Search for MSSM $H \to b\bar{b}$

I. Marfin

on behalf of the DESY Higgs analysis group

DESY

February 19, 2013

DESY CMS group meeting @ Zeuthen
DPG 2013 rehearsal

Outline

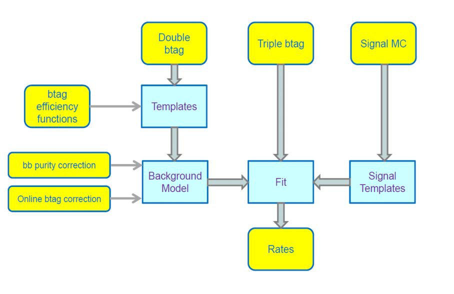
Introduction

- Discovery of a new Boson was announced on the 4^{th} of July
- Latest mass measurement
  $m_X = 125.8 \pm 0.4(stat.) \pm 0.4(syst.)$ GeV
- Consistent with a Higgs boson. But which the Higgs boson have we observed?
  - SM? MSSM? NMSSM?
- Higgs sector of MSSM
  - 5 scalars: 2 **CP-even** (h/H); 1 **CP-odd** (A)**; 2 charged** $(H^{\pm})$
  - **only two parameters:** $tan\beta = v_u/v_d;\ M_A$

- Main channel of neutral MSSM Higgs boson production at LHC
  - b-quark radiation off: $gb \to b\Phi$
  - gluon-gluon fuison: $gg \to bb\Phi$
  - $\Phi = h, H, A$
- $tan^2\beta$ enhanced cross section
- Light *CP-even* h Higgs boson as 'SM' one at large M_A
- Degeneration of A, H states at large M_A
- The current analysis of
  - all-hadronic and semileptonic channels at
  - $\sqrt{s} = 7$ TeV data recorded in 2011 in CMS

All-hadronic channel

- Events with 3 leading (highest-p_T) jets to be **b-tagged**
- Signal as a peak in invariant mass distribution of the two leading jets (M_{12})
- Background mainly from multijet QCD composed of
  - events with 3(2) **b-** jets and 1 mistagged **c/u/d/s/g-** jet
- Contributions from $t\bar{t}$, $Z + jets \leq 1\%$



- Signal yield from **2D** fit on M_{12} versus X_{123}

All-hadronic trigger and event selection

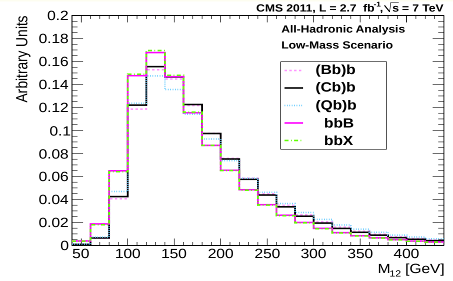
Triggering events

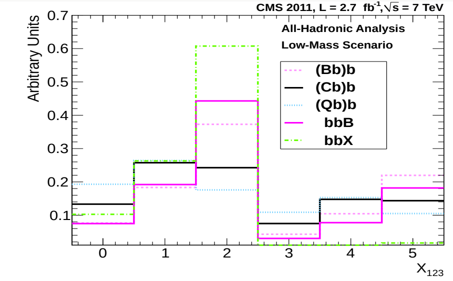
- accept events with
  - 2 leading jets with $|\eta| < 2.6$ and p_T $p_T > 46, 38(60, 53)$ GeV accordingly
  - **b-** tagging jets
  - 2 Higgs boson mass scenarios: Low(Medium)
  - recorded integrated luminosity $2.7 fb^{-1}(4.0 fb^{-1})$

Selection of events

- at least 3 reconstructed jets with $|\eta| < 2.2$, $p_T > 46, 38, 20(60, 53, 20)$ GeV
- with minimal separation $\Delta R > 1$ between two leading jets
- 3(2) leading jets pass tight **CombinedSecondaryVertex** btagging selection

Background model

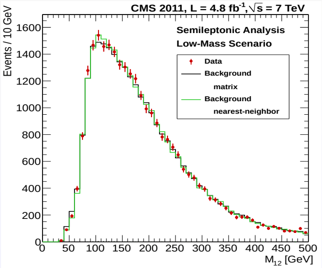


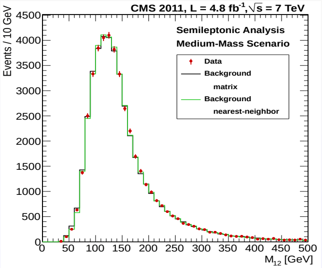


- Modeling background in triple-btagged data sample
  - take selected double-btagged data events
  - **bbX** , **bXb** , **Xbb** with untagged **X** jet
  - weight event-by-event by b-tagging probability of the assumed **X** flavor
    - **2D** templates $\mathbf{M_{12}}$ versus $\mathbf{X_{123}}$ to model background
    - $\mathbf{X_{123}}$ based on Mass of reco Secondary Vertex (SV) to provide additional separation between *b-* , *c-* and *udsg-* jets
  - 5 templates, after merging templates close in shape

Semileptonic channel

- Search of MSSM Higgs Signal in events
  - with 3 leading (highest-p_T) jets to be **b-identified**
  - with at least one non-isolated muon
  - as a peak in invariant mass distribution of the two leading jets (M_{12})
- Background
  - multijet QCD composed of events with 1 or 2 **b-tagged** jets
  - contributions from $t\bar{t}, Z + jets \lesssim 1\%$
  - data-driven estimation based on **double/single-btag-data** sample in control region





Event selection in semileptonic channel

- Triggering requires
  - muon with $p_T > 12$ GeV
  - either 1 or 2 jets with $p_T > 20(30)$ GeV and $|\eta| < 2.6$
  - 1 or 2 leading jets online **b-tagged**
  - total recorded integrated luminosity $4.8 fb^{-1}$
- Offline selection
  - at least 3 reconstructed jets having
    - $|\eta| < 2.6$, $p_T > 30, 30, 20$ GeV
  - at least one muon with $p_T > 15$ GeV. to be in one of 2 first leading jets
  - minimal separation $\Delta R > 1$ for all pairs among 3 leading jets
  - **b-** identification of 2(3) leading jets

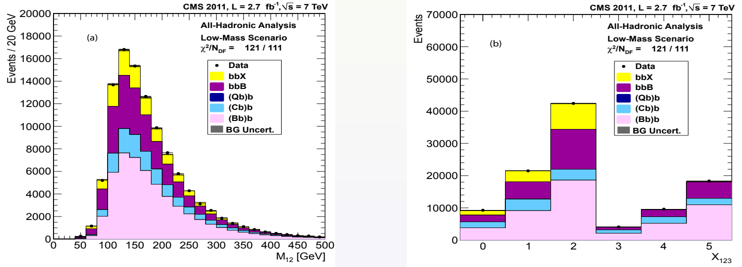
Systematic uncertainties

- systematic uncertainties on the expected signal and background estimates
  - affect the cross section estimation
  - consequently its interpretation within MSSM (*)

Source	All-hadronic	Semileptonic	Type
Trigger efficiency	10%	$3-5\%$	rate
Online b-tagging efficiency	32%	–	rate
Offline b-tagging efficiency	10–13%†	12%	shape/rate
b-tagging efficiency dependence on topology	6%	–	rate
Jet energy scale	1.4–6.8%	3.1%	shape/rate
Jet energy resolution	0.6–1.3%	1.9%	shape/rate
Muon momentum scale and resolution	–	1%	rate
Signal Monte Carlo statistics	1.1–2.6%		rate
Integrated luminosity	2.2%		rate
PDF and α_s uncertainties	3–6%*	2.7–4.7%*	rate
Factorization and renormalization QCD scale	6–28%*		rate
Underlying event and parton showering	4%*		rate

Results. Signal extraction

- All-hadronic channel
- Signal estimation from a fit of $N_{bbb}(f_{sig}T_{sig} + \sum_i f^i_{bgd}T^i_{bgd})$
- $T_{sig(bgd)}$ - corresponded $2D$ templates in $M_{12}-$ and $X_{123}-$ space
- The fit results for low-mass scenario

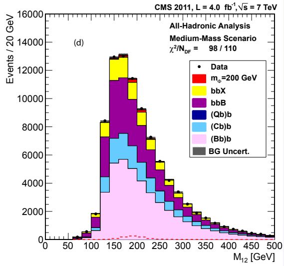


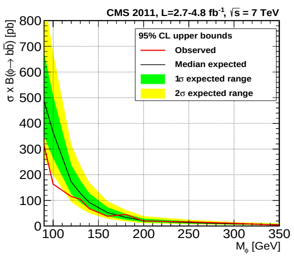
No significant deviation from background

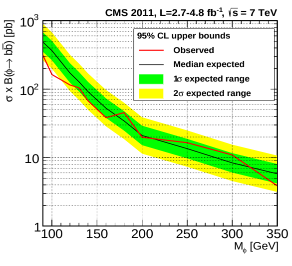
Results. Upper level 95% CL limit on $\sigma \times Br$

- combination of all-hadronic and semileptonic channels
  - all-hadronic case: remove common to semileptonic events
- CL_s criterion to determine $95\% CL$ limit on signal contribution
  - *using* `RooStats` *package*

Fit w/ Signal @ $M_\Phi = 200$ GeV



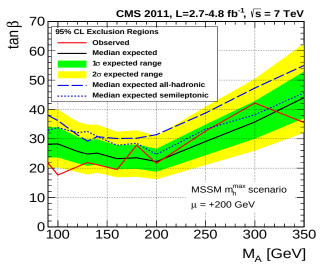


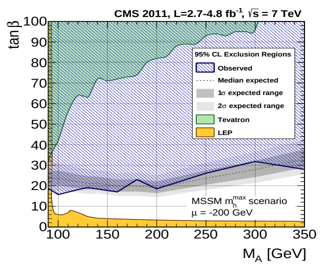


- observed upper limit ranges from 312 pb at $M_A = 90$ GeV to 4 pb at $M_A = 350$ GeV

Results. MSSM interpretation

- MSSM m_h^{max} benchmark scenario.
  - $M_{SUSY} = 1\,TeV,\ X_{t(b)} = 2M_{SUSY},\ \mu = \pm 200\,GeV,\ M_{\tilde{g}} = 800\,GeV,\ M_2 = 200\,GeV,\ M_3 = 800\,GeV$
- Results as functions of the MSSM parameters M_A and $tan\beta$
  - Combination of all-hadronic and semileptonic channels
    - @ $\mu = +200\,GeV$
    - @ $\mu = -200\,GeV$ with limits set by Tevatron and LEP
  - The 95% CL bound on $tan\beta$ varies from 18 to 42





MSSM limits from the analysis supersede Tevatron results

Summary and conclusions

- Search for a Higgs boson decaying into pair of b quarks produced in association with one or more additional b-jets
  - First time at LHC
  - Highest sensitivity for MSSM bounds in $(tan\beta, M_A)$ space
- Total integrated luminosity of $2.7 - 4.8\, fb^{-1}$ collected in 2011 at $\sqrt{s} = 7\, TeV$ in CMS
- No observed signal above SM expectation
- Upper level 95% CL limit on $pp \to b\Phi + X$, $\Phi \to b\bar{b}$ $\sigma \times Br$ in 90-350 GeV range
- MSSM interpretation in m_h^{max} scenario in terms of bound in the space of M_A and $tan\beta$
- Excluding a region of parameter space previously unexplored for this final state

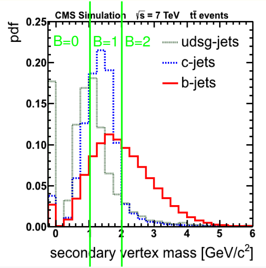
Backup slides

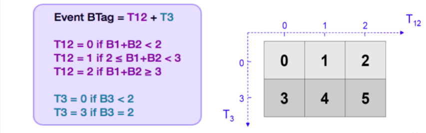
Backup slides

Event reconstruction

- The CMS particle-flow event reconstruction
  - optimized reconstruction and identification of particles in the event
- Reconstructed primary vertex
  - the largest $\sum p_T^2$ of associated tracks
- Reconstructed jets
  - anti- k_T algorithm with $R = 0.5$
  - Pileup subtraction
- Reconstructed muons
  - a global track fit
  - track-candidates from
    - inner silicon tracker
    - outer muon system
- Identification of b- quark
  - Combined secondary vertex (CSV) algorithm
    - Likelihood discriminant based on track impact parameter and secondary verices in a jet
    - $CSV > 0.89$

EventBtag





- EventBtag
  - based on Mass of reco Secondary Vertex(SV)
  - provide additional separation between *b-* , *c-* and *udsg-* jets
- SV distributions divided in 3 bins
  - $M_{SV=}[0,1);\ [1,2);\ [2,\infty)$
  - Defined variables
    - T_{12} for 1^{st} and 2^{nd} leading jets
    - T_3 for 3^{rd} leading jet

Background estimation for semileptonic channel

- Background determination
  - two data-driven methods: **Matrix method** and **Nearest-neighbor method** on exlusive data regions
  - requires control region (CR) of background enriched data obtained with
    - two Likelihood discriminators in mass scenarios: $M_\Phi \leq 180$ GeV and $M_\Phi > 180$ GeV

- Matrix method
- uses double-b-tag (bbj) control data-sample region
- weights M_{12} event-by-event with probability of 3^{rd} jet to be b-tagged:

  $P_3 = \varepsilon_b \cdot f_b + \varepsilon_c \cdot f_c + \varepsilon_q \cdot (1 - f_b - f_c)$
- quark fractions f from the fit of 2 templates in CR

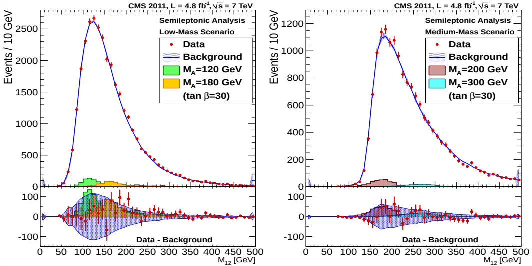
- Nearest-neighbor method
- uses single-b-tag (bjj) and triple-b-tag (bbb) CR(**train**)/SR(**test**) of data-sample
- parametrizes event kinematics by variables x_i with probabilities w_i to pass selection
- weights M_{12} event-by-event with

  $P_{bbb} = \frac{N_{bbb}}{N_{bjj} \cdot D^2}$, $D^2 = \sum_i w_i^2 (x_i^{test} - x_i^{train})^2$

Results. Signal extraction in semileptonic channel

- Semileptonic channel
- Signal extracted as $M_{12}^{sig} = M_{12}^{data} - M_{12}^{bgd}$
- The predicted background with an expacted signal for two Higgs boson masses $M_{\Phi} = 120, 180$ GeV at $tan\beta = 30$ for medium-mass scenario

Low-mass scenario Medium-mass scenario



No significat deviation from background