Search for vector like quarks and heavy resonances decaying to top quarks

Deep Inelastic Scattering conference 2017

Clément Camincher
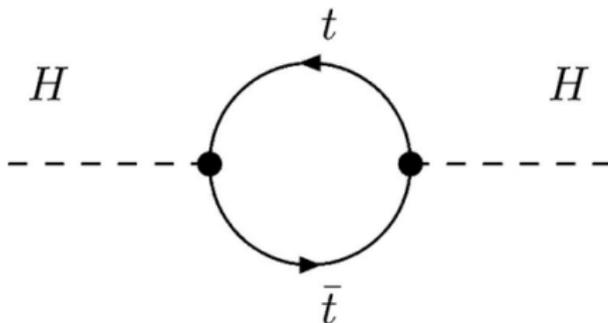
On behalf of the ATLAS collaboration

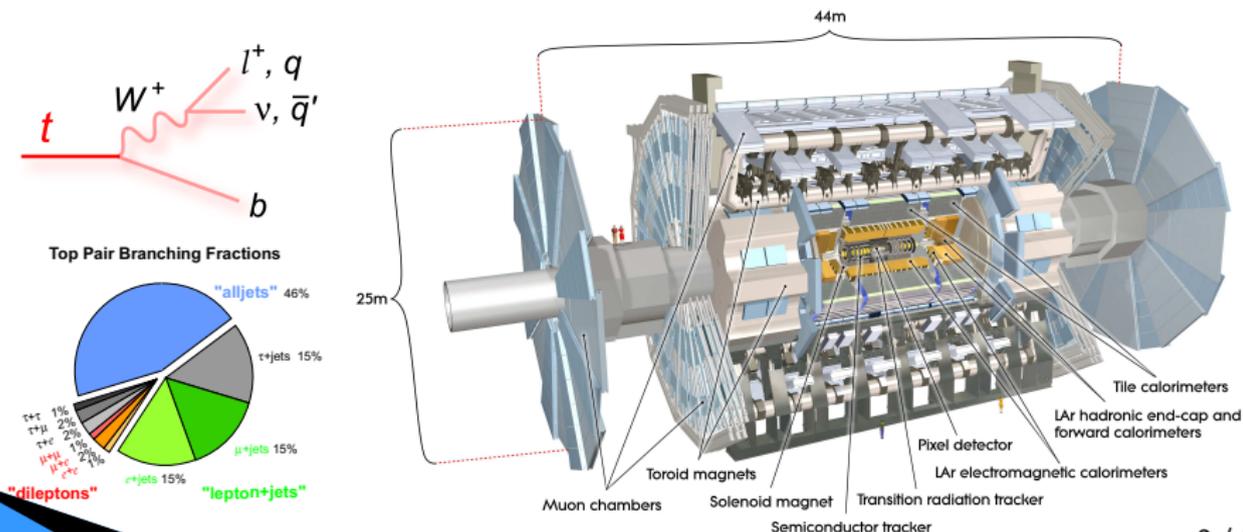5/04/2017

Why using top quarks to probe new physics ?

Top quarks play an important role at the electroweak energy scale

▶ Top quarks are very heavy ≈ 173 *GeV*
⇒ Strong Yukawa coupling with the Higgs boson
⇒ Major piece in the naturalness issue
▶ Many BSM scenarios predict:
  ▶ New couplings involving tops ⇒ $t\bar{t}$ **resonances searches**
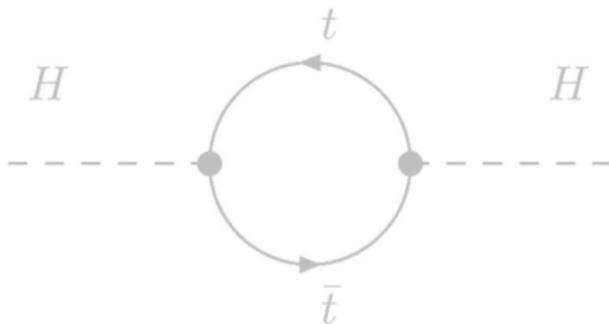 ▶ Top partners ⇒ **VLQ searches**

ATLAS a top detector at the LHC

- At 13TeV, the LHC is a top-quark factory
  (1 $t\bar{t}$ pair produced every second)
- ATLAS detector optimized to find top-quarks
  - Very good electron and muon reconstruction and identification
    Present in 39% of top decays
  - Full azimuthal coverage to measure neutrino transverse component via E_T^{miss}
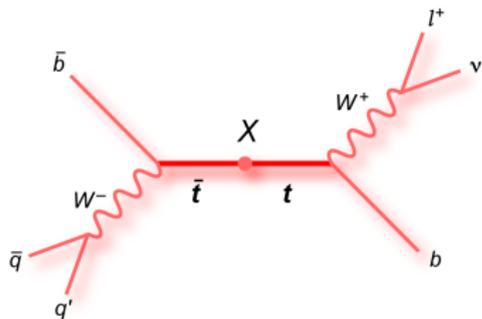 - Close tracker layer (IBL) to select displaced vertexes from b-jets

Top Pair Branching Fractions

"alljets" 46%
τ+jets 15%
μ+jets 15%
e+jets 15%
τ+τ 1%
τ+μ 2%
τ+e 2%
μ+μ 1%
e+μ 2%
e+e 1%
"dileptons"
"lepton+jets"

44m

25m

Tile calorimeters

LAr hadronic end-cap and
forward calorimeters

Pixel detector

LAr electromagnetic calorimeters

Toroid magnets   Transition radiation tracker

Muon chambers   Solenoid magnet

Semiconductor tracker

Top quarks play an important role at the electroweak energy scale

- Top quarks are very heavy ≈ 173 *GeV*
- $\Rightarrow$ Strong Yukawa coupling with the Higgs boson
- $\Rightarrow$ Major piece in the naturalness issue
- Many BSM scenarios predict:
  - New couplings involving tops $\Rightarrow$ $t\bar{t}$ **resonances searches**
  - Top partners $\Rightarrow$ **VLQ searches**

Concept of the $t\bar{t}$ resonances searches

▶ Many BSM scenarios predict new particles (X) decaying into a pair of tops

  ▶ Z' technicolor assisted topcolor
  ▶ Randall Sundrum $\rightarrow$ gluon/graviton: g_{KK}/G_{KK}
  ▶ 2 Higgs Doublet Models $\rightarrow$ heavy pseudoscalar/scalar A/H

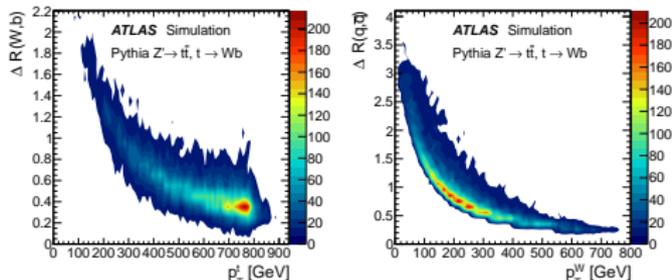Analysis strategy :
Build a model independent analysis

1. Select events
2. Reconstruct $t\bar{t}$ invariant mass
3. Scan $m_{t\bar{t}}$ to find an excess/deficit

▶ Lepton + jet channel choice

  ▶ Lepton $\Rightarrow$ multijet background suppression (trigger more efficient on leptons)
  ▶ Jets $\Rightarrow$ keep good statistics level (high branching ratio)
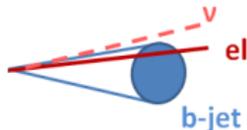
Lepton +jets boosted $t\bar{t}$ search

- ▶ Results using 3.2 fb^{-1} of 13 TeV data
- ▶ Analysis dedicated to probe heavy resonances
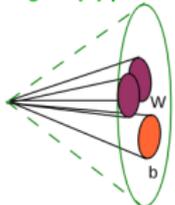- ⇒ Producing boosted top
- ⇒ collimated decay products

- ▶ **Boosted leptonic top selection**
  - ▶ Select exactly 1 muon or 1 electron
  - ▶ b-jet selected as highest pt jet in $\Delta R(el; jet) < 1.5$
  - ▶ Neutrino transverse momentum estimated with E_T^{miss}
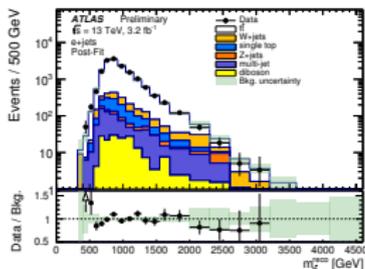 - ▶ Neutrino longitudinal momentum estimated constraining the system with the on-shell W mass.

- ▶ **Boosted hadronic top selection**
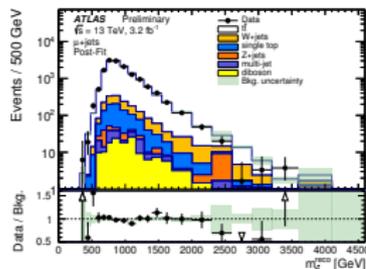 - ▶ 1 top-tagged (80%) fat jet (anti-kt with R=1.0)
  - ▶ Large jets are trimmed by removing kt sub-jets with pt fraction below 5%
  - ▶ Tagging ⇔ cut on jet mass and τ_{32} "nsubjetiness" ratio (describing the internal structure of the jet)

Lepton +jets boosted $t\bar{t}$ analysis

- Dominant background is SM $t\bar{t}$
- Multijet estimated with data
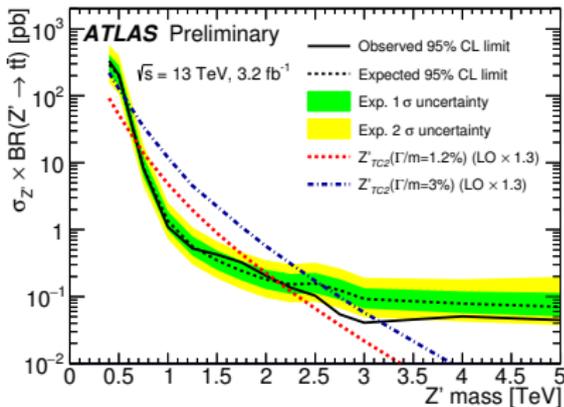- W+jets estimation
  - Shape from MC
  - Scaling from data
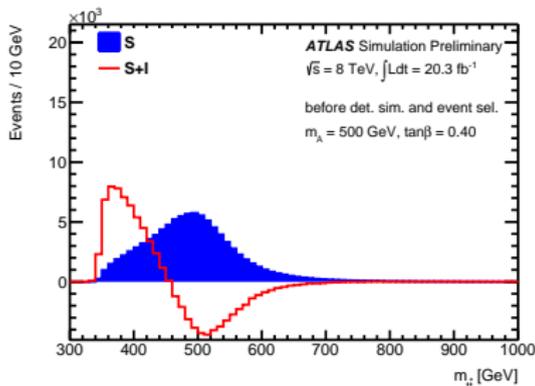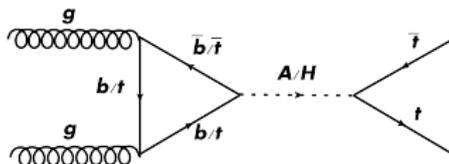- Main systematics from fat jet modeling

Electron selection

Muon selection

- Fit simultaneously both lepton distributions
- No deviation found
- Limit set on a Z' $TC2$
- Z' with $\Gamma/m = 1.2\%$ excluded in the region $0.7 < m_{Z'} < 2\,TeV$

Re-interpretation 8 TeV (20 fb^{-1}**)** $t\bar{t}$ **resonances search**

- 2HDM model predicts Scalar (H) or Pseudoscalar (A) decaying into $t\bar{t}$
- Large interference (I) with SM $gg \rightarrow t\bar{t}$ production
- It would require signal (S)+ (I) + background (B) simulation at parton level for each mass point

- Generator modified to produce only S+I
- $m_{t\bar{t}}$ Signal modeling at parton level versus Signal + Interference modeling.
- Validated with S+I+B generation for some parameter space points

Difference treated as systematics uncertainties (0.4%)

$H/A \to t\bar{t}$ analysis setup and results

Resolved selection

Low top p_T

- Exactly 1 lepton (electron or muon)
- 4 jets (1 or 2 b-tags)
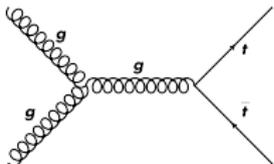- Jet ambiguity solved by minimizing a χ^2 over jet combinations

Invariant mass

muon selection

electron selection

Limit on the signal strength μ versus $\tan(\beta)$, $m_{H/A} = 500\,GeV$

Pseudo scalar

Scalar

Top quarks play an important role at the electroweak energy scale

▶ Top quarks are very heavy ≈ 173 *GeV*
⇒ Strong Yukawa coupling with the Higgs boson
⇒ Major piece in the naturalness issue
▶ Many BSM scenarios predict:

  ▶ New couplings involving tops $\Rightarrow t\bar{t}$ **resonances searches**
  ▶ Top partners $\Rightarrow$ **VLQ searches**

Vector Like Quarks (VLQ) searches

- To solve the hierarchy problem many BSM scenarios predict VLQ
- They don't receive mass through BEH mecanism
- VLQ would contribute to Higgs mass loop correction

ATLAS search for VLQ ($T\bar{T}$) pair production

- T has same electric charge than top-quarks ($Q=+2/3$)
- Considering only T decays :

$$T \rightarrow Zt, \; Ht, \; Wb$$

Three analysis to probe each channel preferentially

Zt+X	**Ht+X**	**Wb+X**
ATLAS-CONF-2017-015	ATLAS-CONF-2016-104	ATLAS-CONF-2016-102

Global Strategy

▶ Preselection (detail per analysis in backup):

  ▶ 1 lepton ($W \to e\nu$) or/and large $\mathrm{E_T^{miss}}$ ($Z \to \nu\nu$)
  ▶ Several jets ($W \to qq$, $H \to bb$, $t \to qqb$)
  ▶ 2 or more b-jets ($H \to bb$, $t \to Wb$)

▶ Background modeling

  ▶ $t\bar{t}$ (+jets) estimated using MC and constrained via Control Regions (CR)
  ▶ Multijet contribution estimated with data driven method

▶ Signal Regions (SR)

  ▶ Optimized to probe the desired channel
  ▶ Additional cuts on high level variables (jet sub-structure, objects combination ...)
  ▶ Several signal bins to probe several final states

Analysis specificities

Zt+X (1-lepton)

- Use am_{T2}, m_T^W to reduce background
- Leading to a single bin SR
- $t\bar{t}$ and W+jets CR used in the fit

Ht+X (1-lepton and 0-lepton)

- Several SR defined:
- For instance cut on $N_{mass\ tagged\ jet}$
- "mass tagged" jet $\Leftrightarrow m(fat\ jet) > 100 GeV$
- For each SR, search for bump in: m_{eff}
  $m_{eff} = p_T^{lep} + p_T^{jets} + E_T^{miss}$

Wb+X (1-lepton)

- Defined a boosted and a resolved SR
- Search for bump in m_T^{lep} distribution
  for $T \to Wb \to l\nu b$

Post-fit distributions

Zt+X   **Ht+X**   **Wb+X**

VR   **SR**

1-lep SR   **0-lep SR**

SR   **CR**

▶ Signal enriched regions (SR) fitted simultaneously with **background enriched regions**

▶ No deviation found

Observed Mass limit vs Branching ratios

Zt+X

Ht+X

- ▶ Main sytematics are jet modeling and heavy-flavours tagging

- ▶ Each analysis well sensitive to the branching ratio (BR) it is dedicated to probe

- ▶ Excluding up to 1.1 TeV for BR=100%

- ▶ Wb+X (BR=100%) limit can also be reinterpreted in terms of $Y_{-4/3}$ production

Wb+X

Search for same sign leptons producion
ATLAS-CONF-2016-032

Probe scenarios with at least two leptons of same sign

- Very rare in SM production
- Allow to probe vector like top(T) and vector like bottom ($B \to Wt$, Zb, Hb)
- Can also probe $T_{5/3} \to W^+ t \to W^+ W^+ b$ pair production
- Selection
  - $\geqslant 2$ leptons with at least 2 with same charge
  - Large E_T^{miss}
  - At least 2 jets with at least 1 b-tagged jet
- Dominant background
  - Charge mis-identification estimated in $Z \to ee$ resonance
  - Fake lepton estimated with data driven method

No excess found
⇒ set limits

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

13 / 16

Observed limits on VLQs mass versus the decay branching ratio

Vector like bottom (B)

Vector like top (T)

- ▶ Main systematics from the fake lepton estimation
- ▶ Limit on VLQ $T_{5/3}$ pair cross section versus its mass
- ▶ The production of $T_{5/3}\bar{T}_{5/3}$ production is excluded for a mass up to 0.99 TeV

Conclusion

$t\bar{t}$ resonance search

- No evidence so far of $t\bar{t}$ resonance at the TeV scale
- Only 10% of available data already analysed
- With 2016 luminosity may probe higher mass regions and lower cross sections
- In addition the resolved channel will also be added
- Keep improving the methods to reconstruct top quarks

VLQ search

- Several channels studied without evidence of VLQ production
- Previous searches have been significantly improved with Run-2 data
- VLQ with mass of $\approx 1\,TeV$ are now excluded in some regions
- The results will be combined

Selection

Var	Boosted electron selection	Boosted Muon selection
Lepton		
N-lepton	==1	
Pt lepton [GeV]	>30	>25GeV
Quality	LHTight	Medium
Isolation	Track based with pt dependant cone size	
η lepton	>2.47 and not in $1.37 < |\eta| < 1.52$ (crack)	
Overlapp removal	$\Delta R(el; jet) > 0.4$	$\Delta R(\mu; jet) > 0.04 + 10/pt_\mu$
Leptonic b-jet	Highest pt jet within $\Delta R(el; jet)1.5$	
Hadronic fat jet		
N top tagged fat jet	>1	
Top-Taging	80% efficiency using mass and τ_{32}	
Trimming	R sub = 0.2 and f cut = 0.05	
b-tagging		
Jets	Track jets antikt R=0.2	
WP	MV2c20 algorithm 70% efficiency	

W+jets estimation

$N_{W^+} + N_{W^-} = \left(\frac{r_{MC}+1}{r_{MC}-1} \right) (D_+ - D_-)$

With $r_{MC} = \frac{N_{MC}^+}{N_{MC}^-}$ and $D_\pm$ the number of event in data with lepton charge of $\pm$

Systematics contribution in $t\bar{t}$ res. search

Generator modification:
The modified generator is MadGraph5_aMC@NLO

Systematics

JES/JER (SM $t\bar{t}$)	6%
JES/JER (Signal S+I)	8%
JES/JER (Signal S)	4%
PDF on signal	12.3%
b-tag (bkg)	2%
b-tag (signal)	1%
$t\bar{t}$ cross section	6.5%
$t\bar{t}$ Parton Showering	5%
multijet bkg norm	20%
ST bkg norm	7.7 %
Z+jet bkg norm	48%
Signal remorm	7.5%
Inerference modeling	0.4%

To find better jet lepton combination
$\Rightarrow$ minimize χ^2

$$\chi^2 = \left[\frac{m_{jj}-m_W}{\sigma_W}\right]^2 + \left[\frac{m_{jjb}-m_{jj}-m_{th-W}}{\sigma_{th-W}}\right]^2 +$$

$$\left[\frac{m_{j\ell\nu}-m_{tl}}{\sigma_{t\ell}}\right]^2 + \left[\frac{(p_{T,jjb}-p_{T,j\ell\nu})-(p_{T,th}-p_{T,t\ell})}{\sigma_{diffp_T}}\right]^2$$

1. term constraint the W jet mass
2. term constraint the top-W mass
3. term constraint the top mass
4. term constraint the pt baance between leptonic and hadronic side

Multijet background estimation

- Multijet background is dominant at LHC
- Analysis design to supress it a maximum
- Not enough MC to run representative amound of QCD bkg
- Fakin lepton is not well modeled (else it would have been corrected!)
- Measure QCD background in data direcly

Matrix method

- define Tight and Loose lepton selection
- Measure ϵ (f): the fraction of real(fake) lepton selected as tight lepton
- ϵ (f) is measured in a signal (background) enriched region

$\epsilon = N_{tight}^{real}/N_{loose}^{real}$
$f = N_{tight}^{fake}/N_{loose}^{fake}$
The total number of Tight event is
$N_{tight} = \epsilon \times N_{prompt} + f \times N_{fakes}$
$N_{loose} = N_{prompt} + N_{fakes}$
By inverting the system it yield N_{prompt} and N_{fakes}

Signal region definition

Variable	SR	TCR	WCR
$E_{\mathrm{T}}^{\mathrm{miss}}$	$> 350\,GeV$	$> 300\,GeV$	
m_{T}^{W}	$> 170\,GeV$	$m_{\mathrm{T}}^{W} \in [30, 90]\,GeV$	
am_{T2}	$> 175\,GeV$	$> 100\,GeV$	
$H_{\mathrm{T,sig}}^{\mathrm{miss}}$	> 12	–	
small-R jet p_{T}	$> 120, 80, 50, 25\,GeV$	$> 120, 80, 50, 25\,GeV$	
number of b-tagged jets	≥ 1	≥ 1	$= 0$
number of large-R jets	≥ 2	≥ 2	
large-R jet mass	$> 80, 60\,GeV$	$> 80, 60\,GeV$	
large-R jet p_{T}	$> 290, 290\,GeV$ if $E_{\mathrm{T}}^{\mathrm{miss}} < 450\,GeV$	$> 200\,GeV$	
	$> 200, 200\,GeV$ if $E_{\mathrm{T}}^{\mathrm{miss}} > 450\,GeV$		

Systematics:

Jet modeling (scale and resolution)	
Also b, c, light-flavor jet tagging.	(1-5%)
MC $t\bar{t}$ uncert in SR	18% (dominat bkg uncert for $t\bar{t}$)
Interference modelling btw $t\bar{t}$ and ST	11% (dominant for ST modelling)
W+jet modeling	6% on backg modelling
PDF	6%

Preselection

Preselection requirements		
Requirement	1-lepton channel	0-lepton channel
Trigger	Single-lepton trigger	E_T^{miss} trigger
Leptons	=1 isolated e or μ	=0 isolated e or μ
Jets	≥5 jets	≥6 jets
b-tagging	≥2 b-tagged jets	≥2 b-tagged jets
E_T^{miss}	$E_T^{miss} > 20$ GeV	$E_T^{miss} > 200$ GeV
Other E_T^{miss}-related	$E_T^{miss} + m_T^W > 60$ GeV	$\Delta\phi_{min}^{4j} > 0.4$

0-lep SRs

Search regions (≥7 jets)			
Mass-tagged jet multiplicity	b-jet multiplicity	$m_{T,min}^b$	Channel name
0	2	-	0J, ≥7j, 2b
0	3	-	0J, ≥7j, 3b
0	≥4	-	0J, ≥7j, ≥4b
1	2	-	1J, ≥7j, 2b
1	3	< 160 GeV	1J, ≥7j, 3b, LM
1	3	> 160 GeV	1J, ≥7j, 3b, HM
1	≥4	< 160 GeV	1J, ≥7j, ≥4b, LM
1	≥4	> 160 GeV	1J, ≥7j, ≥4b, HM
≥2	2	-	≥2J, ≥7j, 2b
≥2	3	< 160 GeV	≥2J, ≥7j, 3b, LM
≥2	3	> 160 GeV	≥2J, ≥7j, 3b, HM
≥2	≥4	-	≥2J, ≥7j, ≥4b
Validation regions (6 jets)			
Mass-tagged jet multiplicity	b-jet multiplicity	$m_{T,min}^b$	Channel name
0	2	-	0J, 6j, 2b
0	3	-	0J, 6j, 3b
0	≥4	-	0J, 6j, ≥4b
1	2	-	1J, 6j, 2b
1	3	-	1J, 6j, 3b
1	≥4	-	1J, 6j, ≥4b
≥2	2	-	≥2J, 6j, 2b
≥2	3	-	≥2J, 6j, 3b
≥2	≥4	-	≥2J, 6j, ≥4b

1-lep SRs

Search regions (≥6 jets)				
Mass-tagged jet multiplicity	b-jet multiplicity	$m_h^{min\Delta R}$	m_{eff}	Channel name
0	3	-	> 400 GeV	0J, ≥6j, 3b
0	≥4	-	> 400 GeV	0J, ≥6j, ≥4b
1	3	< 100 GeV	> 700 GeV	1J, ≥6j, 3b, LM
1	3	> 100 GeV	> 700 GeV	1J, ≥6j, 3b, HM
1	≥4	< 100 GeV	> 700 GeV	1J, ≥6j, ≥4b, LM
1	≥4	> 100 GeV	> 700 GeV	1J, ≥6j, ≥4b, HM
≥2	3	-	-	≥2J, ≥6j, 3b
≥2	≥4	-	-	≥2J, ≥6j, ≥4b
Validation regions (5 jets)				
Mass-tagged jet multiplicity	b-jet multiplicity	$m_h^{min\Delta R}$	m_{eff}	Channel name
0	3	-	> 400 GeV	0J, 5j, 3b
0	≥4	-	> 400 GeV	0J, 5j, ≥4b
1	3	-	> 700 GeV	1J, 5j, 3b
1	≥4	-	> 700 GeV	1J, 5j, ≥4b
≥2	3	-	-	≥2J, 5j, 3b
≥2	≥4	-	-	≥2J, 5j, ≥4b

Detailed Systematics in Ht+X

Systematics:
Dominant source depend on SR but,
In most sensitive SR syst dominated by :

tt+jet modeling :	
X-Section	up to 6.1%
$t\bar{t}$+1c	50%
jet JES/JER	12/16%
ST modelling :	
X-Section	$\approx$ 5%
Interference	24-52%
V+jet	30%
b, c, light jet modelling	
Diboson	48%

Signal regions definitions

Region	S_T	$\Delta R(\text{lep}, \nu)$	Small-R jets	Large-R jets
Boosted signal region	> 1200 GeV	< 0.8	≥ 3	≥ 1
Resolved signal region	> 1200 GeV	< 0.8	≥ 4	$= 0$
$t\bar{t}$ control region 1	> 700 GeV	> 1.0	≥ 3	≥ 1
$t\bar{t}$ control region 2	750 – 1200 GeV	< 1.0	≥ 3	≥ 1

Systematics:
Main source from jet and Large R-jet scale and resolution

lumi.	2.9
jet	24% per jet
VV production	6%
VV+jet production	5%
$t\bar{t}$	6%
single top	6%
Multijet bkg estimation	100%

Systematics sources from background

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

Signal regions defintion

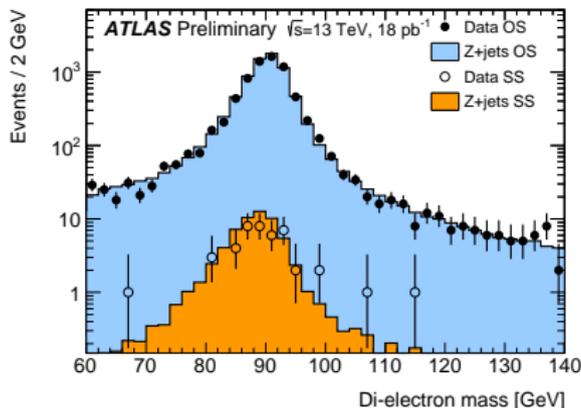| Definition			Name
$e^{\pm}e^{\pm} + e^{\pm}\mu^{\pm} + \mu^{\pm}\mu^{\pm} + eee + ee\mu + e\mu\mu + \mu\mu\mu$, $N_{\text{jets}} \geq 2$			
$400 < H_{\text{T}} < 700 GeV$	$N_b = 1$	$E_{\text{T}}^{\text{miss}} > 40$ GeV	SR0
	$N_b = 2$		SR1
	$N_b \geq 3$		SR2
$H_{\text{T}} \geq 700$ GeV	$N_b = 1$	$40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

Tag and Probe

- negligible for muons
- Using the pure Z->ee resonnance
- Estimate the charge flip ϵ rate in η, pt bins
- For one el in bin i and one in bin j the expected number of SS event is

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

τ_{32} is the ratio of τ_3/τ_2 and represent oth probility that a Fat jet is rather 3-prong than 2-prong

τ_N is the N-subjetiness and is computed as

$\tau_N = \frac{1}{d_0} \sum_k p_{\text{T}k} \times R_k^{min}$, with $d_0 = \sum_k p_{\text{T}k} \times R$

where k run over all the clusters and R is the characteristic size of sub-kt jets

am_{T2} target dilepton $t\bar{t}$ events. For such topology it has an end-point at $m_{top} = 175 GeV$. Derived from m_{T2} variables

m_T^W is teh w trasverse mass $m_T^W = \sqrt{2 p_{\text{T}}^{lep} E_{\text{T}}^{miss}[1 - cos(\Delta\phi)]}$