



中国CMS冬令营, 2021.12.18

数据分析A-Z

肖朦 (浙江大学)



g g fusion :

t t fusion :

- Why: physics motivation
- Where: optimal channel
- How: optimal strategy







What does CMS see



through CMS

What does CMS see



POG: CMS Physics Object Groups PAG: CMS Physics Analysis Groups



Data/MC format



- For analyses, mostly miniAOD and nanoAOD, objects are all reconstructed
- CMSSW and root based

450kb/ev AOD 50kb/ev MiniAOD 1-2kb/ev **AOD** Nano



Taken from induction course here



Data/MC format

• What's the difference between data and MC?









CMS Preliminary





- - Hard scattering: theoretical ME calculation at different orders, POWHEG, Madgraph (MC@NLO)...
 - Parton shower and hadronization: Pythia, Herwig, Sherpa...
 - Detector response: Geant4, Delphes... \bullet

• Tools to simulate real data, usually for one particular process, e.g $H \rightarrow 4I$, $qq \rightarrow 4I$







Data and MC

- Where to look? CMS data aggregation system (DAS)
 - Website: <u>https://cmsweb.cern.ch/das/</u>
 - Terminal: dasgoclient -- query "dataset=/*/*/*"
- Naming conventions
 - Data: /DoubleEG/Run2016H-17Jul2018-v1/MINIAOD / Primary datasets / DataPeriod / Format Primary dataset: combination of a certain collections of HLT paths
 - MC: /QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/ RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TranchelV_v6-v1/MINIAODSIM



Trigger

- Data rates very high, 100 kHz, most of them not interesting
- HLT: high level trigger, 0.5 kHz, only events passing HLT stored
- Primary Datasets: grouped events with similar physical contents
 - Non-exclusive, one event could pass more than 1 HLT, thus 1 PD
- MC: apply the same HLT trigger and estimate scale factors

Primary Dataset	rate [Hz]			
SingleElectron	68			
DoubleElectron	15			
ElectronHad	14			
SinglePhoton	16			
SinglePhotonParked	69			
DoublePhoton	33			
DoublePhotonHighPt	9			
PhotonHad	11			

Numbers from <u>here</u>

MuEG	19
SingleMu	63
MuHad	14
DoubleMu	23
DoubleMuParked	26
MuOnia	38
MuOniaParked	94

HLT path

HLT_Ele17_Ele12_Calo HLT_Ele23_Ele12_Calo HLT_DoubleEle33_Calo HLT_Ele16_Ele12_Ele8_



Multijet	23		Tau	21			
Multijet1Parked	174		TauParkee	40			
JetHT	17		TauPLus	36			
HTMHT	16		BTag	3			
JetMon	6		BTagPlus	3			
VBF1Parked	201			·			
MET	16						
METParked	43						
			prescale prima		ary dataset		
[dL_TrackIdL_IsoVL	1	DoubleEG					
[dL_TrackIdL_IsoVL	1	DoubleEG					
dL_GsfTrkIdVL			1	DoubleEG			
_CaloIdL_TrackIdL	1	DoubleEG					

Event selection

- You will always see backgrounds
 - How to determine what the major bkg are?
 - Look for the same signature with large xsec



Based on the final signatures, design selections and strategies (blind!)

• $H \rightarrow ZZ \rightarrow 4I$: final state 4 leptons, find them and calculate the invariant mass



Event selection

- Trigger
- Pileup reweighting: match MC to data
- Object selection: good quality objects
- Analysis specific selections: enhance signal, reduce bkg
- Observable & category building: differentiate signal and bkg





Multiple protons collide, concentrate on one, others become pileup in the event



Object selection

- Detector has certain acceptance/resolution, follow POG recommendations
- Kinematic selection: pT > xx, $|\eta| < xx$ \bullet
- Object identification (ID): cut-based or machine learning techniques to identify objects from fakes Isolation & Significance of impact parameter (SIP)
- Energy/momentum calibration
- Scale factors of reconstruction/ID/Isolation/SIP: different selection efficiency in data/MC, use tag & probe
- Many analyses require development of object ID!





Muon ID

Analysis selection

• Rely on careful MC based studies



M4I > 70 GeV



M12 > 40 GeV

M34 > 12 GeV

Observables & categories



Cut based/Machine learning/ME based observables to differentiate signal from bkg

CMS 137 fb⁻¹ (13 TeV) Events / 0.1 8 Data $118 < m_{4l} < 130 \text{ GeV}$ H(125), VBF H(125), other VBF-2jet tagged category **qq**→**ZZ**, **Z**γ* **gg→ZZ, Ζ**γ* 6⊢ EW Z+X 5 ŀ 4 3 2 0 0.1 0.2 0.6 0.7 8.0 0.9 0.3 0.4 0.5 $\mathsf{D}_{bkg}^{\mathsf{VBF+dec}}$ 16



Observables & categories

Design categories where S/B ratio is high







Expected #events

- MC could simulate as many as events
- What are the number of events to expect in reality after selection?
- Need to know
 - σ: xsec of the process
 - L: Luminosity of data taking
 - N_{MC}: #events generated
 - N_{sel}: #events after selection

$$N_{exp} = \frac{\sigma \mathscr{L}}{N_{MC}} N_{sel}$$

Reconstructed ev category Untagged-0j- $p_{\rm T}^{4\ell}$ [(Untagged-0j- $p_{\rm T}^{4\ell}$ [10 Untagged-1j- $p_{\rm T}^{4\ell}$ [0 Untagged-1j- $p_{\rm T}^{4\ell}$ [60 Untagged-1j- $p_{\rm T}^{4\tilde{\ell}}$ [120 Untagged-2j- $p_{\rm T}^{4\ell}$ [(Untagged-2j- $p_{\rm T}^{4\ell}$ [60 Untagged-2j- $p_{\rm T}^{4\ell}$ [120 Untagged- $p_{\rm T}^{4\ell}$ > Untagged-2j- m_{ii} > VBF-1jet-tagge VBF-2jet-tagged-m_{ii} VBF-2jet-tagged-m_{ij} VBF-3jet-tagged-m_{ij} VBF-2jet-tagged- $p_{\rm T}^{4\ell}$ VBF-rest VH-hadronic-tagged-*n* VH-rest VH-leptonic-tagged-p VH-leptonic-tagged-p ttH-leptonic-tag tīH-hadronic-tag

Signal					Background				Expected			
ggH	VBF	WH	ZH	tīH	bbH	tΗ	$q\overline{q} \rightarrow ZZ$	gg ightarrow ZZ	EW	Z+X	signal	total
27.7	0.09	0.03	0.03	0.00	0.15	0.00	71.5	3.06	0.01	3.21	27.9±0.1	106±0
96.2	1.69	0.60	0.77	0.01	1.01	0.00	98.1	11.6	0.35	37.8	100 ± 0	$248{\pm}1$
26.8	1.51	0.56	0.48	0.01	0.45	0.01	25.3	3.02	0.64	14.2	29.8±0.1	$72.9{\pm}0.4$
13.5	1.31	0.51	0.41	0.02	0.11	0.01	7.81	0.82	0.62	7.95	$15.9 {\pm} 0.1$	$33.1 {\pm} 0.3$
3.51	0.60	0.17	0.17	0.01	0.02	0.00	1.15	0.19	0.25	1.63	$4.48 {\pm} 0.05$	$7.69 {\pm} 0.16$
3.45	0.29	0.15	0.14	0.08	0.09	0.02	2.14	0.32	0.63	4.75	4.20 ± 0.06	12.1 ± 0.2
5.26	0.56	0.24	0.19	0.12	0.04	0.03	2.19	0.30	0.72	4.14	6.43 ± 0.06	$13.8 {\pm} 0.2$
3.07	0.40	0.16	0.13	0.07	0.01	0.02	0.75	0.14	0.34	1.19	$3.86 {\pm} 0.05$	$6.28 {\pm} 0.14$
2.79	0.62	0.21	0.17	0.07	0.01	0.02	0.43	0.21	0.21	0.73	3.89 ± 0.04	$5.47 {\pm} 0.11$
0.77	0.16	0.06	0.04	0.05	0.01	0.01	0.34	0.06	0.31	1.71	1.12 ± 0.02	$3.54{\pm}0.14$
15.5	3.29	0.22	0.16	0.00	0.13	0.01	6.85	1.53	0.20	2.44	19.3 ± 0.1	30.3 ± 0.2
0.83	1.19	0.01	0.01	0.00	0.01	0.00	0.19	0.07	0.11	0.14	2.05 ± 0.03	2.55 ± 0.05
0.43	1.96	0.00	0.00	0.00	0.00	0.00	0.07	0.05	0.12	0.03	2.40 ± 0.02	2.67 ± 0.03
2.52	2.35	0.06	0.06	0.03	0.03	0.05	0.62	0.21	0.64	2.43	5.11 ± 0.05	9.01 ± 0.17
0.44	0.79	0.01	0.01	0.01	0.00	0.01	0.03	0.03	0.04	0.06	1.26 ± 0.02	1.42 ± 0.03
2.48	0.94	0.13	0.09	0.04	0.04	0.01	0.98	0.20	0.39	2.18	3.74 ± 0.05	7.49 ± 0.17
4.11	0.25	1.09	0.96	0.13	0.06	0.02	1.69	0.22	0.52	2.93	6.62 ± 0.06	12.0 ± 0.2
0.57	0.03	0.09	0.06	0.03	0.01	0.00	0.16	0.02	0.06	0.33	0.79 ± 0.02	1.36 ± 0.06
0.33	0.04	0.85	0.26	0.10	0.03	0.03	2.16	0.36	0.19	1.11	$1.64{\pm}0.02$	5.47 ± 0.13
0.02	0.01	0.21	0.06	0.04	0.00	0.01	0.05	0.01	0.03	0.08	0.35 ± 0.01	0.52 ± 0.03
0.02	0.01	0.02	0.02	0.68	0.00	0.03	0.08	0.01	0.23	0.21	0.79 ± 0.01	1.32 ± 0.07
0.18	0.05	0.03	0.05	0.86	0.01	0.03	0.03	0.01	0.82	1.06	$ 1.22 \pm 0.01$	3.15 ± 0.14
	ggH 27.7 96.2 26.8 13.5 3.51 3.45 5.26 3.07 2.79 0.77 15.5 0.83 0.43 2.52 0.44 2.48 4.11 0.57 0.33 0.02 0.02 0.02 0.18	ggHVBF27.70.0996.21.6926.81.5113.51.313.510.603.450.295.260.563.070.402.790.620.770.1615.53.290.831.190.431.962.522.350.440.792.480.944.110.250.570.030.330.040.020.010.020.010.180.05	ggHVBFWH27.70.090.0396.21.690.6026.81.510.5613.51.310.513.510.600.173.450.290.155.260.560.243.070.400.162.790.620.210.770.160.0615.53.290.220.831.190.010.431.960.002.522.350.060.440.790.012.480.940.134.110.251.090.570.030.090.330.040.850.020.010.210.030.050.03	signalggHVBFWHZH27.70.090.030.0396.21.690.600.7726.81.510.560.4813.51.310.510.413.510.600.170.173.450.290.150.145.260.560.240.193.070.400.160.132.790.620.210.170.770.160.060.0415.53.290.220.160.831.190.010.010.431.960.000.002.522.350.060.060.440.790.010.012.480.940.130.094.110.251.090.960.570.030.090.060.330.040.850.260.020.010.210.060.020.010.020.02	ggHVBFWHZHtTH27.70.090.030.030.0096.21.690.600.770.0126.81.510.560.480.0113.51.310.510.410.023.510.600.170.170.013.450.290.150.140.085.260.560.240.190.123.070.400.160.130.072.790.620.210.170.070.770.160.060.040.0515.53.290.220.160.000.431.960.000.000.010.440.790.010.010.012.480.940.130.090.044.110.251.090.960.130.570.030.090.060.040.020.010.210.060.040.020.010.210.060.04	signalVBFWHZHtTHbbH27.70.090.030.030.000.1596.21.690.600.770.011.0126.81.510.560.480.010.4513.51.310.510.410.020.113.510.600.170.170.010.023.450.290.150.140.080.095.260.560.240.190.120.013.070.400.160.130.070.012.790.620.210.170.070.010.770.160.060.040.050.0115.53.290.220.160.000.010.431.960.000.000.000.010.440.790.010.010.010.010.451.090.010.010.010.010.450.040.050.060.030.010.431.960.010.010.010.010.440.790.010.010.010.010.451.090.060.030.010.460.940.130.060.30.010.470.040.050.260.100.030.430.940.130.060.030.010.440.790.010.060.030.010.570.030.090.26	signalVBFWHZHttHbbHtH27.70.090.030.030.000.150.0096.21.690.600.770.011.010.0026.81.510.560.480.010.450.0113.51.310.510.410.020.110.013.510.600.170.170.010.020.003.450.290.150.140.080.090.025.260.560.240.190.120.040.033.070.400.160.130.070.010.022.790.620.210.170.070.010.020.770.160.060.040.050.010.0115.53.290.220.160.000.010.000.431.960.000.000.000.000.010.431.960.010.010.000.010.010.440.790.010.010.010.010.010.440.790.010.010.010.010.011.480.940.130.090.040.030.030.570.030.090.060.030.010.010.330.040.210.060.040.000.010.330.040.020.020.680.000.030.530.040.02 <t< td=""><td>SignalVBFWHZHtHbbHtHqq \rightarrow ZZ27.70.090.030.000.150.0071.596.21.690.600.770.011.010.0098.126.81.510.560.480.010.450.0125.313.51.310.510.410.020.110.017.813.510.600.170.170.010.020.001.153.450.290.150.140.080.090.022.145.260.560.240.190.120.040.032.193.070.400.160.130.070.010.020.752.790.620.210.170.070.010.020.430.770.160.060.040.050.010.010.3415.53.290.220.160.000.010.000.190.431.960.000.000.000.000.072.522.350.060.060.030.030.050.620.440.790.010.010.010.000.010.032.480.940.130.060.030.010.032.160.570.030.090.060.030.010.050.040.010.530.040.260.100.030.032.160.540.01<t< td=""><td>signalVBFWHZHtHb$\overline{b}H$tH$q \overline{q} \rightarrow ZZ$$gg \rightarrow ZZ$27.70.090.030.030.000.150.0071.53.0696.21.690.600.770.011.010.0098.111.626.81.510.560.480.010.450.0125.33.0213.51.310.510.410.020.110.017.810.823.510.600.170.170.010.020.001.150.193.450.290.150.140.080.090.022.140.325.260.560.240.190.120.040.032.190.303.070.400.160.130.070.010.020.430.210.770.160.060.040.050.010.010.340.0615.53.290.220.160.000.010.000.070.050.431.960.000.000.000.000.070.050.210.431.960.010.010.000.000.000.070.030.431.960.000.000.000.000.070.052.522.350.060.060.330.050.620.210.440.790.010.010.040.010.980.200.440.790.010.</td><td>SignalBackgrountggHVBFWHZHtHbHtH$q\bar{q} \rightarrow ZZ$$gg \rightarrow ZZ$EW27.70.090.030.030.000.150.0071.53.060.0196.21.690.600.770.011.010.0098.111.60.3526.81.510.560.480.010.450.0125.33.020.6413.51.310.510.410.020.110.017.810.820.623.510.600.170.170.010.020.001.150.190.253.450.290.150.140.080.090.022.140.320.635.260.560.240.190.120.040.032.190.300.723.070.400.160.130.070.010.020.430.210.213.790.620.210.170.070.010.020.430.060.311.553.290.220.160.000.010.030.070.050.120.431.960.000.000.000.000.070.050.120.431.960.000.000.000.000.070.050.120.431.960.010.010.000.010.030.030.040.440.790.110.010.01</td><td>SignalBackgroundggHVBFWHZHtHbbHtHqq \rightarrow ZZgg \rightarrow ZZEWZ+X27.70.090.030.030.000.150.0071.53.060.013.2196.21.690.600.770.011.010.0098.111.60.3537.826.81.510.560.480.010.450.0125.33.020.6414.213.51.310.510.410.020.110.017.810.820.627.953.510.600.170.170.010.020.001.150.190.251.633.450.290.150.140.080.090.022.140.320.634.755.260.560.240.190.120.040.032.190.300.724.143.070.400.160.130.070.010.020.430.210.735.260.560.240.190.170.010.020.430.210.341.192.790.620.210.170.070.010.020.430.210.341.192.790.620.210.170.070.010.010.010.010.010.140.441.92.790.620.210.160.000.010.010.030.110.14</td><td>Image: SignalBackgrountExampleggHVBFWHZHtHbhHtH$q\bar{q} \rightarrow ZZ$$gg \rightarrow ZZ$EWZ+Xsignal27.70.090.030.030.000.150.0071.53.060.013.2127.9\pm0.196.21.690.600.770.011.010.0098.111.60.353.78100\pm026.81.510.560.480.010.450.0125.33.020.6414.229.8\pm0.113.51.310.510.410.020.110.017.810.820.627.9515.9\pm0.13.510.600.170.170.010.022.140.320.634.754.20\pm0.605.260.560.240.190.120.040.032.190.010.214.146.43\pm0.603.070.620.210.170.070.010.020.430.210.211.93.89\pm0.040.770.620.210.170.070.010.020.430.210.211.20.231.553.290.220.160.000.010.010.070.110.142.05\pm0.030.431.960.000.000.010.030.030.030.210.210.42.4419.3\pm1.10.553.290.220.160.000.01</td></t<></td></t<>	SignalVBFWHZHtHbbHtHqq \rightarrow ZZ27.70.090.030.000.150.0071.596.21.690.600.770.011.010.0098.126.81.510.560.480.010.450.0125.313.51.310.510.410.020.110.017.813.510.600.170.170.010.020.001.153.450.290.150.140.080.090.022.145.260.560.240.190.120.040.032.193.070.400.160.130.070.010.020.752.790.620.210.170.070.010.020.430.770.160.060.040.050.010.010.3415.53.290.220.160.000.010.000.190.431.960.000.000.000.000.072.522.350.060.060.030.030.050.620.440.790.010.010.010.000.010.032.480.940.130.060.030.010.032.160.570.030.090.060.030.010.050.040.010.530.040.260.100.030.032.160.540.01 <t< td=""><td>signalVBFWHZHtHb$\overline{b}H$tH$q \overline{q} \rightarrow ZZ$$gg \rightarrow ZZ$27.70.090.030.030.000.150.0071.53.0696.21.690.600.770.011.010.0098.111.626.81.510.560.480.010.450.0125.33.0213.51.310.510.410.020.110.017.810.823.510.600.170.170.010.020.001.150.193.450.290.150.140.080.090.022.140.325.260.560.240.190.120.040.032.190.303.070.400.160.130.070.010.020.430.210.770.160.060.040.050.010.010.340.0615.53.290.220.160.000.010.000.070.050.431.960.000.000.000.000.070.050.210.431.960.010.010.000.000.000.070.030.431.960.000.000.000.000.070.052.522.350.060.060.330.050.620.210.440.790.010.010.040.010.980.200.440.790.010.</td><td>SignalBackgrountggHVBFWHZHtHbHtH$q\bar{q} \rightarrow ZZ$$gg \rightarrow ZZ$EW27.70.090.030.030.000.150.0071.53.060.0196.21.690.600.770.011.010.0098.111.60.3526.81.510.560.480.010.450.0125.33.020.6413.51.310.510.410.020.110.017.810.820.623.510.600.170.170.010.020.001.150.190.253.450.290.150.140.080.090.022.140.320.635.260.560.240.190.120.040.032.190.300.723.070.400.160.130.070.010.020.430.210.213.790.620.210.170.070.010.020.430.060.311.553.290.220.160.000.010.030.070.050.120.431.960.000.000.000.000.070.050.120.431.960.000.000.000.000.070.050.120.431.960.010.010.000.010.030.030.040.440.790.110.010.01</td><td>SignalBackgroundggHVBFWHZHtHbbHtHqq \rightarrow ZZgg \rightarrow ZZEWZ+X27.70.090.030.030.000.150.0071.53.060.013.2196.21.690.600.770.011.010.0098.111.60.3537.826.81.510.560.480.010.450.0125.33.020.6414.213.51.310.510.410.020.110.017.810.820.627.953.510.600.170.170.010.020.001.150.190.251.633.450.290.150.140.080.090.022.140.320.634.755.260.560.240.190.120.040.032.190.300.724.143.070.400.160.130.070.010.020.430.210.735.260.560.240.190.170.010.020.430.210.341.192.790.620.210.170.070.010.020.430.210.341.192.790.620.210.170.070.010.010.010.010.010.140.441.92.790.620.210.160.000.010.010.030.110.14</td><td>Image: SignalBackgrountExampleggHVBFWHZHtHbhHtH$q\bar{q} \rightarrow ZZ$$gg \rightarrow ZZ$EWZ+Xsignal27.70.090.030.030.000.150.0071.53.060.013.2127.9\pm0.196.21.690.600.770.011.010.0098.111.60.353.78100\pm026.81.510.560.480.010.450.0125.33.020.6414.229.8\pm0.113.51.310.510.410.020.110.017.810.820.627.9515.9\pm0.13.510.600.170.170.010.022.140.320.634.754.20\pm0.605.260.560.240.190.120.040.032.190.010.214.146.43\pm0.603.070.620.210.170.070.010.020.430.210.211.93.89\pm0.040.770.620.210.170.070.010.020.430.210.211.20.231.553.290.220.160.000.010.010.070.110.142.05\pm0.030.431.960.000.000.010.030.030.030.210.210.42.4419.3\pm1.10.553.290.220.160.000.01</td></t<>	signalVBFWHZHtHb $\overline{b}H$ tH $q \overline{q} \rightarrow ZZ$ $gg \rightarrow ZZ$ 27.70.090.030.030.000.150.0071.53.0696.21.690.600.770.011.010.0098.111.626.81.510.560.480.010.450.0125.33.0213.51.310.510.410.020.110.017.810.823.510.600.170.170.010.020.001.150.193.450.290.150.140.080.090.022.140.325.260.560.240.190.120.040.032.190.303.070.400.160.130.070.010.020.430.210.770.160.060.040.050.010.010.340.0615.53.290.220.160.000.010.000.070.050.431.960.000.000.000.000.070.050.210.431.960.010.010.000.000.000.070.030.431.960.000.000.000.000.070.052.522.350.060.060.330.050.620.210.440.790.010.010.040.010.980.200.440.790.010.	SignalBackgrountggHVBFWHZHtHbHtH $q\bar{q} \rightarrow ZZ$ $gg \rightarrow ZZ$ EW27.70.090.030.030.000.150.0071.53.060.0196.21.690.600.770.011.010.0098.111.60.3526.81.510.560.480.010.450.0125.33.020.6413.51.310.510.410.020.110.017.810.820.623.510.600.170.170.010.020.001.150.190.253.450.290.150.140.080.090.022.140.320.635.260.560.240.190.120.040.032.190.300.723.070.400.160.130.070.010.020.430.210.213.790.620.210.170.070.010.020.430.060.311.553.290.220.160.000.010.030.070.050.120.431.960.000.000.000.000.070.050.120.431.960.000.000.000.000.070.050.120.431.960.010.010.000.010.030.030.040.440.790.110.010.01	SignalBackgroundggHVBFWHZHtHbbHtHqq \rightarrow ZZgg \rightarrow ZZEWZ+X27.70.090.030.030.000.150.0071.53.060.013.2196.21.690.600.770.011.010.0098.111.60.3537.826.81.510.560.480.010.450.0125.33.020.6414.213.51.310.510.410.020.110.017.810.820.627.953.510.600.170.170.010.020.001.150.190.251.633.450.290.150.140.080.090.022.140.320.634.755.260.560.240.190.120.040.032.190.300.724.143.070.400.160.130.070.010.020.430.210.735.260.560.240.190.170.010.020.430.210.341.192.790.620.210.170.070.010.020.430.210.341.192.790.620.210.170.070.010.010.010.010.010.140.441.92.790.620.210.160.000.010.010.030.110.14	Image: SignalBackgrountExampleggHVBFWHZHtHbhHtH $q\bar{q} \rightarrow ZZ$ $gg \rightarrow ZZ$ EWZ+Xsignal27.70.090.030.030.000.150.0071.53.060.013.2127.9 \pm 0.196.21.690.600.770.011.010.0098.111.60.353.78100 \pm 026.81.510.560.480.010.450.0125.33.020.6414.229.8 \pm 0.113.51.310.510.410.020.110.017.810.820.627.9515.9 \pm 0.13.510.600.170.170.010.022.140.320.634.754.20 \pm 0.605.260.560.240.190.120.040.032.190.010.214.146.43 \pm 0.603.070.620.210.170.070.010.020.430.210.211.93.89 \pm 0.040.770.620.210.170.070.010.020.430.210.211.20.231.553.290.220.160.000.010.010.070.110.142.05 \pm 0.030.431.960.000.000.010.030.030.030.210.210.42.4419.3 \pm 1.10.553.290.220.160.000.01



Background estimation

- In most cases, we rely on MC to model signals
- For bkg, if the MC prediction is reliable, use MC
 - Otherwise, e.g. QCD, DY+jets.. data driven methods
- Data-driven methods usually combines data and MC
 - Control region (CR): orthogonal to signal region, where bkg is dominant, trust data in CR
 - Transfer factor: from CR to signal region (SR), relies on MC or some other data (ABCD method)
 - Requires a lot of validations!





Systematics

- MC is not reality
- Theoretical uncertainty
 - QCD scale, EW correction, PDF, parton shower, hadronization
- Experimental uncertainty
 - Luminosity
 - Object energy/momentum scale/resolution
 - Object reconstruction/identification efficiency
 - Data driven method
- Yield/Shape uncertainty, log-Normal nuisances



Statistical interpretation

- Precision measurement / Search
- Roofit/RooStats
- Tool in CMS: Higgs-CombinedLimit package
 - Parameter estimate, test statistics
 - Maximum log likelihood, chi2
 - 68% CL, 95% CL
 - Significance, p-value
 - Upper limit









To start in CMS

- To get familiar with CMS: CMS induction course, link
- To know what's being published: CMS public results, link
- To get a deeper understanding of an analysis: internal CADI line, link
- To get reminders of all internal communication on certain physics/object groups, subscribe to:
 - CMS hyper news, link
 - CMS Talk, <u>link</u>

