Status of H->bb/cc/gg Note

Bai Yu

2017-08-14

Comments From Referee

- The background in current analysis are fixed by theory prediction
    - Updates:
        - Define control regions of the dominant backgrounds
        - Include the control regions in the template fit in the way that the normalization of dominant backgrounds are float

The Dominant Backgrounds

mumuH

Semi-leptonic ZZ->mumuqq

eeH

Semi-leptonic SZ->eeqq

nnH

Semi-leptonic ZZ->nunu qq, SZ->nunu qq, WW->mu nu qq/tau nu qq, qq

qqH

Hadronic ZZ/WW->4 q, qq

In addition H->WW/ZZ->4 q are considered in each channel

Major consideration:

- **Purity**
- **Statistics**
- **Mutually exclusive**

Control Region of muumuH Channel

	Signal	mumuH bkg	zzsl->mumuqq	other background
Event Yields	11.84k	1.23k	3.81k	23.4

- ZZ->mumuqq control region: similar to signal region but different Mjj and M_mumu_recoil
  - 65 GeV<Mjj< 105 GeV
  - 70 GeV<M_mumu_recoil < 105 GeV

Purity: 99.91%
Statistics: 100.7k

- Higgs background control region: same as control region but require y34>0.025 and blikeness<0.5, clikeness<0.5

Purity: 64.9%(mumuH_bkg, 24.3% zzsl->mumuqq, 7.8 mumuH->mumu gg)
Statistics: 446

By applying constrains from control region, no significant difference found from by fixing them directly in template fit.
much improved by completely release them

Control Region of nnh Channel

	Signal	nnH_bkg	swsl_qq(very low stat.)	ww_sl_tauq	ww_sl_mu_up	sznu_nu_up	sznu_up_down	zz_sl_un_up	zz_sl_nu_down
FSClasser	167500	71934	13200	7.11162M	7.12766M	282392	459391	411970	678581
cos θ_{jj}	102186	5078	87	302146	5598	5032	6821	9258	11927
BDT>0.07	82371	1927	1	6637	285	630	931	1097	1609

	ww_cuxx	bbrad	ccard	ssrad	uurad	ddrad	bbnonrad	ccnonrad	ssnonrad
FSClasser	1.70919M	234349	267377	291697	296770	291253	218209	236130	267998
cos θ_{jj}	92	34	27	7	4	5	344	117	26
BDT>0.07	1	1	0	0	0	0	68	10	0

	uunonrad	ddnonrad	bbfilter	ccfilter	ssfilter	uufilter	ddfilter
FSClasser	258736	267416	728308	1.09601M	788760	1.11429M	794210
cos θ_{jj}	7	14	526	239	40	14	23
BDT>0.07	0	0	91	23	4	0	2

	eeh_bb	eeh_cc	eeh_gg	eeh_bkg	mumuh_bb	mumuh_cc	mumuh_gg	mumuh_bkg	tautauh_bb	tautauh_cc	tautauh_gg	tautauh_bkg
FSClasser	55058	2514	8815	29354	54974	2579	7994	28422	57909	2742	8537	31089
cos θ_{jj}	0	0	0	4	1	0	0	0	1333	44	102	150
BDT>0.07	0	0	0	0	0	0	0	0	380	8	38	4

	qqh_bb	qqh_cc	qqh_gg	qqh_bkg	nnh_bb	nnh_cc	nnh_gg	nnh_bkg
FSClasser	417960	19411	61462	223861	140472	6547	20481	71934
cos θ_{jj}	3	0	0	2670	85670	4524	11992	5078
BDT>0.07	2	0	0	357	69820	3029	9522	1927

Background with considerable need contribution in Signal Region

Background with considerable contribution in other background's control region

Control Region of Semi-leptonic WW events

WWSL_Muqq

- **2 jets final states with 1 isolation muon**
- **Muon energy > 50 GeV**
- **Visible mass between 170 and 250 GeV**
- **Missing mass < 50 GeV**

Purity : 98%
Statistics : 70k

WWSL_Tauqq

- **3 jets in final states, isolation lepton veto**
- **at least one tau jet candidate**
  - **3-prong or 1-prong**
  - **95% of energy deposited in a cone**

  θ(etcone95) which cos θ = 0.95
- **tau jets candidate with largest etcone considered**

as the tau, and other 2 jets treated as QCD jets

- **QCD jet pair invariant mass between 60-100 GeV**
- **cos $\theta_{\text{tau-qcd}}$ between -0.5 and 0.5**
- **cos $\theta_{\text{qcd-qcd}}$ between -0.5 and 0.5**
- **Visible mass between 135 and 180(to exclude Signal region)**
- **y34<0.01(optional, to remove ZZ/WW->4jets)**

purity: 97%-98%
Statistics : 600k-640k

qq and ZZ/WW->4jets

- Both of them are important to nnh and qqH channel
- They are still in progress:
  - QQ
    - 3-category of qq sample: 1)without photon and gluon radiation, 2) with gluon radiation but no photon radiation 3) with photon radiation
    - qqH are dominantly affected by 2) and nnH dominantly affected by 1), so define them in 2-jets and 4-jets final states respectively.
  - WW/ZZ hadronic:
    - Not easy to get purified sample with specific flavor components: define control region according to jet pair invariant mass to include them in the control regions

Summary and Plans

- Working toward analysis with dominant background constrained by control regions.
- Template fit code upgraded to incorporate control regions(not necessary to fit flavor tagging information but can also be other distributions)
- Control sample for irreducible background in llH sample done. With high purity and sufficiently large statistics of the control region, fit results seems reasonable
- Control regions for WWSL events studied, the purity and statistic is satisfied.
- The control region for qq and ZZ/WW hadronic sample under development.
- Get higgs background control regions with high purity and statistic seems failed, but seems not affected much on the results. Maybe y_ij value can be added in the variable to be fitted(instead of fitting event number in blikeness-clikness bins).
- All the control regions should be done in this week. These results will be presented in the 2nd version of the notes.