Leptonic MSSM $bb(H \rightarrow bb)$ status update

Antonio Branca, Ugo Gasparini, Tommaso Dorigo, Kostya Kanishev,
Stefano Lacaprara, Paolo Ronchese, Mia Tosi, Alberto Zucchetta

INFN Padova

Hbb meeting,
Hbb, 20 April 2012

Intro

Intro and selections

- Neutral SUSY Higgs, $H \to b\bar{b}$
- Use associate production to reject hadronic background $pp \to b\bar{b}H \to b\bar{b}b\bar{b}$
- Use semi-leptonic b decay for trigger;
- Baseline selections:
  - at least 1 global muon $P_T^{\mu} > 15\ GeV$, no isolation required;
  - at least 3 jets (PFak5, LooseId) $|\eta| < 2.6$, $P_T > (30, 30, 20)\ GeV$
  - $\Delta R_{ij} > 1$ for any pair $_{ij}$ of jets (**new**)
    - ★ to avoid problem in b-tag associator with overlapping jets, very limited loss in signal;
  - the first 2 jets must have b-tag $CSV > 0.8$, the μ inside one of the two;
  - last selection: third jet b-tag $CSV > 0.7$

Strategy

- Major background source is QCD
- Define a control region using a likelihood ratio discriminator using the most discriminating variables (depends on M_H)
- Build B-tagging matrices in control region;
- Estimate *bbb* background for variable x in signal region starting from *bbj* (two b-tagged jets, plus a third one with no b-tagging required):

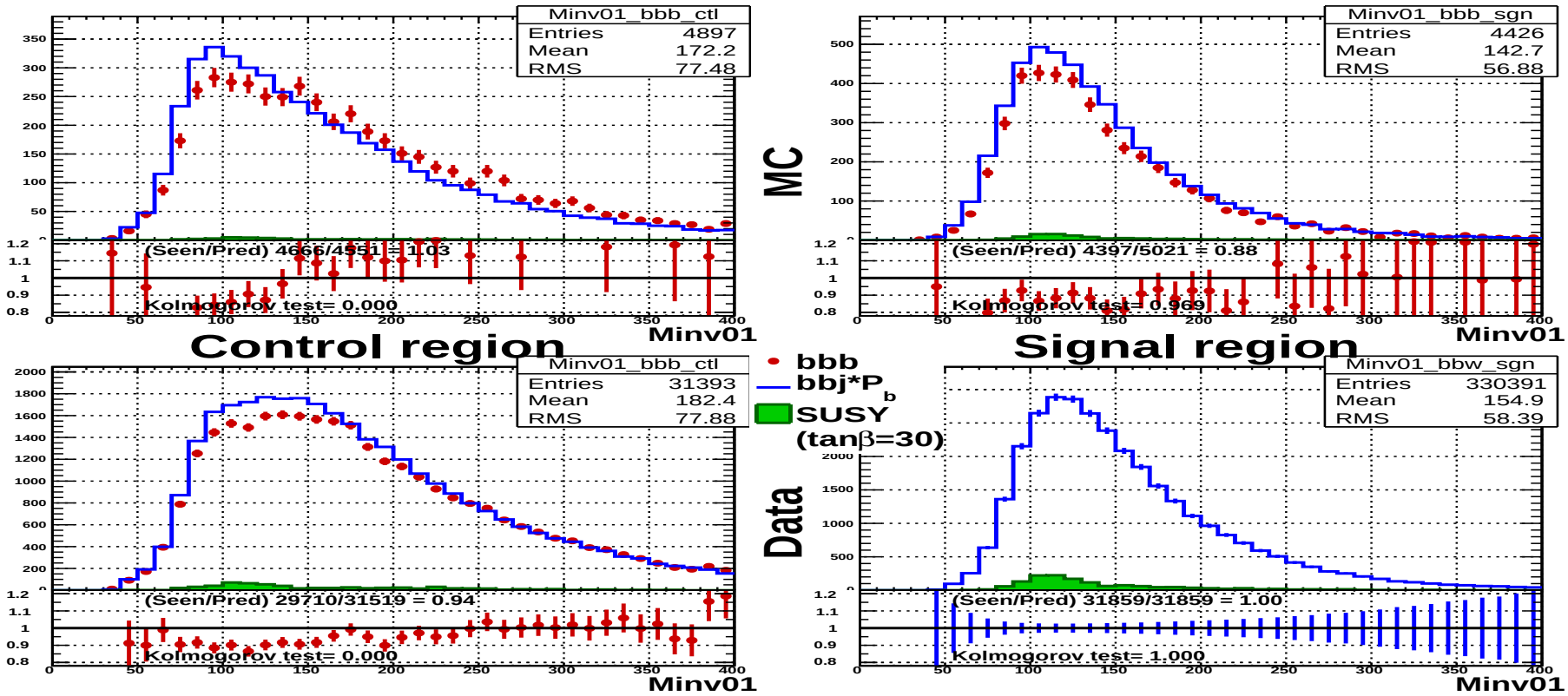
$$F(x; bbb) = F(x; bbj) \otimes P_b^{3rd-j}(j)$$

$$P_b^{3rd-j}(j) = \epsilon_b \cdot f_b + \epsilon_c \cdot f_c + \epsilon_l \cdot f_l$$

- Get $\epsilon's$ from MC and $f_{b,c,l}$ from Data
  - (fit to mass@vertex and BJetProbability);
- Compare invariant mass distribution (M_{bb});

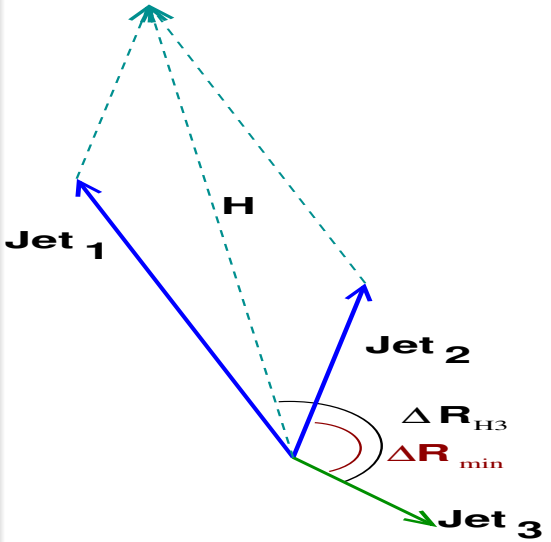
Issues last month (showing for $M_H = 120\ GeV$)

- Problem in shape for MC closure in control region;
- Over prediction for MC in signal region;
- Control region probably with still too much signal;



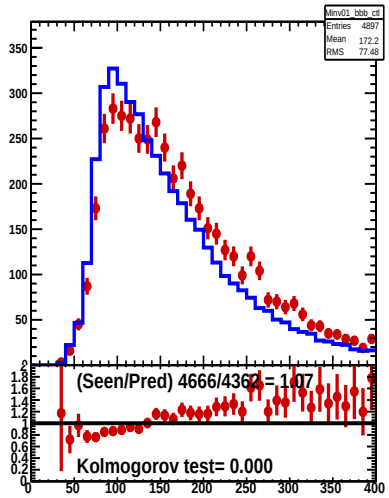
Improvements

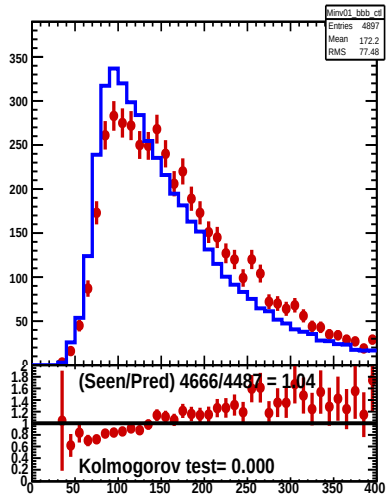
- Further optimization of discriminator
- after cutting on $\Delta R_{ij} > 1$ drop ΔR_{min} usage
- try $F_{B,C}(E_T^{(3)}, \Delta R_{H,3})$
  - See right for $\Delta R's$ definitions $\Rightarrow$
- try $F_{B,C}(E_T^{(j3)}, \Delta R_{H,j3}) \times F_{B,C}(\Delta\eta_{j1,j2})$
- assuming no correlation
  - no statistics to do fit with full correlation
- $F_{B,C}(\Delta\eta_{j1,j2})$ only for shape, normalized to 1.
  $\int_{C\ reg.} \frac{dN}{d\Delta\eta_{12}} \cdot F_{B,C}(\Delta\eta_{12}) d\Delta\eta_{12} = \int_{C\ reg.} \frac{dN}{d\Delta\eta_{12}} d\Delta\eta_{12}$

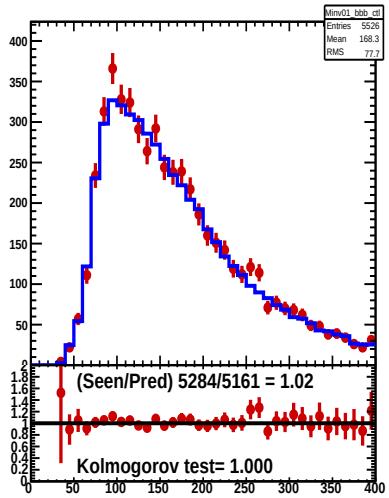


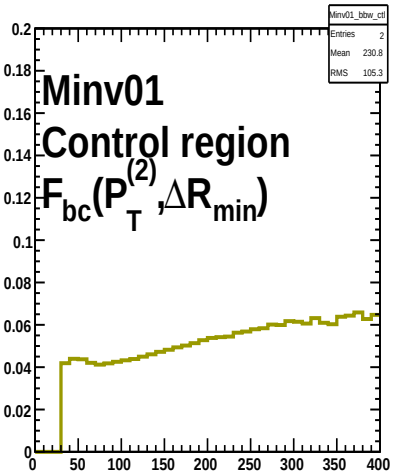
New parametrization

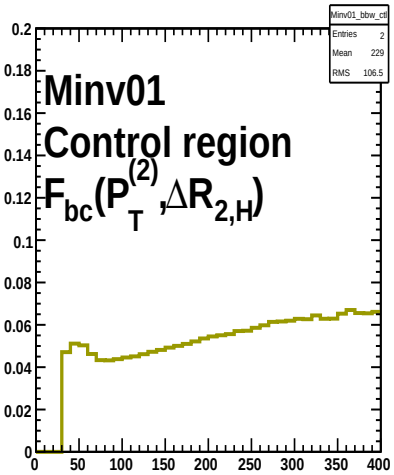
Look only to control region, comparing different parametrizations

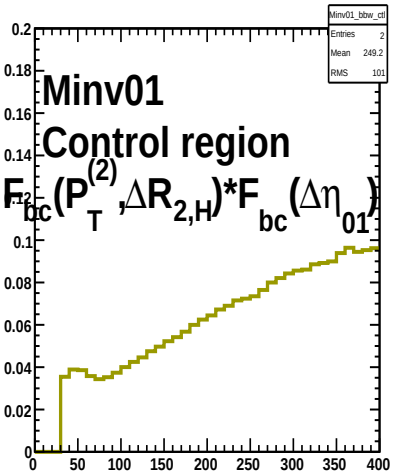






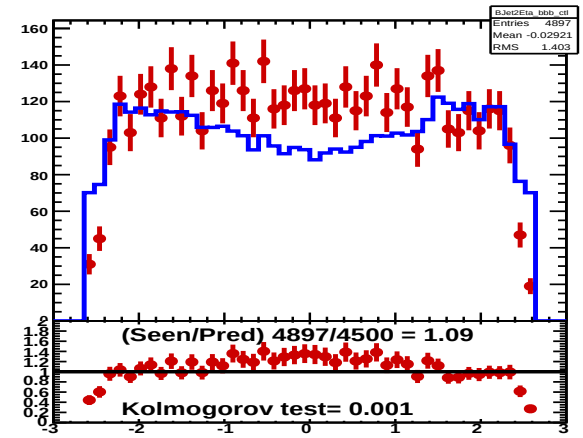


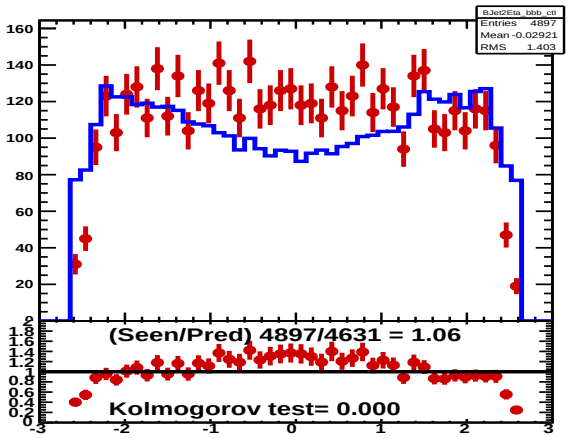


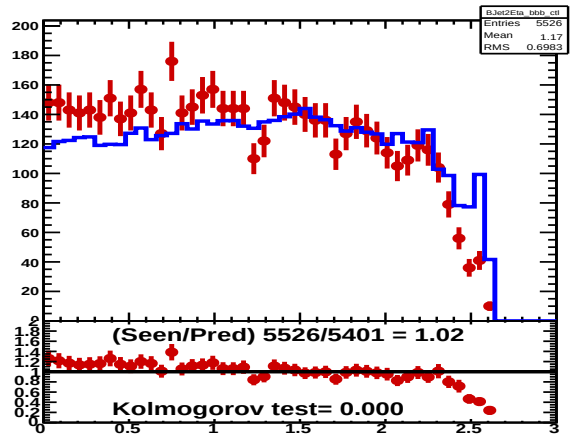


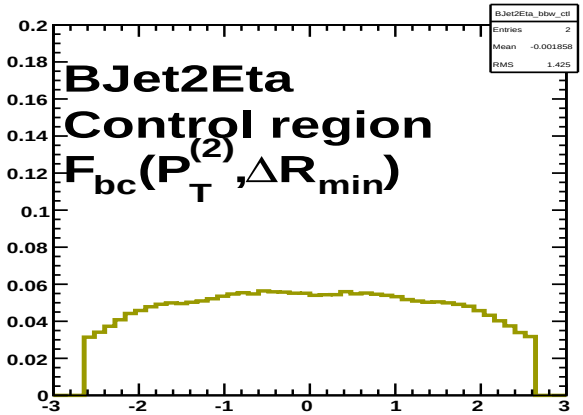
Top:
$(bbj) \cdot P_b^{3rd-j}(j)$ vs (bbb)
bottom: just $P_b^{3rd-j}(j)$
Very good shape and normalization agreement!
Other *good variables* in the backup
Some have still problem...

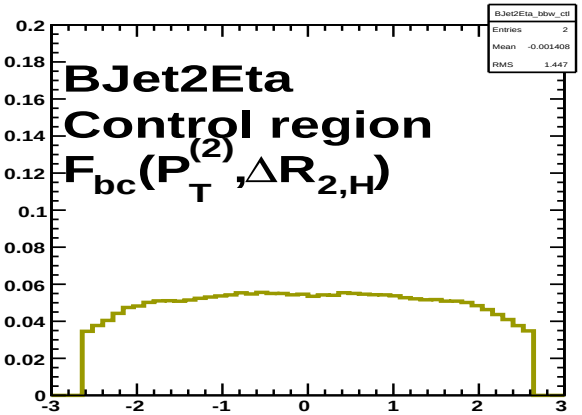
Comparing different parametrizations: $|\eta^{(3rd\ jet)}|$

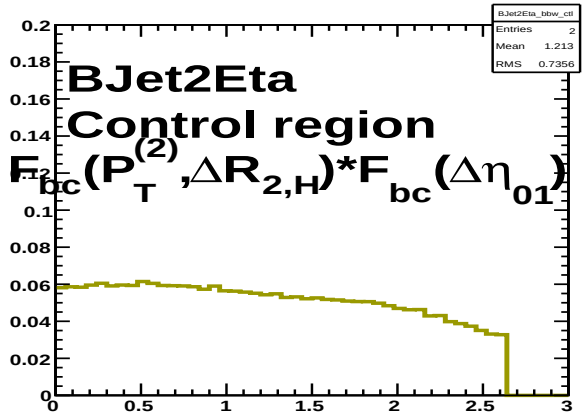








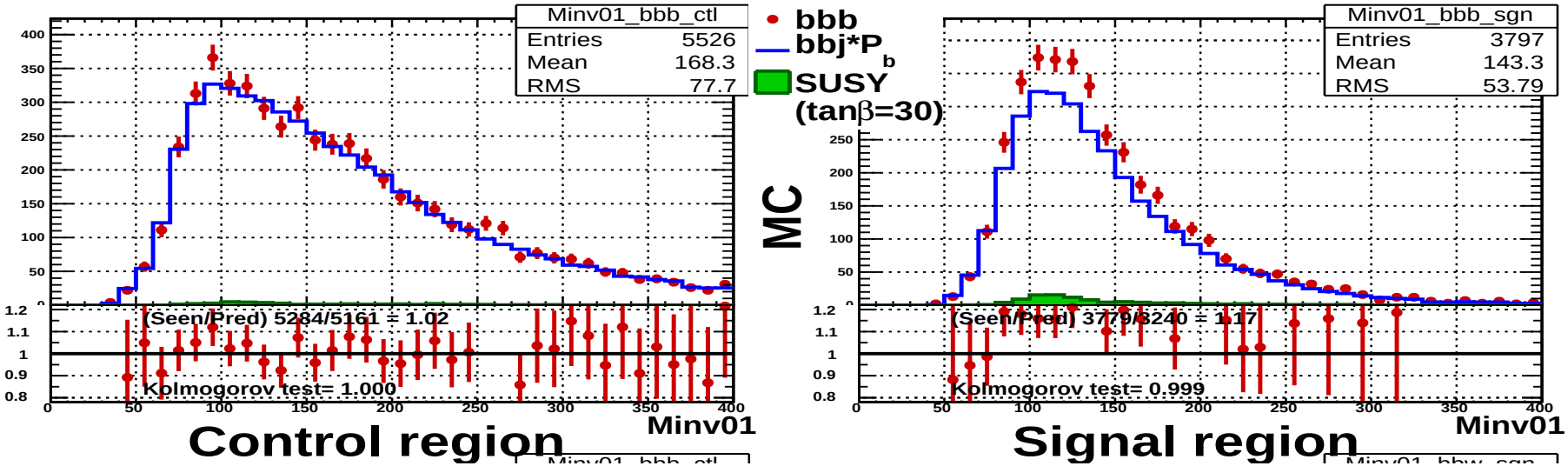




Still some problem

Note η for first two params, $|\eta|$ for last one.

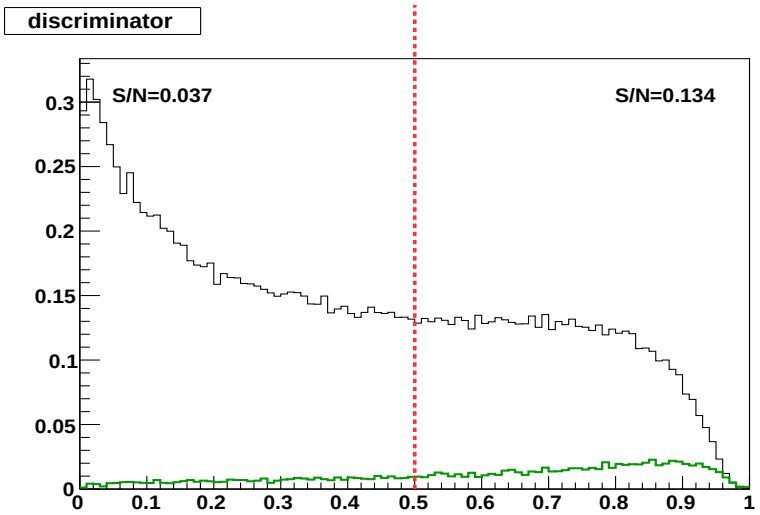
Signal region

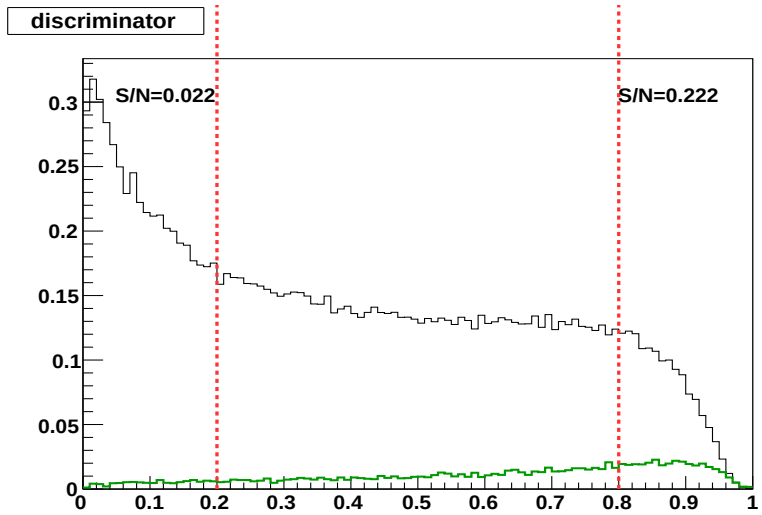


- Bad news: normalization is off by 17%
- Good news: shape is very well predicted! (kolmogorov $P \approx .999$)
- still work to do, but progressing (we have some ideas)
- Can we buy the normalization correction from MC?

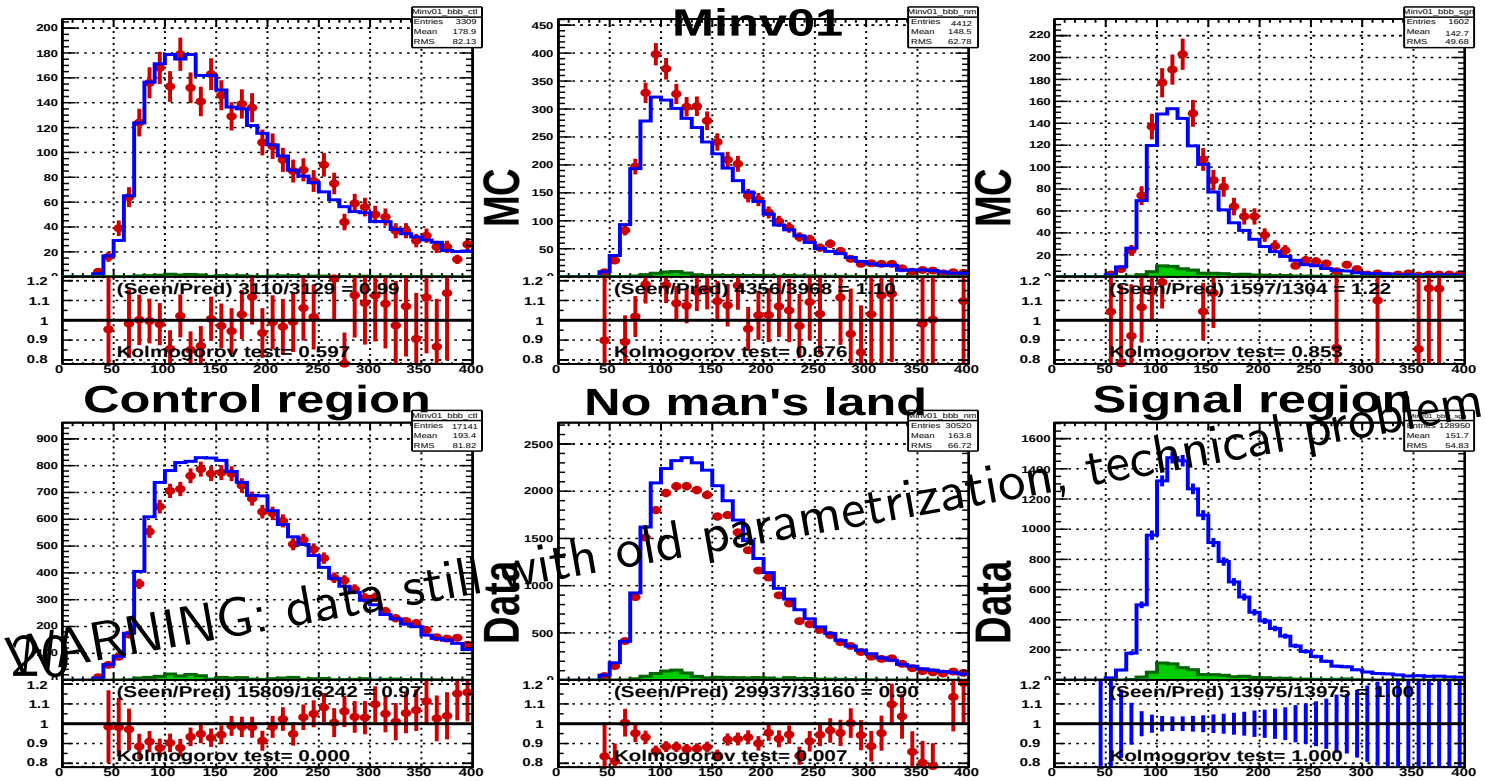
Can we buy normalization from MC? No man's land

- Can we buy normalization from MC? well, NO.
- we can try to check on data, tough.
  - divide phase space not just in two region (control and signal)
  - but in three (control, signal, and no man's land in between)
    - ★ Probably the one shown is too large
- can use control to get $F_{B,C}$
- no man's land to check it and compare behaviour with MC
- then predict on signal region





Plot No man's land



Shape is well predicted in all three regions

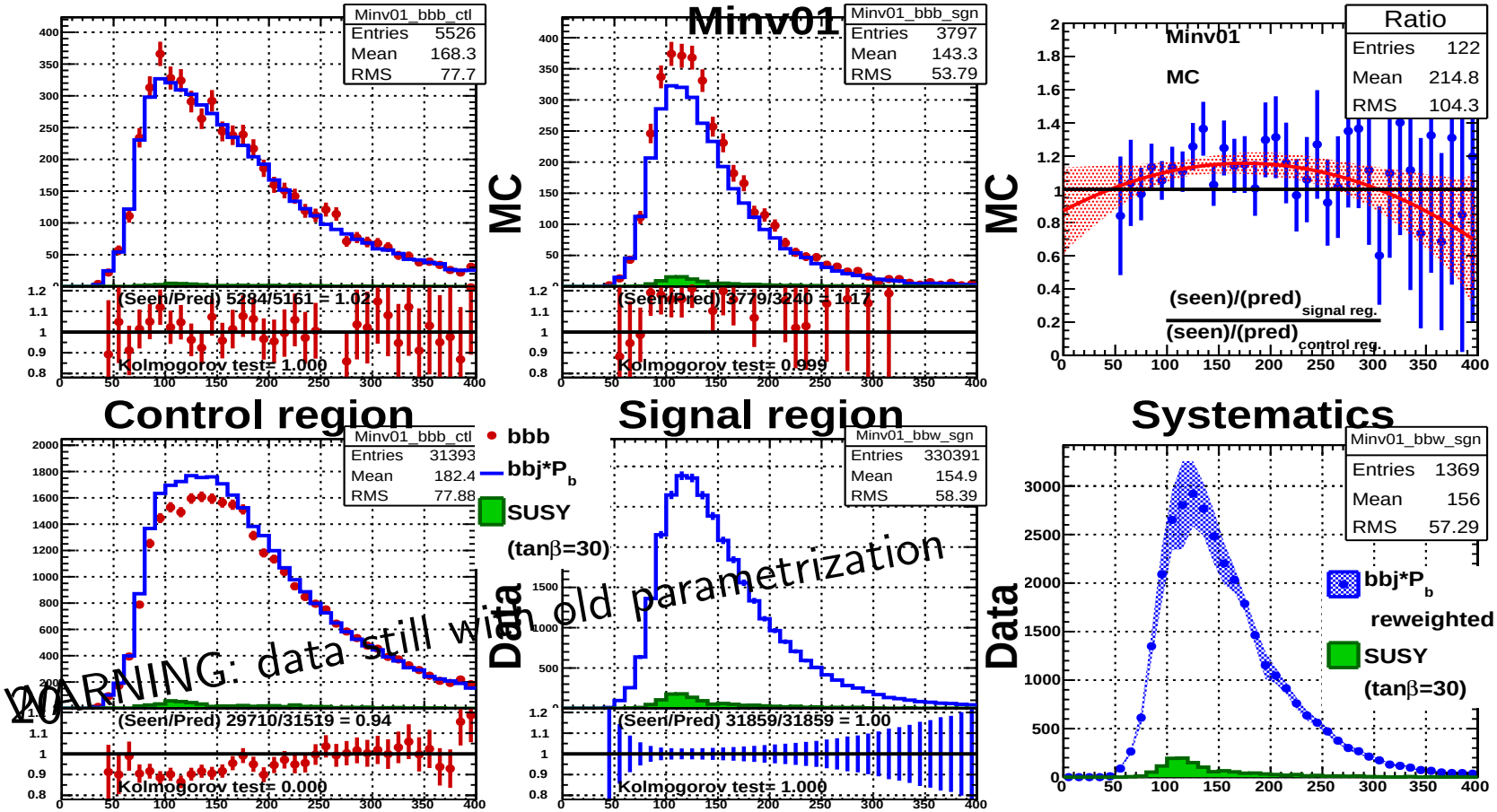
There is a trend for normalization ($0.99 \rightarrow 1.10 \rightarrow 1.22$)

No man's land has too much signal (and too large)

Systematics

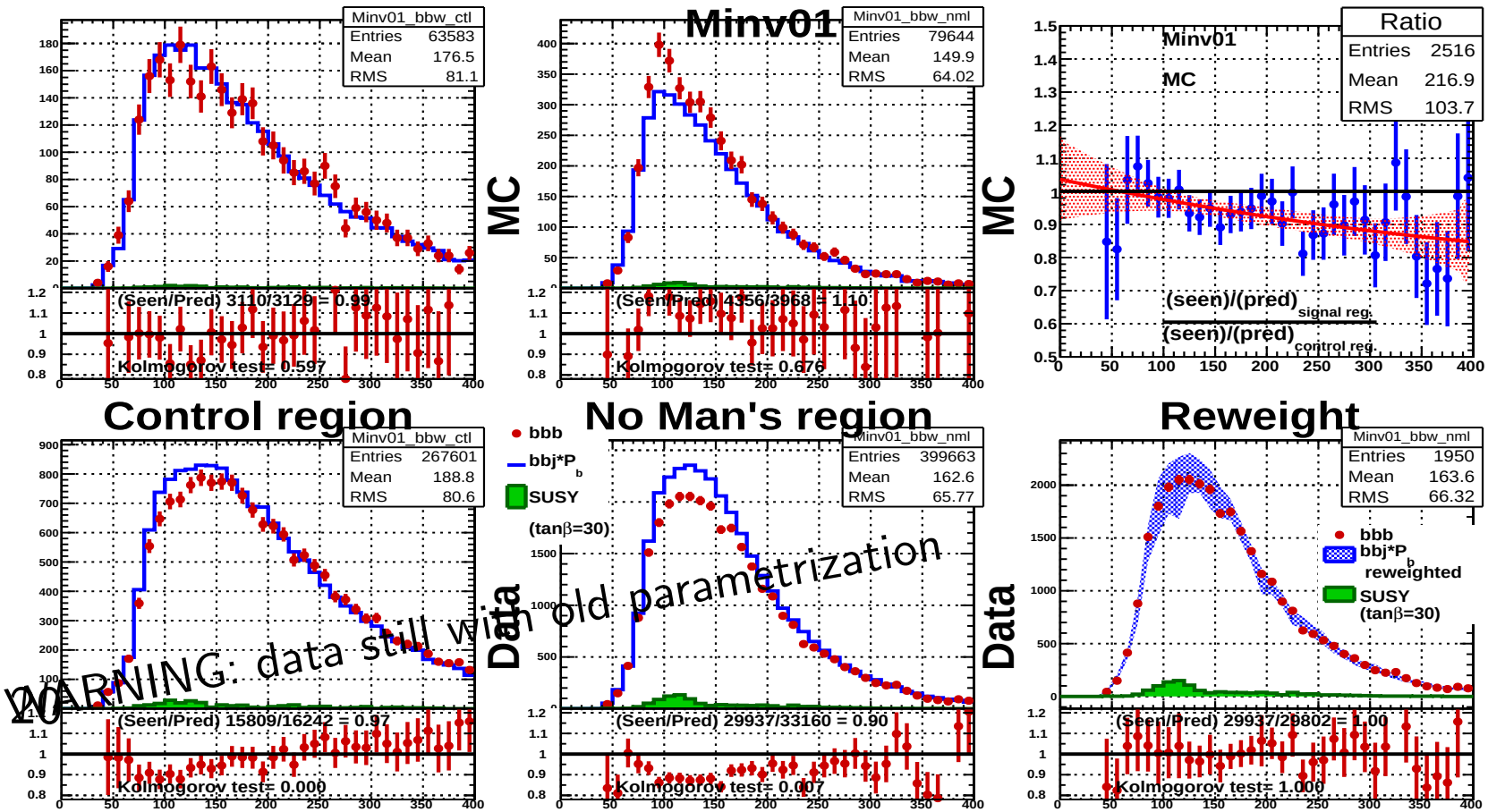
- Two major source of systematics coming for the predicted *bbb* in signal region:
  1. Systematics from *bbb* prediction from DATA control region
     - ★ compare *bbb* and $bbj \times P$ in control region and use the difference as systematics;
  2. Systematics due to extrap. from control to signal region from MC
     - ★ get ratio of ratios from MC (signal/control) and fit it;
     - ★ use fit results to correct extrapolation bias (?);
     - ★ and fit errors to estimate systematics for extrapolation;
- both can be used bin per bin when computing CL's
- we can use *no man's land* in data to check extrapolation systematics and bias in data;

Systematics test



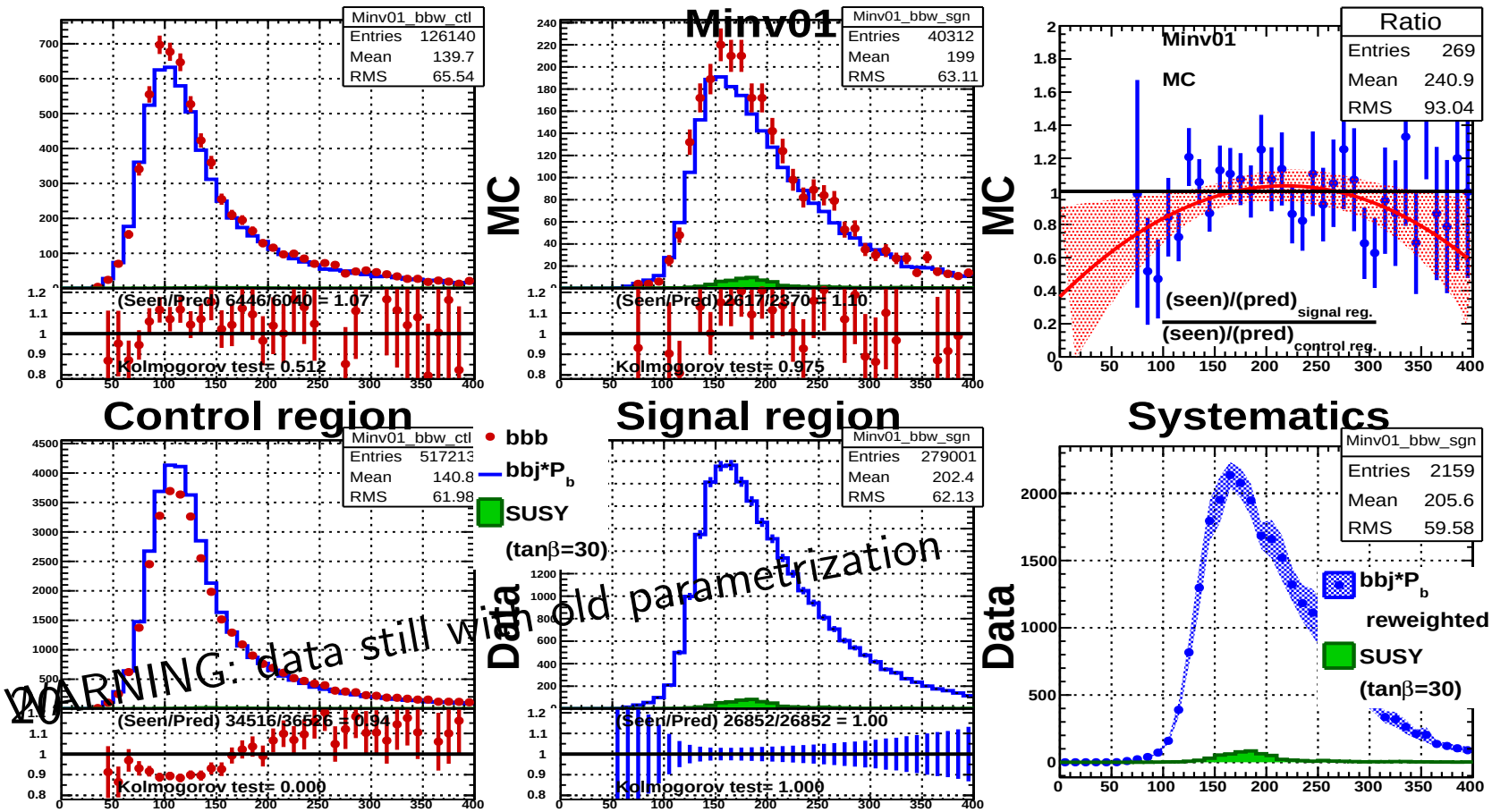
WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**

Check Systematics and bias in NML



WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**

For Higgs Mass 200 GeV

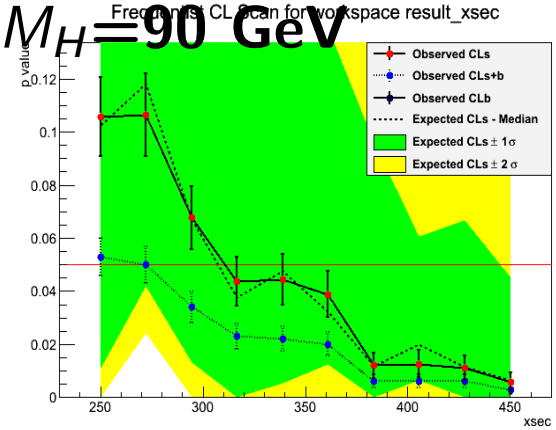


WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**

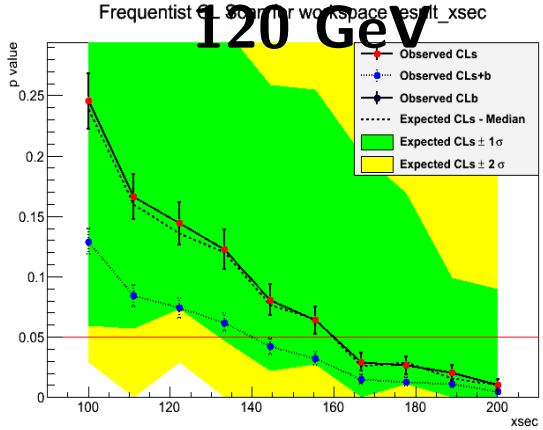
CL for mass points

Still with 5% systematics flat, and old parametrization

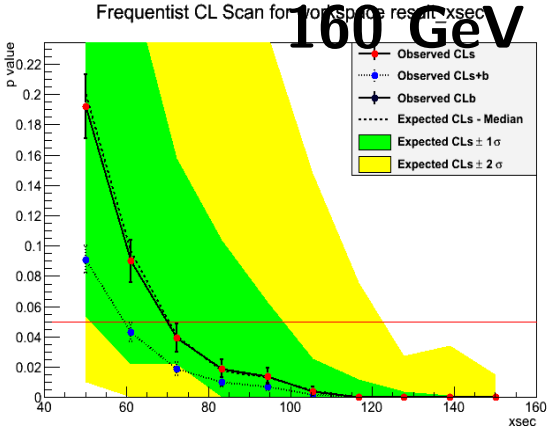
M_H=90 GeV



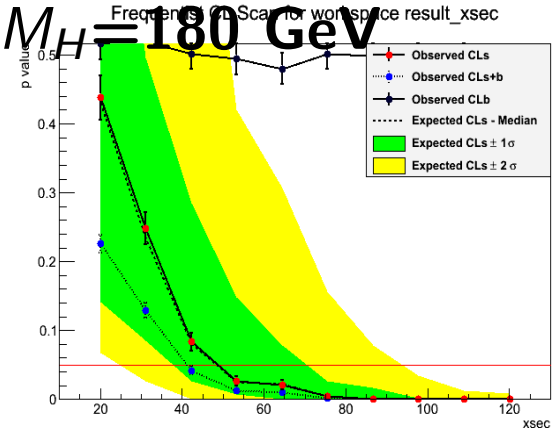
120 GeV



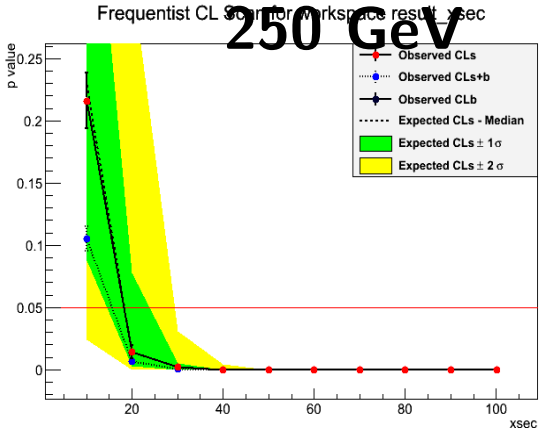
160 GeV



M_H=180 GeV



250 GeV



Exclusion 95% CL

M_H	σ (pb)	tan β
90	310	20
120	160	25
160	70	28
180	50	30
250	18	33

Very Preliminary

Summary

- Progress in MC closure test
  - almost Ok in control region
  - still normalization issue in signal one.
  - no man's land between control and signal to check extrapolation on data
- Systematics uncertainties' determination presented
- CL's for different mass points

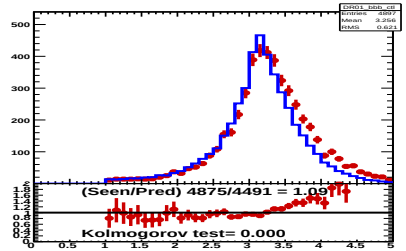
What's next (by May 4th)

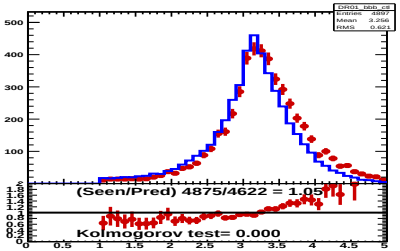
- New parametrization on data
- control region vs no man's land checks finalized for different mass points
- full CL prediction exercise performed with *final* systematics

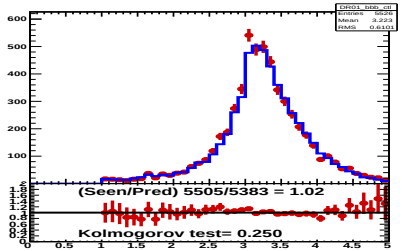
BACKUP

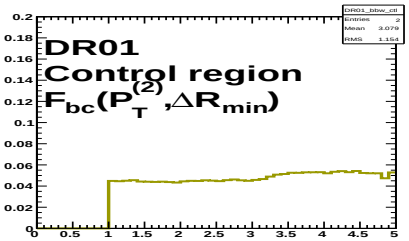
BACKUP

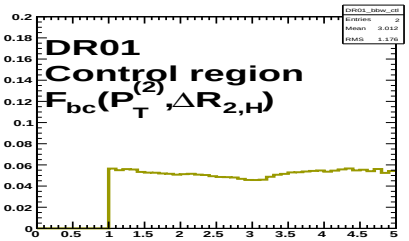
comparing different parametrizations

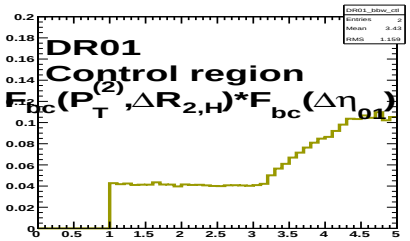


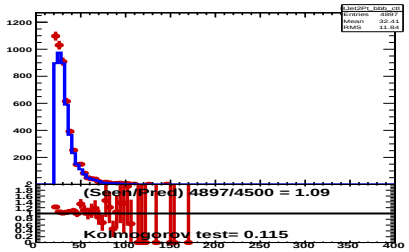


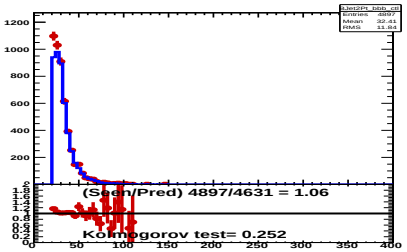


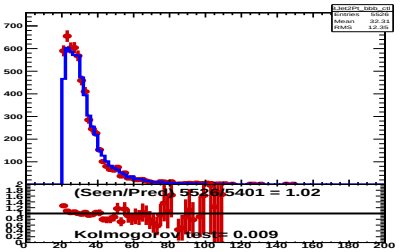


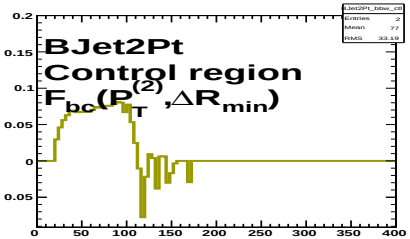


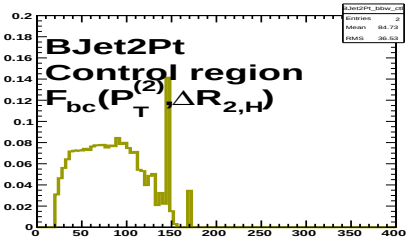


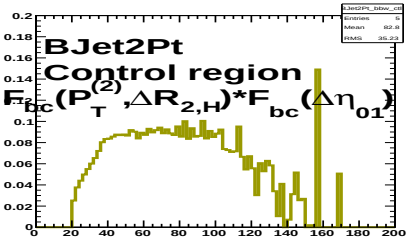




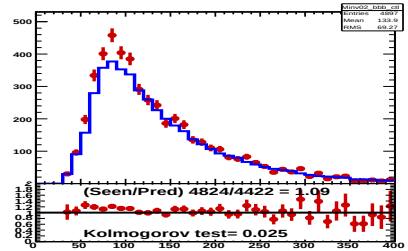


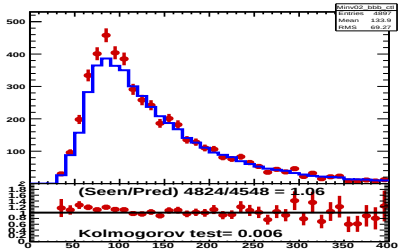


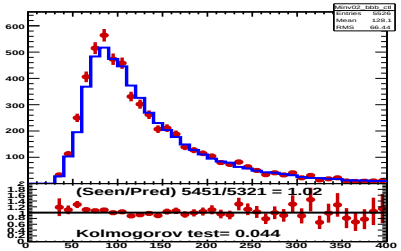


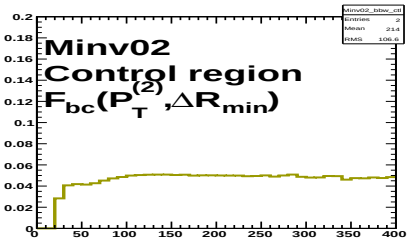


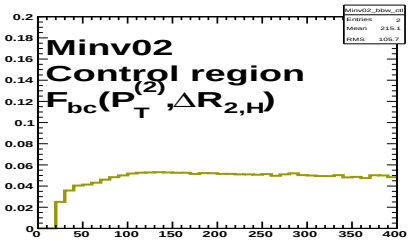
comparing different parametrizations

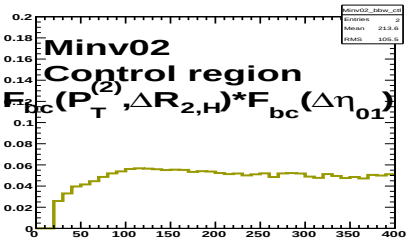


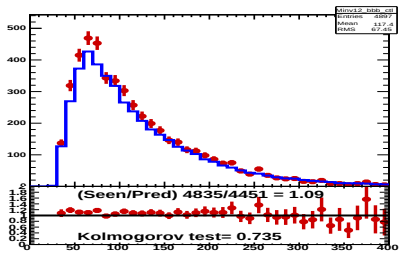


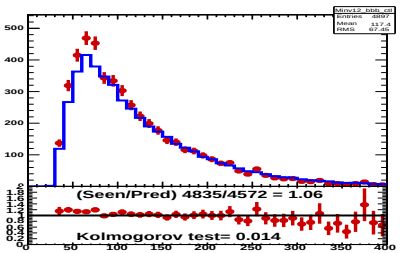


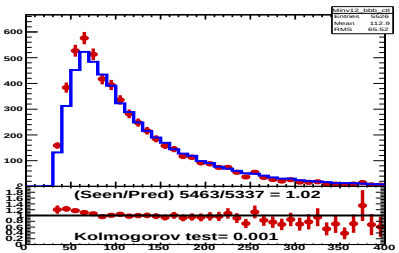


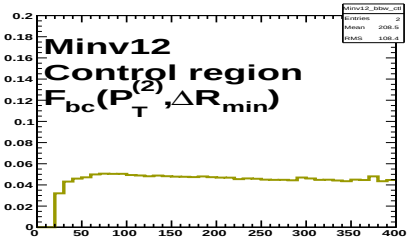


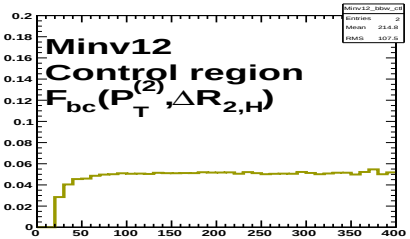


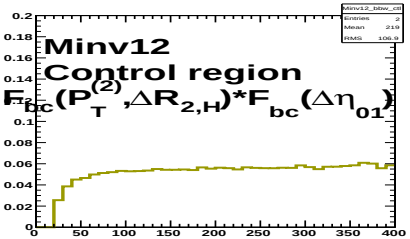




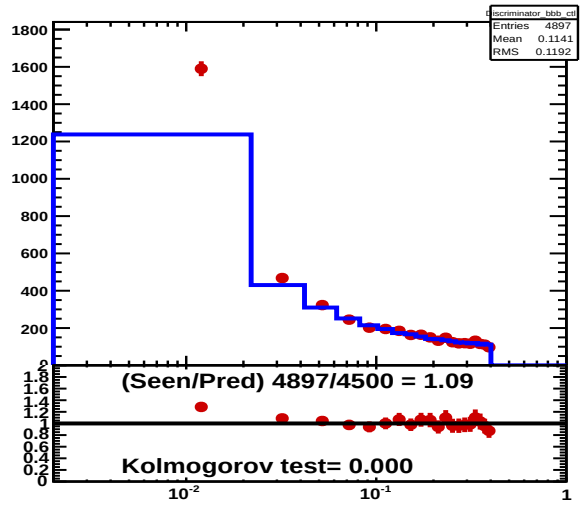


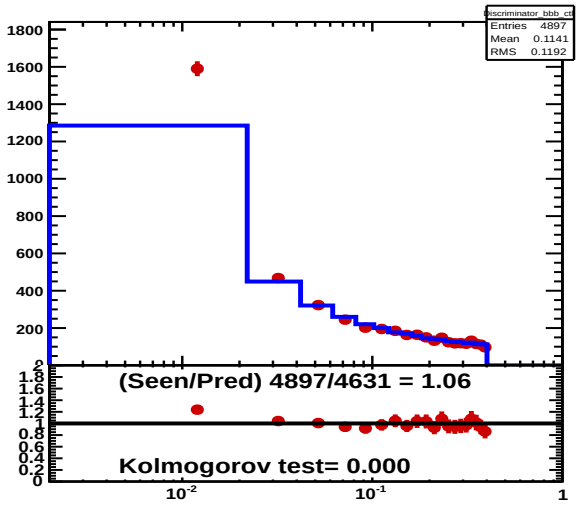


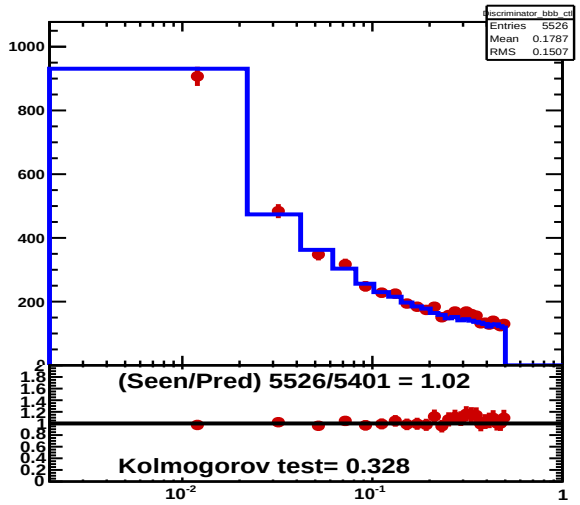


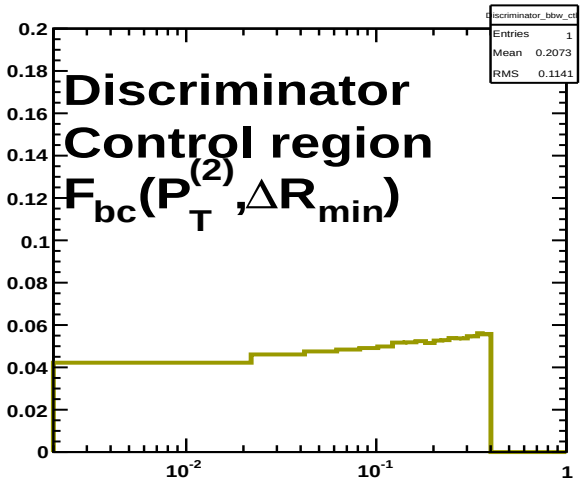


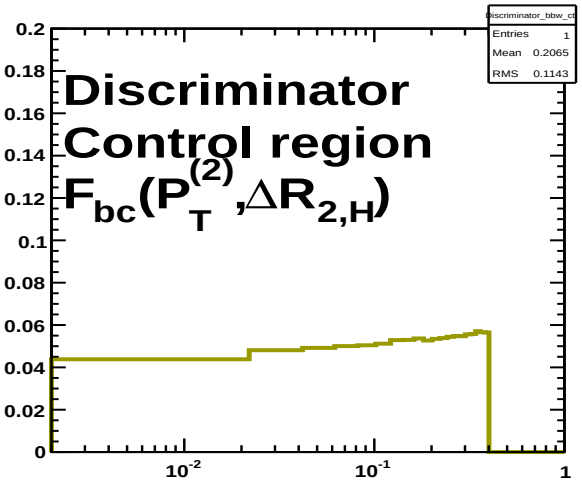
comparing different parametrizations

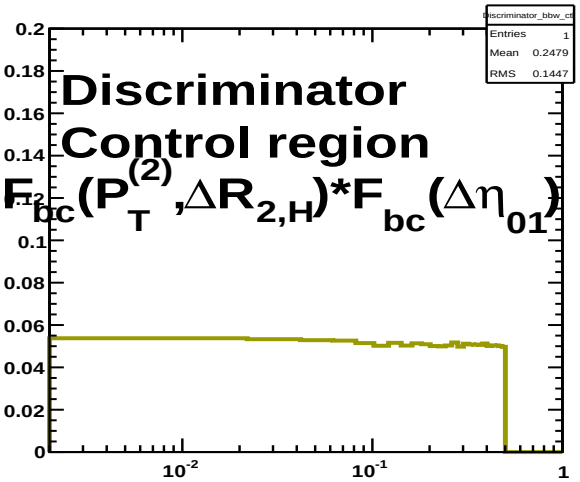




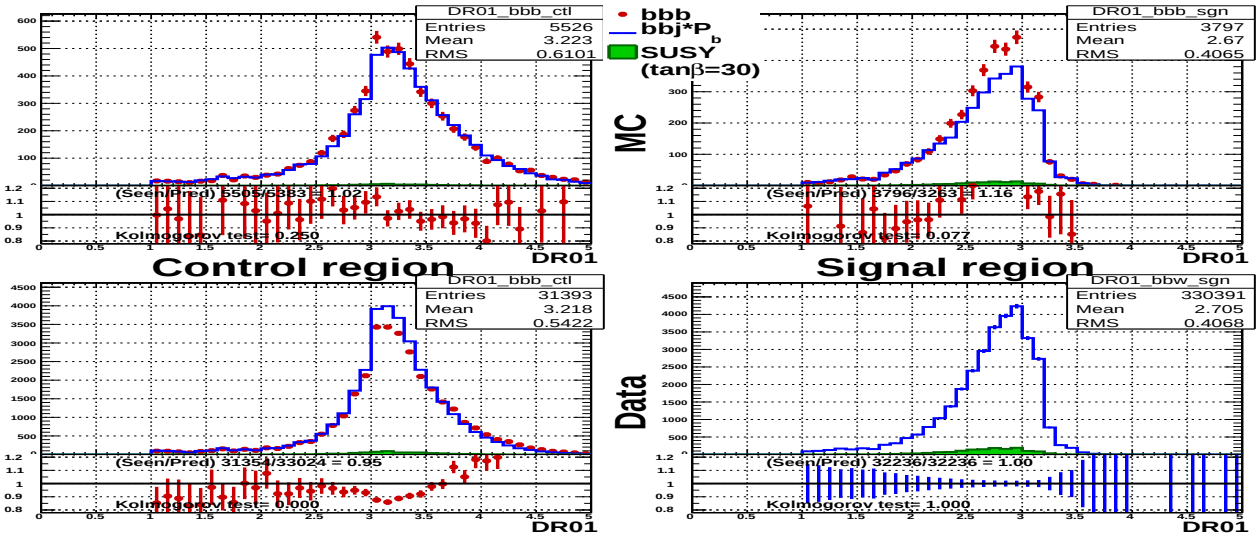


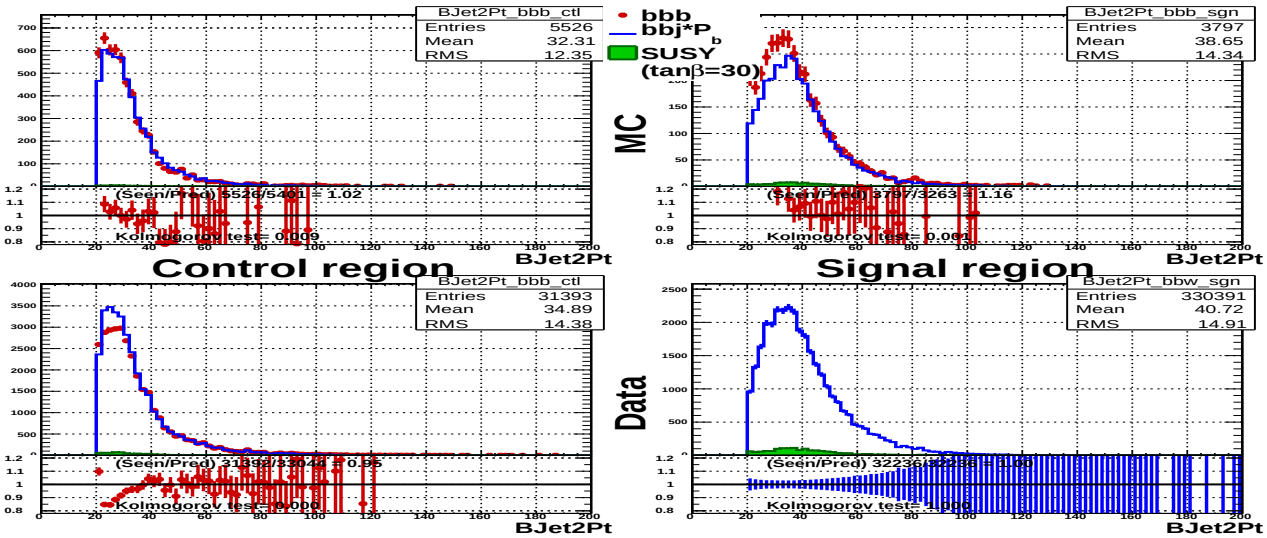




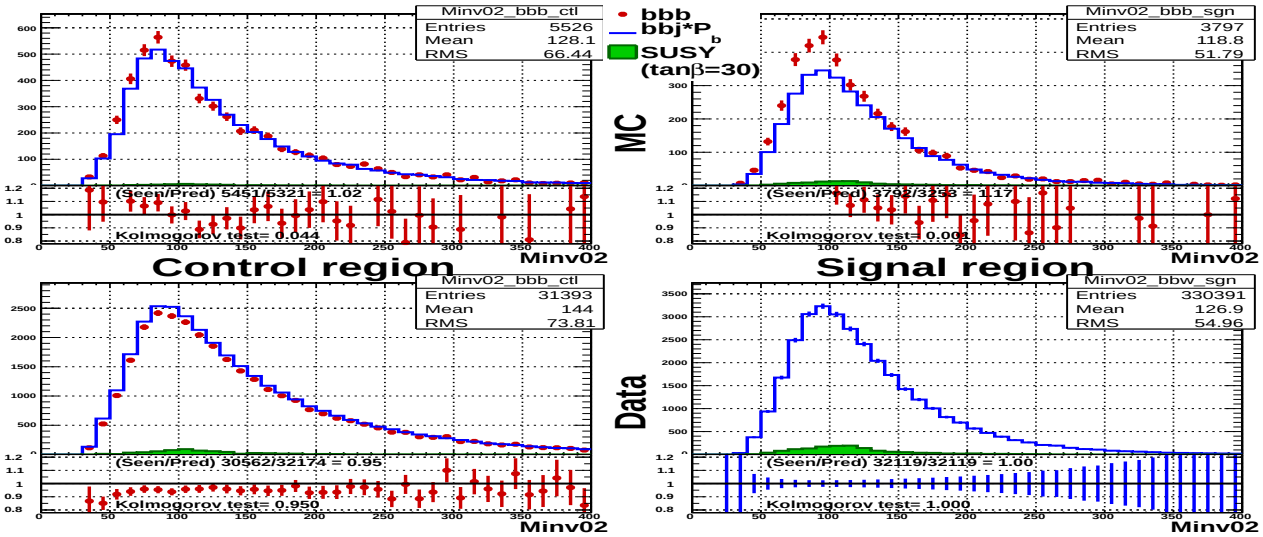


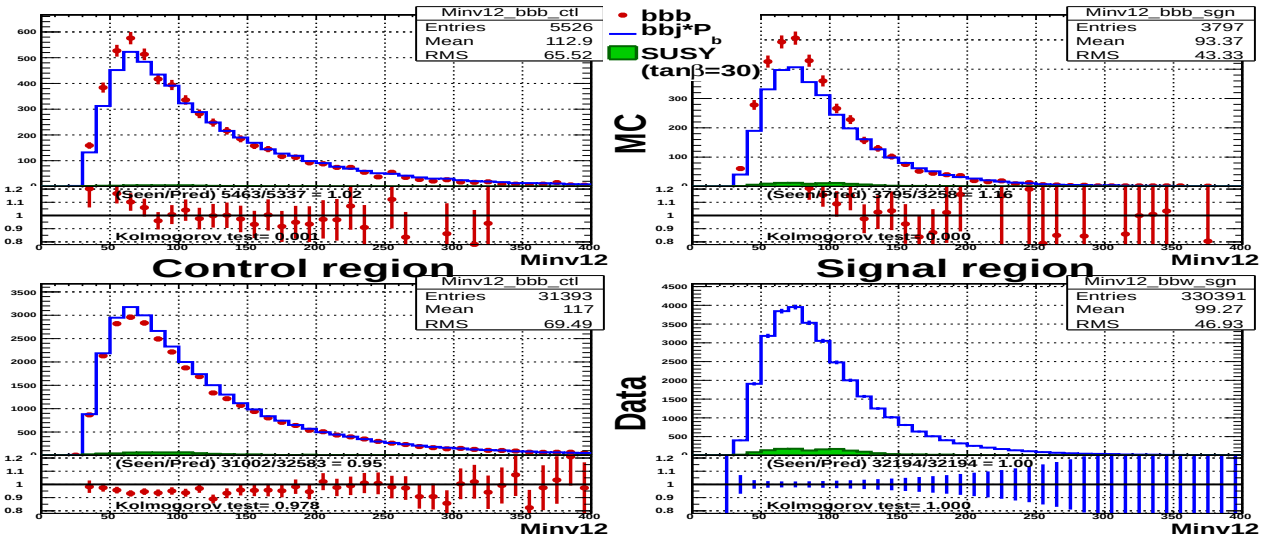
Predictions



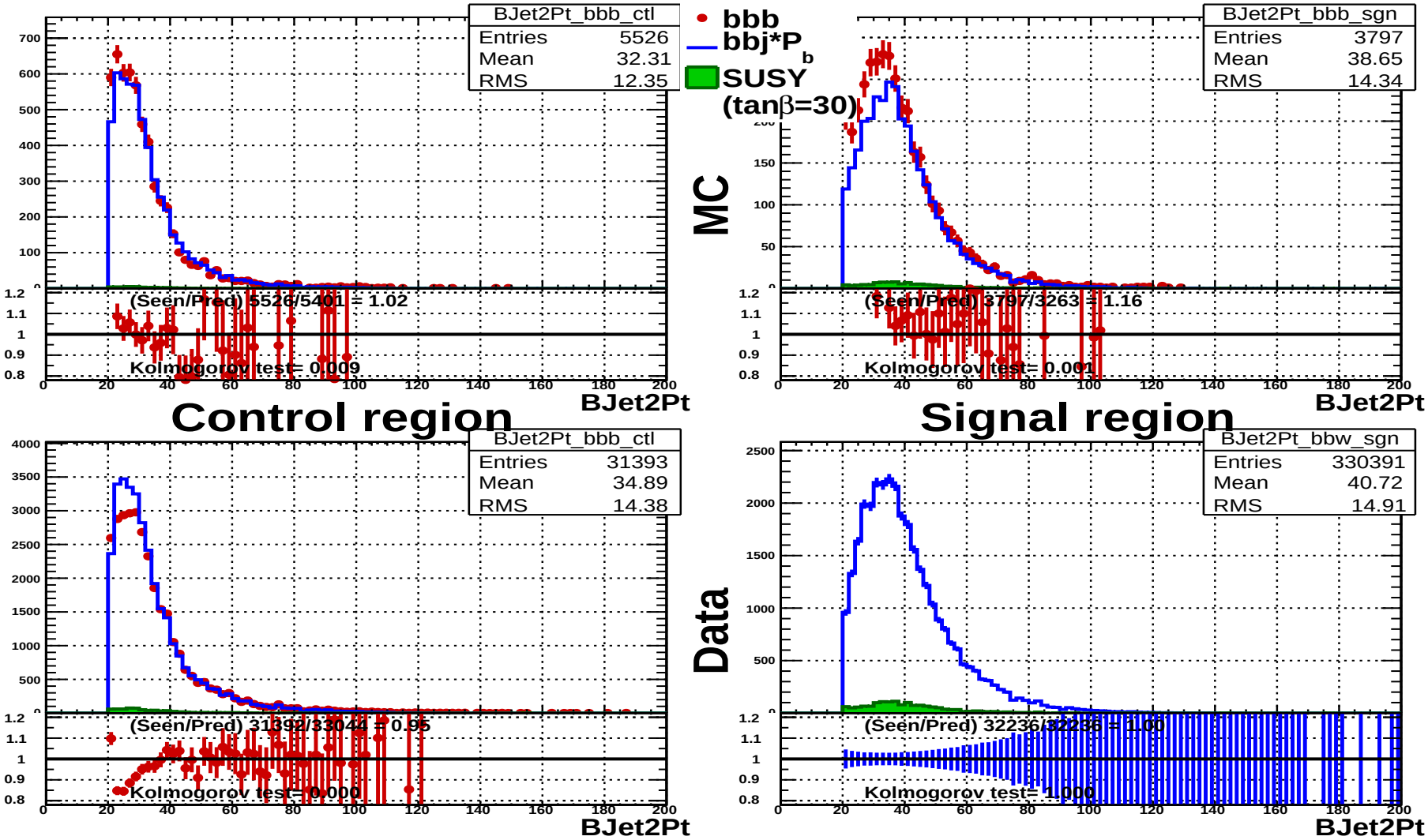


Predictions

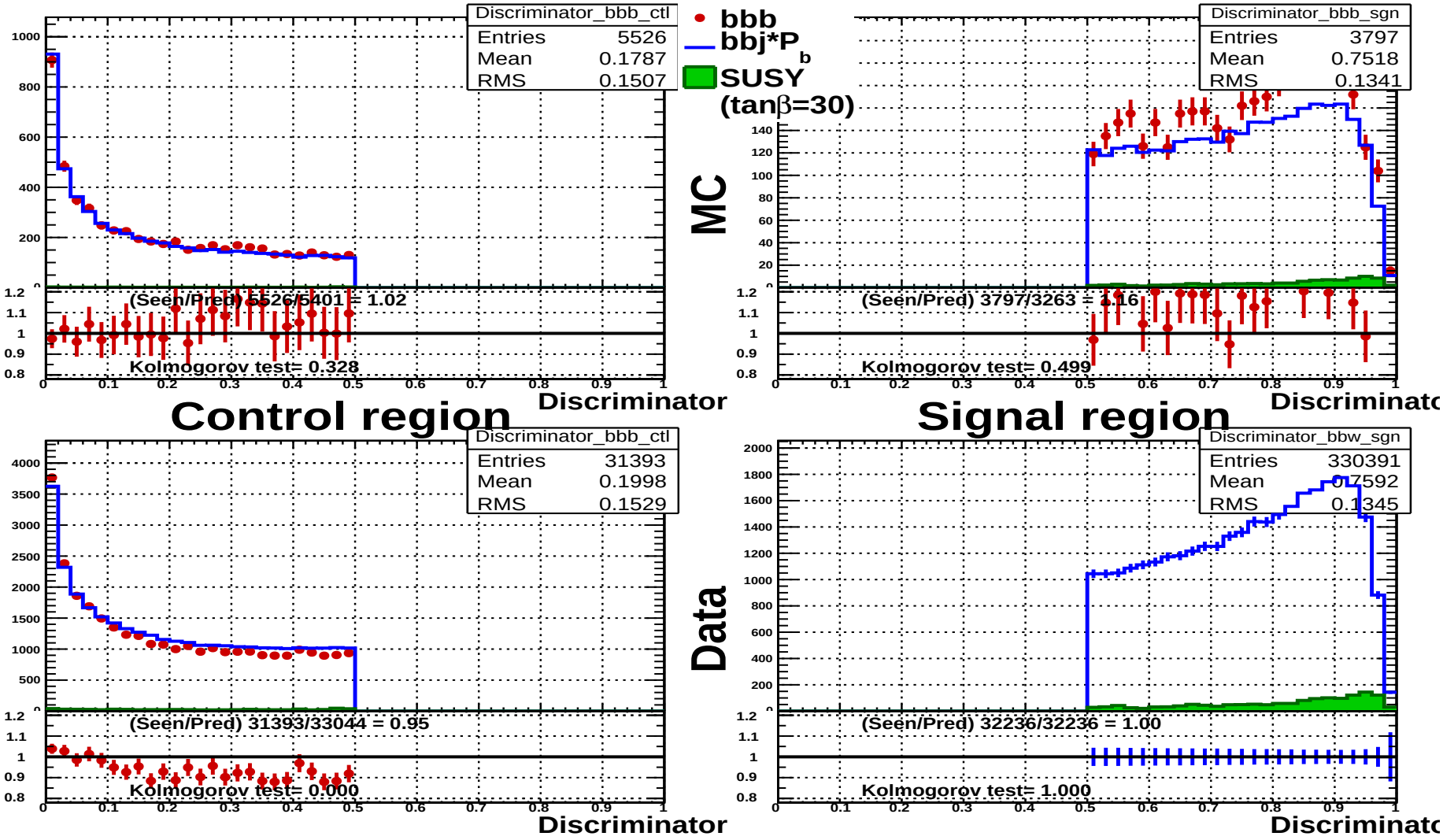




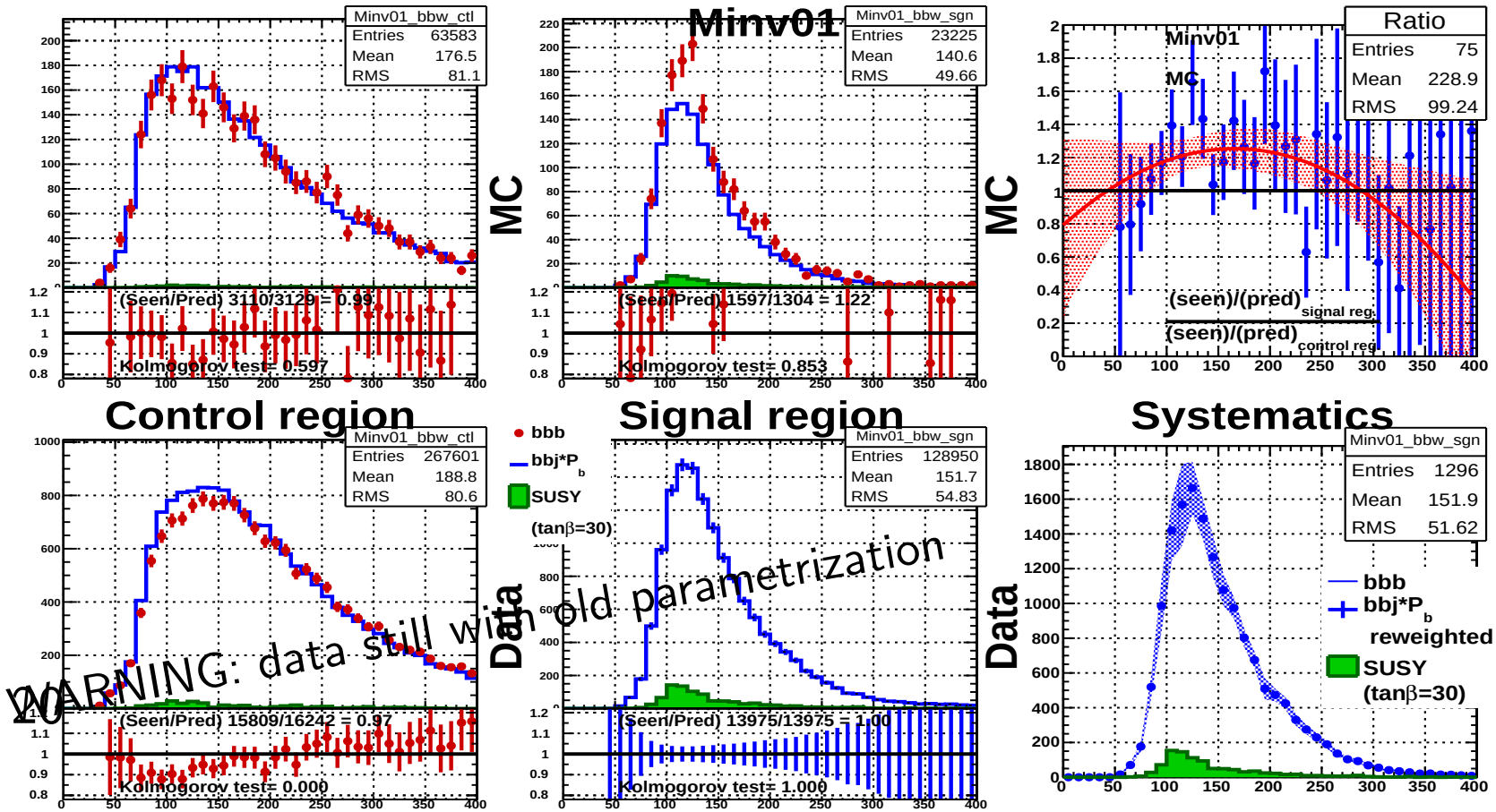
Predictions



Predictions

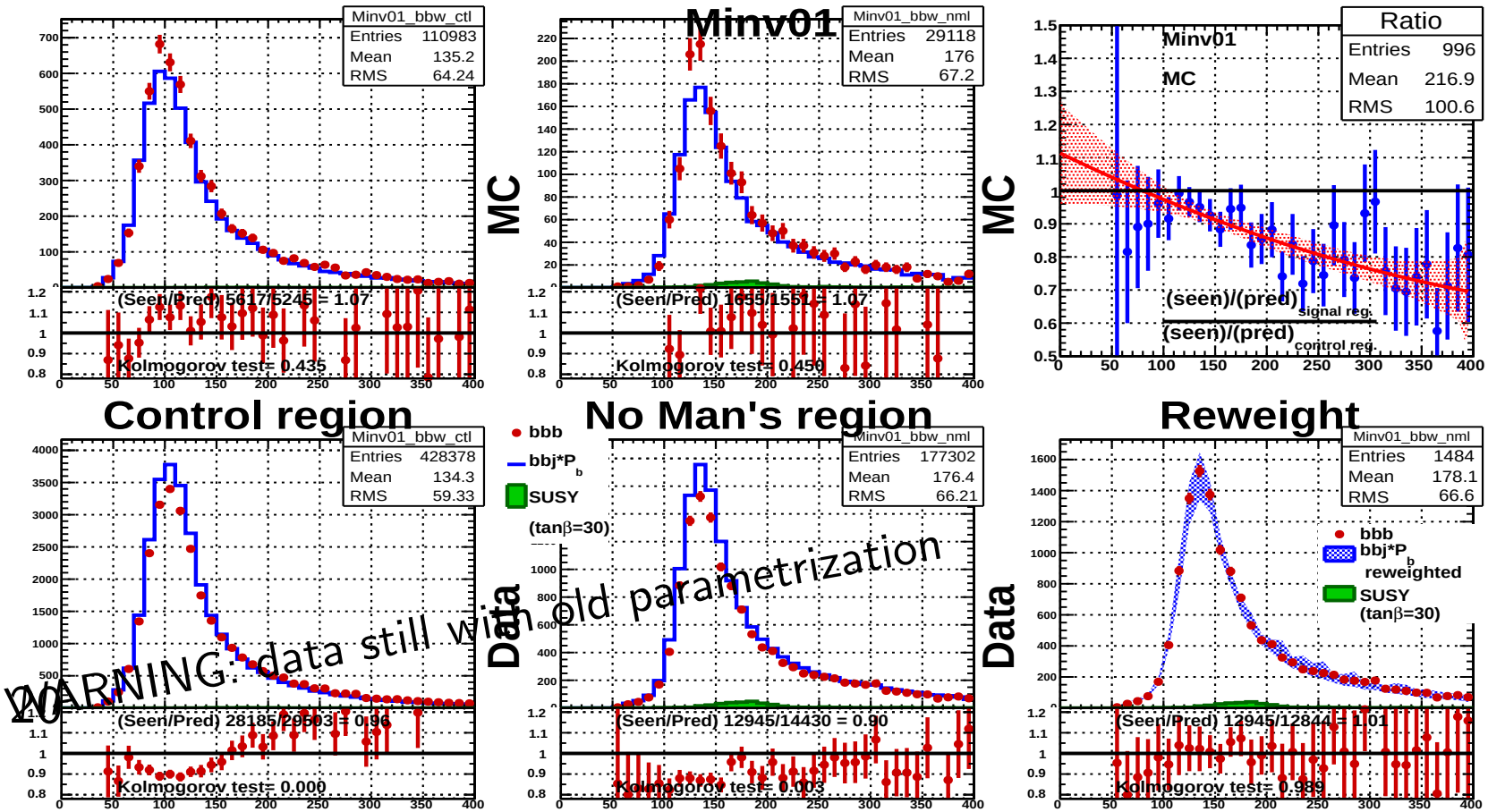


Systematics and bias with no man's land MH=120



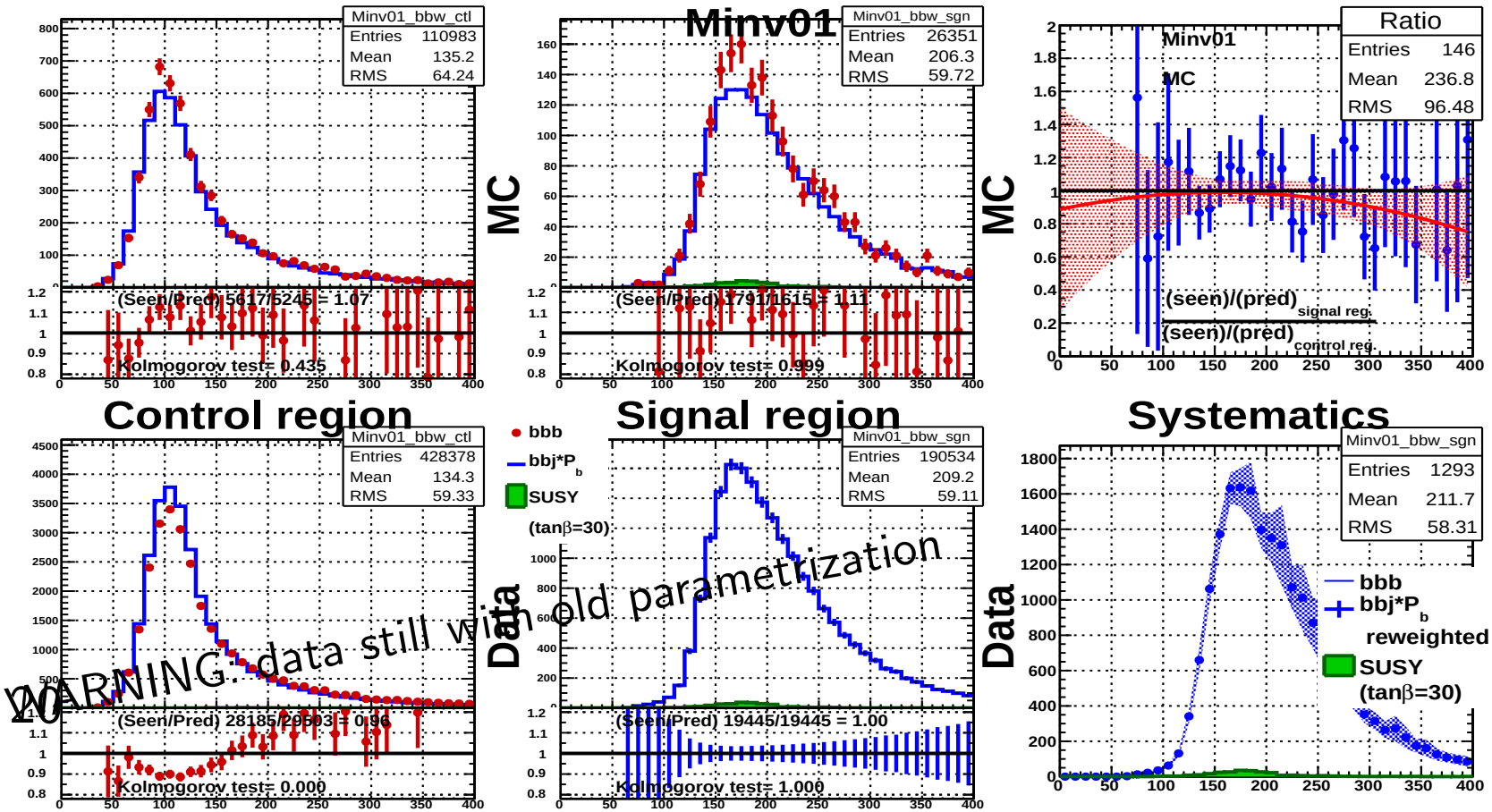
WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**

Systematics and bias with no man's land MH 200 GeV



WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**

Systematics and bias with no man's land MH 200 GeV



WARNING: large systs in above plot from discrepancies seen-predicted in DATA control region, due to old parametrization. **It is better!**