Fully hadronic Higgs decay $H \to WW^* \to qqqq$ in Higgsstrahlung HZ, $Z \to qq$ at 250 GeV CepC

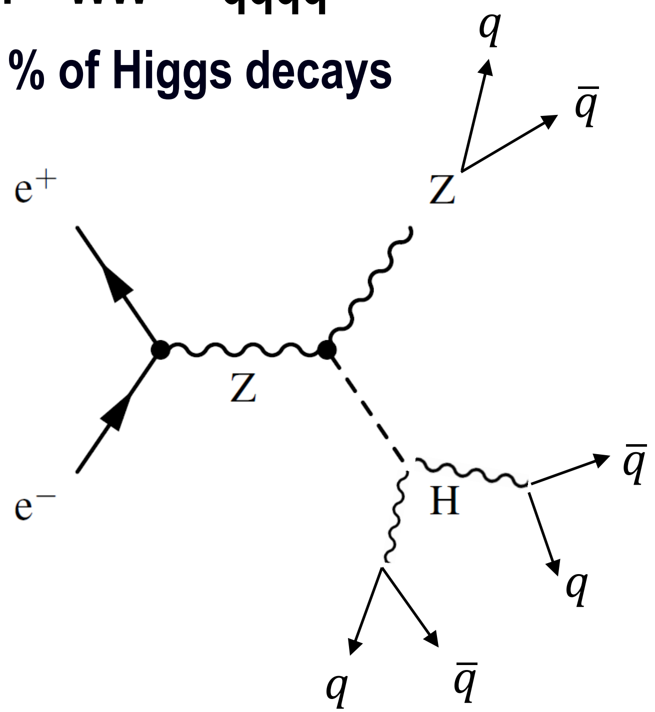
Mila Pandurović

Vinca Institute of Nuclear Sciences
University of Belgrade, Serbia

04 June 2018

Introduction

- **Analyzed HZ fully hadronic decay, signal : Z→qq , H→WW*→qqqq**
- **$BF_{H126\rightarrow WW}$~23.0% , $BF_{WW\rightarrow qqqq}$~45.4% ⇒ signal ~ 10 % of Higgs decays**



- $\sigma_{HZ,\, Z\rightarrow qq}$ ~ 143,39 fb (unpolarized beams)
- $\sigma_{(HZ,\, Z\rightarrow qq\,,\, H\rightarrow WW^*\rightarrow qqqq)}$ ~ 16,12 fb

- Measurement of the relative branching fraction

$$\frac{g^2_{HZZ} \cdot g^2_{HWW}}{\Gamma_H}$$

- **Signal signature: 6 central jets in the final state**
- Goal of the analysis:
  - Calculate the statistical potential for the determination of the specific Higgs couplings
  - Verify the analysis strategy

Two Analysis strategies

Fast Jet : Forcing events into 6 jets

$\Downarrow$

Lcfi Vertexing

$\Downarrow$ $\Downarrow$

Preselection

$\Downarrow$

MV selection

Static cuts analysis

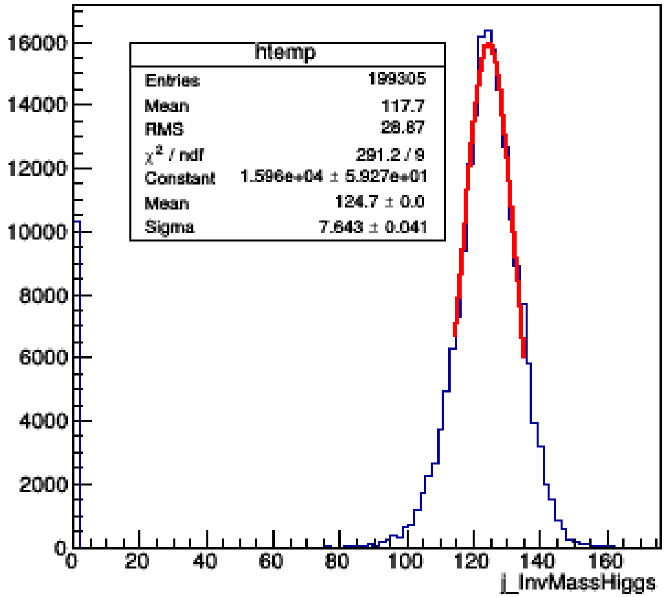
$\Downarrow$ $\Downarrow$

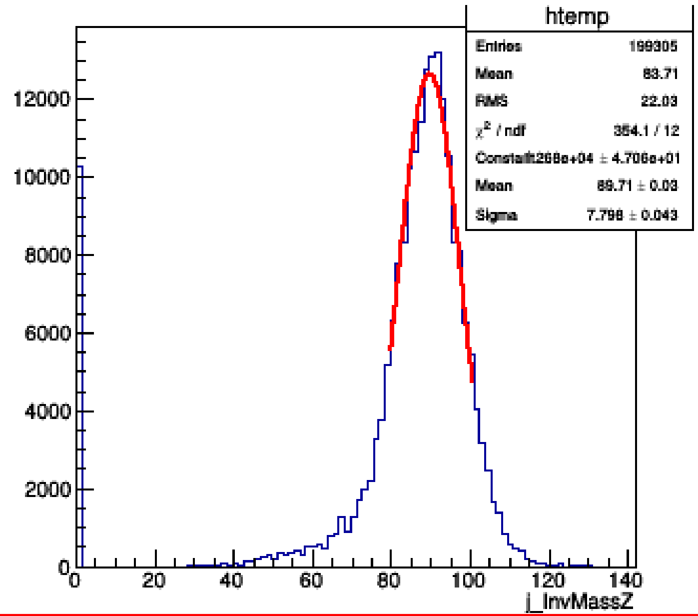
Relative statistical precision

$$\frac{\Delta\sigma}{\sigma} = \frac{\sqrt{S+B}}{S}$$

Signal reconstruction

- k_T **exclusive, particle flow with Arbor v3.1**
- **Jet formation: force events into 6 jets, do the jet pairing to form H (W and W*), Z**
  - Fit in boson the peak vicinity (±10 GeV, ±5 GeV,) for the Higgs and the Z boson for several jet openings **R=0.8, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5**
  - The best result is obtained for R=1.5 (maximal allowed for the Fast Jet 2.4.2 max ~ as used in ilcsoft_150612)
  - m_{Higgs} and m_Z show slight underestimation ⇒ R of the jets slightly smaller (discussed later)





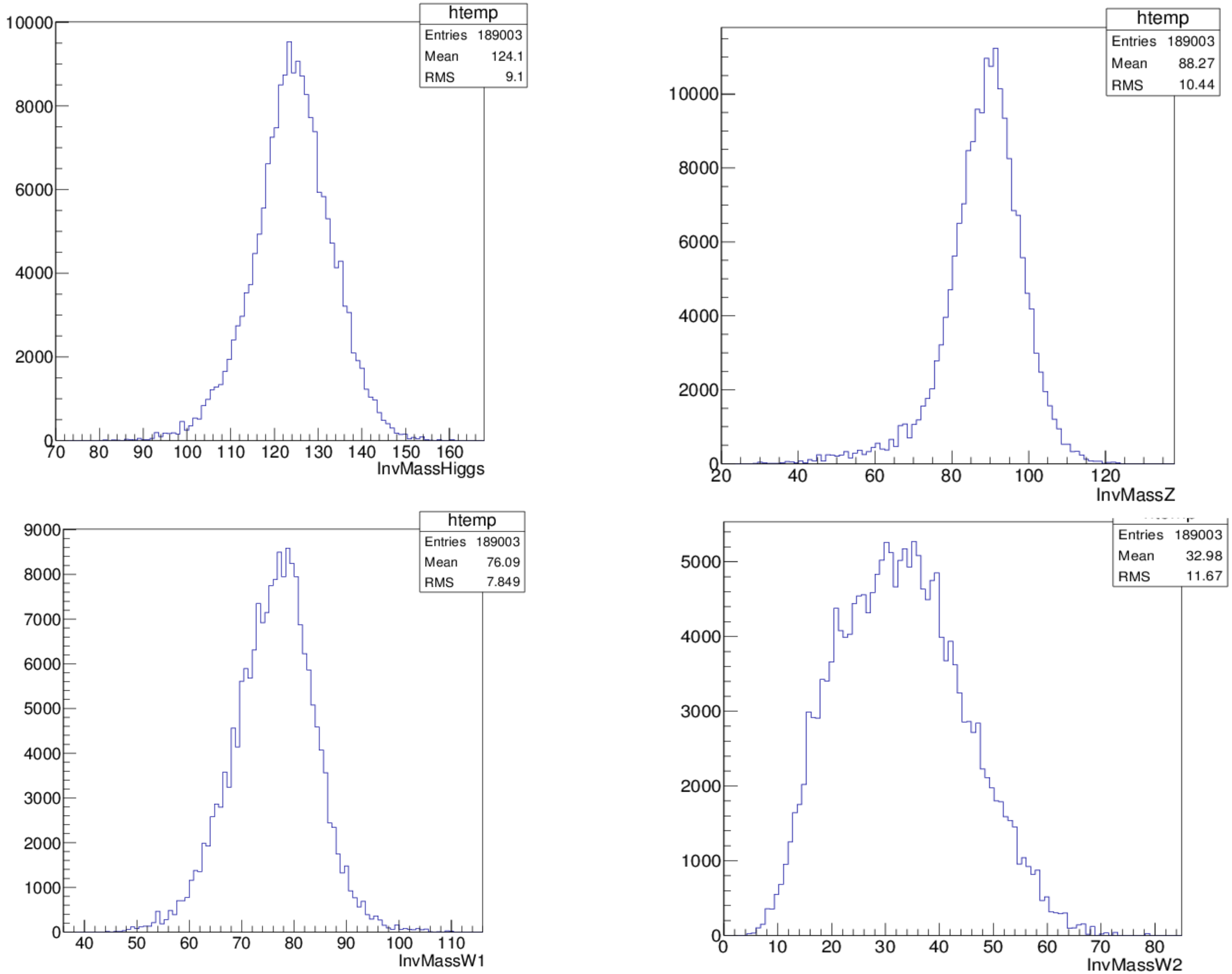
Reconstruction of the Higgs, Z and W bosons

- In order to reconstruct the Higgs, Z, W boson reconstruction the event is forced into six jets
- Obtained jets are grouped into three pairs to form the W, W* and Z bosons
- From WW* pair - the Higgs boson
- The combination which minimizes the χ^2 is chosen :

$$\chi^2 = \frac{(m_{ij} - m_W)^2}{\sigma^2_W} + \frac{(m_{kl} - m_Z)^2}{\sigma^2_Z} + \frac{(m_{ijmn} - m_H)^2}{\sigma^2_H}$$

- For the corresponding σ are the WA width was taken $\sigma^2_{H,W,Z}$

Reconstructed boson invariant masses for signal



Signal and background samples

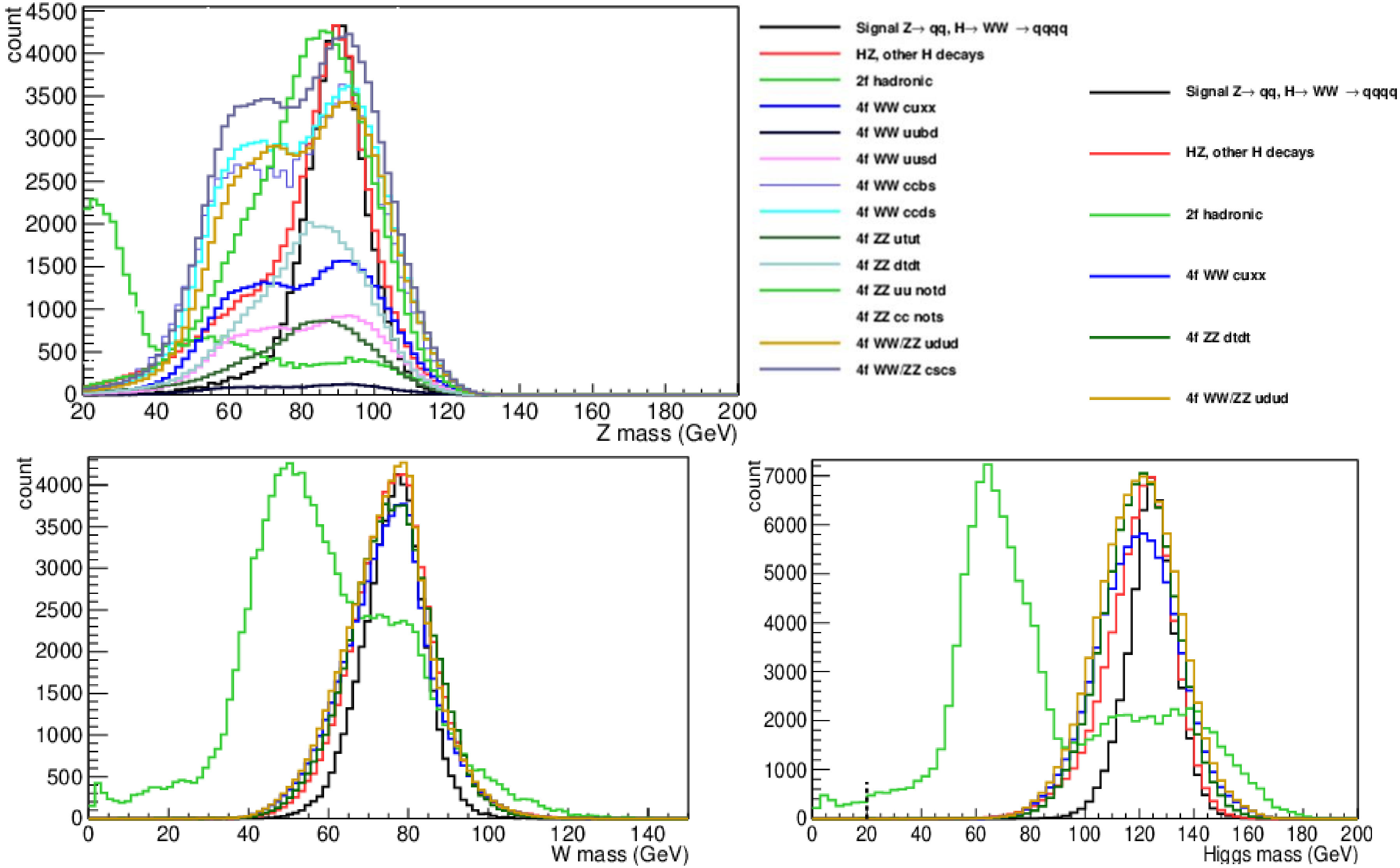
sample	$\sigma[fb]$	$\#evts/5ab^{-1}$	$\#evts\ used$
$qqh \to q\bar{q}WW^{*} \to q\bar{q}q\bar{q}q\bar{q}$	16,12	80600	74342
$other\ Higgs\ decays$ $non\ qqh \to q\bar{q}WW^{*} \to q\bar{q}q\bar{q}q\bar{q}$	127,27	636350	644354
$2f$	49561,30	247806500	100000
$4f\_ww\_cuxx$	3395,48	16977400	1220200
$4f\_ww\_ccbs$	5,74	28700	99400
$4f\_ww\_ccds$	165,57	827850	1343474
$4f\_ww\_uubd$	0.05	250	99800
$4f\_ww\_uusd$	165,94	829700	691057
$4f\_Mix\_udud$	1570,40	7852000	2782962
$4f\_Mix\_cscs$	1568,94	7844700	2375076
$4f\_zz\_utut$	83,09	415450	400000
$4f\_zz\_dtdt$	226,20	1131000	332600
$4f\_zz\_uu\_notd$	95,65	478250	477400
$4f\_zz\_cc\_nots$	96,04	480200	337400

Many thanks to Bingyang and Gang !!! for the production of 4f hadronic large x-sec bck at the end of last year – enabled analysis continuation in 2018

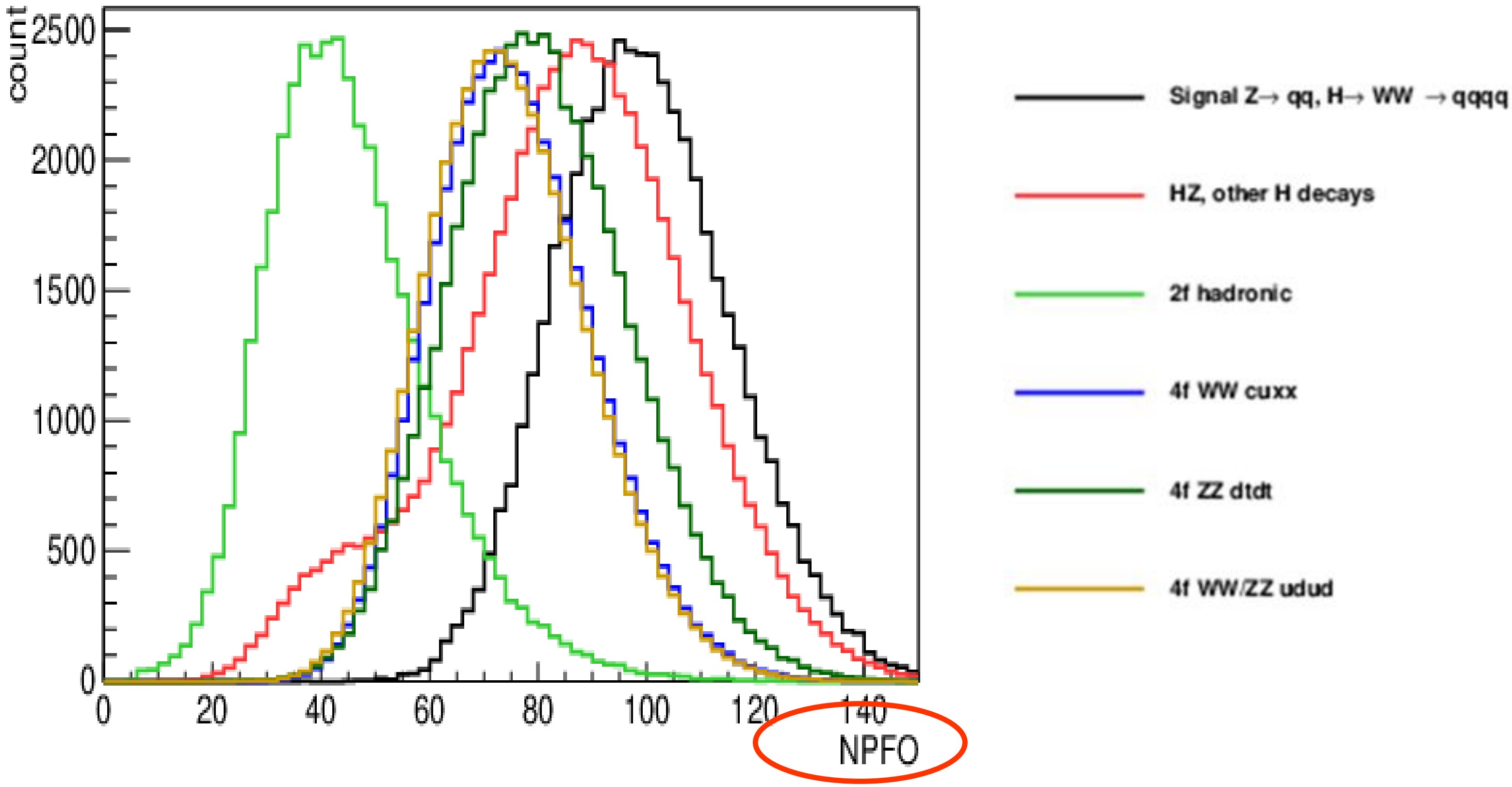
Investigated variables

- Invariant masses: m_{Higgs} m_Z m_W m_{w*}
- Number of particle flow objects NPFO
- Visible energy E_{vis}
- The highest transverse momentum of the jet in the event –highestPtJet
- Transverse momentum of the Higgs boson PtOfHiggsJets
- Event shape variables: thrust, oblatness, sphericity, aplanarity
- Jet transitions: y_{12} y_{23} y_{34} y_{45} y_{56} y_{67}
- Force event into 2 jet: btag1, btag2, btag1*btag2
- ctag1, ctag2
- Force event into 6 jet: $btag_{i,}$ $ctag_i$
- Angle between jets that comprise W boson: ThetaWqq,
- Angle between jets that comprise Z boson: ThetaZqq
- Angle between W and W* that comprise the Higgs boson : ThetaHiggsW1W2
- Arithmetic variable Energy*Theta of the W, Higgs and Z boson

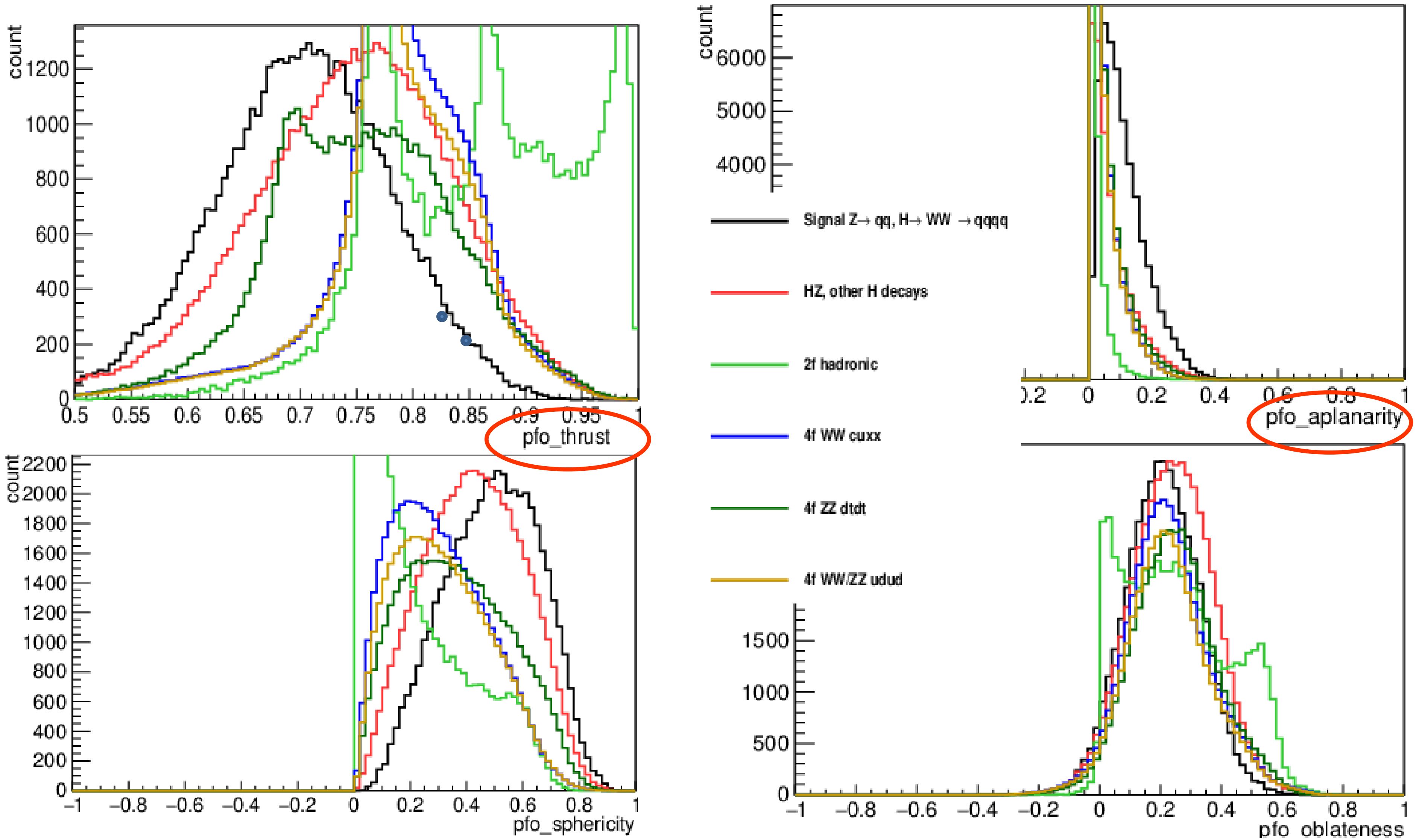
Invariant masses



Number of particle flow objects

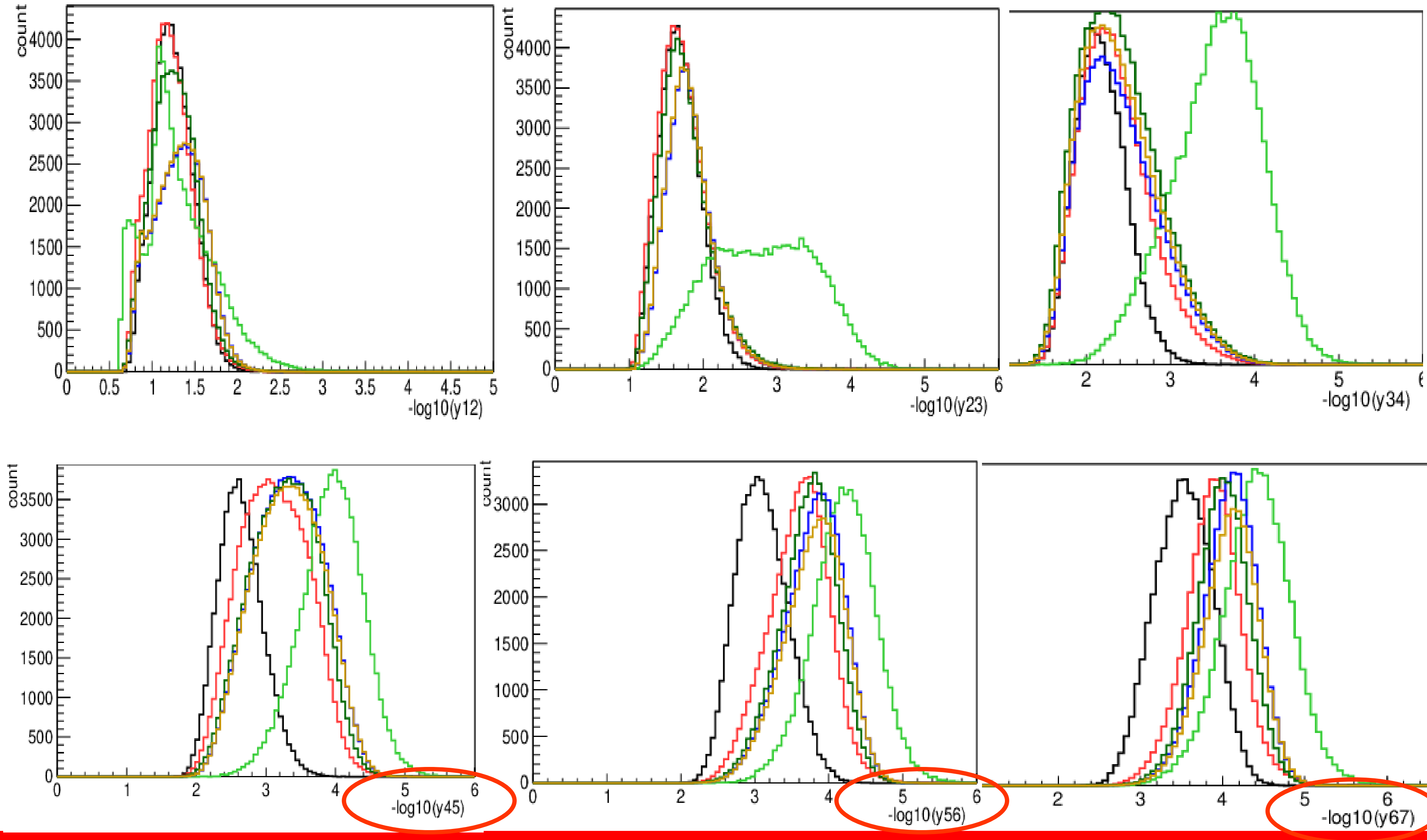


The event shape variables



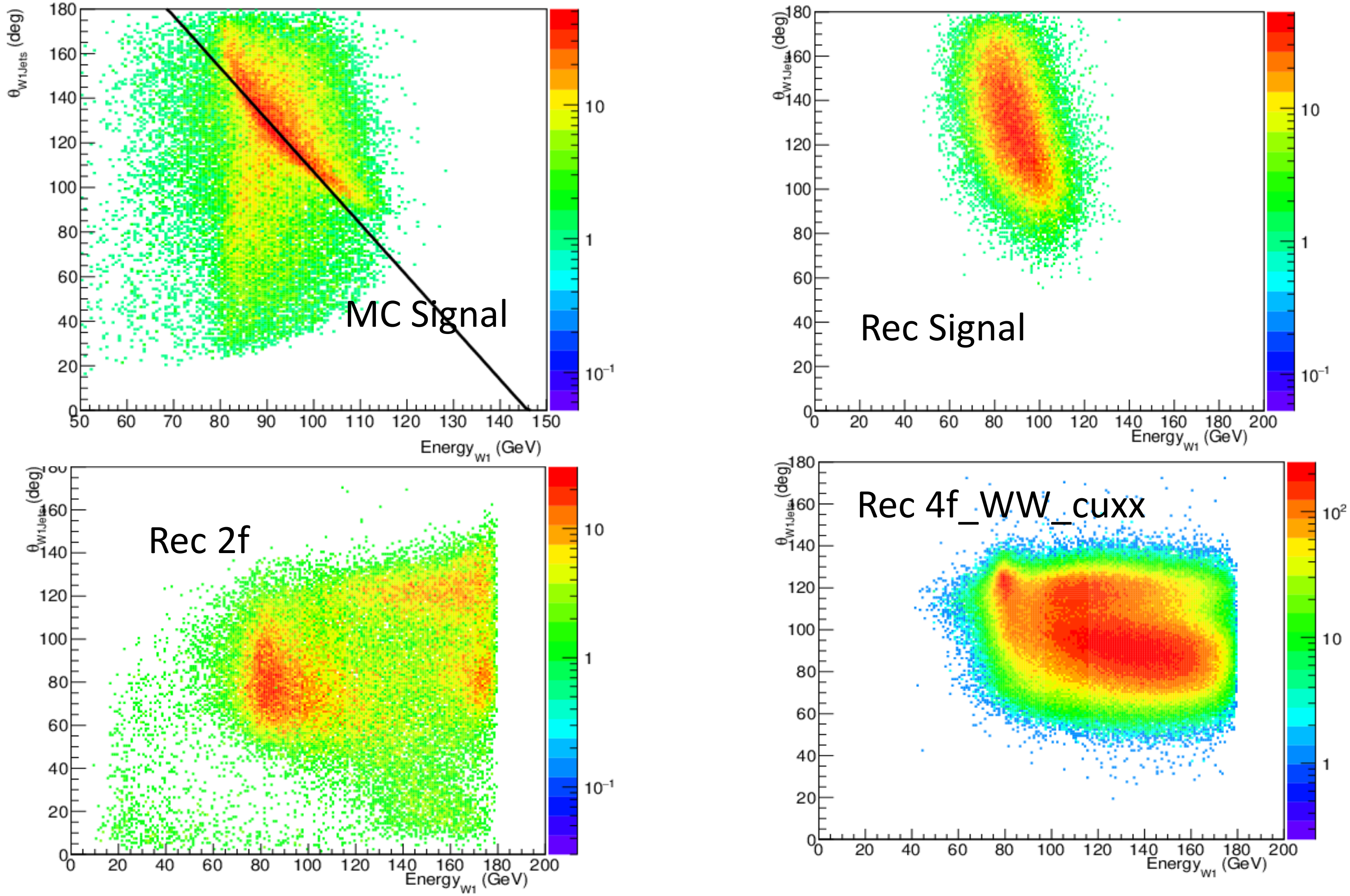
Jet transitions

- The k_t values at which the number of jet goes from $i \rightarrow i + 1$ number of jets



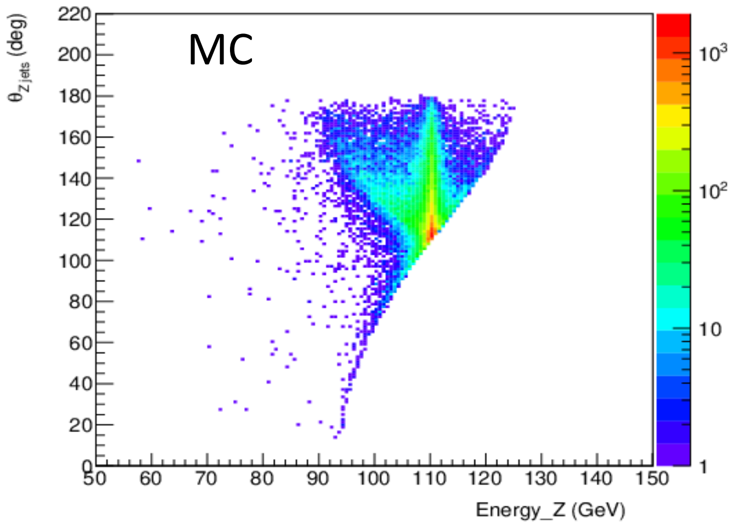
New variable construction based on signal Monte Carlo information

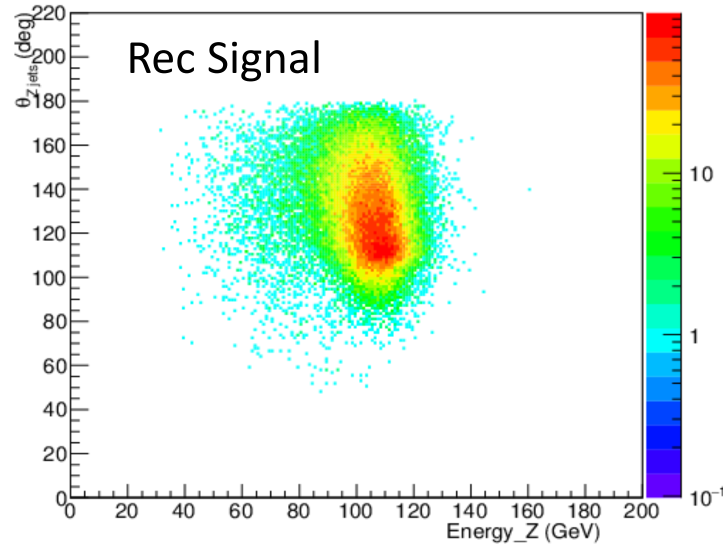
The distribution of the energy of the W real boson versus the angle between jets that comprise it

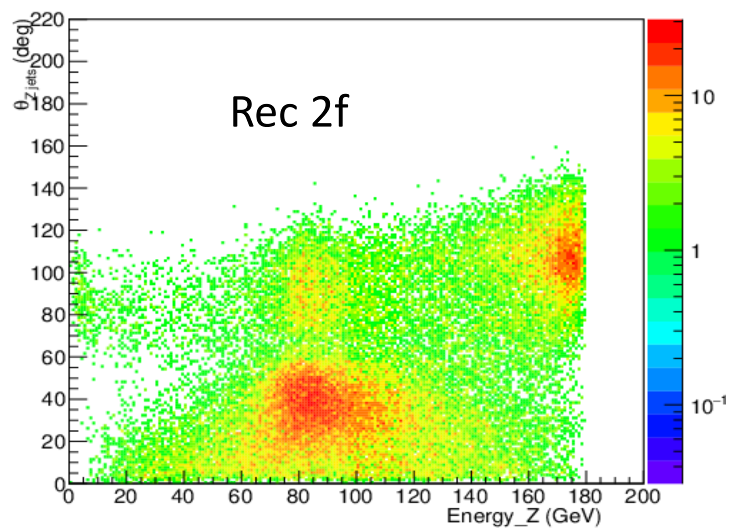


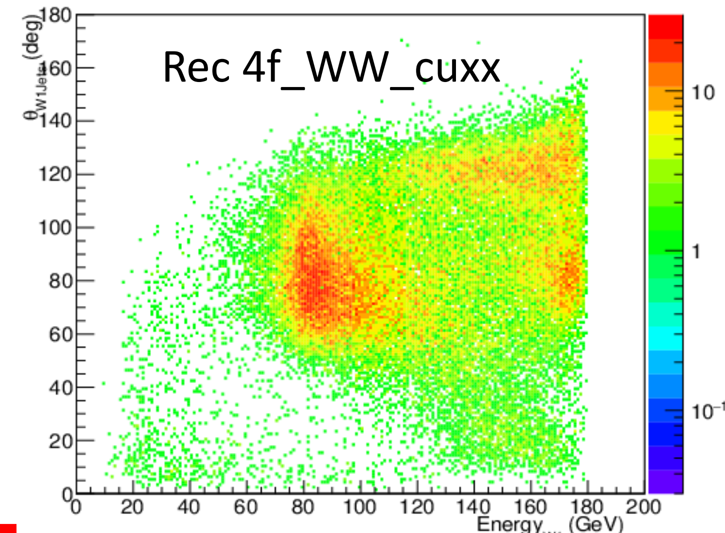
New variable construction energy theta of the Z boson

The distribution of the energy of the W real boson versus the angle between jets that comprise it

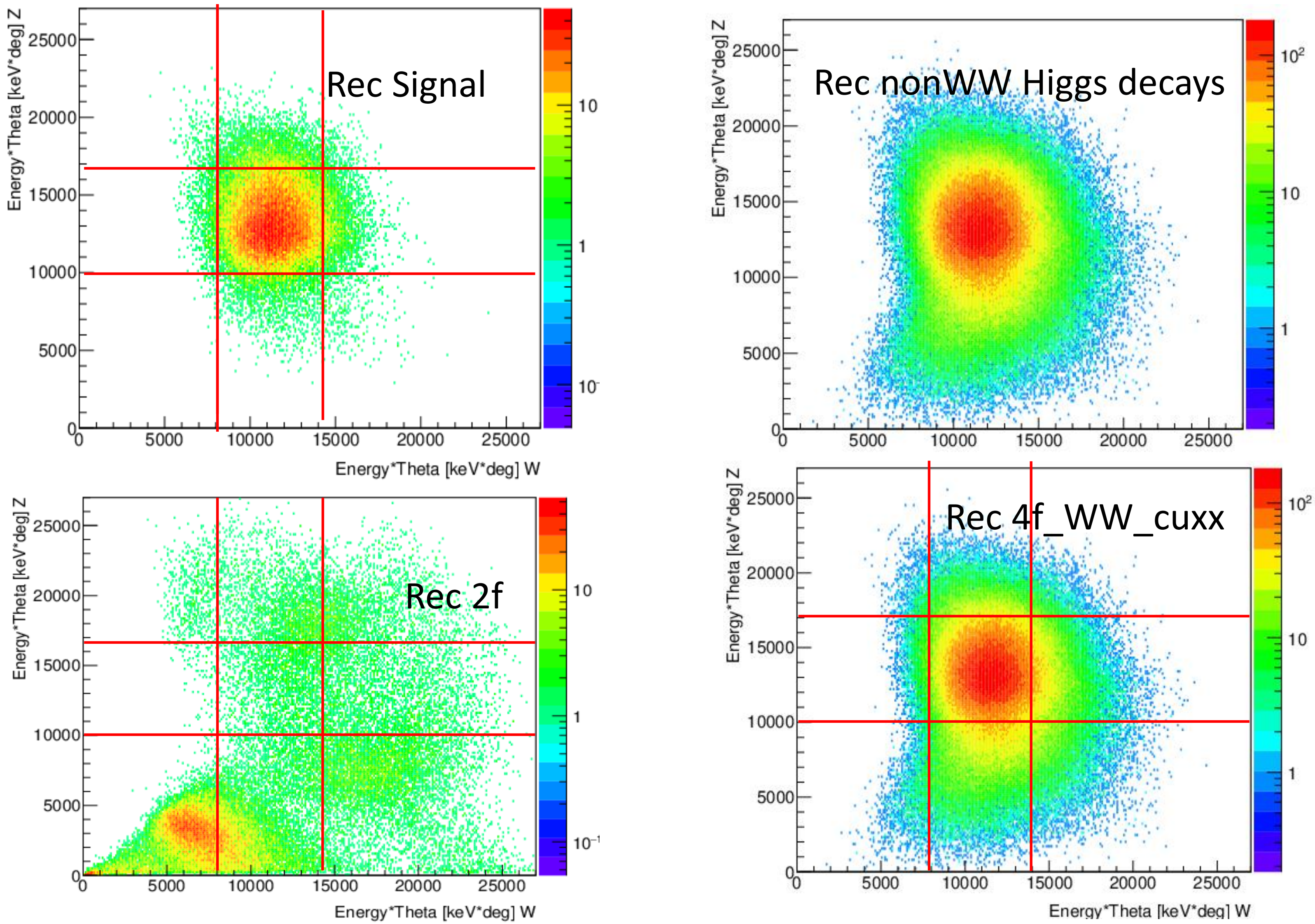








Arithmetic Variables Energy*Theta for W boson



Multivariate approach : preselection

- 8000< EnergyThetaW <14000. 10000<EnergyThetaZ <17000. NPFO>80.

sample	$\sigma[fb]$	$\#evts/5ab^{-1}$	$\varepsilon_{pres}[\%]$	$evts\ after\ preselection$
$qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$	16,12	80600	70.0	56380
$other\ Higgs\ decays$ $non\ qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$	127,27	636350	43.0	273975
$2f$	49561,30	247806500	0.8	1990414
$4f\_ww\_cuxx$	3395,48	16977400	16.7	2838452
$4f\_ww\_ccbs$	5,74	28700	22.5	6453
$4f\_ww\_ccds$	165,57	827850	18.3	151787
$4f\_ww\_uubd$	0.05	250	19.8	50
$4f\_ww\_uusd$	165,94	829700	15.3	127241
$4f\_ww\_zz\_udud$	1570,40	7852000	16.0	1255551
$4f\_ww\_zz\_cscs$	1568,94	7844700	17.9	1406147
$4f\_zz\_utut$	83,09	415450	22.0	91366
$4f\_zz\_dtdt$	226,20	1131000	27.5	311025
$4f\_zz\_uu\_notd$	95,65	478250	23.7	113345
$4f\_zz\_cc\_nots$	96,04	480200	27.8	133496

Multivariate analysis

- The training of BDTG was performed on ten background samples excluding:
  - 2f backgrounds cut down at the preselection level 0.8% - preserved sensitive observables low training statistics (50000 training evts after preselection 0.8%=400evts left)
  - 4f_WW_ccbs and 4f_WW_uubd low cross-section lower training statistics

- The variables set was optimized to a set with the minimal stable relative statistical error (41 variables investigated – 18 final variables)
  - Invariant masses: m_{Higgs} m_Z m_W
  - Number of particle flow objects NPFO
  - Highest PtJet, transverse momentum of jets that comprise Higgs boson - PtOfHiggsJets
  - Event shape variables: thrust, oblatness, aplanarity
  - Jet transitions: y_{12} y_{34} y_{45} y_{56} y_{67}
  - Force event into 2 jet: btag1, btag2
  - ctag1
  - Arithmetic variable Energy*Theta of the Z boson

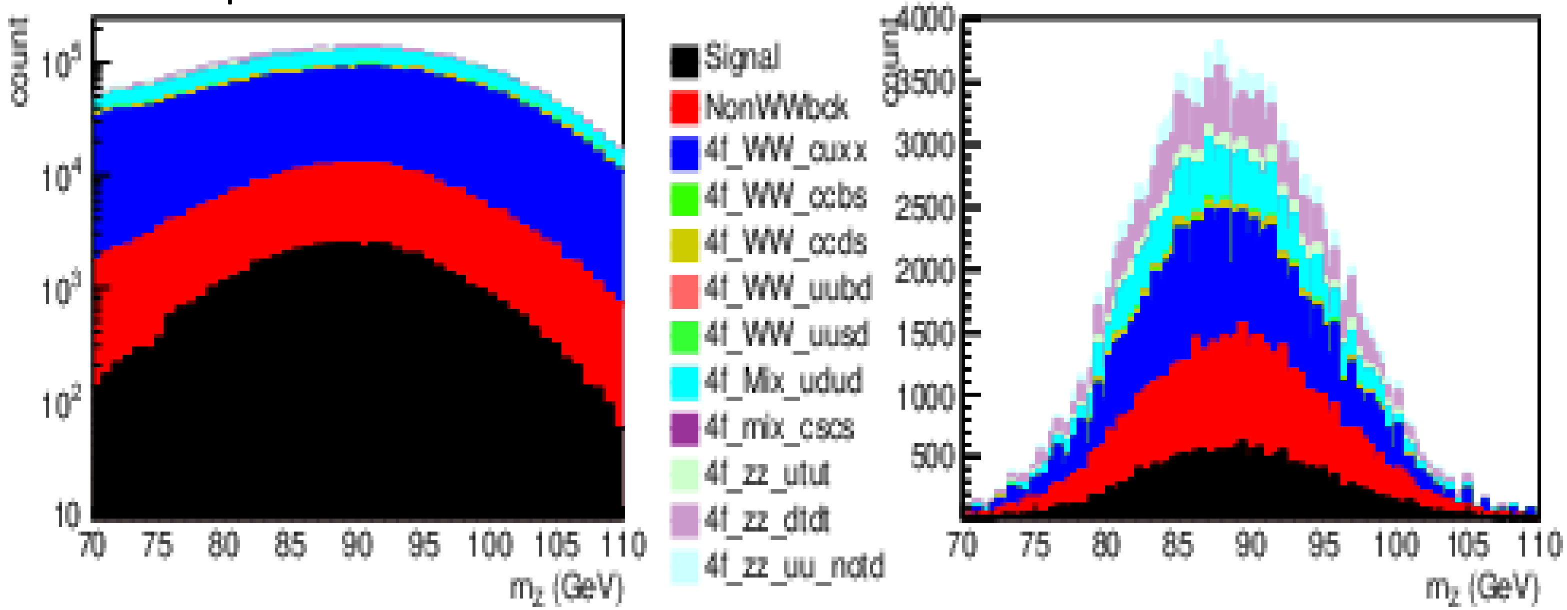
Final selection

- After preselection and multivariate analysis ~99% of the background is reduced

sample	$\sigma[fb]$	$\#evts$ $/5ab^{-1}$	$evts\ after$ $preselection$	ε_{tmva} [%]	ε_{total} [%]	$evts\ after$ $final\ selection$
$qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$	16,12	80600	56380	41.3	28.85	23257
$other\ Higgs\ decays$ $non\ qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$	127,27	636350	273975	14.1	6.1	38629
$2f$	49561,30	247806500	1990414	0.25	0.002	4976
$4f\_ww\_cuxx$	3395,48	16977400	2838452	1.45	0.24	41188
$4f\_ww\_ccbs$	5,74	28700	6453	1.7	0.38	110
$4f\_ww\_ccds$	165,57	827850	151787	1.5	0.28	2294
$4f\_ww\_uubd$	0.05	250	50	2.0	0.4	1
$4f\_ww\_uusd$	165,94	829700	127241	0.8	0.13	1073
$4f\_ww\_zz\_udud$	1570,40	7852000	1255551	1.5	0.24	19102
$4f\_ww\_zz\_cscs$	1568,94	7844700	1406147	1.6	0.29	22514
$4f\_zz\_utut$	83,09	415450	91366	5.5	1.2	4997
$4f\_zz\_dtdt$	226,20	1131000	311025	6.4	1.8	19845
$4f\_zz\_uu\_notd$	95,65	478250	113345	5.9	1.4	6675
$4f\_zz\_cc\_nots$	96,04	480200	133496	6.0	1.7	7949

The relative statistical uncertainty: MVA method

After preselection | After final selection



- The dominant background after final selection are ee →qqqq backgrounds
- The high cross-section $2f \rightarrow q\bar{q}$ background show good response to the preselection and multivariate analysis. The obtained relative statistical precision is 1.9 % with the corresponding signal efficiency of 29%

$$\frac{\Delta\sigma}{\sigma} = \frac{\sqrt{S+B}}{S} \approx 1.9\%$$

Static cut analysis

- The statatic cut variables used:
  - Invariant masses: 80 <m_Z < 100 GeV
  - Invariant masses: 115 <m_H < 135 GeV
  - Number of particle flow objects NPFO > 90 GeV
  - Highest PtJet <90
  - transverse momentum of jets that comprise Higgs boson < 80 GeV
  - Jet transitions: y_{23} <2.4
  - y_{34} <2.4
  - Y_{45} <2.7
  - y_{56} <3.2
  - y_{67} <3.5
  - Arithmetic variable Energy*Theta of the Z boson 8000< EnThW<14000
  - Arithmetic variable Energy*Theta of the Z boson 10000< EnThZ<17000

Static cuts analysis results

sample	$\sigma[fb]$	$\#evts/5ab^{-1}$	$\varepsilon_{tot\ mva}$ [%]	ε_{static} [%]	$evts\ after\ final\ selection$
$qqh \to q\bar{q}WW^* \to q\bar{q}q\bar{q}q\bar{q}$	16,12	80600	28.85	28.9	23293
$other\ Higgs\ decays$ $non\ qqh \to q\bar{q}WW^* \to q\bar{q}q\bar{q}q\bar{q}$	127,27	636350	6.1	8.1	51544
$2f$	49561,30	247806500	0.002	0.02	49561
$4f\_ww\_cuxx$	3395,48	16977400	0.24	1.5	254661
$4f\_ww\_ccbs$	5,74	28700	0.38	1.9	545
$4f\_ww\_ccds$	165,57	827850	0.28	1.6	13246
$4f\_ww\_uubd$	0.05	250	0.4	1.8	5
$4f\_ww\_uusd$	165,94	829700	0.13	1.3	10786
$4f\_ww\_zz\_udud$	1570,40	7852000	0.24	1.4	109928
$4f\_ww\_zz\_cscs$	1568,94	7844700	0.29	1.6	125515
$4f_zz_$utut	83,09	415450	1.2	2.4	9971
$4f\_zz\_dtdt$	226,20	1131000	1.8	2.9	32799
$4f\_zz\_uu\_notd$	95,65	478250	1.4	2.5	11956.
$4f\_zz\_cc\_nots$	96,04	480200	1.7	2.9	13926

- After the static cut analysis ~98% of the background is reduced.
- The obtained relative statistical uncertainty 3.6 % with the corresponding signal efficiency of 29%

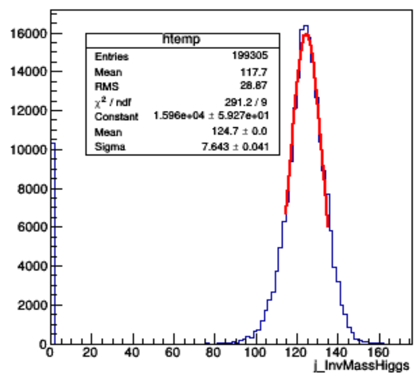
Summary

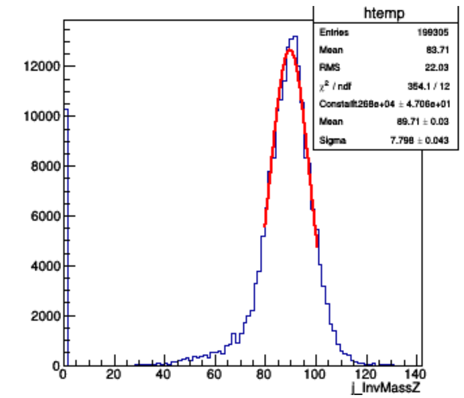
- The fully hadronic decay is most abundant channel in the H→WW* decay
- In Higgsstrahlung, Z →qq, this decay leads complex central six jet final state
- High cross section hadronic backgrounds
- The channel is analysed with two types of analysis flow:
  - multivariate analysis
  - static cut analysis
- The multivariate approach showed better reduction capabilities in comparison to the static cut analysis
- This is due to lack of distinct cut variables for hadronic final state
- The obtained relative statistical precission with the static cut analysis is 3.6% with the signal efficiency of 29 %, while the result obtained with the multivariate analysis is 1.9% with the corresponding signal efficiency of 29% also, for the integrated luminosity of 5 ab^{-1}

- Many thanks to Manqi Ruan for the invitation to do this analysis at CepC!

Fast jet

- Fast Jet 2.4.2. intrinsically limits the jet opening (FastJet 2.4.2/JetDefinition.cc) which results in the slight underestimation of the reconstructed invariant masses of the Z and the Higgs boson





- Overload Fast Jet 2.4.2 with Fast Jet 3.1.2 where R_{max} is not limited – moderate increases in R – (not to interfere with beam jets)
- The mean of the signal with the slightly higher R would reproduce the Higgs and Z mass slightly better – boundary - beam jets
- The possible difference in the distributions would not make any visible or significant impact on the analysis result
- The suggestion is of principal nature

end

Preselection on the Monte Carlo level

- Leptonic, semileptonic final states – make use of final state leptons (e, μ), which are products of boson decays: Higgs, Z

- These leptons are stable before reaching the detector (exc. ISR,FSR) $\Rightarrow$ provide effective cutting variables at the generator level $m_H, m_Z, \ldots$
- Opposite to leptonic and semileptonic final states, fully hadronic final states does not exhibit these property and are characterized by lack of so effective cutting variables on the MC level
- For hadronic final states at generator level (WHIZARD) we can employ the cut on invariant mass fo diquark – $m_{qq} >$ at least 50 GeV (probably could go much higher) in the case of 2f would cut significant generator production and reconstruction time and storage space, but all analyses that use this particular background should be considered
- Also, if necessary– postgenerator moderately effective cuts can be applied on the variables which can be constructed from boson decay product quarks taken before the fragmentation and hadronization
- Todays update from Bingyang -2f reconstructede ~2.000.000! – therefore not necessary

MC variables that can be used in preselection

- Variables that can be used to filter the generated files that are going to be used for the reconstruction

1. The event shape variables can be constructed using the Monte Carlo information:
    - one can use the final state MC particles to construct thrust, oblateness, sphericity, aplanarity

2. Number of final state particles – efficiently can reduce 2f hadronic bck

Table 19: The information of the four fermions background samples

Process	Final states	σ [fb]	ILC result [fb]	Events expected	Events generated
zz_h0utut	up, up, up, up	83.09	83.05	419604	419584
zz_h0dtdt	$down, down, down, down$	226.20	226.42	1142310	1142270
zz_h0uu_notd	$uq, uq, (sq, bq), (sq, bq)$	95.65	96.00	483032	483045
zz_h0cc_nots	$cq, cq, (dq, bq), (dq, bq)$	96.04	96.28	485002	485016
zz_sl0nu_up	$nu_{\mu,\tau}, nu_{\mu,\tau}, up, up$	81.72	81.86	412686	412709
zz_sl0nu_down	$nu_{\mu,\tau}, nu_{\mu,\tau}, down, down$	134.86	135.48	681043	681041
zz_sl0mu_up	mu, mu, up, up	82.38	83.10	416019	416008
zz_sl0mu_down	$mu, mu, down, down$	127.96	128.84	646198	646181
zz_sl0tau_up	tau, tau, up, up	39.78	40.02	200889	200882
zz_sl0tau_down	$tau, tau, down, down$	64.30	64.64	324715	324709
zz_l04tau	$\tau^-, \tau^+, \tau^-, \tau^+$	4.38	4.41	22119	100000
zz_l04mu	$\mu^-, \mu^+, \mu^-, \mu^+$	14.57	14.63	73578	100000
zz_l0taumu	$\tau^-, \tau^+, \mu^-, \mu^+$	17.54	17.73	88577	100000
zz_l0mumu	$\nu_\tau, \bar{\nu}_\tau, \mu^-, \mu^+$	18.17	18.34	91758	100000
zz_l0tautau	$\nu_\mu, \bar{\nu}_\mu, \tau^-, \tau^+$	9.20	9.27	46460	100000
ww_h0cuxx	$uq, cq, down, down$	3395.48	3413.36	17147189	17147188
ww_h0uubd	uq, uq, dq, bq	0.05	0.05	252	100000
ww_h0uusd	uq, uq, sq, bq	165.94	167.21	837997	838010
ww_h0ccbs	cq, cq, sq, bq	5.74	5.75	28987	100000
ww_h0ccds	cq, cq, sq, dq	165.57	166.30	836128	836128
ww_sl0muq	$mu, nu, up, down$	2358.69	2369.79	11911394	11911396
ww_sl0tauq	$tau, nu, up, down$	2351.98	2368.64	11877519	11877519
ww_l0ll	mu, tau, nu_μ, nu_τ	392.96	394.73	1984448	1984437
zzorww_h0udud	uq, uq, dq, dq	1570.40	1579.63	7930514	7930514
zzorww_h0cscs	cq, cq, sq, sq	1568.94	1572.41	7923141	7923140
zzorww_l0mumu	mu, mu, nu_μ, nu_μ	214.81	216.12	1084790	1084777
zzorww_l0tautau	$tau, tau, nu_\tau, nu_\tau$	205.84	206.38	1039492	1039510
sze_l0tau	e^-, e^+, τ^-, τ^+	150.14	150.30	758207	758206
sze_l0mu	e^-, e^+, μ^-, μ^+	852.18	850.70	4303527	4303528
sze_l0nunu	$e^-, e^+, \nu_{\mu,\tau}, \bar{\nu}_{\mu,\tau}$	29.62	29.41	149581	149583
sze_sl0uu	e, e, up, up	195.86	195.37	989093	989109
sze_sl0dd	$e, e, down, down$	128.72	128.20	650036	649940
sznu_l0mumu	$\nu_e, \bar{\nu}_e, \mu^-, \mu^+$	43.33	43.41	218816	218824
sznu_l0tautau	$\nu_e, \bar{\nu}_e, \tau^-, \tau^+$	14.57	14.61	73578	100000
sznu_sl0nu_up	$\nu_e, \bar{\nu}_e, up, up$	56.09	55.95	283254	283254
sznu_sl0nu_down	$\nu_e, \bar{\nu}_e, down, down$	91.28	90.95	460964	460961
sw_l0mu	$e, nu_e, mu, nu_{\mu,\tau}$	429.20	430.64	2167446	2167447
sw_l0tau	$e, nu_e, tau, nu_{\mu,\tau}$	429.42	430.27	2168556	2168556
sw_sl0qq	$e, nu_e, up, down$	2579.31	2581.03	13025535	13025535
szeorsw_l0l	$e^-, e^+, \nu_e, \bar{\nu}_e$	249.34	249.74	1259167	1259165

179