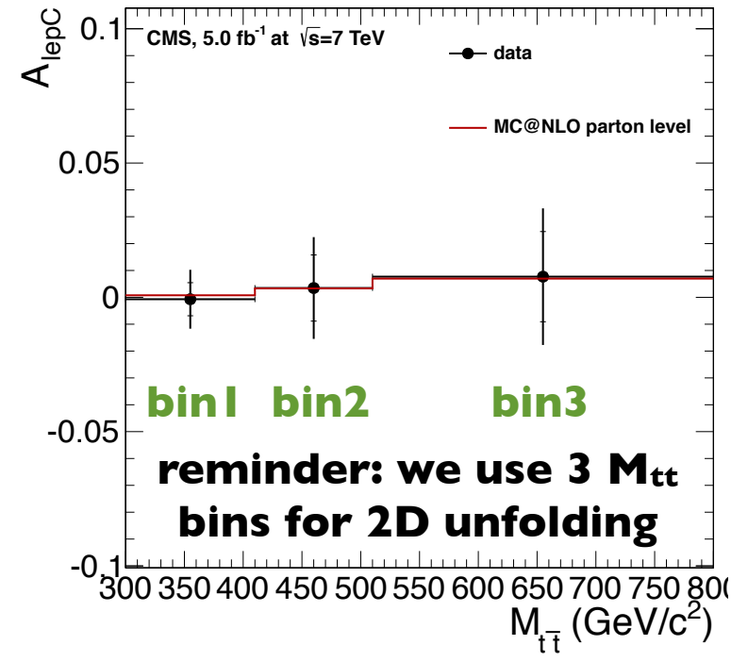
2D unfolding linearity tests

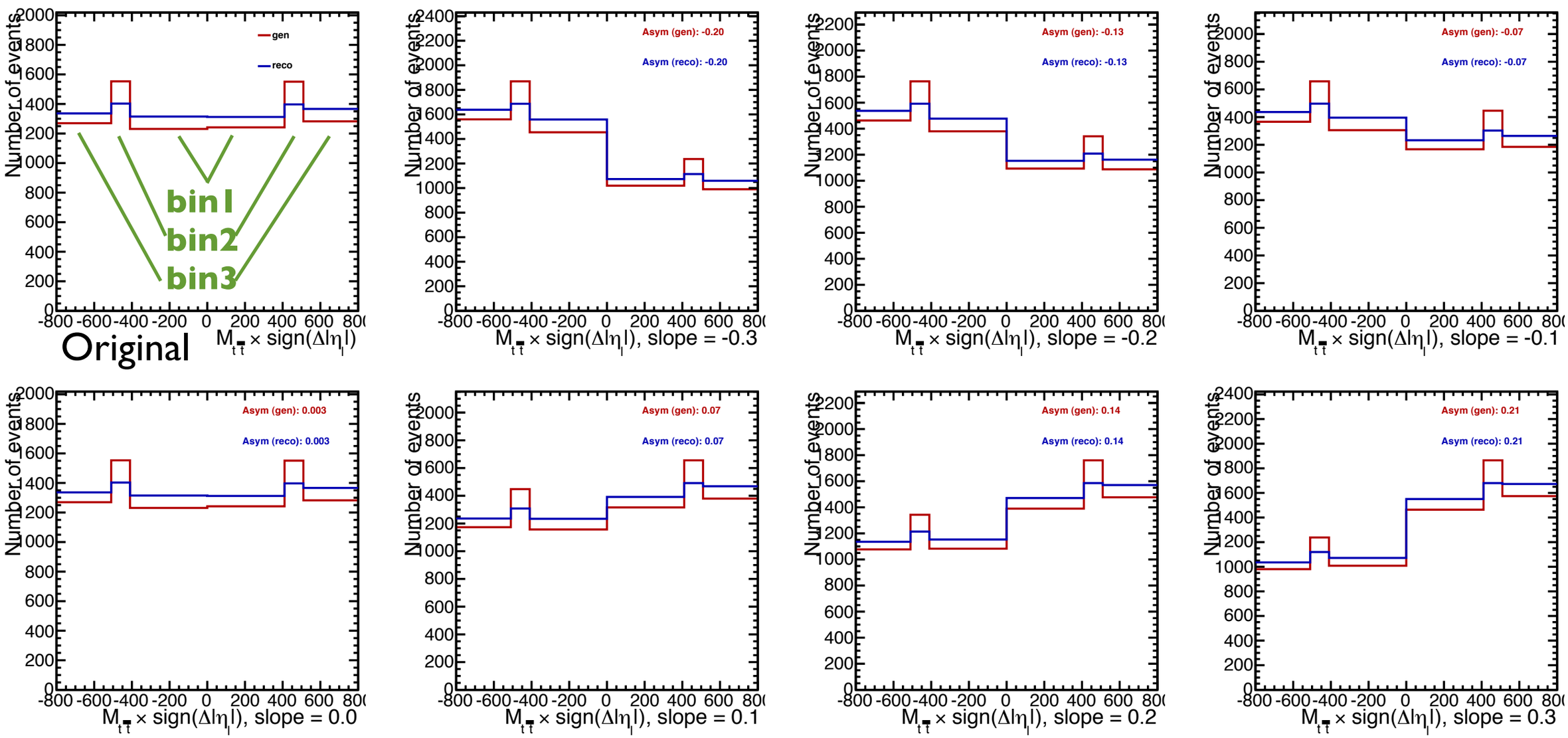
These slides focus on A_{lepC} and A_C vs M_{ttbar}, but the results and conclusions are very similar for the $p_{T,ttbar}$ and y_{ttbar} dependence (and for the polarisation and spin correlation asymmetries too)

Linearity test method

- Add asymmetry by applying weight=1+slope*$\Delta|y_{gen}|$ per event (see plots below), then compare unfolded asymmetry to true asymmetry, inclusively and in the 3 bins we use for the differential measurement



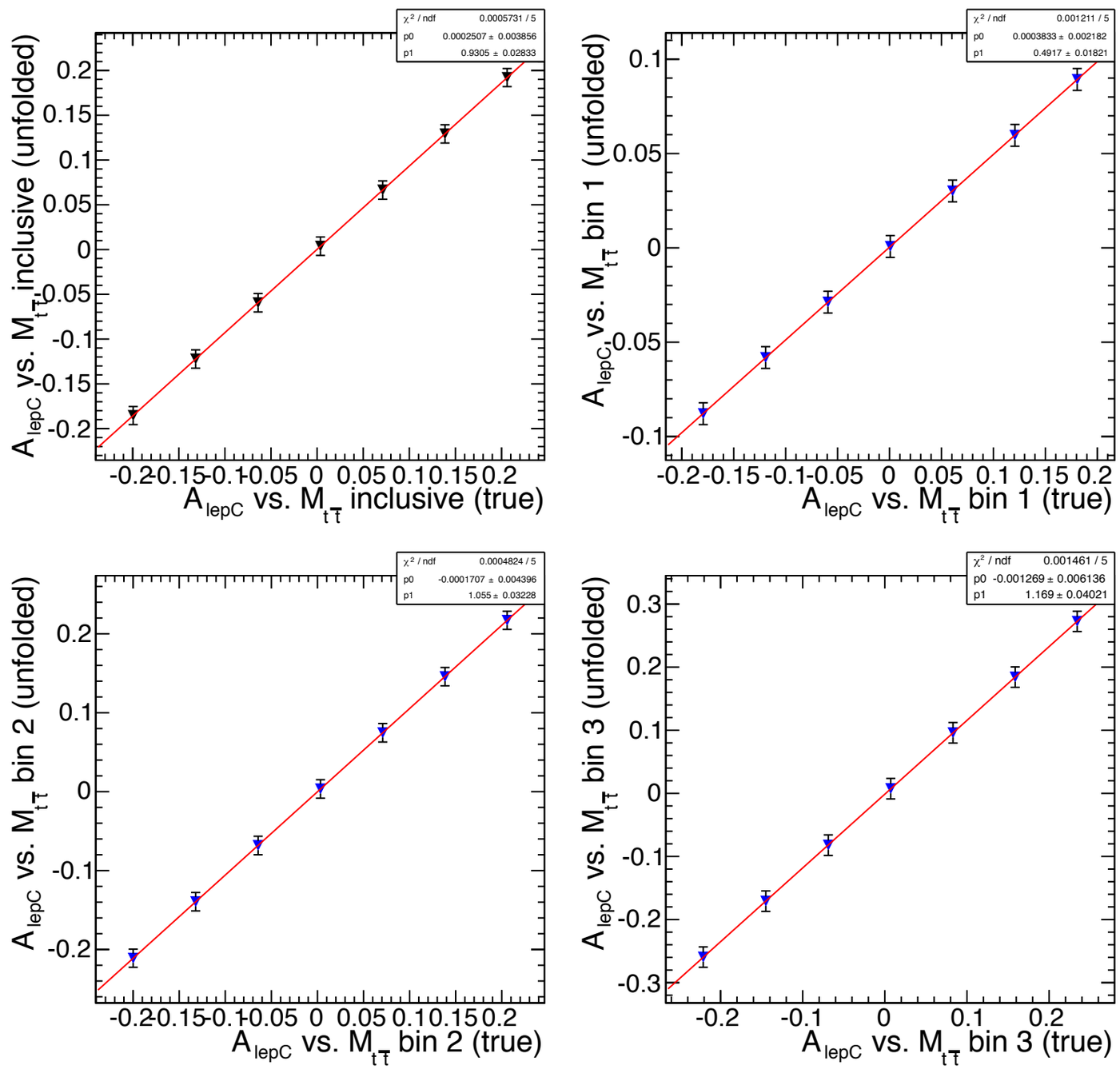
Original and weighted distributions for M_{tt}*sign($\Delta|y_l|$), before unfolding:



Results for A_{lepC} (original)

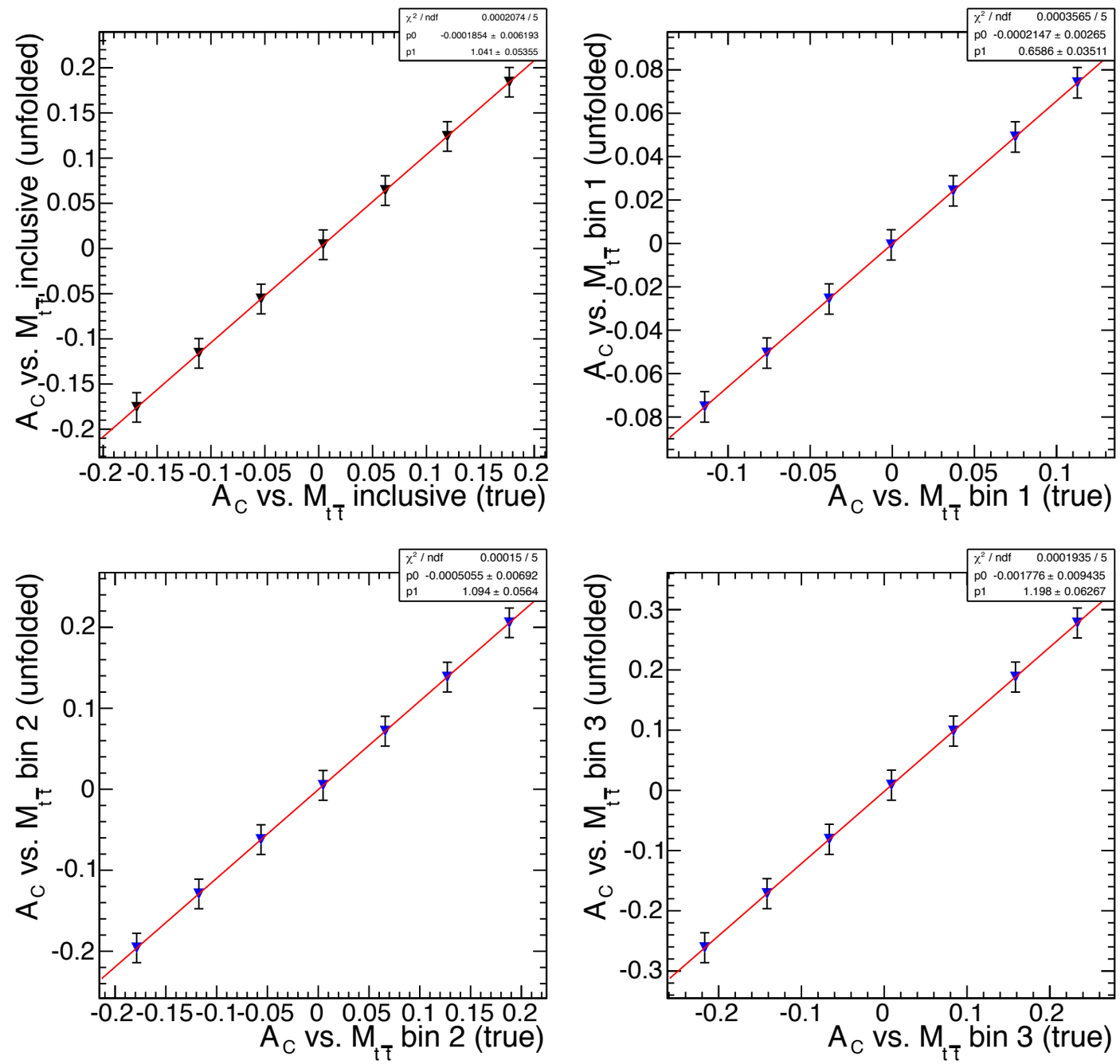
- Slope of linearity plot (p1) significantly away from one for bin1 (0.49) and bin3 (1.17)

- I also checked the pull width for each point, and it's consistently very close to 1



Results for A_C (original)

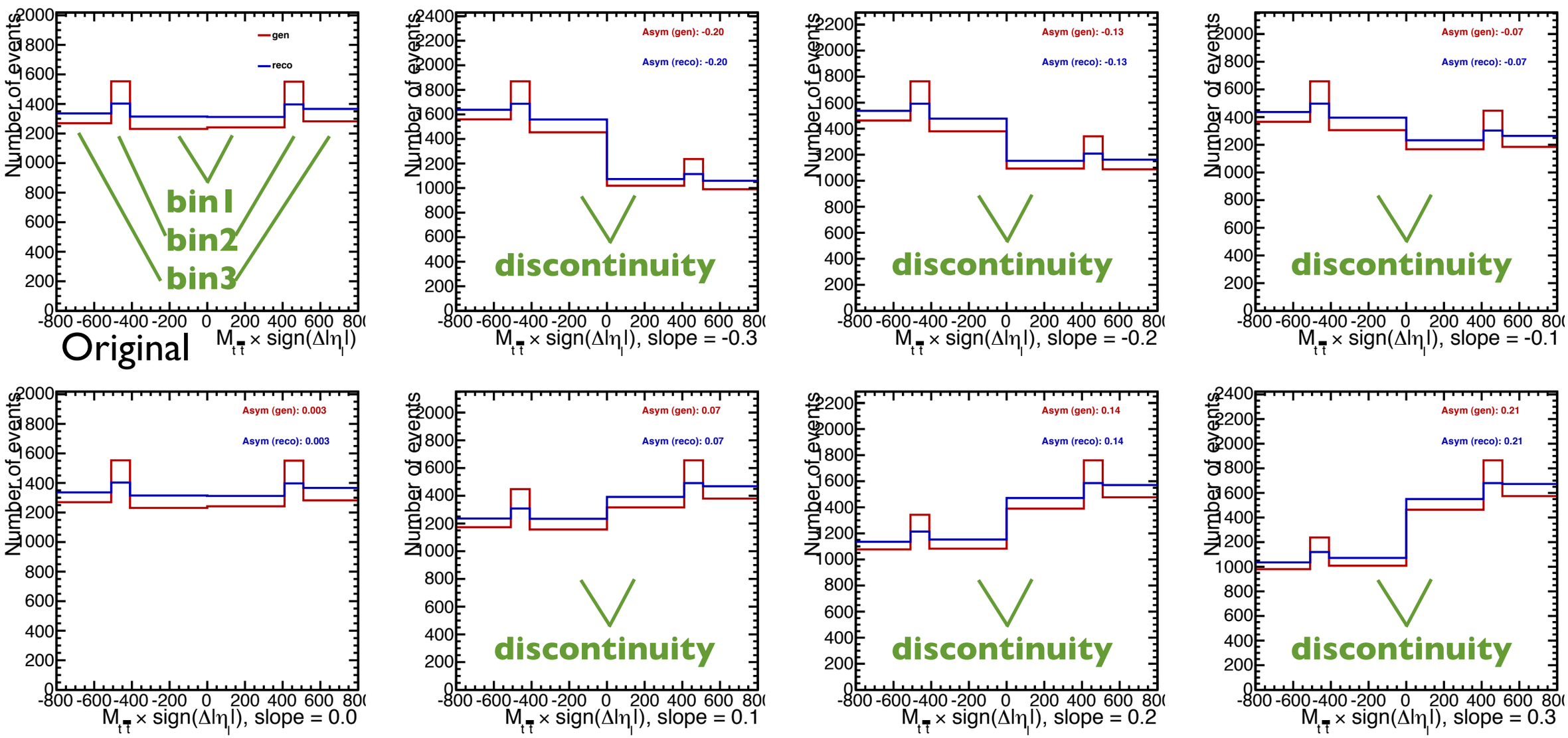
- Slope of linearity plot (p1) significantly away from one for bin1 (0.66) and bin3 (1.20)
- I also checked the pull width for each point, and it's consistently very close to 1



Cause of the bias

- Since the M_{tt}*sign($\Delta|y|$) distribution only has 2 "bins" in $\Delta|y|$, the added asymmetry appears as a step function, with the discontinuity between the two $\Delta|y|$ bins for M_{tt} bin1
- Regularisation reduces the discrepancy between adjacent bins relative to the original distribution
  => the M_{tt} "bin1" asymmetry is reduced
  => "p1"<1 for M_{tt} bin1

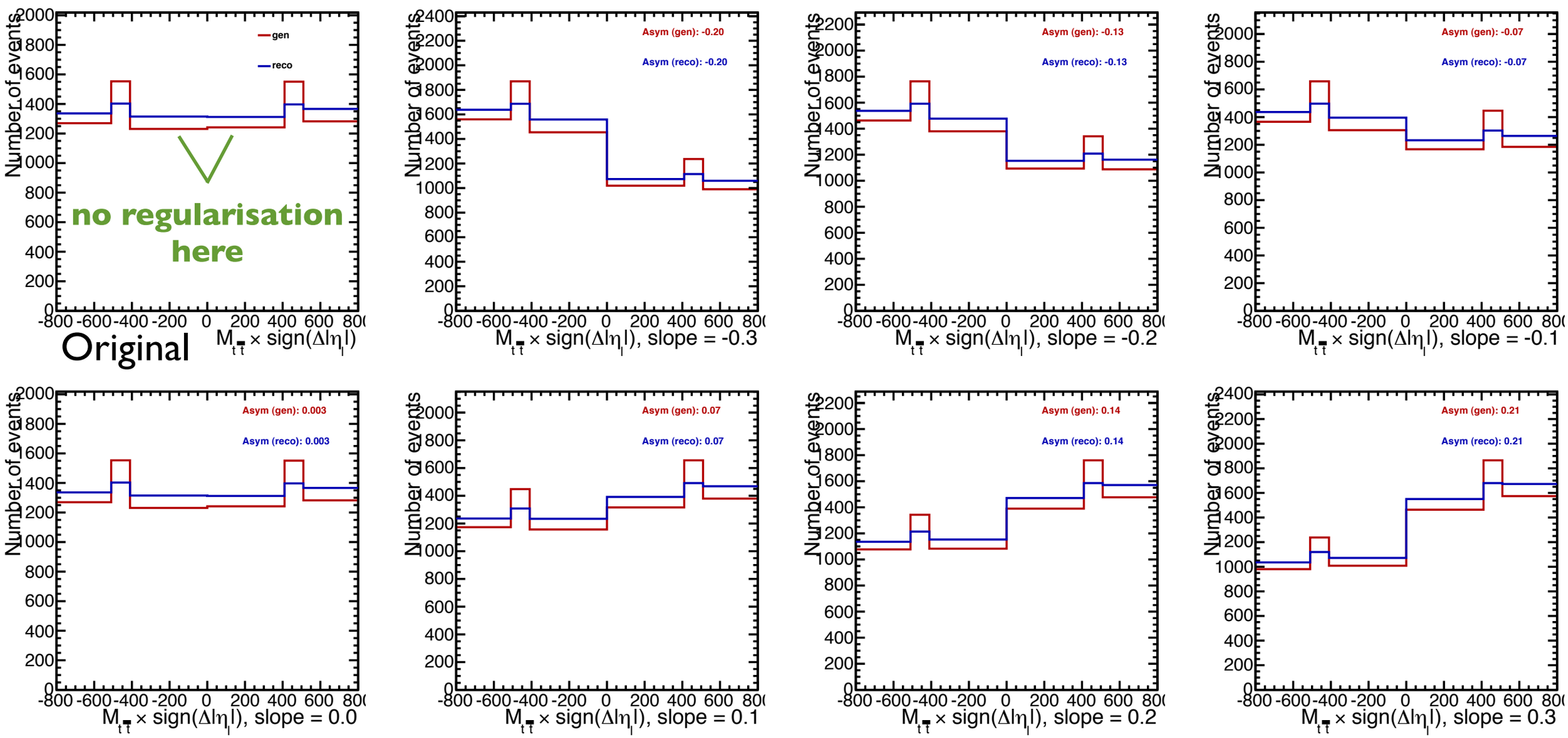
Original and weighted distributions for M_{tt}*sign($\Delta|y_l|$), before unfolding:



Fix for the bias

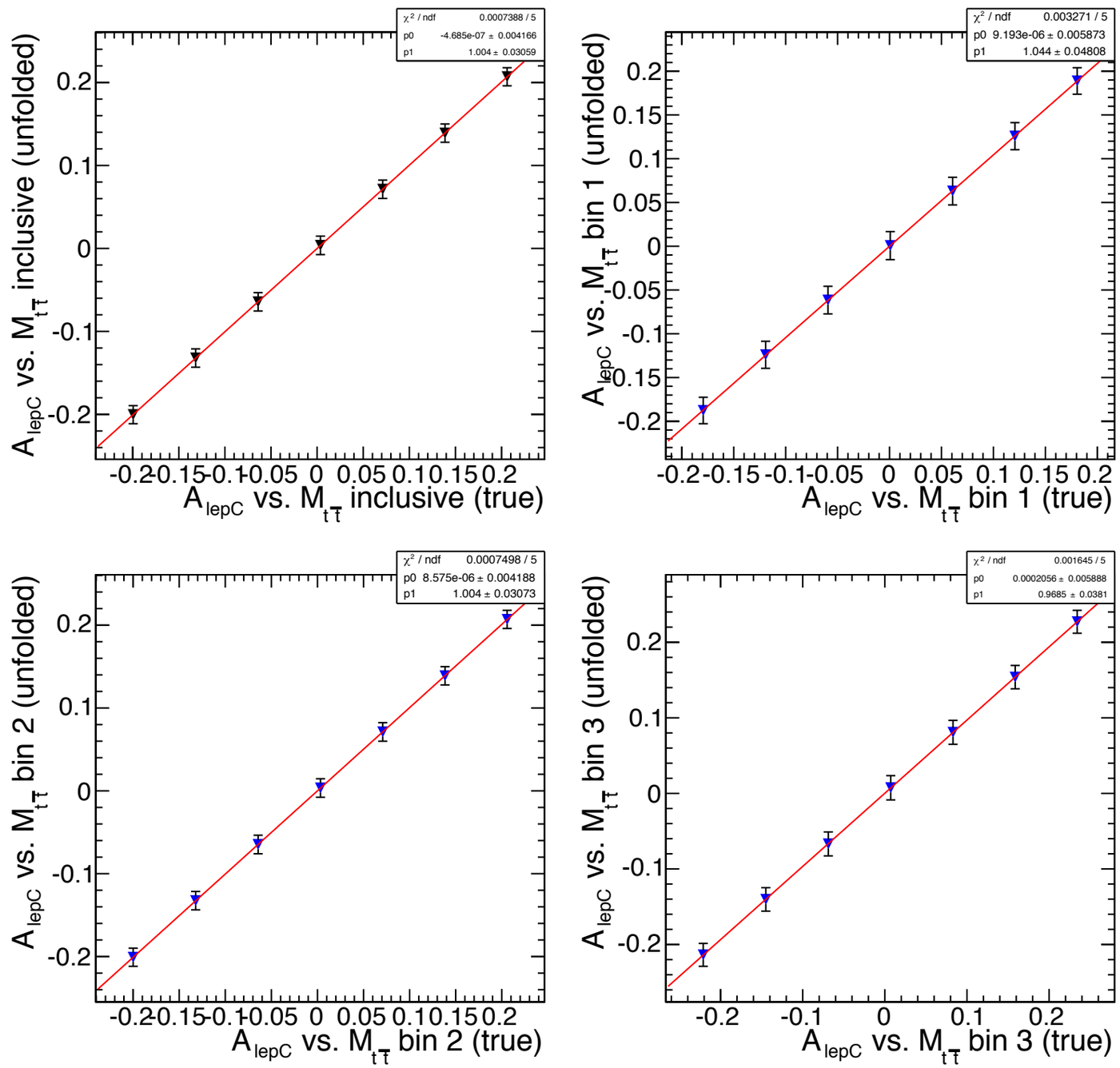
- We remove the regularisation bias simply by disabling the regularisation between the central pair of bins
  - of course this does not affect the weighted distributions (below), only the unfolding

Original and weighted distributions for M_{tt}*sign($\Delta|y_l|$), before unfolding:



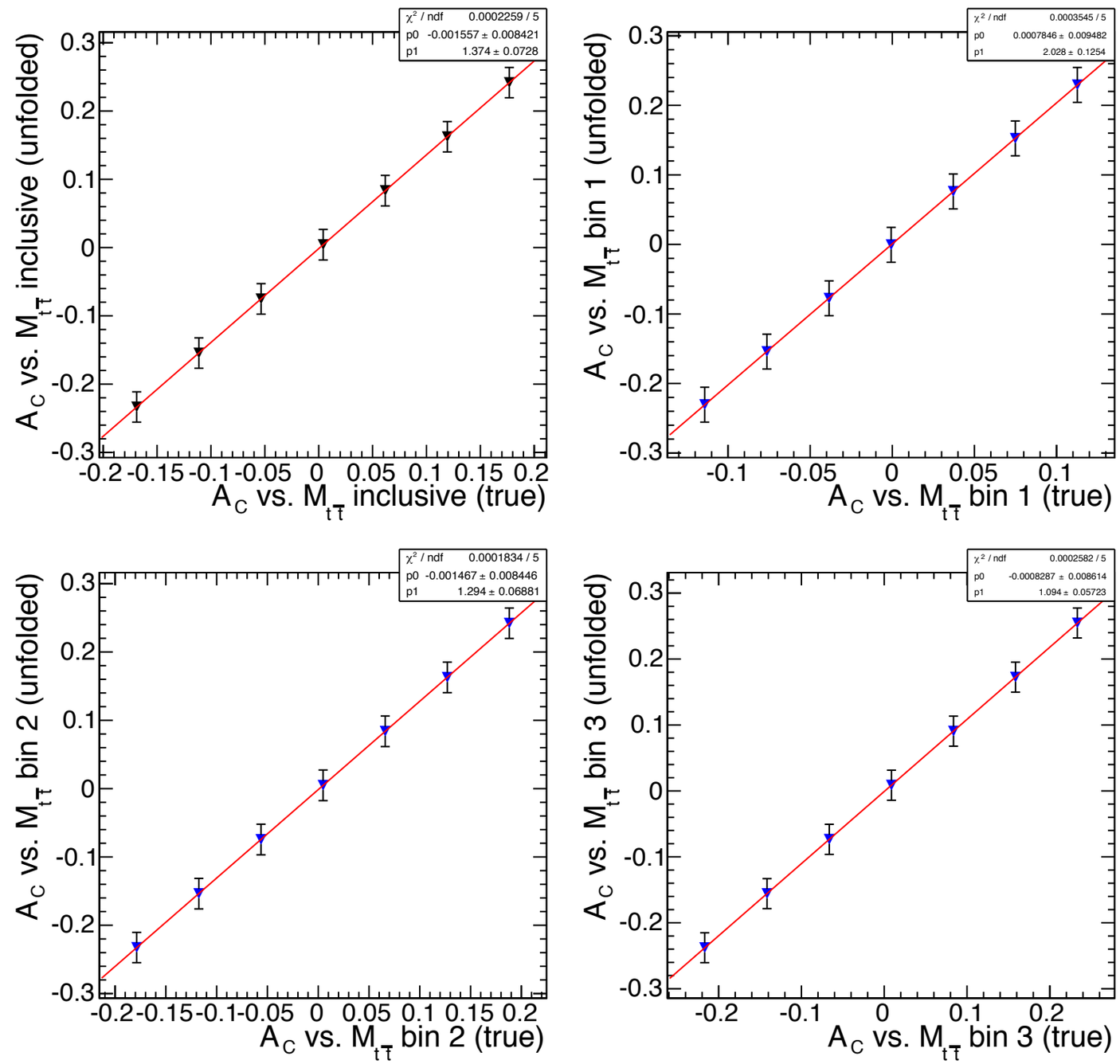
Results for A_{lepC} (fixed)

- Linearity now looks pretty good for all bins (p1~1)



Results for A_C (fixed)

- Slope of linearity plot (p1) significantly away from one for bin1 (2.03) and bin2 (1.29), **as well as inclusively (1.37)**
- I think this is an inherent bias of using 2 bins in $\Delta|y|$ to unfold distributions with a continuous slope in $\Delta|y_{gen}|$ (from weight $1+\text{slope}*\Delta|y_{gen}|$, see slide 2)

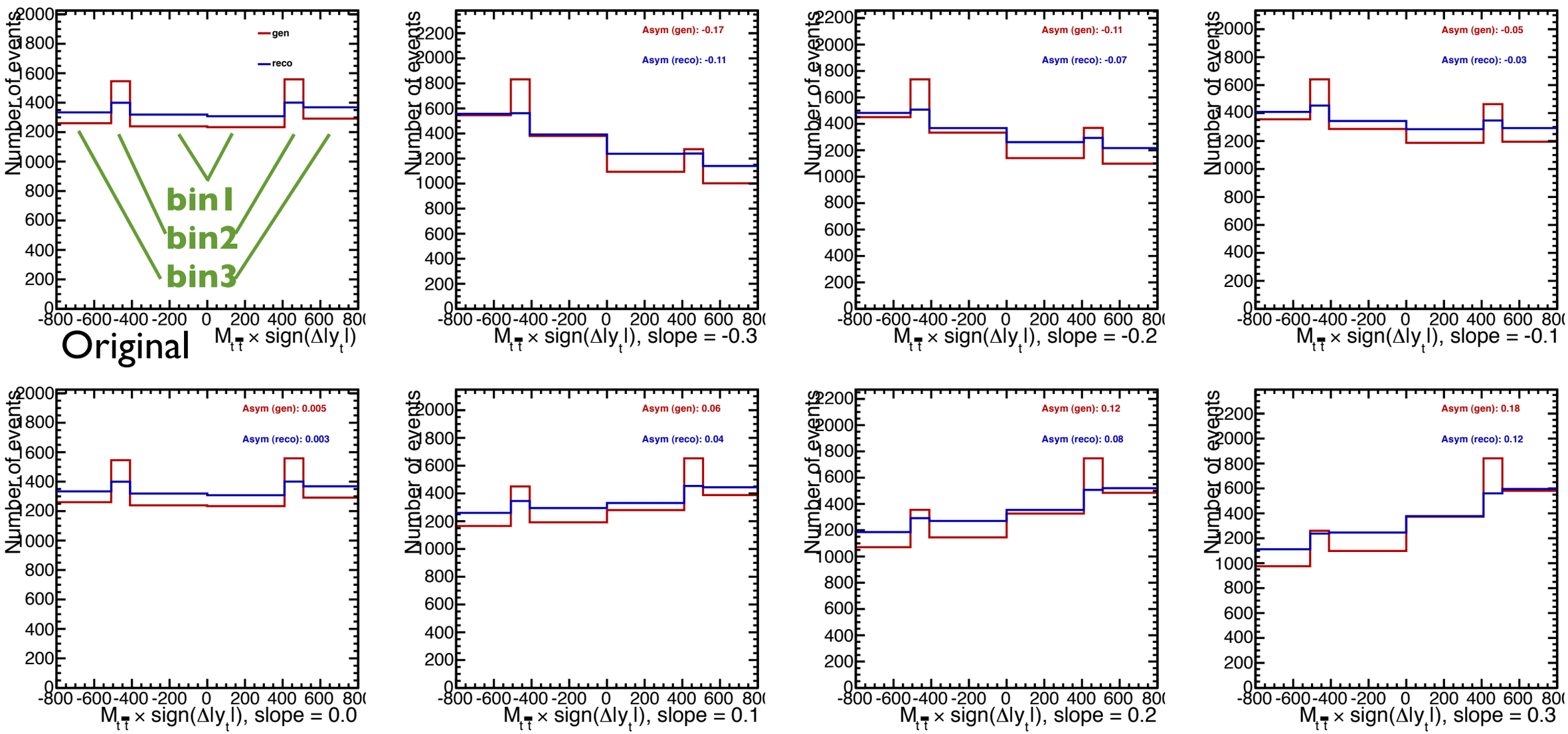


Bias from insufficient $\Delta|y|$ bins

- With a continuous slope, weight = $1+\text{slope}*\Delta|y_{gen}|$
  - => events with large | $\Delta|y_{gen}|$ | get the largest weight
    - but these are also the events least prone to migration
    - however, the response matrix treats them as though they had the average amount of migration => biased unfolding
- With a slope matching the granularity of the response matrix, i.e. weight $1+0.5*\text{slope}*\text{sign}(\Delta|y_{gen}|)$, weighting affects all events equally
  - => should be no bias
    - this is confirmed on the next 3 slides

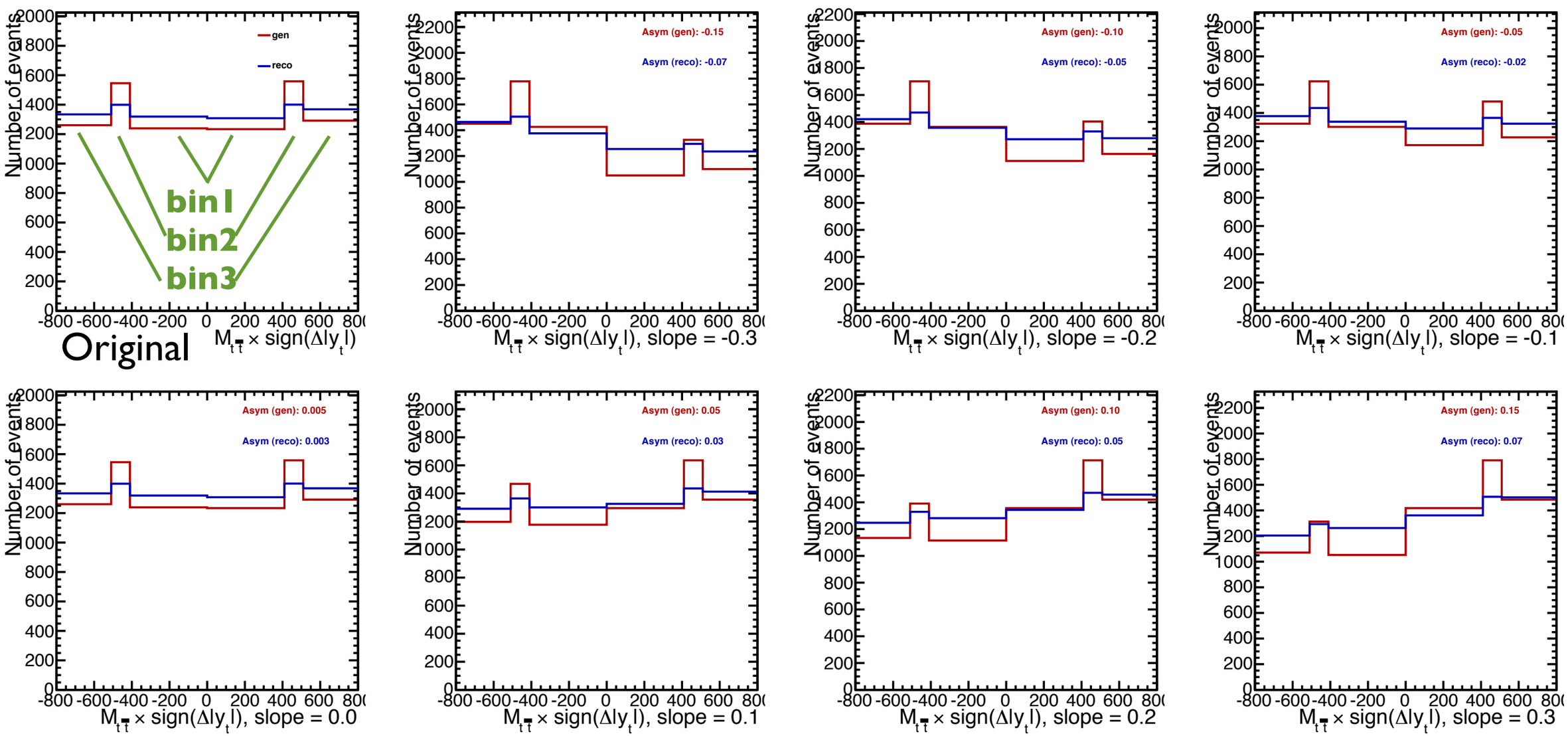
Weighted distributions with continuous slope in $\Delta|y_t|$

- plots show original results with weight = $1+\text{slope}*\Delta|y_{gen}|$
- Reco-level asym is **~2/3***gen-level asym, prior to unfolding (see numbers on plots)



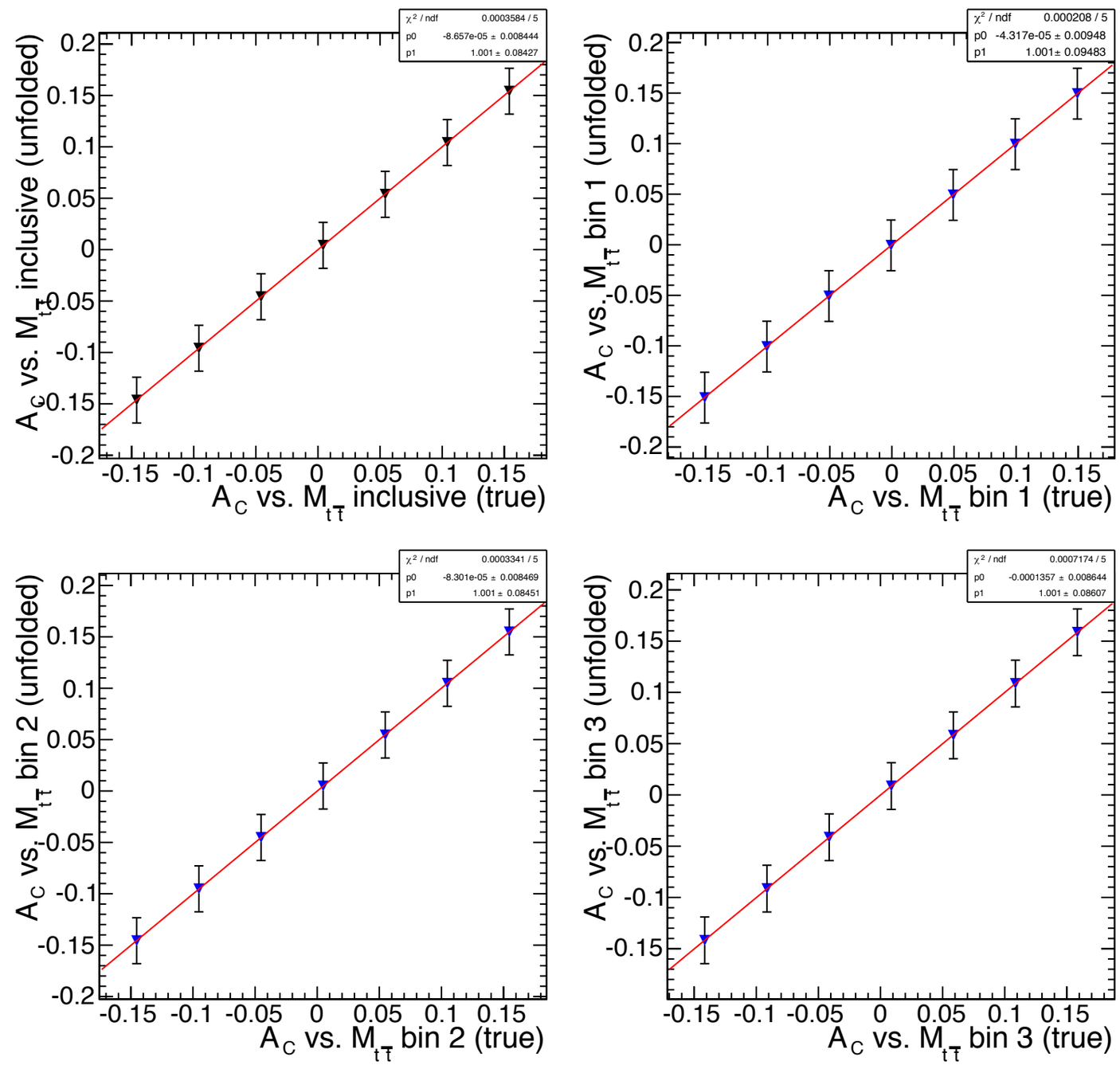
Weighted distributions with slope in sign($\Delta|y_t|$)

- weight 1+0.5*slope*sign($\Delta|y_{gen}|$)
- Reco-level asym is ~**1/2***gen-level asym, prior to unfolding
  - => reco distribution responds very differently to the different types of slope, but the effect of applying our "2-bin" response matrix will be the same in both cases
  - => expect p1 (inclusive) to be (1/2)/(2/3)*1.37 ~= 1.0 when we add a slope in sign($\Delta|y_t|$) instead of $\Delta|y_t|$ (results next slide)



Results for A_C (fixed regularisation, slope only in sign($\Delta|y_t|$))

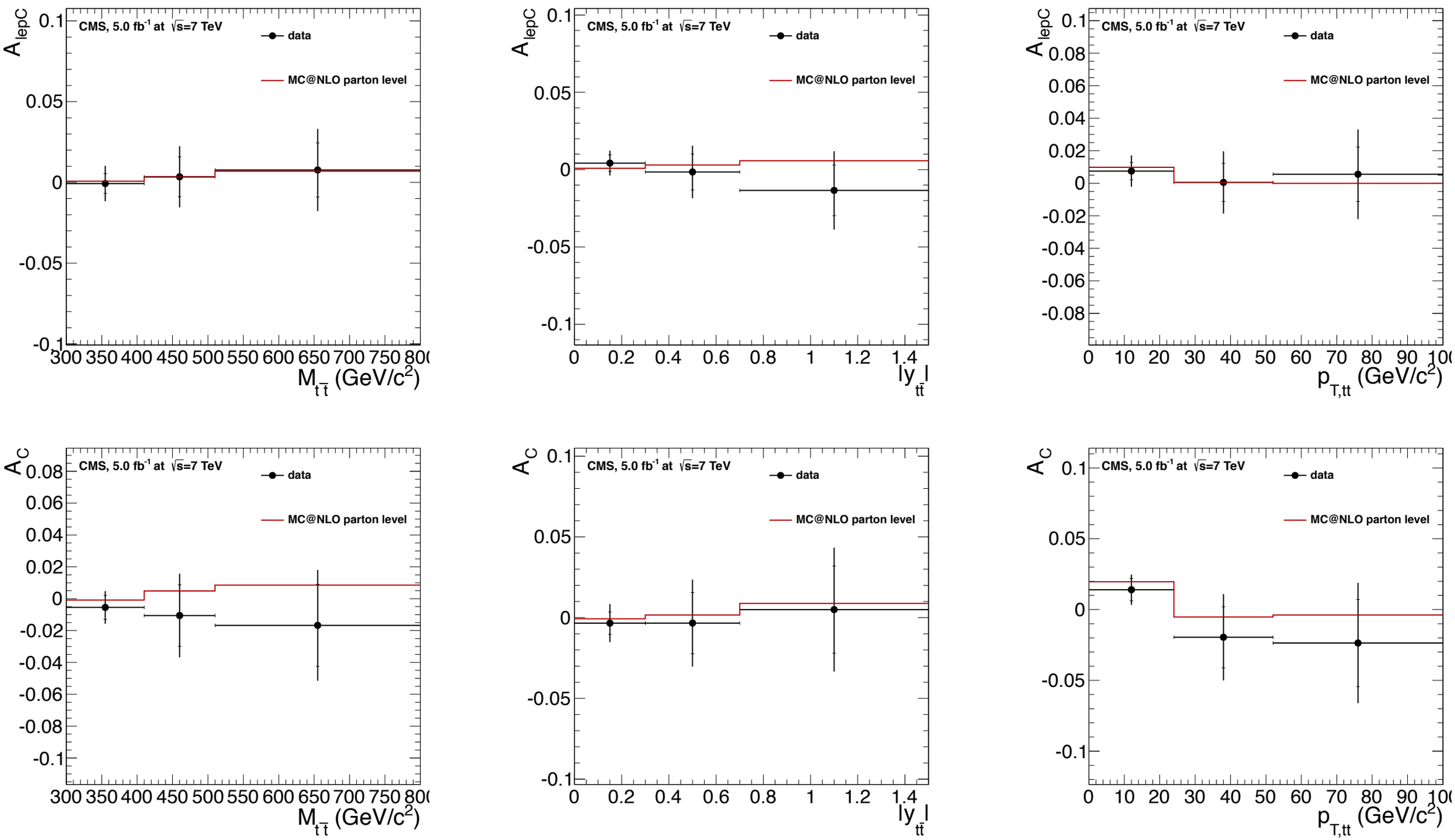
- Slope now very close to p1=1 for all bins
- This result holds for all 5 asymmetry variables, and unfolding vs M_{tt}, $p_{T,tt}$ and y_{tt}
- The only reason we can get away with using 2 $\Delta|y|$ bins for A_{lepC} (slide 7) is that the response matrix basically does nothing here (excellent resolution)



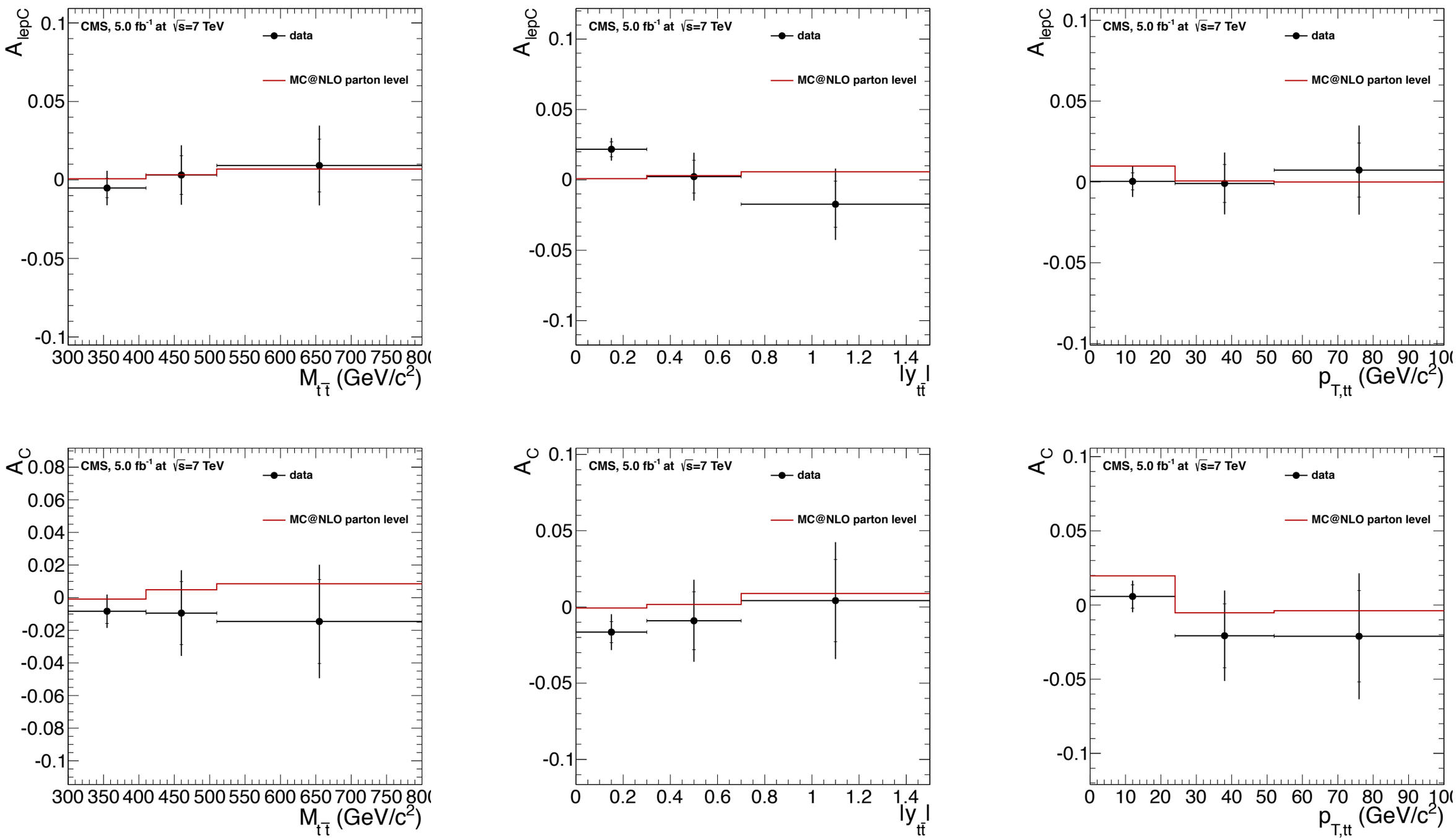
Data results

- Now let's see how the data results change when we "fix" the regularisation (by removing the regularisation between the central pair of bins)

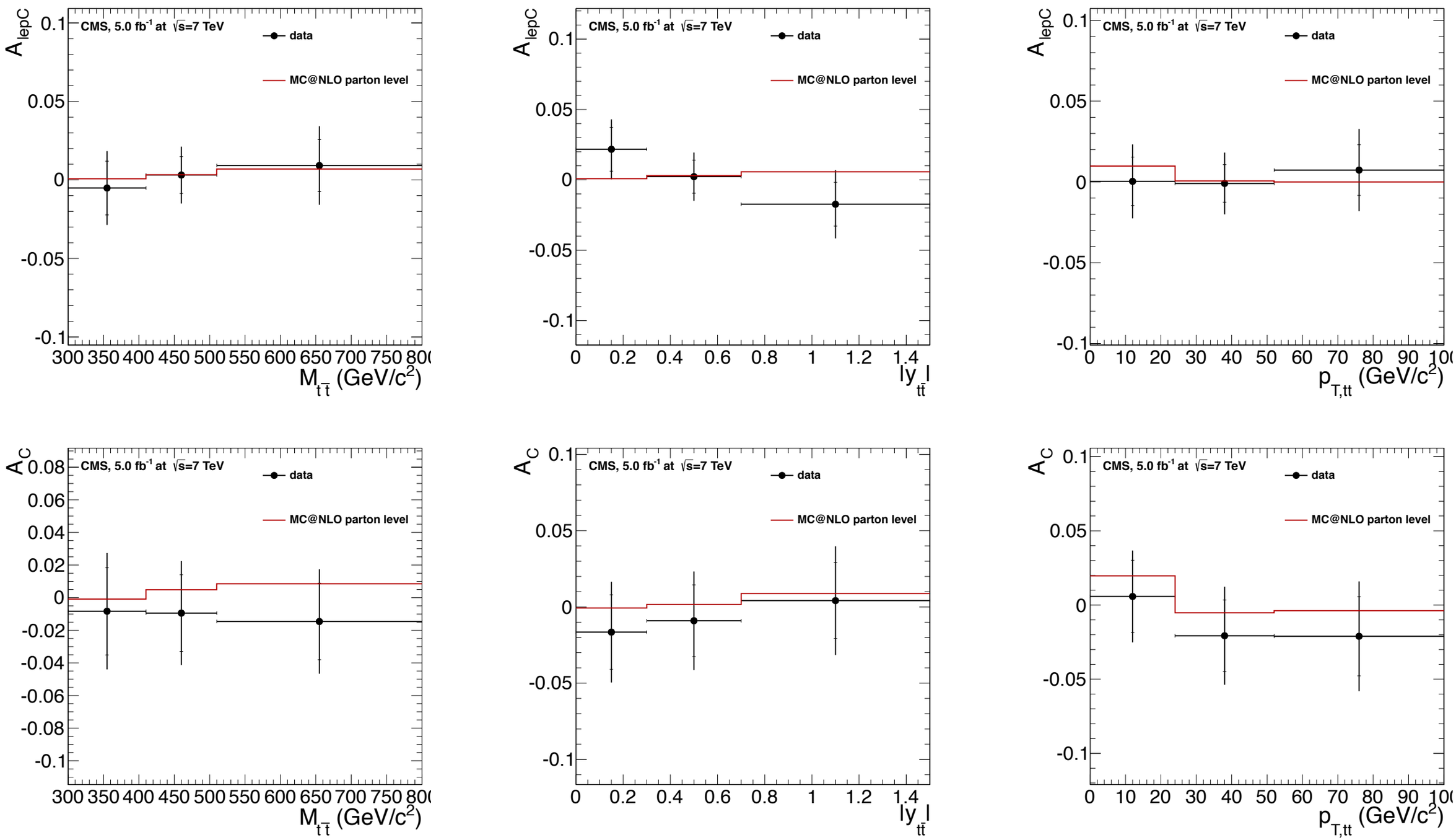
Data results (original)



Data results (fixed regularisation, uncertainties not updated)



Data results (fixed regularisation, updated stat and unfolding errors)



Conclusions

- Fairly small change in data results (almost all in bin1 as expected)
- Linearity of the method is good for A_{lepC} with fixed regularisation (slide 7)
  - probably we have to re-evaluate the systematics (particularly for bin1) because these will also be affected by the over-regularisation (systematics changing the MC distribution will be over-estimated, while systematics changing the data, i.e. JES, lepton energy scale, and background, will be underestimated)
    - maybe it would be reasonable to re-evaluate only the JES, lepton energy scale, background, and unfolding systs, and also top mass (by far the largest MC systematic *)
- I'm not sure we can show the A_C results, although we could argue that they are close enough to the SM that we don't care how bad the linearity is (as demonstrated by the small difference between the previous 2 slides, despite linearity "p1" changing from 0.66 to 2.03 for bin1!)

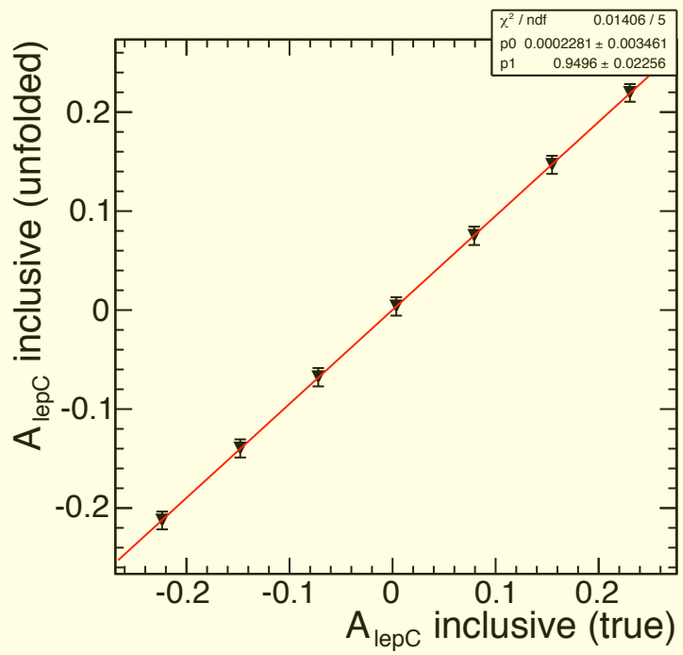
* top mass does not affect the data for A_{lepC}, since it is independent of the AMWT

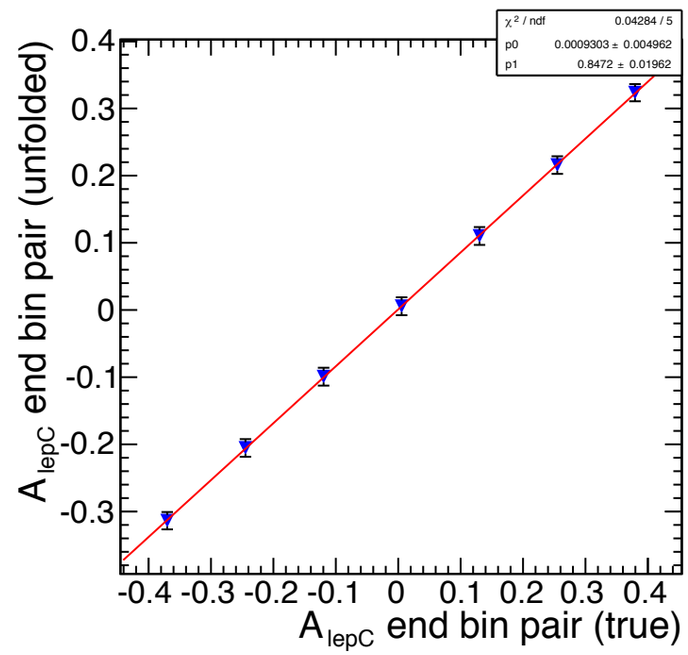
1D results

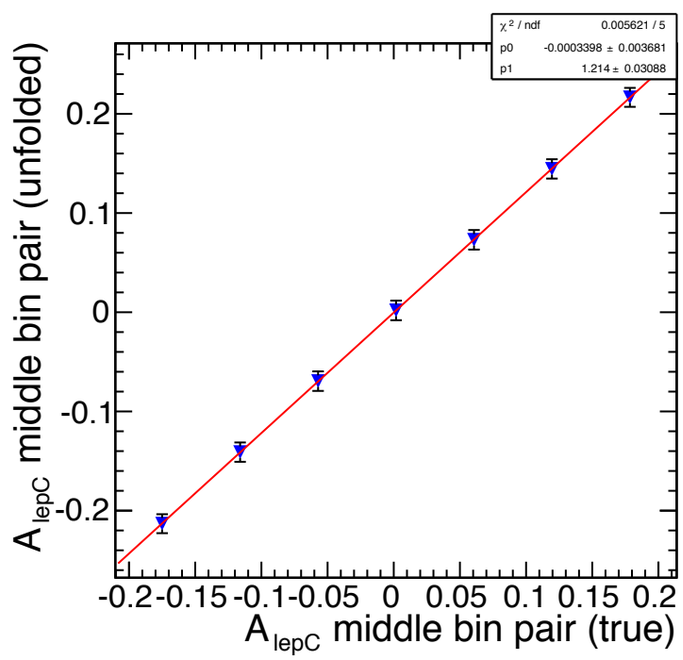
- “p1” is only ~1 for inclusive results due to bias from regularisation
  - => only show inclusive linearity plots
- this is OK since we only quote the inclusive asymmetries, and the bias in our plots of the unfolded distributions will be small since they are so close to the SM

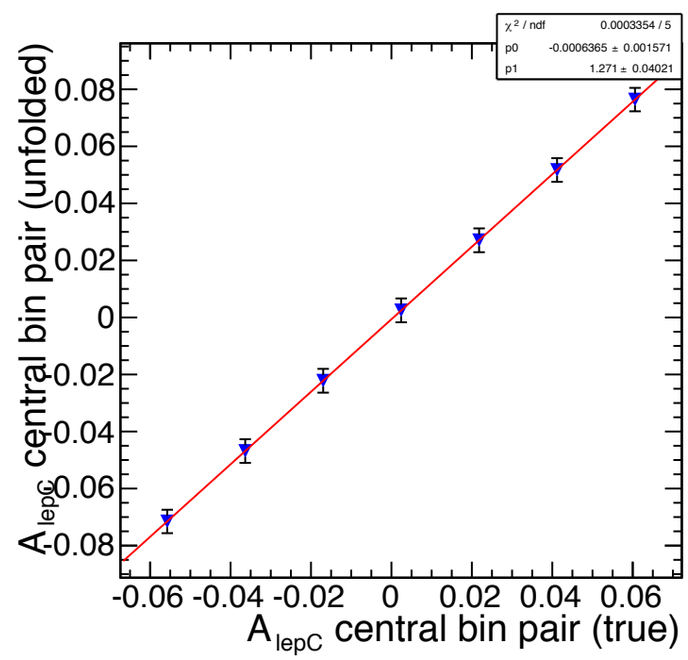
A_{lepC}

show this plot



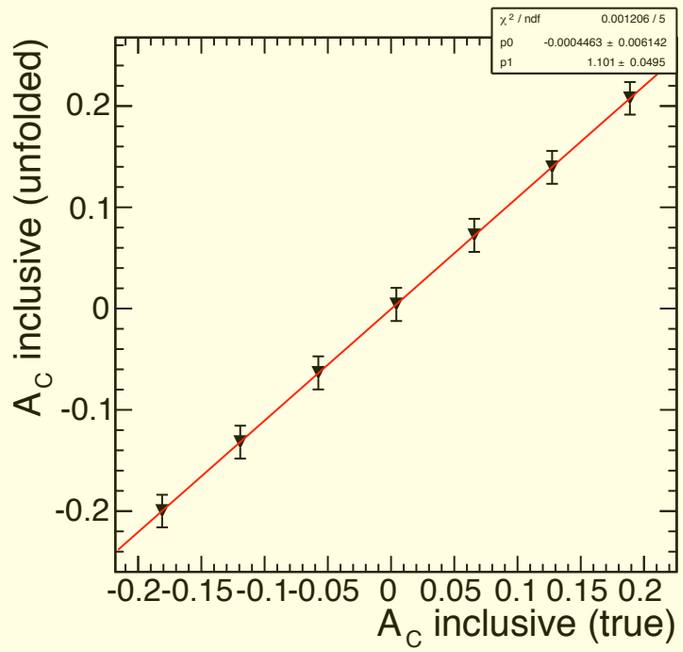


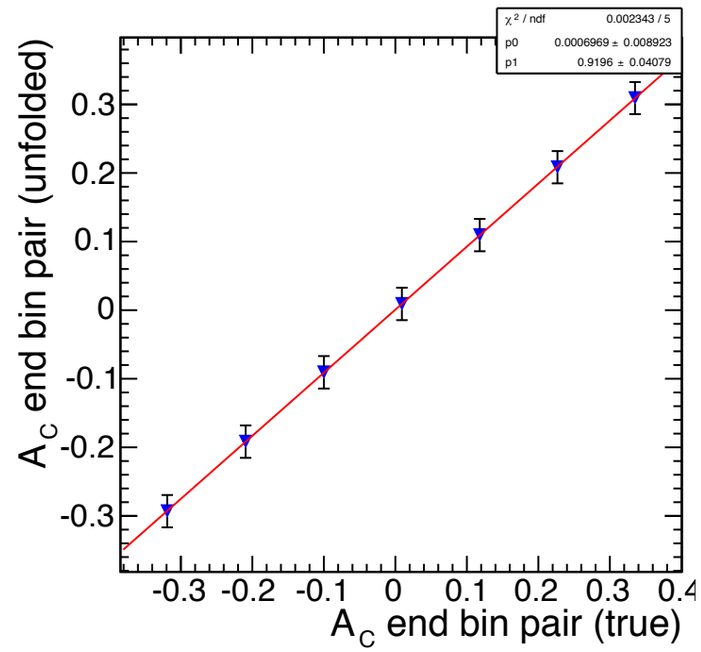


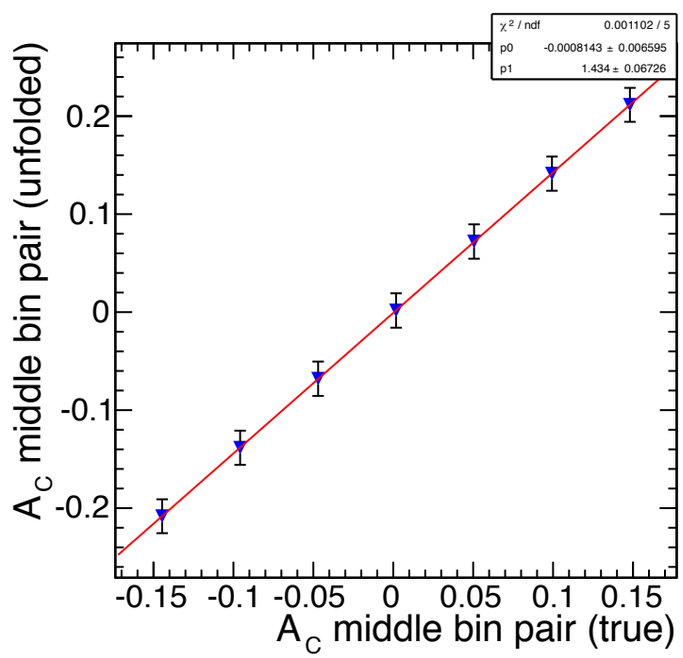


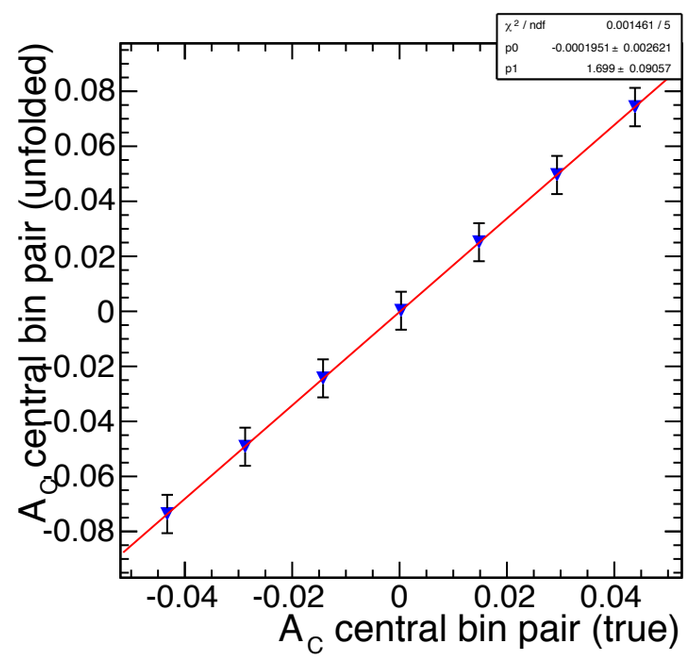
A_C

show this plot



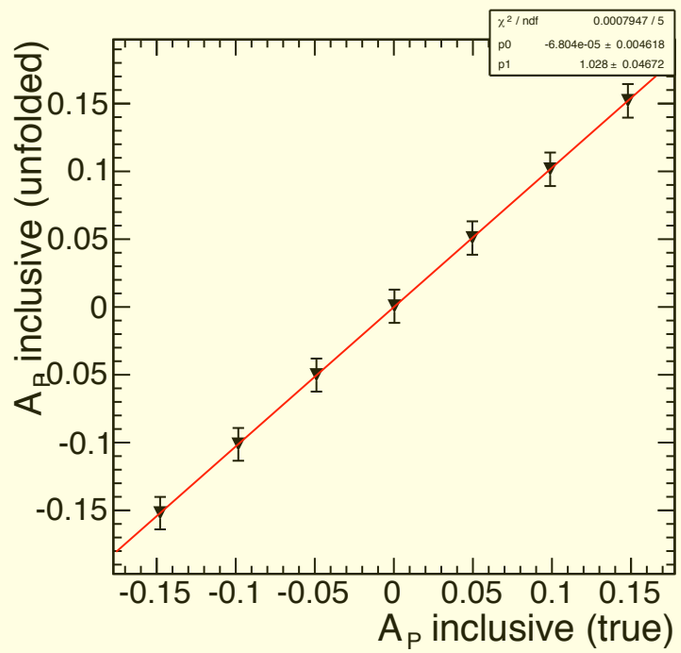


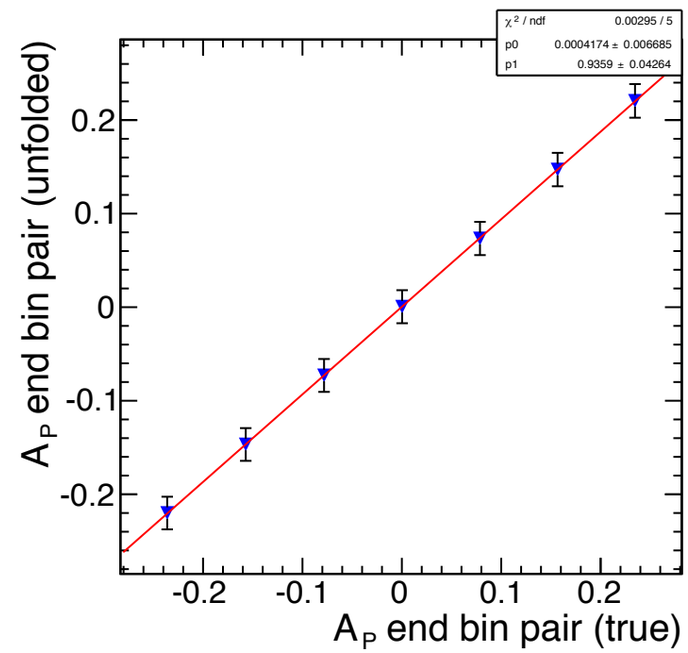


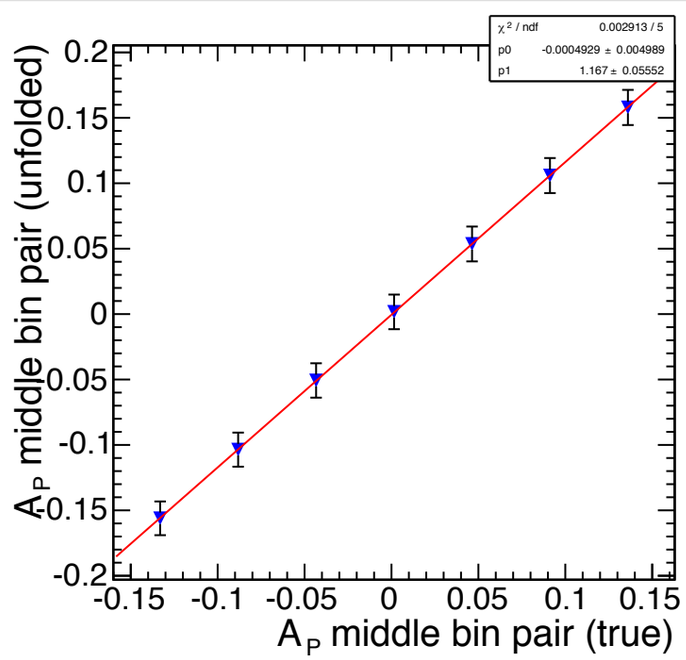


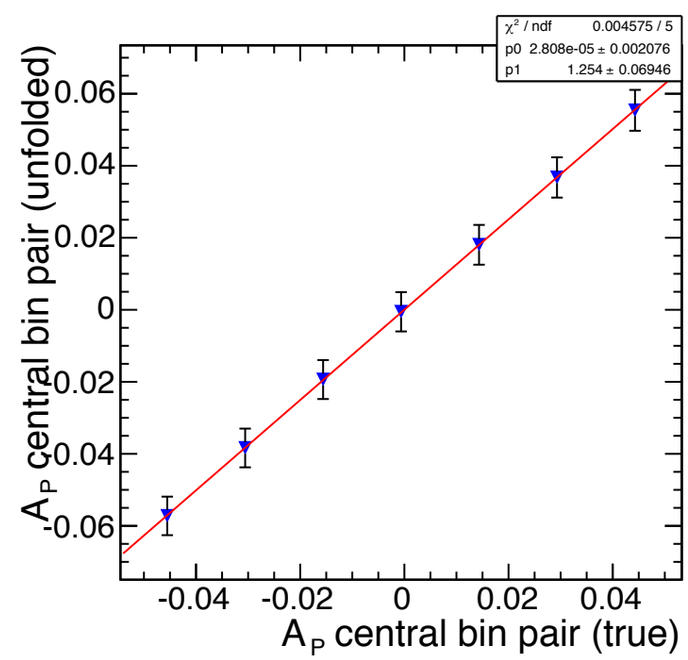
A_P

show this plot



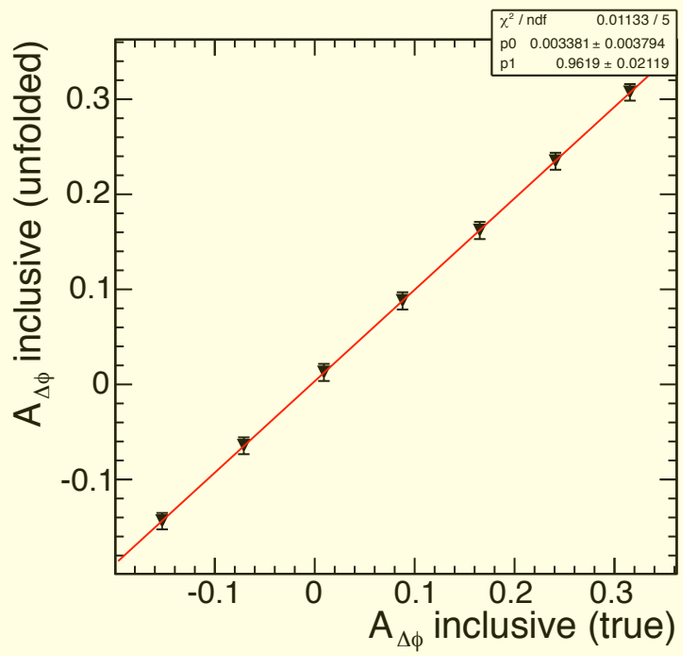


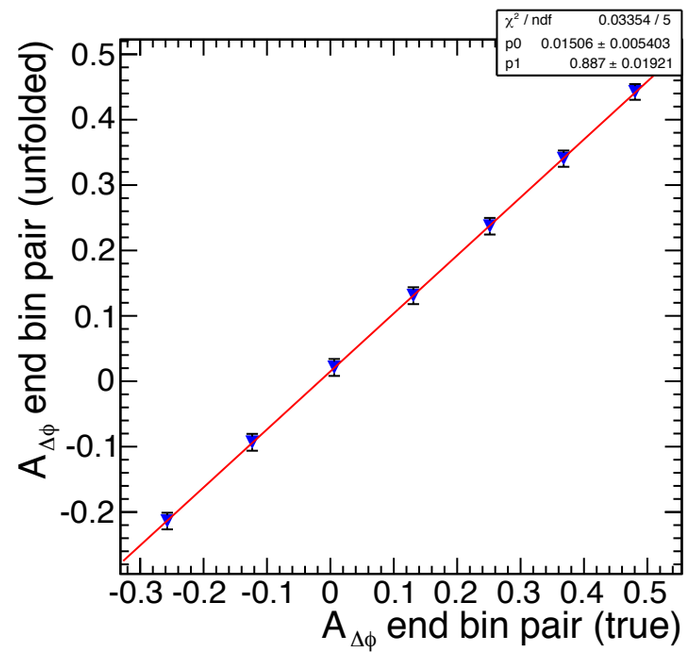


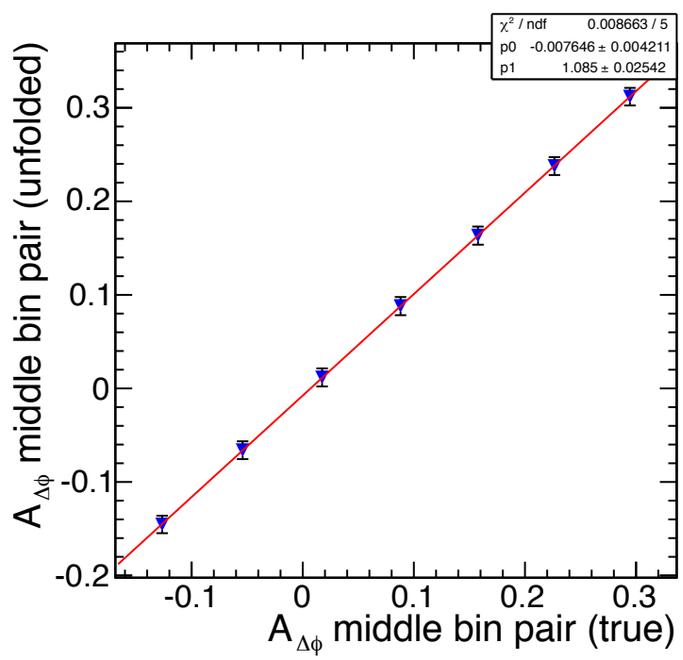


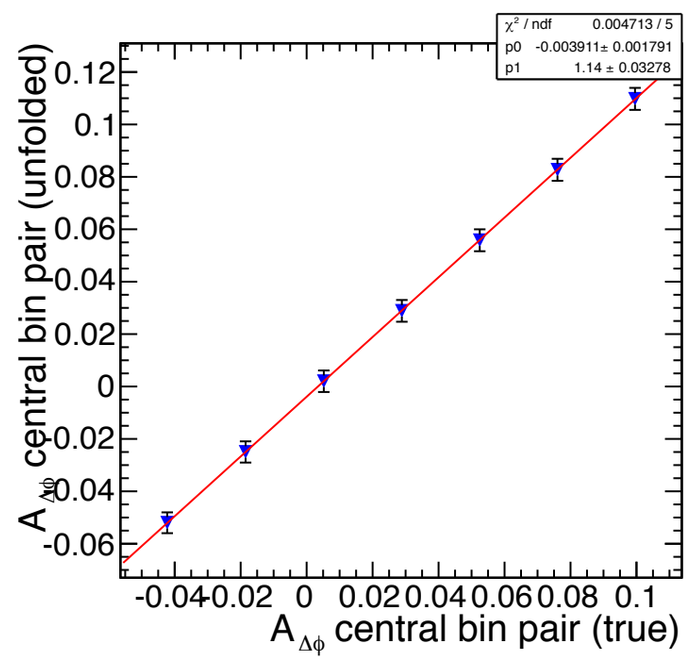
$A_{\Delta phi}$

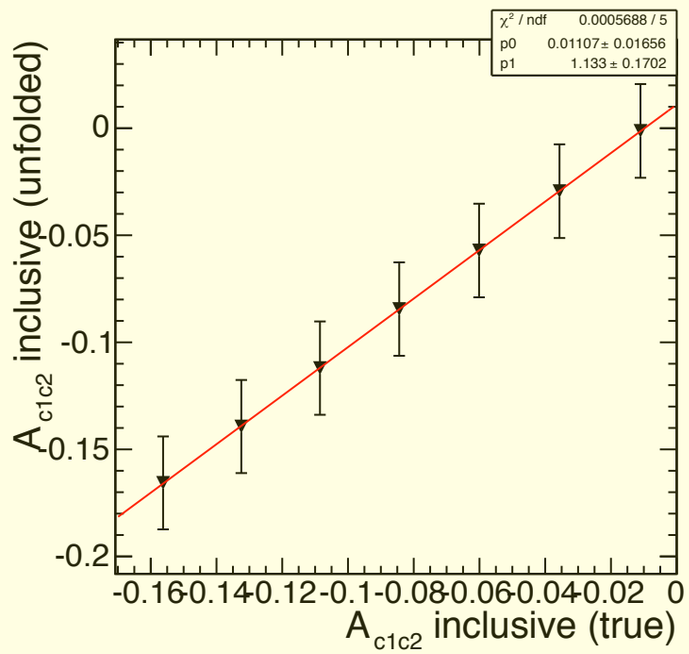
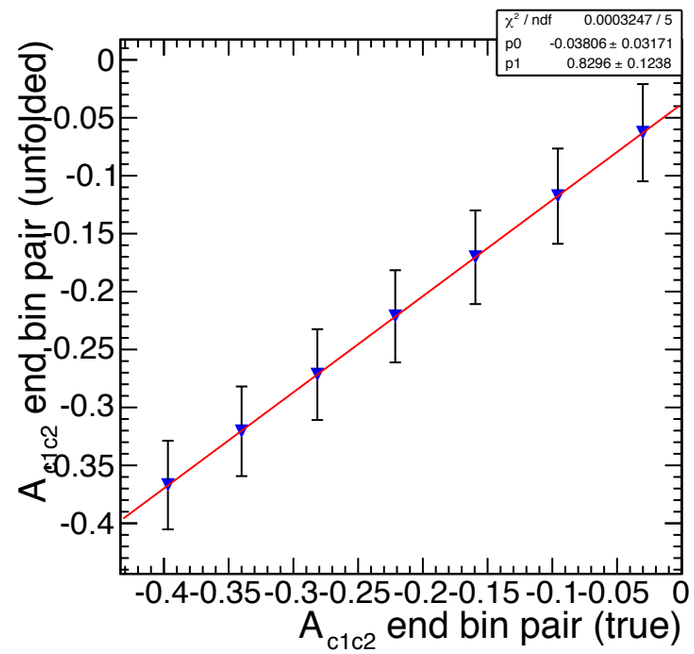
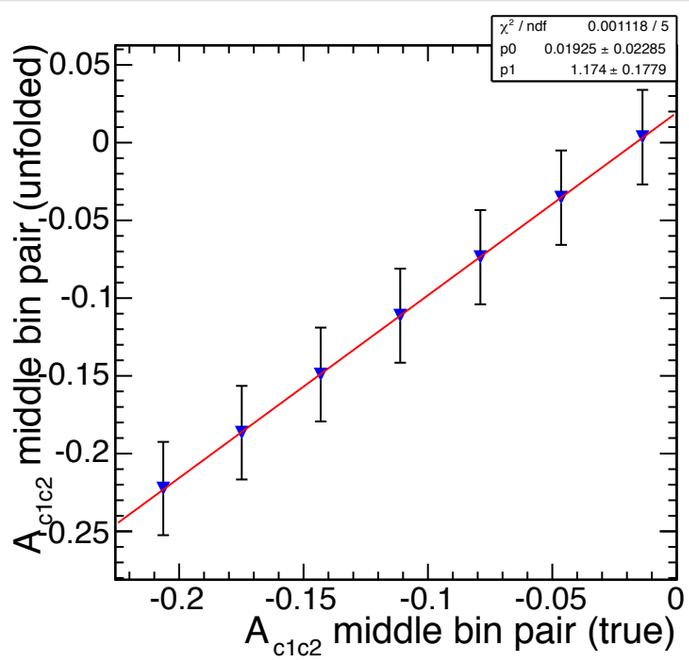
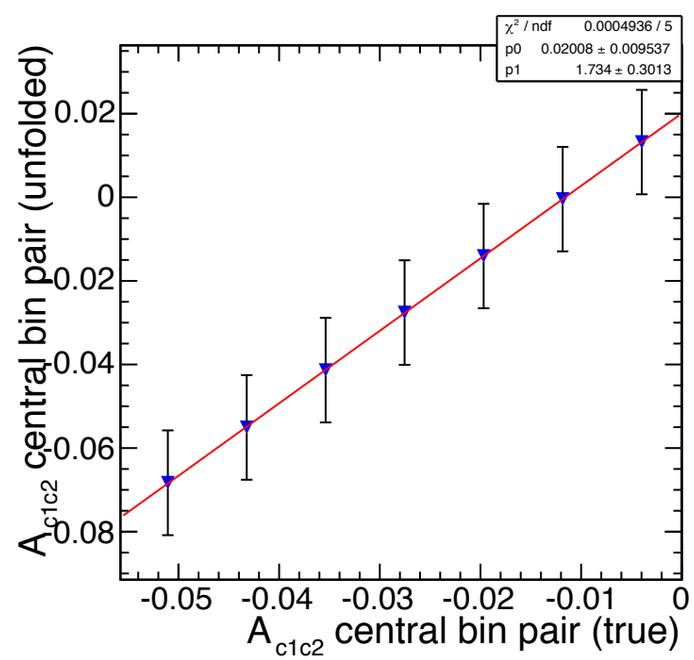
show this plot









A_{c1c2}

show this plot