Summer Programme presentation

$t\bar{t}X$, $X \to \tau\tau$ analysis

BEAUCAMP Jean Yves, with the $t\bar{t}X$ analysis team.
Summer Programme project directors: BARAK, Liron; VANNICOLA, Damiano.

ATLAS Experiment, CERN - 2021.

TEL AVIV UNIVERSITY
UNIVERSIDAD NACIONAL DE LA PLATA

Introduction and motivation

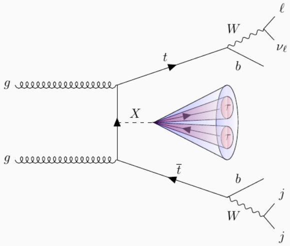
As we've seen during the lectures, there are many open questions not answered by the SM: dark matter, neutrino masses, baryon asymmetry.

- Many SM extensions have been proposed, most requiring expanding the Higgs sector.
- Light resonances can exist in many of those extensions. This is mostly an unexplored sector of the phase space.

The search for light CP-odd scalars with mass up to 100 GeV is very challenging at the LHC, because their mass is almost unconstrained by direct searches.

On this search...

- Searching for a light resonance X in the mass-range of $20-100$GeV, produced in association with a pair of t. Final state is a pair of τ_{had}.
- Because of it's small mass and the heavy $t\bar{t}$, these resonances will be produced with a large boost.
- Leptonic $t\bar{t}$ decays offer handles to trigger the signal events and to reject background.



The boosted pair of τs

Due to large boost in the pair of τ, their angular separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is very small, and thus can't be efficiently reconstructed as separate objects. A new strategy is required.

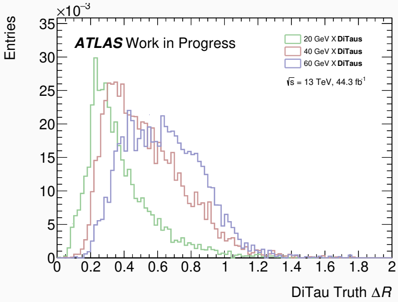


Figure 1: Angular separation ΔR between τ_{had}.

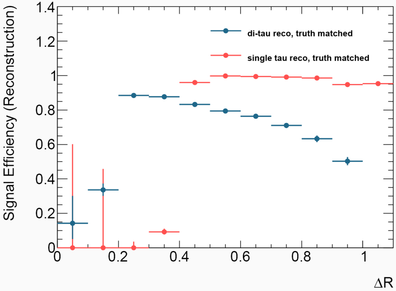
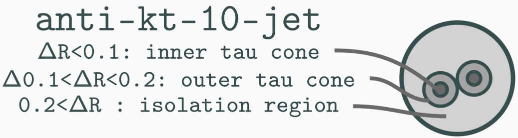


Figure 2: Pair of τ_{had} reconstruction efficiencies[1].

[1] https://cds.cern.ch/record/2105592.

Due to large boost in the pair of τ, their angular separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is very small, and thus can't be efficiently reconstructed as separate objects. A new strategy is required.

Large $R \sim 1$ jets candidates are reconstructed, with a pair of jet candidates (AntiKT algorithm) with $R \sim 0.1$ inside.



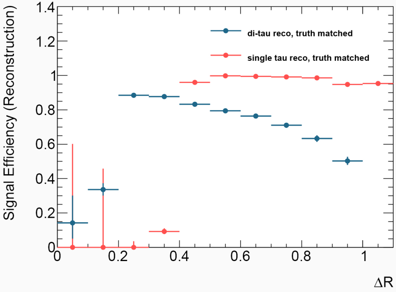
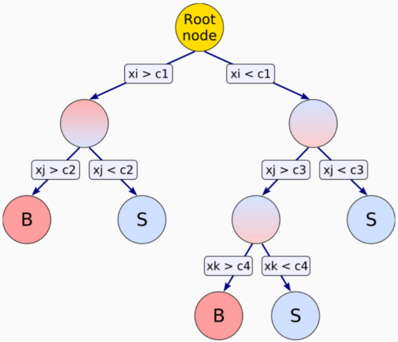


Figure 3: Pair of τ_{had} reconstruction efficiencies[2].

[1] https://cds.cern.ch/record/2105592.

DiTau object identification

A Boosted Decision Tree (BDT) is used for identifying the ”real” DiTau from all the DiTau objects candidates.



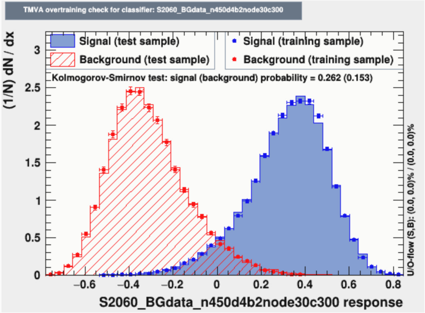


Figure 4: BDT TMVA training. Signal and Background not weighted.

A cut on $BDT > 0.2$ is proposed to separate the signal from the background.

Background estimation

A very good estimation of the Standard Model backgrounds is needed on the same region of the phase space as the light resonance we are searching for. There are many background processes that could contribute, but the main ones we identified are: $t\bar{t}$, $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, tZ, $4t$, $3t$, $Z \to \tau\tau + Jets$, $W \to \mu\nu + Jets$, $W \to e\nu + Jets$ and $W \to \tau\nu + Jets$.

There can also be other background contributions due to other particles being miss-identified as DiTau objects (fakes). Those are studied using data-driven methods.

In order to compare the MC simulations to the data, simulated events need to be weighted by:

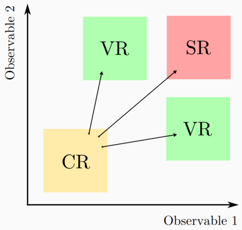
$$w = \frac{\int L\,\mathrm{d}t}{\mathrm{TotalSumW}} \times \mathrm{xSection} \times \mathrm{kFactor} \times \mathrm{FilterEfficiency} \times \mathrm{GenWeight} \times \mathrm{muWeight}$$
$$\times\, (\mathrm{Ele} - \mathrm{Muo})\mathrm{Weight} \times \mathrm{JetWeight},$$

Some weights have to do with the event generation (*GenWeight*). Others raise from differences in the reconstruction and simulation of the objects (*EleWeight*, *MuoWeight*, *JetWeight*).

The different regions

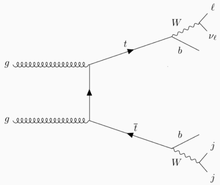
Even if we weight all the MC events correctly, the simulations still don't have the necessary precision to search for new physics. Thus, a Control Region (CR) will be used.

- Signal Regions (SR): where the BSM models predict an excess over the backgrounds (SM).
- Control Regions (CR): designed to normalize the estimation of the different backgrounds to the experimental data. Orthogonal to the SRs (in this case by inverting the *BDT* cut).
- Validation Regions (VR): allow to validate the estimation of the backgrounds, with the corrections introduced by the CRs.



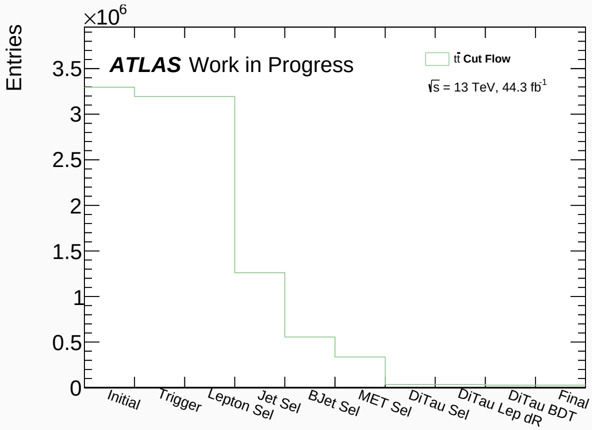
Preselection common to all regions: lepton trigger, kinematic and quality cuts on leptons, Jets, B-Jets, DiTau objects.

Control Region	Signal Region
Preselection && $E_T^{miss} \geq 50$ GeV && $NBJets \geq 2$ && $ditau\_BDT < 0.2$	Preselection && $\Delta R(DiTau, Lepton) \geq 1$ && $NBJets \geq 1$ && $ditau\_BDT \geq 0.2$



Results

All samples generated with the XAMPP framework. Selection and plotting with a custom ROOT code.



Step	Events	
Initial	3295864	(100.00%)
Trigger	3194553	(96.93%)
Lepton Sel	319408	(96.91%)
Jet Sel	1261144	(38.26%)
BJet Sel	555355	(16.85%)
MET Sel	33539	(10.18%)
DiTau Sel	35213	(1.07%)
DiTau Lep dR	35213	(1.07%)
DiTau BDT	26777	(0.81%)
Final	26777	(0.81%)

Table 1: $t\bar{t}$ MC samples cutflow in the Control Region. Work in progress.

Results

All SM Backgrounds.

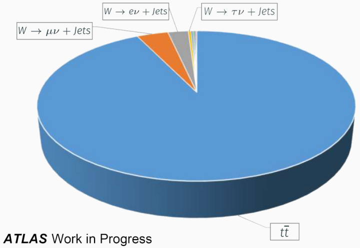


Figure 5: CR background contributions.

Process	CR		SR	
$t\bar{t}$	26776.8	(92.760%)	4844.2	(88.001%)
$W \to \mu\nu + Jets$	1086.4	(3.764%)	315.7	(5.736%)
$W \to e\nu + Jets$	709.6	(2.458%)	224.3	(4.075%)
$W \to \tau\nu + Jets$	93.2	(0.323%)	17.6	(0.321%)
$t\bar{t}H$	54.8	(0.190%)	11.2	(0.204%)
$t\bar{t}Z$	52.1	(0.180%)	9.4	(0.171%)
$t\bar{t}W$	46.1	(0.160%)	12.2	(0.221%)
$Z \to \tau\tau + Jets$	36.6	(0.127%)	67.1	(1.218%)
tZ	7.2	(0.025%)	2.1	(0.038%)
$4t$	3.4	(0.012%)	0.7	(0.013%)
$3t$	0.5	(0.002%)	0.1	(0.002%)

Table 2: MC background contributions. Work in progress.

We also checked that the signal contamination in the CR is negligible.

Next Steps

- Preliminary Control Region already defined.
- We are now waiting to compare and cross-check the results from different members of the team. A comparison with experimental data is next.

I'm looking forward to keep working with the team in the next weeks as part of my Summer Programme project, and later for my Master's thesis!

Thank you!

Backup

Events Preselection

- Leptons: 1 electron or muon.
  - Electrons:
    - Trigger: (TrigHLT_e26_lhtight_nod0_ivarloose || TrigHLT_e60_lhmedium_nod0 || TrigHLT_e140_lhloose_nod0)&& (TrigMatchHLT_e26_lhtight_nod0_ivarloose || TrigMatchHLT_e60_lhmedium_nod0 || TrigMatchHLT_e140_lhloose_nod0),
    - $p_T \geq 27\,\text{GeV}$ && $|\eta| \leq 2.47$.
  - Muons:
    - Trigger: (TrigHLT_mu26_ivarmedium || TrigHLT_mu50)&& (TrigMatchHLT_mu26_ivarmedium || TrigMatchHLT_mu50),
    - $p_T \geq 27\,\text{GeV}$ && $|\eta| \leq 2.5$.
- Jets: At least 3 jets. $p_T \geq 25\,\text{GeV}$ && $|\eta| \leq 2.5$. $JVT > 0.59$ if $p_T < 60\,\text{GeV}$ && $|\eta| \leq 2.4$.
- DiTaus: At least one DiTau object. 1-prong or 3-prong leading and subleading subjets (1 or 3 tracks). $|charge| = 1$ for leading and subleading subjets. $p_T > 40\,\text{GeV}$ && $|\eta| < 2.5$.