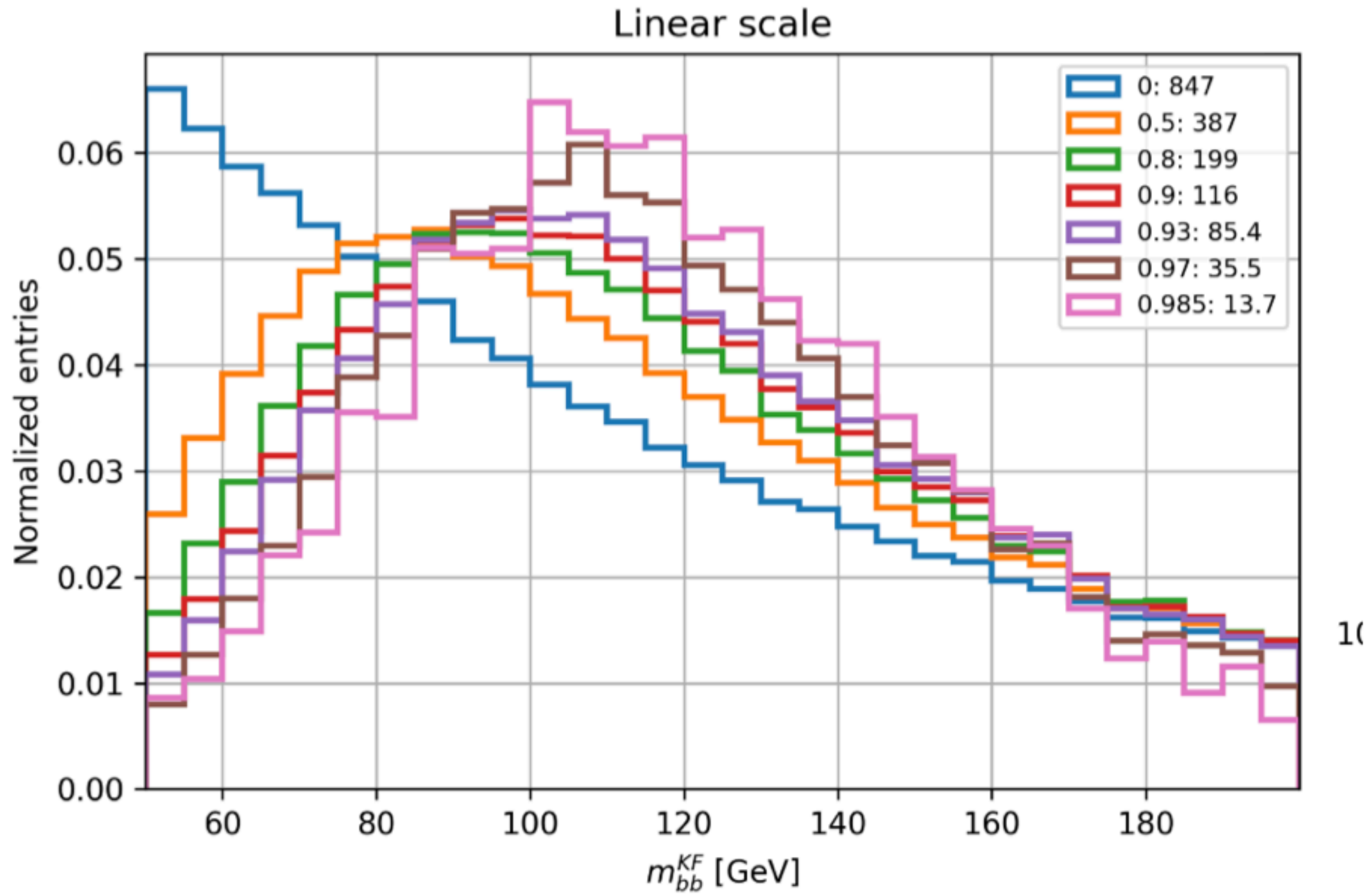
HH>bbyy transformer and toys

Oleksii Kurdysh

Laboratoire d'Annecy de Physique des Particules

28Sept26 / 2026 Japan-France FJPPN HEP_18 meeting

- I designed a model discriminating HH from backgrounds while
  - Not sculpting myy of yyjets
  - Not sculpting mbb of yyjets
- Here focus on High Mass region providing SM significance, LM currently is 1D
- On the plot: yybb sample mbb distribution for increasingly tight cut sof non-mbb-decorrrelated model

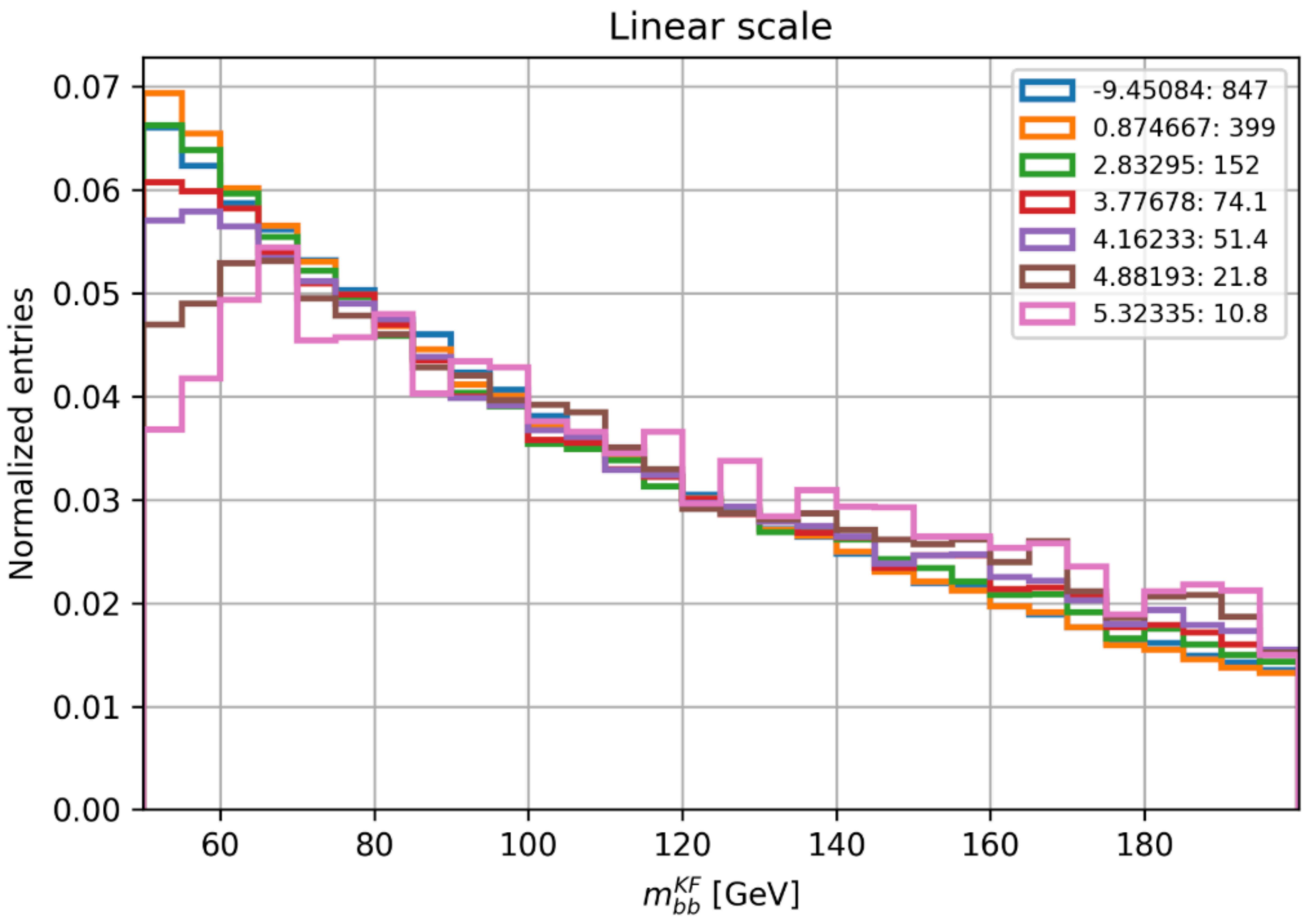


Potential ways to archieve decorrleation

- 1)Distance or other type of decorrelation added, adversarial network, ...
  - Additional component in the network
  - With additional parameters which are to be tuned
  - Overall intended to clean up accidental leakage of info, not to directly fight with the most important variable in 1D which is mbb
  - Tested and getting similar results to 2)
- 2)Usage of variables which are not sculpting
  - Either works or doesn't
  - Simpler and more robust

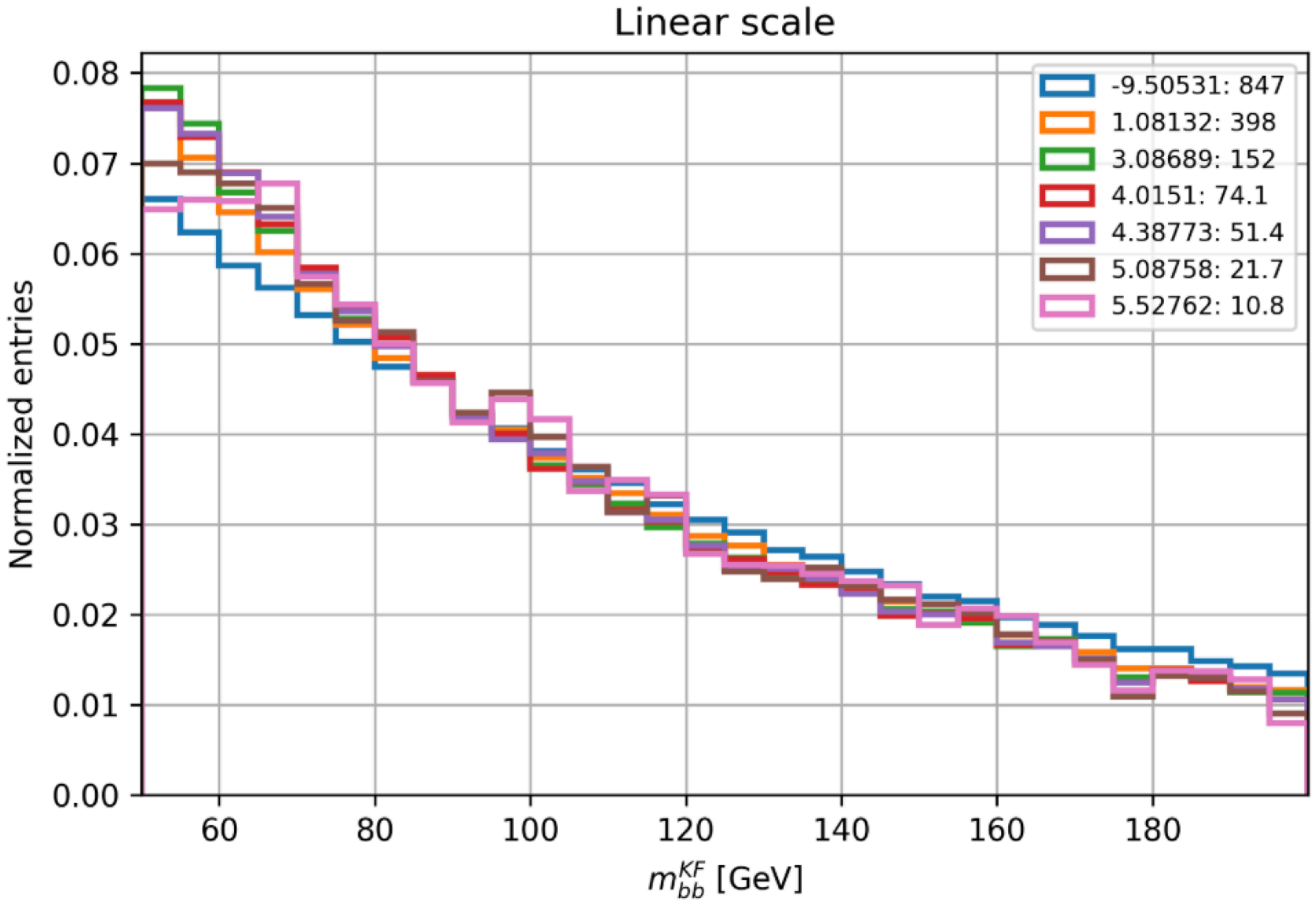
Decorrelation step1,2

- Step 1 - remove all dRbb, pTbb, mbb and their proxies
- Step 2 - train only in region where variation of mbb is less
- Can't pick up corrrelation if it's not there in training
- Pick 100-150 GeV as region with HH mbb core



Decorrelation step3

- Step 3 - as to make mbb need two jets dividing only of jets pt by mbb should enough
  - Divide second jet and keep information from 1st jet for discimination
- Safe to fit from 70, depending on where tightest category lands can fit from 80
- Decorrelation achieved

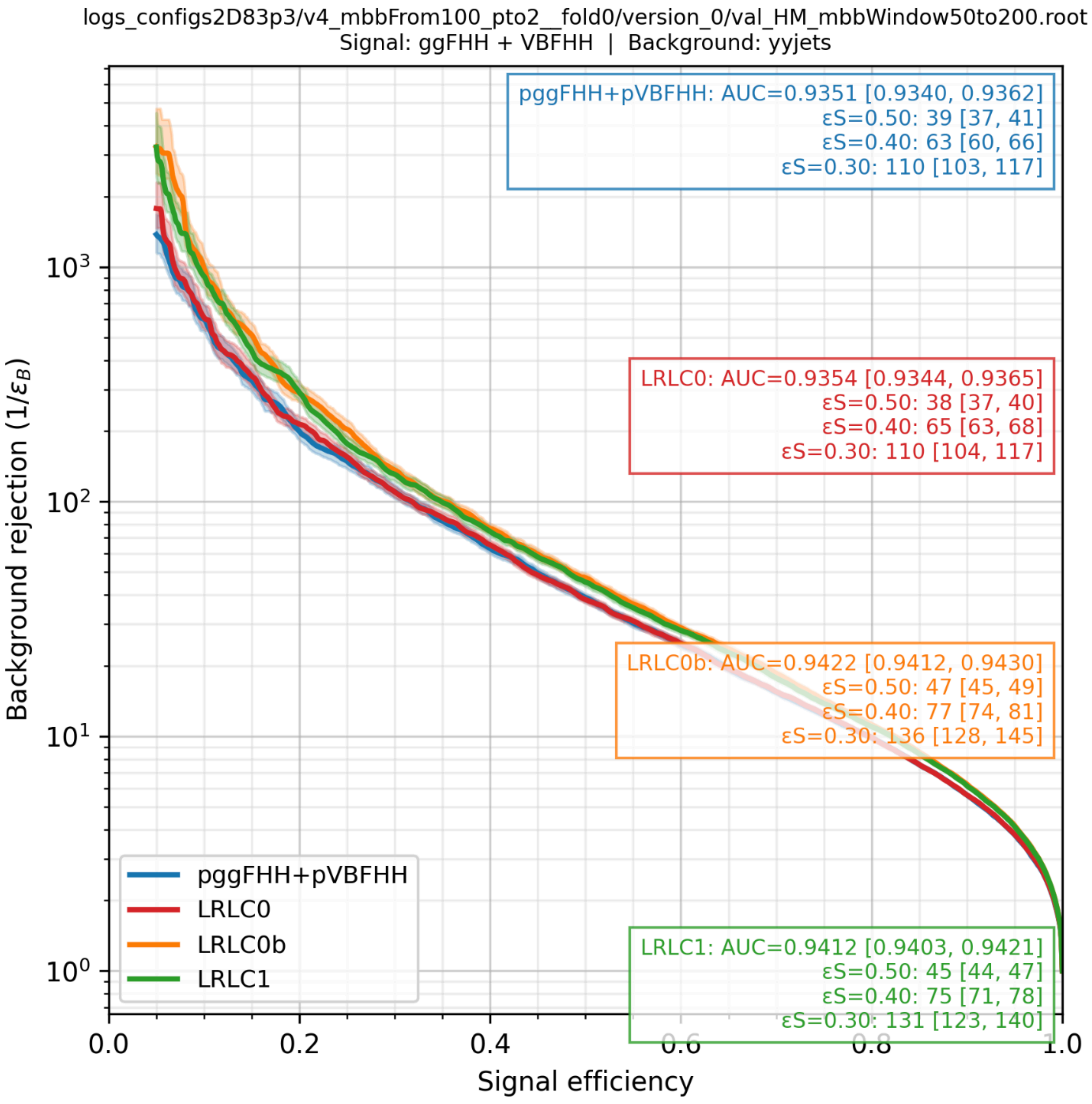


Discriminants involved

- Network provides seven "probabilities"
  - pVBFHH,pggFHH for signal
  - pggF_bbH,pttH,pZH for major resonant backgrounds
  - Pyyjets for yycont
- HH-oriented ways of using it
  - pggFHH+pVBFHH is equivalent of binary score
  - $\mathrm{LRLC1} = \log\left(\frac{0.95 \cdot p_{ggFHH} + 0.05 \cdot p_{VBFHH}}{0.65 \cdot p_{yyjets} + 0.11 \cdot p_{ZH} + 0.11 \cdot p_{ttH} + 0.13 \cdot p_{ggFbbH} + 0 \cdot p_{otherSH}}\right)$ or other tunable fractions

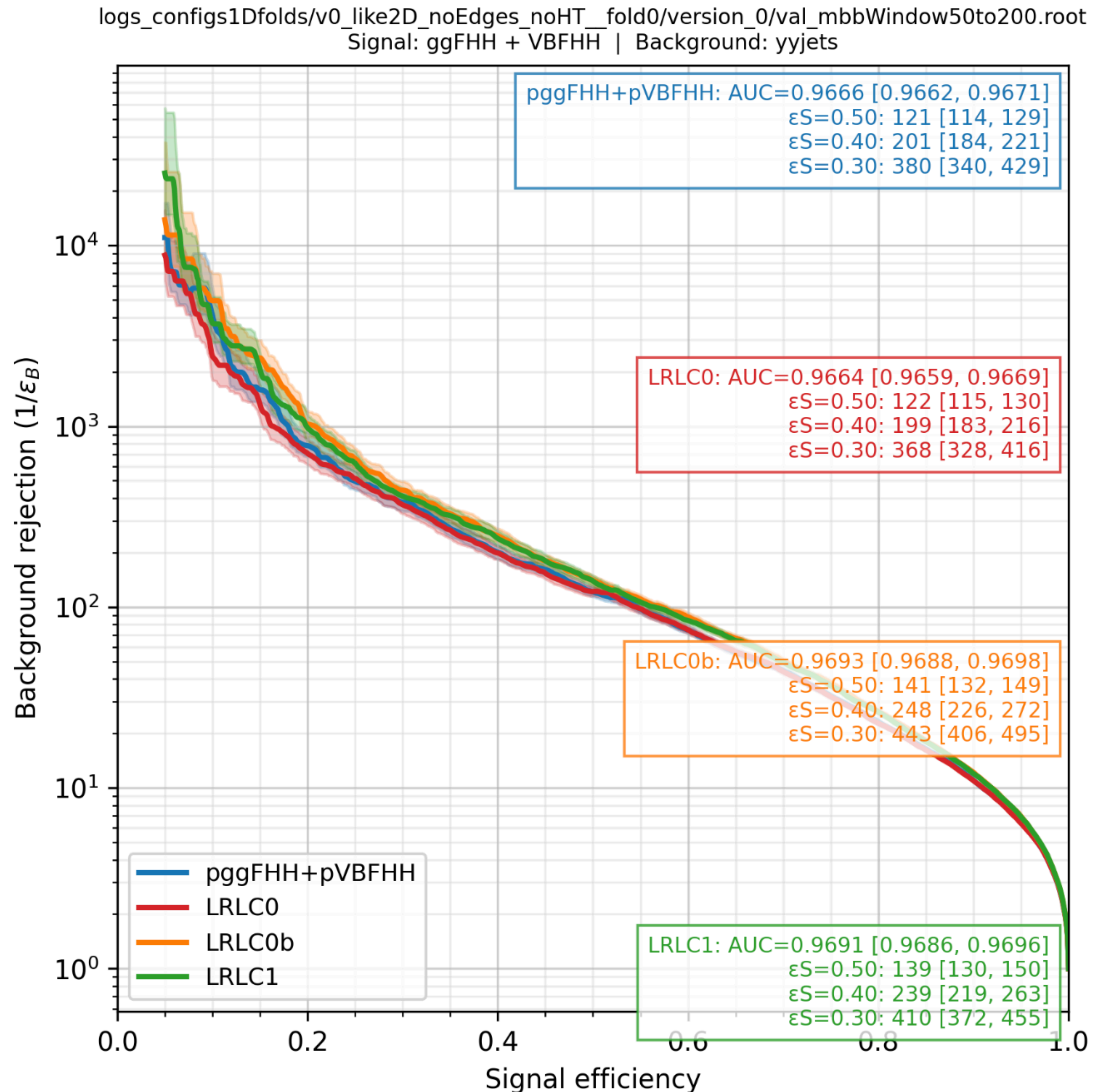
Discriminants involved and 2D ROC

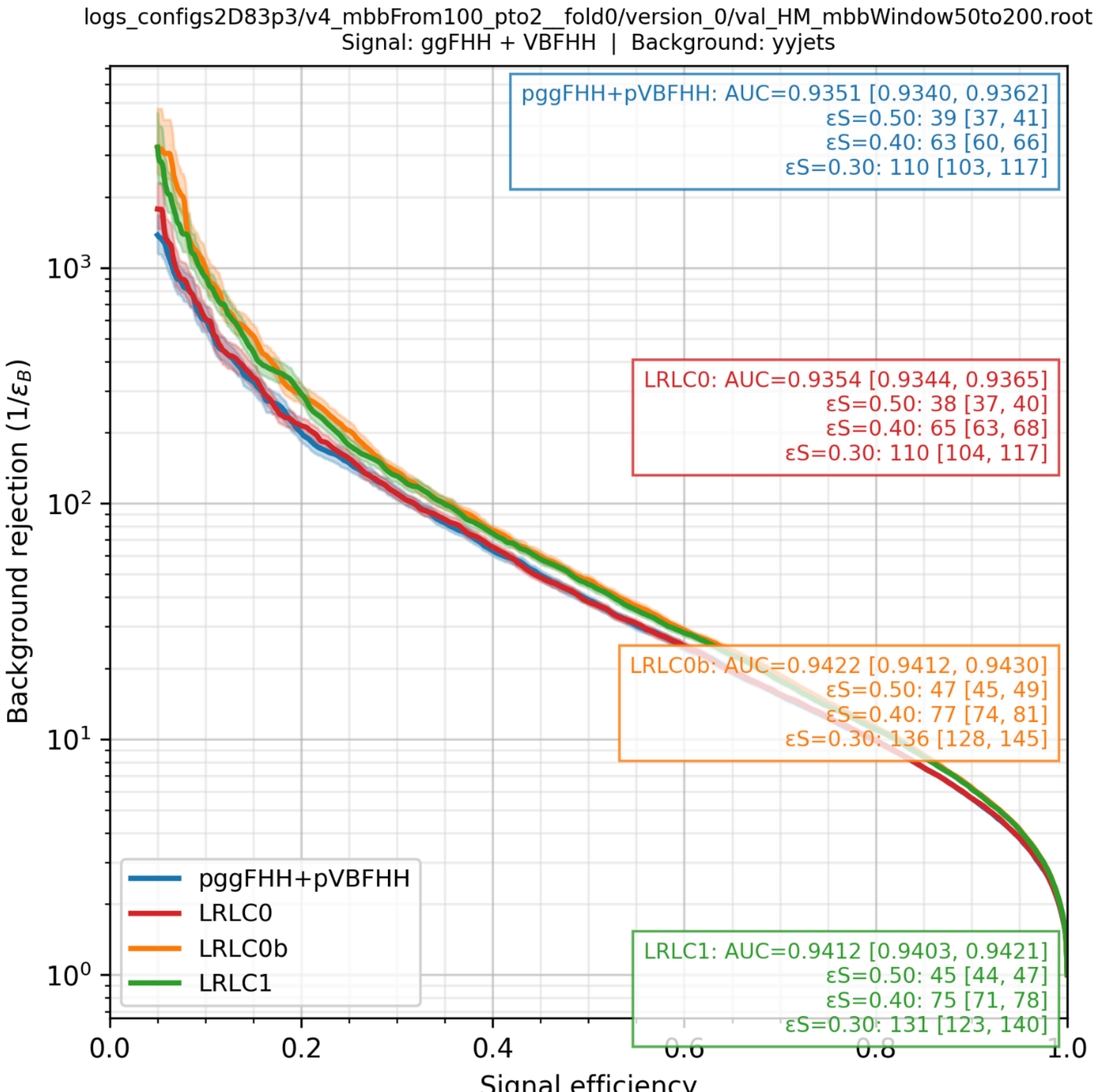
- Evaluated in 50-200 mbb window
- For discriminant rejections, written for reference signal efficincies (previous round tightest category 35%)
- Tuned discriminant provides better rejection



1D and 2D ROC

- Better in 1D but not appropriate metric as it doesn't contain myy-mbb granularity



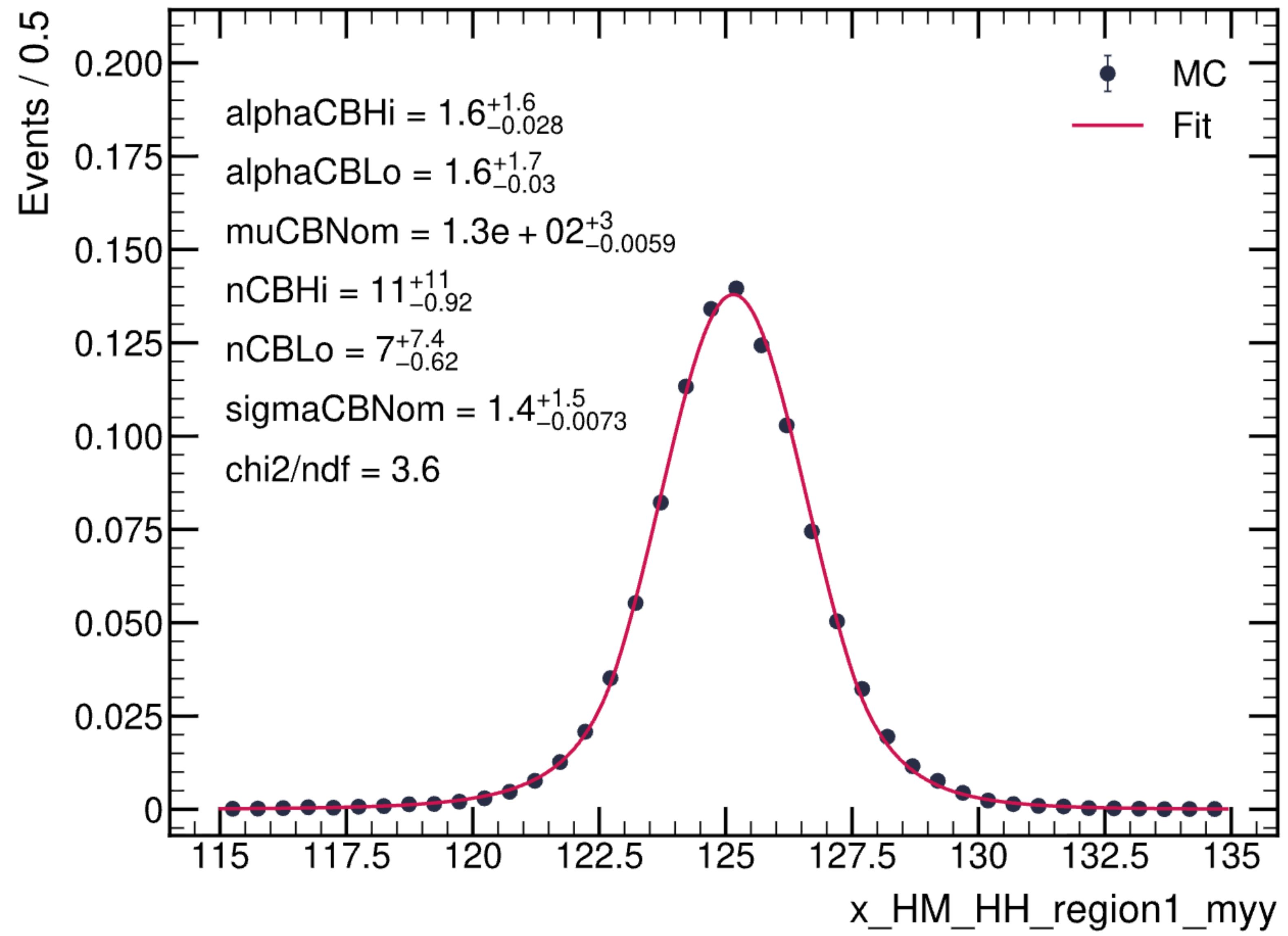


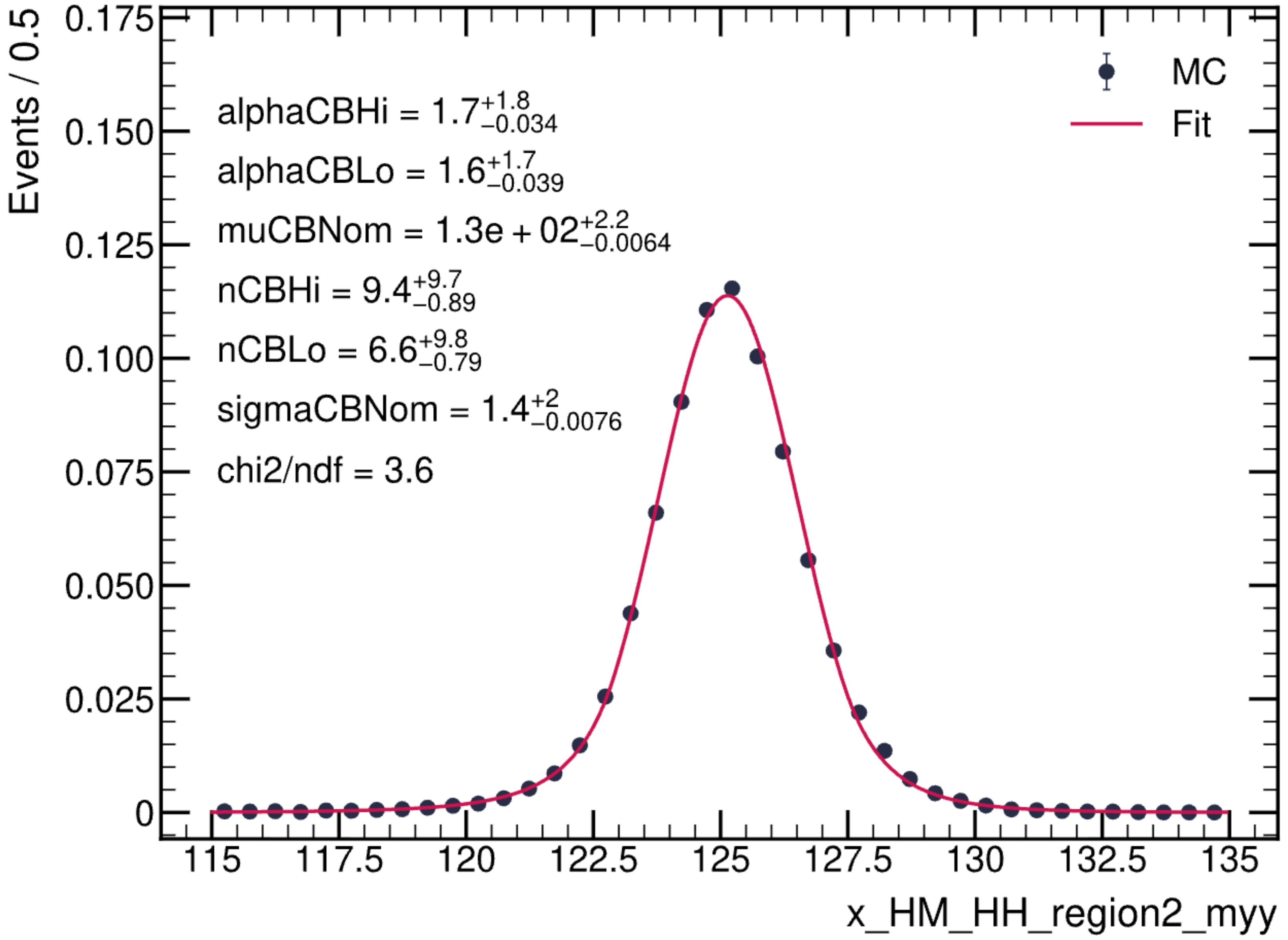
Determining boundaries of categories

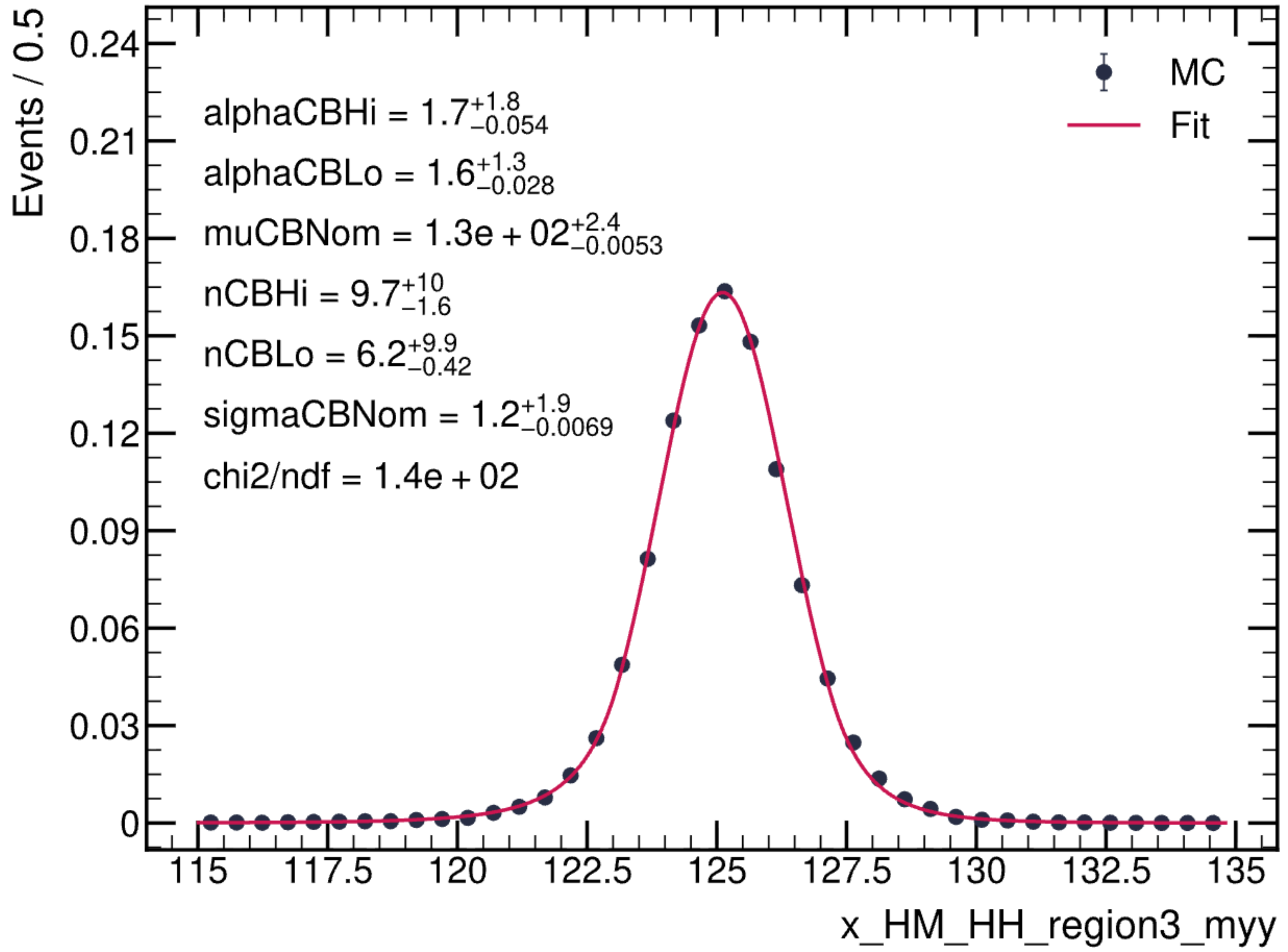
- In 1D
  - H/HH yield in 120-130 window
  - Ycont yield in 105-160 window and multiplied by 10/55 - way to stabilize as this sample is jumpy in category that matters the most
  - Take category edges that best give S/sqrt(B)
- In 2D
  - Necessarily done on myy-mbb grid
  - Stabilization instead is running 300 toys and taking boundary from toys which is closest to median
- In both cases require min sumw 11 for yyjets

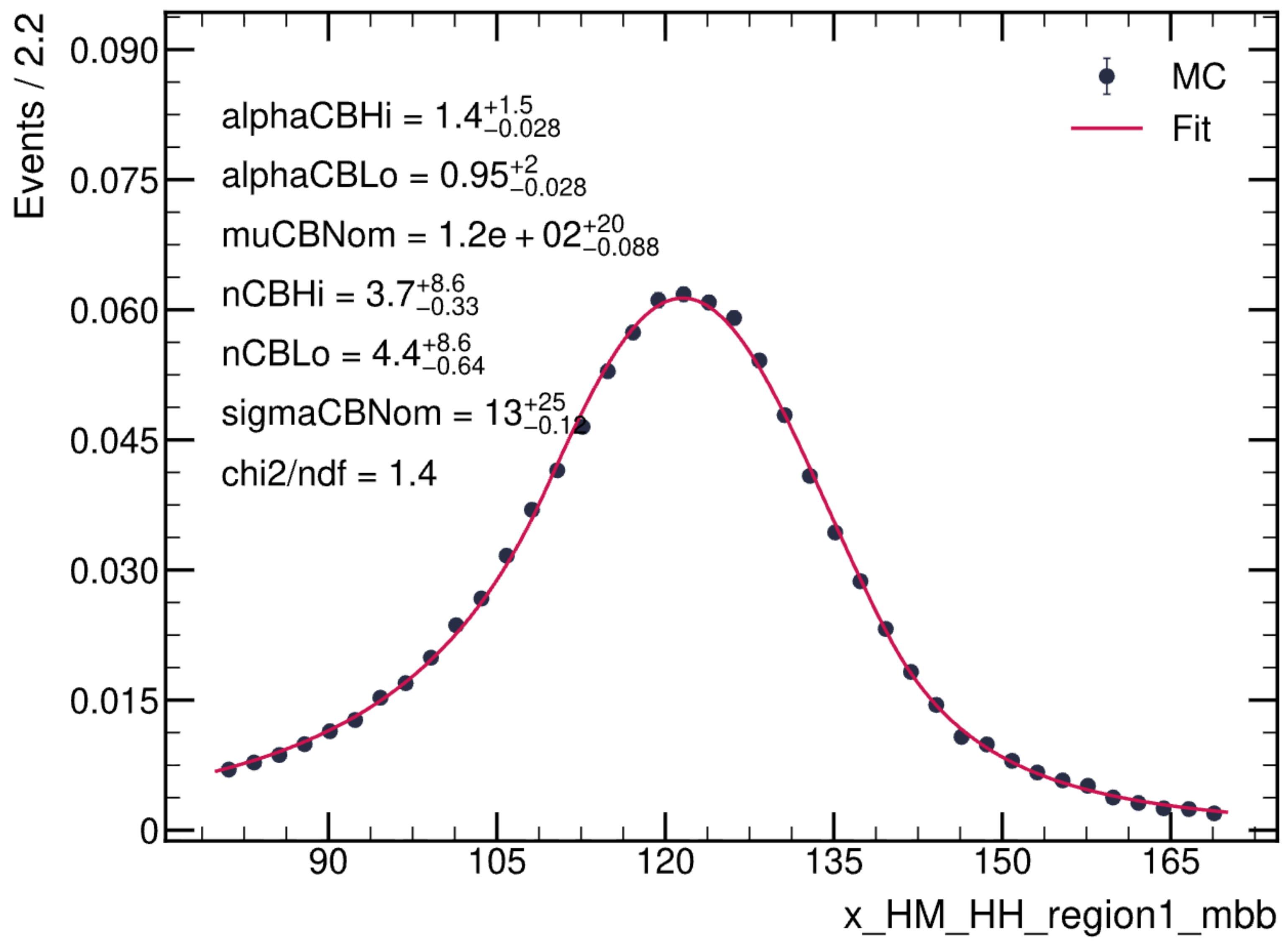
From category boundaries to stat results

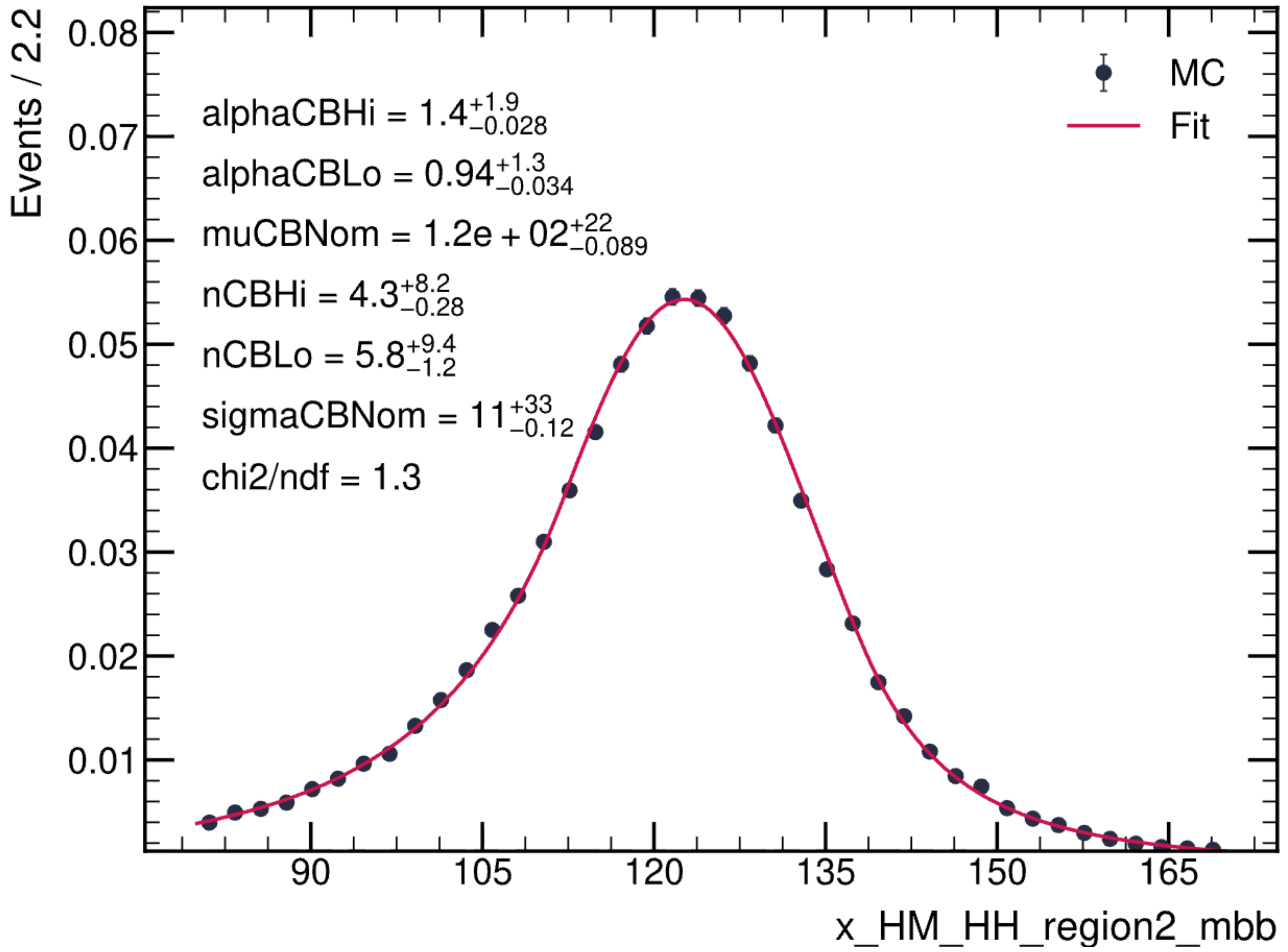
HH modelled with DCSB(myy)xDCSB(mbb)

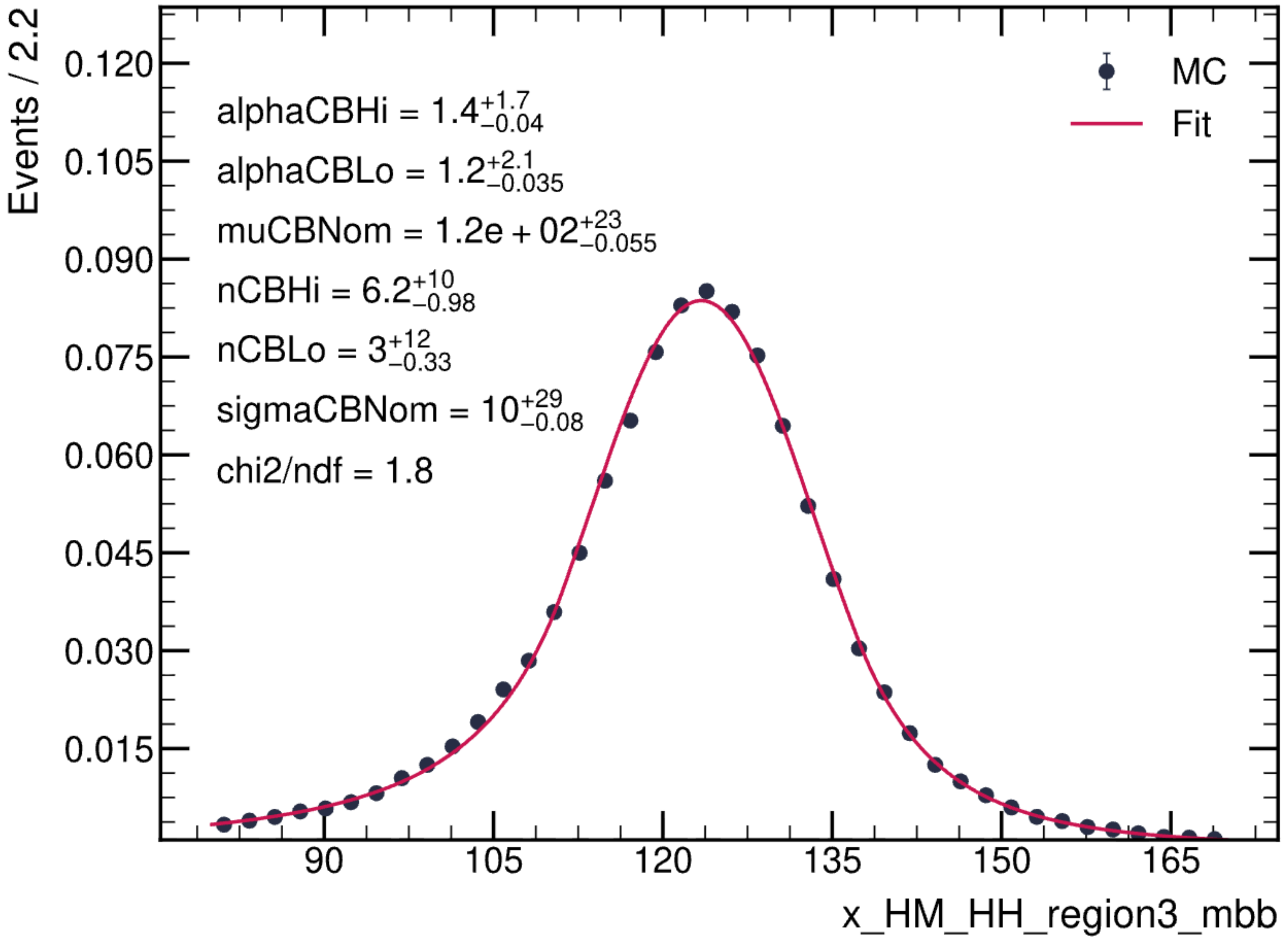






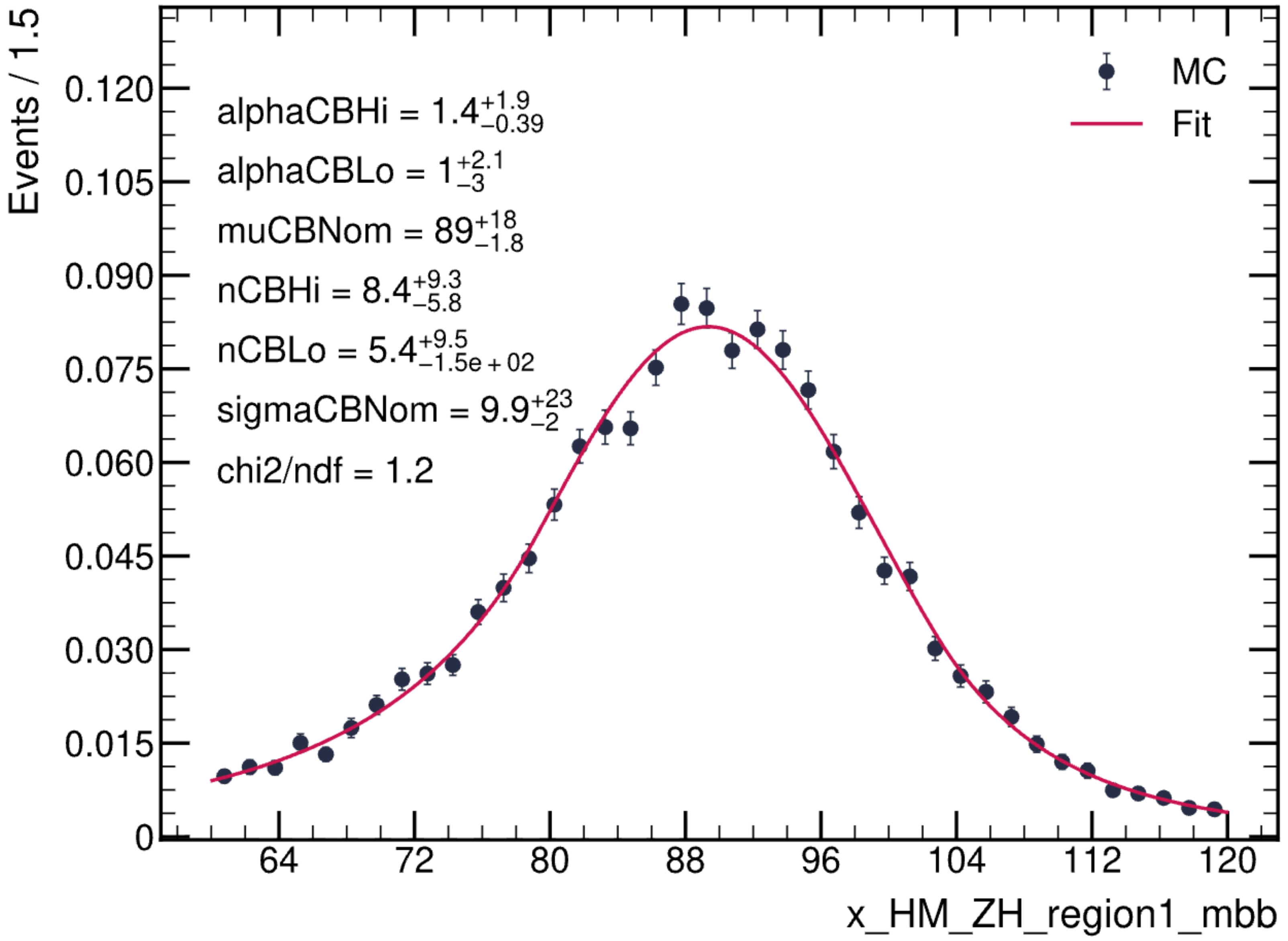


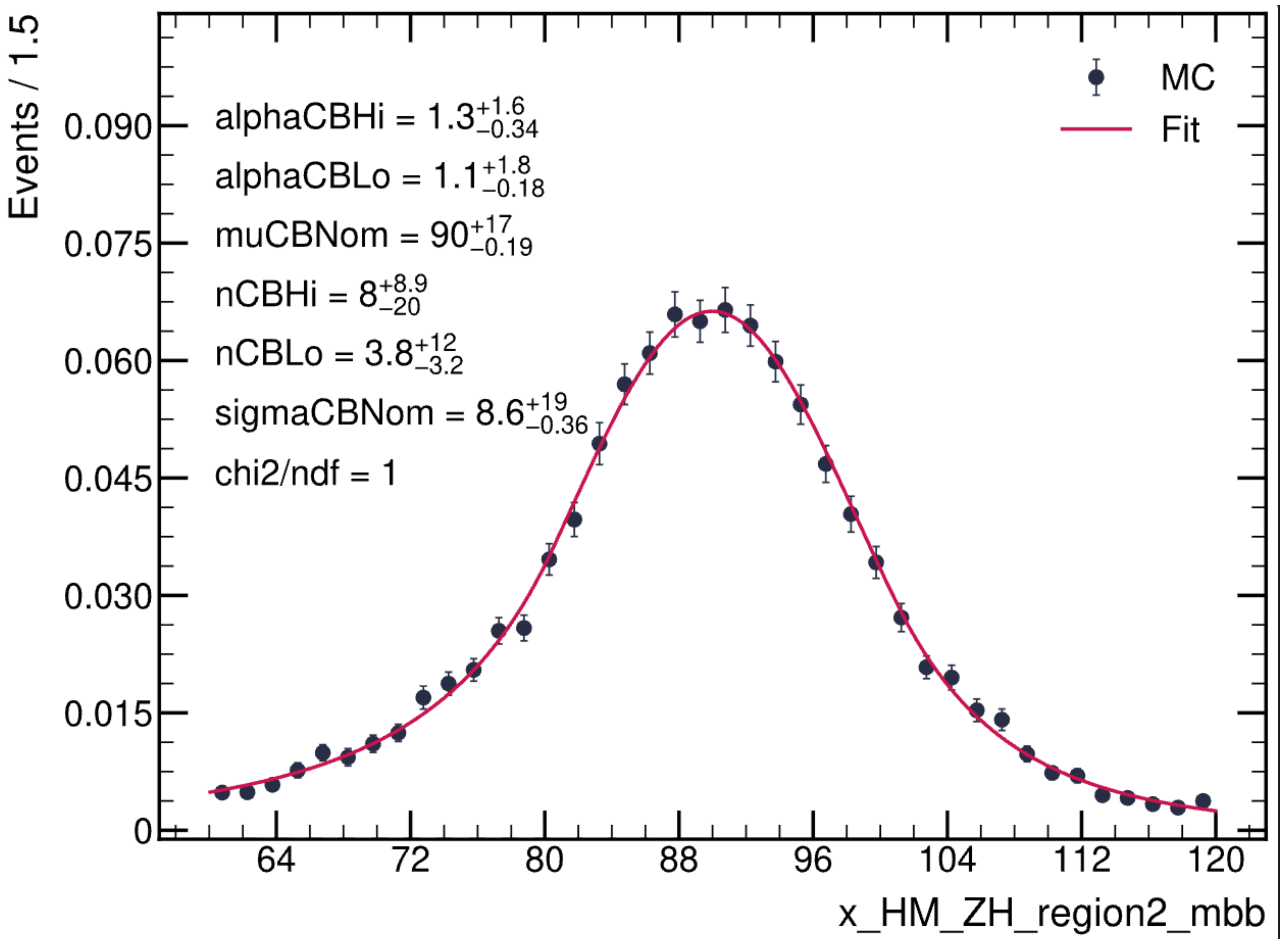


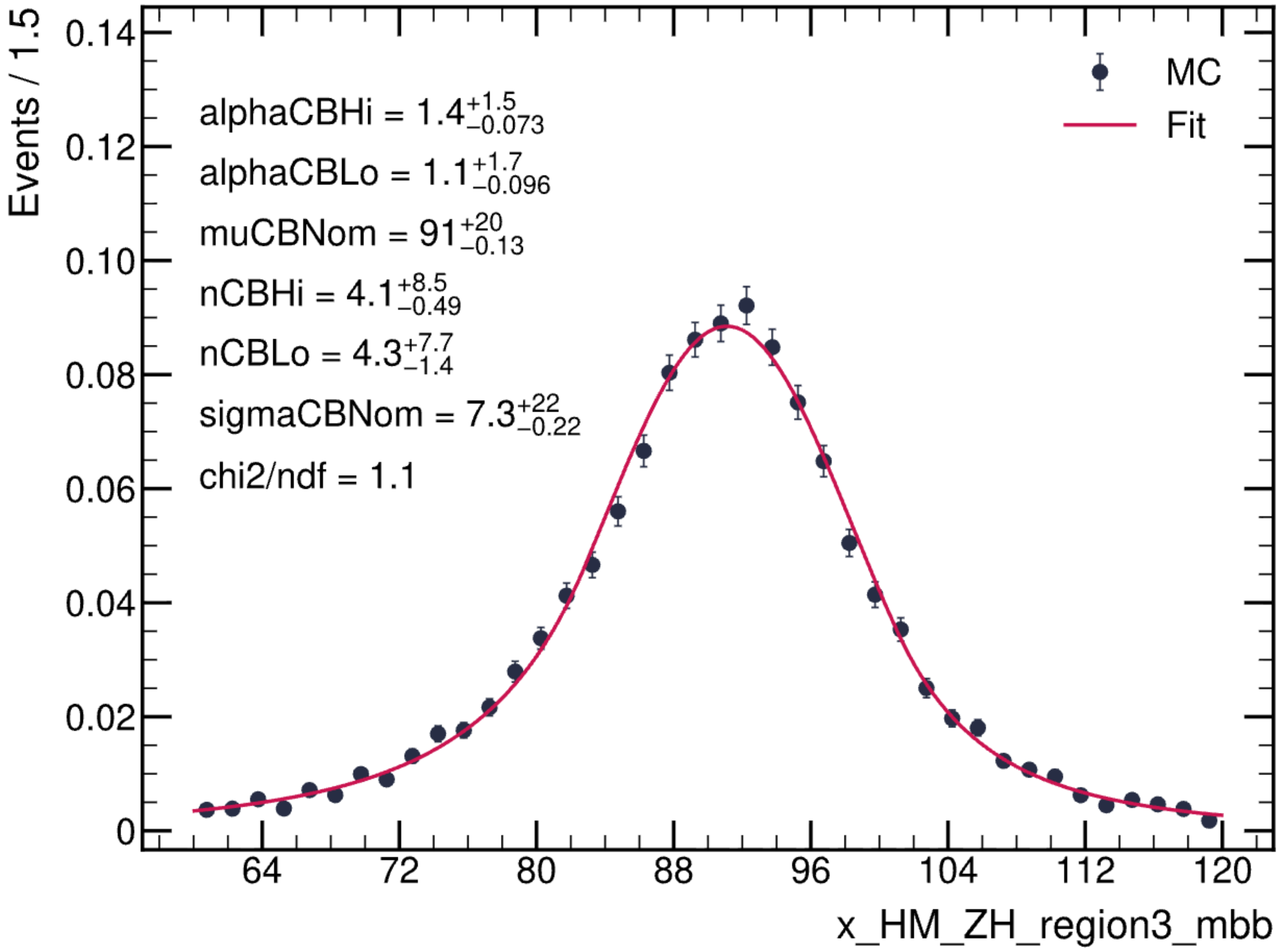


From category boundaries to stat results

- DCSB(mbb) for ZH







Results and stability with data<=24

- After training model, I prepared ntuples and workspace for others to look at
- HH SM 1.3sigma (while 1D gives 1.102) from blinded stat-only fit
- 1.268 with toys (while 1D gives 1.109)
- Unlike 1D, 2D fit allows simultaneous fit with ZH (validation of the analysis) without degrading HH results
- Stability of fit is also investigated with software I wrote
  - 9934 converged for 1D model
  - 9935 converged for 2D model

Results with data<=25

- Data 25 is available, but corresponding MC is not

- Run-3 MC is scaled by 1.69 to emulate

	μHH	μZH	μH	HH expected significance	ZH expected significance
Fit with μHH floating, and μZH=1 and μH=1	1 +/- (-0.728482, 0.97897)	-	-	1.5165091	-
Fit with (μHH, μZH) floating and μH=1	1 +/- (-0.739566, 0.9993)	1.00001 +/- (-0.67483, 0.857299)	-	1.4870104	1.6287418
Fit with (μHH, μH) floating and μZH=1	1 +/- (-0.780662, 1.05001)	-	0.999998 +/- (-0.506505, 0.589593)	1.3747727	-

The End