Status on combinations and BSM gg → H → hh → WWγγ search with jjjjγγ final state

Yaquan Fang, Xiaohu Sun
04-06-2014
IHEP, Beijing

Combination SM HH non-resonant

- With bbyy final state, expected upper limit 1.0 pb (HH) corresponding to **~114.5** times of SM HH production
  - Also, there is **2.4** standard deviation in observation
- With bbbb final state, expected upper limit corresponding to ***~40*** times of SM HH production by using m(lead) vs m(sub) corrected from 6th May, before including systematic uncertainties and before re-optimizing the cuts that are used for resonant search
- Regarding the limited differences on the sensitivities, it is still worthy to combine both results with Run I
  - To obtain a better upper limit on SM HH production
  - As well as, if possible, to extract a “significant” *significance*
  - To serve as a good reference for Run II
- The machinary for combination is in place, all we need to do is
  - Gather the workspaces from both analyses after optimization
  - Converge on the correlated uncertainties (lumi, JES, isr/fsr etc.)

*Checks on the overlapped phase space in two analyses, should be negligible

Combination BSM H → hh

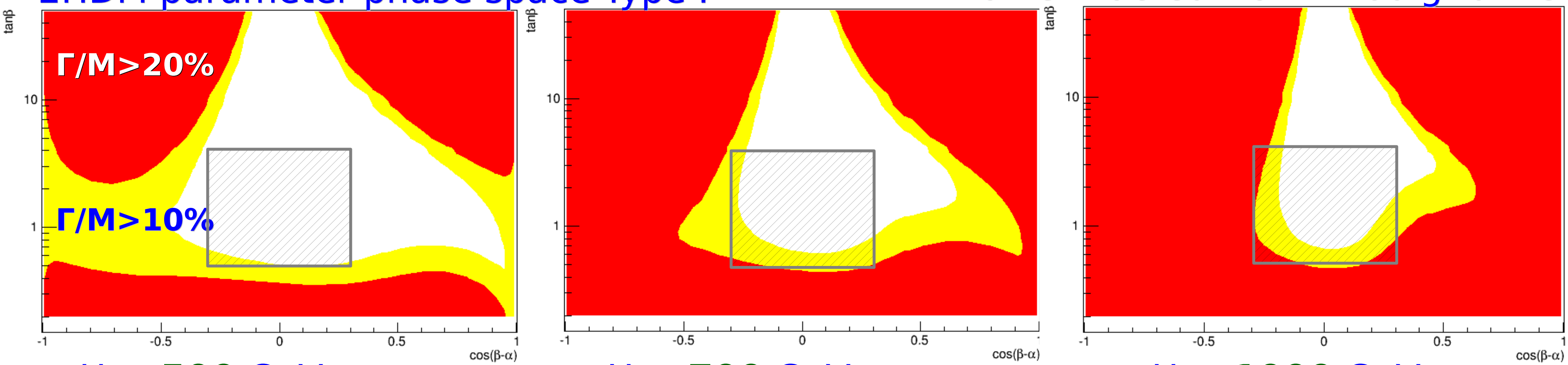
- The scanned mass points: bbyy final state covers from 260 GeV to 500 GeV, while with bbbb final state from 500 GeV 1000 GeV
  - bbbb signal acceptance drops significantly below 500 GeV
  - It seems that we can only combine in high mass region if bbyy final state can extend the search
- Well, if only looking at 500 GeV, by eye catching on the limit plots:
  - Expected upper limit on gg → H → hh (bbyy): **0.8** pb
  - Expected upper limit on gg → H → hh (bbbb): ***0.1*** pb from 6th May without systematic uncertainties
- At least at this joint mass point, both analyses are comparable with respect to the sensitivity
- Low mass (<500GeV), bbyy definitely has more sensitivities
- High mass (>500GeV), it is still hard to say now

Combination BSM H → hh (other issues)

- To interpret for 2HDM, in **high** mass region, one has to re-**check resonant width**, maybe has to redefine a smaller window

2HDM parameter phase space Type I

* all extracted from v160 grid file



mH = **500** GeV mH = **700** GeV mH = **1000** GeV

@ 500 GeV, consistent with what was checked by German
The windows is proposed by German to interpret bbyy result for 2HDM
cos(b-a) ~ [-0.3,0.3]; tanb ~ [0.5,4]

- The VBF, bbH production may also vary in high mass region, need to check when interpreting for 2HDM
- The signal templates are different: for bbyy final state, latest HeavyScalor in MG5; for bbbb final state, 2HDM in MG5
- As mentioned in SM HH comb, the overlapped phase space should be negligible

Combination BSM A → Zh and H → hh

- The basic idea is to combine the measurements from two different production: A → Zh, H → hh, by using one scale μ for both cross section
- Then by using this scale μ, one can provide the upper limits for each production as well as makes constraints in the tanb vs cos(b-a) plane
- If one assumes the cross section of H → hh is μ, then the cross section of A → Zh should be μ***(A/H)**
  - Then the two measurements are correlated in the combination, leading to the possibility of obtaining from the fit the combined upper limits or "combined significance"
  - In the combined fit, the only POI is μ and (A/H) exists as a function of b and a
  - Due to the varying (A/H), one has to extract the upper limit for each point in the phase space to see if this certain point is rejected or not (***quite*** computing-consuming, any other ideas?)

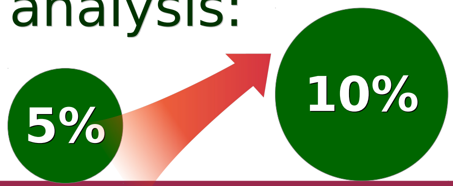
Glance at H → hh → WWyy → jjjjyy

- In parallel, we started to look at gg → H → hh → WWyy with W hadronic decay leading to final state of jjjjyy
  - h→WW has the second largest branching ratio after h→bb
- Apply the same cuts from yy side, then ask njets>=4, estimate roughly the expected upper limit, and then additionally apply MVA cut to see the improvement on the expected upper limit

	Lumi (pb-1)	Branching ratio	Cut eff (yy&Njet>=4)	Upper limit	Cut eff (additionally MVA)	Upper limit
Non-resonant SM HH	20,000	4.48e-4	15%	**18 pb**	15%*93%	**7.2 pb**
Resonant 300GeV	20,000	4.48e-4	9%	**30 pb**	9%*63%	**14 pb**

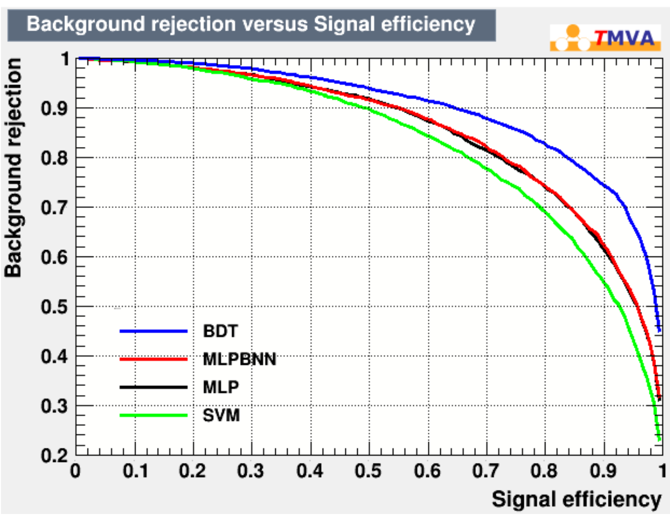
*The MVA is trained with signal MC sample and background from sideband
This leads to signal eff = **63%**, bkg eff = **9%** for resonant,
signal eff **93%**, bkg eff = **13%** for non-resonant

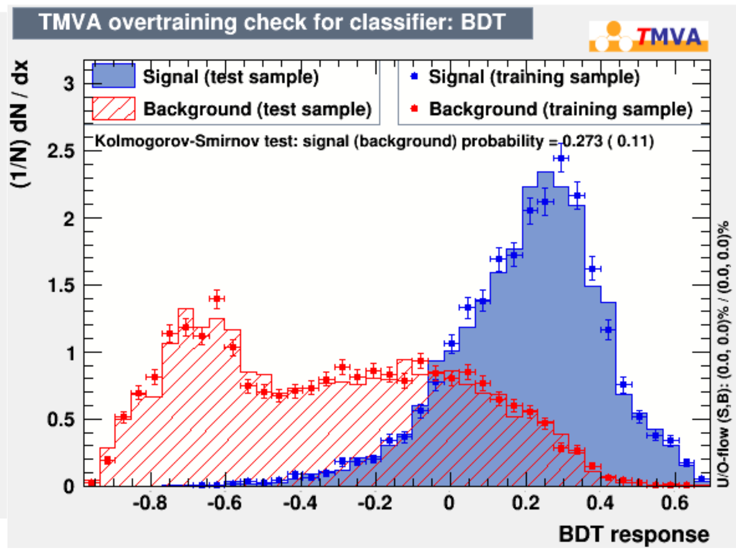
Compared to the expexted upper limits from bbyy analysis:
* Non-resonant: **1.0** pb
* Resonant @ 300GeV: **1.5** pb

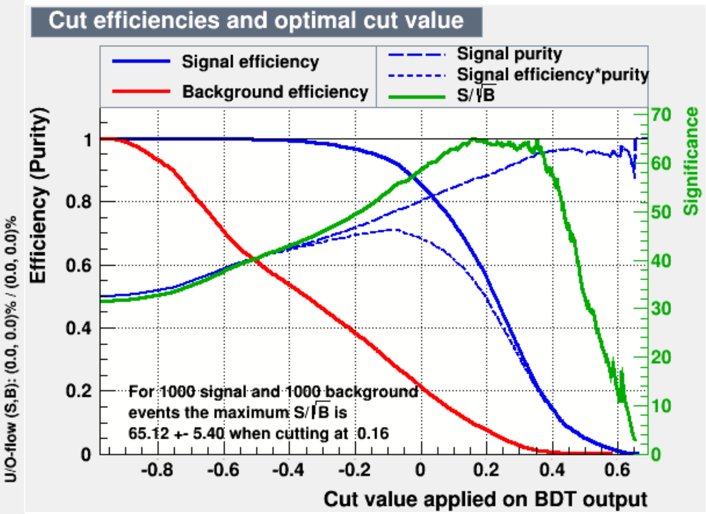


MVA with JxJy

- Train with the variables from JxJy instead of J1J2
- Much more better performance is obtained





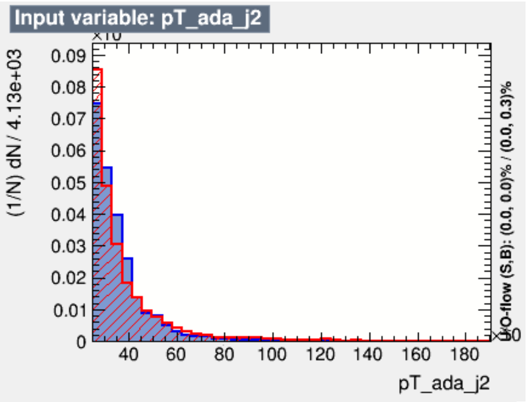


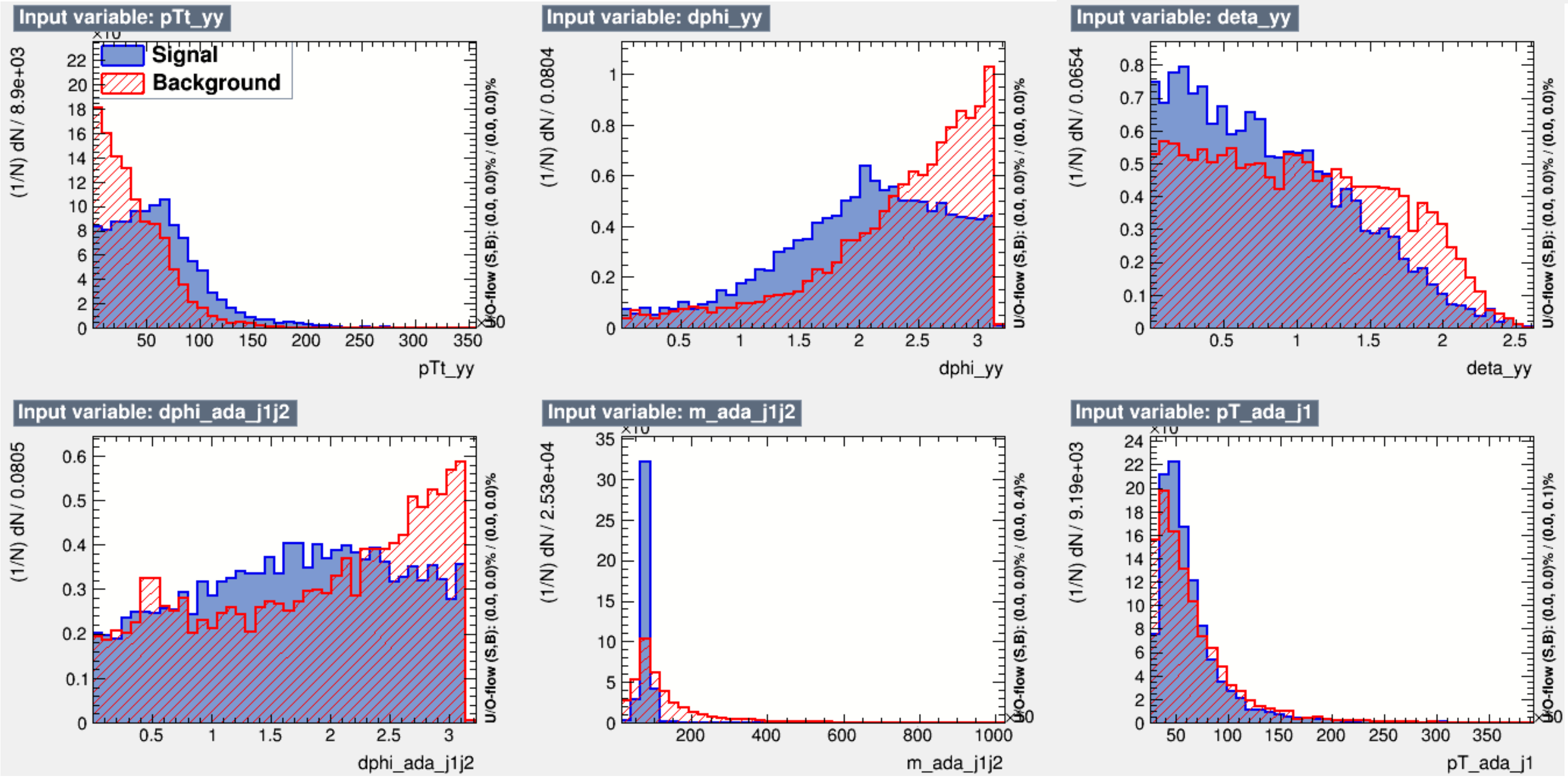
Classifier	(#signal, #backgr.)	Optimal-cut	S/sqrt(B)	NSig	NBkg	**EffSig**	**EffBkg**
MLP:	(1000,1000)	0.6192	(55 +- 4.8)	515.8762	87.92532	0.5159	0.08793
MLPBNN:	(1000,1000)	0.6068	(54.9 +- 4.8)	545.217	98.76543	0.5452	0.09877
BDT:	(1000,1000)	0.1597	(65.1 +- 5.4)	633.2395	94.54983	**0.6332**	**0.09455**

The signal eff is kept 63% while the bkg eff is lower 9%
The relative S/sqrt(B) is leveled up to 65.1

MVA inputs with adaptive method

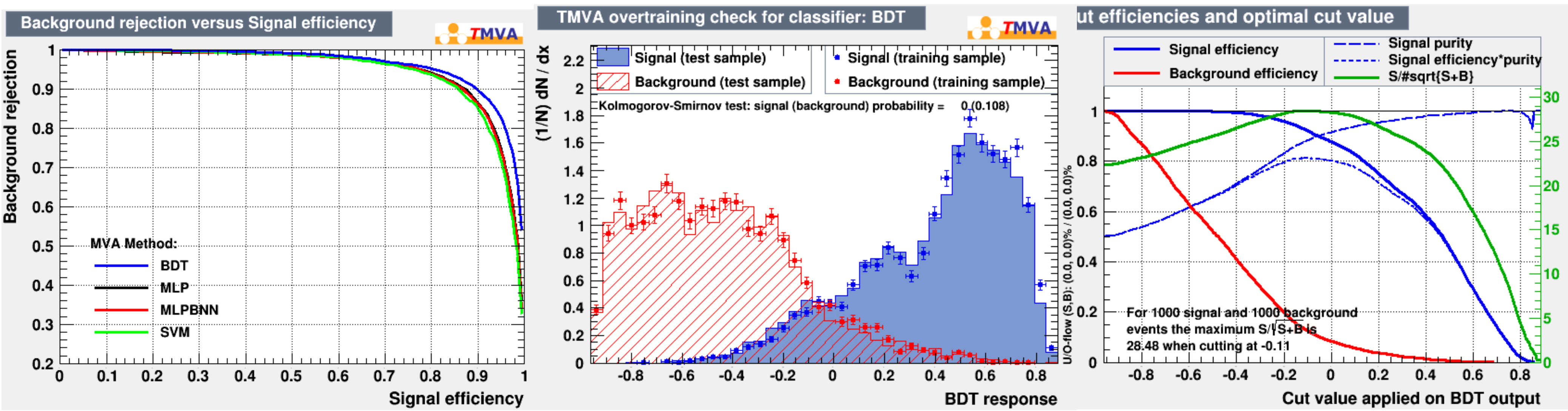
- J1J2 variables are replaced by the Jx Jy variables obtained by adaptive method





MVA with JxJy for **SM HH**

- Train with the variables from JxJy instead of J1J2

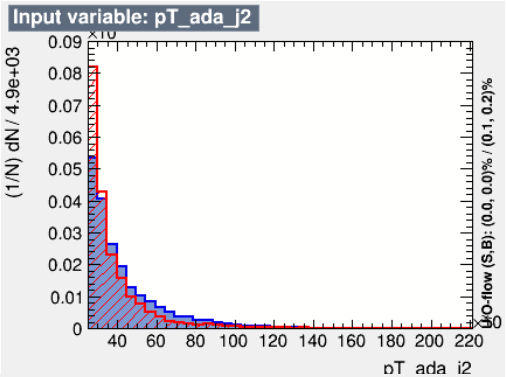


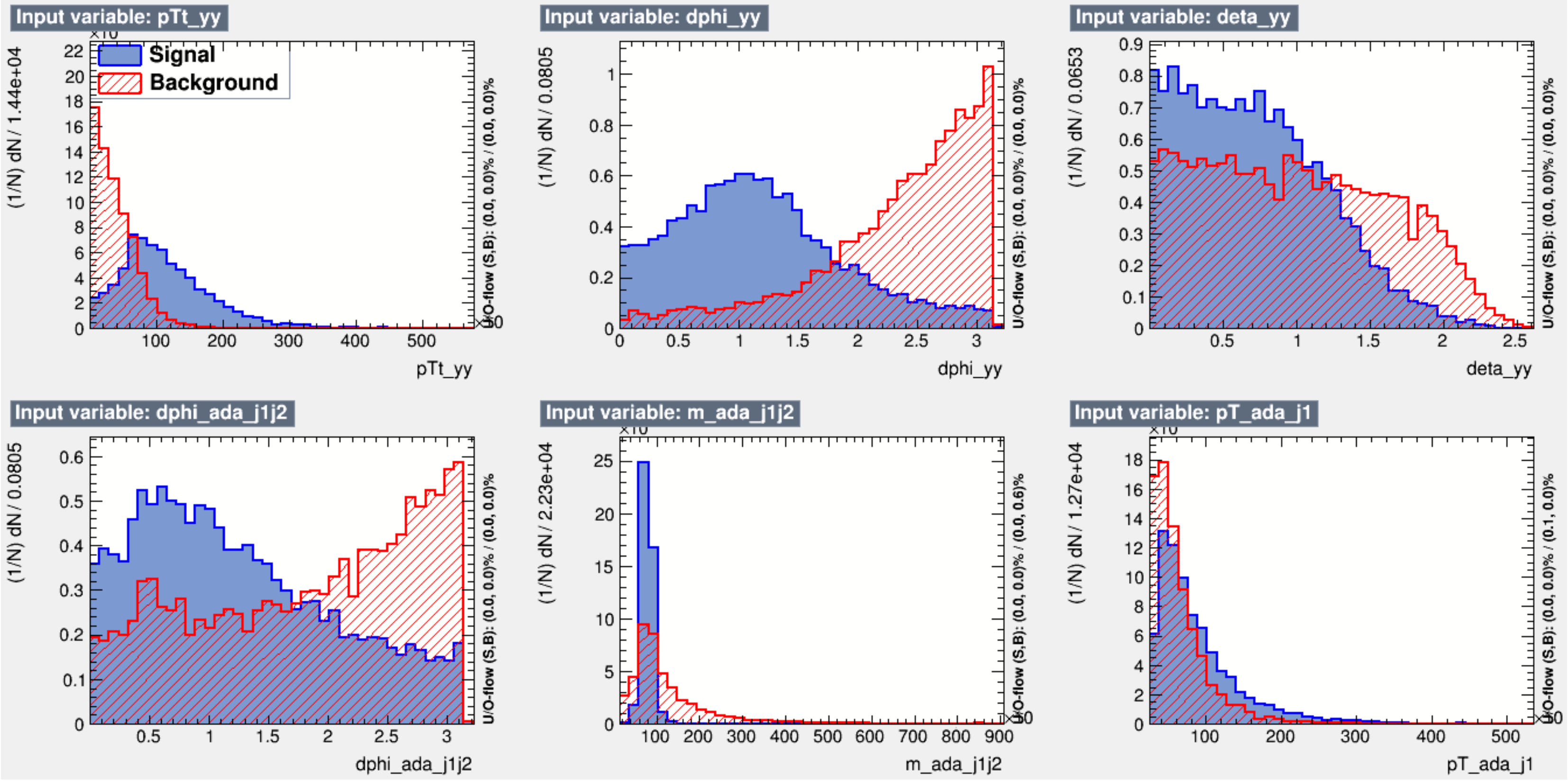
Classifier	(#signal,#backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg	**EffSig**	**EffBkg**
MLP:	(1000, 1000)	0.4930	27.9795	895.2804	128.5757	0.8953	0.1286
MLPBNN:	(1000, 1000)	0.4863	27.9399	901.6102	139.717	0.9016	0.1397
SVM:	(1000, 1000)	0.4642	27.8195	893.3926	137.9103	0.8934	0.1379
BDT:	(1000, 1000)	-0.1115	28.4816	925.4858	130.3824	**0.9255**	**0.1304**

The signal eff is 93% while the bkg eff is 13%
The relative S/sqrt(B) is up to 28

MVA inputs with ada method (**SM HH**)

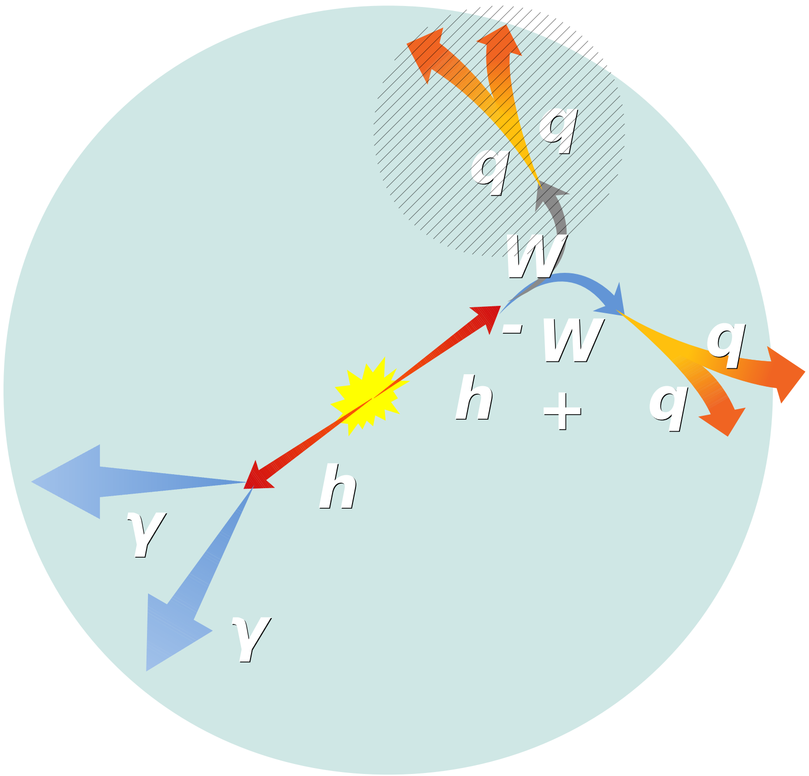
- J1J2 variables are replaced by the Jx Jy variables obtained by adaptive method





A new idea to try

- We all know missingET, which is due to the undetectable neutrino
- In our case, we don't have missingET, instead, we have a missing W since our pT requirements on jet almost remove all the 3rd and 4th jets
- We can try to reconstruct the other W boson by the momentum conservation law in transverse plan!



pT(yy) + pT(jxjy) + pT(missW) + pT(residual)= 0

To be calculated

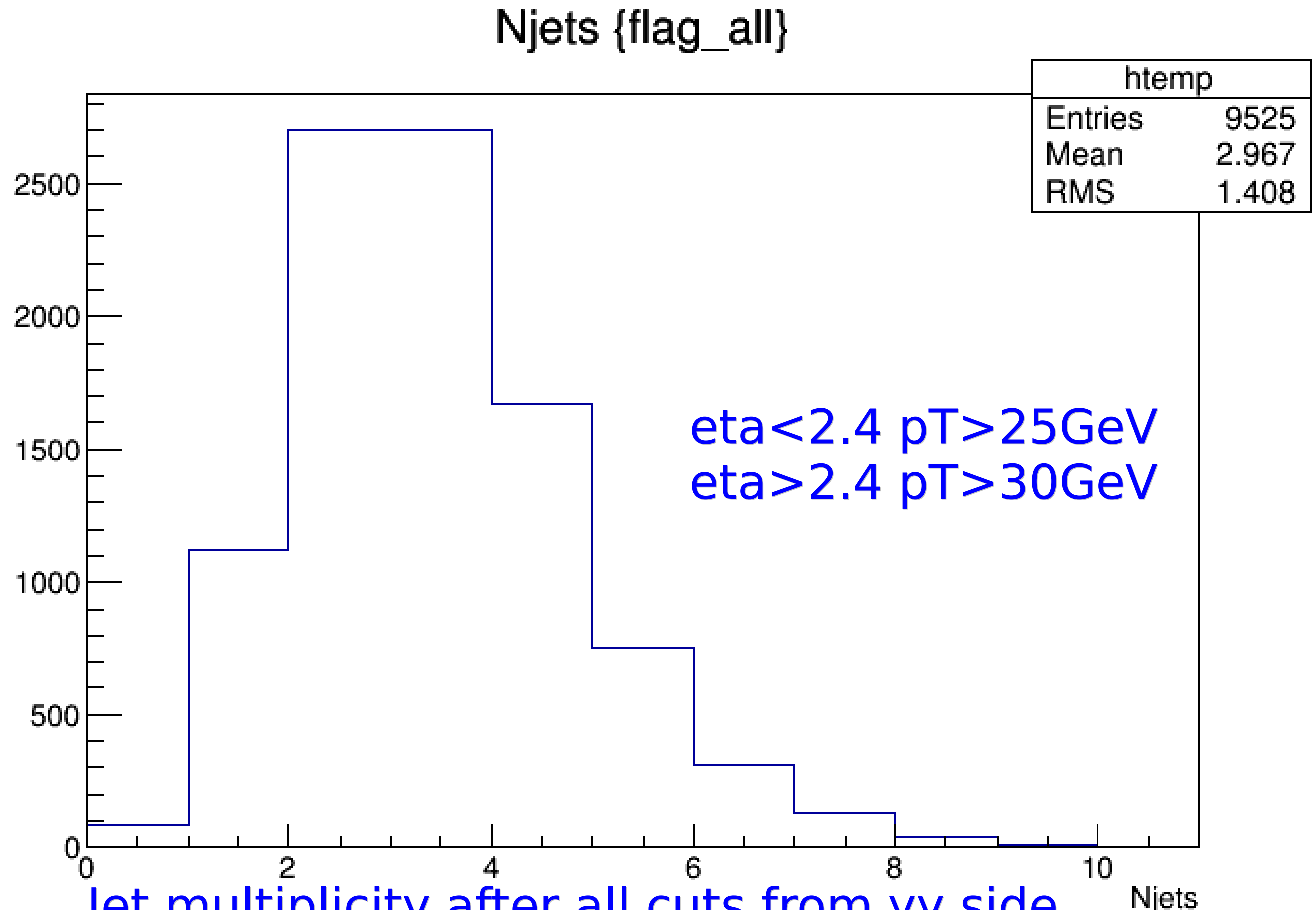
Backup

Introduction

- Sample page

Additional cuts

- Additionally, we will cut on kinematics of the children from the other Higgs boson
  - To do some studies on cuts: deltaPhi(j,j) deltaEta(j,j)
  - To find the more correct combinations of jets originating from the same W boson



Jet multiplicity after all cuts from yy side

Cutflow (mainly yy-cuts)

- Use h → yy trigger and get the gamma pair
- Follow the same preselection in HSG1, check on cutflow (private MC production 28k):

	Unweighted evt	Cut eff (%)
generated	28000	100%
trigger	19953	71%
GRL	19953	71%
detector errors	19953	71%
vertex tracks	19953	71%
pre-selection	15547	56%
photon pT	13996	50%
photon ID	12041	43%
photon isolation	9525	34%
diphoton mass	9516	34%

Additionally	Njets>=4	2913	10%
And	XXX	-	5%

Descriptions of cuts

	descriptions
generated	Generated events
trigger	EF g35 loose g25 loose (8TeV)
GRL	Good Run List
detector errors	LAr error, TileError, BadJet
vertex tracks	Primary vertex track requirements
pre-selection	At least two loose photons
photon pT	1^{st} > 40 GeV, 2^{nd} > 30 GeV
photon ID	IsEM & 0x45fc01 == 0
photon isolation	Calo (etcone40) < 6GeV && track < 2.6GeV
diphoton mass	[100GeV, 160GeV]

BSM gg → H → hh → bbγγ

	SM Benchmark	Resonant NWA 260 GeV	300 GeV	350 GeV	500 GeV	1000 GeV
Generated	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Trigger	73.1%	72.5%	71.6%	71.8%	73.6%	81.0%
Preselection	57.3%	56.7%	56.1%	56.2%	57.7%	65.1%
Photon p_T	51.6%	51.6%	49.8%	49.2%	52.5%	62.4%
Photon Identification	45.3%	44.2%	42.8%	42.6%	46.4%	56.2%
Isolation	39.1%	33.1%	33.8%	35.9%	40.6%	47.4%
$105 < m_{\gamma\gamma} < 160$ GeV	39.0%	33.0%	33.8%	35.9%	40.5%	47.4%
2 Central Jets	33.9%	25.5%	26.9%	29.8%	36.2%	45.1%
Tagging	12.5%	8.4%	8.9%	10.0%	14.1%	19.1%
b p_T Cuts	10.1%	4.8%	5.6%	7.2%	12.0%	18.1%
$95 < m_{b\bar{b}} < 135$ GeV	7.4%	4.0%	4.3%	5.3%	8.6%	14.2%

Reconstructed jets >= 4

- By matching with truth information of jets, one find calculate the correctness of the well-reconstructed jet pairs

specific	#evts	%
none (all cuts form yy side)	2914	-
any pairs from W	1887	65%
1 pair from W	1704	**58%**
2 pairs from W	183	**6%**
the pair from 1st W	1243	43%
the pair from 2nd W	827	28%

*all numbers except the case of 'none' are calculated by asking njets>=4
pT > 25 GeV when |eta| < 2.4
pT > 30 GeV when |eta| > 2.4

So, instead of asking for >=4 jets and trying to reconstruct both Ws,
one may keep higher statistics and obtain better sensitivities by
Looking also 2,3,>=4 jets and reconstruct only one of the Ws from 2 jets

Reconstructed jets **>=2**

- Redo the cutflow and the correctness tables

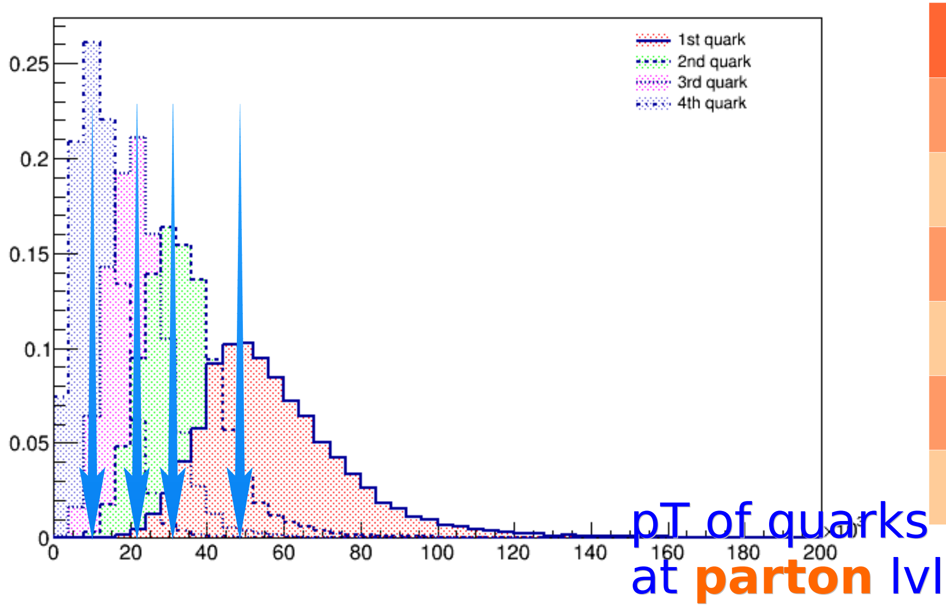
	Unweighted evt	Cut eff (%)
generated	28000	100%
trigger	19953	71%
GRL	19953	71%
detector errors	19953	71%
vertex tracks	19953	71%
pre-selection	15547	56%
photon pT	13996	50%
photon ID	12041	43%
photon iso	9525	34%
diphoton mass	9516	34%

Njets>=2	**8312**	**30%**

specific	#evts	%
none (all cuts form yy side)	8312	-
any pairs from W	3909	47%
1 pair from W	3726	**45%**
2 pairs from W	183	**2%**
the pair from 1^{st} W	2631	32%
the pair from 2^{nd} W	1461	18%

BSM gg → H → hh → WWyy → jjjjyy

- Triggered by the recent results on H → hh → bbyy search
  - There is a significance of 3.0 standard deviations at mH=300GeV
- We look at second largest decay: H → hh → WWyy → jjjjyy
- Look at the pT of jets in parton level, find it quite difficult to reconstruct all four jets



pT threshold	evts(parton lvl)	efficiencies
non	19430	100%
5 GeV	17168	88%
10 GeV	11389	59%
15 GeV	5437	28%
20 GeV	1963	10%
25 GeV	**594**	**3%**

A first look at sideband region

Njets>=4

- Sideband region:
  - mass(yy) within [100,160] GeV
  - mass(yy) is excluded from $|m_h - \Delta m_h - m_{\gamma\gamma}| < 2\sigma_{\gamma\gamma}$
  - where mh=125.6, deltamh=0.15, sigma=1.6

sideband	# of evt
ggH	0.467175
VBF	0.123474
WH	0.0638113
ZH	0.0405459
ttH	0.138622
Continuum	?
In data	1170

Continuum?

There are large components in backgrounds not yet clear
Need to at least introduce pp→jjjjyy and pp→jjyy

Bkg samples, use bbyy continuum samples?

A first look at signal region

Njets>=4

- Signal region:
  - mass(yy) is required by $|m_h - \Delta m_h - m_{\gamma\gamma}| < 2\sigma_{\gamma\gamma}$
  - where mh=125.6, deltamh=0.15, sigma=1.6

sideband	# of evt
ggH	4.91724
VBF	1.0963
WH	0.570564
ZH	0.374228
ttH	1.34295
Continuum	?
est bkg*	143

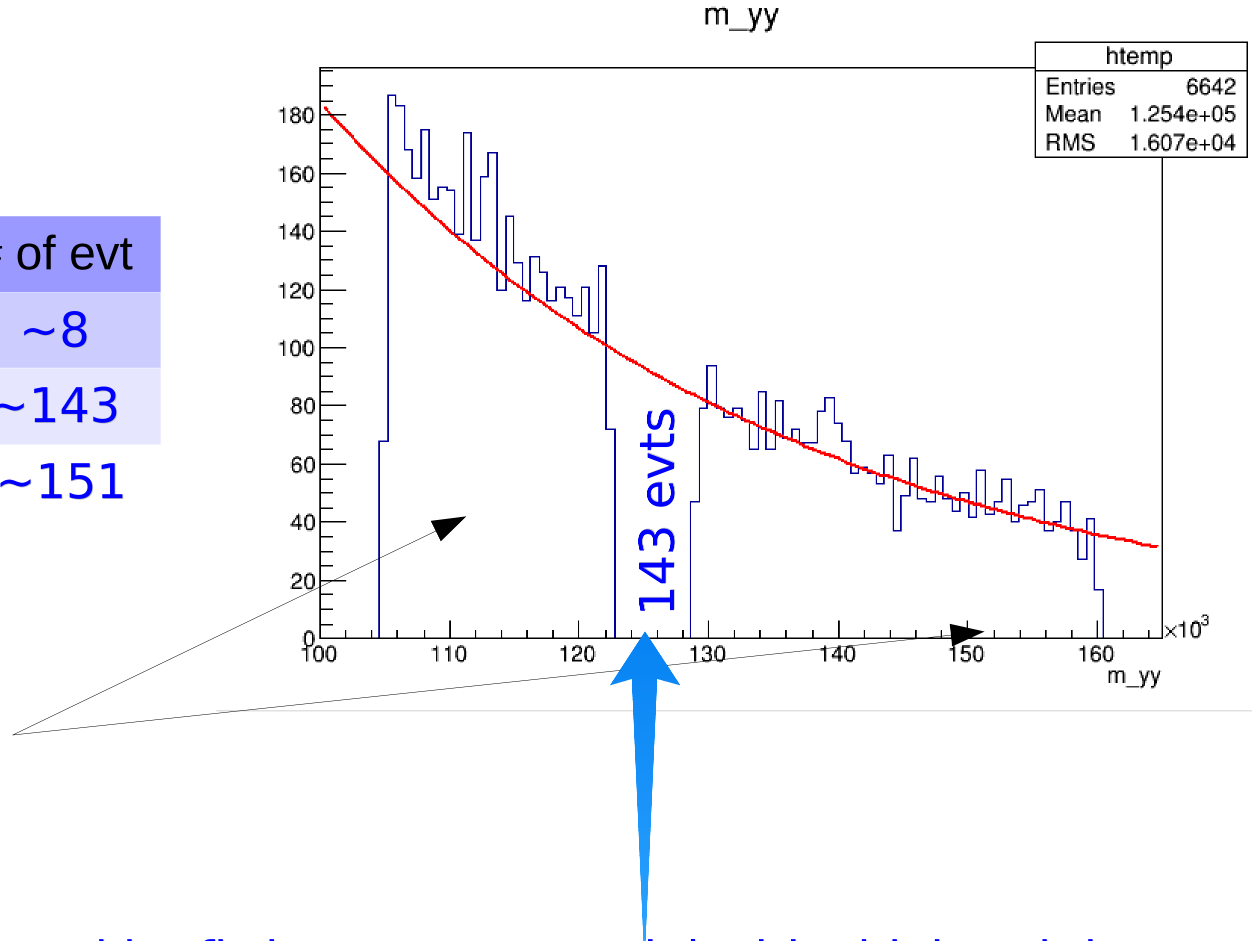
* bkg in signal region estimated by fitting to exponential with sideband data

Signal region event yields [EXP]

- Signal region (ask yy cuts && njets>=4):
  - mass(yy) is required by $|m_h - \Delta m_h - m_{\gamma\gamma}| < 2\sigma_{\gamma\gamma}$
  - where mh=125.6, deltamh=0.15, sigma=1.6

SIGNAL REGION

bkg components	# of evt
SM H (ggH,VBF,VH,ttH)	~8
Continuum	~143
	~151



* bkg in signal region estimated by fitting to exponential with sideband data

JxJy not has to be j1j2

- Till now I only assume j1 j2 are from a real W boson and use them to reconstruct it for simplicity
- An adaptive method is used to improve the correctness of finding the two jets from a same W boson here
  - By asking the invariant mass of jx jy, and choose the pair with the mass closest to W boson mass from PDG

6%

Using j1j2

The correctness of finding jet pair from the same real W boson

24%

Adaptive jxjy

*these correctness are calculated by using signal MC only @ mH=300GeV

JxJy not has to be j1j2

- Compare the invariant mass of jxjy in adaptive method and the one of j1j2 in the fixed method