Fake Lepton Estimation in ttH Multi-Lepton Analysis

$$t\bar{t}H(H \to WW/ZZ/\tau\tau \to\, \geq 1\ell)$$
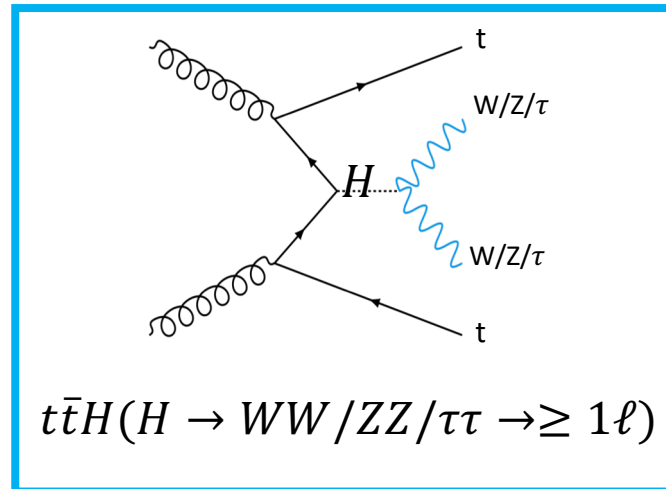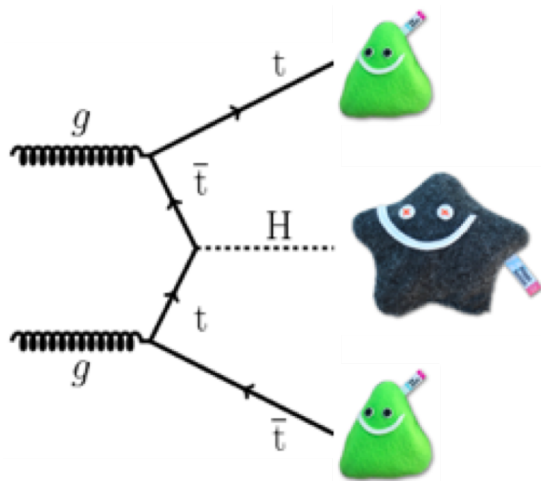
Zhi Zheng (University of Michigan)
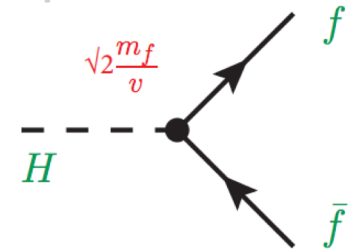
2018 US LHC Users Association Meeting Oct. 26

Top Quark: Bull in a China Shop
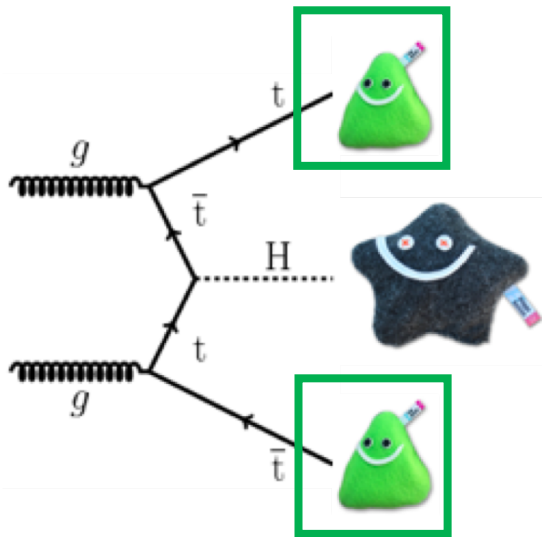
H.Mildner

Post-Higgs discovery:

- ## Measuring Higgs Yukawa couplings at the LHC
  - Top is heaviest fermion in the SM→largest Yukawa coupling $\lambda_t \approx 1$
- $t\bar{t}H$ production: best direct way to measure top quark Yukawa coupling
  - $t\bar{t}H$ production cross section at 13TeV:
    $0.507pb^{+5.8\%}_{-9.2\%}$(QCD scales)$\pm 3.6\%$(PDF,$\alpha_s$)
  - Only ~1% of the total Higgs cross-section

$\sqrt{2}\frac{m_f}{v}$

H f $\bar{f}$
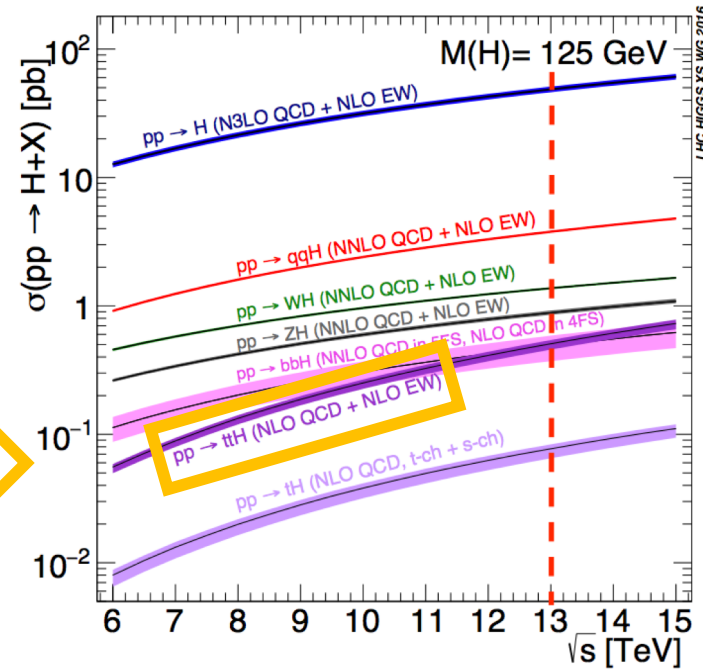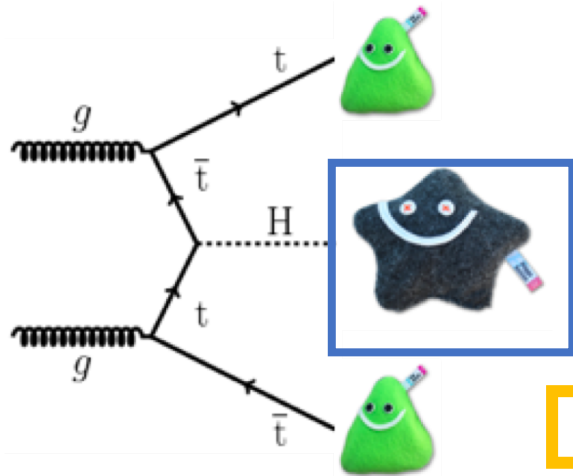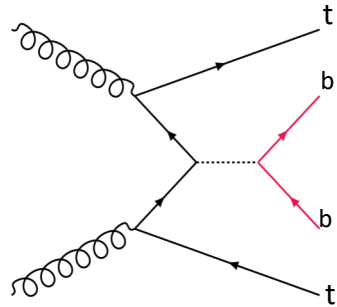
Post-Higgs discovery:

- Measuring Higgs Yukawa couplings at the LHC
  - Top is heaviest fermion in the SM→largest Yukawa coupling $\lambda_t \approx 1$

- $t\bar{t}H$ production: best direct way to measure top quark Yukawa coupling
  - $t\bar{t}H$ production cross section at 13TeV:
    $0.507pb^{+5.8\%}_{-9.2\%}$(QCD scales)$\pm 3.6\%$(PDF,$\alpha_s$)
  - Only ~1% of the total Higgs cross-section
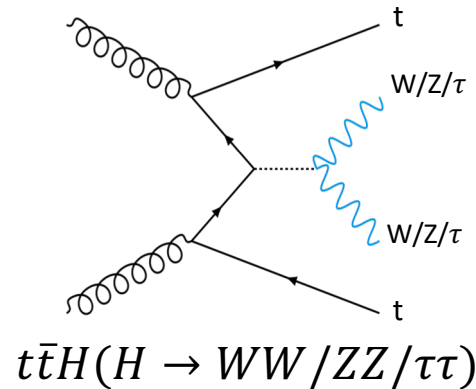
$t\bar{t}H$ production

Large Branch Ratio

Cleaner Signal
Smaller Background

| ttH(bb) | ttH Multi-lepton | ttH($\gamma\gamma$) |

$t\bar{t}H(H \rightarrow b\bar{b})$

$t\bar{t}H(H \rightarrow WW/ZZ/\tau\tau)$
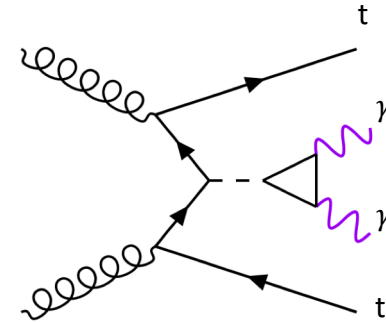
$t\bar{t}H(H \rightarrow \gamma\gamma)$

$t\bar{t}H$ production

Large Branch Ratio

Cleaner Signal Smaller Background

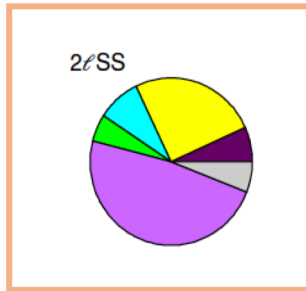**ttH(bb)** | **ttH Multi-lepton** | **ttH($\gamma\gamma$)**
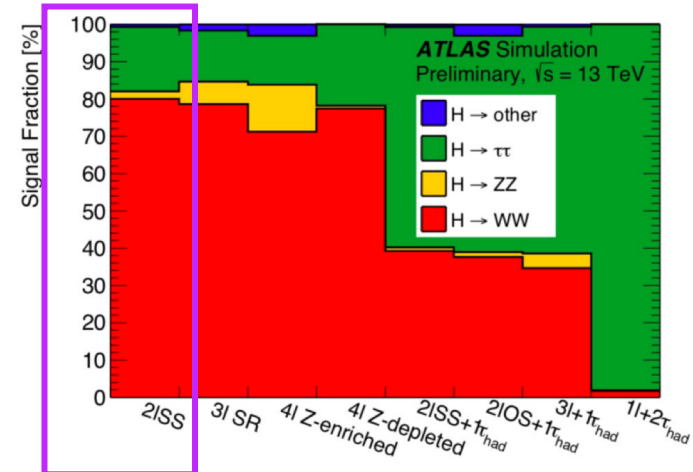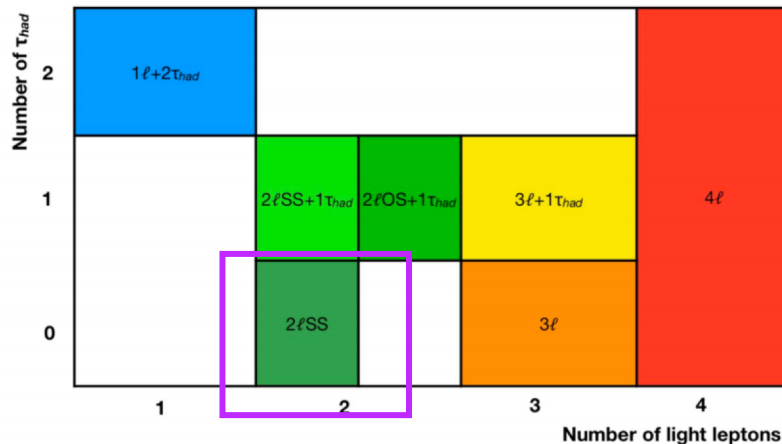
- High multiplicity final state: leptons, jets, bjets
  - Have 7 categories based on number of lepton and number of τ_{had}

- Light-lepton channels are more sensitive to $H \rightarrow WW$ decays
- τ_{had} channels are more sensitive to $H \rightarrow \tau\tau$

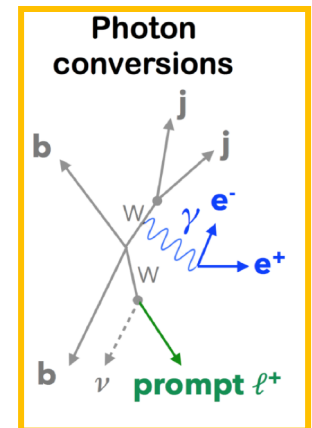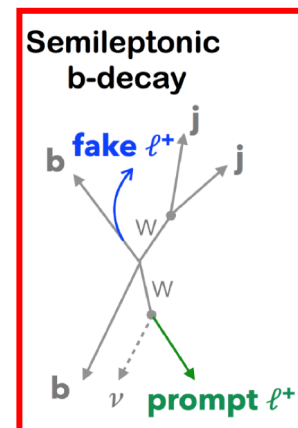# $t\bar{t}H$ Multilepton Same Sign Dilepton Region

ATLAS Work in process
$\sqrt{s} = 13$ TeV

2ℓSS

- q mis-id
- $t\bar{t}Z$
- Fake τ_{had}
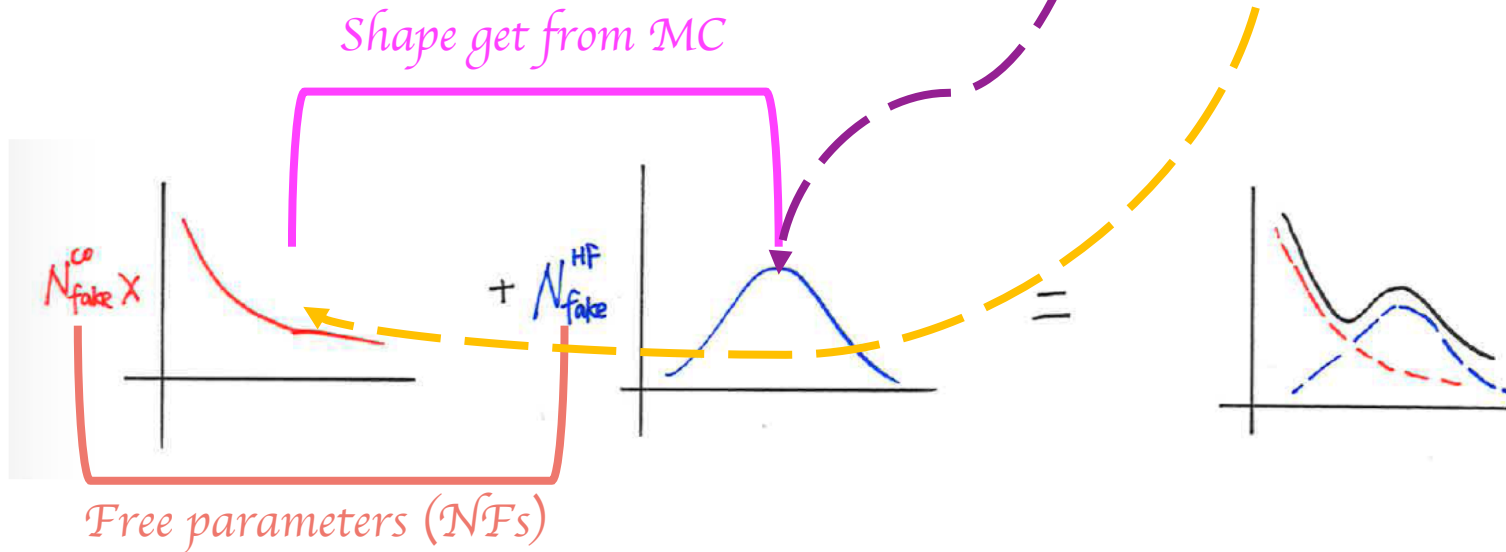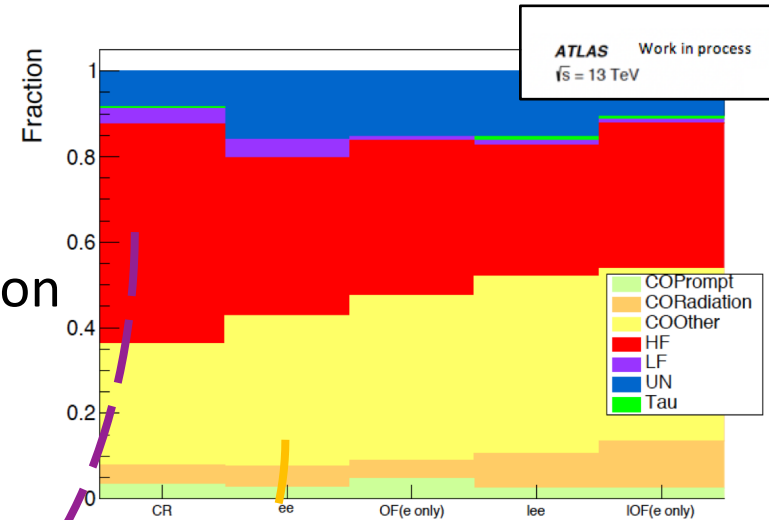- Other
- $t\bar{t}W$
- Diboson
- Non-prompt

- Main background sources:
  - Standard model processes with all prompt lepton
    - $t\bar{t}W$, $t\bar{t}Z$ other rare SM processes
  - Electron charge mis-identification
  - Events with fake/non-prompt light leptons
    - Semileptonic b-decay, photon conversion

Semileptonic b-decay

Photon conversions

Fake Lepton Estimation => Template Fit

Assumption:

- Rely on Monte Carlo (MC) to describe desired observables (jet multiplicity ect.)

- Build templates based on the truth classification of events containing a non-prompt lepton, a free floating normalization factor (NF) is assigned to each of them

- Overall normalization same in control regions(CRs) and signal regions(SRs)

Shape get from MC

Free parameters (NFs)

Fake Lepton Estimation => Template Fit
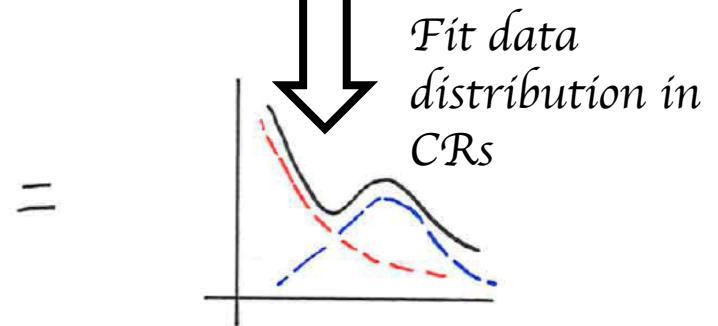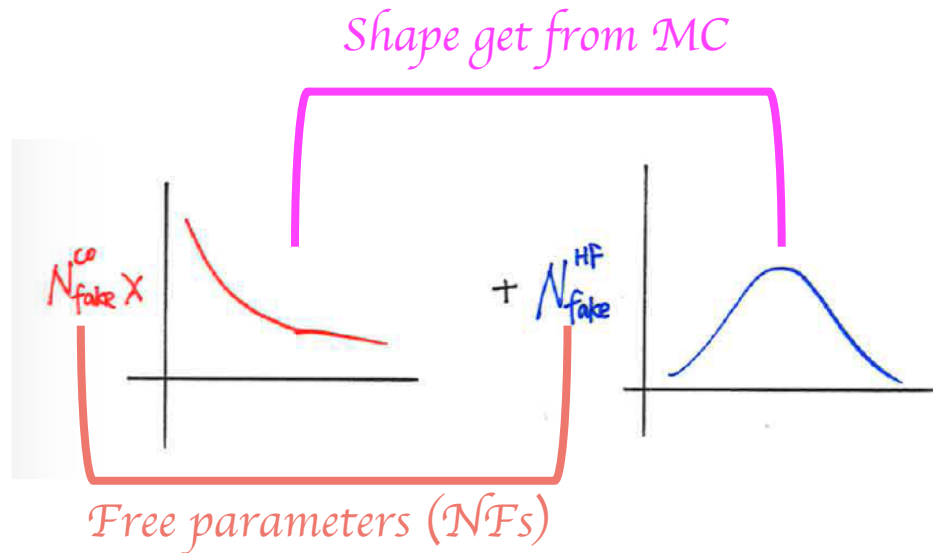
Assumption:

- Rely on Monte Carlo (MC) to describe desired observables (jet multiplicity ect.)

- Build templates based on the truth classification of events containing a non-prompt lepton, a free floating normalization factor (NF) is assigned to each of them

- Overall normalization same in control regions(CRs) and signal regions(SRs)

Shape get from MC

Free parameters (NFs)

Fit data distribution in CRs

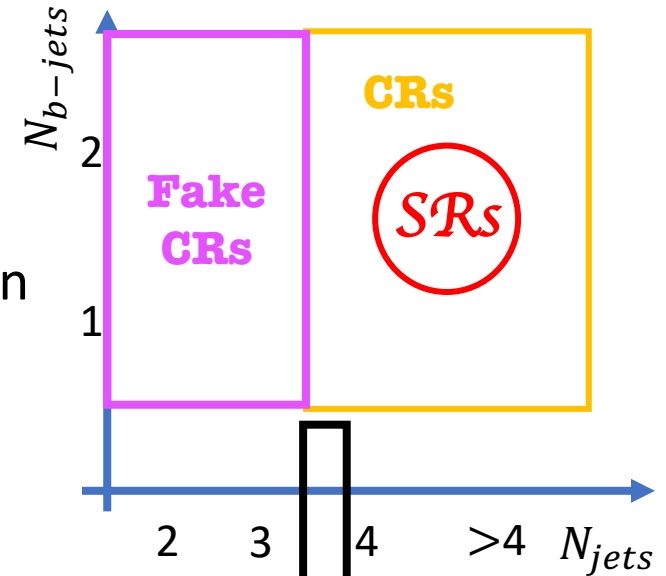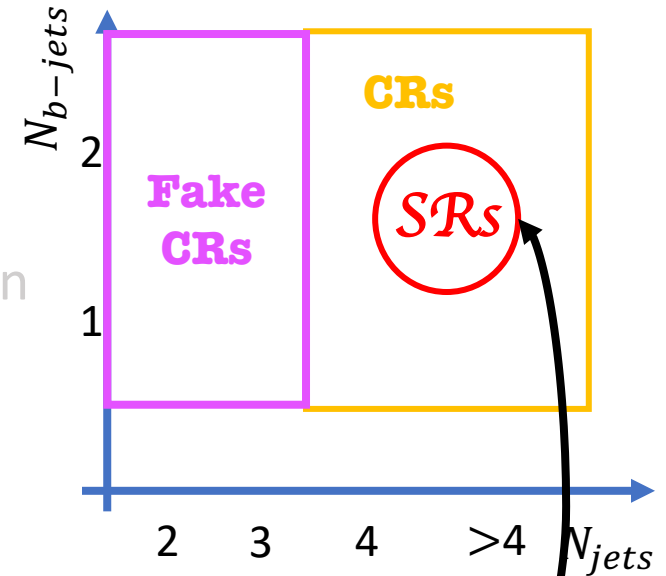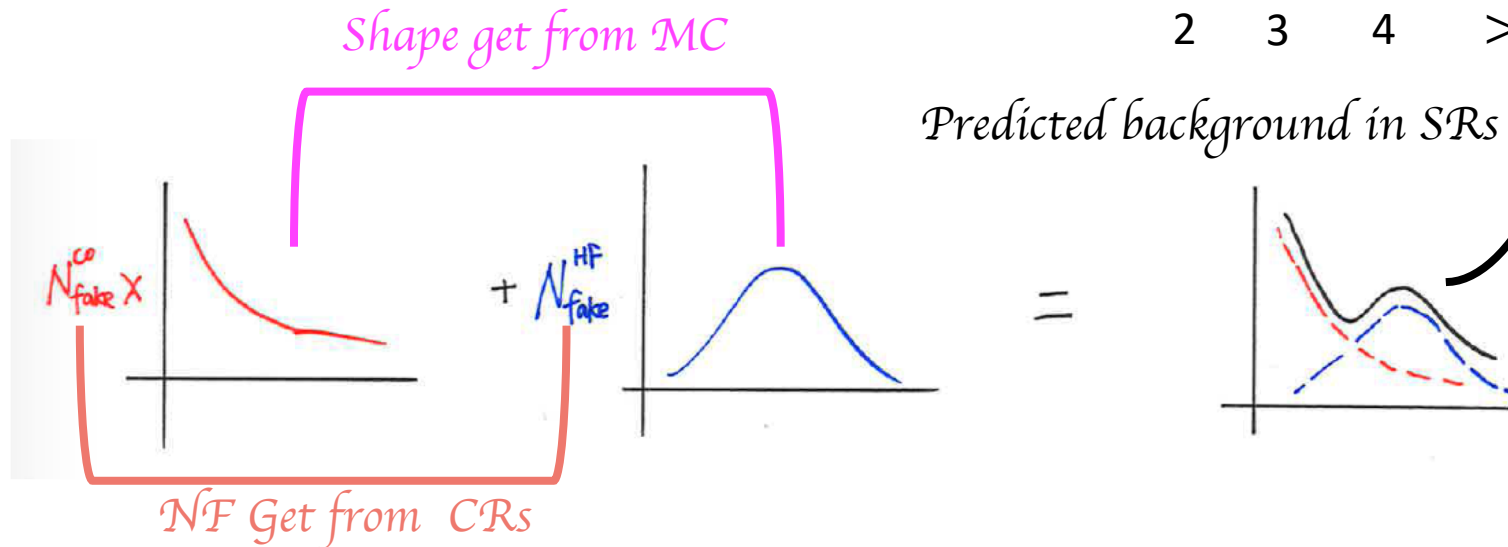# Fake Lepton Estimation => Template Fit
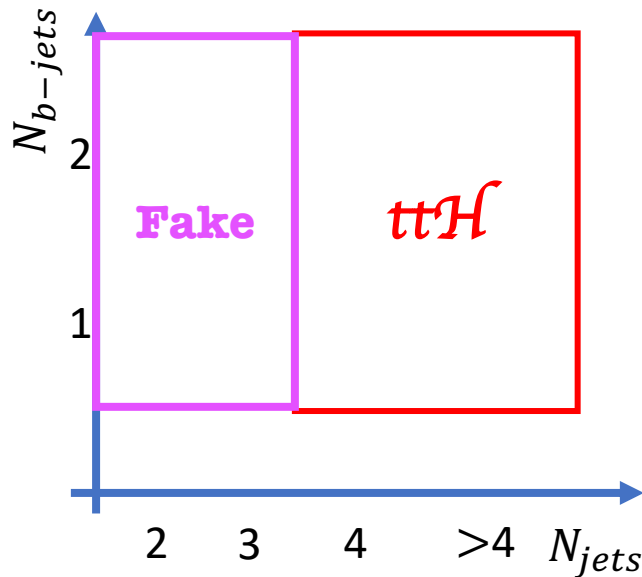
Assumption:

- Rely on Monte Carlo (MC) to describe desired observables (jet multiplicity ect.)

- Build templates based on the truth classification of events containing a non-prompt lepton, a free floating normalization factor (NF) is assigned to each of them

- Overall normalization same in control regions(CRs) and signal regions(SRs)

Shape get from MC

Predicted background in SRs

N_{b-jets}

2

1

CRs

Fake CRs

SRs

2   3   4   >4 N_{jets}

$N_{fake}^{co} \times$ + N_{fake}^{HF} =

NF Get from CRs

Fake Lepton Estimation => Template Fit

Challenge:

- Very complex fake compositions and limited MC stats.
  - Adding $N_{jets} \in [2,3]$ as fake control region
  - CRs chose based on MC stats.

- Large ttW contribution in the fake lepton enriched region and ttH enriched region
  - Free float ttW normalization
  - Build ttW enriched region in High NJ regions

- High correlations between those NFs
  - Variables are carefully chose to decorrelate NFs
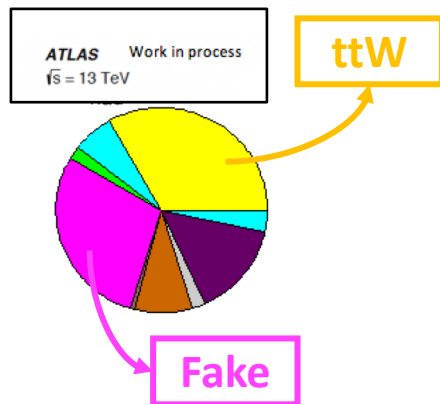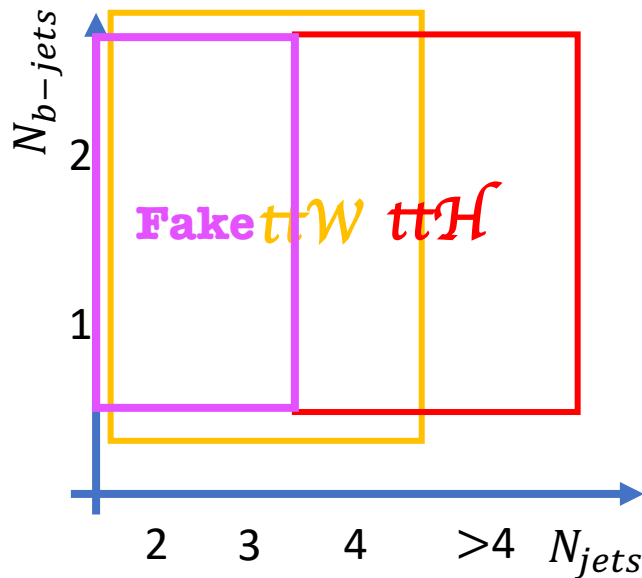
Fake Lepton Estimation => Template Fit


Challenge:

- Very complex fake compositions and limited MC stats.
  - Adding $N_{jets} \in [2,3]$ as fake control region
  - CRs chose based on MC stats.

- Large ttW contribution in the fake lepton enriched region and ttH enriched region
  - Free float ttW normalization
  - Build ttW enriched region in High NJ regions   Can measure ttW as well

- High correlations between those NFs
  - Variables are carefully chose to decorrelate NFs

Application of Template Fit

Input Bins:

$\mu\mu + e\mu$ **HighNJ CR** $ee + \mu e$ **HighNJ CR**   **LowNJ CR**
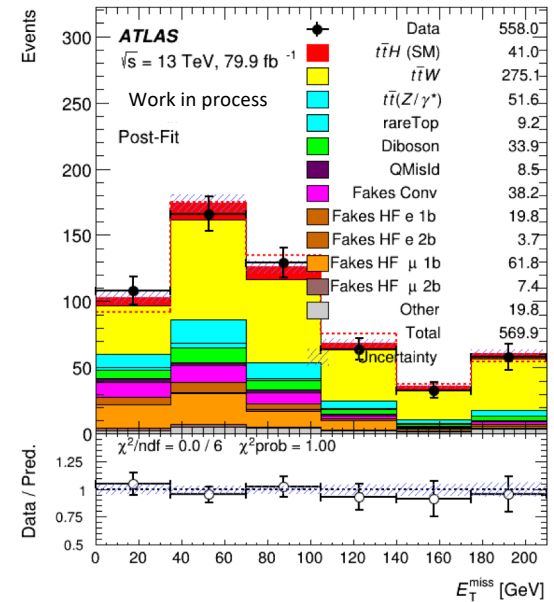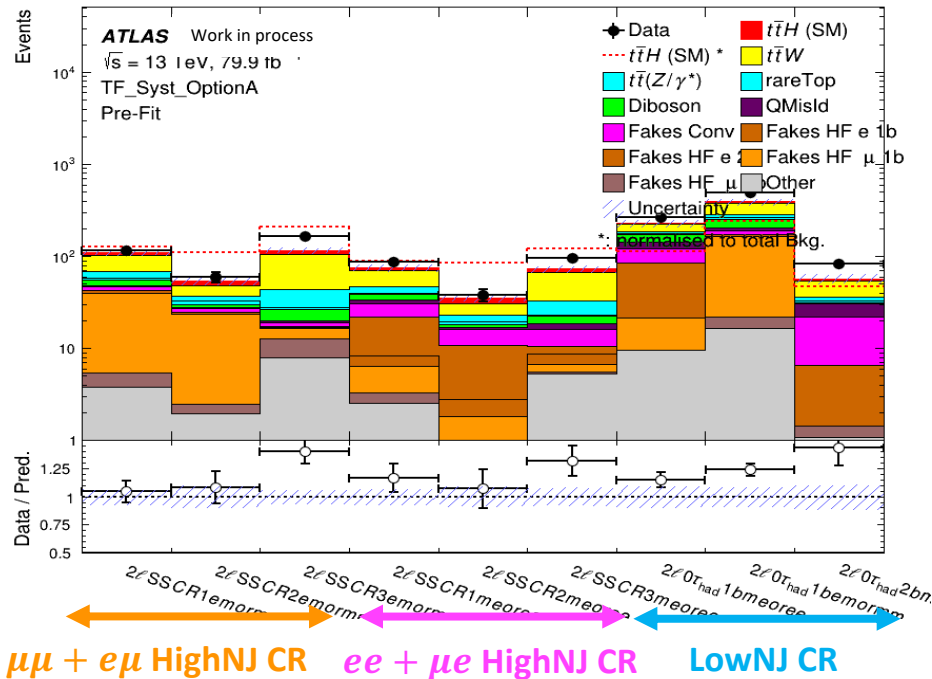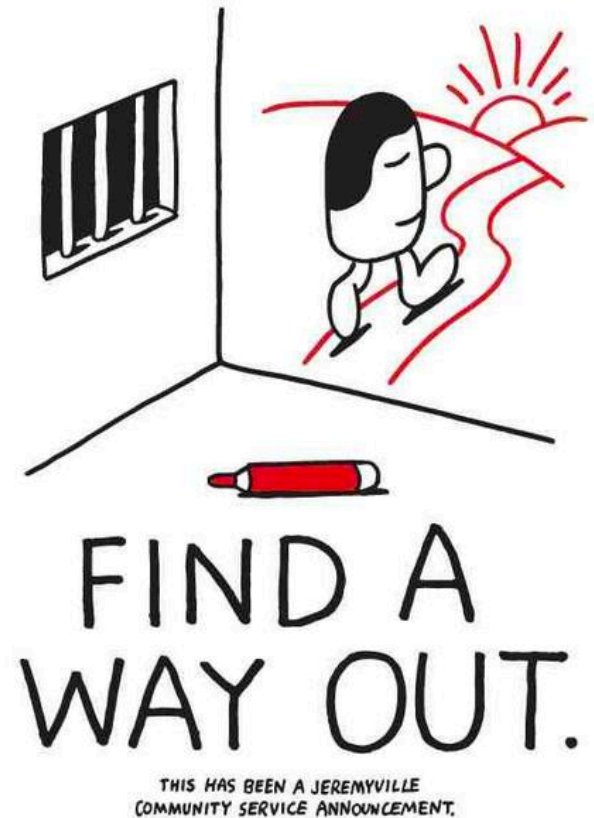
Validation:
Good Agreement

Conclusion and Future Plan

- We study Higgs top yukawa coupling through $t\bar{t}H$ multi-lepton channel

- We find a good way to estimate non-prompt/fake lepton background

- As we continue to explore this Higgs decay mode and production mode:
  - Need understand different sources of fake leptons
  - Need better understanding of ttW process

Thank you ☺