Angular correlations

Event classification study

Overview

Based on questions received during QGP France presentation, e.g.

- Studies are not minimum-bias. Doing equivalent studies with MB will be about a factor 100 for charm and 1000 for beauty. This is achievable to a certain extent, but should be considered carefully.
- Consider effect of jets in high multiplicity pp. This means we might have to consider using different event classifications than just the multiplicity.

Investigate other event classification variables based on two papers:

- About Nch as nMPI proxy: https://journals.aps.org/prd/pdf/10.1103/PhysRevD.102.076014
- About the plateau at high Nch, Fig:10: https://inspirehep.net/files/0fc2ffef7a5e4a6073e599630b3fd2d8

Method

Event classification variables:

- Multiplicity (number of prompt primaries per event)
- Number of MPIs
- Number of jets (using jet finder algorithm in PYTHIA, considering double pTjetMin = 5.0 GeV)
- Average pT of all particles in event (note that this was the motivation to develop colour reconnections in PYTHIA; already in MONASH, because the trends seen as a function of multiplicity in e+e- collisions were not predicted by earlier versions of PYTHIA)
- Sphericity (jet variable basically measuring how ‘spherical’ the event is. Potentially usable for jet identification)
- Only B+B- and D+D- pairs (junctions effect might be larger with baryons..)

Method

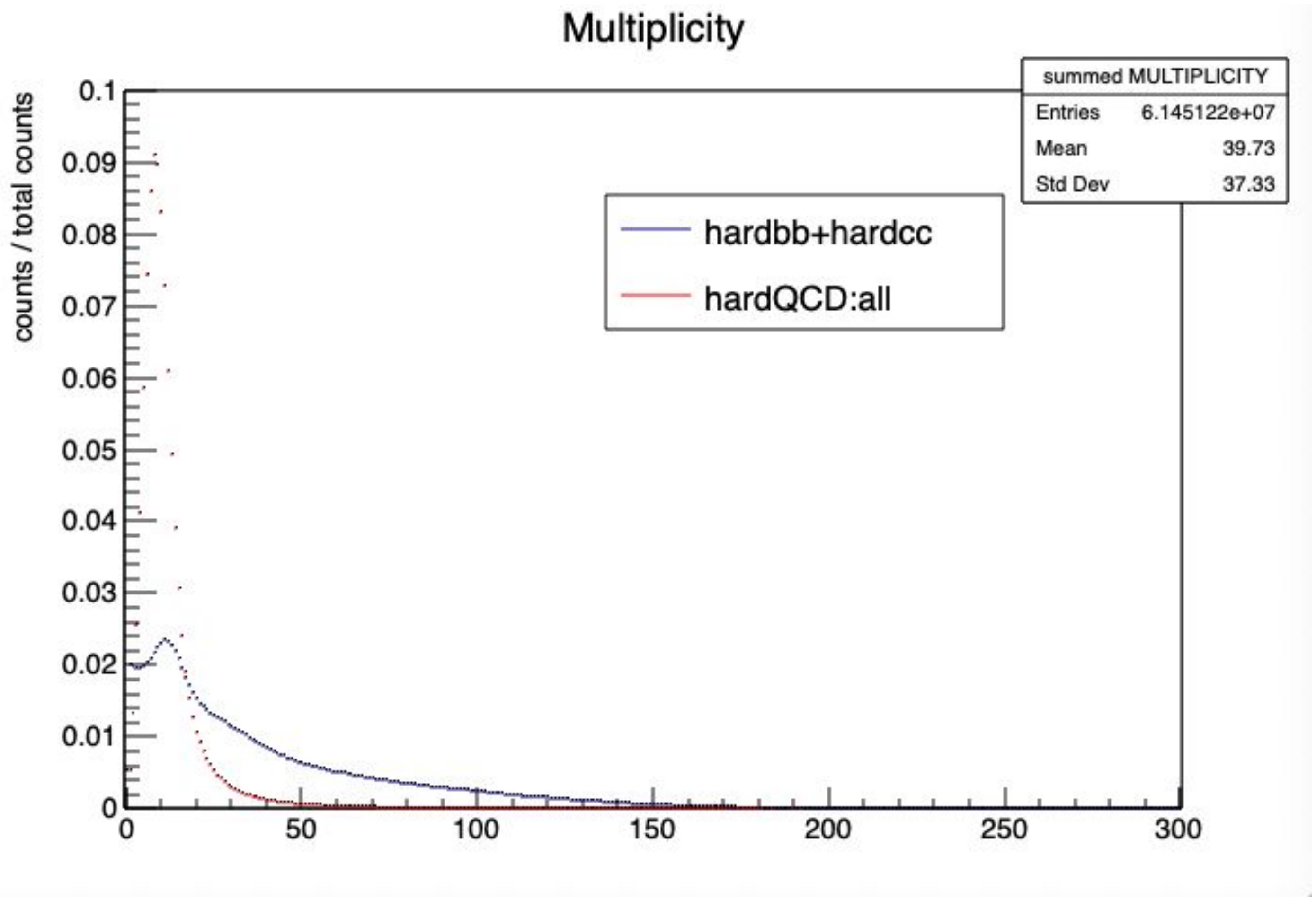
The simulations used to only save b or c hadrons, changed now to save b or c hadrons and anything else.

- This means the b and c hadrons do not decay, so there is a slight lack of final-state particles, as well as typically higher pT for the b and c hadrons, etc. These effects I take into account by splitting the histograms into a ‘all’ and a ‘soft’ part (the soft part ignores all b or c hadrons). The effect seems to be small, expect for the <pT> plots.
- In that sense, even the minimum bias events are not really minimum-bias, but it will be a tiny effect in those (as they don’t produce so many charm or beauty hadrons). It is a slightly larger effect for the biased simulations, but we don’t care about it anyways there.

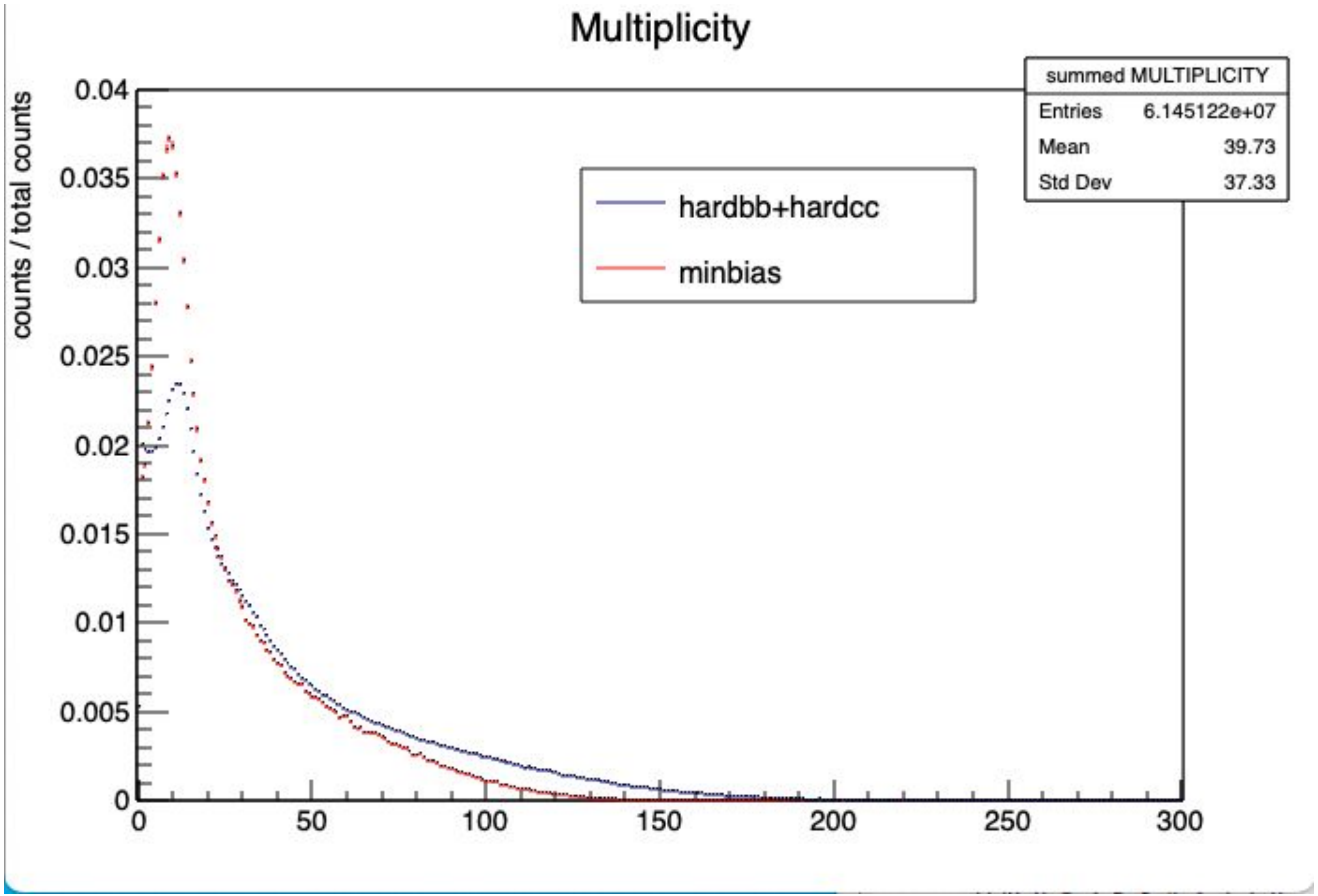
Method

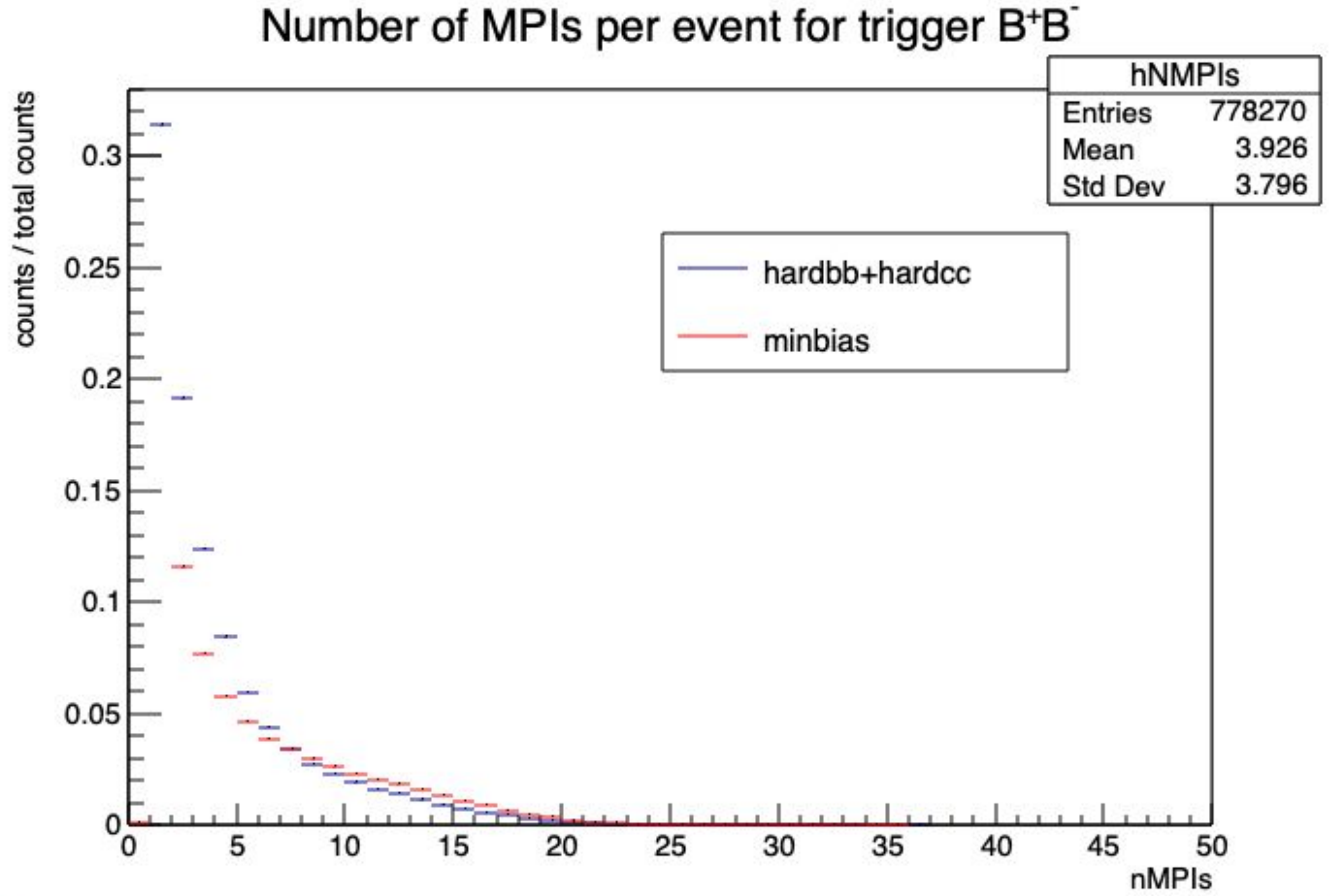
Statistics is really limited for now: 1e4 (junctions) or 1e5 (monash) events. This can be increased if it is necessary, without too many problems. For a first study, what I did now should be okay.

05/10/2025. Comparisons using hardbbbar + hardccbar at the same time

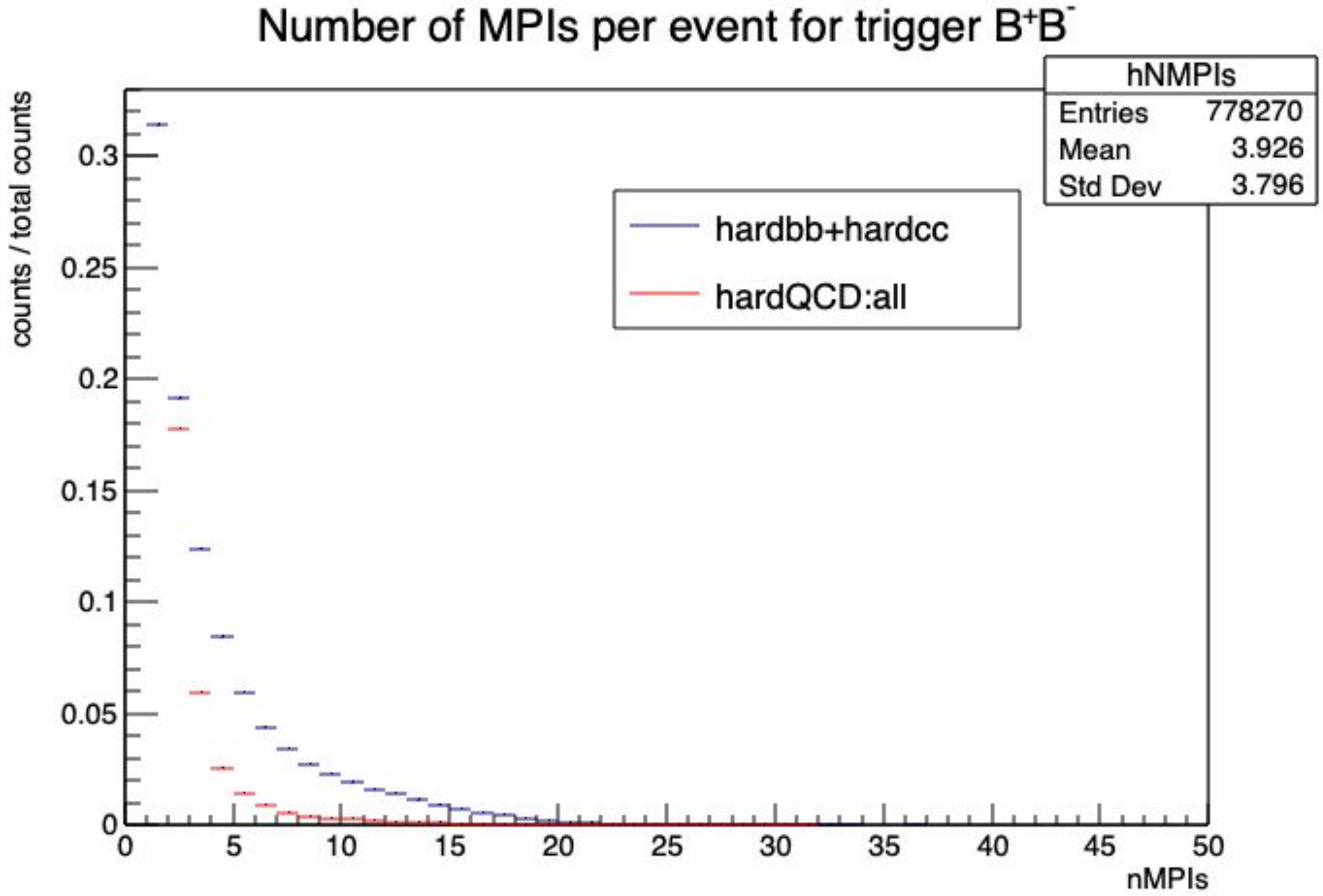


Multiplicity
summed MULTIPLICITY
Entries 6.145122e+07
Mean 39.73
Std Dev 37.33
hardbb+hardcc
hardQCD:all
counts / total counts
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
250
300

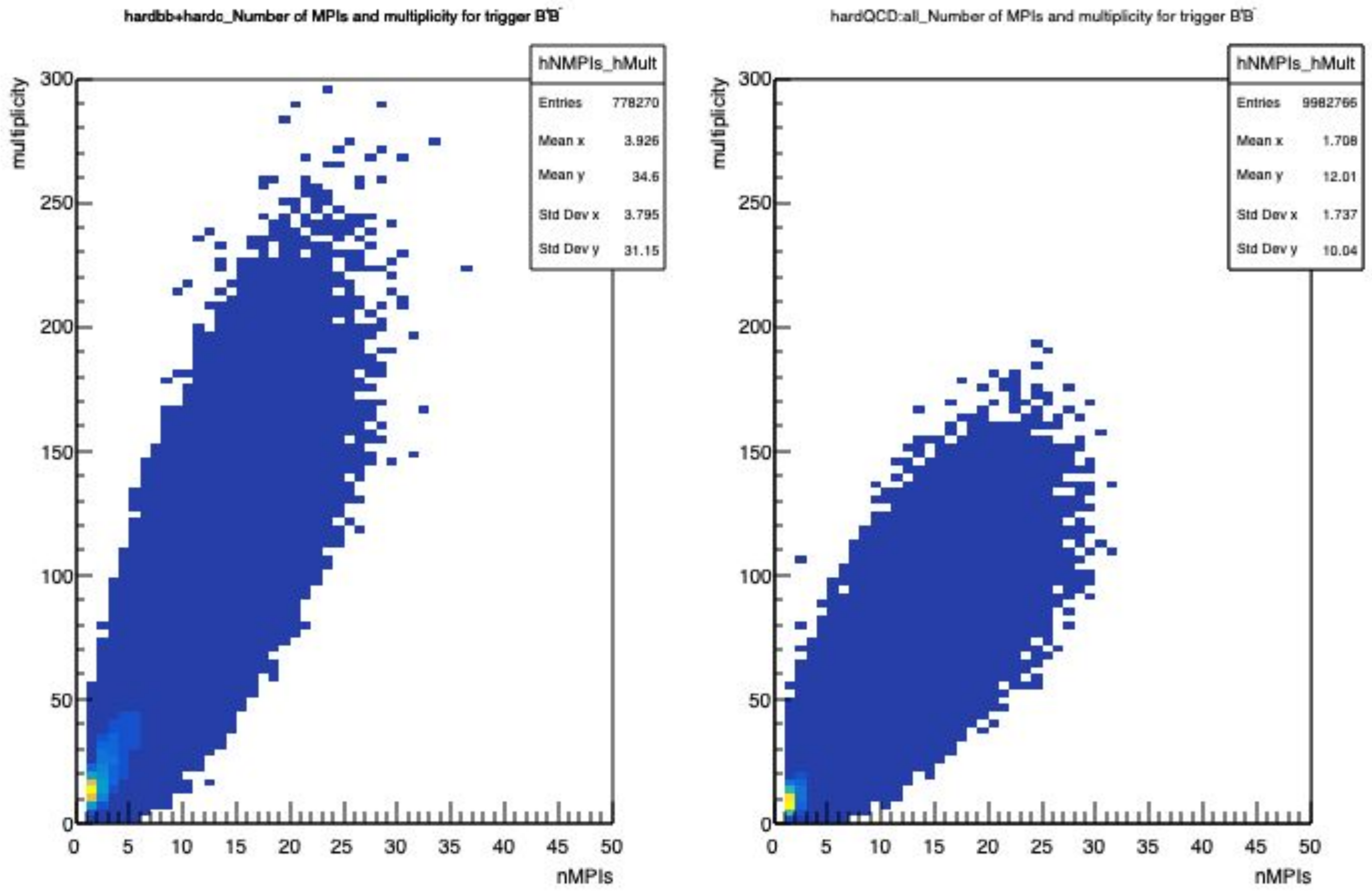

Multiplicity
summed MULTIPLICITY
Entries 6.145122e+07
Mean 39.73
Std Dev 37.33
hardbb+hardcc
minbias
counts / total counts
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
50
100
150
200
250
300

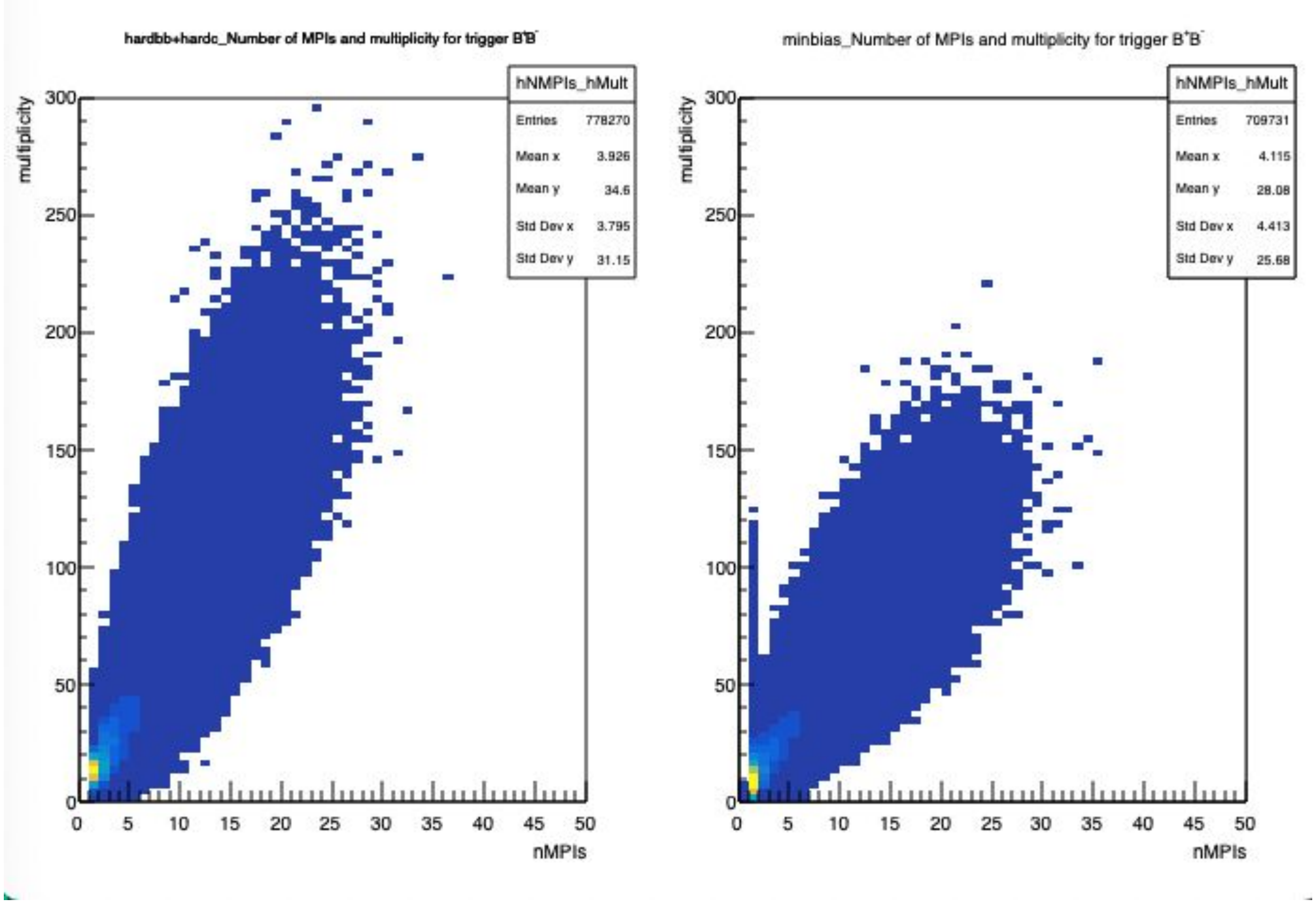


Number of MPIs per event for trigger B^+B^-
hNMPIs
Entries 778270
Mean 3.926
Std Dev 3.796
hardbb+hardcc
minbias
counts / total counts
0.3
0.25
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40 45 50
nMPIs



Number of MPIs per event for trigger B^+B^-
hNMPIs
Entries 778270
Mean 3.926
Std Dev 3.796
hardbb+hardcc
hardQCD:all
counts / total counts
0.3
0.25
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25 30 35 40 45 50
nMPIs


hardbb+hardcc_Number of MPIs and multiplicity for trigger B⁺B⁻
hNMPIs_hMult
Entries 778270
Mean x 3.926
Mean y 34.6
Std Dev x 3.795
Std Dev y 31.15
multiplicity
nMPIs
hardQCD:all_Number of MPIs and multiplicity for trigger B⁺B⁻
hNMPIs_hMult
Entries 9982766
Mean x 1.708
Mean y 12.01
Std Dev x 1.737
Std Dev y 10.04
multiplicity
nMPIs


hardbb+hardcc_Number of MPIs and multiplicity for trigger B+B-
hNMPIs_hMult
Entries 778270
Mean x 3.926
Mean y 34.6
Std Dev x 3.795
Std Dev y 31.15
multiplicity
nMPIs
minbias_Number of MPIs and multiplicity for trigger B+B-
hNMPIs_hMult
Entries 709731
Mean x 4.115
Mean y 28.08
Std Dev x 4.413
Std Dev y 25.68
multiplicity
nMPIs

Conclusions

Interestingly enough, it seems like the hardbbbar + hardccbar is more close to minimum bias than it is to hardQCD:all

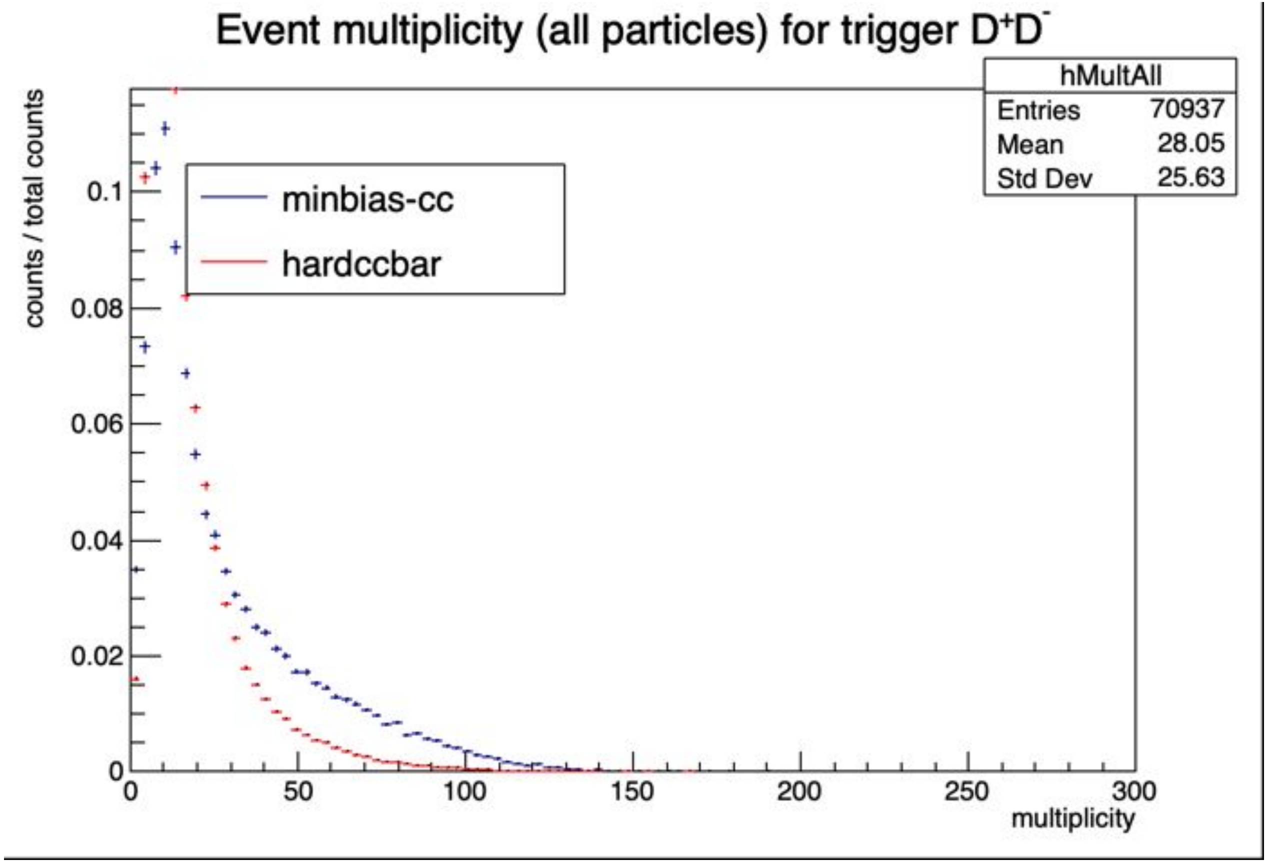
- Except for a trend of an increased amount of lower multiplicities -> makes sense due to absence of low pT jets in hardccbar + hardbbbar

Another huge advantage is that we would now use the same event structure and same multiplicity distribution for both beauty and charm -> fair comparison, less bias

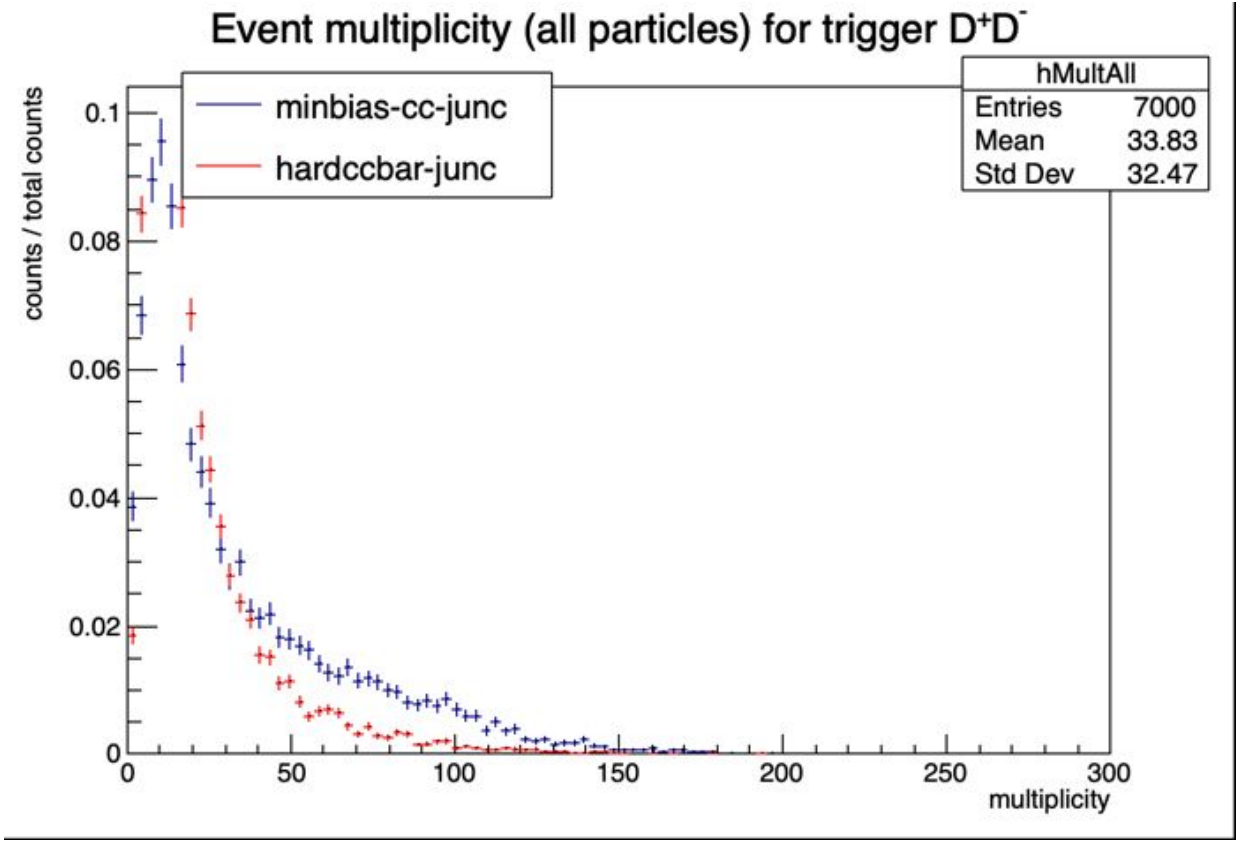
Comparison might even be better than using hardQCD:all and gives us way more statistics (factor 10^4 difference for beauty)

Translation from nMPIs to multiplicity looks about the same (and better than the MB, as before)

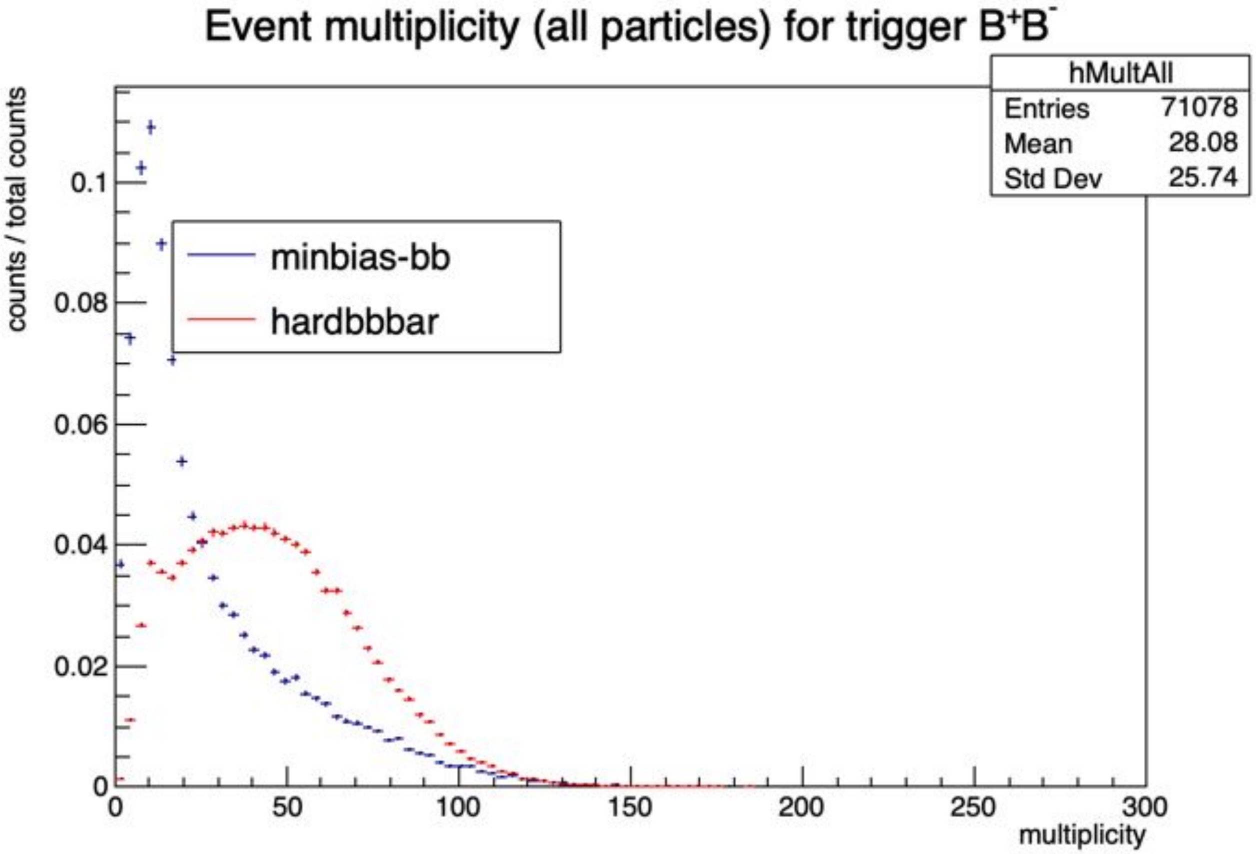
Charm Monash: multiplicity



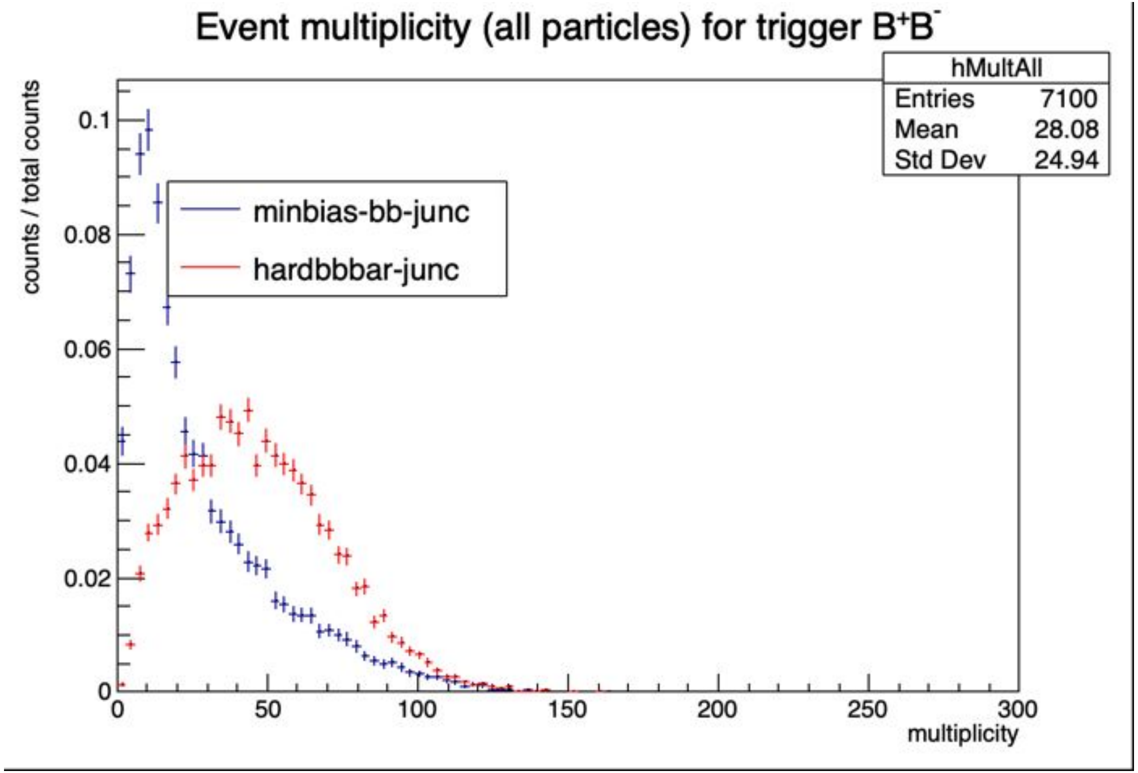
Charm Junctions: multiplicity



Beauty Monash: multiplicity



Beauty Junctions: multiplicity



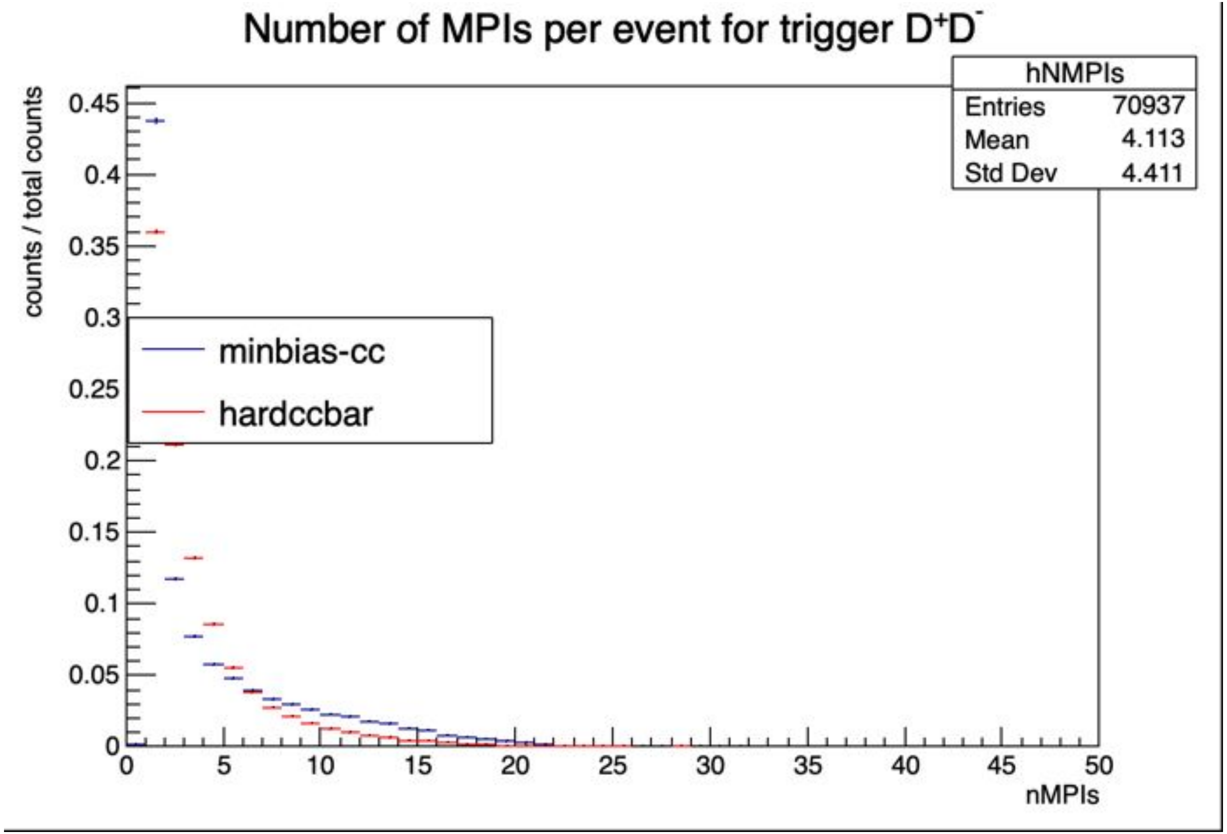
Conclusion: multiplicity

Big differences between hardccbar and hardbbbar.. In particular, the hardbbbar has a very particular shape.

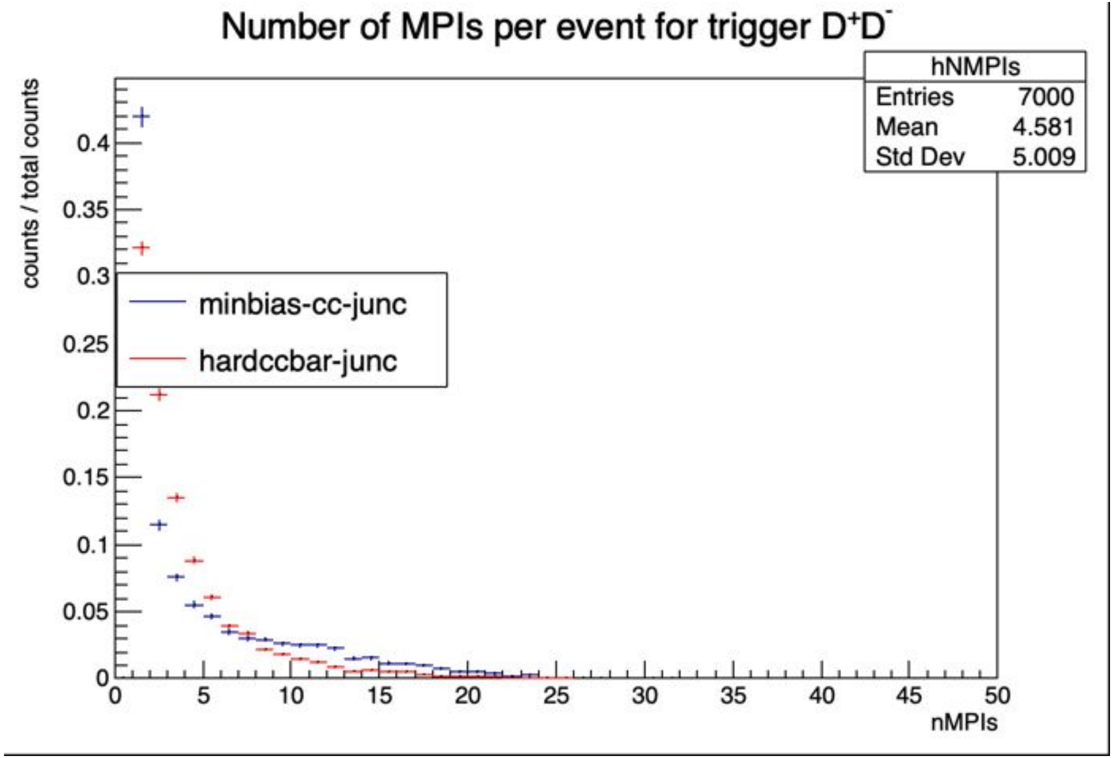
Also unexpected that hardccbar seems to have lower multiplicities than the minbias-cc

Monash and junctions are consistent with each other

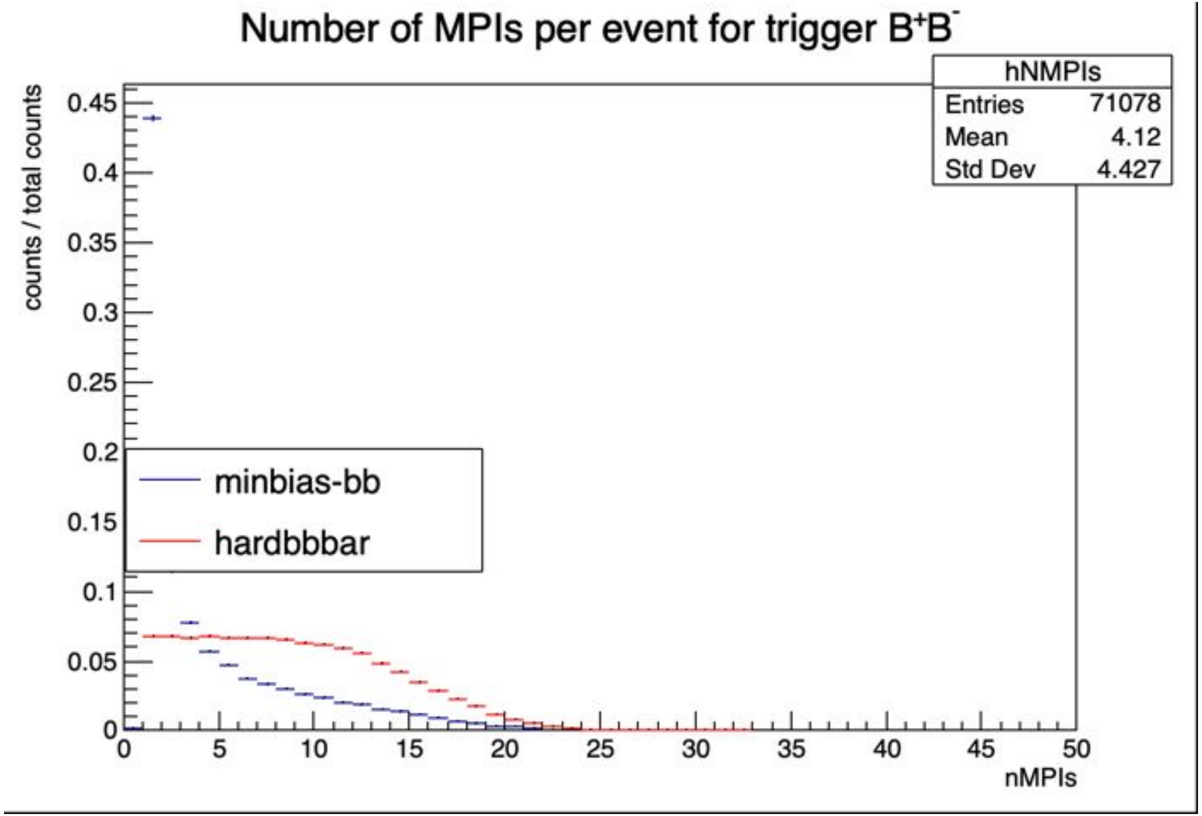
Charm Monash: number of MPIs



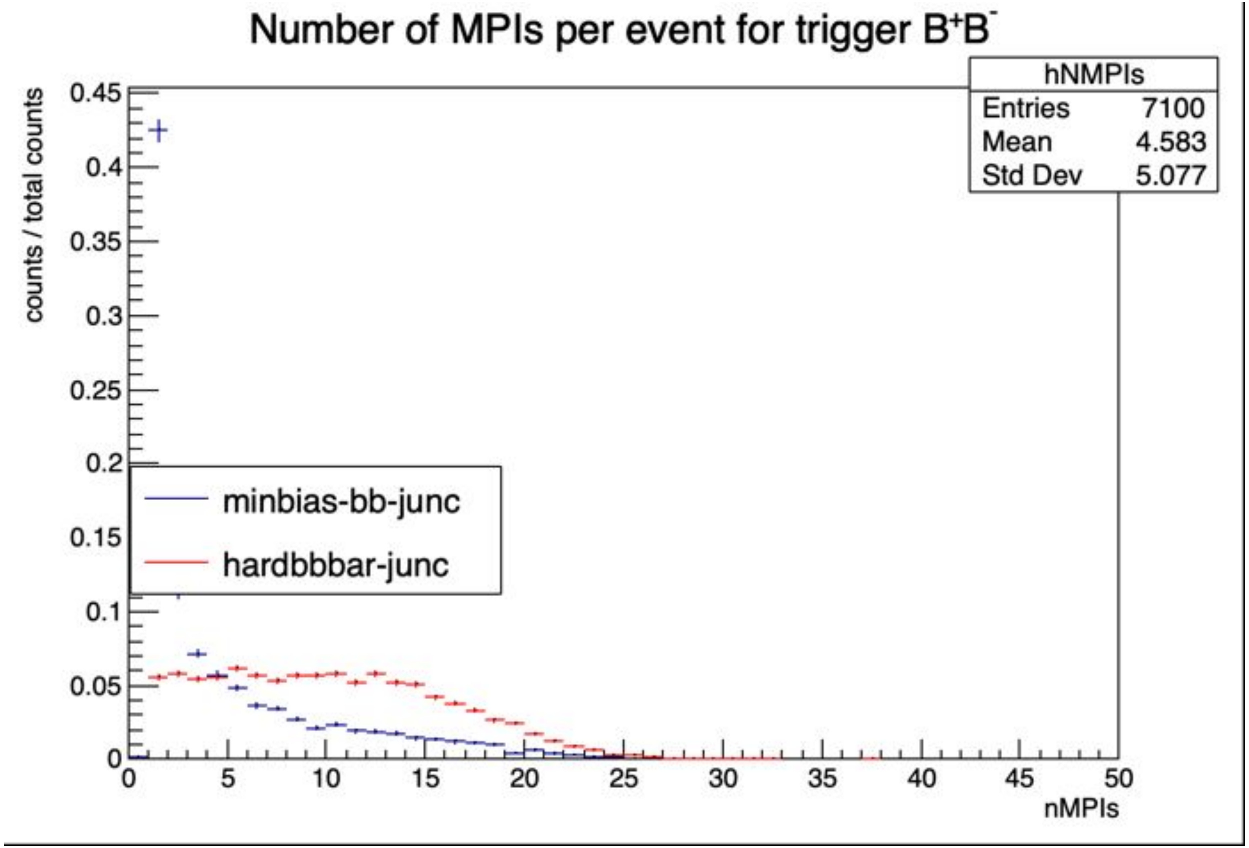
Charm Junctions: number of MPIs



Beauty Monash: number of MPIs



Beauty Junctions: number of MPIs

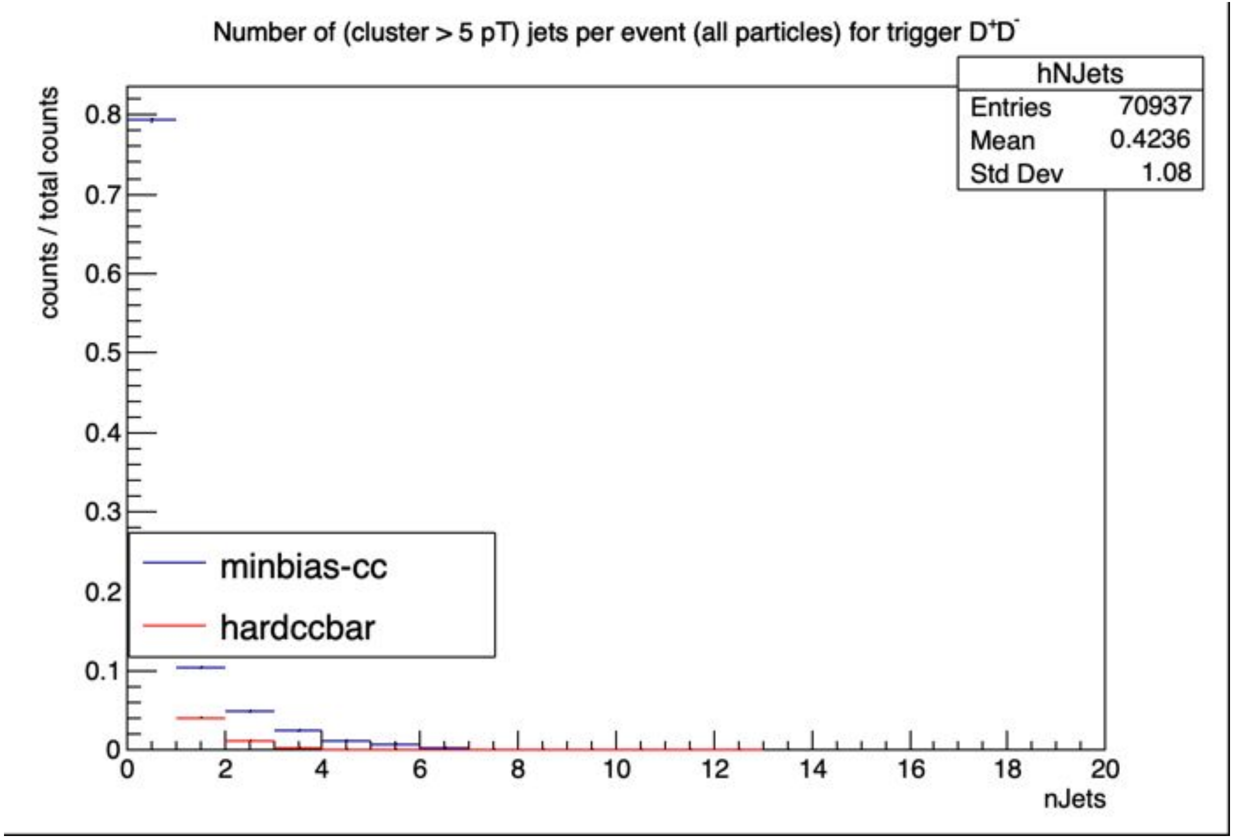


Conclusion: number of MPIs

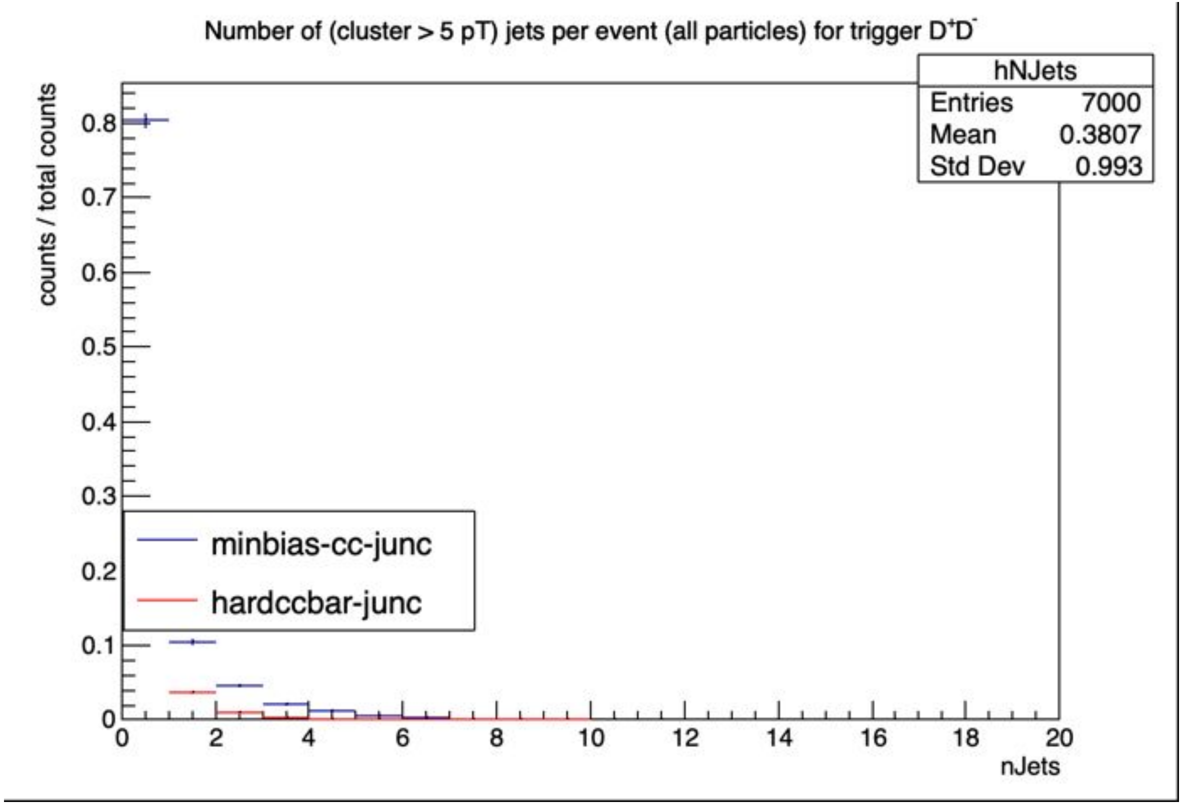
Similar behaviour as for multiplicity, see also the 2D-histograms later for their correlation

Variable: number of jets

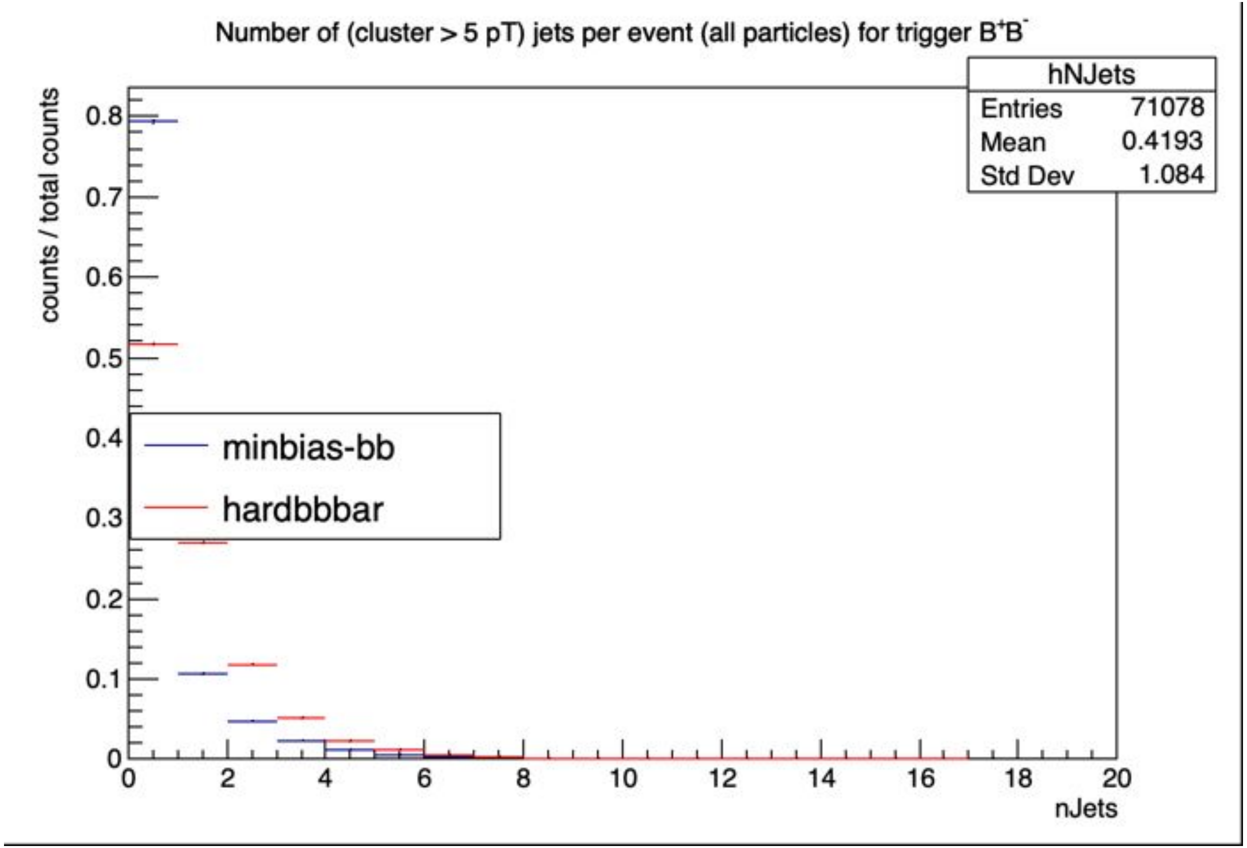
Charm Monash: number of jets



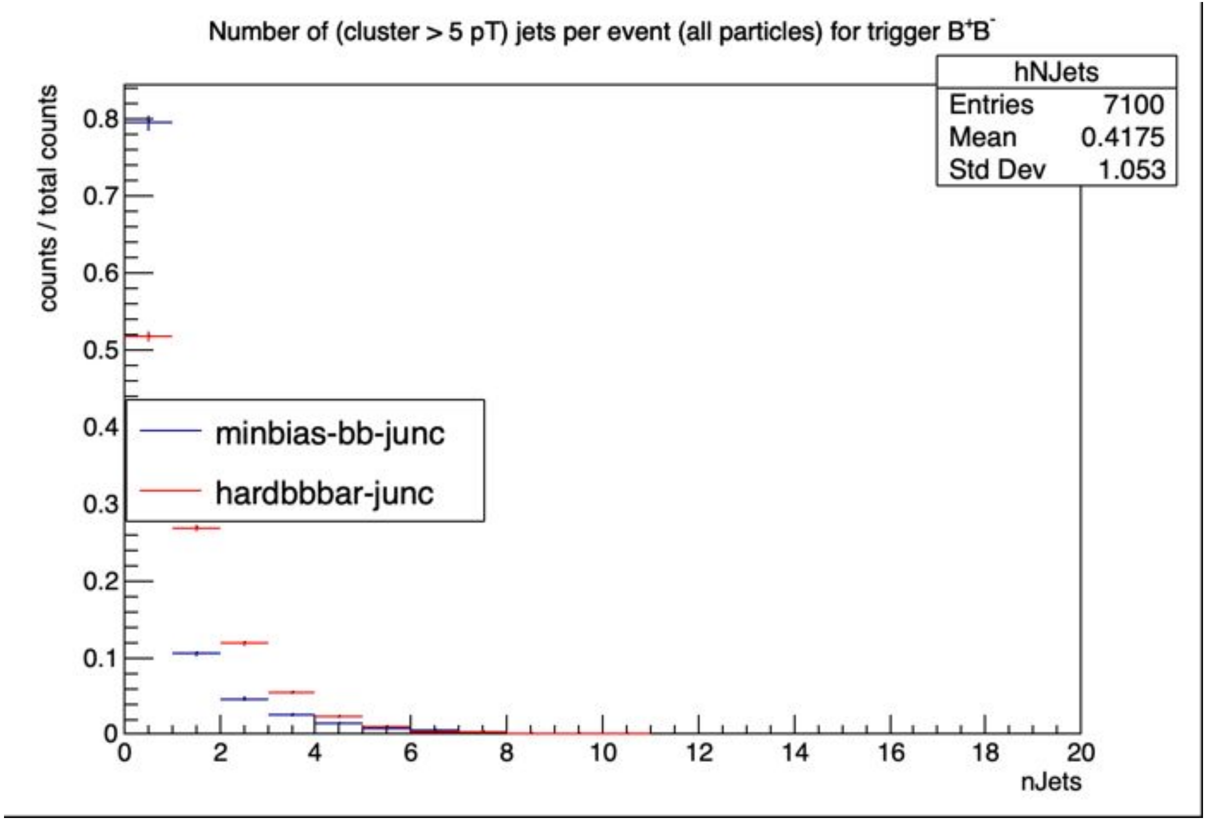
Charm Junctions: number of jets



Beauty Monash: number of jets



Beauty Junctions: number of jets



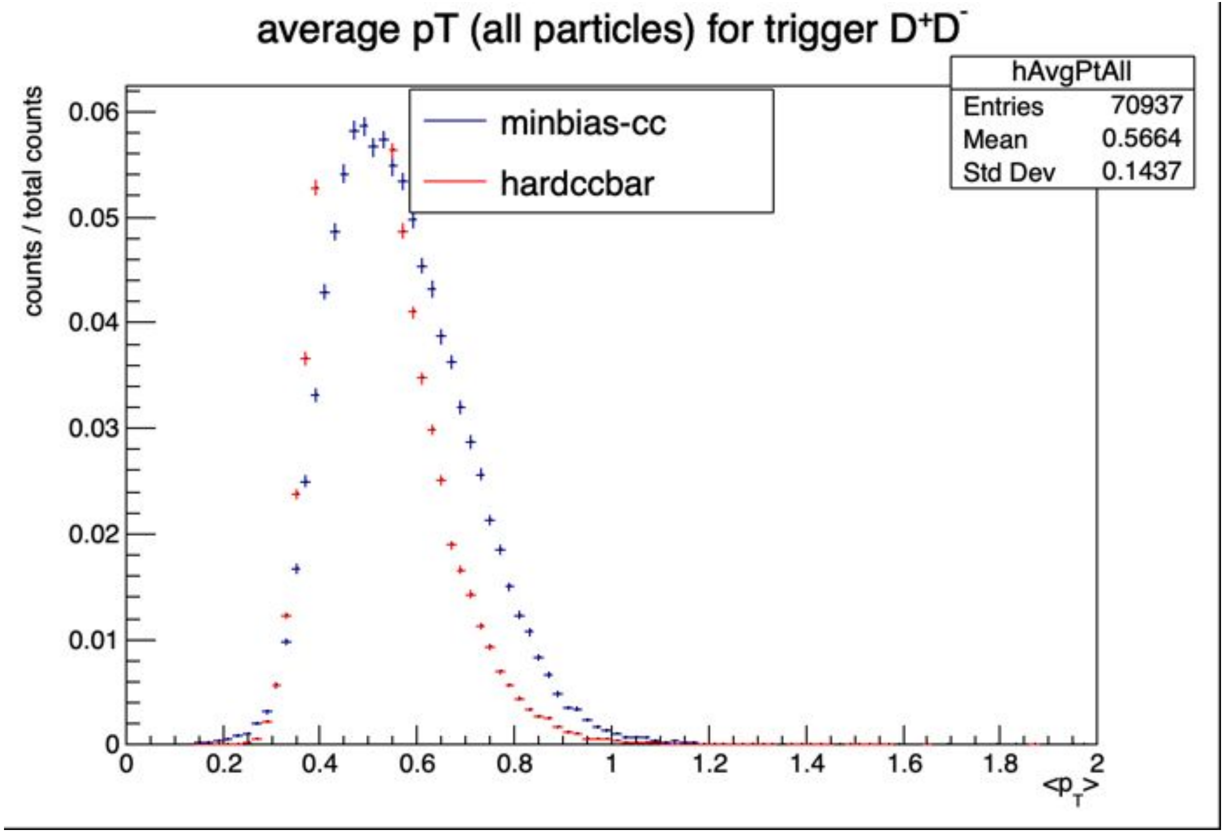
Conclusion: number of jets

MB has a lot of 0 jet events, in general more events with fewer jets than biased simulations (makes sense).

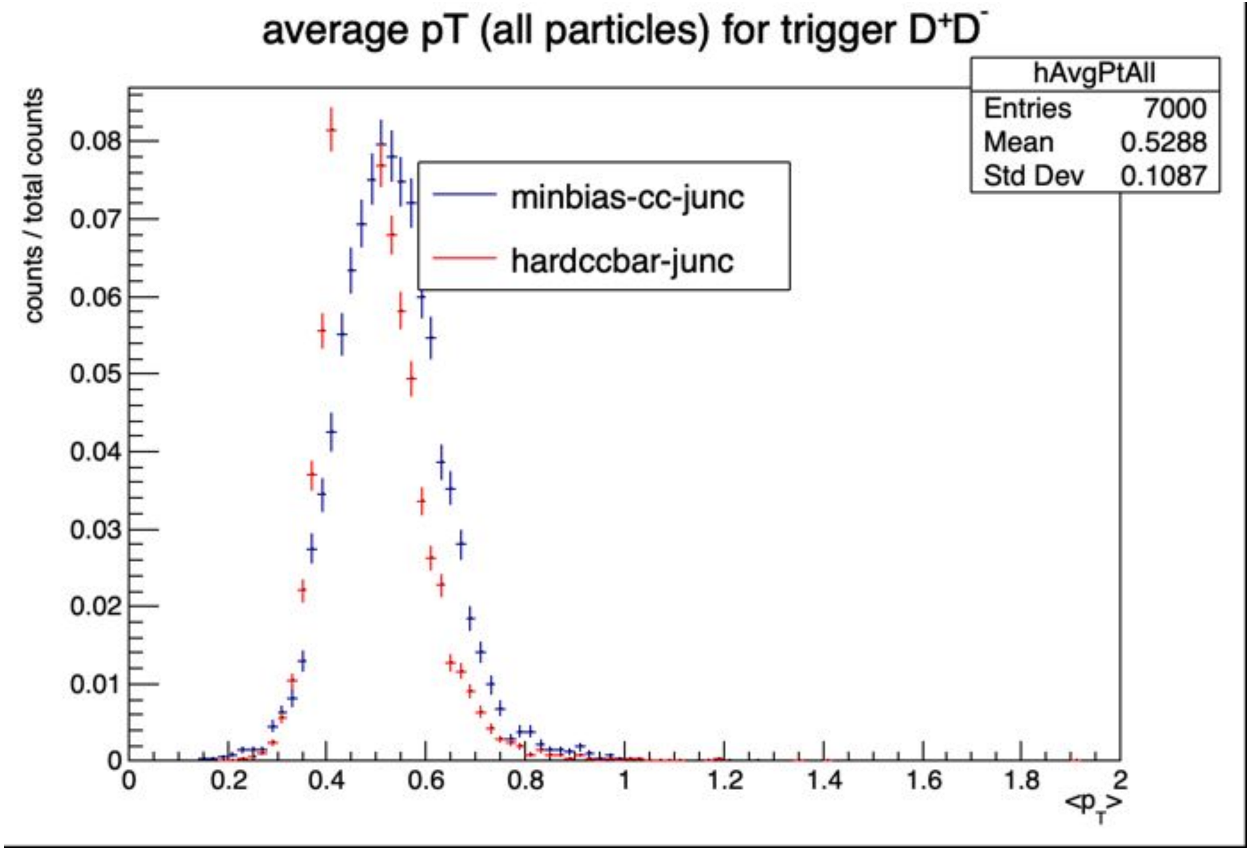
- Somehow for beauty this trend is mostly inverted. It seems like the hardccbar simulations are more ‘natural’ and closer to MB, while the bb-bar are more artificial. Maybe a hard bb jet is usually formed in the biased bb simulations, while for charm it’s less ‘forced’?
- At high number of jet events this seems to ‘converge’ to the same trend for beauty and for charm

Monash and junctions show similar trends

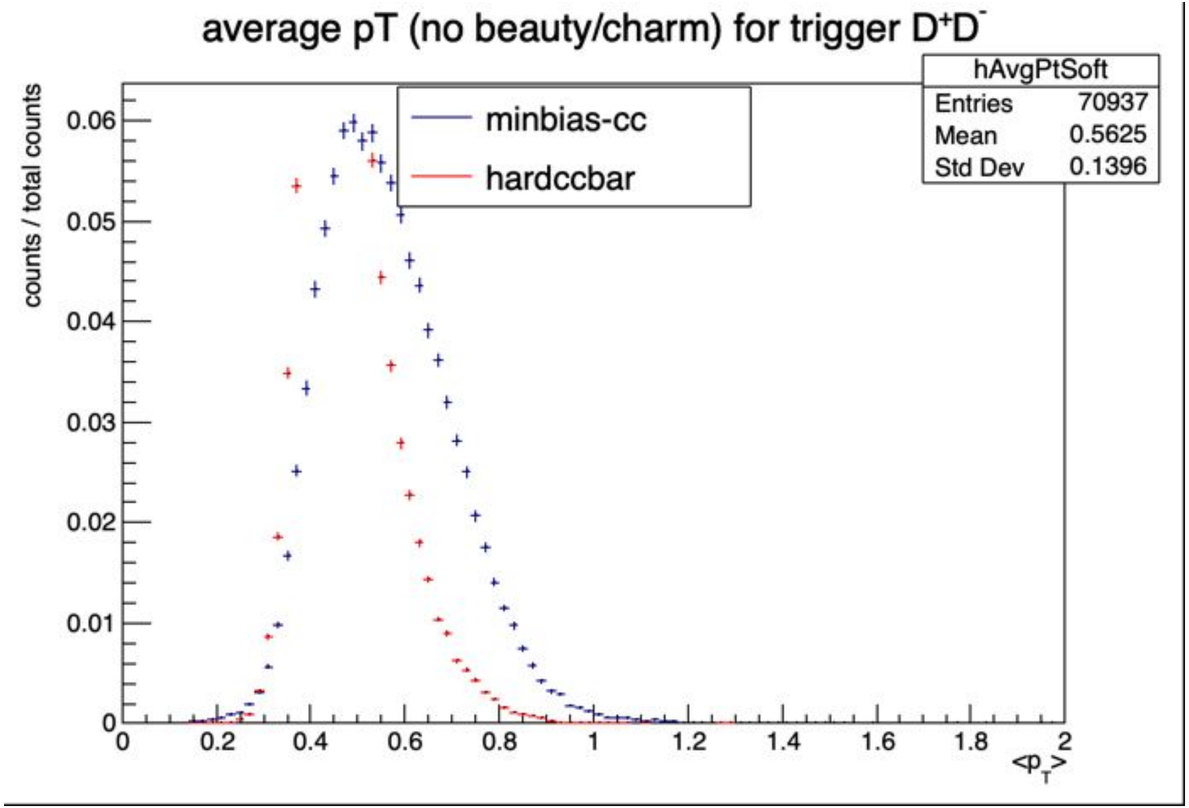
Charm Monash: <pT> (all)



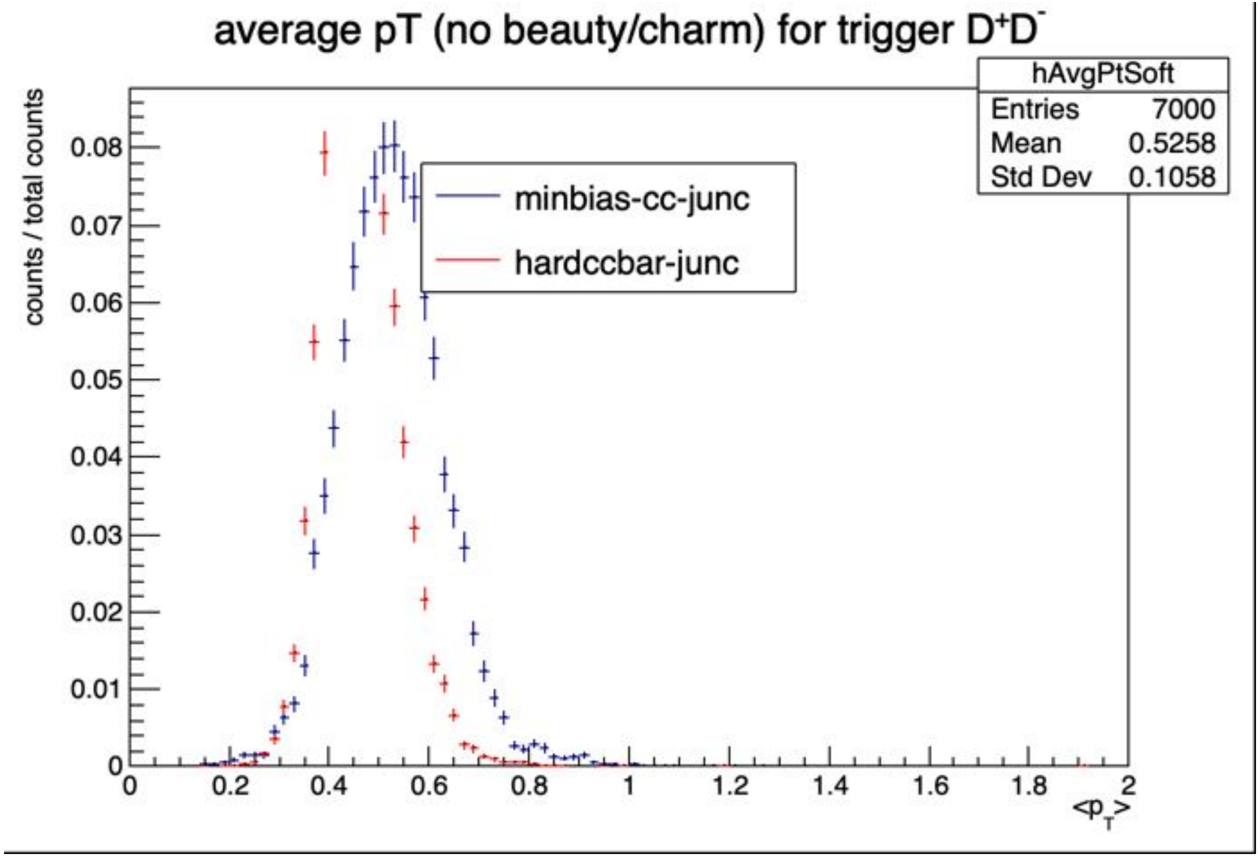
Charm Junctions: <pT> (all)



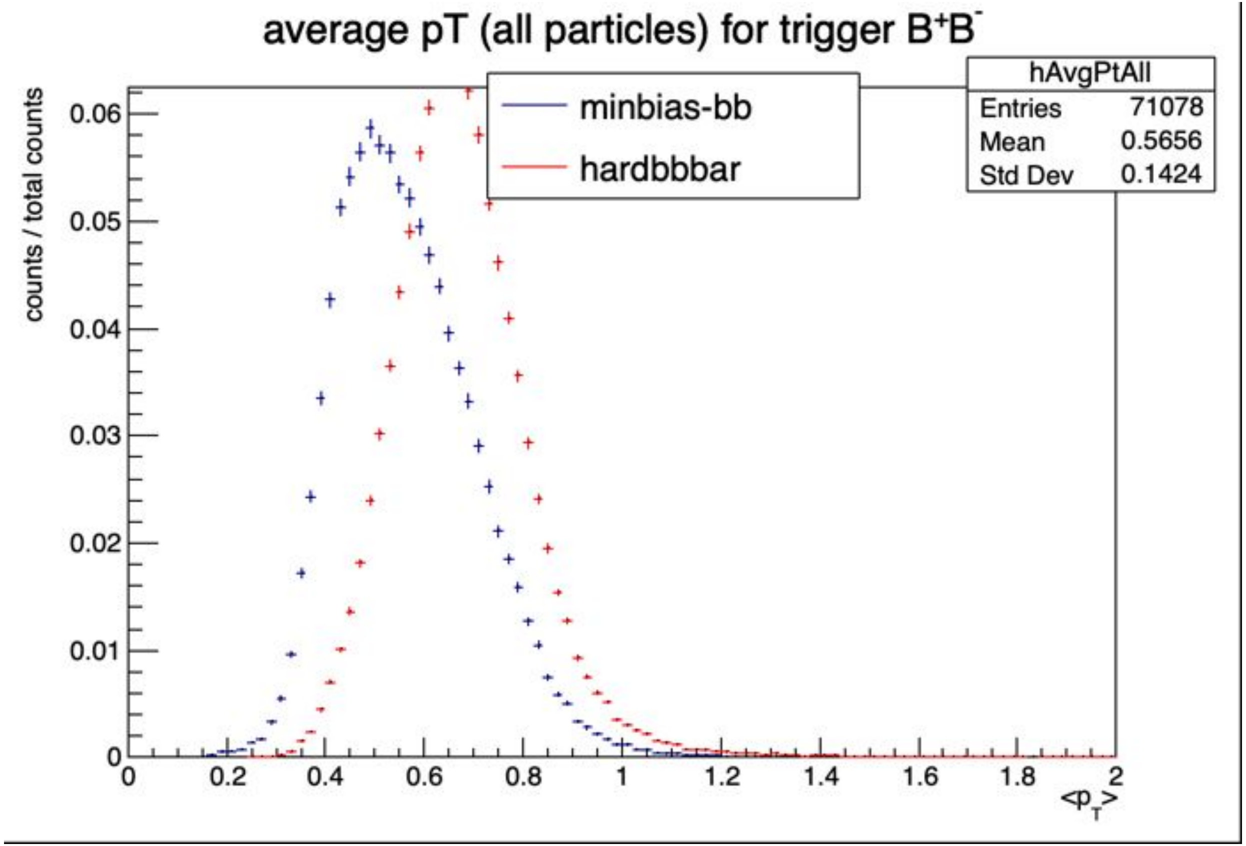
Charm Monash: <pT> (soft)



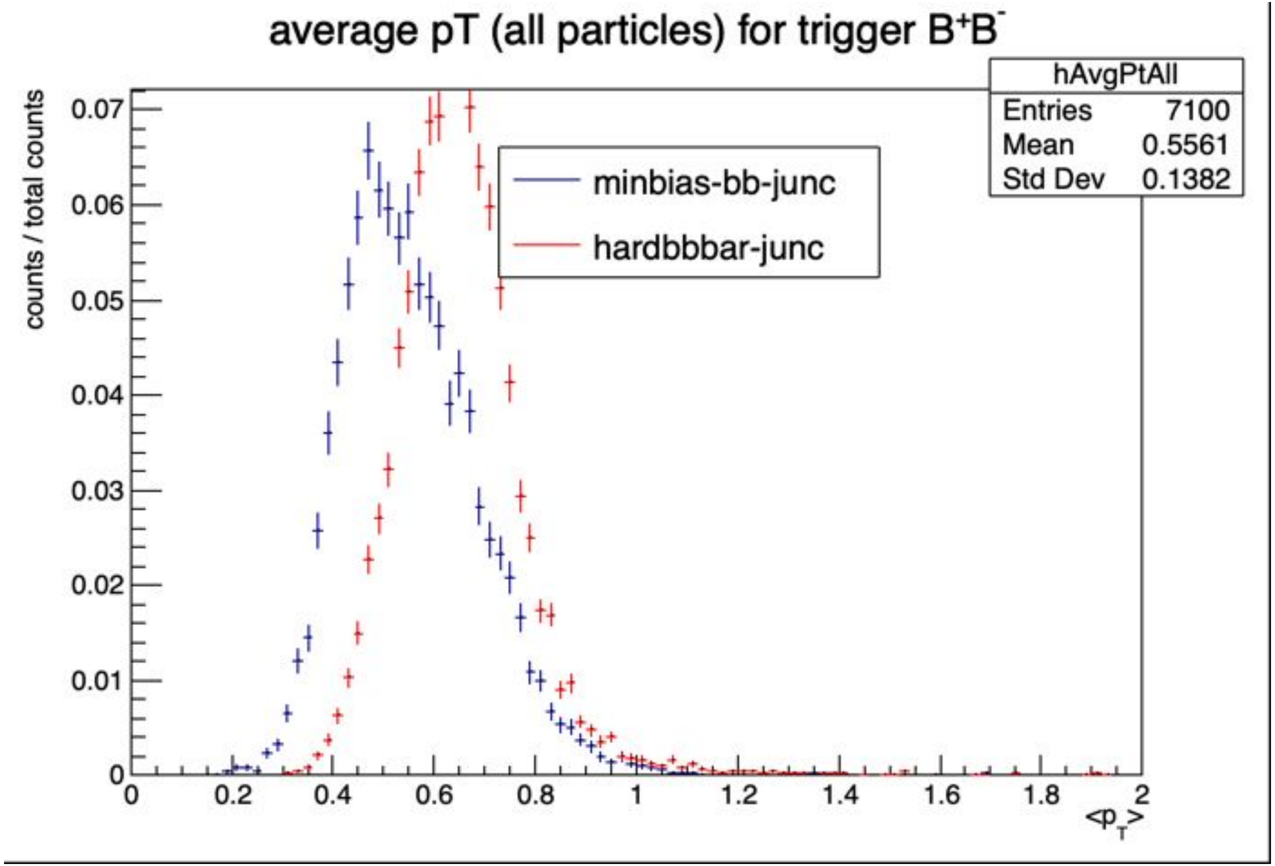
Charm Junctions: <pT> (soft)



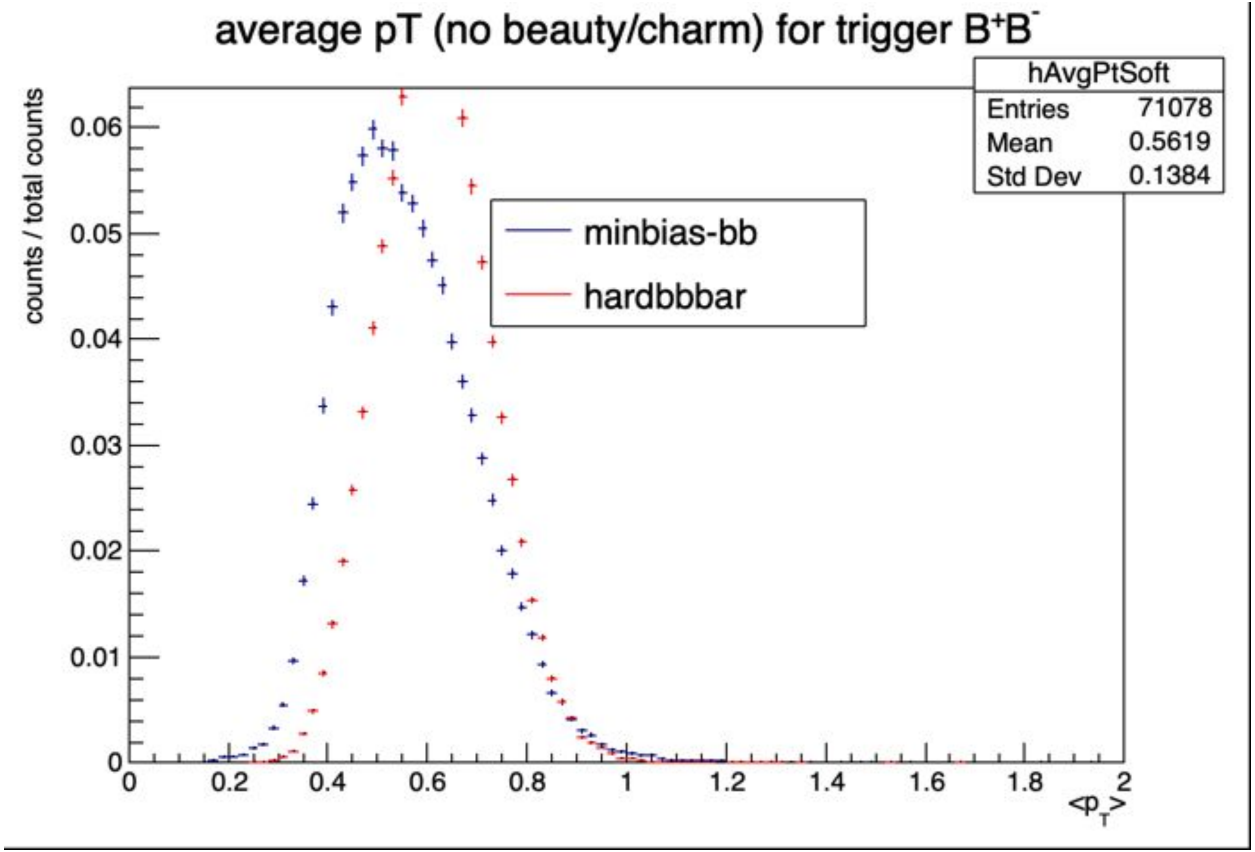
Beauty Monash: <pT> (all)



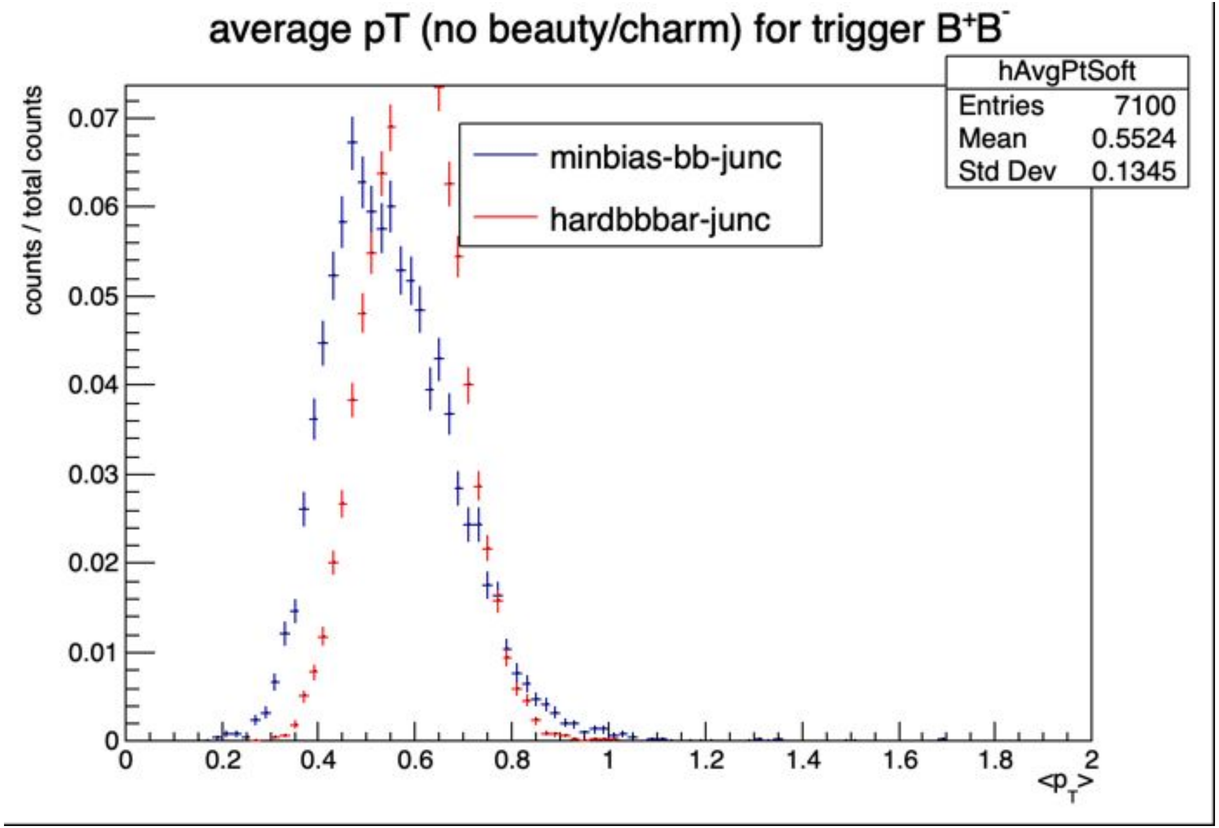
Beauty Junctions: <pT> (all)



Beauty Monash: <pT> (soft)



Beauty Junctions: <pT> (soft)



Conclusion: <pT>

Almost no differences between hard and soft? (not necessarily expected?)

Charm: average pT seems to be lower for hardccbar than MB

Beauty: the opposite

Naively I would expect the average pT to be higher for the boosted simulations, although this effect would be more noticeable between hard and soft? But there is no differences between hard and soft.. Need to think more about it.

Variable: sphericity

Sphericity basically describes how ‘spherical’ the event is

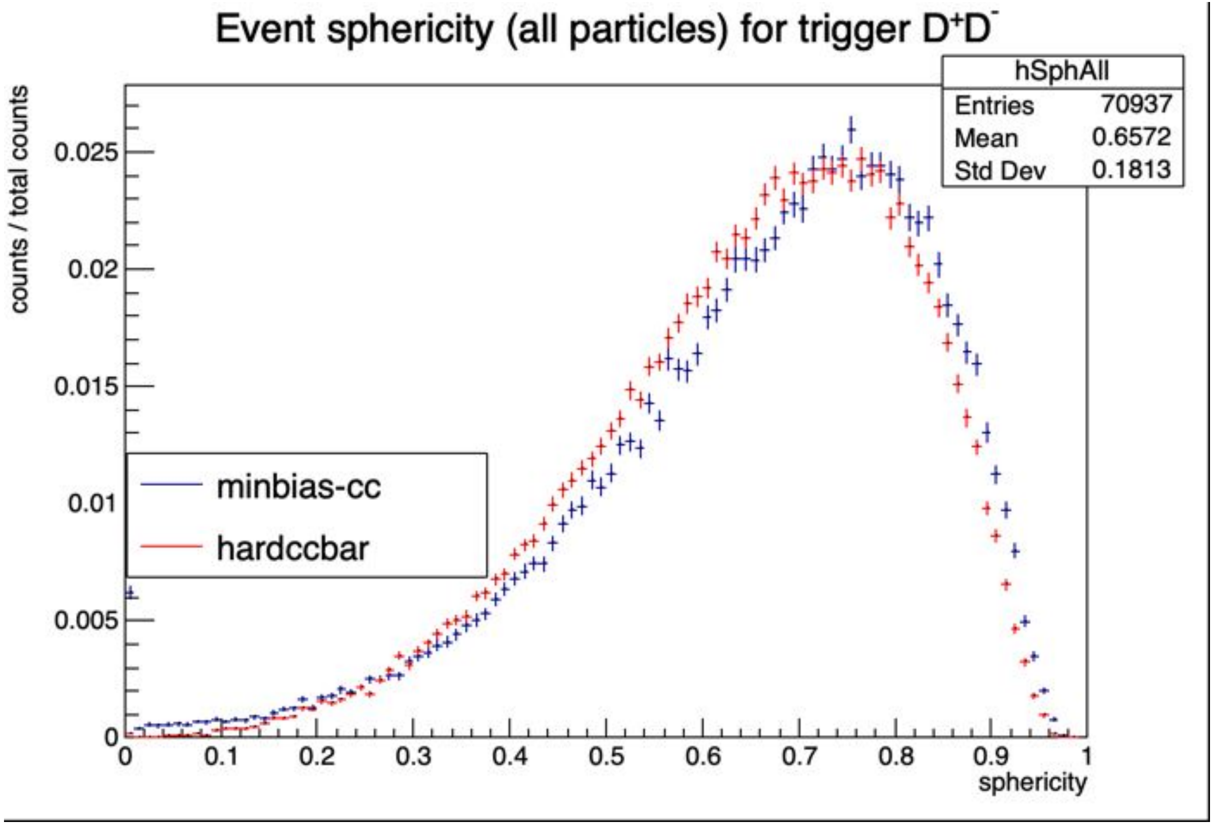
- A value close to 0 is less spherical, i.e. more like an e+e- event with a characteristic di-jet.
- A value close to 1 is more spherical, like the typical hadron collision events.

I wanted to see if, like it is used in the paper, the sphericity can be used to check if events have jets, or are more jet-like.

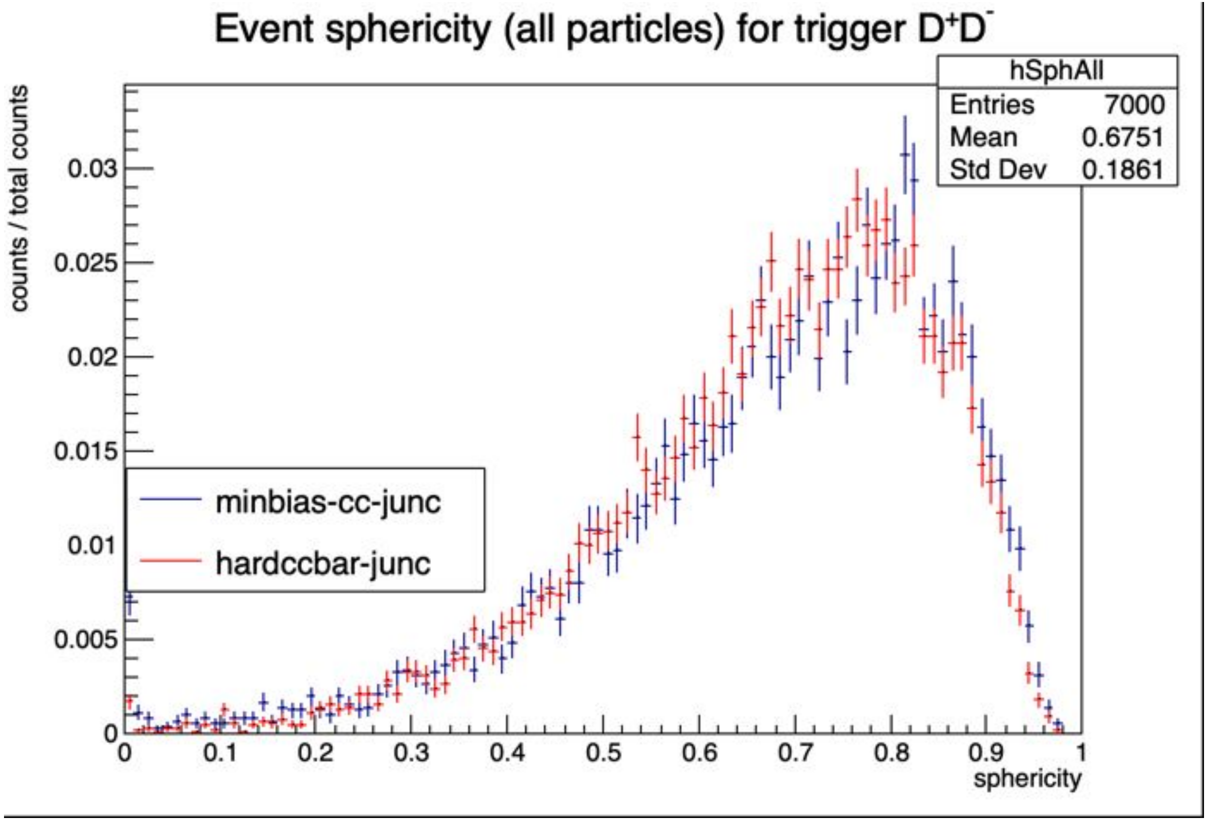
(i) Transverse spherocity: this quantity allows one to know whether a dijetlike structure is present in the event [33]. It is defined for a unit vector $\hat{\mathbf{n}}_s$ which minimizes the ratio:

$$S_0 \equiv \frac{\pi^2}{4} \min_{\hat{\mathbf{n}}_s} \left(\frac{\sum_i |\vec{p}_{T,i} \times \hat{\mathbf{n}}_s|}{\sum_i p_{T,i}} \right)^2, \qquad (1)$$

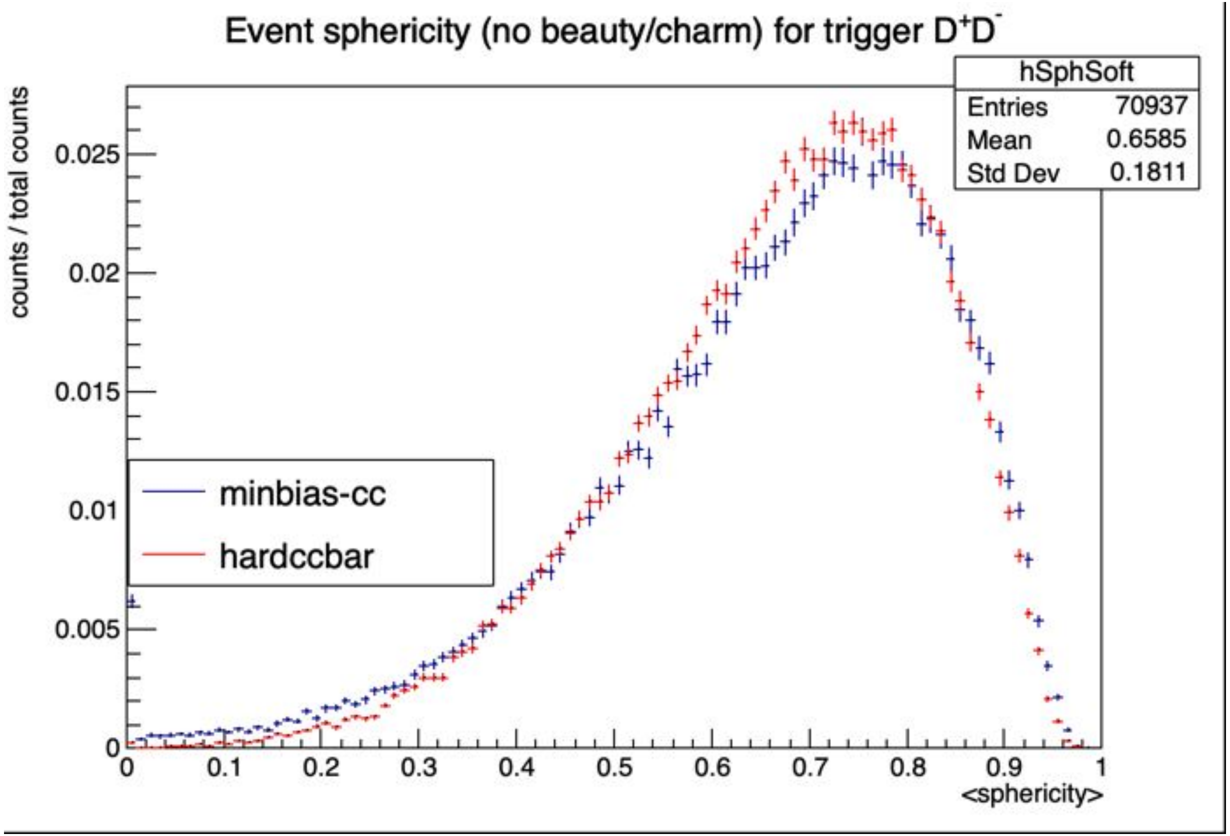
Charm Monash: sphericity (all)



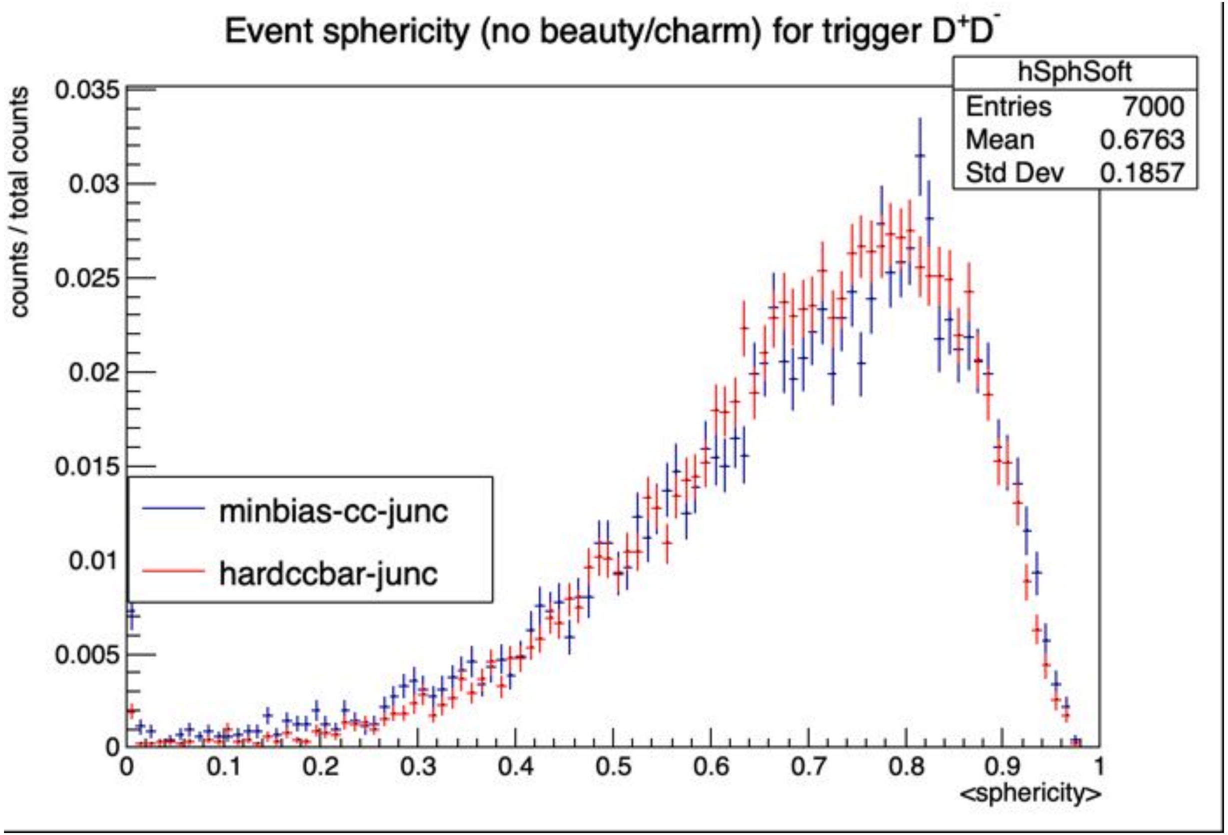
Charm Junctions: sphericity (all)



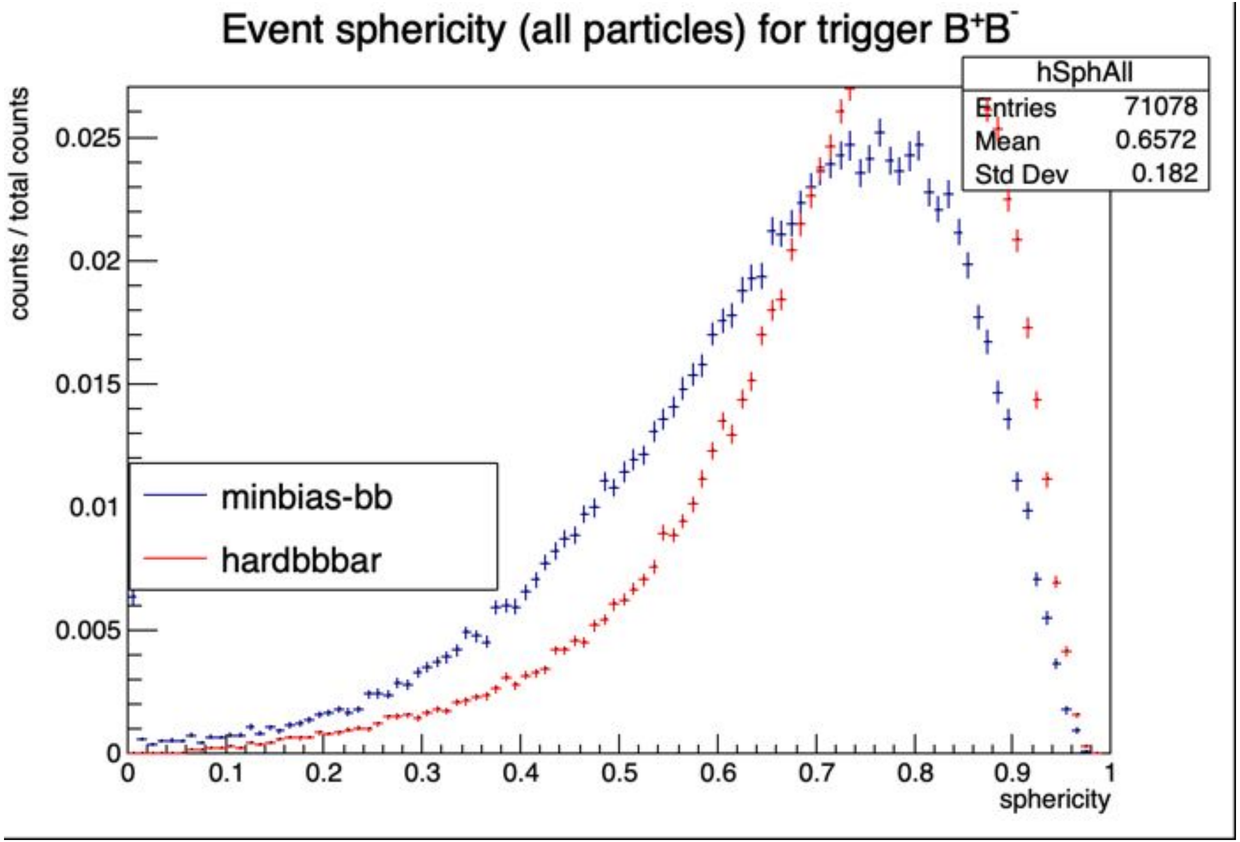
Charm Monash: sphericity (soft)



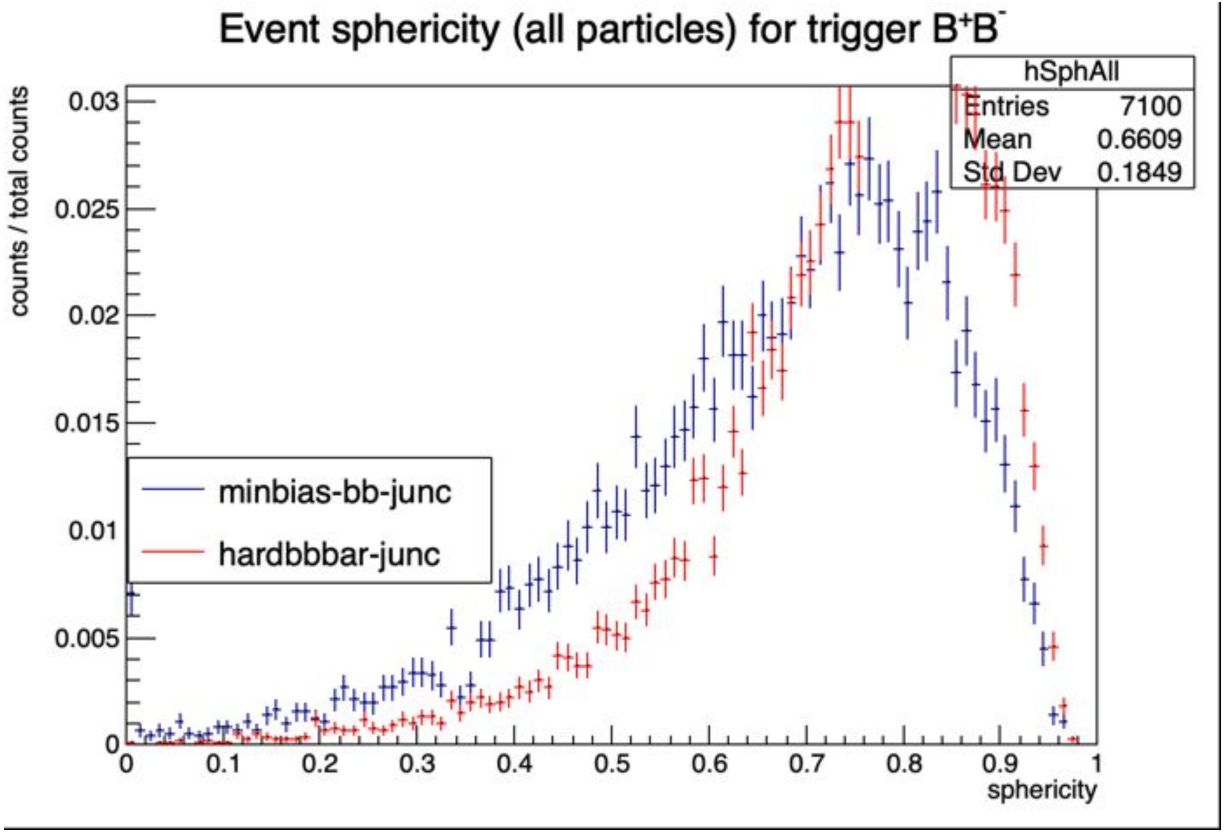
Charm Junctions: sphericity (soft)



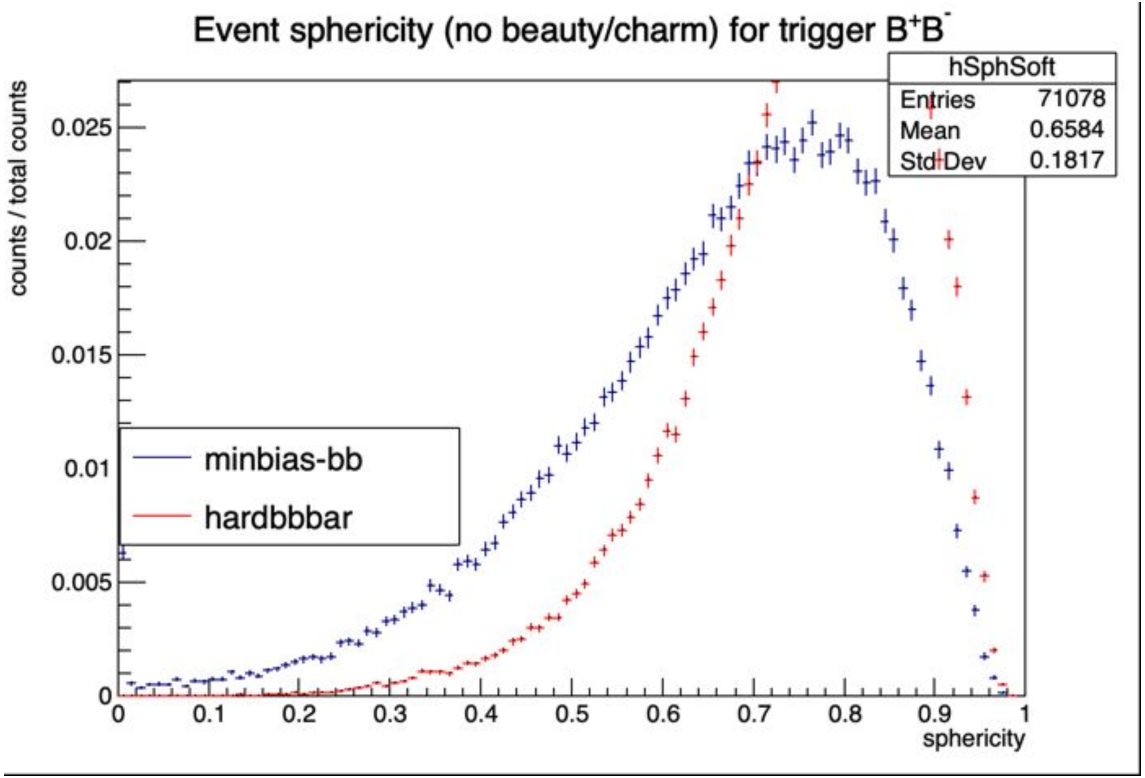
Beauty Monash: sphericity (all)



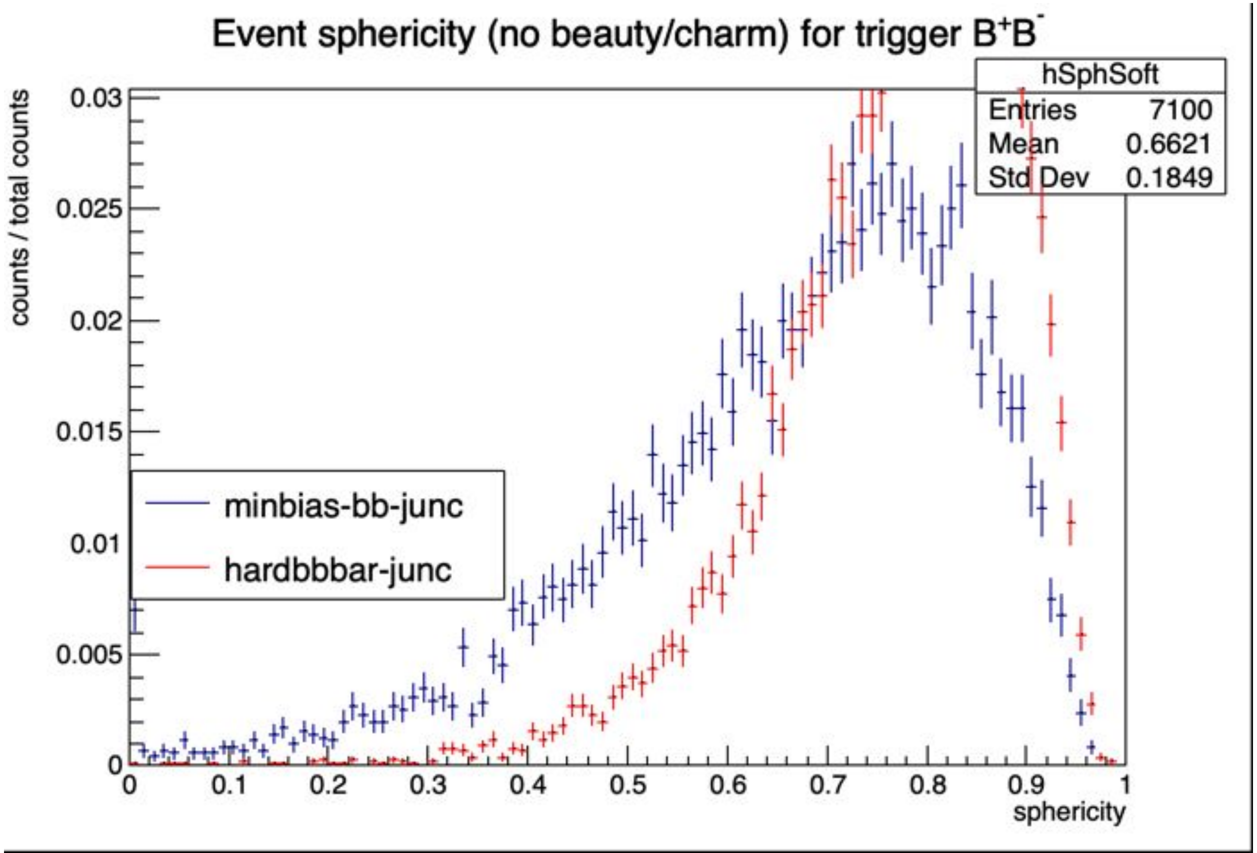
Beauty Junctions: sphericity (all)



Beauty Monash: sphericity (soft)



Beauty Junctions: sphericity (soft)



Conclusion: sphericity

Soft and hard very similar.

In charm only marginal differences between all categories. Needs more statistics.

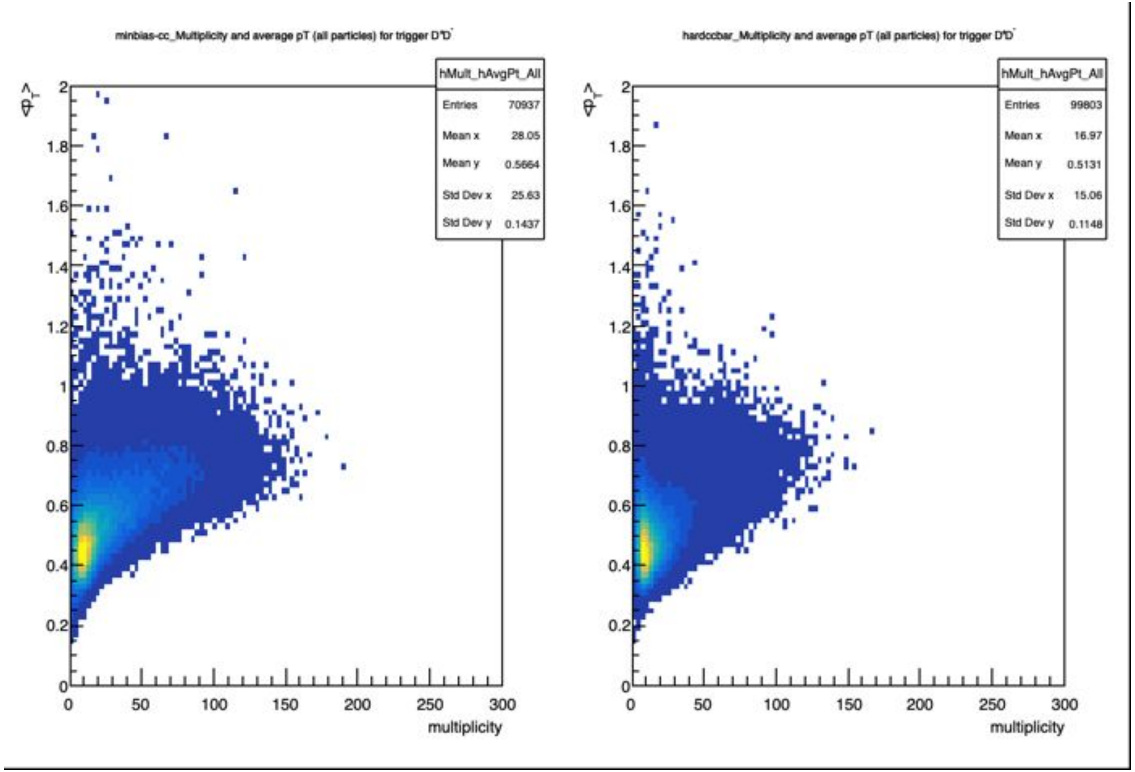
For beauty, something interesting: it seems like the hardbbbar events are a bit more spherical, i.e. less jet-like, which was my hypothesis from the number of jets earlier.

Two-dimensional histograms

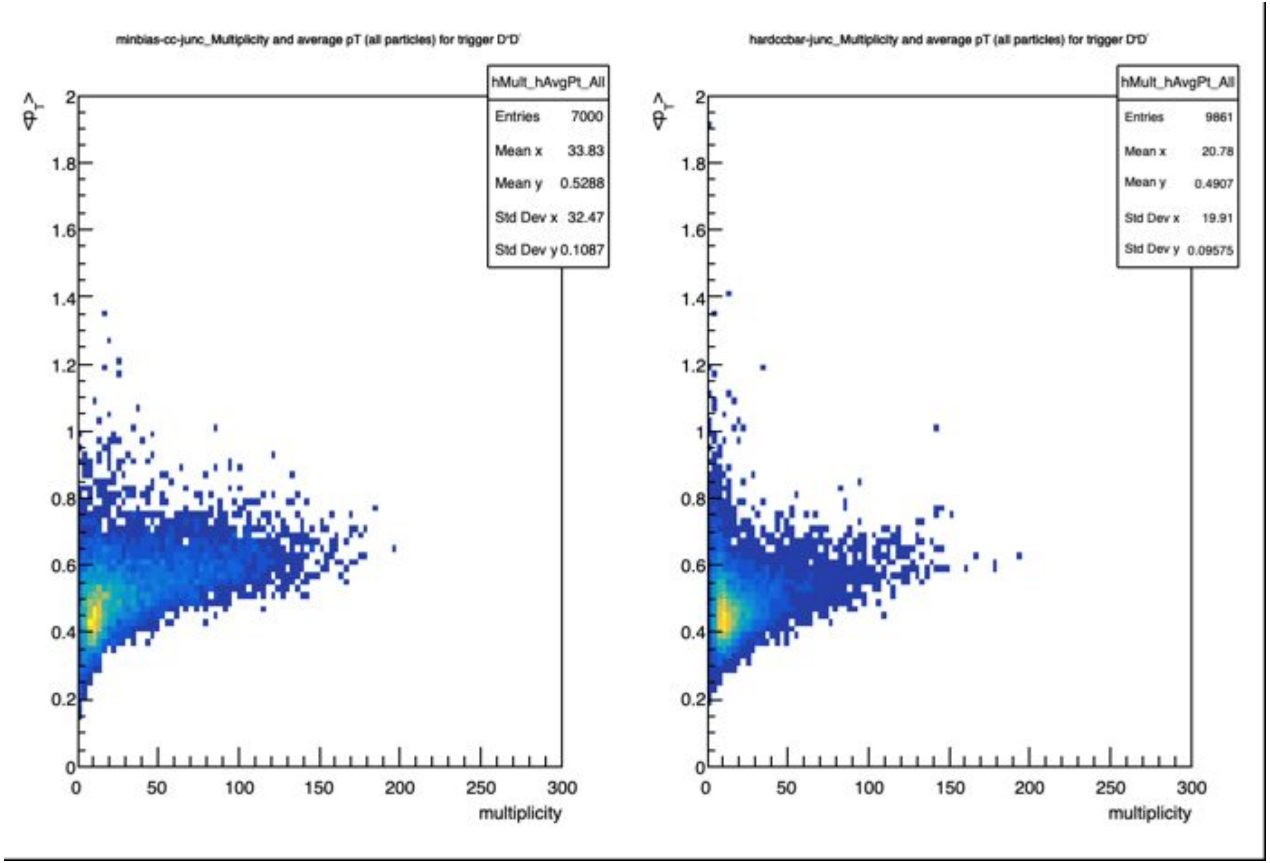
Always

- left: minbias
- right: hardccbar/bbbar

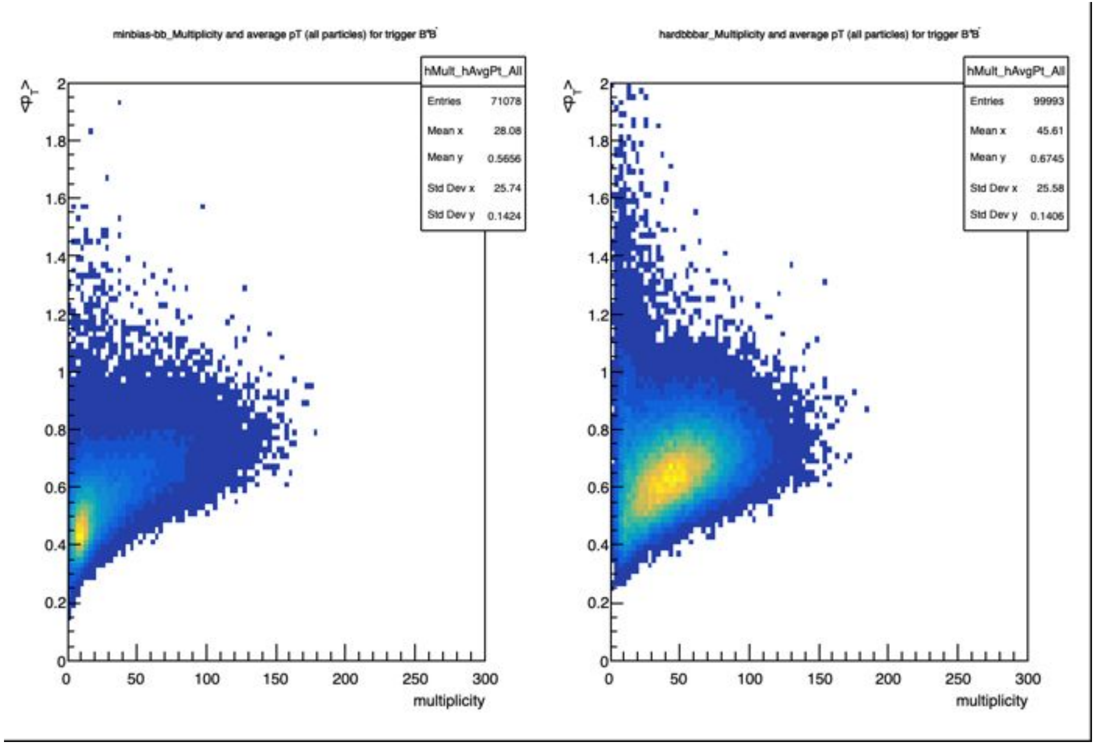
Charm Monash: multiplicity vs <pT> (all)



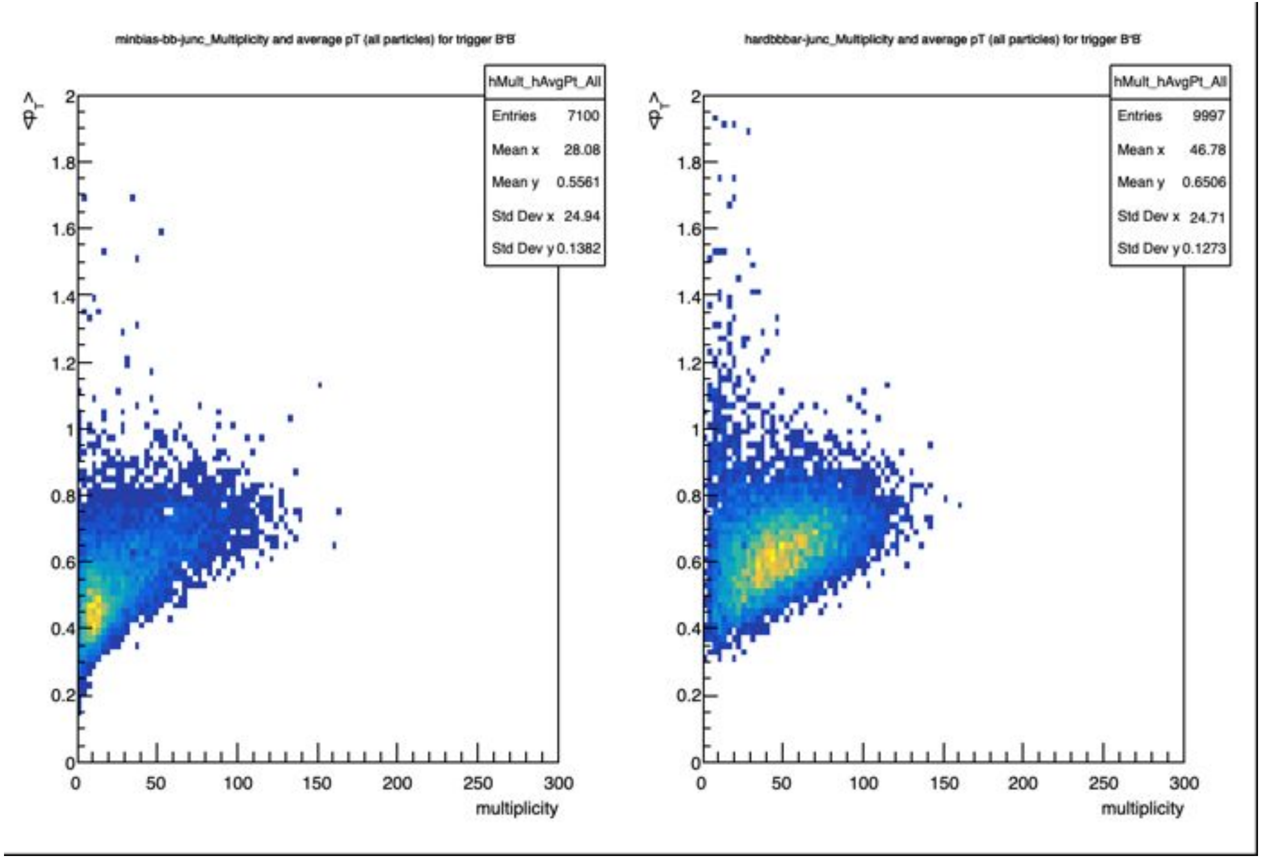
Charm Junctions: multiplicity vs <pT> (all)



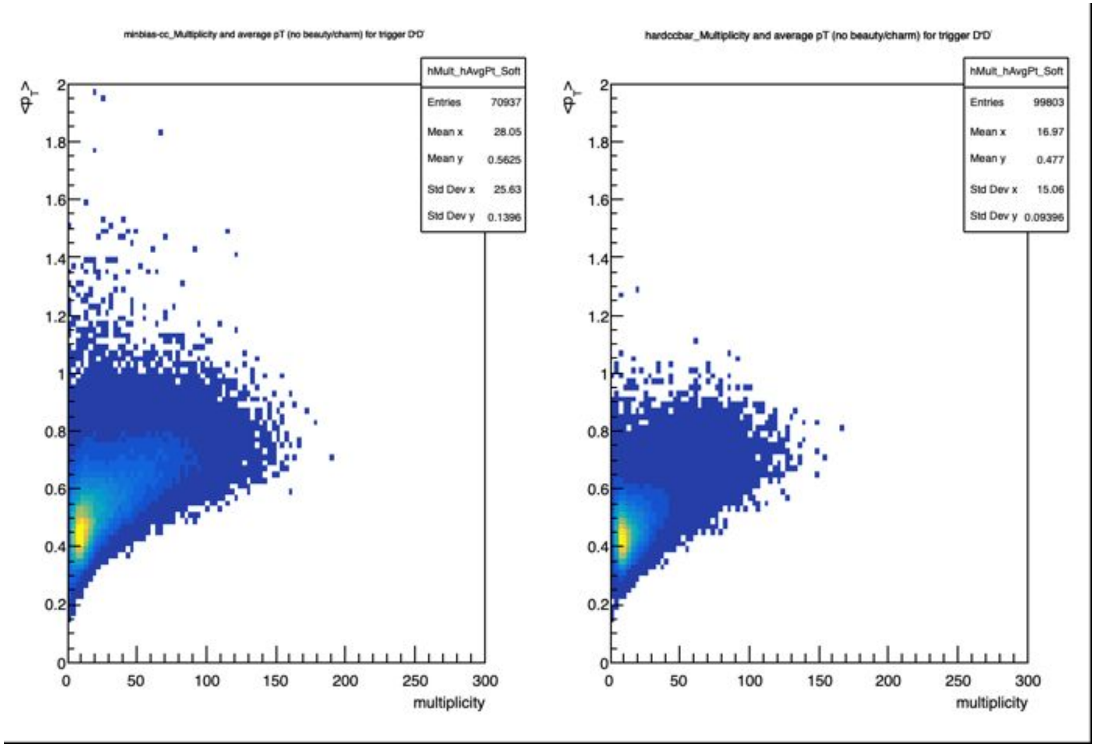
Beauty Monash: multiplicity vs <pT> (all)



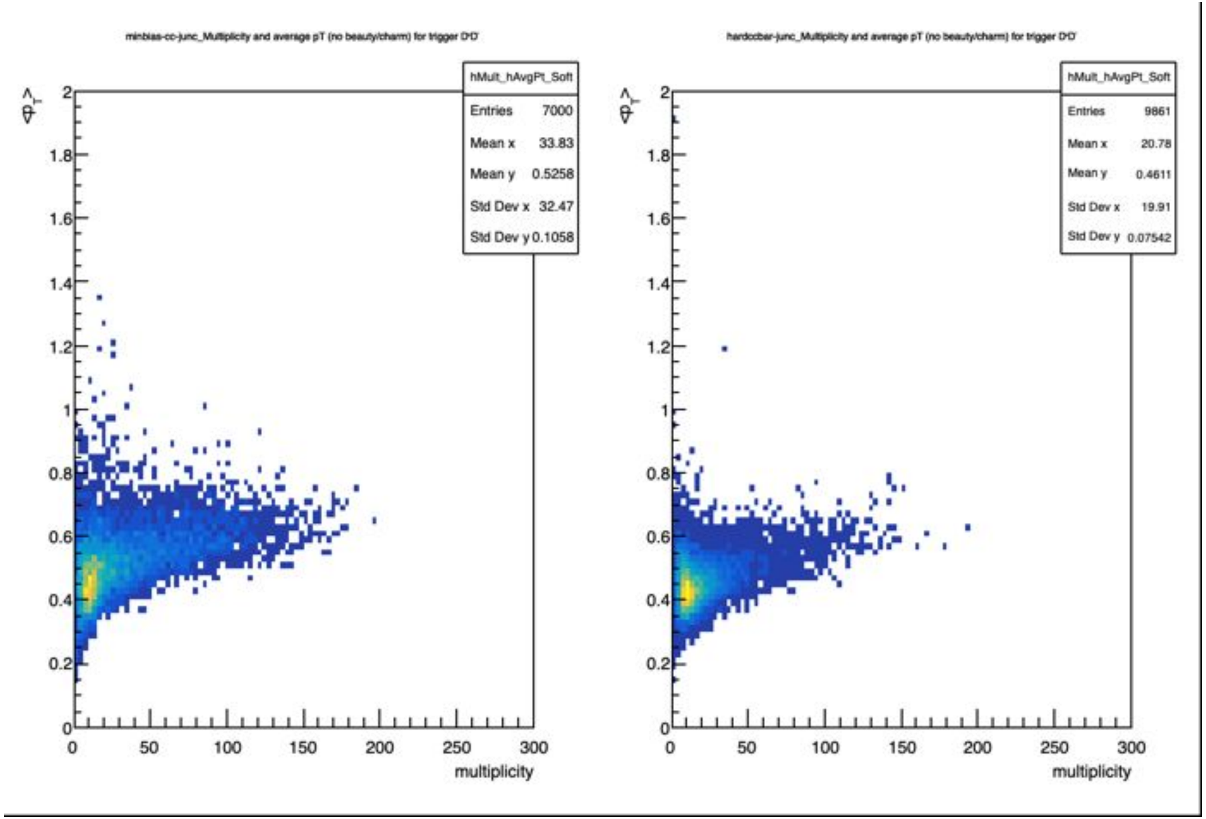
Beauty Junctions: multiplicity vs <pT> (all)



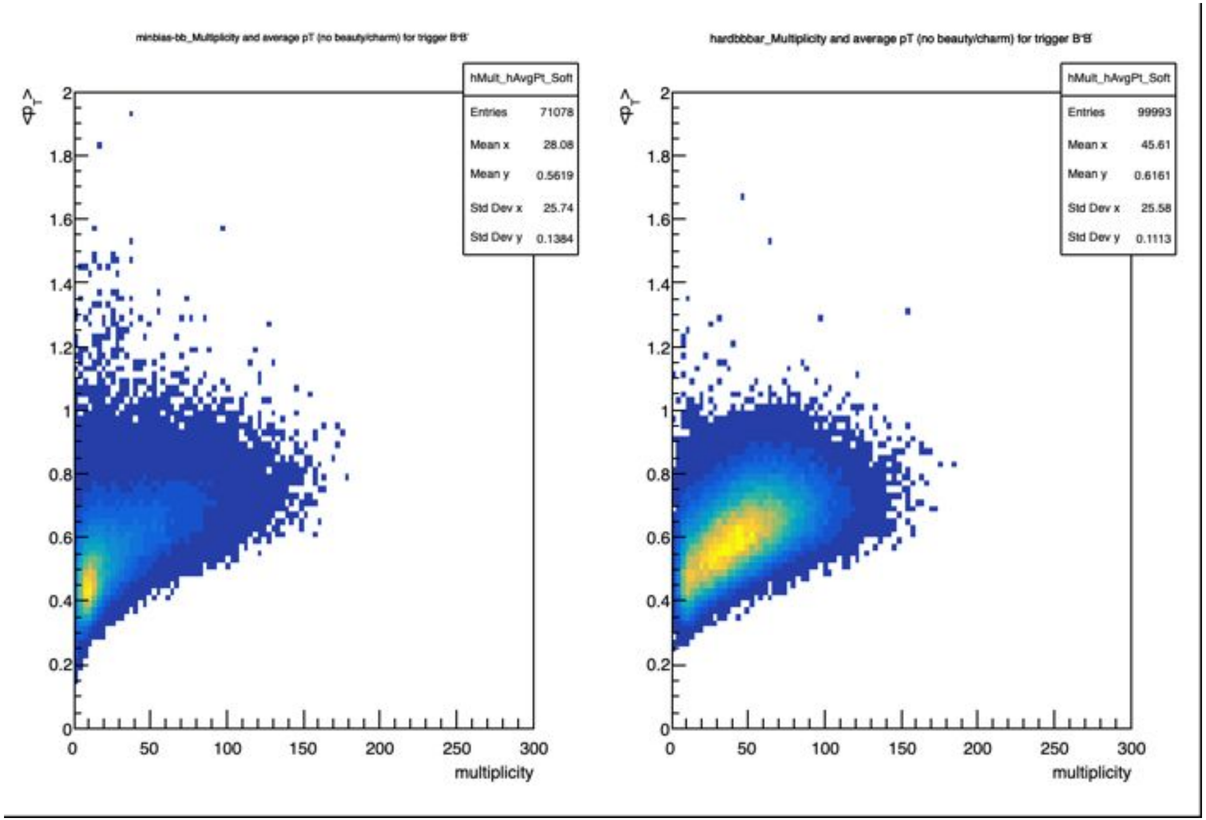
Charm Monash: multiplicity vs <pT> (soft)



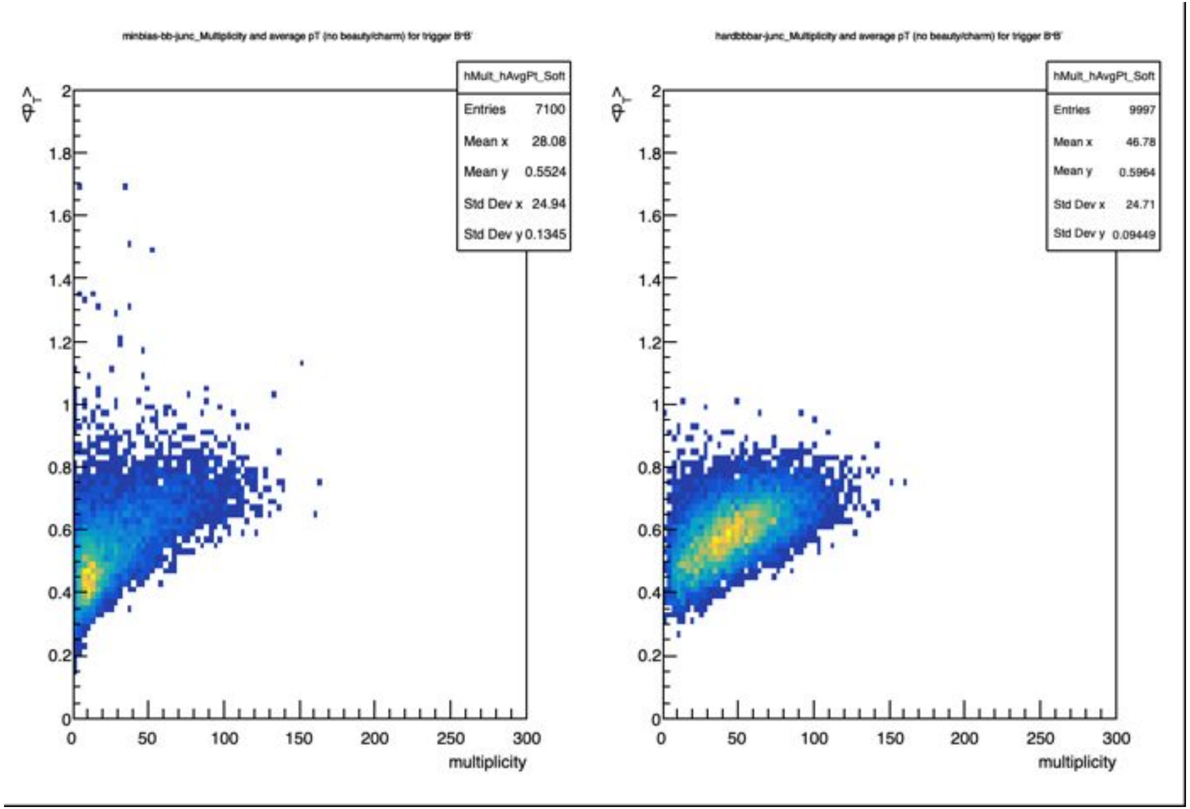
Charm Junctions: multiplicity vs <pT> (soft)



Beauty Monash: multiplicity vs <pT> (soft)



Beauty Junctions: multiplicity vs <pT> (soft)



Conclusion: multiplicity vs <pT>

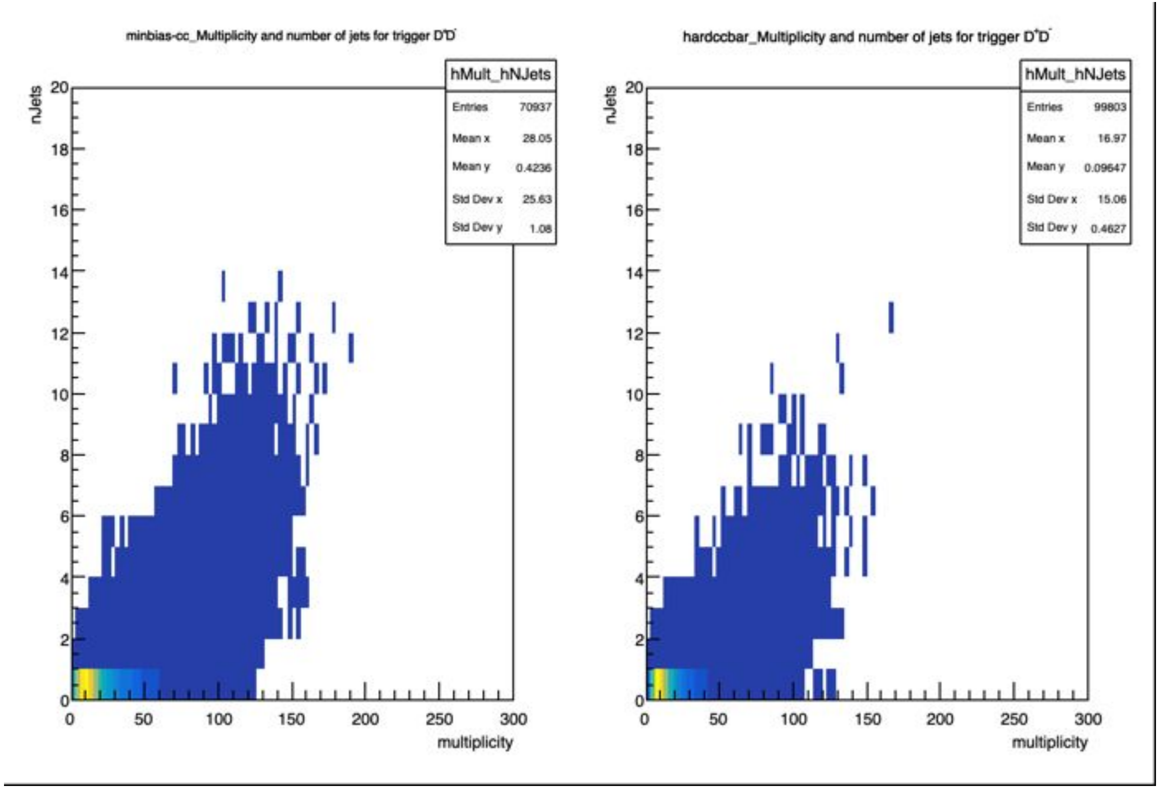
In junctions the trends seem to be sharper than in monash (can also be due to difference in statistics?)

In general, <pT> shows increasing trend with increasing multiplicity but saturates rather quickly too. Exactly what I expected and similar to what is seen in the balancing studies.

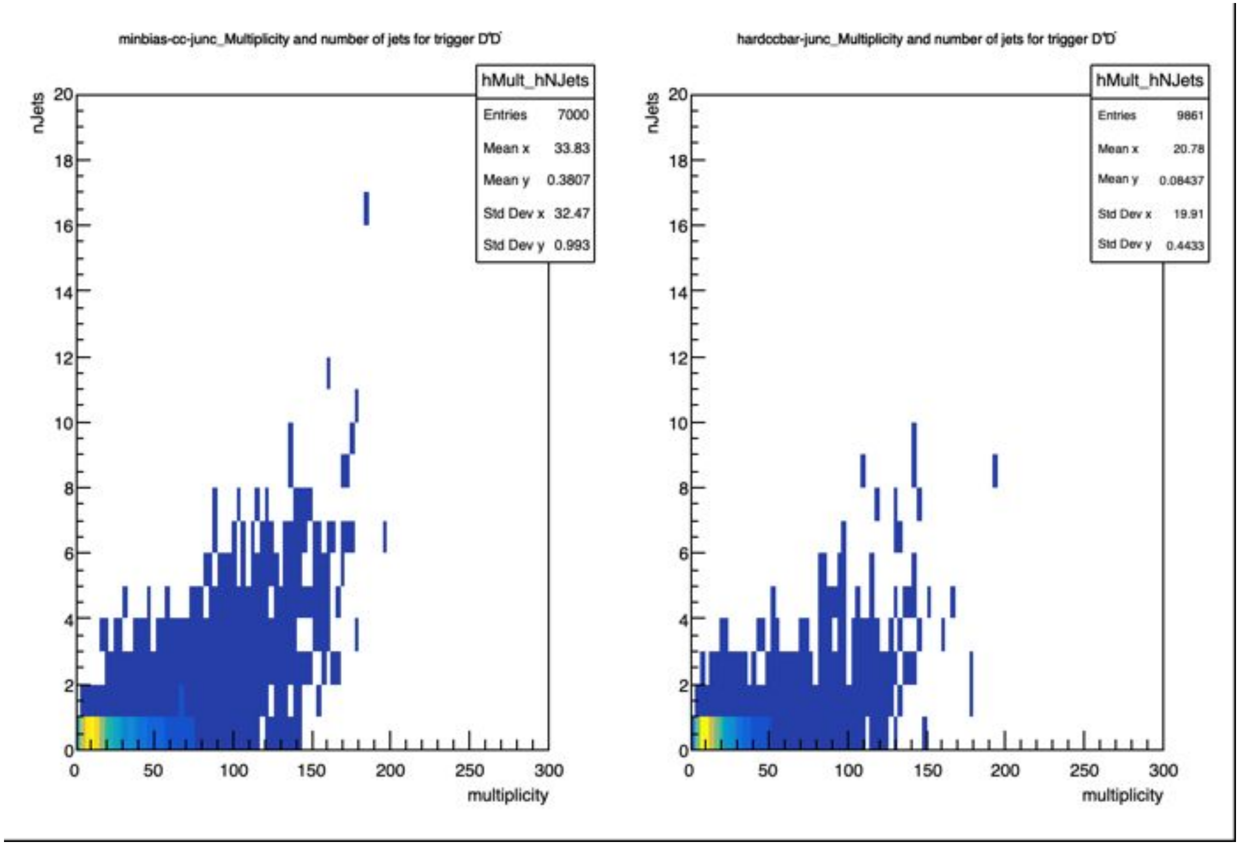
It is not so clear to my what conclusion to draw from the MB and boosted simulations. Needs more statistics. Though it seems like the trends saturate a bit faster in MB, as I would expect.

Soft and all have only minor effects due to ‘ignoring’ particles

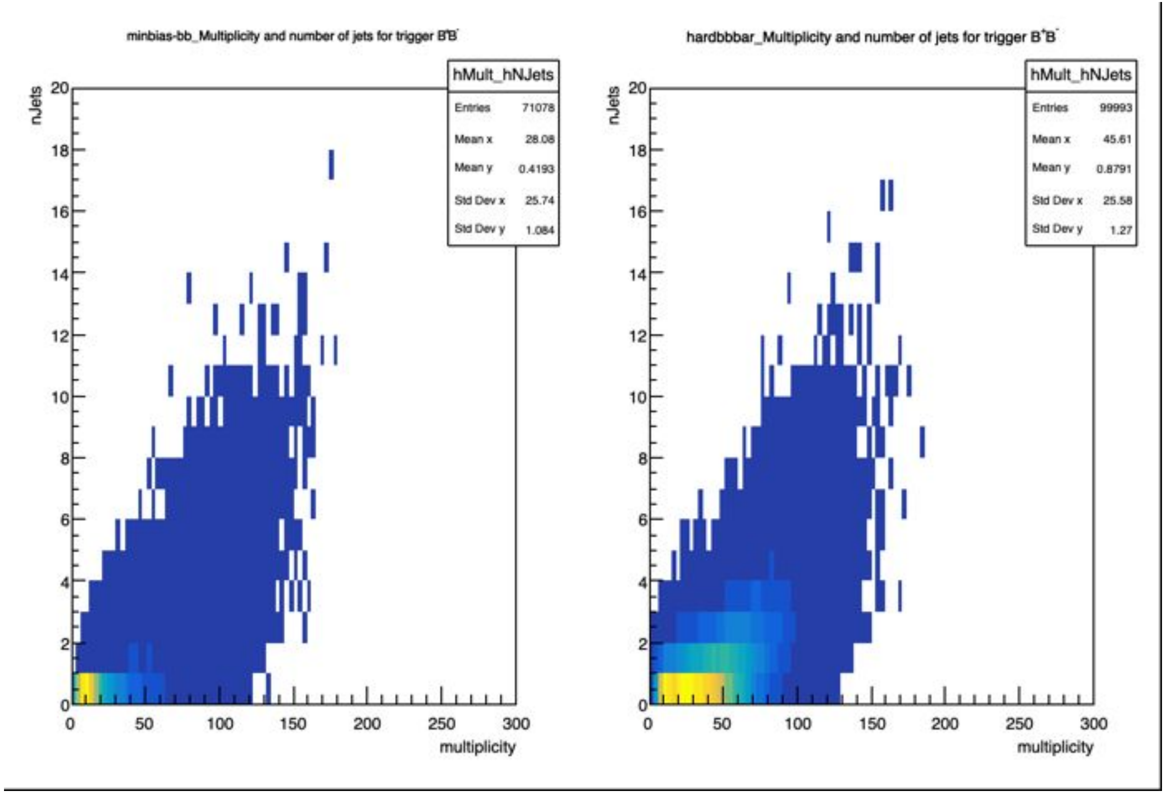
Charm Monash: multiplicity vs number of jets



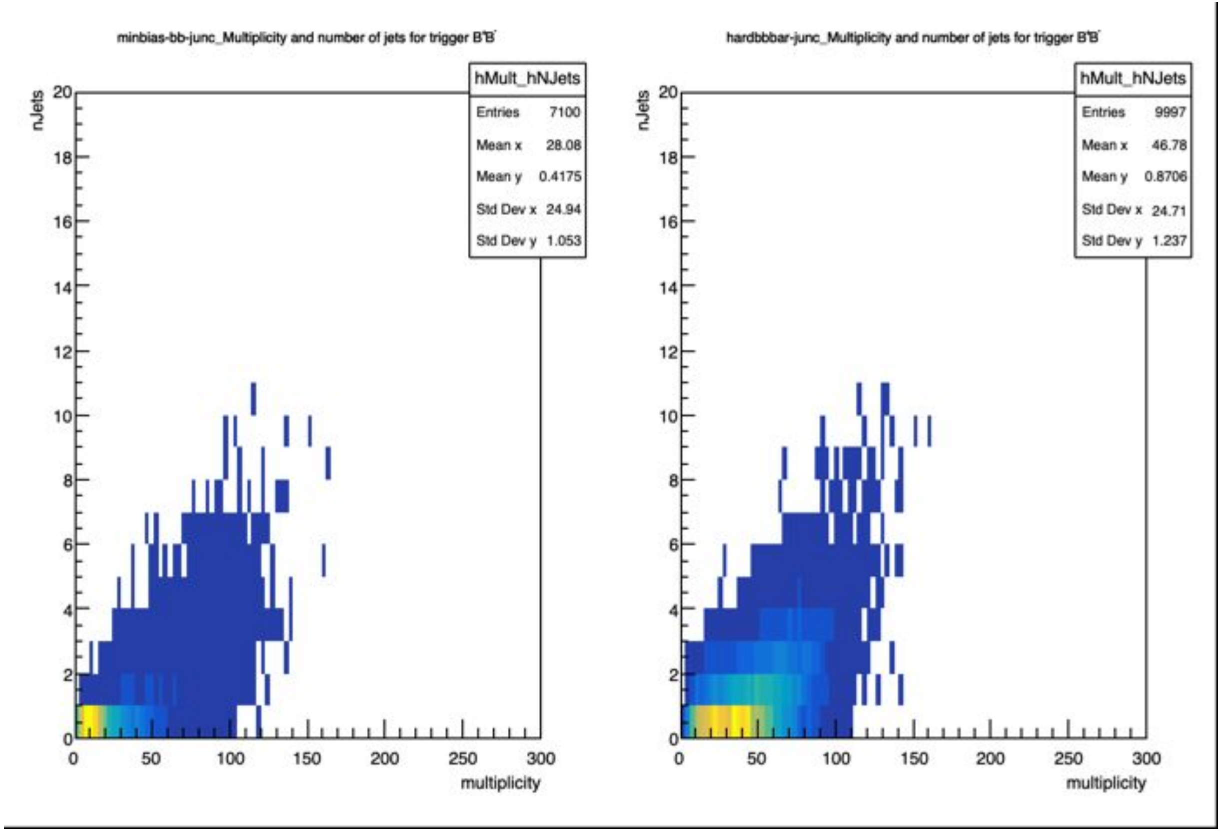
Charm Junctions: multiplicity vs number of jets



Beauty Monash: multiplicity vs number of jets



Beauty Junctions: multiplicity vs number of jets



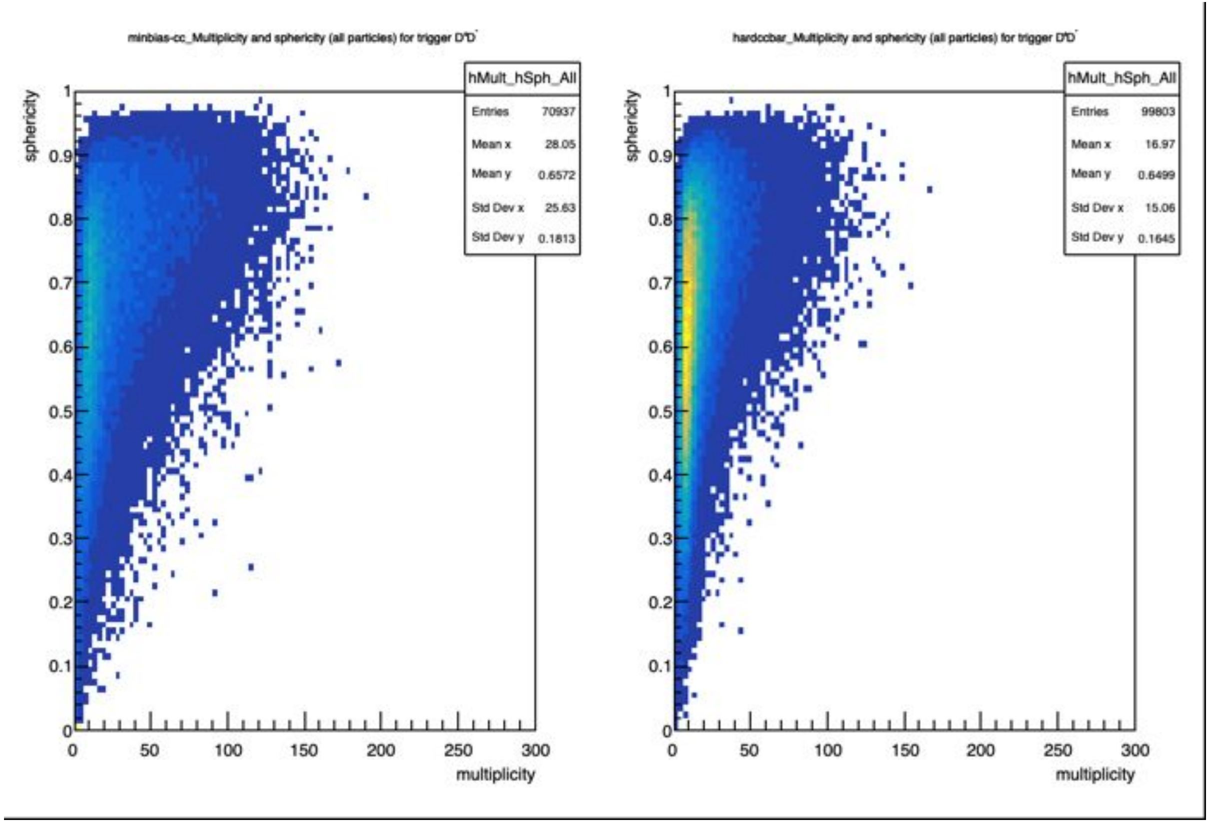
Conclusion: multiplicity vs number of jets

To me it seems like in all categories, there are no big differences.

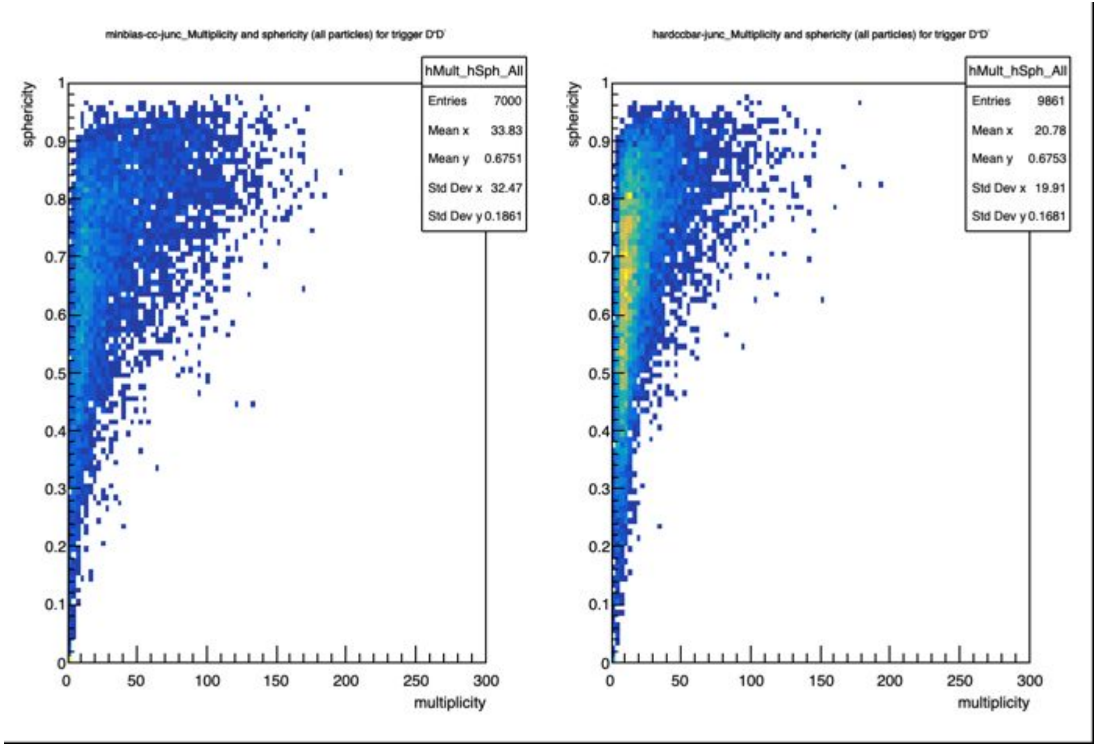
The main trend is that the number of jets increases (sharply) with increasing multiplicity.

Maybe it’s statistics, but it could be that monash has on average more jets than junctions.

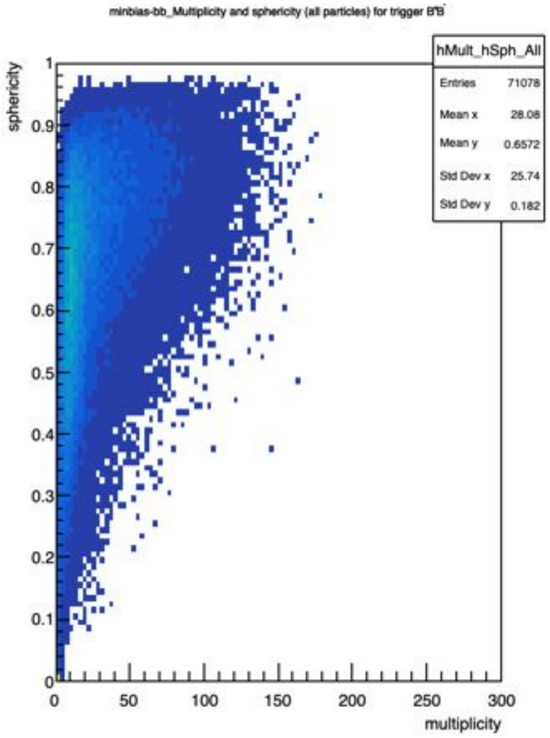
Charm Monash: multiplicity vs sphericity (all)

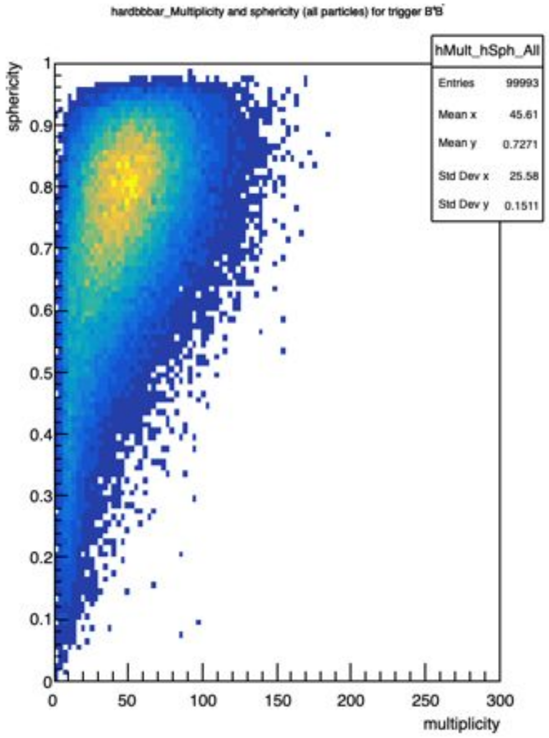


Charm Junctions: multiplicity vs sphericity (all)

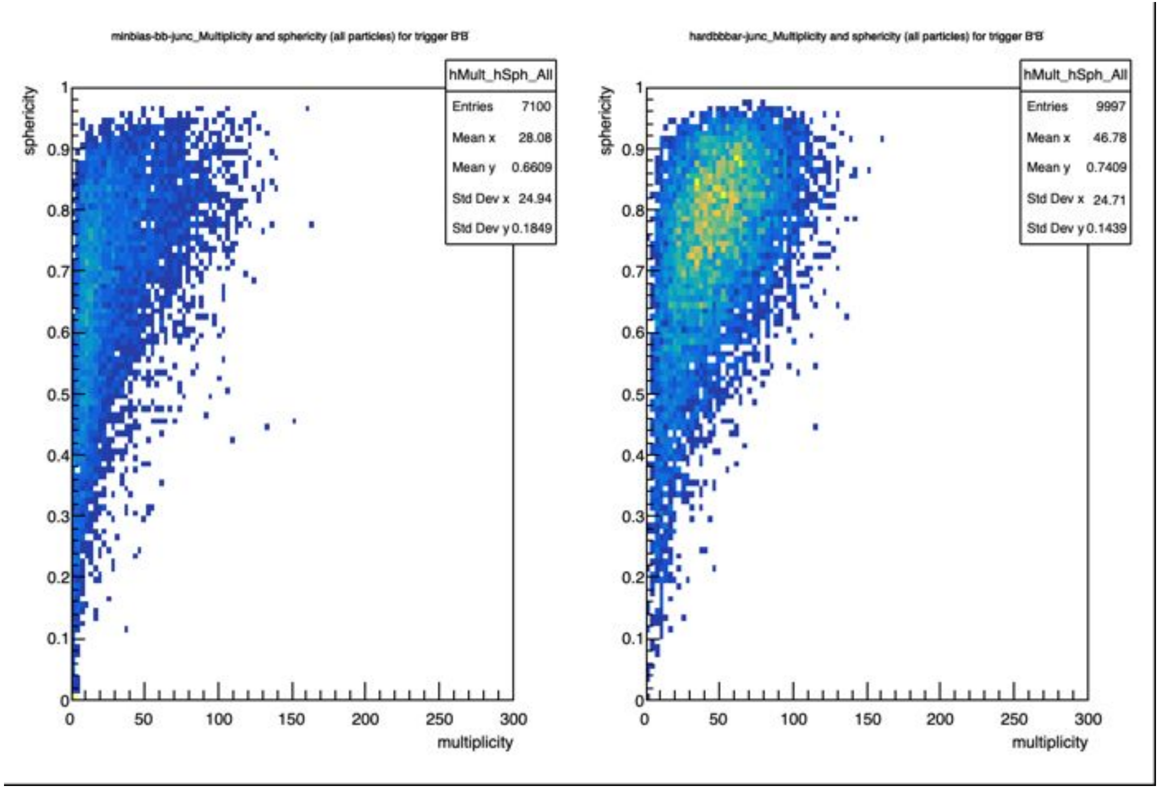


Beauty Monash: multiplicity vs sphericity (all)

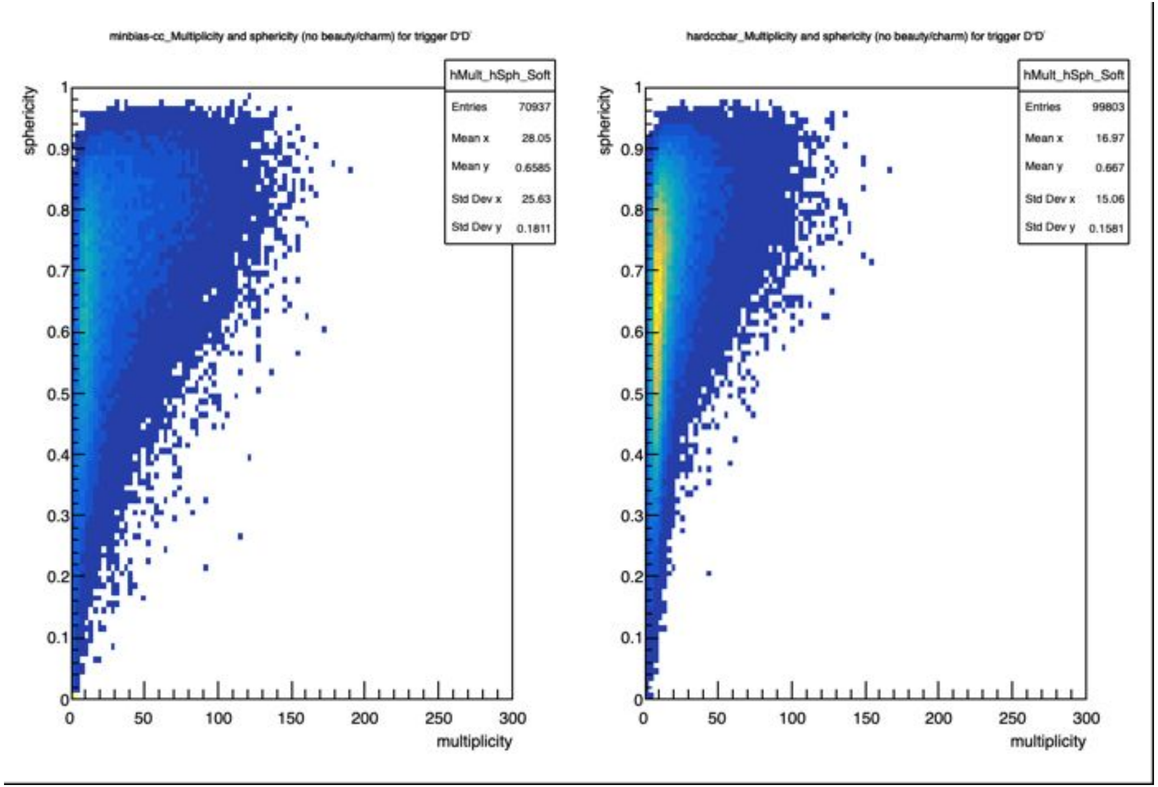




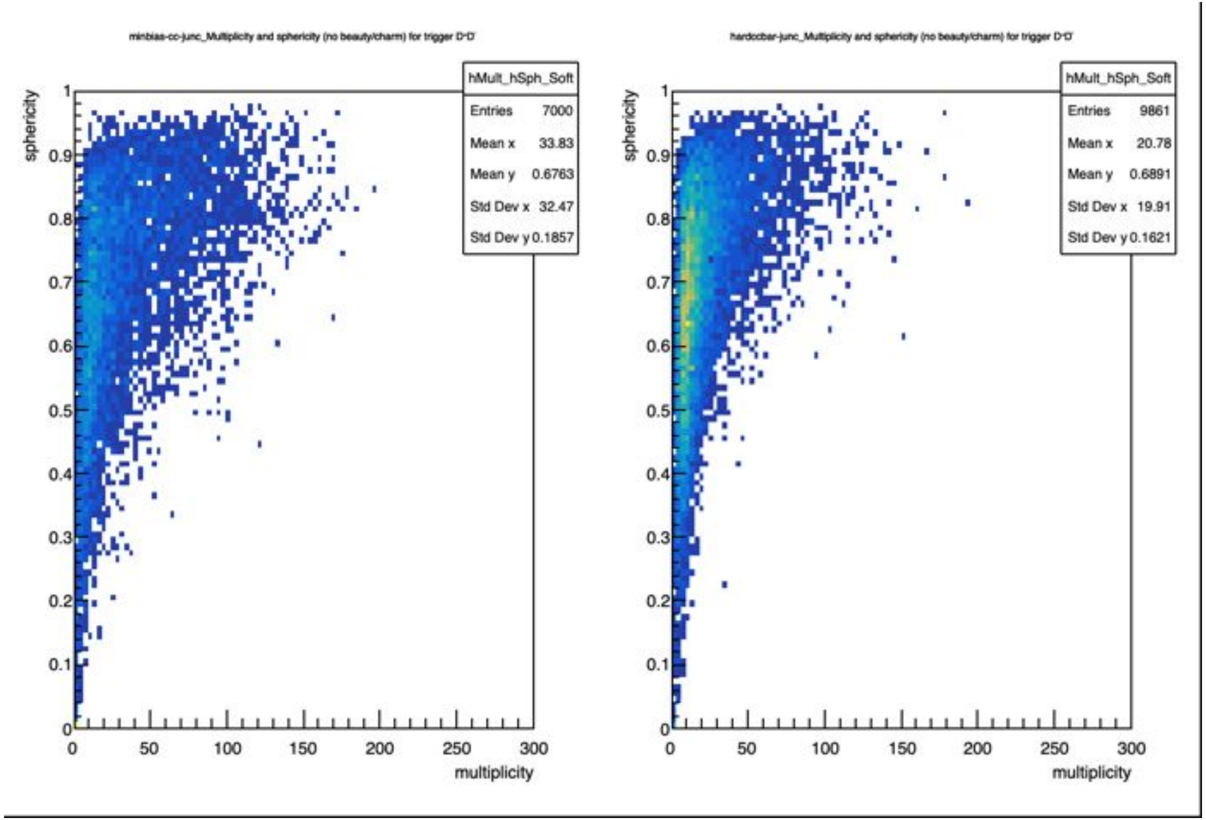
Beauty Junctions: multiplicity vs sphericity (all)



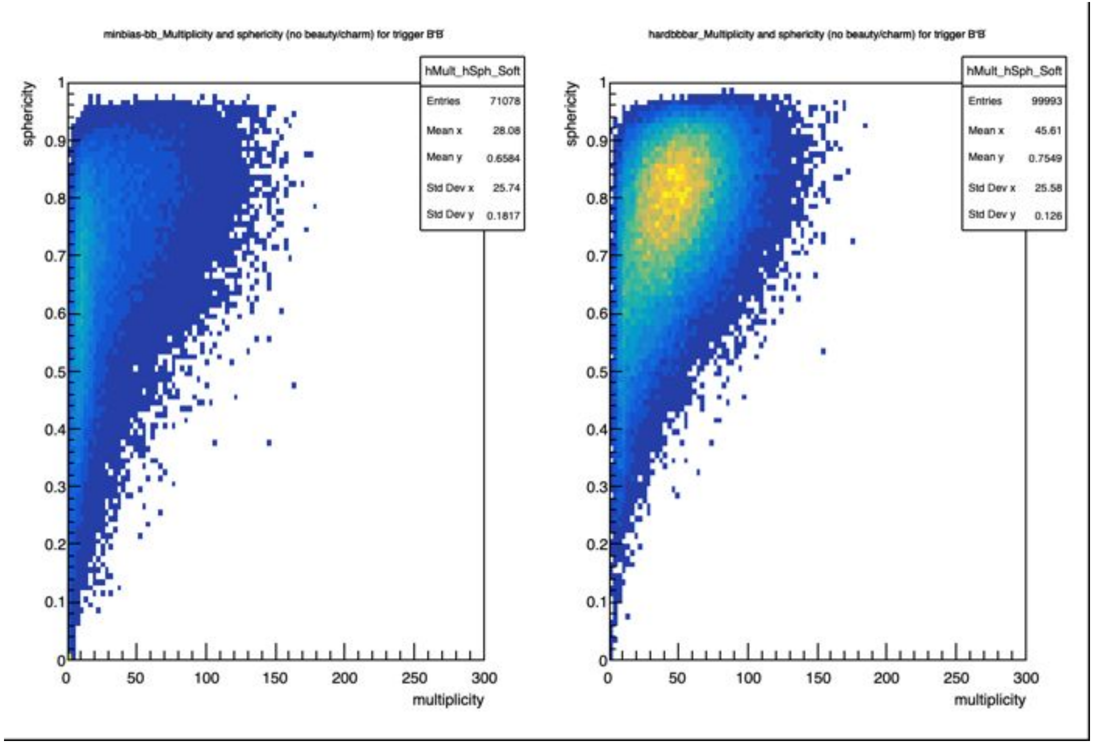
Charm Monash: multiplicity vs sphericity (soft)



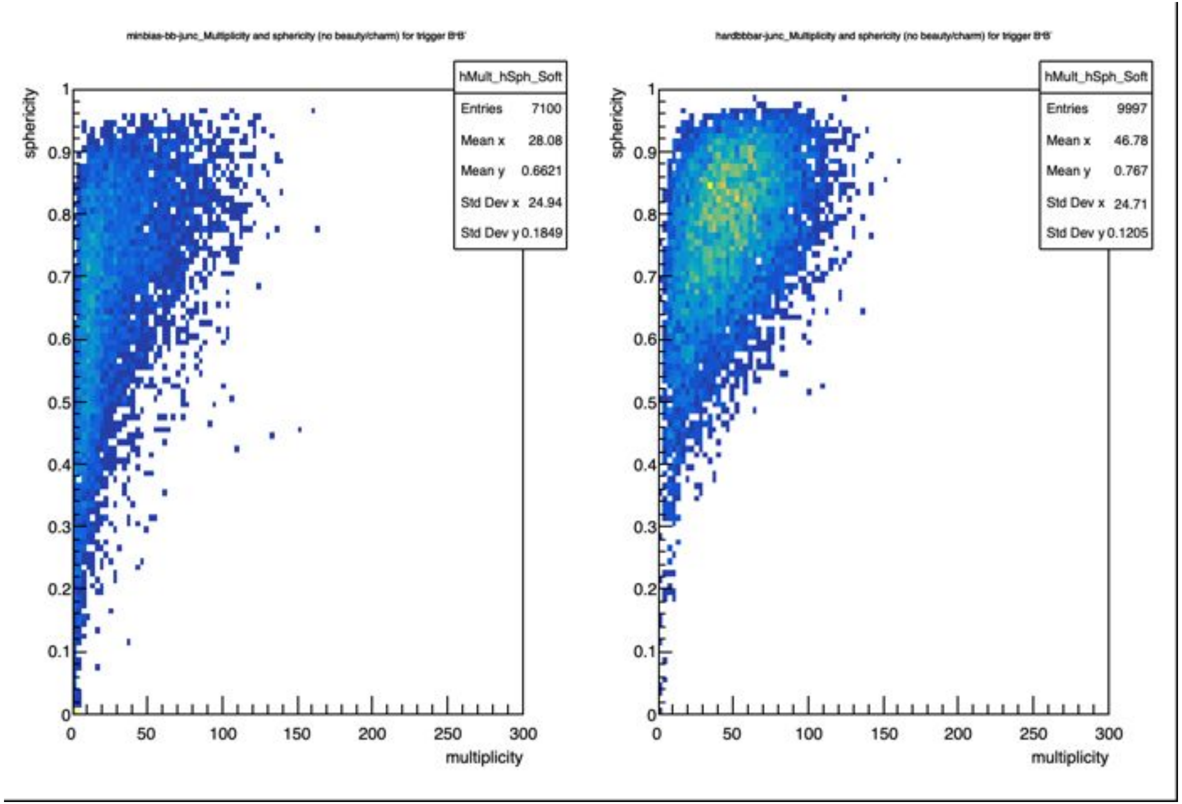
Charm Junctions: multiplicity vs sphericity (soft)



Beauty Monash: multiplicity vs sphericity (soft)



Beauty Junctions: multiplicity vs sphericity (soft)



Conclusion: multiplicity vs sphericity

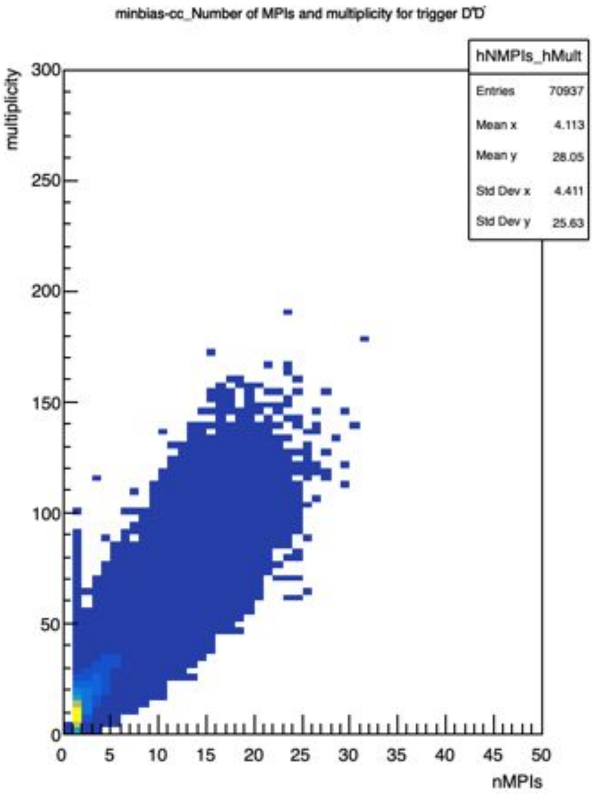
For me sphericity is the most difficult variable to interpret, which is why I included these correlation plots.

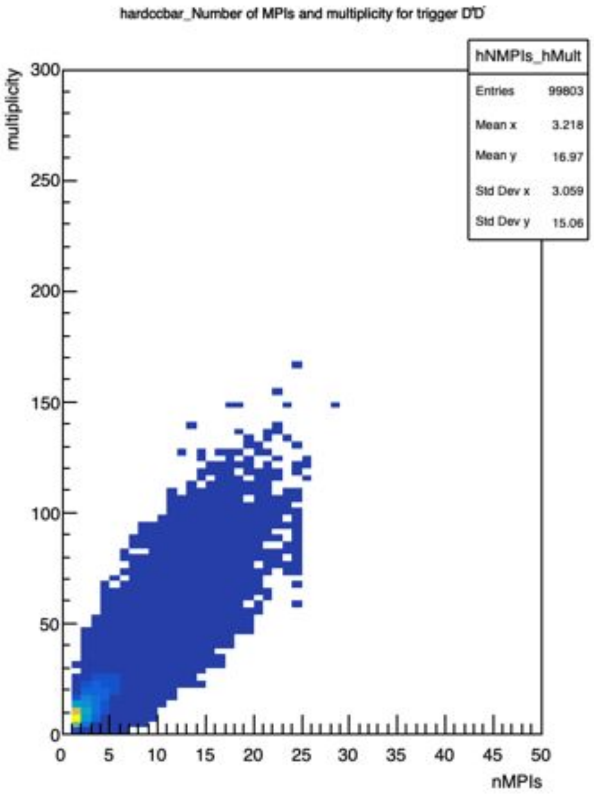
The clearest trend is that with higher multiplicity the sphericity gets closer to 1, which makes sense, as events will be more ‘busy’ and hence more ‘spherical’ (hadronic-like).

There don’t seem to be large differences in Monash and Junctions, but it seems like the beauty boosted simulations are much more centered at higher multiplicity (but the same sphericity) compared to charm. This is also what we saw in the multiplicity histograms.

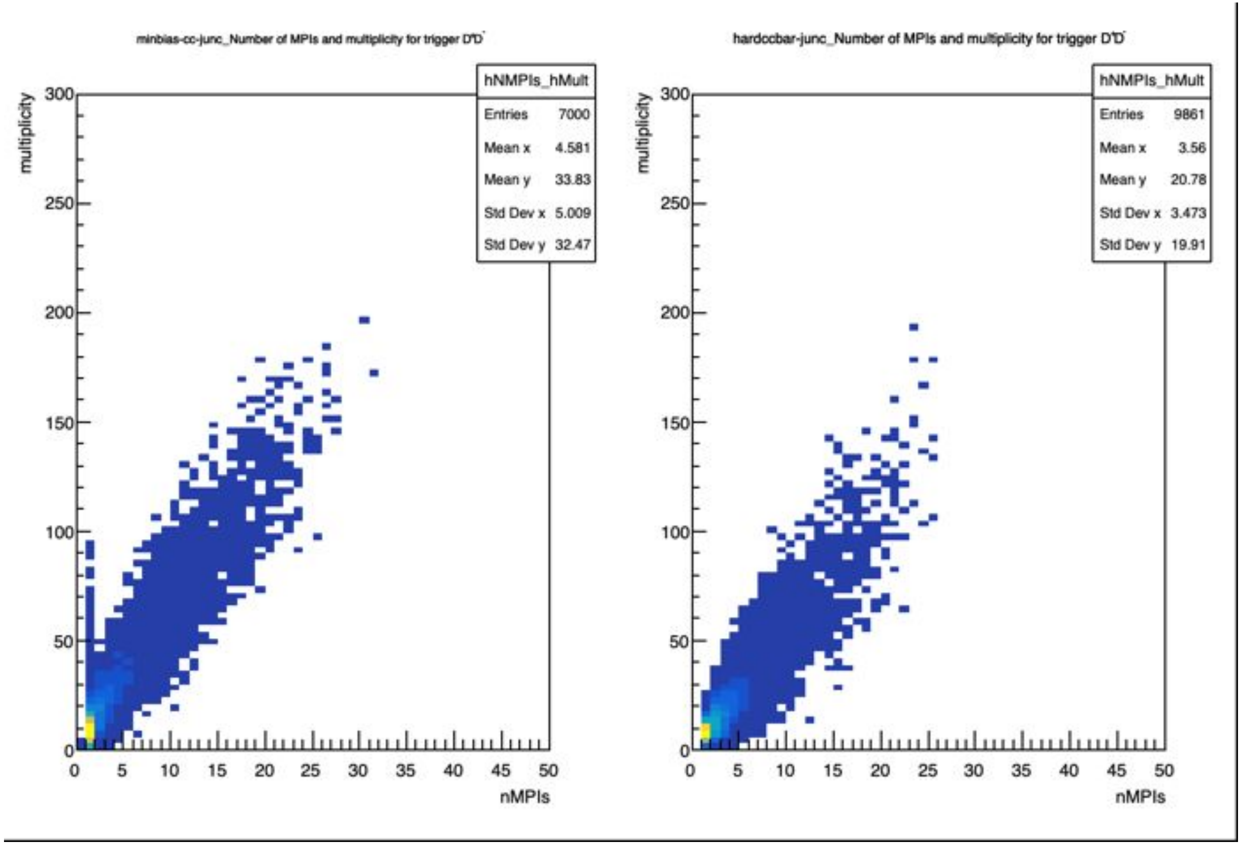
- The conclusion is the same: beauty boosted is quite different, while charm MB, charm boosted and beauty MB are similar.

Charm Monash: number of MPIs vs multiplicity

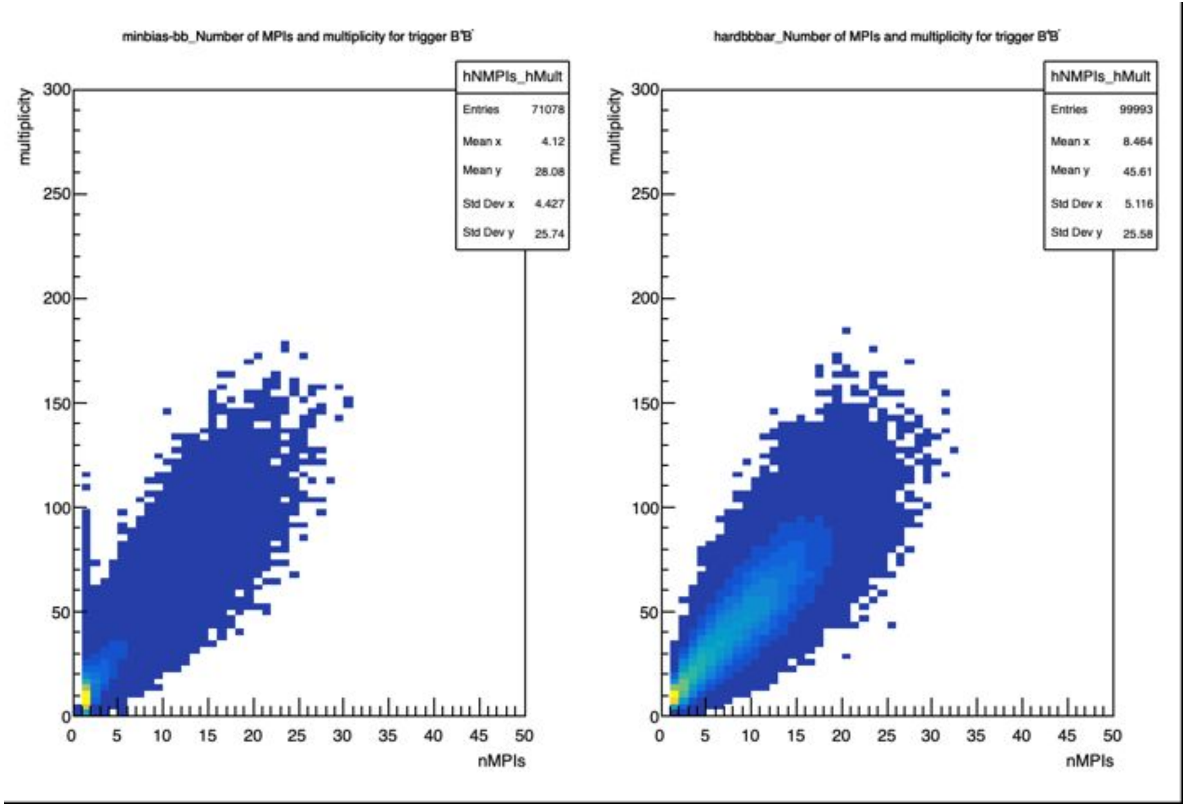




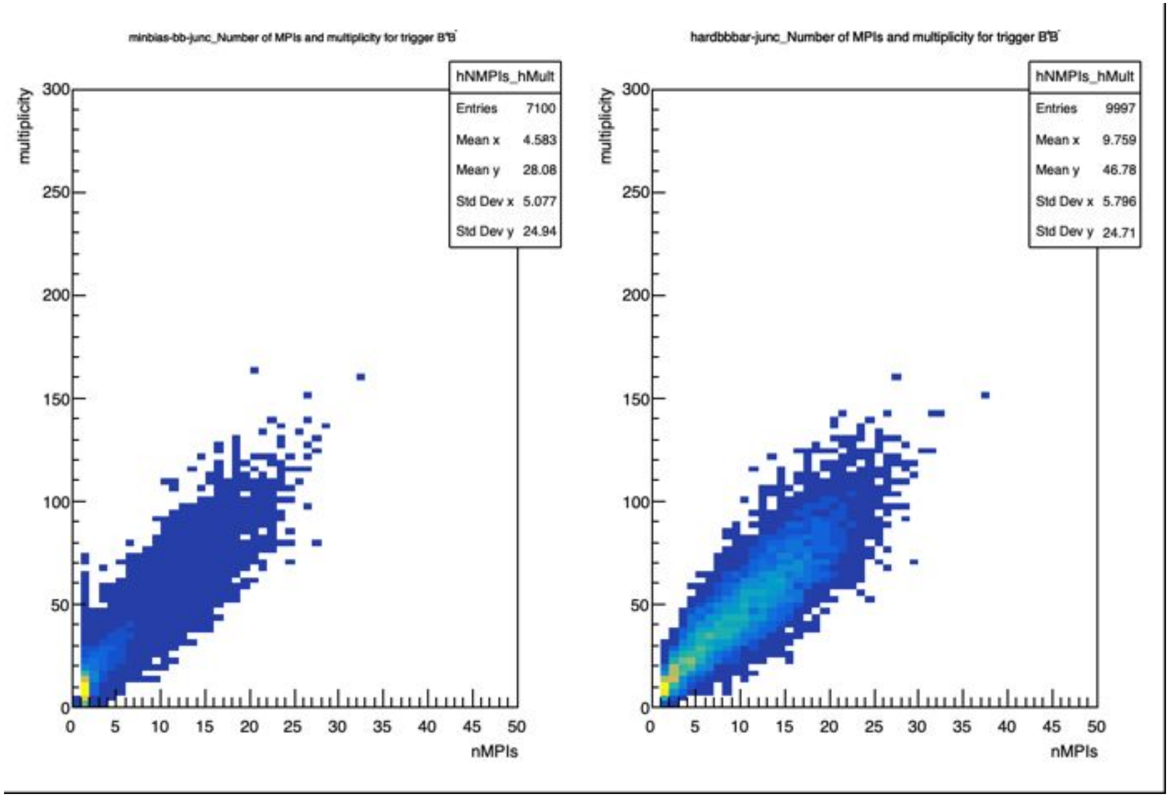
Charm Junctions: number of MPIs vs multiplicity



Beauty Monash: number of MPIs vs multiplicity



Beauty Junctions: number of MPIs vs multiplicity



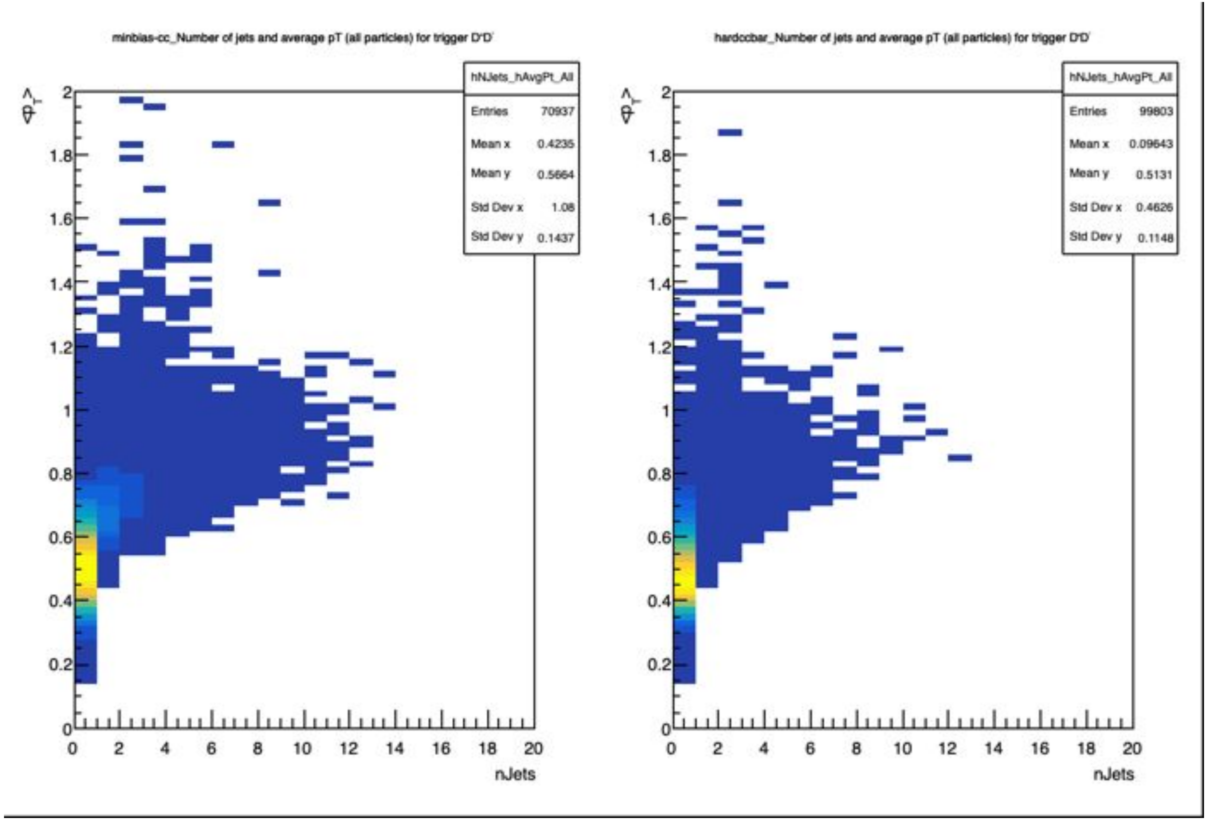
Conclusion: number of MPIs vs multiplicity

These plots are more of a consistency check: is multiplicity a good probe for MPIs and how similar are all the simulations?

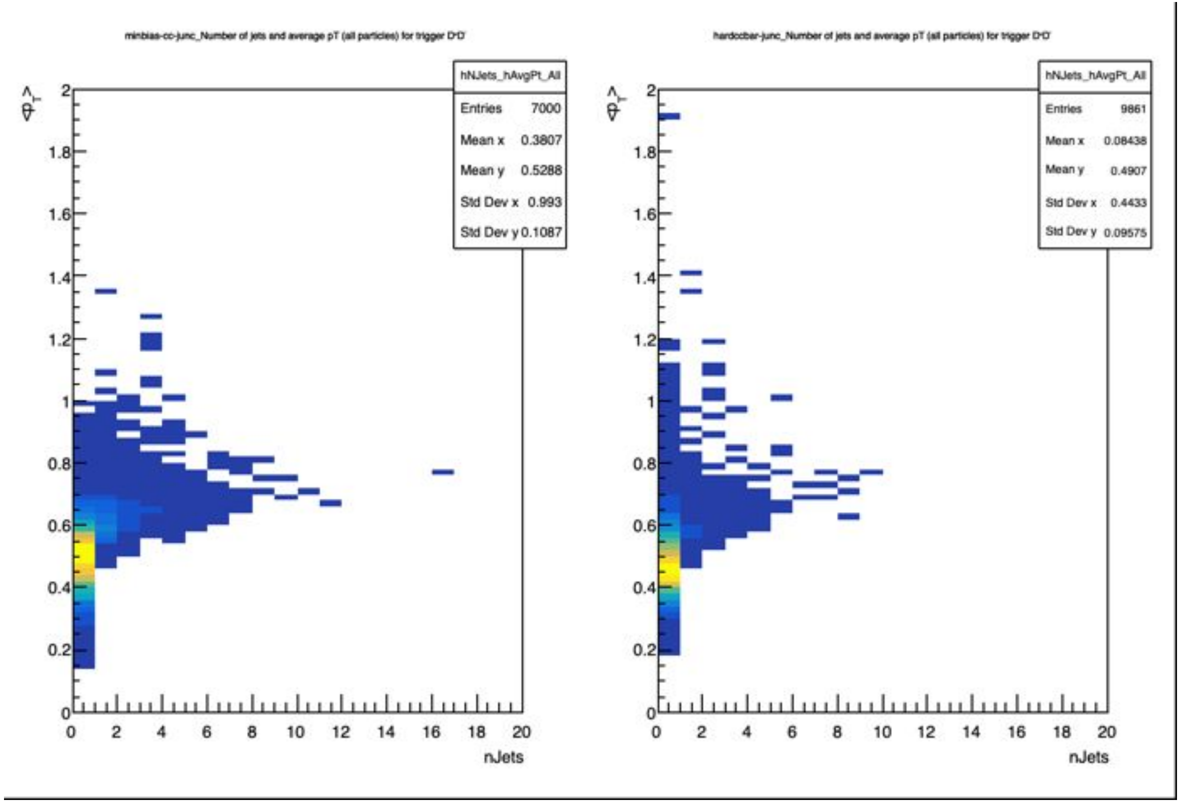
- They all have the (expected) trend that the multiplicity increases as the number of MPIs increases, and all trends in all simulation types are similar.

It might be worth noting the peak at nMPIs = 1 for the MB simulations. This ‘peak’ is absent in the boosted simulations.

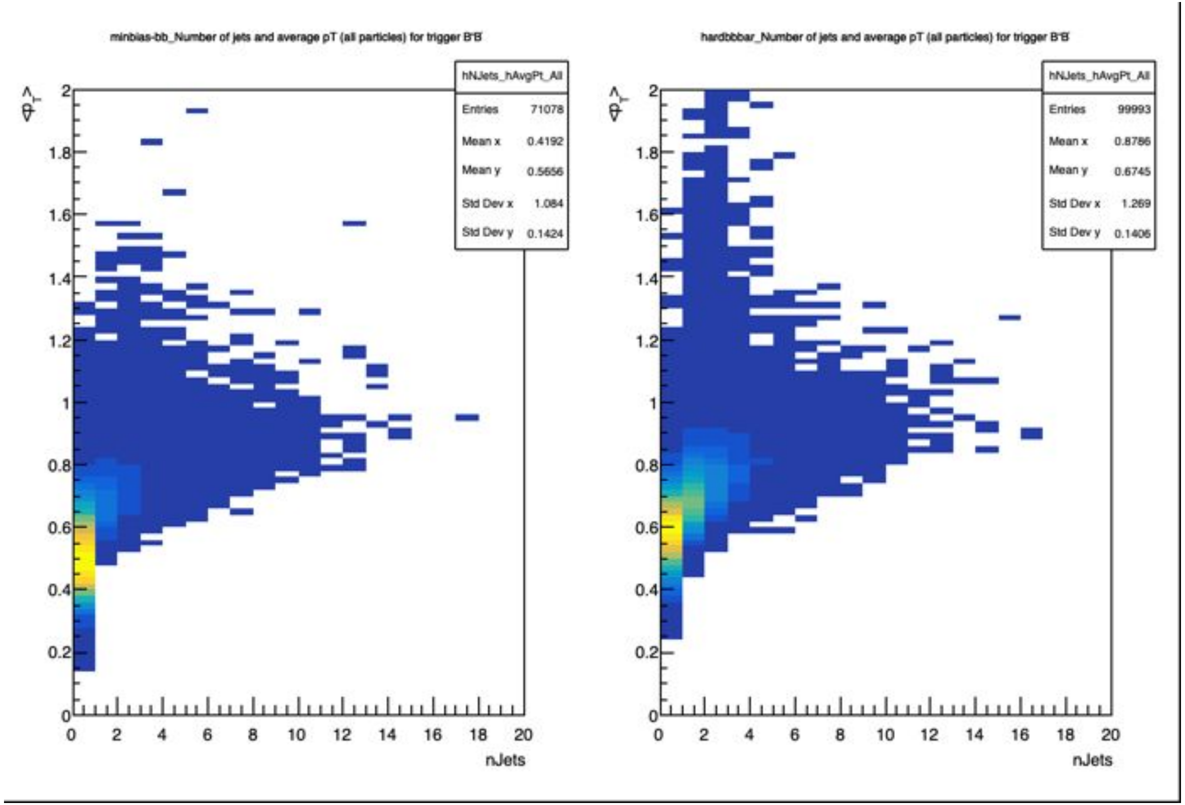
Charm Monash: number of jets vs <pT> (all)



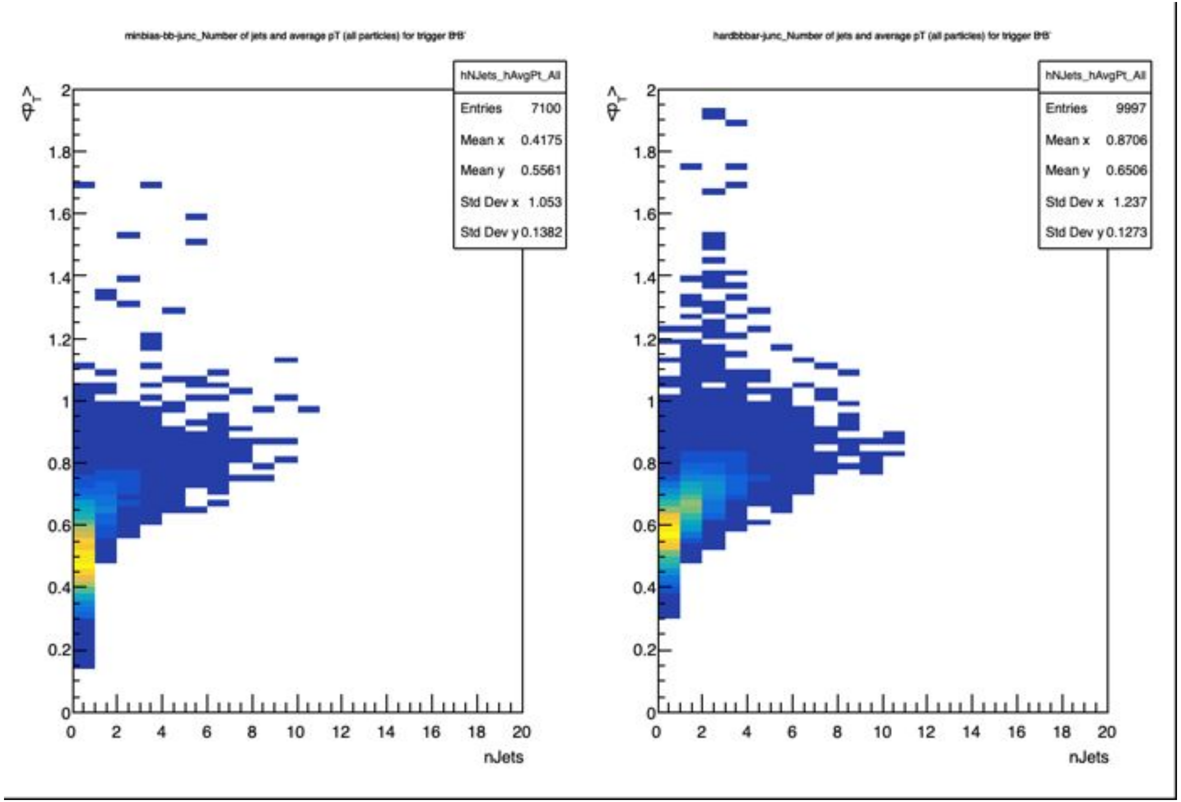
Charm Junctions: number of jets vs <pT> (all)



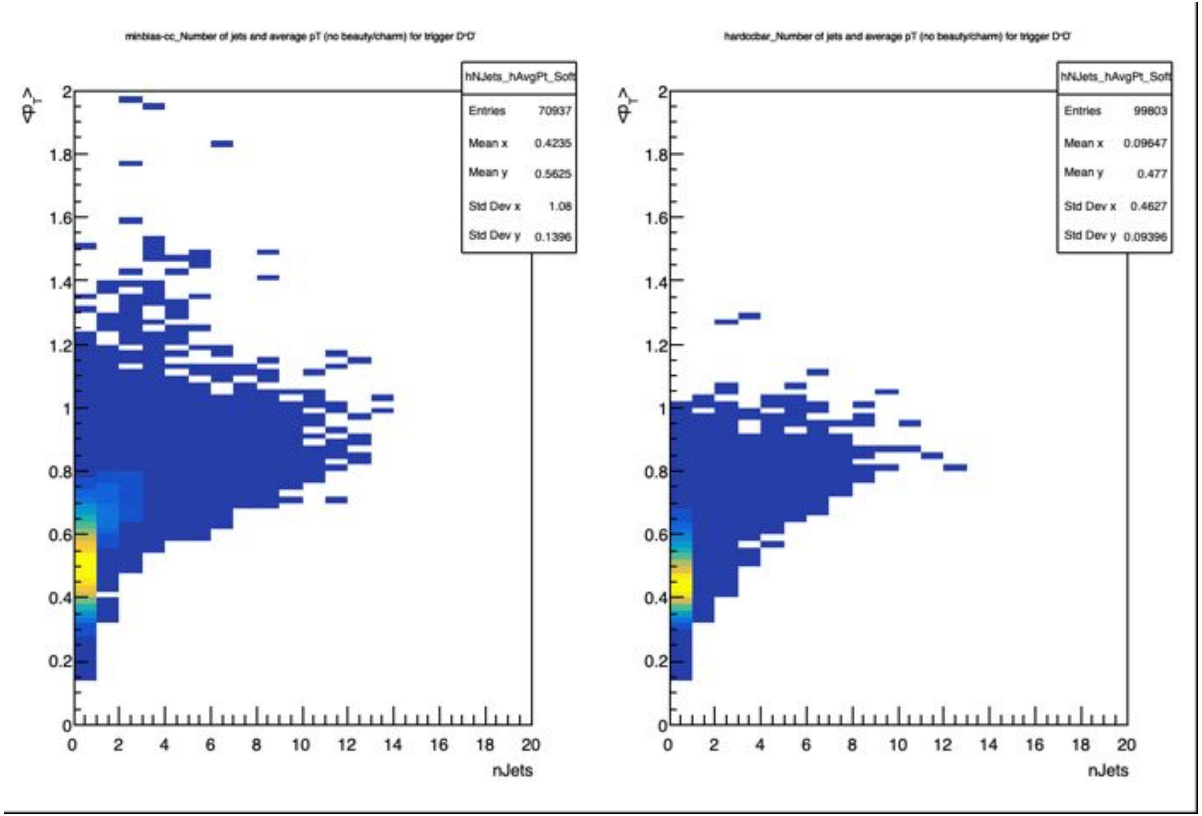
Beauty Monash: number of jets vs <pT> (all)



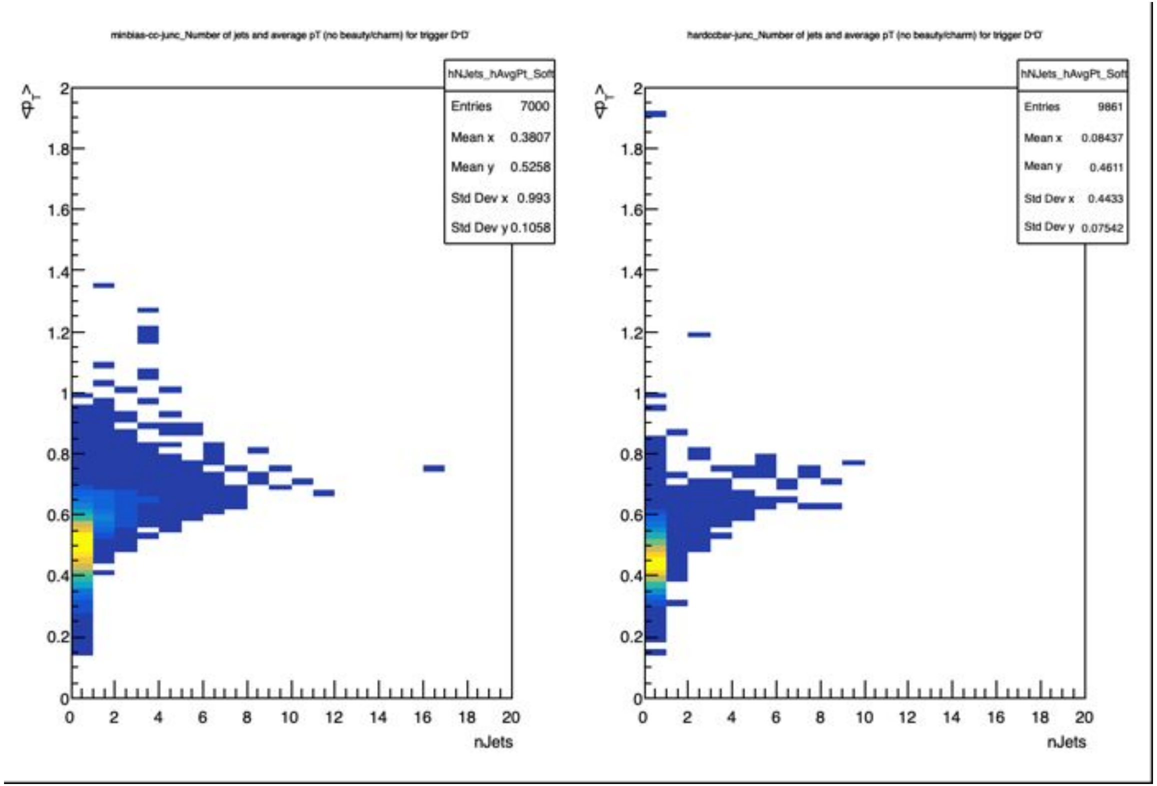
Beauty Junctions: number of jets vs <pT> (all)



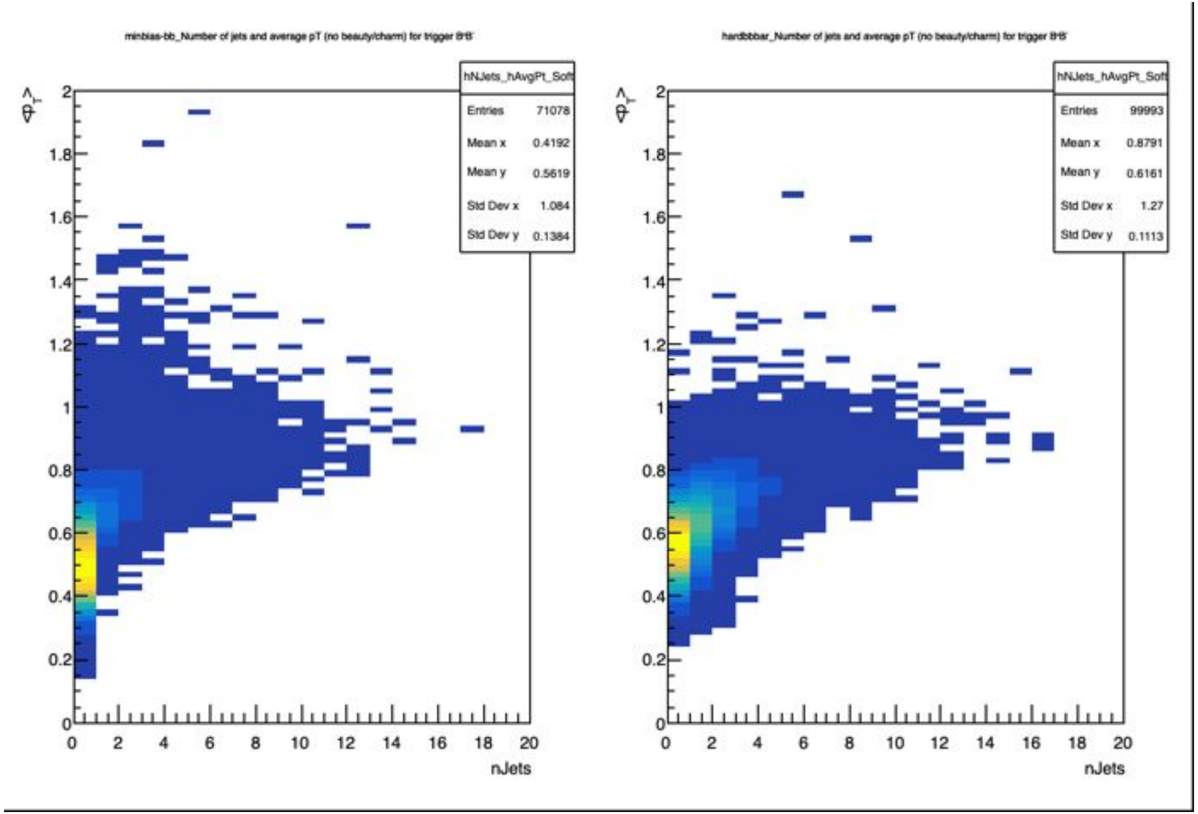
Charm Monash: number of jets vs <pT> (soft)



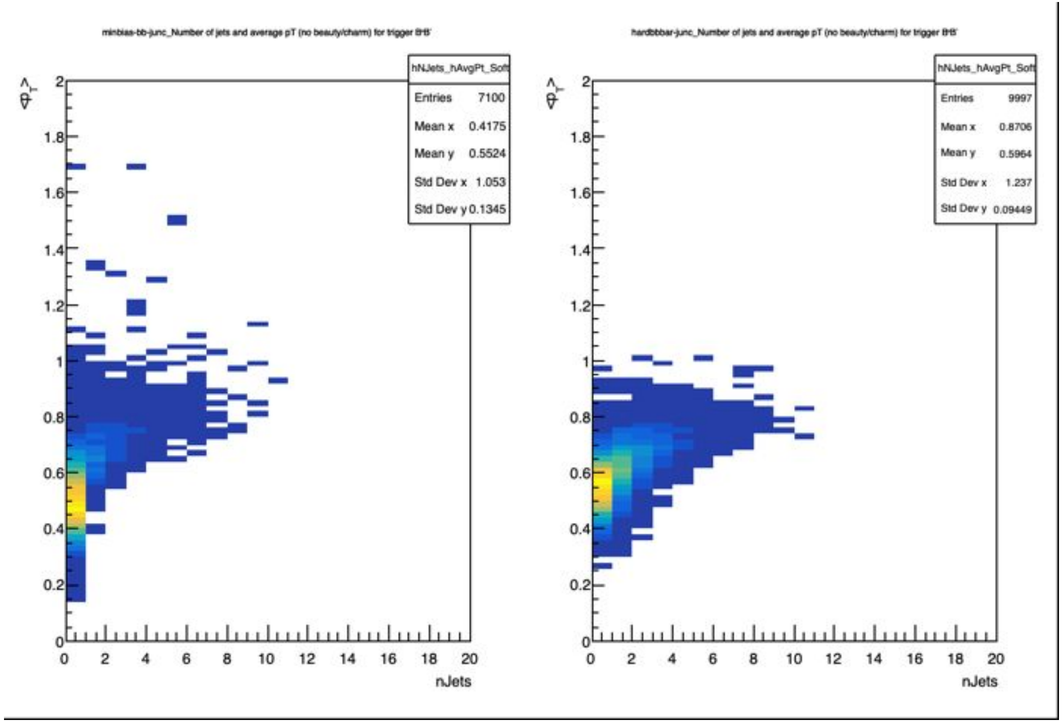
Charm Junctions: number of jets vs <pT> (soft)



Beauty Monash: number of jets vs <pT> (soft)



Beauty Junctions: number of jets vs <pT> (soft)



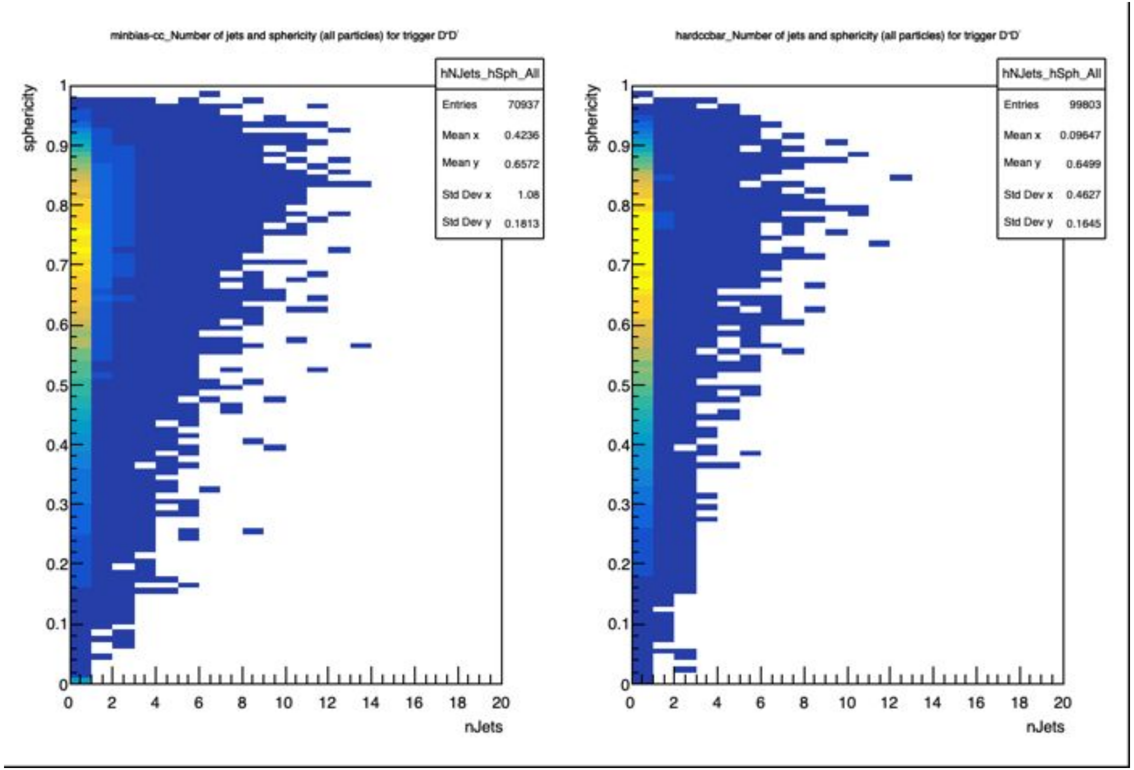
Conclusion: number of jets vs <pT>

I don't see many differences in trends between all the simulation types.

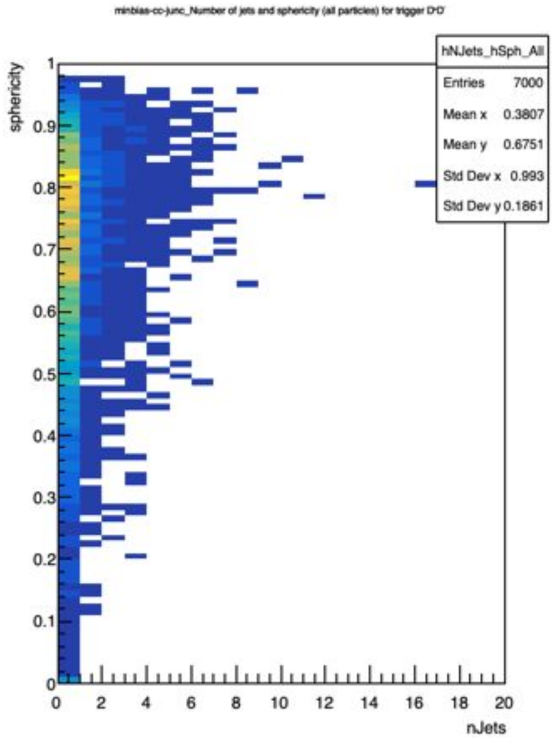
The general trend is what I expected, the <pT> increases as the number of jets increases, but it saturates quickly.

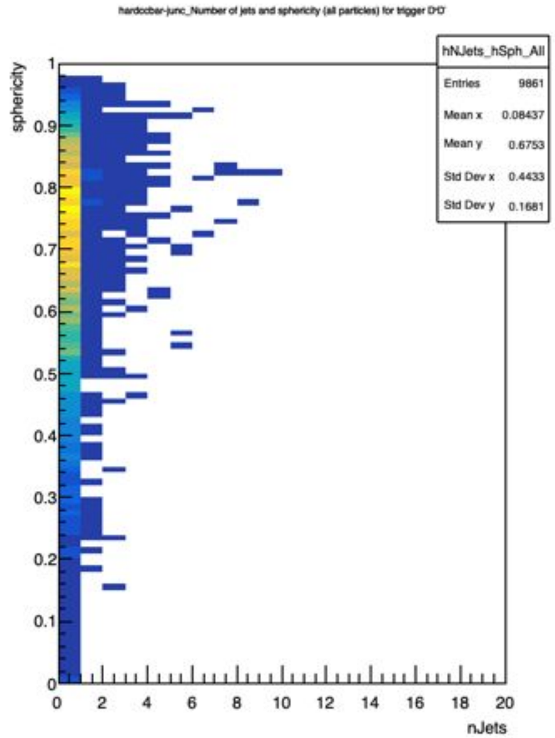
- It is interesting that we don't see the (quite) different behaviour for beauty-charm in the <pT> here.

Charm Monash: number of jets vs sphericity (all)

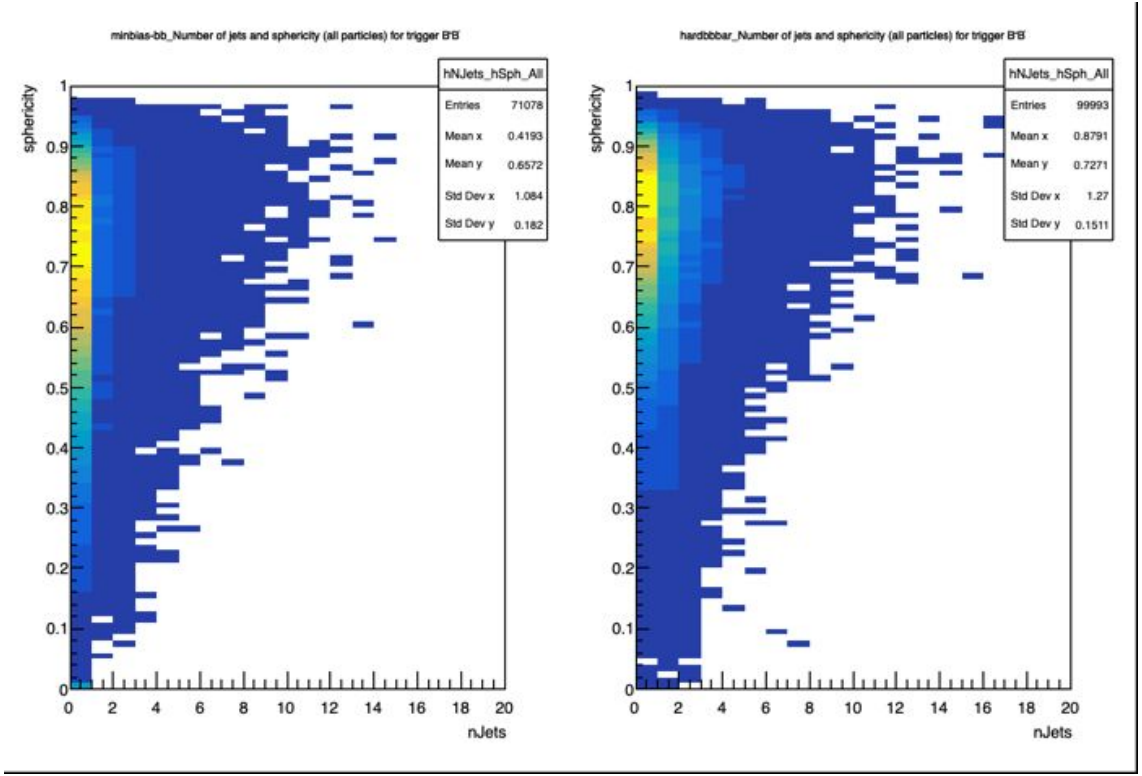


Charm Junctions: number of jets vs sphericity (all)

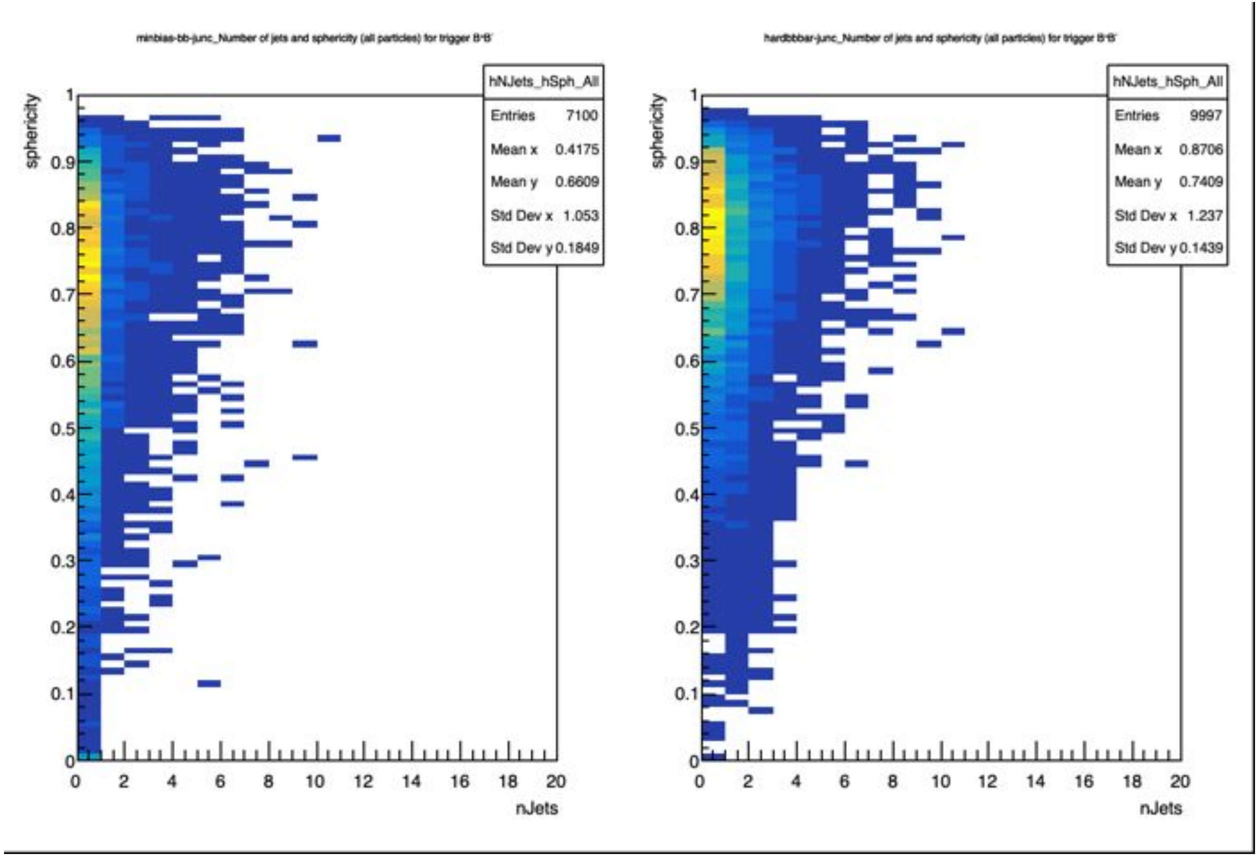




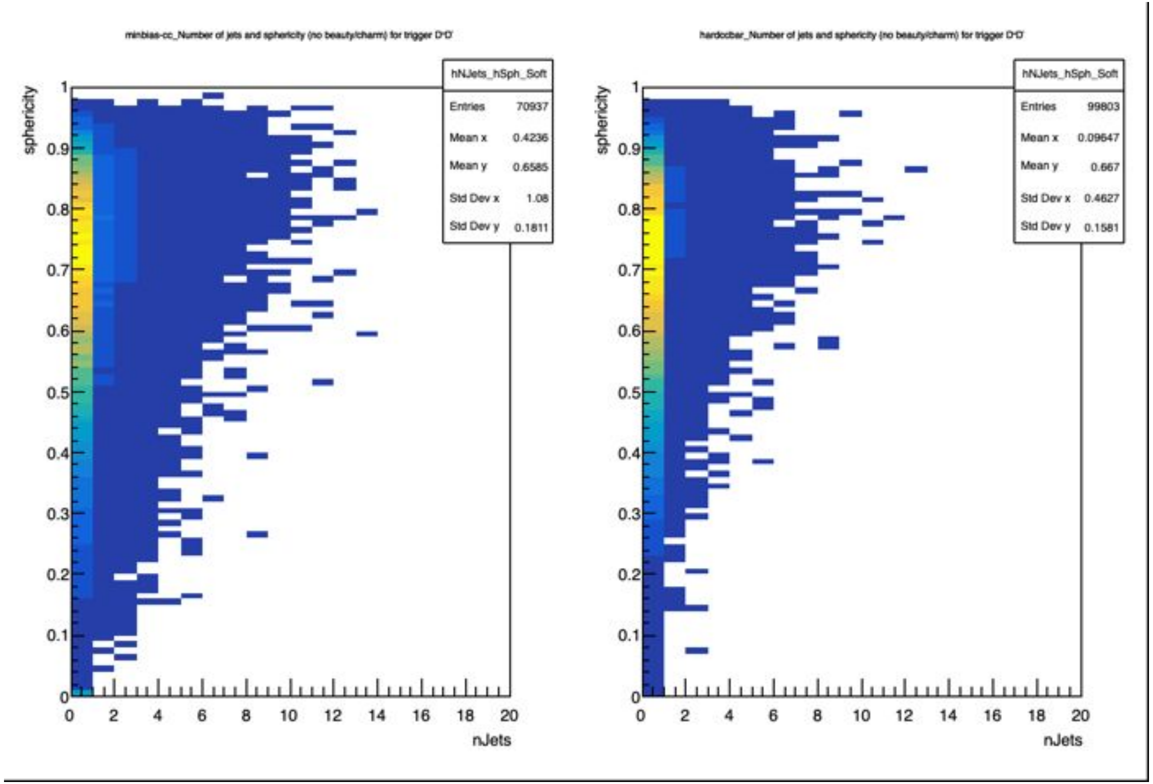
Beauty Monash: number of jets vs sphericity (all)



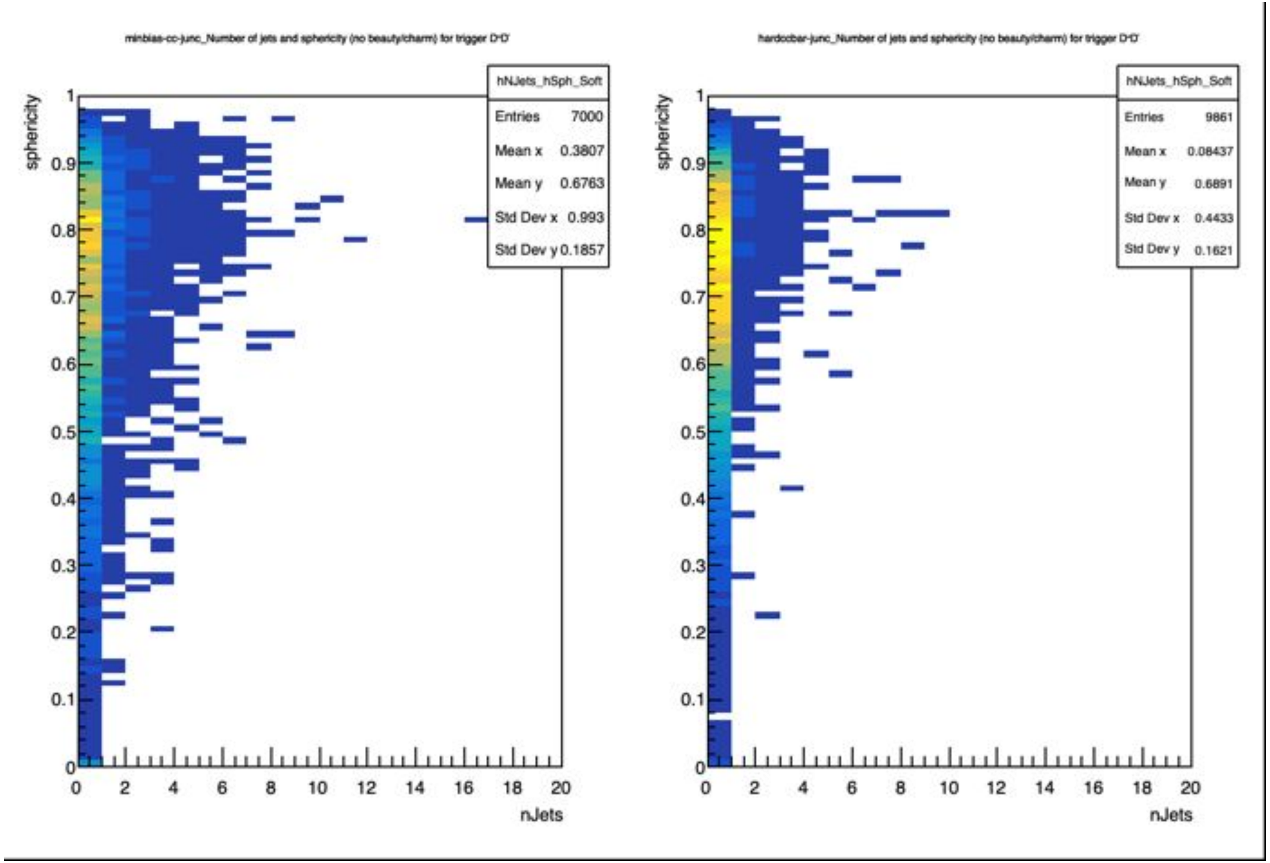
Beauty Junctions: number of jets vs sphericity (all)



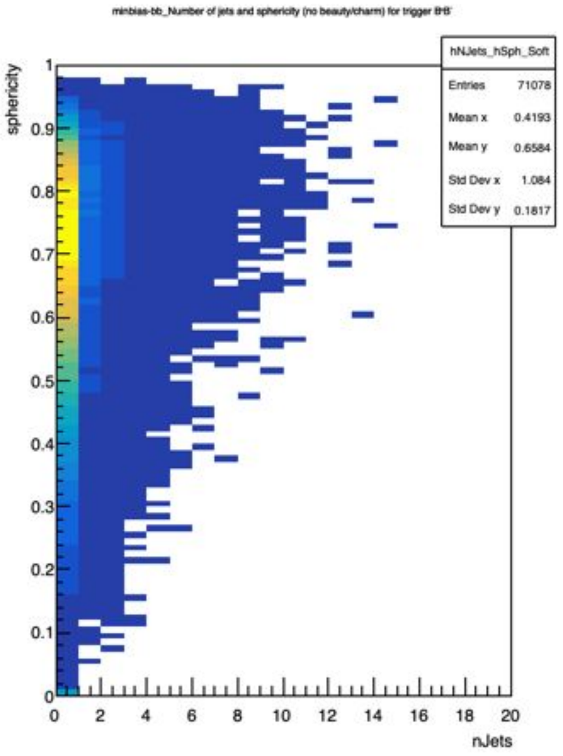
Charm Monash: number of jets vs sphericity (soft)

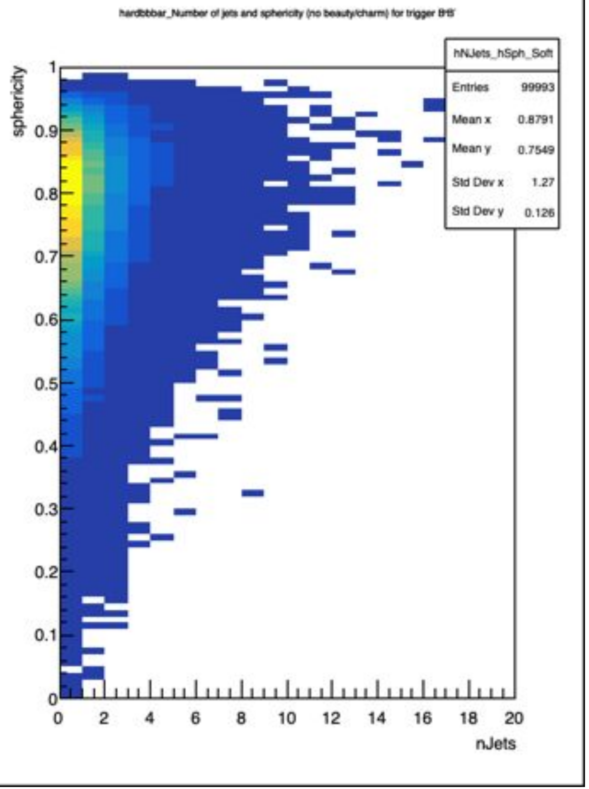


Charm Junctions: number of jets vs sphericity (soft)

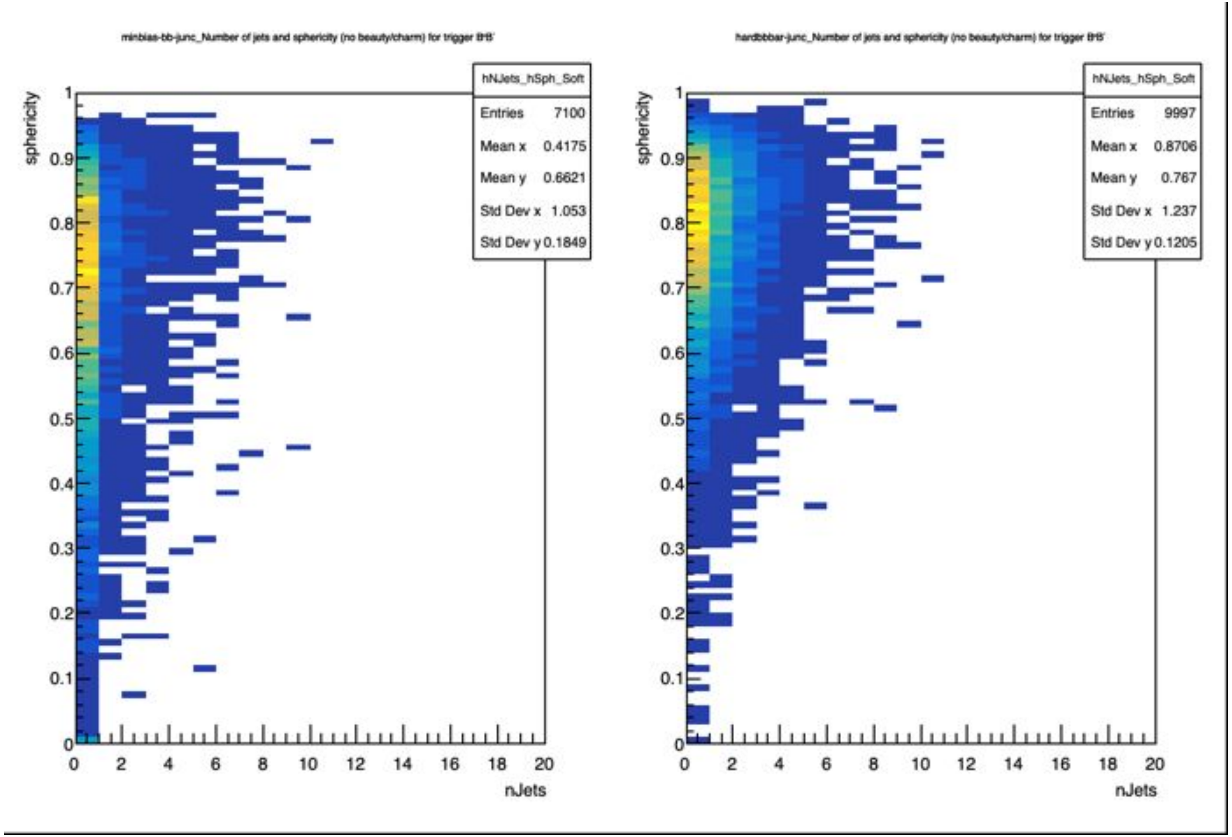


Beauty Monash: number of jets vs sphericity (soft)





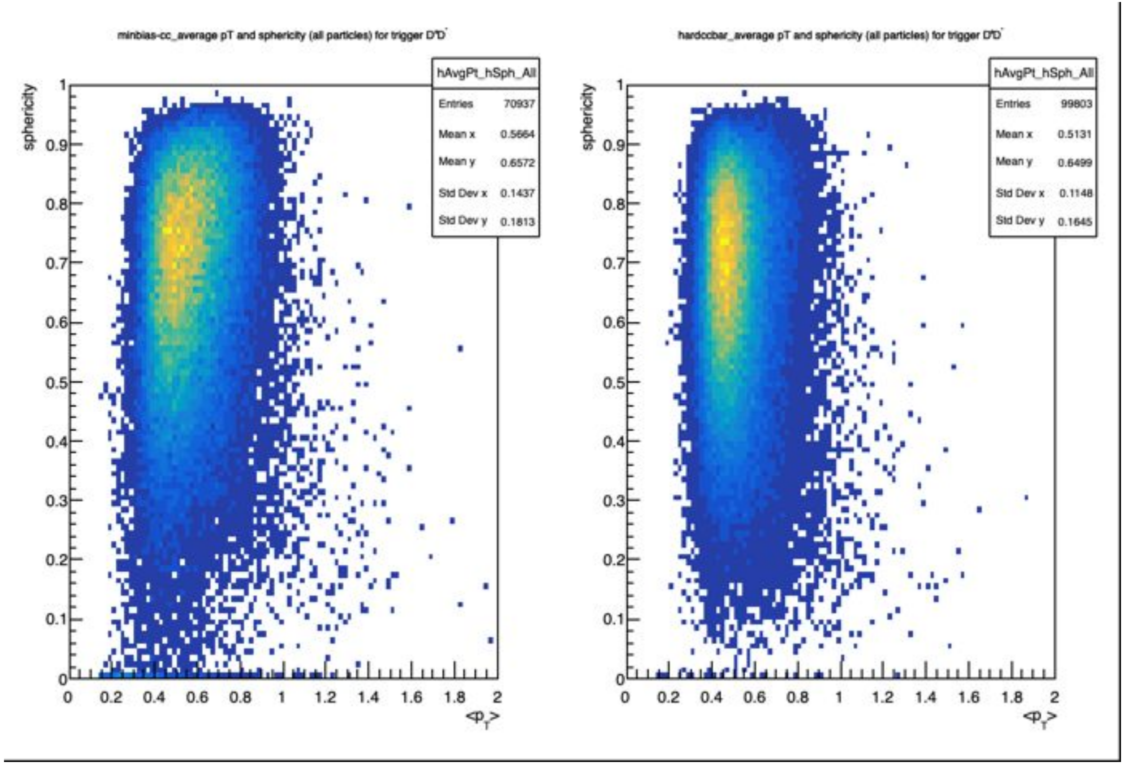
Beauty Junctions: number of jets vs sphericity (soft)



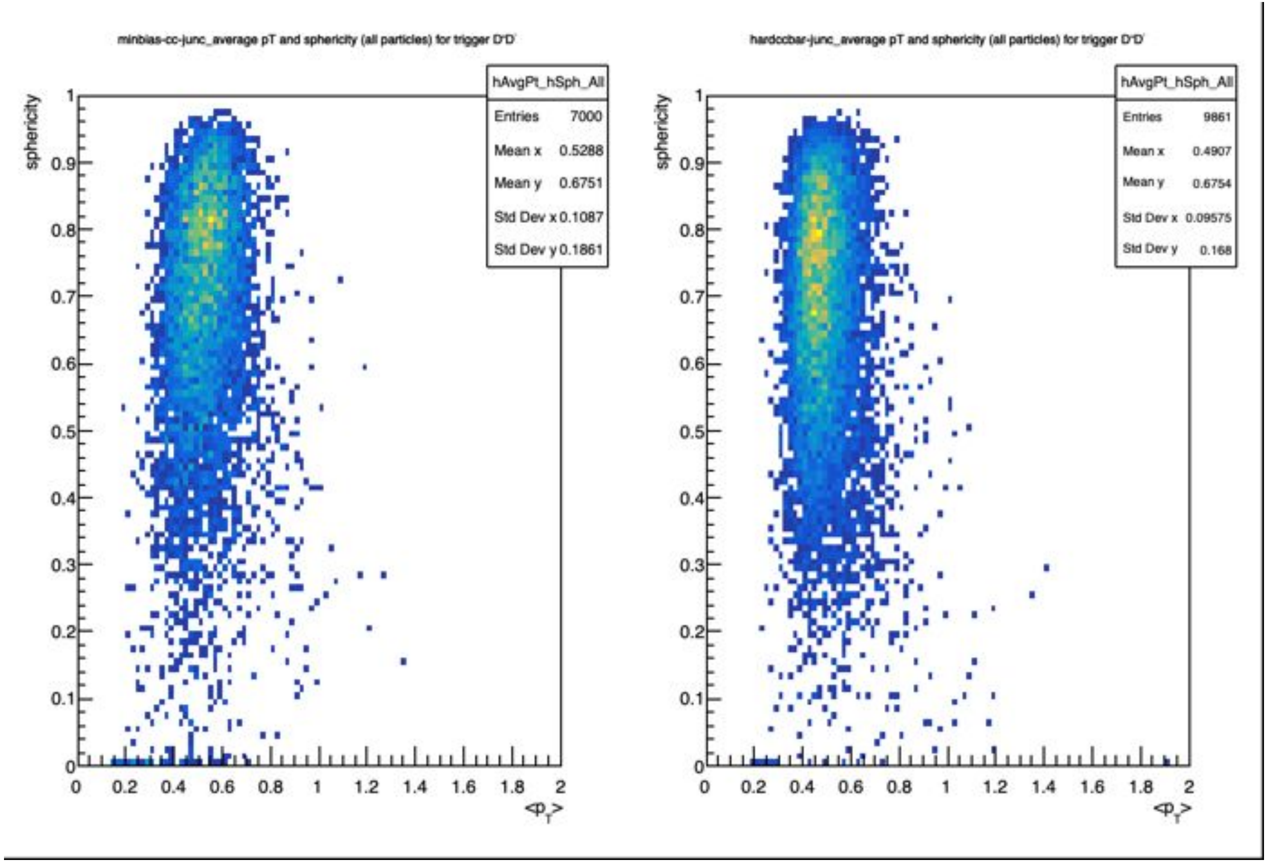
Conclusion: number of jets vs sphericity

Basically this doesn't give new information, because we have the correlation of number of jets and multiplicity. The same sphericity trends as before are seen here.

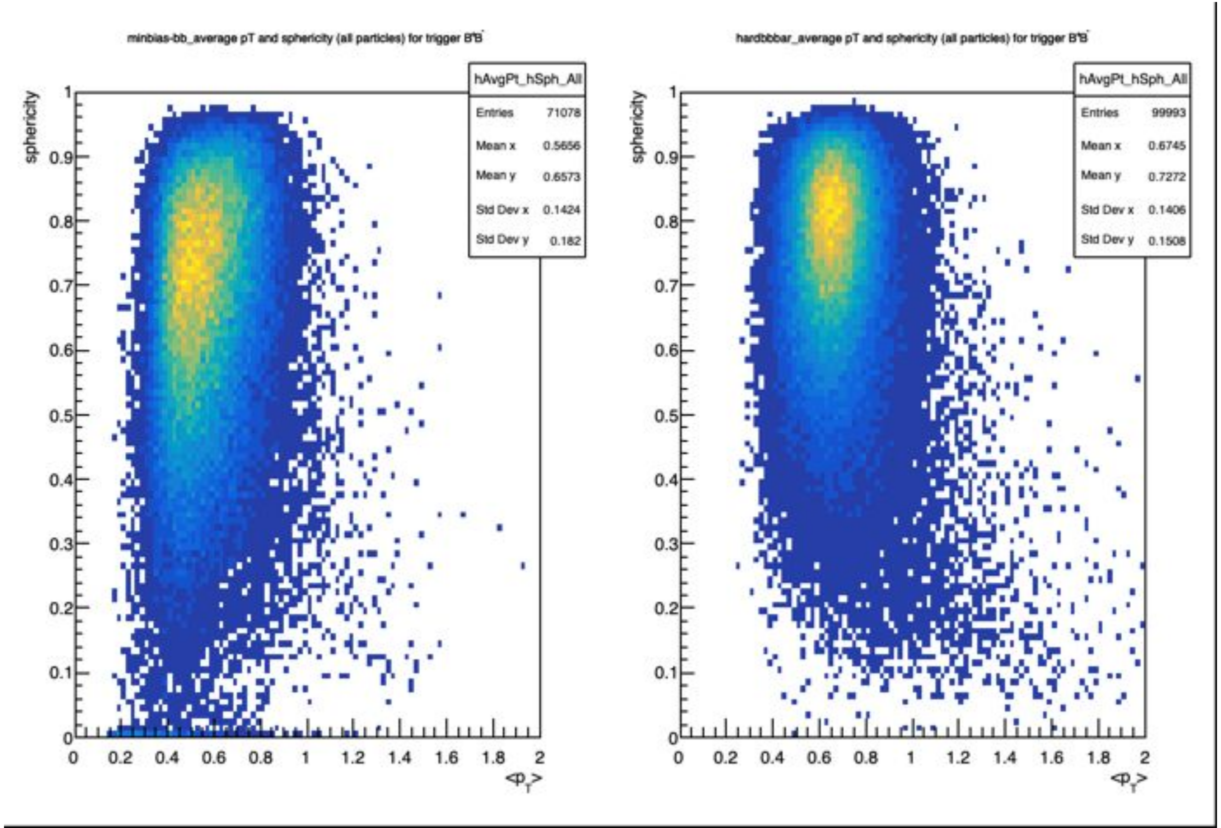
Charm Monash: <pT> vs sphericity (all)



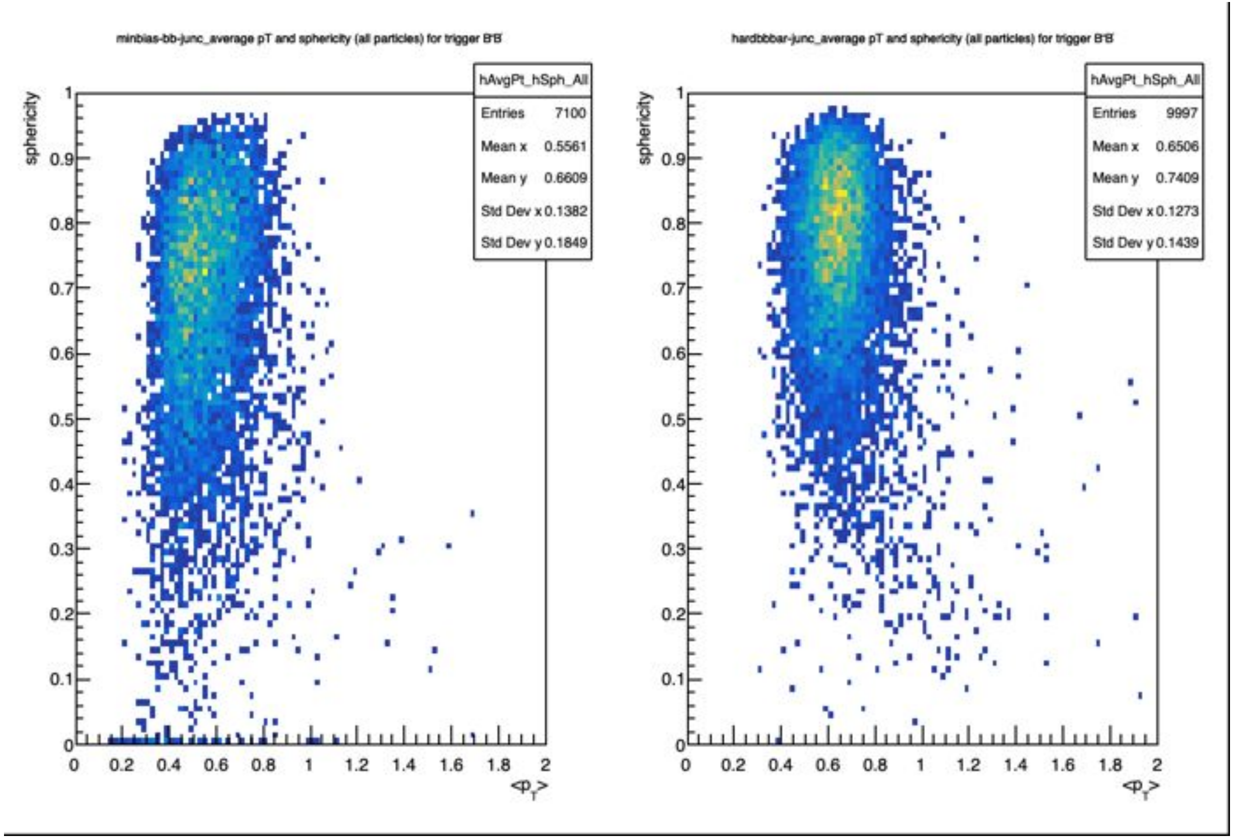
Charm Junctions: <pT> vs sphericity (all)



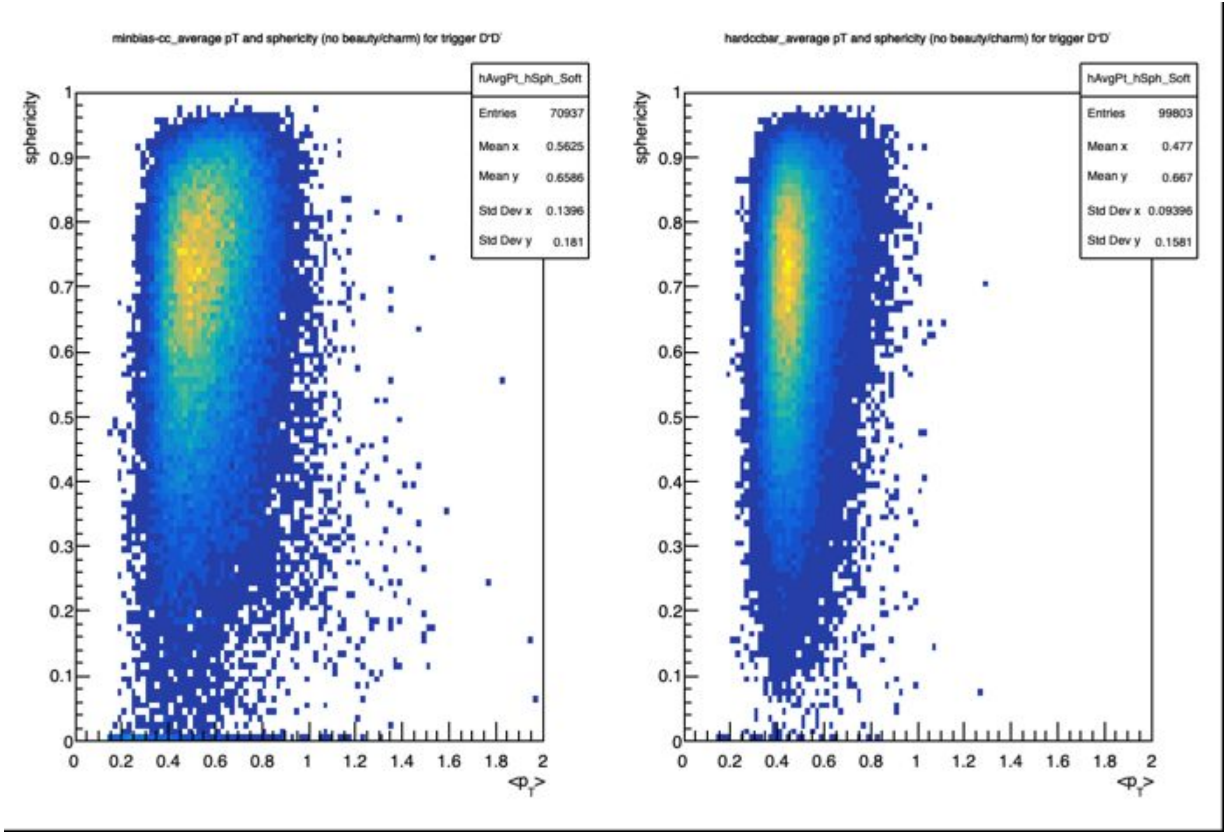
Beauty Monash: <pT> vs sphericity (all)



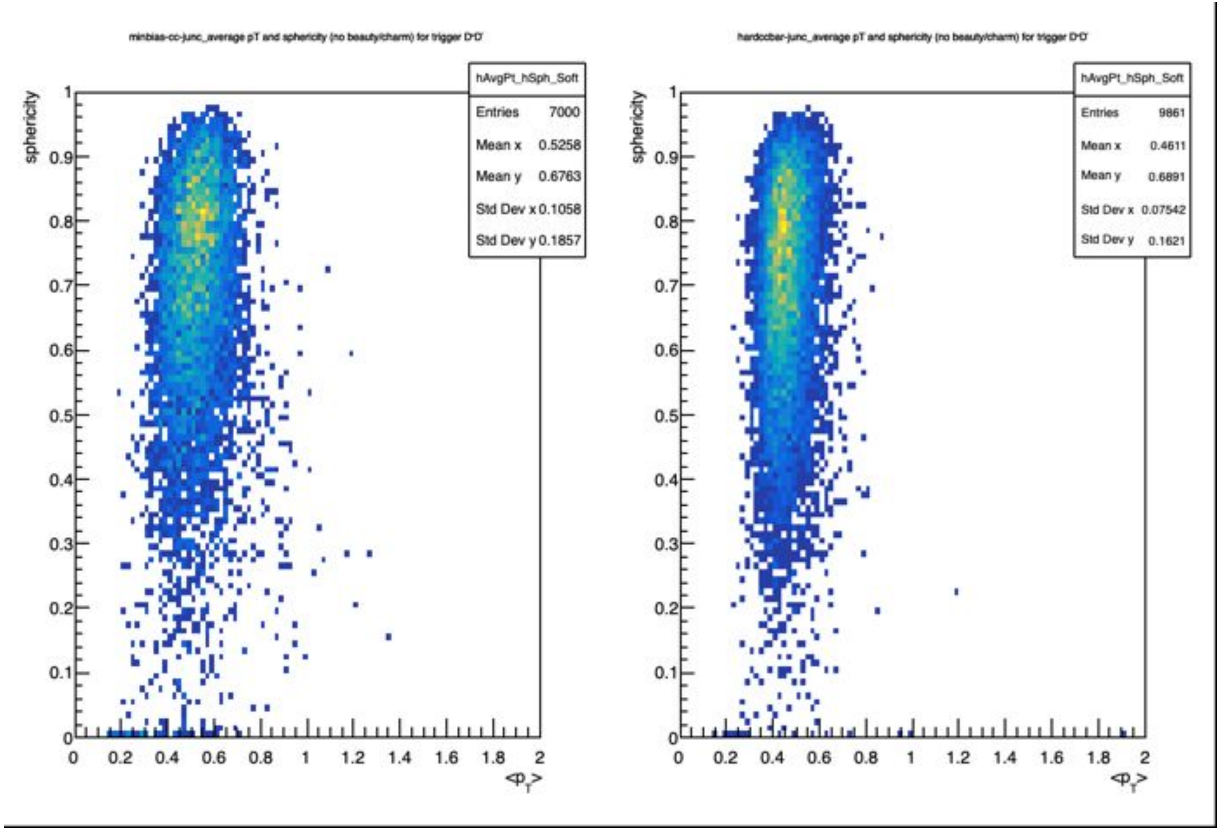
Beauty Junctions: <pT> vs sphericity (all)



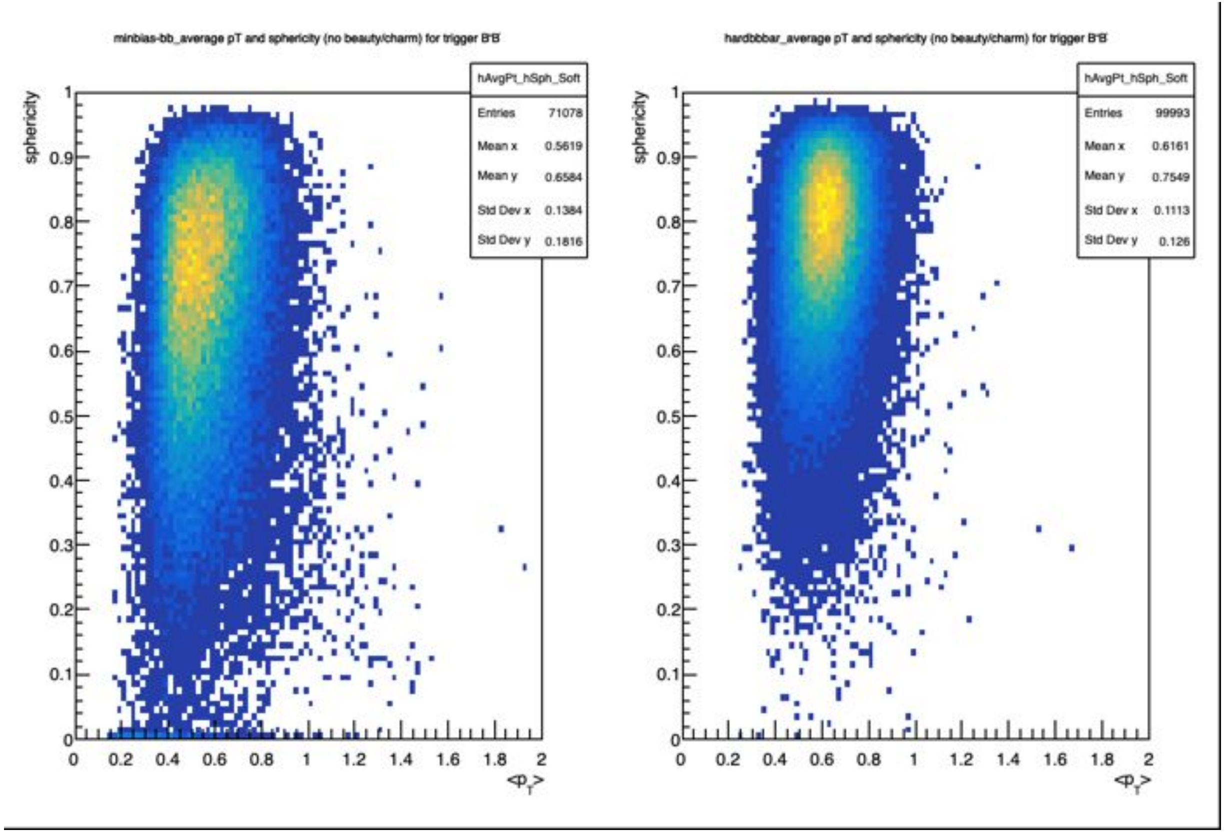
Charm Monash: <pT> vs sphericity (soft)



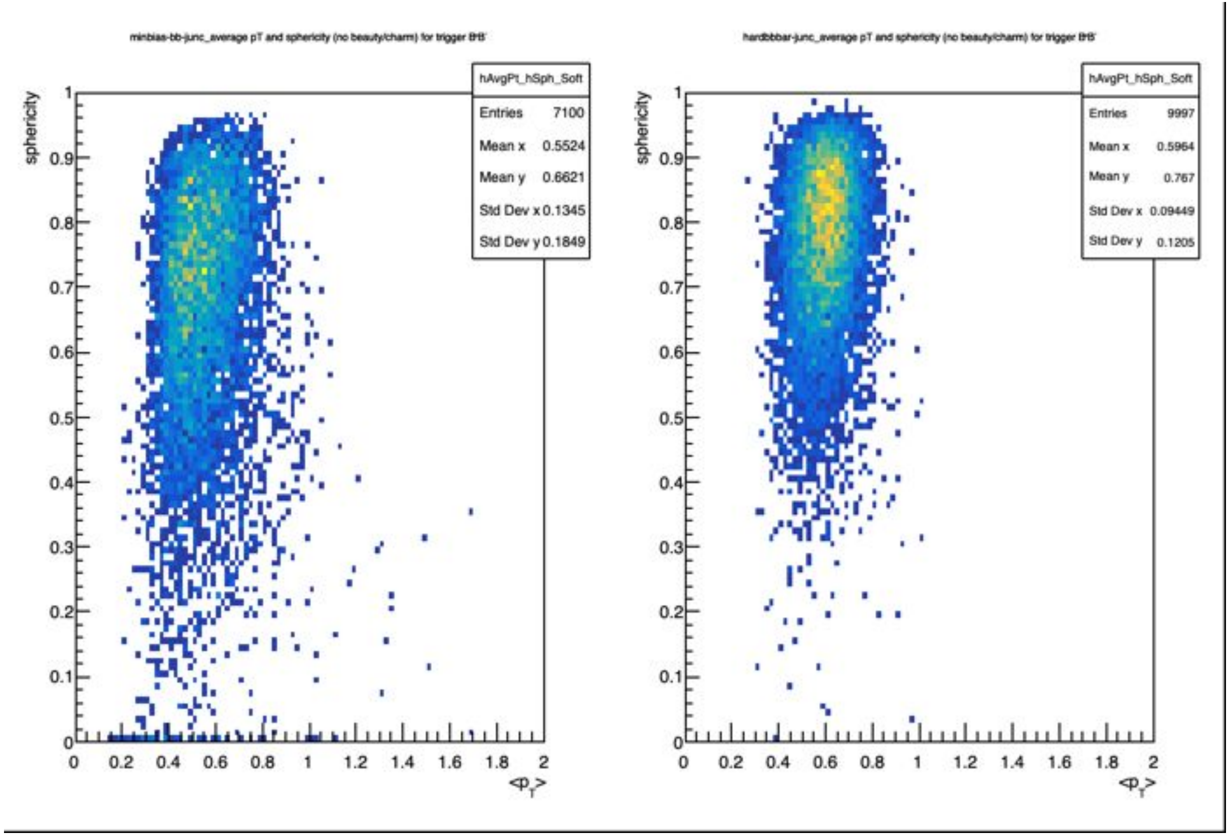
Charm Junctions: <pT> vs sphericity (soft)



Beauty Monash: <pT> vs sphericity (soft)



Beauty Junctions: <pT> vs sphericity (soft)



Conclusion: <pT> vs sphericity

I think the most noticeable effect here is the ‘escaping tail’ for the boosted beauty all simulations. It seems like in those simulations the events with higher <pT> seem to be more ‘jet-like’ and that it occurs more often than in the other ones.

- This does seem to agree with the hypothesis from earlier.

General conclusion

I think the most important conclusion is that charm MB, charm boosted and beauty MB are similar, while beauty boosted is quite different.

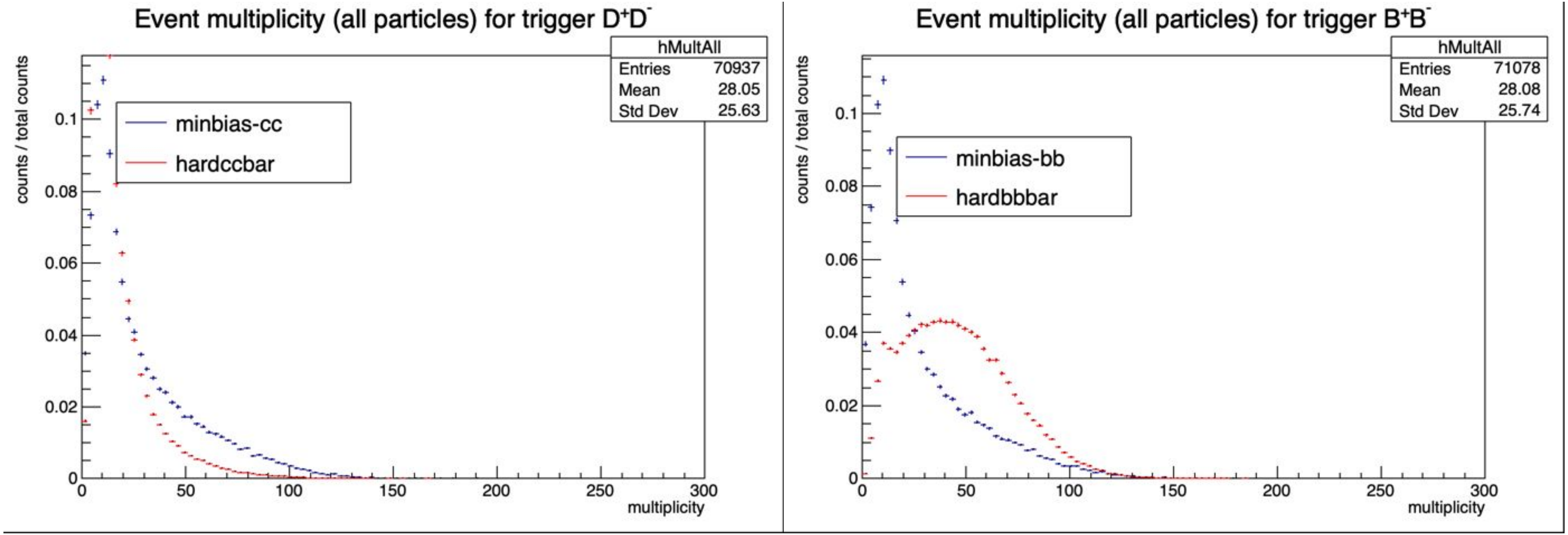
- The question is to understand why this is the case and how it can be ‘equalised’?

Some hints are found in the multiplicity spectra, the number of jets per event and the event sphericities (see next slides for easy reference to plots). It seems like the hardccbar simulations are more ‘natural’ and closer to MB, while the bb-bar are more artificial. Maybe a hard bb jet is usually formed in the biased bb simulations, while for charm it’s less ‘forced’?

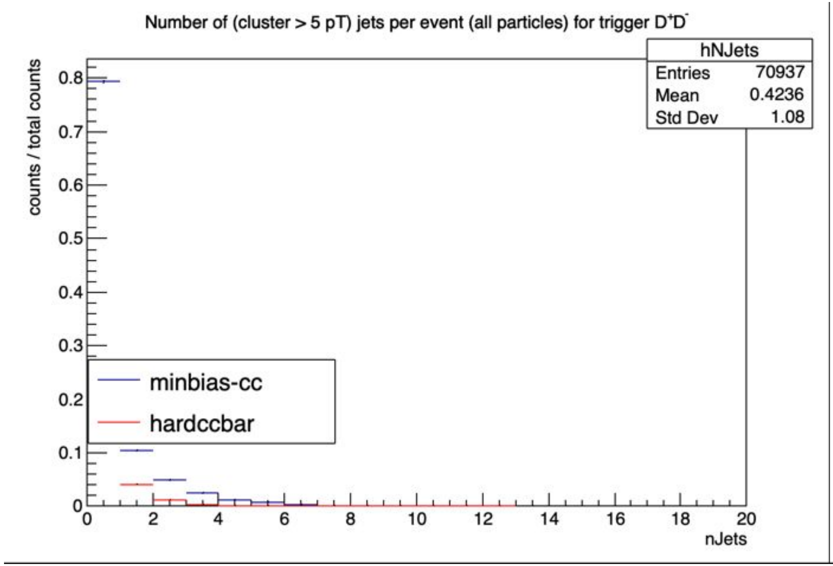
At high number of jet events this seems to ‘converge’ to the same trend for beauty and for charm

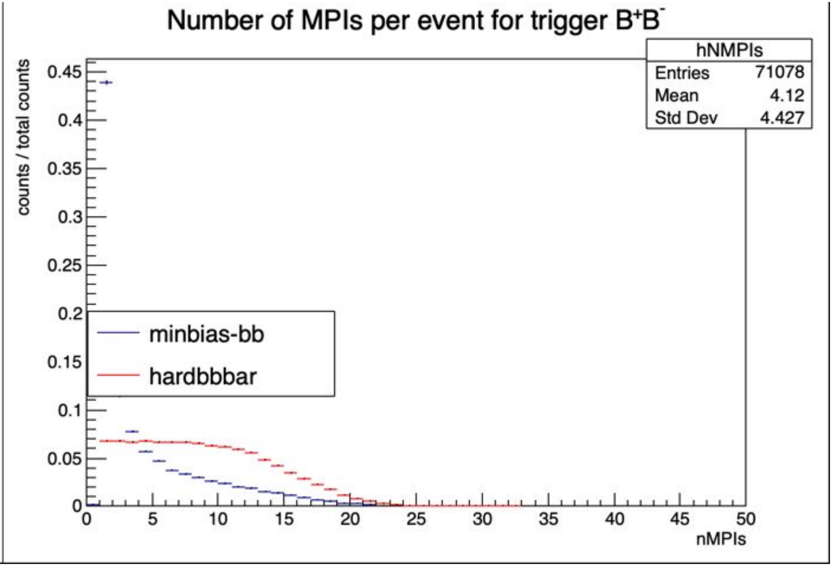
- But some results show the opposite, where hardbbbar events seem to be more ‘spherical’ (i.e. less ‘jet-like’) (e.g. in the 1D sphericity histograms).

Monash multiplicity

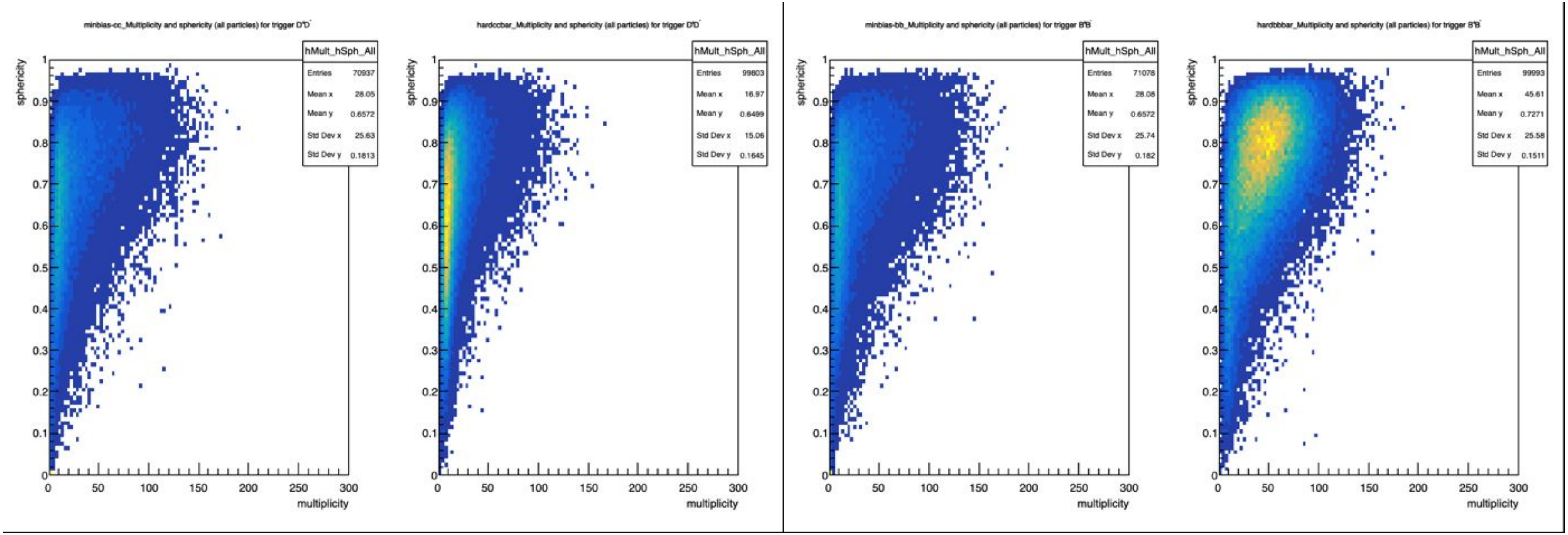


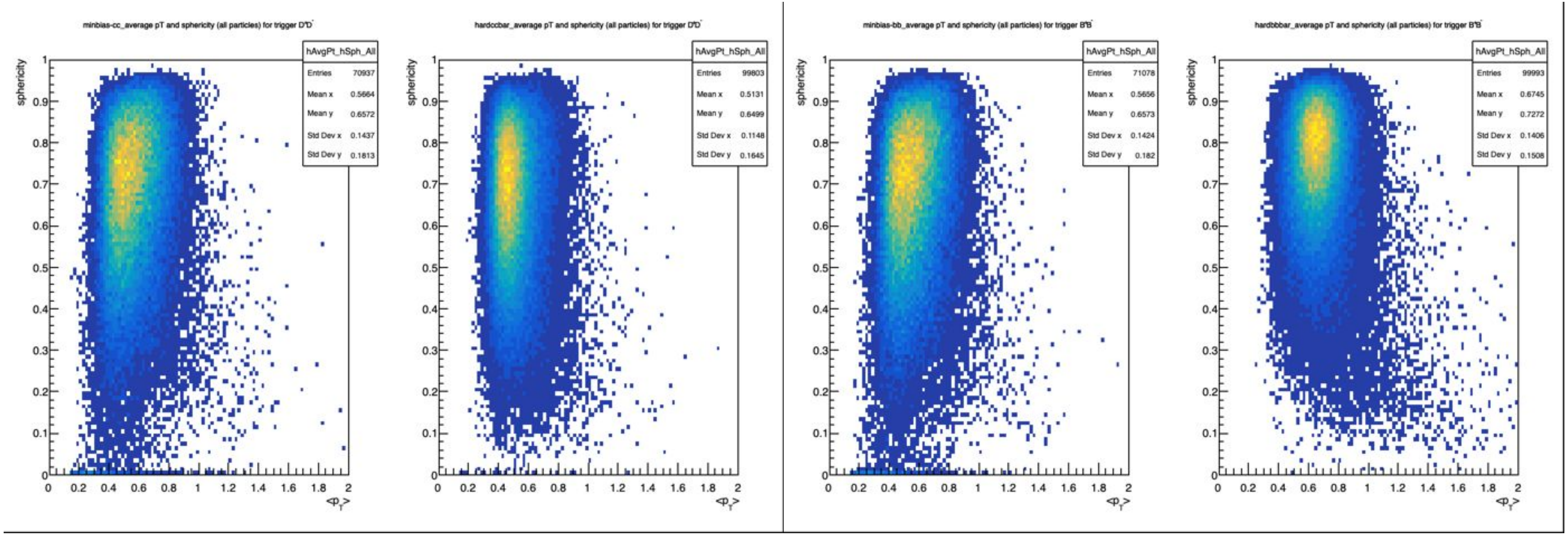
Monash number of jets





Monash multiplicity vs shpercity





minbias-cc_average pT and sphericity (all particles) for trigger D⁰D̄*
hAvgPt_hSph_All
Entries 70937
Mean x 0.5664
Mean y 0.6572
Std Dev x 0.1437
Std Dev y 0.1813
sphericity
$\langle p_T \rangle$
hardccbar_average pT and sphericity (all particles) for trigger D⁰D̄*
hAvgPt_hSph_All
Entries 99803
Mean x 0.5131
Mean y 0.6499
Std Dev x 0.1148
Std Dev y 0.1645
sphericity
$\langle p_T \rangle$
minbias-bb_average pT and sphericity (all particles) for trigger B⁰B̄*
hAvgPt_hSph_All
Entries 71078
Mean x 0.5656
Mean y 0.6573
Std Dev x 0.1424
Std Dev y 0.182
sphericity
$\langle p_T \rangle$
hardbbbar_average pT and sphericity (all particles) for trigger B⁰B̄*
hAvgPt_hSph_All
Entries 99993
Mean x 0.6745
Mean y 0.7272
Std Dev x 0.1406
Std Dev y 0.1508
sphericity
$\langle p_T \rangle$

General conclusion

Other comparisons:

- Only marginal differences between “soft” and “all”. Easier to focus on one of the two (easiest one is “all”).
- Essentially no differences between monash and junctions -> comparison between the two is “fair” and the same categorisation can be used

Note on statistics

Code seems to run much faster with the following changes

- hardQCD instead of hardbbbar or hardccbar
- veto events with 0 multiplicity (we're not interested in them anyways..)
- no QC checks (no 'test' plots created for correlations)

Charm/event = 1.2 (hardQCD) vs 1.3 (hardccbar)

Beauty/event = 0.068 (hardQCD) vs 1.6 (hardbbbar)

-> 'loss' of about factor 100 for beauty.. Can we live with that? (consider also that hardQCD might be up to 10x faster..)

Considerations

Jet finding algorithm: many parameters

More statistics?

Pick variables to cut with? Repeat studies with those cuts to see the effect on the balancing?

Noticeable differences in bb-bar and cc-bar boosted simulations. Perhaps consider using both of them in each simulation, instead of just bb for beauty and cc for charm?

Repeat study with baryons?