

Search for the SM production of 4 tops

Thibault CHEVALÉRIAS

CEA Saclay - IRFU

Top LHC France
May 4, 2017



Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

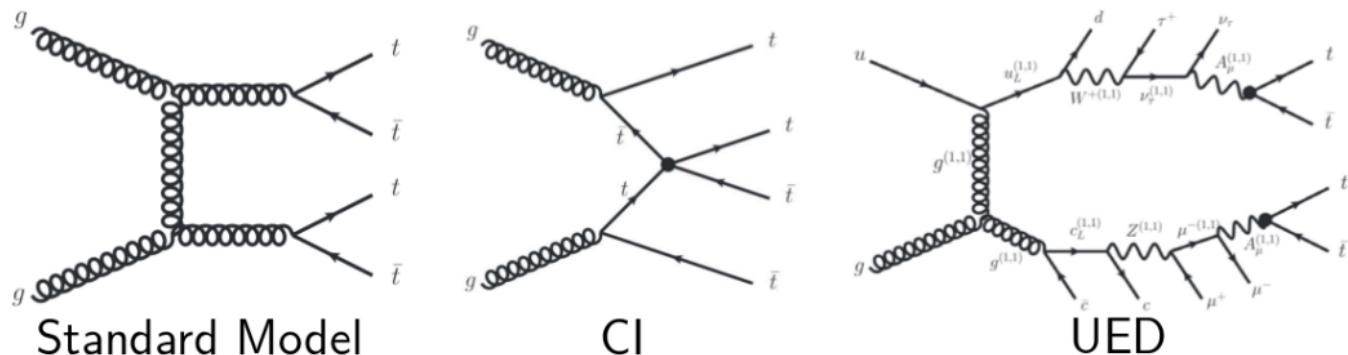
5 Conclusion

Why four tops?

- The top quark plays a key role in many BSM scenarios

⇒ Some of which predict an **enhancement of the $t\bar{t}t\bar{t}$ cross-section**

- Contact Interaction model (CI)
- Universal Extra Dimensions (UED)
- ...



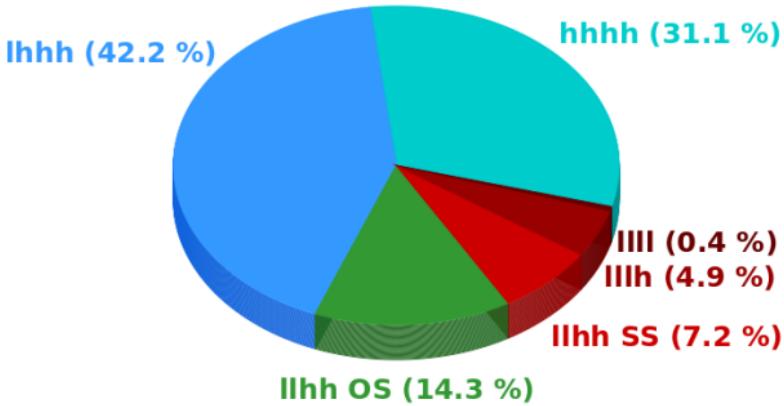
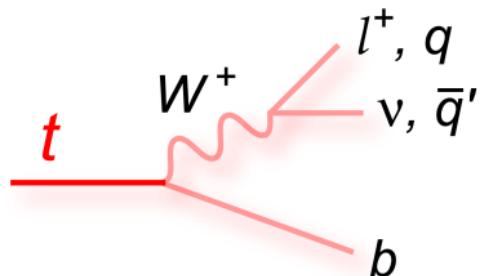
- $\sigma_{t\bar{t}t\bar{t}}^{SM}$ not yet measured experimentally

→ theory (NLO): $9.2 \text{ fb}^{+30.8 \% \atop -25.6 \% \atop +5.5 \% \atop -5.9 \%}$ @ 13 TeV¹ (was $\sim 1 \text{ fb}$ @ 8 TeV)

¹J. Alwall et al., 10.1007/JHEP07(2014)079

The $t\bar{t}t\bar{t}$ final state

- $4t \rightarrow 4W + 4b$
- Channels:
 - single lepton $1\ell\nu_\ell$, 6 light jets, 4 b-jets
 - dilepton $2\ell\nu_\ell$, 4 light jets, 4 b-jets
 - trilepton $3\ell\nu_\ell$, 2 light jets, 4 b-jets



- Most sensitive analyses use:
 - single lepton channel
 - same-sign dilepton channel
(including trilepton)

Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

SM 4top Analyses @ 13 TeV

- Most of them use only 2015 data, and 2 of them 2015+2016 data

ATLAS

- **single lepton - 4top**
ATLAS-CONF-2016-020 3.3 fb^{-1}
- **single lepton - VLQ + 4top**
ATLAS-CONF-2016-013 3.3 fb^{-1}
- **single lepton - SUSY**
ATLAS-CONF-2017-013 36.1 fb^{-1}
- **same-sign dilepton - VLQ + 4top**
ATLAS-CONF-2016-032 3.3 fb^{-1}

CMS

- **single lepton + opposite-sign dilepton - 4top**
CMS-TOP-16-016 2.6 fb^{-1}
- **same-sign dilepton - SUSY**
CMS-SUS-16-035 35.9 fb^{-1}

- Few analyses dedicated to 4top
 - Indeed, those signatures can be used to study VLQ, SUSY, ...
 - ⇒ See Sergei's talk tomorrow on same-sign top production
- CMS analyses have combined their results on 4top for 2015 data

Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

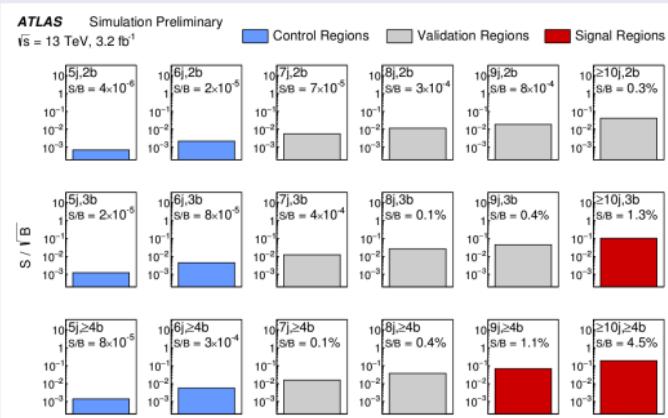
Single lepton analyses

- **Main bkg:** $t\bar{t} + \text{jets}$
- **Discriminant variables:** N_j total jet multiplicity and N_b b-jet multiplicity

ATLAS-CONF-2016-020

3.3 fb^{-1}

- **CR:** $N_j = 5, 6$ $N_b = 2, 3, \geq 4$
- **VR:** $N_j = 7, 8, 9, \geq 10$ $N_b = 2, 3, \geq 4$
- **SR:** $N_j = 9, \geq 10$ $N_b = 3, \geq 4$



- VR to check if $t\bar{t} + \text{jets}$ bkg (estimated from CR) can be extrapolated to SR

CMS-TOP-16-016

2.6 fb^{-1}

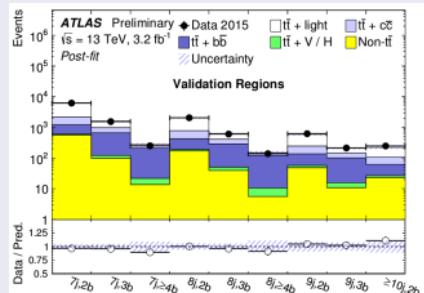
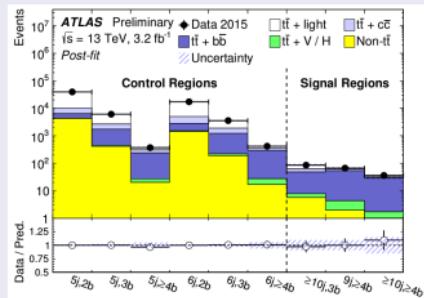
- Split into e and μ channels
- Regions splitted in N_j and N_b
 - $N_j = 6, 7, 8, \geq 9$
 - $N_b = 2, 3, \geq 4$

Signal extraction strategy and data/MC comparisons

ATLAS-CONF-2016-020

3.3 fb^{-1}

- Simultaneous fit of the H_T^{had} distributions in the 6 CR and 3 SR.



CMS-TOP-16-016

2.6 fb^{-1}

- Use a **BDT**, trained on $t\bar{t} +$ jets and $t\bar{t}t\bar{t}$. Variables used:
 - Event activity: N_j , H_T^b , ...
 - Event topology
 - N_b
- Simultaneous fit of BDT distributions in each region and channel.
- If fit of H_T distributions instead: results are 20 % worse \implies **BDT is better**.

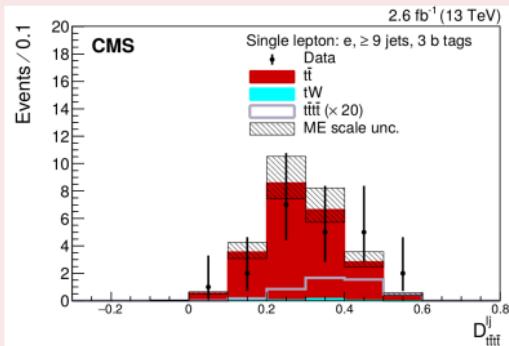


Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

Same-sign dilepton analyses

- **Main physics backgrounds:** $t\bar{t}V$, VV
- **Instrumental backgrounds:** charge mis-identification, fake/non-prompt leptons

ATLAS-CONF-2016-032

3.3 fb^{-1}

Definition			Name
$e^\pm e^\pm + e^\pm \mu^\pm + \mu^\pm \mu^\pm + eee + ee\mu + e\mu\mu + \mu\mu\mu$, $N_{\text{jets}} \geq 2$			
$400 < H_T < 700 \text{ GeV}$	$N_b = 1$		SR0
	$N_b = 2$	$E_T^{\text{miss}} > 40 \text{ GeV}$	SR1
	$N_b \geq 3$		SR2
$H_T \geq 700 \text{ GeV}$	$N_b = 1$	$40 < E_T^{\text{miss}} < 100 \text{ GeV}$	SR3
		$E_T^{\text{miss}} \geq 100 \text{ GeV}$	SR4
	$N_b = 2$	$40 < E_T^{\text{miss}} < 100 \text{ GeV}$	SR5
		$E_T^{\text{miss}} \geq 100 \text{ GeV}$	SR6
	$N_b \geq 3$	$E_T^{\text{miss}} > 40 \text{ GeV}$	SR7

CMS-SUS-16-035

35.9 fb^{-1}

- 100 signal regions!
- Splitted in:
 - H_T = sum of jets p_T
 - N_b = b-jet multiplicity
 - N_j = jet multiplicity
 - E_T^{miss}
 - charge: ++ or --
 - high (H) and low (L) lepton momentum: HH, LL and HL
- Normalizations for $t\bar{t}V$ and VV backgrounds **estimated from data** in a control region.
- Simultaneous fit of the yields in several regions to extract the signal

- Regions splitted in:

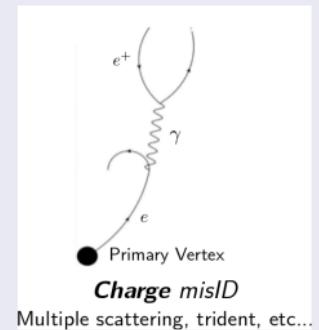
- H_T = sum of jets and leptons p_T
- N_b = b-jet multiplicity
- E_T^{miss}

- Simultaneous fit of the yields in several regions to extract the signal

Data driven instrumental background (mostly from $t\bar{t}$ events)

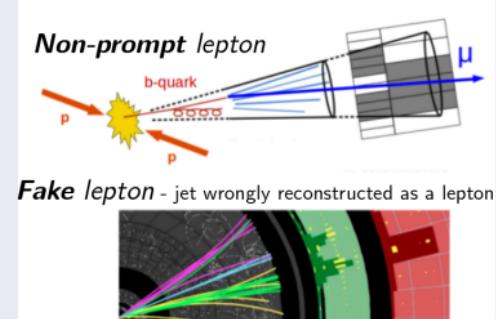
Charge mis-identification

- **Mis-ID rates** are calculated from a $Z \rightarrow ee$ sample, then extrapolated to SR
ATLAS: data driven **CMS**: estimation from MC
- The **mis-ID background** is estimated from opposite-sign CR for each SR, using these rates



Fakes/non-prompt (same method for both)

- **2 lepton definitions:** Loose (l) and Tight (t) = Loose + extra isolation
- **Calculate probabilities** that a loose real/fake lepton is tagged as tight in CR
- **Fake yields** are then estimated from loose data events in SR



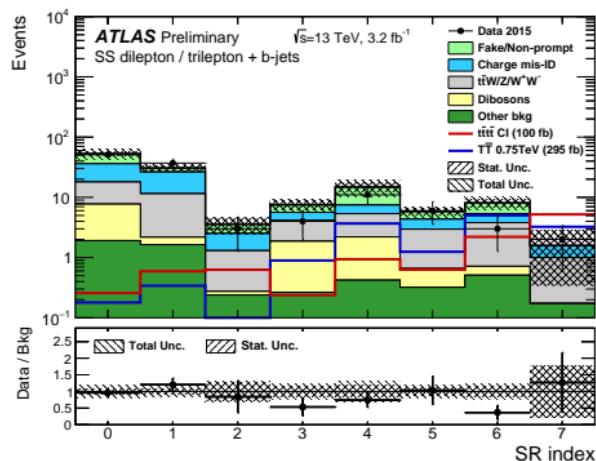
Data/MC comparison

ATLAS-CONF-2016-032

3.3 fb^{-1}

	SR0	SR1	SR2	SR3	SR4
Fake/Non-prompt	16.3 ± 9.5	4.2 ± 3.3	1.0 ± 0.9	1.8 ± 1.4	7.1 ± 4.5
Charge mis-ID	18.1 ± 4.1	14.9 ± 3.5	1.2 ± 0.3	1.5 ± 0.4	2.1 ± 0.5
$t\bar{t}W/Z/W^+W^-$	10.1 ± 1.4	9.2 ± 1.3	1.0 ± 0.3	2.2 ± 0.3	3.1 ± 0.5
Dibosons	5.8 ± 1.0	0.5 ± 0.2	0.03 ± 0.07	1.6 ± 0.4	1.8 ± 0.4
Other bkg.	2.0 ± 1.0	1.7 ± 0.9	0.3 ± 0.2	0.3 ± 0.2	0.5 ± 0.3
Total bkg.	52 ± 11	31 ± 5	3.6 ± 1.0	7.4 ± 1.5	15 ± 5
$t\bar{t}t\bar{t}$ (SM)	0.5 ± 0.1	0.8 ± 0.1	0.9 ± 0.1	0.2 ± 0.1	0.5 ± 0.1
Data	51	37	3	4	11

	SR5	SR6	SR7
Fake/Non-prompt	1.4 ± 0.9	2.6 ± 1.8	0.0 ± 0.6
Charge mis-ID	1.4 ± 0.4	1.6 ± 0.5	0.6 ± 0.2
$t\bar{t}W/Z/W^+W^-$	2.3 ± 0.6	3.0 ± 0.7	0.8 ± 0.4
Dibosons	0.3 ± 0.1	0.2 ± 0.1	0.0 ± 0.1
Other bkg.	0.4 ± 0.2	0.7 ± 0.4	0.5 ± 0.3
Total bkg.	5.8 ± 1.2	8.1 ± 2.0	1.9 ± 0.8
$t\bar{t}t\bar{t}$ (SM)	0.7 ± 0.1	1.8 ± 0.2	3.6 ± 0.4
Data	6	3	2



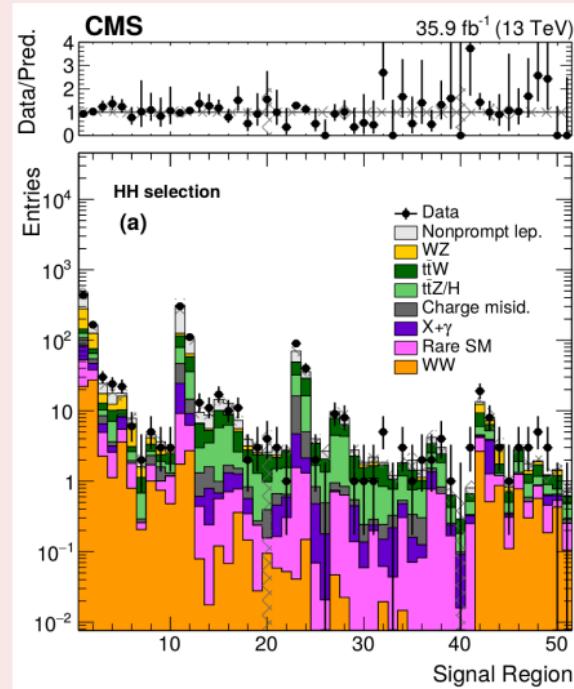
For these yields, $\sigma_{t\bar{t}t\bar{t}}^{SM}$ normalized to 100 fb

Data/MC comparison

CMS-SUS-16-035

35.9 fb^{-1}

Only HH signal regions shown



Systematics

Main systematics:

- normalization + scale and pdf variations of ttV and VV backgrounds
- fakes/non-prompt and misID systematics

ATLAS-CONF-2016-032

3.3 fb⁻¹

Systematics on bkg

Source	Signal region							
	SR0	SR1	SR2	SR3	SR4	SR5	SR6	SR7
Cross section	8	11	26	13	9	27	23	57
Jet energy scale	1	1	3	1	1	3	2	4
Jet energy resolution	<1	2	2	2	<1	1	<1	3
b-tagging efficiency	1	2	5	3	1	2	2	7
Luminosity	1	1	1	1	1	1	1	1
Fake/non-prompt leptons	17	7	15	13	26	13	17	17
Charge misID	8	3	7	5	3	6	5	8

Systematics on signal

Source	Signal region							
	SR0	SR1	SR2	SR3	SR4	SR5	SR6	SR7
Jet energy scale	2	12	2	6	4	3	3	3
Jet energy resolution	16	6	7	16	14	11	1	2
b-tagging efficiency	8	5	5	21	14	15	5	5
Lepton ID efficiency	1	1	1	4	2	2	2	1
Luminosity	2	2	2	2	2	2	2	2

CMS-SUS-16-035

35.9 fb⁻¹

Source	Typical uncertainty (%)
Integrated luminosity	2.5
Lepton selection	4 – 10
Trigger efficiency	2 – 7
Pileup	0 – 6
Jet energy scale	1 – 15
b tagging	1 – 15
Simulated sample size	1 – 10
Scale and PDF variations	10 – 20
WZ (normalization)	12
t̄tZ (normalization)	30
Nonprompt leptons	30 – 60
Charge misidentification	20

Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

Upper limits on $\sigma_{t\bar{t}t\bar{t}}^{SM}$ at 95 % CL

- **SM value** computed at NLO precision is 9.2 fb

Channels		ATLAS		CMS	
		3.3 fb ⁻¹	36.1 fb ⁻¹	2.6 fb ⁻¹	35.9 fb ⁻¹
1 - Single lepton	dedicated	190 (143)		158 (151)	
	VLQ+4top	370 (180)			
	SUSY		60 (84)		
2 - Dilepton OS				134 (227)	
1 & 2				94 (118)	
3 - Dilepton SS		95 (107)		119 (101)	42 (27)
1 & 2 & 3				69 (71)	

Table: **Observed** (expected) upper limits on $\sigma_{t\bar{t}t\bar{t}}^{SM}$ at 95 % CL (**in fb**)

- ATLAS and CMS analyses have **very similar sensitivities** in each channel.
- **Best limit** to date from CMS dilepton SS analysis: **42 (27) fb**, or **4.6× (3.0×) the theory** prediction.

Table of Contents

1 Introduction

2 SM 4top Analyses @ 13 TeV

3 Analyses Strategies

- Single lepton analyses
- Same-sign dilepton analyses

4 Results

5 Conclusion

Conclusion

- **SM 4top production** is studied by ATLAS and CMS in single lepton and same-sign dilepton channels
 - both experiments achieve **similar sensitivities** with 2015 data
 - **no excess** found yet
 - **best limit** from CMS same-sign dilepton analysis 2016:
$$42 \text{ (27) fb, or } 4.6 \times (3.0 \times) \sigma_{t\bar{t}t\bar{t}, \text{ theory}}^{SM}$$
- These searches are **very sensitive** to a wide variety of BSM signals
- Sensitivities will be **very close** to $\sigma_{t\bar{t}t\bar{t}}^{SM}$ with the **full Run 2 dataset**

⇒ Looking forward to all the results with 2016 data, and then using the full Run 2 dataset

BACKUP

Table of Contents

6 More info

- Mis-ID
- Fakes

7 Pull plots for ATLAS same-sign analysis

Table of Contents

6 More info

- Mis-ID
- Fakes

7 Pull plots for ATLAS same-sign analysis

Charge Mis-identification for ATLAS same-sign dilepton analysis

- Wrong identification of a particle's charge (typically an electron)
→ some opposite-sign (OS) events are taken for SS events.

Estimation: data driven method

- Measure charge flip rate ϵ in $Z \rightarrow ee$ data events (81 - 101 GeV)
- Binned in electron η , p_T : Poisson likelihood minimization in each bin

$$N_{SS}^{ij} = N_{OS(true)}^{ij}(\epsilon_i(1 - \epsilon_j) + \epsilon_j(1 - \epsilon_i))$$

- Weight observed OS events to predict QMisID background:

$$\text{weight} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j}$$

Remarks

- Only consider **ee and $e\mu$ events** (Mis-ID rate is negligible for muons)
- Does not concern trilepton channels.

Likelihood for Mis-ID (ATLAS same-sign dilepton analysis)

(ref <https://cds.cern.ch/record/1635458/>)

$$N_{SS}^{ij} = (\epsilon_i(1 - \epsilon_j) + \epsilon_j(1 - \epsilon_i)) N_{OS(true)}^{ij} = \underbrace{\frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j}}_{\text{weight}} N_{OS(observed)}^{ij}$$

Poisson distribution: $f(k, \lambda) = \frac{\lambda^k \exp^{-\lambda}}{k!}$

k Observed number of occurrences, i.e. N_{SS}^{ij}

λ Expected mean, i.e. $N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)$, where N^{ij} is $N_{OS(true)}^{ij}$.

So, the probability to get (ϵ_i, ϵ_j) , given N_{SS}^{ij} and N^{ij} is:

$$P(\epsilon_i, \epsilon_j | N_{SS}^{ij}, N^{ij}) = \frac{[N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)]^{N_{SS}^{ij}} \exp(-N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j))}{N_{SS}^{ij}!}$$

So, the likelihood to maximize (log-likelihood to minimize) is:

$$L(\epsilon_i, \epsilon_j | N_{SS}^{ij}, N^{ij}) = \prod_{i,j} \frac{[N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)]^{N_{SS}^{ij}} \exp(-N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j))}{N_{SS}^{ij}!}$$

$$-\ln L(\epsilon_i, \epsilon_j | N_{SS}^{ij}, N^{ij}) \approx \sum_{i,j} \ln[N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)] N_{SS}^{ij} - N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)$$

where we removed the terms that do not depend on the ϵ 's in the log-likelihood

Table of Contents

6 More info

- Mis-ID
- Fakes

7 Pull plots for ATLAS same-sign analysis

Fake/non-prompt leptons for ATLAS same-sign dilepton analysis

- Fake lepton: jet wrongly reconstructed as lepton.
- Non-prompt lepton: secondary lepton (radiated), wrongly reconstructed as coming from the primary vertex.

Estimation: data driven method (Likelihood Matrix Method)

- 2 lepton definitions: Loose (l) and Tight (t), used in signal regions
- Tight = Loose + extra isolation criteria
- Probability that a loose real/fake lepton is tagged as tight: ϵ_r/ϵ_f (binned in η , p_T , $\Delta R(\ell, 1^{st} \text{jet})$)
- e.g. for 1 lepton:

$$N^l = N_r^l + N_f^l$$
$$N^t = \epsilon_r N_r^l + \epsilon_f N_f^l$$

Likelihood Matrix Method (ATLAS same-sign dilepton analysis)

For 2 leptons:

$$\begin{pmatrix} N^{tt} \\ N^{t\bar{t}} \\ N^{\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}} \end{pmatrix} = \mathbf{M} \begin{pmatrix} N_{rr}^{ll} \\ N_{rf}^{ll} \\ N_{fr}^{ll} \\ N_{ff}^{ll} \end{pmatrix} \quad \mathbf{M} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1 \bar{r}_2 & r_1 \bar{f}_2 & \bar{f}_1 r_2 & \bar{f}_1 \bar{f}_2 \\ \bar{r}_1 r_2 & \bar{r}_1 f_2 & \bar{f}_1 r_2 & \bar{f}_1 f_2 \\ \bar{r}_1 \bar{r}_2 & \bar{r}_1 \bar{f}_2 & \bar{f}_1 \bar{r}_2 & \bar{f}_1 f_2 \end{pmatrix}$$

For 3 leptons:

$$\begin{pmatrix} N^{ttt} \\ N^{t\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \\ N^{\bar{t}\bar{t}\bar{t}} \end{pmatrix} = \begin{pmatrix} r_1 r_2 r_3 & r_1 r_2 f_3 & r_1 f_2 r_3 & r_1 f_2 f_3 & f_1 r_2 r_3 & f_1 r_2 f_3 & f_1 f_2 r_3 & f_1 f_2 f_3 \\ r_1 r_2 \bar{r}_3 & r_1 r_2 \bar{f}_3 & r_1 \bar{f}_2 \bar{r}_3 & r_1 \bar{f}_2 \bar{f}_3 & f_1 r_2 \bar{r}_3 & f_1 r_2 \bar{f}_3 & f_1 \bar{f}_2 \bar{r}_3 & f_1 \bar{f}_2 \bar{f}_3 \\ r_1 \bar{r}_2 r_3 & r_1 \bar{r}_2 f_3 & r_1 f_2 r_3 & r_1 f_2 f_3 & f_1 \bar{r}_2 r_3 & f_1 \bar{r}_2 f_3 & f_1 \bar{f}_2 r_3 & f_1 \bar{f}_2 f_3 \\ r_1 \bar{r}_2 \bar{r}_3 & r_1 \bar{r}_2 \bar{f}_3 & r_1 \bar{f}_2 \bar{r}_3 & r_1 \bar{f}_2 \bar{f}_3 & f_1 r_2 \bar{r}_3 & f_1 r_2 \bar{f}_3 & f_1 \bar{f}_2 \bar{r}_3 & f_1 \bar{f}_2 \bar{f}_3 \\ \bar{r}_1 r_2 r_3 & \bar{r}_1 r_2 f_3 & \bar{r}_1 f_2 r_3 & \bar{r}_1 f_2 f_3 & \bar{f}_1 r_2 r_3 & \bar{f}_1 r_2 f_3 & \bar{f}_1 f_2 r_3 & \bar{f}_1 f_2 f_3 \\ \bar{r}_1 r_2 \bar{r}_3 & \bar{r}_1 r_2 \bar{f}_3 & \bar{r}_1 \bar{f}_2 r_3 & \bar{r}_1 \bar{f}_2 f_3 & \bar{f}_1 r_2 \bar{r}_3 & \bar{f}_1 r_2 \bar{f}_3 & \bar{f}_1 \bar{f}_2 r_3 & \bar{f}_1 \bar{f}_2 f_3 \\ \bar{r}_1 \bar{r}_2 r_3 & \bar{r}_1 \bar{r}_2 f_3 & \bar{r}_1 \bar{f}_2 \bar{r}_3 & \bar{r}_1 \bar{f}_2 \bar{f}_3 & \bar{f}_1 \bar{r}_2 r_3 & \bar{f}_1 \bar{r}_2 f_3 & \bar{f}_1 f_2 \bar{r}_3 & \bar{f}_1 f_2 \bar{f}_3 \\ \bar{r}_1 \bar{r}_2 \bar{r}_3 & \bar{r}_1 \bar{r}_2 \bar{f}_3 & \bar{r}_1 \bar{f}_2 \bar{r}_3 & \bar{r}_1 \bar{f}_2 \bar{f}_3 & \bar{f}_1 \bar{r}_2 \bar{r}_3 & \bar{f}_1 \bar{r}_2 \bar{f}_3 & \bar{f}_1 f_2 \bar{r}_3 & \bar{f}_1 f_2 \bar{f}_3 \end{pmatrix} \begin{pmatrix} N_{rrr}^{lll} \\ N_{rrf}^{lll} \\ N_{rfl}^{lll} \\ N_{rlf}^{lll} \\ N_{fff}^{lll} \\ N_{ffr}^{lll} \\ N_{rff}^{lll} \\ N_{lff}^{lll} \end{pmatrix}$$

Table of Contents

6 More info

- Mis-ID
- Fakes

7 Pull plots for ATLAS same-sign analysis

Pull plots for ATLAS same-sign analysis

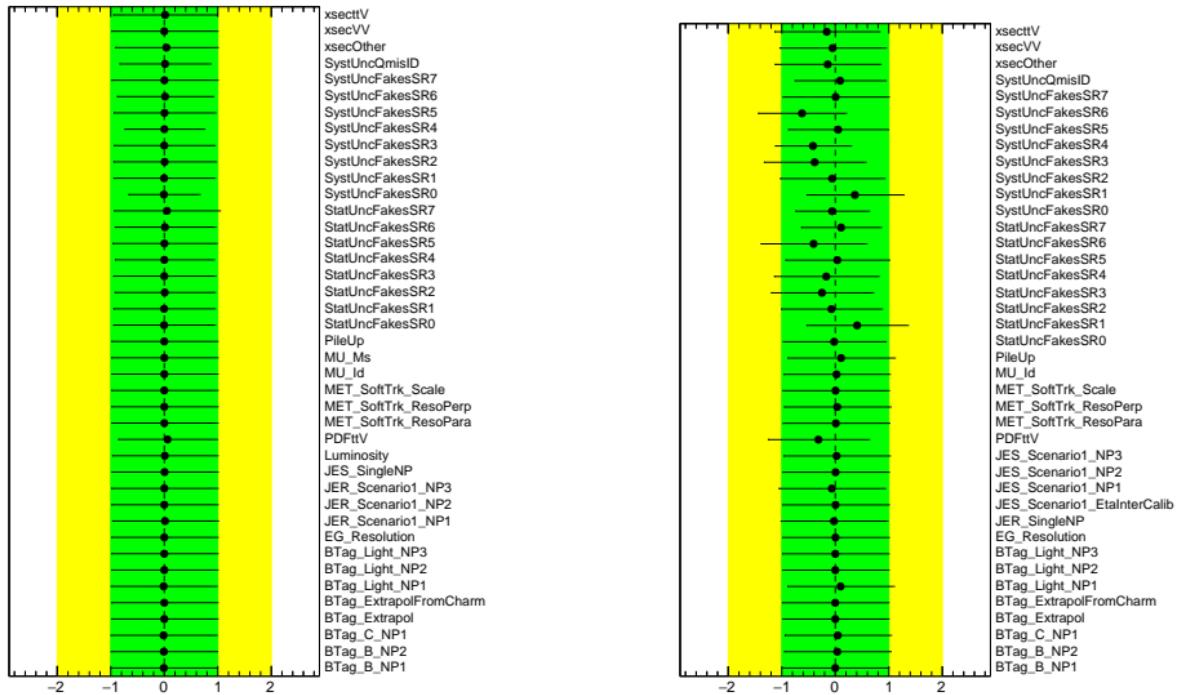


Figure: Post-fit nuisance parameters for a fit the expected background (left) and for a fit on observed data (right).