Response to CMS Questions on Z→4l Analysis

Bing Zhou
The University of Michigan

1) *Looks like CMS and ATLAS are using very similar cuts (60-120 vs 66-116), but predicted x-section is 7.7 vs 7.2. How to compare ATLAS results with NNLO (1405.2219v2)?*

→ This is the question for on-shell ZZ analysis. For Z→4l analysis, we required m4l mass to be in a window [80, 100] GeV

2) *Your isolation cone 0.2 is very small compare to CMS 0.4, and the cut 0.15 on isolation seem to be very loose - how did you optimize your choice?*

→In Z→4l analysis we used isolation cone size 0.2 for both tracker and calorimeter isolation requirements.

The cuts for e/m are <0.30 for calorimeter isolation, and <0.15 for tracker-isolation requirements; at 8 TeV for electrons, the calo-isolation for electrons is tightened to < 0.2; for standalone muons, the tracker isolation cut is <0.15. These cuts were tuned to maximize signal/background ratio for H→ZZ*→4l detections

3) *The fiducial volume is defined with the requirement that any two leptons are separated by dR>0.2. In event selection you reject only electrons which are closer then dR<0.1 to any muon or another electron (which has higher pT). Occasionally there will be signal events where two of the leptons are within the distance of 0.1<dR<0.2 (and hence not in the fiducial region), but will contribute to the selected events. What is the reason for this difference in the events selection and fiducial volume definition? How is this component treated in the fit (e.g. this component is a priori unknown - can be different in case of some BSM models)*

In $Z \rightarrow 4l$ analysis, we have applied the same lepton separation requirements in fiducial selection and in event selection:

$\Delta R(e,\mu) > 0.1$, and $\Delta R(e,e)$ and $\Delta R(\mu,\mu) > 0.2$

4) *How do you treat events with 5 or more leptons (in the event selection and fiducial volume definition)?*

→ We allow more than 4 leptons in an event as long as we can select a 'Quadralet' with two pairs of the same-flavor opposite-charge leptons and with their mass satisfying our di-lepton mass requirement (m_{12} > 20 GeV, m_{34}>5 GeV) only these selected four leptons are used in the analysis. In fact, at the end, we only have one 4muon events with 5 muons in this event.

5) *How much does analysis gain by using eta region 2.5-2.7 and special treatment for eta<0.1 ?*

→Our analysis strategy is to maximize the signal (particularly Higgs→4l) acceptance. Our MC simulations show that the gain for H→4muon detection is about 10% from eta region 2.5-2.7.

6) *Did you try di-lepton trigger? Does you single lepton trigger include isolation? Is it similar to the offline one?*

→ We have used both single and di-lepton triggers as shown below; the single muon trigger has some loose isolation requirement, not the same as used in offline

Table 2: Triggers used to select data. In each data taking period, the OR of single- and di-lepton triggers of one flavor is used to select $4e$ and 4μ events. The $2e2\mu$ channel uses the OR of all triggers.

Period	Electron triggers Single-e	Di-e	Muon triggers Single-μ	Di-μ	Electron-Muon triggers
2011 B-I	e20_medium	2e12_medium	mu18_MG	2mu10_loose	EF_e10_medium_mu6
2011 J	"	"	mu18_MG_medium	"	"
2011 K	e22_medium	2e12T_medium	"	"	"
2011 L-M	e22vh_medium1	2e12Tvh_medium	"	"	"
2012	e24vhi_medium1 OR e60_medium1	2e12Tvh_loose1 OR 2e12Tvh_loose1_L2StarB	mu24i_tight OR mu36_tight	2mu13 OR mu18_tight_mu8_EFFS	e12Tvh_medium1_mu8 OR e24vhi_loose1_mu8

Table 3: Trigger efficiency for $Z \to 4\ell$ signal events (selected as described in Section 6) evaluated with MC.

$\sqrt{s}$	$eeee$	$ee\mu\mu$	$\mu\mu ee$	$\mu\mu\mu\mu$
7 TeV	99.057%	95.958%	95.738%	95.009%
8 TeV	98.073%	98.208%	94.146%	94.063%

7) *In the paragraph on the reconstruction acceptance factors (CZZ) and its uncertainties it is stated: "The uncertainties are estimated by varying the data-driven correction factors applied to simulation by their systematic and statistical uncertainties." Can you please elaborate on this?*

→If we have n sources for uncertainties for lepton IDs, energy/momentum scale/resolutions, isolations, triggers…We basically re-run MC 2n times (+- 1σ) of event selections to determine the fractional efficiency changes. The final quoted ΔC/C for each channel is evaluated by sum over all the fractional changes quadratically.

8) The theoretical uncertainties on the correction factors and acceptances (CZZ and AZZ) are very different between CMS and ATLAS in case of "PDF & Scale" uncertainties (and similar in case of the "MC Generator Difference"). Can you elaborate how exactly "PDF & Scale" uncertainties are computed?

→ We vary scales (μR, μF) from 0.5 to 2 respect to minimal scale (m_{4l}) independently and determine the scale uncertainty by sum over uncertainties quadratic ally; MC events are produced with different scales for this study. For PDF, we count for the differences between different set of PDF (CT10, MSTW) and the uncertainties from CT10 PDF Eigen vectors (at 68%)

$$\sigma^{+} = \sigma^{-} = \frac{\sqrt{\sum_{i=0}^{N}[max(A_i - A_0, 0)]^2} + \sqrt{\sum_{i=0}^{N}[max(A_0 - A_i, 0)]^2}}{2A_0}$$

where A_0 is the acceptance predicted with the nominal CT10 PDF set, A_i is the acceptance predicted with the i^{th} eigenvector memeber in the CT10 PDF set, N is the total number of eigenvector memebers in the CT10 PDF set. The uncertainty produced through Eq 4 is essentially the average of the average upward fluctuation and average downward fluctuation in acceptance predicted with different eigenvector memebers. In addition to uncertainty due to different eigenvector memebers in the CT10 PDF set, the

9) *The fake factors are computed in a Z+"lepton-like jet" sample, with requirement 20 GeV around the Z mass. This sample will contain events Z+"true electron from asym. Photon conversion", where this additional lepton will have very high reconstruction efficiency and hence its inclusion will lead to an increase of the measured fake factor. If contribution to the "fakes" from these conversion electrons is different in the sample where the fake factor is measured and in the control sample where it is applied, it can easily lead to over/under-estimate of the "fakes" background. Did you explicitly check how large is this effect? (Comment: in CMS we require +/-5GeV mass window around the Z mass to reduce this effect.)*

→ In Z→4l analysis, more than 90% of fake lepton objects come from b decays. To determine the factor-factors, we used both ttbar and Z+jet control samples. For the Z+jets sample, we required Z-mass in +- 15 GeV Z-window. Indeed, the fake-factor from our Z+jets control sample is higher than that from the ttbar control sample.

We use MC to check our signal control regions to find the predicted fraction of b-jet fakes and applied the b-jet fake rate derived from ttbar sample and combined small portion of jet fake rate derived from Z+jets samples (mainly for light jet fakes).

The fake-factor method is cross-checked by simultaneous fit method for background estimations (the default method used in Higgs→4l analysis).

10) The estimation of the uncertainty by comparing the nominal data-driven estimation and the estimation using the average fake factor assumes implicitly the following:
a) that the pT/eta spectra of "lepton-like jets" are identical in the sample where the fake factor is measured and in the control sample where it is applied.
b) that the composition of the sources of fakes are identical in the sample where the fake factor is measured and in the control sample where it is applied. Did you check these assumptions explicitly?

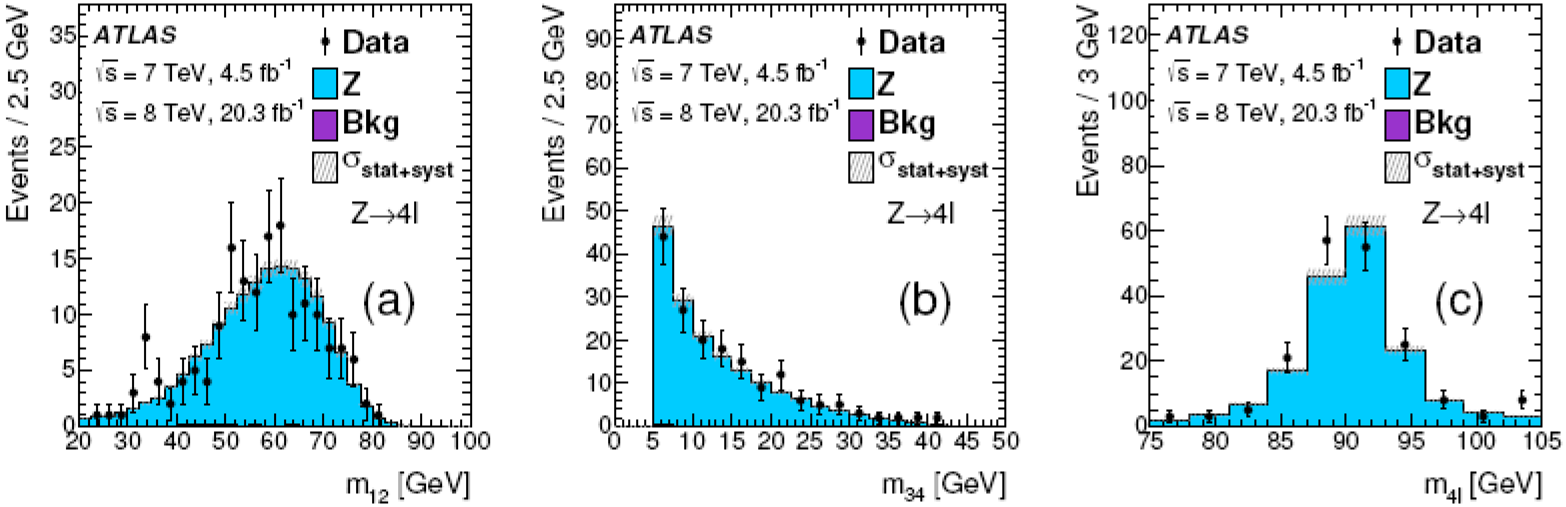
→a) We determine the fake-factor as a function of fake-object pT and eta, as well as for different jet component (b-jet vs. light jet), and cross check with different jet-enriched samples, and taking the difference as systematic uncertainties

b) We did not assume the fake-able lepton object compositions in the signal control regions are the same as the jet-enriched samples. So we checked these composition with MC in the signal region, and derived fake-factors from different jet-enriched samples (i.e. ttbar and Z+jets). Final application of the fake-factor has taken into account the composition of fake-able lepton objects.

11) *Fig.3c has more backgrounds then Fig.3a - why?*

→ Figures you referred?

Below are figures published in our paper



12) How exactly is ZZ->4tau & ZZ->2tau2l contribution treated/subtracted in the analysis? Is the contribution fixed to the SM expected yield and subtracted at the reconstruction level (hence being fixed in the fit), or is it kept proportional to the signal cross section which one fits for?

→ *In Z→4l analysis, leptons decay from ll+tautau or 4 taus are treated as background. For 8 TeV data analysis, we estimated the total number of 4l events selected from ll+tautau and 4 taus is 0.39 eevent compared to total expected 4l events from promptly decays from ZZ, 145. These tau decays events are from full simulations, and using PowHeg+Pythia8 simulation and modeling. We simply subtracted these tau decay events from data when calculating the cross-sections.*

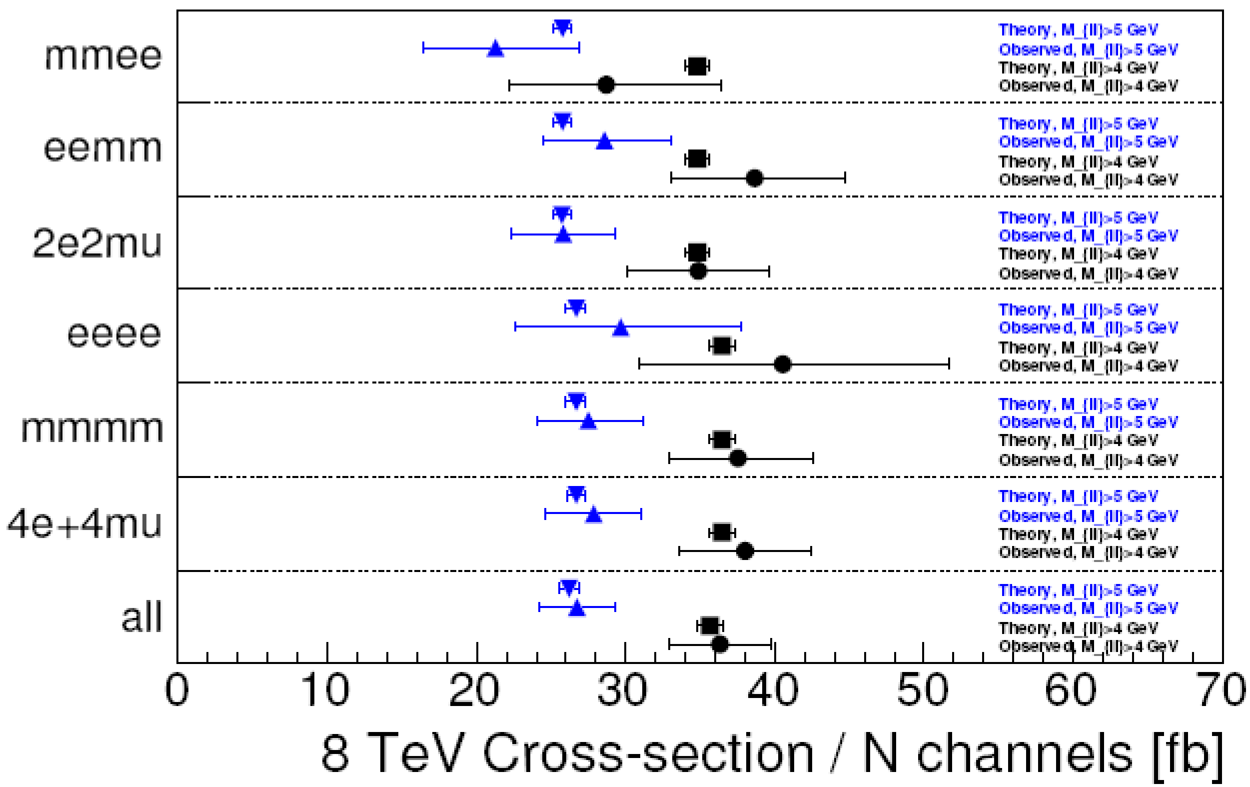
13) *Is the total cross section extracted from the fit simultaneously for all three final states (or separately and then summed up)?*

→ In final phase space, we first determine individual channel cross sections; then we combine 4e with 4m using 2X2 covariance error matrix, the same for the eemm and mmee channels;

→Finally we used 4x4 matrix to combine four channel cross-sections fit (chi-sq) and handle the uncertainties.

→We also performed likelihood fit (-LogL) for individual channels, and combined fit. We obtained very consistent results from two methods

comparisons



14*) Cross section is measured using likelihood fit. Is a binned fit used? Which distribution or distributions were used?*

→ We only used event counting for fit, not use distributions for fitting for cross section measurement