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

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

INFN Padova

Hbb meeting,
CERN, May, the 4^{th} *, 2012

* *be with you*

Intro

- Search for Neutral SUSY Higgs, $H \to b\bar{b}$;
- Use associate production to reject hadronic background $pp \to b\bar{b}H \to b\bar{b}b\bar{b}$

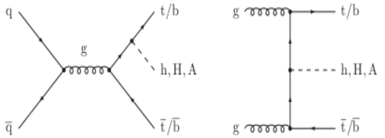


Fig. 17: Typical diagrams contributing to $q\bar{q}/gg \to Q\bar{Q} + h/H/A$ $(Q = t, b)$ at lowest order.

- in MSSM, production σ can be quite large for large $\tan\beta$
  - *e.g.* for $M_A = 120$ GeV and $\tan\beta = 30$: $\sigma \times BR \approx 130\ pb$.
  - Also, a factor 2 is gained from the degeneration of two of the three MSSM neutral higgs (h, H, A);

Analysis Strategy

- three b final state, two b-jets build the h;
- Major background source is QCD
- trigger is critical:
  - Use semi-leptonic (muon) b decay for trigger: muon+jets+b-tagging
- Estimate QCD background from bbj sample
  - define control sample signal poor;
  - estimate b/c-fraction of 3rd jet from mass & lifetime fits
  - combine with MC b-tagging efficiency to derive B-tag probability
  - weight untagged events to estimate number of 3-b-tags in signal region
- Use reconstructed mass of leading jet pair as signal-sensitive variable in final fit

Analysis selection

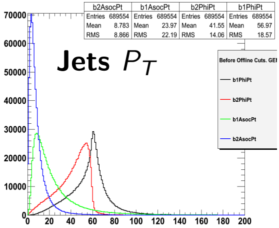
Selections

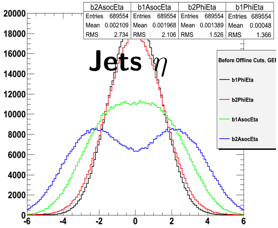
Baseline selections:

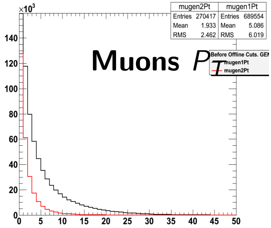
- at least 1 global muon $P_T^{\mu} > 15\ GeV$, no isolation required;
- at least 3 jets (PFak5, LooseId) $|\eta| < 2.6$, $P_T > (30, 30, 20)\ GeV$
- $\Delta R_{ij} > 1$ for any pair $_{ij}$ of jets
- the first 2 jets must have b-tag $CSV > 0.8$, the μ inside one of the two;
- last selection: third jet b-tag $CSV > 0.7$

Trigger Strategy

- *b*-rich final state:
  - 2 central b-jets (from h decay): E_T up to $M_H/2$;
  - 2 forward b-jets (associated prod.) low E_T;
    - ★ fourth one is typically low pt and outside tracker.
  - One muon from b fragmentation







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

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

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

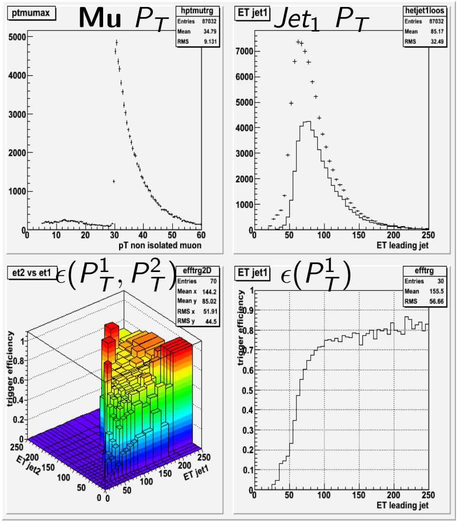
Warning: still `lumiCalc2.py`

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:

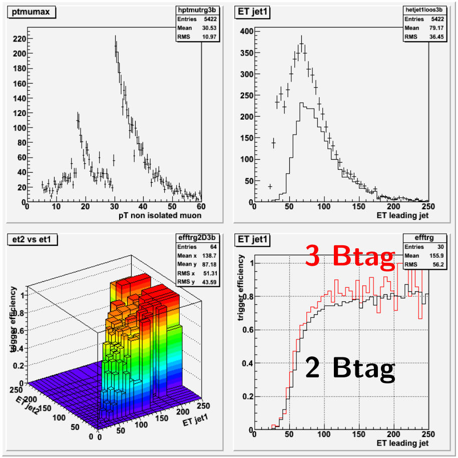
$$\epsilon = \frac{\text{Hbb path \& presel \& SingleMuHLT}}{\text{presel \& SingleMuHLT}}$$

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

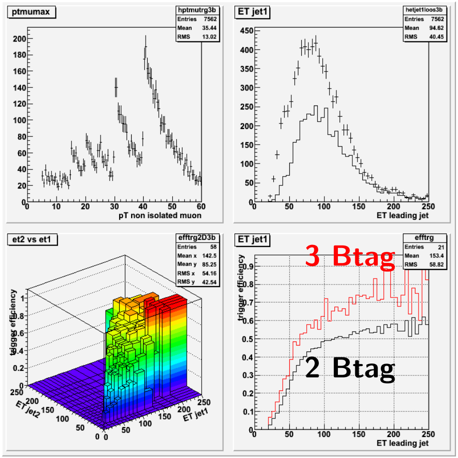


Trigger Efficiency (II) 2 btags vs 3 btags

HLT_Mu12_DiCentralJet30_BtagIP3D

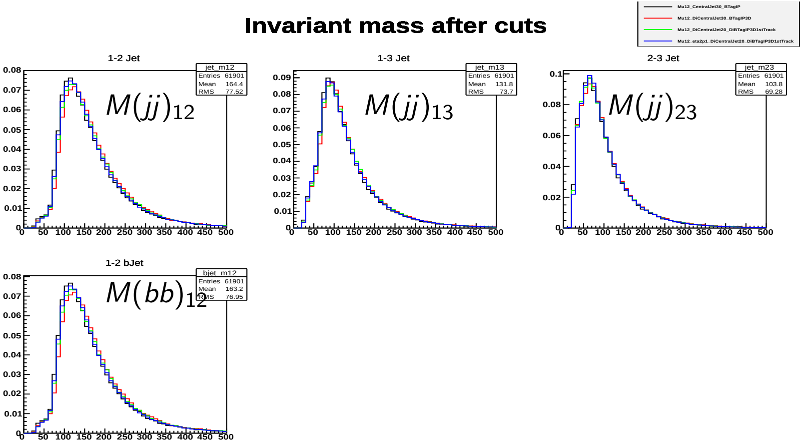


HLT_Mu12_DiCentralJet20_DiBtagIP3D1stTrack



Statistics bit low but still affordable

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



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

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;

Background determination

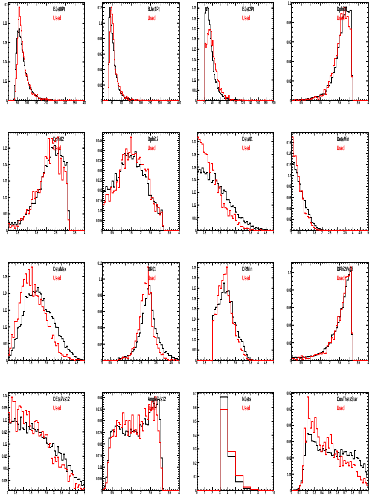
- Major background source is QCD
- Define a control region using a likelihood ratio discriminator using the most discriminating variables (depends on M_H)
- Build B-tagging probability matrices in control region for third jet;
- Estimate *bbb* background for variable x in signal region starting from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

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

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

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

Discriminator

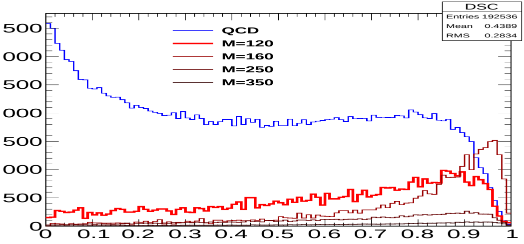


Likelihood ratio

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

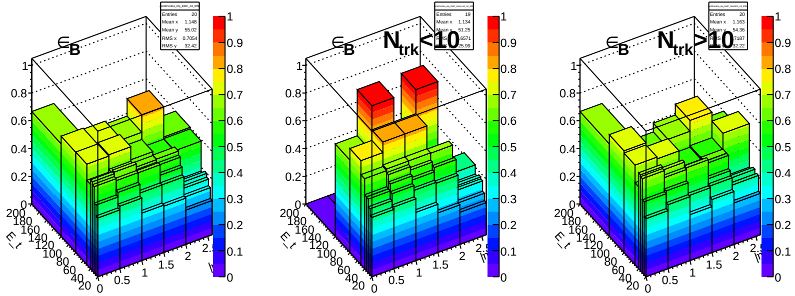
Actual variable choice depends on M_H

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



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

Parametrization vs $|\eta|$, E_t and for two range of charged multiplicity ($N_{Trk} \lessgtr 10$)



Left to Right:

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

$F_{b,c}$ parametrization

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

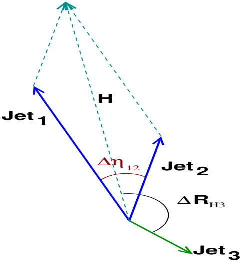
assuming no correlation.

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

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

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

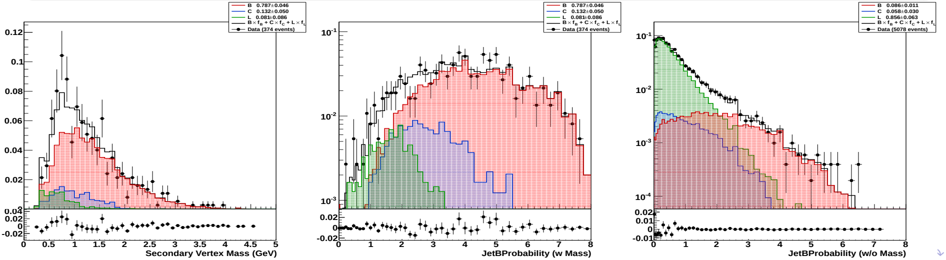
$$\int_{C\ reg.}\frac{dN}{d\Delta\eta_{12}d\Delta R_{H,j3}}\cdot F_{B,C}\left(\Delta R_{H,j3},\Delta\eta_{12}\right)d\Delta\eta_{12}d\Delta R_{H,j3}=\int_{C\ reg.}\frac{dN}{d\Delta\eta_{12}d\Delta R_{H,j3}}d\Delta\eta_{12}d\Delta R_{H,j3}$$



Heavy quark fraction

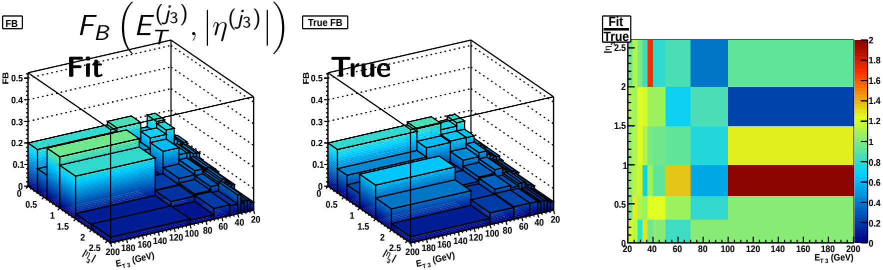
Consider Mass@Vertex and JetBProbabilty for third jet;

- build distribution templates from MC QCD, for each flavour (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 DiB-Tag HLT paths.

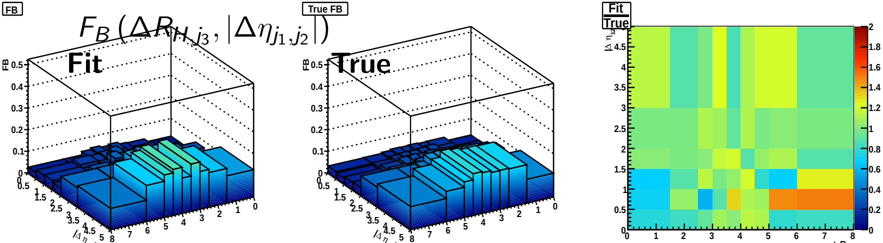


$F_{B,C}$ Fit results vs MC truth

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



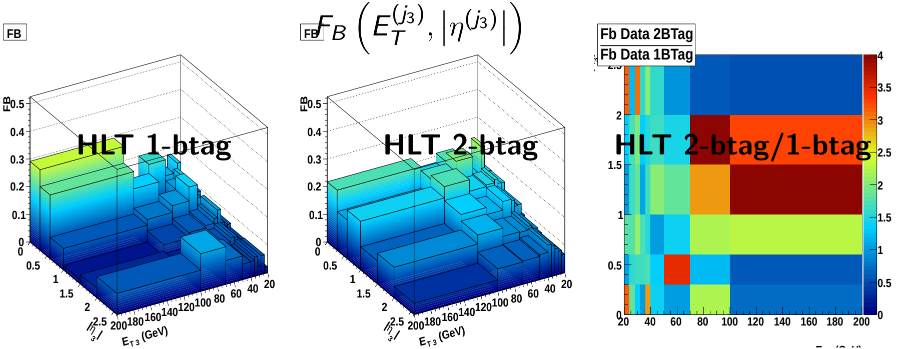
$F_B\left(\Delta R_{H,j_3}, |\Delta\eta_{j_1,j_2}|\right)$



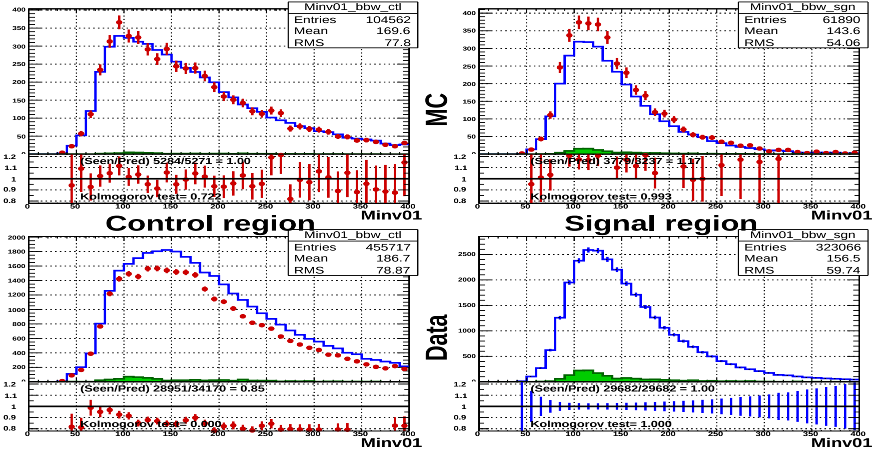
Overall bias for F_b around $\approx +6\%$

$F_{B,C}$ Fit results Data

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



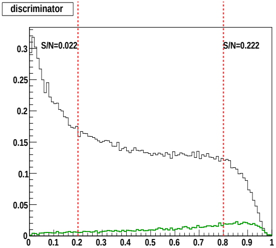
Prediction vs *bbb* in MC and Data

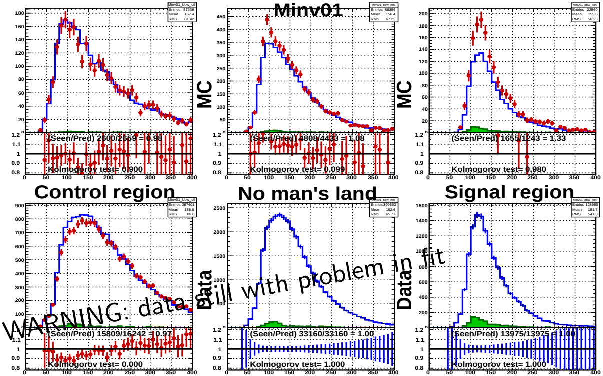


Clearly still problem in MC closure in signal region and in Data control region!

Prediction with No Man's Land

Or 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.





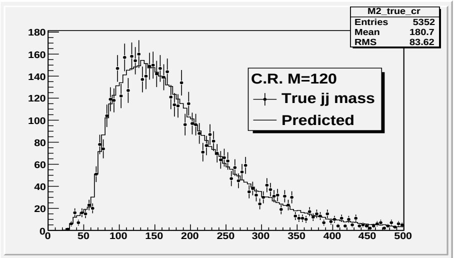
Hyper Ball

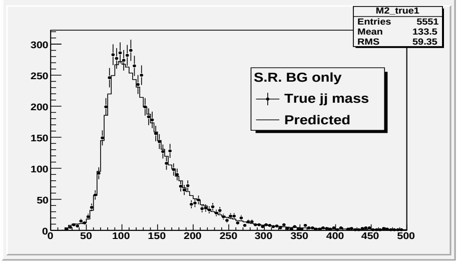
Alternative approach, can be useful to cross check results, study sistematics or for combination: described in detail in Paolo's presentation on Hbb meeting 27/4/2012

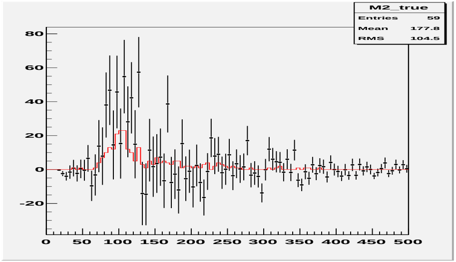
General idea

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

Hyper Ball: prelim results on QCD MC







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

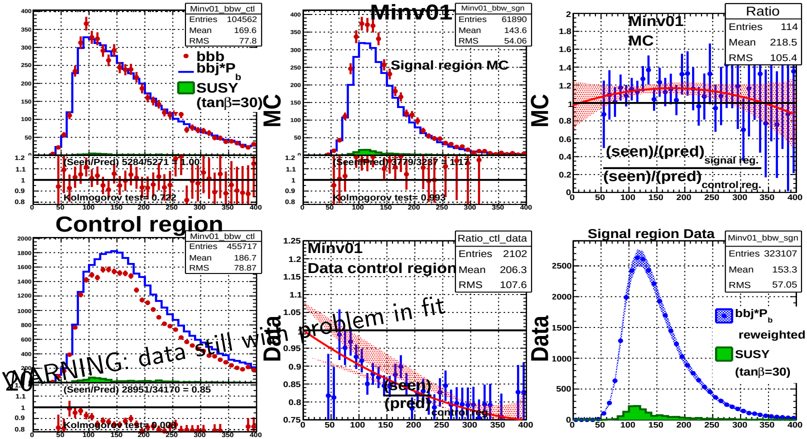
Systematics

- Trigger syst: $\approx 3 - 4\%$ from data driven ϵ estimate;
- Physics object syst:
  - B-tagging efficiency and scale factor from `BTV-12-001` $\approx 1\%$ per BJet
  - Jet Energy Scale from JEC group studies (see JetMET presentation 9/1/12 and JEC Twiki) $\approx 2 - 3\%$ (?)
  - Mu momentum scale $\approx 0.2\%$ and resolution $\approx 0.6\%$ negligible
  - Mu efficiency (`MUO-10-004`) (small?)
- Background determination syst:
  - Bin per bin syst comparing prediction $bbj \cdot P$ vs bbb in data control region
  - Bin per bin syst for extrapolation of prediction from control to signal region from MC
- Integrated Lumi syst: $\approx 2.2\%$

Background 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 DATA control region, fit ratio, use fit results as bias and 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;

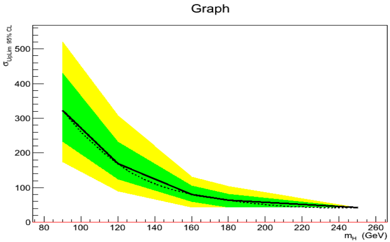
Systematics test



Sensitivity including systematics

Upper limit on x-sec: CLs Asymptotic:

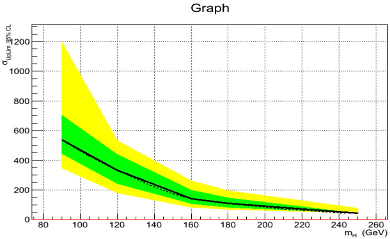
- mass hypothesis: 90, 120, 160, 180, 250 GeV



Lint = 5 /fb

Still 5% flat sys for bkg and eff.

4.5% sys lumi



Lint = 5 /fb

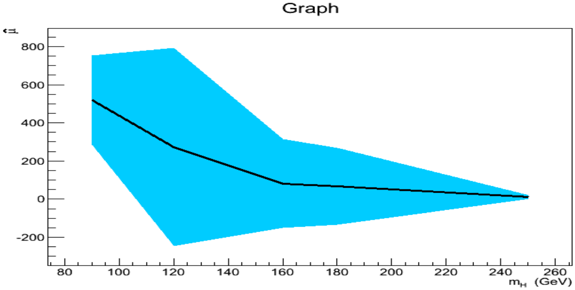
7% flat sys bkg;
5% flat sys eff;

4.5% sys lumi

signal reduction: factor 2 to take into account hlt eff

Signal injection test
(very prelim, error band to be checked)

Fitted xsection vs mass for a given signal injection:



x-section of the Injected Signal for each mass point:

- mA = 90 GeV – x-sec = 662.1 /pb;
- mA = 120 GeV – x-sec = 259.8 /pb;
- mA = 160 GeV – x-sec = 88.5 /pb;
- mA = 180 GeV – x-sec = 56.5 /pb;
- mA = 250 – x-sec = 14.5 /pb;

Fitted x-section:

- mA = 90 GeV – fitted x-sec = 519 /pb
- mA = 120 GeV – fitted x-sec = 273.5 /pb
- mA = 160 GeV – fitted x-sec = 81.5 /pb
- mA = 180 GeV – fitted x-sec = 67 /pb
- mA = 250 GeV – fitted x-sec = 11.4 /pb

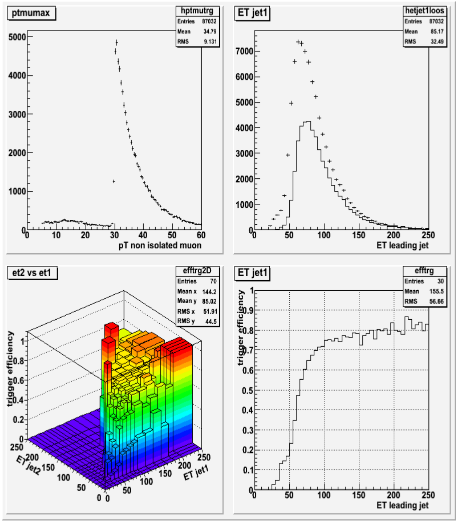
Work plan

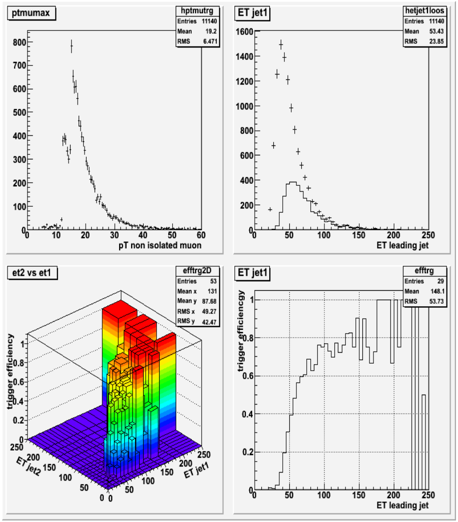
ma che state pazziando? (trad. “Are you kidding?”)

- May 15
  - trigger efficiency studies;
  - parametrization;
  - systematics (all sources) for final sensitivity prediction;
  - HB aligned and tested on data control region;
  - AN in advanced state;
- June 1
  - full analysis on all mass points;

BACKUP

Trigger Efficiency (I)

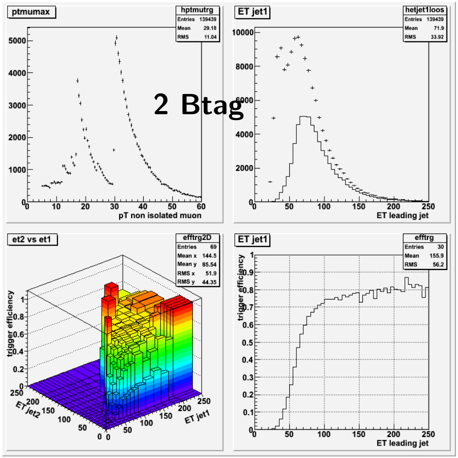




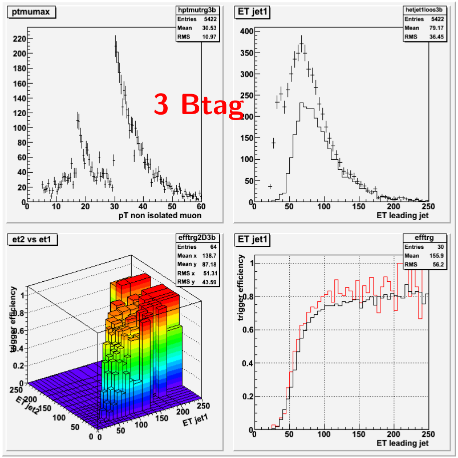
Turn-on stable for different SingleMu threshold (left Mu30, right Mu12 &

Trigger Efficiency (II) 2 btags vs 3 btags

HLT_Mu12_DiCentralJet30_BtagIP3D



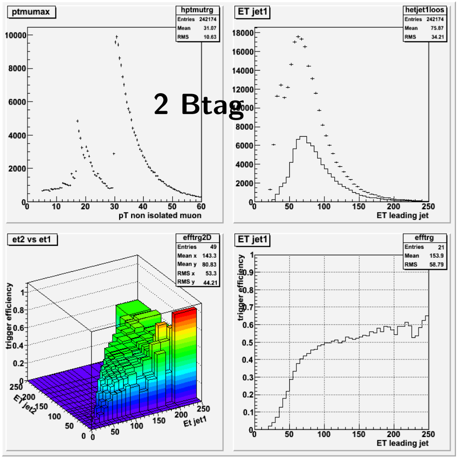
2 btags vs 3 btags

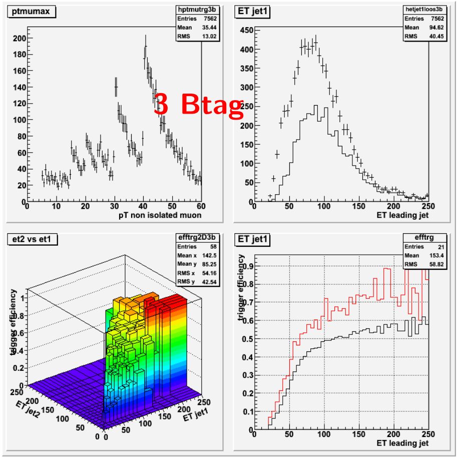


Statistics bit low but still affordable

Trigger Efficiency (III) Double online BTag: RunA

HLT_Mu12_DiCentralJet20_DiBtagIP3D1stTrack **2 btags vs 3 btags**

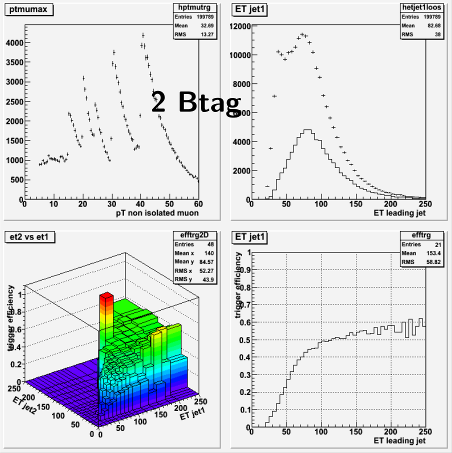


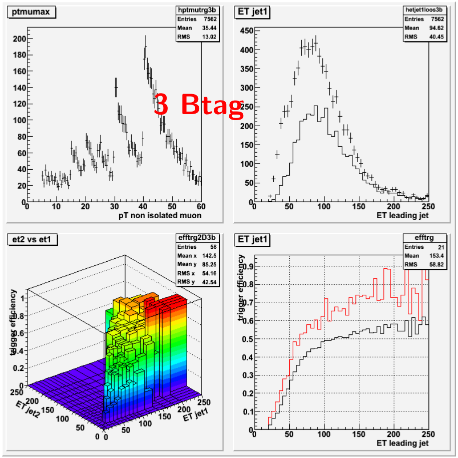


Ok also for double b-tag

Trigger Efficiency (IV) Double online BTag: RunB

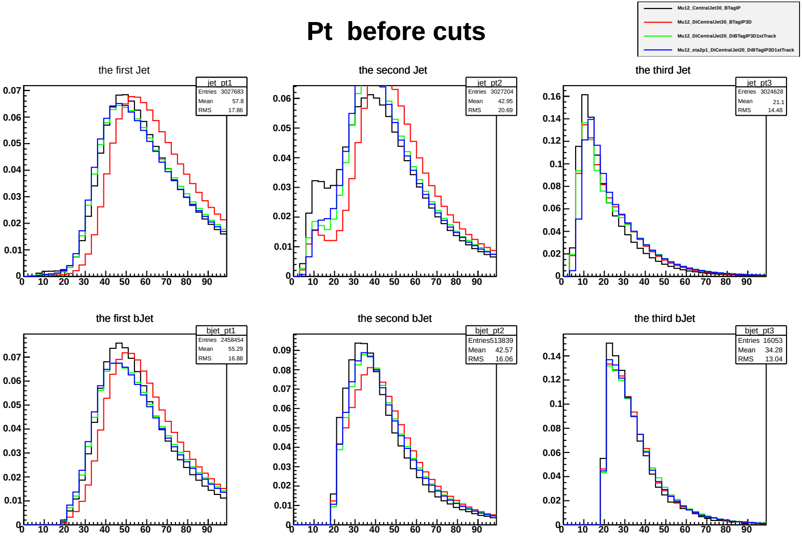
HLT_eta2p1_Mu12_DiCentralJet20_DiBtagIP3D1stTrack 2 btags vs 3 btags



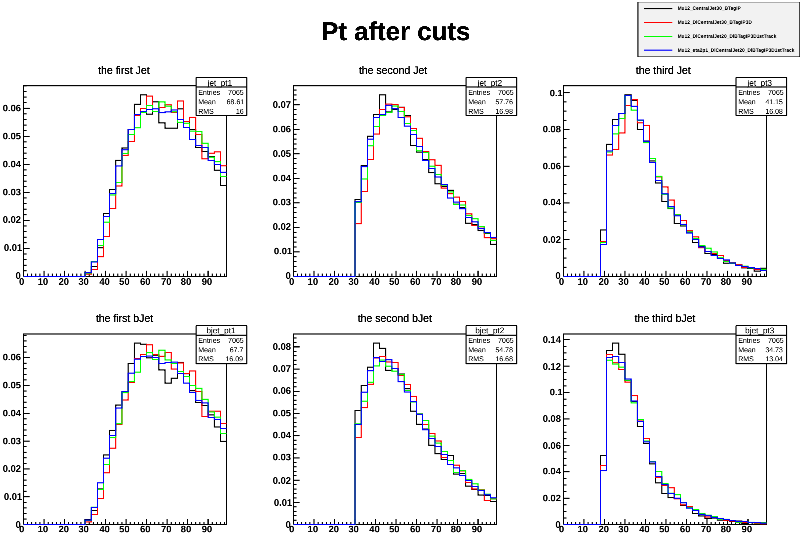


Ok also for RunB

Offline cuts and HLT paths



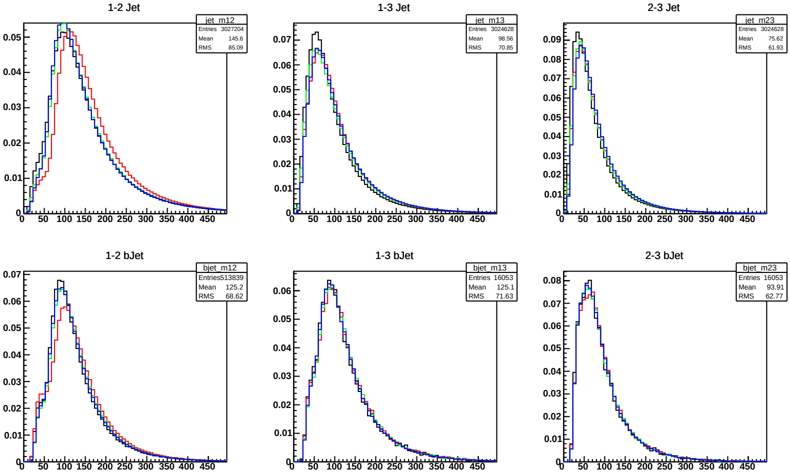
Offline cuts and HLT paths



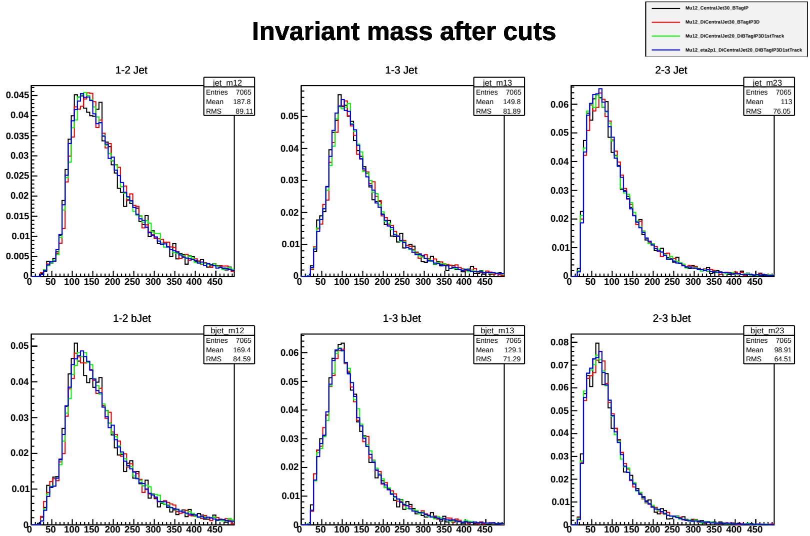
Offline cuts and HLT paths

Invariant mass before cuts

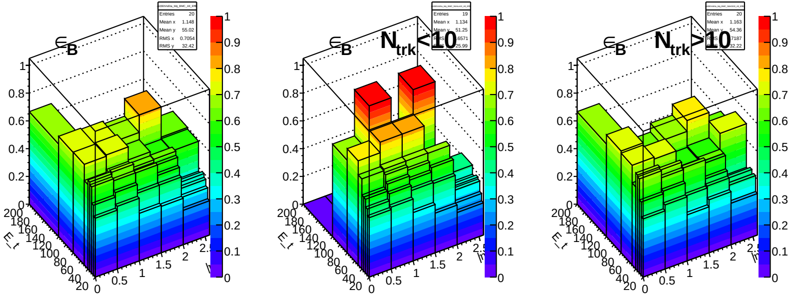
Mu12_CentralJet30_BTagIP
Mu12_DiCentralJet30_BTagIP3D
Mu12_DiCentralJet20_DiBTagIP3D1stTrack
Mu12_eta2p1_DiCentralJet20_DiBTagIP3D1stTrack



Offline cuts and HLT paths



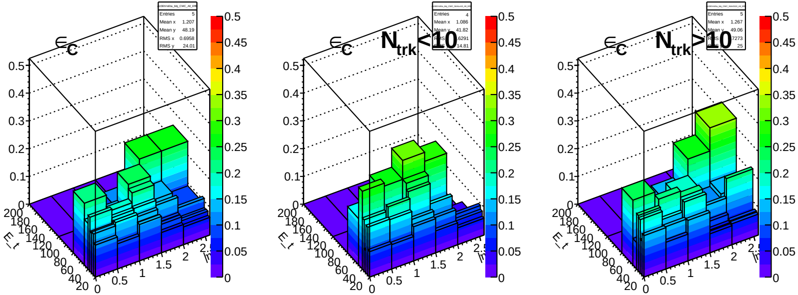
ϵ_b for bbj vs $|\eta|$, E_t vs N_{trk}



Left to Right:

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

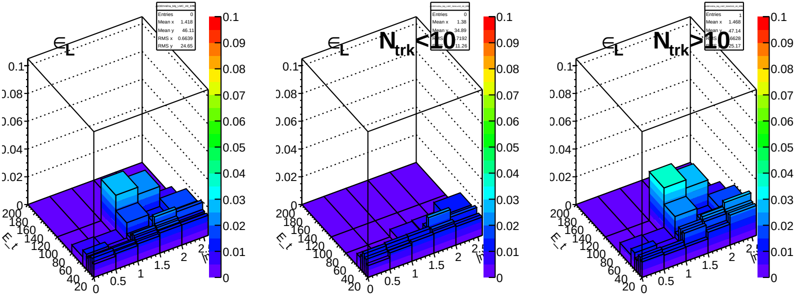
ϵ_c for bbj vs $|\eta|$, E_t vs N_{trk}



Left to Right:

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

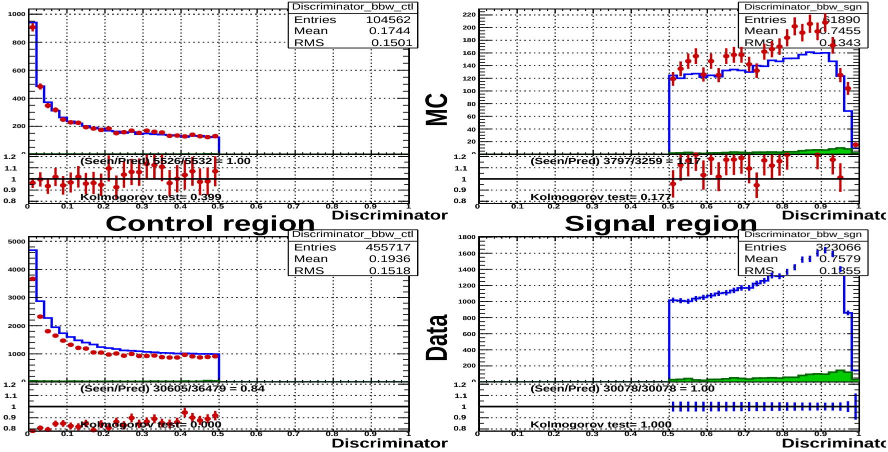
ϵ_{light} for bbj vs $|\eta|$, E_t vs N_{trk}



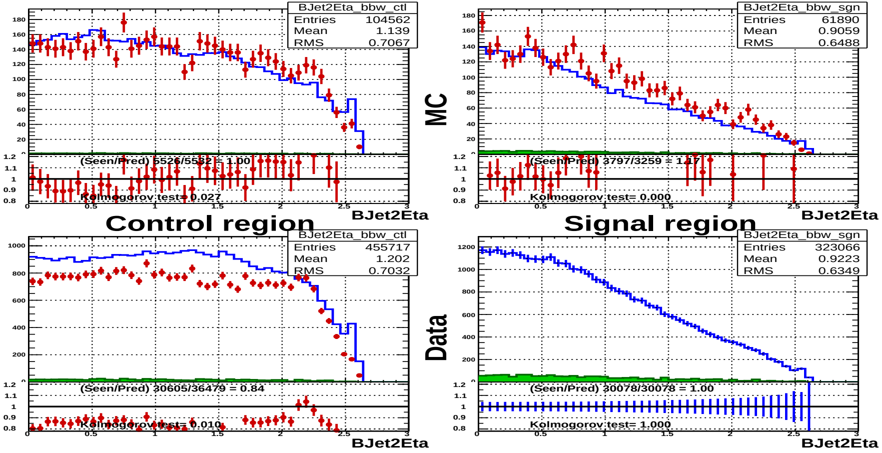
Left to Right:

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

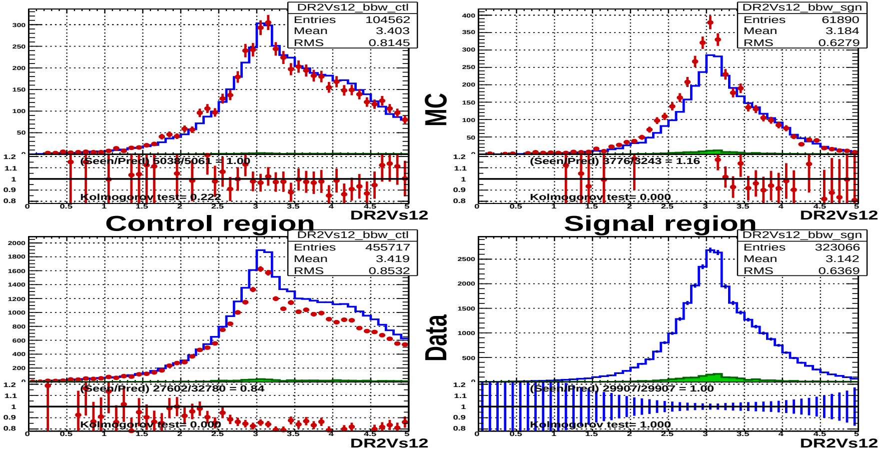
Prediction vs *bbb* in MC and Data



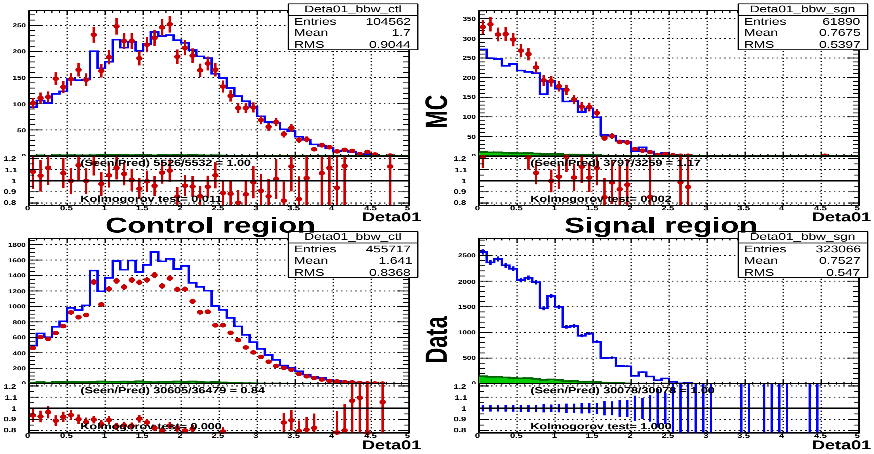
Prediction vs *bbb* in MC and Data



Prediction vs *bbb* in MC and Data



Prediction vs *bbb* in MC and Data

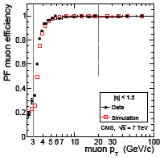


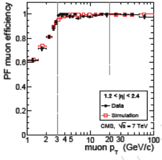
Phys.objects: muon inside jets

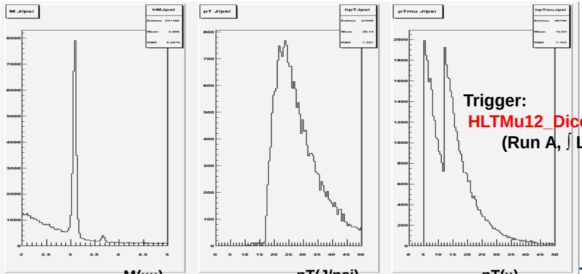
-Detailed mu reco studies exist in CMS on 2010 data (e.g. tag&prob on J/psi), see paper MUO-10-004

-For non-iso muons, not the same for 2011 at moment (to our knowledge)

- However, sizable number of J/psi's selected by our Hbb trigger, with relatively hard pT spectrum; tag&probe method feasible (not for June ?)







M(μμ) pT(J/psi) pT(μ)

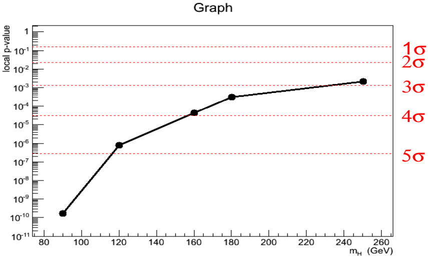
Trigger:
HLTMu12_DicentralJet30_BtagIP3D
(Run A, ∫ Ldt~0.525 pb-1)

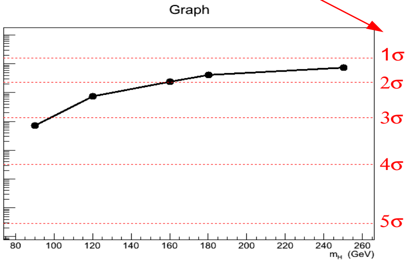
5/4/2012 1

Local P-value: probability of the background to fluctuate and give an excess as large or larger than the observed one

- is local, so no *Look Elsewhere Effect* included;

Significance of the observed excess





x-section of the Injected Signal for each mass point:

- mA = 90 GeV – x-sec = 662.1 /pb;
- mA = 120 GeV – x-sec = 259.8 /pb;
- mA = 160 GeV – x-sec = 88.5 /pb;
- mA = 180 GeV – x-sec = 56.5 /pb;
- mA = 250 – x-sec = 14.5 /pb;