$H \rightarrow WW \rightarrow 2l2\nu$ in 0 and 1-jet Final States at CMS

Jae Hyeok Yoo(UC San Diego)
jaehyeokyoo@gmail.com
on behalf of CMS collaboration

DPF 2013, Santa Cruz, USA, Aug 13-17, 2013

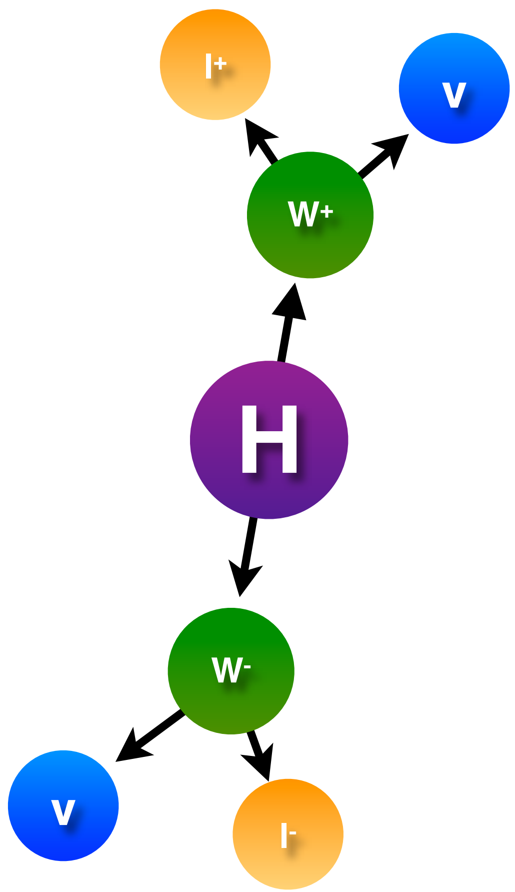
Overview of talk

- The $H \to WW \to 2l2\nu$: signal and background
- Signal extraction
- Background estimation
- Fit validation for shape analysis
- Search results
- Spin-parity hypothesis test
- Summary

The H→WW→2l2ν channel

What is different from other modes?

- No mass peak due to neutrinos in the final state
  - measure overall excess on top of backgrounds
  - very important to understand backgrounds
  - measure signal strength at the measured M_H



from ZZ→4l and γγ

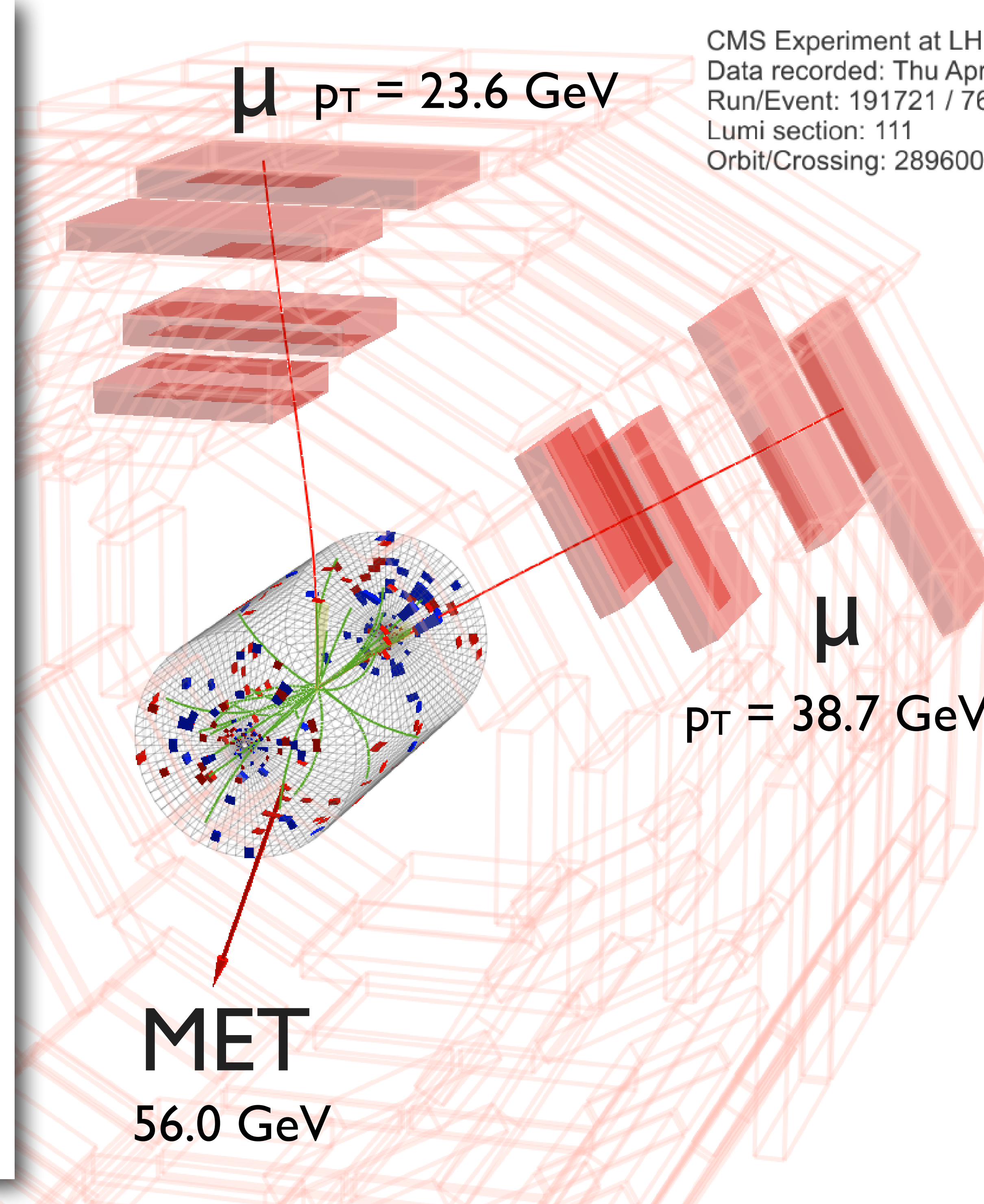
- Large signal yields
  - good statistical power to measure signal strength
  - This channel measures the signal strength with the best precision with current data

The H→WW→2l2ν channel

Signature and analysis strategy

- Signature
  - Two energetic, identified/isolated, opposite-sign leptons (e or μ)
  - large missing transverse energy(MET)
- Background composition depends on
  - number of jets : 0 and 1
  - lepton flavor : ee/μμ and eμ
- Analysis optimized in 4 categories

0-jet ee/μμ	0-jet eμ
1-jet ee/μμ	1-jet eμ

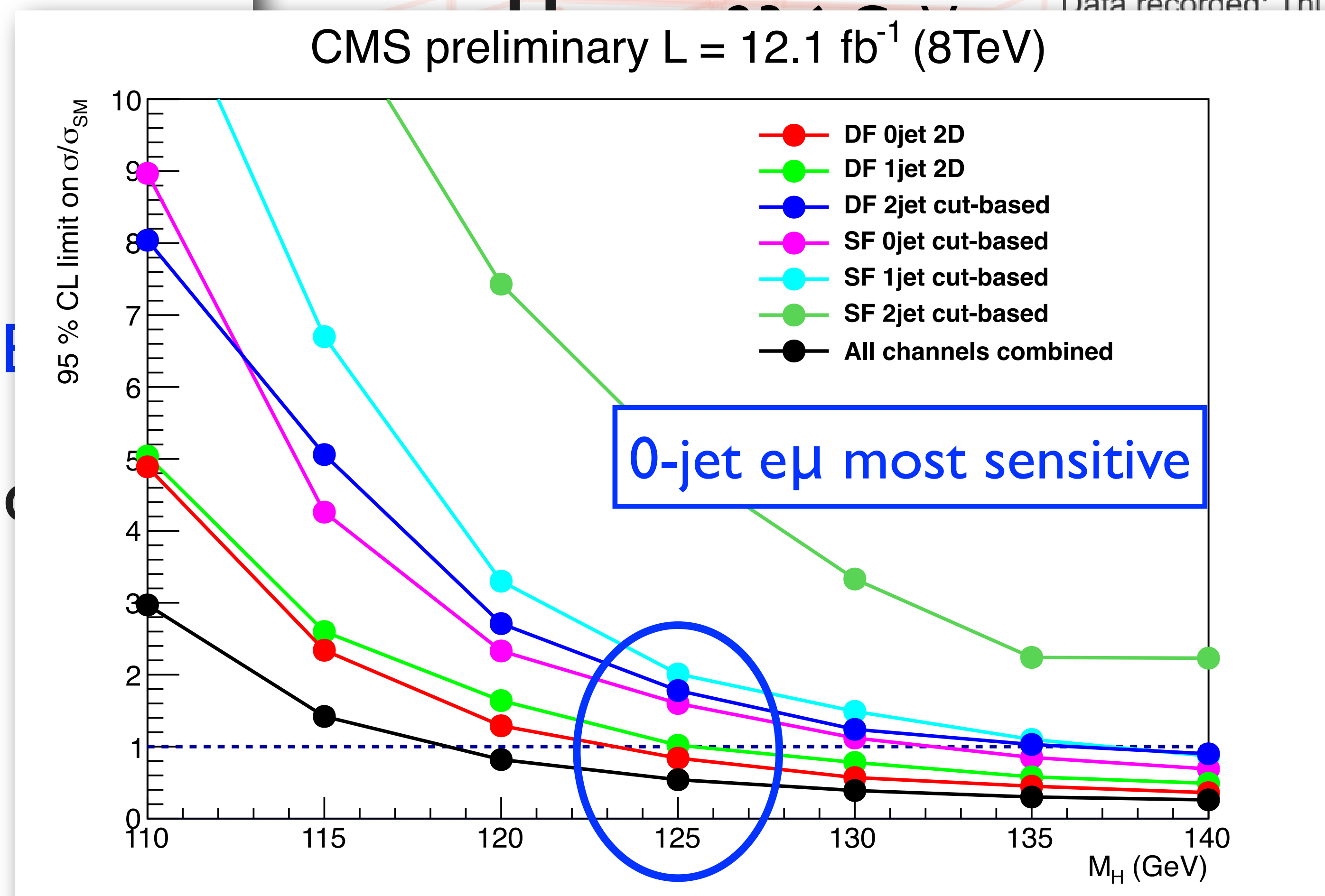


The H→WW→2l2ν channel

Signature and analysis strategy

- Signature
  - Two energetic, identified/isolated, opposite-sign leptons (e or μ)
  - large missing transverse energy(ME
- Background composition depen
  - number of jets : 0 and 1
  - lepton flavor : ee/μμ and eμ
- Analysis optimized in 4 categories

0-jet ee/μμ	0-jet eμ
1-jet ee/μμ	1-jet eμ

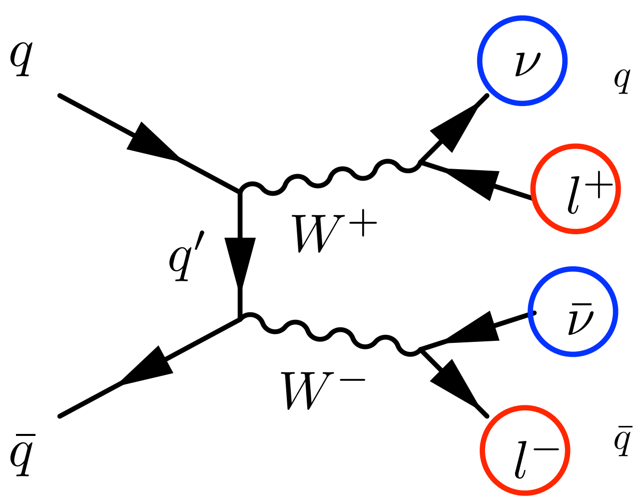


Backgrounds

How they fake signal and how to suppress them

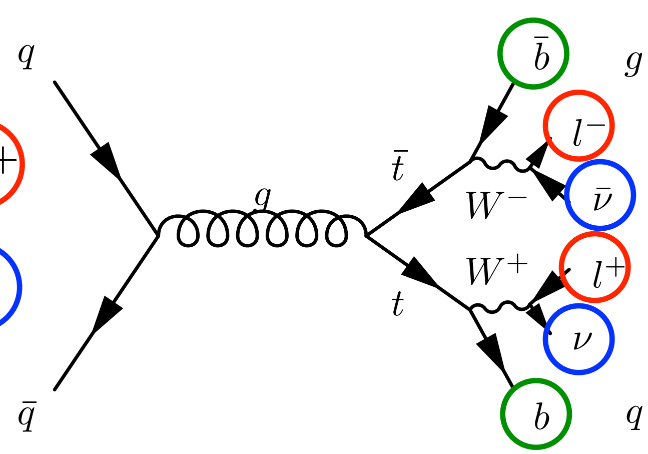
Color code : two opposite-sign leptons MET handles for suppression

WW



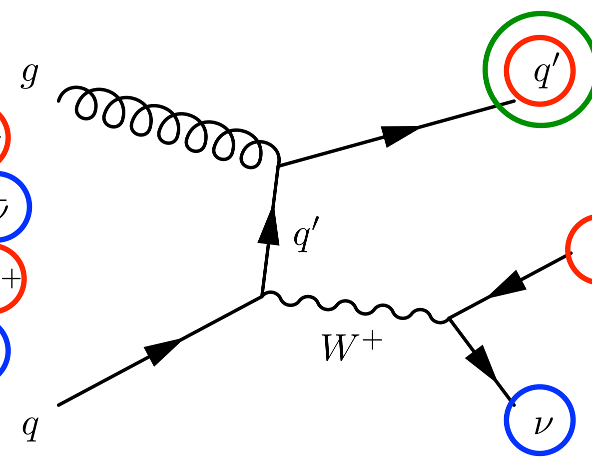
Irreducible
Use difference in kinematics

Top



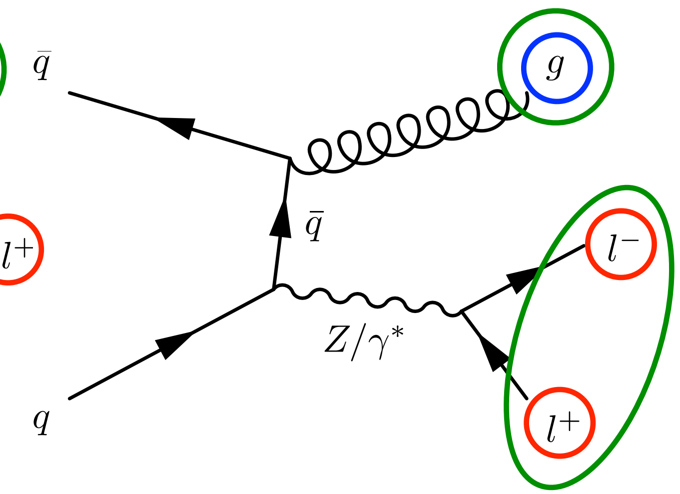
B-tagging
Soft-muon tagging
Number of jets

W+jets



Tight lepton ID & ISO
Tight $p_T(ll)$ cut

Drell-Yan



Z mass veto
Tight MET cut
(MVA technique)

Wγ

conversion rejection

WZ/ZZ/Wγ*

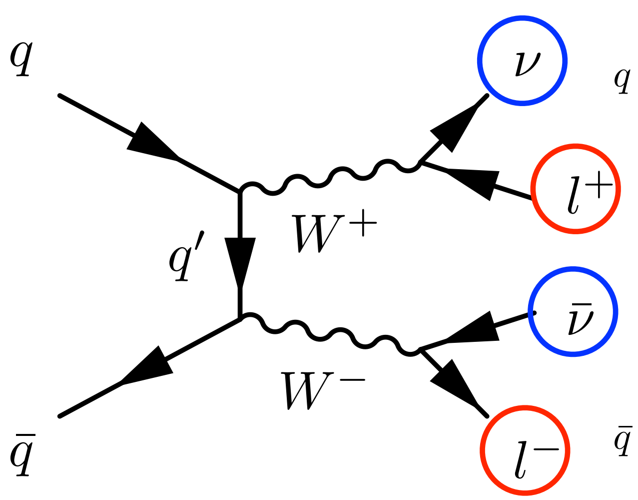
extra lepton veto

Backgrounds

How they fake signal and how to suppress them

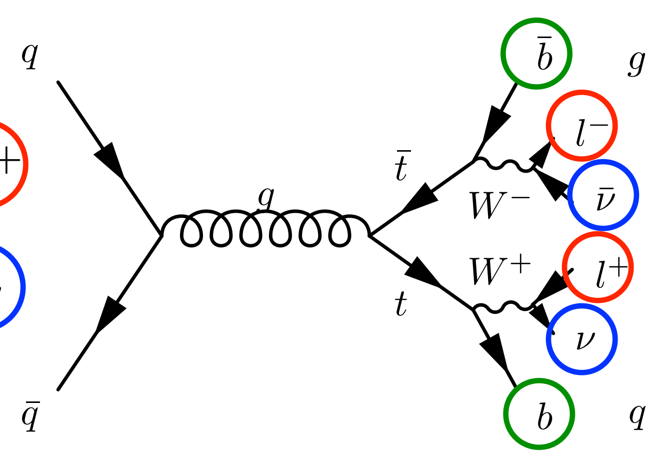
Color code : two opposite-sign leptons MET handles for suppression

WW



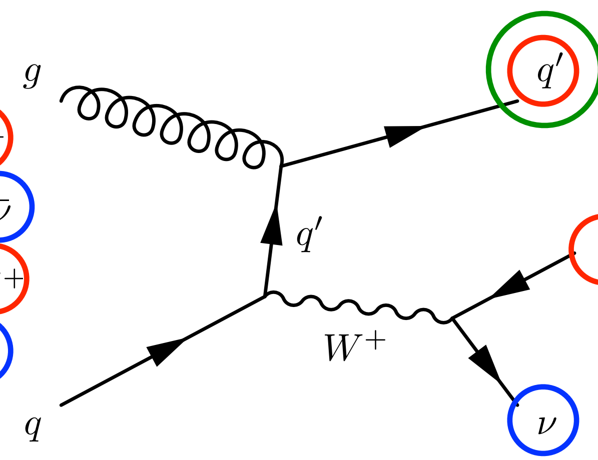
Irreducible
Use difference in kinematics

Top



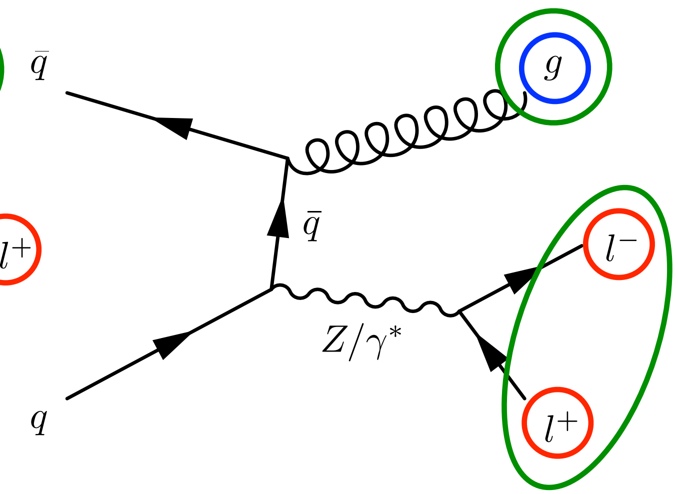
B-tagging
Soft-muon tagging
Number of jets

W+jets



Tight lepton ID & ISO
Tight $p_T(ll)$ cut

Drell-Yan



Z mass veto
Tight MET cut
(MVA technique)

Wγ

conversion rejection

WZ/ZZ/Wγ*

extra lepton veto

WW selection

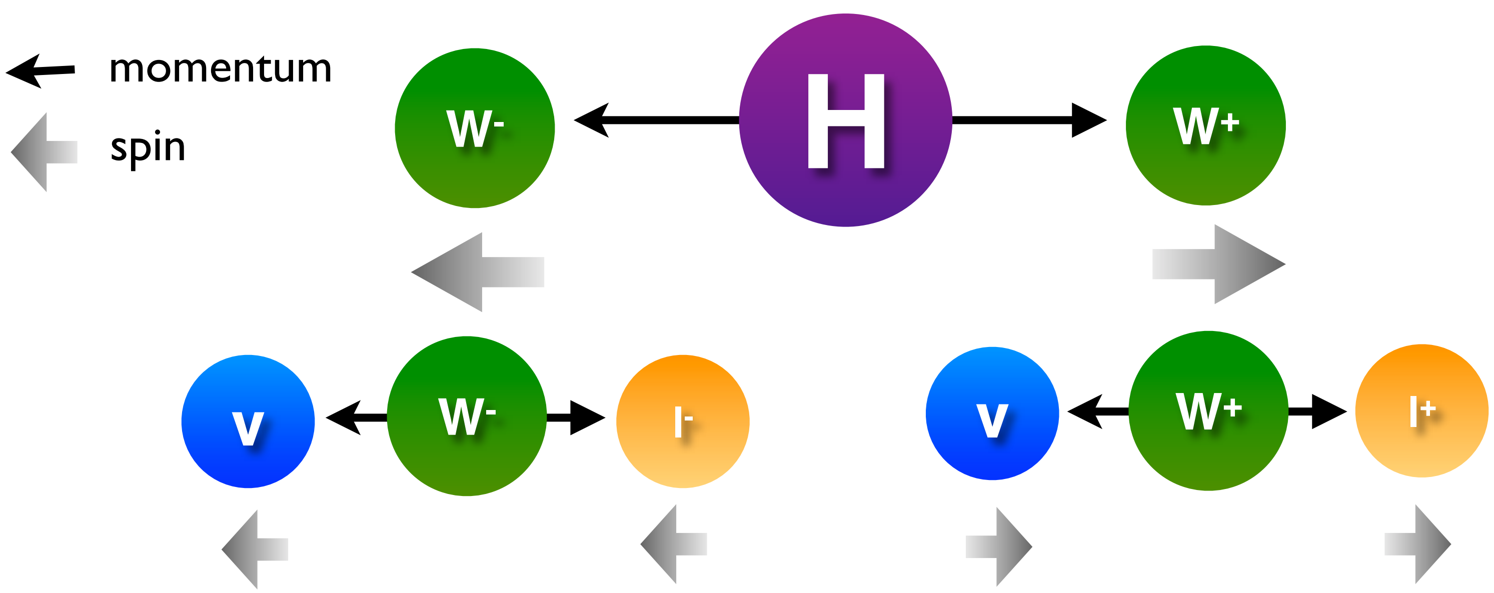
Signal Extraction

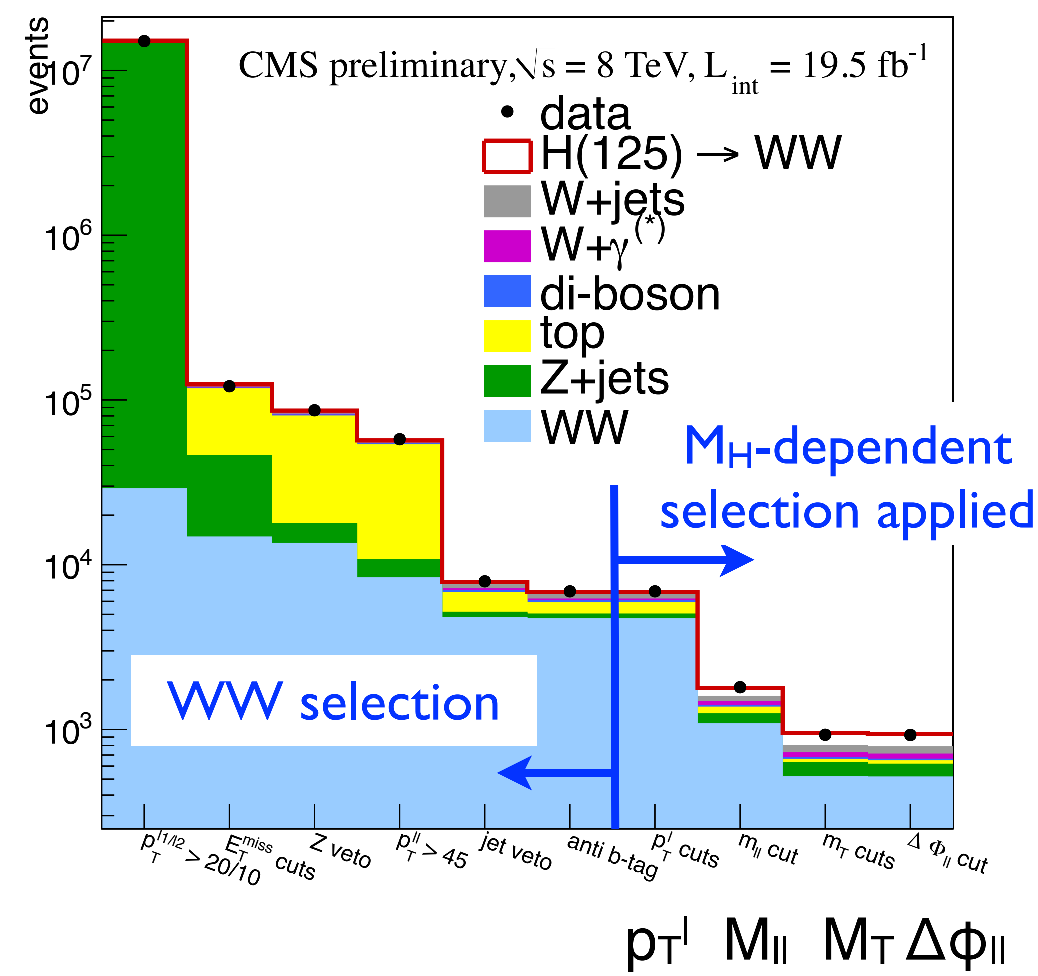
How to extract signal yields : cut-based

- Baseline selection to reject backgrounds : WW selection
- Two approaches : cut-based(ee/μμ/eμ) and shape-based(eμ)

Cut-based

- M_H-dependent selection taking advantage of event kinematic difference due to helicity conservation





- Low M_H : small $\Delta\phi_{ll}$, small M_{ll}

Signal Extraction

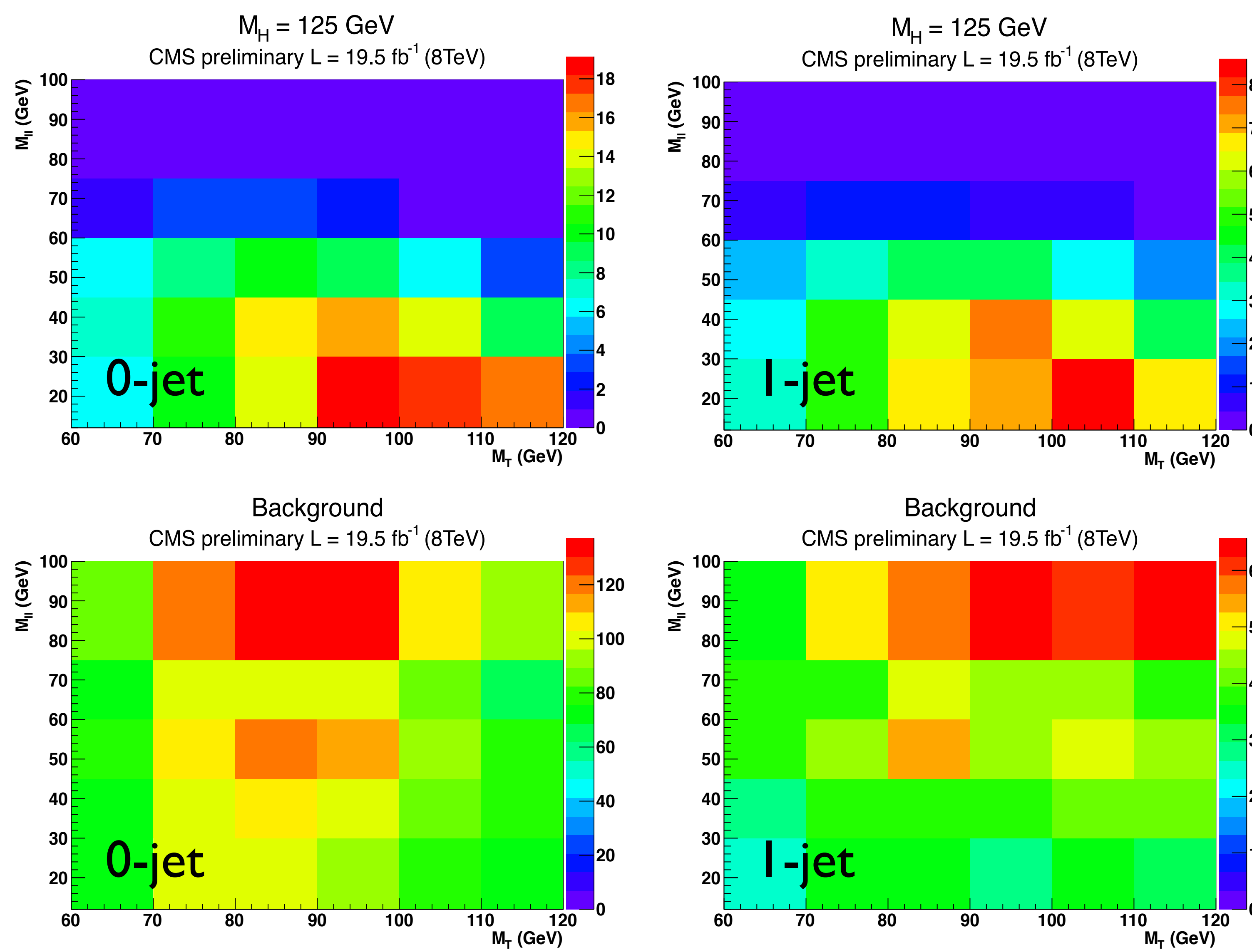
How to extract signal yields : shape-based

Shape-based

- Use binned 2D templates of $[M_T, M_{ll}]$ and fit the full shape

$$M_T = \sqrt{2p_T^{ll} \cdot \mathrm{MET} \cdot (1 - cos(\Delta\phi_{ll-\mathrm{MET}}))}$$

- Applied to only eμ channel
- Two templates : for low (<300 GeV) and high(≥300 GeV) Higgs mass
- Large signal-free region to constrain backgrounds : especially WW in 0-jet
- More sensitive than cut-based



Low M_H templates zoomed in signal-populated region
full range : $60<M_T<280$GeV, $12<M_{ll}<200$ GeV

Background Estimation

Overview of background estimation

will be discussed

Background	Method
WW	Data-driven
Top	Data-driven
Drell-Yan	Data-driven
W+jets	Data-driven
Wγ*	Data-driven
Wγ	from MC
WZ/ZZ	from MC

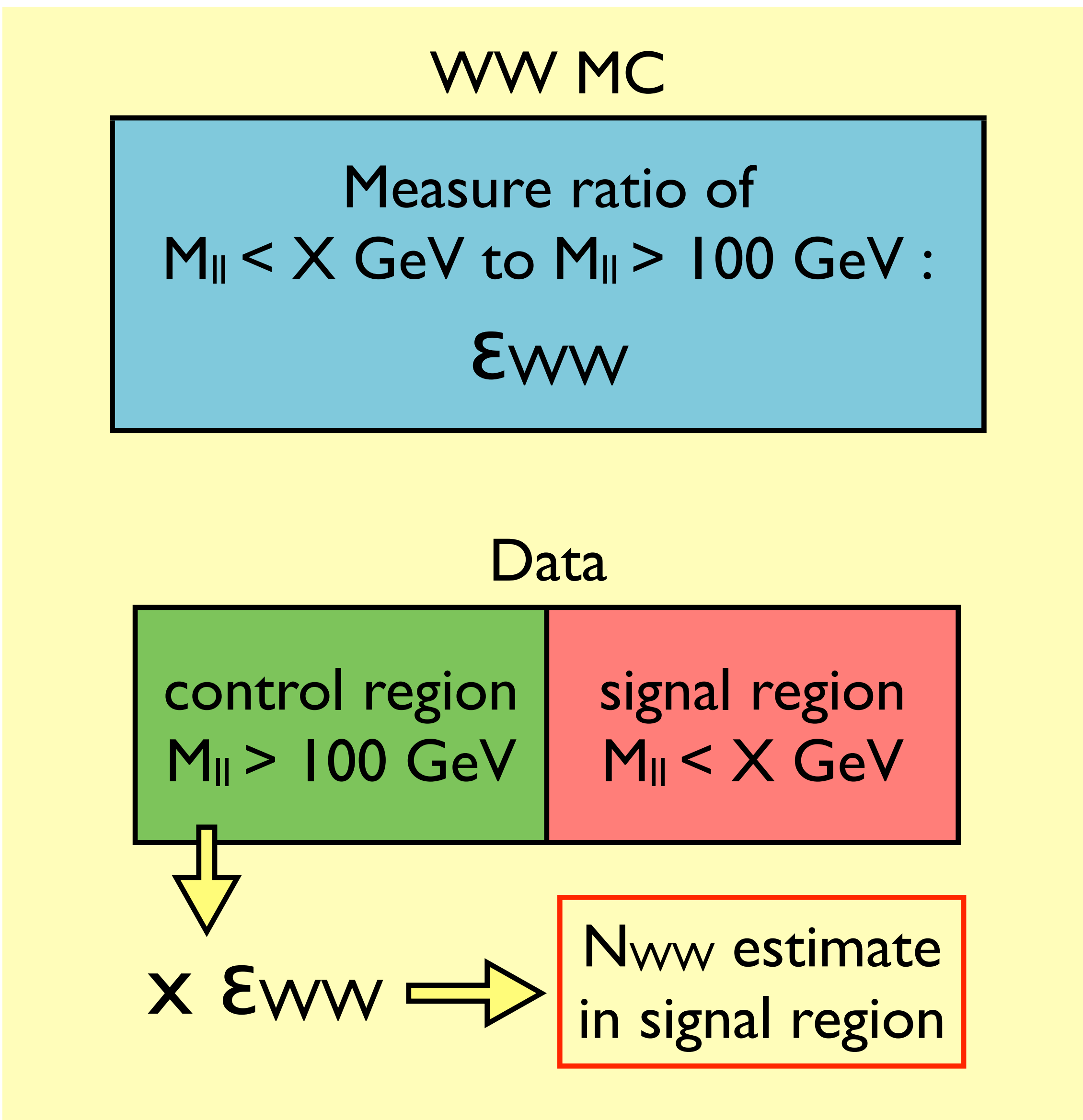
- WW selection applied
- Data-driven methods for dominant backgrounds
  - Measure the ratio(ε) of yields in signal region(SR) to control region(CR) in an independent sample(data or MC), and apply ε to CR

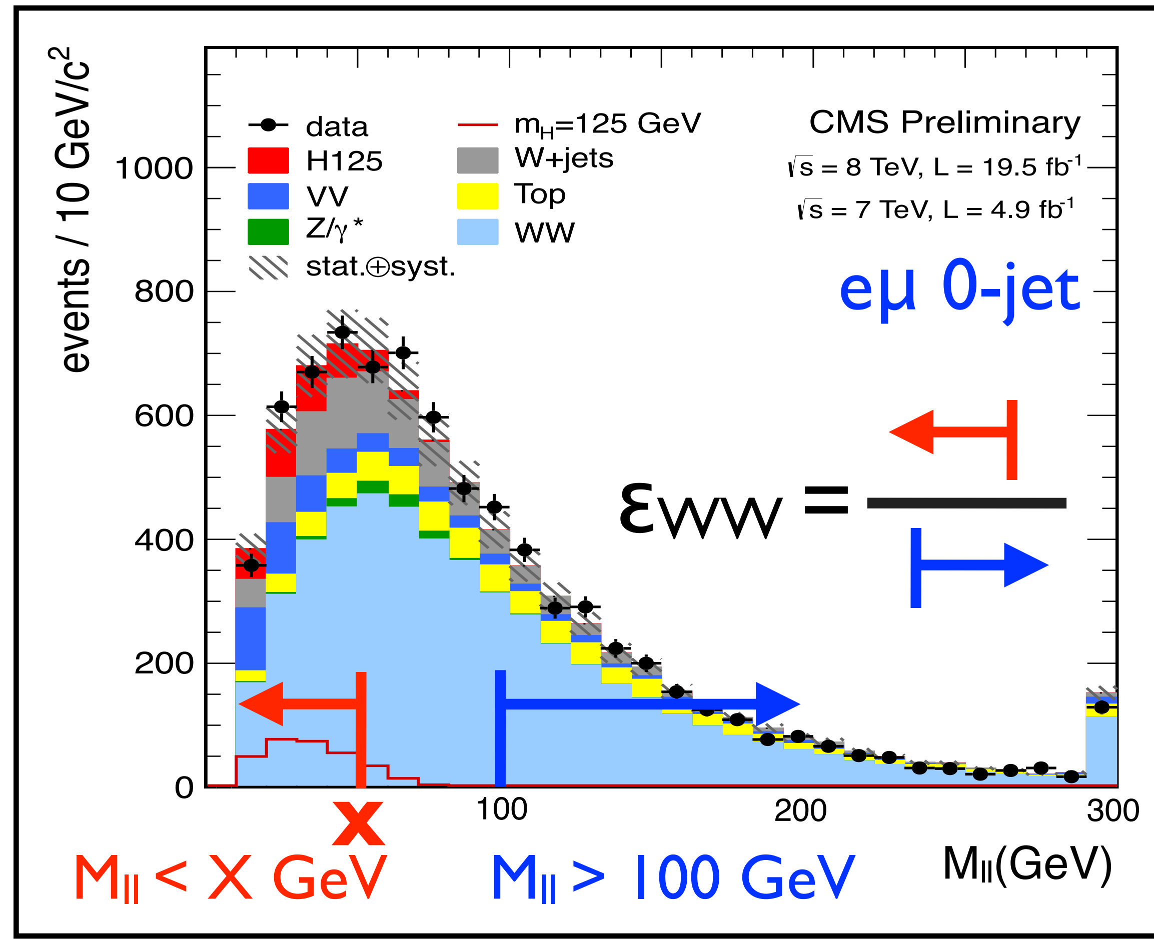
$$N_{SR} = N_{CR} \times \varepsilon$$

- Others are taken from MC

WW Estimation

- Main background in the 0-jet category
- Cut-based analysis : extrapolation from high M_{ll} to low M_{ll} region



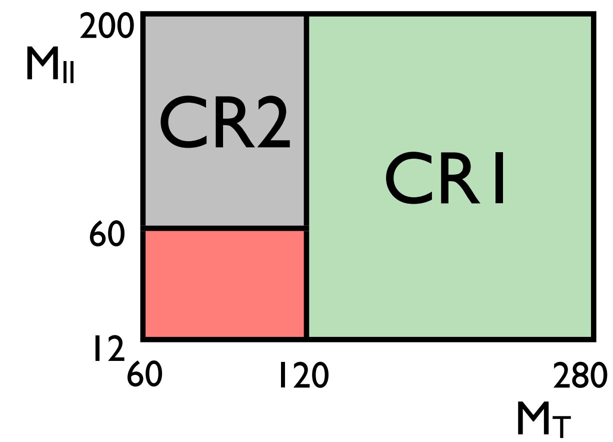


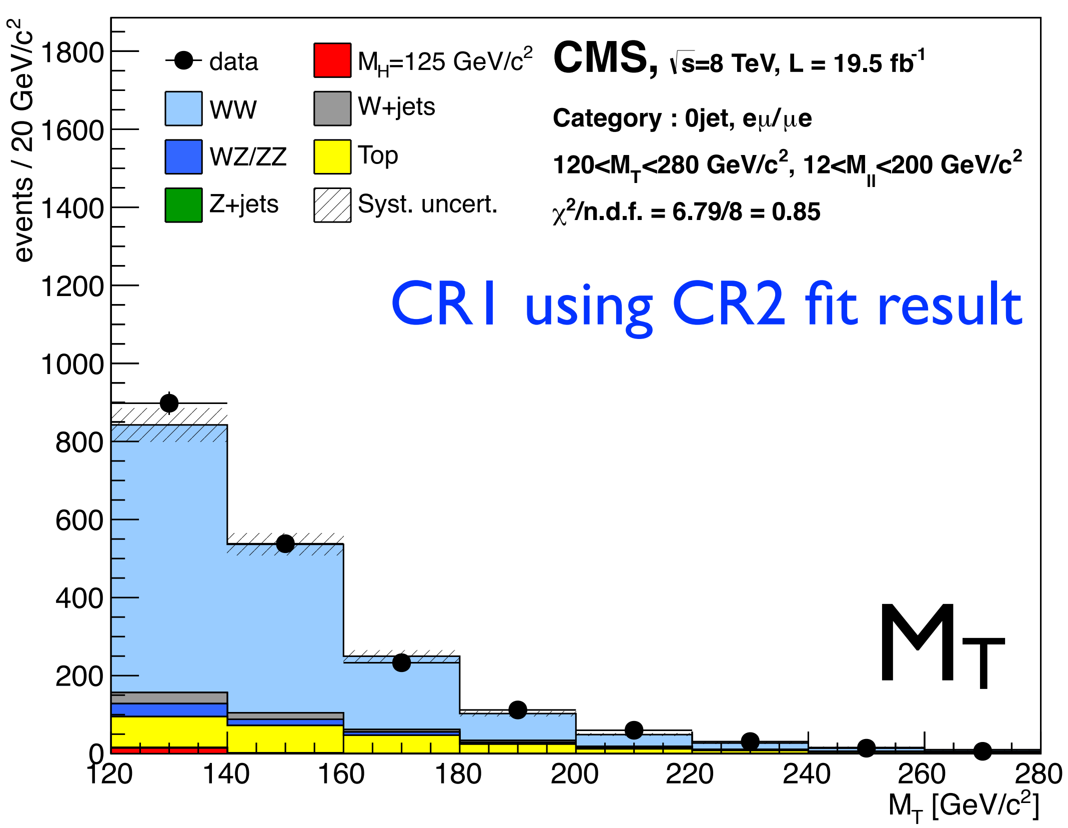
- Shape-based analysis : Data/MC in whole M_{ll} region is taken → fit is able to constrain WW using high M_{ll}, high M_T regions

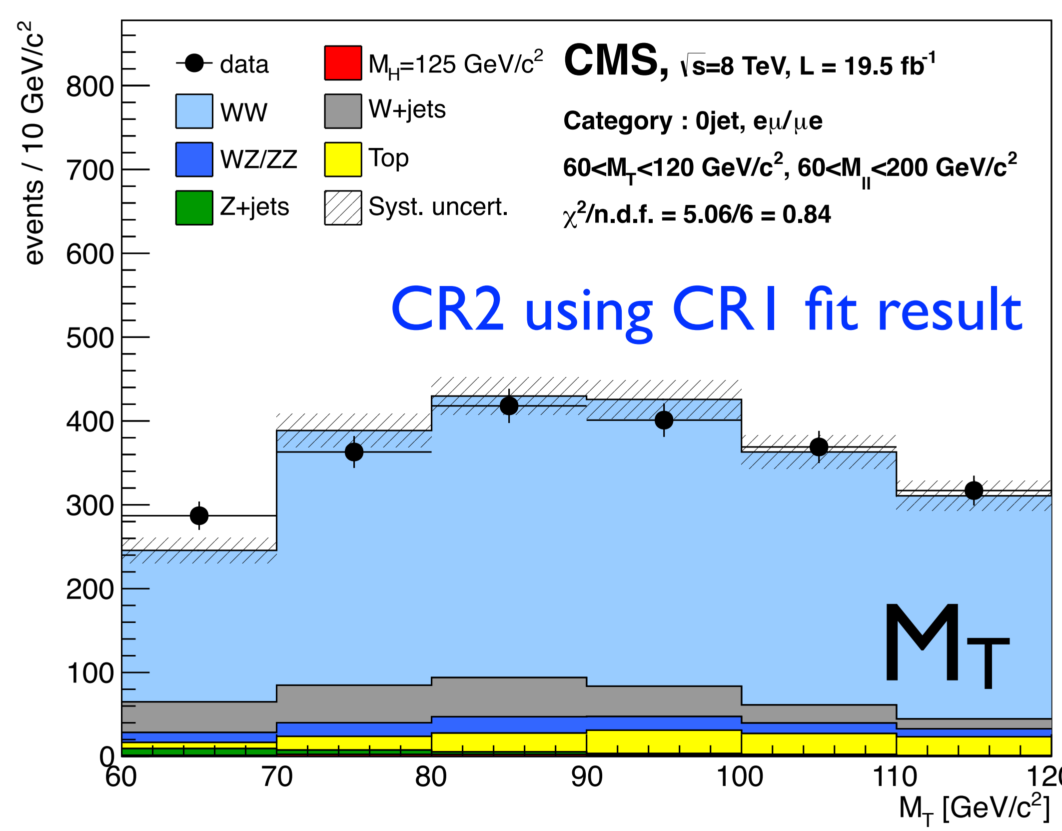
Fit Validation in Data

Is the WW fit model correct ?

- Need to make sure fit model fits data correctly : WW, Top, W+jets, Wγ(*), ...
- WW template is taken from MC normalized by data-driven estimation and shapes are allowed to move to match data in the fit
- Test WW fit model using WW sideband in eμ 0-jet
  - Divide signal-free region into two control regions (CR1 and CR2)
  - Predict CR1(2) from the fit result using only CR2(1)
  - All other backgrounds are fixed by nominal fit to test only WW



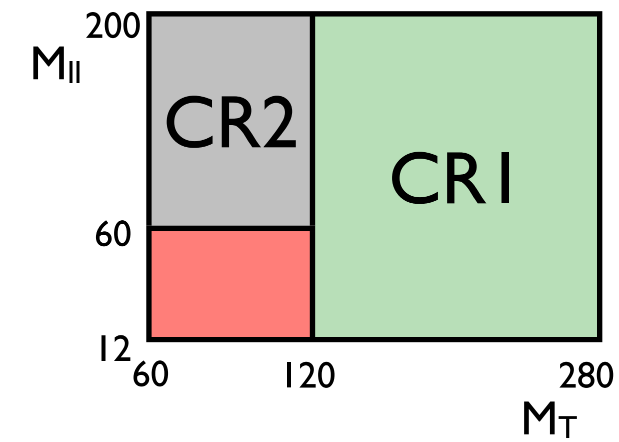


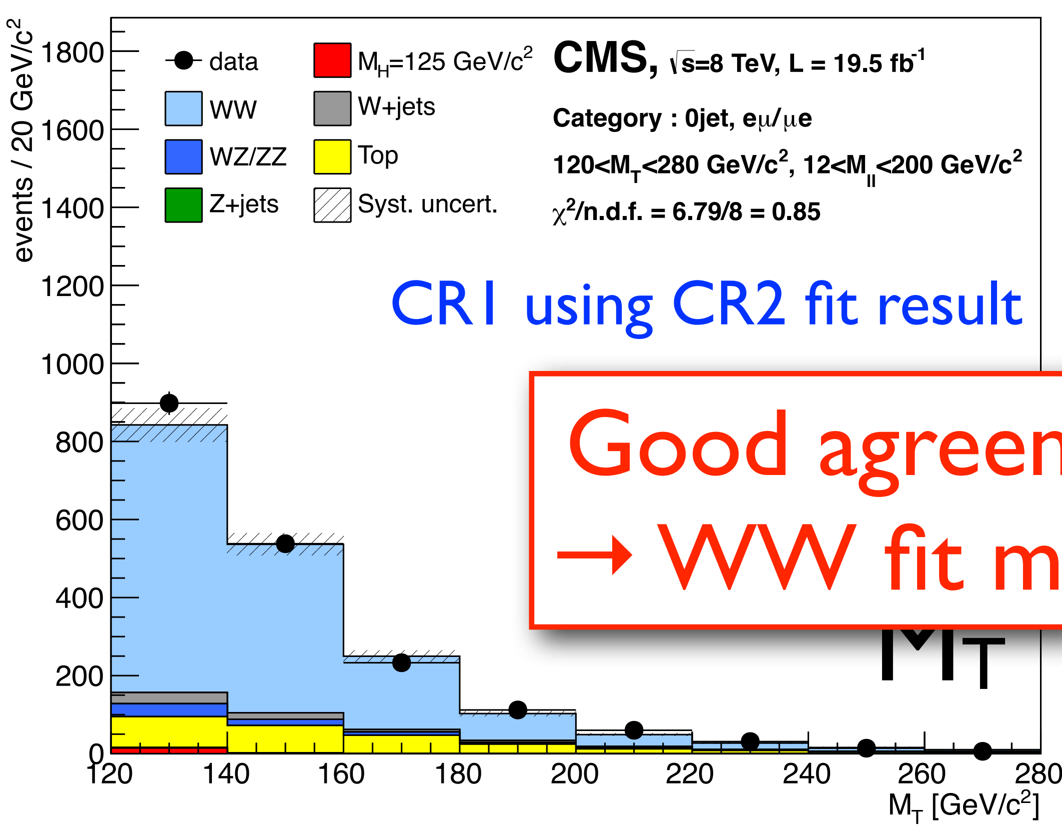


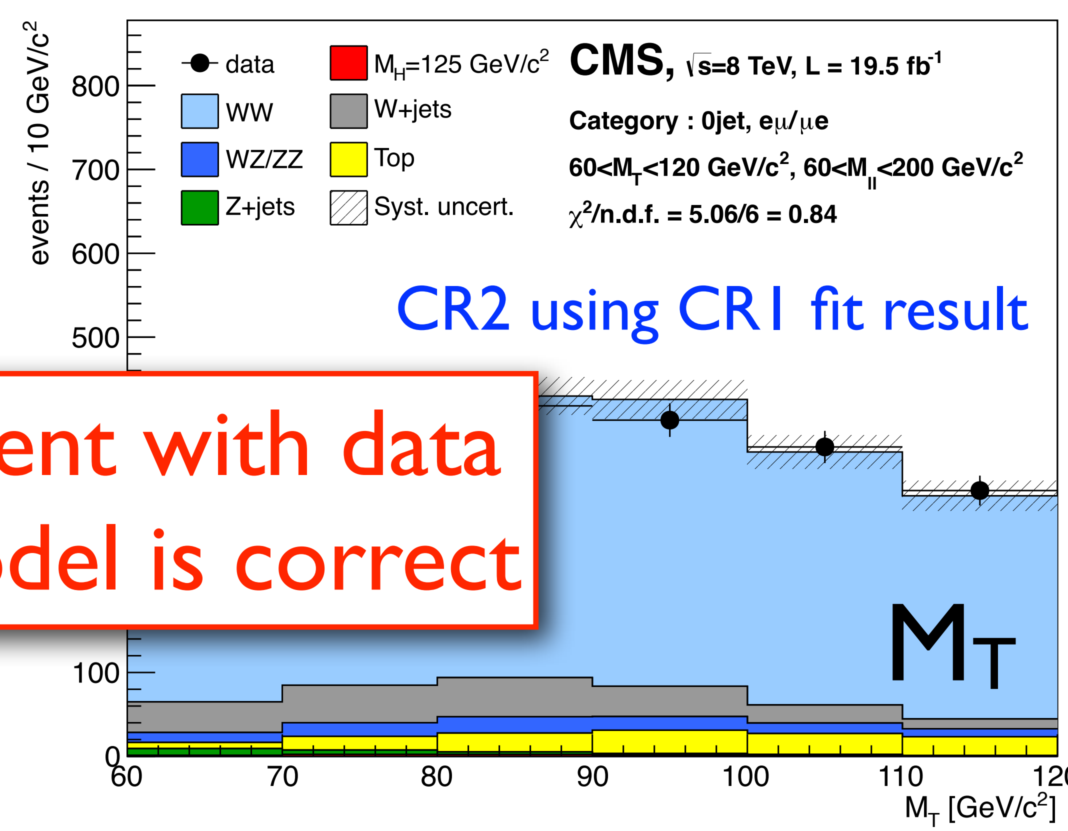
Fit Validation in Data

Is the WW fit model correct ?

- Need to make sure fit model fits data correctly : WW, Top, W+jets, Wγ(*), ...
- WW template is taken from MC normalized by data-driven estimation and shapes are allowed to move to match data in the fit
- Test WW fit model using WW sideband in eμ 0-jet
  - Divide signal-free region into two control regions (CR1 and CR2)
  - Predict CR1(2) from the fit result using only CR2(1)
  - All other backgrounds are fixed by nominal fit to test only WW







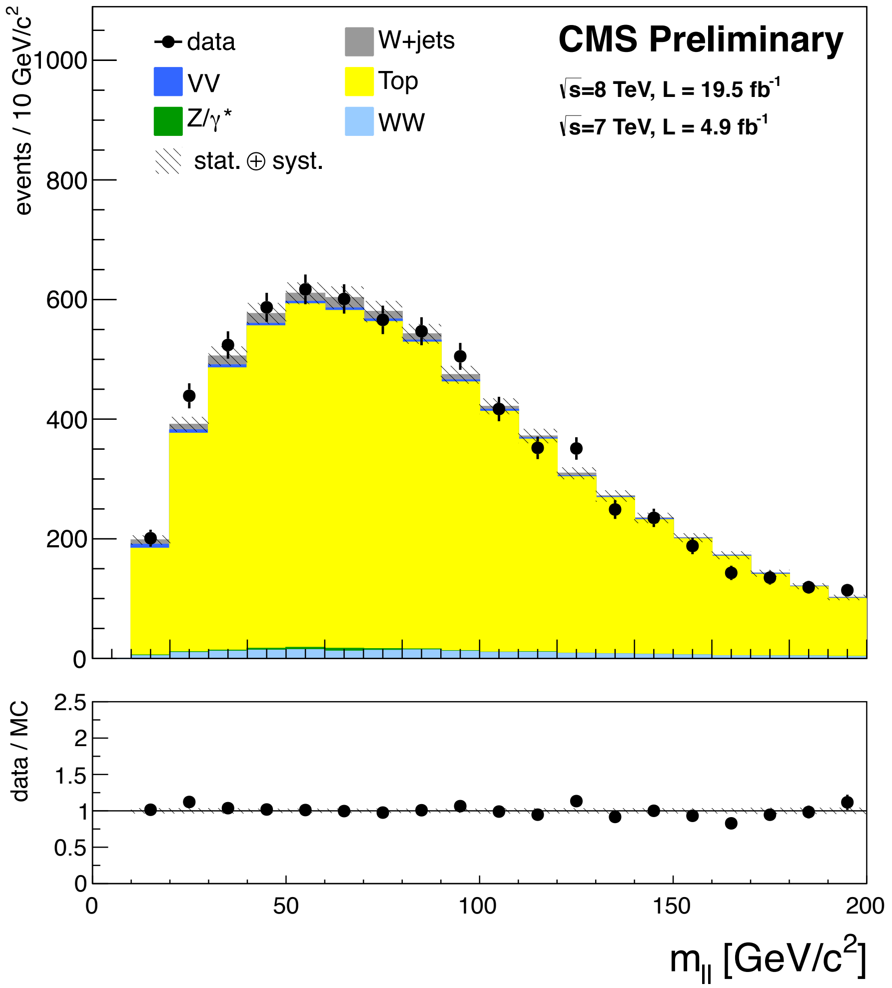
Good agreement with data
→ WW fit model is correct

Fit Validation in Data

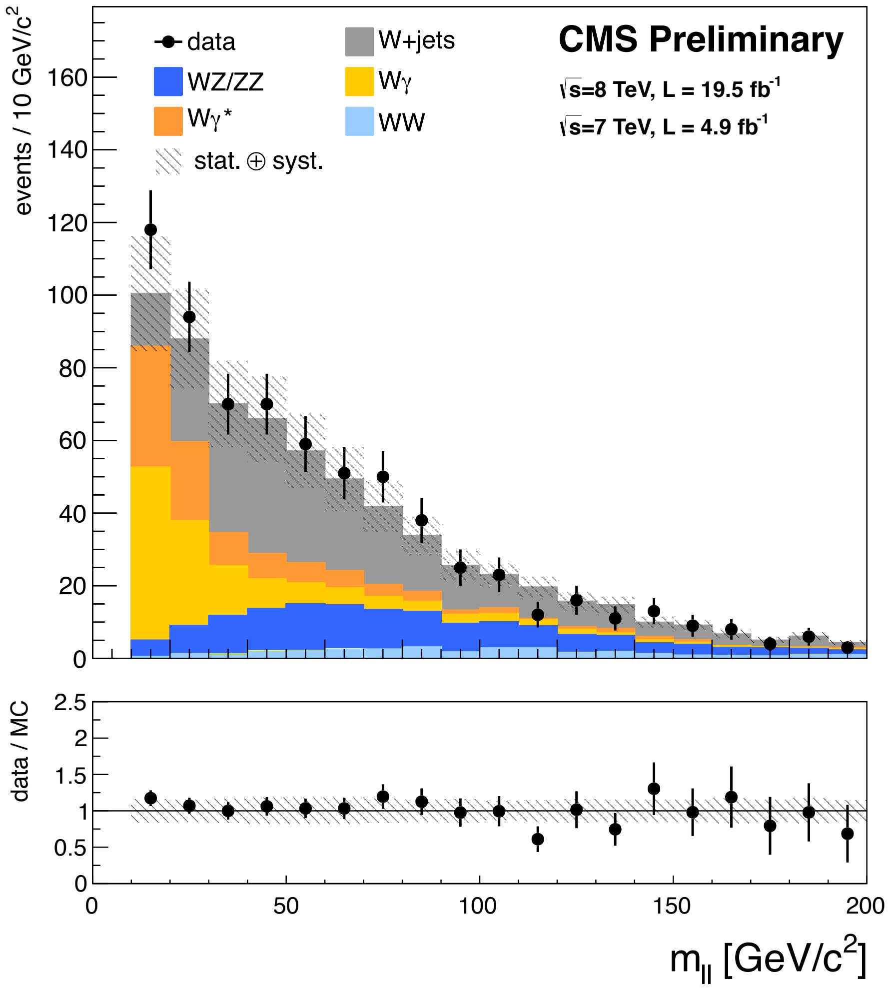
Are the Top and W+jets/Wγ(*) models correct?

- Fit two control regions populated by Top and W+jets/Wγ(*)
- Same selections as 2D analysis except for inverting top-veto and opposite-sign requirements

Top-tagged events eμ 1-jet



Same-sign events eμ 0-jet

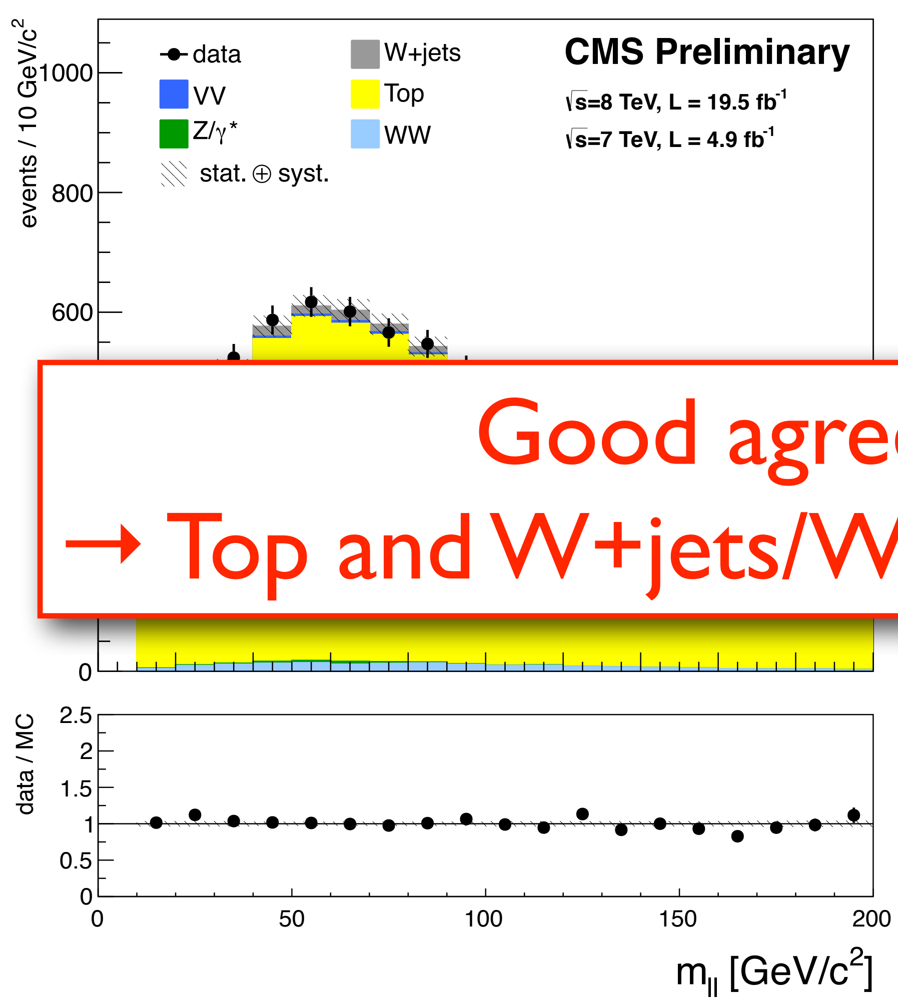


Fit Validation in Data

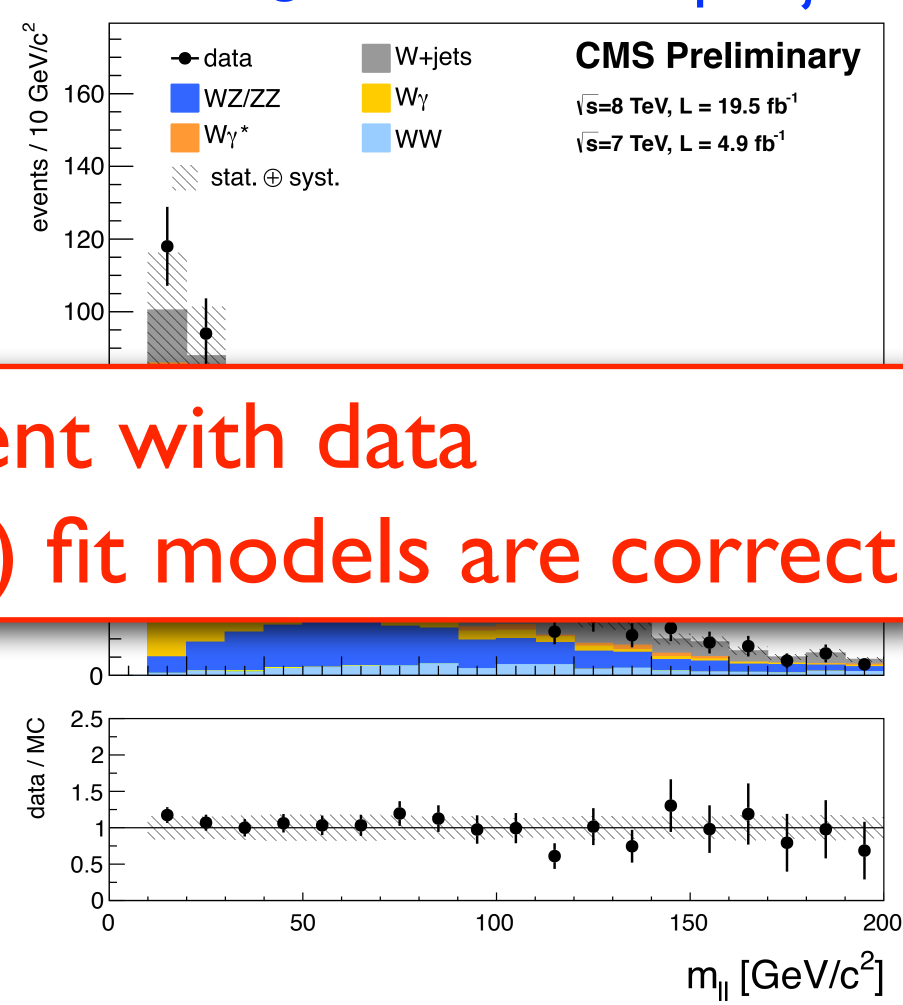
Are the Top and W+jets/Wγ(*) models correct?

- Fit two control regions populated by Top and W+jets/Wγ(*)
- Same selections as 2D analysis except for inverting top-veto and opposite-sign requirements

Top-tagged events eμ 1-jet



Same-sign events eμ 0-jet



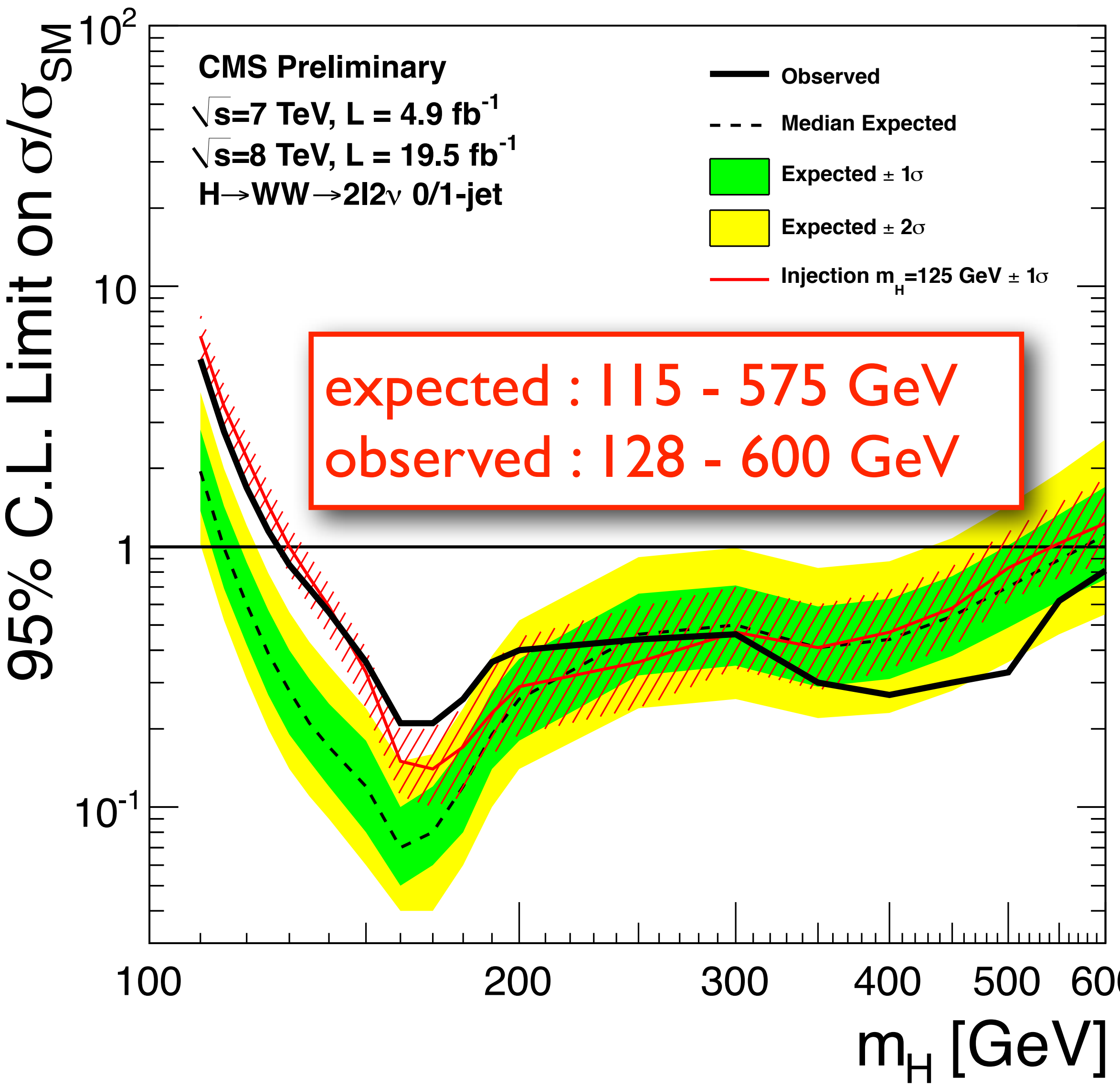
Good agreement with data
→ Top and W+jets/Wγ(*) fit models are correct

Search Results

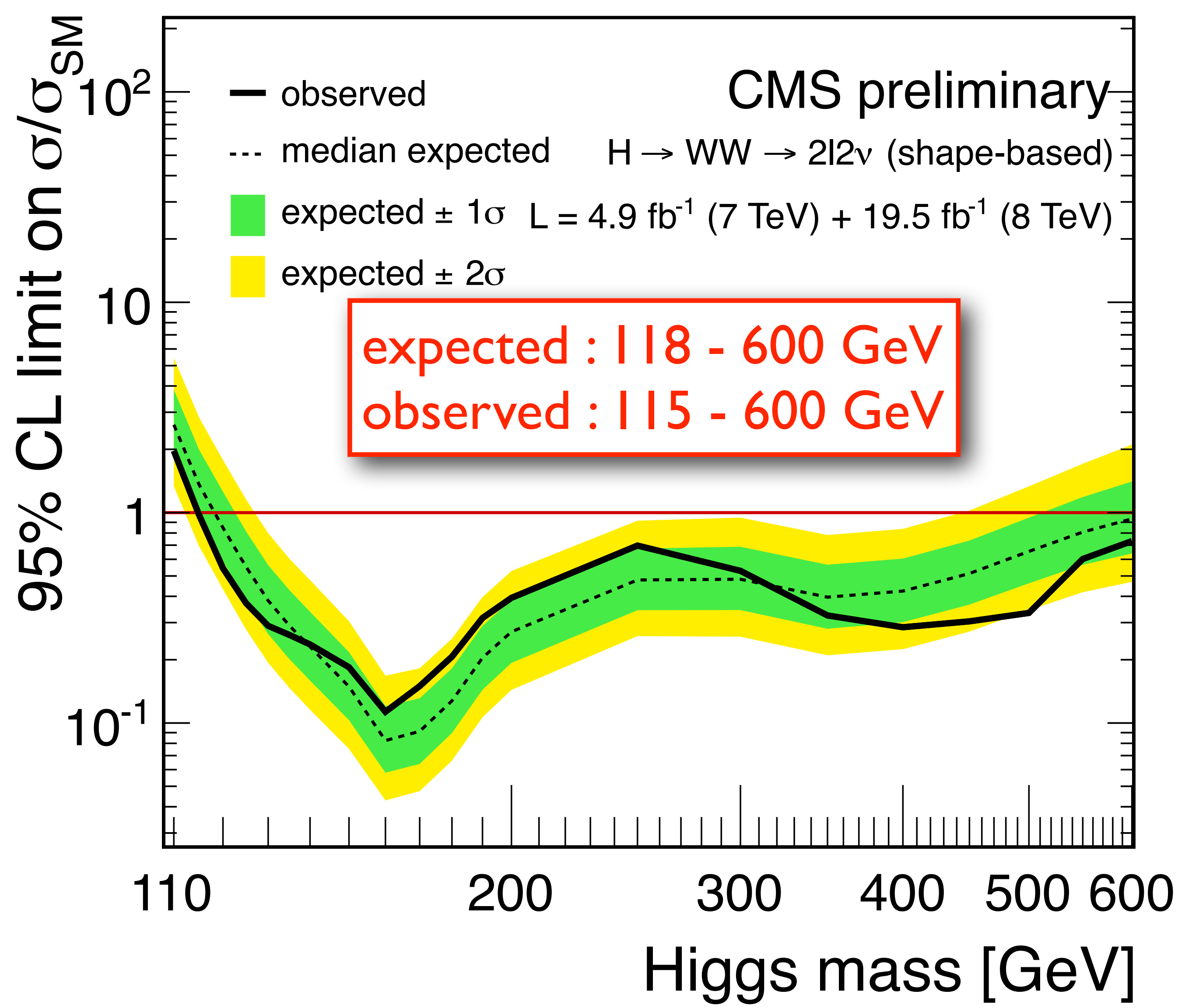
Exclusion : Compatible with SM Higgs hypothesis?

2D method is used in eμ channel and cut-based method is used in ee/μμ channel

SM Higgs exclusion



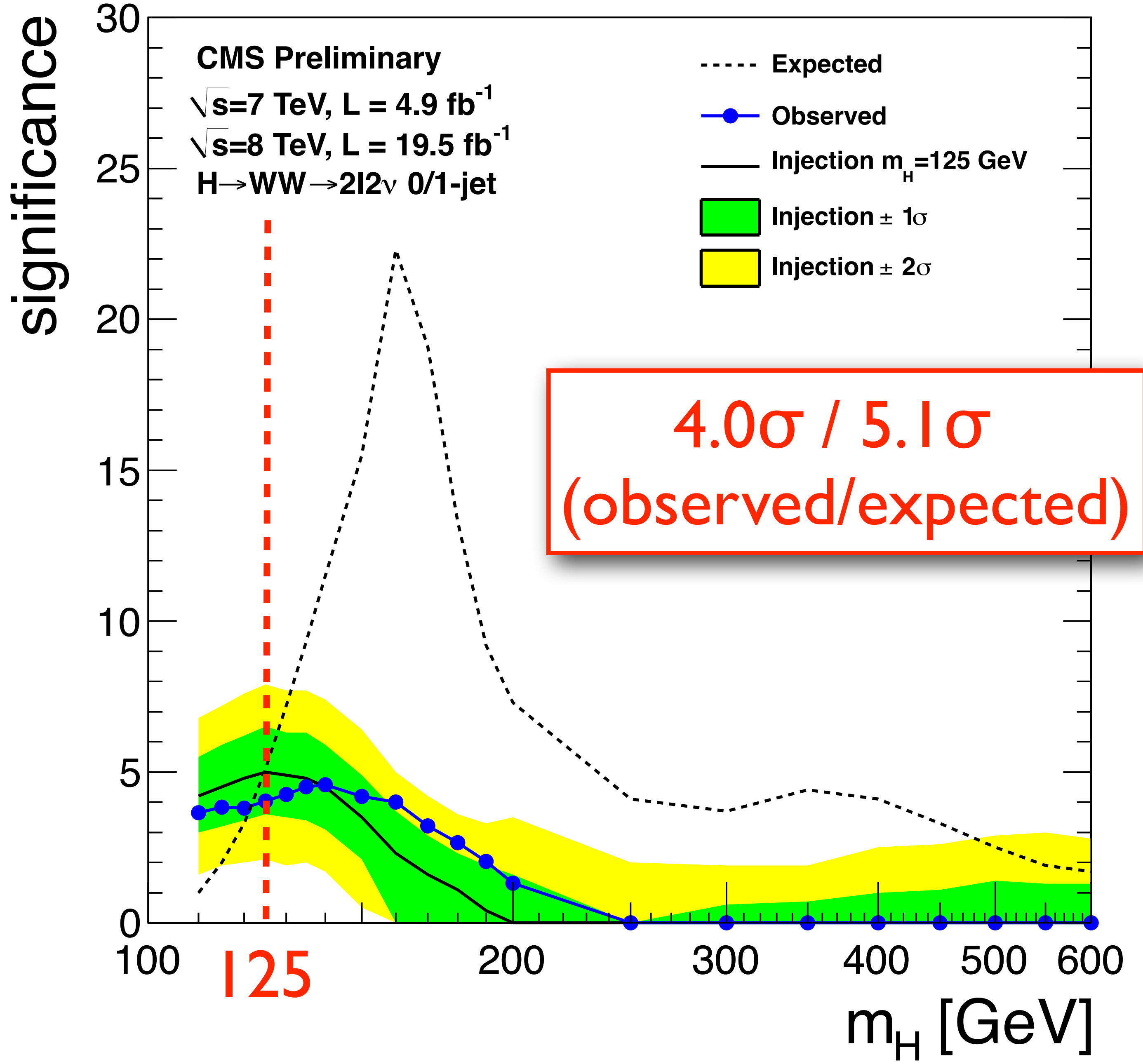
Second SM-like Higgs exclusion



Search Results

Significance : Compatible with bkgd-only hypothesis?

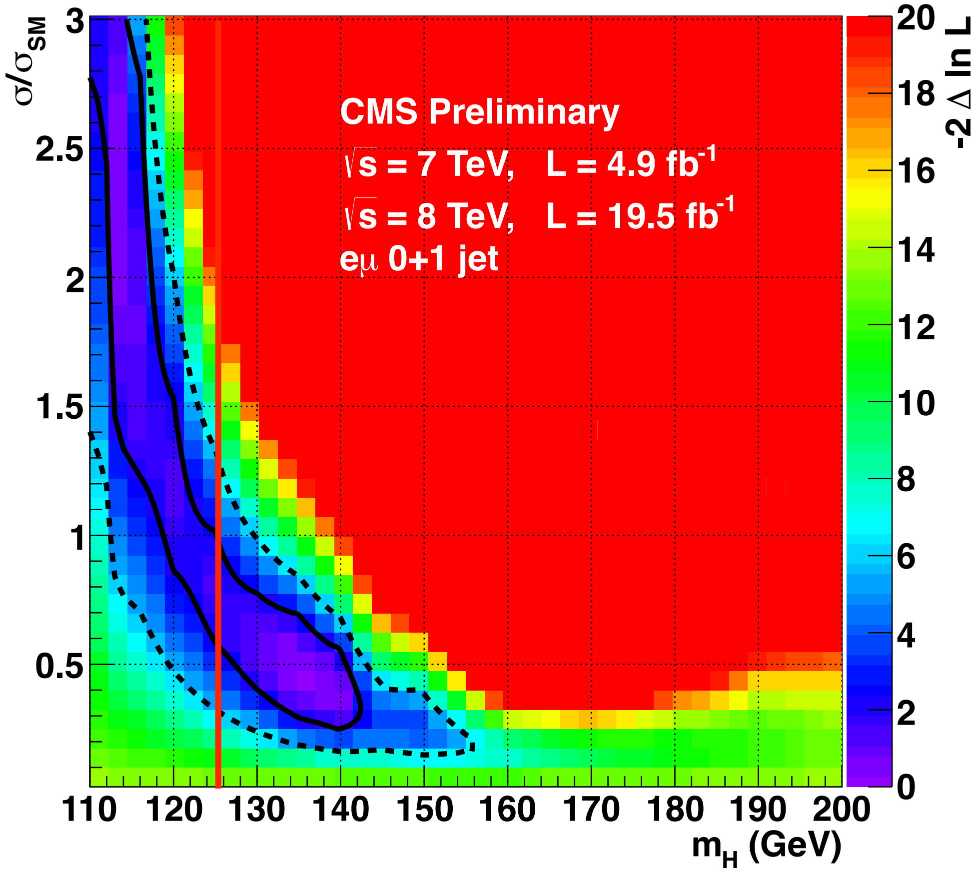
2D method is used in eμ channel and cut-based method is used in ee/μμ channel



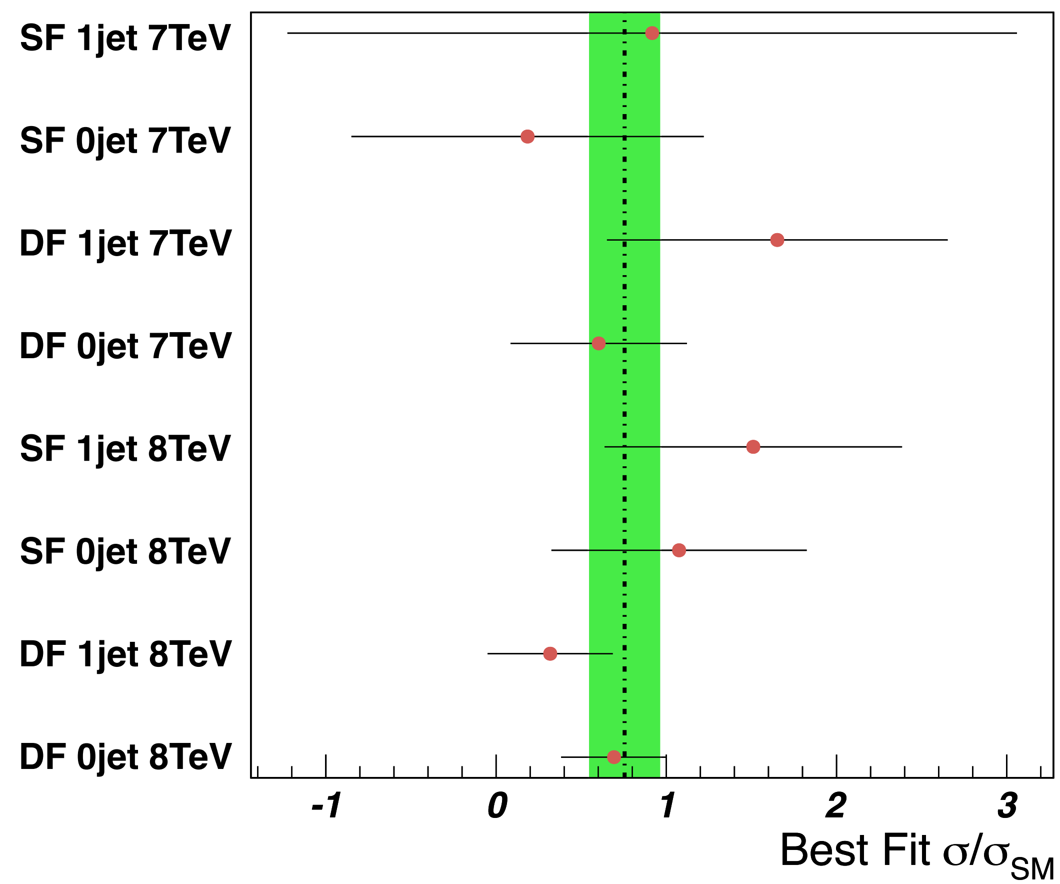
Search Results

Is signal strength consistent with SM Higgs?

Confidence intervals in (M_H,μ) plane



Signal strength(μ) in each channel



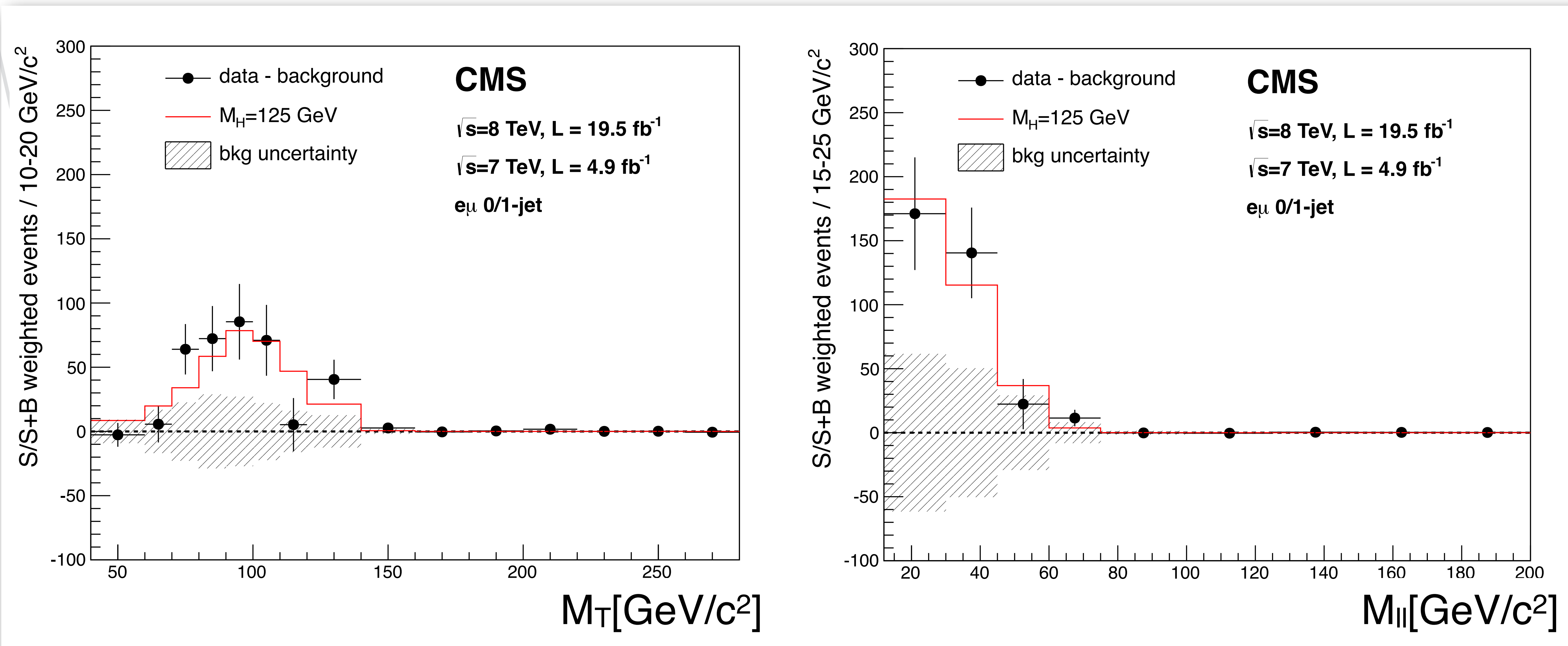
Signal strength (μ)	
μ = 0.76 ± 0.13 (stat.) ± 0.16 (syst.) = 0.76 ± 0.21 (stat.+syst.)	μ(H→γγ) = 0.78 ± 0.27 (stat.+syst.) μ(H→ZZ→4l) = $0.91^{+0.30}_{-0.24}$ (stat.+syst.)

Search Results

Is "Data − background" consistent with SM Higgs?

- Data - background plots in 0/1-jets eμ with S/(S+B) weighting
  - S/(S+B) weighting at each bin of 2D template
  - Post-fit normalization and uncertainties

Important plots to show consistency of data with SM Higgs

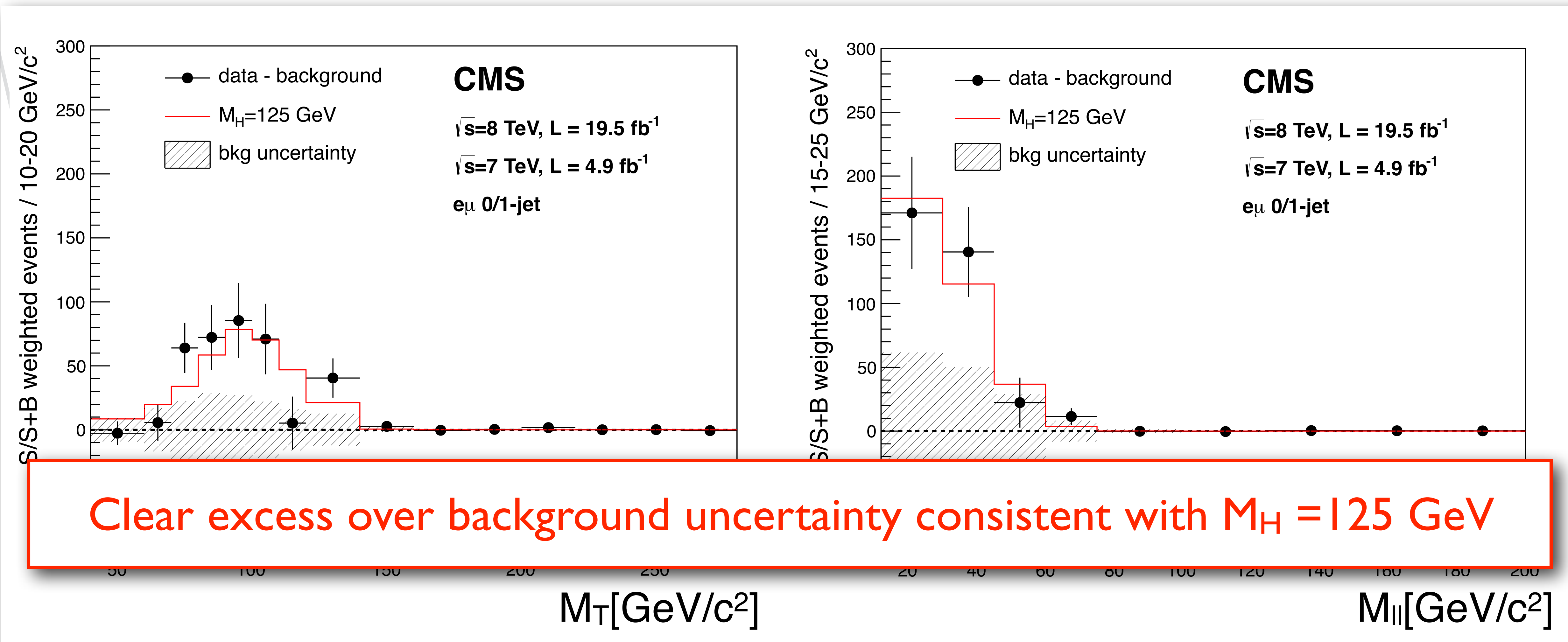


Search Results

Is "Data − background" consistent with SM Higgs?

- Data - background plots in 0/1-jets eμ with S/(S+B) weighting
  - S/(S+B) weighting at each bin of 2D template
  - Post-fit normalization and uncertainties

Important plots to show consistency of data with SM Higgs



Clear excess over background uncertainty consistent with M_H =125 GeV

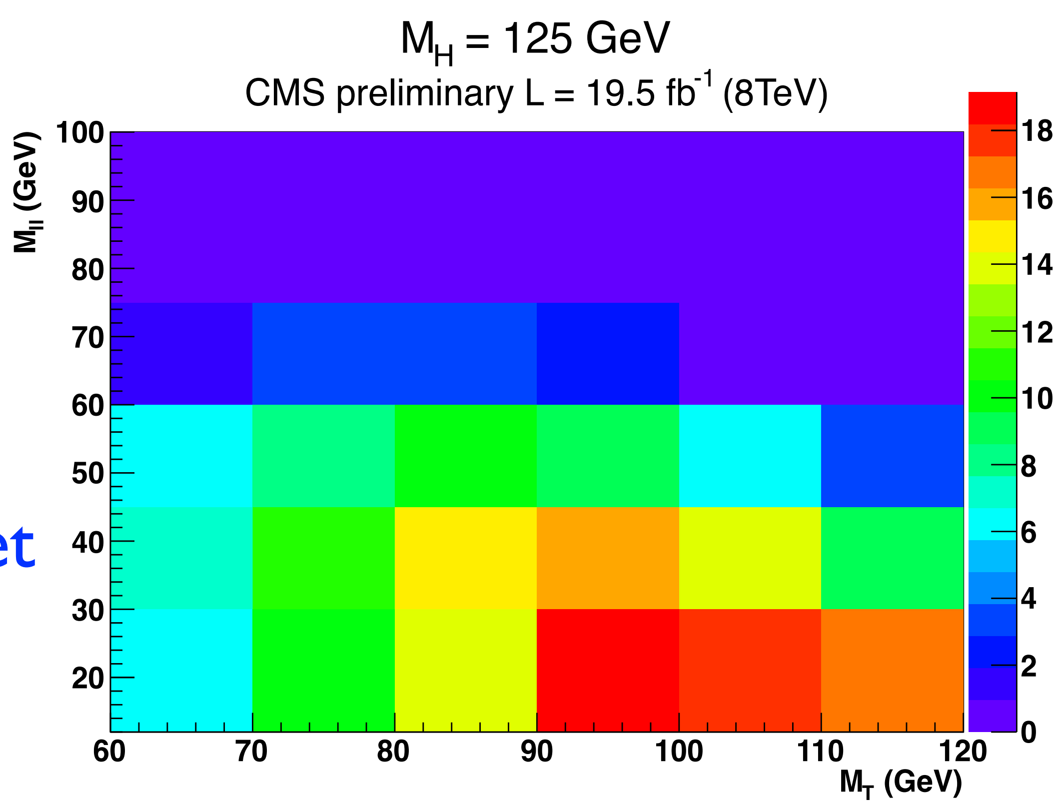
Spin-parity Test

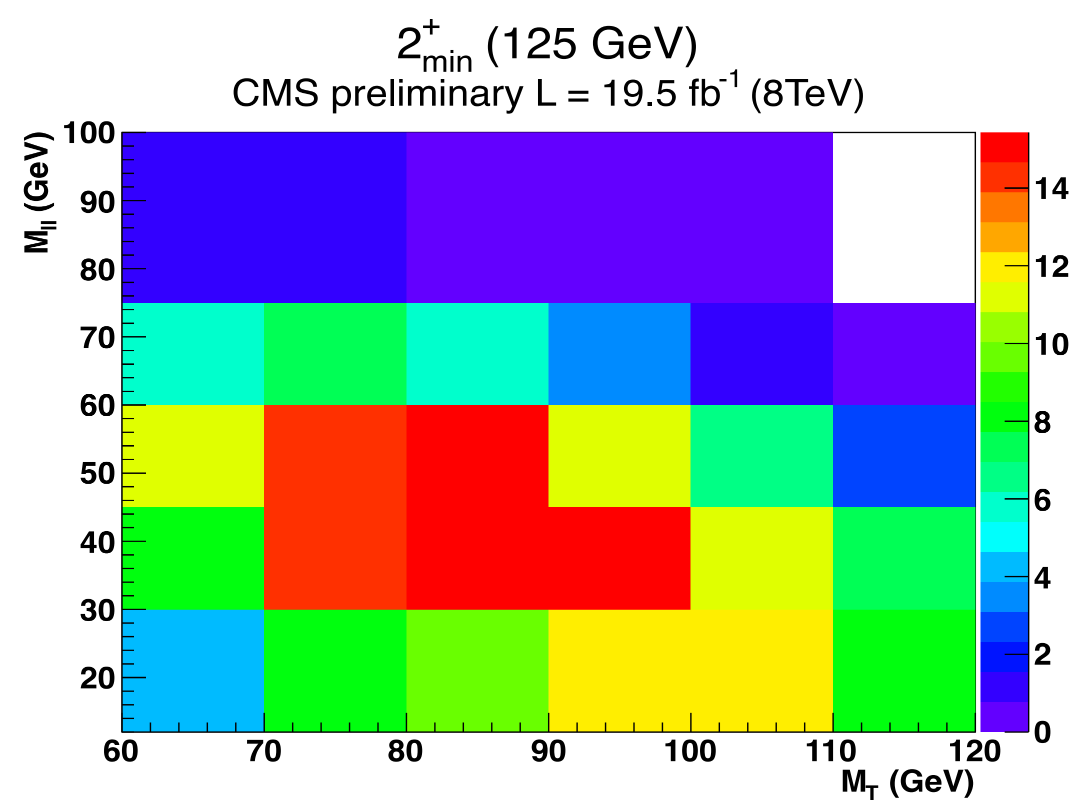
Model to test and method

- $H \to WW \to 2l2\nu$ has good sensitivity to distinguish SM Higgs($J^P=0^+$) from a spin-2 resonance which couples to di-boson through minimal couplings($J^P=2^+$)
- Use the same 2D templates and background estimation as the main analysis in 0/1-jets eμ categories

0^+

eμ 0-jet





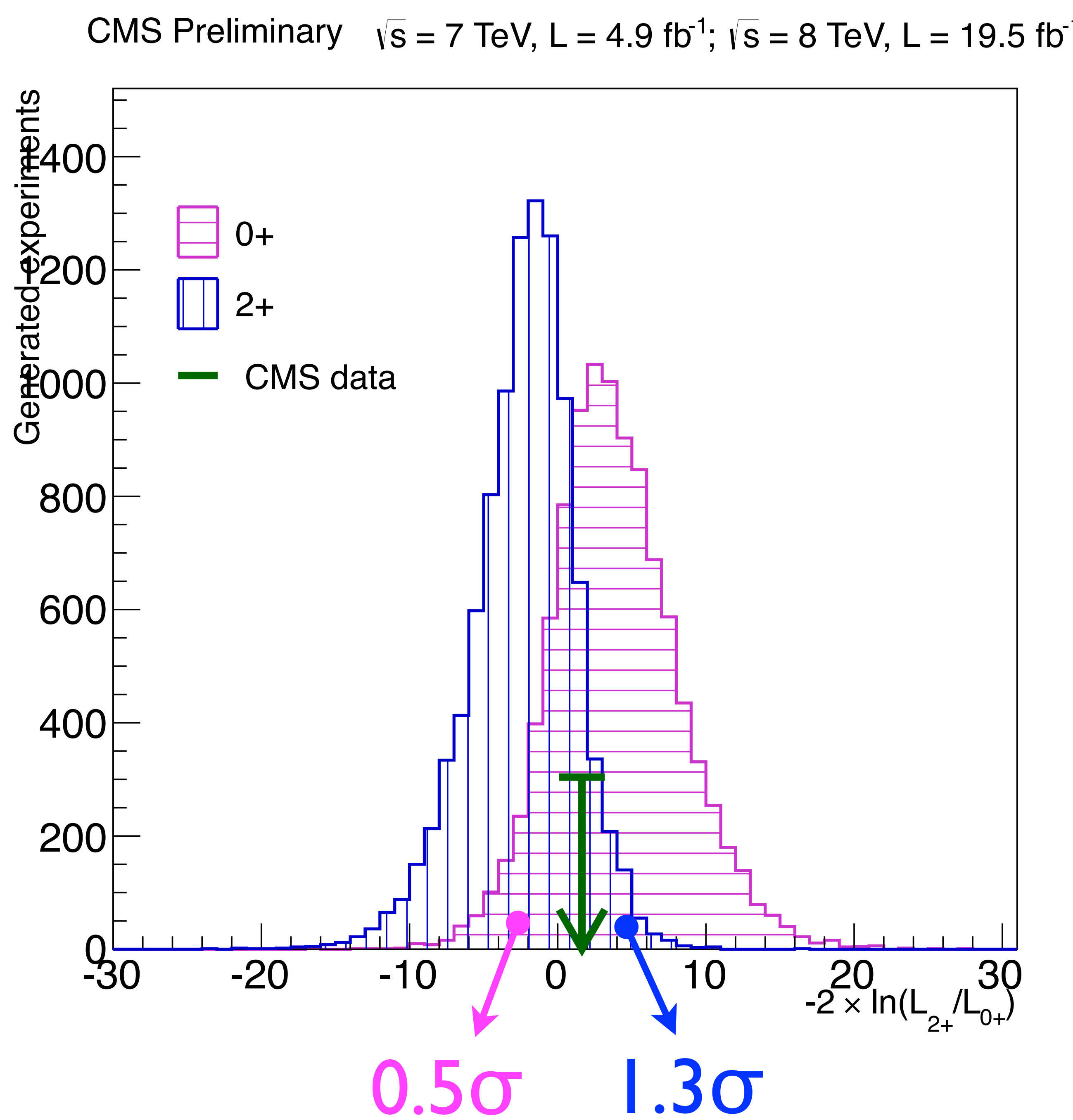
2^+

eμ 0-jet

- Test $gg \to H/X$ ($gg \to X$ normalized to $gg \to H$)

Spin-parity Test

Is data consistent with 2^+ model?



- Test statistic

$$q = -2\ln\left(\mathcal{L}_{2^+}/\mathcal{L}_{0^+}\right)$$

same likelihood used for search

- Result using the best fit values (μ^{0+}=0.76 and μ^{2+}=0.83)

Assumed Model (J^P)	Separation of other model	
	Expected	Observed
0^+	1.5σ	0.5σ
2^+	1.8σ	1.3σ

Summary

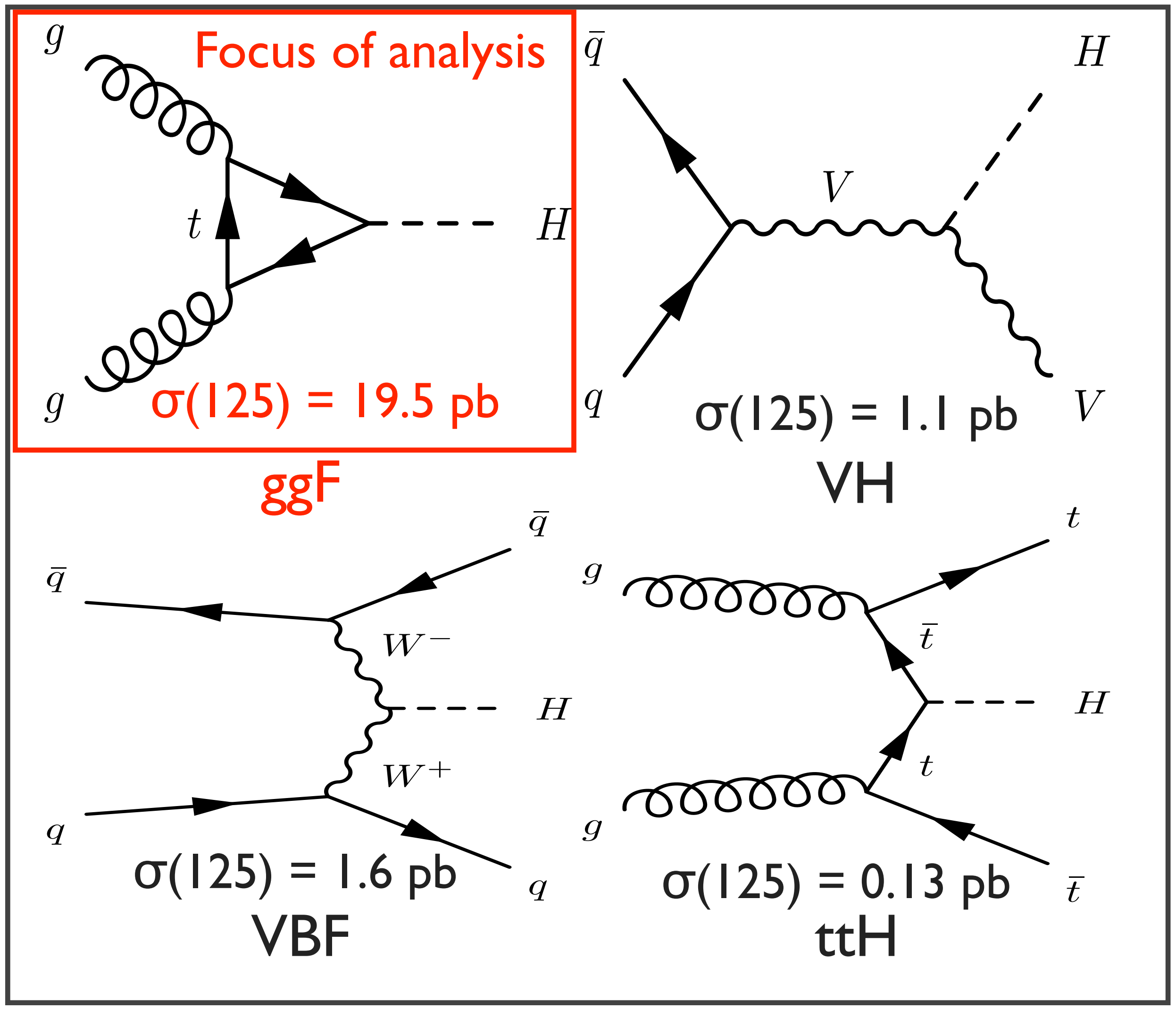
- The whole Run1 LHC data of 4.9 + 19.5 fb^{-1} analyzed for SM Higgs boson search in $H \rightarrow WW \rightarrow 2l2\nu$ 0/1-jet channel
  - Fit model validated using data
  - Significance : 4.0σ / 5.1σ (observed/expected)
  - Signal strength : μ = 0.76 ± 0.13 (stat.) ± 0.16 (syst.)
- Spin-parity hypothesis test performed in 0/1-jet eμ categories
  - Inconsistency with 2+ model : 1.3σ
- All results are consistent with SM Higgs at M_H =125 GeV
- Future plan
  - Publication in progress including VBF and other leptonic channels(WH/ZH)

Extra slides

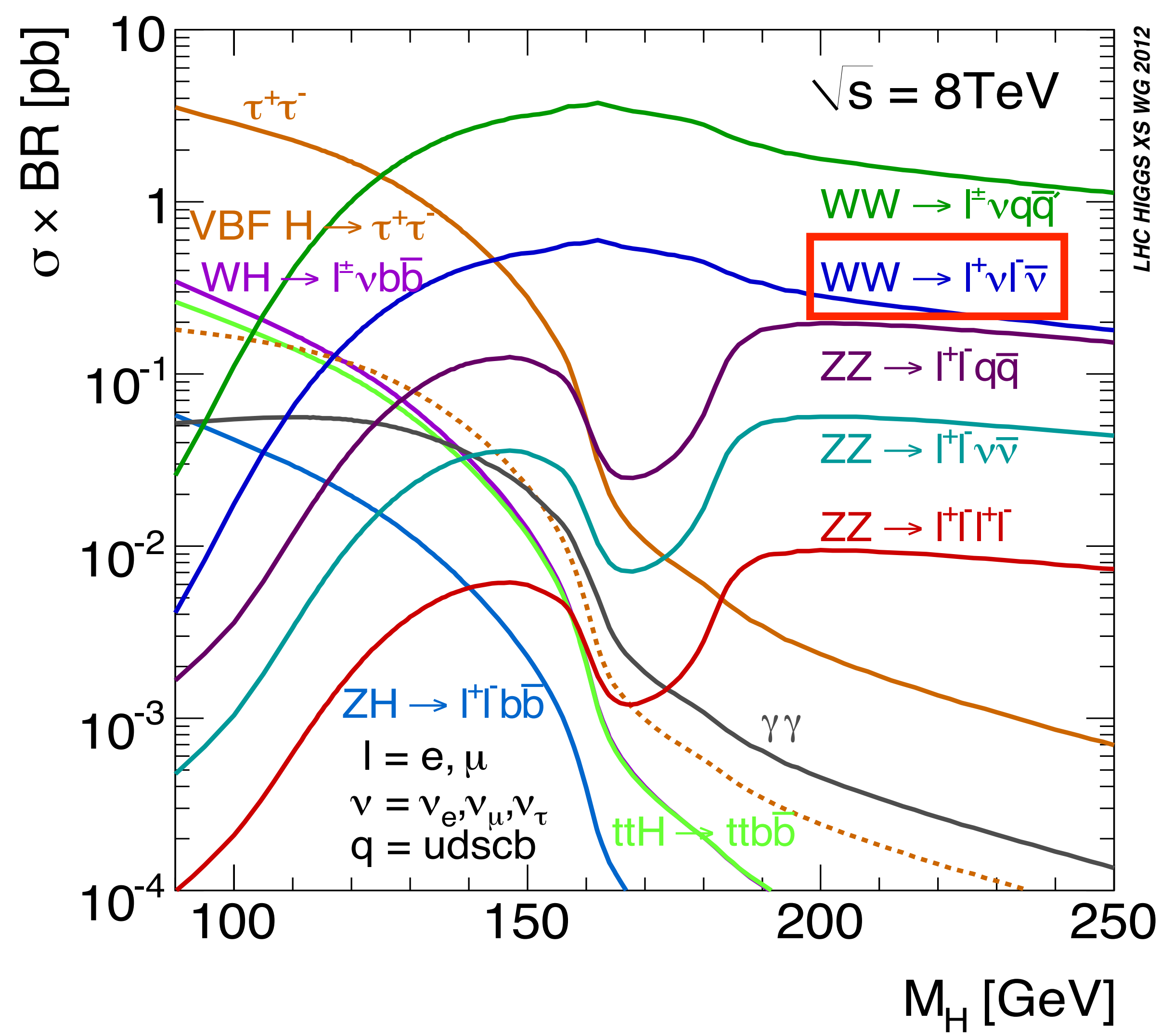
Standard Model Higgs Boson

Production and decay

Production modes



$\sigma \times$ BR



- Standard Model Higgs : charge = 0 and spin = 0
- Mass is a free parameter → task for experimentalists

WW selection

Selection [units]	√s = 7 TeV		√s = 8 TeV	
	ee,μμ	eμ	ee,μμ	eμ
pTmax [GeV/c]	20	20	20	20
pTmin [GeV/c]	15	10	10	10
third lepton veto	applied	applied	applied	applied
opposite-sign requirement	applied	applied	applied	applied
mll [GeV/c2]	20	12	12	12
projected MET [GeV]	$37 + N_{vtx}/2$	20	20	20
Drell-Yan MVA	---	---	applied	---
Z mass veto	applied	---	applied	---
Δφ(ll-jetmax) [dg.]	165	---	---	---
top veto	applied	applied	applied	applied
pTll [GeV/c]	45	30 [*]	45	30 [*]

[*] For the cut and count analysis, p_T^{ll} is required to be larger than 45 GeV.

To suppress WZ/ZZ

To suppress low M_{ll} resonance

To suppress DY, QCD

To suppress DY

To suppress Top

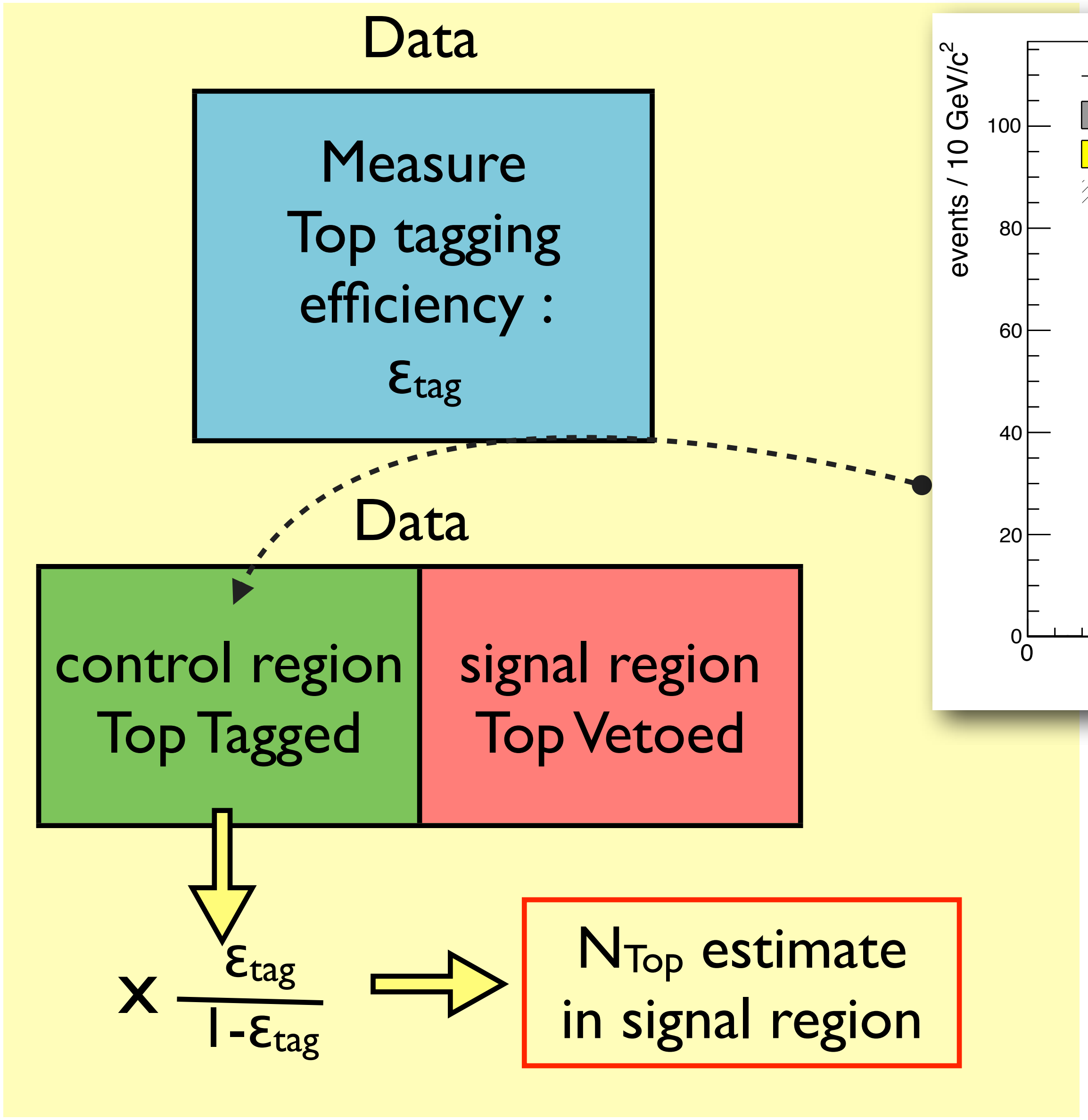
To suppress W+jets

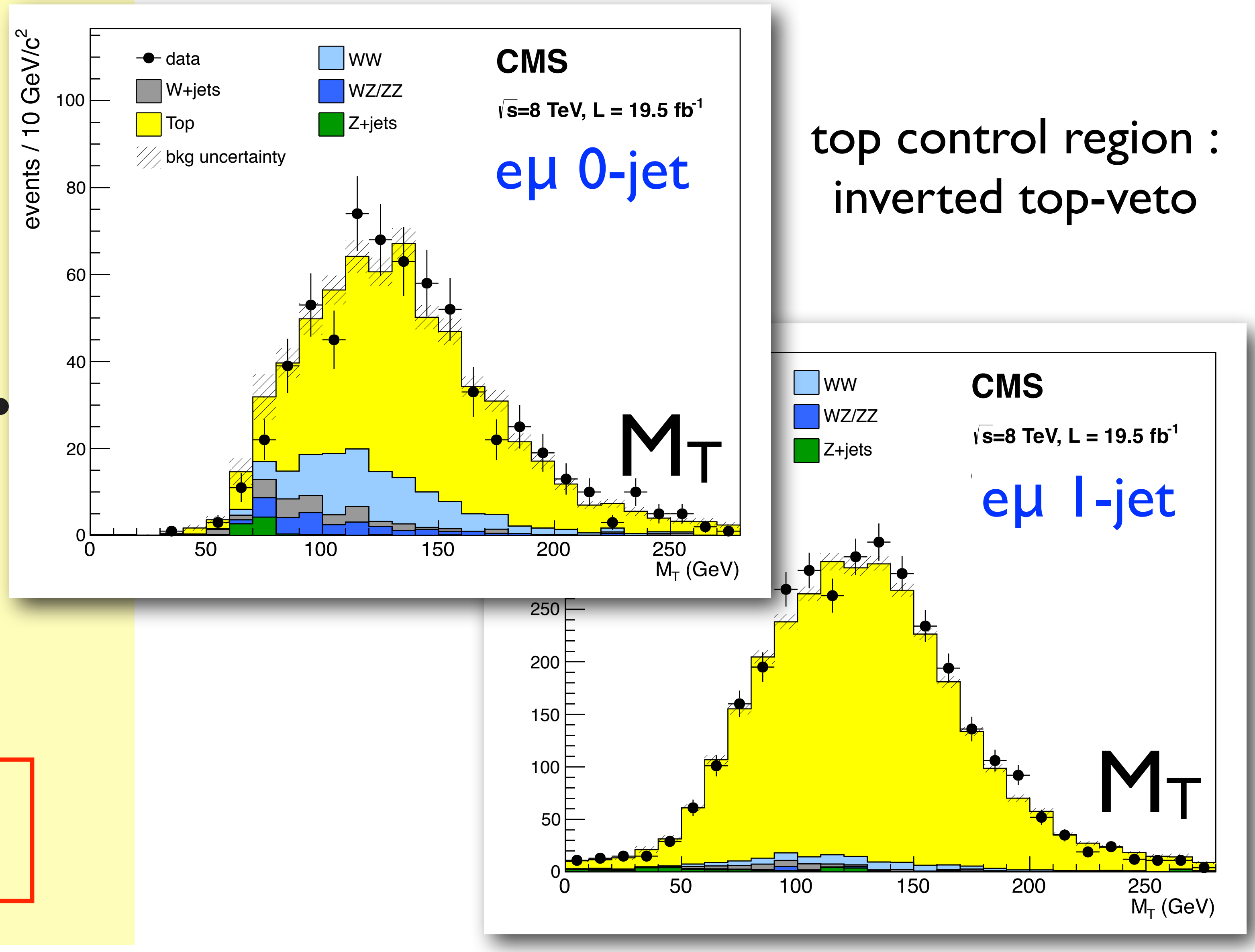
M_H-dependent selection for cut-based analysis

m_{H}	$p_{\mathrm{T}}^{\ell,\mathrm{max}}$	$p_{\mathrm{T}}^{\ell,\mathrm{min}}$	$m_{\ell\ell}$	$\Delta\phi_{\ell\ell}$	m_{T}
[GeV]	[GeV]	[GeV]	[GeV]	[°]	[GeV]
	$>$	$>$	$<$	$<$	[,]
120	20	10	40	115	[80,120]
125	23	10	43	100	[80,123]
130	25	10	45	90	[80,125]
160	30	25	50	60	[90,160]
200	40	25	90	100	[120,200]
250	55	25	150	140	[120,250]
300	70	25	200	175	[120,300]
400	90	25	300	175	[120,400]

Top Estimation

- Main background in 1-jet category
- Handle : presence of b-quarks → top tagging with b-tagged jets and soft muons
- Extrapolation from top-tagged region to top-vetoed region

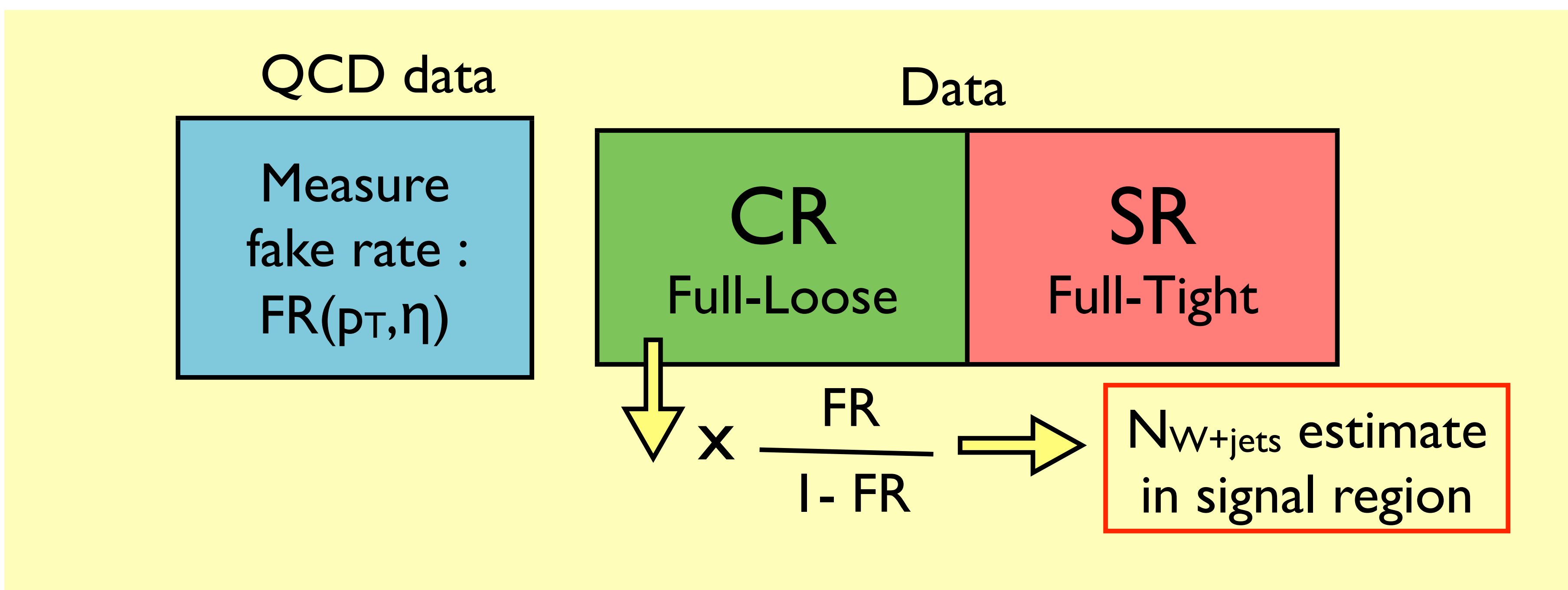




top control region :
inverted top-veto

W+jets Estimation

- Jets can be mis-identified as leptons
- Measure the rate(Fake Rate) for a lepton with loose selection to pass the full requirement in data events dominated by QCD
- Apply FR to the control region where one lepton passes the full selection and the other passes loose selection but not full selection

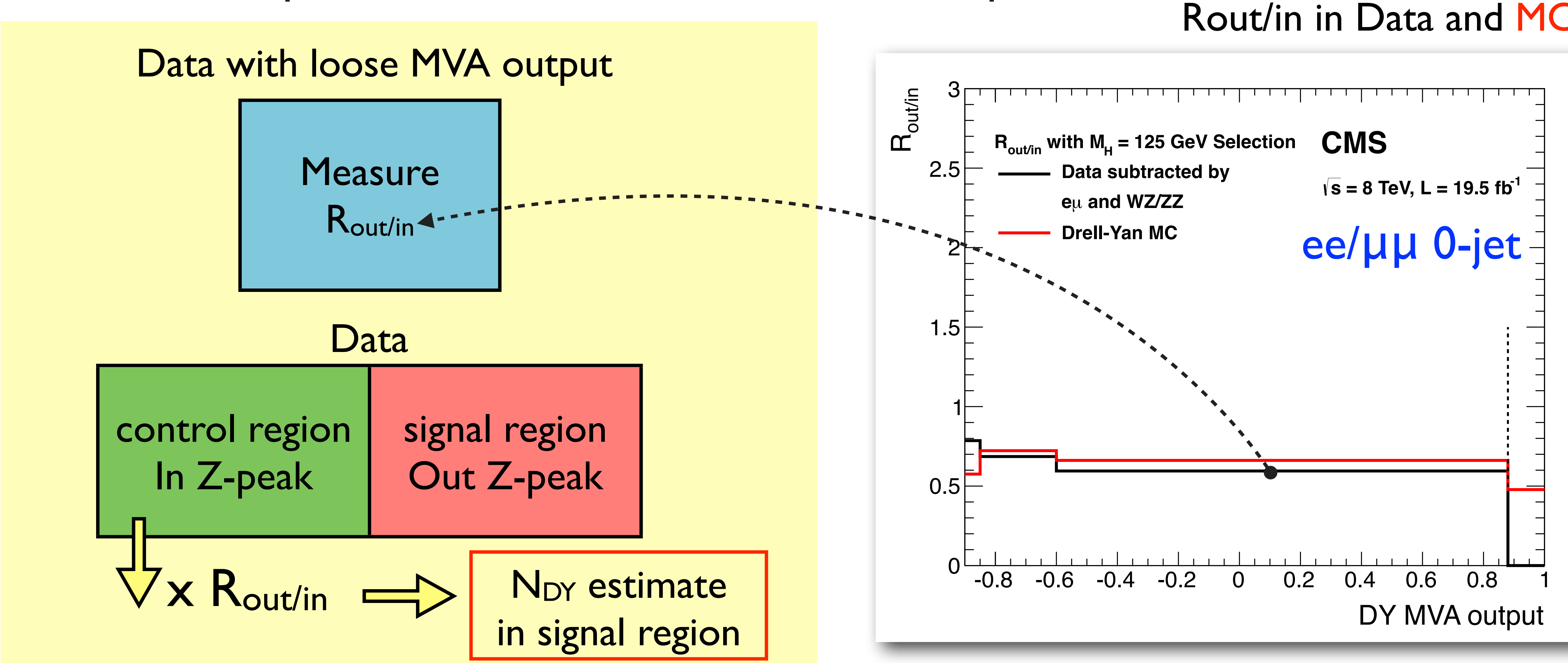


- Systematics : ~ 40 %

Drell-Yan Estimation

- Main background in ee/μμ final states
- Handle : Z mass veto and MVA-based Drell-Yan suppression technique
  → worse sensitivity than eμ channel
- Extrapolation from inside to outside of Z peak

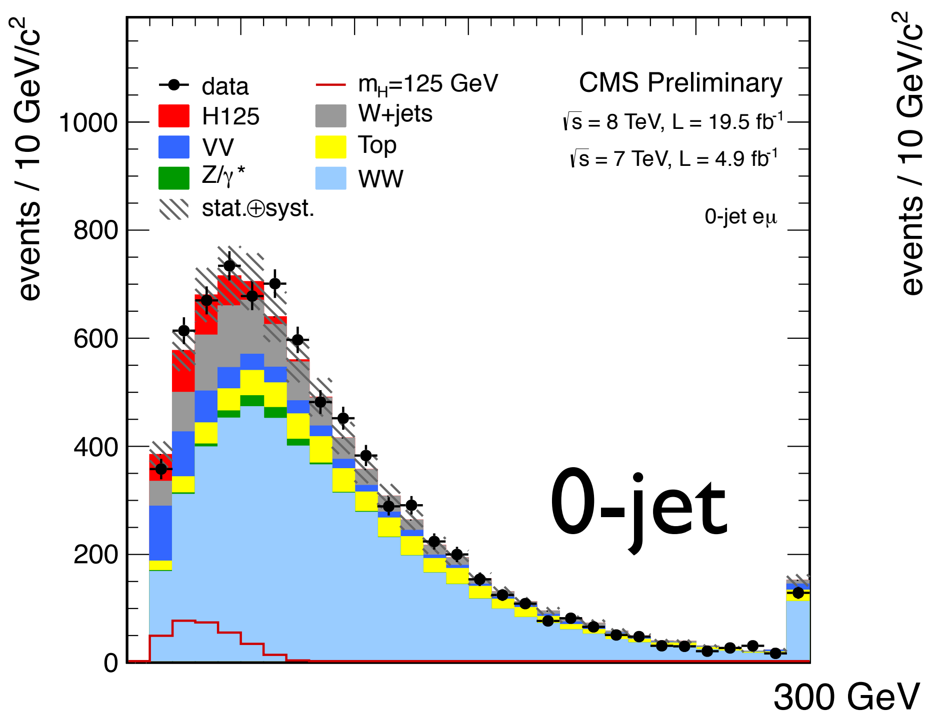
Rout/in in Data and MC

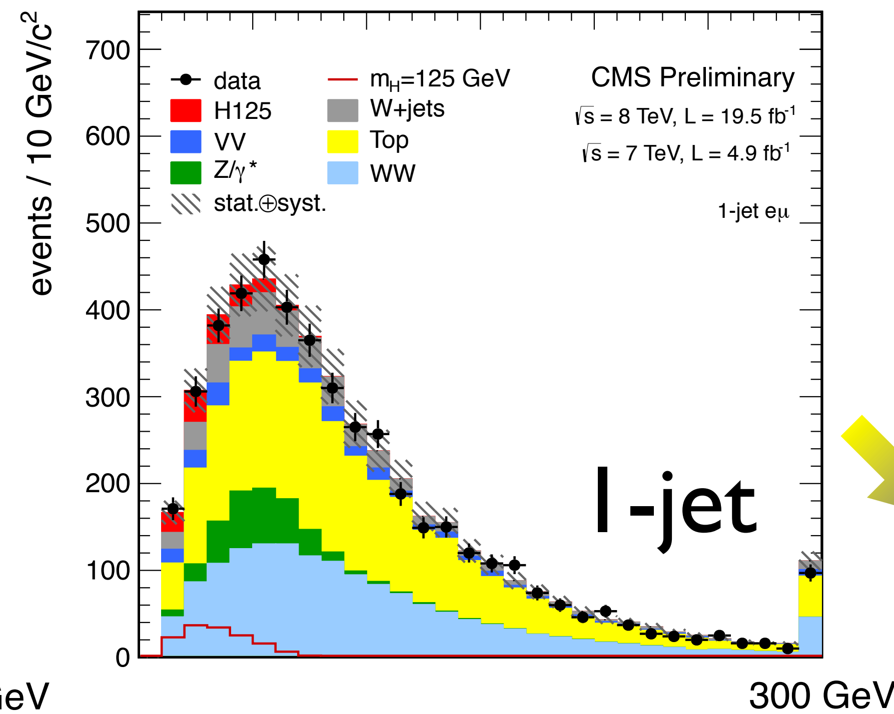


WW Selection Results

Putting all these together

M_{ll}

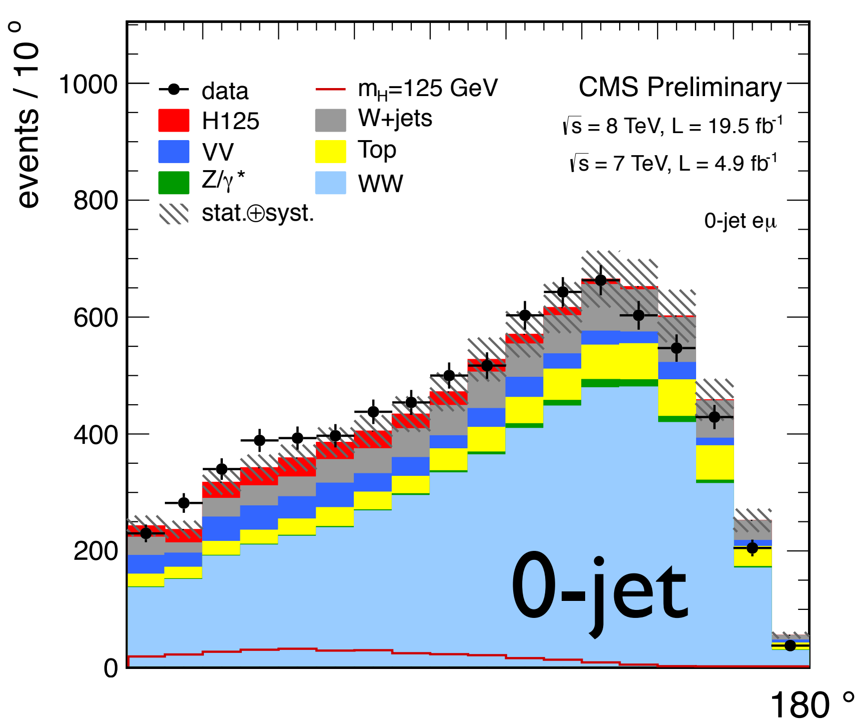


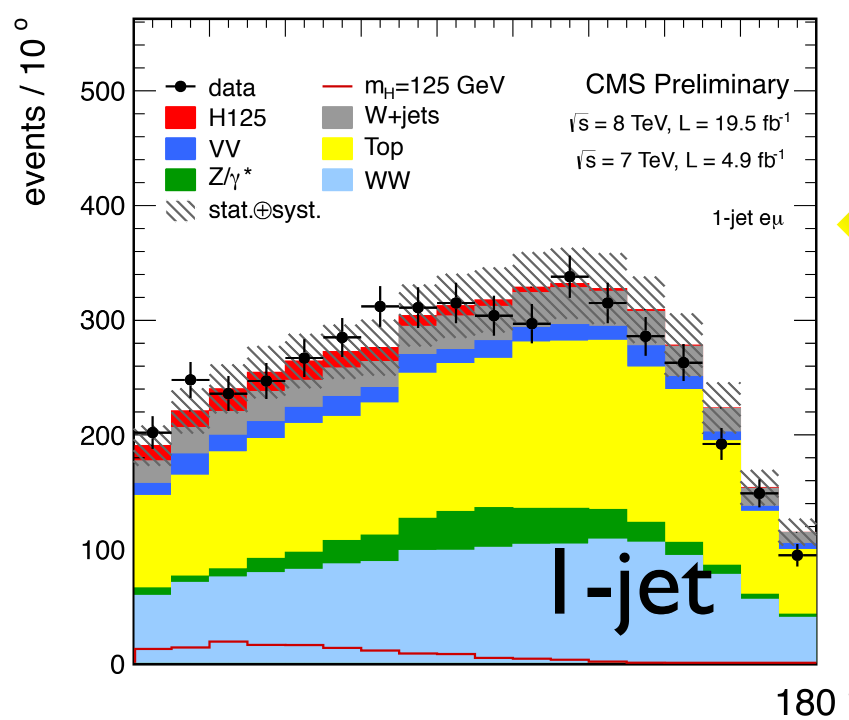


WW selection
eμ channel

Overall good agreement between Data and MC

$\Delta\phi_{ll}$



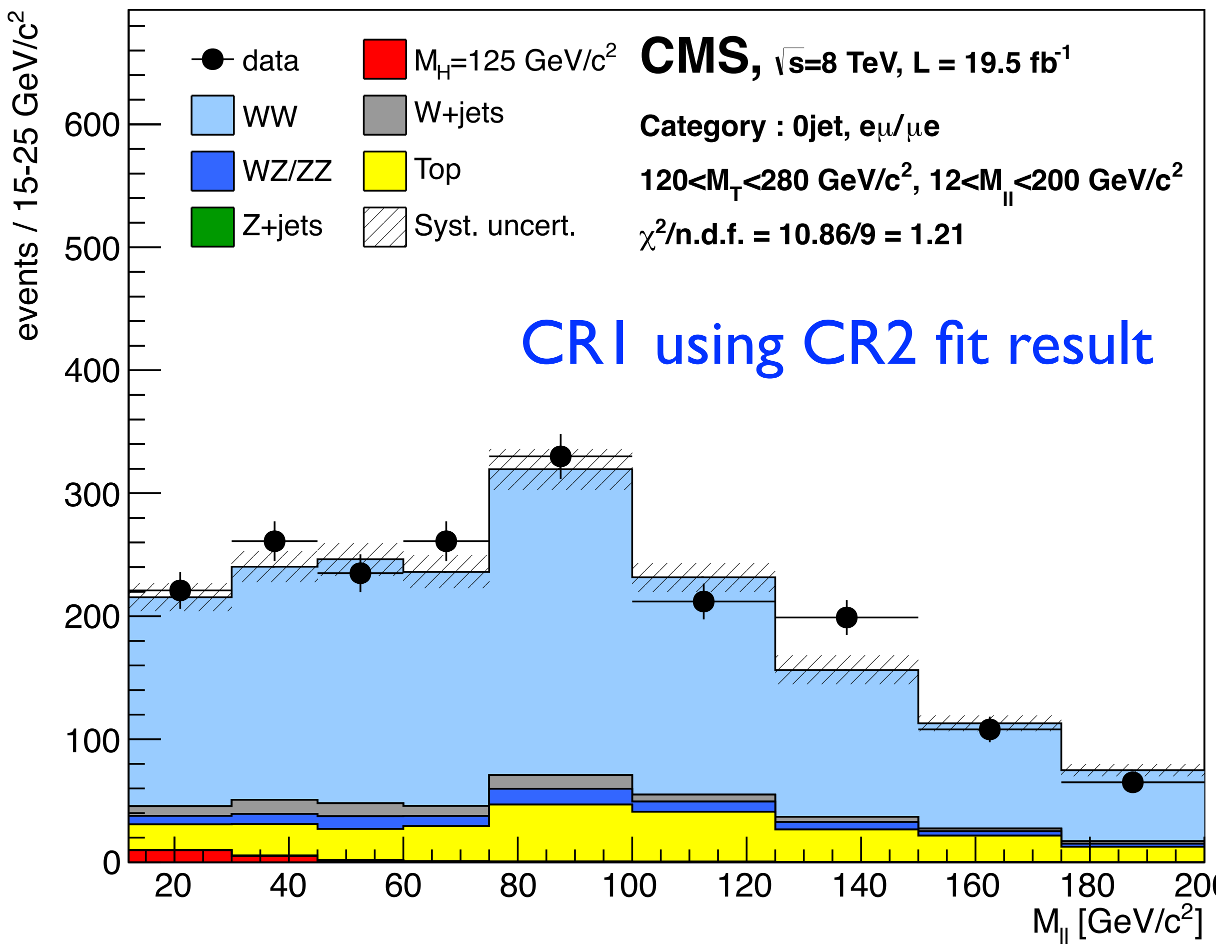


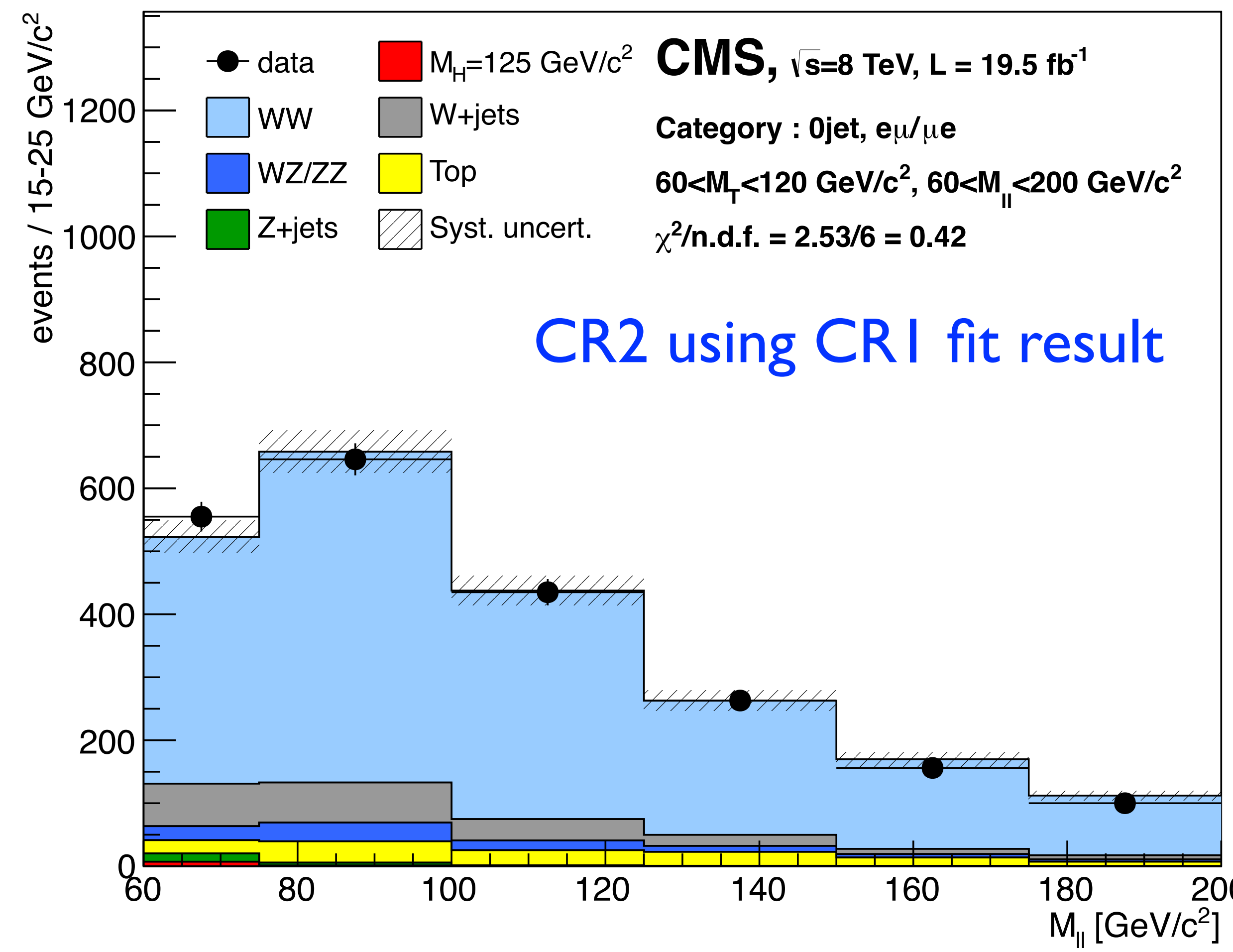
Systematics

- Luminosity : 4.4 % (8TeV), 2.2% (7 TeV)
- Theoretical uncertainties on signal following LHC cross section recommendation
  - PDF + higher order effects + UEPS : 20 - 30 %
- Background normalization
  - WW : 5/10 % for cut-based, W+jets : 36 %, Top : 20/5 %, DY : 30 - 200 %, Wγ(*) : 30 - 40 %
- Instrumental
  - Lepton identification and trigger efficiency : 3(4) % for muon(electron)
  - Lepton Energy/Momentum scale : 1.5 % for muon, 2 % (5 %) for electron in barrel (endcap)
  - MET resolution : 2 %, Jet energy scale : 2 - 10 %
- Shape variations
  - Instrumental variation : list same as above
  - Backgrounds :
    - WW : QCD scale variation and different generators(Madgraph vs MC@NLO)
    - Top : different generators(Madgraph vs Powheg)
    - W+jets : difference away jet p_T thresholds

More WW fit validation plots

M_{ll} is shown here instead of M_T

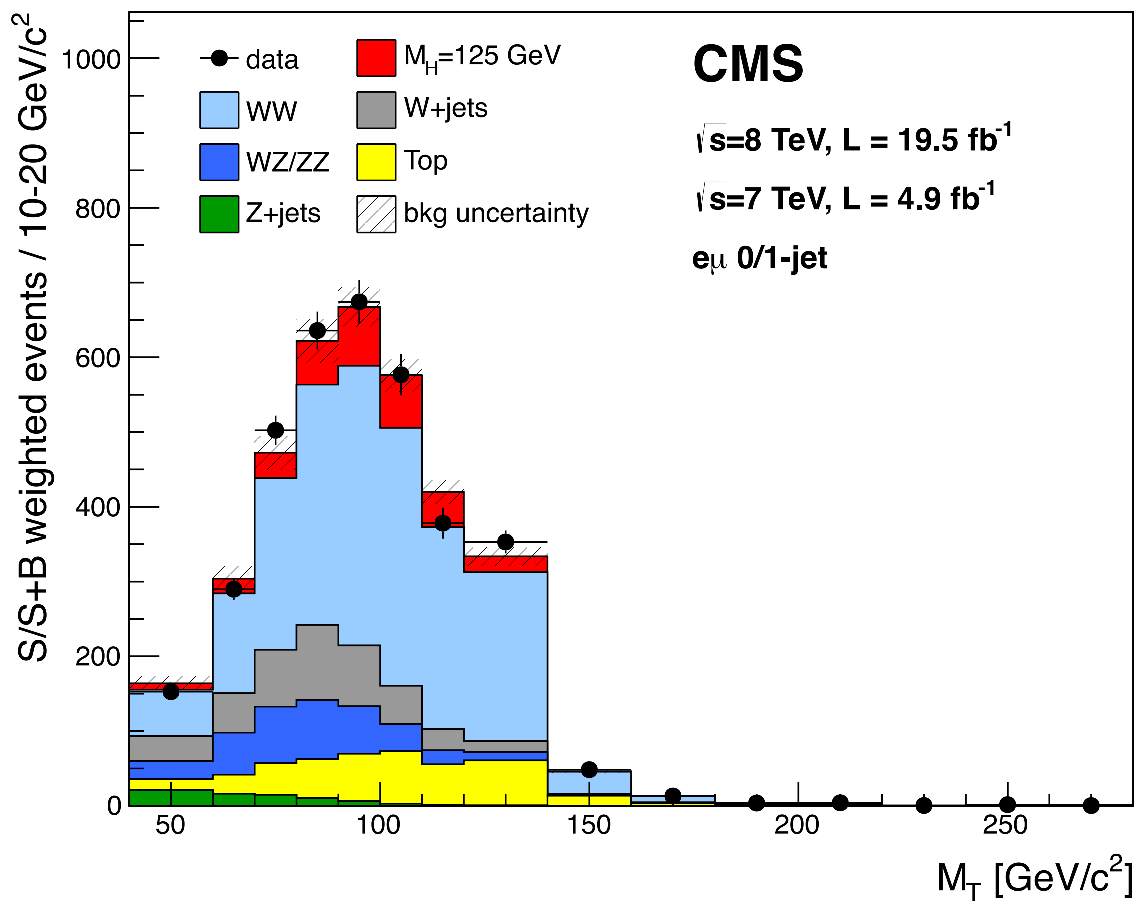


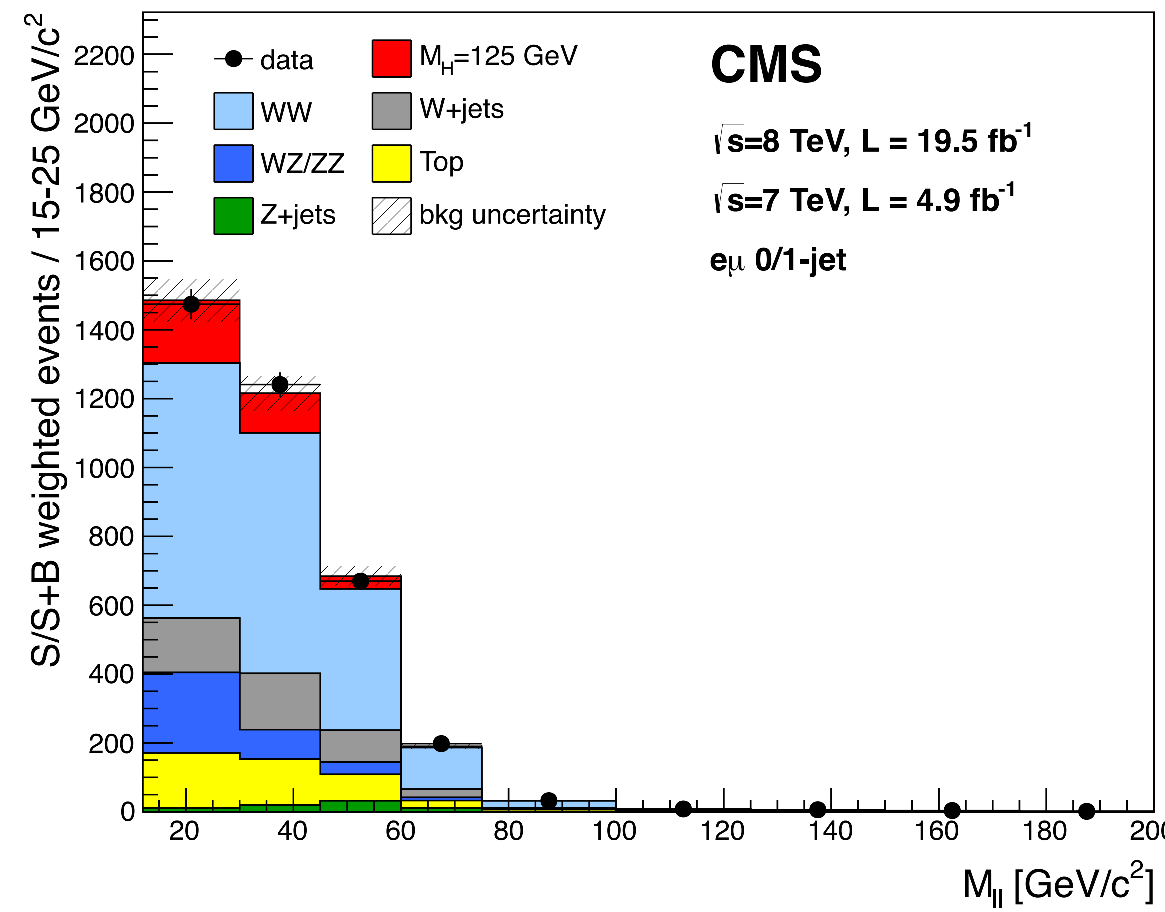


Good agreement with data
→ WW fit model is correct

Stacked S/(S+B) weighted plots

used to make data - bkgd plots





Significance and signal strength

Divided by energy and analysis method

cut-based : cut-based ee/μμ + cut-based eμ
shape-based : cut-based ee/μμ + shape-based eμ

7 TeV expected/observed significance		8 TeV expected/observed significance		7+8 TeV expected/observed significance	
cut-based	shape-based	cut-based	shape-based	cut-based	shape-based
1.7/0.8	2.5/2.2	2.6/2.1	4.7/3.5	2.7/2.0	5.1/4.0

7 TeV expected/observed significance		8 TeV expected/observed significance		7+8 TeV expected/observed significance	
best fit value		best fit value		best fit value	
cut-based	shape-based	cut-based	shape-based	cut-based	shape-based
0.46 ± 0.57	0.91 ± 0.44	0.79 ± 0.38	0.71 ± 0.22	0.71 ± 0.37	0.76 ± 0.21