

simple boosted objects and novel views of top

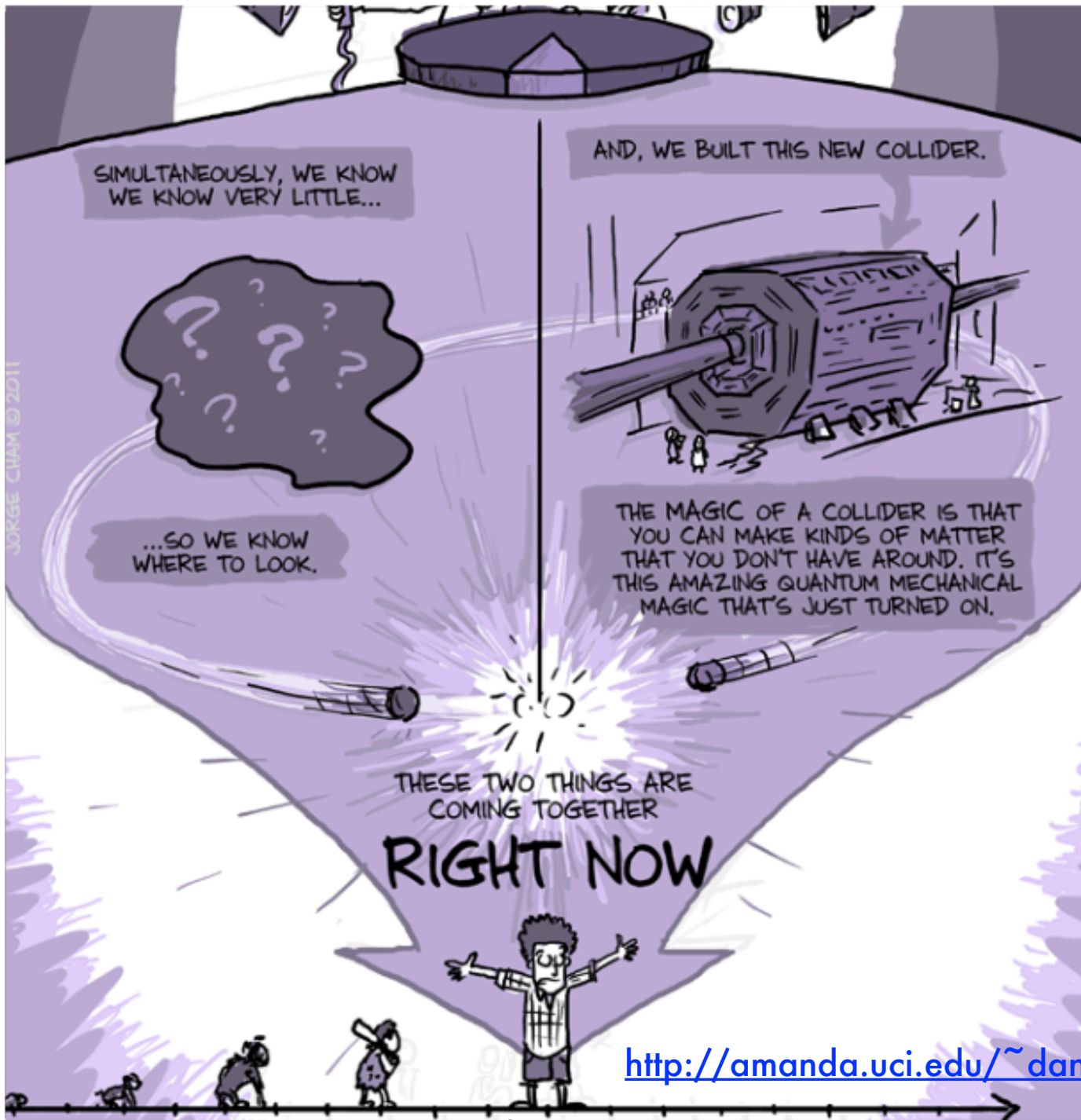
Featuring brand new
ATLAS, CDF results!

Daniel Whiteson
UC Irvine



Seminar, UC Davis, Feb 2012

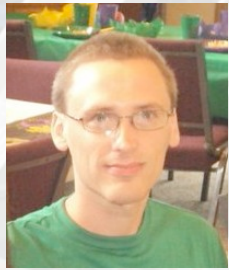






▶ Whiteson, ZZ → 4l, 4l searches

Postdocs



Ning Zhou, SS 2l, SS tops, γ +MET

▶ Andy Nelson, ZZ + searches



Grads



▶ Robert Porter, CDF SS dileptons

Michael Werth, t' in OS dileptons 34/pb

t' in OS dileptons 1/fb



▶ Kanishka Rao, CDF top+MET

b' in l+jets

$t+t$ resonance (CDF+ATLAS)

resonances in top



Eric Albin, γ +MET

Undergrads



▶ Johnny Ho, $t+t$ resonance (ATLAS)

Jared Vasquez, CDF ZZ+met, t' to OS dilep 5/fb



Four younger students Z' to gluons³, resonances in top

Color Key

CONF

Published

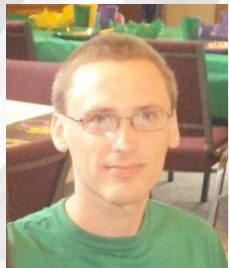
In Review

Planned



▶ Whiteson, ZZ->4l, 4l searches

Postdocs



Ning Zhou, SS 2l, SS tops, γ +MET



▶ Andy Nelson, ZZ + searches

Grads



▶ Robert Porter, CDF SS dileptons

Michael Werth, t' in OS dileptons 34/pb

t' in OS dileptons 1/fb



▶ Kanishka Rao, CDF top+MET

b' in l+jets

t +j resonance (CDF+ATLAS)

resonances in top



Eric Albin, γ +MET

Undergrads



▶ Johnny Ho, t +j resonance (ATLAS)

Jared Vasquez, CDF ZZ+met, t' to OS dilep 5/fb



Four younger students Z' to gluons, resonances in top

Color Key

CONF

Published

In Review

Planned

Outline

I. Introduction: experimental innovation

II. Simple boosted objects

III. Novel views of top quark events

VI. Concluding rant: open data

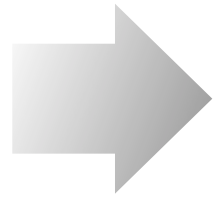
History of **biology** PhDs

1980s

Choose a
gene

Manually
sequence <5

Crude
comparisons

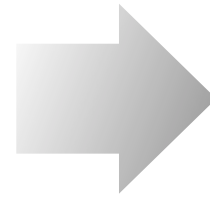


1990s

Choose a
set of genes

Automatically
sequence 10s

Compare to
databases of
genes



2000s

Choose a
genome

Cheaply sequence
1M in < 1 week

Invent new analytics
to compare
genomes

*Still don't understand
correlation vs causation*

History of **experimental HEP** PhDs

1980s

Choose a
channel/signal

Enumerate
backgrounds

Estimate
backgrounds

Estimate
uncertainties

Set limits

1990s

Choose a
channel/signal

Enumerate
backgrounds

Estimate
backgrounds

Estimate
uncertainties

Set limits

2000s

Choose a
channel/signal

Enumerate
backgrounds

Estimate
backgrounds

Estimate
uncertainties

Set limits

Of course advances have been made for each step

So what is modern?

New experimental tools

Sophisticated triggering

Multi-variate analysis/ Machine learning

jet flavor tagging

monopoles

lepton-jets

track-less jets

boosted objects

new resonances

So what is modern?

New experimental tools

Sophisticated triggering

Multi-variate analysis/ Machine learning

jet flavor tagging

monopoles

lepton-jets

track-less jets

boosted objects

new resonances

Outline

I. Introduction: experimental innovation

II. Simple boosted objects

III. Novel views of top quark events

VI. Concluding rant: open data

Theory papers

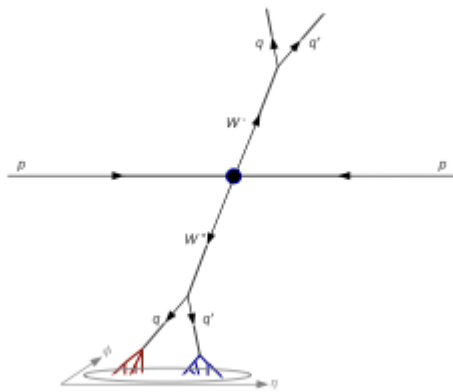
Identifying Boosted Objects with N-subjettiness

Jesse Thaler and Ken Van Tilburg

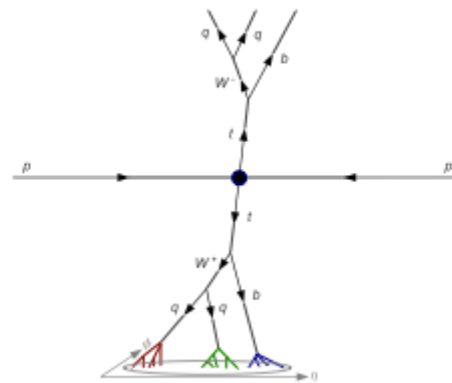
*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA
02139, USA*

E-mail: jthaler@jthaler.net, kvt@mit.edu

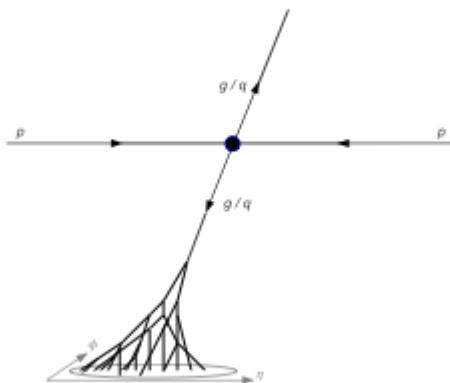
Counting sub-jets



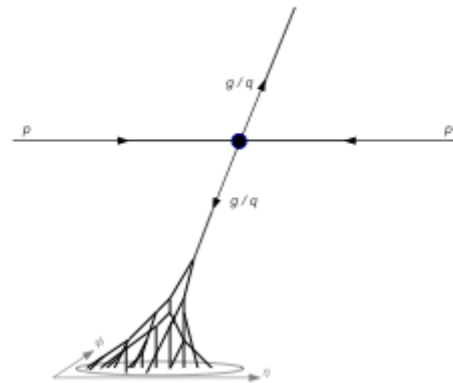
(a)



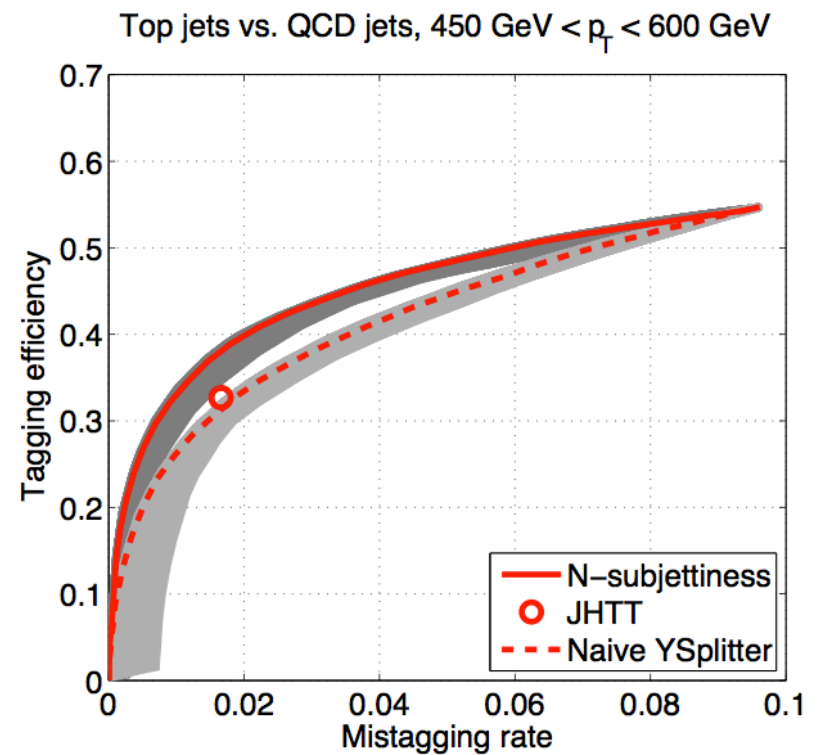
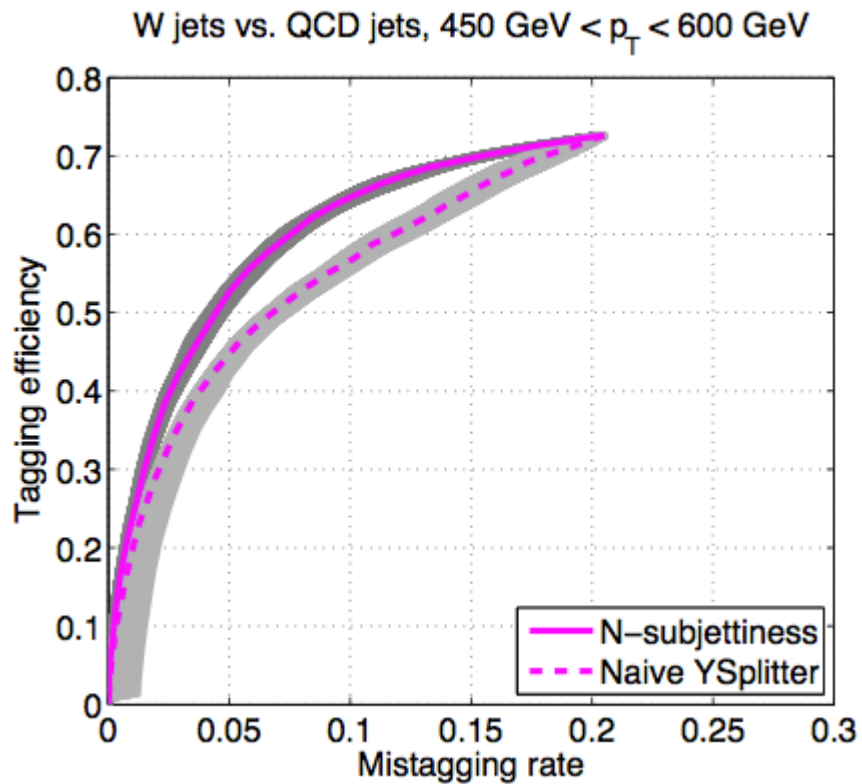
(a)



(c)



Performance



How realistic?

Hard process

→ Pythia

→ Toy Calorimeterization

→ FastJet

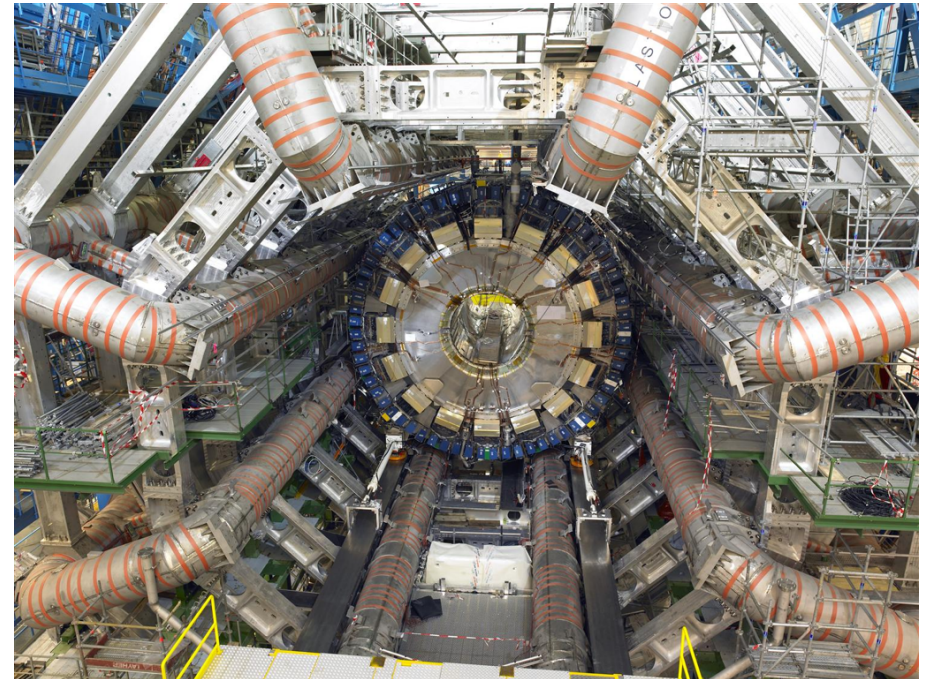
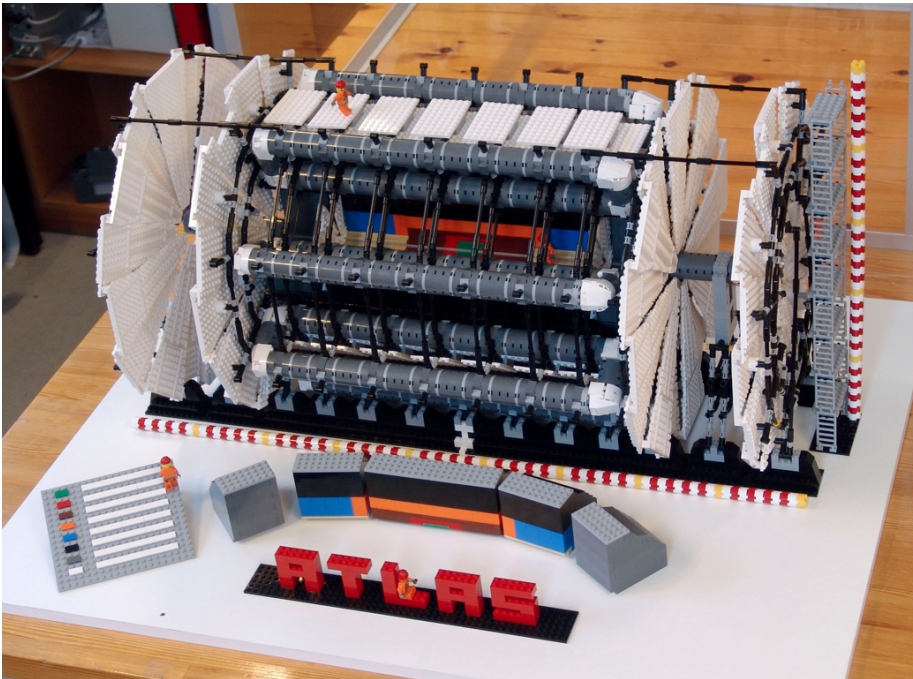
To partially simulate detector effects and to speed up jet reconstruction, observable final-state particles with $|\eta| < 4$ are collected into “calorimeter” cells arranged in a rectangular lattice with 80 rapidity (η) and 64 azimuth (ϕ) bins (corresponding to approximately 0.1×0.1 sized cells). The calorimeter momenta are interpreted as massless pseudo-particles with energy given by the calorimeter energy.

Play time

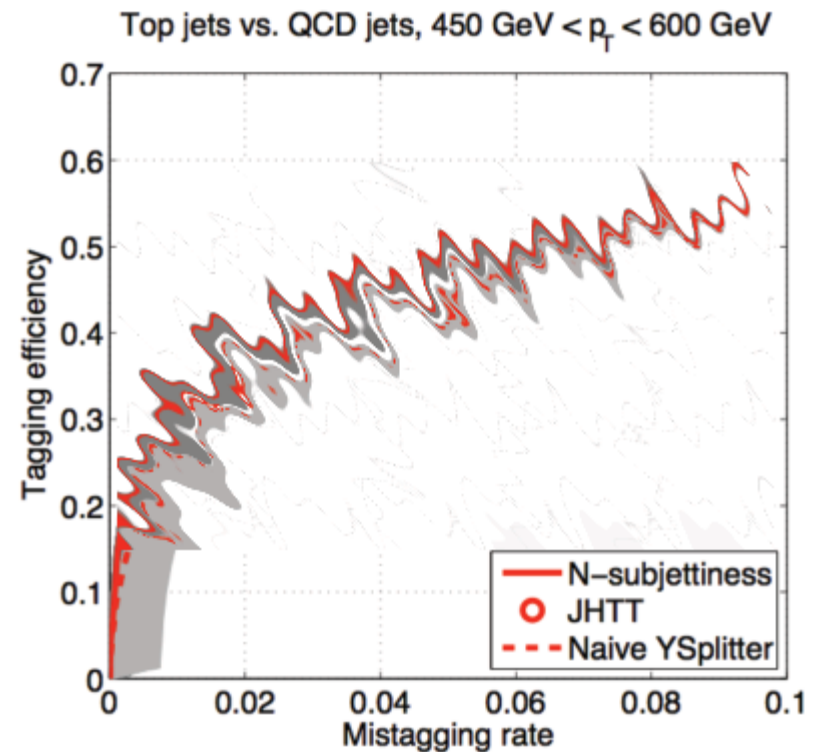
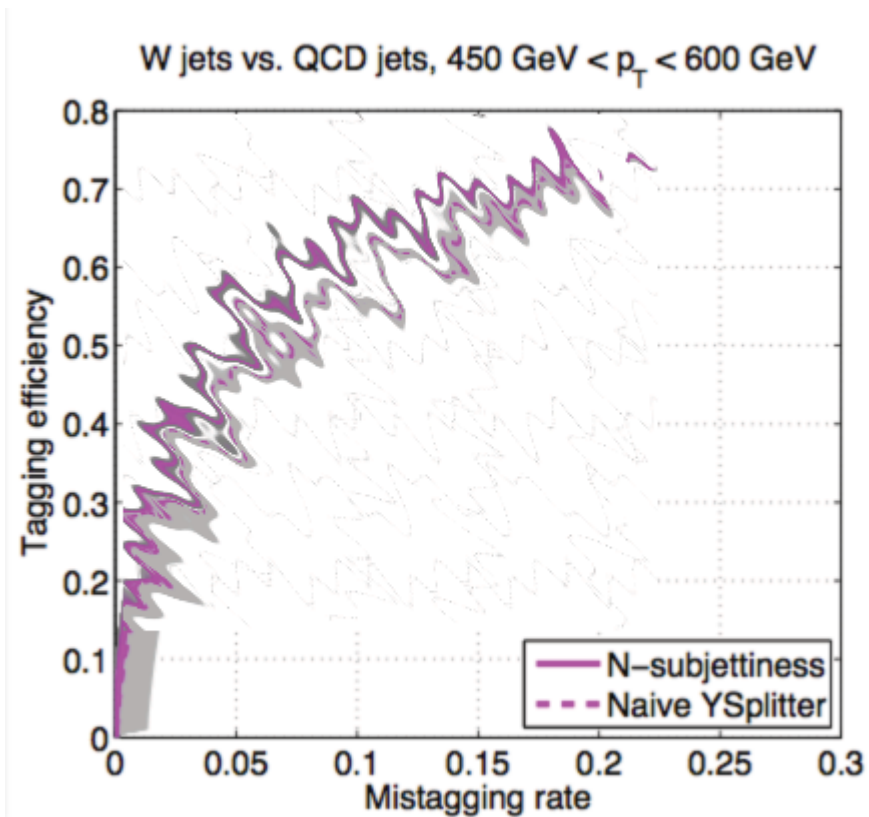
theory performance : results in data

=

LEGO ATLAS : real ATLAS



How I read these plots

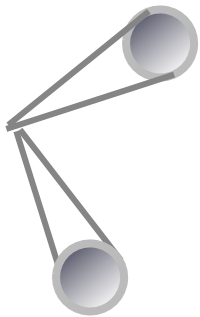


Summary: a lot of work for minor improvements
which may be irrelevant

Simple approach

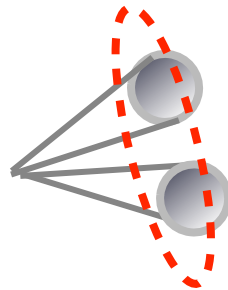
Do we need fatjets and subjets at all?

Small pT



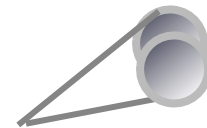
Use narrow jets.

Medium pT



Use narrow jets.
Group jets in
fat cones

Large pT

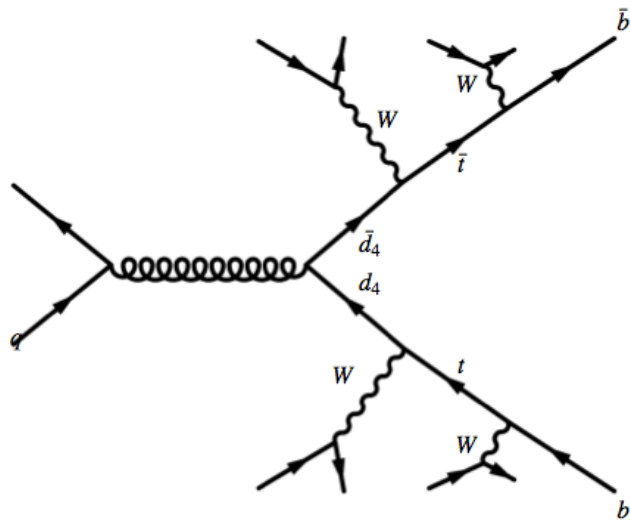


Use narrow jets.
Measure jet mass

$b' \rightarrow tt\bar{b} \rightarrow l + \text{jets}$



Heavy particles decaying to
top+jets



Selection

One lepton (e or μ)

MET > 20

At least **five** jets

Dataset

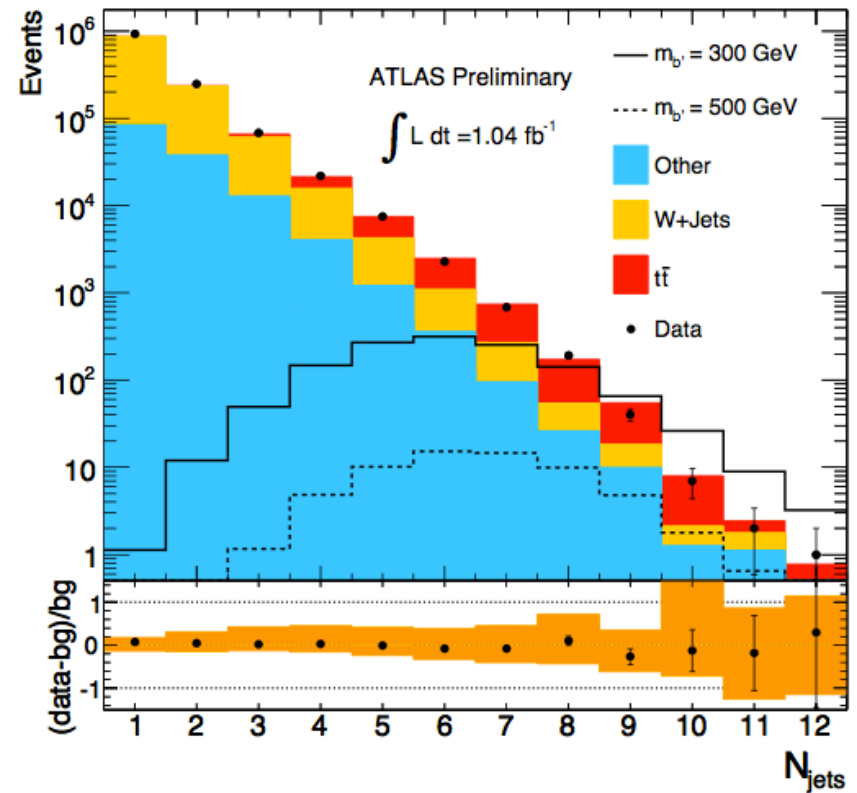
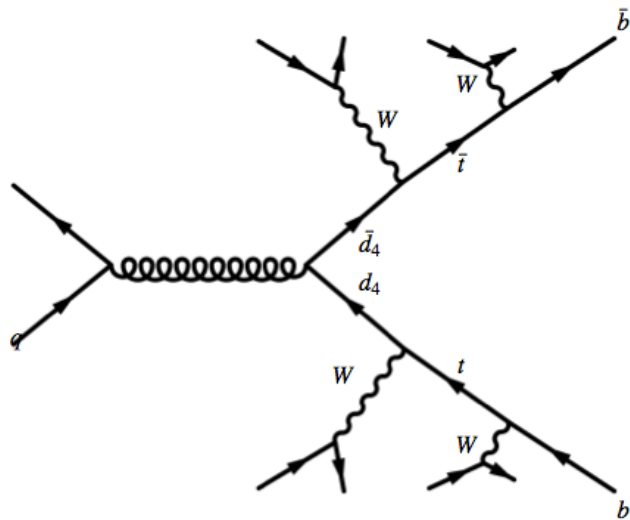
ATLAS 1/fb

Analysis done largely by a single
grad student: **Kanishka Rao, UCI**

$b' \rightarrow tt\bar{b} \rightarrow l + \text{jets}$



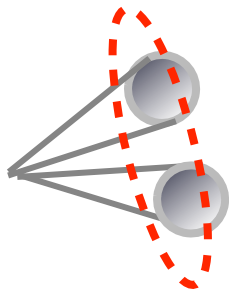
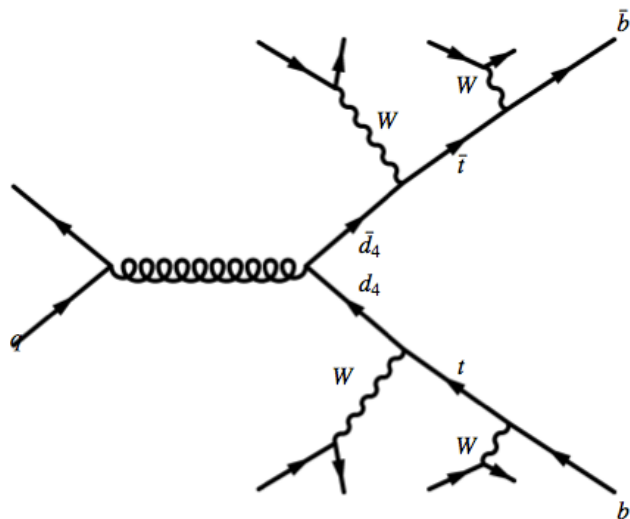
Heavy particles decaying to
top+jets



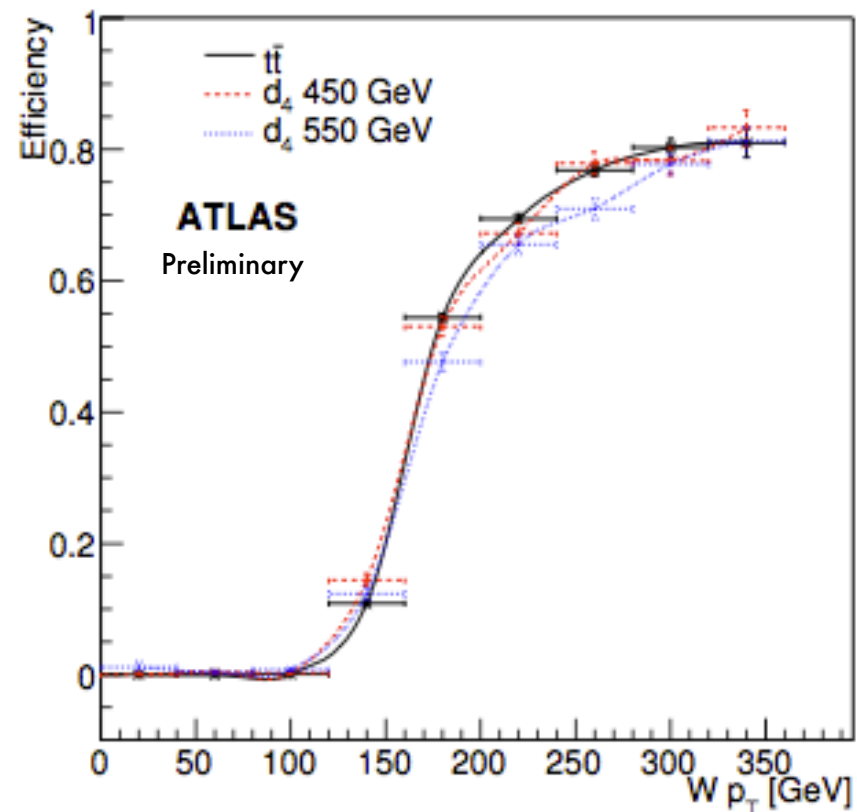
$b' \rightarrow tt\bar{b} \rightarrow l + \text{jets}$



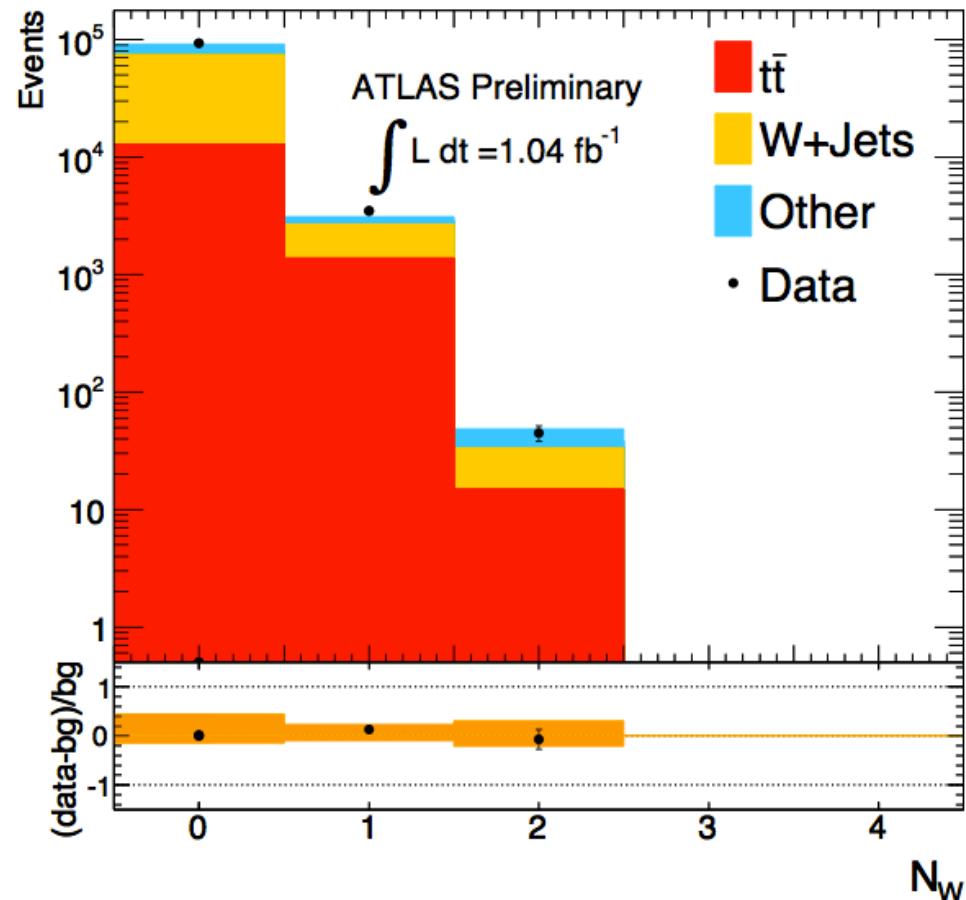
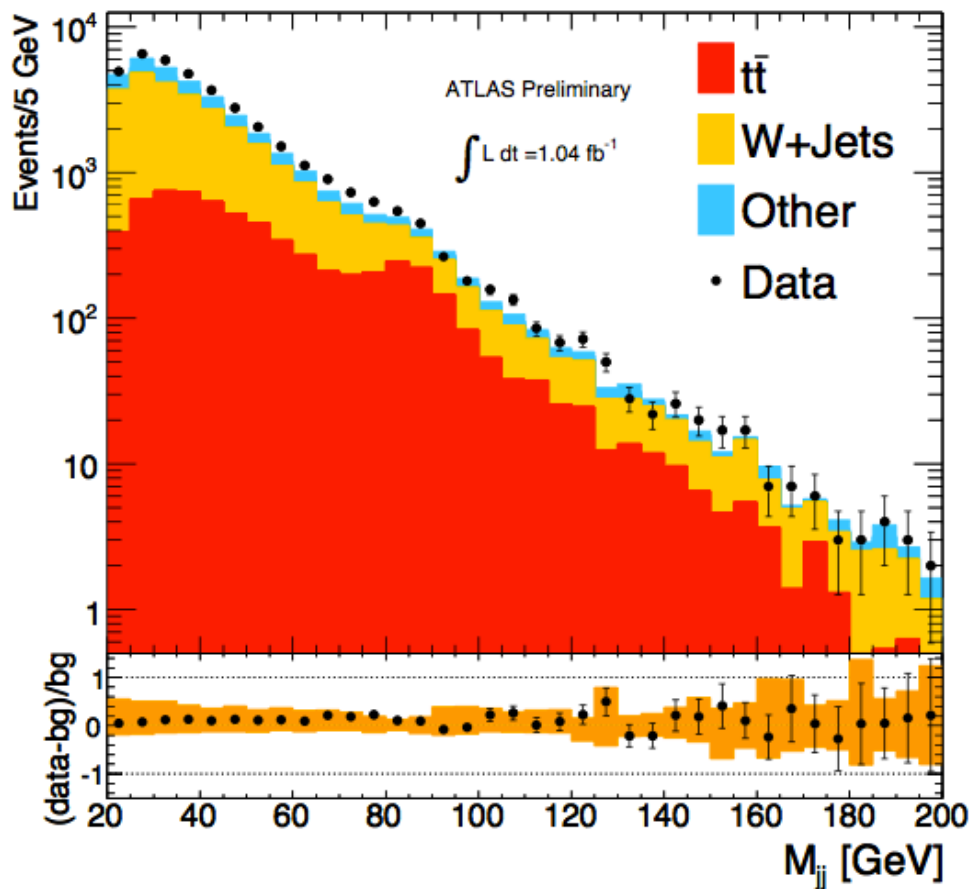
Heavy particles decaying to
top+jets **with boosted W's**



Find **standard jets**
within a cone
examine dijet mass

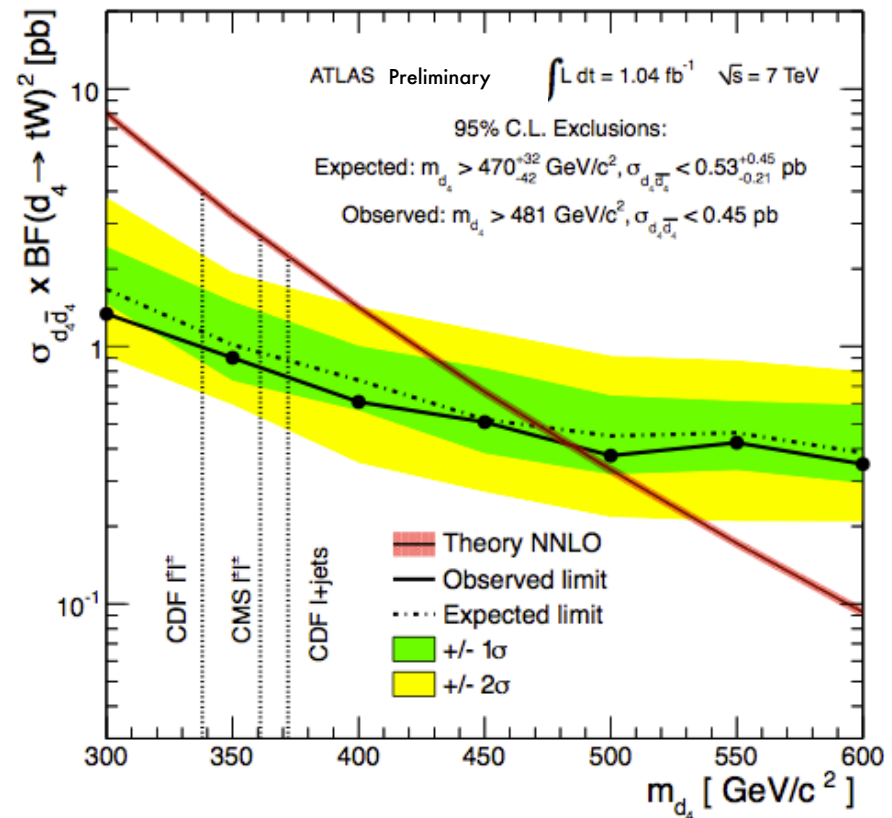
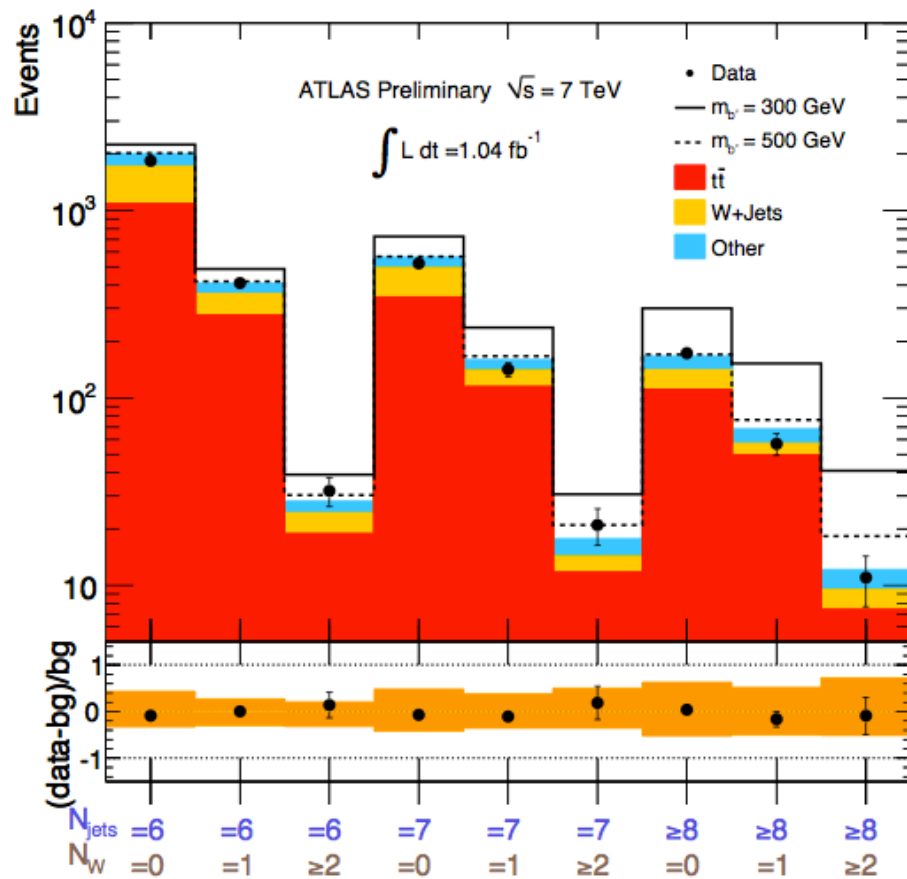


Validate in $t\bar{t}$

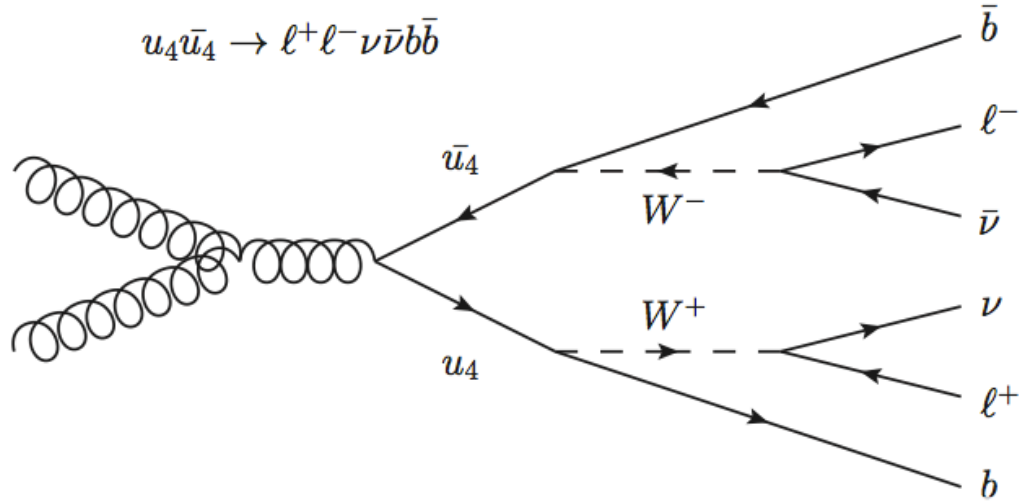


events with 3 or 4 jets

results



t'



Selection

Two OS lepton (e or μ)

$MET > 60$

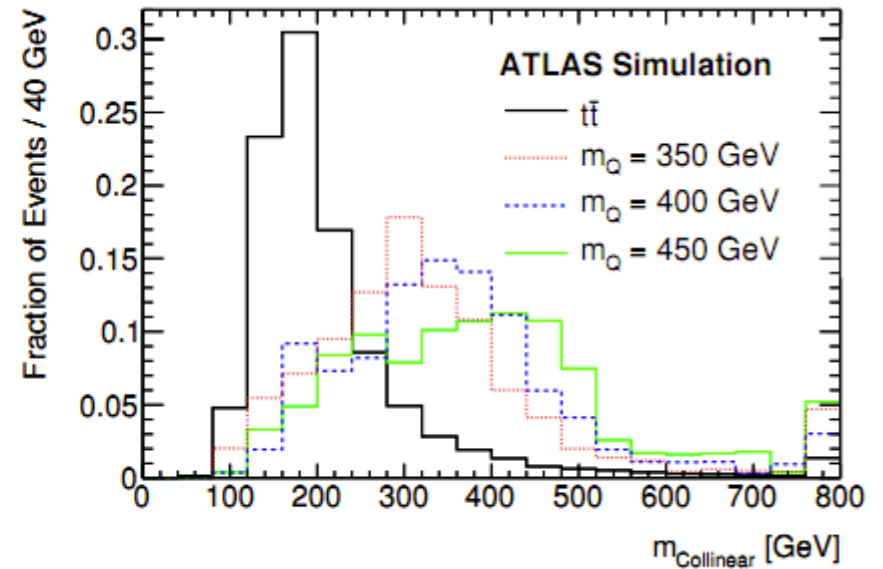
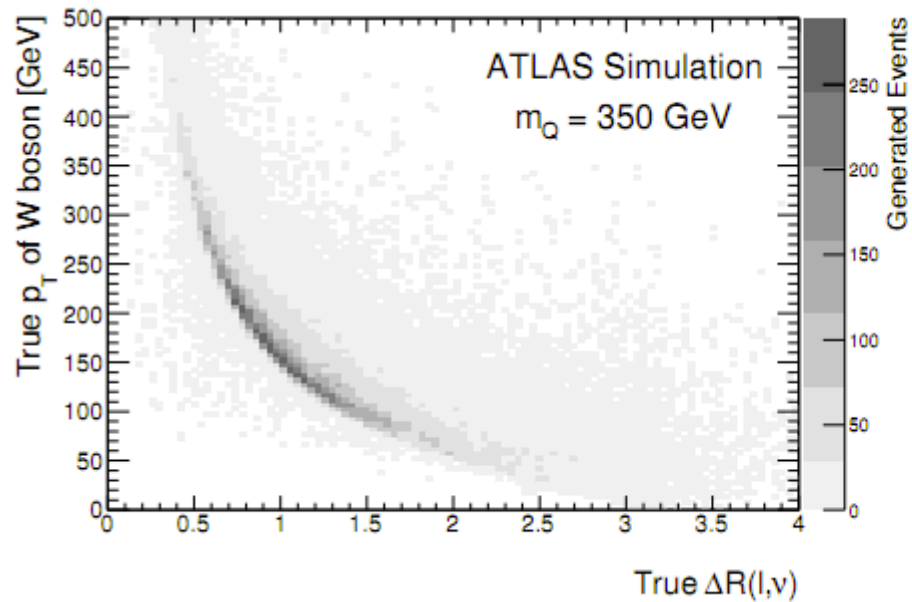
At least two jets

Dataset

ATLAS 1/fb

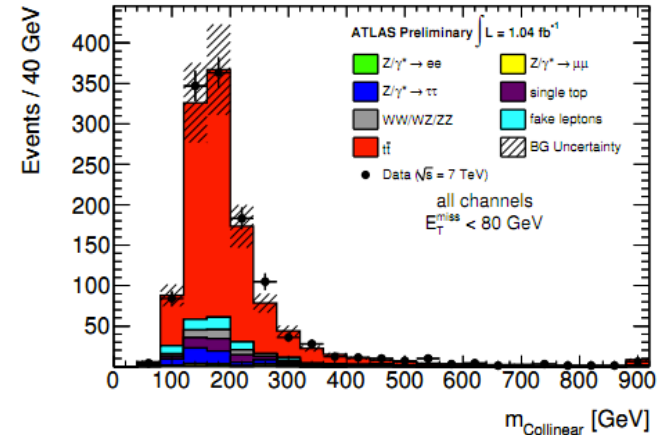
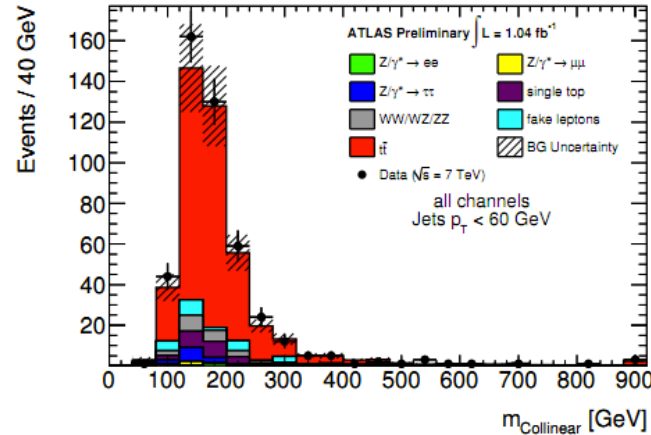
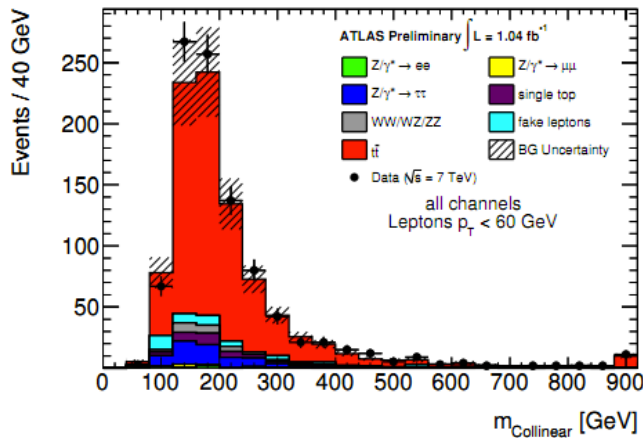
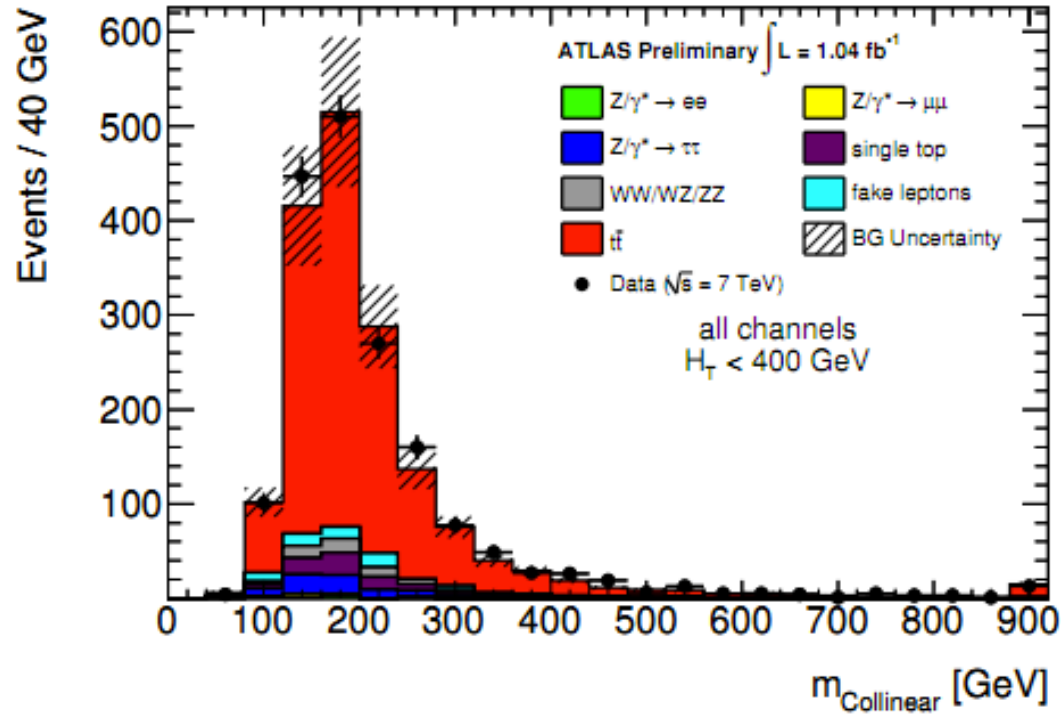
Analysis done largely by a single grad student: **Michael Werth, UCI**

OS $QQ \rightarrow WqWq$

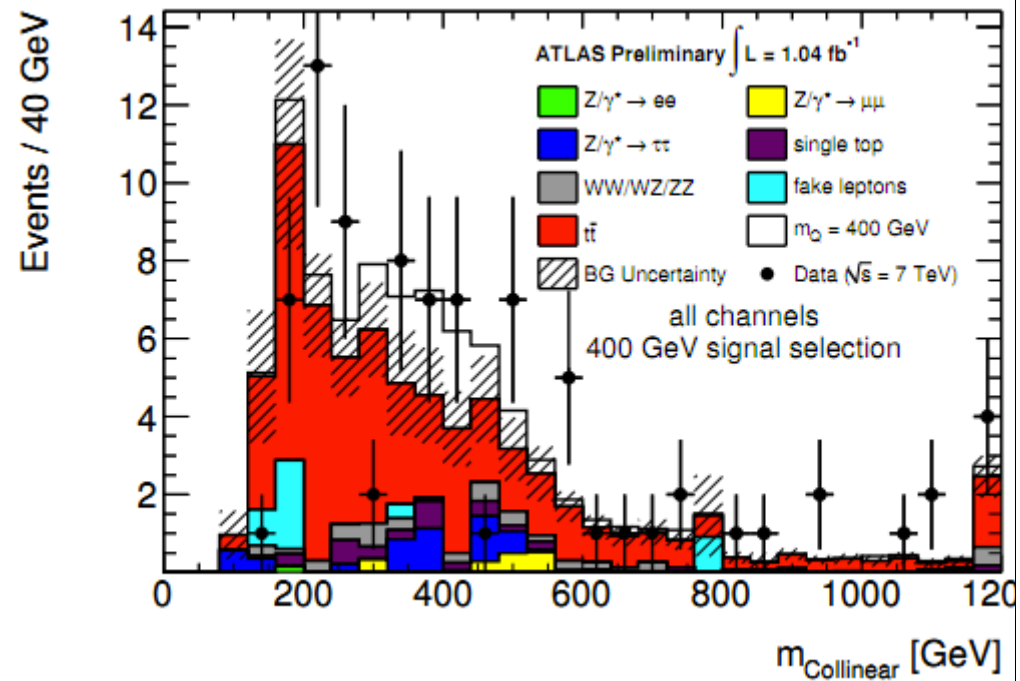
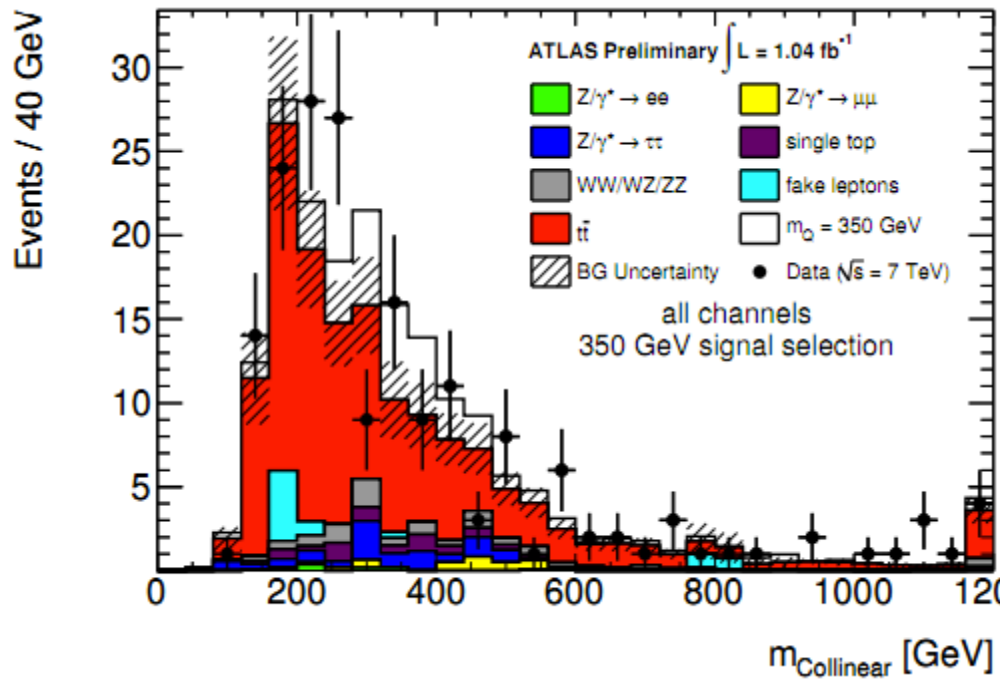


More W p_T means smaller opening angle between lepton and neutrino

Validation



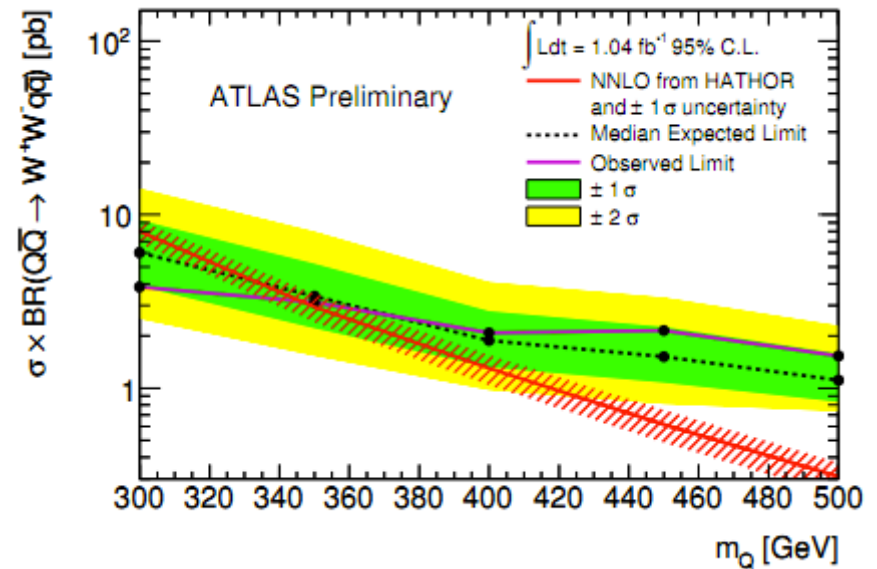
OS $QQ \rightarrow WqWq$



Results

TABLE V. Expected background, expected signal and observed data in ee , $\mu\mu$, and $e\mu$ channels for $m_Q = 300 - 500$ GeV after final selection. The uncertainties shown include both statistical and systematic contributions.

m_Q (GeV)	Expected Background	Expected Signal	Observed Data
300	300 $^{+40}_{-40}$	95 $^{+14}_{-12}$	315
350	148 $^{+22}_{-18}$	35 $^{+5}_{-4}$	180
400	75 $^{+11}_{-10}$	17.1 $^{+2.5}_{-2.1}$	89
450	49 $^{+8}_{-6}$	8.4 $^{+1.2}_{-1.0}$	57
500	30 $^{+5}_{-4}$	4.4 $^{+0.6}_{-0.5}$	36



Outline

I. Introduction: experimental innovation

II. Simple boosted objects

III. Novel views of top quark events

VI. Concluding rant: open data

Tops

Studied to death?

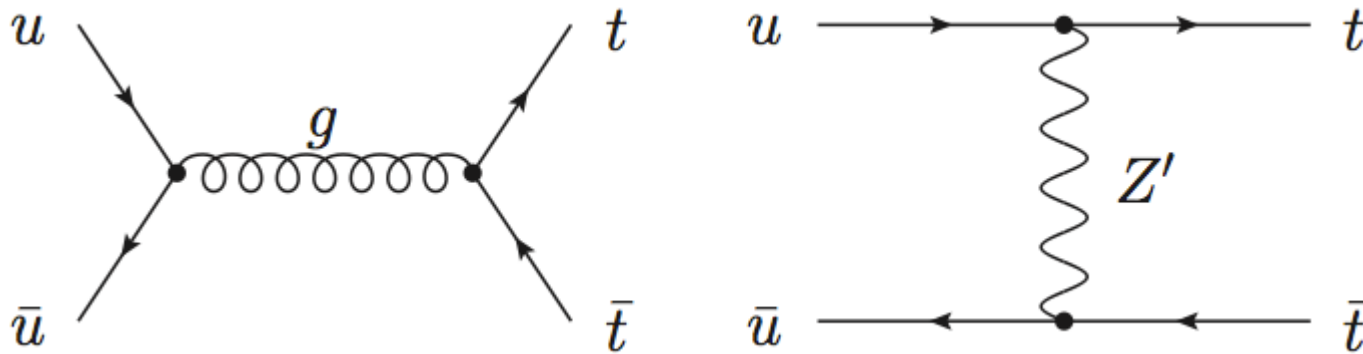
- M_t , V_{tb} , BRs, charge, total cross-section, W helicity. top width & lifetime, gg vs qq , $X \rightarrow tt$, etc...

What is left?

- Understand A_{FB} anomaly
- Resonances inside top-like events

Afb

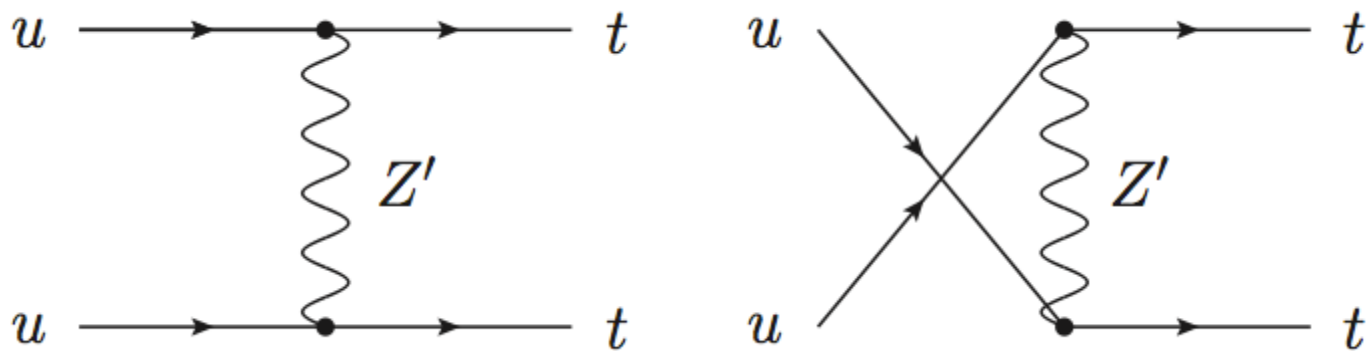
CDF's top A_{FB} can be explained



if there is a **new $t\bar{t}$ production**
mechanism via t-channel Z'

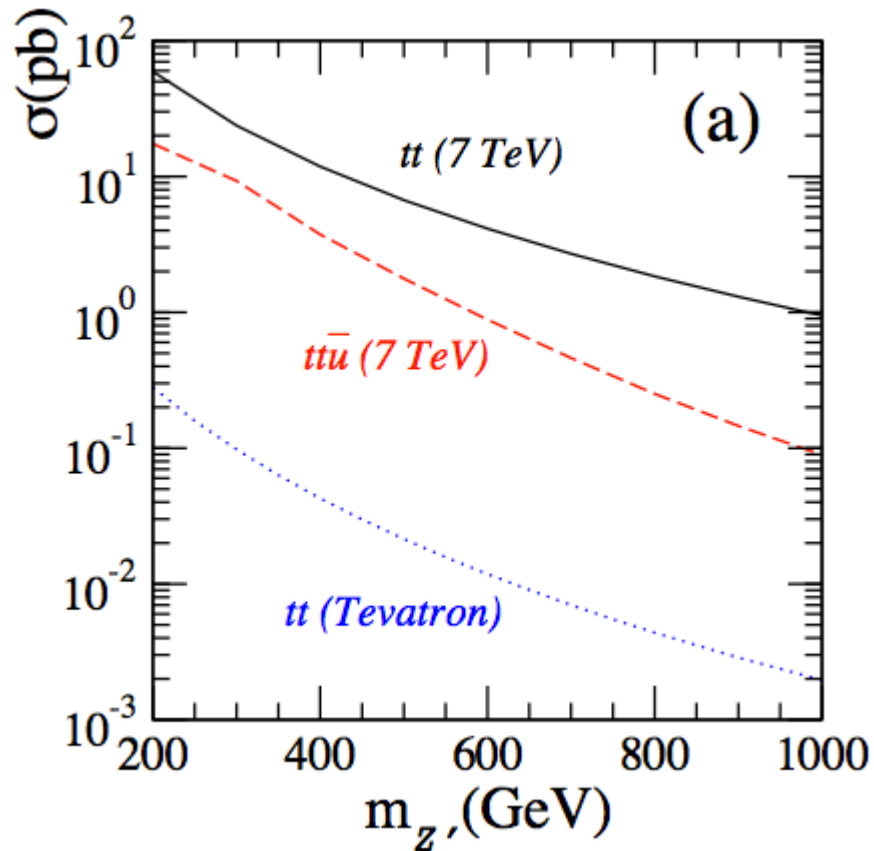
Testable hypothesis

Also predicts



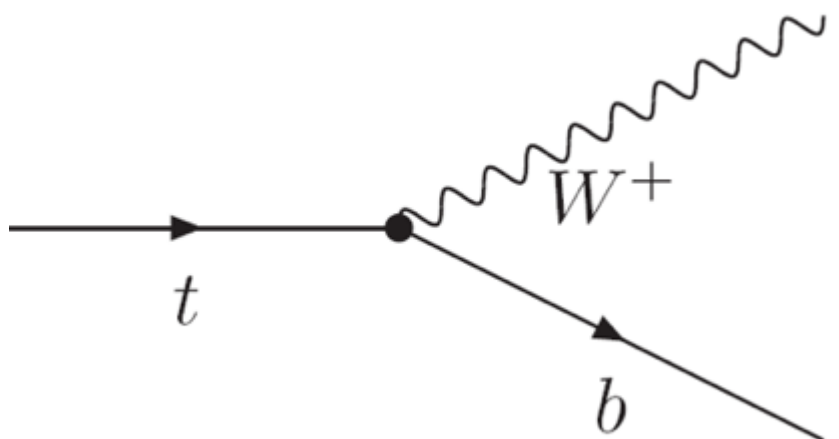
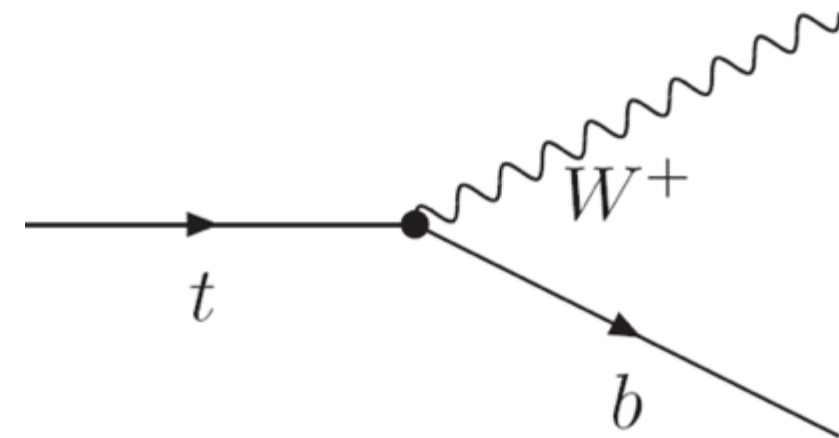
same-sign tops through $M \times FV$ coupling

Rates



Testable at the Tevatron, enormous at the LHC

Same sign tops



Same-sign tops

Gives ss W^+ 's

Gives ss dileptons

LHC

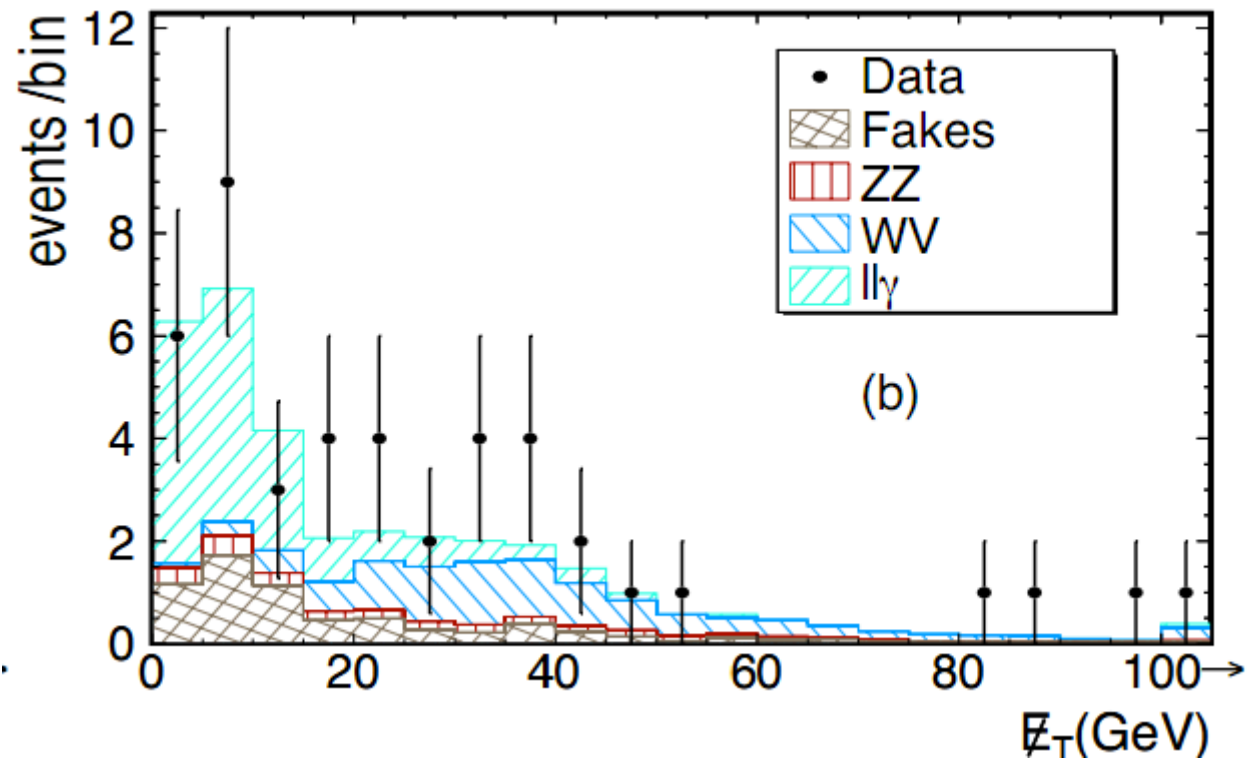
pp means $++ > --$

Smaller backgrounds

$t\bar{t}$ jets charge asymmetry

CDF like-sign dileptons

	n_{obs}	n_{pred}
$e_{\text{si}}e_{\text{si}}$	11	6.3 ± 1.0
ee	3	1.3 ± 0.3
$e_{\text{si}}e$	9	9.1 ± 1.8
$e_{\text{si}}\mu$	11	6.8 ± 0.8
$e\mu$	5	6.4 ± 1.2
$\mu\mu$	5	3.2 ± 0.3
Total	44	33.2 ± 4.7



1/fb PRL 2007

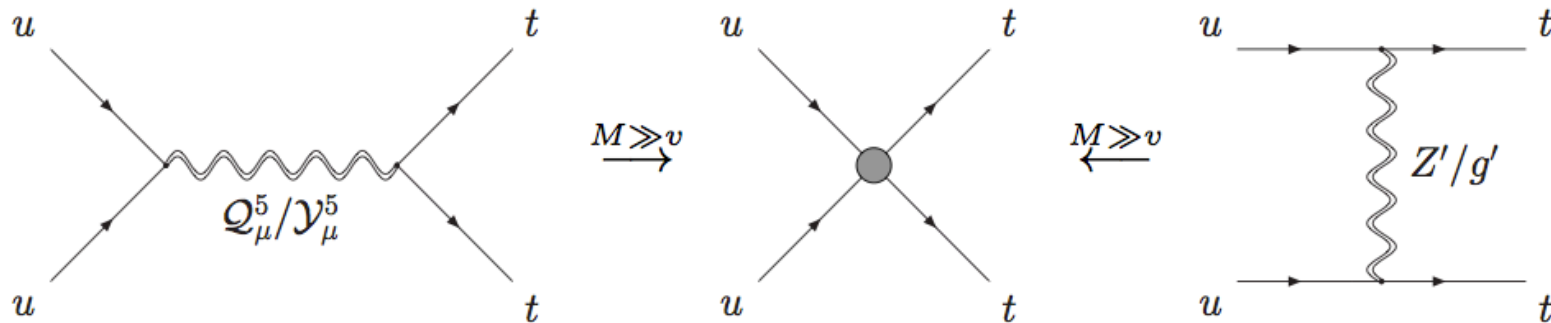
Is dileptons 6.1/fb



<u>CDF RunII Preliminary</u>	<u>6.1/fb</u>
top quark pairs	0.1 ± 0.1
Z	26.6 ± 3.4
WW, WZ, ZZ	28.4 ± 2.0
W+gamma	16.2 ± 2.4
Fakes	51.6 ± 24.2
<u>Total</u>	<u>123.0 ± 24.6</u>
Data	145

same-sign tops

Many models



Use 4f effective operators
(LL,LR,RR) modes

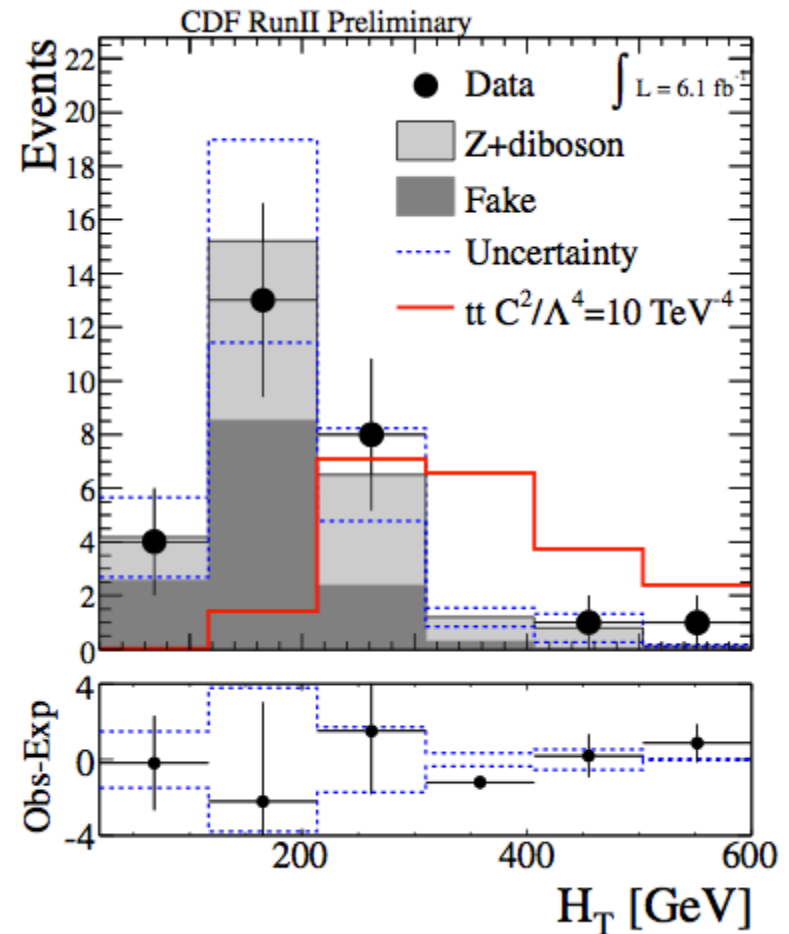
same-sign leptons+2jets



Process	Total ll
$t\bar{t}$	0.1 ± 0.0
$Z \rightarrow ll$	5.9 ± 1.7
WW, WZ, ZZ	7.2 ± 0.5
$W(\rightarrow l\nu)\gamma$	0.9 ± 0.7
Fakes	13.8 ± 7.2
Total	28.0 ± 7.5
Data	27

coupling $|C|/\Lambda^2$

cross-section $\propto C^2/\Lambda^4$



CMS

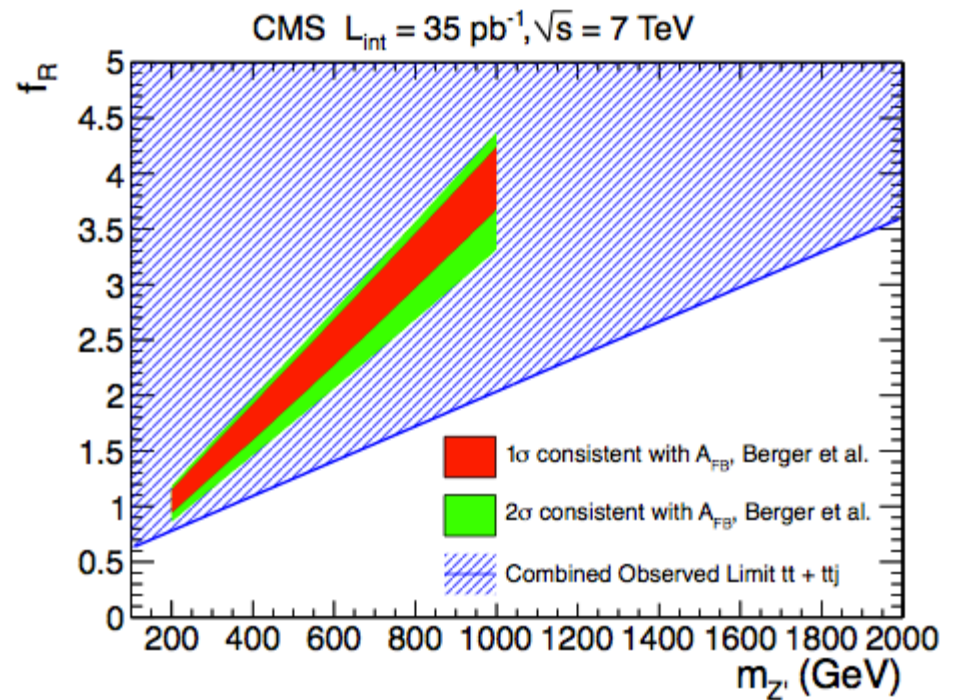
Expect

3.4 ± 1.8

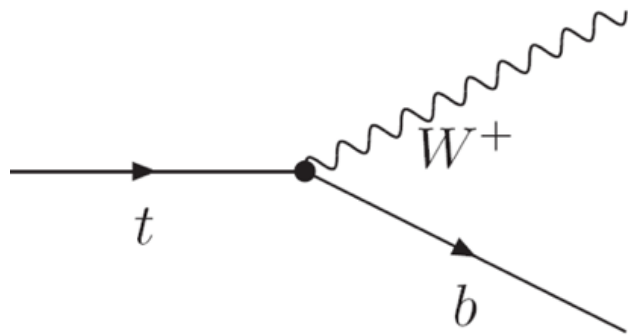
Observe

3

Rule out a bunch of Z models



ATLAS

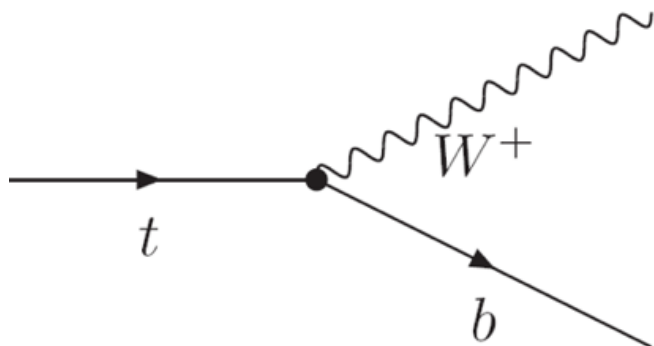


Selection

Two SS lepton (e or μ)

MET > 40

At least two jets



Dataset

ATLAS 1/fb

Analysis led by
UCI postdoc : **Ning Zhou**

ATLAS



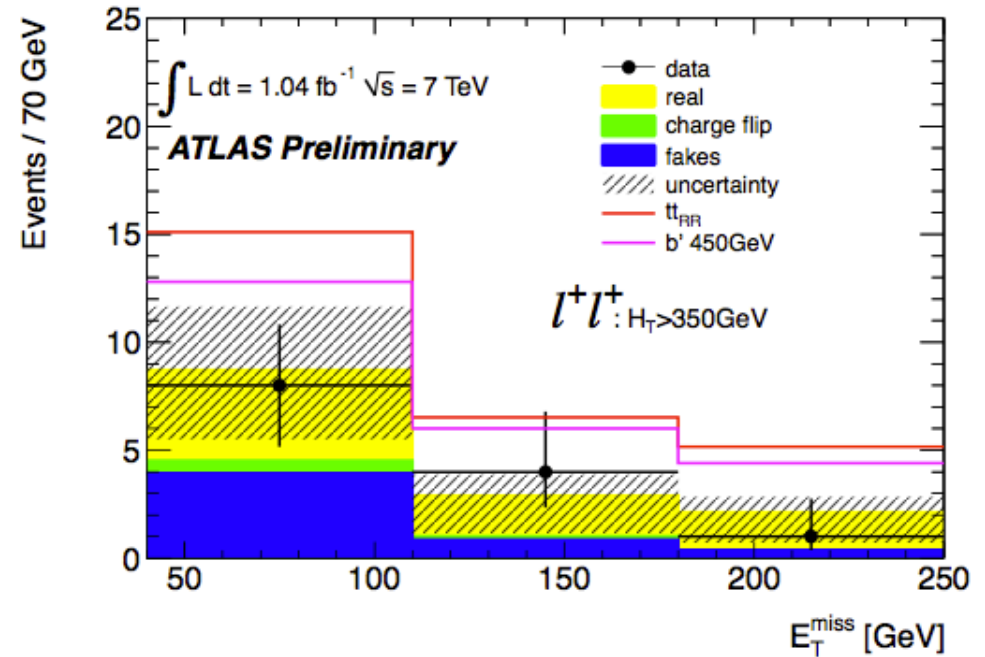
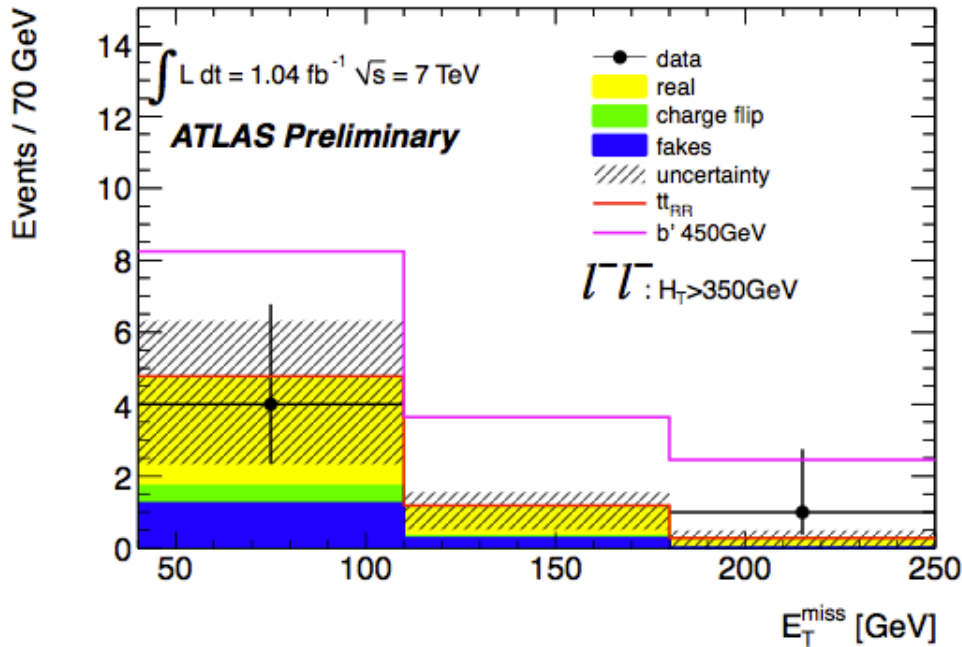
negative
charge

	e^-e^-	$\mu^-\mu^-$	$e^-\mu^-$
Fake	$0.2 \pm 0.3 \pm 0.1$	$0.7 \pm 0.3^{+0.6}_{-0.3}$	$0.5 \pm 0.2^{+0.7}_{-0.3}$
Charge flip	$0.3 \pm 0.1^{+0.3}_{-0.1}$	$0 \pm 0^{+0.01}_{-0}$	$0.3 \pm 0.1^{+0.2}_{-0.1}$
Real	$0.8 \pm 0.3^{+0.2}_{-0.5}$	$1.0 \pm 0.3^{+0.3}_{-0.6}$	$2.3 \pm 0.5^{+0.6}_{-1.8}$
Total	$1.4 \pm 0.4^{+0.3}_{-0.6}$	$1.7 \pm 0.4 \pm 0.7$	$3.1 \pm 0.6^{+1.0}_{-1.8}$
Data	1	2	2
tt_{LL}	$0.2 \pm 0.1 \pm 0.03$	$0.2 \pm 0.1 \pm 0.02$	$0.5 \pm 0.3 \pm 0.1$
tt_{LR}	$0.02 \pm 0.01 \pm 0.01$	$0.001^{+0.01}_{-0.001} \pm 0.001$	$0.02 \pm 0.02 \pm 0.01$
tt_{RR}	$0.5 \pm 0.3 \pm 0.1$	$0.1 \pm 0.1 \pm 0.2$	$0.8 \pm 0.3 \pm 0.1$
b' 450 GeV	$1.8 \pm 0.2 \pm 0.2$	$2.1 \pm 0.2 \pm 0.2$	$4.3 \pm 0.3 \pm 0.4$

positive
charge

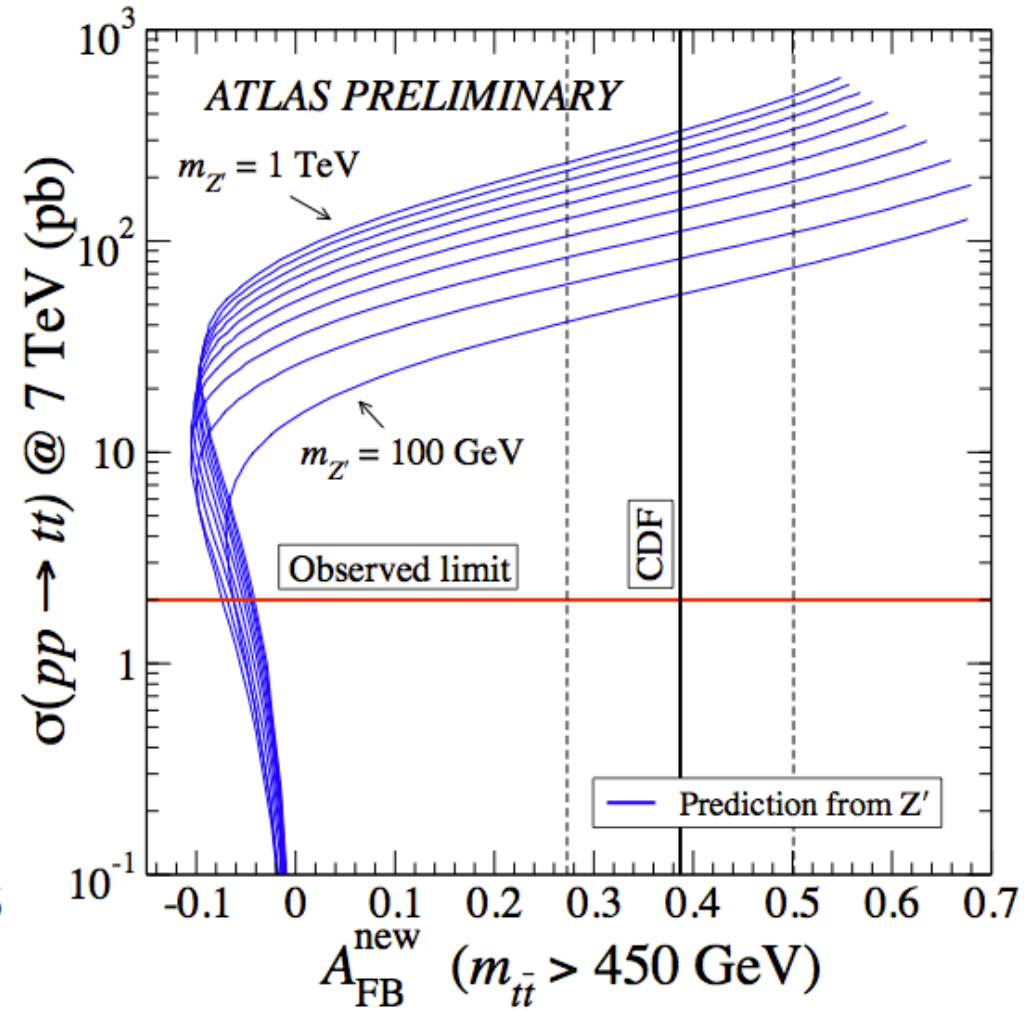
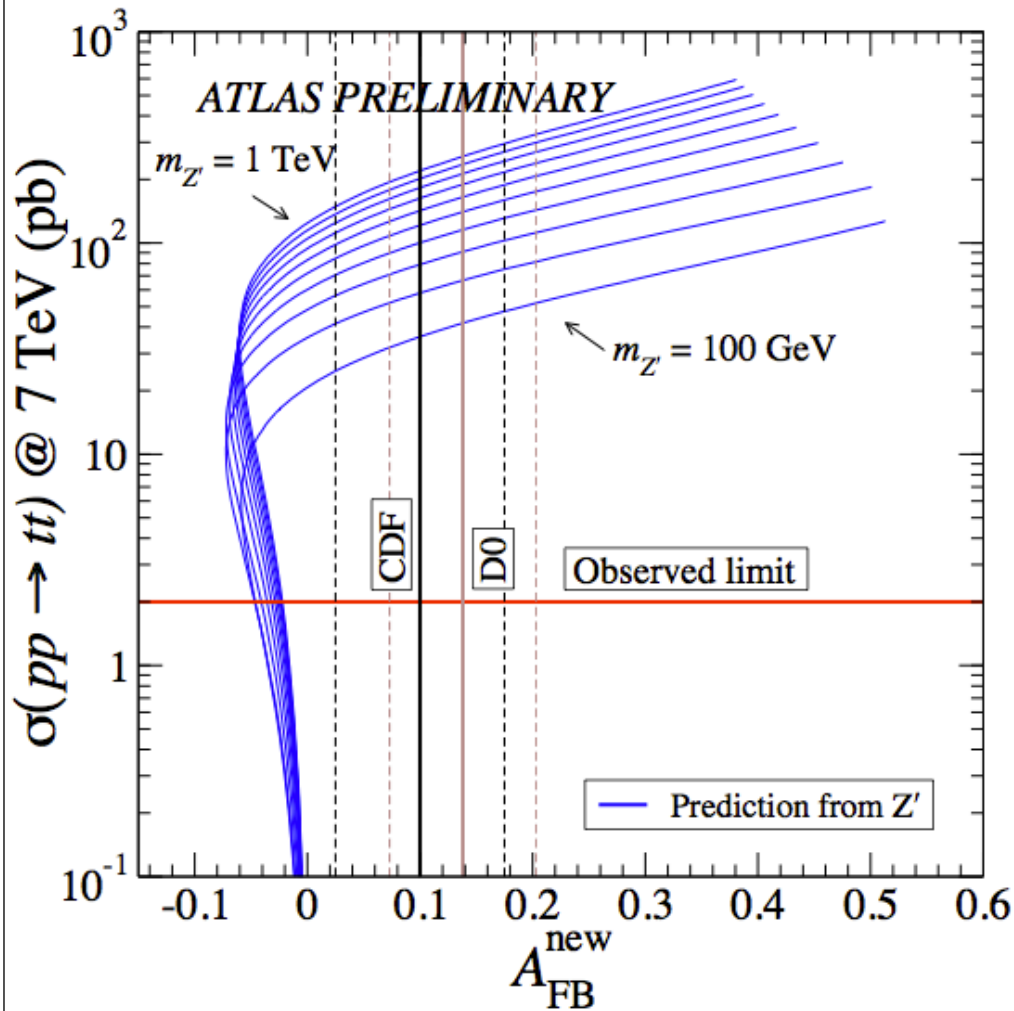
	e^+e^+	$\mu^+\mu^+$	$e^+\mu^+$
Fake	$0.8 \pm 0.6^{+0.2}_{-0.4}$	$1.0 \pm 0.3^{+0.6}_{-0.4}$	$3.3 \pm 1.1^{+1.6}_{-1.4}$
Charge flip	$0.3 \pm 0.1^{+0.3}_{-0.1}$	$0 \pm 0^{+0.01}_{-0.0}$	$0.4 \pm 0.1^{+0.3}_{-0.1}$
Real	$1.9 \pm 0.5^{+0.5}_{-1.4}$	$1.6 \pm 0.3^{+0.6}_{-0.9}$	$4.4 \pm 0.7^{+1.1}_{-3.0}$
Total	$3.0 \pm 0.8^{+0.6}_{-1.4}$	$2.6 \pm 0.4^{+0.9}_{-1.1}$	$8.1 \pm 1.3^{+2.1}_{-3.3}$
Data	2	1	10
tt_{LL}	$30.1 \pm 1.9 \pm 4.6$	$30.4 \pm 1.9^{+4.3}_{-4.4}$	$64.2 \pm 2.8^{+9.8}_{-9.9}$
tt_{LR}	$3.8 \pm 0.2 \pm 0.6$	$4.2 \pm 0.3 \pm 0.6$	$8.3 \pm 0.4 \pm 1.2$
tt_{RR}	$35.5 \pm 2.1 \pm 5.6$	$29.5 \pm 1.9 \pm 4.2$	$65.7 \pm 2.8 \pm 10.0$
b' 450 GeV	$1.8 \pm 0.2 \pm 0.2$	$2.7 \pm 0.2 \pm 0.3$	$5.0 \pm 0.3 \pm 0.6$

ATLAS

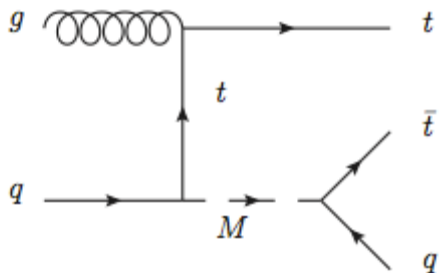


ATLAS sees a few events, but consistent with background.

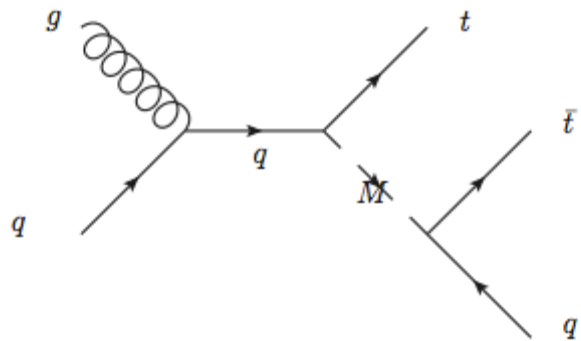
Interpretation



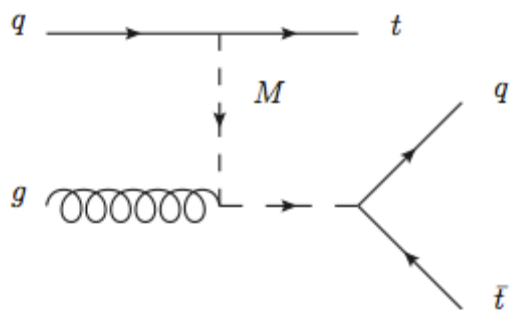
$t\bar{t}+U$



M not self-conjugate,
so no ss tops!



$qg \rightarrow Mt \rightarrow$
 $(t\bar{t} + q) t \rightarrow t\bar{t} + q$



Look for resonance in $t+j$

But...

s- channel: can't generate asymmetry to explain top A_{FB} measurement

t- channel:

light mediators constrained by overall $t\bar{t}$ cross-sec

heavy mediators constrained by CDF invariant mass distribution

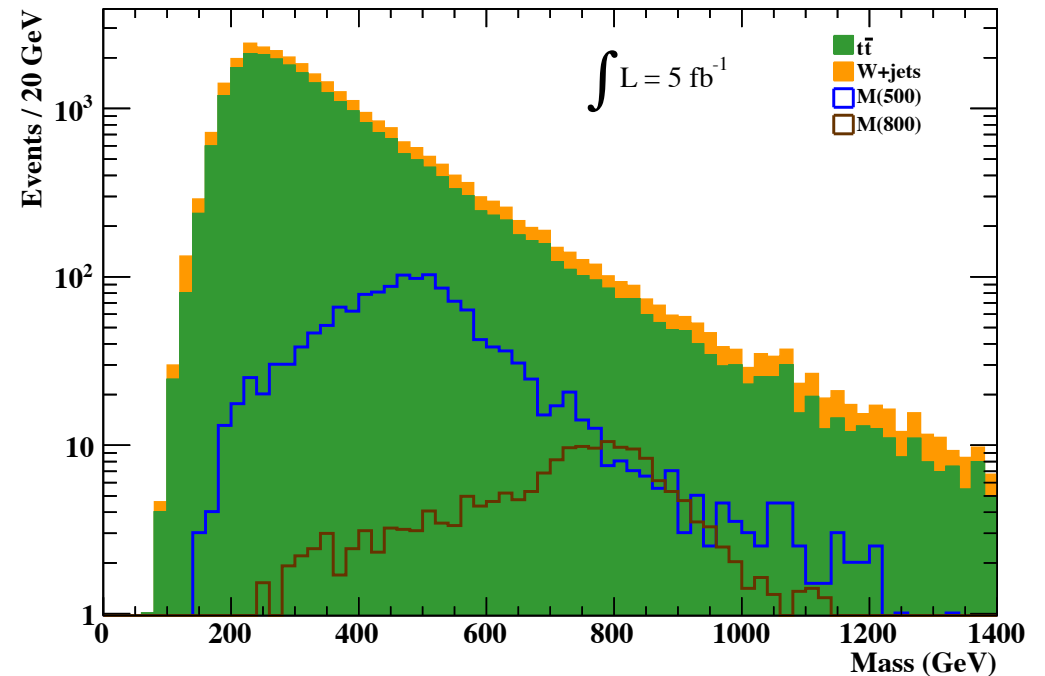
$t\bar{t}j$ resonance



Resonance

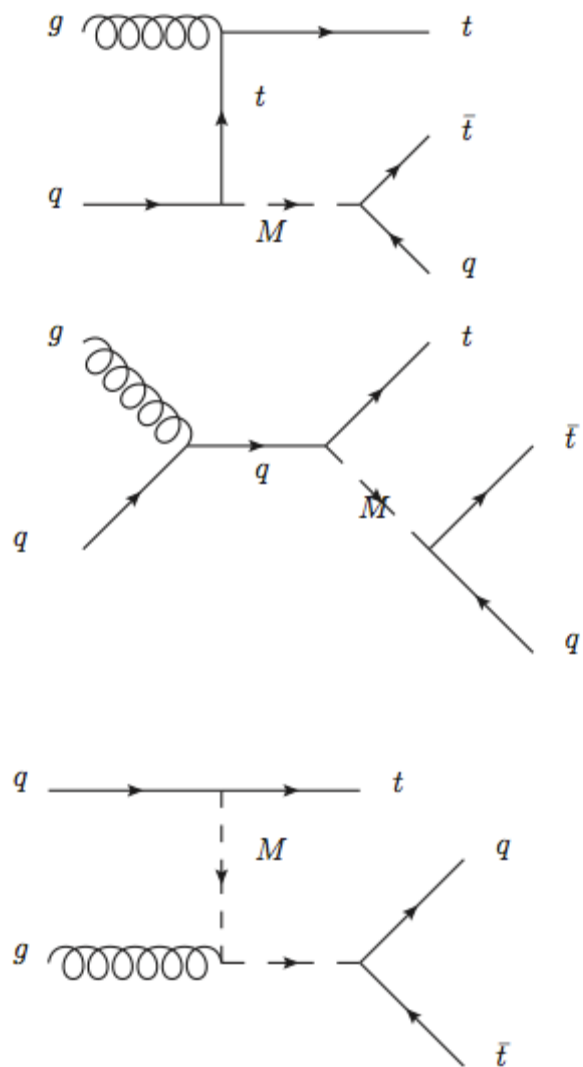
$t\bar{t}j$ resonance has never been examined

$t\bar{t}j$ mass at 500 and 800 GeV



Analysis done largely by a single grad student: **Kanishka Rao, UCI**

Selection & dataset



Selection

One lepton (e or μ)

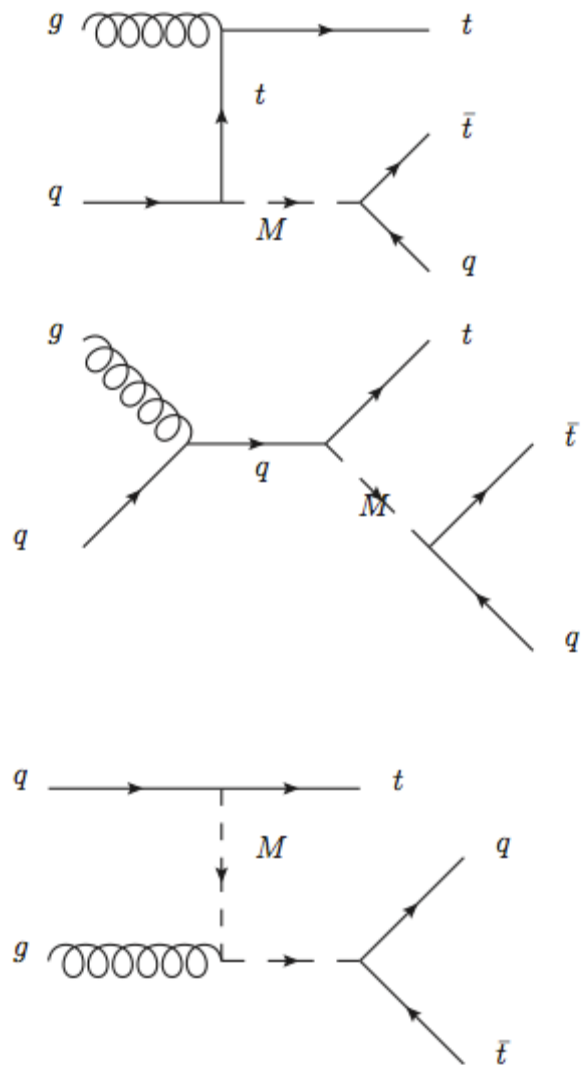
$MET > 20$

At least five jets

Dataset

CDF 8.7/fb

Reconstruction



Identify top, antitop

use standard kinematic fitter

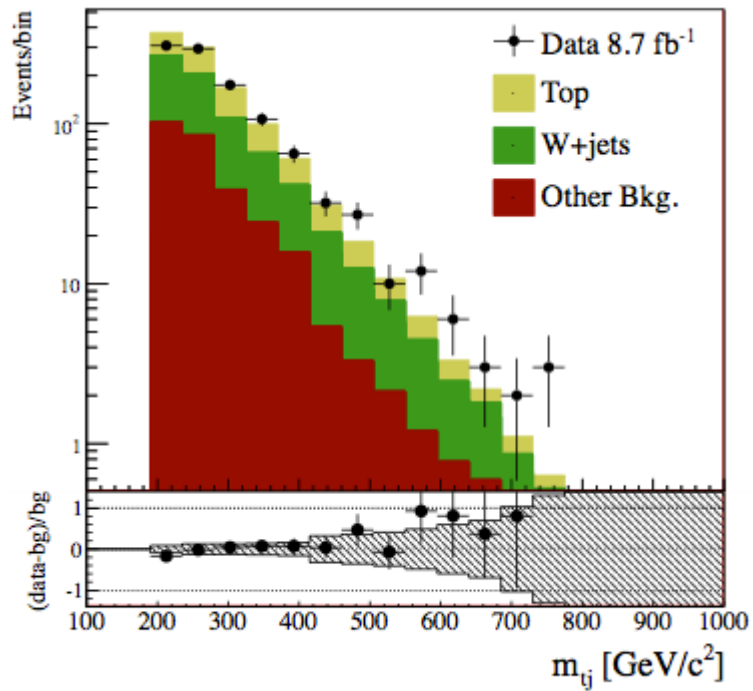
Choose additional jet

maximize m_{tj}

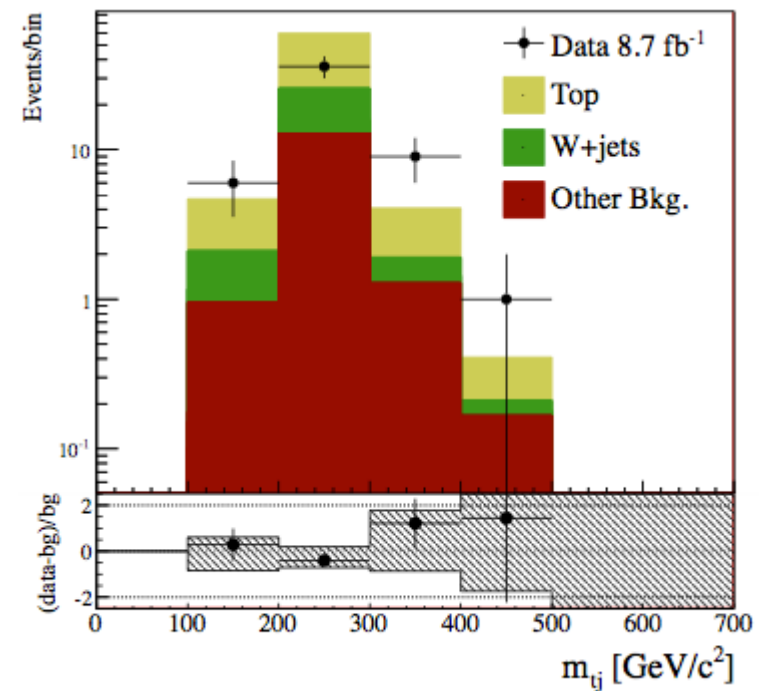
Validation



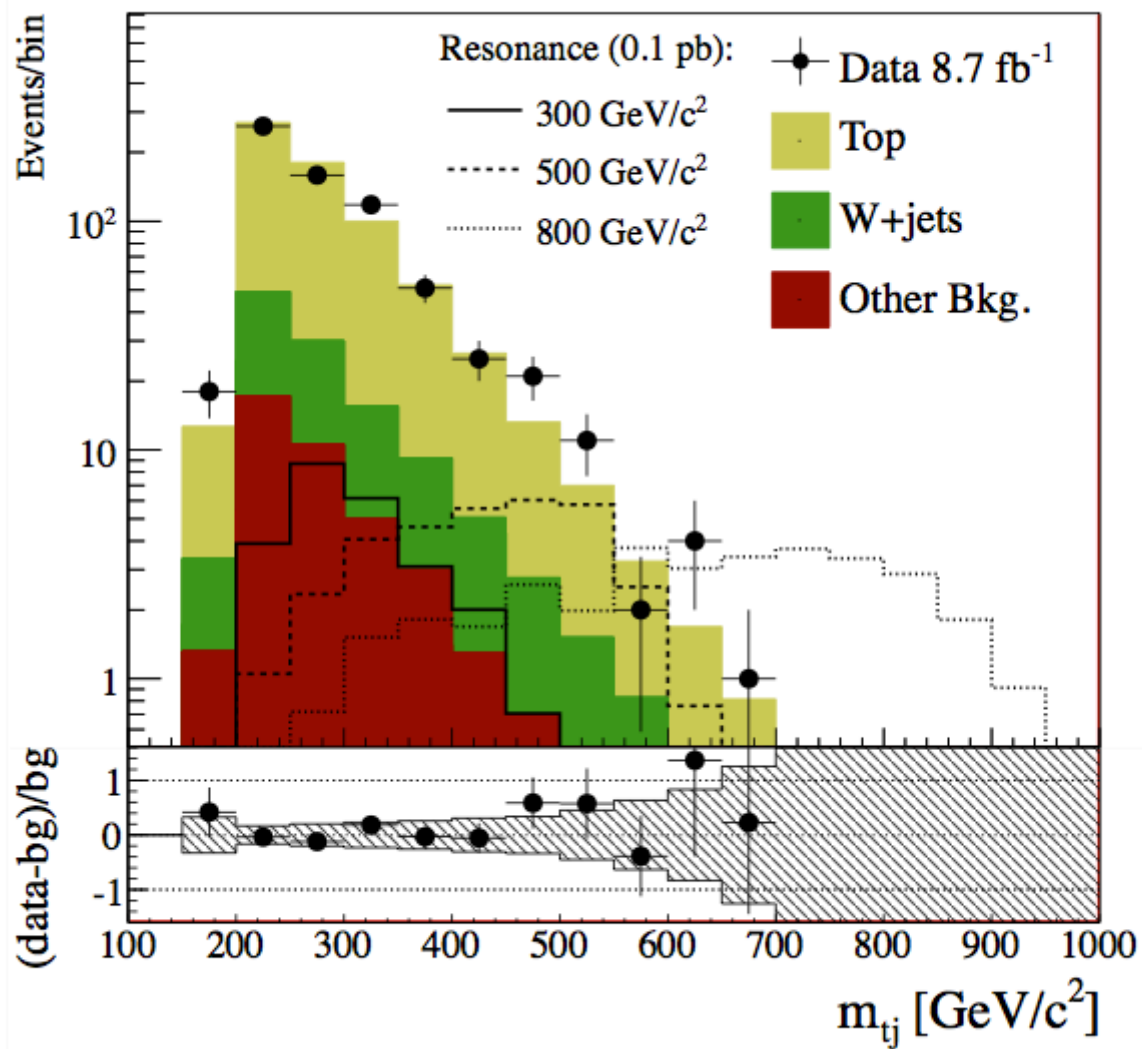
5-jet, 0tag



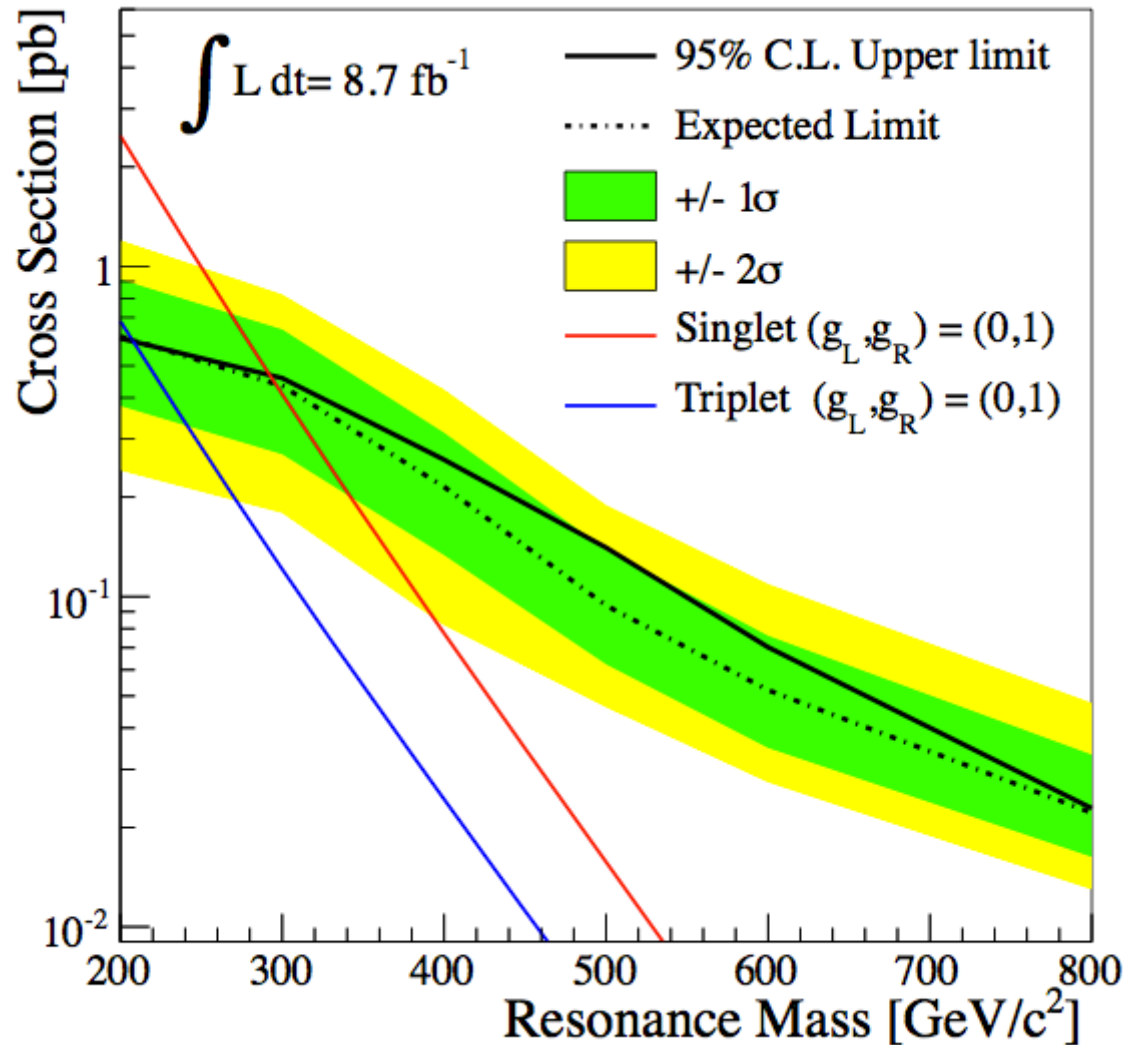
5-jet, 1tag, low HT



CDF $t\bar{t}j$

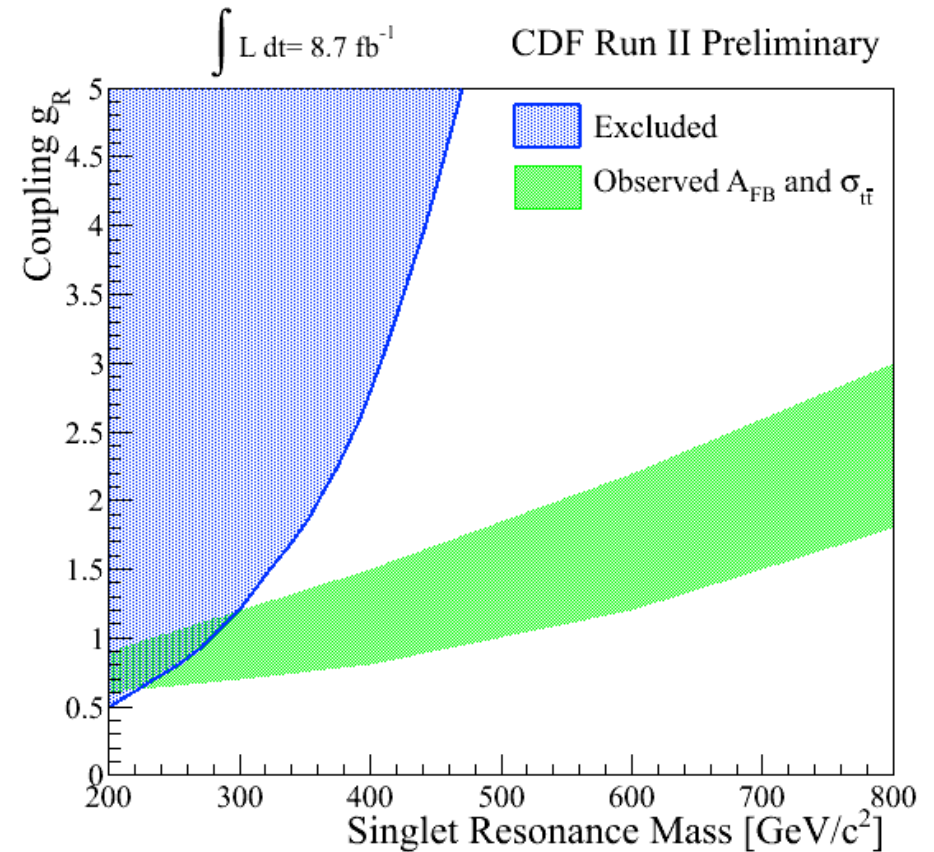
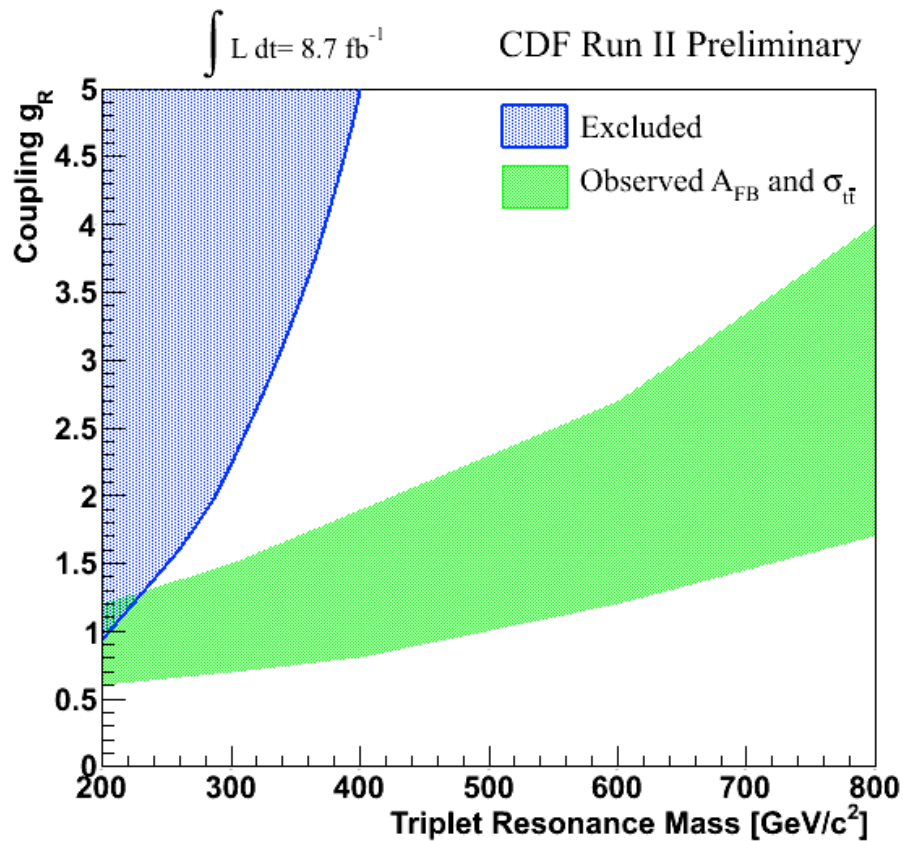


CDF $t\bar{t}$

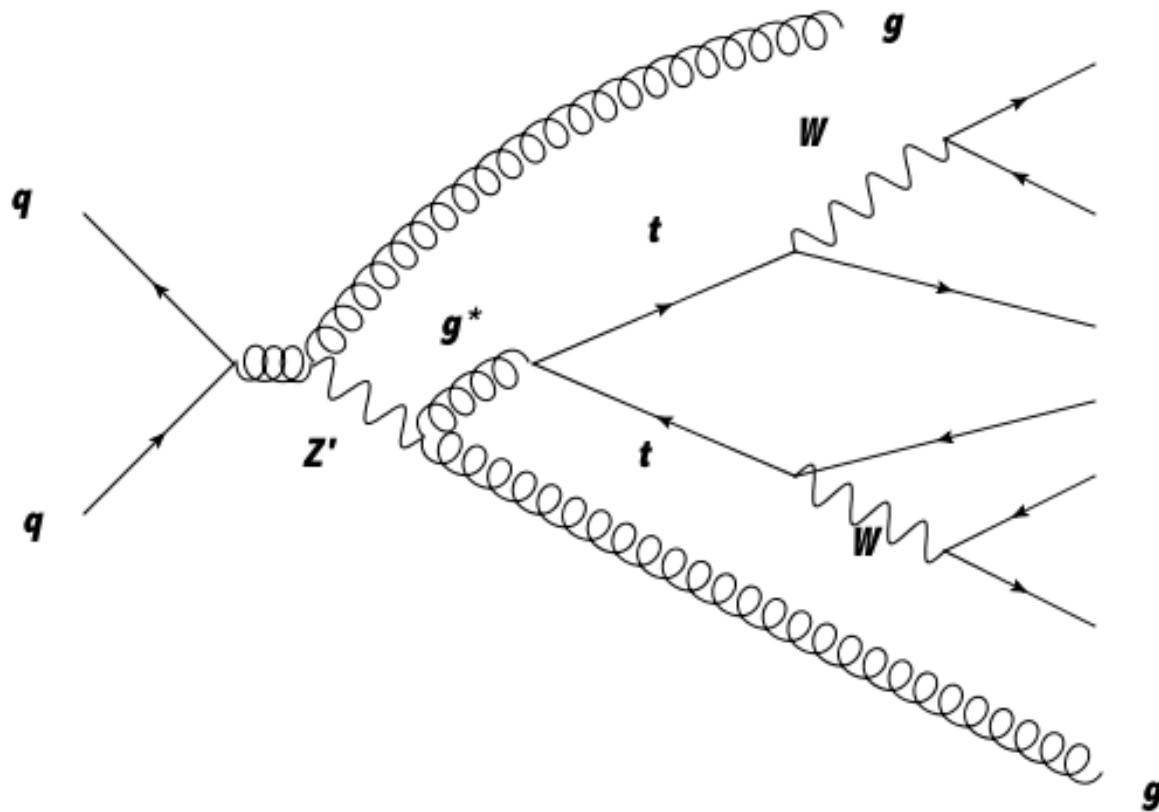


Analysis done largely by a single grad student: **Kanishka Rao, UCI**

Interpretation

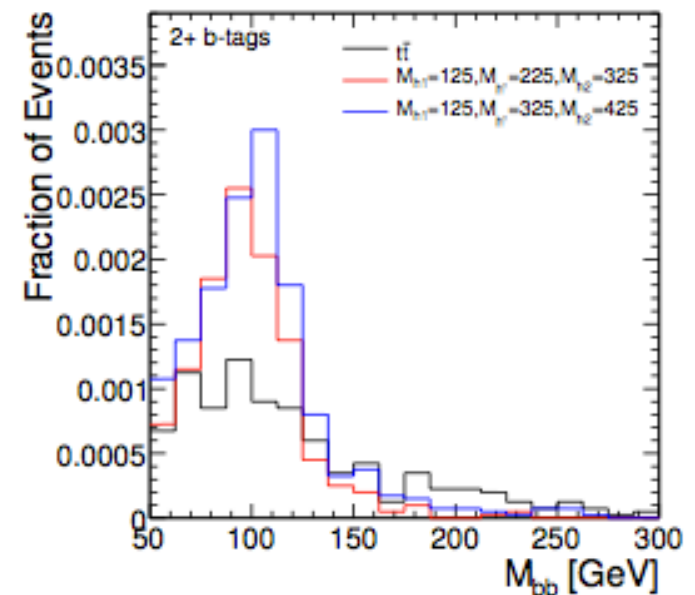
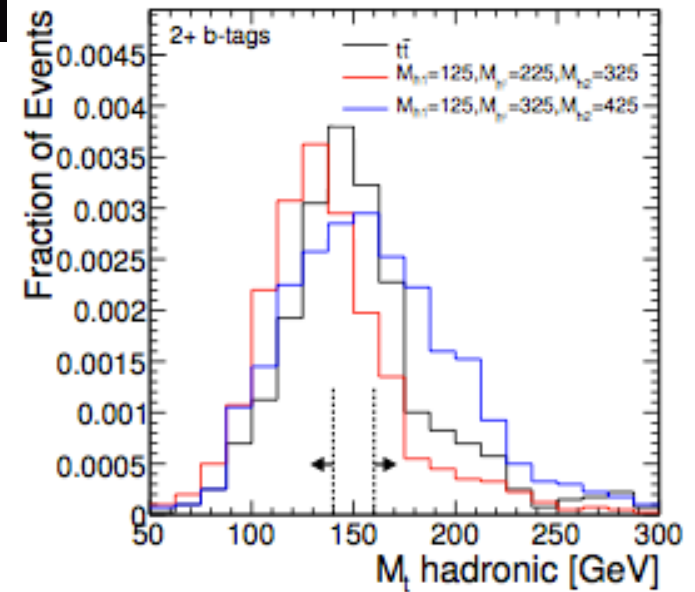
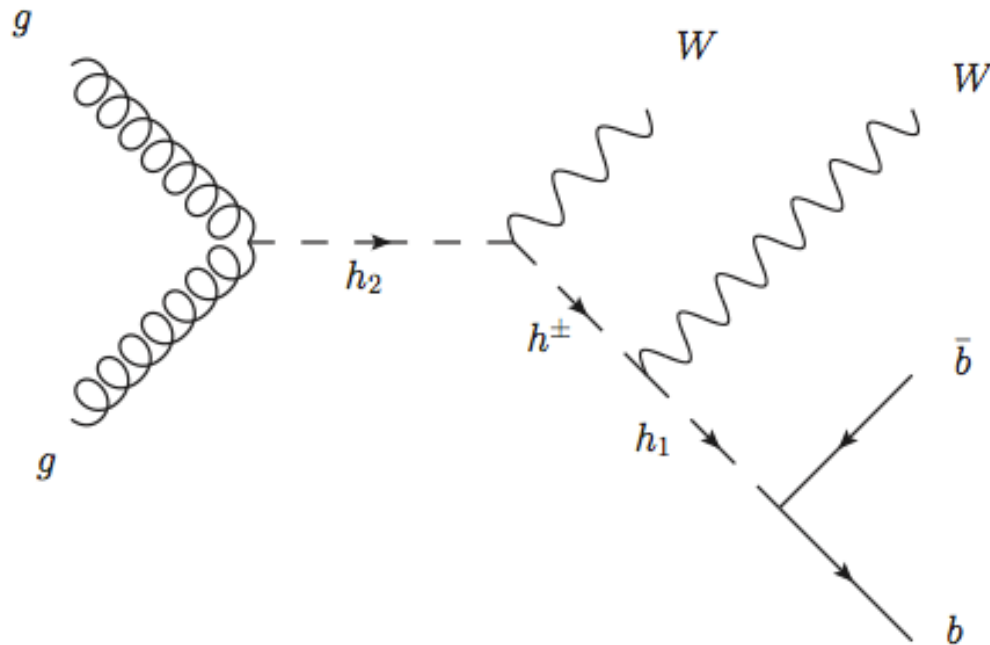


$Z' \rightarrow gg$ giving ttj resonance



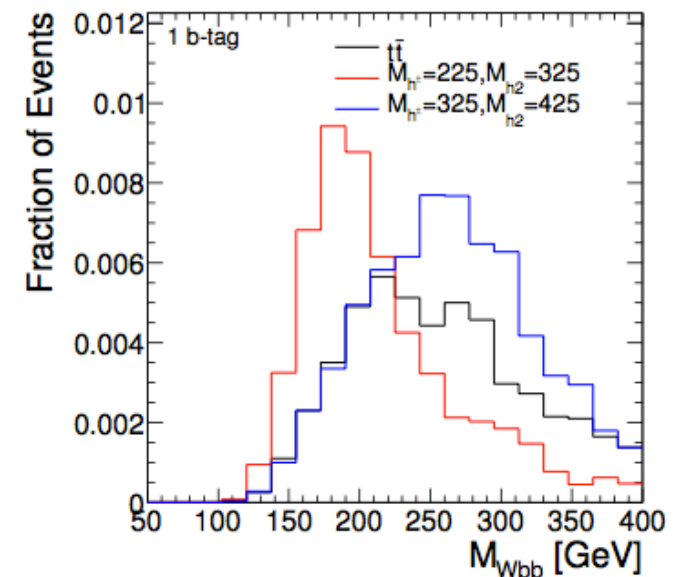
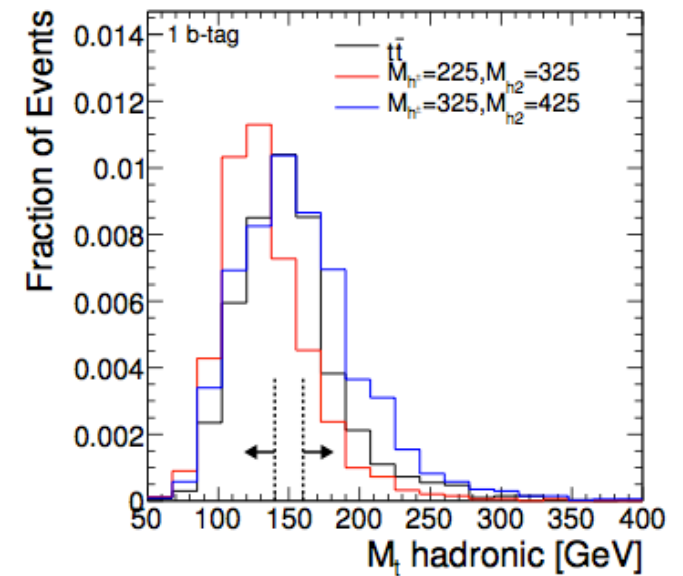
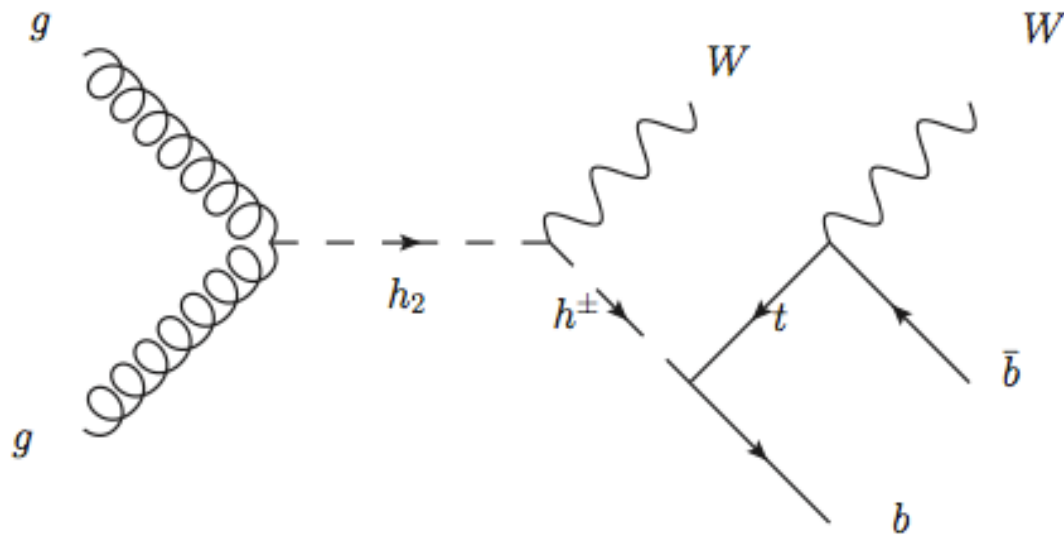
With Alwall, Rajaraman (to appear)

Higgs in WWbb



With Evans, Luty, Kilminster (arXiv:1201.3691)

Higgs in WWbb



With Evans, Luty, Kilminster (arXiv:1201.3691)

Outline

I. Introduction: experimental innovation

II. Simple boosted objects

III. Novel views of top quark events

VI. Concluding rant: open data

Open data

“big learn”: machine learning in big data

Lots of smart people & good ideas

Few real applications (Netflix)

Number of people who wanted to work on HEP

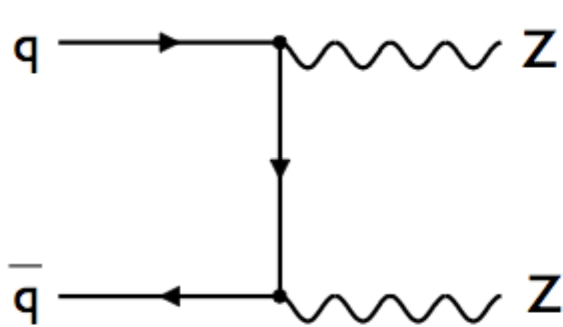
after my talk: > 25

after I said the data is not public: 0

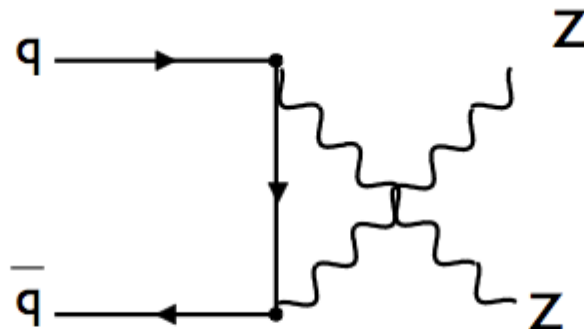


That's it!

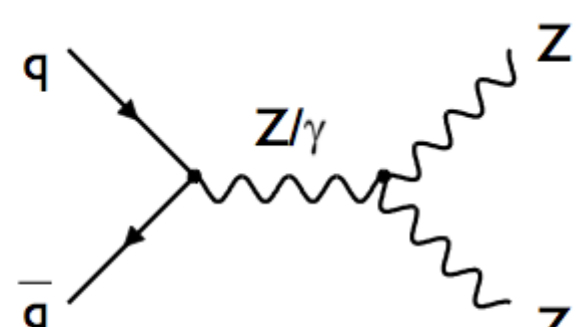
4 charged leptons



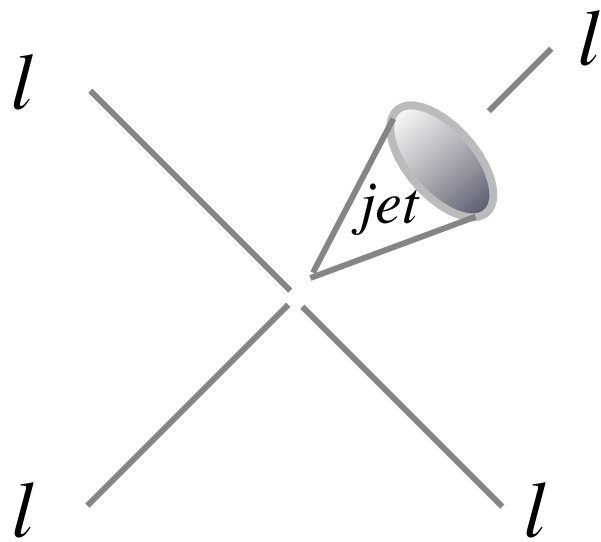
Standard Model Production



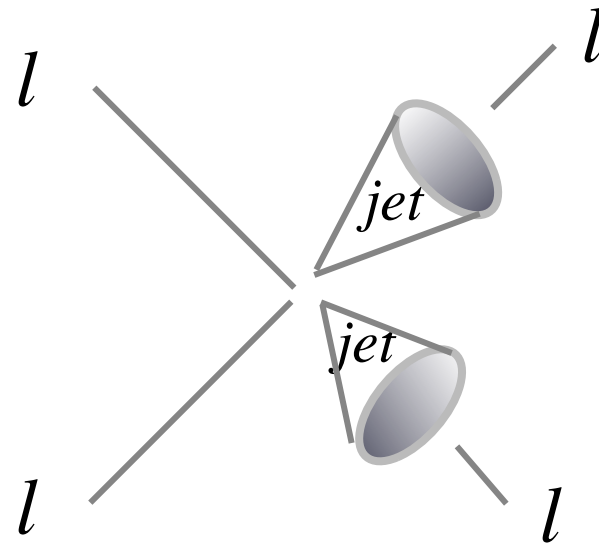
SM Forbidden



Backgrounds?



$WZ+jet$



$Z+jet$
 $tt \rightarrow ll bb vv$

Lepton types

Final State	$e^+e^-e^+e^-$	$\mu^+\mu^-\mu^+\mu^-$	$e^+e^-\mu^+\mu^-$	$\ell^+\ell^-\ell^+\ell^-$
Observed	2	8	2	12
Bkg(data-driven)	$0.01^{+0.03+0.05}_{-0.01-0.01}$	$0.3^{+0.9}_{-0.3} \pm 0.3$	$< 0.01^{+0.03}_{-0.01}$	$0.3^{+0.9+0.4}_{-0.3-0.3}$
Expected ZZ	$1.57 \pm 0.03 \pm 0.11$	$3.09 \pm 0.04 \pm 0.06$	$4.5 \pm 0.1 \pm 0.2$	$9.1 \pm 0.1 \pm 0.3$

Table 2: Summary of observed events, total background contributions and expected signal in the individual four-lepton and combined channels. The first error is statistical while the second is systematic. The uncertainty on the luminosity is not included. The errors on the background estimates span the 68% confidence interval, which is not symmetric about the best estimate because the background cannot be negative.

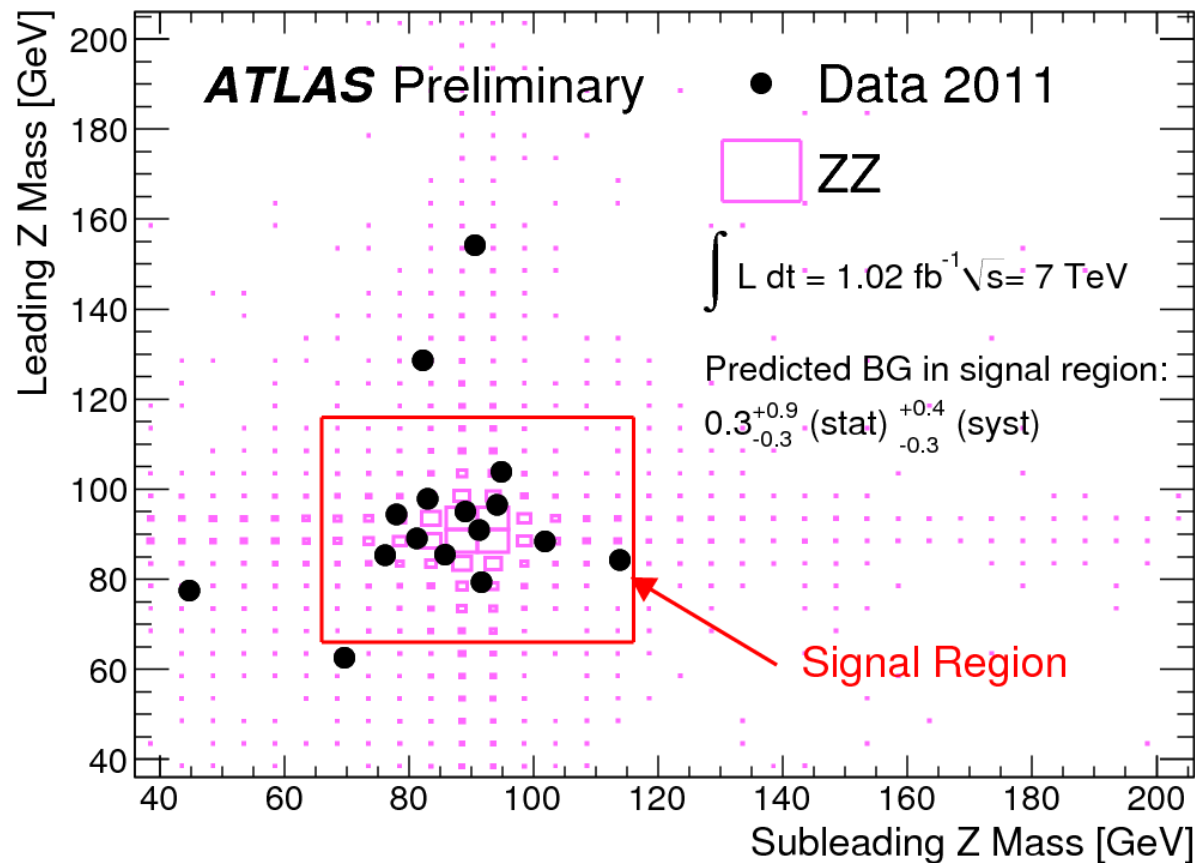


Lepton types

Final State	$e^+e^-e^+e^-$	$\mu^+\mu^-\mu^+\mu^-$	$e^+e^-\mu^+\mu^-$	$\ell^+\ell^-\ell^+\ell^-$
Observed	2	8	2	12
Bkg(data-driven)	$0.01^{+0.03+0.05}_{-0.01-0.01}$	$0.3^{+0.9}_{-0.3} \pm 0.3$	$< 0.01^{+0.03}_{-0.01}$	$0.3^{+0.9+0.4}_{-0.3-0.3}$
Expected ZZ	$1.57 \pm 0.03 \pm 0.11$	$3.09 \pm 0.04 \pm 0.06$	$4.5 \pm 0.1 \pm 0.2$	$9.1 \pm 0.1 \pm 0.3$

Table 2: Summary of observed events, total background contributions and expected signal in the individual four-lepton and combined channels. The first error is statistical while the second is systematic. The uncertainty on the luminosity is not included. The errors on the background estimates span the 68% confidence interval, which is not symmetric about the best estimate because the background cannot be negative.

Mass distribution

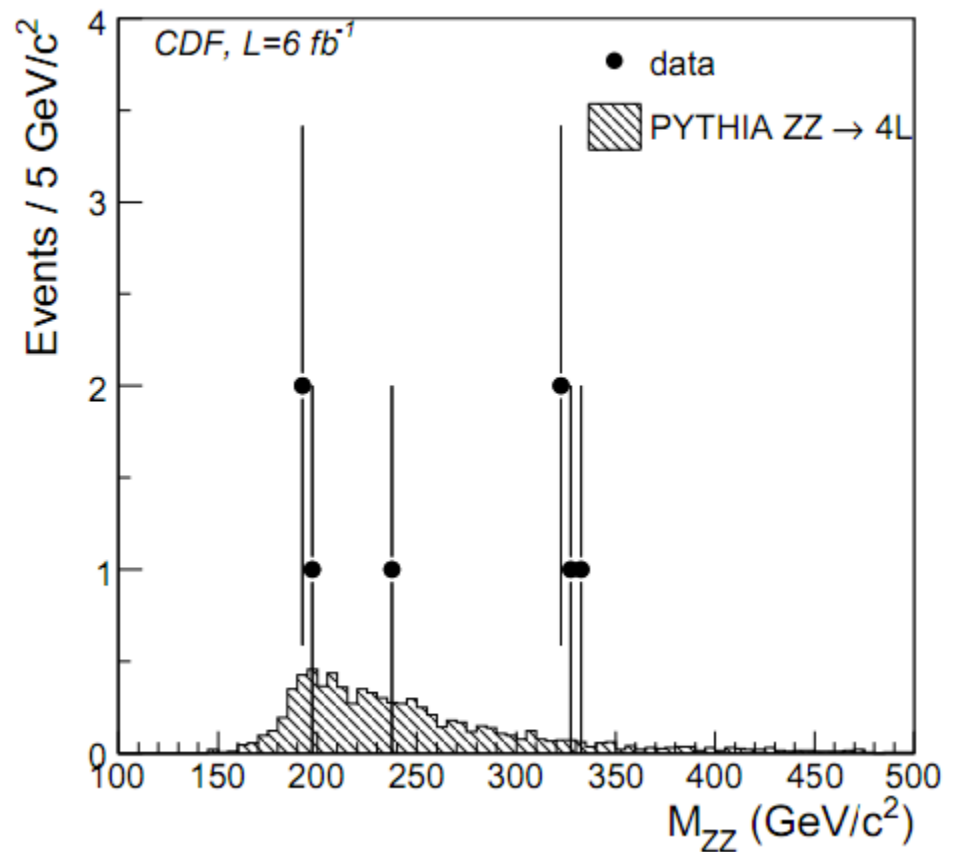
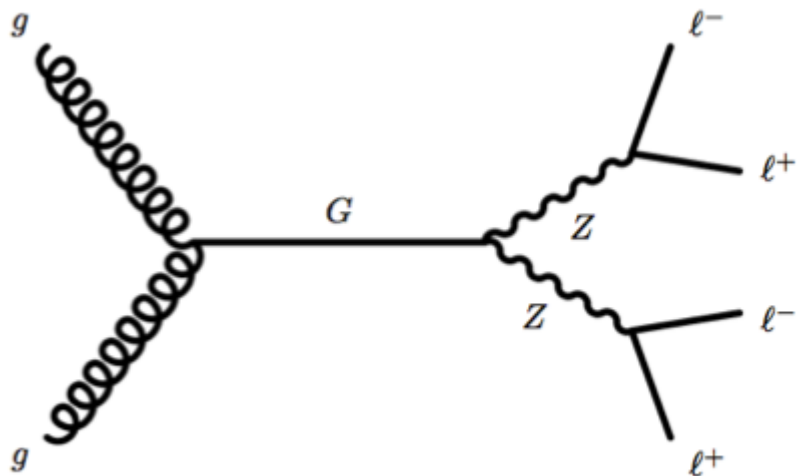


$$\sigma_{ZZ}^{\text{tot}} = 8.5_{-2.3}^{+2.7} \text{ (stat.)} \quad ^{+0.4}_{-0.3} \text{ (syst.)} \pm 0.3 \text{ (lumi.) pb.}$$

$$\sigma^{\text{NLO}} = 6.5_{-0.2}^{+0.3}$$

PRL 2011

ZZ resonances?

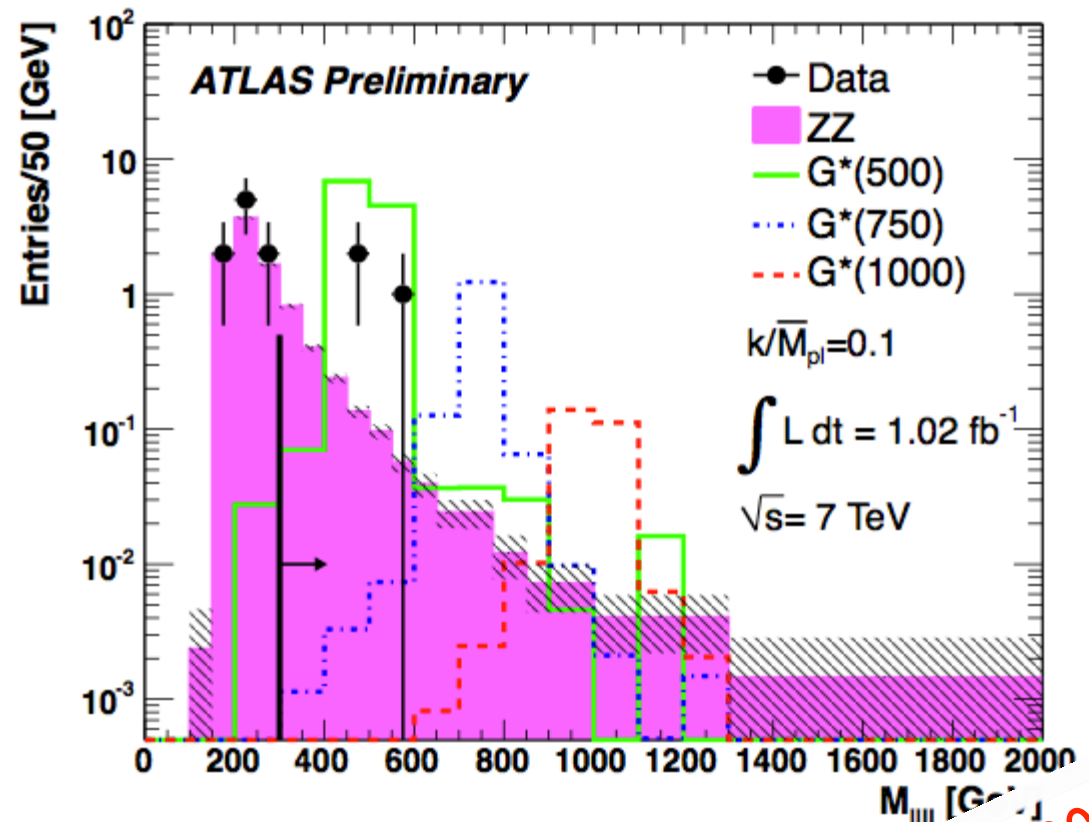


Exp: 5.5

Obs: 8

ATLAS ZZ resonances

Process	Total
ZZ	$1.85 \pm 0.11 \pm 0.09$
Fakes	$0.02^{+1.03}_{-0.01} \pm 0.75$
Total Bkg.	$1.87^{+1.04}_{-0.11} \pm 0.75$
Data	3
$G(350 \text{ GeV})$	$71 \pm 3 \pm 4$
$G(500 \text{ GeV})$	$12 \pm 0.5 \pm 0.6$
$G(750 \text{ GeV})$	$1.5 \pm 0.08 \pm 0.07$
$G(1000 \text{ GeV})$	$(2.7 \pm 0.2 \pm 0.1) \times 10^{-1}$
$G(1250 \text{ GeV})$	$(6.6 \pm 0.4 \pm 0.3) \times 10^{-2}$
$G(1500 \text{ GeV})$	$(1.9 \pm 0.1 \pm 0.1) \times 10^{-2}$



CONF 2011
Paper in review

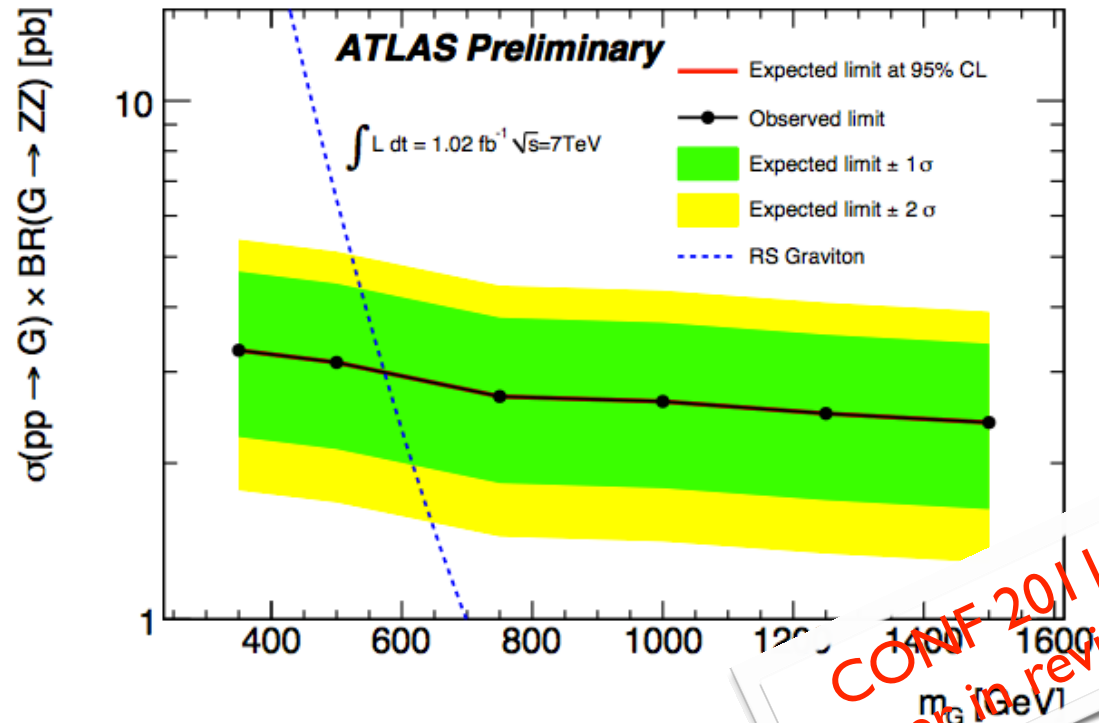
fiducial limits

$$\sigma_{ZZ \text{ fid}} < \frac{N_{ZZ}}{\epsilon_{ZZ} \times BF(ZZ \rightarrow \ell\ell\ell\ell) \times \mathcal{L}} = \frac{5.7}{0.61 \times 0.010 \times 1.02 \text{ fb}^{-1}} = 0.92 \text{ pb}$$

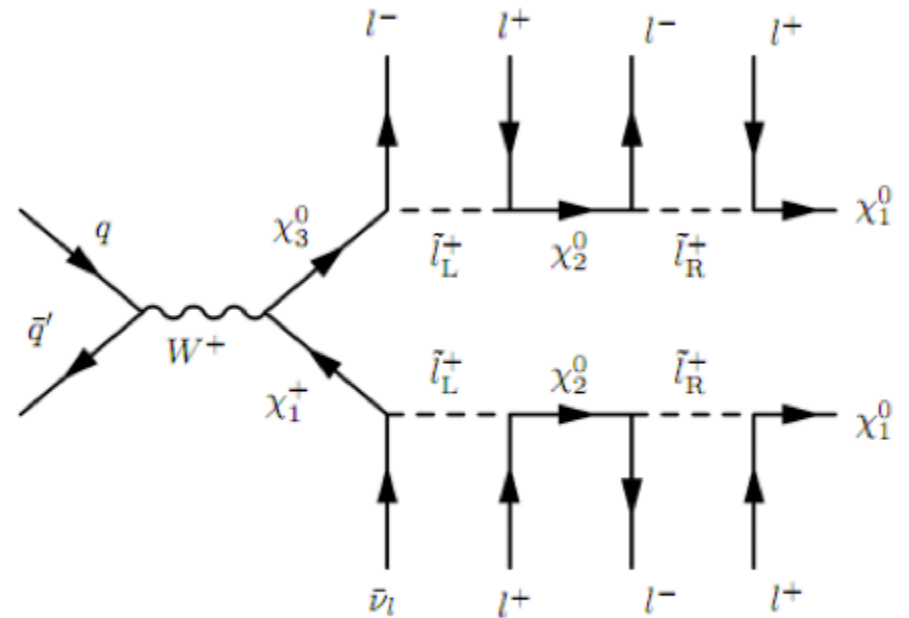
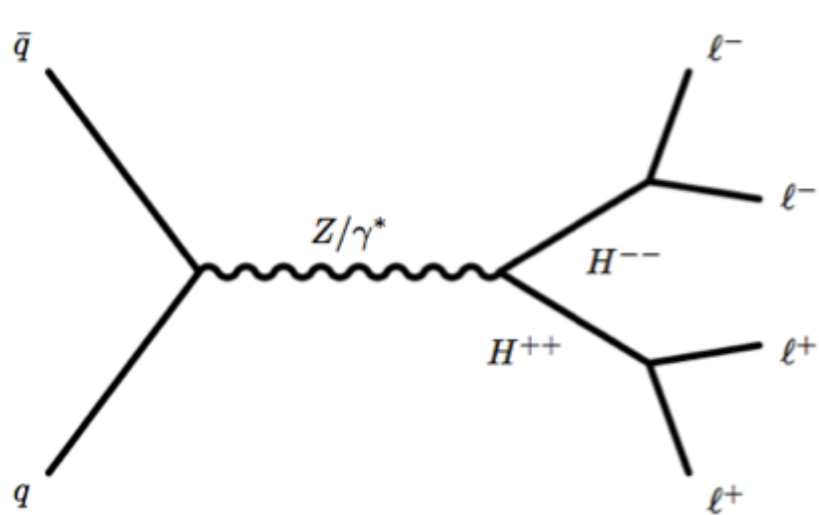
fiducial limits

$$\sigma_{ZZ \text{ fid}} < \frac{N_{ZZ}}{\epsilon_{ZZ} \times BF(ZZ \rightarrow \ell\ell\ell\ell) \times \mathcal{L}} = \frac{5.7}{0.61 \times 0.010 \times 1.02 \text{ fb}^{-1}} = 0.92 \text{ pb}$$

Graviton Mass [GeV]	Theory [pb]	Fiducial Acceptance	Selection Efficiency
350	41.70	27%	61%
500	6.45	28%	63%
750	0.69	31%	66%
1000	0.13	32%	66%
1250	0.03	33%	67%
1500	0.01	35%	66%



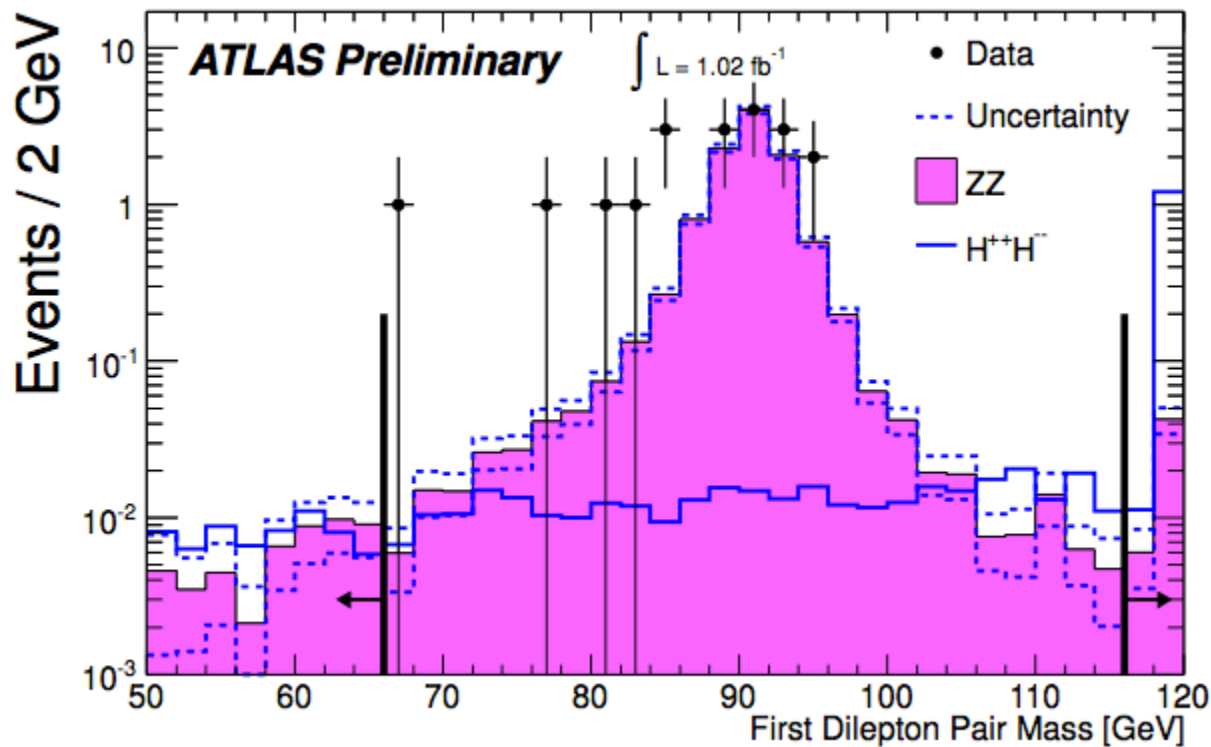
Anomalous 4l



Many Leptons at the LHC from the NMSSM

Vernon Barger,¹ Gabe Shaughnessy,^{2,3} and Brian Yencho¹

Anomalous 4l



Process	Total
ZZ	$0.10 \pm 0.01 \pm 0.01$
Fakes	$0.61^{+1.25}_{-0.61} \pm 0.91$
Total Bkg.	$0.71^{+1.25}_{-0.61} \pm 0.91$
$H_{200}^{\pm\pm}$	$1.27 \pm 0.04 \pm 0.03$
Data	0

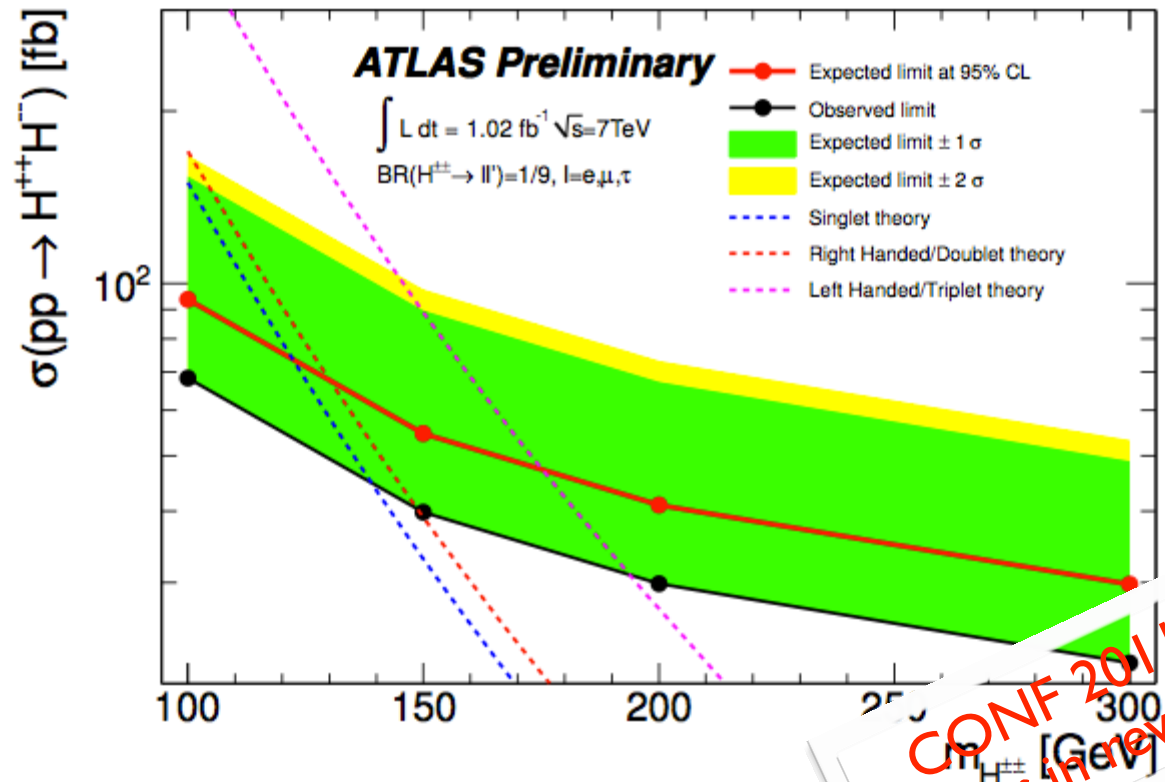
fiducial limits

$$\sigma_{4l \text{ fid}} < \frac{N_{4\ell}}{\epsilon_{\text{non-ZZ}} \times \mathcal{L}} = \frac{3.0}{0.62 \times 1.02 \text{ fb}^{-1}} = 4.7 \text{ fb}$$

fiducial limits

$$\sigma_{4l \text{ fid}} < \frac{N_{4l}}{\epsilon_{\text{non-ZZ}} \times \mathcal{L}} = \frac{3.0}{0.62 \times 1.02 \text{ fb}^{-1}} = 4.7 \text{ fb}$$

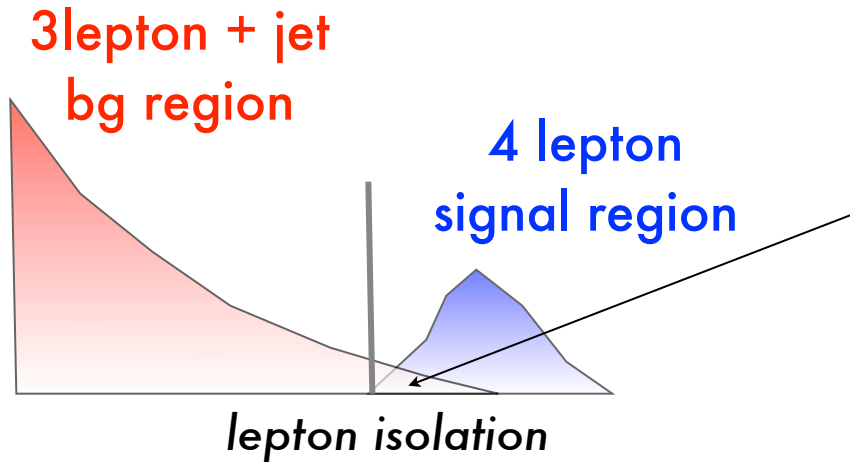
$H^{\pm\pm}$	Fiducial	non-ZZ
Mass [GeV]	Acceptance	Selection efficiency
100	7%	62%
150	12%	66%
200	16%	67%
300	22%	67%



CONF 2011
 Paper in review

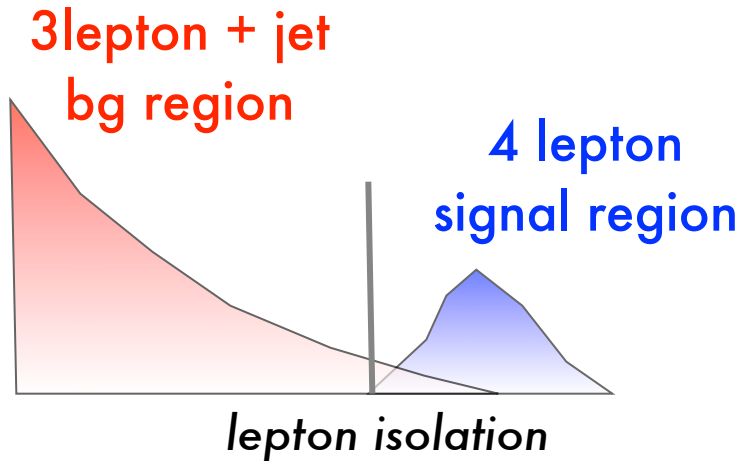
Backup

Background strategy

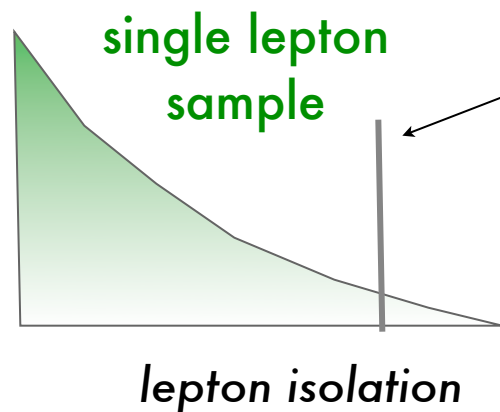


Need to know
the size of this tail

Background strategy



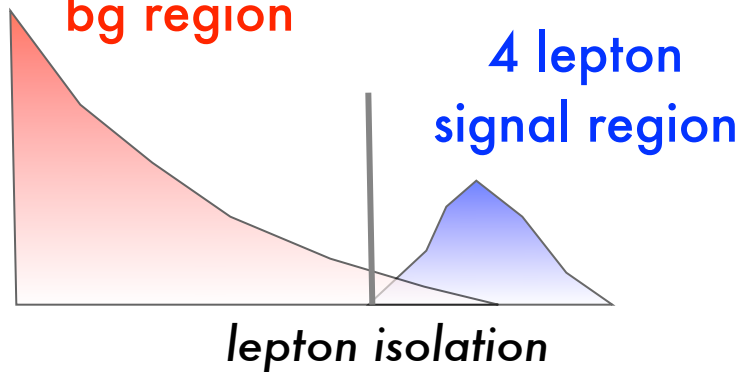
Find the same
shape in a different sample



Measure fraction that
pass threshold

Background strategy

3lepton + jet
bg region

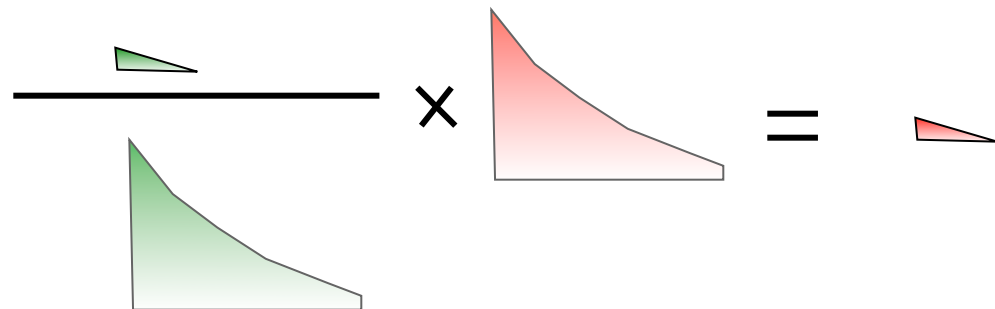
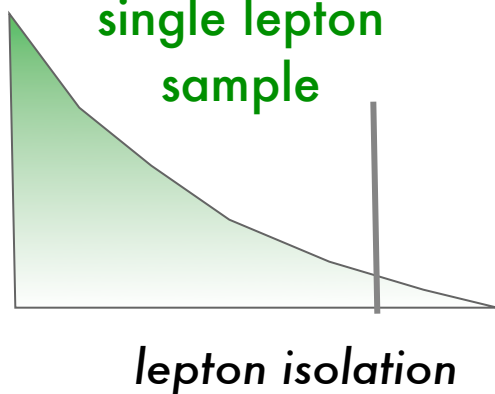


4 lepton
signal region

Use fake rate
to extrapolate



single lepton
sample



FR

$\times 3l+j = 4l \text{ bg}$

systematic uncertainties

Lepton efficiencies

electron efficiencies (7% in eeee)

muon efficiencies (2% in $\mu\mu\mu\mu$)

Fake background (up to 80%)

Light/heavy flavor composition

quark/gluon composition

Z momenta

