Searching for Higgs bosons in $t\bar{t}H$, $H \to b\bar{b}$ with ATLAS in LHC Run-2

a brief introduction to the $t\bar{t}H$, $H \to b\bar{b}$ analysis, its infrastructure and discovery prospects

Will Breaden Madden

IOP 2016

2016-03-22

University of Glasgow | Experimental Particle Physics

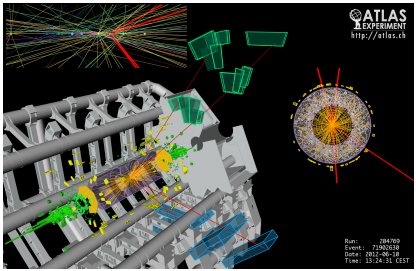
Higgs search in $t\bar{t}H$ process with ATLAS

- interesting process in which Higgs bosons have not been observed
- very difficult analysis, but done because extremely useful for testing SM description of reality
- this talk (first of three):
  - a very brief introduction to the $t\bar{t}H$, $H \to b\bar{b}$ analysis
  - some of the tools we have developed to do the analysis
  - me: l + jets channel
- following: dilepton channel, boosted analysis

University of Glasgow | Experimental Particle Physics

Why search for the $t\bar{t}H$ process?

- Given the existence of a Higgs boson,
  - is it a Standard Model Higgs boson?
  - what is its character?
- The t is the most massive of all observed elementary particles $\rightarrow Y_t \sim 1$.
- The $t\bar{t}H$ process is a way to measure directly the t-H Yukawa coupling Y_t.

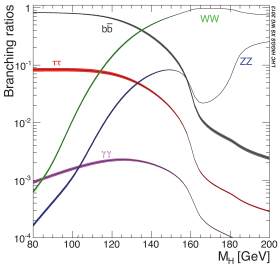


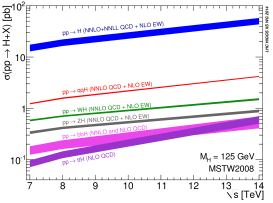
- So, the $t\bar{t}H$ process is interesting and useful because its discovery and measurement would be powerful evidence for the SM while deviations would indicate BSM physics.

University of Glasgow | Experimental Particle Physics

Higgs boson processes in $t\bar{t}H$, $H \to b\bar{b}$ search

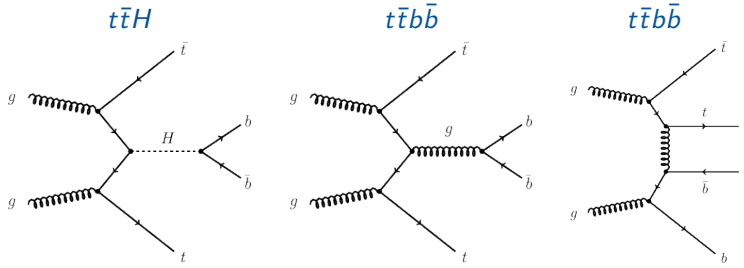
- for $m_H \simeq 125$ GeV, $H \to b\bar{b}$ dominant decay mode (not observed)
- direct search from g fusion precluded by overwhelming multijet background
- however: search in association with vector boson (VH) or t-quark pair ($t\bar{t}$) significantly improves signal-to-background ratio
- so: Our search for is for Higgs bosons in the $t\bar{t}H$ production mode and it's designed to be sensitive primarily to the $H \to b\bar{b}$ decay.
- For $t\bar{t}H$, the cross section is low and the process is quite supressed, but the branching fraction to two b quarks is very high. So, relatively, we'd get a lot of Higgs bosons from the $t\bar{t}H$, $H \to b\bar{b}$ process.





signal and background processes in $t\bar{t}H$ search

- main background source: t pairs produced in association with additional jets
- dominant source: $t\bar{t} + b\bar{b}$ production – resulting in same final-state as signal!

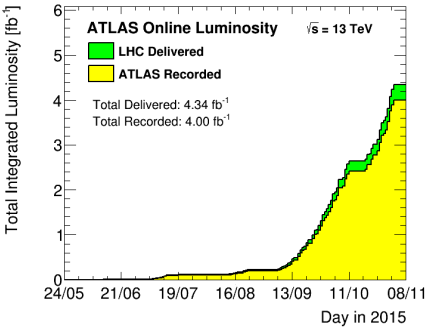


all look similar $\Rightarrow$ irreducible backgrounds

- can't use *simple* observations to discriminate between processes

challenges in $t\bar{t}H$ search

- suppressed with respect to other Higgs modes
- $H \to b\bar{b}$ has the largest branching ratio (0.577 for m_H 125 GeV)
- irreducible background from $t\bar{t}b\bar{b}$
- other backgrounds: $t\bar{t}$ production in association with light quarks (u, d, s) or gluon jets (called $t\bar{t}$ + light), and $t\bar{t} + c\bar{c}$



$\sqrt{s}$ (TeV)	7	8	13	14
$t\bar{t}H$ ($m_H = 125$ GeV) (pb)	0.086	0.130	0.5085	0.611
$t\bar{t}$ (pb)	177	253	832	950
$S/\sqrt{B}$	0.00646	0.0082	0.0176	0.0198

7 TeV $\to$ 13/14 TeV: $S/\sqrt{B}$ changes by factor of $\simeq$3

University of Glasgow | Experimental Particle Physics

How can we address this large, irreducible background?

analysis strategy:

- irreducible background

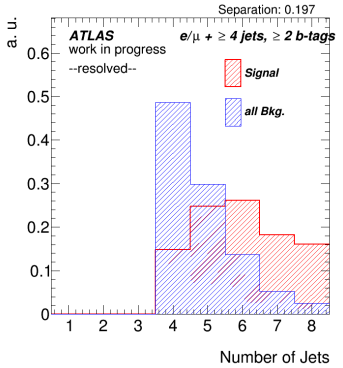
  $\Rightarrow$ no clear discriminating variables

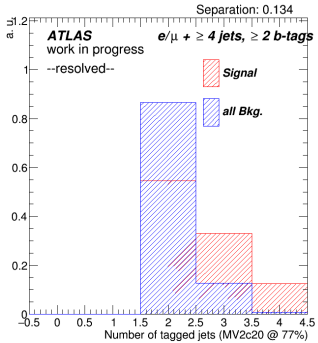
  $\Rightarrow$ multivariate approach to get best possible S/B separation (NN, BDT, DNN in due course)

- define analysis 'regions' using jet multiplicities
- use signal-depleted regions to constrain backgrounds and uncertainties
- make a combined nuisance parameter fit to all regions

University of Glasgow | Experimental Particle Physics

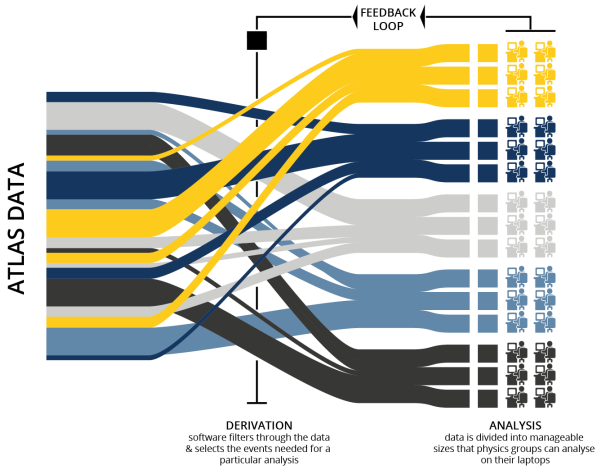
motivation for jet multiplicity analysis regions





- goal: define signal-enriched and signal-depleted regions
- good separation provided by jet multiplicity and b-tagged jets

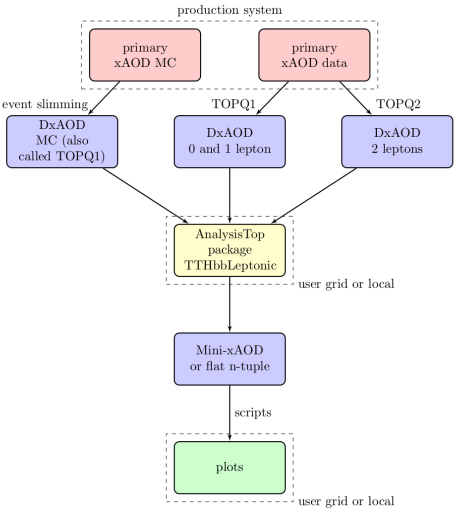
ATLAS analysis data flow



infographic by Nicola Quadri

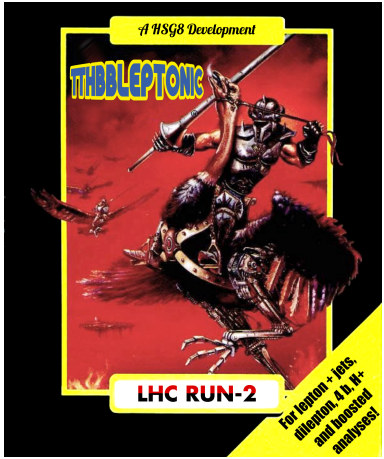
University of Glasgow | Experimental Particle Physics

$t\bar{t}H$ analysis data flow

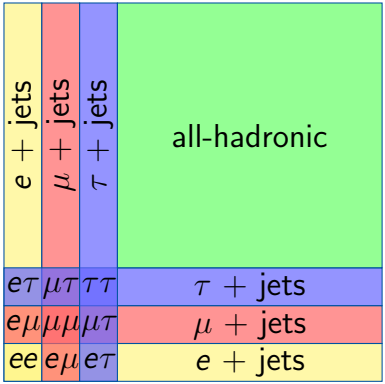


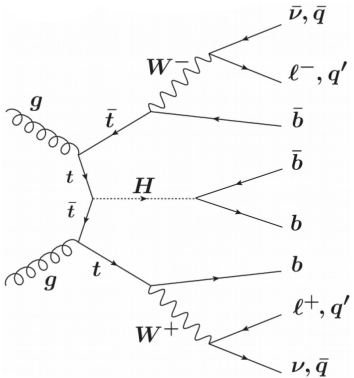
enter: TTHbbLeptonic!

- TTHbbLeptonic: analysis package for the $t\bar{t}H$, $H \to b\bar{b}$ analysis
- formed from the union of the analysis efforts of the l + jets, dilepton, boosted $t\bar{t}H$, $H \to 4b$, $H^+ \to tb$ and, recently, all-hadronic analyses
- good, clear code with validation procedures, branch-based development, good, detailed, per-version documentation, welcoming to new analyses, release early and often



$t\bar{t}H\,(b\bar{b})$ channels



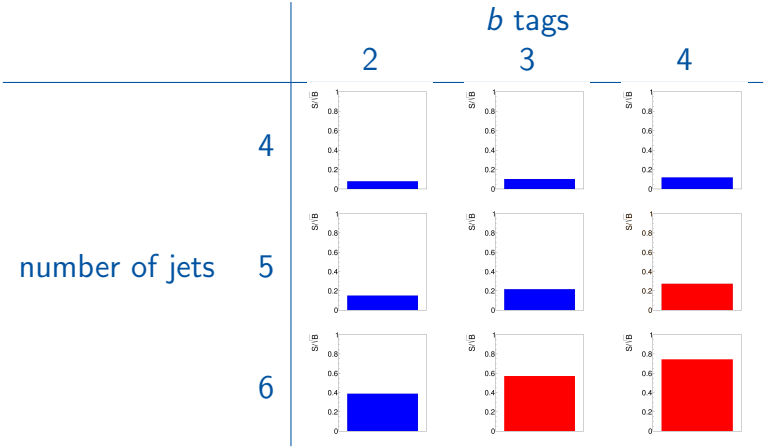


$t\bar{t}H\left(b\bar{b}\right)$ l + jets analysis

- data recording
- data derivations (remove information unnecessary to specific analyses)
- event selection and calculation of variables – analysis framework
  - single lepton triggers
  - single lepton channel cuts:
    - exactly 1 isolated high–p_T (i.e. 25 GeV) lepton (e or μ),
    - at least 4 jets,
    - at least 2 of which are b-tagged
- event categorisation by jet characteristics
  - constrain systematic uncertainties with low S/B regions
  - study data-MC agreement in inclusive regions

University of Glasgow | Experimental Particle Physics

$S/\sqrt{B}$ for the various analysis regions

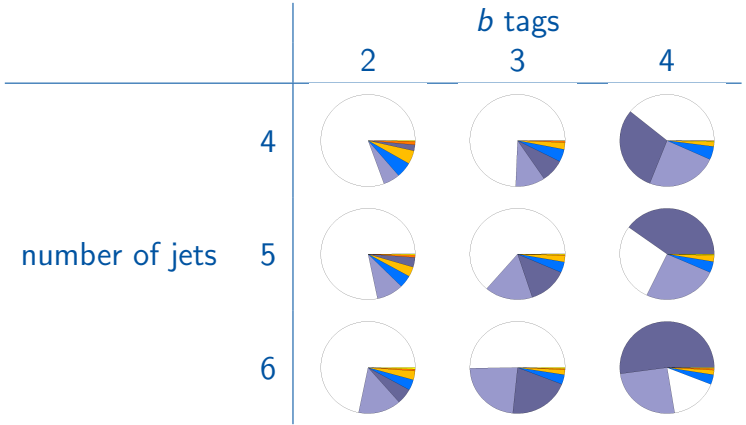


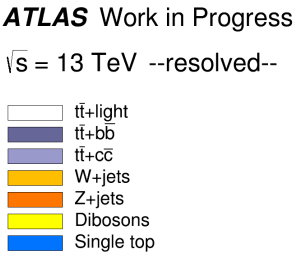
ATLAS Work in Progress

$\sqrt{s}$ = 13 TeV --resolved--

*colours defined by collective consideration of S/B and $S/\sqrt{B}$

background compositions of various analysis regions

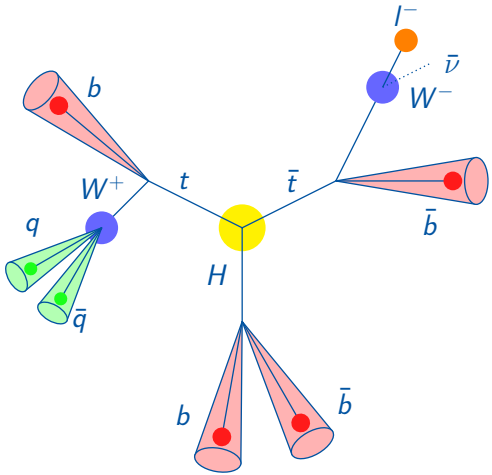




major contributions from processes involving pairs of t quarks

some human-defined MVA discriminating variables

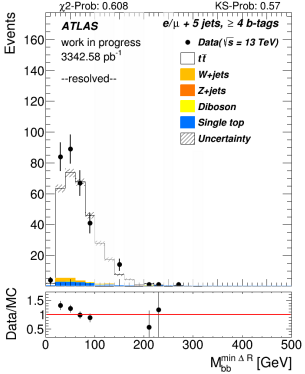
- obvious discriminant: masses of particles decaying to two b quarks
  (the difference between the Higgs and gluon masses)
- full kinematic reconstruction would work, but it would be computationally difficult with 6 jets...
- (see next talks)

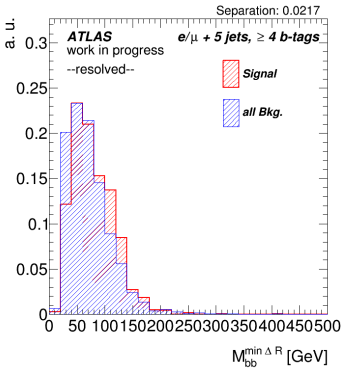


University of Glasgow | Experimental Particle Physics

MVA discriminating variable: $M_{b\bar{b}}^{min\Delta R}$

$M_{b\bar{b}}^{min\Delta R}$: the mass of the b quark pair with the minimum ΔR between them

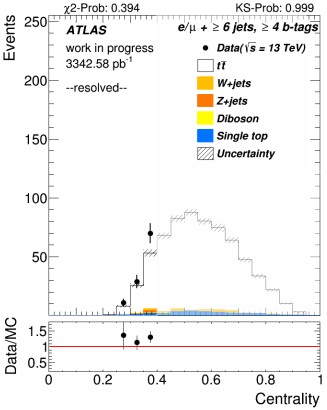


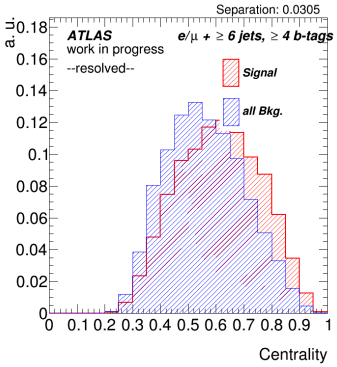


University of Glasgow | Experimental Particle Physics

MVA discriminating variable: centrality

Centrality: the scalar sum of the p_T divided by the sum of the E for all jets and leptons





University of Glasgow | Experimental Particle Physics

event yields for two signal-enriched regions

6 jets, 3 b-tags

Process	Events	Stat. unc.
$t\bar{t}$	5145	20
W+jets	94.1	9.9
Z+jets	23.0	2.5
Diboson	20.9	1.2
Single top	186.3	2.3
Total bkg	5469.4	6.0
ttH exp.	42.23	0.32
S/Bkg	0.0077	
S/sqrt(Bkg)	0.57	

Run-1: 40 $t\bar{t}H$ expected events

6 jets, 4 b-tags

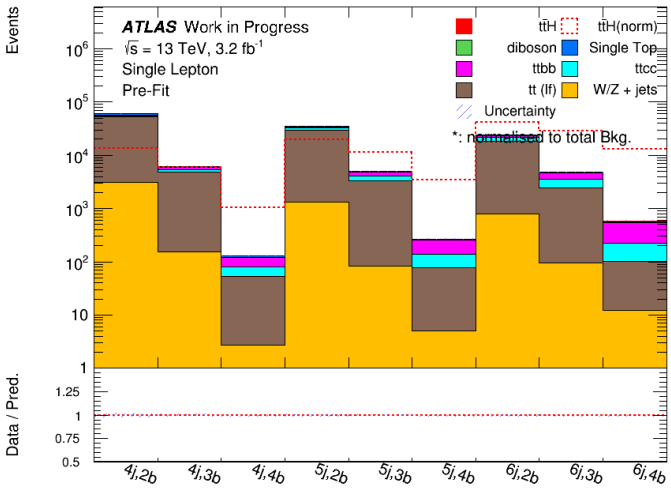
Process	Events	Stat. unc.
$t\bar{t}$	622.6	7.4
W+jets	10.7	2.7
Z+jets	3.7	2.2
Diboson	2.17	0.37
Single top	22.98	0.83
Total bkg	662.2	3.7
$t\bar{t}H$ exp.	19.13	0.21
S/Bkg	0.029	
S/sqrt(Bkg)	0.74	

Run-1: 16 $t\bar{t}H$ expected events

ATLAS Work in Progress

University of Glasgow | Experimental Particle Physics

event yields for all regions



- This plot shows the number of events in the standard resolved analysis in each resolved category for signal and background.

What's next?

- $t\bar{t} + V$ estimates, QCD estimates
- MVA training to estimate expected limit
- restart imminent
  (beam: $\sim$2016-03-28; stable beams and physics: $\sim$2016-04-25)
- first Run-2 combined $t\bar{t}H$ result
  – aim: ICHEP 2016 (2016-08-03–2016-08-10)
- improve on sensitivity of Run-1 results

University of Glasgow | Experimental Particle Physics

When might we expect to discover a $t\bar{t}H$ SM Higgs boson?

- expected significance of $\sim$3 σ for the full $t\bar{t}H$ combination with $\sim$ 20 fb^{-1} data
- expected significance of $\sim$5 σ for the full $t\bar{t}H$ combination with $\sim$ 100 fb^{-1} data
- have 4 fb^{-1} from 2015 when LHC Run-2 started
  – we hope to have much more this year!

University of Glasgow | Experimental Particle Physics

What's next?

- There are *many* parts to the combined $t\bar{t}H$, $H \to b\bar{b}$ analysis.
- I work on the single lepton part.
  - Will Breaden Madden (single lepton analysis)
- In the next two talks, you'll have the pleasure of hearing about two other parts of the analysis from my colleagues:
  - Ben Sowden (dilepton analysis)
  - Sam Crawley (boosted analysis)

University of Glasgow | Experimental Particle Physics

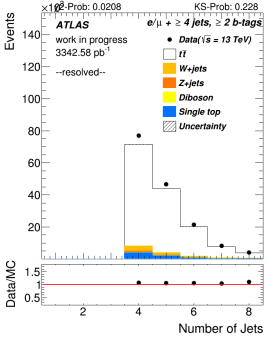
questions?

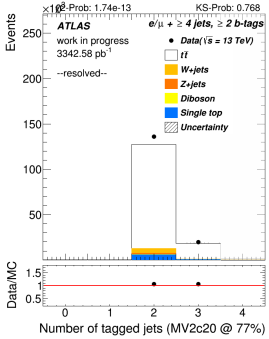
thanks!
questions?

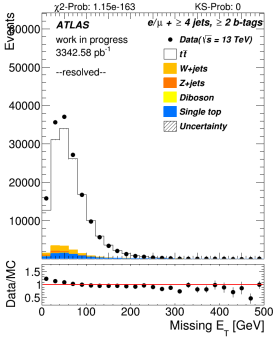
University of Glasgow | Experimental Particle Physics

BACKUP

jet multiplicities and MET

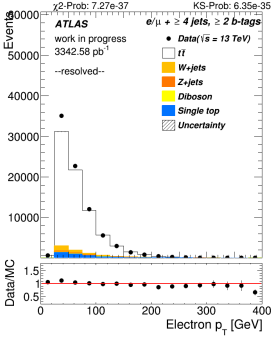


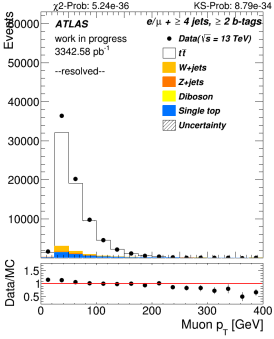


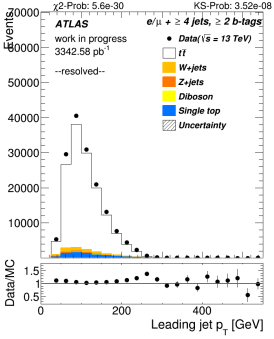


*working on lepton fakes estimate currently ⇒ less MC-data discrepancy at low p_T

leptons and jet p_T







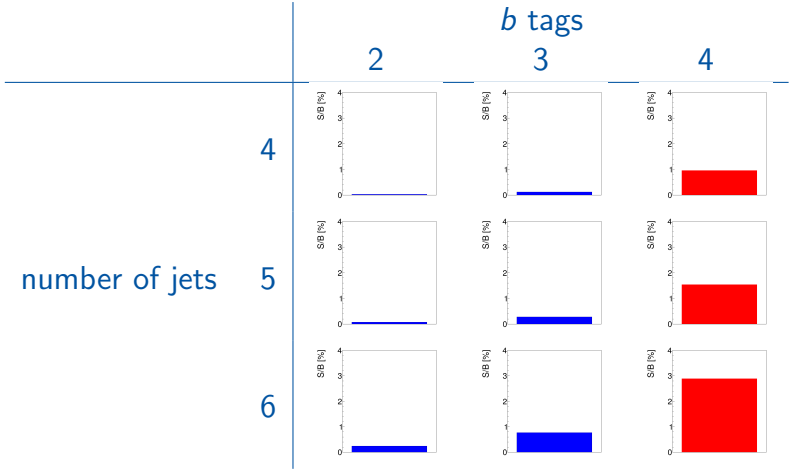
*working on lepton fakes estimate currently ⇒ less MC-data discrepancy at low p_T

derivation framework and data reductions

- The Run-2 Derivation Framework defines the conversion of xAOD data to derived DxAODs, which groups and users can process. The Derivation Framework can produce one or several output DxAODs for a specified input xAOD. The target size for a derivation is 1–10 TB.
- data reduction by
  - skimming (removing events)
  - slimming (removing per-event information)
- A target upper limit threshold of 1% of the xAOD per derivation is defined as a workable fraction of xAOD input. Assuming 50 derivations in total, this allows for $\sim$4 versions of each derivation.
- target CPU usage: < 40 ms per derivation or ≤ 160 ms per group with the total budget of 2 s per event

University of Glasgow | Experimental Particle Physics

S/B



ATLAS Work in Progress

$\sqrt{s}$ = 13 TeV --resolved--

*colours defined by collective consideration of S/B and $S/\sqrt{B}$

END