ATLAS EXPERIMENT

NATIONAL RESEARCH NUCLEAR UNIVERSITY NRNU MEPhI MEPhI

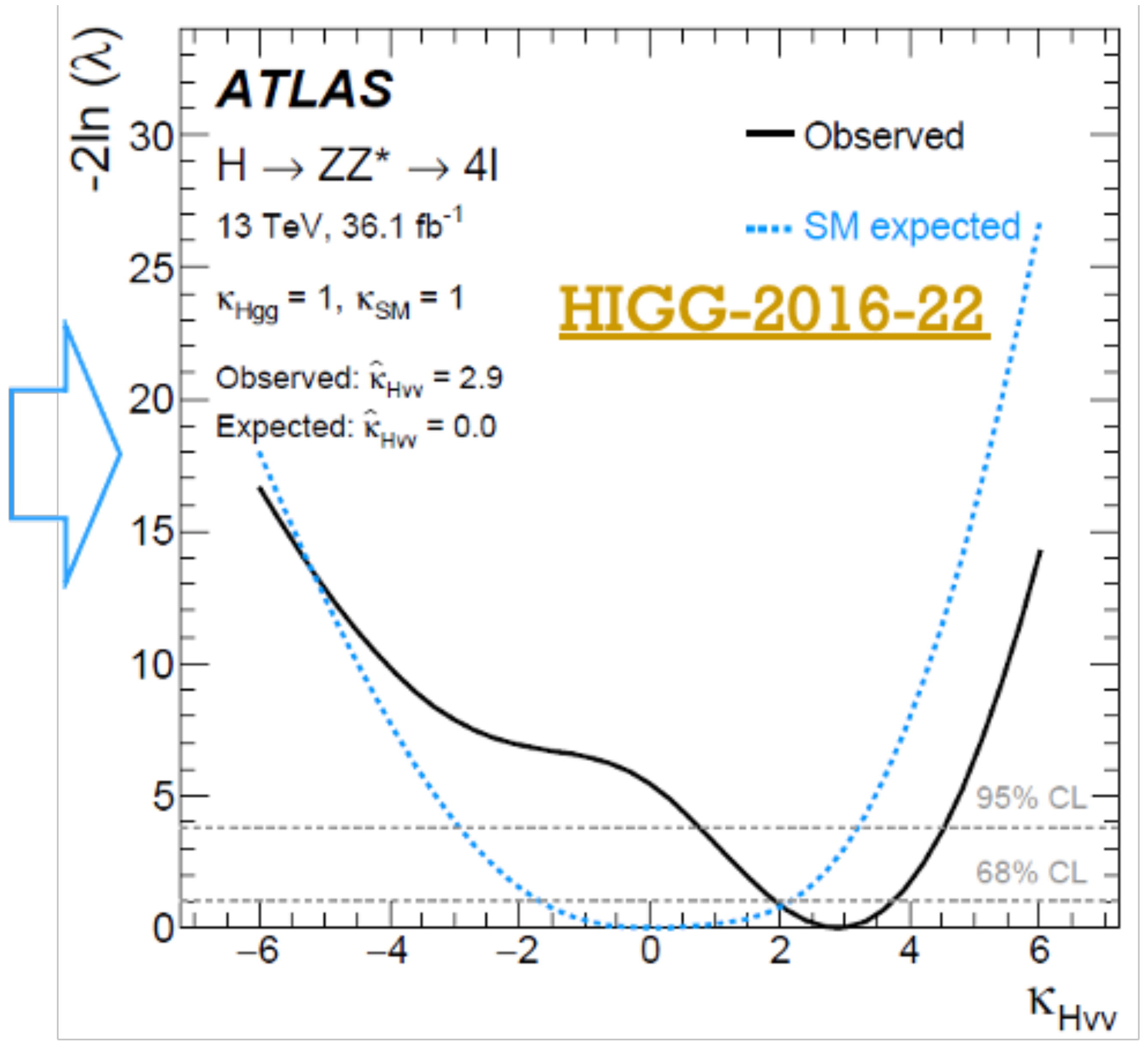
EFT analysis & categorization optimization status

Serafima Nechaeva

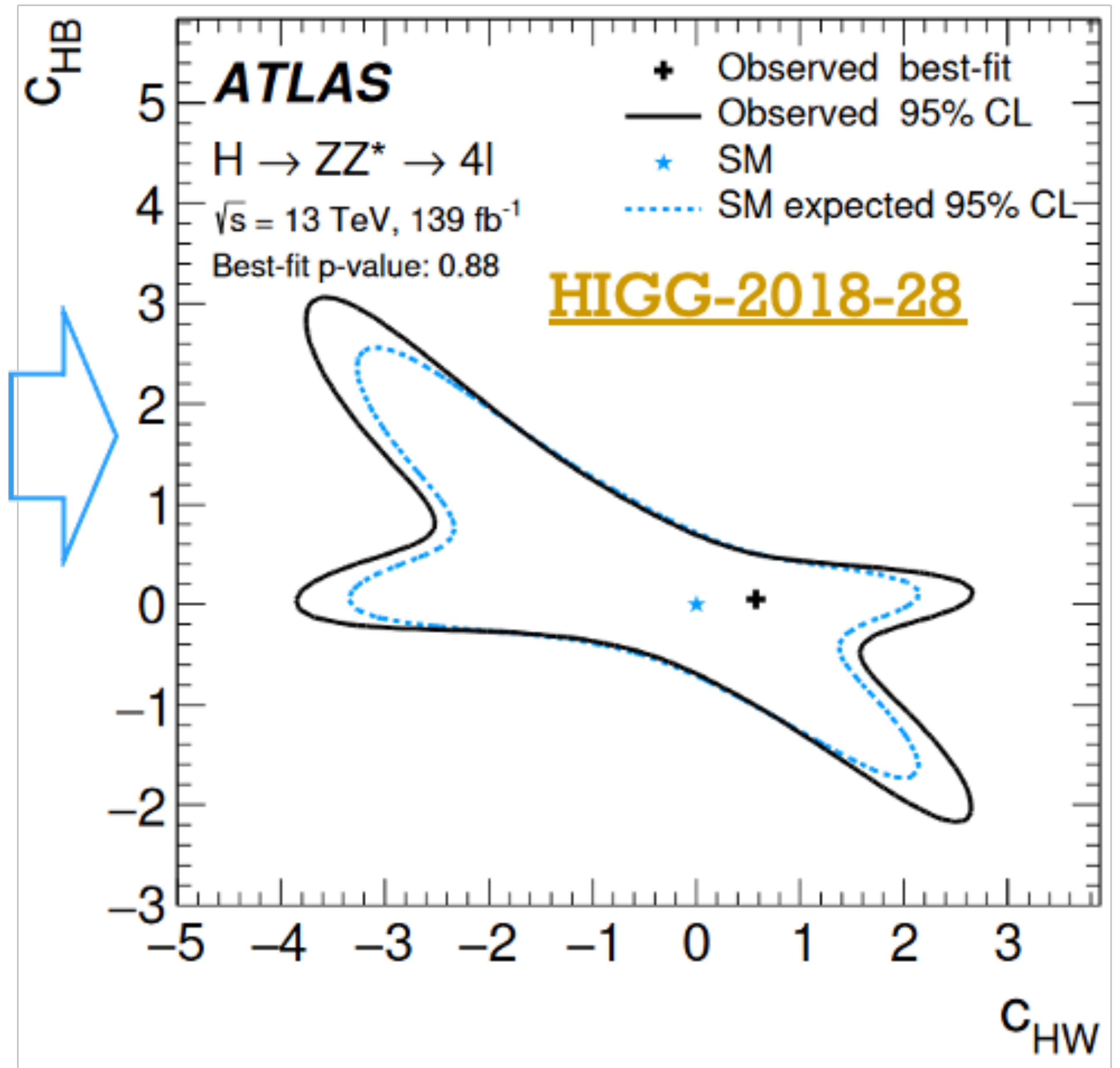
09.03.2022

Previous EFT measurements

- Rate-only (event yield counting) 36.1 fb^{-1}analysis.
- Signal modeling: BSM contributions from VBF, VH and ggF processes were obtained by morphing.
- The results were interpreted in the Higgs Characterization framework.



- Rate-only (event yield counting) 139 fb^{-1}analysis.
- Dedicated STXS measurements.
- The results were interpreted in the SMEFT framework.



Motivation for the EFT analysis

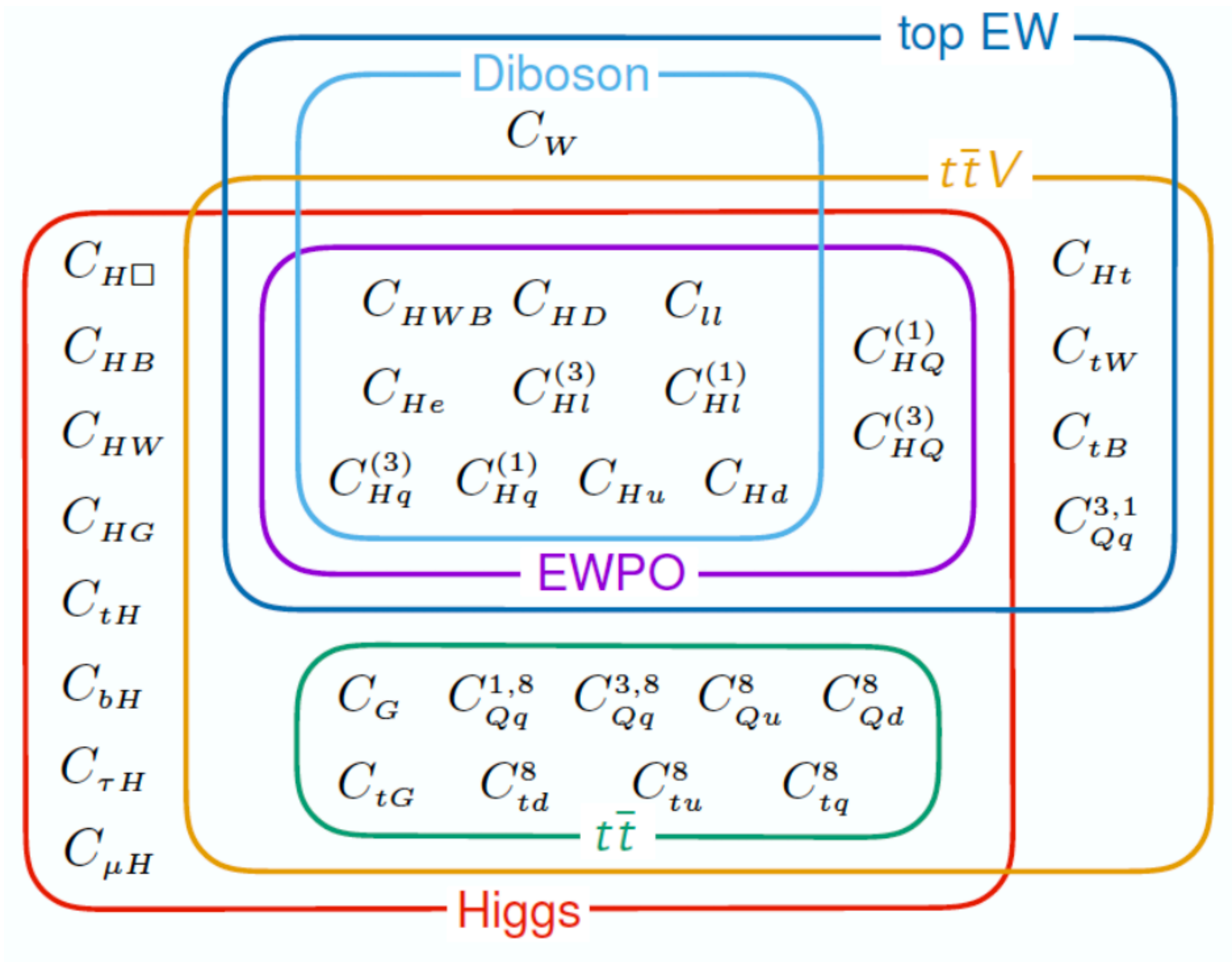
The original motivation for the EFT analysis was to increase the measurements sensitivity compared to the STXS analysis by using the following features:

- Shape+rate measurements instead of rate only.
- Dedicated BSM-sensitive kinematic observables.
- Optimized categorization.
- Natural accounting of the acceptance.

EFT framework

The BSM effects are parameterized by the dimension-6 SMEFT model

$$\mathcal{L}_{\text{SMEFT}}^{\text{Warsaw}} \supset \frac{C_{eH}}{\Lambda^2} y_e (H^\dagger H)(\bar{l}eH) + \frac{C_{dH}}{\Lambda^2} y_d (H^\dagger H)(\bar{q}dH) + \frac{C_{uH}}{\Lambda^2} y_u (H^\dagger H)(\bar{q}u\widetilde{H})$$
$$+ \frac{C_G}{\Lambda^2} f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu} + \frac{C_{H\Box}}{\Lambda^2}(H^\dagger H)\Box(H^\dagger H) + \frac{C_{uG}}{\Lambda^2} y_u (\bar{q}\sigma^{\mu\nu} T^A u)\widetilde{H}\, G_{\mu\nu}^A$$
$$+ \frac{C_{HW}}{\Lambda^2} H^\dagger H\, W_{\mu\nu}^I W^{I\mu\nu} + \frac{C_{HB}}{\Lambda^2} H^\dagger H\, B_{\mu\nu} B^{\mu\nu} + \frac{C_{HG}}{\Lambda^2} H^\dagger H\, G_{\mu\nu}^A G^{A\mu\nu}\,.$$



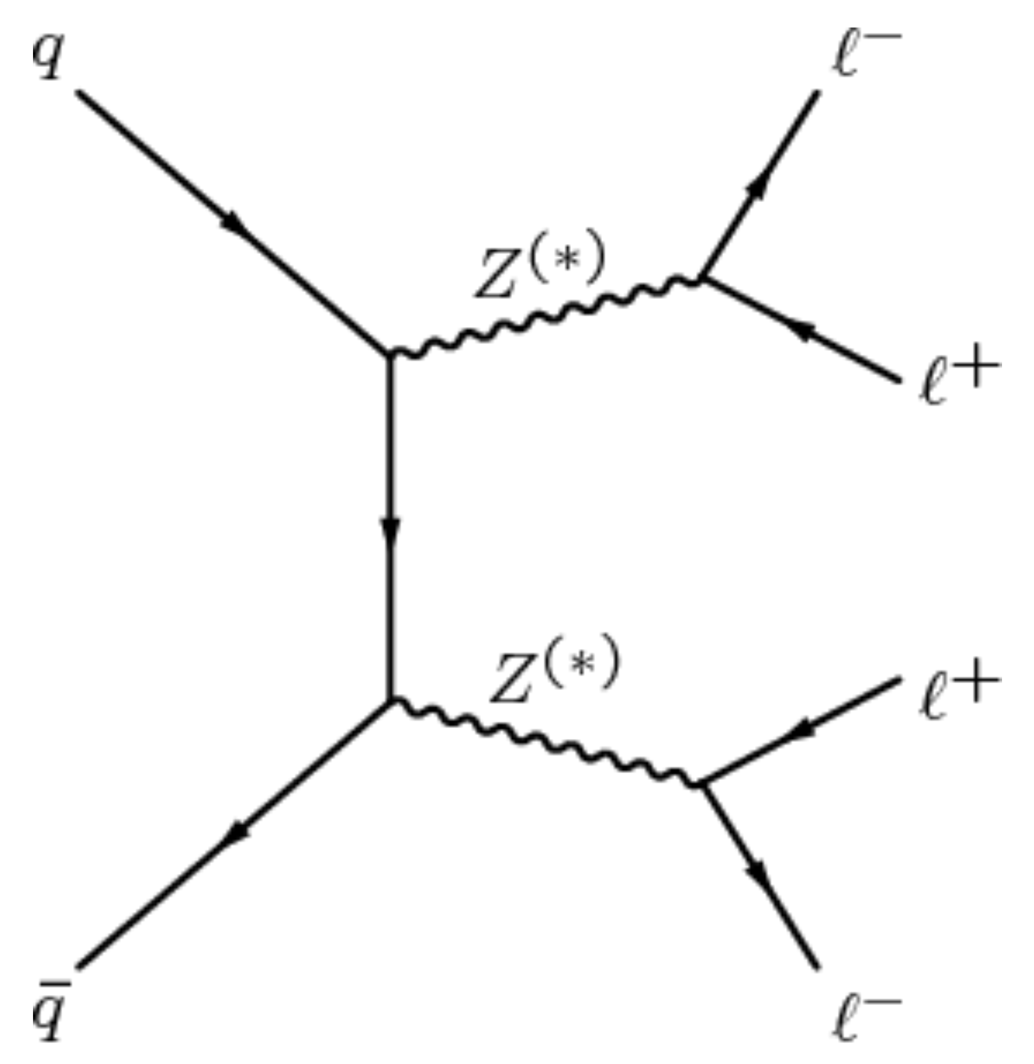
[Ellis et al. 2012.02779]

- The MadGraph5 generator is used to simulate the tree-level processes.
- Focusing on the CP-even C_{HW} and C_{HB} EFT Higgs couplings.

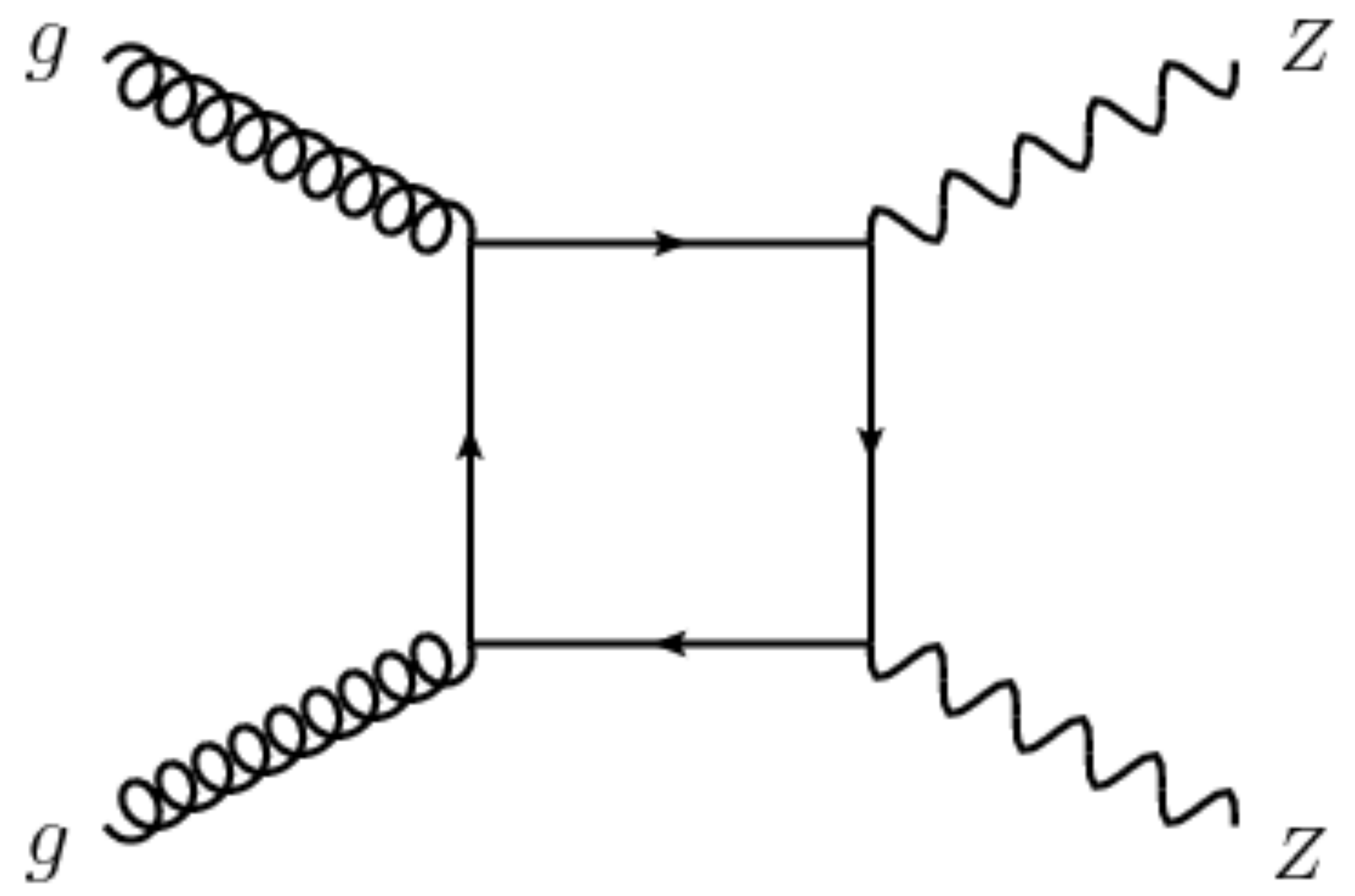
Background processes

- The following background processes are now considered: qqZZ, ggZZ, bbH, WH, ZH, tHjb, ttH, tWH and ttV.
- For now, the backgrounds are represented by the MC samples.
- The contribution from the backgrounds are calculated for each category and added to the Asimov data.

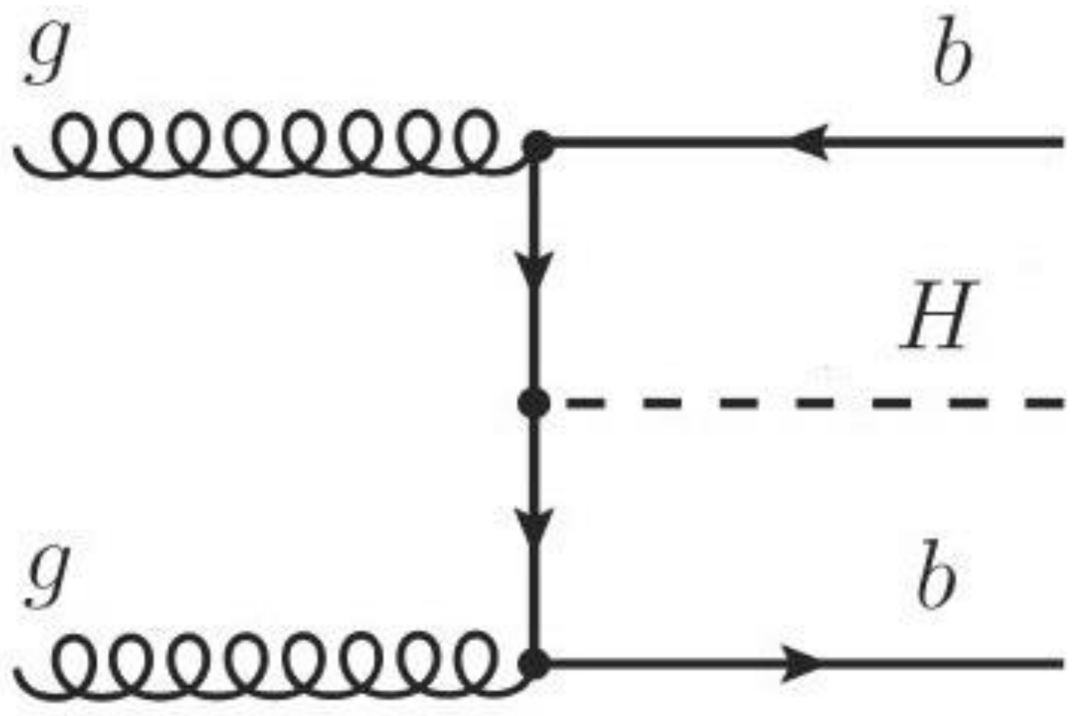
qqZZ



ggZZ



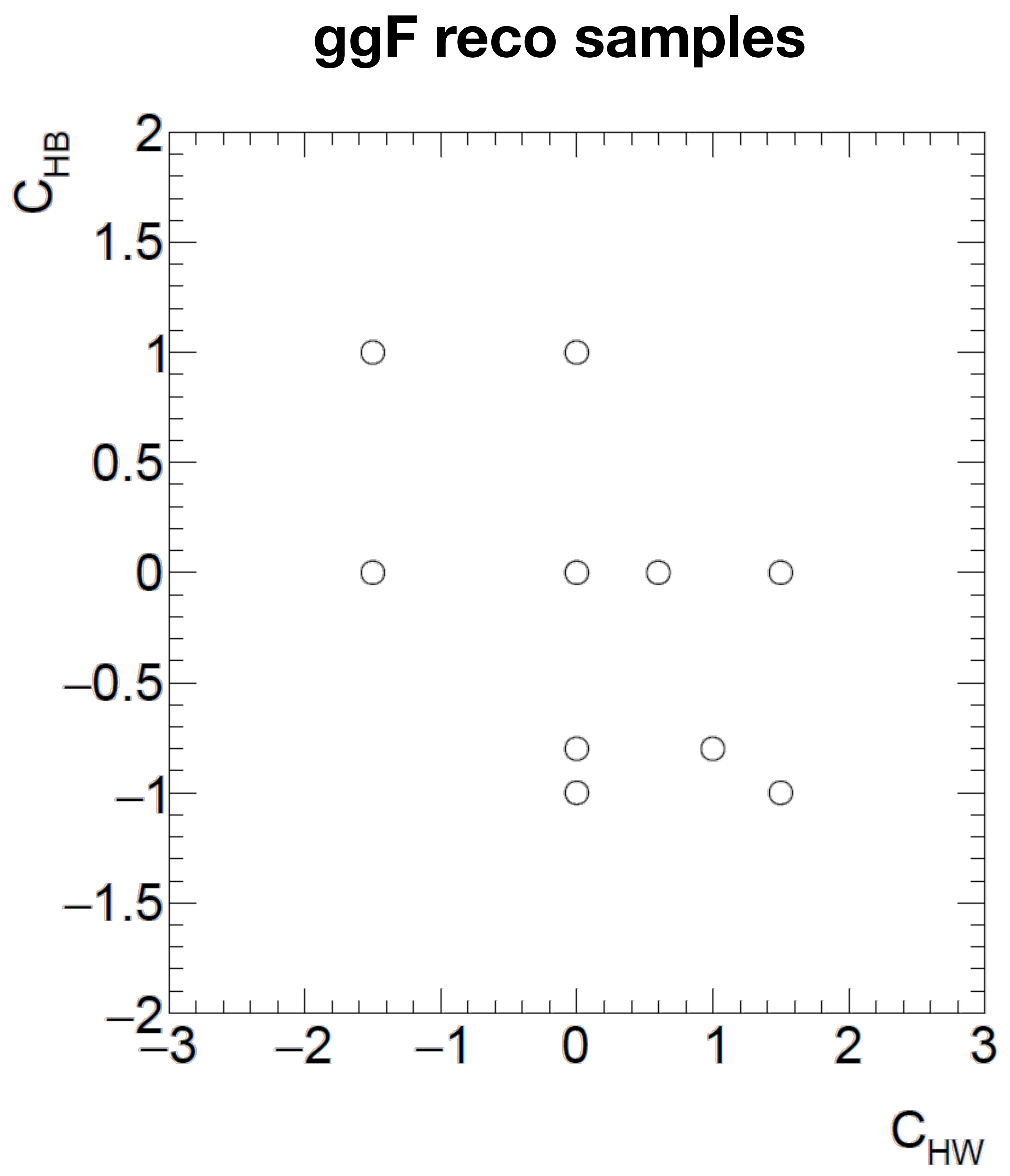
bbH

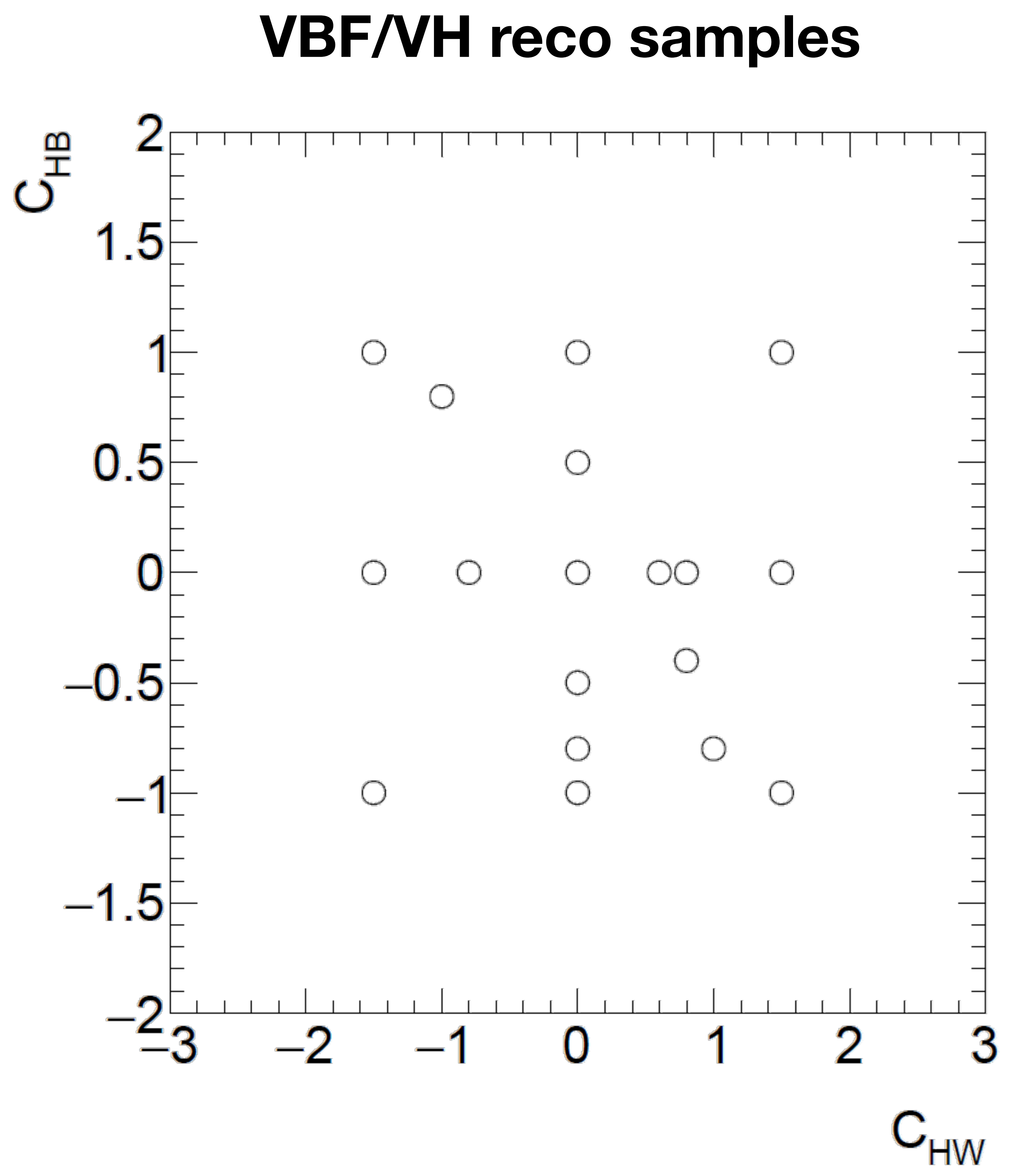


Signal MC samples

- The signal events are represented by the ggF and VBF Leading-order reco MC samples (20-130k events each, ggF SM is the highest).

- The NLO corrections are taken into account by the k-factors.





Event selection

The standard HZZ selection is used to select the signal events at the minitrees stage.

Leptons and Jets	
Electrons	$E_T > 7$ GeV and $\lvert\eta\rvert < 2.47$
Muons	$p_T > 5$ GeV and $\lvert\eta\rvert < 2.7$, calorimeter-tagged: $p_T > 15$ GeV
Jets	$p_T > 30$ GeV and $\lvert\eta\rvert < 4.5$

Quadruplets
All combinations of two same-flavour and opposite-charge lepton pairs
- Leading lepton pair: lepton pair with invariant mass m_{12} closest to the Z boson mass m_Z
- Subleading lepton pair: lepton pair with invariant mass m_{34} second closest to the Z boson mass m_Z
Classification according to the decay final state: 4μ, $2e2\mu$, $2\mu 2e$, $4e$

Requirements on each quadruplet	
Lepton reconstruction	- Three highest-p_T leptons must have p_T greater than 20, 15 and 10 GeV - At most one calorimeter-tagged or stand-alone muon
Lepton pairs	- Leading lepton pair: $50 < m_{12} < 106$ GeV - Subleading lepton pair: $m_{min} < m_{34} < 115$ GeV - Alternative same-flavour opposite-charge lepton pair: $m_{\ell\ell} > 5$ GeV - $\Delta R(\ell, \ell') > 0.10$ for all lepton pairs

- Four-lepton invariant mass window in the signal region: $115 < m_{4\ell} < 130$ GeV

Kinematic observables

- Different kinematic observables have different sensitivity to the BSM effects.

- The idea is to determine the most sensitive observables which will provide us the best limits.

- Mass-related observables:

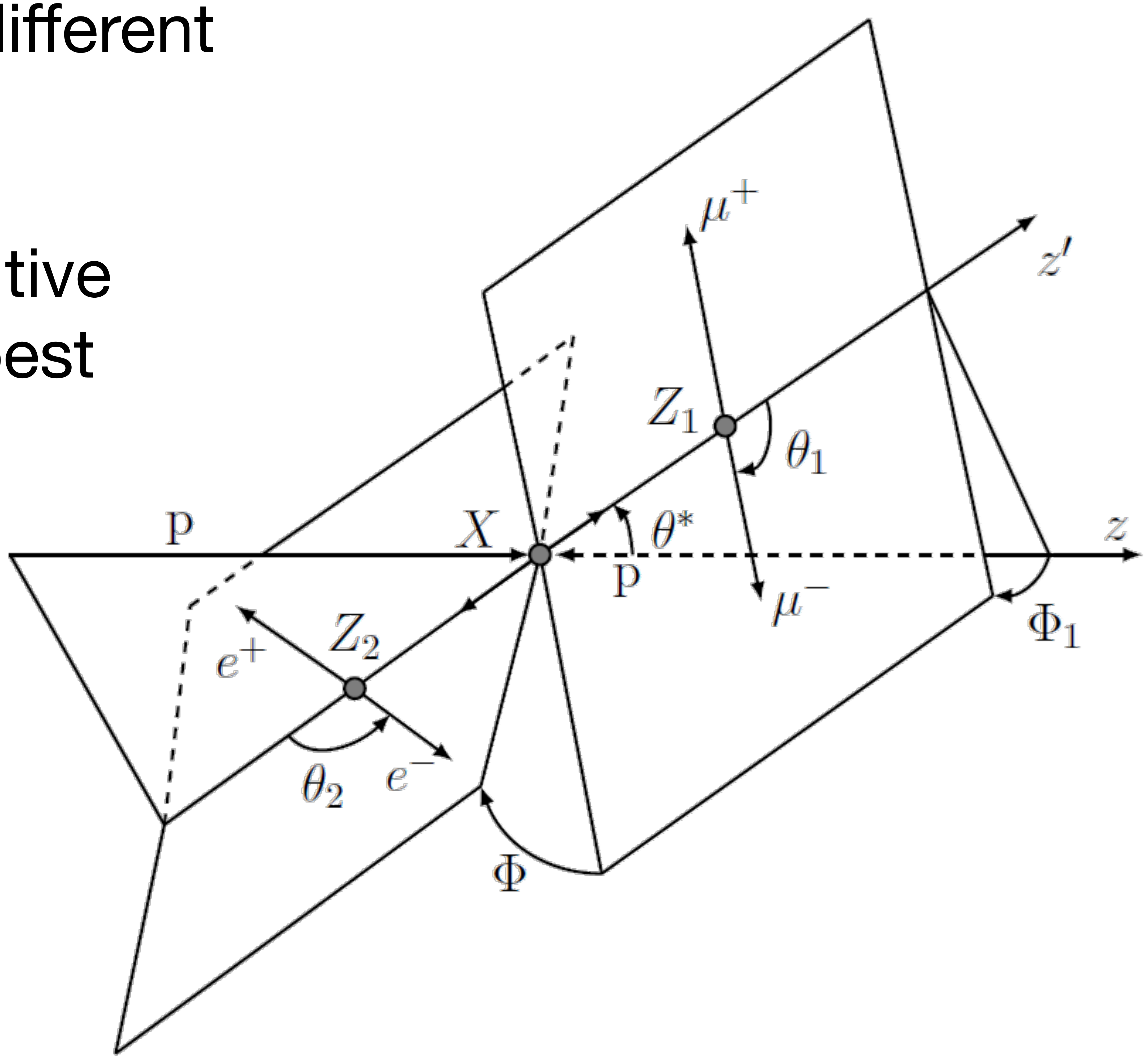
  $m_{Z_1}, m_{Z_2}, m_{4l}, m_{4lj}, m_{4ljj}, m_{jj}, m_{4l}^{err};$

- Momentum-related observables:

  $p_{4l}^{T}, p_{4lj}^{T}, p_{4ljj}^{T}, p_{jj}^{T}, y_{4l}, \eta_{4l}, \eta_{Zepp}^{ZZ}, \Delta\eta_{jj};$

- Angular observables:

  $cos\theta_1, cos\theta_2, cos\theta^*, \Phi, \Phi_1, \varphi_{4l}, \Delta\varphi_{jj}.$



Categorization

- The idea is to keep sensitive categories and merge the others.
- BSM samples suffer from low stats due to low acceptance at far BSM regions (less than 20k at some extreme cases).
- Optimization is currently on-going: pt4l_fsr-based categorizations are under investigation

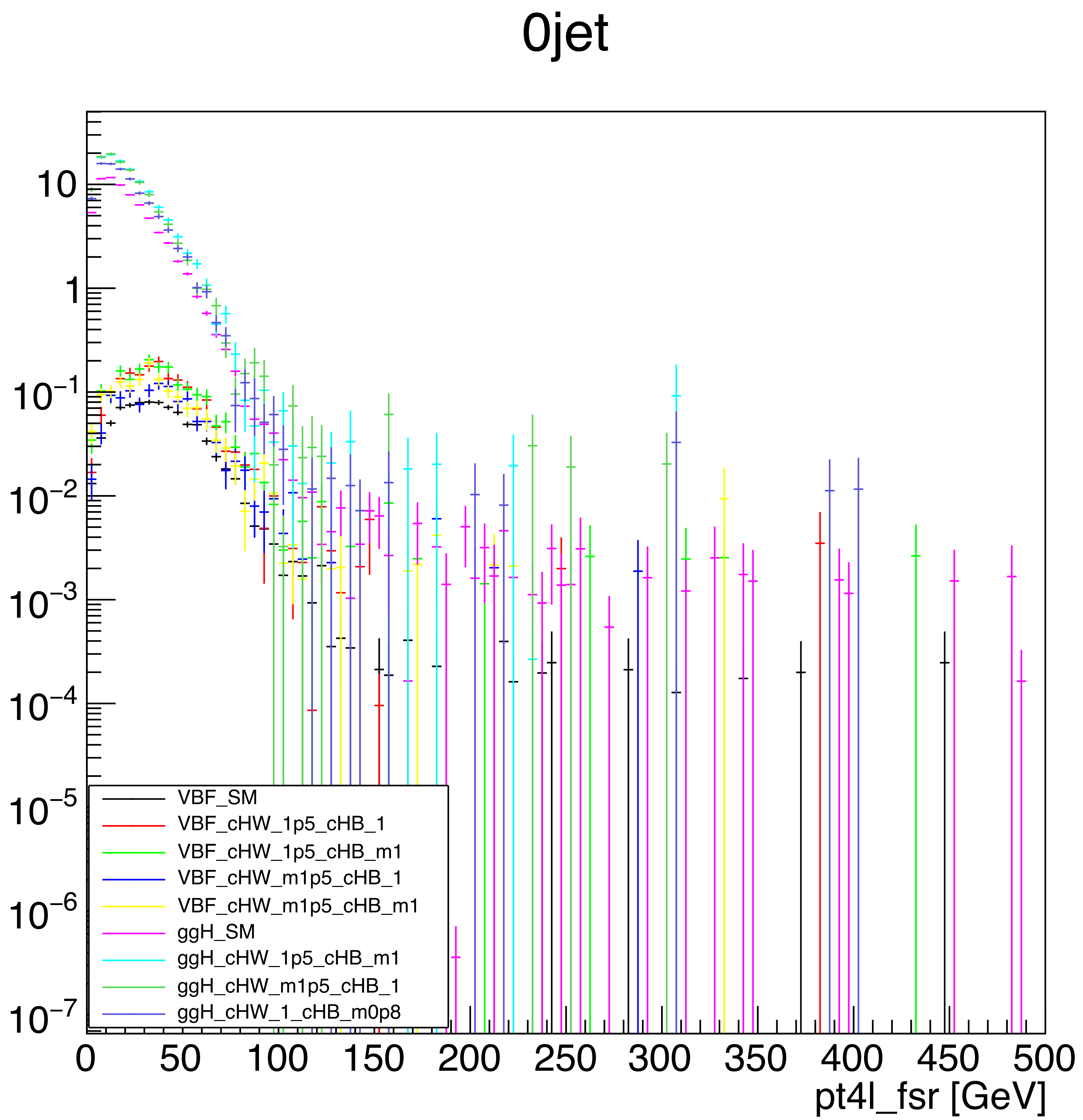
Previous categorization (phi_fsr - based)

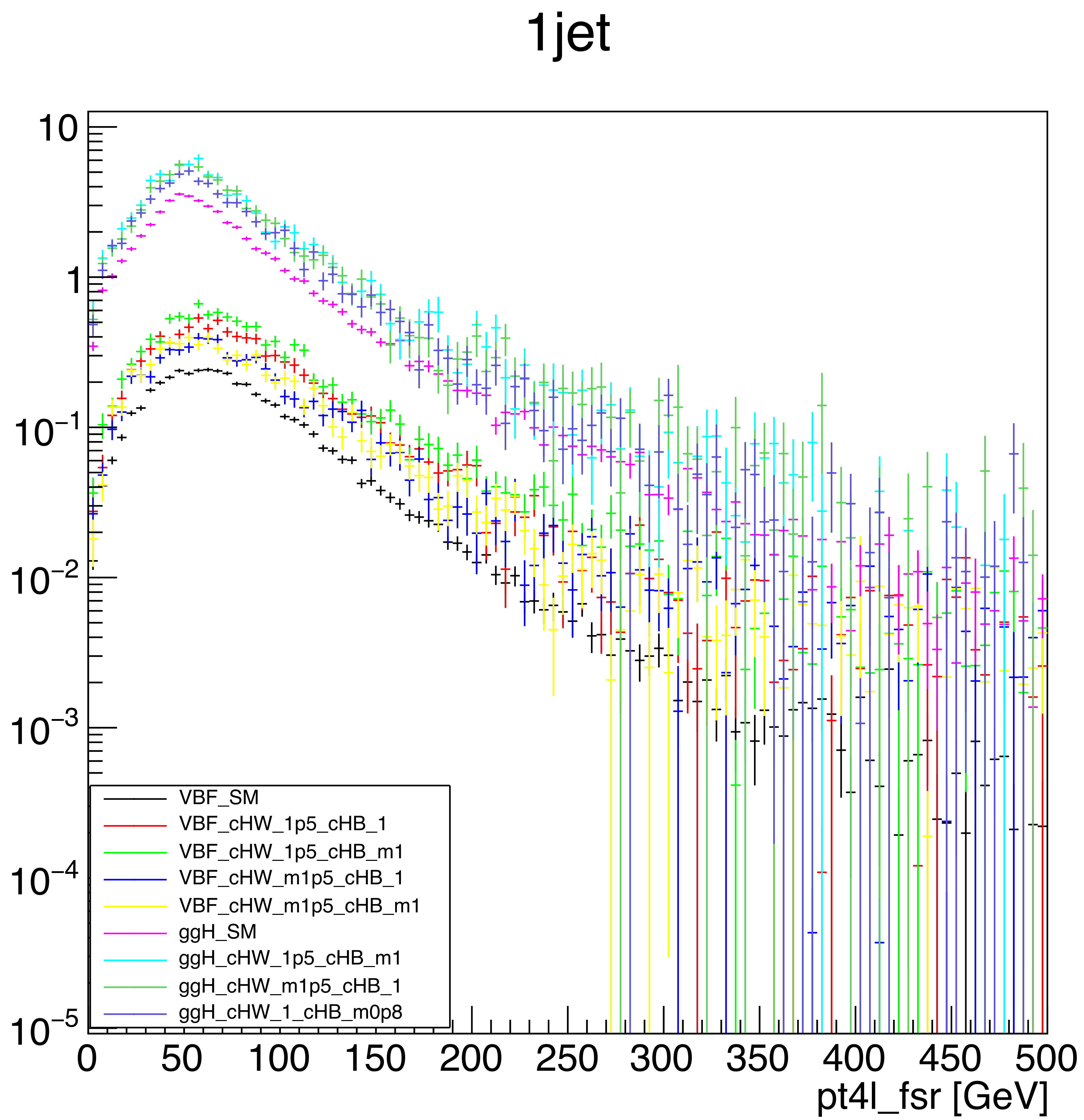
Categories
0jet_phi_low
0jet_phi_medium
0jet_phi_high
1jet
2jet_VBF
2jets_others

Observables with $n_{jets} > 0$

Observables with $n_{jets} > 1$

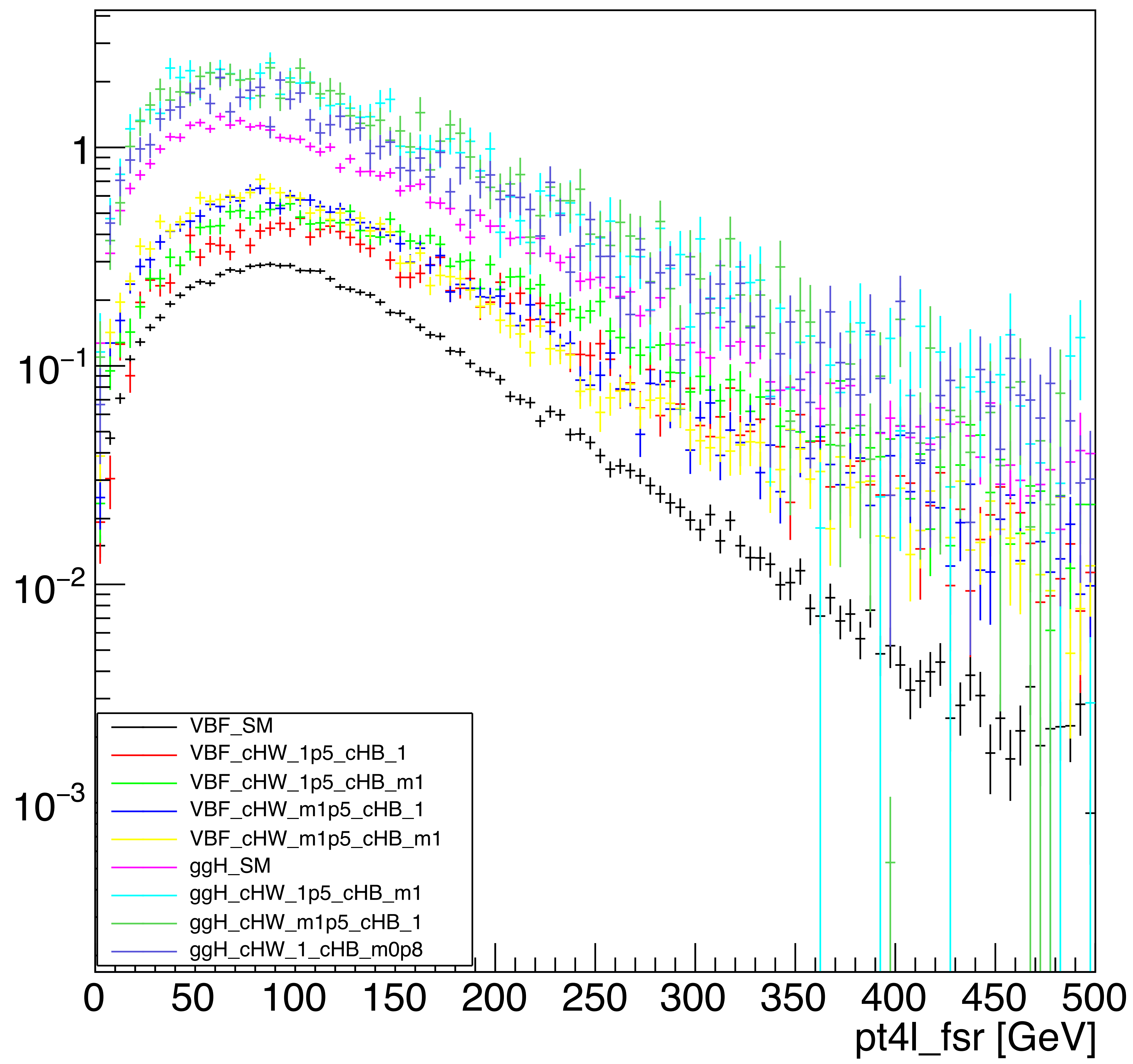
pt4_fsr distributions for several channels (I)



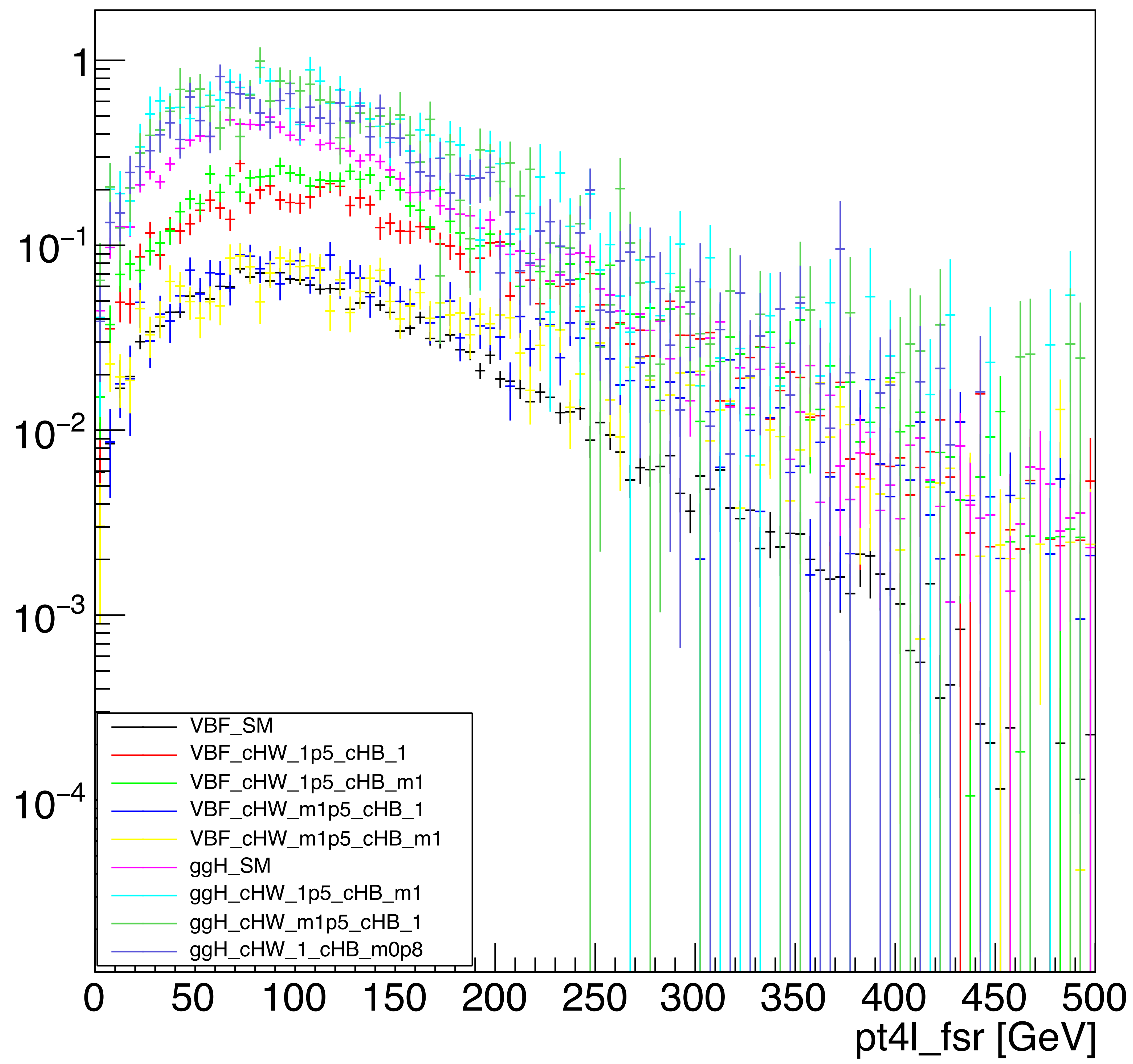


pt4_fsr distributions for several channels (II)

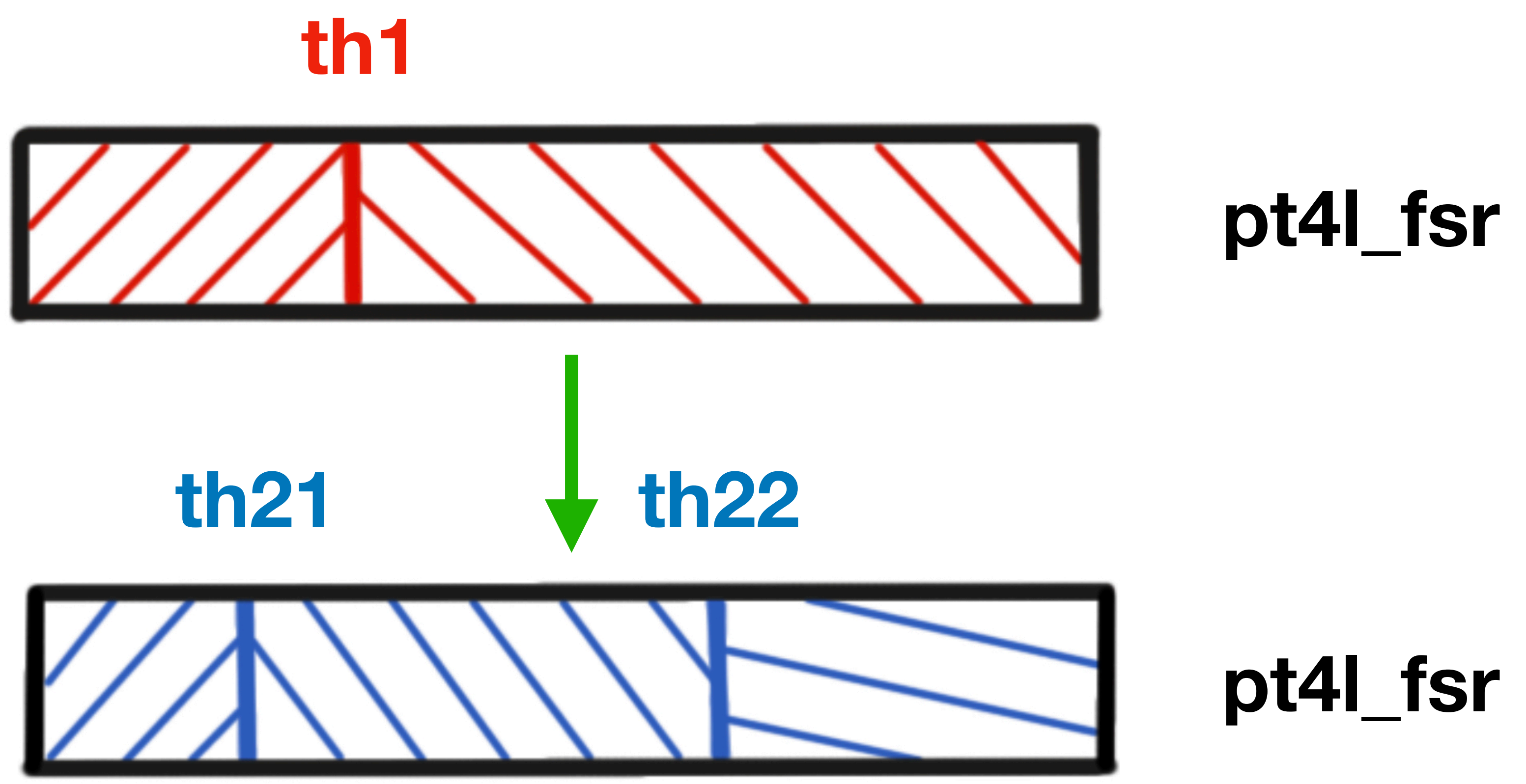
2jet_VBF

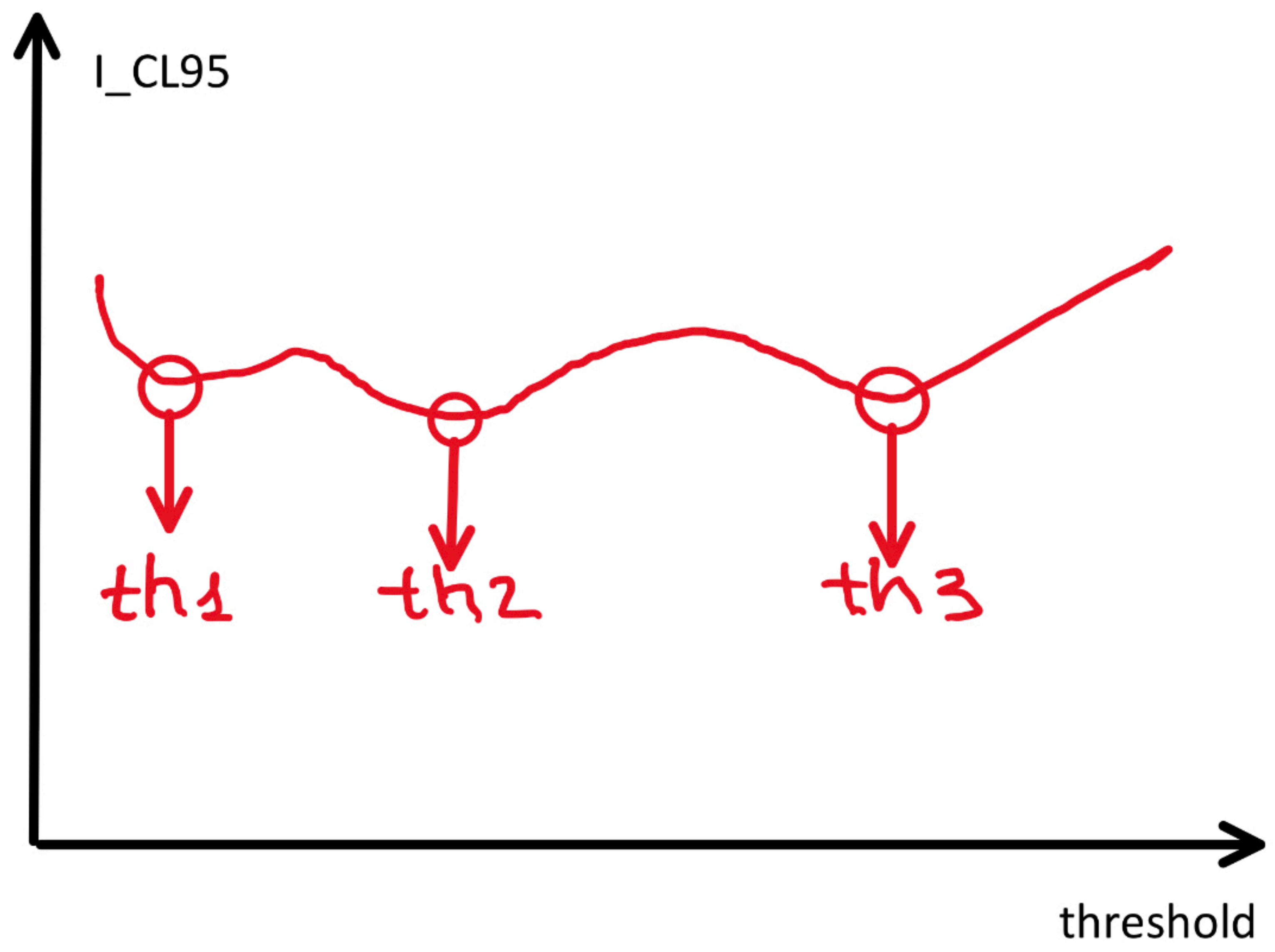


2jet_others



Brief description of categorization choice algorithm



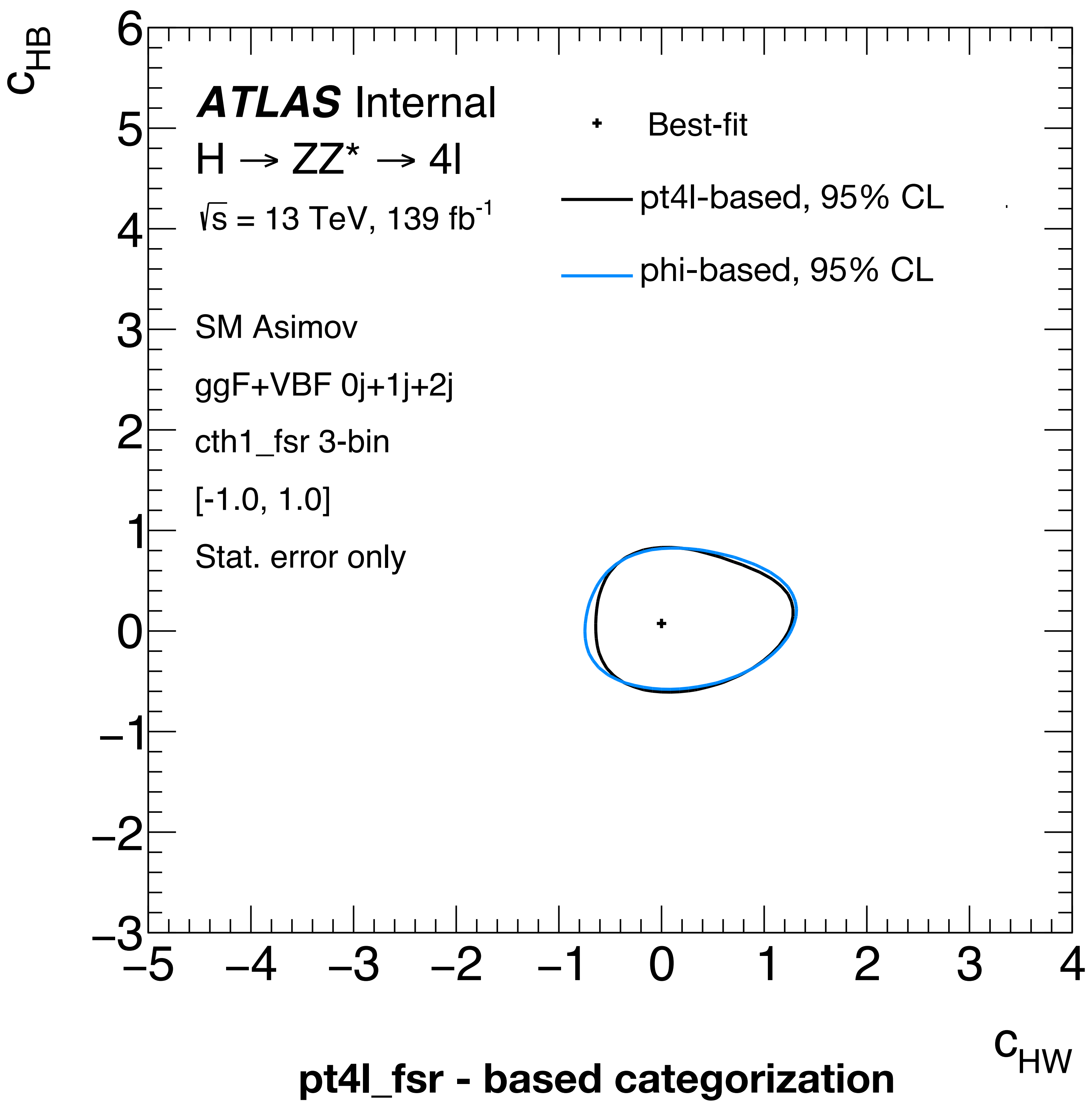


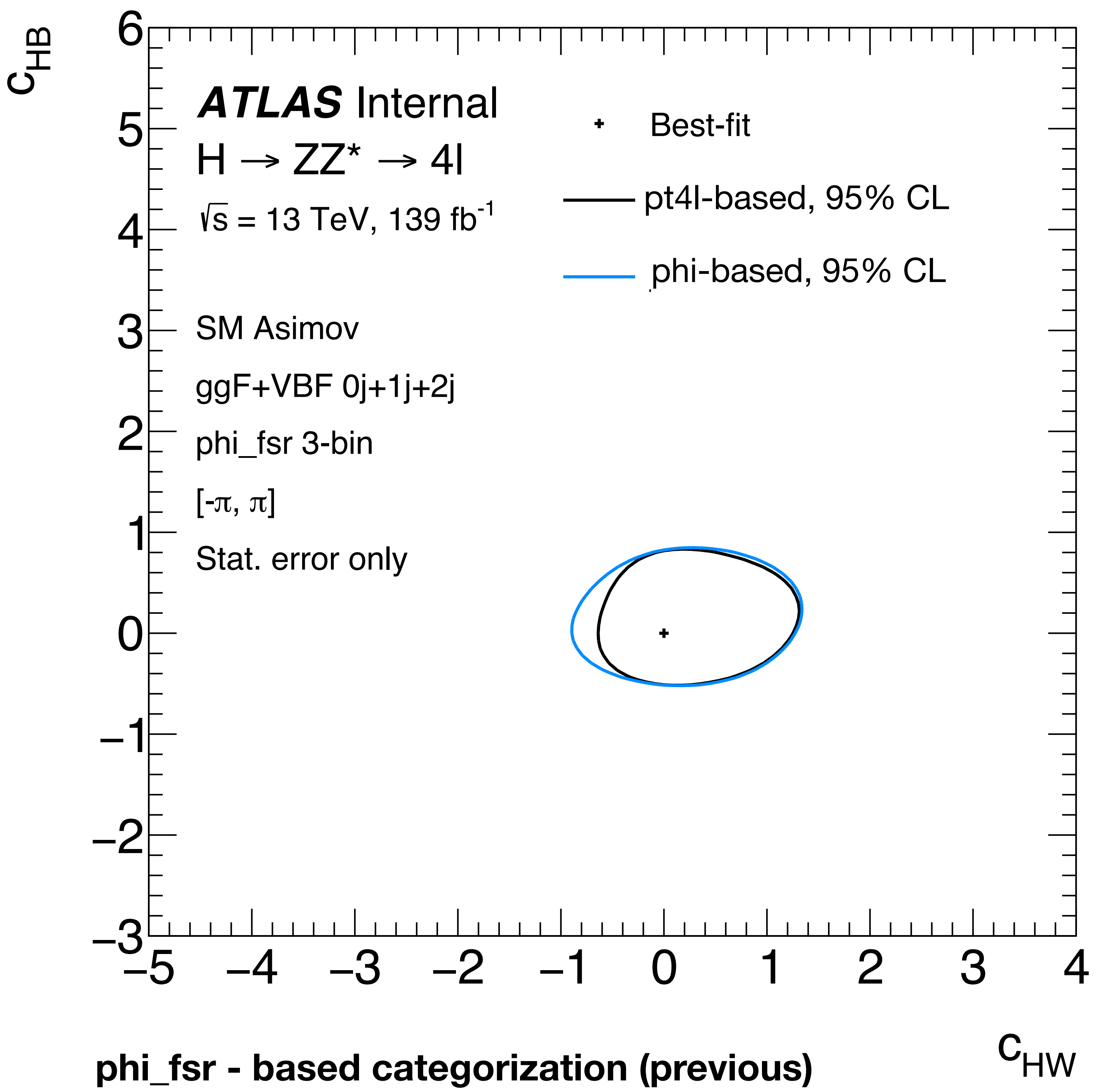
- Each category (per prod_type) is divided into 2 subcategories
- Local minimums are found
- 2 local minimums' values are used in categories, which are divided into 3 subcategories

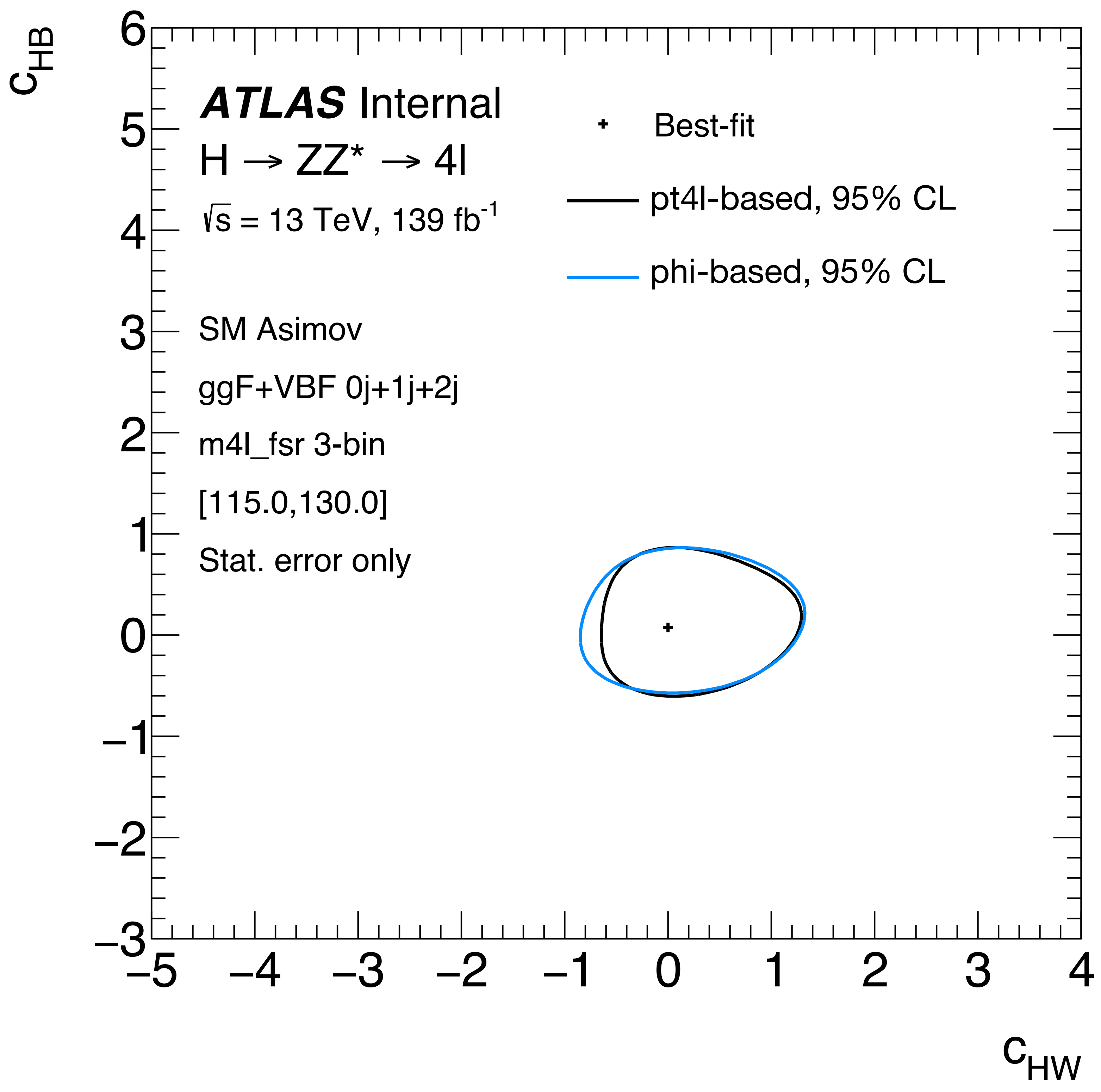
SN_test_pt4l_v3

	prod_type	pt4l_fsr
0jet_pt4l_fsr_low_v1		<=8 GeV
0jet_pt4l_fsr_medium_v1	0	8 GeV<pt4l_fsr<=24 GeV
0jet_pt4l_fsr_high_v1		>24 GeV
1jet_pt4l_fsr_low_v1		<=42 GeV
1jet_pt4l_fsr_medium_v1	1	42 GeV<pt4l_fsr<=178 GeV
1jet_pt4l_fsr_high_v1		> 178 GeV
2jet_VBF_pt4l_low_v1		<=70 GeV
2jet_VBF_pt4l_medium_v1	3	70 GeV<pt4l_fsr<=200 GeV
2jet_VBF_pt4l_fsr_high_v1		> 200 GeV
2jet_others	2 or >=3	

pt4l_fsr-based categorization vs phi_fsr-based categorization







pt4l_fsr - based categorization

	prod_type	pt4l_fsr
0jet_pt4l_fsr_low_v1	0	<=8 GeV
0jet_pt4l_fsr_medium_v1		8 GeV<pt4l_fsr<=24 GeV
0jet_pt4l_fsr_high_v1		>24 GeV
1jet_pt4l_fsr_low_v1	1	<=42 GeV
1jet_pt4l_fsr_medium_v1		42 GeV<pt4l_fsr<=178 GeV
1jet_pt4l_fsr_high_v1		> 178 GeV
2jet_VBF_pt4l_low_v1	3	<=70 GeV
2jet_VBF_pt4l_medium_v1		70 GeV<pt4l_fsr<=200 GeV
2jet_VBF_pt4l_fsr_high_v1		> 200 GeV
2jet_others	2 or >=3	

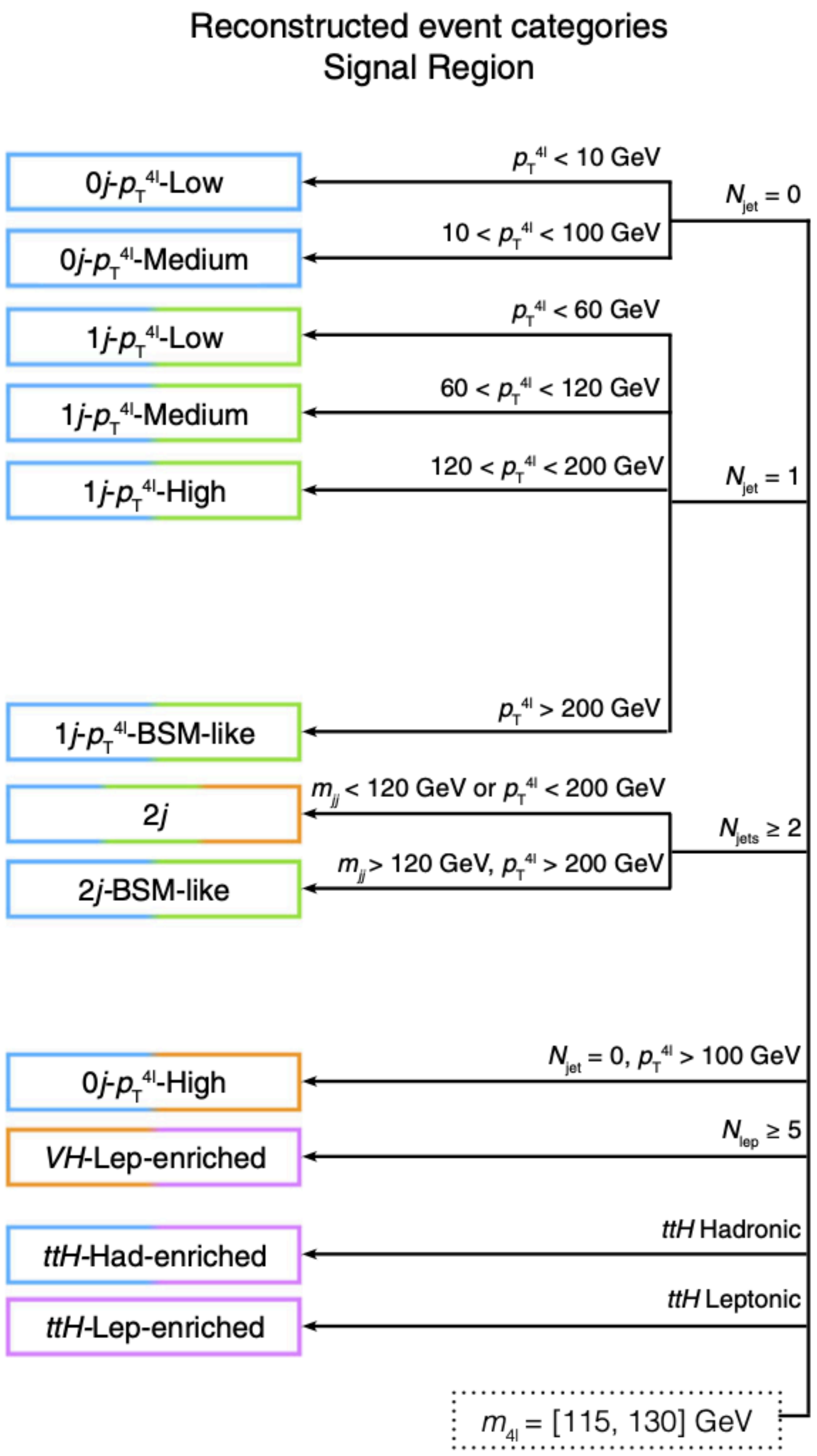
phi_fsr - based categorization (previous)

	prod_type	phi_fsr
0jet_phi_low	0	-π<phi_fsr<=-π/3
0jet_phi_medium		-π/3<phi_fsr<=π/3
0jet_phi_high		π/3<phi_fsr<=π
1jet	1	
2jet_VBF	3	
2jet_others	2 or >=3	

	cth1_fsr	phi_fsr	m4l_fsr
S_CL95_{pt4l} / S_CL95_{phi}	0.95	0.87	0.87

δ S = 0.07%

STXS*-like categorization



stxs-like_v1

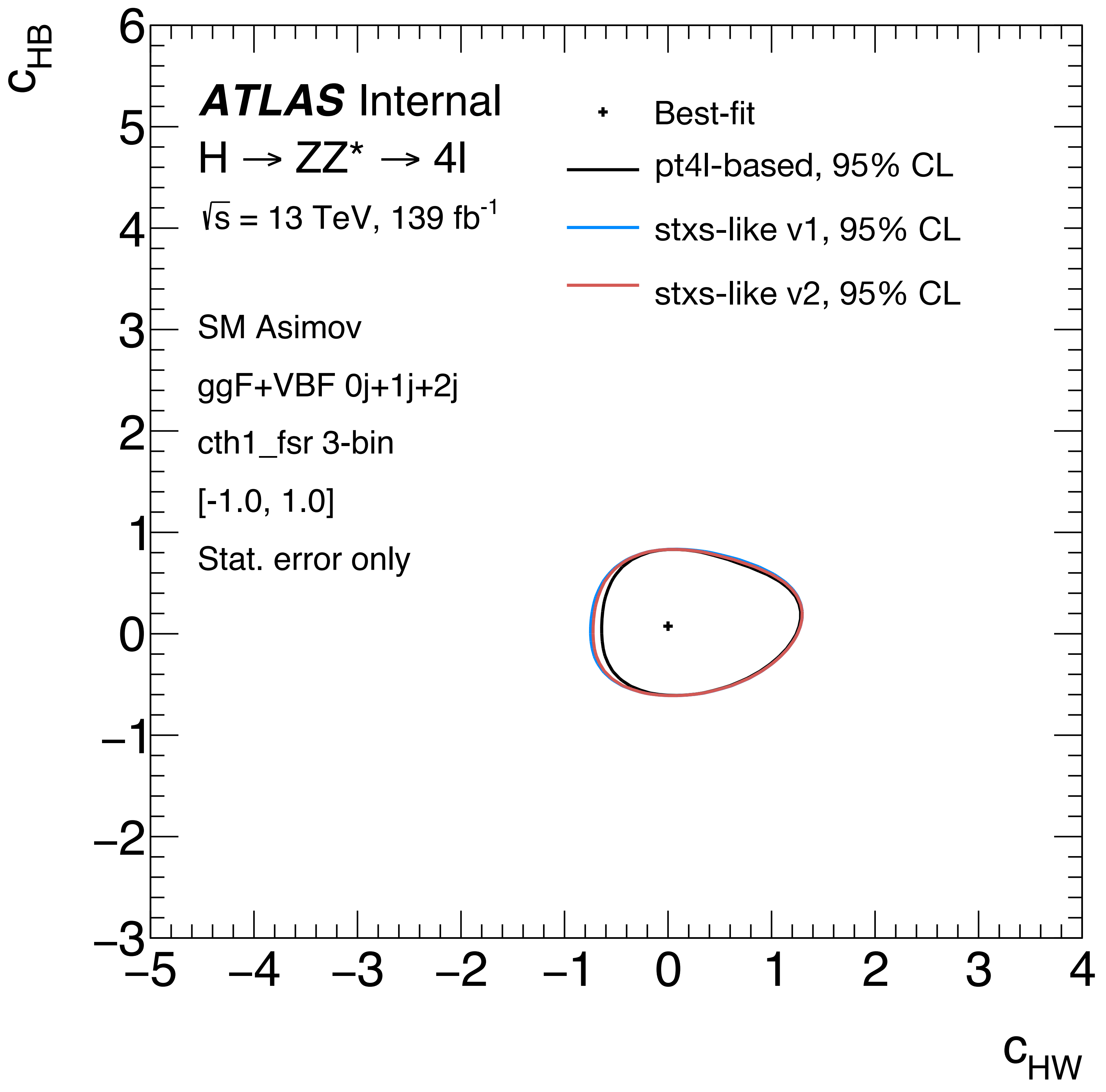
	prod_type	n_jets_fidDres_4lsel	pt4l_fsr
0jet_stxs_pt4l_fsr_low		0	<=10 GeV
0jet_stxs_pt4l_fsr_medium			10 GeV<pt4l_fsr<=100 GeV
0jet_stxs_pt4l_fsr_high			> 100 GeV
1jet_stxs_pt4l_fsr_low		1	<=60 GeV
1jet_stxs_pt4l_fsr_medium			60 GeV<pt4l_fsr<=120 GeV
1jet_stxs_pt4l_fsr_high_v1			> 120 GeV
2jet_stxs		>= 2	dijet_m_fidDres_4lsel<120 GeV or pt4l_fsr<200 GeV
2jet_stxs_bsm_like			dijet_m_fidDres_4lsel>120 GeV or pt4l_fsr>=200 GeV
2jet_stxs_others	>=4		

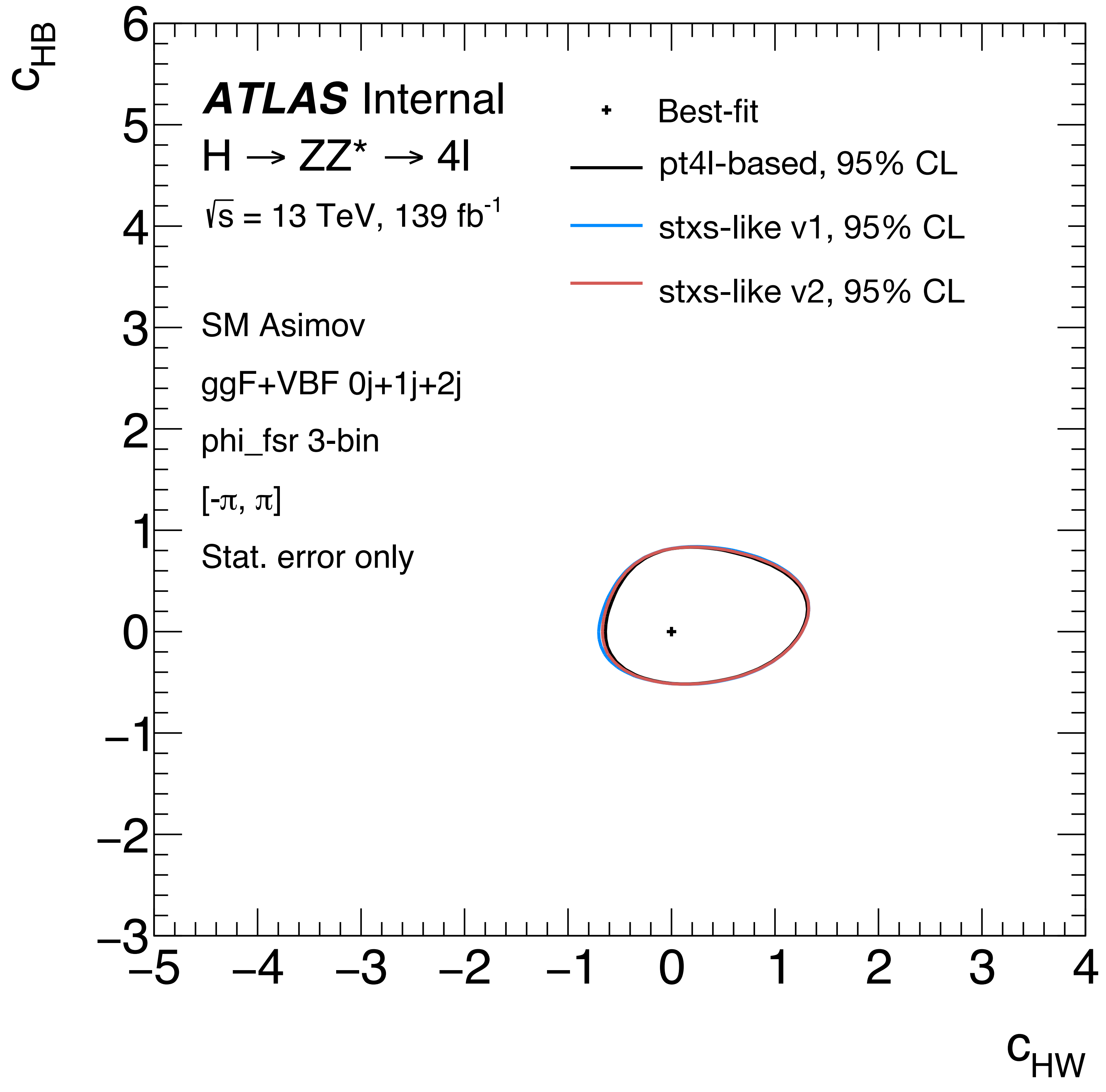
stxs-like v3

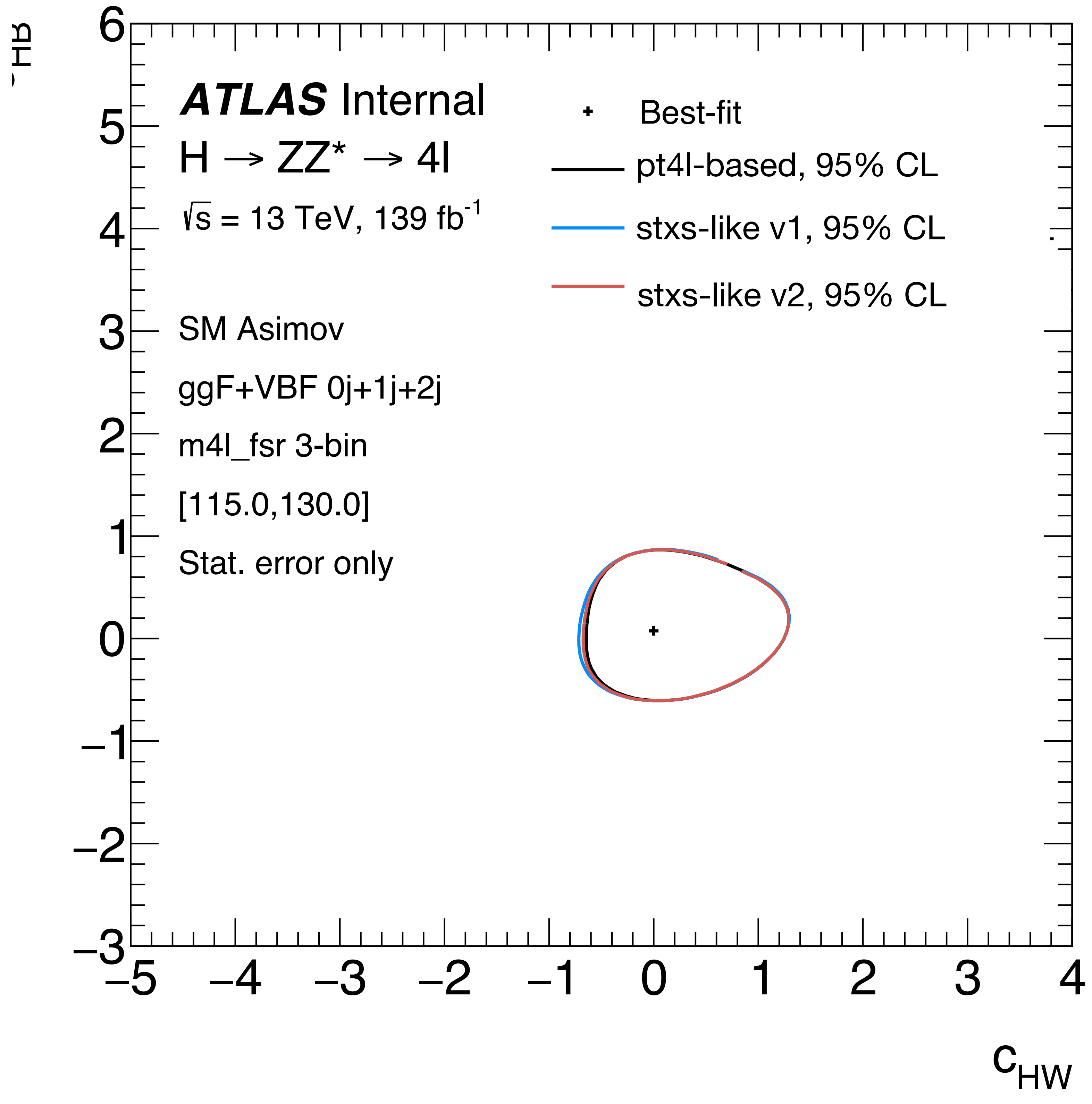
	prod_type	n_jets_fidDres_4lsel	pt4l_fsr
0jet_stxs_pt4l_fsr_low		0	<=10 GeV
0jet_stxs_pt4l_fsr_medium			10 GeV<pt4l_fsr<=100 GeV
0jet_stxs_pt4l_fsr_high		1	> 100 GeV
1jet_stxs_pt4l_fsr_low			<=60 GeV
1jet_stxs_pt4l_fsr_medium			60 GeV<pt4l_fsr<=120 GeV
1jet_stxs_pt4l_fsr_high_v3			120 GeV<pt4l_fsr<=200 GeV
1jet_stxs_pt4l_fsr_bsm_like			>200 GeV
2jet_stxs		>= 2	dijet_m_fidDres_4lsel<120 GeV or pt4l_fsr<200 GeV
2jet_stxs_bsm_like			dijet_m_fidDres_4lsel>120 GeV or pt4l_fsr>=200 GeV
VHlep	4		
ttHlep	5		
ttHhad	6		

*https://doi.org/10.1140/epjc/s10052-020-8227-9

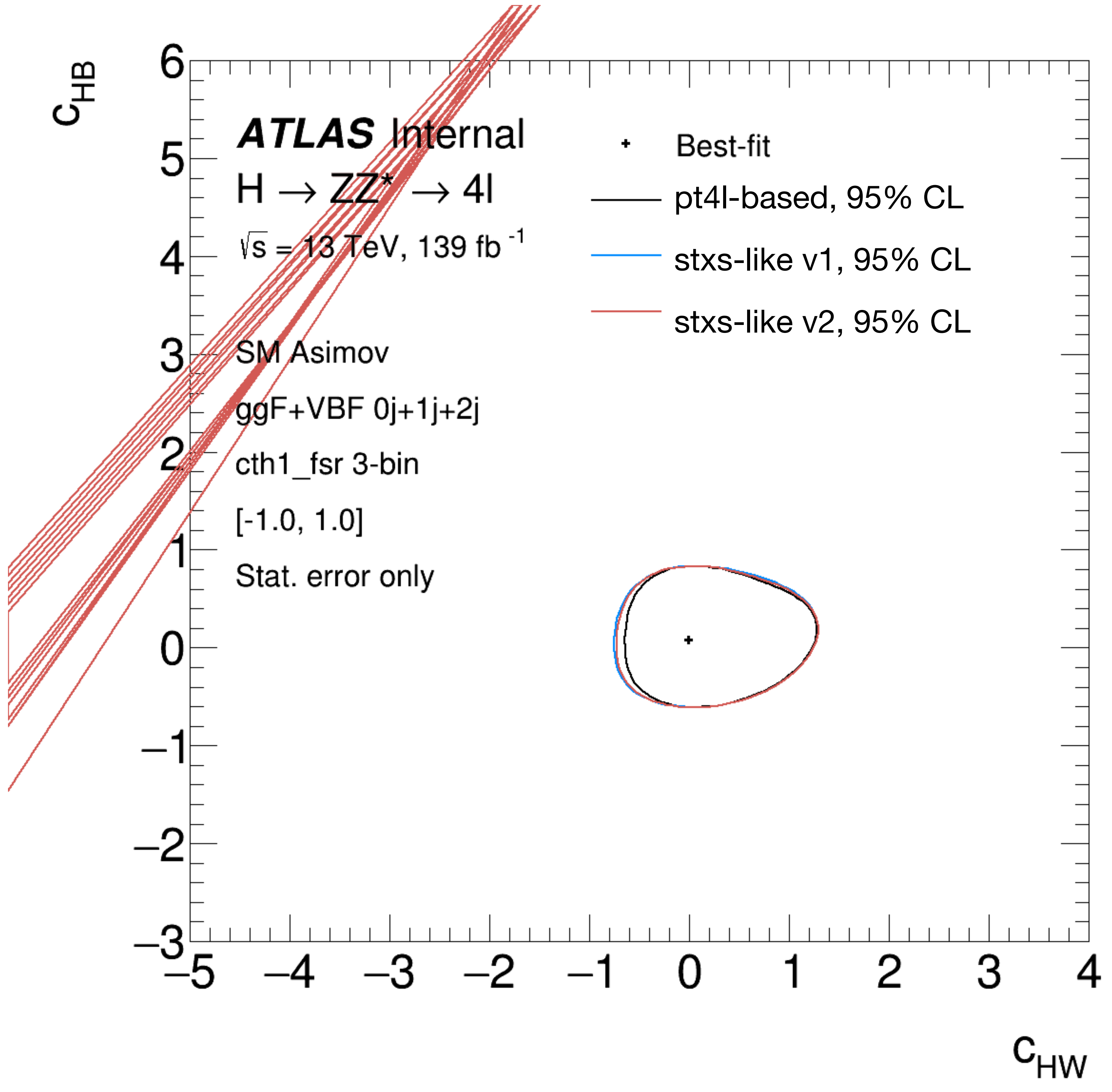
STXS-like categorization vs pt4l_fsr-based categorization

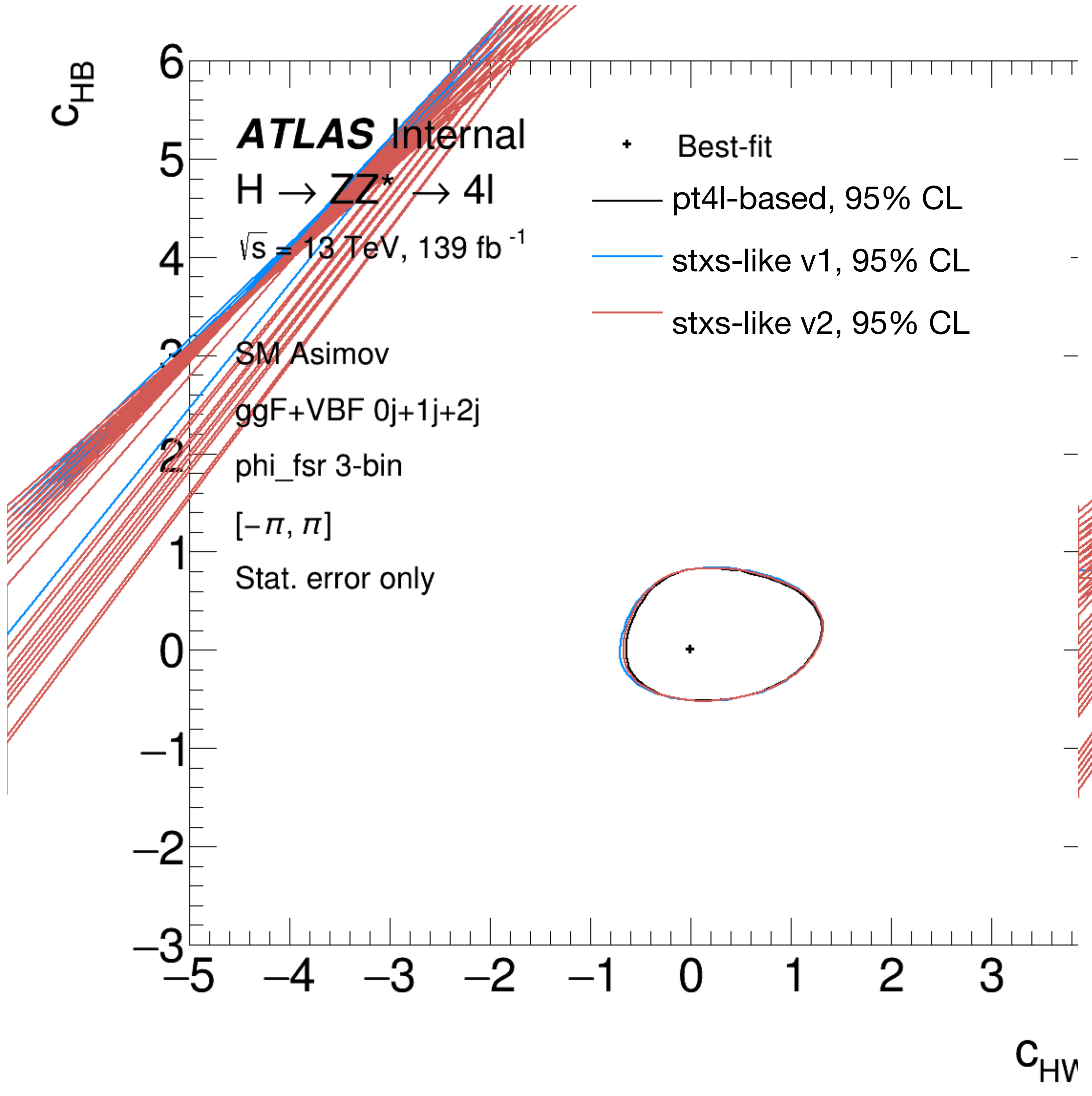


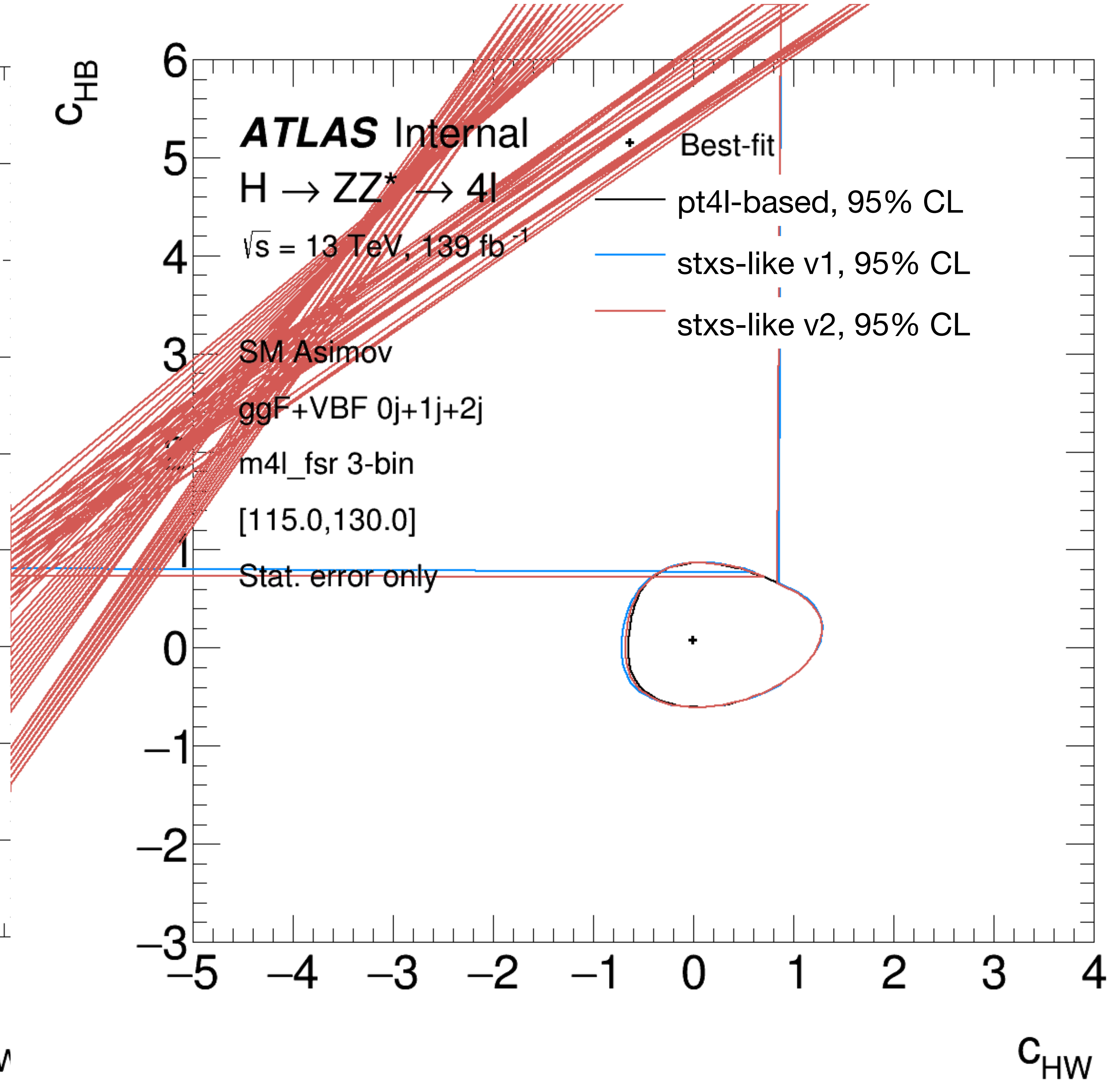




Problems with STXS-like categorization







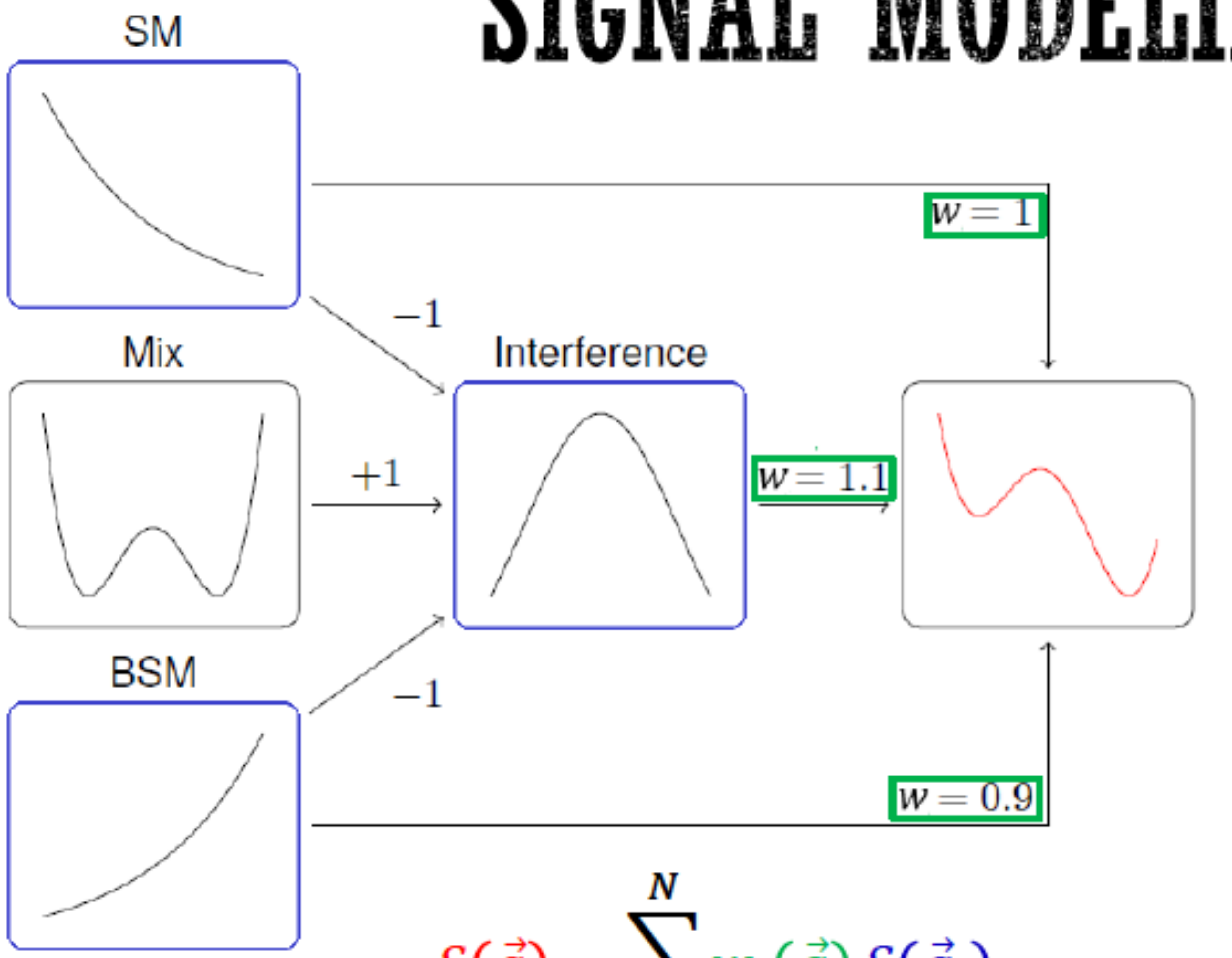
Preliminary results

- The overview of the EFT analysis is presented.
- The 2D LLH scans over CP-even C_{HW} and C_{HB} couplings are produced.
- The algorithm for categorization optimization has been created.
- The best values for the thresholds of pt4l-based categories were found
- Chosen p4l_fsr-based categorization shows itself to be more sensitive than phi_fsr-based categorization.

Plans:

- Investigate cth1_fsr-based categorizations
- Check the 1D LLH scans over CP-even C_{HW} and C_{HB} couplings for optimized pt4l_fsr-based categories
- To take systematics into account to make comparison with STXS results.

SIGNAL MODELING: MORPHING



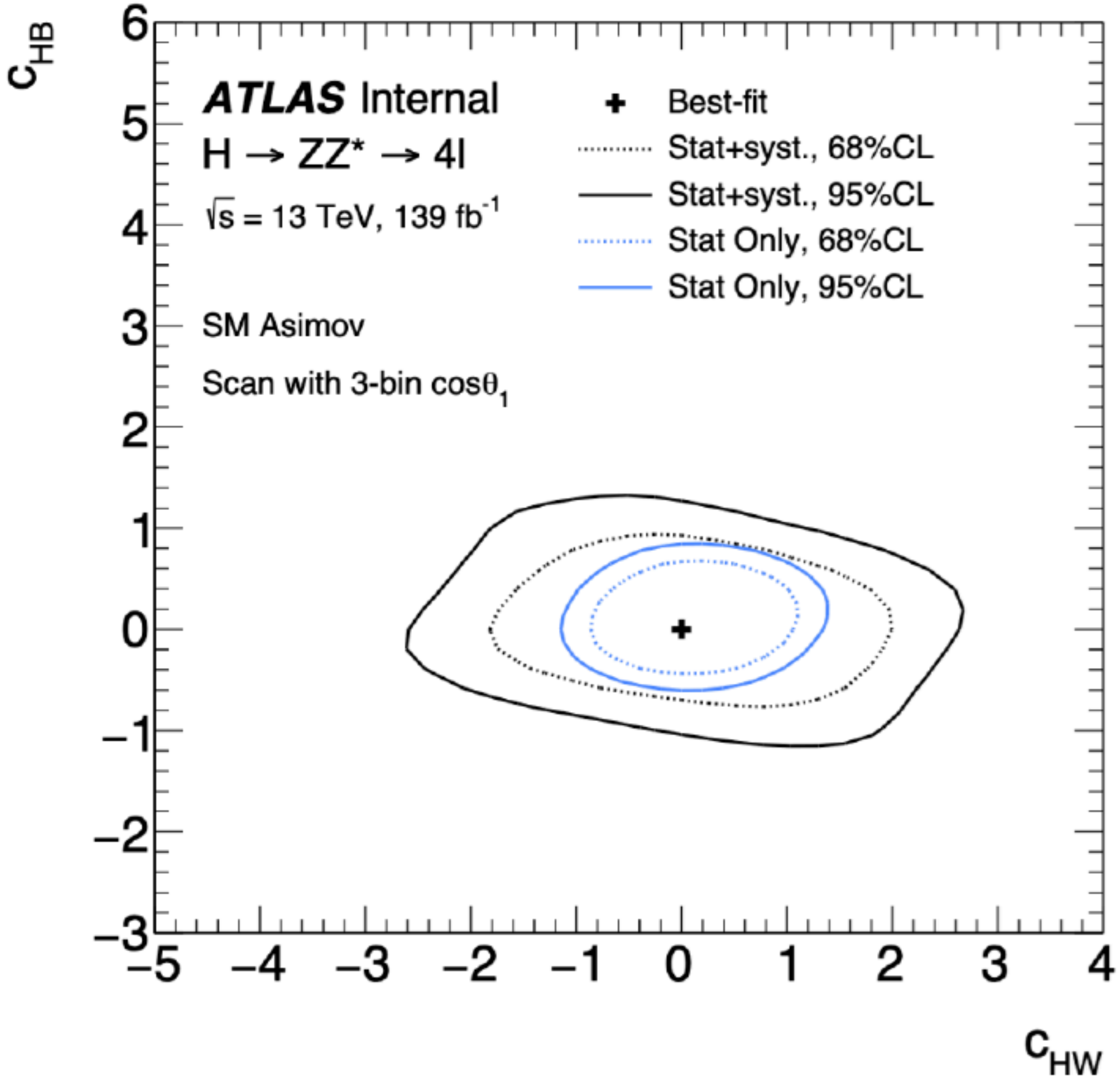
$$S(\vec{g}) = \sum_{i=1}^{N} w_i(\vec{g})\, S(\vec{g}_i)$$

- The signal model is created with the morphing method.
- In case of ggF and VBF production mechanisms with 2 BSM couplings, 6 and 15 samples required for the morphing.
- Advanced morphing technique (overdetermined morphing) was also implemented in the analysis package.

Production	VBFVHHad			ggF		
# of BSM couplings	1	2	3	1	2	3
n_s	2	3	4	1	1	1
n_p	0	0	0	0	0	0
n_d	0	0	0	1	2	3
N_{min}	5	15	35	3	6	10

- Several morphing-related topics are still in progress: basis optimization, statistical uncertainties and systematics.

CONTRIBUTION OF THE SYSTEMATICS: 2D



CATEGORIZATION

- The goal was to estimate the sensitivity of different observables.
- Scans with $n_{bins} = 3$ were produced based on each observable.
- Scans produced with all backgrounds and w/o systematics.

Categories
0jet_phi_low
0jet_phi_medium
0jet_phi_high
1jet
2jet_VBF
2jets_others

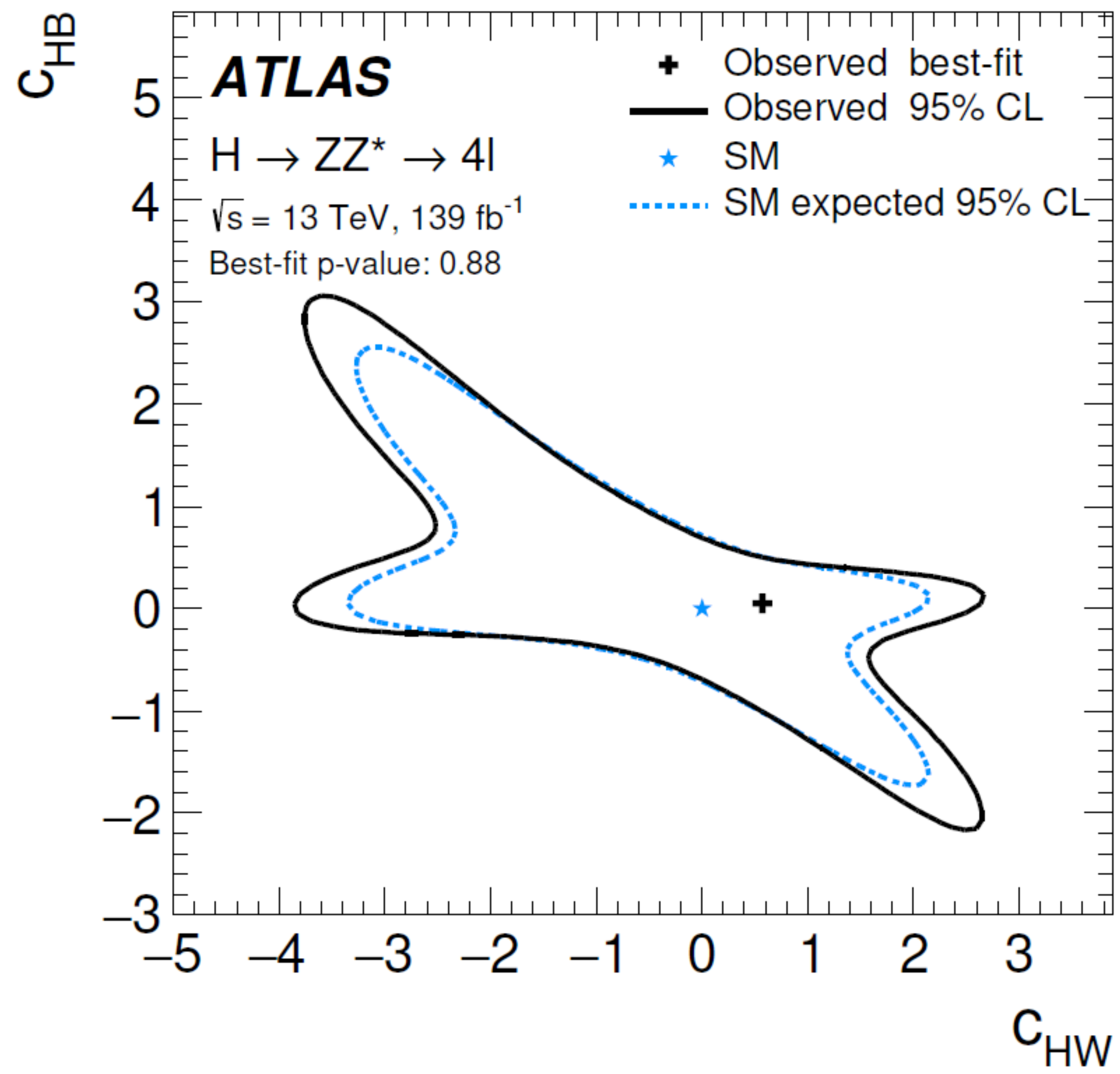
Observables with $n_{jets} > 0$ (1jet, 2jet_VBF, 2jets_others)

Observables with $n_{jets} > 1$ (2jet_VBF, 2jets_others)

m_{Z_1}	m_{Z_2}	m_{4l}	m_{4lj}	m_{4ljj}	m_{jj}	m_{4l}^{err}
[50; 100]	[12; 60]	[110; 135]	[160; 1200]	[220; 2200]	[0; 2000]	[1; 4]

p_{4l}^{T}	p_{4lj}^{T}	p_{4ljj}^{T}	p_{jj}^{T}	y_{4l}	η_{4l}	η_{Zepp}^{ZZ}	$\Delta\eta_{jj}$
[0; 180]	[0; 200]	[0; 150]	[0; 500]	[-2; 2]	[-4; 4]	[0; 4]	[0; 8]

$cos\theta_1$	$cos\theta_2$	$cos\theta^*$	Φ	Φ_1	φ_{4l}	$\Delta\varphi_{jj}$
[-1; 1]	[-1; 1]	[-1; 1]	[$-\pi$; π]	[$-\pi$; π]	[$-\pi$; π]	[0; 2π]


ATLAS
H → ZZ* → 4l
√s = 13 TeV, 139 fb-1
Best-fit p-value: 0.88
Observed best-fit
Observed 95% CL
SM
SM expected 95% CL
c_HB
c_HW

STXS result