted downstairs (i.e., after exponent expansion). This implies that the two approaches differ

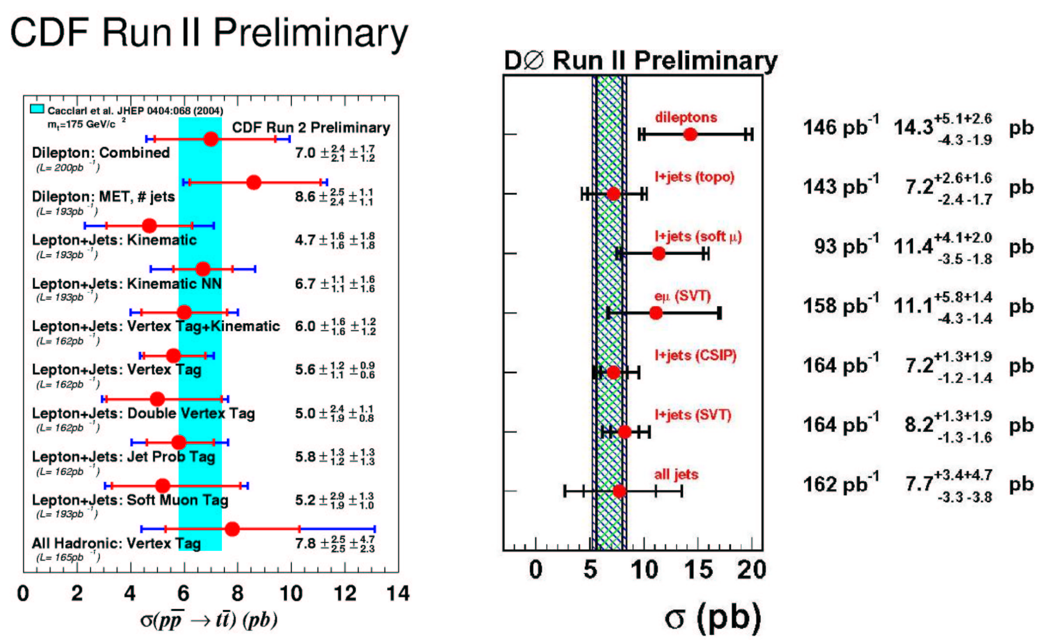

FIGURE 1. Comparison between NLO+NLL QCD predictions [11] and Run II data (see e.g. ref. [12]) for top production at the Tevatron.

by terms of NNLO, and of NNLL (in the exponent), which are in any case beyond control, and therefore that the perturbative accuracies of these results are the same.

Predictions for bottom and charm cross sections are much more involved than those for top, and therefore they represent a tougher test for perturbative QCD. As is well known, bottom data have received a lot of attention in the past. Such data are extremely abundant at hadron colliders, especially so for single-inclusive observables that are relatively well measured also in absence of vertex detectors. It was therefore worrisome that basically all Run I measurements overshot NLO QCD predictions by sizable factors (see e.g. ref. [1]). In general, the shape of single-inclusive p_T spectrum given by QCD is consistent with the data, but if one uses "default" values for the inputs of the computations (such as $m_b = 4.75$ GeV, $\mu = \sqrt{p_T^2 + m_b^2}$, and central PDF sets) the measured rates are larger than NLO results by a factor of 2.84 (for CDF) and 2.12 (for D0) on average. The last data available from Run I confirmed this trend: for the B^+ p_T spectrum, CDF found that the average data/theory ratio is $2.9 \pm 0.2 \pm 0.4$.

The disagreement between b production data at the Tevatron and QCD predictions has been one of the most compelling problems in hadronic physics. It triggered studies on possible beyond-the-SM explanations, which to the best of my knowledge have not survived the tests of LEP data. Although the BSM scenario is still an option, at present it seems premature to adopt it without first reassessing carefully all possible sources of mistakes in the past comparisons between theory and data, and without considering the uncertainties that so-far uncalculated SM contributions can give. To start with, it appears that a better comparison with Tevatron data should be obtained by using FONLL rather than pure NLO predictions, with the fragmentation function obtained with an $N = 2$ fit. The result of such a comparison is shown in the left panel of fig. 2 (taken from ref. [8]), where the data are the same B^+ ones quoted above. The average data/theory ratio is now $1.7 \pm 0.5 \pm 0.5$ i.e. data are within 1σ from the default theoretical prediction. A factor of 20% of the reduction from the formerly claimed discrepancy of 2.9 to the current 1.7 is due to the use of FONLL in place of NLO results; the remaining 45% to the correct treatment of the NPFF.

These findings suggest to recompute the theoretical predictions upon which the vast majority of the past comparisons between theory and data were based. Unfortunately, this would not help much, since most of those data are relevant to b quarks, rather

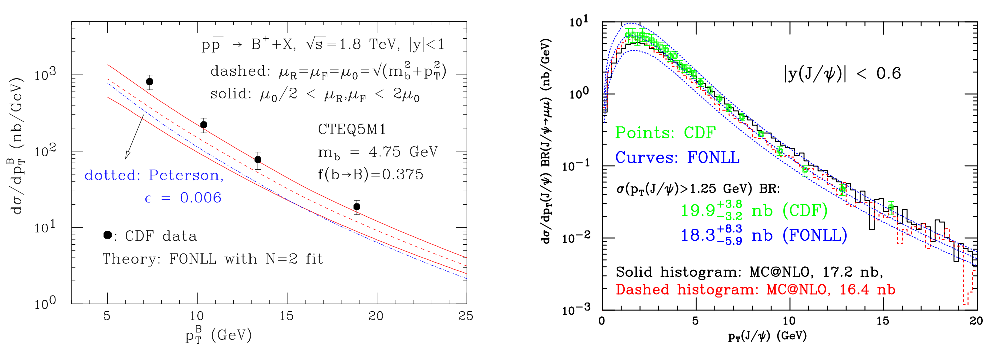

FIGURE 2. Comparisons between b data at the Tevatron, and QCD predictions. See the text for details.

than to B mesons; in other words, experimental collaborations deconvoluted the $b \to B$ fragmentation. This has been typically done using PSMC models, and in general it appears rather difficult to recover the data for B mesons, which is a practical example of a theoretical bias affecting data. On the other hand, there are other ways to understand whether we are on the right track. One interesting possibility consists in considering jets containing b quarks (i.e., any b-hadron species) rather than tagging a specific b-hadron: in this case, the NPFF simply doesn't enter the cross section, and the theoretical predictions are also less prone to develop large p_T logs, since the p_T of the b is not involved in the definition of the observable. The comparison between NLO predictions for b-jets [14] and D0 measurements [15] is indeed satisfactory: data are consistent with theory in the range $25 < E_T^{b-jet} < 100$ GeV. More measurements relevant to b-jets are being performed in Run II.

To fully test the ideas relevant to the NPFF treatment one still needs heavy-flavoured meson data. Fortunately, a lot of these are expected to become available in the near future, thanks to the ongoing Run II. The first results on single-inclusive b-hadron p_T spectrum have been presented by CDF [16], and they are particularly interesting in view of the fact that for the first time they probe the region of $p_T \simeq 0$, which is fairly sensitive to the description of the fragmentation. The comparison of data with FONLL (dotted lines) is presented in the right panel of fig. 2 and displays the best-ever agreement between theoretical predictions and b data at colliders. The same pattern of agreement is obtained by using MC@NLO (histograms), which constitutes a very powerful check on the theoretical results: both the resummation and the $b \to B$ transition are performed in vastly different ways in FONLL and MC@NLO. I should stress that the plot presents the p_T spectrum of the J/ψ's emerging from b-flavoured hadron decays, and thus involves a highly non-trivial combination of short- and long-distance dynamics (see ref. [17] for a more detailed discussion).

Although the tests presented above prove that when comparing modern data sets with up-to-date theoretical predictions heavy flavour production appears to be fairly well predicted by QCD, we must keep in mind that the uncertainties affecting the QCD results are still quite large; thus, the possibility remains that such uncertainties hide physics effects not included in the computations. Within non-BSM physics, one of the most interesting questions is whether b production at the Tevatron is small-x physics. According to ref. [18], small-x resummation would increase the b cross section by a mere 30%. On the other hand, a good description of Tevatron data is obtained by using CASCADE [19], a Monte Carlo code that implements CCFM evolution equations. It should be noted that, since the (LO) matrix elements convoluted with unintegrated PDFs have off-shell partons, they include part of the contributions which are of NLO in the standard collinear approach. This renders the interpretation of the results more complicated, since it is impossible to tell the *pure* small-x effects apart from higher-order corrections to the matrix elements. It seems therefore necessary to compute the NLO corrections rigorously in the context of the small-x approach in order to achieve firmer conclusions. I should also mention the fact that unintegrated PDFs, exactly like standard PDFs, cannot be computed from first principles, and need to be extracted from data. Such an extraction is at present affected by large uncertainties (especially for the gluon density), which must be systematically reduced in order for small-x computations to be as reliable phenomenologically as those based on collinear factorization.

Thanks to the much improved performances in terms of luminosity of the HERA collider, QCD predictions for b production can also be sensibly tested in γp (photoproduction) and ep (DIS) collisions. The first HERA results on b physics gave ambiguous indications. H1 photoproduction measurement [20] was a factor 3.26 ± 0.74 higher than (i.e., 3σ away from) NLO QCD predictions [21]. On the other hand, in the case of ZEUS [22] the ratio data/theory was 2.50 ± 1.18, i.e. a mere 1.3σ away from NLO QCD. Experiments understand much better now these very complicated analyses; furthermore, the increased statistics allow them to quote cross sections in the visible regions, and therefore to be more independent of the (mainly PSMC) biases implicit in the extrapolations from the visible to the invisible regions. In ref. [23], H1 re-analyzed the data of ref. [20]; the theory/data ratio went down to 1.99 ± 0.47. Furthermore, new data presented in that paper are larger than NLO QCD predictions only by a factor 1.61 ± 0.51, i.e. little more than 1σ. As far as ZEUS is concerned, newer photoproduction results [24] (summarized here in the left panel of fig. 3) are also in better agreement with QCD than those of ref. [22]. By a closer inspection of the experimental results, one can see that the bulk of the discrepancy wrt NLO QCD predictions observed by H1 in ref. [23] is concentrated at low transverse momenta, in both photoproduction and DIS regimes. The same effect occurs in a recent ZEUS DIS measurement [25]. A few more glitches can be seen in the comparison between data and theory, which is summarized for total (visible) rates in the right panel of fig. 3. These kind of discrepancies certainly deserve more attention in the future, and improved vertexing techniques and statistics will help to clarify the situation. On the other hand, it appears to me that at present there is a satisfactory agreement between HERA measurements and QCD predictions for b observables. It is worth noting that small-p_T regions have the highest sensitivity to PSMC biases, since standard (i.e. based on a leading-order picture) parton shower Monte Carlo's cannot give sensible predictions for such momenta (see ref. [6] for a discussion on this point). One should not forget the lesson learnt at the Tevatron, that these kind of biases can lead to overestimating the discrepancy between theory and data. An example of (possible) large biases can be found in ref. [26], in which the b sample is obtained entirely through MC methods: it is perhaps not a coincidence that the ratio data/theory turns out to be extremely large in that paper.

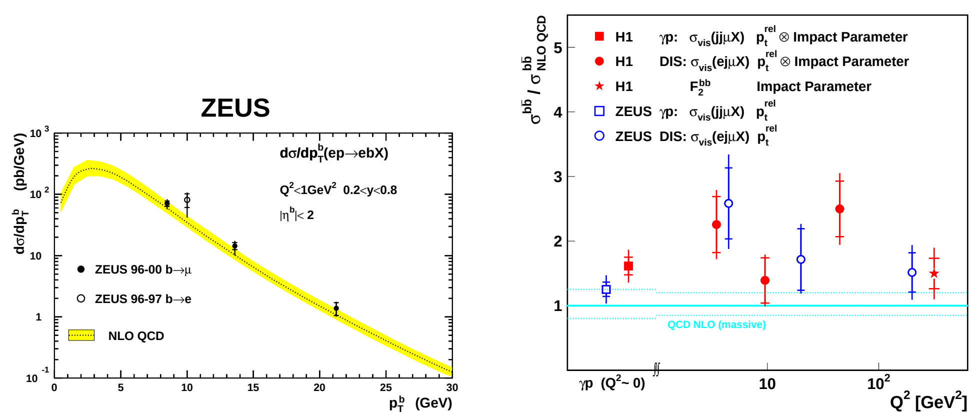

FIGURE 3. Comparisons between b data at HERA, and QCD predictions. See the text for details.

Let me now turn to the case of charm production. Being the lightest of the heavy quarks, QCD predictions for charm cross sections are affected by the largest perturbative uncertainties and non-perturbative effects. That notwithstanding, the overall agreement between charm data and QCD predictions is basically satisfactory, although a few disturbing facts remain. In particular, at HERA data on D^* production cannot be described by QCD in the region $\eta > 0$. Also, the $p_T(D^*)$ photoproduction spectrum measured by ZEUS is harder than FONLL predicts, while that of H1 is softer. In spite of these discrepancies, I believe that open charm production will not pose any serious problem for QCD in the near future. A fairly different situation is that of the $c\bar{c}$ bound states, whose theoretical understanding is less developed than that for open-heavy-quark states, and whose phenomenology poses at present quite a challenge to QCD. I'll give the briefest of the reviews on quarkonium production in what follows.

The hadroproduction cross section of a $c\bar{c}$ or $b\bar{b}$ quarkonium state H can be written in a manner analogous to eq. (5)

$$d\sigma_{H_1H_2\to H}(S) = \sum_{ij}\int dx_1 dx_2 f_i^{(H_1)}(x_1) f_j^{(H_2)}(x_2) d\hat{\sigma}_{ij\to H}(\hat{s} = x_1x_2S). \tag{16}$$

The forms of the short-distance cross sections that appear in eq. (16) depend on the theory or the model adopted to compute them. Among the various approaches, the *only one* that can be mathematically derived from QCD is non-relativistic QCD (NRQCD [27]), an effective theory in which the heavy quarks move at non-relativistic velocities. Within NRQCD, the partonic cross sections are [28]

$$d\hat{\sigma}_{ij\to H} = \sum_n d\hat{\sigma}(ij \to Q\overline{Q}[n])\langle\mathscr{O}^H[n]\rangle, \tag{17}$$

where $d\hat{\sigma}(Q\overline{Q}[n])$ is the cross section for the production of a $Q\overline{Q}$ pair in a given spin and colour state (symbolically, $n = \{c = (1,8); {}^{2S+1}L_J\}$), and the NRQCD matrix elements $\langle\mathscr{O}^H[n]\rangle$ are analogous to PDFs and NPFFs, since they cannot be computed in perturbation theory, and are universal; loosely speaking, they are proportional to the probability for the $Q\overline{Q}$ pair in the state n to fragment into the quarkonium state H.

Given the fact that NRQCD is derived from QCD, and that pQCD can describe open-Q data, we expect that NRQCD does a good job too. A difficulty, however, is immediately apparent by looking at eq. (17), that features an infinite sum in which each term contains a non-calculable long-distance parameter ($\langle\mathscr{O}^H[n]\rangle$); this implies a complete loss of predictive power. Fortunately, the NRQCD matrix elements obey a (velocity) scaling rule [29]

$$\langle\mathscr{O}^H[n]\rangle \propto v^{f(n,H)}, \quad v^2 \simeq 0.3, 0.1 \;\text{ for }\; c\bar{c}, b\bar{b}, \tag{18}$$

where v is the relative velocity of the $Q\overline{Q}$ pair in the quarkonium state, and f is a function that depends in a complicated manner on the quantum numbers of the states $Q\overline{Q}[n]$ and H (see e.g. ref. [30]). Eq. (18), combined with the usual expansion in α_S of the short-distance cross sections, implies that the r.h.s. of eq. (17) can be rewritten as a double series

$$d\hat{\sigma}_{ij\to H} = \sum_{m,k} s_{m,k}\alpha_S^m v^k. \tag{19}$$

This systematic expansion in α_S and v provides a computational framework similar to that relevant to open-Q production. Still, eq. (19) poses some non-trivial computational problems, given the fact that the double series is slowly "convergent", particularly so for charm (owing to eqs. (4) and (18)), and thus one needs to determine a large number of NRQCD matrix elements (some of them can be expressed in terms of others, for example using heavy quark spin symmetry and vacuum saturation approximation). Furthermore, the same problems that affect the short distance coefficients of eq. (6) are also relevant to the present case; most notably, the occurrence of large logs can render the expansion (19) useless.

The first stringent experimental tests of NRQCD predictions at colliders have become available only relatively recently, thanks to the use by CDF of microvertices that allow the precise measurements of the *direct* cross sections (those in which the observed quarkonium state is not obtained through feeddown from more massive bound states). This immediately led to ruling out the so-called colour-singlet model (CSM), which can be obtained from eq. (17) by dropping all but the leading colour-singlet contribution there. In fact, it turns out that no amount of tuning of the NRQCD matrix elements can bring the LO CS contribution in agreement with data, and CO contributions are therefore necessary; this fact holds both for J/ψ and for Υ production.

Although the dominance of CO contributions in quarkonia production at the Tevatron has to be regarded as a highly successful prediction of NRQCD, we must keep in mind that the NRQCD matrix elements are (in part) fitted to the data that the theory is supposed to predict. Such a determination, furthermore, is not only affected by fairly large uncertainties (see ref. [31]), but is also biased by the fact that a definite choice for the PDFs of the colliding hadrons must be made. Within these uncertainties, the NRQCD matrix elements do obey the scaling rules of eq. (18). On the other hand, a more convincing test of NRQCD predictions is the check of the universality of the matrix elements, whose values must be independent of the hard process and therefore, if fitted at the Tevatron, can be used at HERA to predict the quarkonium cross sections there. Unfortunately, the comparison between ep data and NRQCD predictions is largely inconclusive as far as the CO contributions are concerned. The J/ψ energy distribution in photoproduction does not support the growth towards the endpoints of the spectrum which is a consequence of the CO terms; on the other hand, it is known that such regions are strongly affected by large higher-order corrections (in v), and thus a resummation would be necessary in order to draw definite conclusions. On the other hand, the data are in good agreement, for both the energy distribution and the p_T spectrum, with the pure CS prediction if NLO effects [32] are taken into account. The situation appears to be more consistent with the findings at the Tevatron in the case of DIS data, although a few glitches remain there as well (especially in the z distribution). The bottom line is that it is hard to draw any conclusion at present; data of larger statistics must be obtained in the present HERA run phase in order to test NRQCD more thoroughly.

It is interesting to observe that a good agreement with the Tevatron J/ψ and Υ data is also obtained in the context of the colour evaporation model (CEM). When using such a model, the short-distance cross sections in eq. (16) are written as follows

$$d\hat{\sigma}^{(\mathrm{CEM})}_{ij\to H} = F_H \int_{4m_Q^2}^{4m_M^2} dm^2_{Q\overline{Q}} \frac{d\hat{\sigma}(ij \to Q\overline{Q})}{dm^2_{Q\overline{Q}}}, \tag{20}$$

where m_M is the mass of the lowest-lying Q-flavoured meson state, and F_H is a universal (long-distance) constant. The colour evaporation model can be formally written in the same form as NRQCD, by replacing the original expression for the NRQCD matrix elements

$$\mathscr{O}^H[n] = \chi^* \kappa_n \psi \left(\sum_X |H+X\rangle\langle H+X| \right) \psi^* \kappa_n' \chi \tag{21}$$

with

$$\begin{aligned}\mathscr{O}^H[n] \;\; = \;\; & F_H \sum_n \chi^* \kappa_n \psi \sum_X |Q\overline{Q}(m^2_{Q\overline{Q}} < 4m_M^2) + X\rangle \\ & \times \langle Q\overline{Q}(m^2_{Q\overline{Q}} < 4m_M^2) + X| \psi^* \kappa_n' \chi.\end{aligned} \tag{22}$$

Eq. (22) basically entails a change of the scaling rules, $v^{f(n,H)} \to v^{2L}$; this is interesting, since it implies that a re-organization of the double expansion in eq. (19) can also give satisfactory phenomenological results. It should be noted that in general CEM predictions must be supplemented with a k_T-kick in order to describe the data, which decreases the predictive power of the model. The fact that the only information concerning the quarkonium state is contained in the constant F_H is also troublesome, since for example the ratio $\sigma(\chi_c)/\sigma(J/\psi)$ turns out to be different if measured in hadron-hadron and photon-hadron collisions at fixed-target. Furthermore, a weak p_T dependence is found for the J/ψ decay fractions.

In conclusions, it appears that at present perturbative QCD is doing a satisfactory job in predicting open heavy flavour cross sections measured at colliders. The most substantial improvement in recent years occurred in b physics; the long-standing discrepancy between single-inclusive b data at colliders and theoretical predictions has been settled mainly thanks to a better understanding of the non-perturbative phenomena, since the backbone of the computations is still the NLO result of refs. [2, 3], which is essential to get anywhere close to the measurements. On the other hand, the careful re-analysis of the computations motivated their improvements, through the matching with resummed results (FONLL) or with Monte Carlo techniques (MC@NLO); the flexibility of the latter guarantees that studies with the same accuracy of those performed so far only for single-inclusive observables can be repeated for basically any type of variable. It is reassuring that, by increasing the collected statistics and with a better understanding of the analyses, the measurements performed at HERA are also in a much better agreement with QCD than they used to be. At present, I don't see any serious discrepancies between data and theory there; some problems remain, however, and this will serve as a powerful motivation for obtaining data of yet better quality in the last years of HERA running.

The situation is not as bright in the case of quarkonium production. We do have a theory, NRQCD, whose elegant and compact formulation allows the systematic computation of any observable relevant to quarkonium production. Being a direct consequence of QCD, it appears to be fairly solid. On the other hand, problems remain in the comparisons with collider data, the most serious of which is that of the J/ψ polarization. It should be noted that NRQCD computations are not at the same level of accuracy as those used to predict open heavy flavour cross sections; both the v and the α_S expansions must be considered, and the computation of the observables beyond LO (which appears to be a necessity in heavy flavour physics) is extremely complicated. The success of the CEM may suggest that, at least for the case of charm, velocity scaling rules may not be adequate, and more theoretical work is needed in this direction.

REFERENCES

1. S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, Adv. Ser. Direct. High Energy Phys. **15** (1998) 609 [arXiv:hep-ph/9702287].
2. P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B **303** (1988) 607.
3. W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler and J. Smith, Nucl. Phys. B **351** (1991) 507.
4. M. Cacciari, M. Greco and P. Nason, JHEP **9805** (1998) 007 [arXiv:hep-ph/9803400].
5. S. Frixione and B. R. Webber, JHEP **0206** (2002) 029 [arXiv:hep-ph/0204244].
6. S. Frixione, P. Nason and B. R. Webber, JHEP **0308** (2003) 007 [arXiv:hep-ph/0305252].
7. P. Nason *et al.*, arXiv:hep-ph/0003142.
8. M. Cacciari and P. Nason, Phys. Rev. Lett. **89** (2002) 122003 [arXiv:hep-ph/0204025].
9. R. Bonciani, S. Catani, M. L. Mangano and P. Nason, Nucl. Phys. B **529** (1998) 424 [arXiv:hep-ph/9801375].
10. J-F. Arguin, these proceedings.
11. M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0404** (2004) 068 [arXiv:hep-ph/0303085].
12. K. Ranjan, these proceedings.
13. N. Kidonakis, E. Laenen, S. Moch and R. Vogt, Phys. Rev. D **64** (2001) 114001 [arXiv:hep-ph/0105041].
14. S. Frixione and M. L. Mangano, Nucl. Phys. B **483** (1997) 321 [arXiv:hep-ph/9605270].
15. B. Abbott *et al.* [D0 Collaboration], Phys. Rev. Lett. **85** (2000) 5068 [arXiv:hep-ex/0008021].
16. D. Acosta *et al.* [CDF Collaboration], arXiv:hep-ex/0412071.
17. M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0407** (2004) 033 [arXiv:hep-ph/0312132].
18. J. C. Collins and R. K. Ellis, Nucl. Phys. B **360** (1991) 3.
19. H. Jung and G. P. Salam, Eur. Phys. J. C **19** (2001) 351 [arXiv:hep-ph/0012143].
20. C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **467** (1999) 156 [Erratum-ibid. B **518** (2001) 331] [arXiv:hep-ex/9909029].
21. S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. B **412** (1994) 225 [arXiv:hep-ph/9306337].
22. J. Breitweg *et al.* [ZEUS Collaboration], Eur. Phys. J. C **18** (2001) 625 [arXiv:hep-ex/0011081].
23. A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **41** (2005) 453 [arXiv:hep-ex/0502010].
24. S. Chekanov *et al.* [ZEUS Collaboration], arXiv:hep-ex/0312057.
25. S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **599** (2004) 173 [arXiv:hep-ex/0405069].
26. A. Aktas *et al.* [H1 Collaboration], arXiv:hep-ex/0503038.
27. W. E. Caswell and G. P. Lepage, Phys. Lett. B **167** (1986) 437.
28. G. T. Bodwin, E. Braaten and G. P. Lepage, Phys. Rev. D **51** (1995) 1125 [Erratum-ibid. D **55** (1997) 5853] [arXiv:hep-ph/9407339].
29. G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea and K. Hornbostel, Phys. Rev. D **46** (1992) 4052 [arXiv:hep-lat/9205007].
30. M. Kramer, Prog. Part. Nucl. Phys. **47** (2001) 141 [arXiv:hep-ph/0106120].
31. M. Beneke and M. Kramer, Phys. Rev. D **55** (1997) 5269 [arXiv:hep-ph/9611218].
32. M. Kramer, Nucl. Phys. B **459** (1996) 3 [arXiv:hep-ph/9508409].