MSSM $bb(H \to bb)$ semileptonic

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

INFN Padova

Higgs Review,
CERN, 16 May 2012

Intro: motivations

- Search for Neutral SUSY Higgs, $H \to b\bar{b}$;
- Large $BR(H \to b\bar{b}) \approx 90\%$
- huge multijet QCD background:
  - Use associate production to reject hadronic background
  - $pp \to b\bar{b}H \to b\bar{b}b\bar{b}$
- Compete with $H \to \tau\tau$ channel:
  - larger yield,
  - larger background (QCD),
  - different channel.

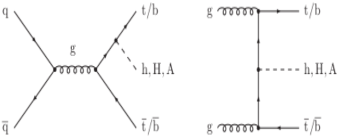
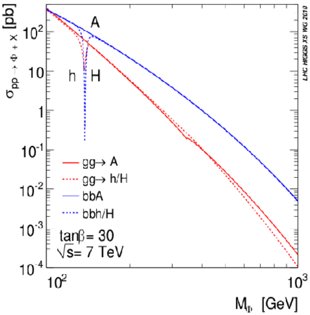


Fig. 17: Typical diagrams contributing to $q\bar{q}/gg \to Q\bar{Q} + h/H/A$ $(Q = t, b)$ at lowest order.

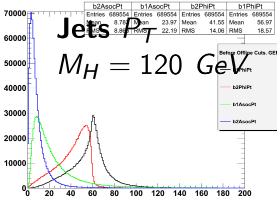


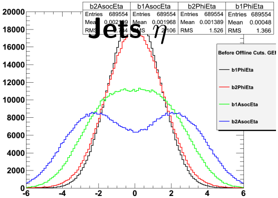
Analysis Strategy

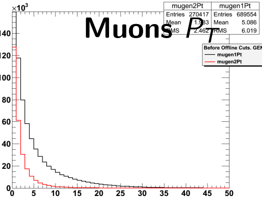
- Three b final state: $H \rightarrow bb$ plus additional associated b
- trigger is critical:
  - Use semi-leptonic (muon) b decay for trigger: muon+jets+b-tagging
- Major irreducible background source is multijet QCD
- Data driven background estimate from bbj sample
  - define signal-poor control sample;
  - get b/c-fraction of 3rd jet from mass & lifetime fits
  - combine with MC b-tagging efficiency to derive B-tag probability
  - weight bbj events to estimate number of 3-b-tags in signal region
  - Alternative approach with near neighbours method: hyperball.
- Use reconstructed mass of leading jet pair as signal-sensitive variable in final fit
- Use only 2011 data

Trigger Strategy

- Use semi-leptonic b decay for trigger: **muon + jets + b-tagging**







Use different trigger Path in 2011 to cope with increasing $\mathcal{L}$

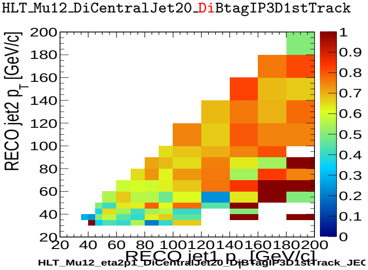
HLT paths	runs	triggers	$\int \mathcal{L}dt$ [pb^{-1}]
`Mu12_CentralJet30_BtagIP`	163738-165633	3 027 717	183.815
`Mu12_DiCentralJet30_BtagIP3D`	165970-172952	4 532 555	524.904
`Mu12_DiCentralJet20_DiBtagIP3D1stTrack`	167039-173198	1 873 247	780.378
`Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack`	173236-178380	4 762 858	1944.527
`Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack`*	178420-180252	2 164 634	811.861
All		16 361 011	4245.485

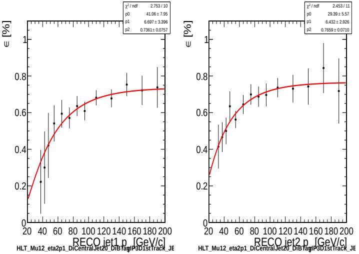
Trigger Efficiency vs Analysis

- $\epsilon_{trigger} = \epsilon(\mu) \times \epsilon(b - jets)$
- **SingleMu PD**, select all events passing a single muon path.
- Apply selection
- Build Turn on curves vs first and second B-jets Pt:

$$\epsilon = \frac{\text{Hbb path \& (pre)sel \& SingleMuHLT}}{\text{(pre)sel \& SingleMuHLT}}$$

- Turn-on stable wrt SingleMu threshold
- `Mu12` turn-on indipendently computed from prescaled SingleMu





Analysis selection

Selections

Baseline selections:

- Trigger
- 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

bbj the first 2 jets must have b-tag $CSV > 0.8$

- the μ inside one of the two leading jets;

bbb last selection: third jet b-tag $CSV > 0.7$

Background determination: b-tag parametrization

- Define a control region using a likelihood ratio discriminator using the most discriminating variables (depends on M_H)
- Build B-tagging probability matrices $P^{3^{rd}jet}_{b-tag}(\ldots)$ in control region for third jet, as a function of 3^{rd} jet and event parameters;

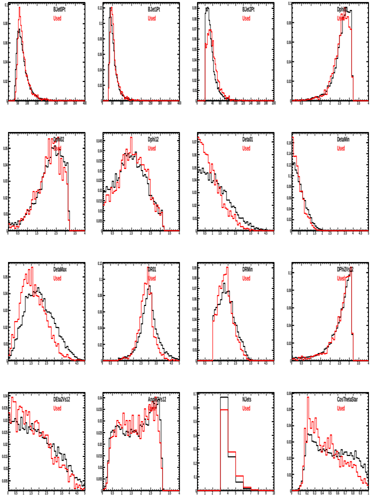
$$P^{3^{rd}jet}_{b-tag}(\ldots) = \epsilon_b \cdot f_b + \epsilon_c \cdot f_c + \epsilon_l \cdot f_l$$

  - b-tagging efficiencies $\epsilon's$ from MC $\epsilon = \epsilon(E_T, |\eta|, N_{trk})$
  - flavour fractions $f_{b,c,l}$ from Data parametrization see next slides

- Estimate any bbb distribution $F(x; bbb)$ for variable x in signal region starting from same distribution for bbj;

$$F(x; bbb) = F(x; bbj) \otimes P^{3^{rd}jet}_{b-tag}(\ldots)$$

Discriminator: example for $M_H = 120\ GeV$

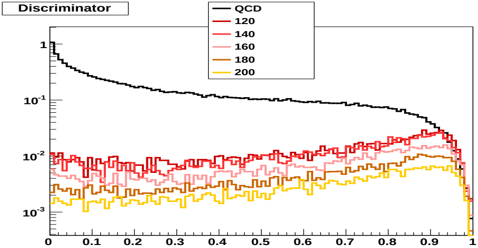


Likelihood ratio

$$Discr = \frac{\Pi_i p_i^{(signal)}(x_i)}{\Pi_i p_i^{(signal)}(x_i) + \Pi_i p_i^{(QCD)}(x_i)}$$

Actual variable choice depends on M_H

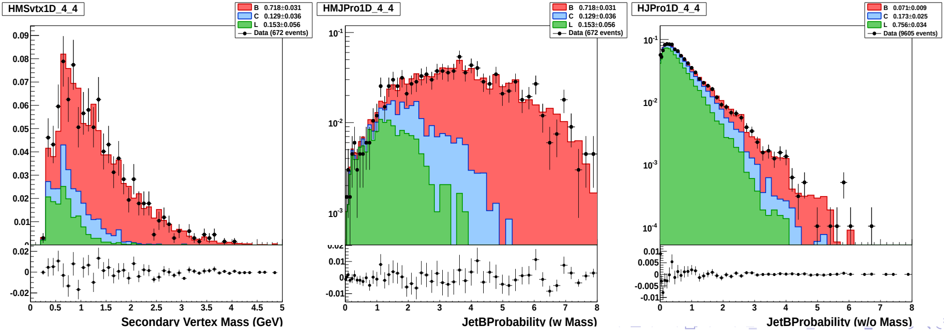
Optimized for $M_H = 120$ but shown for several masses



Heavy quark fraction

Consider Mass@Vertex and JetBProbabilty for third jet;

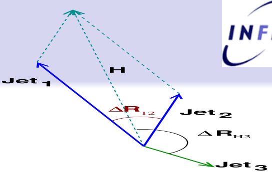
- build distribution templates from MC QCD, for B, C and Light;
- fit third jet distributions using templates and get $F_{b,c}$;
  - used only JetBProbabilty if Mass@Vertex not available.
- for Data, fit separately single B-tag and double B-Tag HLT paths.



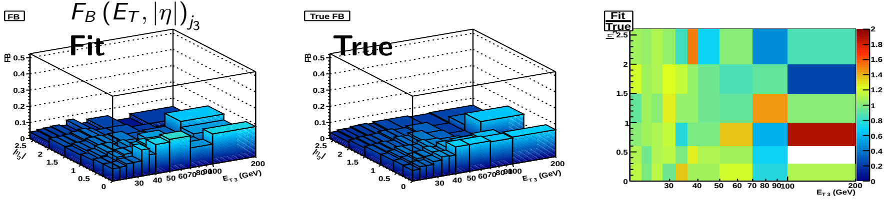
F_B Fit results vs MC truth

$$F_B = F_{B,C}\left(E_T^{(j_3)}, \left|\eta^{(j_3)}\right|\right) \times F_{B,C}\left(\Delta R_{H,j_3}, \Delta R_{j_1,j_2}\right)$$

assuming no correlation. Second factor only for shape.

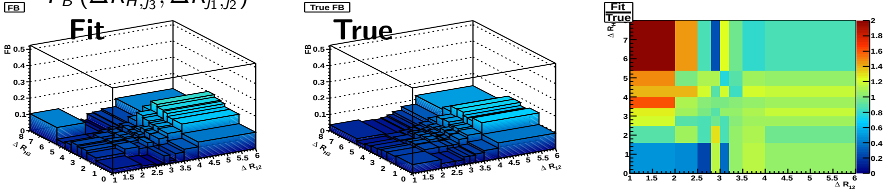


$F_B\left(E_T, |\eta|\right)_{j_3}$



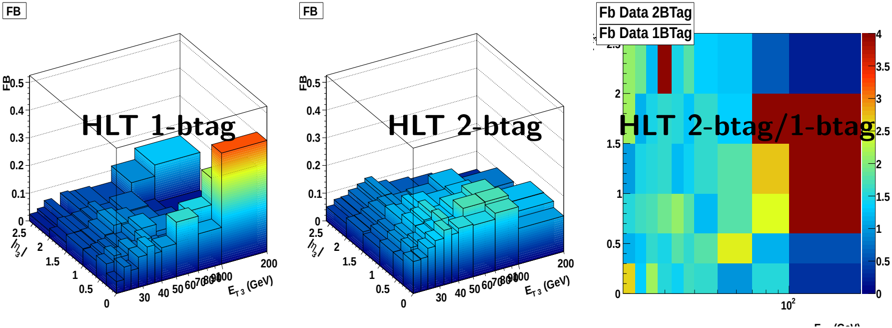
Overall bias for F_b around $\approx +6\%$

$F_B\left(\Delta R_{H,j_3}, \Delta R_{j_1,j_2}\right)$



F_B Fit results Data

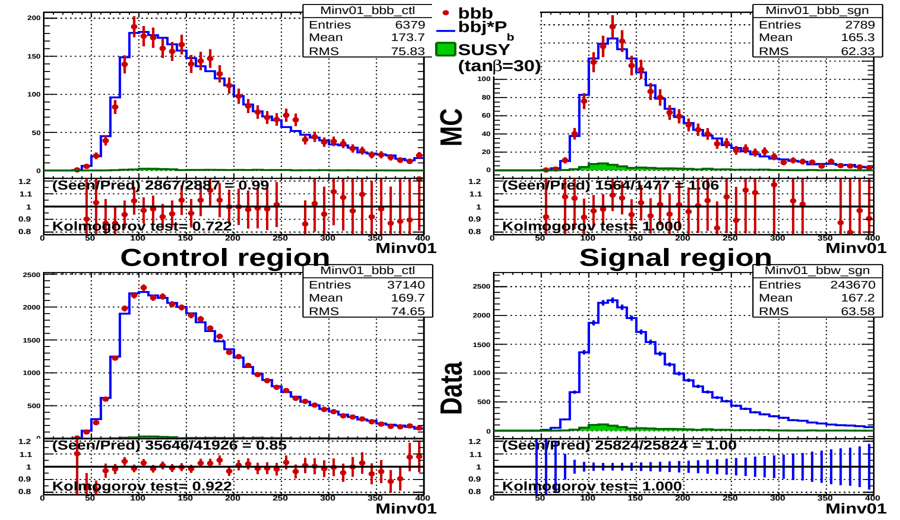
$$F_B\left(E_T^{(j_3)}, \left|\eta^{(j_3)}\right|\right)$$



Enanchement of *b* in third jets, due to the online double b-tag trigger, is clearly visible.

M_{jj} prediction vs *bbb* in MC and Data

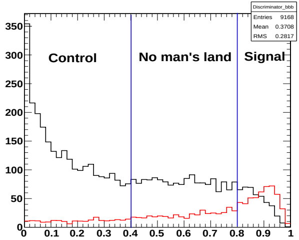
in control and signal region: susy $M_H = 120\ GeV$

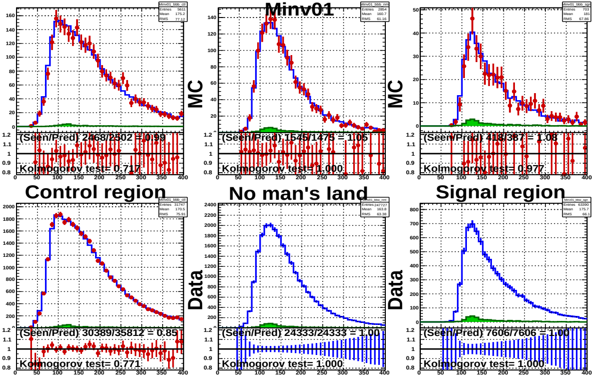


For Data (control region) shape fine, normalization still ~ 0.85

Prediction with No Man's Land

Can divide phase space in three regions: control, no man's land and signal, and use (blind) NML to check prediction consistency on Data before opening signal region.





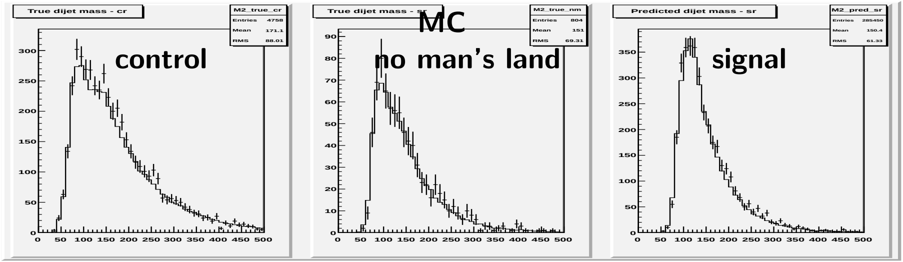
Hyper Ball

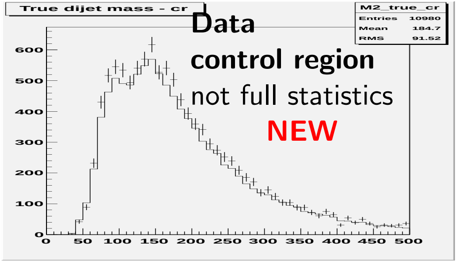
Alternative approach, can be useful to cross check results, study sistematics or for combination: *see Ronchese's talk 27/4/2012*

General idea

- Start from *bjj* sample, control region;
- For each event in $(bjj)_{CR}$ select a set of *similar* events $\mathcal{O}(1000)$
- Compute the fraction of these events passing full selection (bbb);
- *similarity* is defined by *distance* between events in hyperspace $d = \Sigma_i^n \left(w_i \left(x_i - y_i\right)\right)^2$
  - with x_i, y_i jet or event variables (p_T, η, $\Delta\phi_{ij}$, ...);
  - w_i weight to account for variability of *fraction* vs a given variable;
- Use the average fraction to weight events and predict *bbb* distribution.

Hyper Ball: prelim results on QCD MC and Data





- M_{jj} prediction (line) vs signal bbb (crosses)
- MC scale factor: CR : 1.07 NML : 1.12 SR : 1.13;
  - NML and SR plots rescaled for CR scale-factor
- Data scale factor: CR : 1.07

very preliminary!

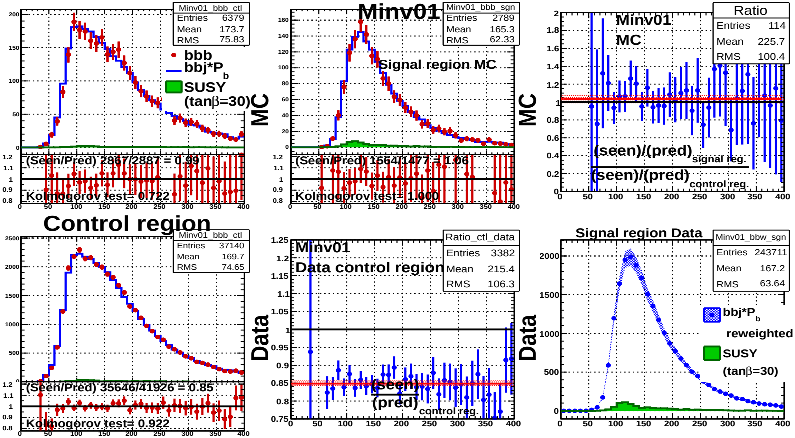
Systematics (preliminary)

- Trigger syst: $\approx 3 - 5\%$ from data driven ϵ estimate;
- Physics object syst:
  - B-tagging eff. `BTV-12-001` $\approx 4\%$ per BJet
  - JEScale $^{+2.5}_{-3.1}\%$
  - JEResolution $\pm 1.9\%$
  - Mu momentum scale $\approx 0.2\%$ and resolution $\approx 0.6\%$ negligible
  - Mu efficiency (`MUO-10-004`) (small?)
- Background determination syst $\approx 5\%$ next slide :
- Integrated Lumi syst: $\approx 2.2\%$
- ...

Background Systematics

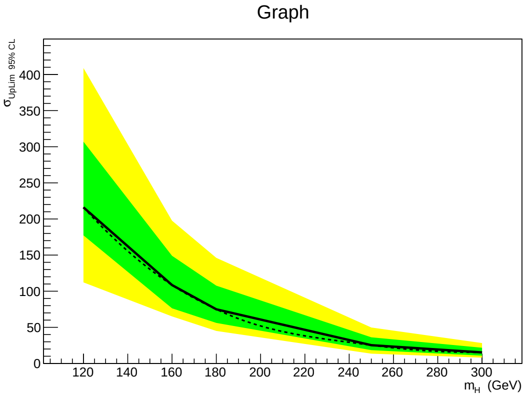
- Two major source of systematics for the predicted *bbb* in signal region:
  1. Systematics from *bbb* prediction from DATA control region
     - ★ compare *bbb* and *bbj* $\times$ *P* in DATA control region;
     - ★ use normalization in signal region;
     - ★ use fit error as systematics;
  2. Systematics 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



Sensitivity including systematics

Expected upper limit on cross section: CLs Asymptotic



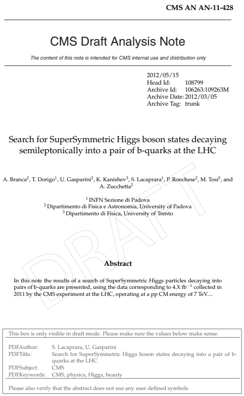
Preliminary: not yet optimized for $M_H > 120\ GeV$!

Summary

- Hbb semileptonic analysis presented;
- Data driven background prediction working quite well;

ToDo

- trigger efficiency studies; in progress ✓
- parametrization;✓
- systematics (all sources) for final sensitivity prediction; in progress ✓
- HyperBall aligned and consolidate; in progress ✗
- Documentation (AN, etc); in progress ✗
- full analysis on all mass points; in progress ✗



CMS AN AN-11-428

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2012/05/15
Head Id: 108799
Archive Id: 106263:109263M
Archive Date: 2012/03/05
Archive Tag: trunk

Search for SuperSymmetric Higgs boson states decaying semileptonically into a pair of b-quarks at the LHC

A. Branca[2], T. Dorigo[1], U. Gasparini[2], K. Kanishev[3], S. Lacaprara[1], P. Ronchese[2], M. Tosi[2], and A. Zucchetta[2]

[1] INFN Sezione di Padova
[2] Dipartimento di Fisica e Astronomia, University of Padova
[3] Dipartimento di Fisica, University of Trento

DRAFT

Abstract

In this note the results of a search of SuperSymmetric Higgs particles decaying into pairs of b-quarks are presented, using the data corresponding to 4.X fb^{-1} collected in 2011 by the CMS experiment at the LHC, operating at a pp CM energy of 7 TeV....

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	S. Lacaprara, U. Gasparini
PDFTitle:	Search for SuperSymmetric Higgs boson states decaying into a pair of b-quarks at the LHC
PDFSubject:	CMS
PDFKeywords:	CMS, physics, Higgs, beauty

Please also verify that the abstract does not use any user defined symbols

BACKUP

Physics Objects

- CMSSW 4_2_7 including JetMet suggested tags.
- AK5 ParticleFlow Jets, JEC applied:
  - L1FastJet, L2Relative, L3Absolute, L2L3Residual (only for Data)
  - Global Tag: `FT_R_42_V20A` and `START42_V17` for Data and MC
  - PU treatment: PF Charged Hadron Subtraction and Area Method;
  - Loose JetId selections;
- Jet b-tagging used is Combined Secondary Vertex (CSV);
- Standard Global Muon (no isolation requirements)

- JES and JER from POG (CERN-PH-2011/102 and update JetMET presentation 9/1/12);
- BTag efficiency studies on top samples (BTV-12-001);
- Muon (non isolated) efficiency on J/ψ MUO-10-004;

Numerology

HLT paths (L1 seed)	run range	triggered events	$\int \mathcal{L}dt$ [pb^{-1}]
`HLT_Mu12_CentralJet30_BtagIP`			
`L1_SingleMu7`	163738-165633	3 027 717	183.815
`HLT_Mu12_DiCentralJet30_BtagIP3D`			
`L1_SingleMu10`	165970-172952	4 532 555	524.904
`HLT_Mu12_DiCentralJet20_DiBtagIP3D1stTrack`			
`L1_SingleMu10`	167039-173198	1 873 247	780.378
`HLT_Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack`			
`L1_Mu10_Eta2p1_DoubleJet_16_8`	173236-178380	4 762 858	1944.527
`HLT_Mu12_eta2p1_DiCentralJet20_DiBtagIP3D1stTrack`*			
`L1_Mu10_Eta2p1_DoubleJet_16_8`	178420-180252	2 164 634	811.861
All		16 361 011	4245.485

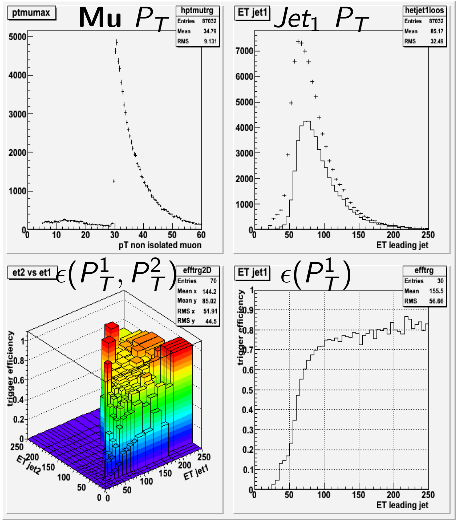
Warning: still `lumiCalc2.py`

Trigger Efficiency vs Analysis

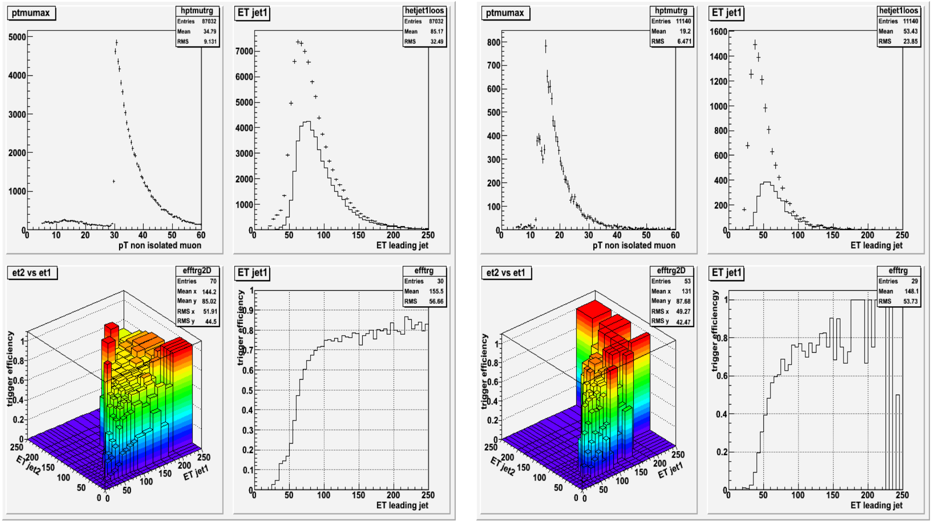
- Use **SingleMuon PD**, select all events passing a single muon path.
- Apply preselection (2 bjets)
- Build Turn on curves vs first and second B-jets Pt:

$$\epsilon = \frac{\text{Hbb path \& presel \& SingleMuHLT}}{\text{presel \& SingleMuHLT}}$$

- Here for `HLT_Mu12_DiCentralJet30_BtagIP3D` and `SingleMu30`
- Turn-on stable wrt SingleMu threshold



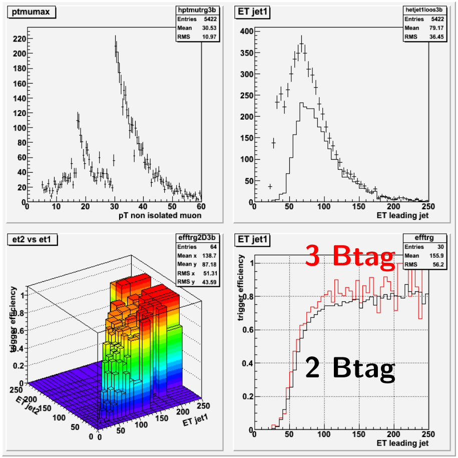
Trigger Efficiency (I)



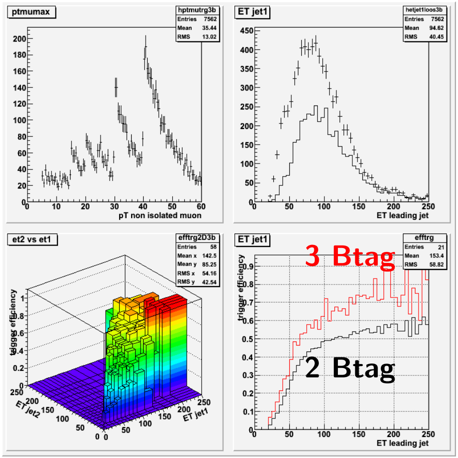
Turn-on stable for different SingleMu threshold (left Mu30, right Mu12 &

Trigger Efficiency (II) 2 btags vs 3 btags

HLT_Mu12_DiCentralJet30_BtagIP3D

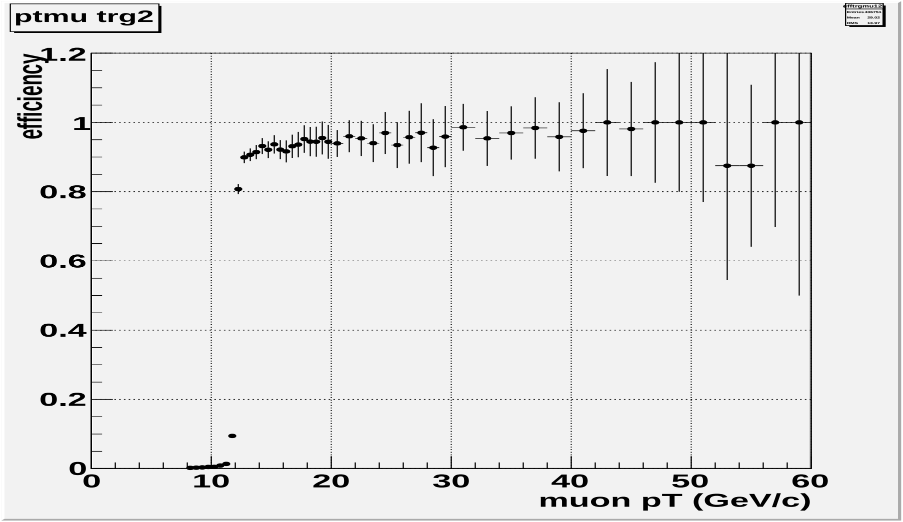


HLT_Mu12_DiCentralJet20_DiBtagIP3D1stTrack



Statistics bit low but still affordable

HLT Mu12 trigger efficiency wrt Analysis

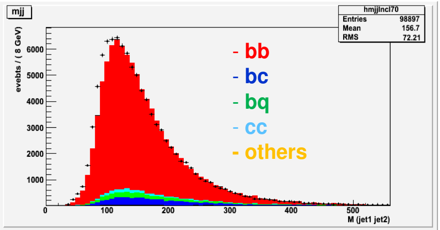


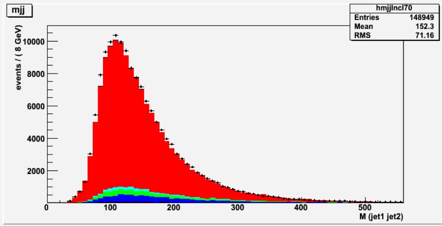
Mass spectra, trigger corrected

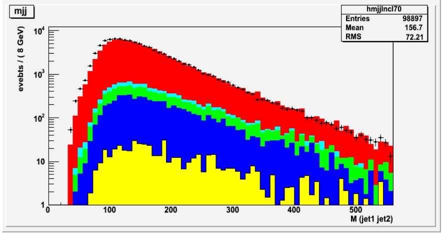
HLTMu12_DicentralJet30_BtagIP3D
L = 525 pb-1

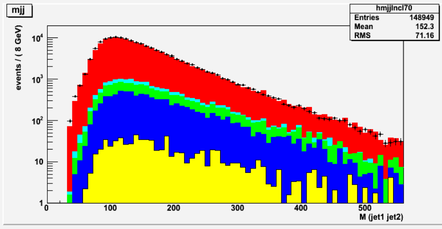
2 btags

HLTMu12_DicentralJet20_DiBtagIP3D1stTk
L = 1002 pb-1

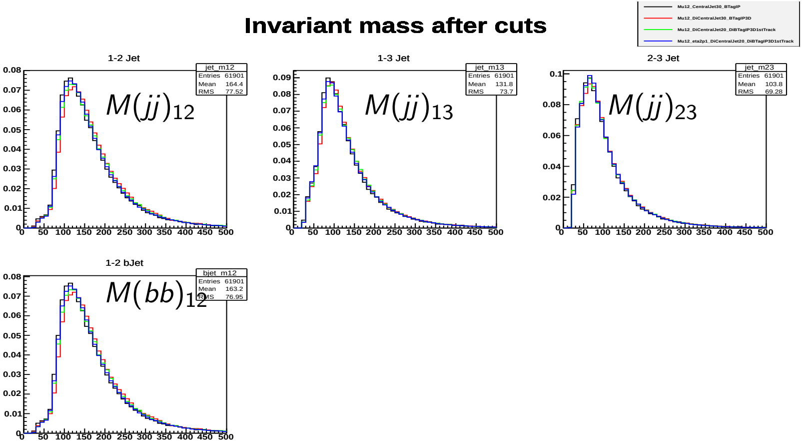






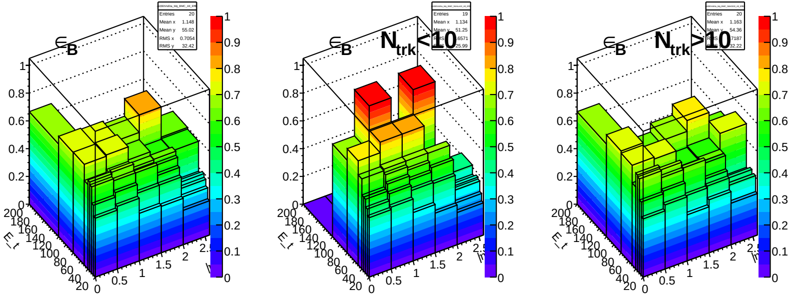


M_{jj} and M_{bb} for different HLT paths



Applying pre-selections (2 b-tag) only: no bias in M_{jj} nor M_{bb}

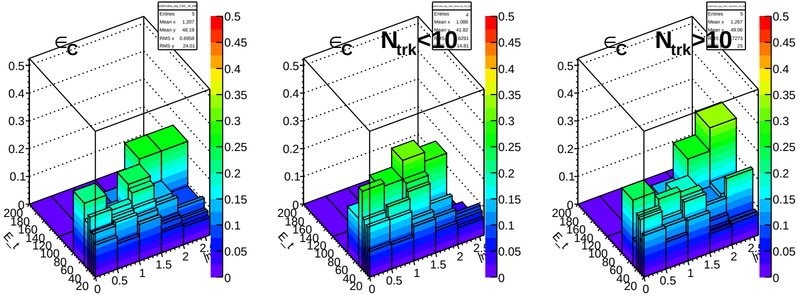
ϵ_b for bbj vs $|\eta|$, E_t vs N_{trk}



Left to Right:

ϵ_B All, $N_{trk} < 10$, $N_{Trk} \geq 10$

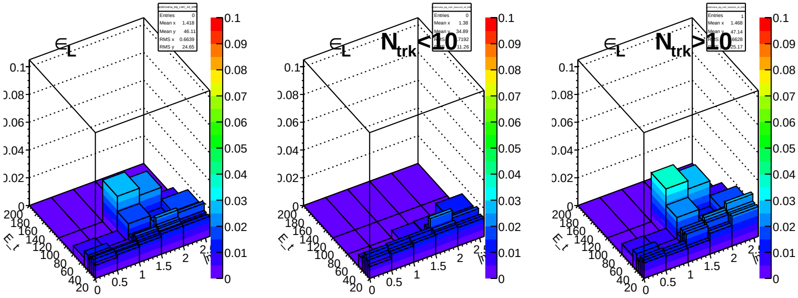
ϵ_c for bbj vs $|\eta|$, E_t vs N_{trk}



Left to Right:

ϵ_C All, $N_{trk} < 10$, $N_{Trk} \geq 10$

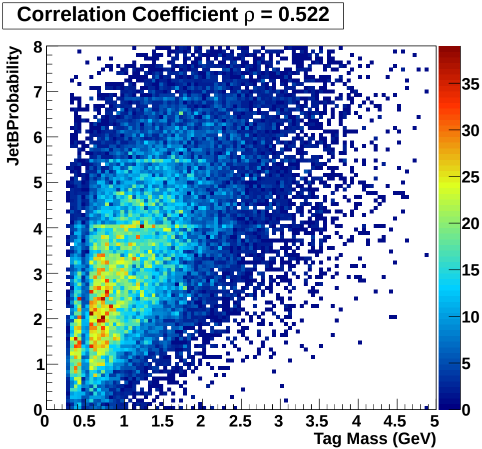
ϵ_{light} for bbj vs $|\eta|$, E_t vs N_{trk}



Left to Right:

ϵ_{Light} All, $N_{trk} < 10$, $N_{Trk} \geq 10$

Mass@Vertex and JetBProbabiliy correlation



$F_{b,c}$ parametrization

$$F_{B,C}\left(E_T^{(j_3)}, \left|\eta^{(j_3)}\right|\right) \times F_{B,C}\left(\Delta R_{H,j_3}, |\Delta R_{j_1,j_2}|\right)$$

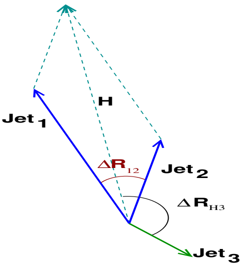
assuming no correlation.

One factor for the third jet features, one for the event topologies.

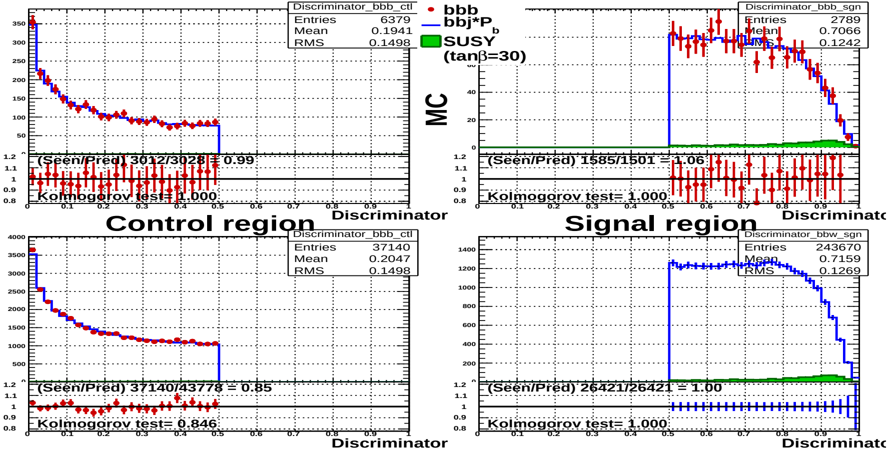
Use $F_{B,C}(\Delta R_{H,j_3}, |\Delta R_{j_1,j_2}|)$ only for shape:

$F_{B,C}$ average, weighted to *bbj* distribution, is normalized to unity.

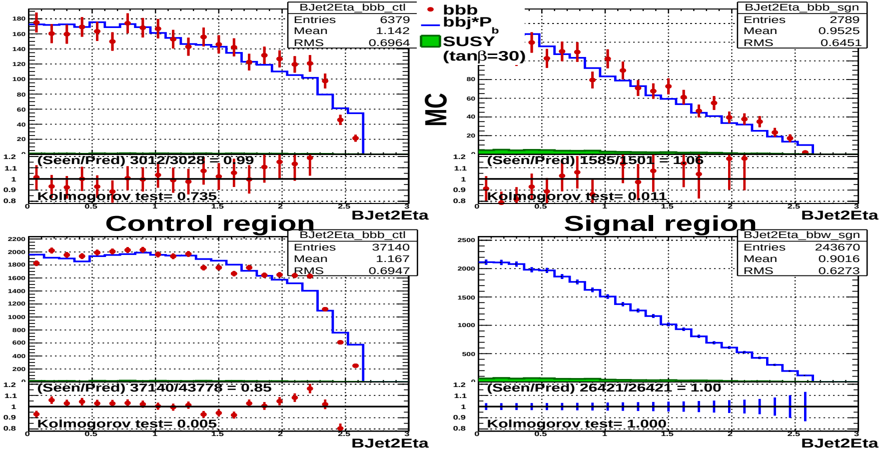
$$\int_{C\ reg.} \frac{dN}{d\Delta R_{12} d\Delta R_{H,j3}} \cdot F_{B,C}\left(\Delta R_{H,j3}, \Delta R_{12}\right) d\Delta R_{12} d\Delta R_{H,j3} = \int_{C\ reg.} \frac{dN}{d\Delta R_{12} d\Delta R_{H,j3}} d\Delta R_{12} d\Delta R_{H,j3}$$



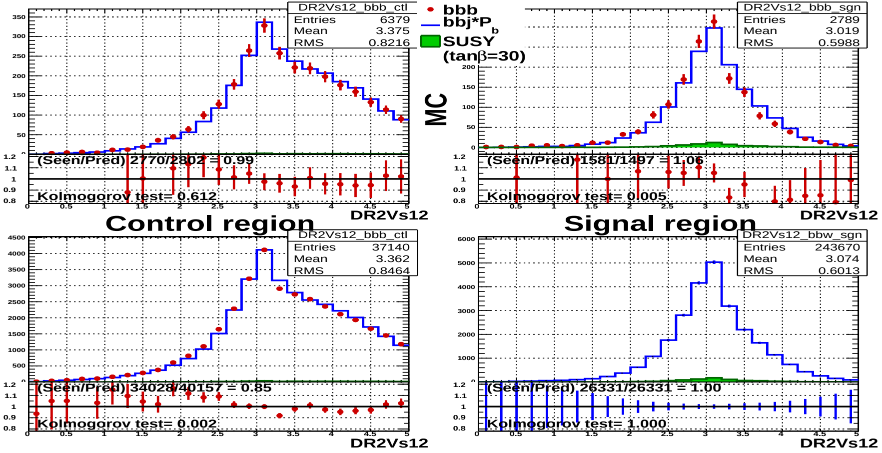
Prediction vs *bbb* in MC and Data



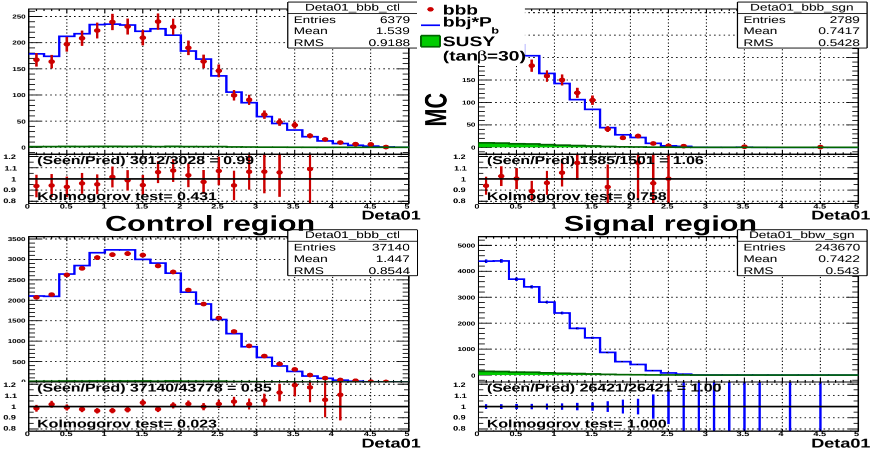
Prediction vs *bbb* in MC and Data



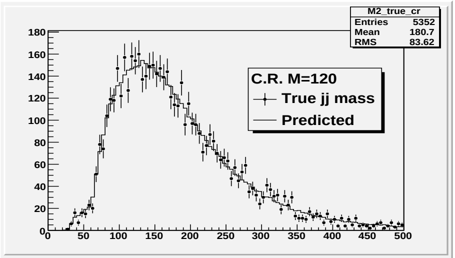
Prediction vs *bbb* in MC and Data

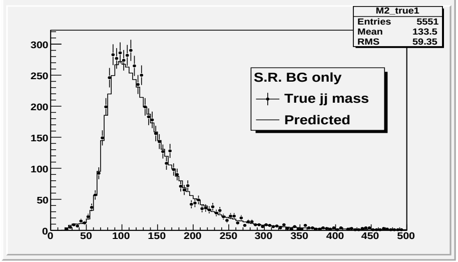


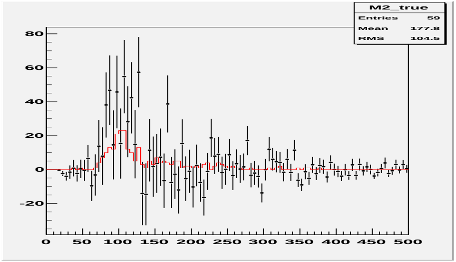
Prediction vs *bbb* in MC and Data



Hyper Ball: prelim results on QCD MC







Example of signal injection and extraction $M_H = 120$ GeV