A few basic facts theorists might want to know about an LHC experiment

Using the Search for lepton jets from Hidden Valley models as an example

Pauline Gagnon
CERN / Indiana University

Purpose of this talk:

- Use this opportunity to give you a chance to ask all those little questions you might not feel comfortable asking in a conference when 95% of the attendees are experimentalists (and not >95% theorists as here!)
- Give you an idea on how ATLAS has looked for lepton jets from Hidden Valley models so far
- Get ideas from you on specific aspects you would like us to test
- In case you don't know, ATLAS offers short term associations for theorists to work on specific issues

Before I get started…..

Let me clarify a few points that might not always be clear

Why does it take so much time to analyze all the available data?

- Many analyses based on 2010 data only became available for the winter or even summer conferences
- Many are only preliminary results
  - Publications require a stringent and long review process in large collaborations
  - With more than 3000 physicists on ATLAS or CMS, not easy to get everybody happy
- Before completing any analysis, each sub-group working on data preparation has to provide their part:
  - Data quality group decides which data can be used
  - Performance groups give calibration, energy scale corrections, etc
  - Needed from each group: e/gamma, jet and Et miss, muon, b-tagging, trigger etc)
- Even with 3000 people, we are lacking people everywhere!

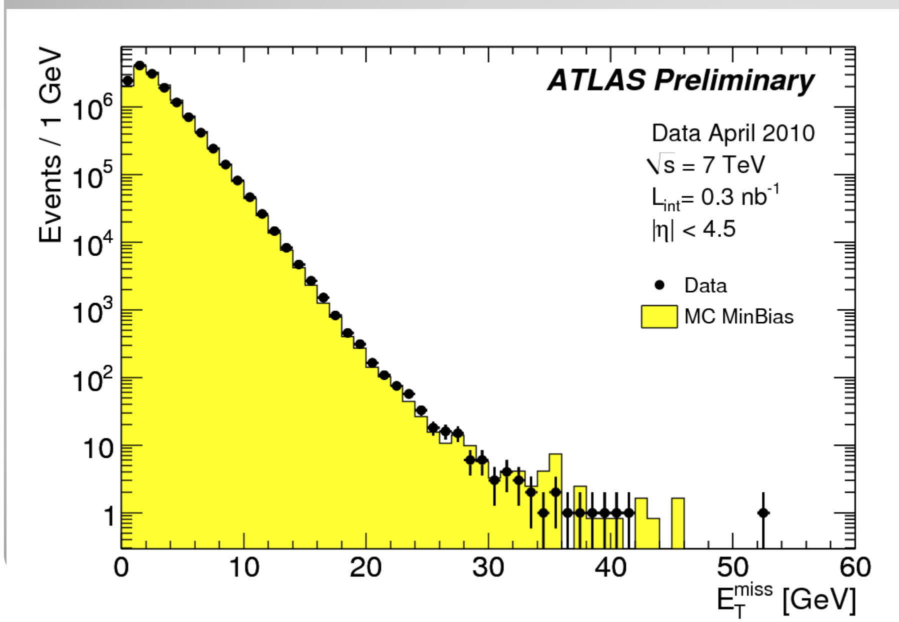
Data quality

- Data taking period is divided in small time interval of 1-3 minutes where the detector and LHC conditions are stable
- These are called "**luminosity blocks**" in ATLAS
- For each lumi block, the detector sub-groups check if their detector was fully or partially operational,
- This information is given to the performance groups. They determine which lumi block can be use to search for each separate physics object (electrons, muons, jets, missing E_T)
- Data quality group then issues "**good run lists**" which are used for each analysis
- This explains why we sometimes publish results on 35 pb^{-1}, 37 pb^{-1} or 40 pb^{-1} depending on which physics objects are needed for each analysis

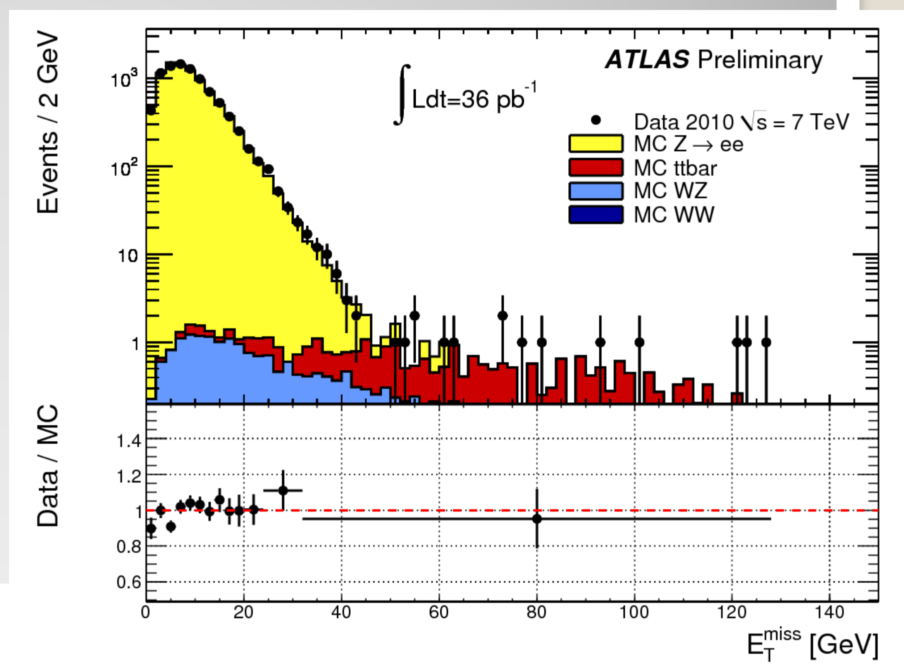
Other reasons for delays

- The more data we have, the less forgiving it gets
- More data means the error bars go down
- We need more precise and more specialized cross-checks
- We are looking at hundreds of distributions! But even missing ET looks good

Very general (minimum bias)



Very specific (Z->ee events)

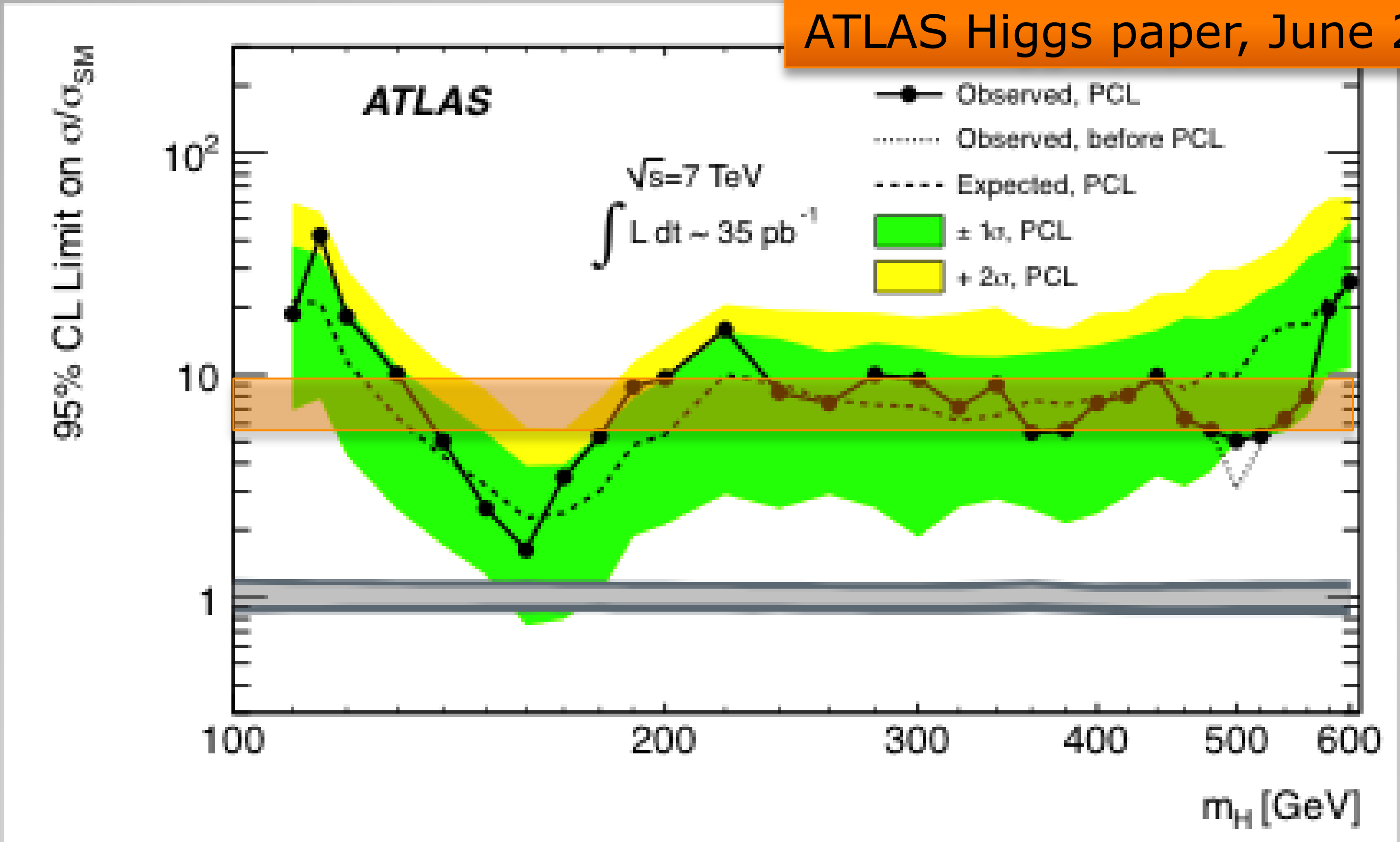


How much better will we do with the 2011 data?

- In 2010, CMS and ATLAS both had ~ 40 pb^{-1} of usable data
  - Most papers so far and preliminary results for Moriond
  - ATLAS just started showing some 2011 data at PLHC
- In 2011, we expect 2-3 fb^{-1} in total (1fb^{-1} already in)
- Assuming 2 fb^{-1} by the end of 2011 (conservative)
  - That'd be 50 times more data than in 2010
  - Signal will increase by 50 but so will the background
  - Uncertainty on the background would decrease by a factor of $\sqrt{50}$
  - The significance, $S/\sqrt{B}$, will increase by a factor of ~7
- With 3 fb^{-1} by the end of 2011 (slightly optimistic)
  - That'd be 75 times more data than in 2010, improve limits by ~8.5
- Combining ATLAS and CMS is like having twice as much data

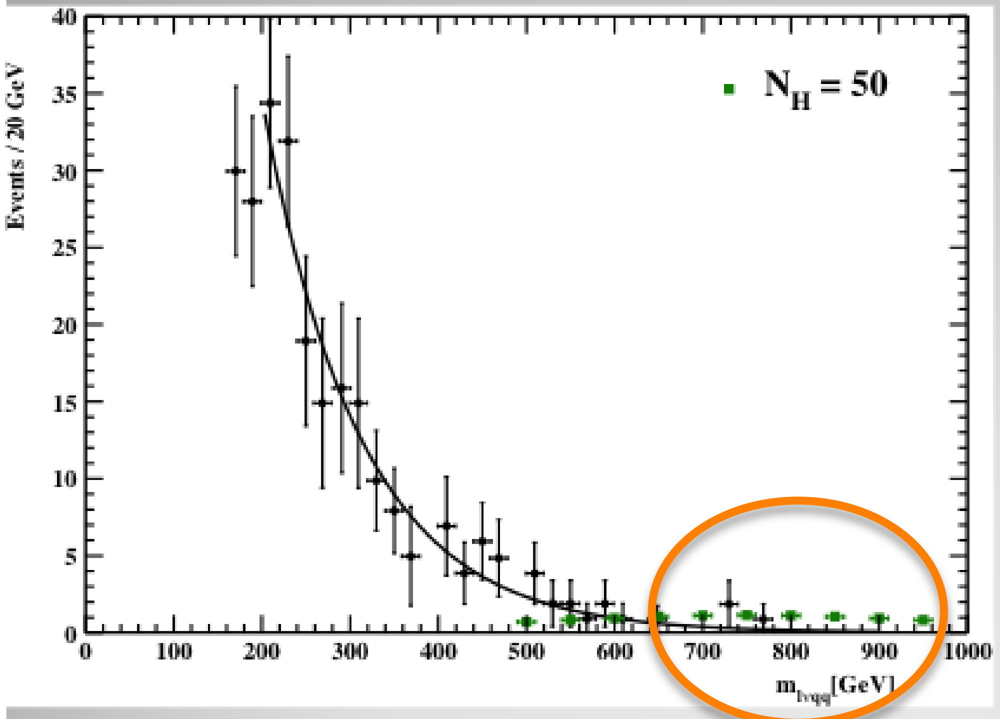
Impact on Higgs searches

ATLAS Higgs paper, June 2011



With 2-3 fb^{-1} per experiment, and combining CMS and ATLAS, we could exclude a SM Higgs in most of this whole range by end of 2011

Easy mistakes to be avoided (1)



"It is remarkable that the experimental data do show an excess of three events in this region."

The trivial Higgs boson: first evidences from LHC
arXiv:1106.4178

Their previous paper was rejected by the referee for not being solid enough to corroborate the great claims about the SM Higgs mass.

"We decided to leave to LHC the reply to the anonymous referee. Indeed, we feel that the time is coming to undertake a profound revision of the peer review process.

Easy mistakes to be avoided(2)

How often should we see a 4.8σ fluctuation in our career?

- **Exclusion**: no signal observed at a 2σ-level (95% CL)
- **Observation**: signal observed at a 3σ-level (99% CL)
- **Discovery**: signal observed at a 5σ-level (99.9995% CL)

What could go wrong?

system	# of channels	Operation efficiency
pixels	80 M	96.9%
Silicon tracker	6.3 M	99.1%
Transition radiation tracker	350 k	97.5%
EM calorimeter barrel	170 k	99.5%
EM calorimeter end-cap	3.5 k	99.8%
Hadronic calorimeter	9.8 k	97.9%
Hadronic end-cap	5.6 k	99.6%
Muon system	1071 k	97.0-99.8%

Monte Carlo simulations

- Every time the LHC changes its running conditions, we need to generate a whole new set of MC events (it takes 1-2 months)
  - For each background and each signal
- Already, we had to go for a new production when:
  - LHC changed the center-of-mass energy
    - Higgs cross-sections go down by a factor of 4 between E_{CM} = 14 TeV and 10 TeV
  - Bunch spacing: going from 75 ns to 50 ns
    - meant more luminosity but also
    - more pile-up (# of low energy events occurring at the same time)
- We are constantly improving our reconstruction algorithms
  - 2-3 times a year, we do a new reconstruction software release when major bugs are found and fixed, or new features have been added
  - The entire data set and MC samples are reprocessed after a new release
  - We try to keep the same release for several months to avoid driving people insane

Estimating background and efficiency using data

- Much more reliable than using Monte Carlo simulations

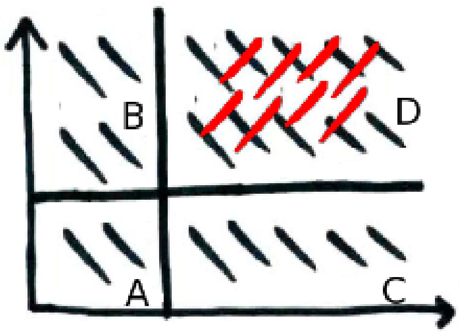
Tag & probe method:

- Using for example Z-> e^+e^- or Z -> $\mu^+\mu^-$
- Select a very clean sample of di-leptons under the Z peak
- Impose selection criteria to only one lepton (**tag**)
- Get the efficiency from the second lepton (**probe**):
  ➔ completely unbiased

- We often use events from J/ψ, Y or Z for various calibrations and cross-checks

Other tricks we like to play

ABCD-method



background

signal

x-axis: missing transverse energy
y-axis: reconstructed *transverse* W-mass
A B C D: regions

n_A n_B n_C n_D: event counts

Have 4 regions in two variables so that:

- background is independent for the variables:

$$\frac{n_A}{n_B} = \frac{n_C}{n_D}$$

- signal sitting in one region

Then it is possible to estimate number of background events in D via:

$$n_D^{est} = \frac{n_B \cdot n_C}{n_A}$$

This can be done completely without Monte-Carlo information
-> data briven background estimation

Trigger issues

- We are limited by how much data we can take in:

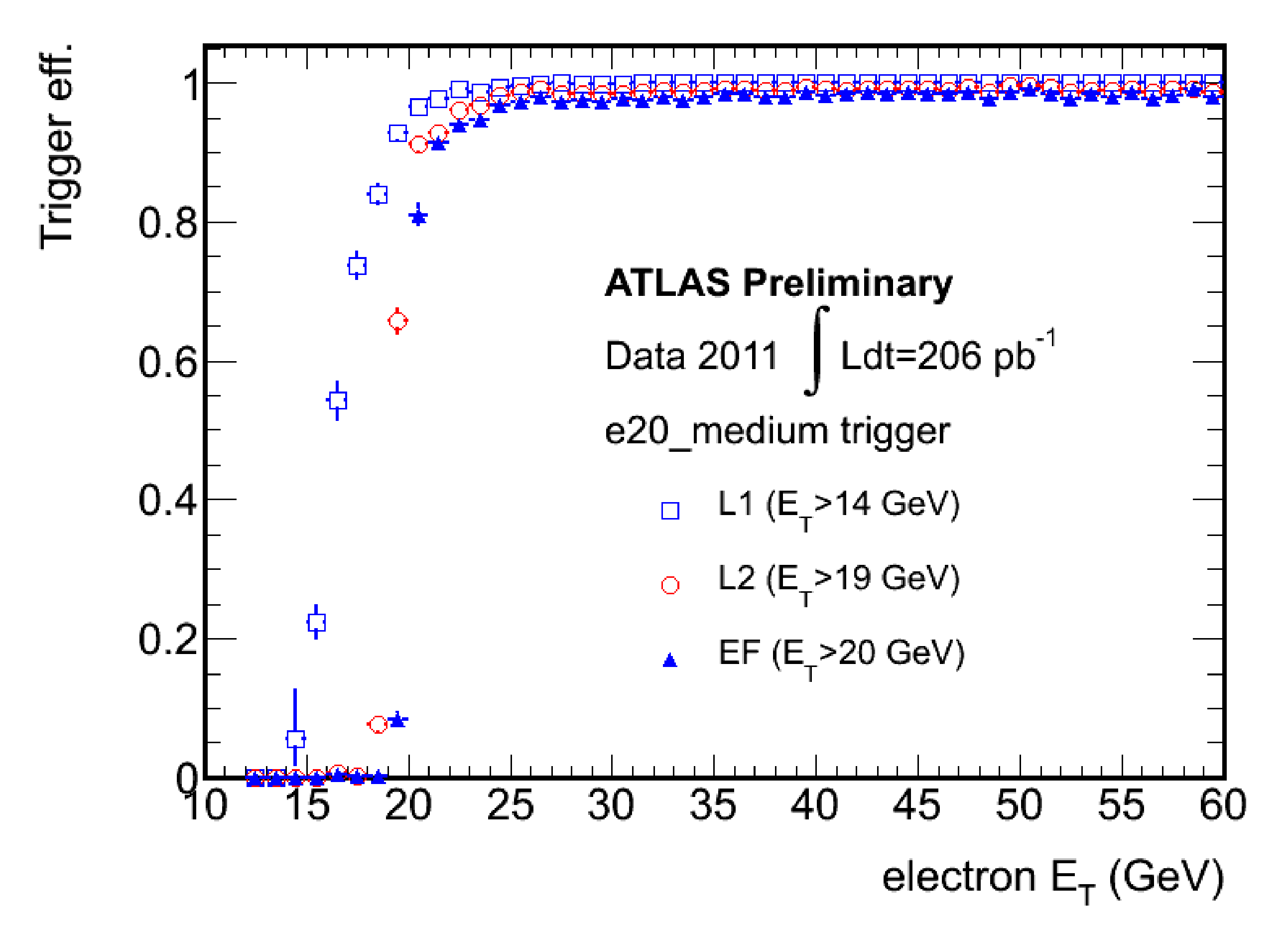
Level 1	**~75 KHz**	**Level 2**	**5 KHz**	**Level 3**	**400 Hz**

- As the luminosity goes up, we must tighten our selection criteria – increase the threshold or quality criteria
- Then we are forced to prescale our triggers:
  - Prescale factor of 100 means retain only 1 out of 100 such events

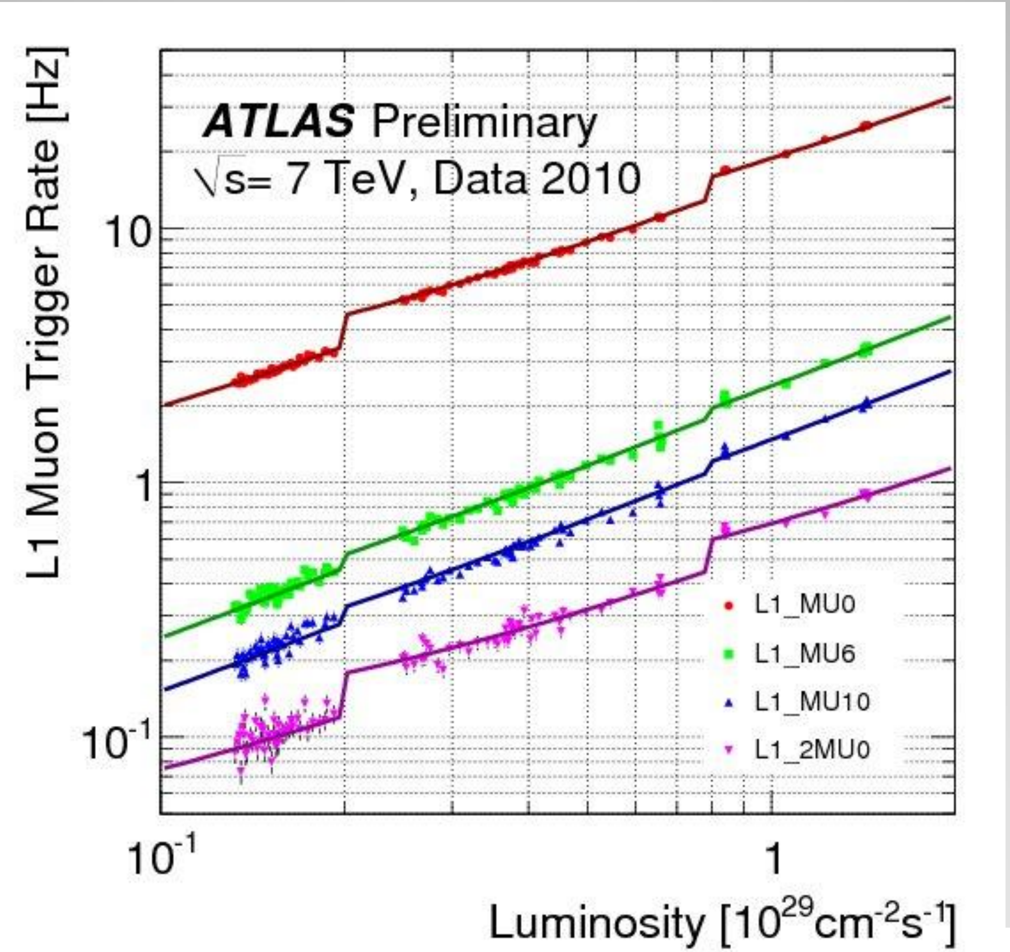
Lowest unprescaled triggers

Minimum p_T	**2010**	**now**
Single electrons	15 GeV	20 GeV
Di-electrons	5 GeV each	12 GeV each
muons	13 GeV	18 GeV
Di-muons	6 GeV each	10 GeV each
photons	40 GeV	60 GeV
Di-photon	15 GeV each	20 GeV each

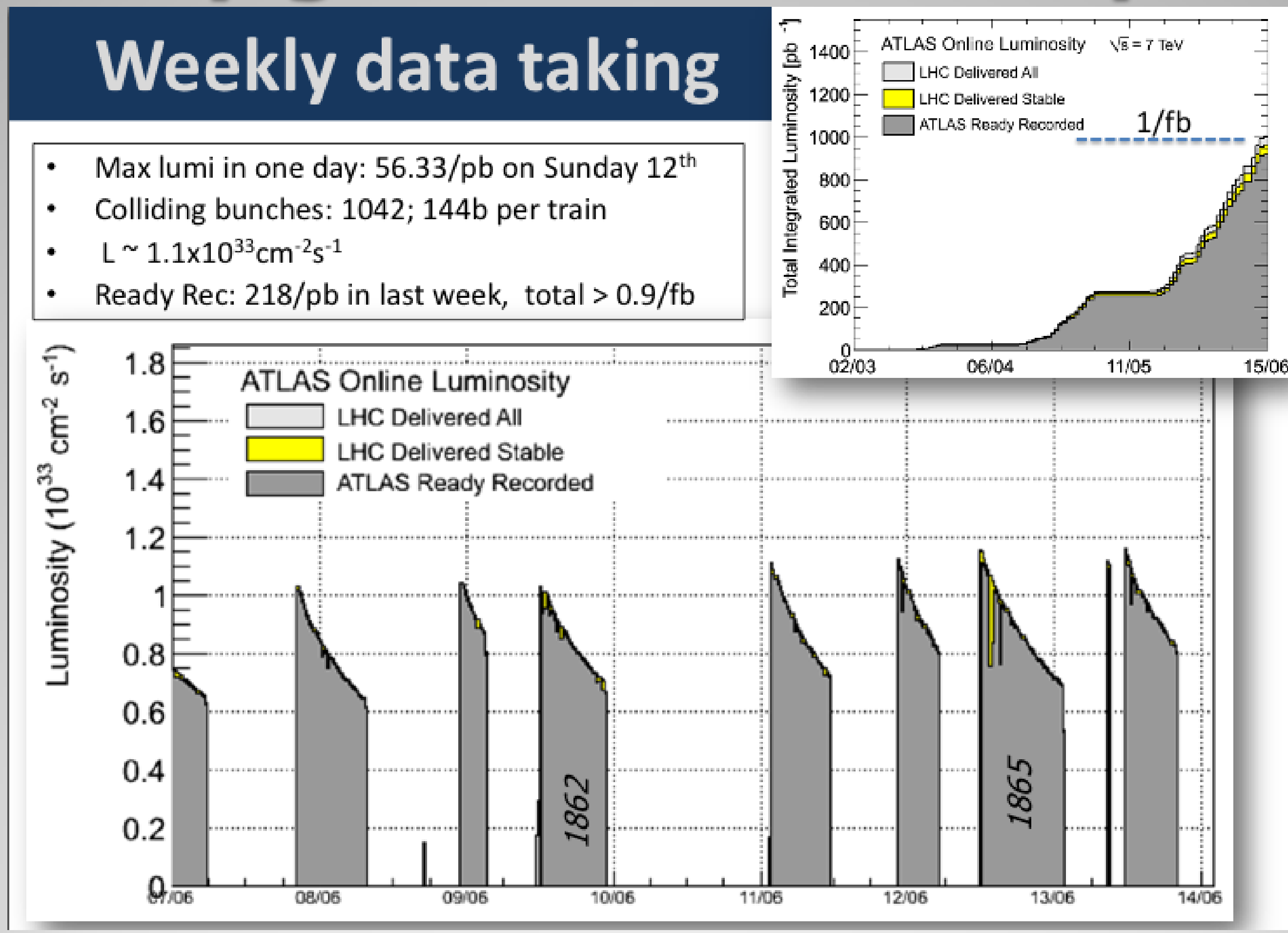
Trigger efficiency: should be stable measured using Z->ee (tag&probe)



Muon trigger rate as a function of luminosity



Trigger prescales are lowered as luminosity goes down to keep 400 Hz



from Claudia Gemme

Various LHC modes

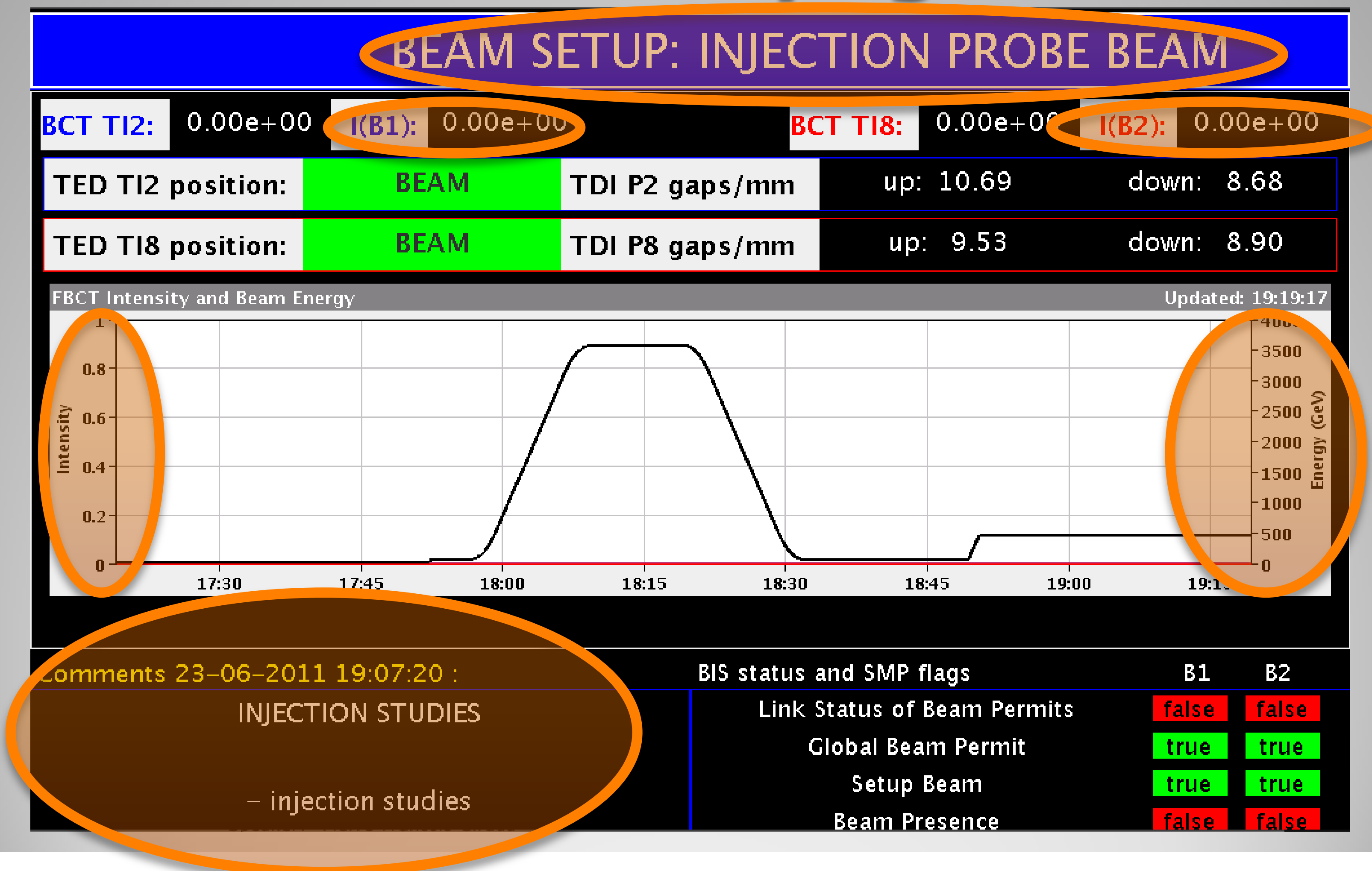
Main steps:

- Setup
- Injection probe beam
- Injection
- Ramping
- Adjust
- Stable beams
- Preparing to dump
- Beam dump

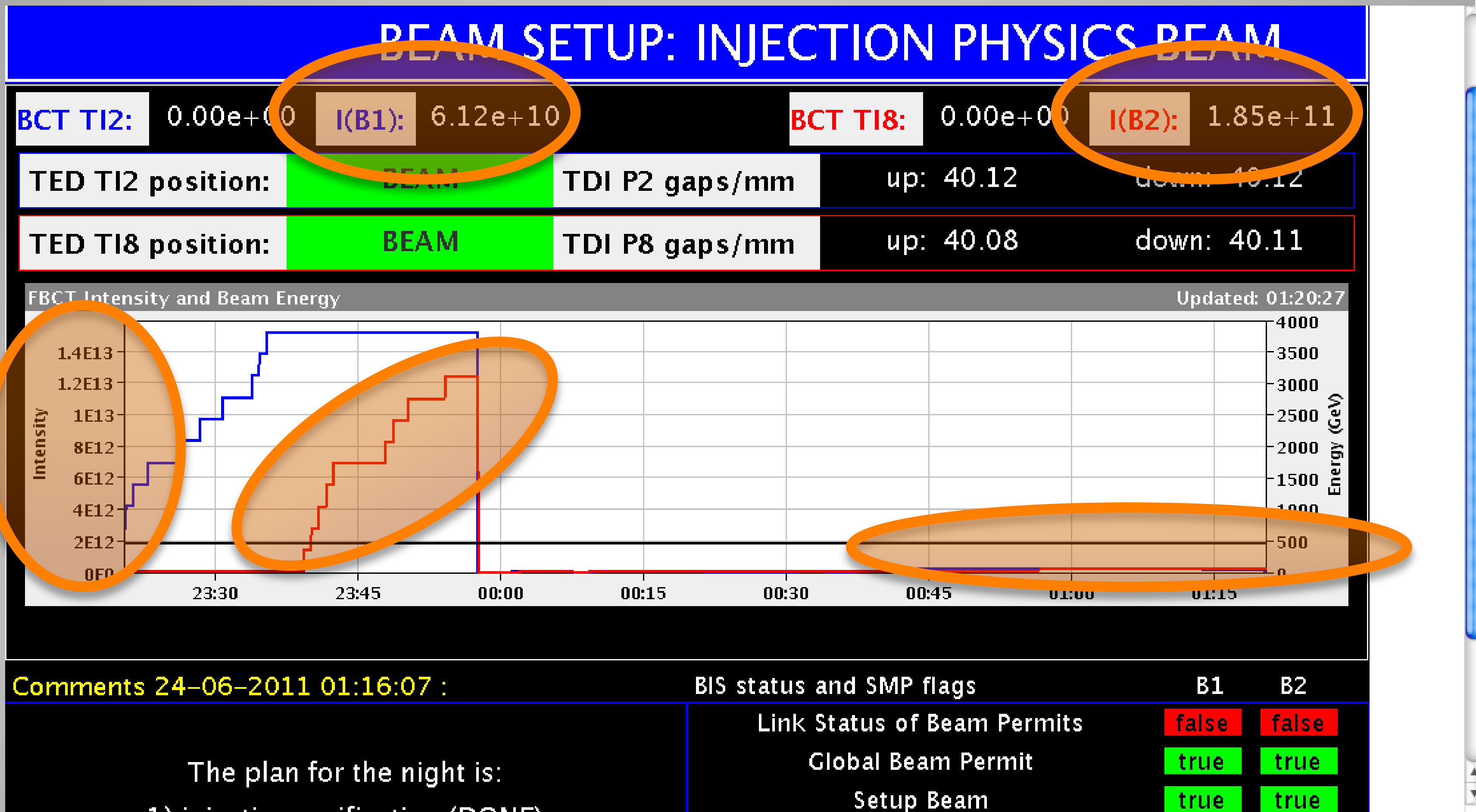
Additional steps occurring at any time

- Beam loss
- Cryo problems
- Recovering
- Access
- Power glitch
- Injection studies
- Beam studies

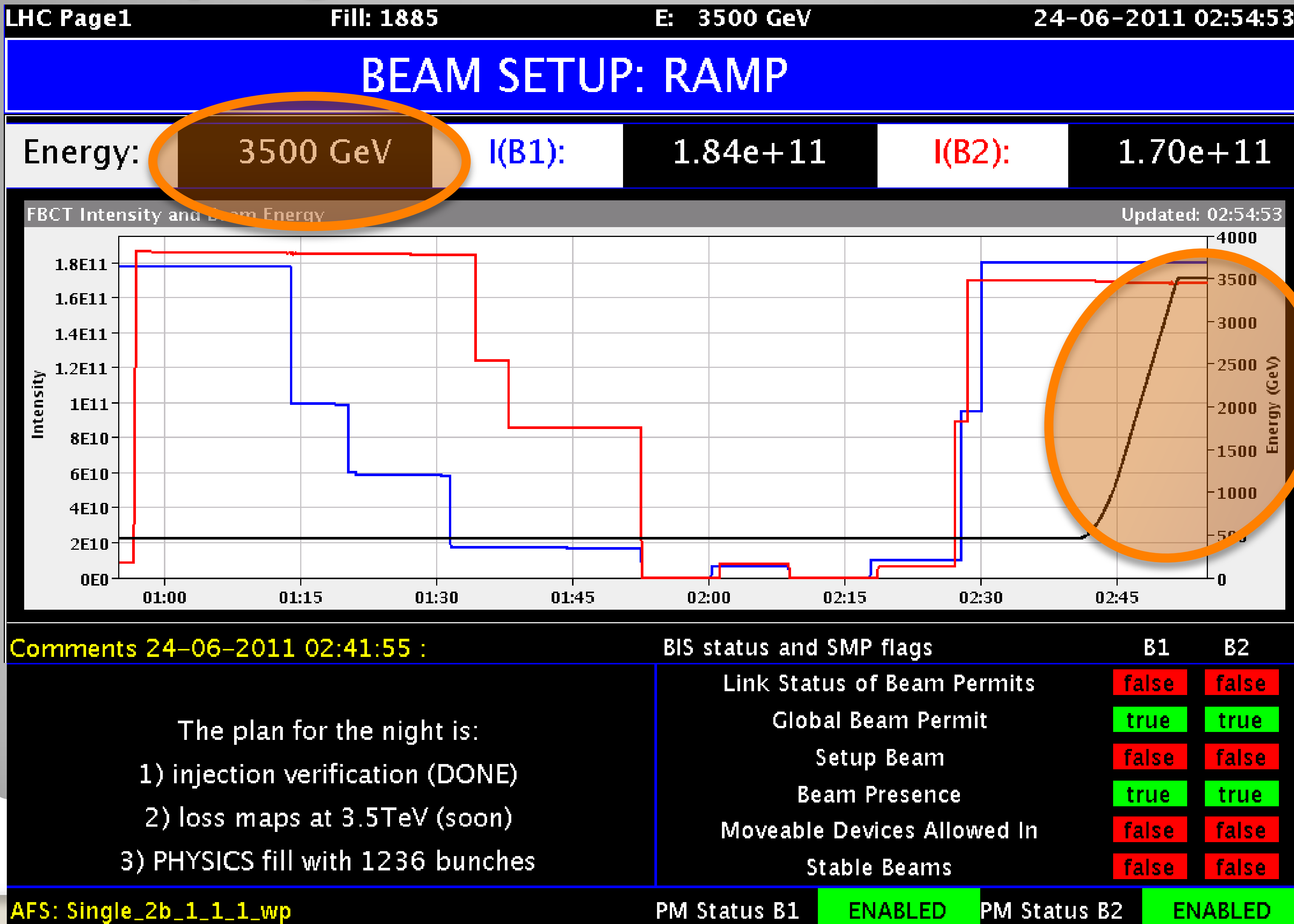
How to read LHC page 1



Injection


BEAM SETUP: INJECTION PHYSICS BEAM
BCT TI2: 0.00e+00
I(B1): 6.12e+10
BCT TI8: 0.00e+00
I(B2): 1.85e+11
TED TI2 position:
TDI P2 gaps/mm
up: 40.12
TED TI8 position:
BEAM
TDI P8 gaps/mm
up: 40.08
down: 40.11
FBCT Intensity and Beam Energy
Updated: 01:20:27
Intensity
Energy (GeV)
Comments 24-06-2011 01:16:07 :
The plan for the night is:
BIS status and SMP flags
B1
B2
Link Status of Beam Permits
false
false
Global Beam Permit
true
true
Setup Beam
true
true

Ramping from 450 GeV to 3.5 TeV



Human aspect

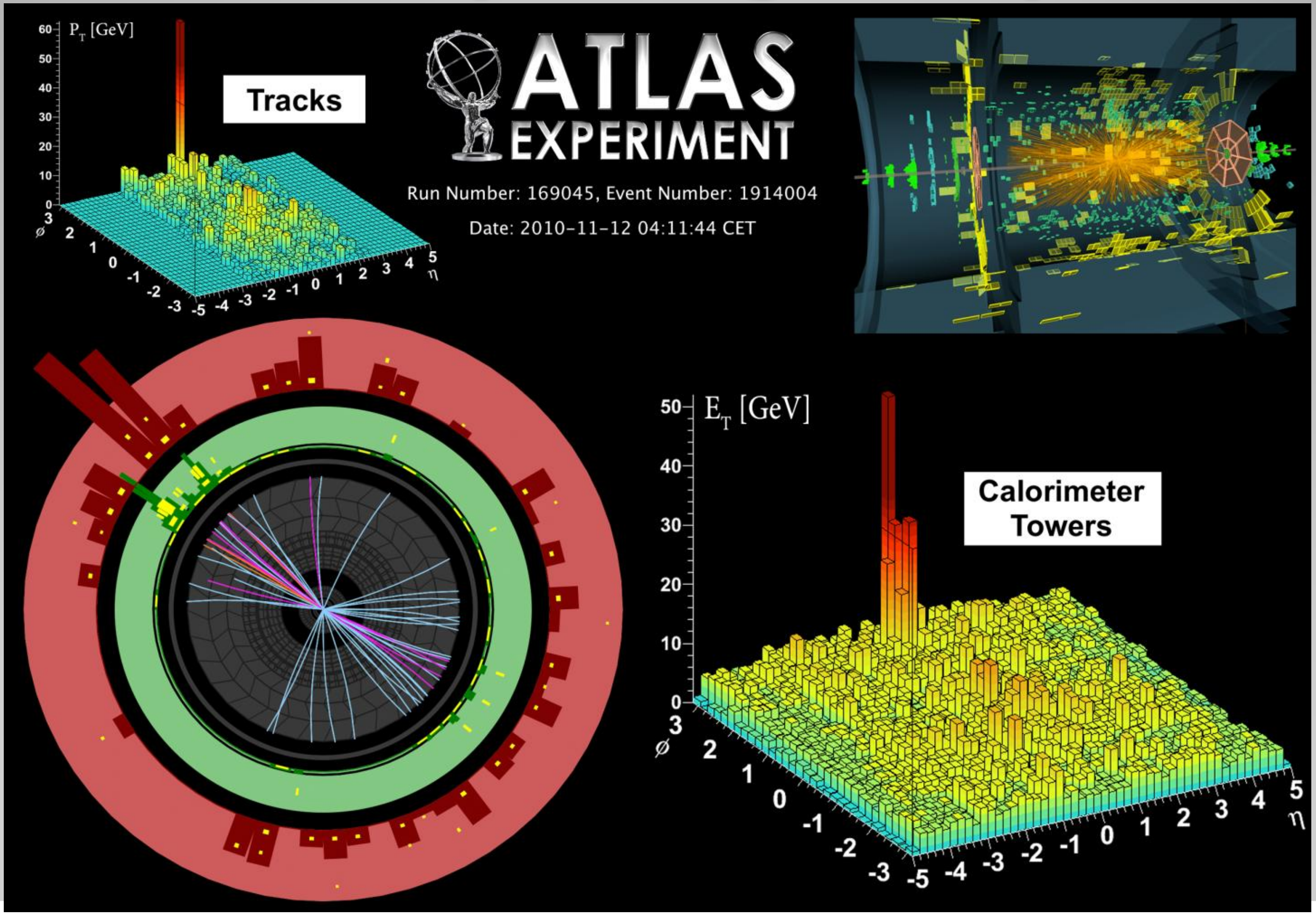
- There are >3000 physicists on ATLAS or CMS alone
  - 138 institutes
  - 37 participating countries
  - 70 nationalities
  - About 18% women
- Alice and LHCb have about 1000 physicists each
- It takes ~20 people at all times to staff the ATLAS control room
- Even more people are on call as system experts at all times

ATLAS control room on March 8, 2010 International Women's Day

2010 at the LHC in a nutshell

- In 2010 alone, LHC experiments produced:
  - 53 publications
  - 1700 conference talks
  - 179 student theses
- Initial focus on performance and calibration
- Confirmed the Standard Model solidity above 2 TeV
- Already put limits or excluded several models for new physics (excited quarks, axigluons, leptoquarks…)
- Limits on long-lived gluinos by CMS using quiescent time between fills
- Already in October, we measured ttbar cross-section
  - This brings into play all the analysis tools: lepton id, b-tagging, MET
- Observation of the quark-gluon plasma

Biggest surprise so far: asymmetric jets due to quark-gluon plasma

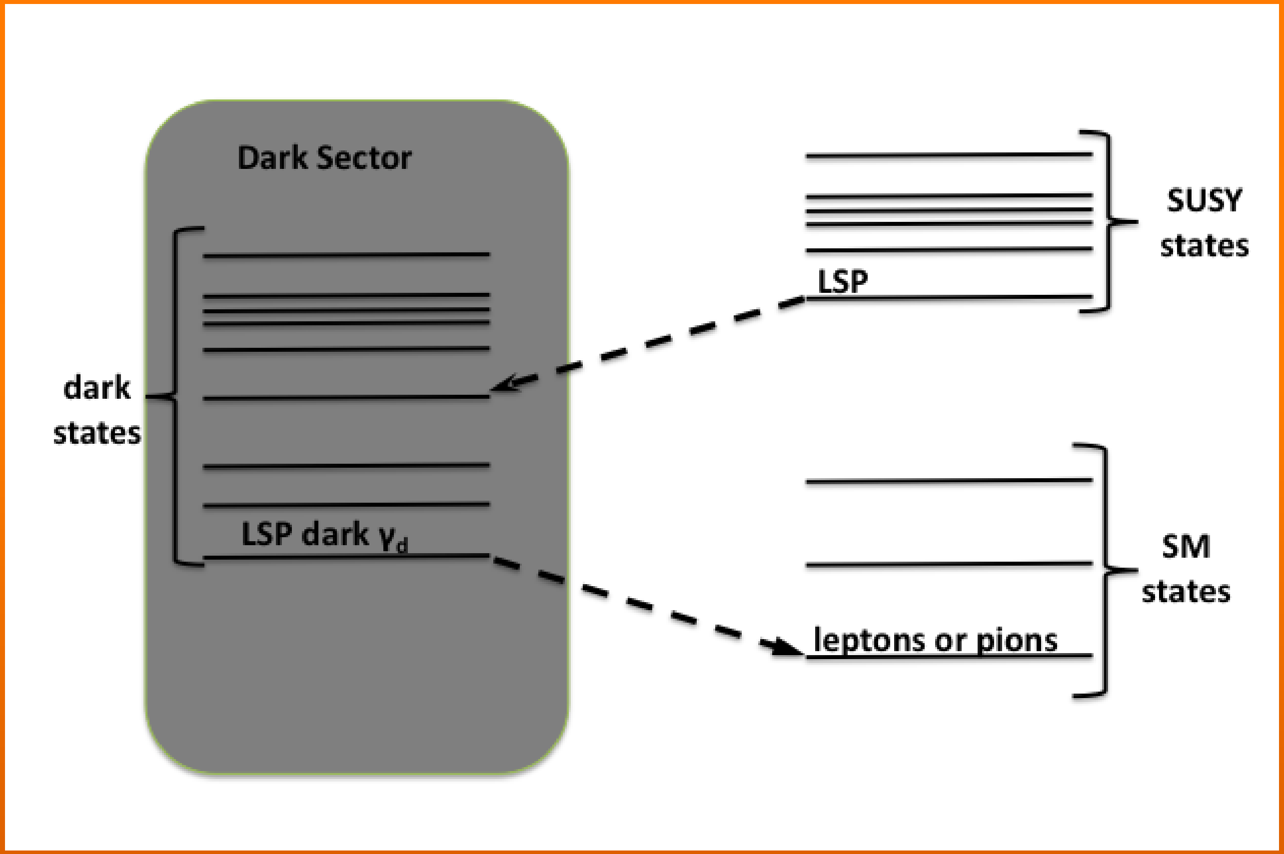


Switching gears

Searching for lepton jets from Hidden Valley models

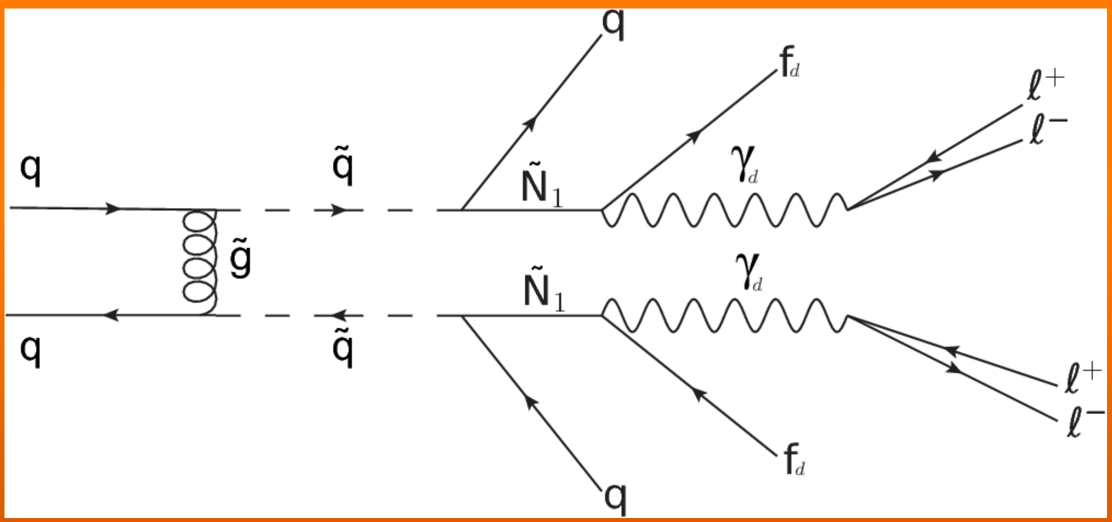
(I also work on invisible Higgs if someone is interested)

Several models predict lepton jets in the framework of Hidden Valley models when a light gauge boson decays into leptons



Based on N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, *"A Theory of Dark Matter", Phys.Rev. D 79 (2009) 015014.*

- ATLAS studied a particular class of such models using events generated by Itay Yavin (JHEP 1004 (2009) 116.)
- Weakly-interacting dark-sector with its own gauge bosons
- The LSP, here a neutralino (N_1) decays to a dark photon (γ_d) and a dark fermion (f_d)
- SPS1a SUSY parameters (Snowmass Points and Slopes from Snowmass 2001)

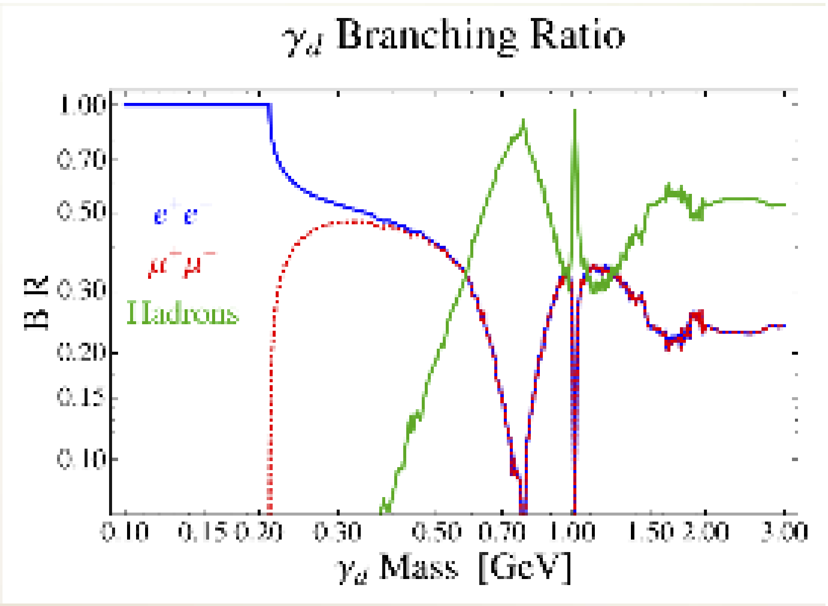


SPS1a	mSUGRA
m_0	100
$m_{1/2}$	250
A_0	-100
tan β	10

Model specificities

- Motivated in part by Pamela:
  - dark photons are assumed to be ~ 1 GeV
- Can decay only to e^+e^-, $\mu^+\mu^-$ or pion pairs depending on $m_{\gamma D}$
- Dark photon is light and comes from cascade decay of heavy squarks: leptons are boosted and collimated, hence the lepton jet name
- The radiation parameter, α_d, determines how many dark photons are produced

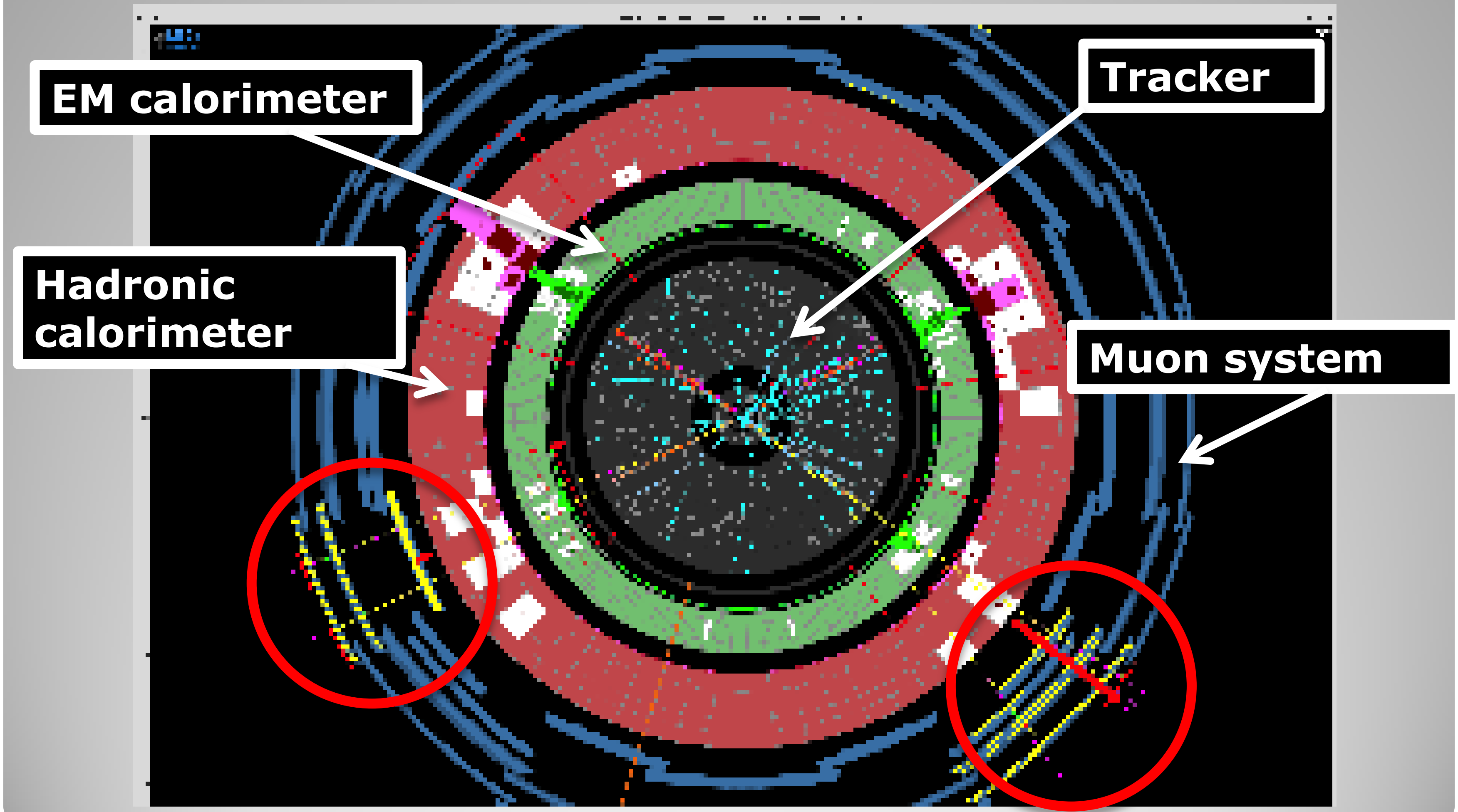
→ **Look for events with many collimated lepton pairs**



Current ATLAS analysis status

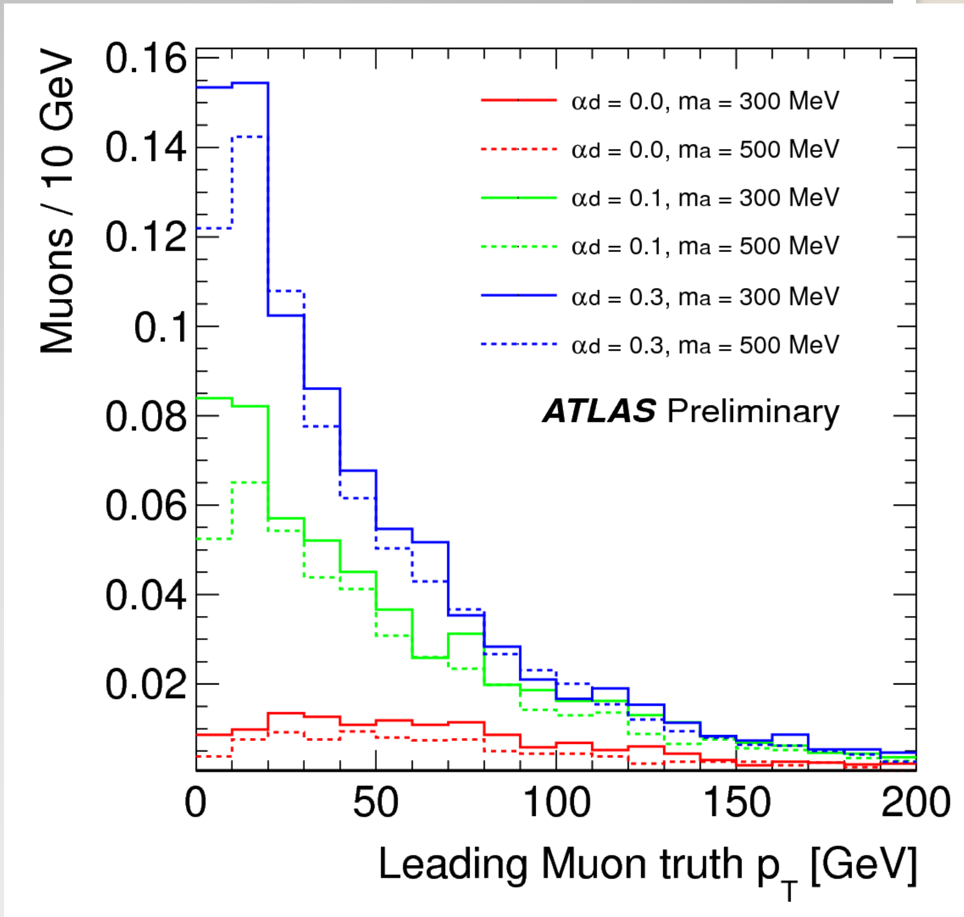
- Preliminary results obtained for the ~40 pb^{-1} 2010 data using only the muon channel
- Updates with ~1 fb^{-1} of 2011 data in progress
  - Lepton jets studies using electrons should be shown later this summer using 1 fb^{-1} hopefully by lepton-photon
- Electrons are more difficult to reconstruct in ATLAS:
  - Calorimeter closer to interaction point than muon spectrometer
  - Electromagnetic clusters tend to overlap, especially when coming from high p_T electrons ➔ broader EM clusters
  - Standard electron identification techniques fail

MC event with 2 muon Lepton Jets



Lepton jets Kinematics

- The heavier the dark photon, the higher the lepton p_T
- The more radiated lepton jets (α_d), the softest the leptons
- # of muons and their separation depends on m_γ and α_d
- Some events have electrons or pion jets, so less muons



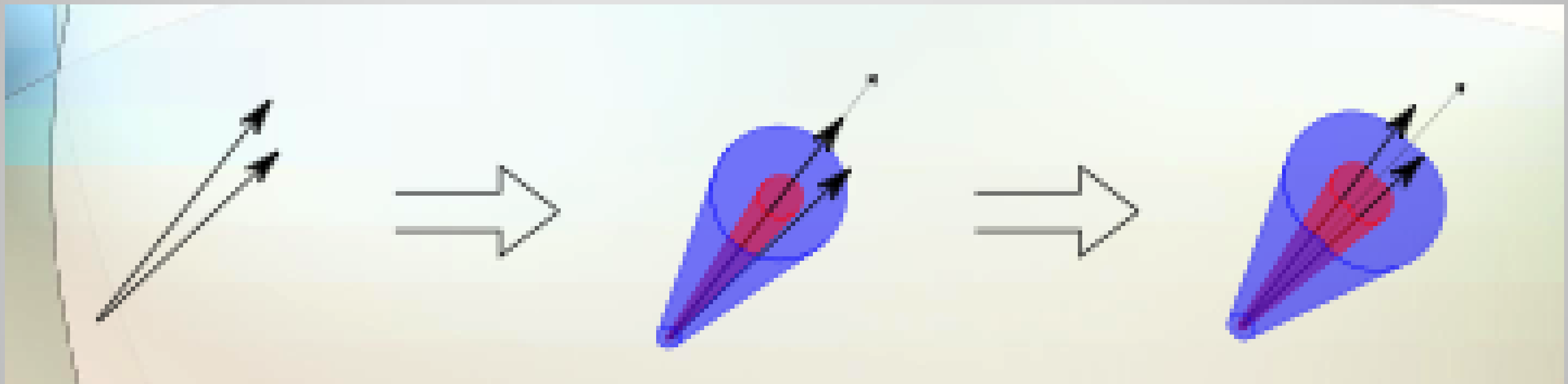
Event selection criteria

4 muons

- Events must pass the di-muon trigger with $p_T > 6$ GeV
- Request at least 4 muons with $p_T > 7$ GeV
- 3 muons passing higher track quality criteria to reduce background from fakes

2 lepton jets

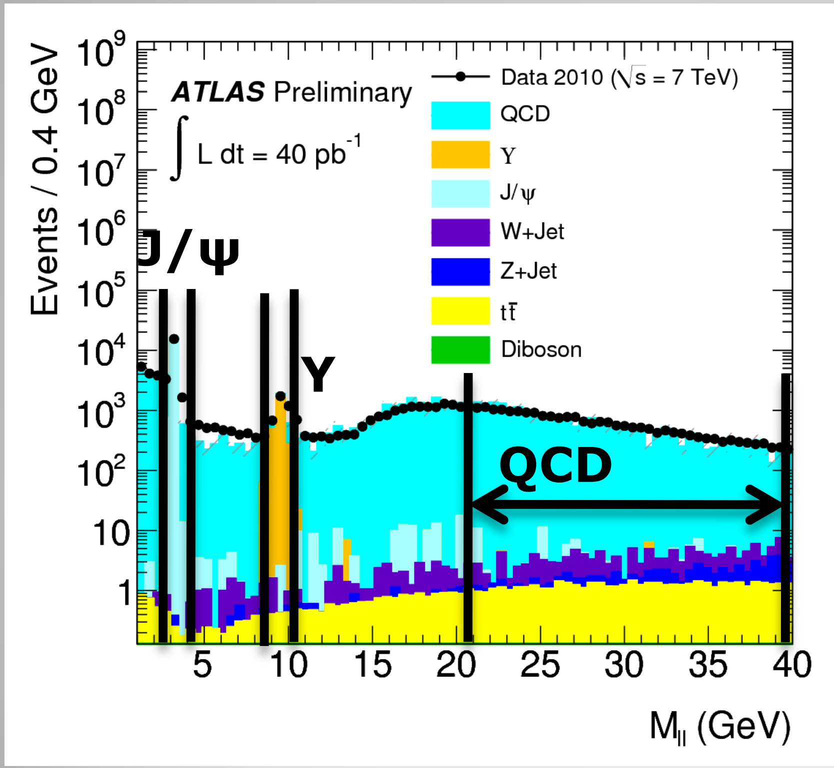
- Lepton jet is built from 2 muons found within $\Delta r < 0.1$ rad
- Lepton jet scaled isolation $E_T^{cone}/p_T < 0.7$
- Event must have at least 2 such lepton jets

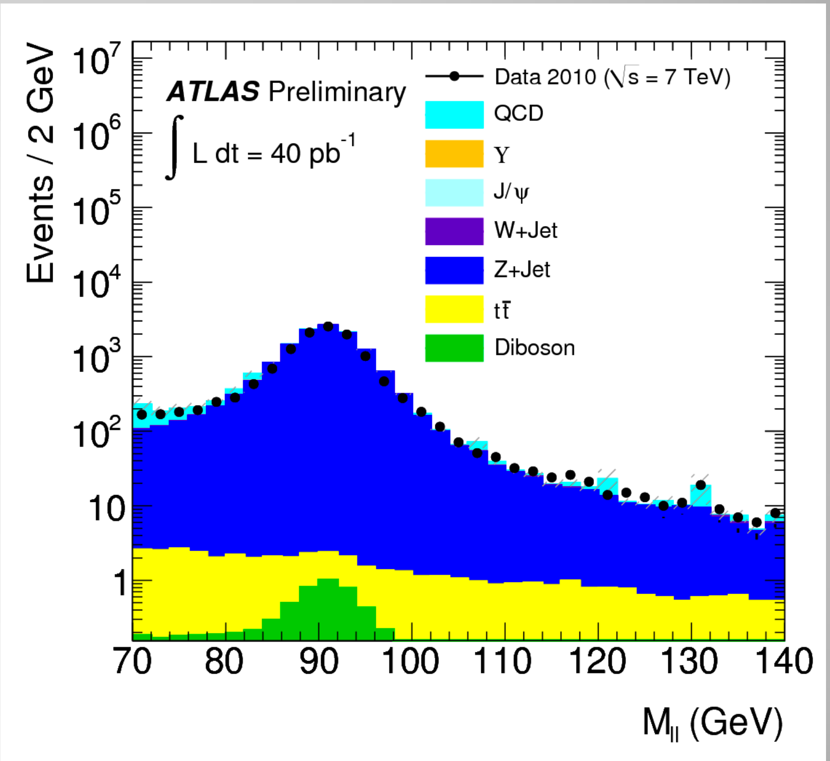


Emanuel Strauss, SLAC

QCD background normalization

- QCD production cross-section not precisely known
- Compare data and MC in 3 separate regions to extract scaling factors for J/ψ, Υ, and QCD from simultaneous fit





Cutflow:

main challenge is the QCD background

	≥ 2 muon	≥ 4 muons	≥ 4 muons w/ ≥ 3 HQ	2 LJets	2 Isolated LJets
data	174450	246	84	3	0
all bkg	200000 ± 15000	200 ± 50	81 ± 20	1.74 ± 0.48	0.20 ± 0.19
QCD	160000 ± 14000	188 ± 50	73 ± 20	1.46 ± 0.42	0.19 ± 0.19
Υ	2100 ± 120	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
J/Ψ	22100 ± 3700	3.4 ± 1.9	0.95 ± 0.43	0.24 ± 0.23	0.00 ± 0.00
W+Jet	332 ± 11	0.40 ± 0.40	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Z+Jet	14420 ± 42	2.00 ± 0.50	1.37 ± 0.41	0.00 ± 0.00	0.00 ± 0.00
$t\bar{t}$	357 ± 1.4	4.31 ± 0.16	3.47 ± 0.14	0.041 ± 0.016	0.012 ± 0.008
Diboson	16.577 ± 0.070	1.640 ± 0.013	1.557 ± 0.013	0.00033 ± 0.00019	0.00033 ± 0.00019
Squark Signal Samples					
$\alpha_d = 0.0, m_\alpha = 300$	8.26 ± 0.27	3.52 ± 0.18	2.38 ± 0.15	1.76 ± 0.12	1.38 ± 0.11
$\alpha_d = 0.0, m_\alpha = 500$	6.90 ± 0.25	2.62 ± 0.15	1.87 ± 0.13	1.35 ± 0.11	1.04 ± 0.10
$\alpha_d = 0.1, m_\alpha = 300$	15.16 ± 0.37	9.14 ± 0.28	7.58 ± 0.26	4.77 ± 0.21	2.90 ± 0.16
$\alpha_d = 0.1, m_\alpha = 500$	15.97 ± 0.38	8.38 ± 0.27	6.99 ± 0.25	4.08 ± 0.19	2.33 ± 0.14
$\alpha_d = 0.3, m_\alpha = 300$	9.60 ± 0.38	6.89 ± 0.32	5.99 ± 0.30	3.28 ± 0.22	1.25 ± 0.14
$\alpha_d = 0.3, m_\alpha = 500$	11.75 ± 0.32	7.88 ± 0.26	7.01 ± 0.25	3.29 ± 0.17	1.11 ± 0.10

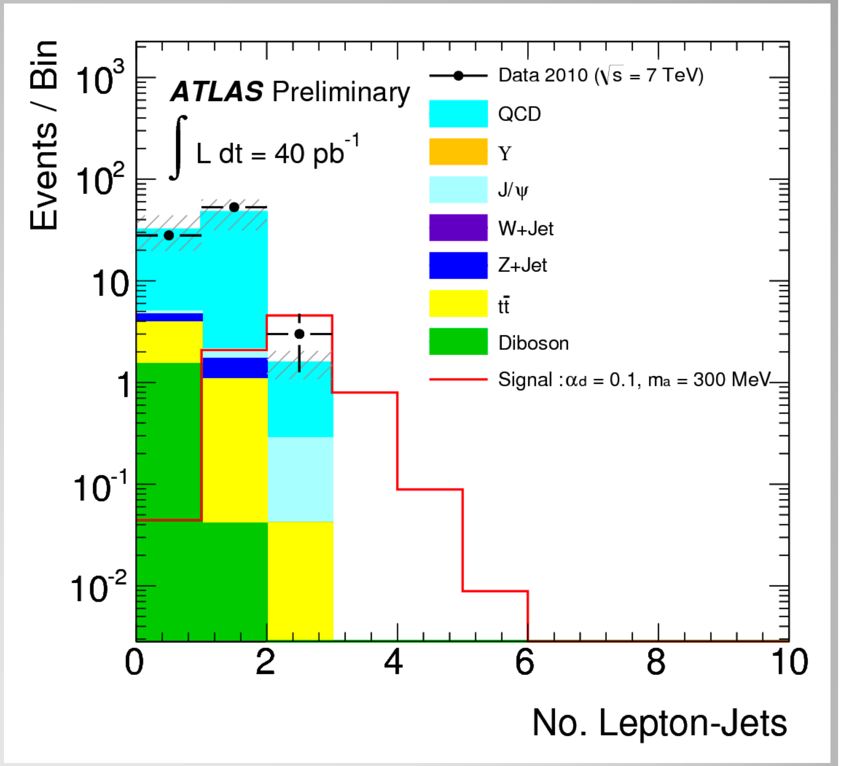
QCD background checks with data

<u>Main analysis</u>: at least 3 high quality muons

Here, relax this cut and instead assign an event weight p(m|n): probability that m high-quality muons are found when n muons are present

Predicts **0.19 ± 0.19** QCD background event

compared to **0.20 ± 0.19** from MC estimates



Second cross-check: ABCD method

- Define 4 separate regions:

A. Signal region:
   - 4 muons with pT > 7 GeV
   - Jet isolation

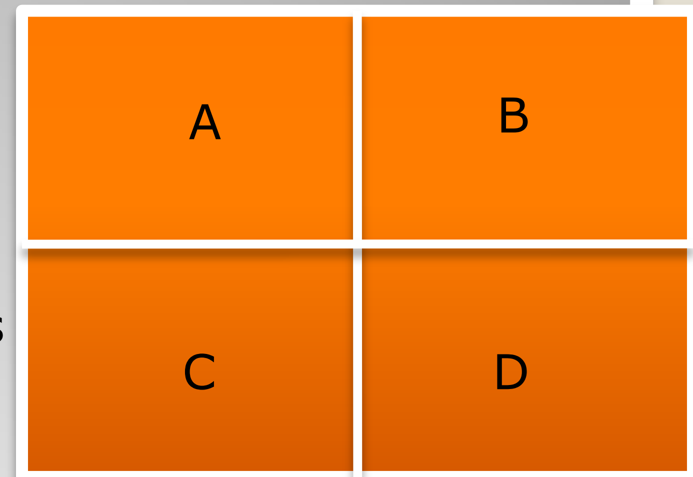
B. 4 > p_T > 7 GeV for 3rd and other muons

C. no isolation for one lepton jet

D. Reverse pT and isolation cuts

- Isolation and pT cuts are uncorrelated
- Assume A/C = B/D

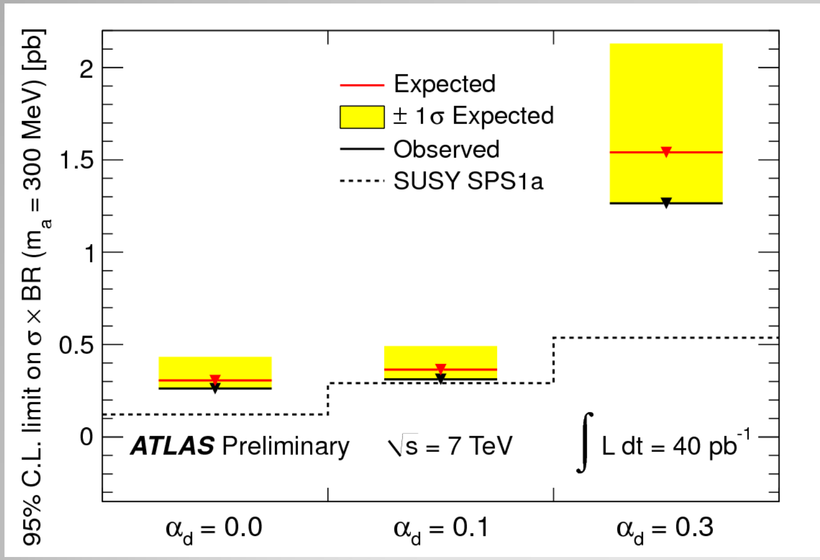
Predicts **0.11 ± 0.11** background events in region A

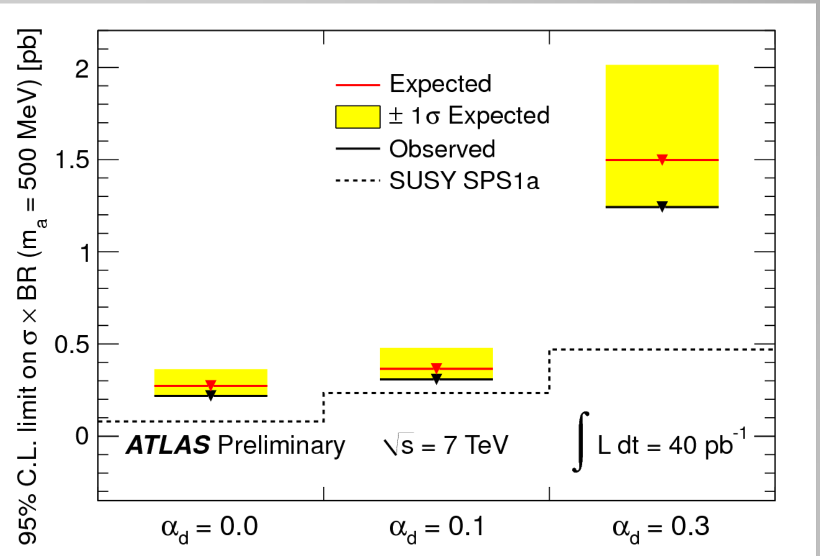


	# data events
B: anti-pT	1
C: anti-isolation	3
D: anti both	26
A: signal region	0

95% CL limits using CLs

- Given no events are found in ~40 pb-1 of 2010 data, we can set limits on production cross-section*BR





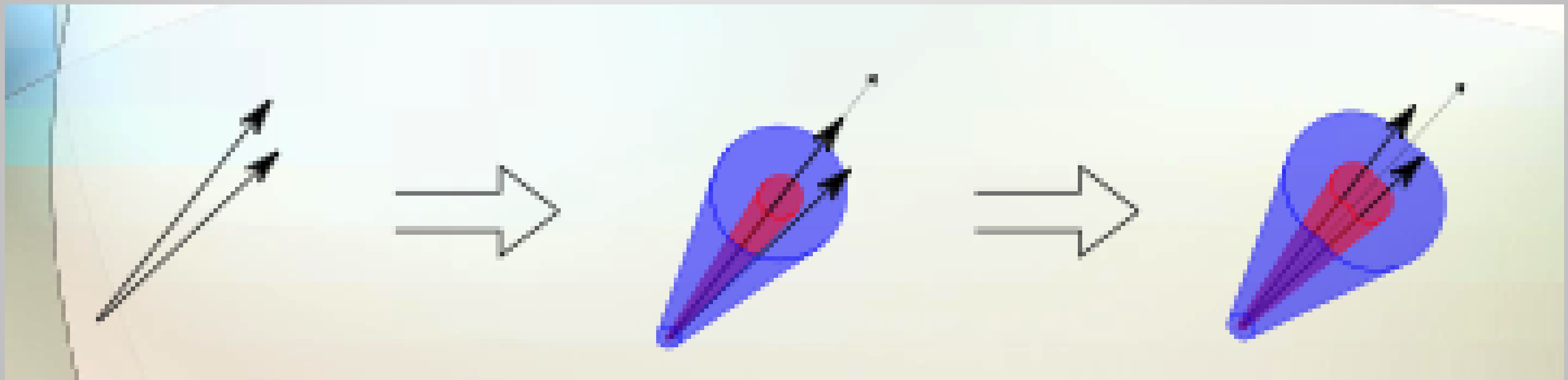
- CMS obtains model-independent limits using a search for a low mass resonance for cross-sections*BR*acceptance between 0.1 and 0.5 pb with ~35 pb^{-1} of data
- This model seems like it'll be ruled out soon

Stay tune! With the 2011 data, it should get really interesting!

Extra details

Event selection criteria

- Events must pass the di-muon trigger with $p_T > 6$ GeV
- Request at least 4 muons with $p_T > 7$ GeV
- 3 muons passing higher track quality criteria to reduce background from fakes
- Lepton jet is built from 2 muons found within $\Delta r < 0.1$ rad
- Lepton jet scaled isolation $E_T^{cone}/p_T > 0.7$
- Event must have at least 2 such lepton jets



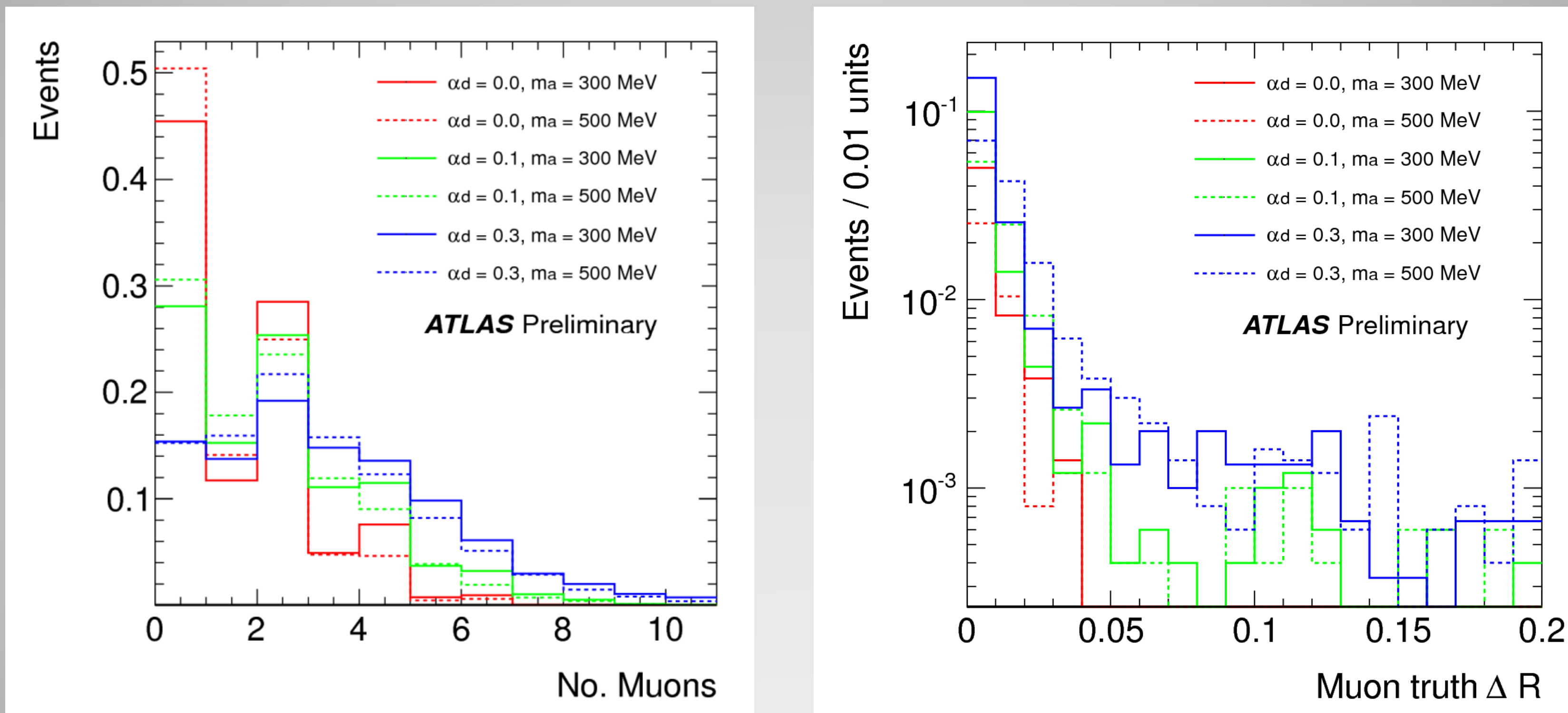
Emanuel Strauss, SLAC

Systematic uncertainties

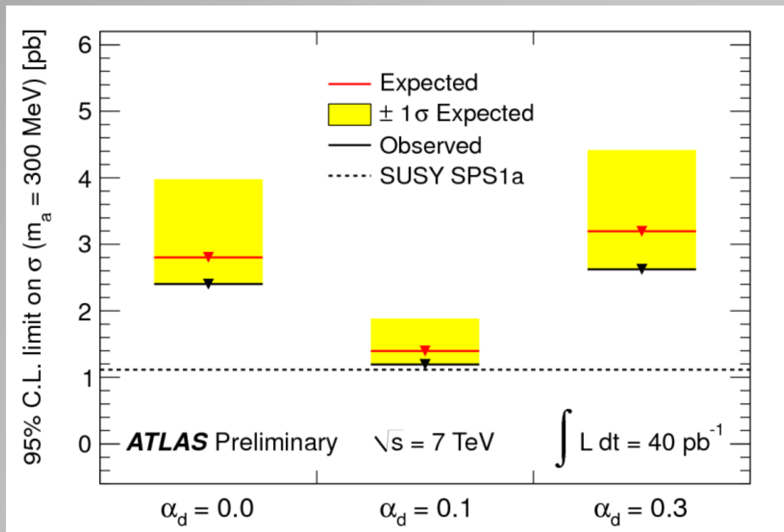
Systematic	Signal	QCD	J/Ψ	Υ	W+Jet	Z+Jet	$t\bar{t}$	Di-boson
Luminosity	3.4%				3.4%	3.4%	3.4%	3.4%
Trigger	1%				1%	1%	1%	1%
Reconstruction	2.9%				2.9%	2.9%	2.9%	2.9%
ΔR Efficiency	8%							
Muon Smearing	1%	1%	1%	1%	1%	1%	1%	1%
σ W					12%			
σ Z						1%		
σ $t\bar{t}$							7%	
σ Di-boson								4%

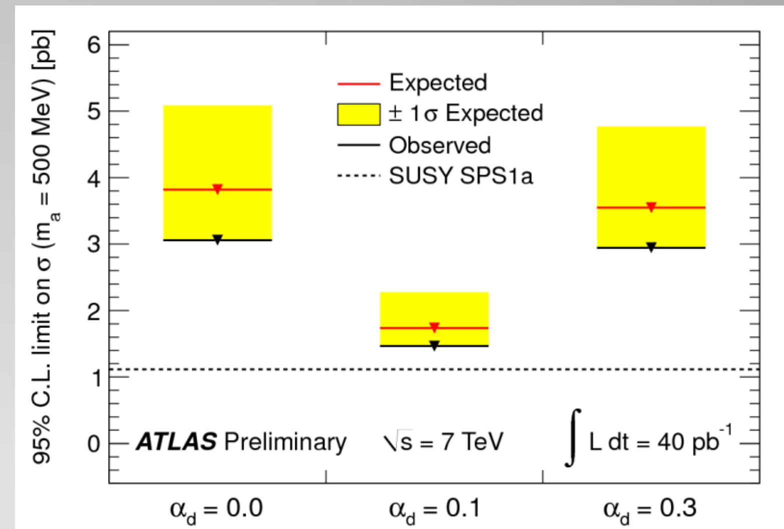
Lepton jets Kinematics (2)

- # of muons and their separation depends on m_{γ} and α_d
- Some events have electrons or pion jets, so less muons



Limits on production cross-section





- Given no events are found in ~40 pb-1 of 2010 data, we can set limits on production cross-section