

Open charm production and spectroscopy at ATLAS and CMS

George W.S. Hou on behalf of the ATLAS and CMS Collaborations*

National Taiwan University E-mail: wshou@phys.ntu.edu.tw

The central and general purpose experiments at the LHC have contributed to open charm physics, complementing LHCb in B_c property studies, and ATLAS has discovered a B_c excitation that is consistent with B_c(2S). ATLAS has studied charged D_(s) meson production, while CMS has studied D⁰ production. In particular, CMS measured the nuclear modification factor R_{AA} in PbPb collisions at 5.02 TeV, finding strong medium suppression in PbPb compared with pp collisions over a broad range of p_T , and consistent with the ALICE result scaled to similar energy. An important tool has been developed to identify charm jets, complementing b-tagging algorithms. CMS has developed a 2D c-tagger to discriminate c-jet from light jet and b-jet, respectively. After training on simulated data, the c-tagger has been validated with W+c and tt events using 2015 data at 13 TeV, with extracted scale factor SF_c close to 1. A similar c-tagger has been developed by ATLAS, but the 2D version for Run 2 is not yet publicly available.

VIII International Workshop On Charm Physics 5-9 September, 2016 Bologna, Italy

*Speaker.

[©] Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

1. What this talk is not

To conform with the charge given me, let me up-front state what this talk is not about:

- This talk does not cover LHCb. ATLAS [1] and CMS [2] are central detectors, which lack K/π identification ability. Instead, the relevant strengths are: tracking (jets, hadron and vertex reconstruction) and muons. There is a dedicated LHCb talk [4], plus a general talk that also covers ALICE [5].
- Although a B⁺ cross section (13 TeV) paper by CMS just appeared [3], which utilizes $J/\psi K^+$ final state, this talk is not about B;
- Although e.g. CMS has once again rediscovered usual neutral vector mesons, from ω to $\Upsilon(nS)$ (and Z boson) at 13 TeV, this talk is not about onia;

But 2016 has been a very good year into LHC Run 2. As of Sept. 1st, LHC delivered close to 28 fb⁻¹ (41 fb⁻¹ by end of 2016 pp run), with ATLAS and CMS more than 90% efficient. In the following, we start from open charm spectroscopy, then turn to open charm production. Reflecting the interest even in the opening session of this CHARM conference, a good fraction of our time is spent discussing charm-jet tagging as an emerging new tool, before we close in a summary.

2. B_c studies

Though a little dated, the only contribution to open charm spectroscopy is the discovery, by ATLAS with both 7 and 8 TeV data, of an excitation of the B_c meson at $6842 \pm 4 \pm 5$ MeV that is consistent with the B_c(2S) state [6]. Detection is through reconstruction of B[±]_c \rightarrow J/ $\psi\pi^+$ that resonates with a $\pi^+\pi^-$ pair. Compared with ψ (2S) and Υ (2S), a B_c(2S) state that decays to B_c $\pi^+\pi^-$ has to exist, but this ATLAS state needs confirmation from LHCb and CMS.

Other than the excited B_c state, both ATLAS and CMS have measured some B_c properties. In a slightly dated paper, CMS measured [7]

$$R_{\rm C/u} = \frac{\sigma({\rm B}_{\rm C}^+)\mathscr{B}({\rm B}_{\rm C}^+ \to {\rm J}/\psi\pi^+)}{\sigma({\rm B}^+)\mathscr{B}({\rm B}^+ \to {\rm J}/\psi K^+)} = [0.48 \pm 0.05 \text{ (stat.)} \pm 0.03 \text{ (syst.)} \pm 0.05 \text{ (}\tau_{\rm B_{\rm C}}\text{)}]\%, (2.1)$$

complementary in kinematic region vs LHCb, i.e. for higher $p_T > 15$ GeV and more central |y| < 1.6. CMS also measured

$$R_{\rm B_{\rm C}} = \frac{\mathscr{B}({\rm B}_{\rm C}^+ \to {\rm J}/\psi\pi^+\pi^+\pi^-)}{\mathscr{B}({\rm B}_{\rm C}^+ \to {\rm J}/\psi\pi^+)} = 2.55 \pm 0.80 \; ({\rm stat.}) \pm \; 0.33 \; ({\rm syst.}) \stackrel{+0.04}{_{-0.01}} \; (\tau_{\rm B_{\rm C}}), \qquad (2.2)$$

confirming the result of LHCb, which has better statistics.

More recently, ATLAS has studied $B_c \rightarrow J/\psi D_S^{(*)}$ decays [8], including transverse polarization fraction for $B_c \rightarrow J/\psi D_S^*$. and compared with LHCb (which has better errors) and theory. The results are basically consistent.

So one can see that the standard bearer is LHCb for open charm spectroscopy and properties. Note also that, for all the studies mentioned, one relies on J/ψ as part of final state reconstruction. We therefore happily recall that this year is the 40th anniversary for the Nobel prize to Burt Richter and Sam Ting, while J/ψ itself is 42 years and going, and would remain an important tool for years to come.

3. Open charm production

Before considering open charm production, let us contrast with open beauty production. A bjet is typically tagged by a secondary vertex, and B-hadron production studies are often conducted with non-prompt $J/\psi \rightarrow \mu^+\mu^-$ as a tool, with the advantage of $c\tau(B) \sim 400$ to 500 μ m. In comparison, c-jet tagging is more difficult (next section), and D hadron production has lower multiplicity, without the luxury of J/ψ assistance, and has shorter $c\tau(D) \sim 100$ to 300 μ m. In the following, we report on $D^{(*)+}$ and D_8^+ production studies by ATLAS in pp collisions, and D^0 production in pp and PbPb collisions by CMS. Charge conjugate states are always implied.

3.1 $D^{(*)+}$ and D^+_S production in pp by ATLAS

ATLAS has studied $D^{(*)+}$ and D_{s}^{+} production in pp collisions at 7 TeV with 280 nb⁻¹ data [9], with the reconstructions $D^{*+} \rightarrow D^{0}(\rightarrow K^{+}\pi^{-})\pi_{s}^{+}$, where π_{s}^{+} is the "soft" pion descendent from D^{*+} ; $D^{+} \rightarrow K^{+}\pi^{-}\pi^{+}$; $D_{s}^{+} \rightarrow \phi(\rightarrow K^{+}K^{-})\pi^{+}$. These were for the kinematic range of $p_{T} \in$ (3.5, 100) GeV and $|\eta| < 2.1$ for the reconstructed *D* meson. Altogether, about 2900, 3700, 500 D^{*+}, D^{+} , D_{s}^{+} mesons, respectively, were reconstructed.



Figure 1: $d\sigma/dp_T$ and $d\sigma/d|\eta|$ for [9] the p_T regions of (3.5, 20) and (20 100) GeV, for D^{*+} (top) and D⁺ (bottom), compared with various NLO QCD calculations, with the FONLL range marked in green.

The differential p_T and rapidity distributions, $d\sigma/dp_T$ and $d\sigma/d|\eta|$, are given in Fig. 1 for the two p_T regions of (3.5, 20) and (20 100) GeV. Comparison is made with various NLO QCD calculations. While there is general agreement, taking the up to date fixed-order next-to-leadinglogarithm (FONLL) results [9, 10], data appear to be at upper limits of theory. Assuming FONLL to extrapolate to full kinematic space, ATLAS finds,

$$\sigma_{c\bar{c}}^{tot} = 8.6 \pm 0.3 \text{ (stat)} \pm 0.7 \text{ (syst)} \pm 0.3 \text{ (lum)} \pm 0.2 \text{ (ff)} ^{+3.8}_{-3.4} \text{ (extr) mb} \text{ (ATLAS)}, (3.1)$$

where "ff" marks the uncertainty due to fragmentation fraction, and "extr" for extrapolation procedure. This is consistent with the ALICE finding of [11]

$$\sigma_{C\bar{C}}^{\text{tot}} = 8.5 \pm 0.5 \text{ (stat)} ^{+1.0}_{-2.4} \text{ (syst)} \pm 0.3 \text{ (lum)} \pm 0.2 \text{ (ff)} ^{+5.0}_{-0.4} \text{ (extr) mb} \text{ (ALICE)}.$$
(3.2)

Note that the visible kinematic ranges for the two experiments are different, resulting in different extrapolation uncertainties.

3.2 Prompt D⁰ production in pp and PbPb by CMS

CMS has made the first measurement of prompt D^0 cross section in pp collisions at 5.02 TeV [12], and corresponding production in PbPb collisions at the same energy of $\sqrt{S_{NN}} = 5.02$ TeV. We illustrate in Fig. 2 the reconstruction of $D^0 \rightarrow K^+\pi^-$ events in pp and PbPb collisions, as well as the differential $d\sigma/dp_T$ distribution for the pp case, compared with theory calculations. One can see that, while there is agreement with theory, similar to D⁺ production studied by ATLAS [9], data is at upper side of FONLL predictions.



Figure 2: D⁰ candidate invariant mass distribution in pp (left) and PbPb (center) collisions at 5.02 TeV and the differential $d\sigma/dp_T$ distribution (right) for the former [12], compared with FONLL predictions.

The real aim of the CMS study, however, is to extract the nuclear modification factor R_{AA} for D⁰ production, by comparing PbPb to pp collisions. As discussed in talk by Geurts [13], one probes energy loss as the D⁰ travels through, and strongly interacts with, the nuclear medium. To quantify medium effects, CMS defines

$$R_{\rm AA}(p_T) = \frac{1}{T_{\rm AA}} \frac{dN_{\rm PbPb}^{\rm D^0}}{dp_T} / \frac{dN_{\rm pp}^{\rm D^0}}{dp_T},$$
(3.3)

for |y| < 1.0, where T_{AA} is the nuclear overlap function, which is related to N_{coll} (number of incoherent nucleon-nucleon collisions) defined through the ratio of double-differential cross sections [13]. In an earlier study by CMS at $\sqrt{S_{NN}} = 2.76$ TeV [14], the nuclear modification factor of prompt D⁰ cross section in PbPb collisions, called R_{AA}^* in that study (because the pp reference is extrapolated through FONLL), was found quite suppressed for $p_T \sim 5-10$ GeV. As can be seen from Fig. 3, the suppression appears more prominent for the centrality class of 0–10%, compared with the 0–100%. The result is consistent with the ALICE 7 TeV result [11], rescaled to 2.76 TeV using FONLL.

To further study this effect with less theory dependence, CMS took 25.8 pb⁻¹ and 404 μ b⁻¹ data for pp and PbPb collisions, respectively, both at 5.02 TeV [12], and made multiple checks and



Figure 3: Nuclear modification factor R_{AA}^* (pp reference scaled by FONLL) vs p_T for prompt D⁰ production in PbPb data at 2.76 TeV [14], for centrality classes 0–10% (left) and 0–100% (right).

comparisons, including the measurement of differential prompt D^0 production p_T distribution in pp collisions given in Fig. 2(right). We show in Fig. 4 the nuclear modification factor R_{AA} [12] for centrality classes 0–10% (left) and 0–100% (right), which are rather busy plots. Let us not comment on the details of comparison with theory and other concerns, which populate the plot, but offer some generic remarks.

First, the suppression of R_{AA} , or medium effect, is generic for p_T above a few GeV to more than 10 GeV. Even for the centrality range 0–100%, there is a factor of 4–5 suppression at $p_T \sim$ 6–7 GeV. Second, this suppression weakens for higher p_T , decreasing to only a factor of 1.5 for p_T in range of 60–100 GeV for the centrality range 0–100%. The large measured p_T range gives some challenge to theory, where none could fully capture the observed features. Third, within uncertainties, the D⁰ nuclear modification factor is consistent with R_{AA} of inclusive charged hadron



Figure 4: Nuclear modification factor R_{AA} vs p_T for D⁰ production in PbPb at 5.02 TeV [12], for centrality classes 0–10% (left) and 0–100% (right), compared with various theories (see Ref. [12] for references). The orange boxes with data marked in red give R_{AA} for inclusive charged hadron production.

production [15], shown as the orange boxes in Fig. 4. Finally, within errors, the 5.02 TeV study is compatible with the results of the 2.76 TeV study [14] given in Fig. 3, where the pp reference result was scaled through FONLL. Similarly, the 5.02 TeV result of CMS is compatible with the 2.76 TeV result of ALICE (scaled with FONLL by CMS [14]) with smaller |y| range.

4. Charm-jet tagging: tool development

If D meson production is a purpose and a means in itself, charm-jet tagging is a new frontier. As Lenz has raised [16] in his theory overview several times, and echoed by good interest in the audience, the $H \rightarrow c\bar{c}$ decay has not yet been probed. Imagine what a c-tagger could do besides probing $H \rightarrow c\bar{c}$: one could study $\tilde{t} \rightarrow \tilde{\chi}^0 + c$ in SUSY, or FCNC $t \rightarrow c + Z/\gamma$ and FCNH $t \rightarrow cH$ processes. The latter is a new frontier in itself, as nothing forbids it from first principles, despite popular *ad hoc* assumptions [17] such as discrete symmetries in two Higgs doublet models.

But c-jet tagging is considerably more difficult than b-tagging, as alluded to at the beginning of the previous section. A c-jet presents characteristics which are in between a b-jet and light jet (originating from light partons: u, d, s quark or gluon), such as having tracks and vertices less displaced than a b-jet, but more than a light jet. The b-tagging algorithms will therefore select part of the c-jets. While b-tagging is quite well developed and exploited, a dedicated c-tagger needs to exploit detailed c-jet properties versus b- and light jets.

4.1 Charm-tagged jet production in pPb and pp collisions by CMS

It is interesting that a study of charm-jets was performed on actual data involving heavy ions. CMS has made a first measurement of charm-jets in pPb collisions at 5.02 TeV [18], and compared it with pp collisions at 2.76 TeV. Charm-jets are identified by requiring a secondary vertex comprised of three or more charged tracks that are significantly displaced from the primary vertex. A variant of the secondary vertex mass is used to extract the relative contributions of jet flavors. It was found that jet energy modification in pPb collisions is consistent with pp collisions, even though comparison was made at different collision energies. Furthermore, the charm-jet fraction is



Figure 5: Simple secondary vertex (SSV) tagger for both b (left) and light (right) jets, as a function of c-jet tagging efficiency [18].

consistent with PYTHIA within uncertainties for both pPb and pp collisions. The results are being updated towards a paper.

Our point of discussion, however, is to compare the charm-jet tagging with a new c-tagger discussed in Sec. 4.3. The study of Ref. [18] uses a variant of the b-tagger, namely [19] simple secondary vertex (SSV), high purity (HP), and is not a dedicated c-jet tagger. To elaborate, we plot [18] SSV for b-jet and light quark jet vs c-jet tagging efficiency in Fig. 5. From the left plot we see the charm-to-bottom discrimination is basically unchanged. However, light jet mistag is reduced by a factor of 3 at the marked point on the blue curve of the right plot.

4.2 Charm-jet tagging at ATLAS

Before we present developments at CMS on a dedicated c-tagger, let us comment on the status at ATLAS. ATLAS has been designing [20] a charm-tagging algorithm, called JetFitterCharm, based on 8 TeV studies for both simulation and data. For 13 TeV at LHC Run 2, the IBL (Insertable B-Layer) would enhance performance, and ATLAS is developing a 2D optimization towards a dedicated c-tagger. Basically, it uses two BDTs where one is trained for c-jet against light jet, and one for c-jet against b-jet. While there is no public update yet, most ingredients can be found on ATLAS Flavour Tagging public pages. For example, b-jet (and c- and light jet) performance plots at 13 TeV can be seen in Ref. [21], while both c-jet and light jet calibrations at 8 TeV are described in Ref. [22].

One important issue is the validation of charm-jet mistag. ATLAS has measured [23], using the b-tagger, charm mistag with W + c production at 7 TeV. Based on this work, a poster [24] was presented at ICHEP2016, where ATLAS promises that a similar measurent with 13 TeV data is "expected soon".

4.3 Charm-jet tagging developments at CMS

In the following, we present the CMS 13 TeV c-tagging tool development based on 2015 data. A poster and a parallel session talk were also presented [25].

For Run2 [26], CMS provides two main b-taggers: CSVv2 and cMVAv2. The first one is an optimised version with respect to Run1, and combines information about tracks and secondary vertices (SV). The second one adds another layer and combines with a multivariate analysis approach the output of CSVv2 with other taggers like Jet Probability (based only on track information) or the Soft Lepton Taggers (based on the presence of a soft lepton (SL) within the jet). For the c-tagger developed [27] for Run 2, which uses 2015 data for calibrating the algorithm, one simplifies and makes a straightforward combination of tracks, SV and SL info.

The issue at hand is to separate c-jet from b-jet as well as light jet backgrounds. The solution is to have *two* BDTs: c- vs light (CvsL) jets and c- vs b- (CvsB) jets [27]. The c-tagger is trained on simulated QCD multijets, but the performance is also validated on simulated tt samples. The corresponding performance [27] are given in Fig. 6. The spikes in discriminator shapes on the left figures appear for jets in which no track passes the specific selection criteria [27]. For the final performance displayed in Fig. 6(right), results from current cMVAv2 and CSVv2 b-tagger algorithms [26] are also shown for comparison. Here, the blue curves correspond to the scales in blue on the left for light-jet efficiency, while the red curves correspond to b jet efficiencies and refer



Figure 6: Left [27]: CvsL (top) and CvsB (bottom) discriminators, normalized for each flavour. Right [27]: final performance of CvsL (blue solid line and axis) and CvsB (red solid line and axis) validated on tt, and compared with using various b-tagging algorithms (dash and dotted lines).



Figure 7: Left [27]: 2D scatter for b (red), c (green), and light (blue) jets, in the plane of CvsL and CvsB. Right [27]: contours in plane of bottom and light mistag efficiencies for different values of constant charm efficiency. The markers L, M and T are explained in text.

to the scale on the right. Just like in Fig. 5, closer to the lower right means better performance. One sees that, compared with current b-tagger tools, CvsL discriminates better between light and charm jets, while CvsB performs worse. The latter is due to the fact that this first implementation of the tagger focused on rejecting light flavor jets. An improved version of the tagger, with a stronger separation between c-jets and b-jets, is being developed.

We show in Fig. 7 (left) the distribution of jets of different flavours in the plane formed by the two discriminators [27]. The two classifiers output a value close to 1 (-1) for signal-like (background-like) jets. Thus c jets will be located towards the upper right corner, while b and light jets populate mostly the bottom right and the top left corners, respectively. Contours of constant charm efficiency are shown in Fig. 7(right) in the plane of light and b-jet mistag efficiencies. To separate c-jets from the background, a rectangular cut is placed to isolate the upper right corner in Fig. 7(left). Three such cuts, L(oose), M(edium) and T(ight) are marked, which correspond to the

working points (WP) marked in Fig. 7(right). In particular, for c-tagger T (ctagT), one has CvsL > 0.45, CvsB > -0.35, and ε^{c} , ε^{b} , $\varepsilon^{\text{light}} = 0.2$, 0.24, 0.02, respectively.

We use ctagT to illustrate the validation on data [27] of the charm tagger algorithm. Two methods are used. The first uses the very pure sample of c-jets from $g + s(d) \rightarrow W^- + c$. Since the signal W and c charges are correlated, while QCD background of $c\bar{c}$ pair production in association with a W is not, subtracting same-sign from opposite-sign events effectively eliminates the latter. A second validation utilizes semileptonic decay of a tt pair, where roughly 25% of jets are c-jets. Utilizing weak decay properties that boost u-type quarks, one can infer the efficiency of the charm-tagging algorithm. From these data vs simulation validations, one measures the scale factors, or ratio of efficiencies,

$$SF_{\rm C} = \varepsilon_{\rm C}^{\rm data} / \varepsilon_{\rm C}^{\rm MC}.$$
 (4.1)

as depicted in Fig. 8 for the working point ctagT. Validation uses 13 TeV data of 2015. The W + c validation provides four momentum bins, while the tt validation is not binned in transverse momentum of the jet due to statistical limitations. The extracted scaling factor is close to 1 [27].



Figure 8: Data-to-simulation scale factor SF of the c-tagging efficiency for working point c-tagger Tight [27]. Upper panel gives measured values by two methods, with (thick error bar) statistical error and (narrow error bar) combined statistical and systematic uncertainties. The hatched area is the combined SF value with overall uncertainty, displayed again in lower panel with solid line as the linear fit.

5. Summary

Although ATLAS and CMS are central, general purpose detectors, they have contributed to open charm spectroscopy, such as finding an excited B_c state, and explored B_c meson properties. Both experiments have studied prompt D meson production in pp collisions, while CMS has measured the Nuclear Modification Factor R_{AA} . For prompt D⁰ production in PbPb vs pp collisions, strong suppression is observed for p_T ranging from a few GeV to over 10 GeV. This medium-induced suppression for prompt D⁰ is consistent with what is observed for inclusive charged parti-

cle production. The high p_T capability allows one to probe up to 100 GeV for D⁰, and hundreds of GeV for charged hadrons.

Following the great impact of b-tagging algorithms, ATLAS and CMS are developing dedicated charm-jet taggers for LHC Run 2. For example, for the first time at CMS, a 2D c-tagger is validated with 13 TeV data of 2015, with room for further improvement. ATLAS has similar developments, but the 2D c-tagger for Run 2 is not yet public. Let's hope we would soon hear from actual applications in future analyses.

References

- [1] ATLAS Collaboration, JINST 3 S08003 (2008).
- [2] CMS Collaboration, JINST 3 S08004 (2008).
- [3] V. Khachatryan et al. [CMS Collaboration], arXiv:1609.00873 [hep-ex].
- [4] P. Spradlin, this proceedings.
- [5] A. Mischke, this proceedings.
- [6] ATLAS Collaboration, Phys. Rev. Lett. 113, 212004 (2014) [arXiv:1407.1032 [hep-ex]].
- [7] V. Khachatryan et al. [CMS Collaboration], JHEP 1501, 063 (2015) [arXiv:1410.5729 [hep-ex]].
- [8] ATLAS Collaboration, Eur. Phys. J. C 76, 4 (2016) [arXiv:1507.07099 [hep-ex]].
- [9] ATLAS Collaboration, Nucl. Phys. B 907, 717 (2016) [arXiv:1512.02913 [hep-ex]].
- [10] M. Cacciari et al., JHEP 1210, 137 (2012); for FONLL webpage update, see Ref. [9].
- [11] B. Abelev et al. [ALICE Collaboration], JHEP 1201, 128 (2012) [arXiv:1111.1553 [hep-ex]].
- [12] CMS PAS HIN-16-001 [CMS Collaboration].
- [13] F. Geurts, this proceedings.
- [14] CMS PAS HIN-15-005 [CMS Collaboration].
- [15] V. Khachatryan et al. [CMS Collaboration], arXiv:1611.01664 [nucl-ex].
- [16] A. Lenz, this proceedings.
- [17] S. L. Glashow and S. Weinberg, Phys. Rev. D 15, 1958 (1977).
- [18] CMS PAS HIN-15-012 [CMS Collaboration].
- [19] S. Chatrchyan et al. [CMS Collaboration], JINST 8, P04013 (2013) [arXiv:1211.4462 [hep-ex]].
- [20] ATL-PHYS-PUB-2015-001 [ATLAS Collaboration].
- [21] See http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2016-001/
- [22] ATLAS-CONF-2014-046 [ATLAS Collaboration].
- [23] ATLAS-CONF-2013-109 [ATLAS Collaboration].
- [24] Poster presented by A. Lapertosa at ICHEP2016, August 2016, Chicago, USA.
- [25] Poster and parallel session talk by S. Moortgat, this proceedings.
- [26] CMS PAS BTV-15-001 [CMS Collaboration].
- [27] CMS PAS BTV-16-001 [CMS Collaboration].