

A comparison of different Zbb dataset

Stefano Lacaprara

INFN Padova

No meeting,
PD, 3 March 2014



Intro



Intro

- We have a new Zbb dataset
/Zbb_4F_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V19-v1/AODSIM:
 - ▶ 7 551 580 events
 - ▶ generated with 4 Fermions (4F) pdf
 - ▶ with no threshold on B-partons
- We have also the inclusive samples, from which we extracted the $Z + bb$ contribution
 - ▶ they have a threshold of 10 GeV on b-partons
- Can we use them both?
- if not, which one is better?
 - ▶ Compare MC vs MC to see how different they are for our studies;
 - ▶ Compare them with data to see which agrees better.



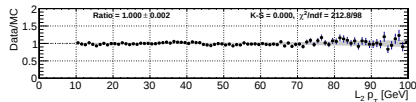
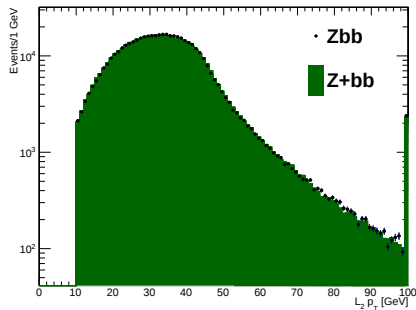
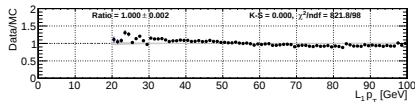
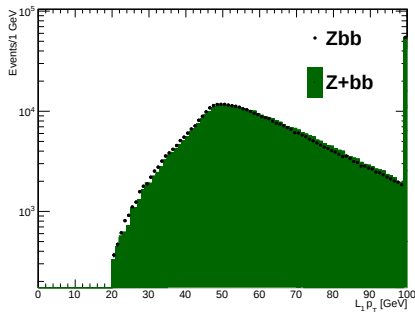
MC vs MC



- Comparing distribution at reco level for the two MC
 - ▶ Zbb (black dots): new dataset;
 - ▶ Z+bb (dark green histo): inclusive DY+jets with events with 2 b-partons selected;
- Normalized to same area
- showing comparison at pre-selection level
- for leptonic, hadronic, and lep+had part of the events.

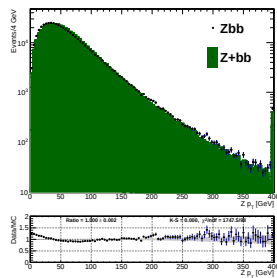
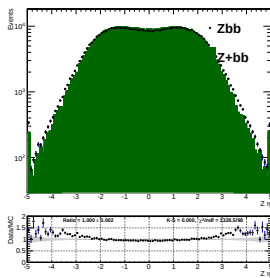
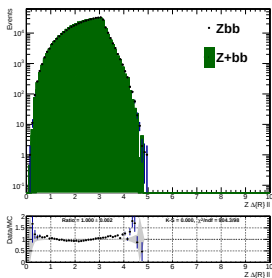


L pt



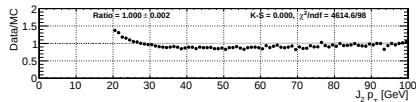
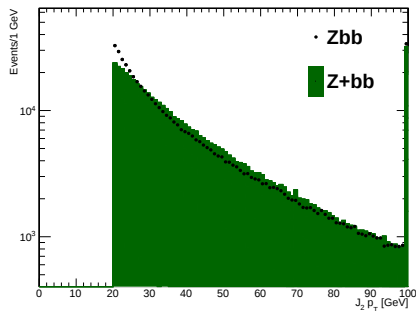
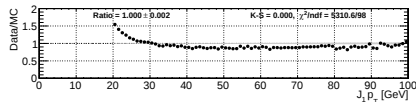
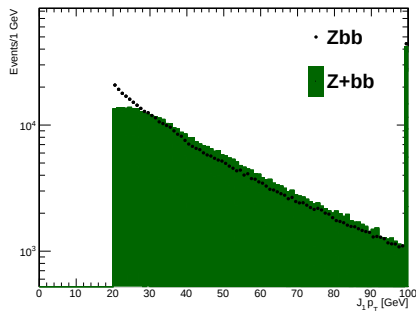


Z



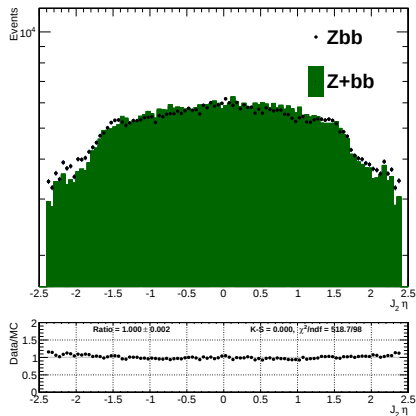
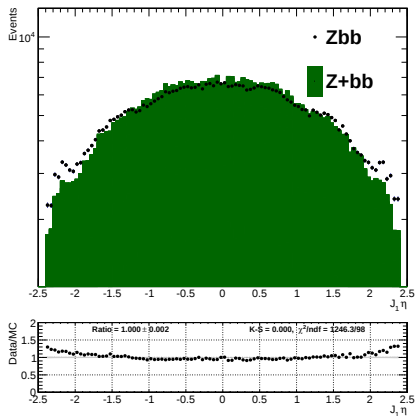


Jet pt



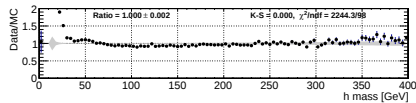
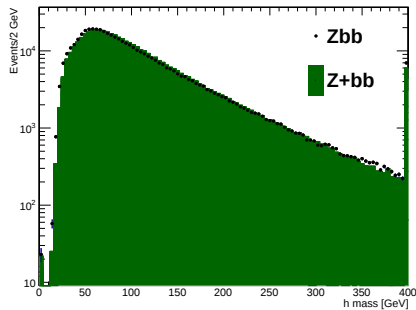
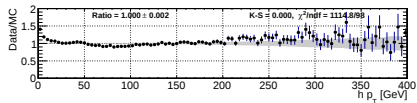
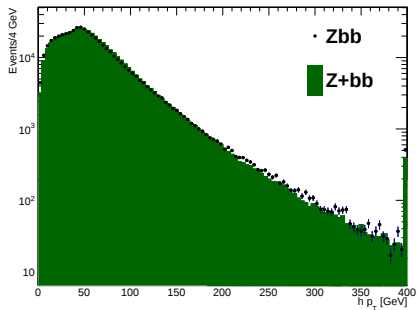


Jet eta



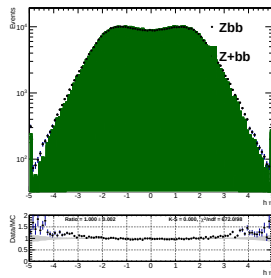
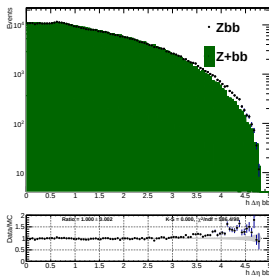
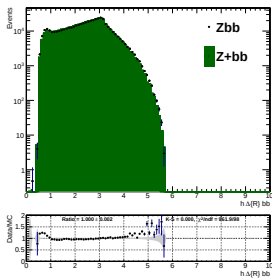


bb



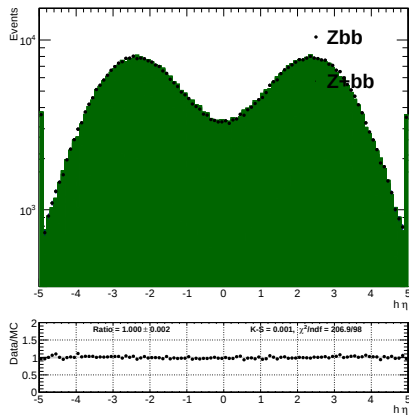
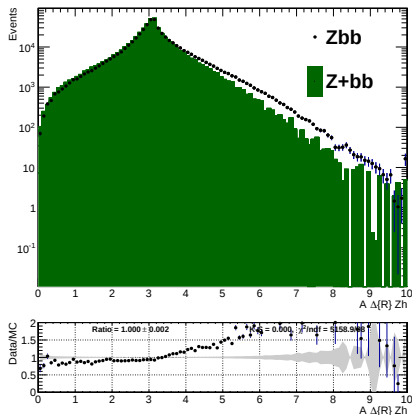


bb



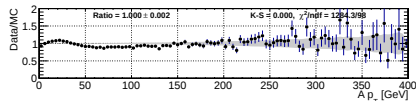
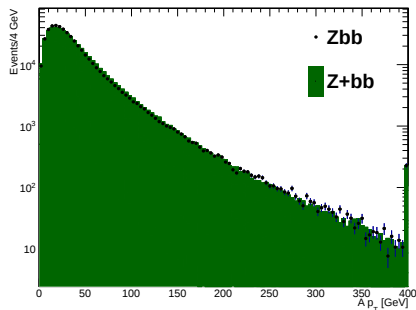
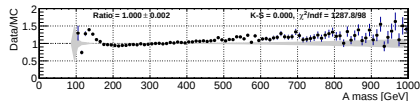
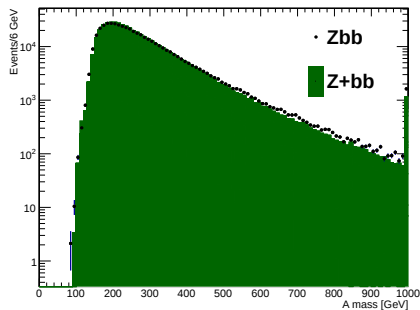


A



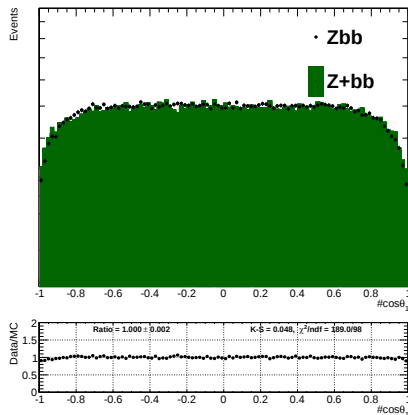
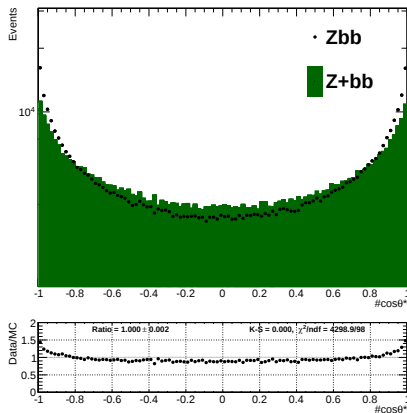


A





Angular



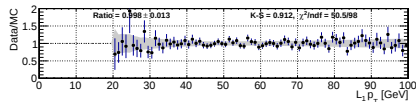
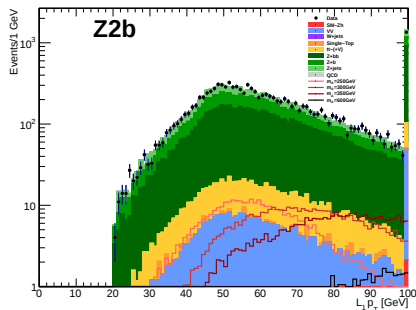
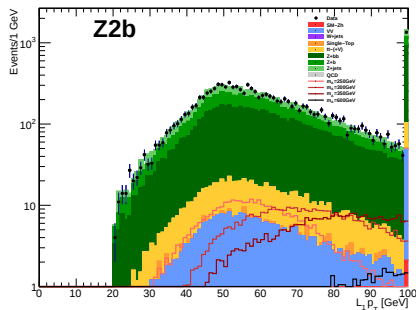


Data vs MC Z2b CR

- Comparing distribution at reco level for each MC vs Data
- **Showing comparison for Zbb control region**
 - ▶ Zbb (New dataset): left column
 - ▶ Z+bb (inclusive DY+jets dataset): right column
- DY+jets is normalized via contro region fit (as usual);
- Zbb is normalized to expected yield of DY+jets;
- for leptonic, hadronic, and lep+had part of the events.

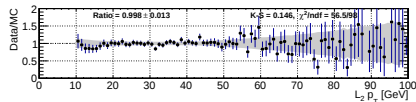
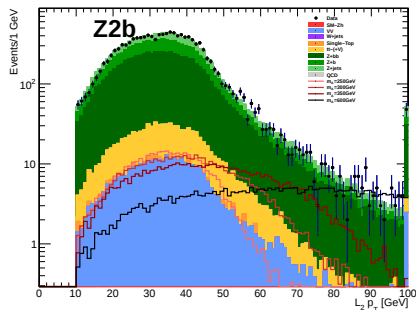
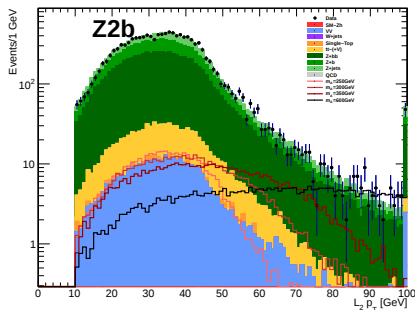


hL1pt Z2b



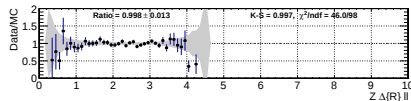
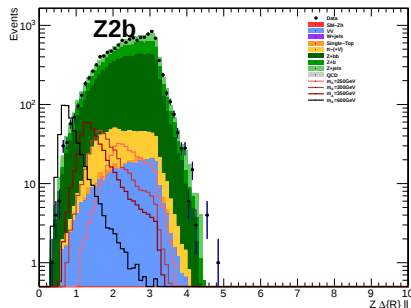
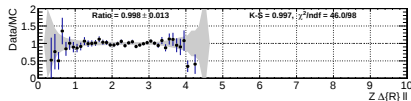
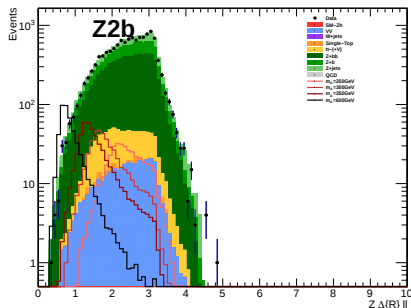


hL2pt Z2b



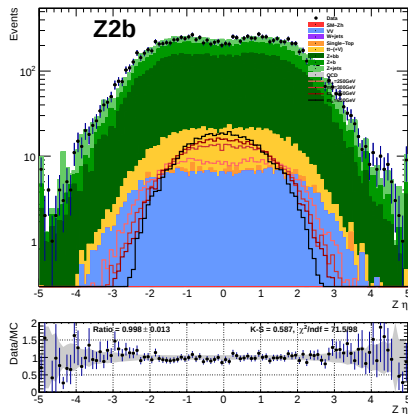
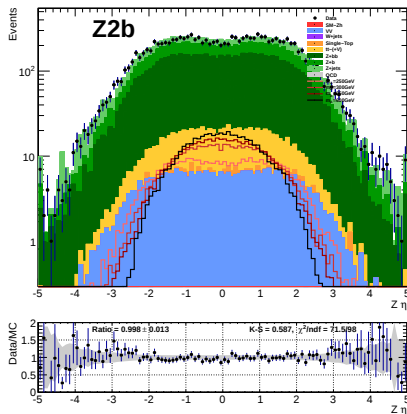


hZdR Z2b



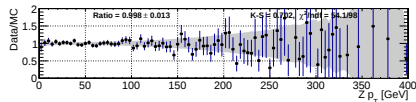
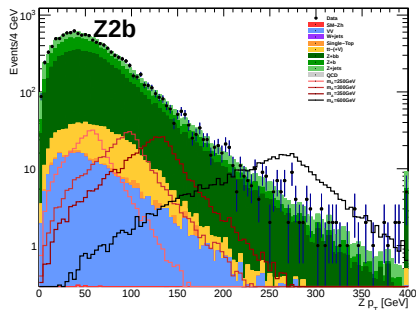
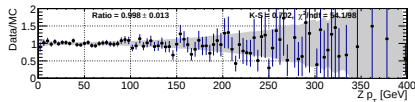
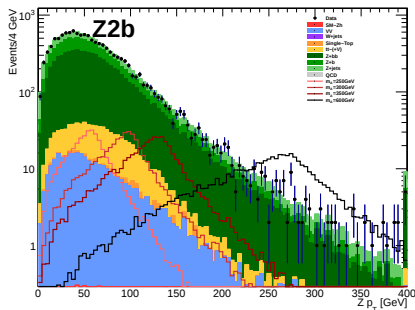


hZeta Z2b



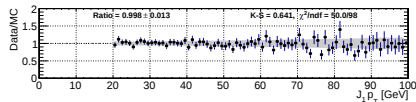
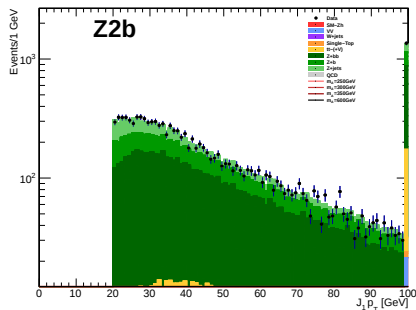
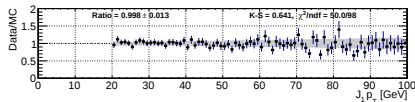
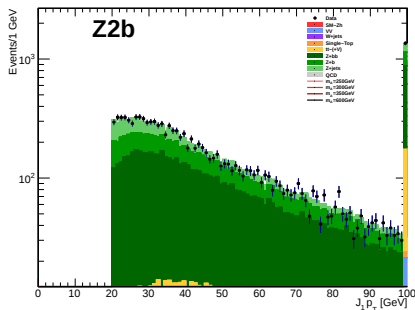


hZpt Z2b



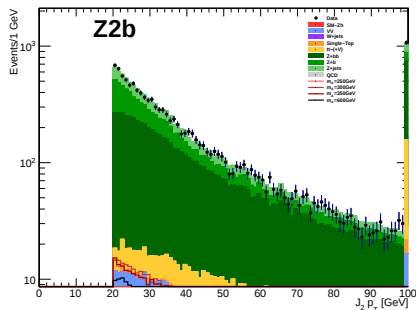
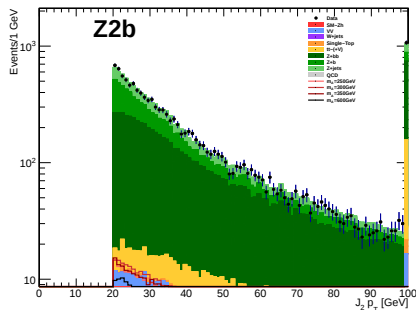


hJ1pt Z2b



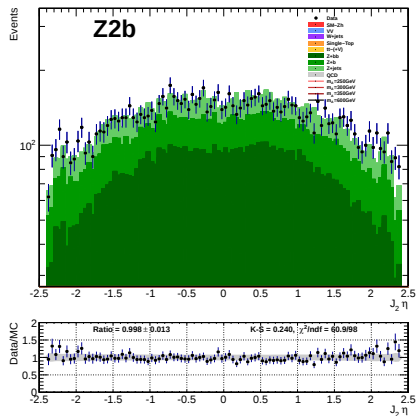
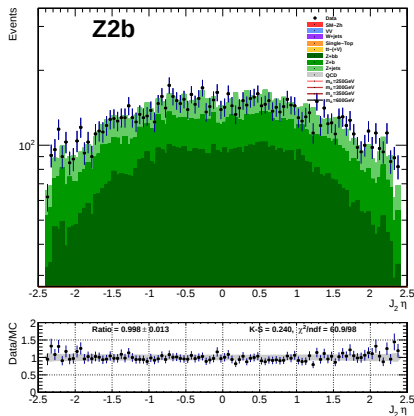


hJ2pt Z2b



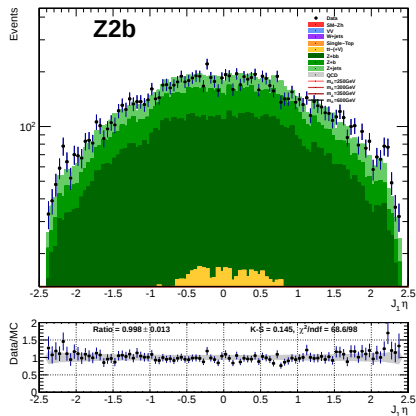
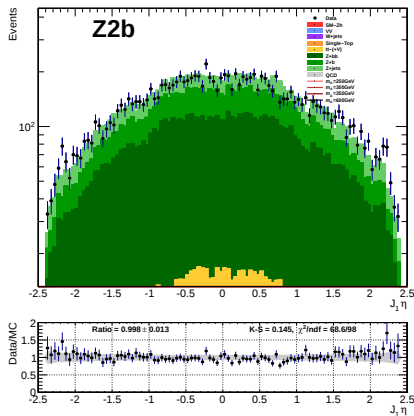


hJ2eta Z2b



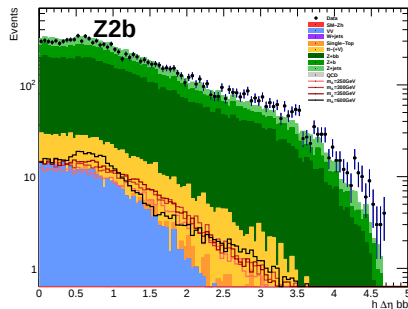
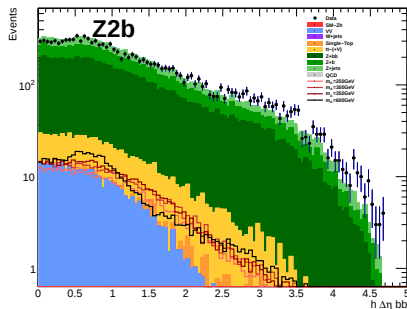


hJ1eta Z2b



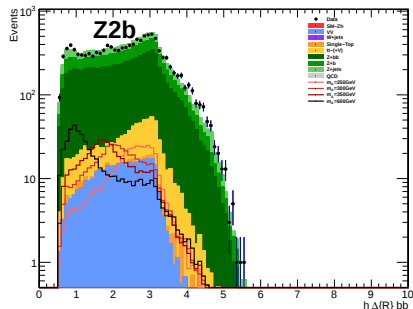
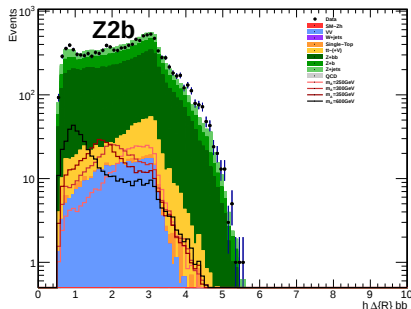


hhdEta Z2b



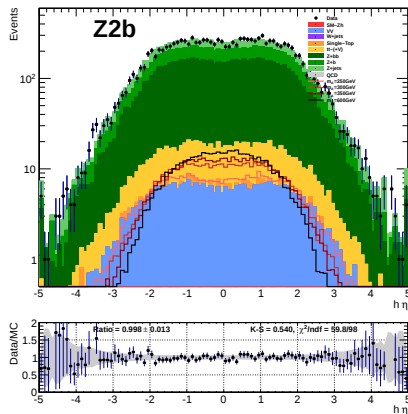
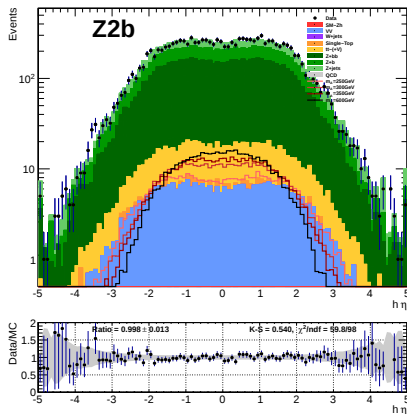


hhdR Z2b



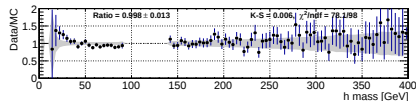
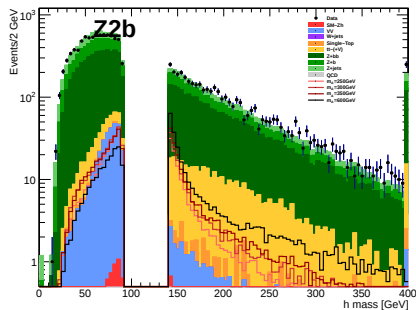
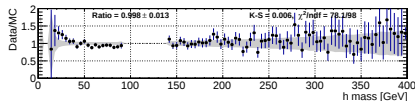
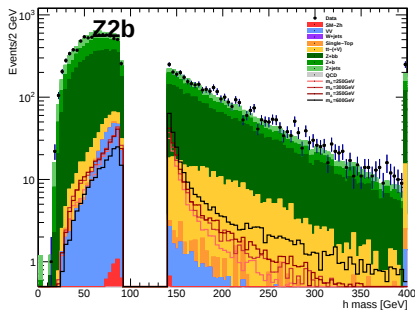


heta Z2b



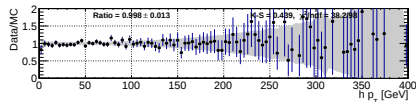
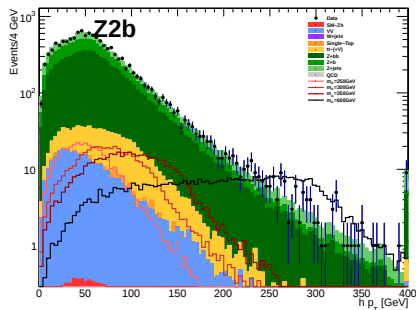
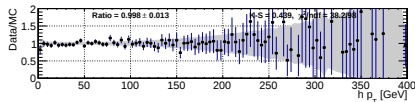
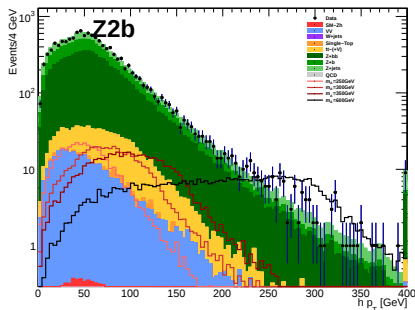


hhmass Z2b



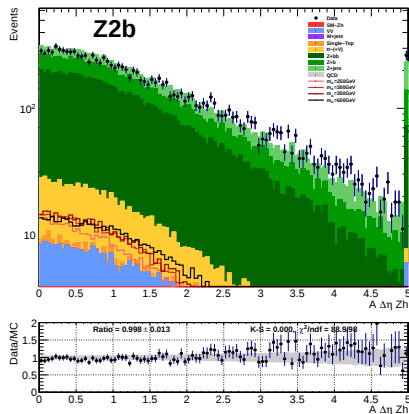
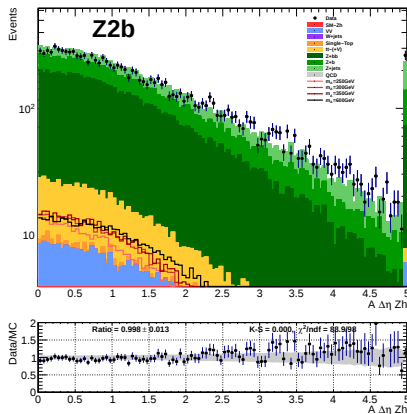


hhpt Z2b



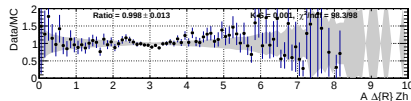
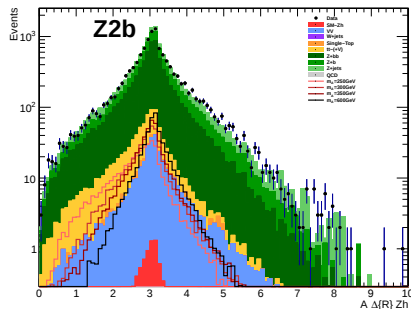
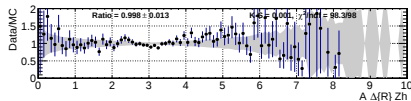
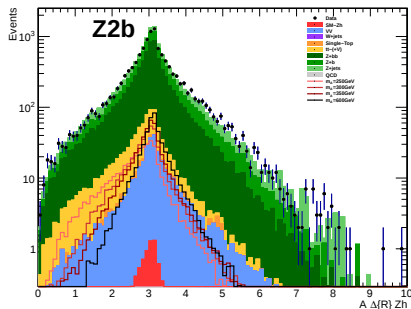


hAdEta Z2b



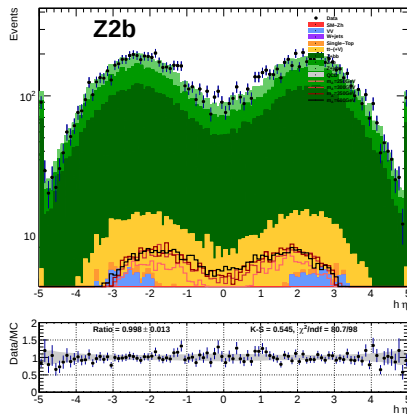
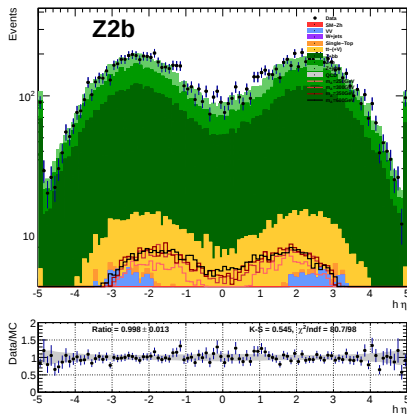


hAdR Z2b



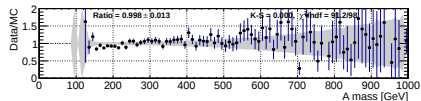
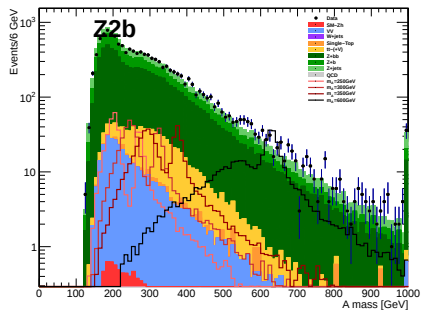
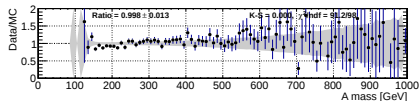
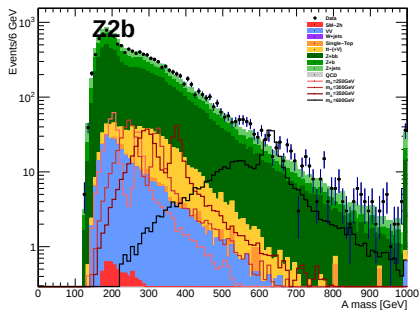


hAeta Z2b



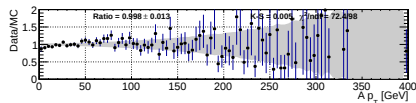
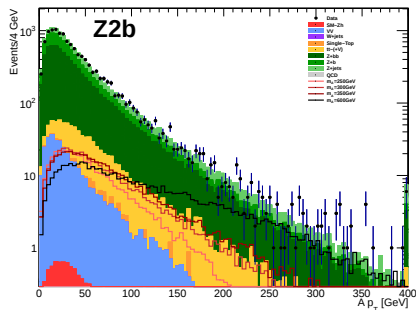
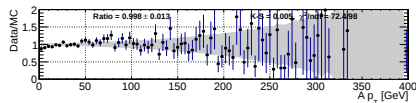
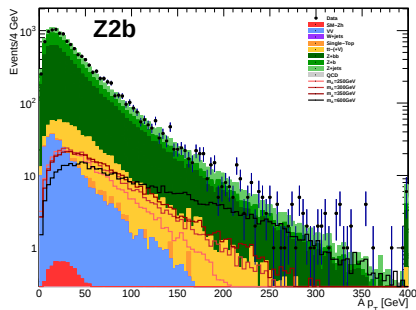


hA_{mass} Z2b



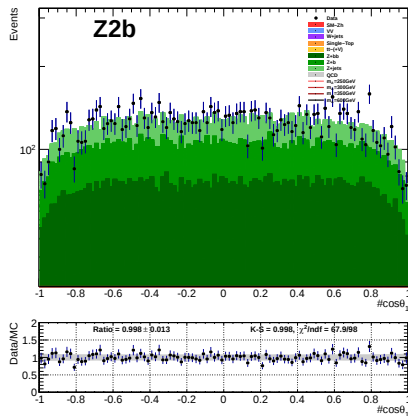
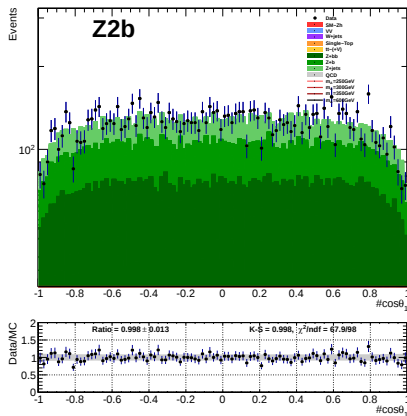


hApt Z2b



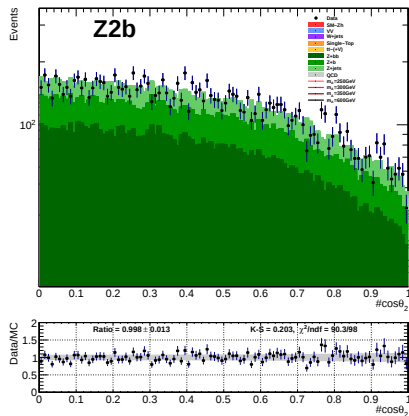
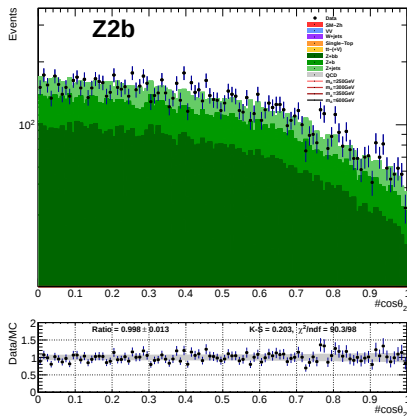


hCosTheta1 Z2b



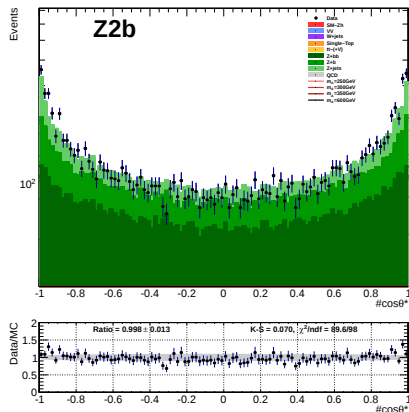
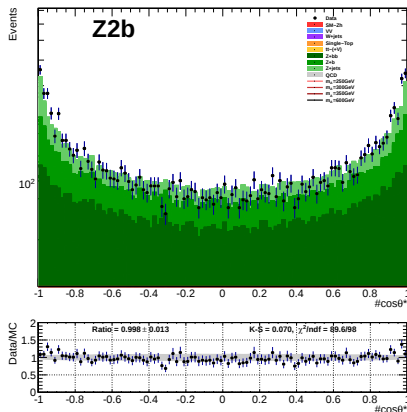


hCosTheta2 Z2b



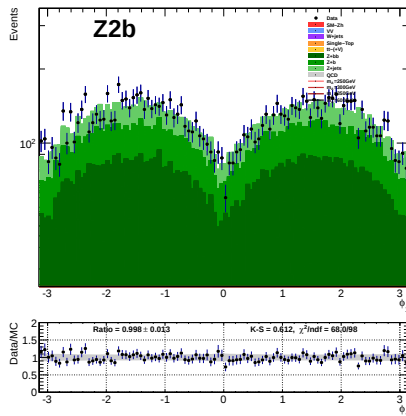
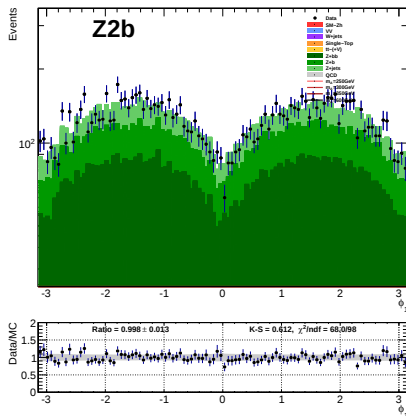


hCosThetaStar Z2b



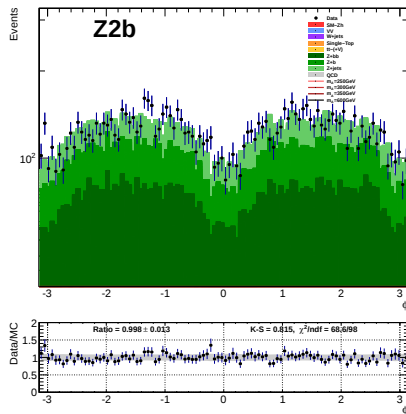
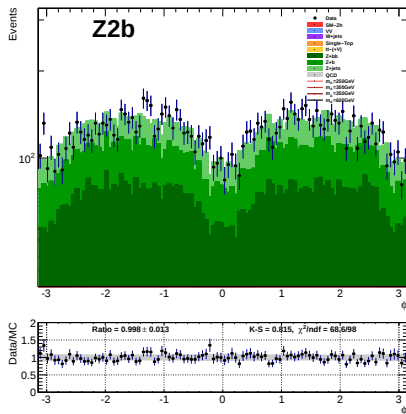


hPhi1 Z2b





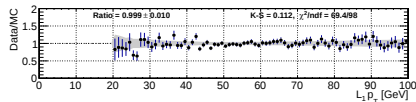
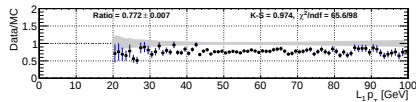
hPhi Z2b





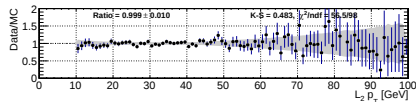
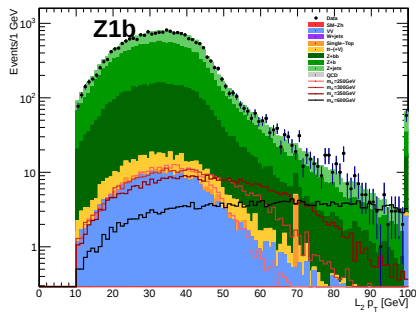
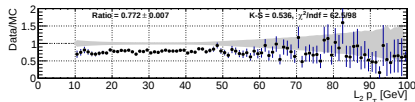
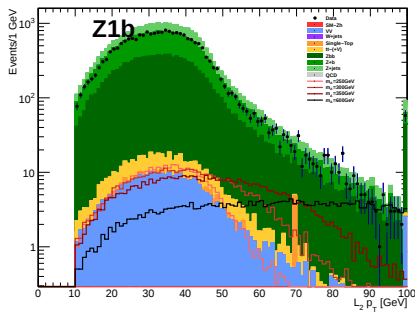
Data vs MC Z1b CR

- Comparing distribution at reco level for each MC vs Data
- Showing comparison for **Z1b** control region
 - ▶ Zbb (New dataset): left column
 - ▶ Z+bb (inclusive DY+jets dataset): right column
- DY+jets is normalized via contro region fit (as usual);
- Zbb is normalized to expected yield of DY+jets;
- Normalization as before, but too much MC in Zbb!
- I suspect double counting for MC Z1b due to different B-partons thresholds



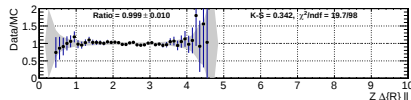
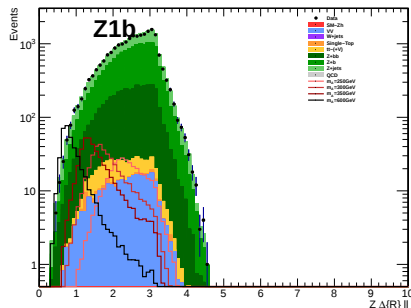
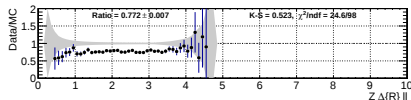
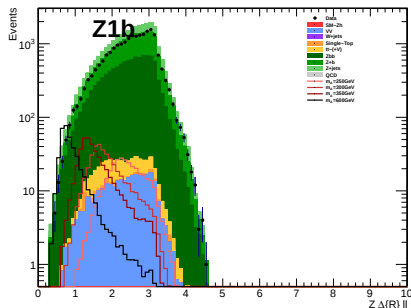


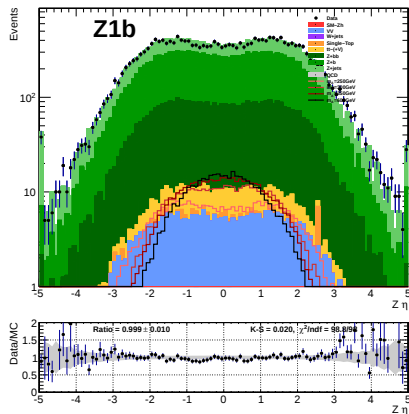
hL2pt Z1b





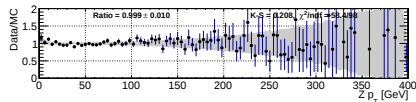
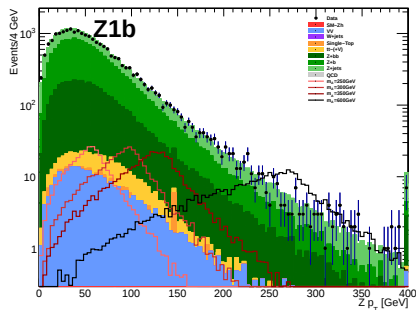
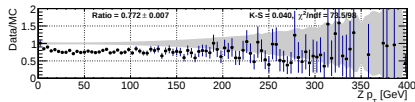
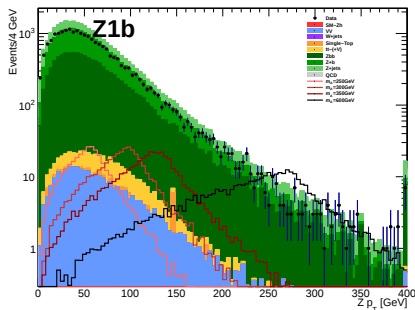
hZdR Z1b





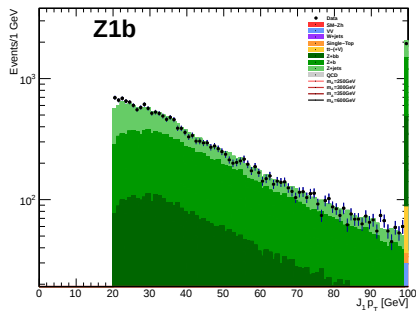
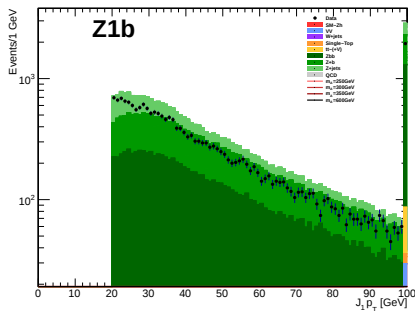


hZpt Z1b



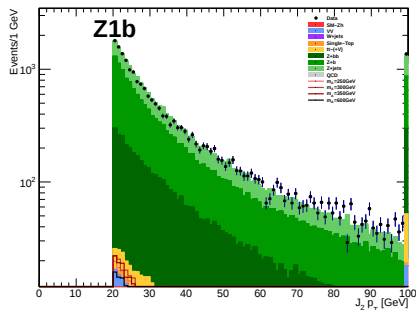
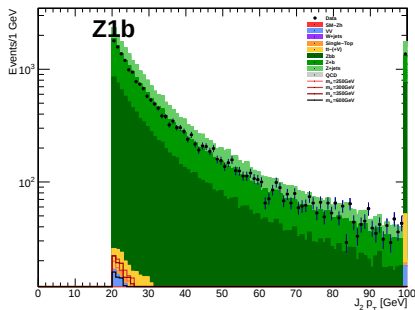


hJ1p_T Z1b



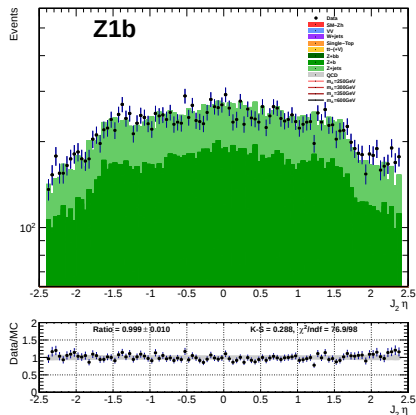
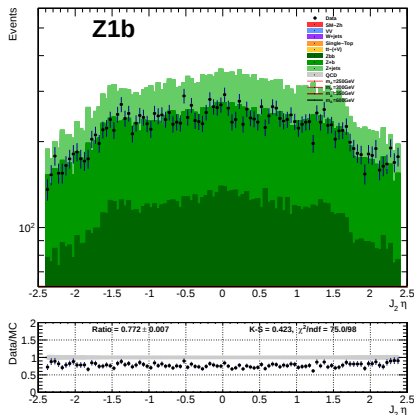


hJ2pt Z1b



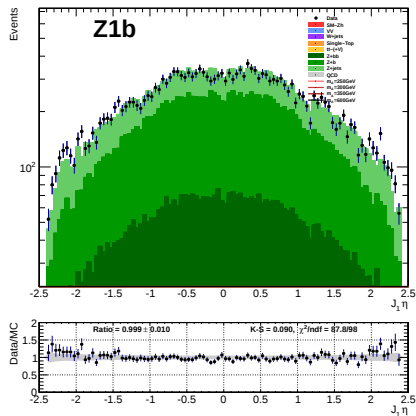
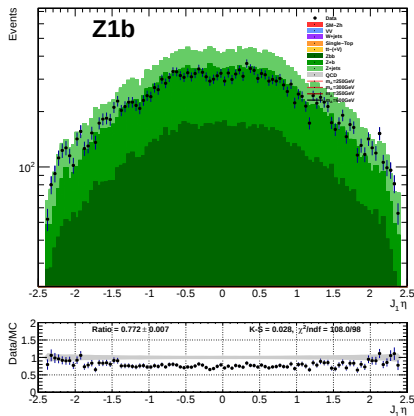


hJ2eta Z1b



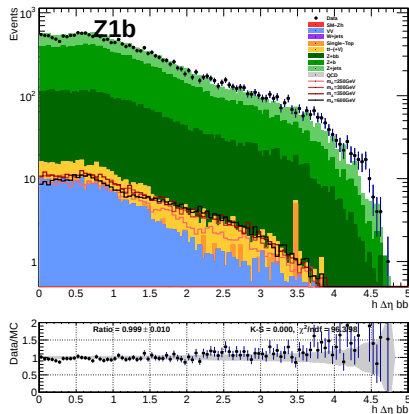
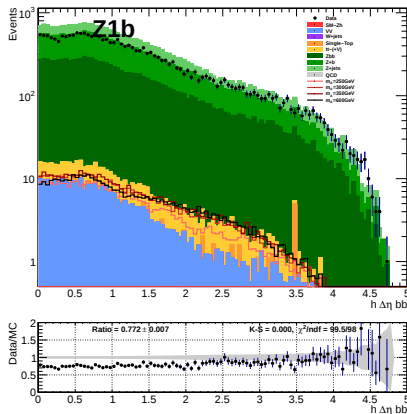


hJ1eta Z1b



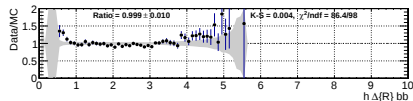
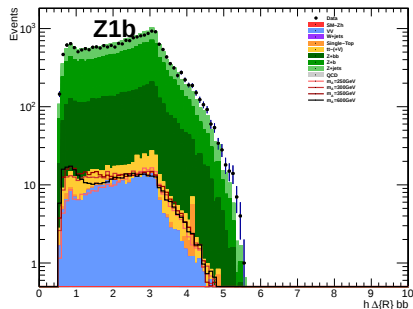
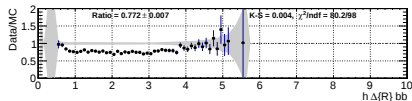
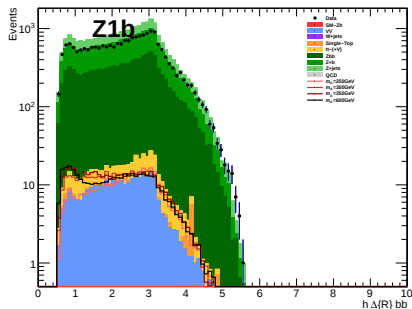


hhdEta Z1b



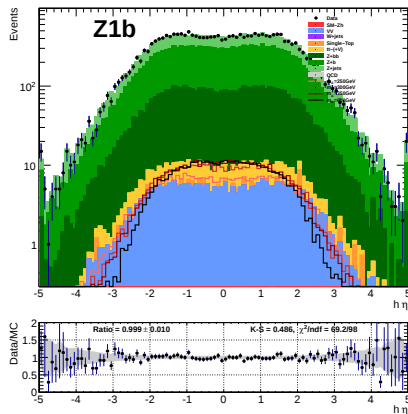
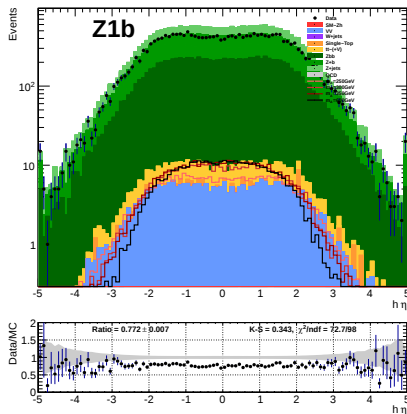


hhdR Z1b



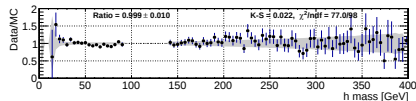
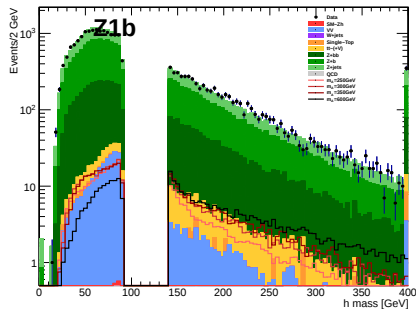
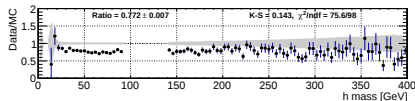
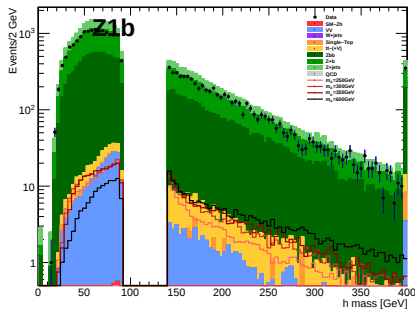


h η Z1b



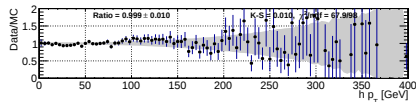
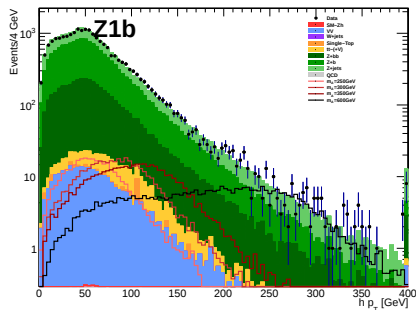
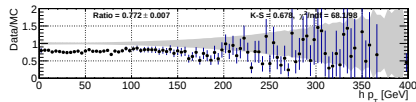
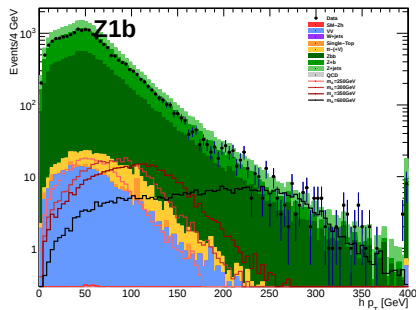


hhmass Z1b



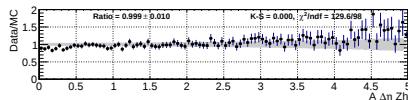
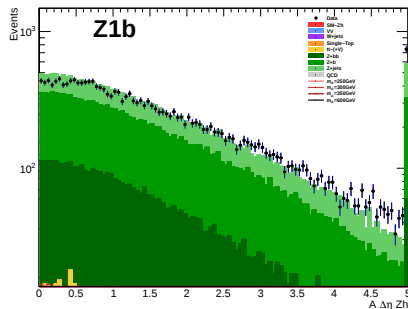
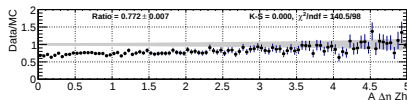
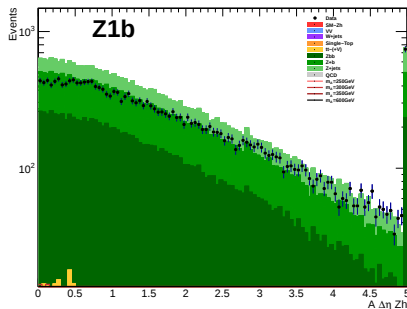


hhpt Z1b



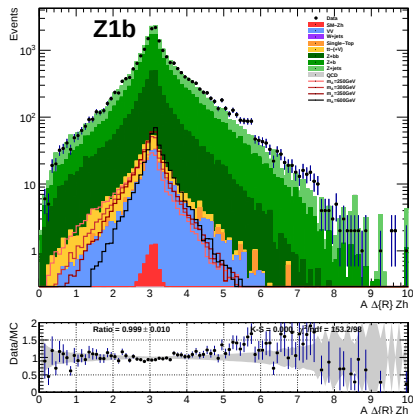
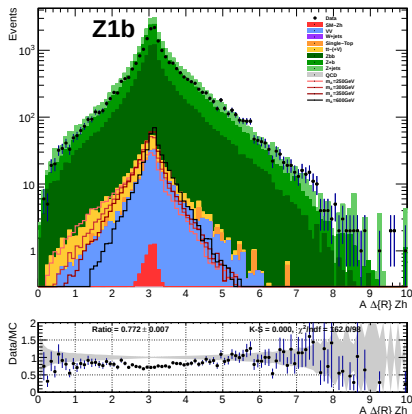


hAdEta Z1b



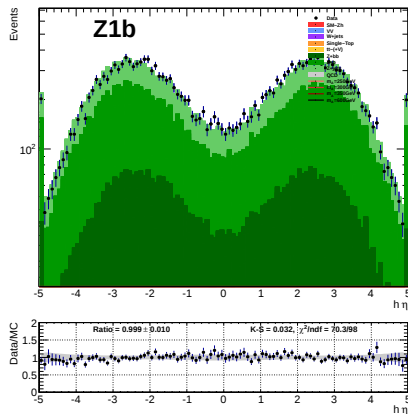
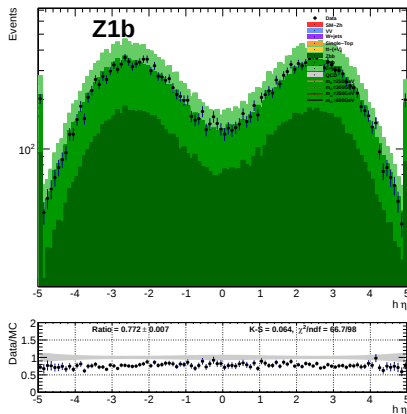


hAdR Z1b



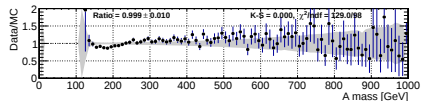
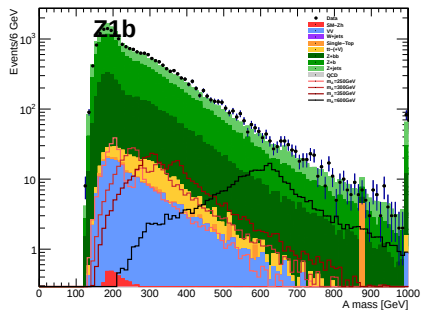
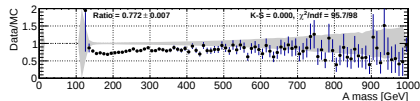
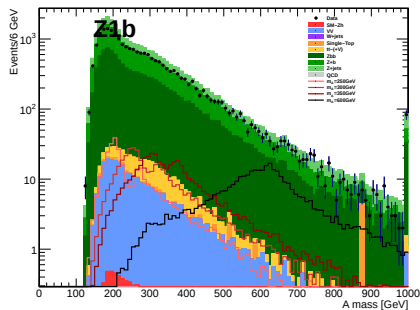


hAeta Z1b



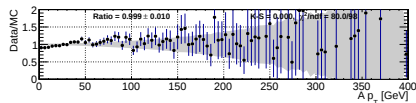
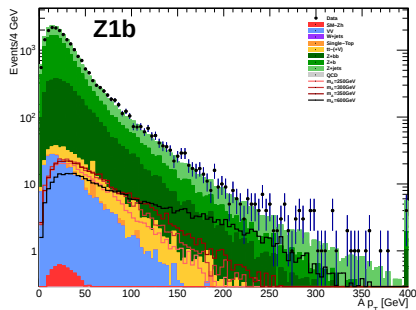
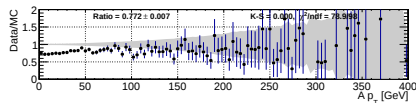
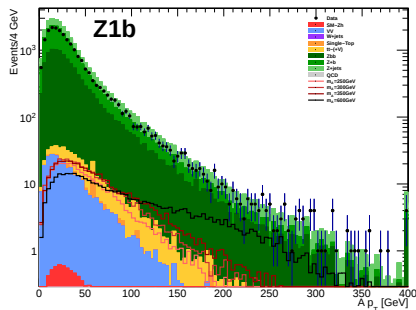


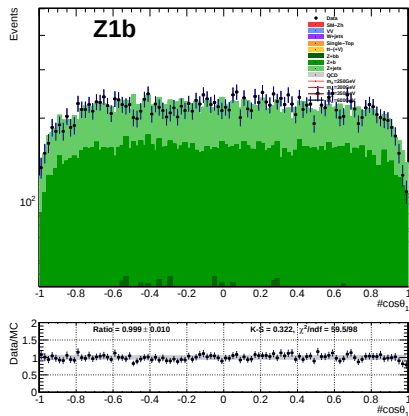
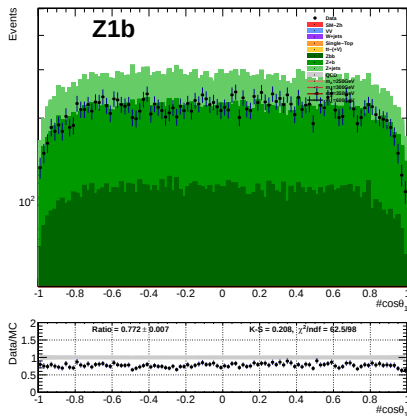
hA_{mass} Z1b





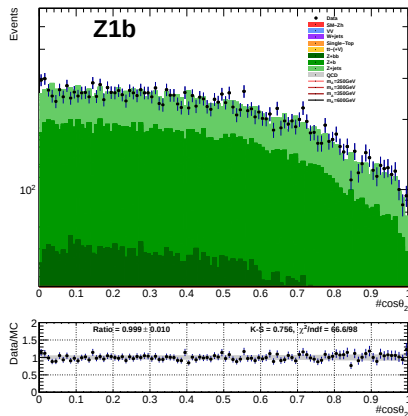
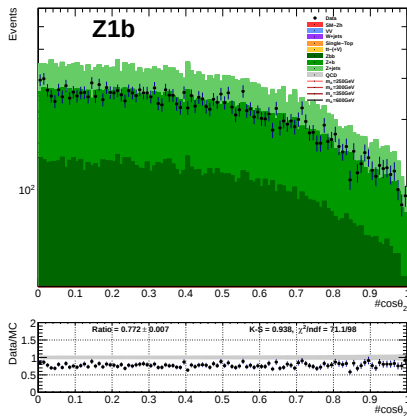
hApt Z1b





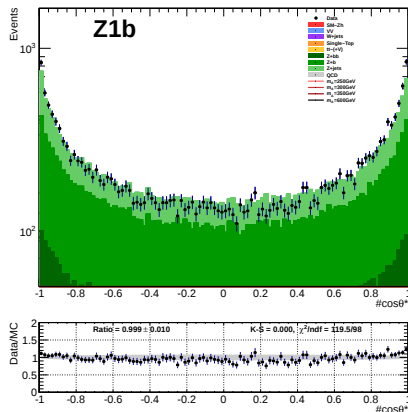
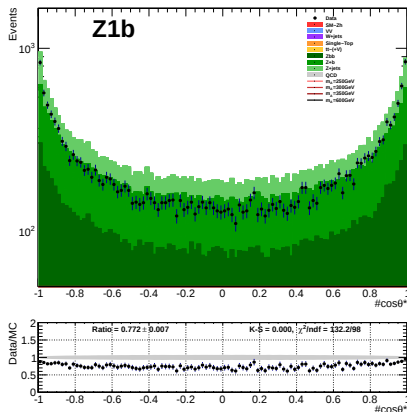


hCosTheta2 Z1b



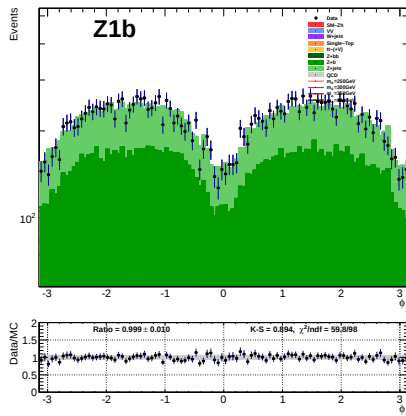
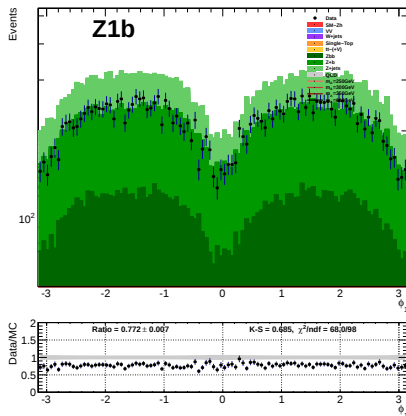


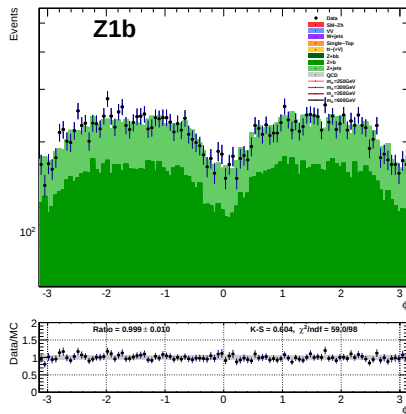
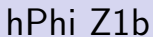
hCosThetaStar Z1b





hPhi1 Z1b







Conclusion

- At preselection level, the difference of the two dataset is important, especially for the lower part of the jets p_T spectrum.
- However, the comparison with data in the Zbb control region shows a very good agreement for both dataset
- the selection for the Z2b CR reduces any difference in the Jet p_T spectrum to a negligible amount;
- on the other hand, any difference present at preselection level is hidden underneath the huge DY0b;
- Same comparison on Z1B CR show no significant discrepancies in shape, either
- but a problem in normalization, likely due to double counting in Z+1b sample definition, due to lack of b-parton threshold in Zbb sample



Conclusion (II)

- In summary, we can use either dataset for the signal region as far as the distribution are concerned
- the issue in normalization suggests to stay with the actual DY inclusive/exclusive dataset
- Moreover, in the SR: Zbb has 18 988 events, DY+bb has 38 956
- maybe use the new Zbb one to possibly study MC systematics which might be already under control from the comparison of data and MC in the Z2b CR.