Gianni Serra - Porto di Alghero

Searching for Dark Matter at the LHC

XIV Seminar on Software for Nuclear, Subnuclear and Applied Physics

Emma Tolley

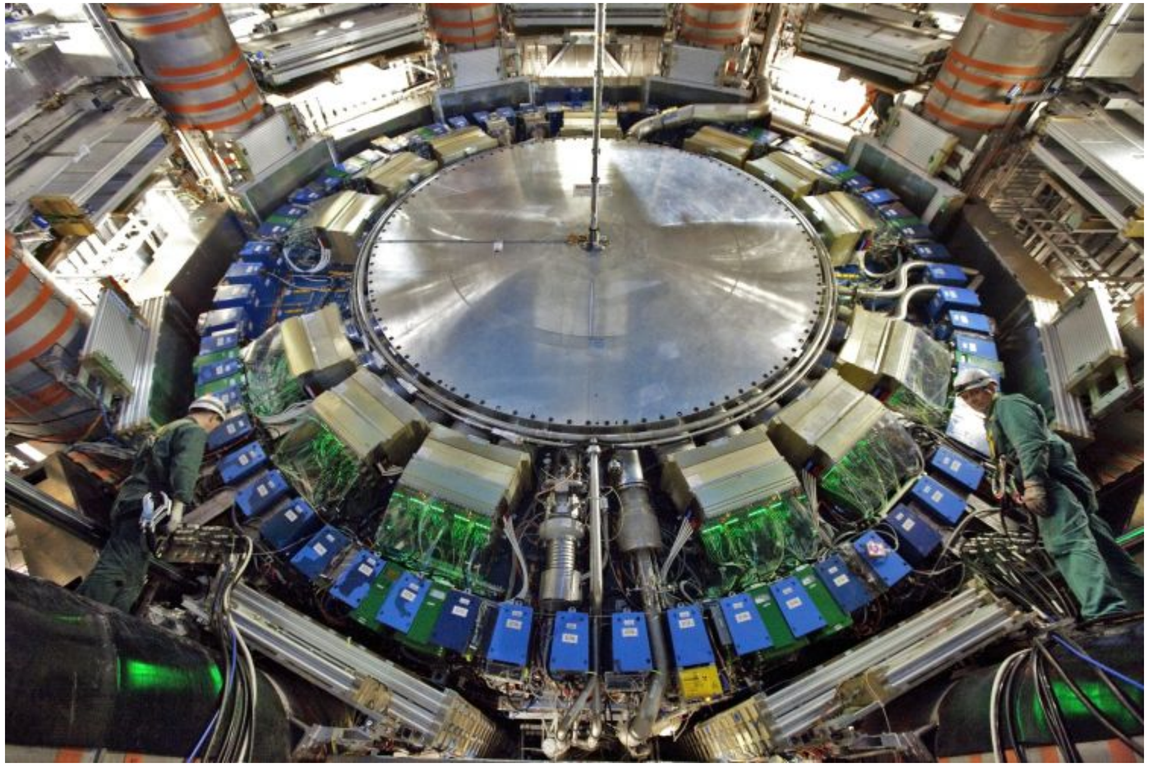
6 June 2017

The Ohio State University

ATLAS Experiment

Outline

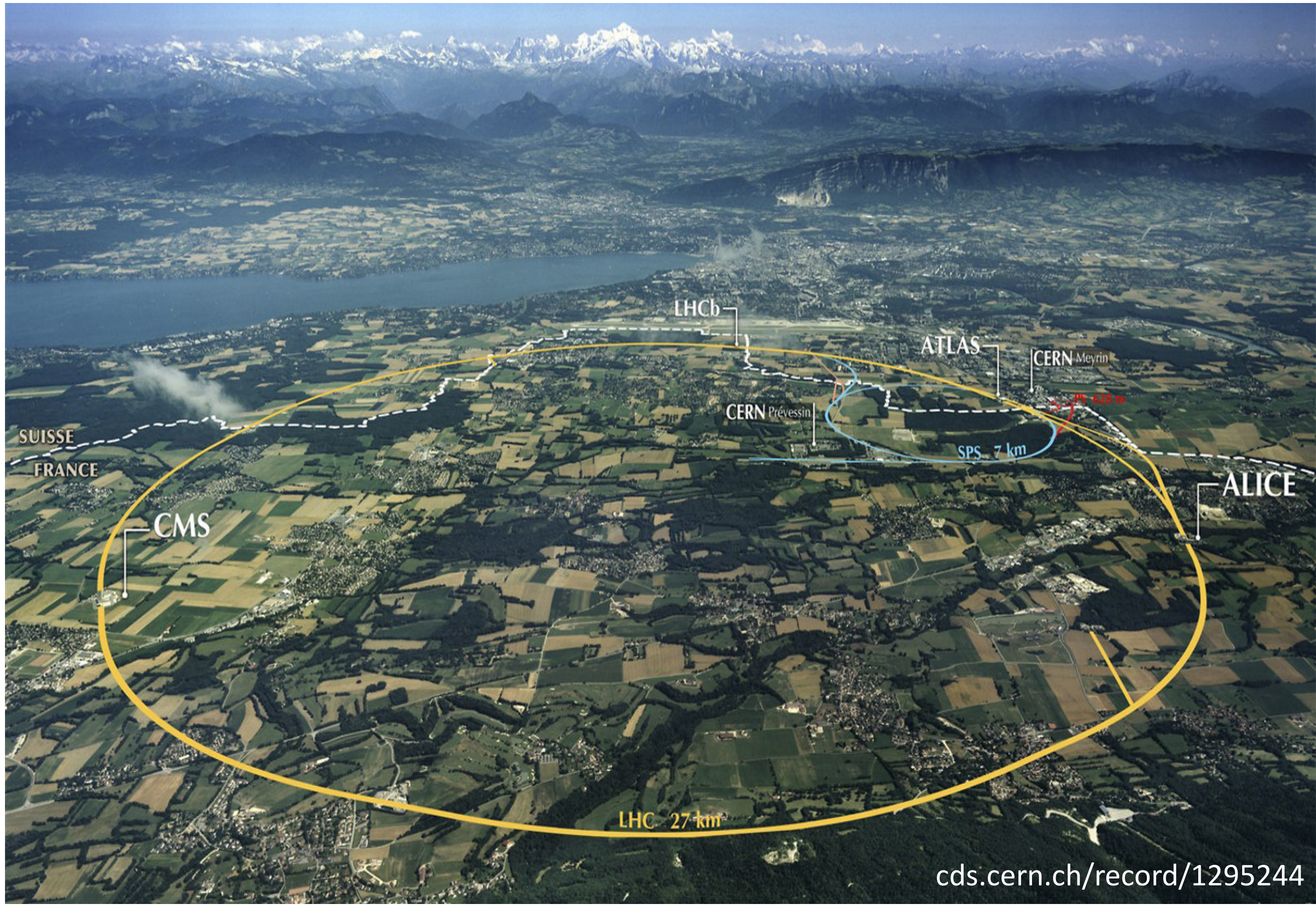
- Introduction
  - The LHC & the ATLAS detector
    - Computing at the LHC
  - Detecting dark matter
- Searches for dark matter production
  - “MET + X” searches
  - The MET + jet topology
    - Estimating backgrounds
  - Software for statistical analysis in high energy physics
- Searches for related dark sector physics
  - Dijet resonances
  - Software for recording partial events
- Complementarity of different dark matter searches



The ATLAS Detector

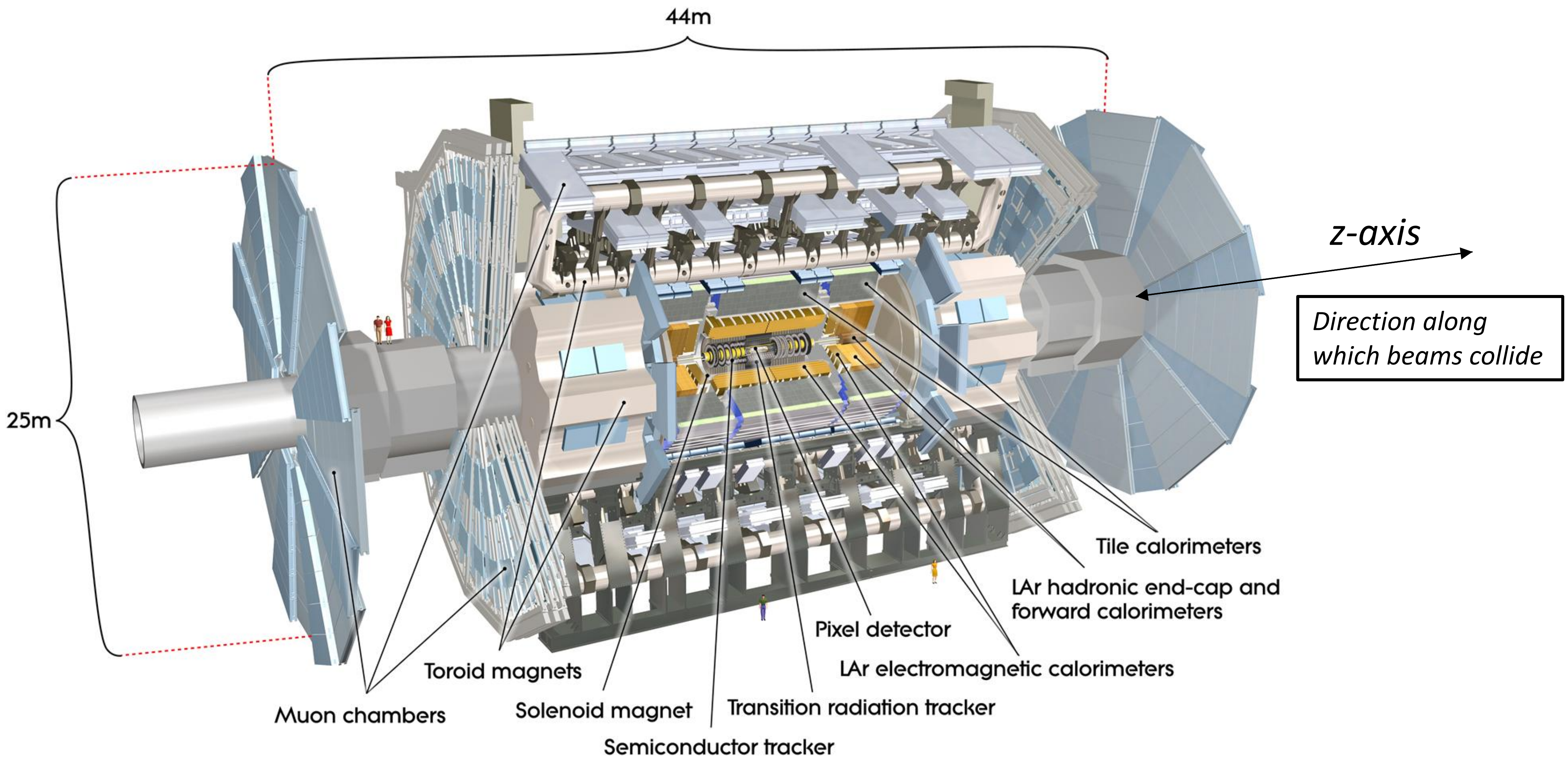
The Large Hadron Collider

- The Large Hadron Collider (LHC)
  - World's largest and most powerful particle collider
  - 27 kilometers of over 8000 superconducting magnets and accelerating structures
  - Proton-proton collision center of mass energy of 14TeV
  - Instantaneous luminosity of 10^{34} cm^{-2} s^{-1}
- Discovery of the Higgs Boson
- Precision measurements of the Standard Model (SM)
- Search for and constrain beyond-the-Standard-Model (BSM) particles
  - Such as dark matter (DM)



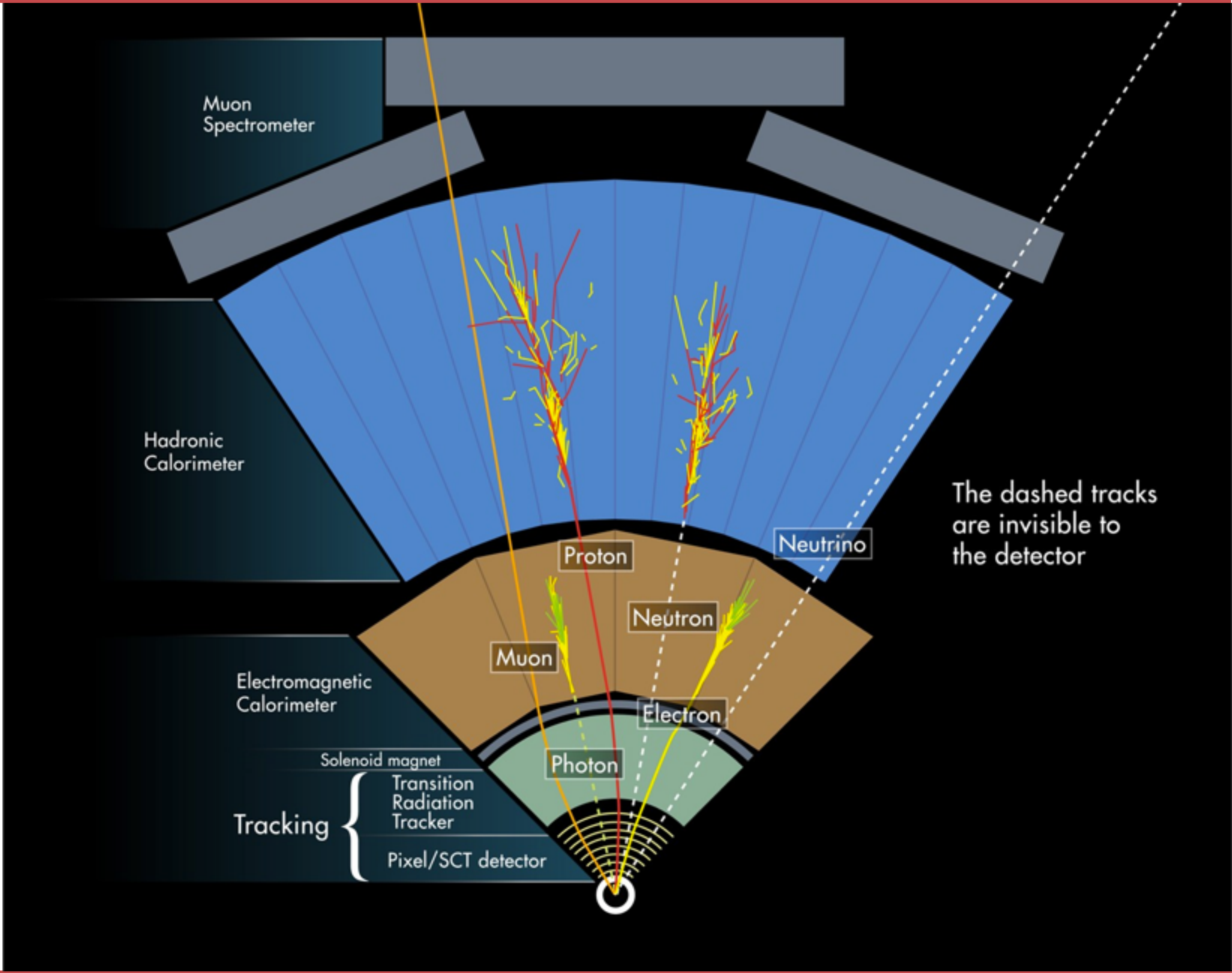
A Toroidal LHC Apparatus (ATLAS)

A particle detector built to study proton-proton collisions at the LHC



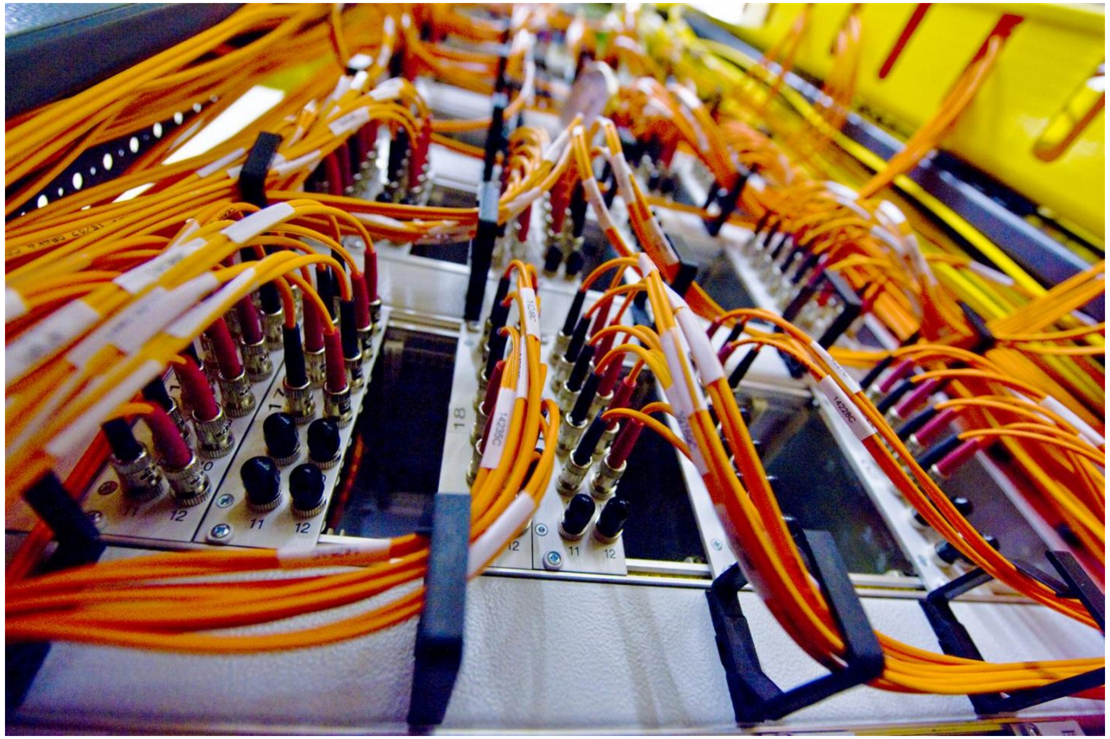
A Toroidal LHC Apparatus (ATLAS)

- Particles created in proton–proton collisions can pass through four detector systems
- Inner detector
  - Particle tracking
  - Measures the momentum charged particles
- Electromagnetic calorimeter
  - Measures the energy of electrons and photons
  - Contributes to the energy measurement of hadronic showers
- Hadronic calorimeter
  - Measures the energy of hadronic showers
- Muon spectrometer
  - Measures the momentum of muons



ATLAS Computing

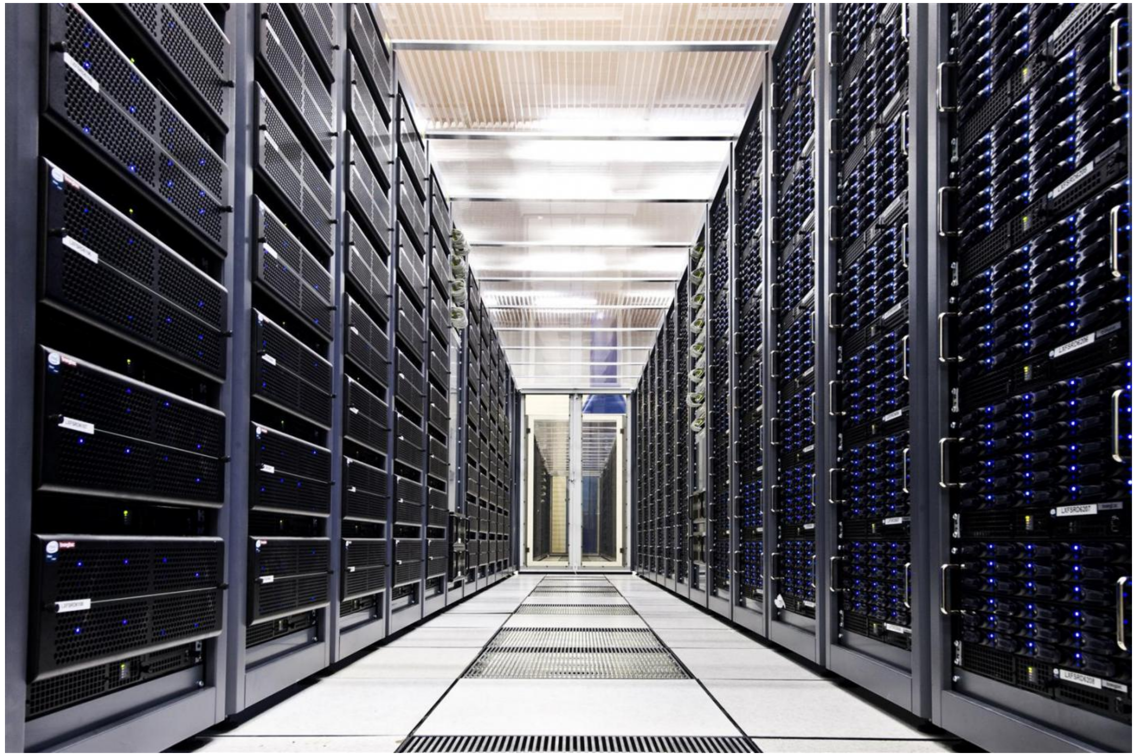
- The LHC produces a lot of data
- **Event:** a collision at the LHC
  - Bunches containing ~10^{11} protons collide
  - Dozens of proton—proton collisions
  - Hundreds of particles that interact with the detector systems
  - Need to record this information to look for interesting events
- However: ~600 million events/s!
- Only want to save interesting events
  - Hardware-level algorithms filter events based on interesting features
    - Energetic photons, multiple muons, etc
  - 100k events/s sent to digital reconstruction
  - More event processing reduces this to 100-200 events/second
- Raw data recorded onto servers at ~1050 megabytes/s
  - 15 petabytes/year



Cabling in the CERN Data Center

ATLAS Computing

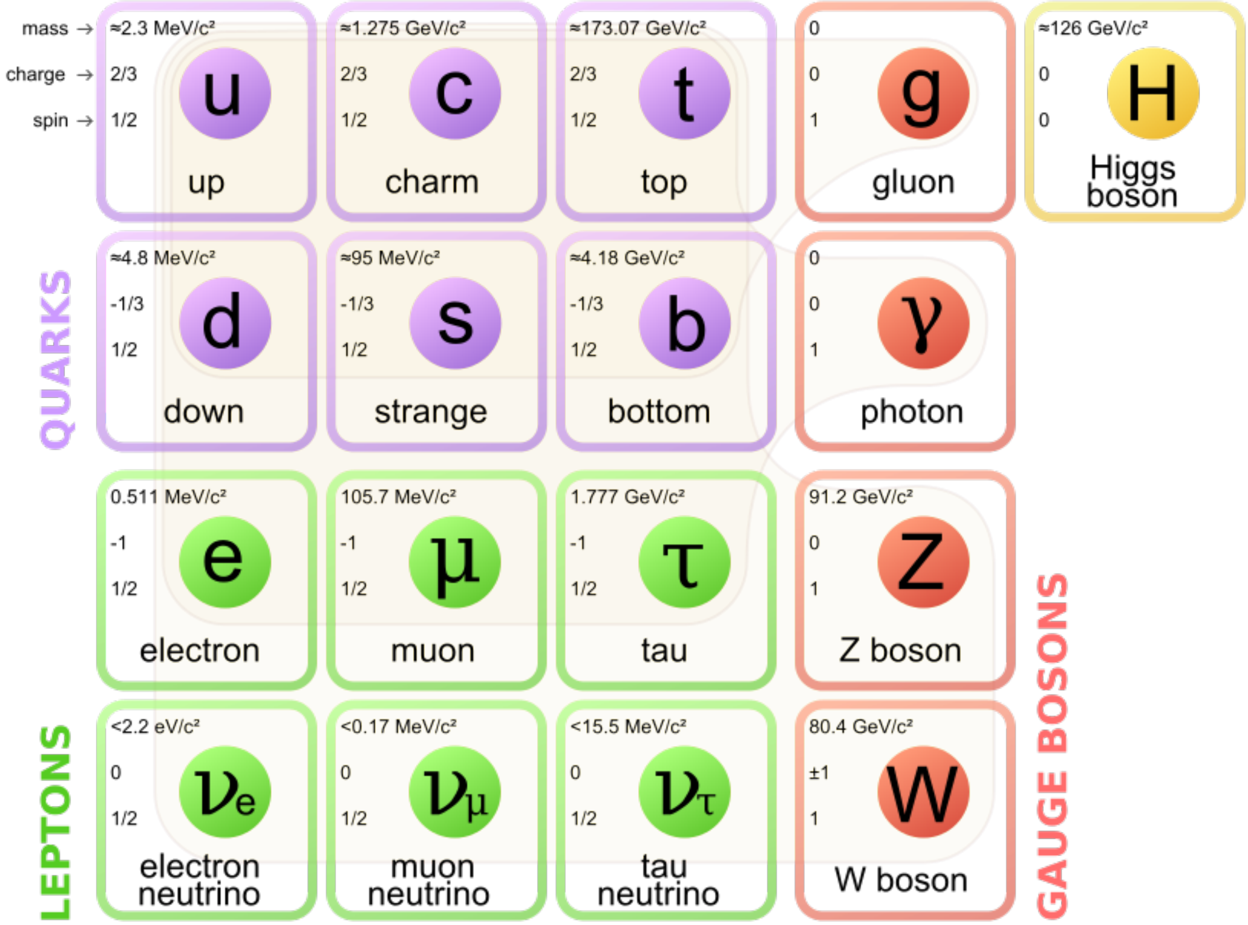
- Distributed computing & parallel processing are necessary to analyze the data
- The Grid
    - System linking thousands of computers and storage systems in over 170 centers across 41 countries
    - Data storage, processing, and analysis
    - Tier 0: CERN Data Centre
        - Raw data storage and run initial reconstruction
            - Detector hits => muon track
    - Tier 1: 13 computer centers around the world
        - Store raw & reconstructed data
        - Run data reprocessing
        - Store simulated data
    - Tier 2: universities and other scientific institutes
        - Production and reconstruction of simulated events
- Can do many studies with these data
    - Search for evidence of DM in LHC events



Servers at the CERN Data Center

What is Dark Matter?

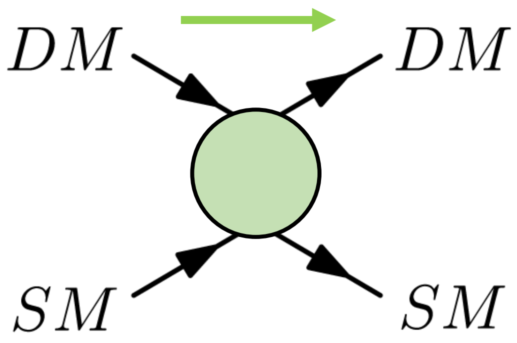
The Standard Model (SM) of particle physics
=> All known elementary particles



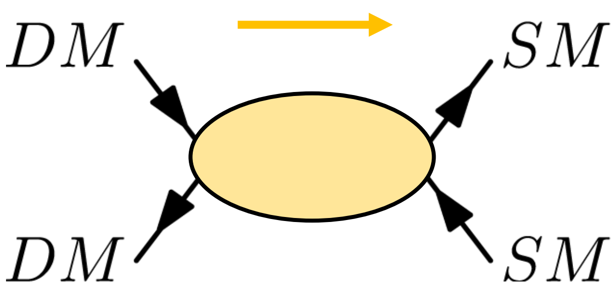
- What do we know about dark matter?
  - Dark: Nonluminous, noninteracting
  - Cold: nonrelativistic during structure formation
  - Long-lived
  - Non-baryonic
- Need to look beyond the Standard Model for particle dark matter…
  - One possibility among many: DM is a weakly interacting massive particle (WIMP)
  - Weakly interacting => DM interacts with the SM in some way

Detecting Dark Matter

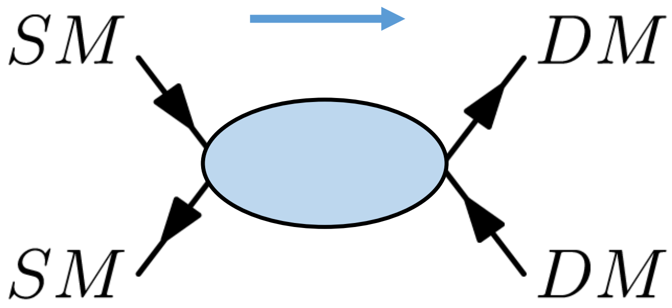
Many different strategies for detecting DM:



Direct Detection



Indirect Detection



Production at Colliders

Look for local DM halo scattering off of terrestrial nuclei

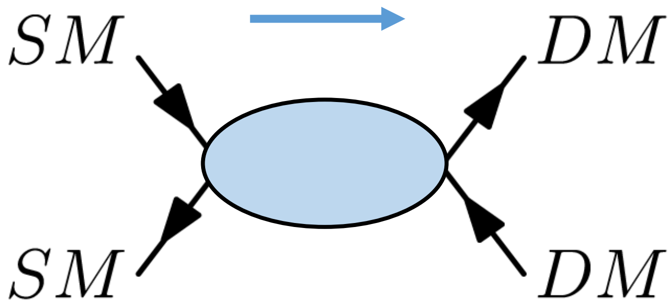
Look for signature of cosmological DM-DM annihilation

Try to make DM in a particle collider

Detecting Dark Matter

CONS:

- Dark matter is dark
  - No interaction with detector
  - Need to identify DM production in association with other objects
  - Lots of fake backgrounds
  - No clean, distinct signature
- Requires that DM interacts with some component of a proton
- If we do find something, how do we know it's cosmological DM?



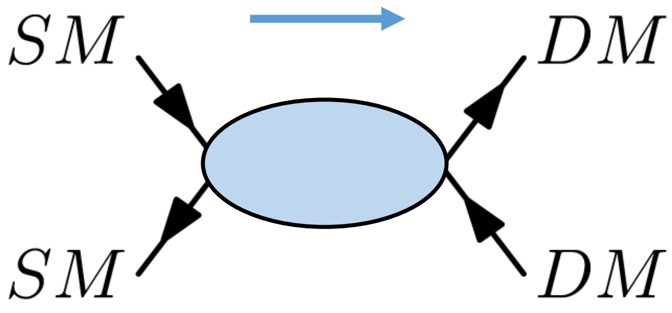
Production at Colliders

Try to make DM in a particle collider

Detecting Dark Matter

PROS:

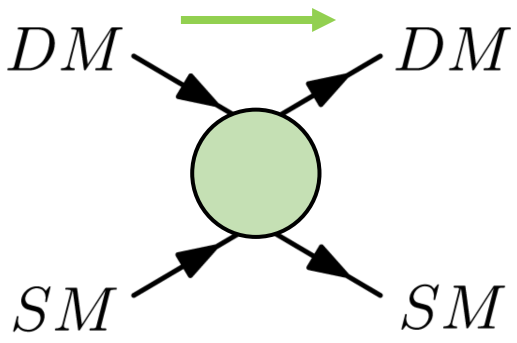
- Independent of astrophysical uncertainties
  - Local density & velocity
  - Galactic densities
- Can probe a wide range of interaction types
  - Equal sensitivity for spin-dependent or spin-independent



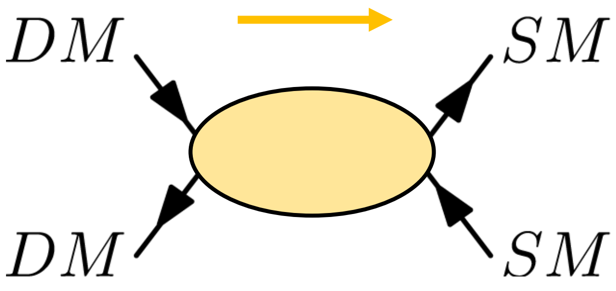
Production at Colliders

Try to make DM in a particle collider

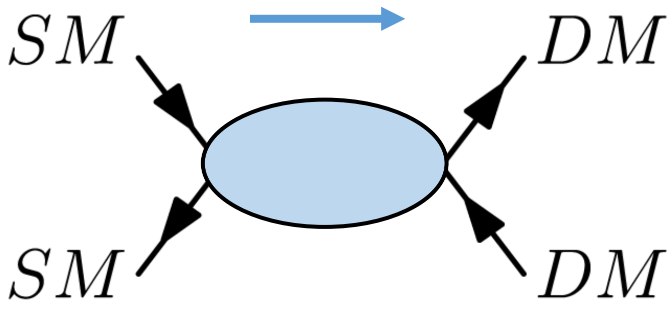
Detecting Dark Matter



Direct Detection



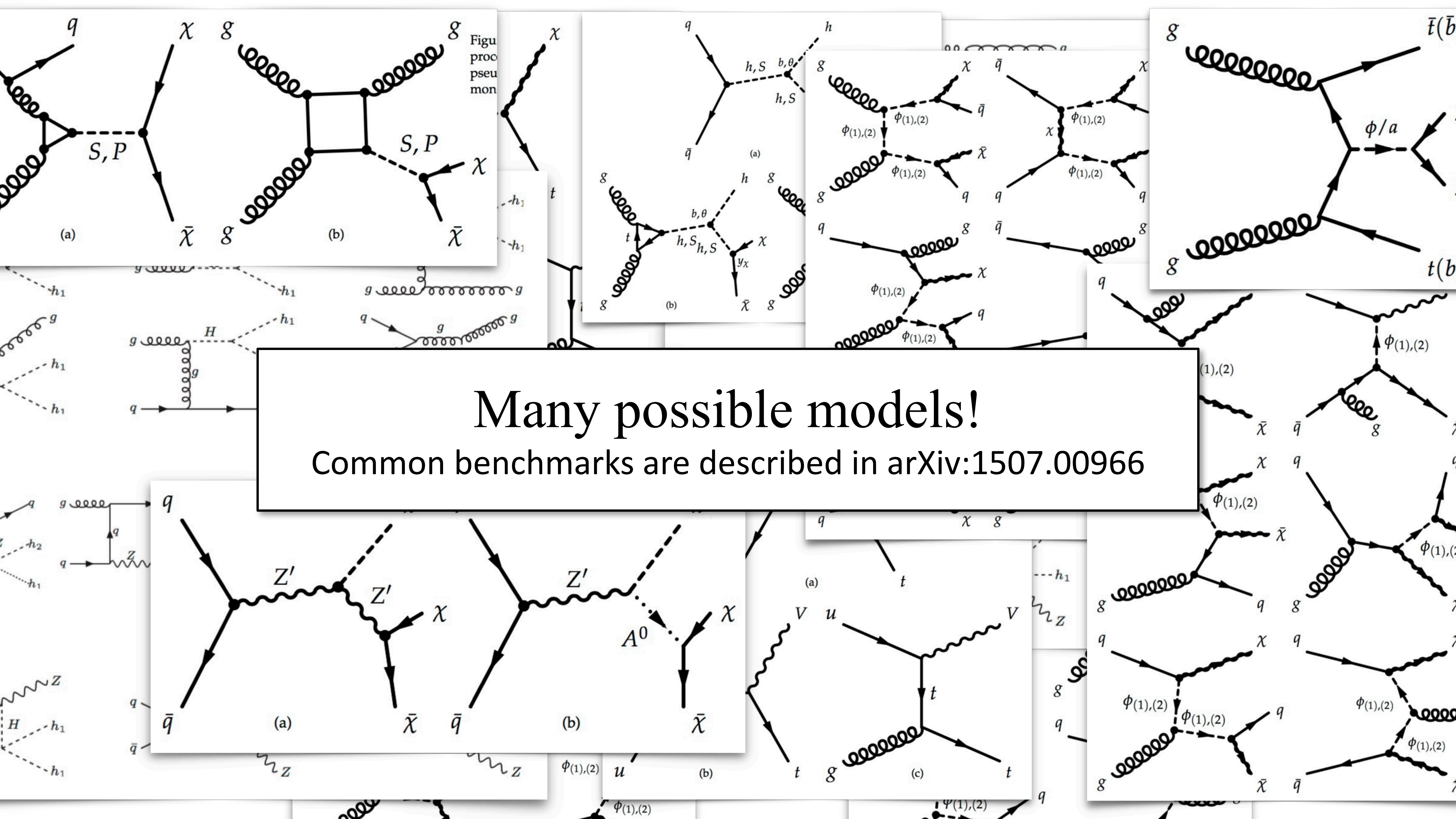
Indirect Detection



Production at Colliders

Need a common theoretical framework to compare different limits

➥ would prefer to remain agnostic to physics beyond DM



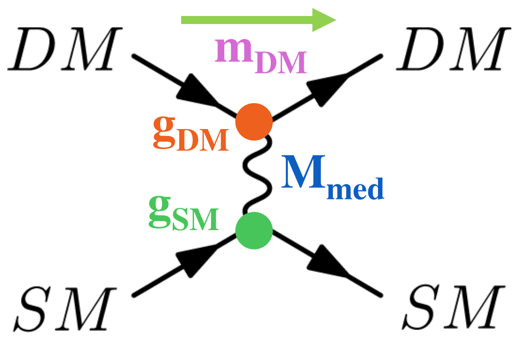
Many possible models!

Common benchmarks are described in arXiv:1507.00966

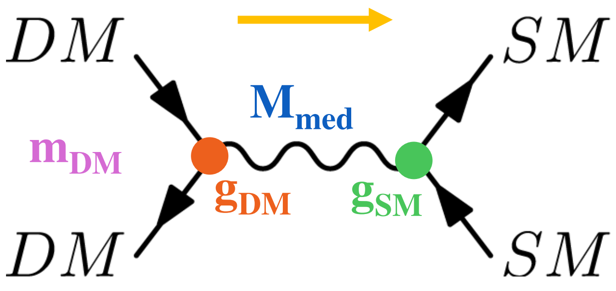
Detecting Dark Matter

One possibility among many: new mediator couples SM and DM

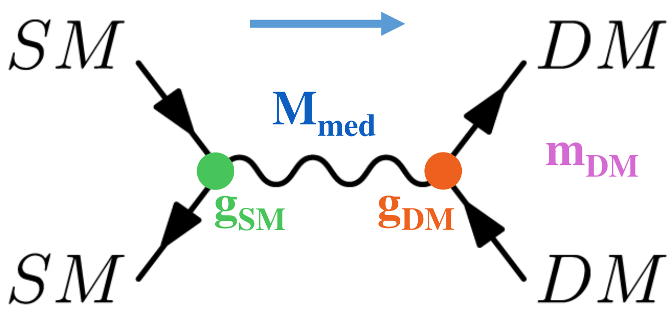
➥ Vector (spin-independent) or Axial Vector (spin-dependent)



Direct Detection



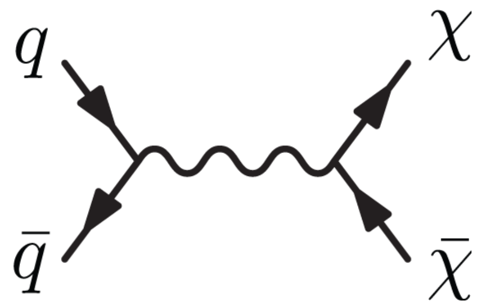
Indirect Detection



Production at Colliders

Model depends on four parameters:

- DM mass
- Mediator mass
- SM-mediator coupling
- DM-mediator coupling

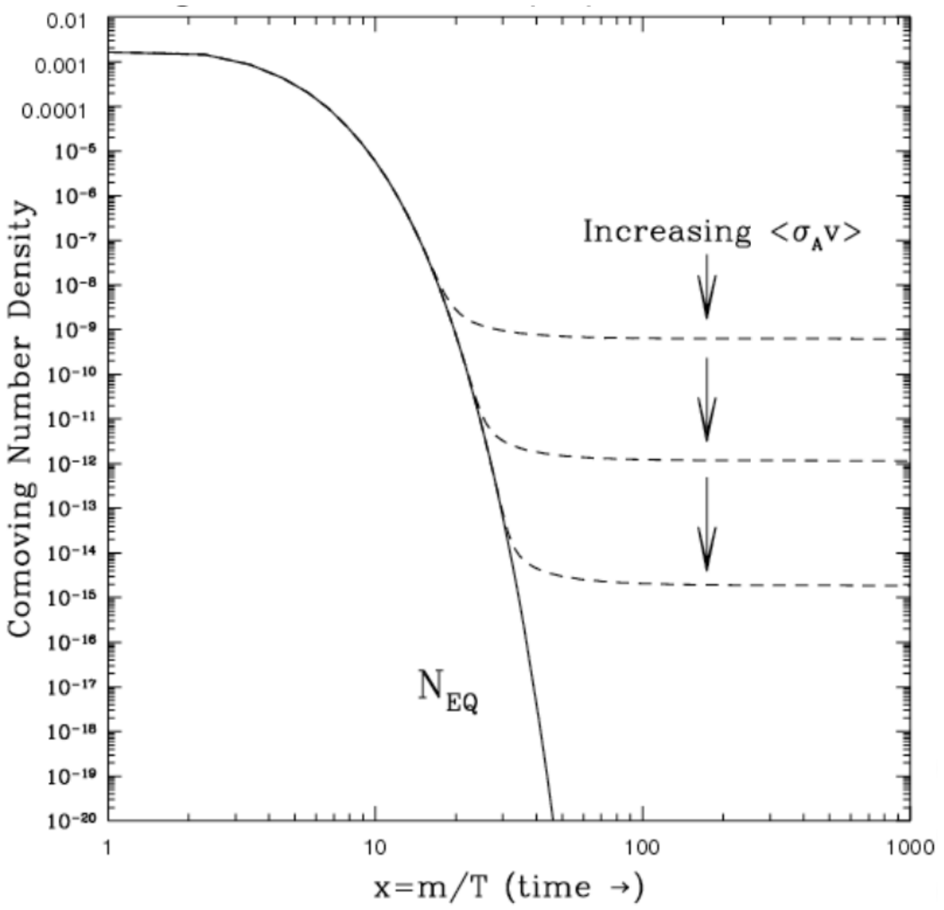


initial state:
Quark & anti-quark

Final state: pair of
DM particles χ

WIMPs and Relic Density

- WIMP dark matter
  - DM in thermal equilibrium with SM in early universe
  - Relic density "freezeout" as universe expands
- The Planck experiment measures the cosmological DM relic density $\Omega_{DM}h^2 \sim 0.12$
  - Interesting to compare this to the relic density predicted by signal models



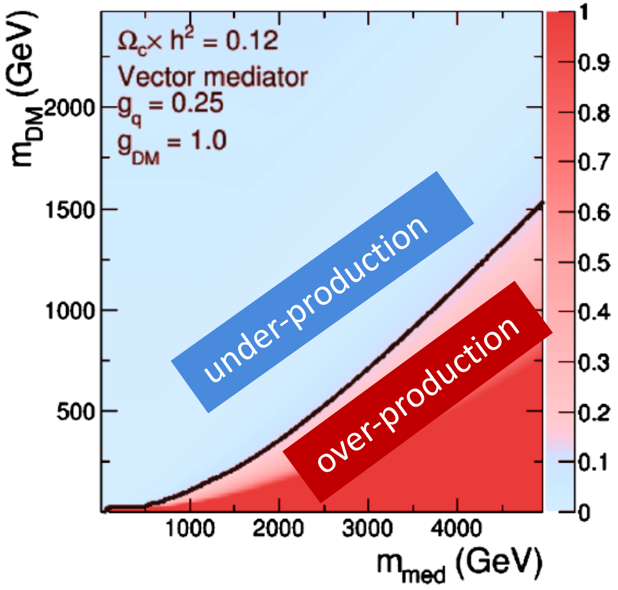
Larger annihilation cross section => later freezeout => lower relic density

Calculation of Relic Density

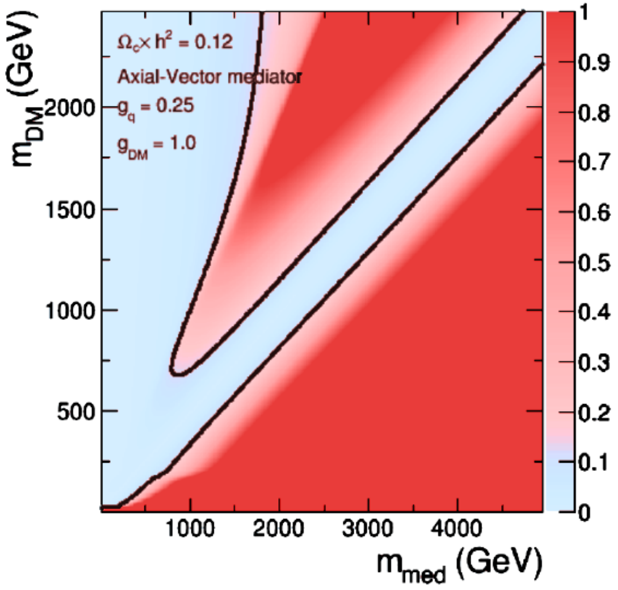
- Numerical calculation of the relic density with MadDM (alternative: MicrOMEGA)
  - arXiv:1505.04190 & arXiv:1308.4955
  - DM phenomenology framework built on MadGraph5 aMC@NLO
  - Calculates thermal relic density for a given model of DM
- Calculation of the dark matter relic abundance for the simplified models relies on the following assumptions:
  - DM only couples to the DM via the mediator.
  - No additional BSM particles couple to the mediator
  - No additional BSM particles couple to DM

Numerical calculation of relic density with MadDM

Vector Mediator



Axial Vector Mediator

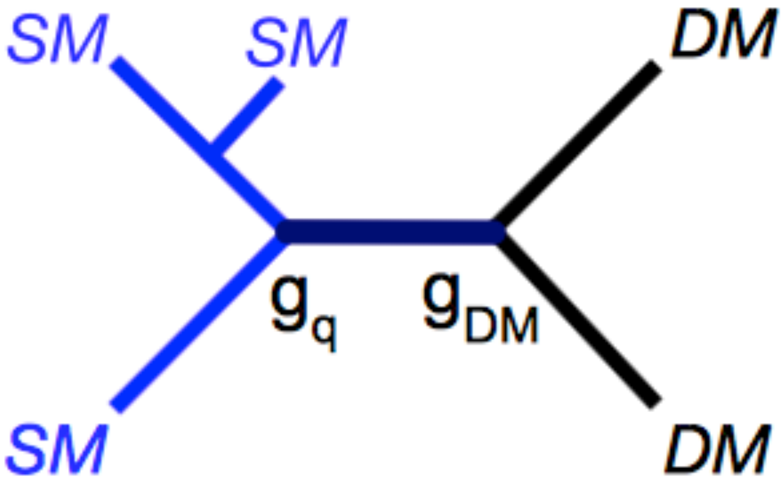


Higher predicted relic abundance for axial vector mediators => s-channel annihilation is helicity suppressed

Figures adapted from arXiv:1703.05703v2

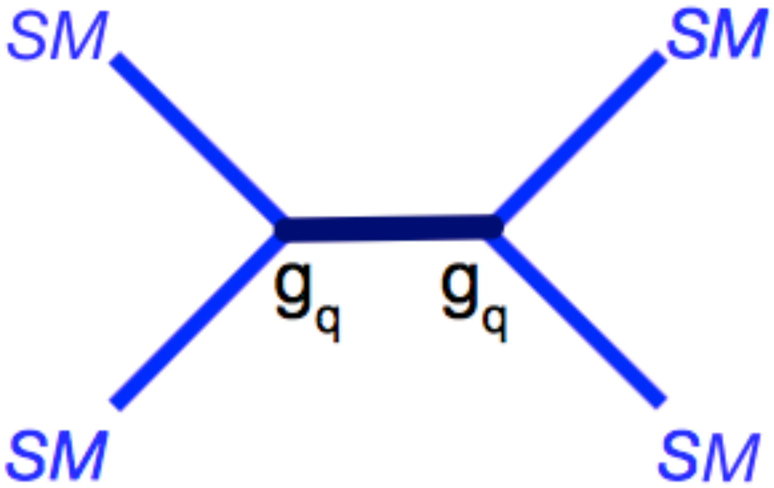
DM Searches in this Lecture

Look for DM particles



Discussing software for simulation of models, statistical analysis

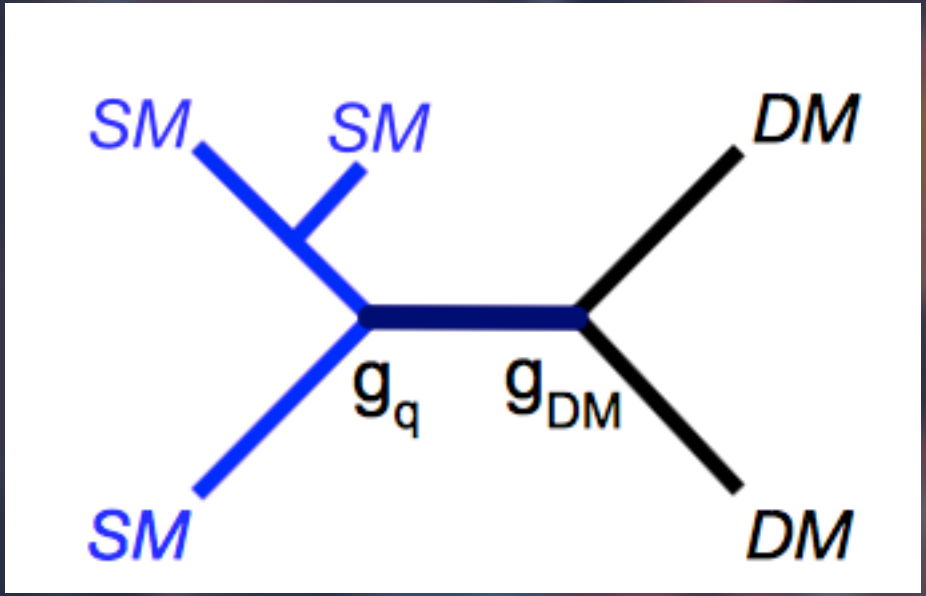
Look for visible decays of the mediator



Discussing software for online data selection, interpretation of results

These are just examples => other searches at the LHC look for different kinds of dark matter (SUSY neutralino, dark sector, etc)

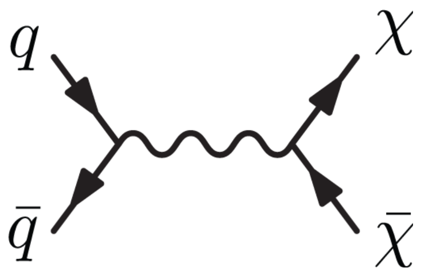
Searching for Dark Matter Particles

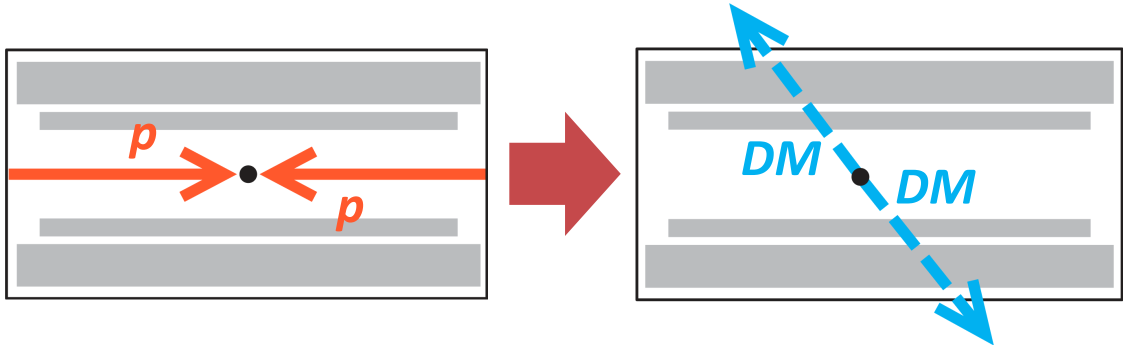


- Introduction to a DM search at the LHC
- Software to simulate DM signals at the LHC
- Software to discover DM in a quantitative way

Dark Matter at the LHC

What would this process look like in a detector?



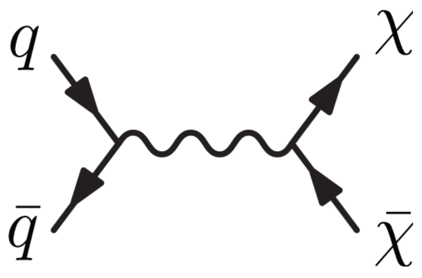


Protons collide…

…creating invisible particles that do not interact with any detector systems

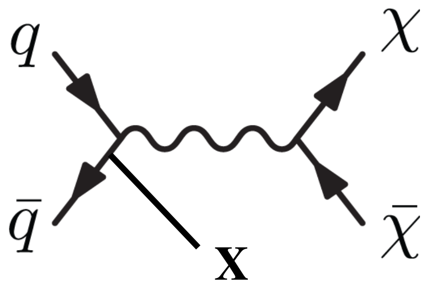
Dark Matter at the LHC

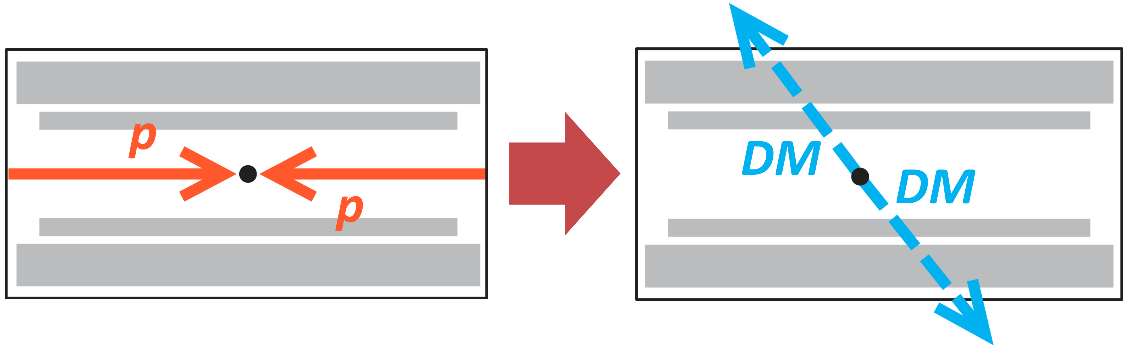
What would this process look like in a detector?



Instead: look for DM produced in association with another particle

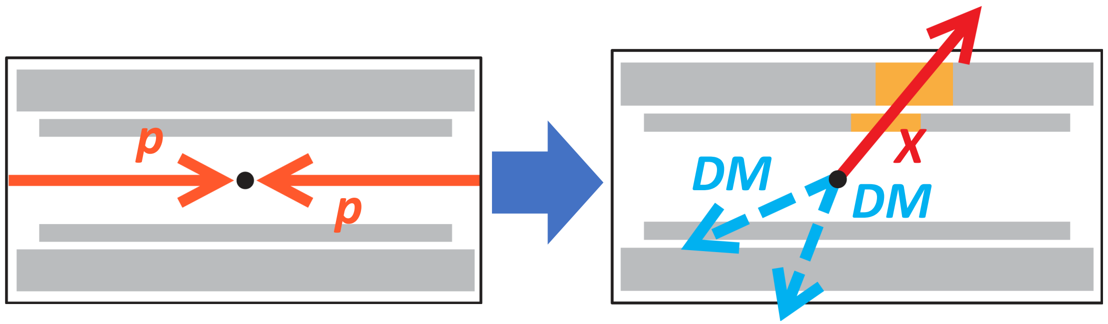
Radiating from a particle in the initial state:
initial state radiation (ISR)





Protons collide…

…creating invisible particles that do not interact with any detector systems

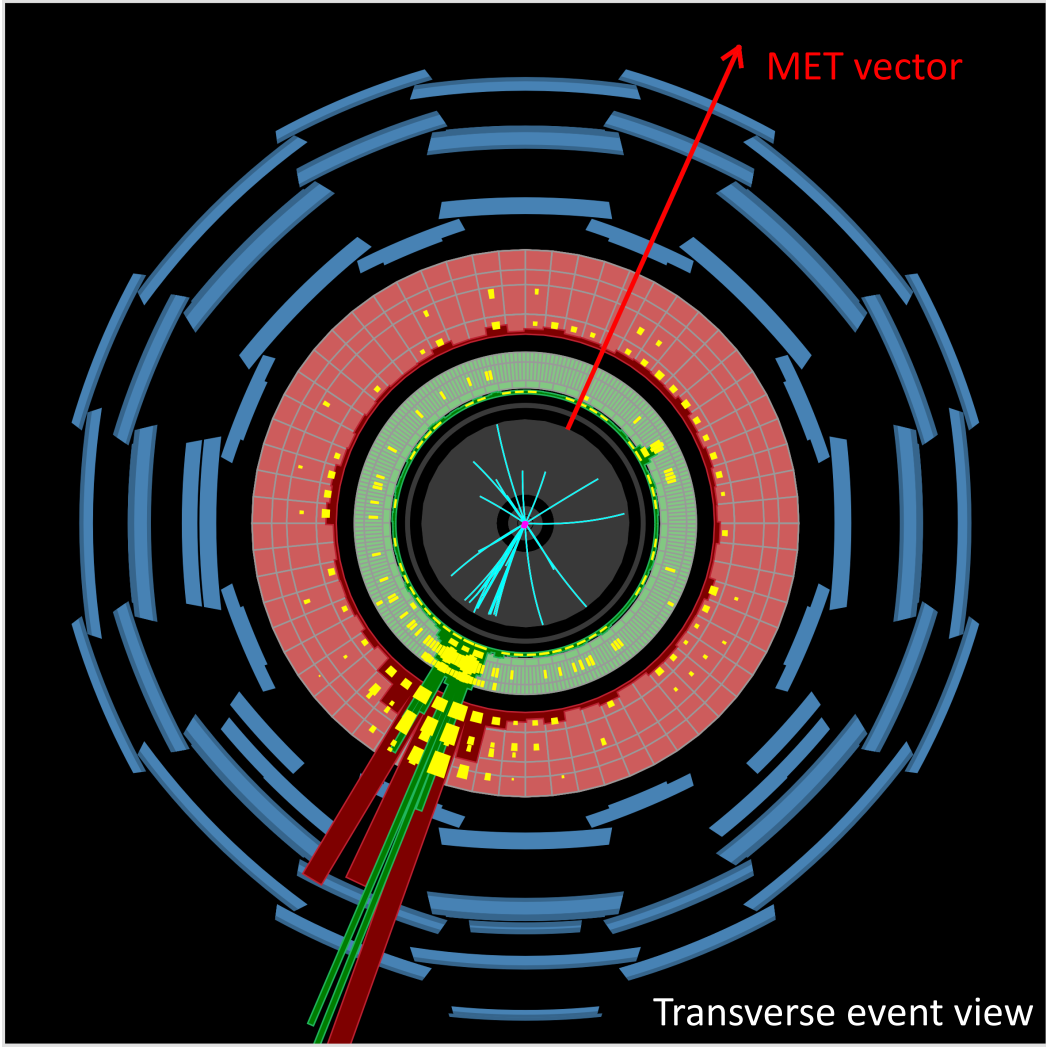


Protons collide…

…creating DM and visible particles that interact with detector systems

Missing Transverse Momentum

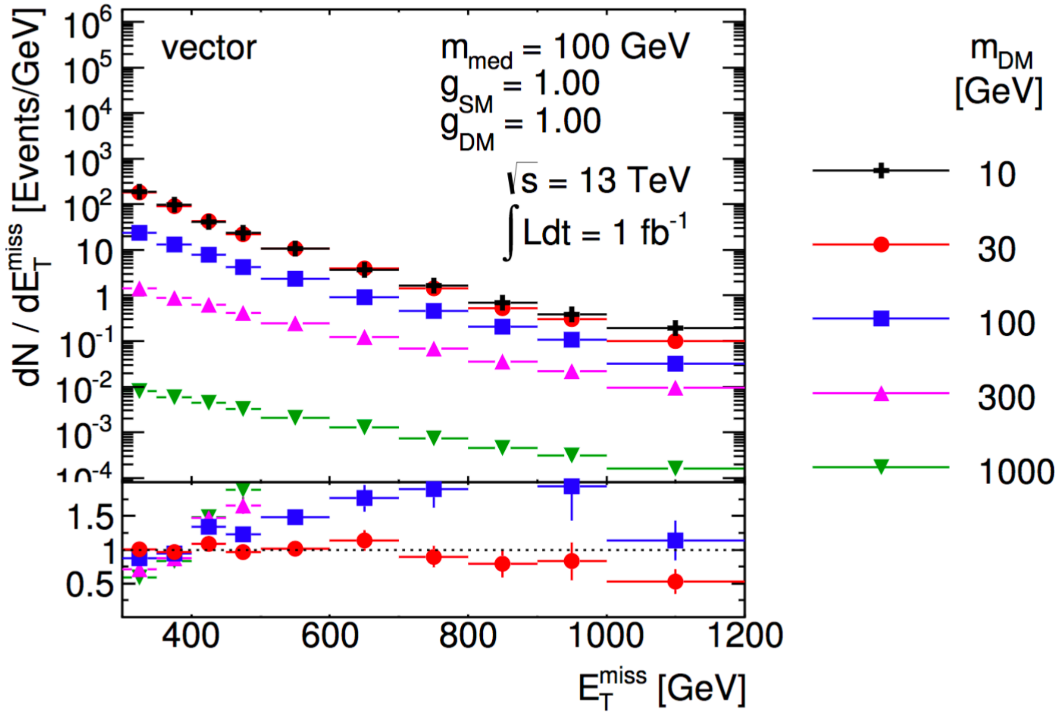
- How to find invisible particles:
- Proton-proton collision => colliding partons
  - Proton is a composite particle!
    - At LHC collision energy, the particles within the proton collide
    - Gluon-gluon, quark-gluon, quark-antiquark, etc
  - Colliding particles have different (unknown) momentum along z-axis
  - However, zero momentum in transverse plane p_T
- Conservation of momentum
  - If the sum of p_T of final-state objects is imbalanced, final state likely includes invisible particles
- Missing transverse momentum (MET)
  - -1 * (vector sum of p_T of all reconstructed particles)
  - Should be ~ vector sum of p_T of all invisible particles



Generation of Dark Matter Signal Samples

- Not sure which model might describe DM
  - Therefore try to test many models with many different sets of a parameters
- Want to know predicted distribution of observables
  - MET, jet energy, etc
  - Usually cannot be formulated analytically
    - Physics model ⊗ soft physics ⊗ detector response ⊗ reconstruction
- Generate many Monte Carlo events in order to work out the consequences of the free parameters of each model
  - Use MadGraph: http://madgraph.physics.illinois.edu/
  - Generate events from “UFO” file describing the particles and interactions within the model
- Most models described in this talk are implemented with DMSimp (simplified dark matter models)
  - http://feynrules.irmp.ucl.ac.be/wiki/DMsimp
- Generated events are then passed through detector simulation based on Geant4

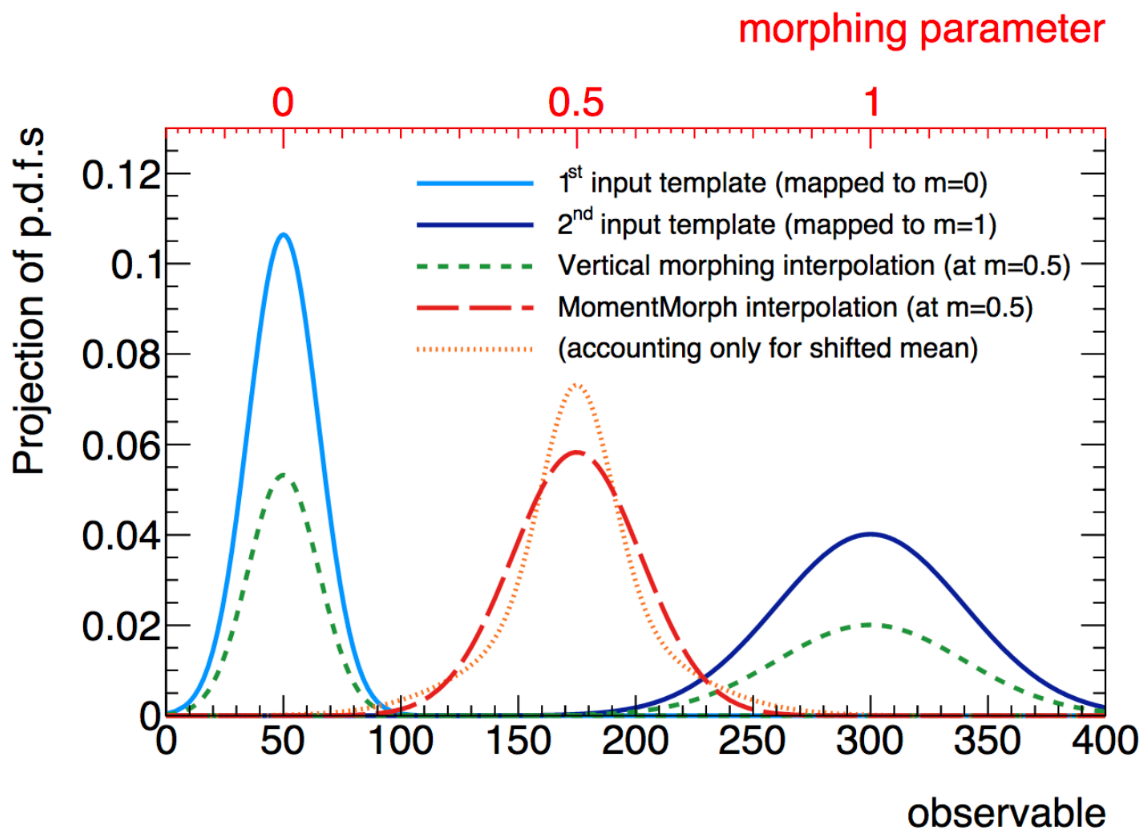
arXiv:1507.00966



Scan over DM mass

Signal Samples

- Impossible to generate enough MC to sufficiently populate every point in the large parameter space
  - Need to decide which parameters are the most interesting to scan
- Can use morphing to extrapolate between points in parameter space
  - Vertical morphing
    - Piecewise linear interpolation between each bin of existing signal distributions
    - However, shifting distributions poorly described
  - Moment morphing: arXiv:1410.7388
    - Extension of vertical morphing
    - Also adjust mean and width of signal distributions
  - Other techniques under study

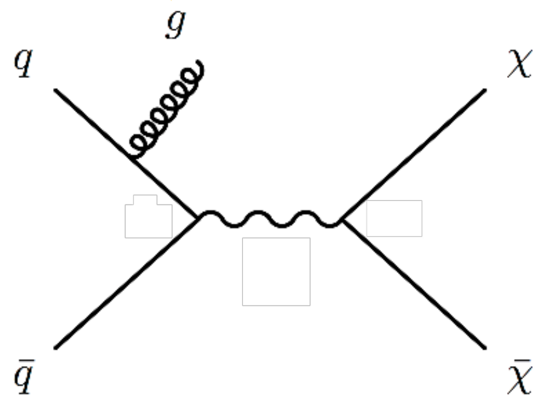


MET+ X

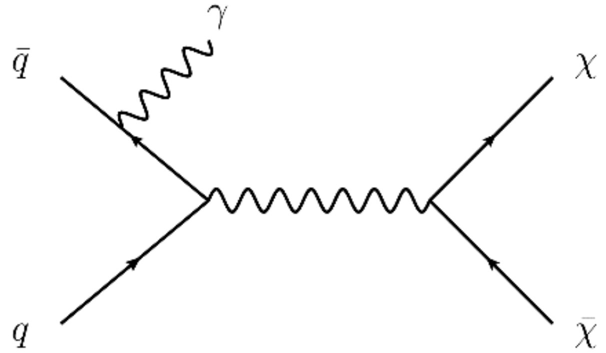
- Search topology: MET + X
  - If we expect X to come from ISR, can look for a jet, photon, W or Z boson
- Not all ISR is created equally
  - Production cross section scales with (q-X coupling)^2
  - Largest cross section for gluon ISR
    - A gluon is reconstructed as a hadronic shower or “jet” in ATLAS
- Next slides: some details of the statistical analysis used for the 2015 Search for DM in jet + MET
  - Phys. Rev. D 94 (2016) 032005

Gluon has α_s ~ 0.1 coupling to initial state quarks

Photon has α_{EM} ~ 0.01 coupling to initial state quarks



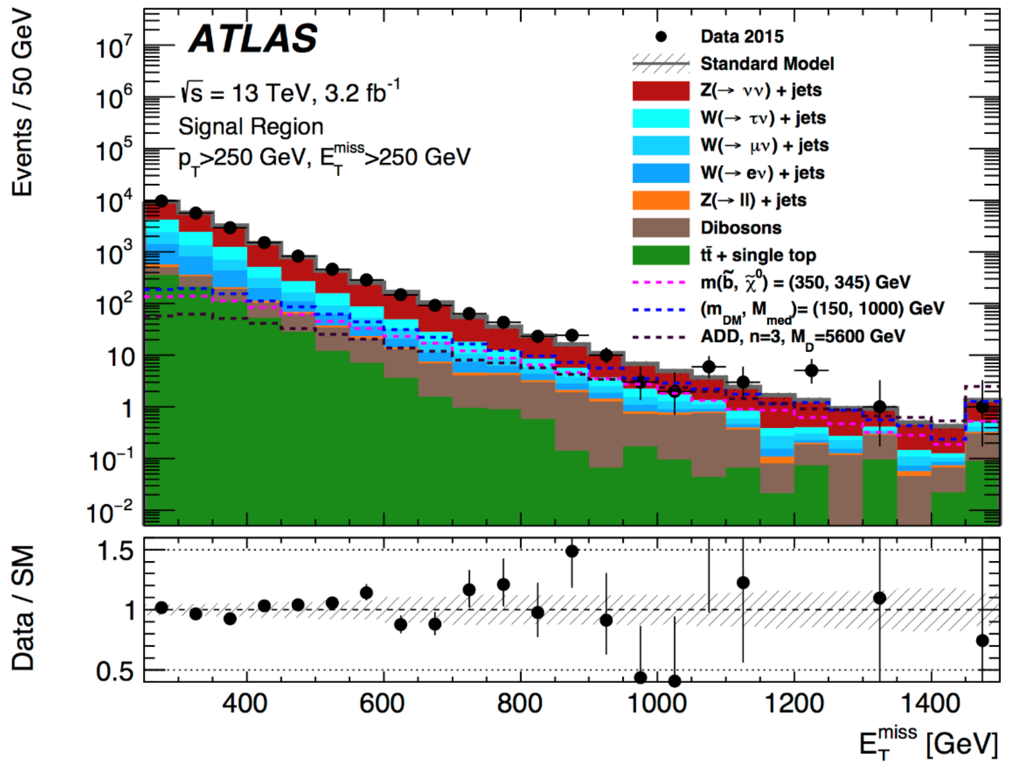
DM production in the MET+ jet topology



DM production in the MET+ photon topology

Analyzing the Jet + MET Topology

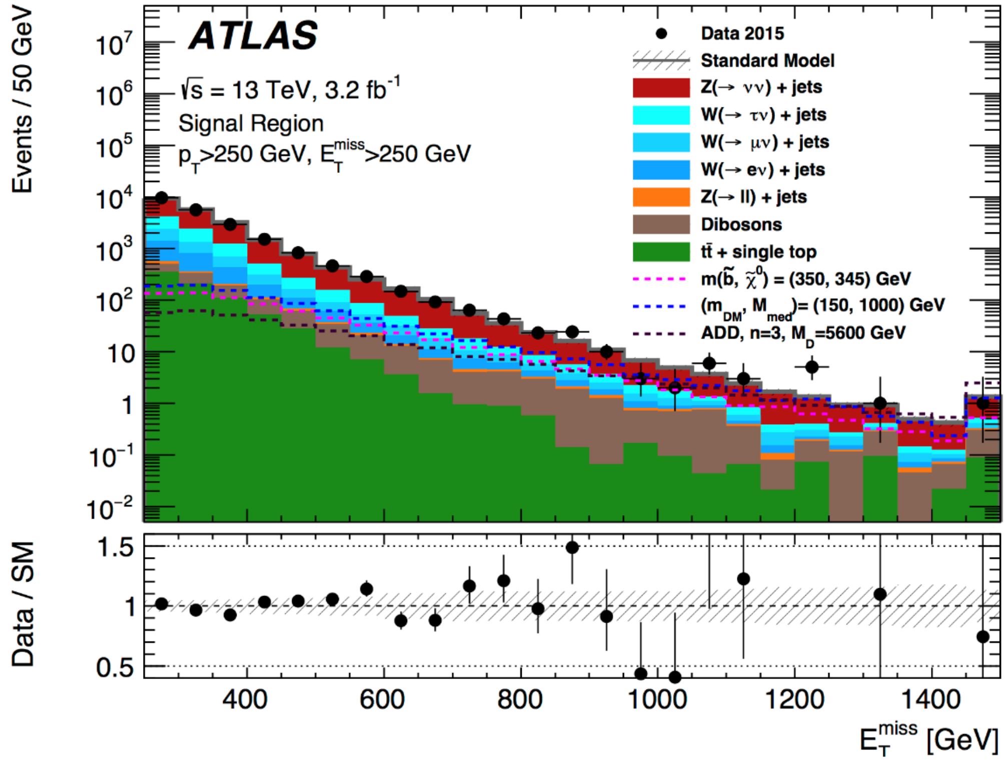
- Expect DM production to manifest as jet + MET events
- However, many physics processes look like a jet + MET
  - Z boson produced in association with jets and decaying to ν ν
    - Looks like MET (ν ν) and a jet (ISR)
  - W boson produced in association with jets
    - Decaying to e ν or μ ν, where the lepton is lost
      - Looks like MET (e/μ + ν) and a jet (ISR)
    - And other decay modes…
  - And many other processes…



Histogram of MET in the jet+MET topology
- Background distributions stacked as solid colors
- Signal distributions overlaid as dotted lines
- Generally expect signal to have “harder” MET

Analyzing the Jet + MET Topology

- Need to know how many jet + MET events are expected to be produced from background processes
  - An excess of observed events could be a sign of new physics!!
  - …or poorly-modeled background
- As for signal, generate Monte Carlo simulation for most background processes
  - However, modeling is not necessarily reliable
  - Correct Monte Carlo predictions using data

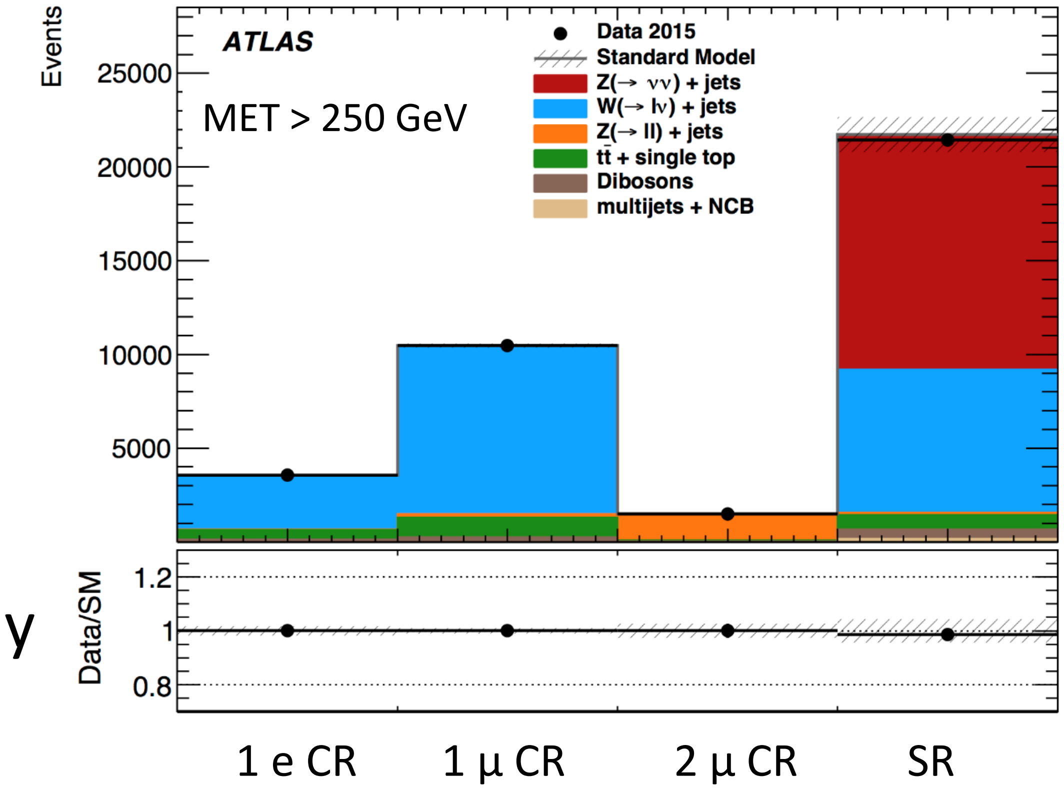


Histogram of MET in the jet+MET topology

- Background distributions stacked as solid colors
- Signal distributions overlaid as dotted lines
- Generally expect signal to have "harder" MET

Analyzing the Jet + MET Topology

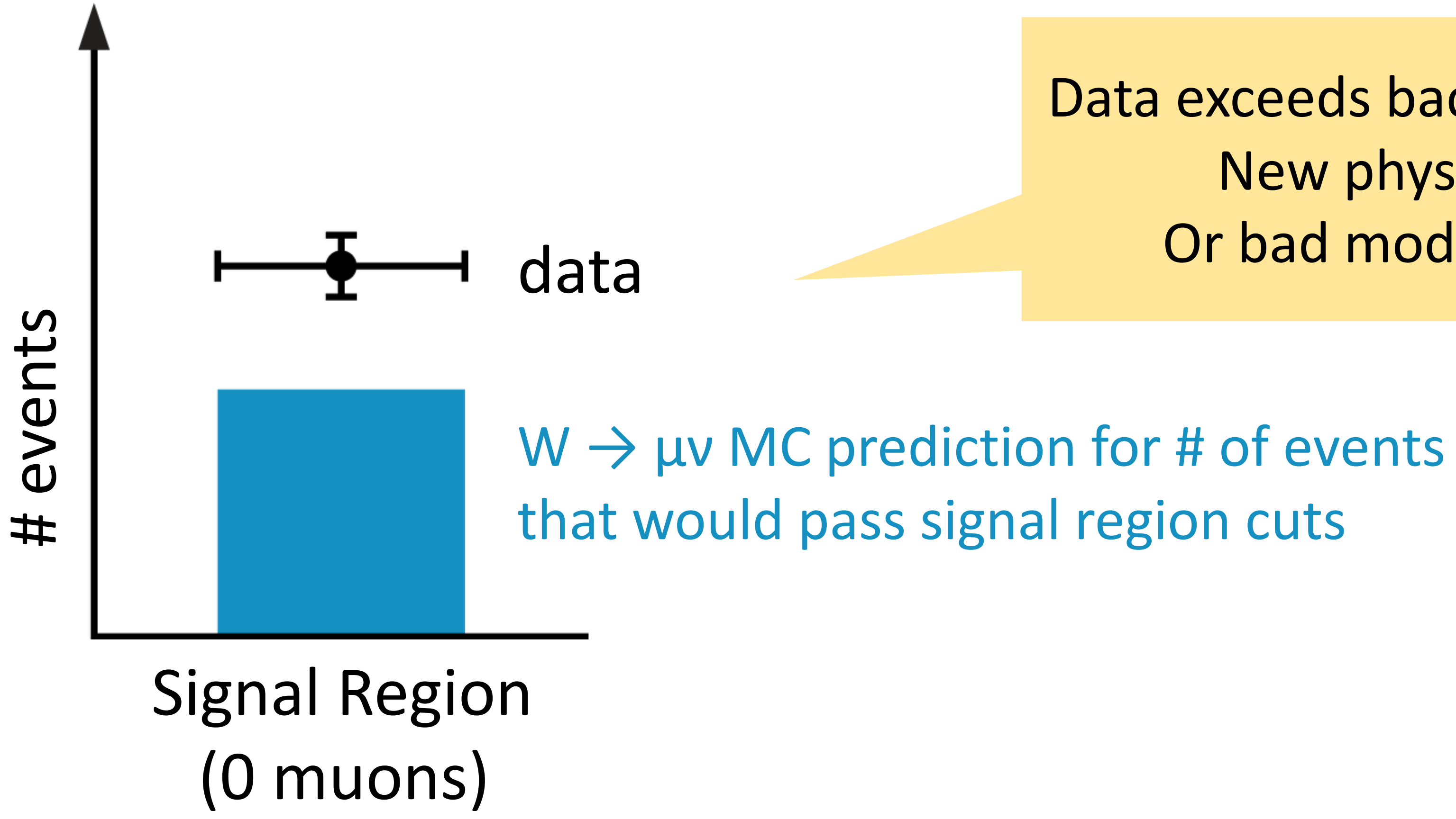
- General Strategy
  - Similar for most MET+X searches
- Define a signal region (SR) in which one might expect to see DM:
  - Require MET (MET > 250 GeV)
  - Select for X (at least one jet with p_T > 250 GeV)
  - Veto other objects (no e or μ)
  - Additional cuts to suppress backgrounds
- Define control regions (CRs) that are kinematically similar to SR, but devoid of DM signals
  - Same selection as SR, but invert lepton vetoes
    - Example: jet + MET + 1 muon => lots of W →μ ν events!
- Run a simultaneous fit across signal & background



Signal & control region yields and compositions

Simplified Background Estimation

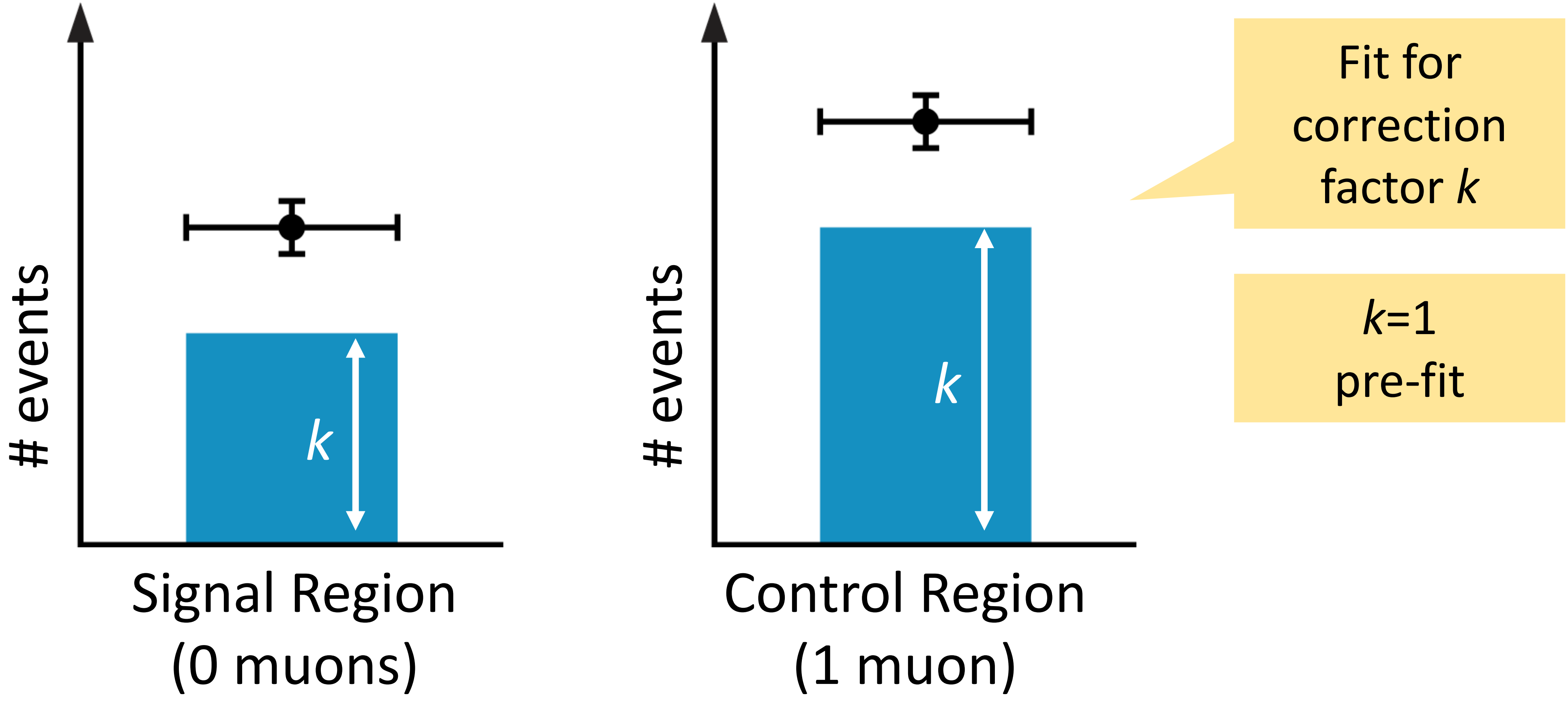
A very simple example for an alternate universe in which $W \to \mu\nu$ is the only background:



Data exceeds background…
New physics?
Or bad modeling?

Simplified Background Estimation

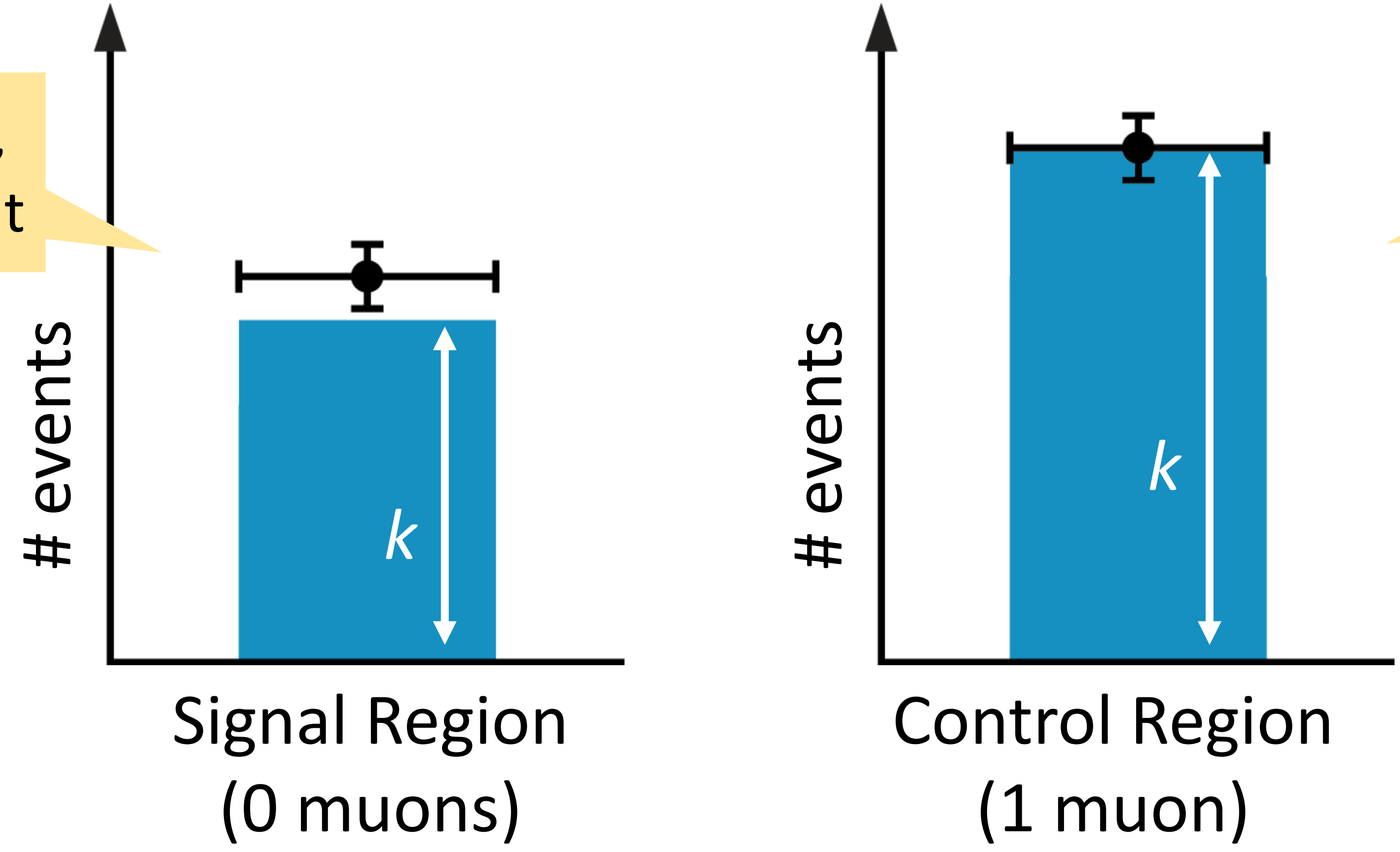
A very simple example for an alternate universe in which $W \rightarrow \mu\nu$ is the only background:



Simplified Background Estimation

A very simple example for an alternate universe in which $W \to \mu\nu$ is the only background:

Excess still there, but less significant

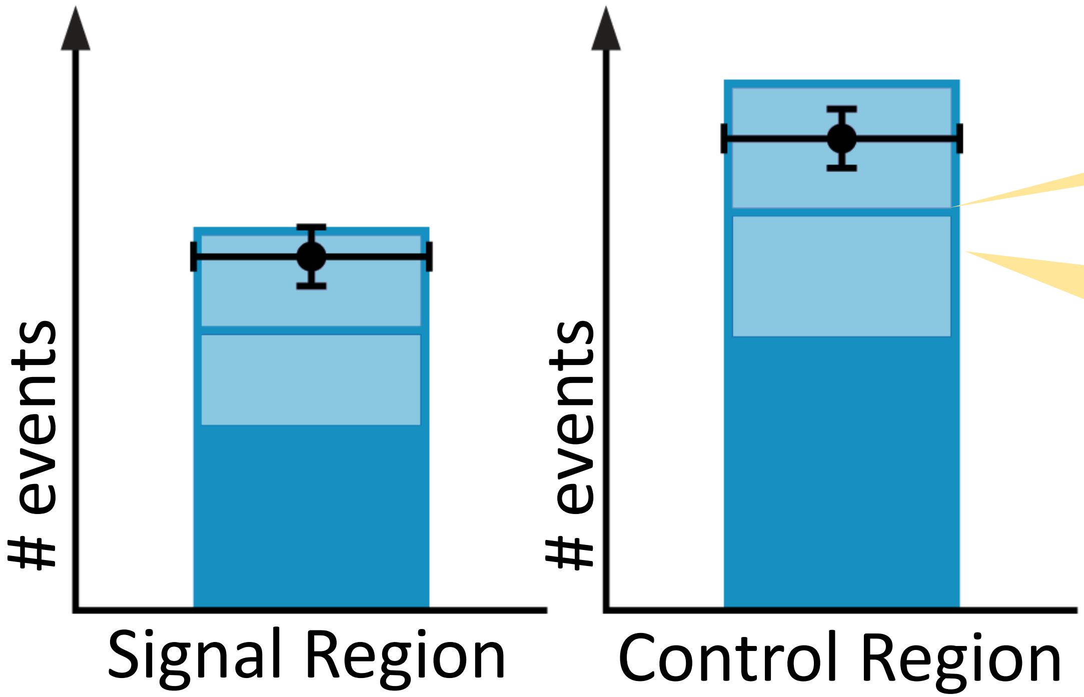


Fit for correction factor *k*

k=1.3 post-fit

Simplified Background Estimation

This technique also strongly constrains correlated systematic uncertainties



Nominal MC prediction

Systematic uncertainty on yield (added as nuisance parameters in fit)

pre-fit

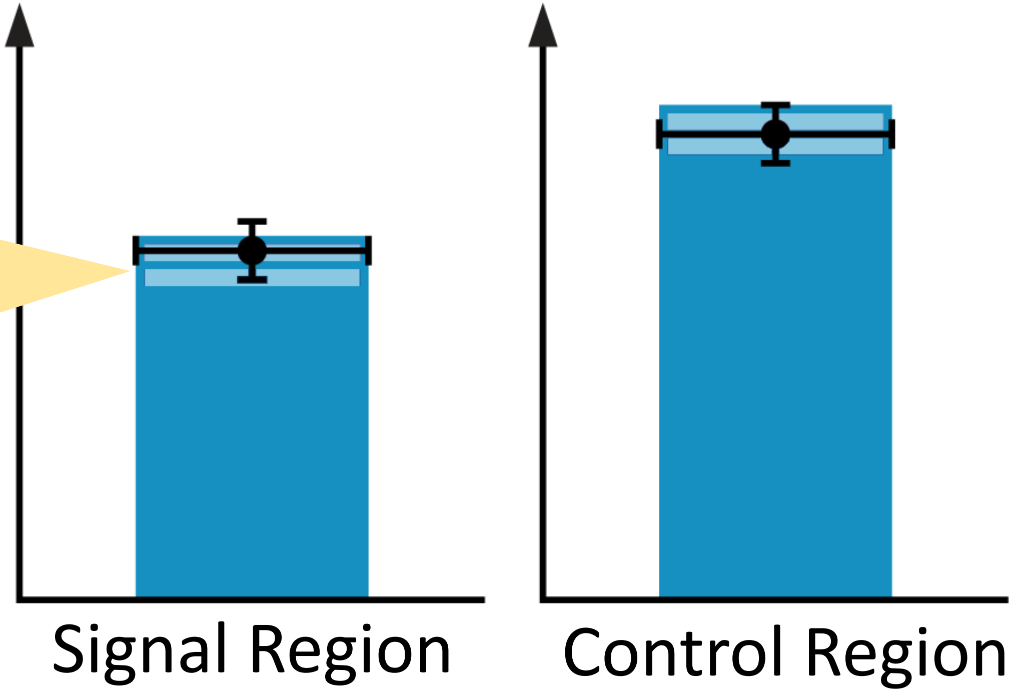
Example of a systematic uncertainty: Jet energy
How accurately do we measure the energy of a jet? What if calibration is off by 10 GeV?
Jet energy shifted up (down) => more (fewer) high-energy jets passing selection => increase (decrease) expected yield

Simplified Background Estimation

This technique also strongly constrains correlated systematic uncertainties

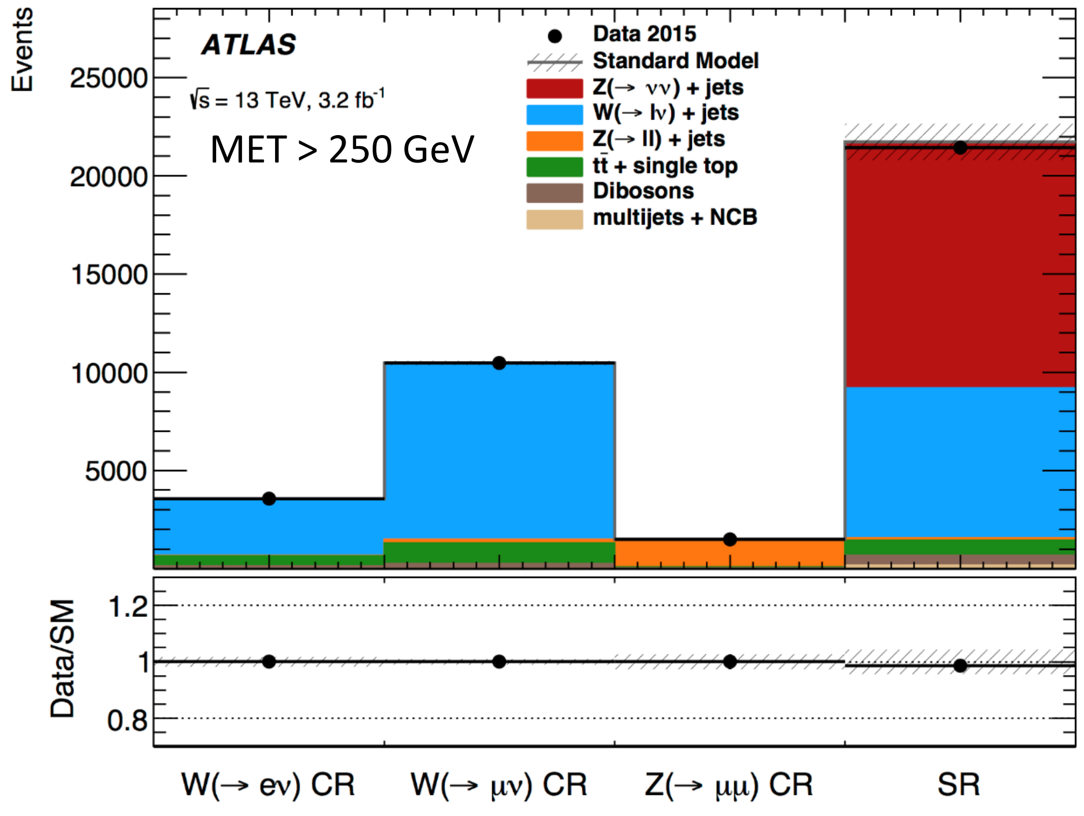
Yield dependence on uncertainty compensated by correction factor *k* in fit
➥ Impact of uncertainty on yield greatly reduced
➥lower total background uncertainty

Full analysis is more complicated
- More background processes
- More control regions
- More uncertainties

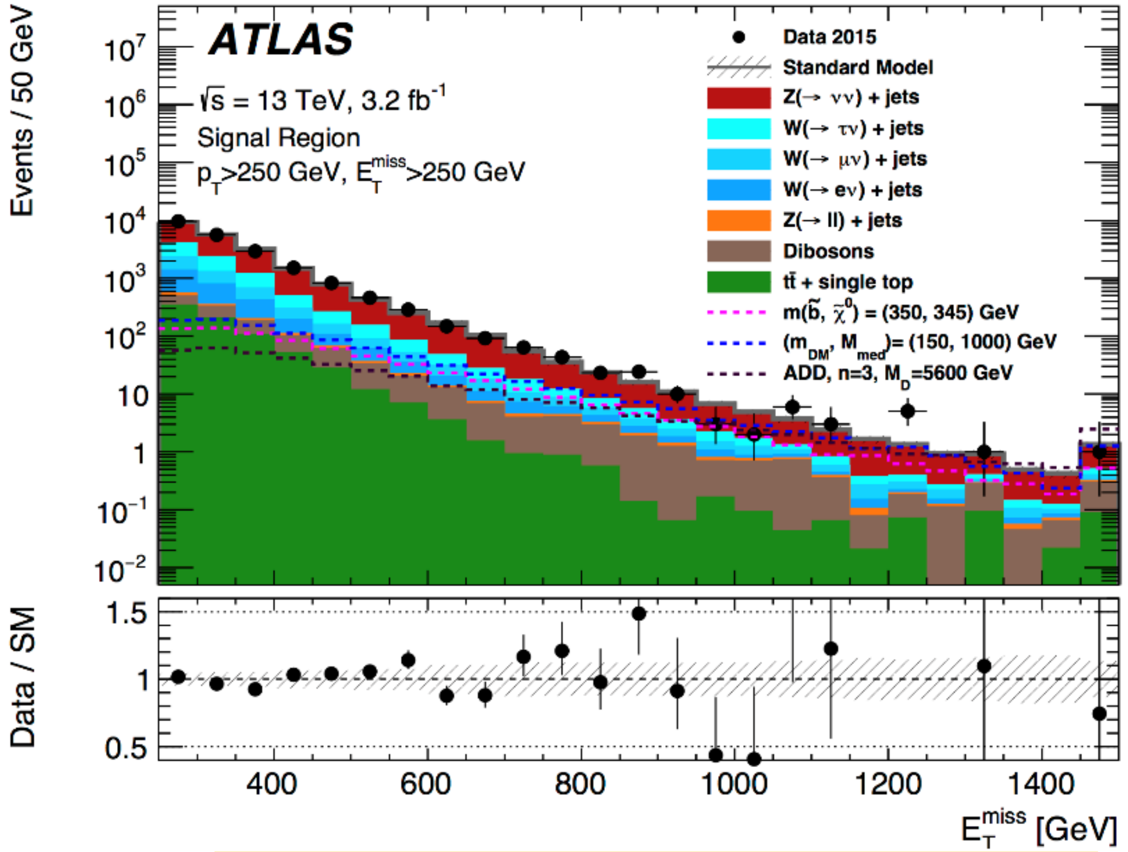


post-fit

Results



Signal & control region yields and compositions



Signal region MET: No excess observed

Time to constrain DM models (set limits)!

Statistical Analysis in Particle Physics

- MET + X searches: a counting experiment
  - Determine expected number of background events and uncertainty
    - Compare to observed number of events: consistent with background-only hypothesis?
    - Is there new physics?
  - Calculate expected number of signal events and uncertainty
    - Is a signal excluded?
- **Need a mathematically rigorous way to make these statements**
  - Quantify results in terms of probabilities for theoretical models and their parameters
  - Hypothesis testing – vary parameter of signal model:
    - $P(\text{data}|\text{theory}(x=35)) = 3\%$
    - $P(\text{data}|\text{theory}(x=20)) = 5\%$
    - $P(\text{data}|\text{theory}(x=10)) = 20\%$
    - $P(\text{data}|\text{theory}(x=5)) = 70\%$
  - Confidence interval constructed from hypothesis testing:
    - $X < 20$ at 95% C.L.
    - Equivalent statement: less than 5% chance that $X > 20$

Counting Experiment

- Simple counting experiment
  - Data: event count collected (n) in a fixed time frame
  - Theory: the expected distribution for n for repeated measurements
- The observed number n will follow a Poisson distribution
- Two relevant hypotheses:

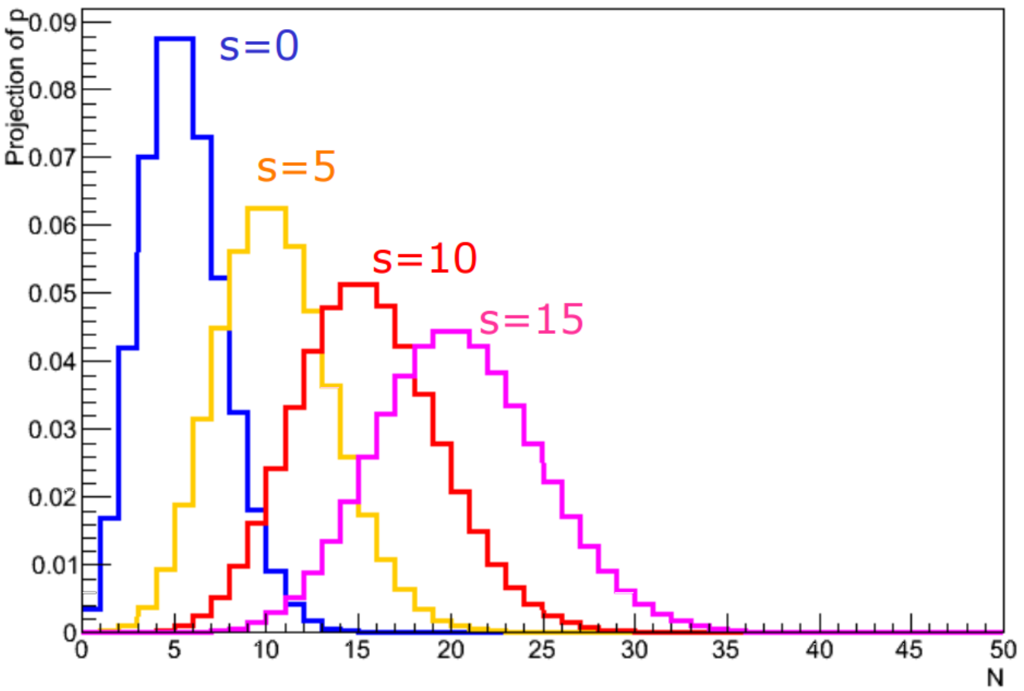
$$P(n|b) = \frac{b^n}{n!}e^{-b}$$

$$P(n|s+b) = \frac{(s+b)^n}{n!}e^{-(s+b)}$$

Background-only: all events are from background

Signal + background : events are a mixture of signal an background

For the following example assume an exact prediction for the number of background events: b = 5

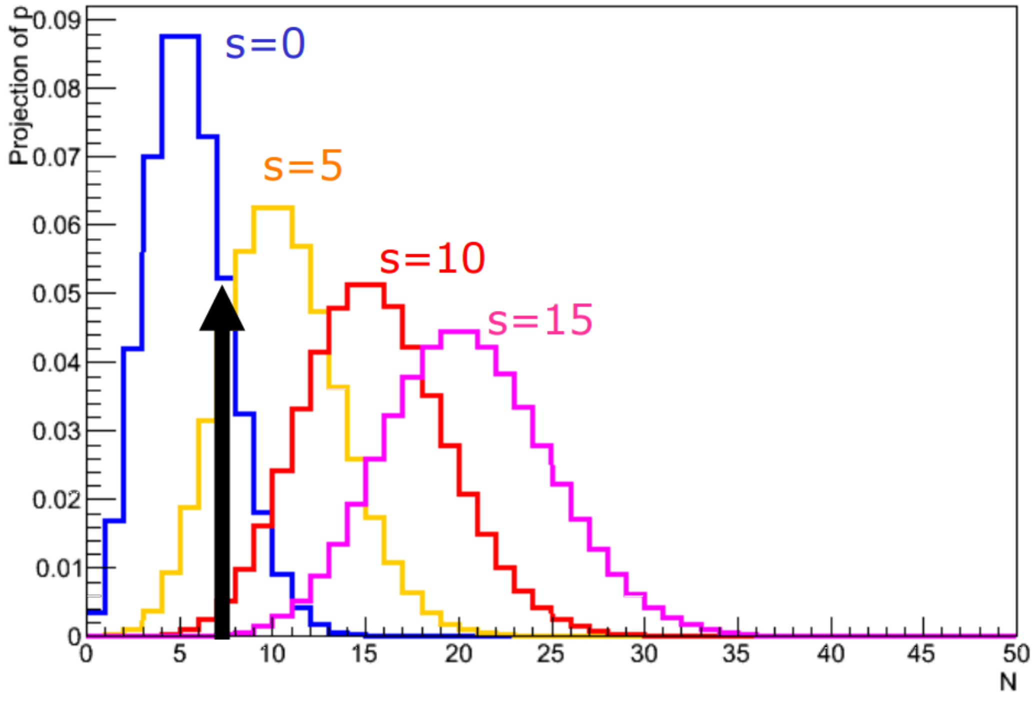


Probability distributions for different signal models

Counting Experiment

Make a measurement:

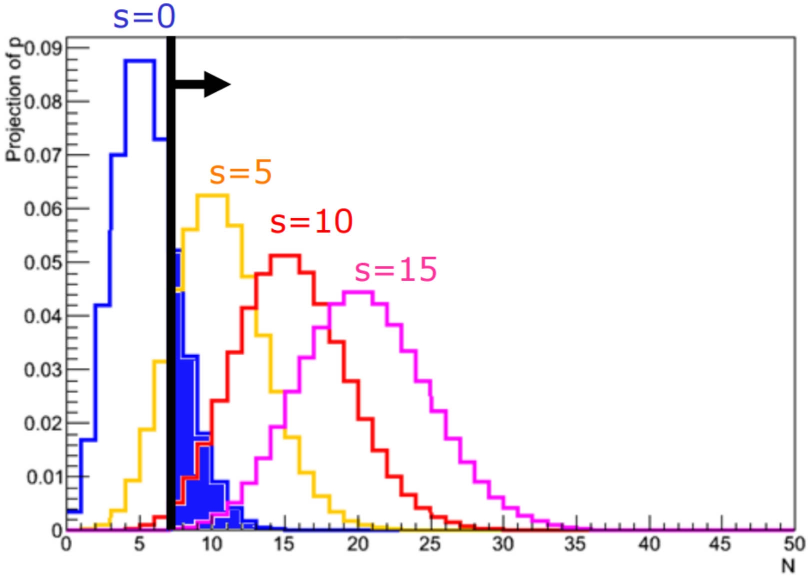
- $n = n_{obs} = 7$
- In this example P(D|T) can be trivially calculated:
  - $P(n_{obs}=7|T(s=0)) = \text{Poisson}(7;5) = 0.104$
  - $P(n_{obs}=7|T(s=5)) = \text{Poisson}(7;10) = 0.090$
  - $P(n_{obs}=7|T(s=10)) = \text{Poisson}(7;15) = 0.010$
  - $P(n_{obs}=7|T(s=15)) = \text{Poisson}(7;20) = 0.001$

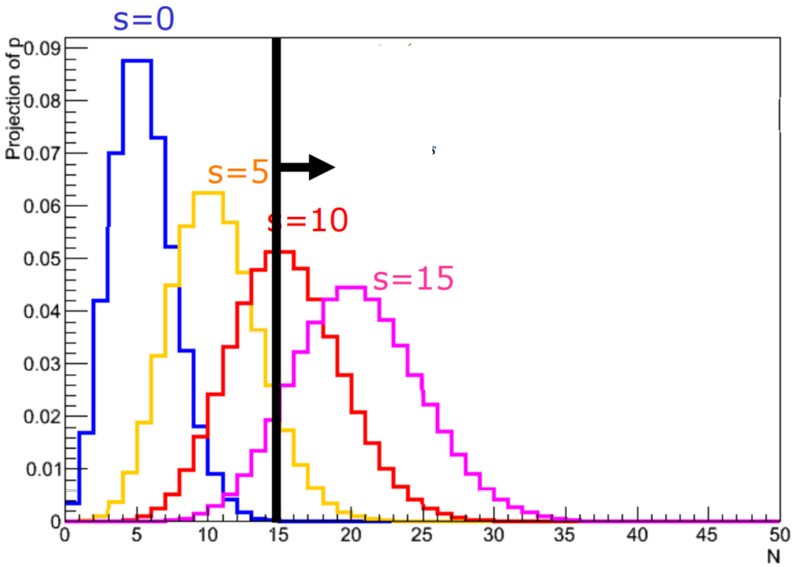


Probability distributions for different signal models

Discovery

p-value of background hypothesis p_b => mathematical formulation for discovery





- If the background-only hypothesis is true, what fraction of future measurements would result in 7 or more events?
- Integrate PDF from n_{obs} to infinity => 0.23

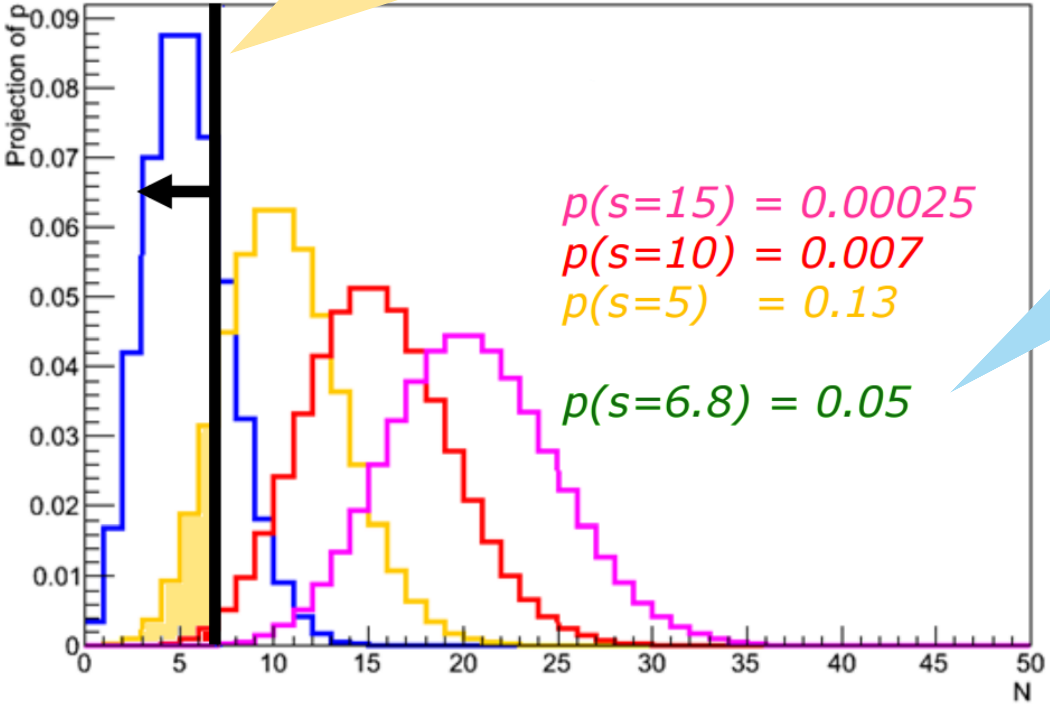
- What if n_{obs} = 15?
- Integrate PDF from n_{obs} to infinity => 0.00022
- This number is usually re-expressed as the odds of a Gaussian fluctuation with equal p-value:
  - 0.00022 => 3.5 sigma

n_{obs}=22 gives $p_b < 2.8\ 10^{-7}$ => 5 sigma => discovery!
Statement that observation is incompatible with bkg-only hypothesis

Upper Limits

Can also define p-values for hypothesis with signal p_{s+b} => used to define exclusion

NB: integration range flipped



Procedure of scanning for the value of s so that p(s+b) = 0.05 is called "Hypothesis Test Inversion"

Express result as value of s for which p_{s+b} = 0.05 => **"s > 6.8 at 95% C.L."**

Actual analysis is much more complicated!
=> Multiple observations in control regions and signal regions, many (often correlated) uncertainties, etc

Software for Statistical Analysis

- LHC uses ROOT for data analysis
  - ROOT: object-oriented analysis environment
- RooFit: Add-on package to ROOT
  - Designed as a particle physics data analysis tool
  - Toolkit for modeling the expected distribution of events in a physics analysis
  - Object-oriented
    - Every variable, data point, function, PDF, etc is represented in a C++ object
  - Models can be used to perform unbinned maximum likelihood fits, produce plots, generate "toy Monte Carlo" samples for various studies, etc

Mathematical concept		RooFit class
variable	x	RooRealVar
function	$f(x)$	RooAbsReal
PDF	$f(x)$	RooAbsPdf
space point	$\vec{x}$	RooArgSet
integral	$\int_{x_{min}}^{x_{max}} f(x)dx$	RooRealIntegral
list of space points		RooAbsData

Software for Statistical Analysis

Example analysis framework

On the Grid

Data

Monte Carlo simulation + Variations due to systematic uncertainties

ROOT-based software designed to run on grid sites and combine output

- Apply calibration & corrections
- Veto “bad” events
- Reduce event size => only need events going into signal and control regions

On clusters or local machines

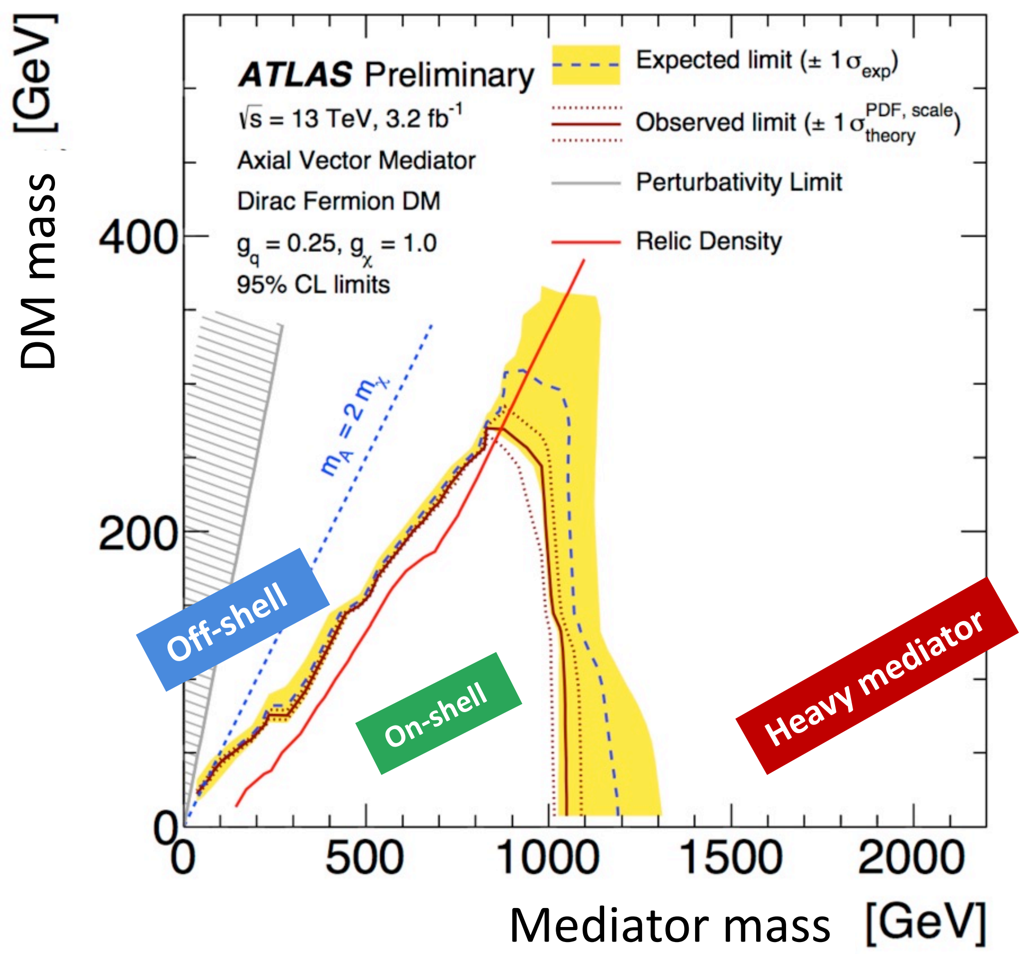
HistFitter & RooFit build the likelihood

- Likelihood describes the probability of observing the number of events seen in the signal region for some set of parameters
  - Used to perform statistical tests
- Inputs: signal and control region yields, systematic uncertainties, signals, etc
- Build mathematical model & fit correction factors to data
- Output: RooFit “workspace”

Hypothesis test inversion with RooFit

- Use likelihood to determine if different signal models are excluded

MET+Jet Dark Matter Limits

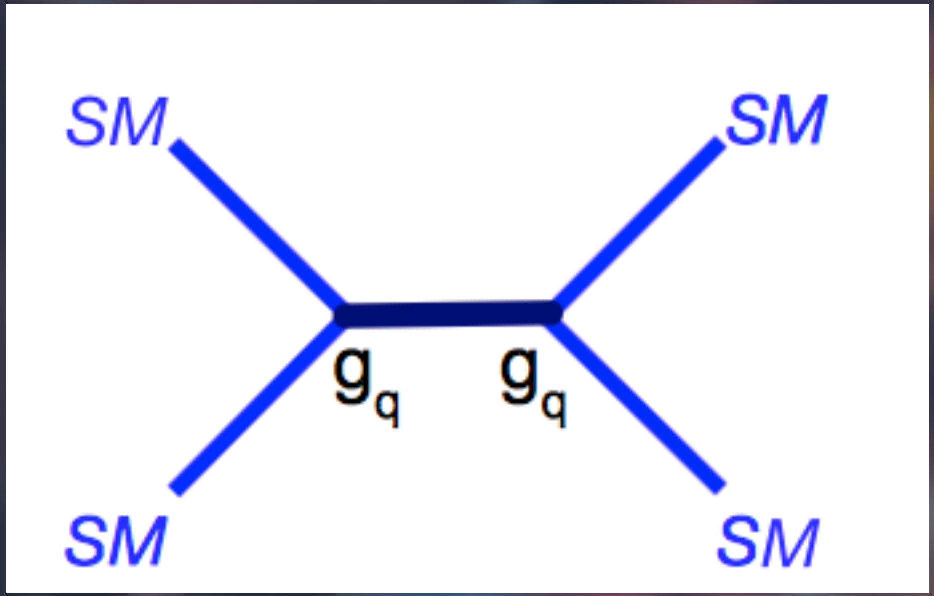


- Present limits as a function of DM & mediator mass
  - Axial vector mediator
  - Fixed values of g_{DM} & g_{SM}
- DM excluded up to 250 GeV for 1 TeV mediator
- On-shell region
  - Mediator light enough to be produced in p-p collisions
  - DM light enough to decay from mediator
  - LHC sensitivity
- Off-shell region
  - 2 mχ > mA ➡ Mediator decay to DM suppressed
  - Relic DM underproduced ➡ annihilation xsec too high
- Heavy Mediator region
  - Mediator too heavy for LHC production
  - Relic DM overproduced ➡ annihilation xsec too low

Recap I: Searching for DM Particles

- Dark matter at the LHC: MET + X
  - Cannot see invisible particles => look for DM + "X"
    - X = gluon/jet, photon, etc
  - Event topology: missing transverse momentum and X
- Signal samples
  - Need to know characteristics of DM signal in order to find it
  - True model of DM unknown => generate as many models as possible
- Background estimation
  - Many SM processes also make MET + X events!
  - Need to understand backgrounds in order to find (or exclude) signal
  - Fit background processes to data using signal and control regions
- Statistical analysis
  - Build analysis likelihood
    - How likely was our observation N, given our understanding of signal and background?
  - Did we find new physics? Can we say that new physics is excluded?
- Questions?

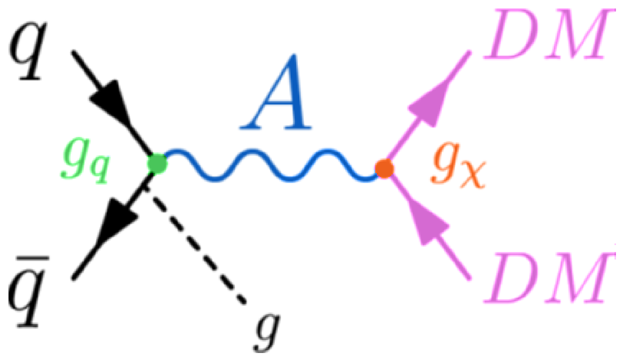
Searching for Dark Matter Mediators

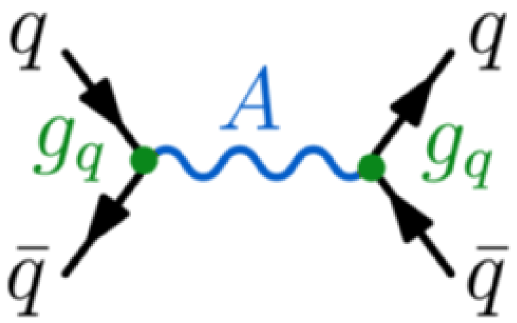


- Introduction to a search for new particle resonances at the LHC
- Software to record and reconstruct data at the LHC

Dark Matter Mediators

- The mediator coupling quarks to DM may also decay back into SM particles
  - No invisible particles in final state => full kinematics can be reconstructed
    - Can look at invariant mass distribution of dijet, dielectron, dimuon, etc
    - Signal should appear as "bump" at ~ mediator mass
- Decay to quarks => "dijet" limits can also constrain DM simplified models
  - Exclusion power depends on the model couplings:



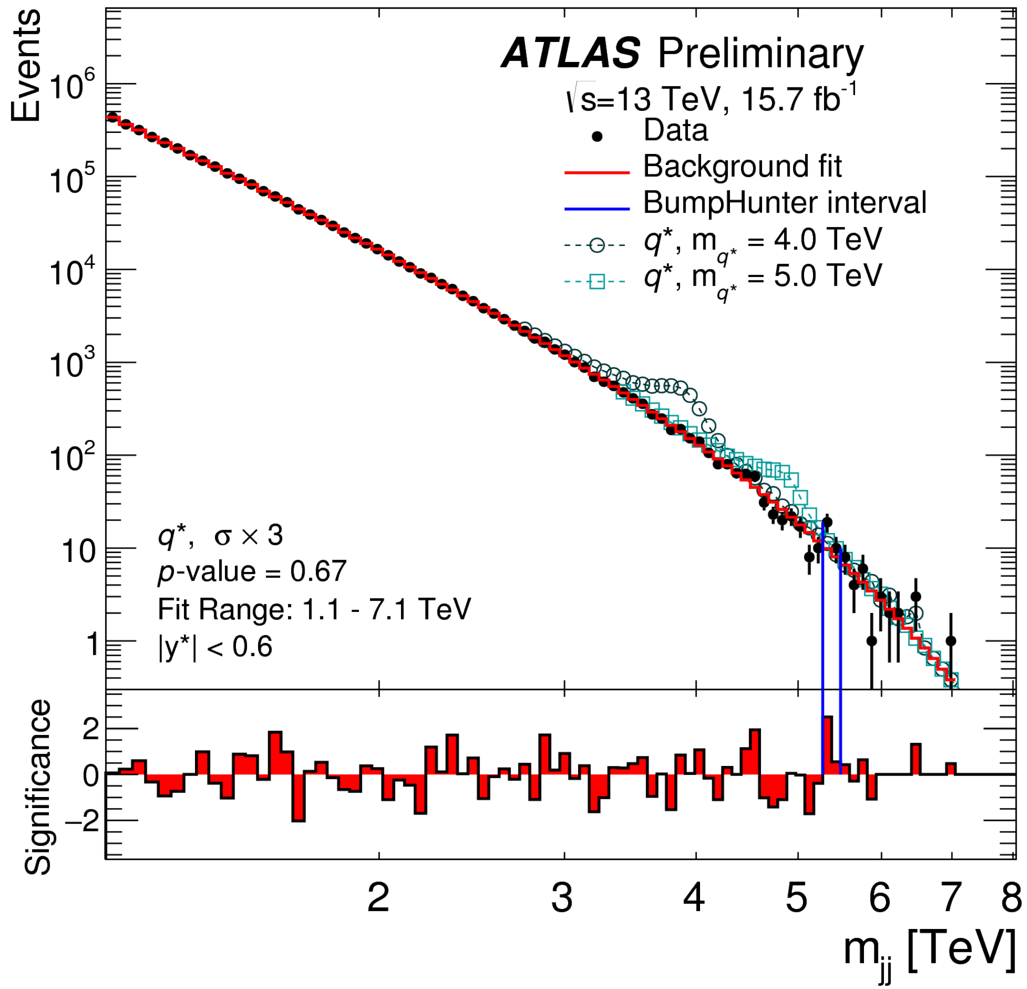


MET+jet cross section scales with $g_q^2 g_x^2$

Dijet cross section scales with g_q^4

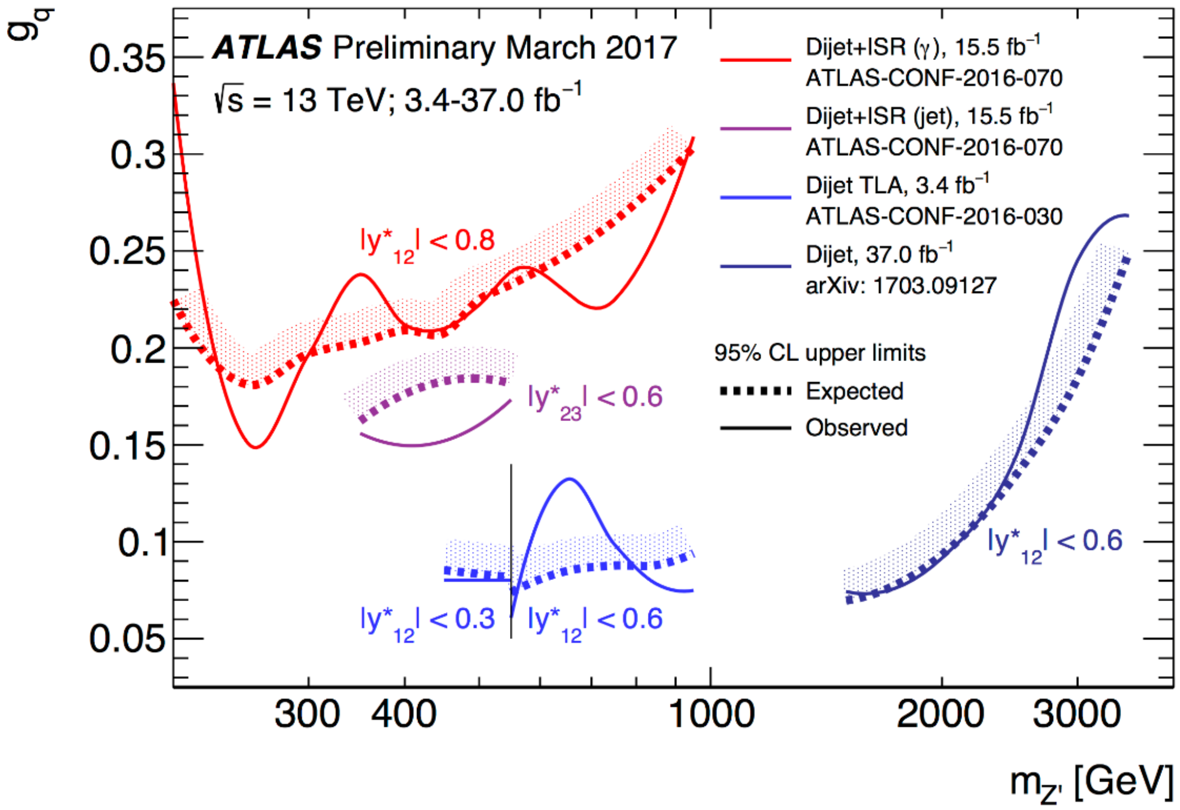
Dijet Resonances

- Search for resonances in dijet invariant mass spectrum
- Completely data-driven analysis
  - Parameterize smooth qcd background
  - Search for signal– deviation from expected smoothly falling background
- High end of mass range limited by statistics
  - Fewer jets at higher energies
- Low end of mass range is limited by the trigger selection
  - Filtering to reduce event rate
  - Only keep all events with jets above a certain energy threshold (~500 GeV)



Dijet Resonances

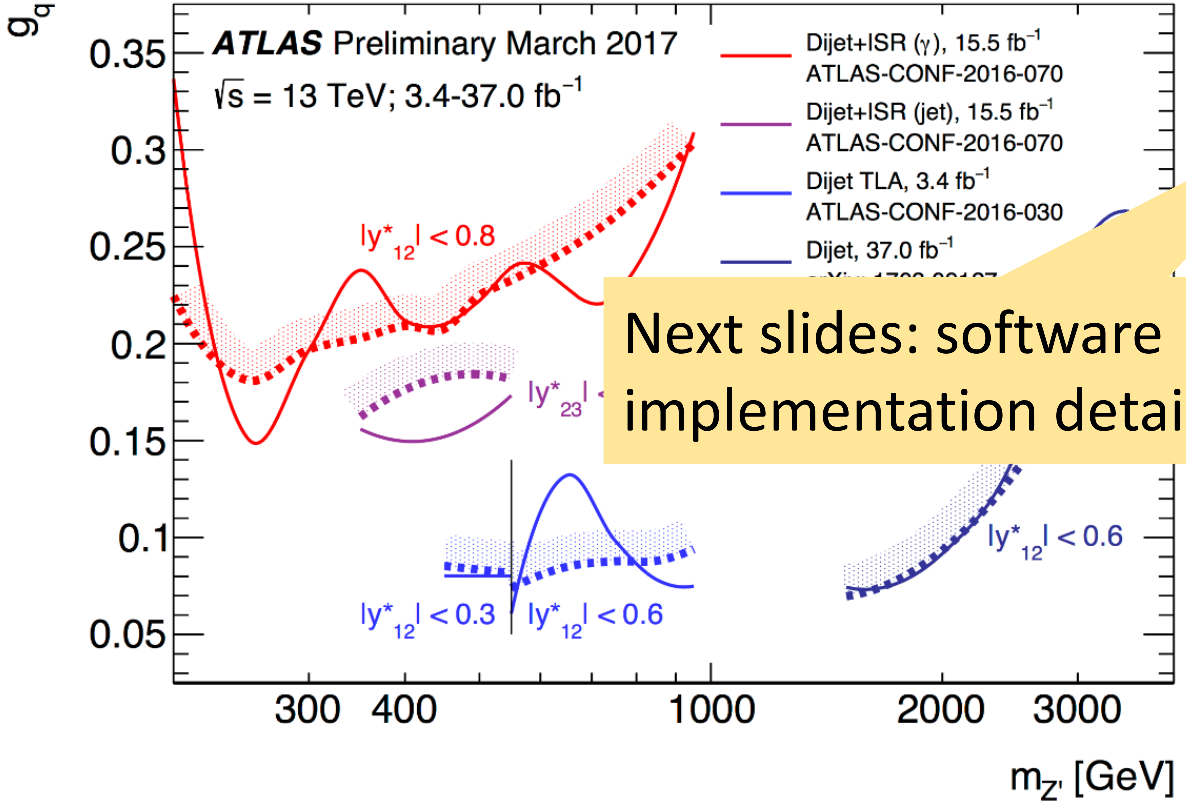
Different classes of dijet searches targeting different mass ranges



- Nominal dijet search
  - Search for resonances in dijet invariant mass spectrum with both 8 TeV and 13 TeV data
- Trigger-level dijet search
  - Data scouting: only output partial event information
  - For jets only need high-level calorimeter information
    - No need for tracking, muons, etc
  - Lower event size => smaller bandwidth => lower trigger threshold
- Dijet + ISR search
  - Jet pair recoiling against an ISR photon or gluon
  - ISR triggers event => jet pair invariant mass can be smaller

Dijet Resonances

Different classes of dijet searches targeting different mass ranges

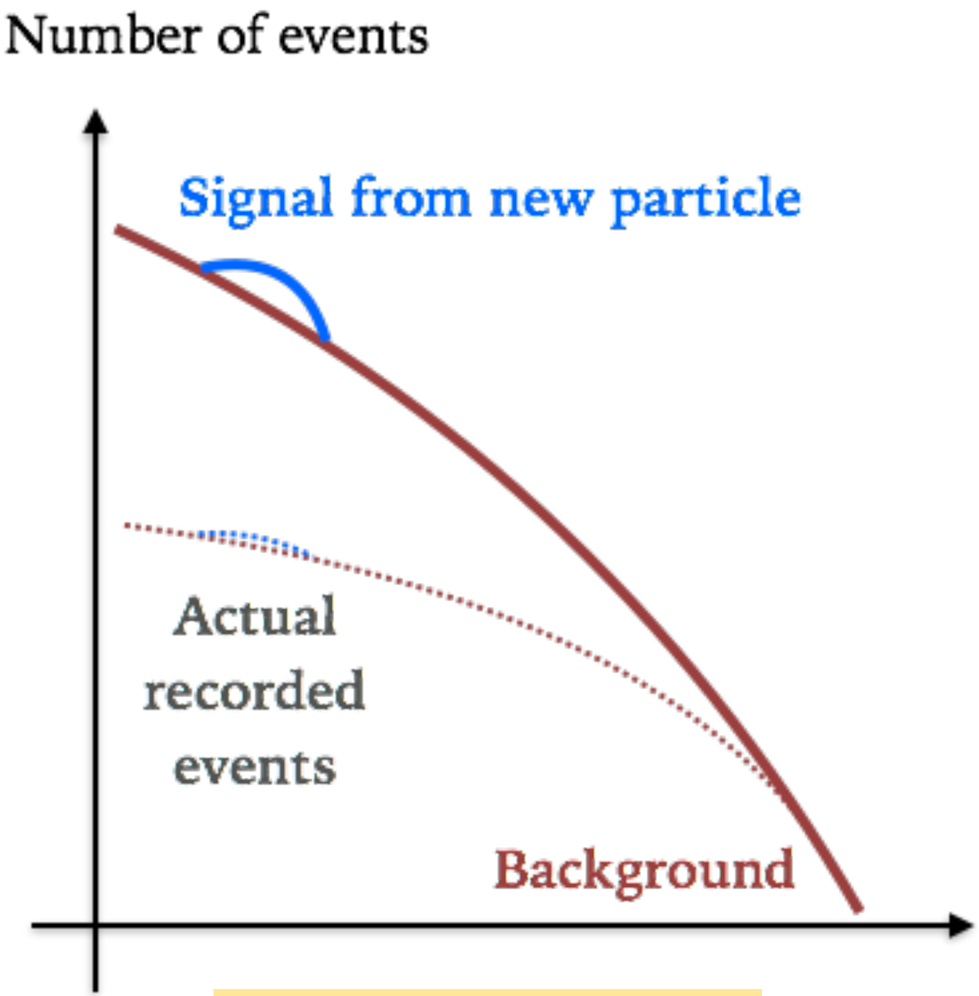


Next slides: software implementation details

- Nominal dijet search
  - Search for resonances in dijet invariant mass spectrum with both 8 TeV and 13 TeV data
- Trigger-level dijet search
  - Data scouting: only output partial event information
  - For jets only need high-level calorimeter information
    - No need for tracking, muons, etc
  - Lower event size => smaller bandwidth => lower trigger threshold
- Dijet + ISR search
  - Jet pair recoiling against an ISR photon or gluon
  - ISR triggers event => jet pair invariant mass can be smaller

Low-Mass Resonances

- Cannot record all data from the LHC
  - Would not fit in storage
- Would like to save events with interesting physics
  - DM signals
- End up discarding signal & background equally in high-rate topologies
  - No handle to distinguish signal & background event-by-event
  - Below some energy threshold, “prescale” events
    - Record one event every N, with N large
- Discovery potential for DM mediators reduced in topologies with less energetic objects
  - Relevant example: low-mass jet resonances

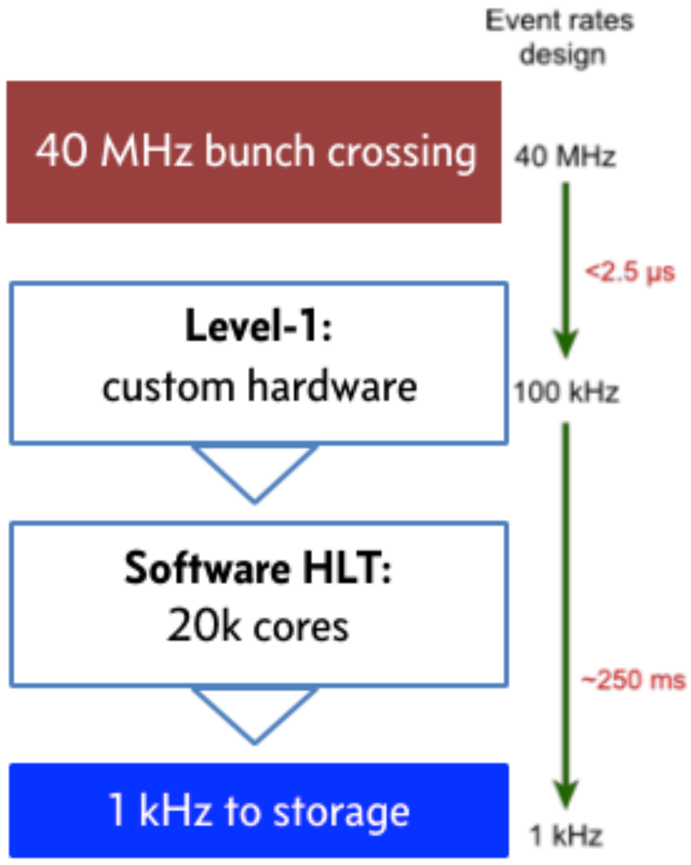


The effect of prescaling

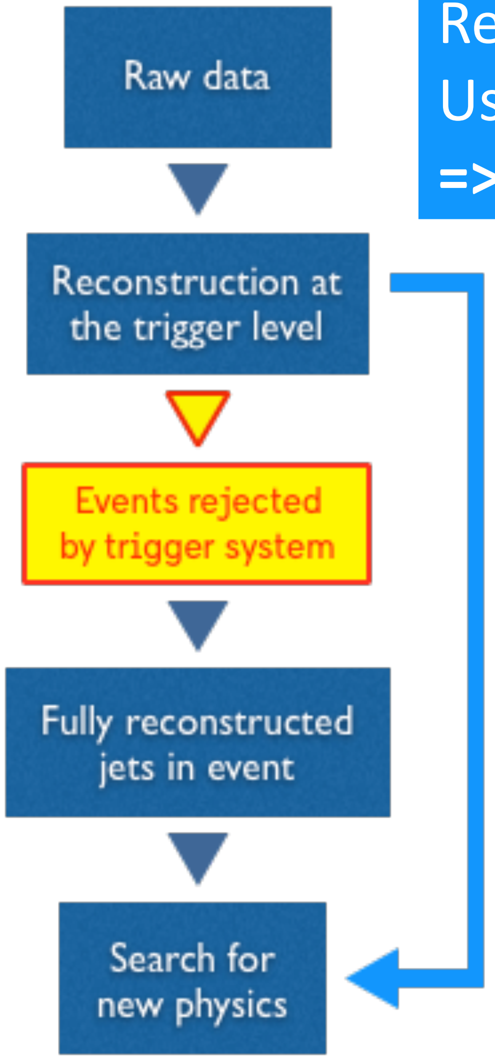
The ATLAS Trigger system

- Data selection at the LHC: multiple levels of event filtering
- Software-based event filtering with the High-Level Trigger (HLT)
- HLT algorithms
  - Fast reconstruction: avoid expensive computations like tracking
  - Accurate reconstruction: full detector data available
- Resources
  - Output rate ~1kHz for full event information
  - Processing time ~300 ms
  - Can’t exceed this! Need a new strategy to save all information in high-rate event topologies
- Partial event building
  - Partial events with data from a subset of the detectors
  - Trigger-level analysis
    - Only write objects created by the HLT
    - Allows higher output rates due to smaller event sizes

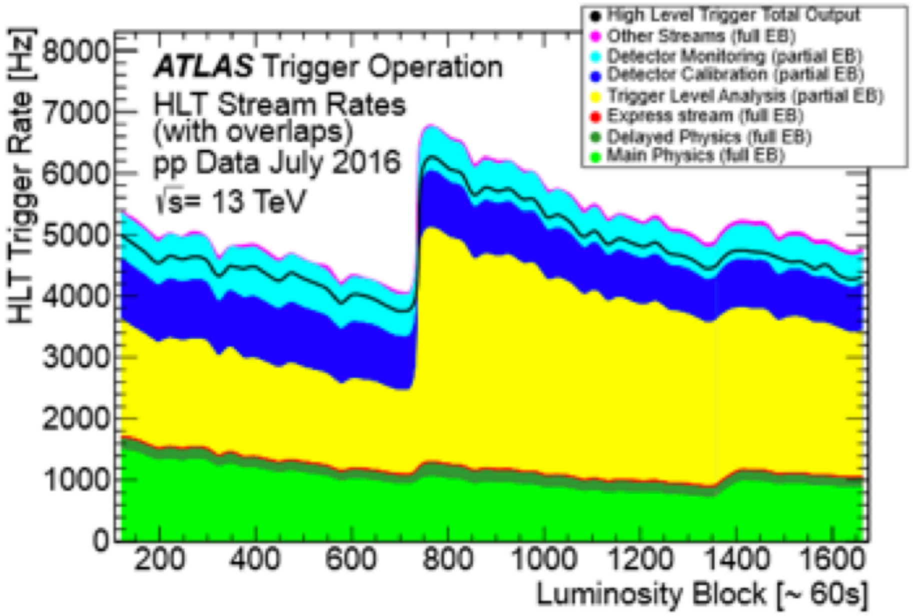
ATLAS trigger system



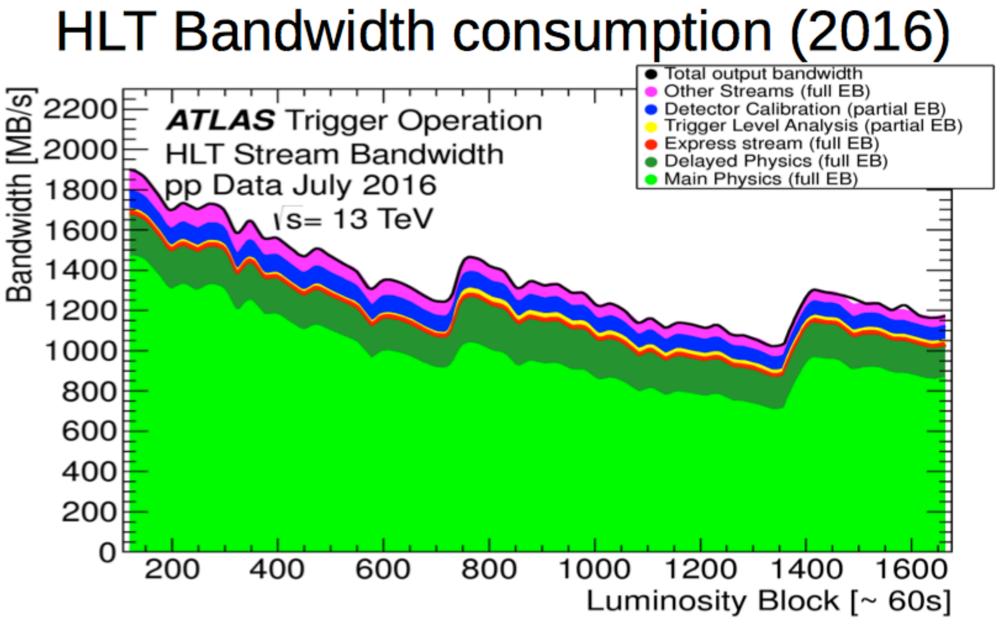
The Trigger-Level Analysis (TLA)



Record only information needed for jet search: **jets**
Use information already available to make decision: **HLT jets**
=> TLA event size reduced to 5% of fully recorded events



Can output more events with smaller size



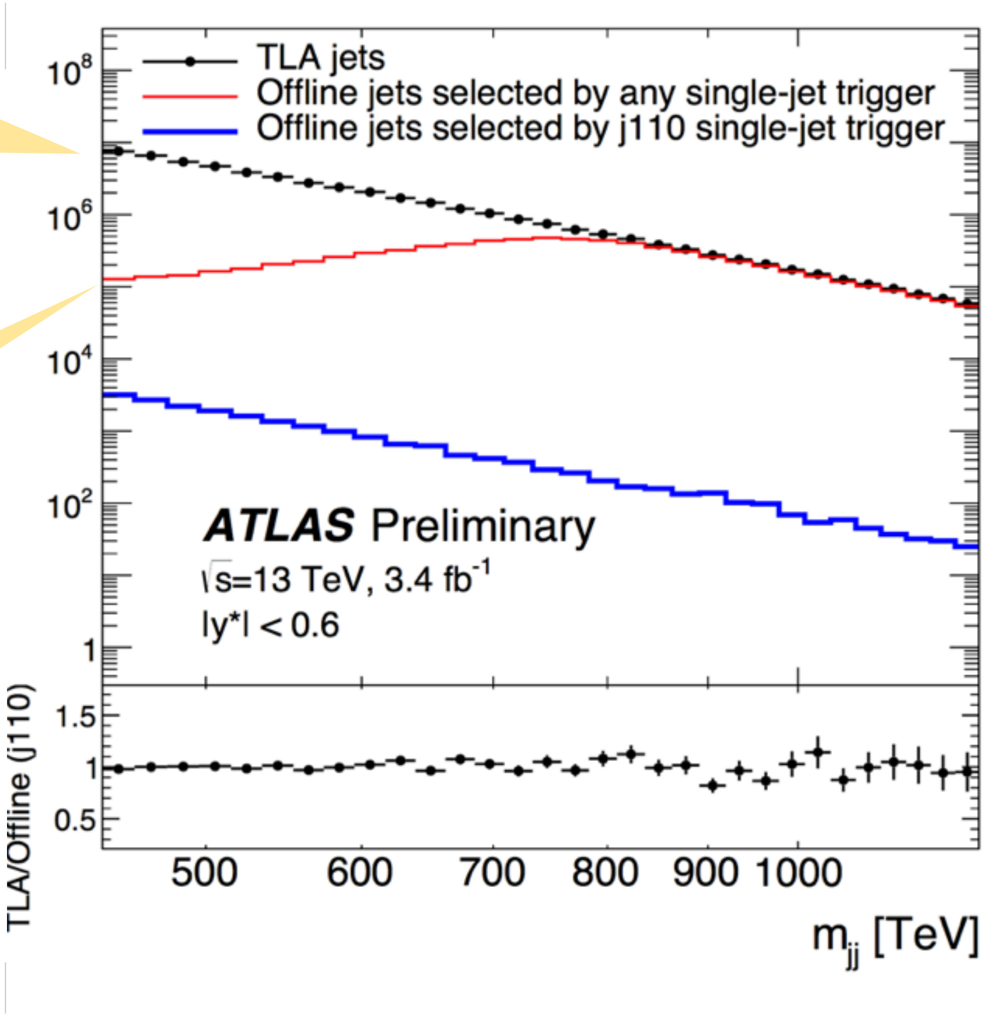
Fraction of total bandwidth remains small

The Trigger-Level Analysis (TLA)

More stats at low mass when using partial events

Data lost from prescaling

Compare TLA jets to jets built with full detector info => unity!

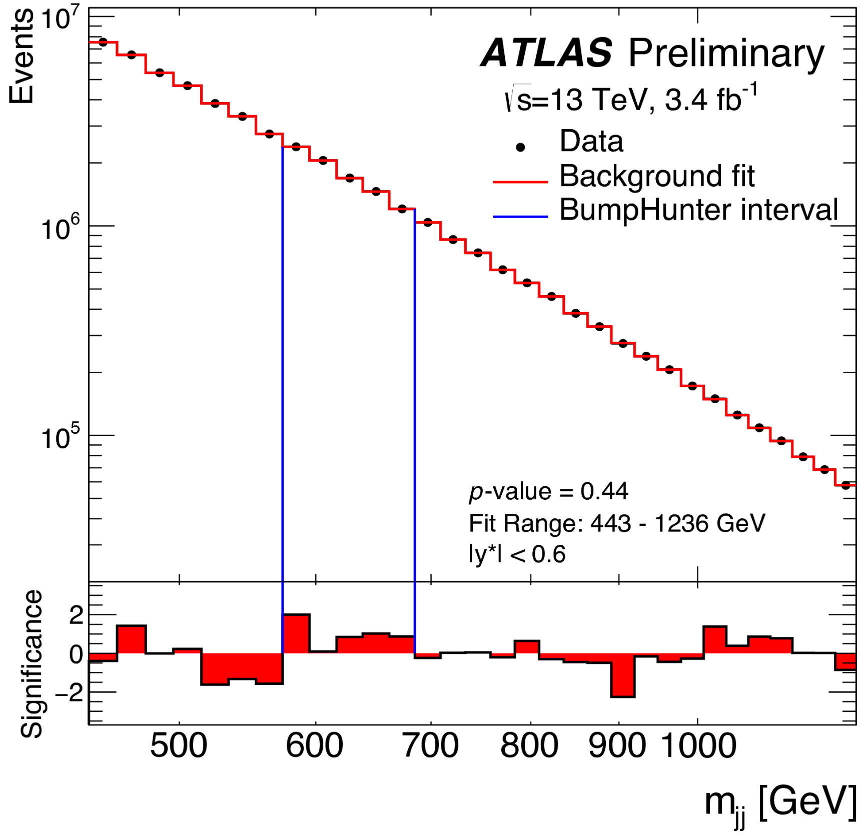


TLA done with specialized reconstruction and calibration and software

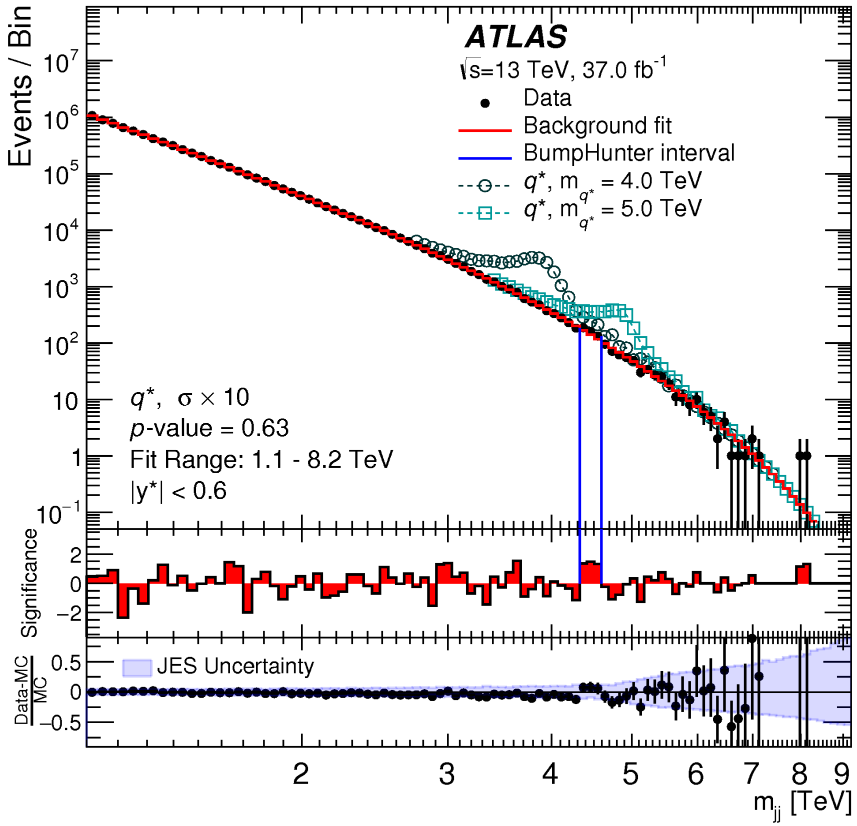
- Data needs to be reconstructed fast by HLT software
- Needs thorough validation before being used for a DM search

The Trigger-Level Analysis (TLA)

TLA dijet analysis



Nominal dijet analysis

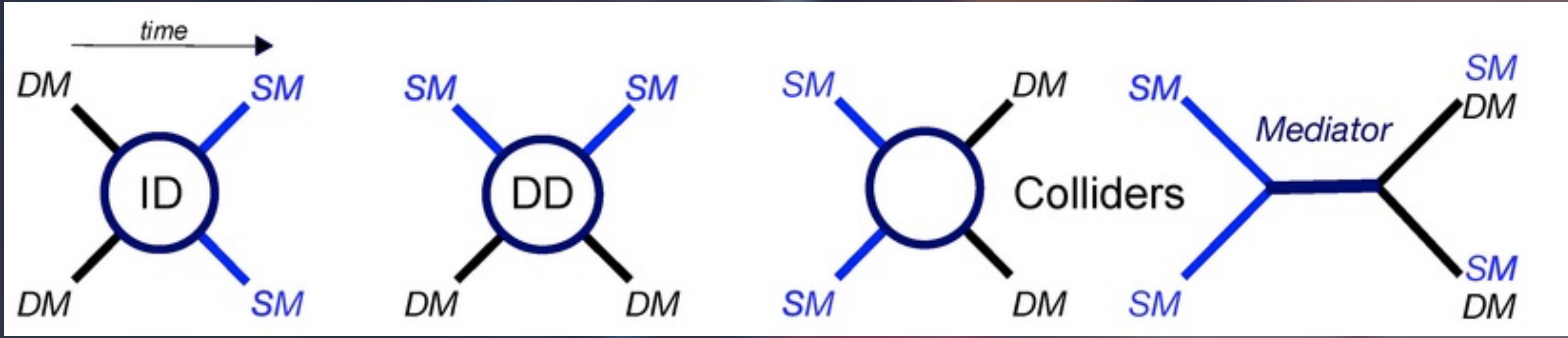


Data scouting extends search range

Recap II: Searching for DM Mediators

- Can also look for evidence of DM by searching for the mediator decaying into SM particles
  - Possible signature: dijet resonance
- Dijet search range limited by event rate & filtering
- Trigger-level dijet analysis
  - Bypass event rate limit by only saving partial event information
  - More events at a fraction of the bandwidth
  - Greatly extend search range
- Questions?

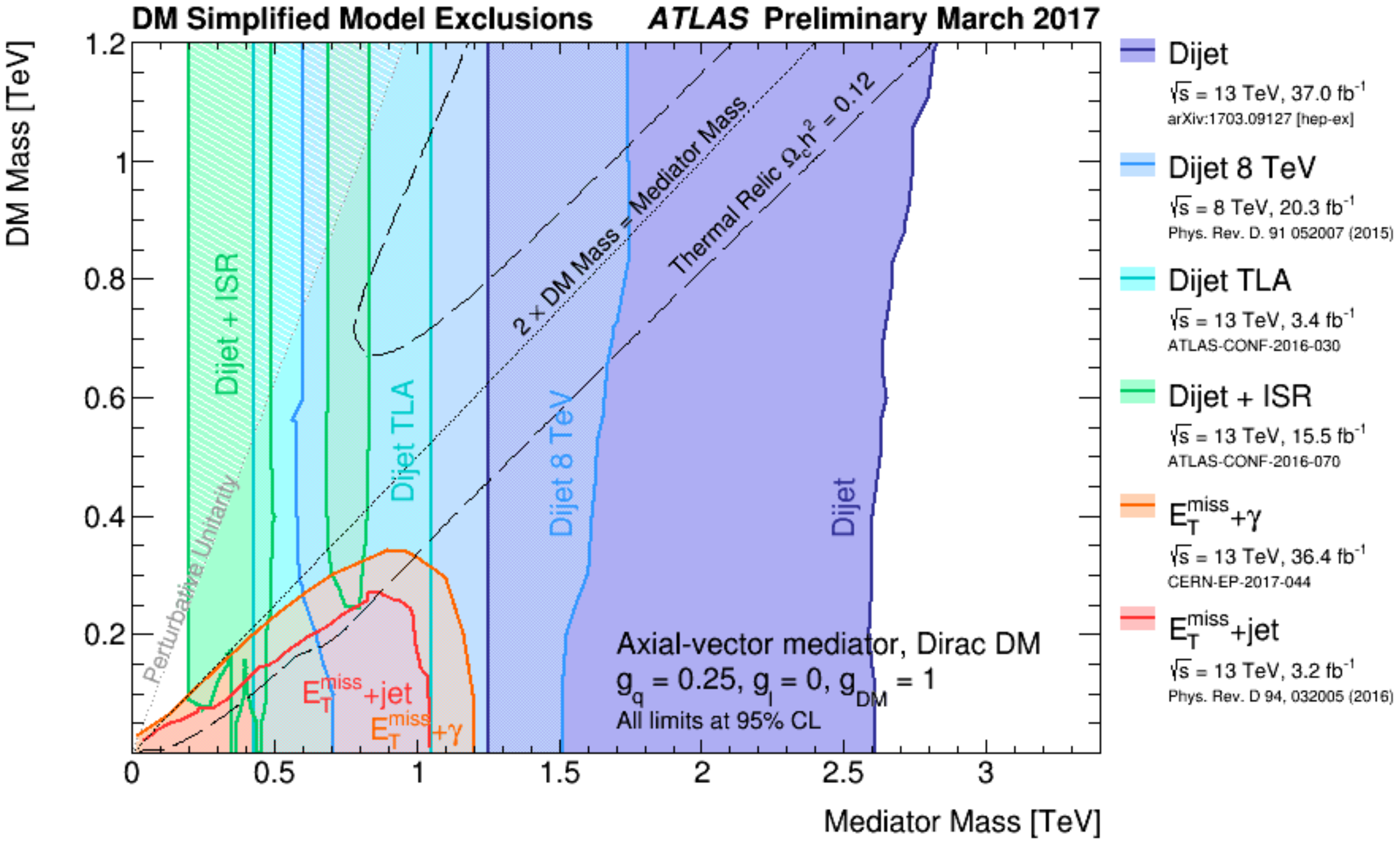
Contextualizing Dark Matter Searches



- How to combine and compare results

LHC DM Exclusions

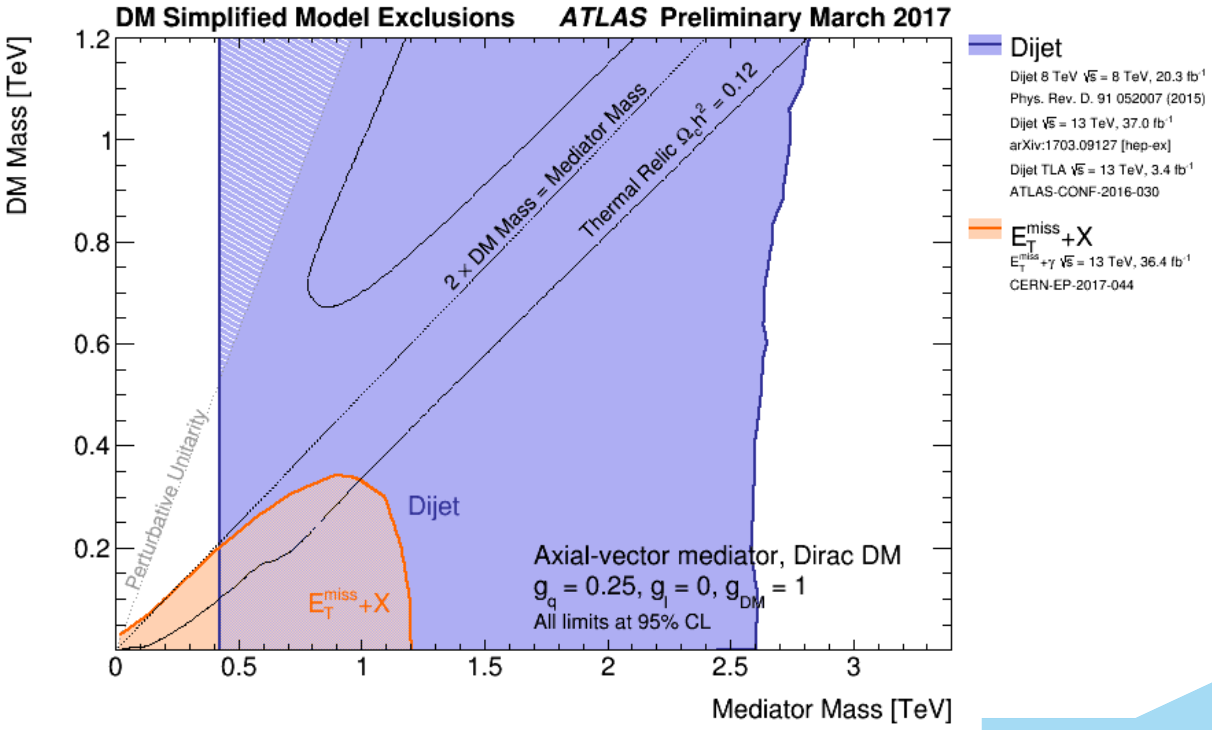
- Can compare constraints from different LHC searches
  - Mediator mass-DM mass plane
  - Dijet searches
    - Strong limits on mediator mass
    - Independent of DM mass in the off-shell region
  - MET+X searches
    - Sensitivity at low masses
- However, exclusion power of different kinds of searches depend strongly on the coupling values in the model…



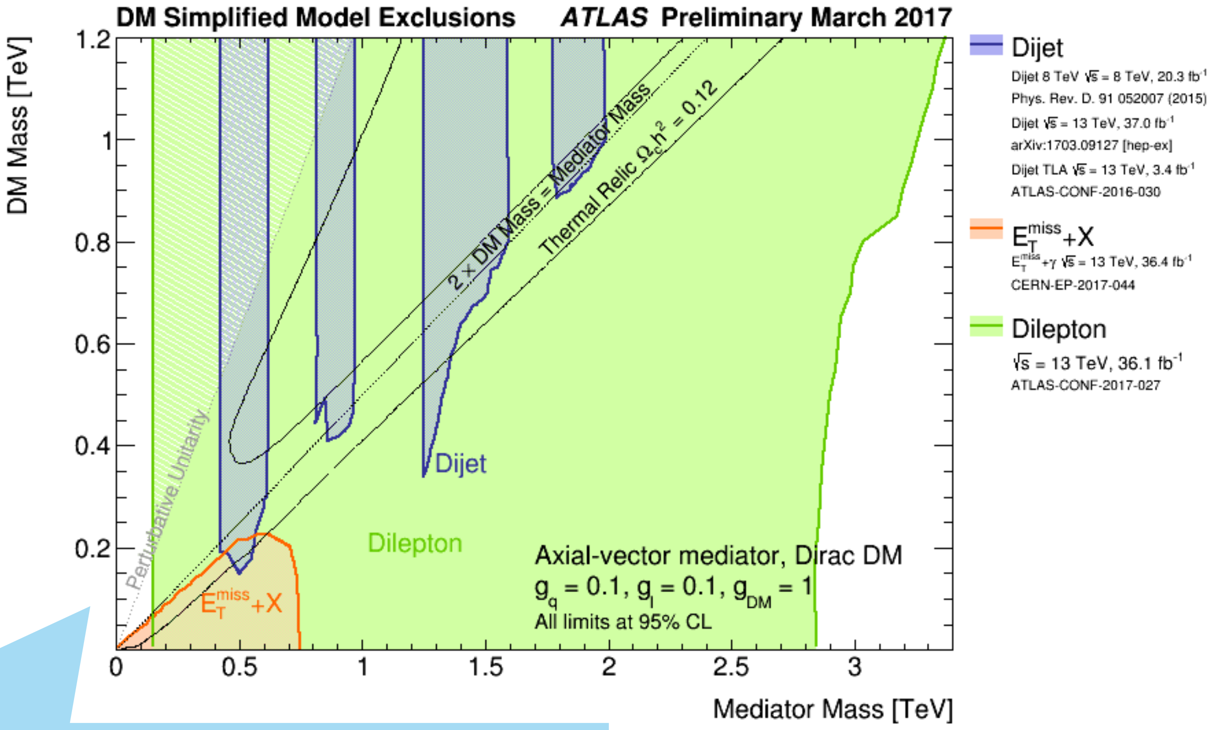
LHC DM Exclusions

- Can look at limits for different couplings
  - Also consider nonzero lepton couplings

Axial vector mediator
$g_q = 0.25$, $g_l = 0$, gDM = 1



Axial vector mediator
$g_q = 0.1$, $g_l = 0.1$, gDM = 1



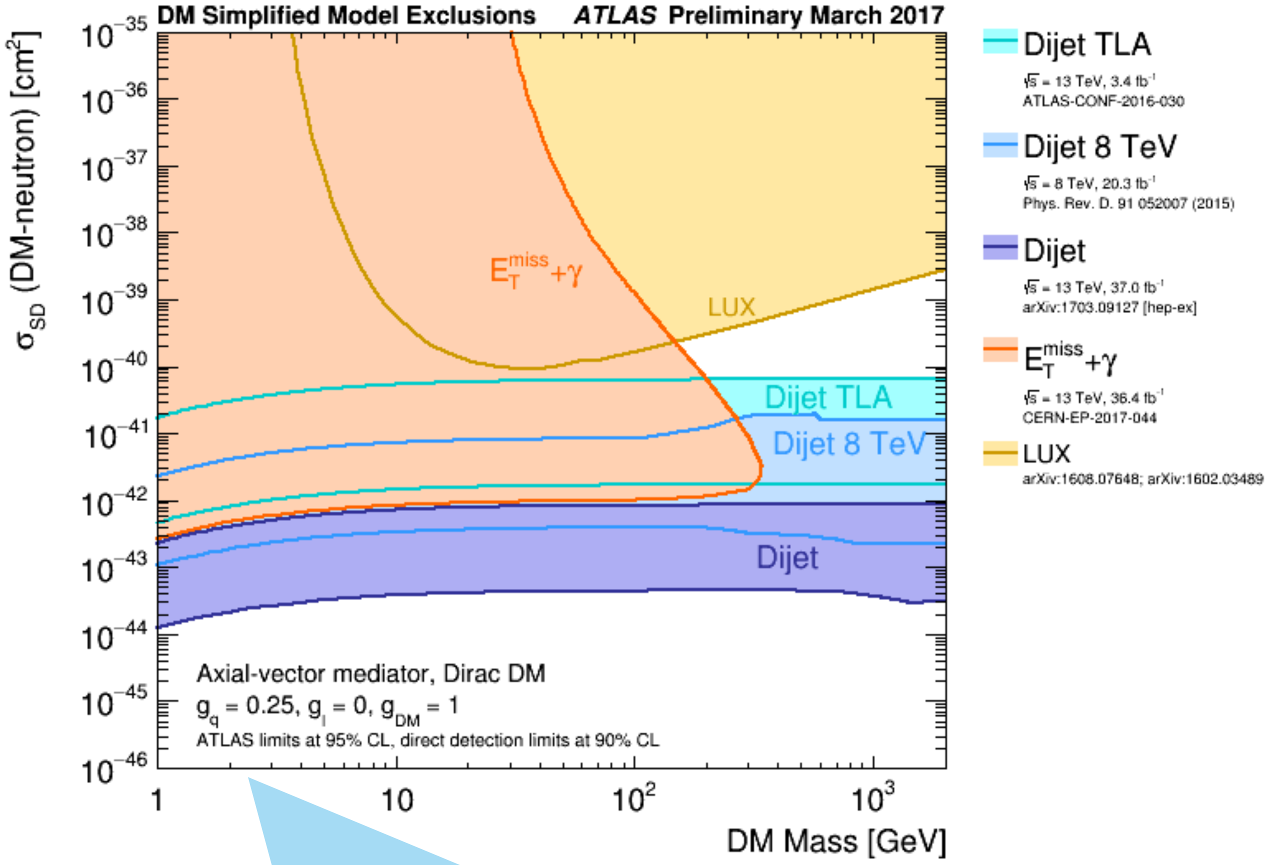
Discontinuous exclusion because coupling value is at limit of dijet search sensitivity

DM-Nucleon Scattering

- Direct detection (DD) of DM
  - DD experiments set a limit on the rate of interactions between local DM halo and atomic nuclei
  - Strong sensitivity to spin-independent interactions
- Translate simplified model collider limits into limits on DM-nucleon effective vertex
  - Constraints on DM & mediator production constrain DM-nucleon scattering
  - Numerically convert (m_{DM}, M_{med}) limit contours to (m_{DM}, $\sigma_{SI/SD}$):

$$\sigma_{\rm SI} \simeq 6.9 \cdot 10^{-41}\ {\rm cm}^2 \left(\frac{g_q g_{\rm DM}}{0.25}\right)^2 \left(\frac{1\ {\rm TeV}}{M_{\rm med}}\right)^4 \left(\frac{\mu_{n\chi}}{1\ {\rm GeV}}\right)^2$$

Limits on spin-dependent (SD) DM-neutron scattering

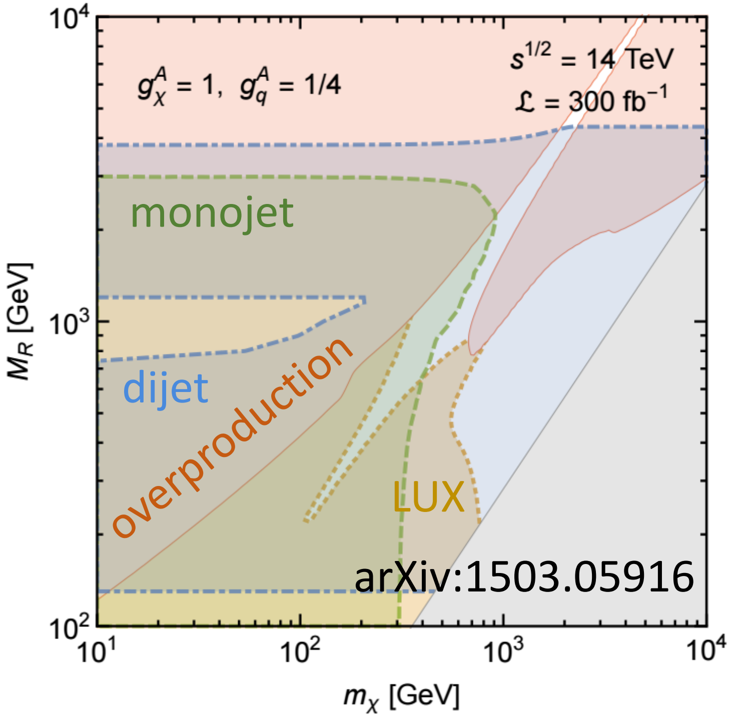


LHC limits are complimentary at low DM mass, where DM is too light to create a nuclear recoil signal

Looking Forward

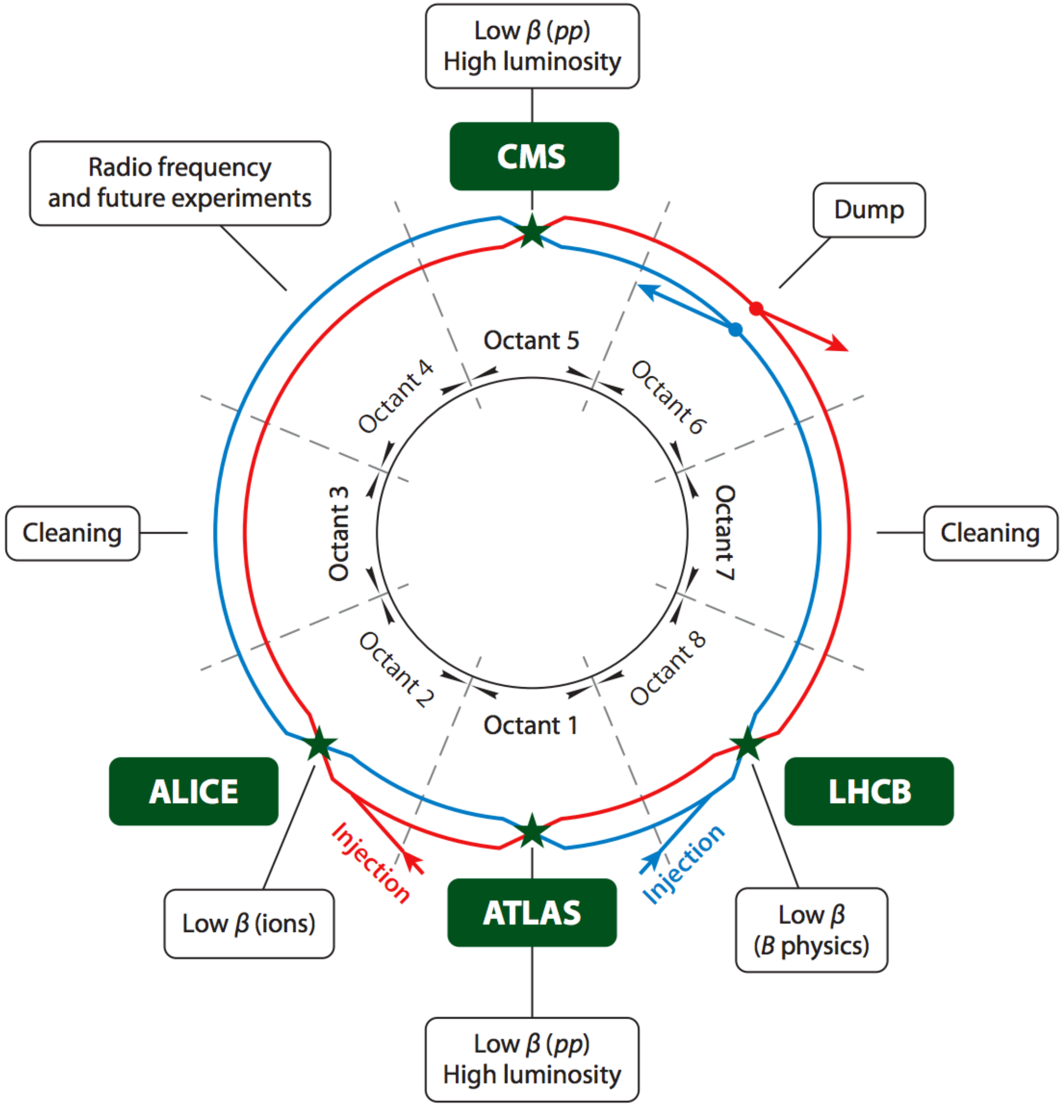
- We expect to gain sensitivity across most of the benchmark simplified model phase space in the next few years
  - If the dark sector interacts with quarks, it will be discovered here
  - If not, constraining this sector will have important ramifications for the future of dark matter detection efforts
- What's next for DM at the LHC?
  - Discovery!?
  - Investigating more complete models of DM?
    - More "correct" theories predicting more nuanced topologies
  - More exotic models of DM?
    - Coannihilating Dark Matter
    - Dark photon/Z'-like mediators

10-year projected limits



Backup

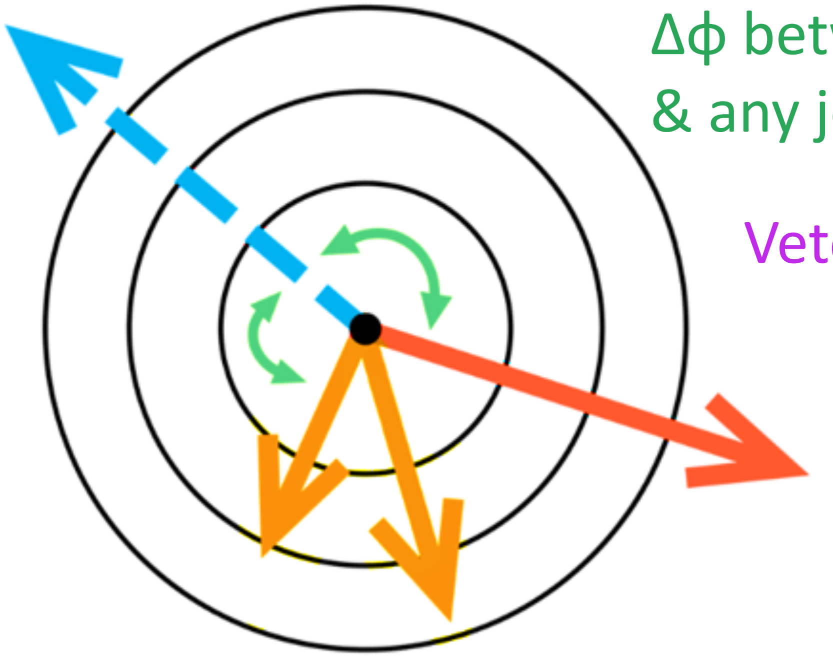
The Large Hadron Collider



MET+Jet Signal Region Selection

MET > 250 GeV

$\Delta\phi$ between MET & any jet > 0.4



Veto e & μ → invert lepton veto to construct three control regions: 1e, 1μ, 2μ

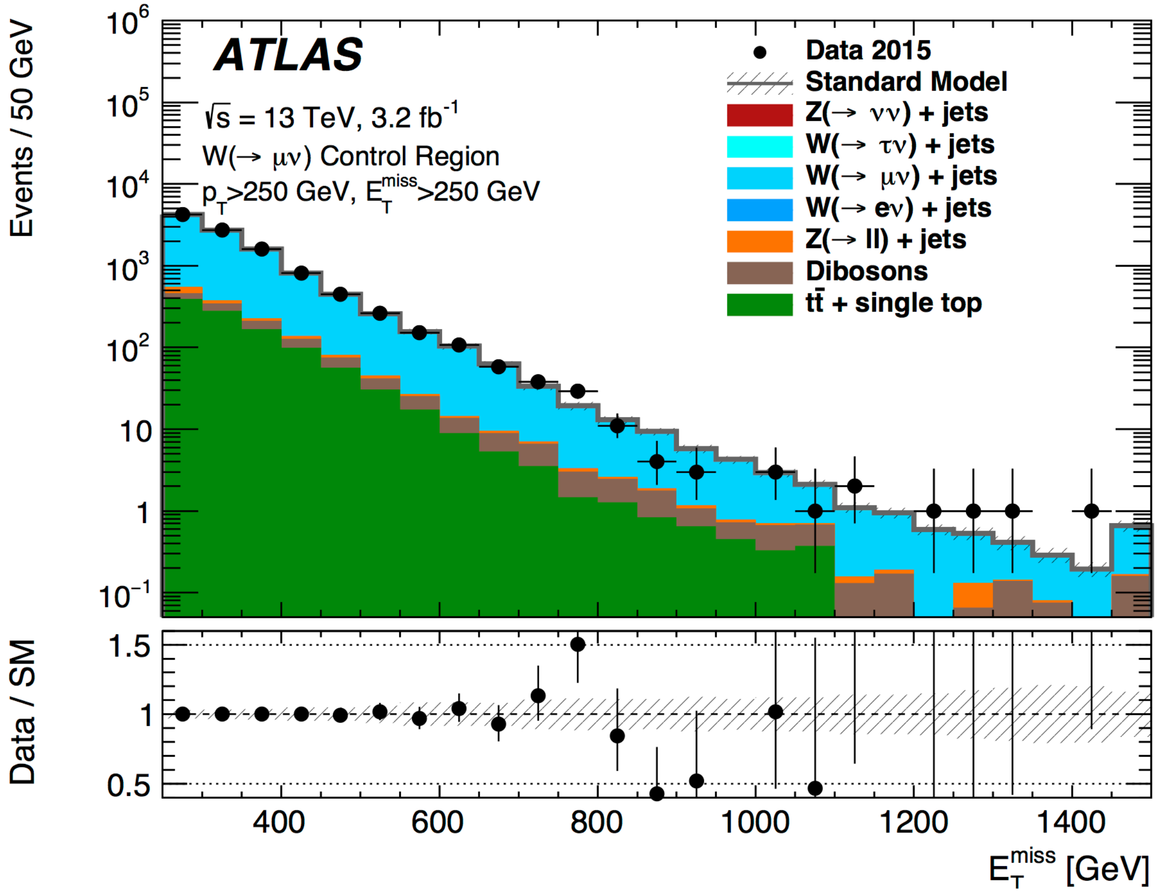
Jet pT > 250 GeV
|eta| < 2.4
Additional quality requirements

Up to 3 other jets
With p_T > 30 GeV

MET + Jet Control Regions

Muon Control Region

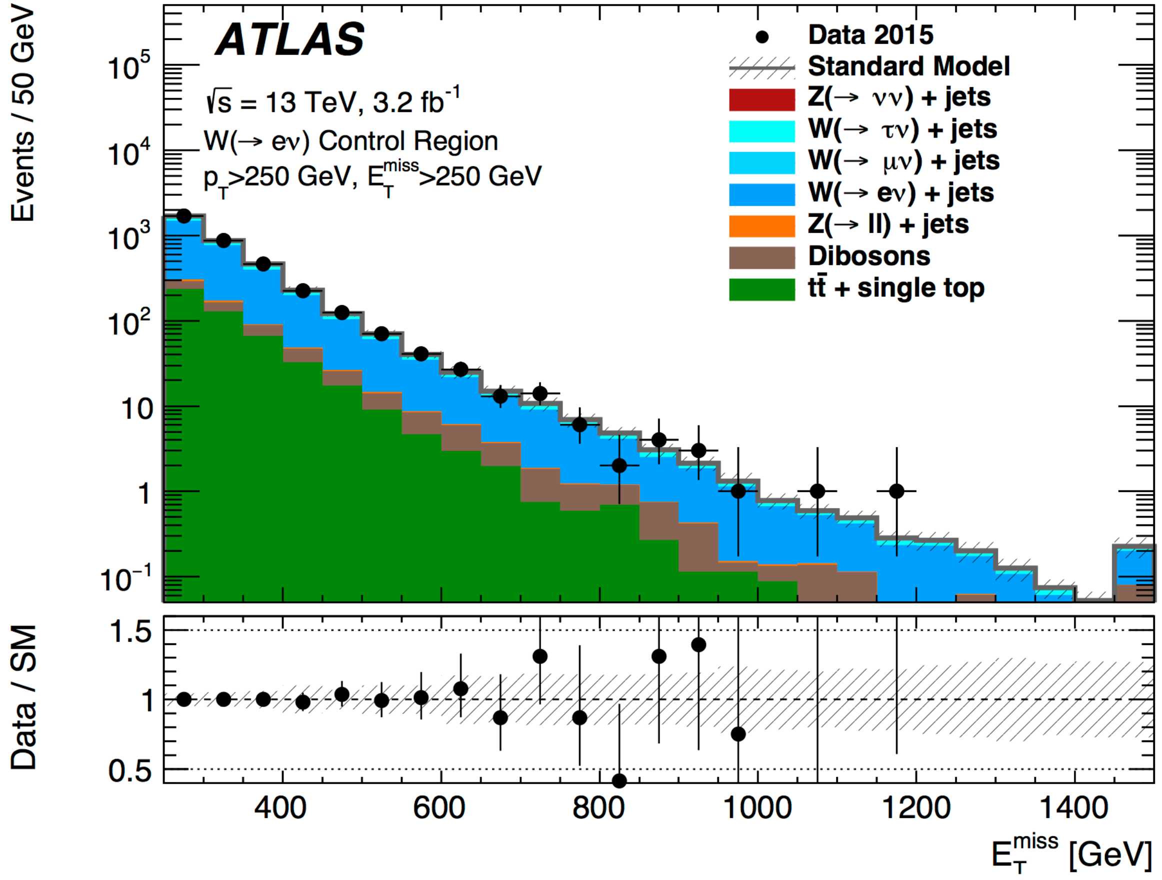
- Optimized to select W → μν+jets events
  - 1 muon, invisible in MET calculation
    - MET ~ W boson p_T
  - $30 < m_T < 100$ GeV
    - Consistent with W
- Constrains W→μν +jets *and* Z→νν+jets



MET+Jet Control Regions

Electron Control Region

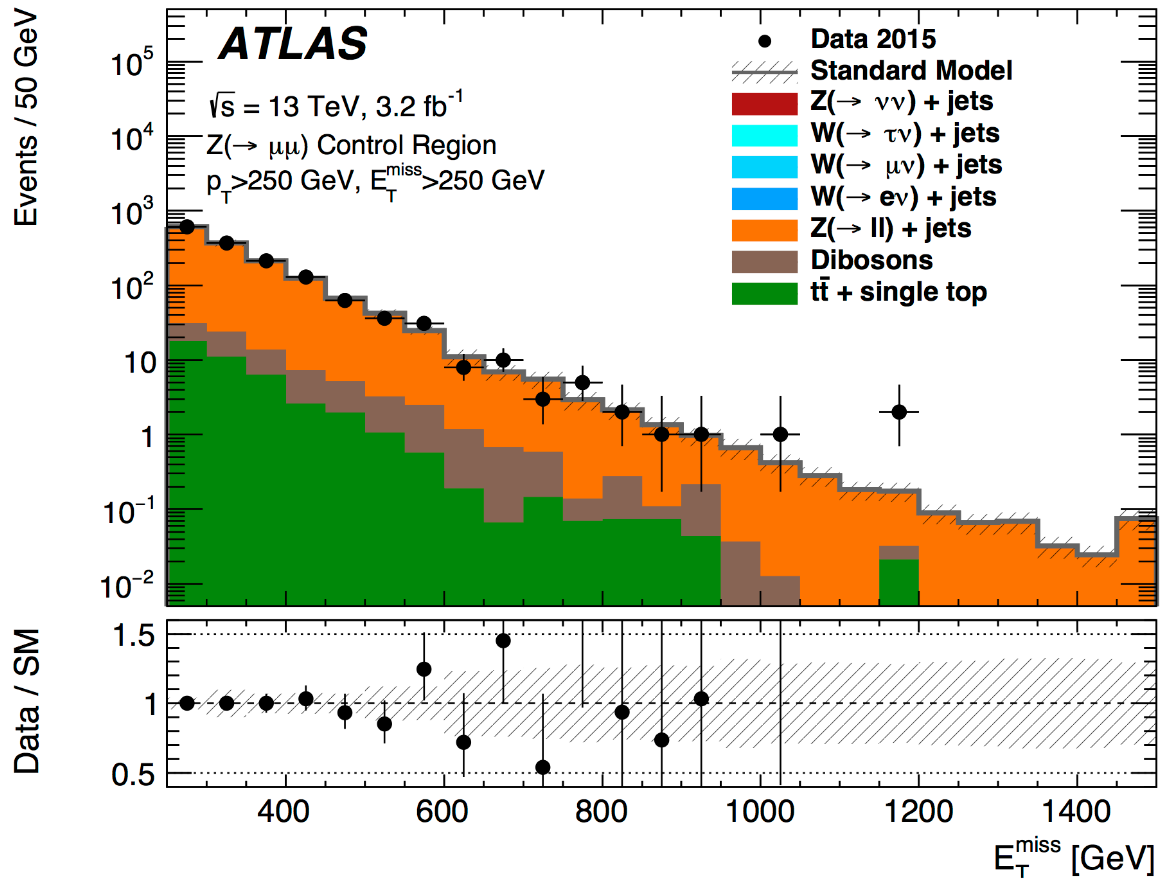
- Optimized to select W→eν+jets and W→τν+jets
  - 1 electron, visible in MET calculation
    - MET resembles ν p_T rather than W p_T
    - Better suppression of multijet background
  - No m_T cut
    - Better Wτν acceptance
- Constrains W→eν+jets and W→τν+jets
  - Difficult to construct dedicated W→τν control region with similar topology
  - Hadronic W→τν can look like jet+MET without ISR Jet



MET+Jet Control Regions

Dimuon Control Region

- Optimized to select Z→μμ+jets
  - 2 muons, invisible in MET calculation
  - 66 GeV < $m_{\mu\mu}$ < 116 GeV (select Z mass)

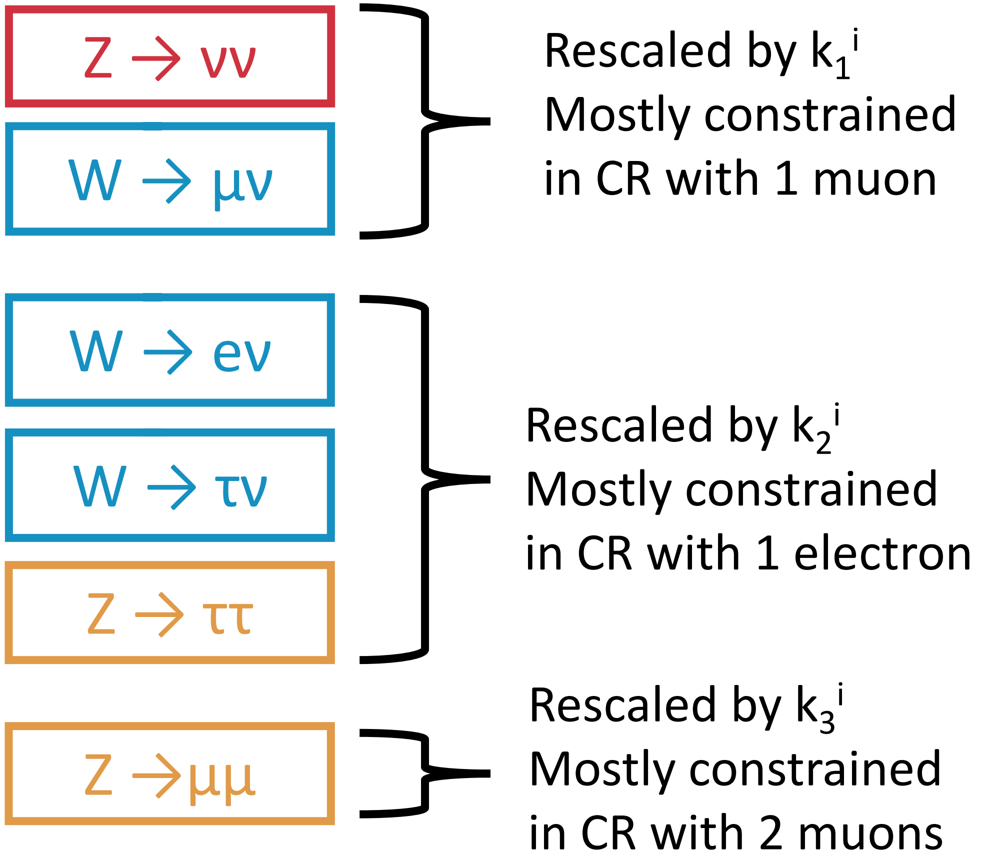


Normalization of $Z \rightarrow \nu \ \nu$

- Cannot construct Z → ν ν control region
  - Instead, use **W → μ ν** to model **Z → ν ν**
    - More statistics than Z → ℓℓ
- MET-dependent W/Z transfer factor uncertainty for Z → νν+jets yield to cover W+jets/Z+jets difference
  - MC modeling differences between W+jets and Z+jets processes
  - NLO electroweak corrections
- Uncertainty on Z → νν yield of ~4-6%
- Uncertainty on total background of ~2-4%

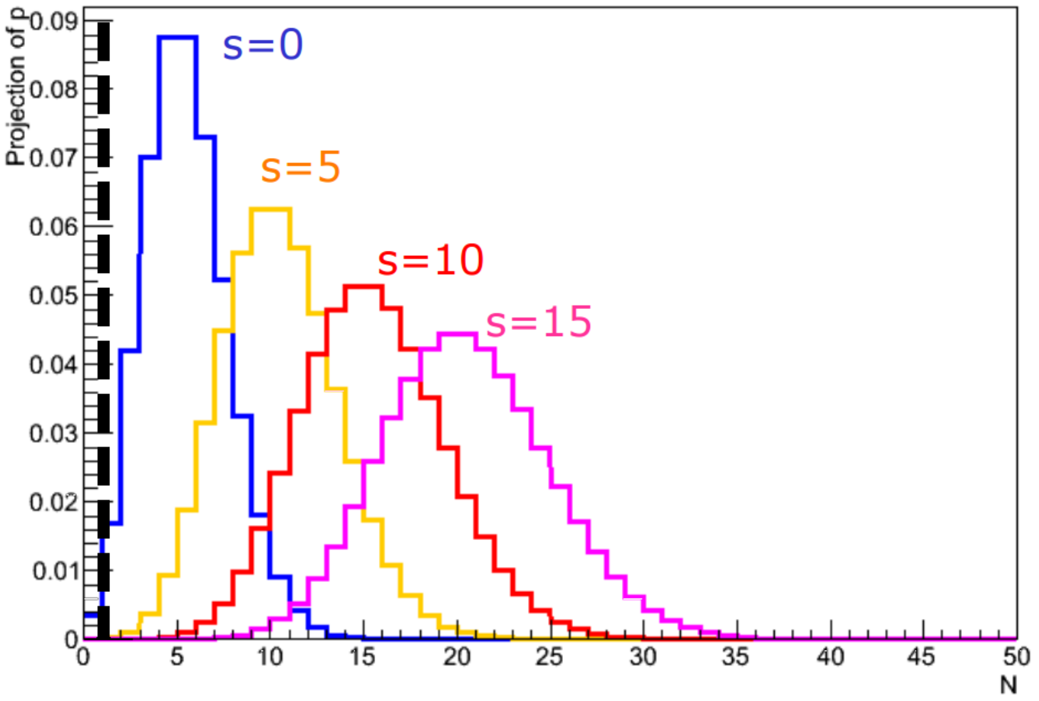
Statistical Analysis in MET+Jet

- Simultaneous shape fit:
  - Fit for three correction factors *k* rescaling six background processes in three control regions and the signal region
  - Additionally, make correction factors MET-dependent => k^i
    - One *k* per bin of MET: [250, 300, 350, 400, 500, 600, 700, ∞]
    - Reduces sensitivity to boson p_T modeling
    - Total of 3x7 = 21 normalization factors
- Many experimental uncertainties (jet energy & MET uncertainties, 5% uncertainty on luminosity, etc) are strongly constrained by the fitting scheme and are negligible
- Primary sources of uncertainty:
  - 2-4% from W/Z transfer uncertainty
  - ~3% uncertainty from modeling of top quark processes
  - 3-10% statistical uncertainty
- Total uncertainty on background of 4-12%
  - Controlling background uncertainty important for setting limits



Modified Frequentist Approach

- What if $n_{obs} = 1$?
  - $p_{s+b}(s=0) = 0.04$
  - s≥0 excluded at >95% C.L….
    - But s *must* be ≥0
  - Spurious exclusion due to weak sensitivity
  - At low s, distributions for s and s+b are very similar
- Bayesian approach:
  - Construct P(t|d)=P(d|t)P(t)
  - Include prior knowledge on s in p(t)
    - p(theory)=0 for s < 0
- Modified frequentist approach
  - Instead of p_{s+b}, use:
  - $CLs = p_{s+b}/(1-p_b) = 0.05$
  - With CLs: for $n_{obs} = 1$ exclude s > 3.4 at 95%



The Likelihood

Actual analysis much more complicated!

Multiple observations in control regions and signal regions, many (often correlated) uncertainties, etc

Signal models parameterized with signal strength μ, the ratio of the signal cross section σ to the cross section expected from theory σ_{theory}

$$\mu = \frac{\sigma}{\sigma_{\text{theory}}}$$

Need to define a likelihood for statistical tests => describes the probability of observing the number of events seen in the signal region for some value of μ

$$\mathcal{L}(\mu) = P\left(N|\mu S + B\right)$$

Simple counting experiment: likelihood is Poisson probability
N: number of observed events in SR
B: expected number of background events in SR
S: expected number of signal events in SR (rescaled by μ)

The Likelihood

Actual analysis much more complicated!

Multiple observations in control regions and signal regions, many (often correlated) uncertainties, etc

However, the estimated yields for certain background processes entering the SR and CRs are scaled by a normalization factor. With just one CR the likelihood would become:

$$\mathcal{L}(\mu, \theta) = P\left(N|\mu S + \theta B\right) \cdot P\left(N_{\mathrm{CR}}|\theta B_{\mathrm{CR}}\right)$$

N_{CR}: number of observed events in CR
B_{CR}: expected number of background events in CR
θ: background normalization factor

Adding multiple CRs *l*, bins of MET *b*, and background processes *k* to the likelihood:

$$\mathcal{L}\left(\mu, \boldsymbol{\theta}\right) = \prod_{\text{bins b}} P\left(N_b \middle| \mu S_b + \sum_{\text{bkg k}} \theta_{kb} B_{kb}\right) \prod_{\substack{\text{CRs l} \\ \text{bins b}}} P\left(N_{lb} \middle| \sum_{\text{bkg k}} \theta_{kb} B_{klb}\right)$$

The Likelihood

Actual analysis much more complicated!

Multiple observations in control regions and signal regions, many (often correlated) uncertainties, etc

- No perfect estimation of the expected background! Many potential systematic uncertainties:
  - Energy: How well do we know the energy of a jet? What's the typical variance?
  - Luminosity: How much data did we actually collect?
  - Reconstruction efficiency: How often do we lose leptons in the detector?
  - Etc
- Want to encode this in the likelihood
  - A mis-measured jet is not new physics!
  - These are not free parameters
    - Know the uncertainty σ on quantities such as the jet energy from other studies
- Uncertainties included as nuisance parameters θ in two terms:

$$\nu(\theta) = (1+\epsilon)^{\theta}$$

Term to normalize background with θ
ϵ: expected change in background for 1σ shift in θ's associated process

$$g(\delta|\theta) = e^{-(\delta-\theta)^2/2}/\sqrt{2\pi}$$

Gaussian constraint on nuisance parameter θ: central value (generally zero)
=> likelihood gets smaller as θ deviates from δ

The Likelihood

Actual analysis much more complicated!

Multiple observations in control regions and signal regions, many (often correlated) uncertainties, etc

Complete likelihood:

Background normalization factors

$$\mathcal{L}(\mu,\boldsymbol{\theta}) = \prod_{\text{SRs i}} P\left(N_b \middle| \mu S_b \cdot \prod_{\text{sig}} \nu_{br}(\theta_r) + \sum_{\text{bkg k}} \theta_{kb} B_{kb} \cdot \prod_{\substack{\text{bkg.}\\ \text{syst.}\\ s}} \nu_{bs}(\theta_s)\right)$$

Need to modify formula to add CR NPs and remove mc stats terms (a lot to explain)

Constrained background normalization factors for systematic uncertainties

$\prod_{\substack{\text{CRs l}\\ \text{bins b}}} \ldots \sum_{\text{bkg k}} \ldots_b$

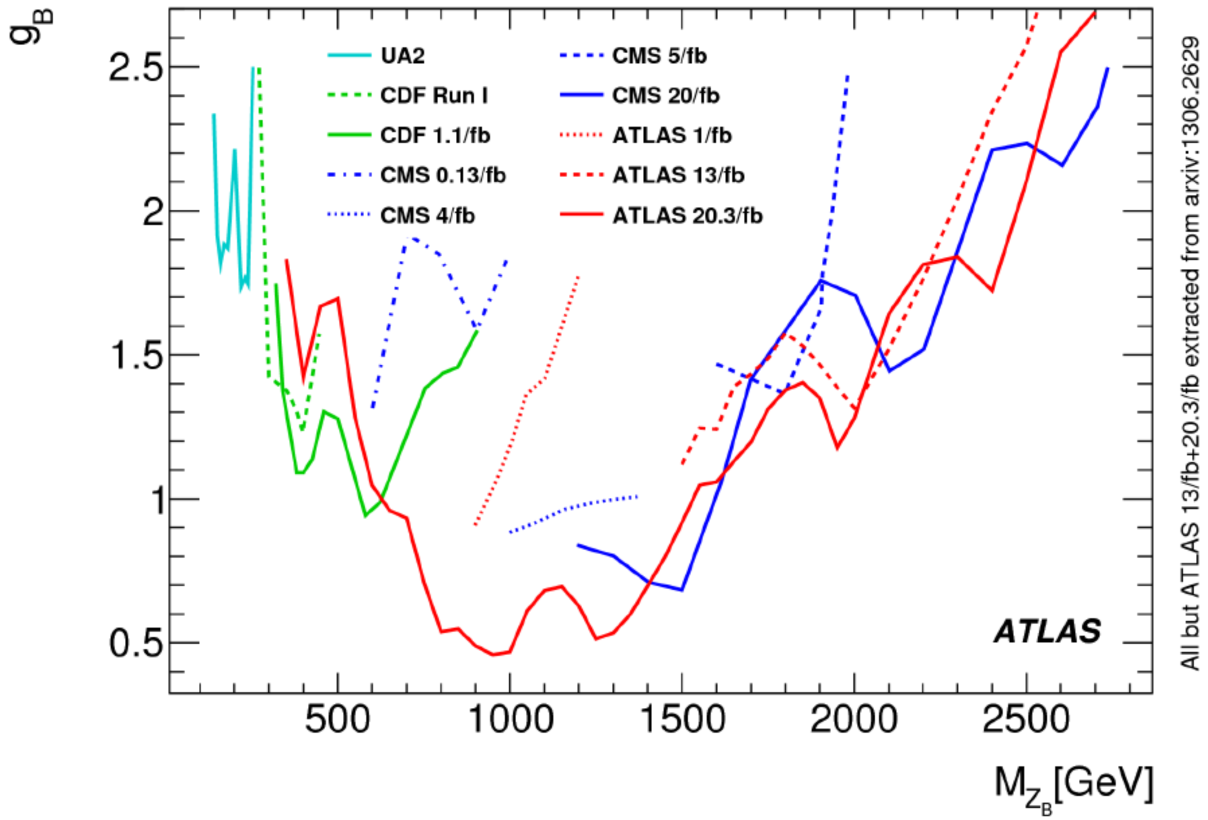
Constraints on systematic uncertainties

$$\cdot \prod_{\substack{\text{syst}\\ t}} g(\delta_t|\theta_t) \cdot \prod_{\text{bkg k}} P(\xi_k|\zeta_k\theta_k) \, .$$

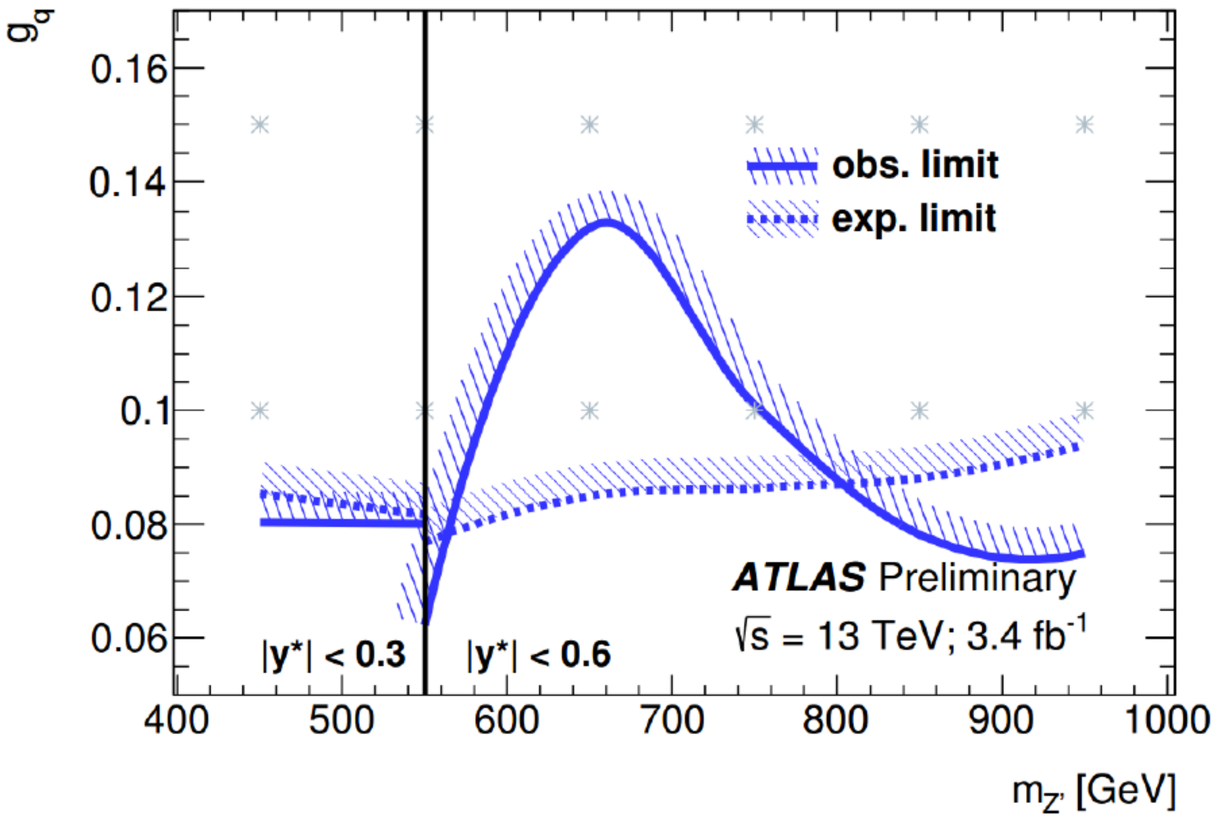
Use this formula to construct test statistics and set limits

Dijet Limits

8 TeV Dijet ($g_B = 6g_q$)



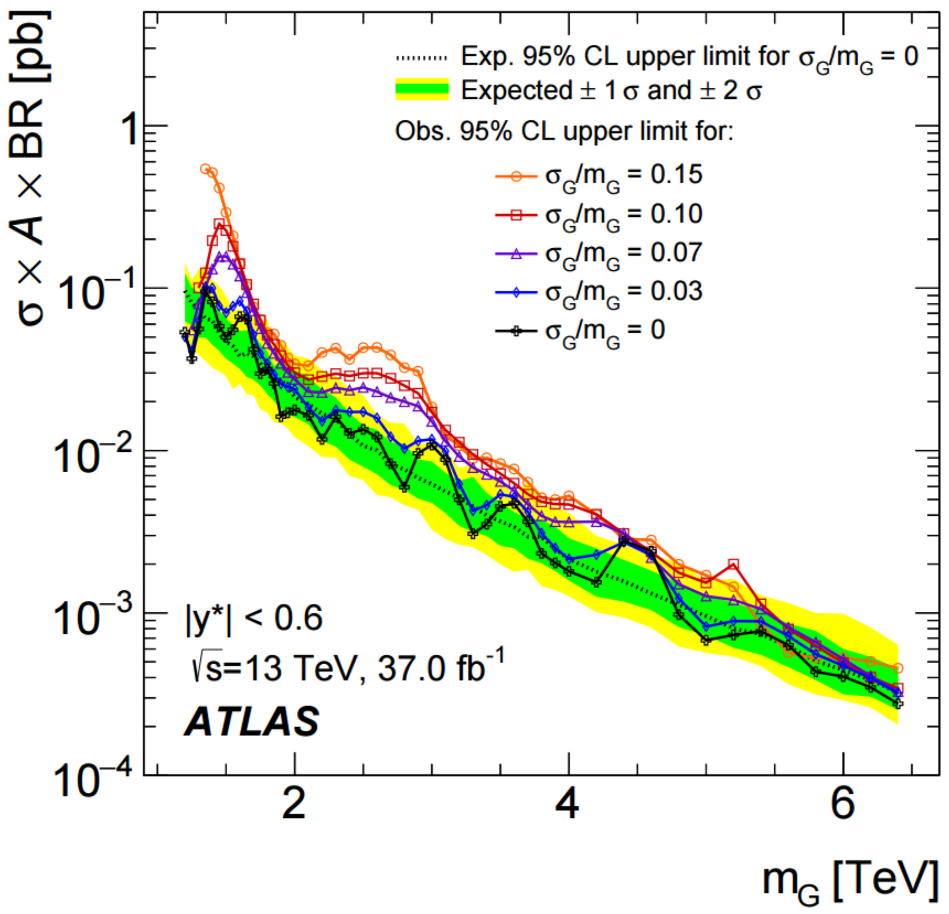
13 TeV Dijet TLA



Dijet Interpretation

- Reinterpretation following the procedure of Phys. Rev. D91 052007 (2015)
- Dijet searches set limits on generic Gaussian signal templates
  - Different masses and widths
  - Can use these generic limits to determine if a given DM model would be excluded or not
- Step 1: generate signal templates
  - Axial vector mediator decaying to light quarks or b quarks
    - With MadGraph DMSimp:
    - `define p => g u c d s b u~ c~ d~ s~ b~` (same for j)
    - `generate p p > xi, xi > j j`
  - Results in truth-level ntuples of events
    - Can make distributions of m_{jj}, etc
    - However, limits are set on resonances as they look in the detector!

arXiv:1703.09127



Dijet Interpretation

- Step 2: adjust signal templates
  - Apply dijet resolution smearing to mass template
    - Define a Gaussian random number generator with width set to resolution
    - For each jet in each event:
      - Draw random number ϵ
      - Multiply the jet energy by 1+ϵ
  - Truncate tails of signal distribution
    - Tails not considered in generic Gaussian limits
    - This removes signal events => need to adjust signal acceptance accordingly
      - Acceptance: fraction of signal events expected to be reconstructed/identified
      - Limits are set on visible cross section
      - $\sigma_{vis} = \sigma$ * acceptance

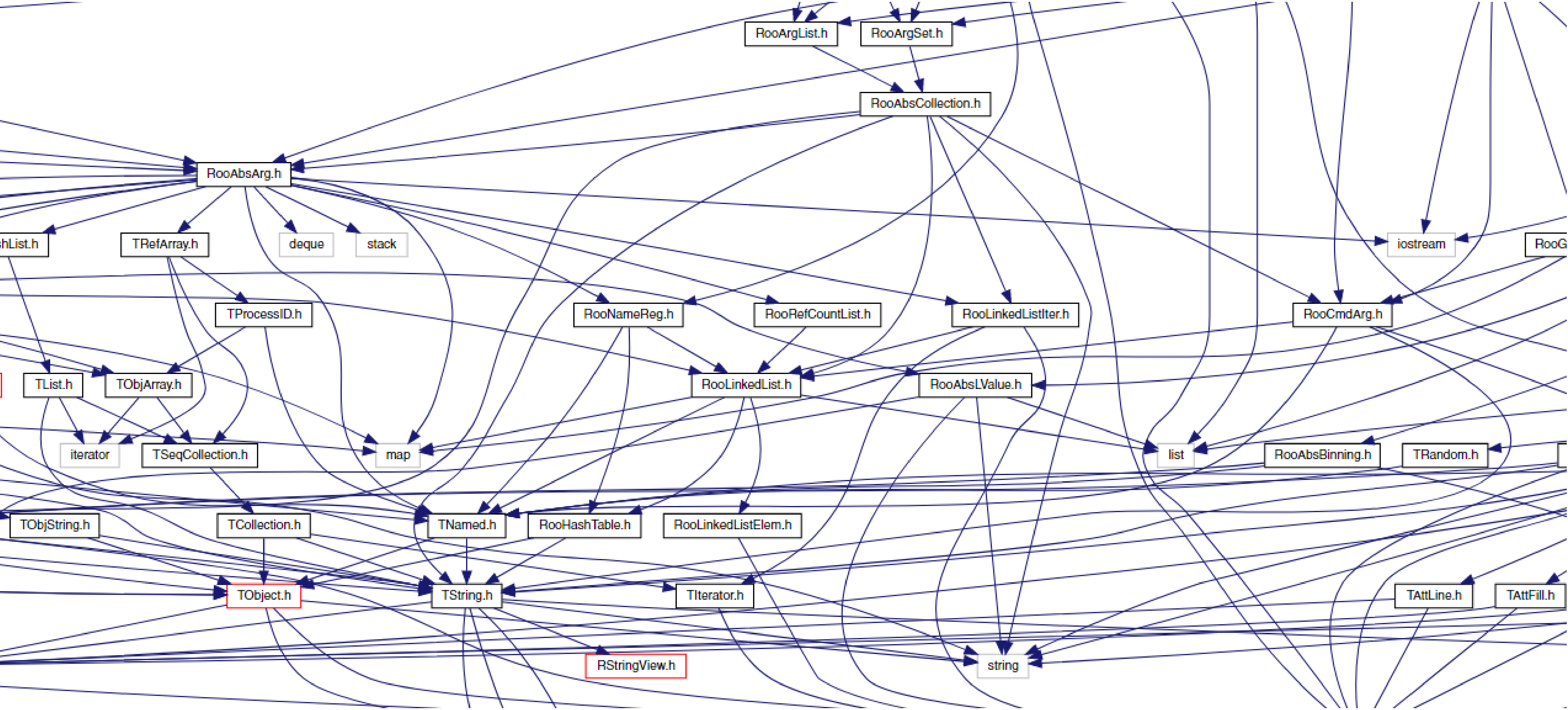
Dijet Interpretation

- Step 3: Fit
  - Need to compare smeared DM signal distributions to the generic Gaussian limits
  - What are the widths and masses of our templates?
  - Fit Gaussian function to signal templates and extract best-fit parameters
- Step 4: Determine exclusions
  - For each signal template:
    - Mass, width, σ_{vis}^{theory}
  - For each generic Gaussian:
    - Mass, width, 95% CL limit on σ_{vis} (σ_{vis}^{limit})
  - If $\sigma_{vis}^{theory} > \sigma_{vis}^{limit}$, signal is excluded!

EFT vs Simplified Models

- Effective field theories used for 8 TeV ATLAS searches
  - Most general model possible
  - Only parameters are m_{DM} and Λ
- However, EFT is only valid if $Q^2 < 4\pi\Lambda$
  - At the LHC $Q^2 \sim$ jet $p_T \sim$ 200 GeV
  - Can truncate when $Q^2 > \Lambda$ but requires knowing masses, couplings, etc
- Simplified models
  - Basic model, not necessarily complete
  - As few assumptions as possible
  - More parameters than EFT, but valid in all regions of phase space

RooFit

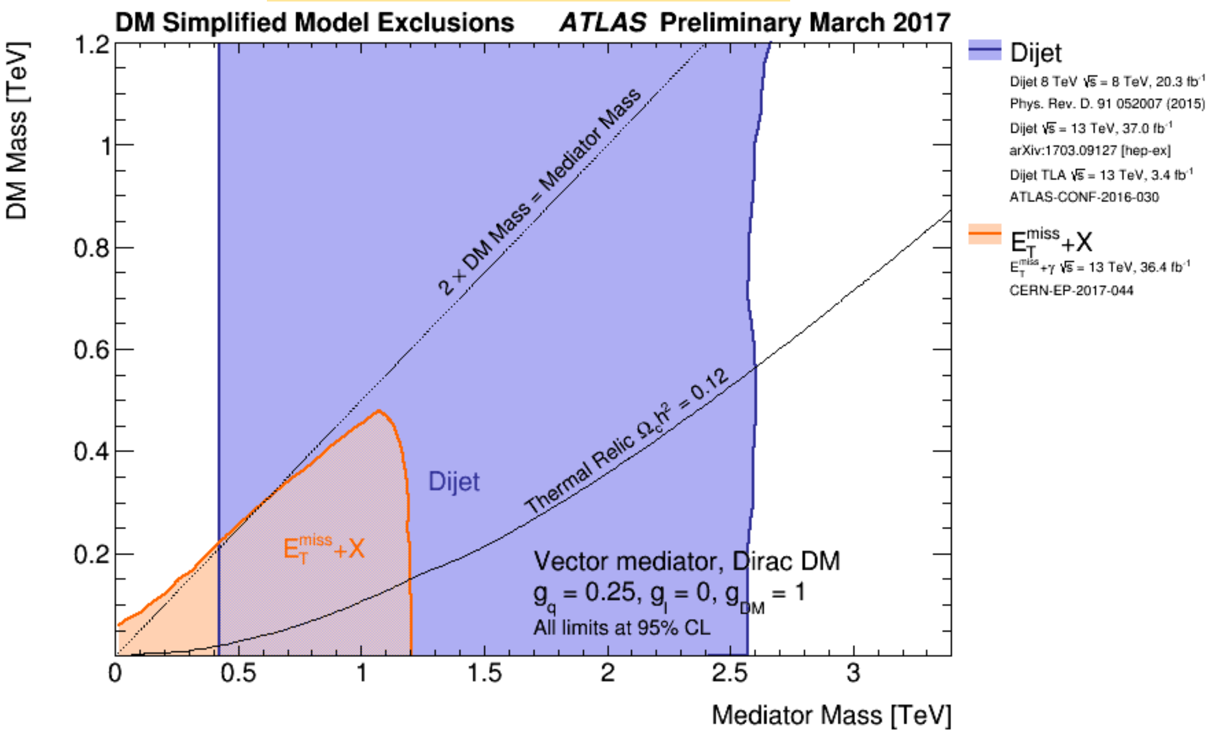


RooArgList.h
RooArgSet.h
RooAbsCollection.h
RooAbsArg.h
hList.h
TRefArray.h
deque
stack
iostream
RooG
TProcessID.h
RooNameReg.h
RooRefCountList.h
RooLinkedListIter.h
RooCmdArg.h
TList.h
TObjArray.h
RooLinkedList.h
RooAbsLValue.h
iterator
TSeqCollection.h
map
list
RooAbsBinning.h
TRandom.h
TObjString.h
TCollection.h
TNamed.h
RooHashTable.h
RooLinkedListElem.h
TObject.h
TString.h
TIterator.h
TAttLine.h
TAttFill.h
RStringView.h
string

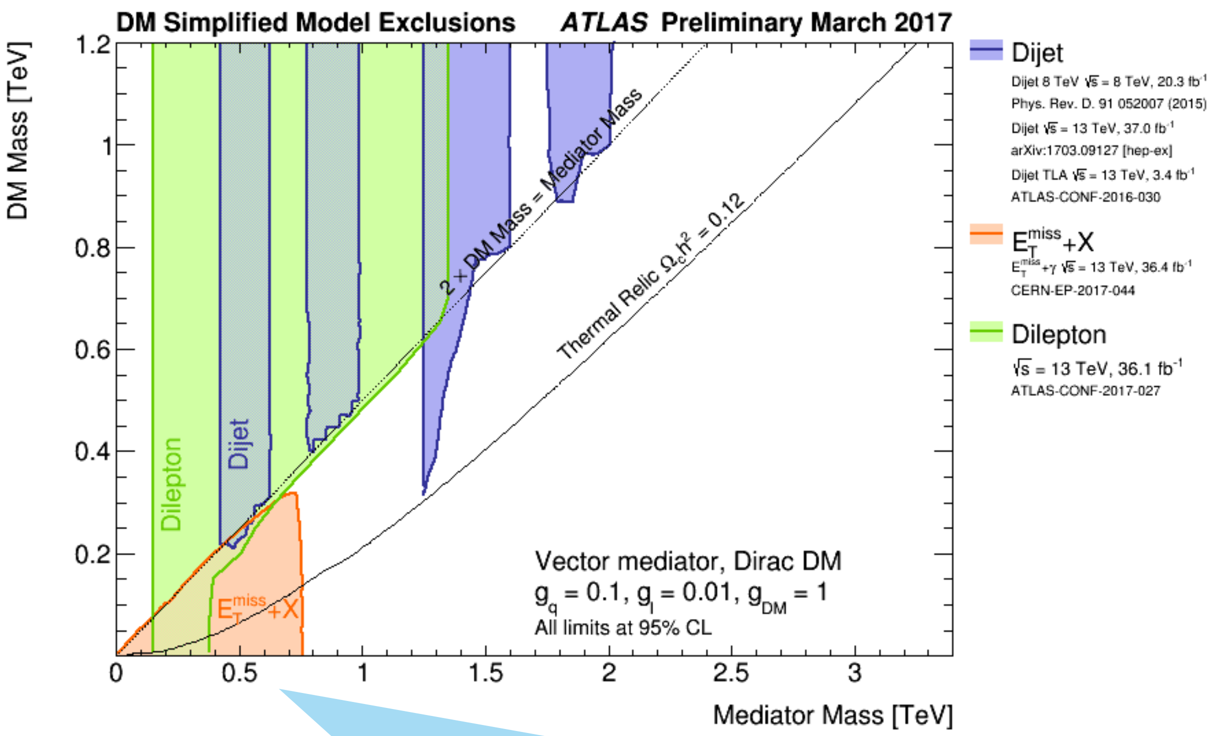
LHC DM Exclusions

- Can look at limits for different couplings …and mediator types
  - Also consider nonzero lepton couplings

Vector mediator
$g_q = 0.25$, $g_l = 0$, gDM = 1



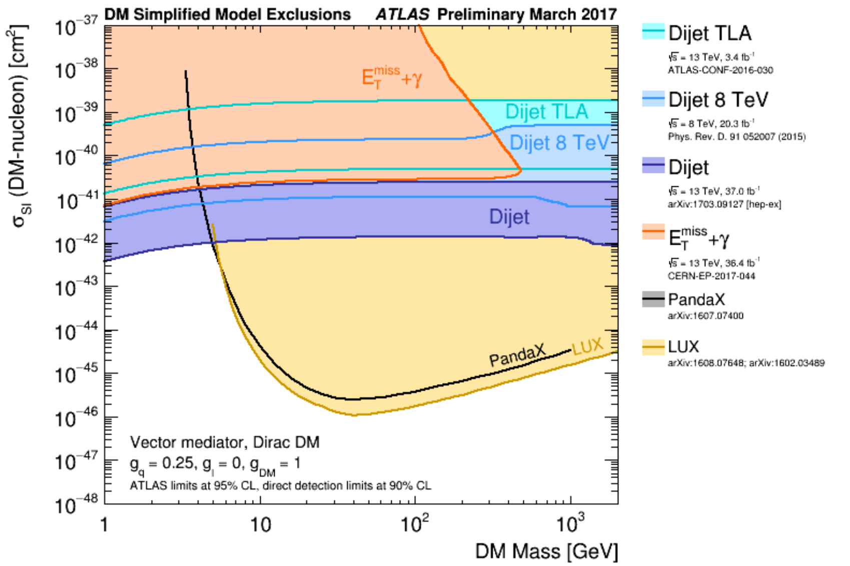
Vector mediator
$g_q = 0.1$, $g_l = 0.01$, gDM = 1



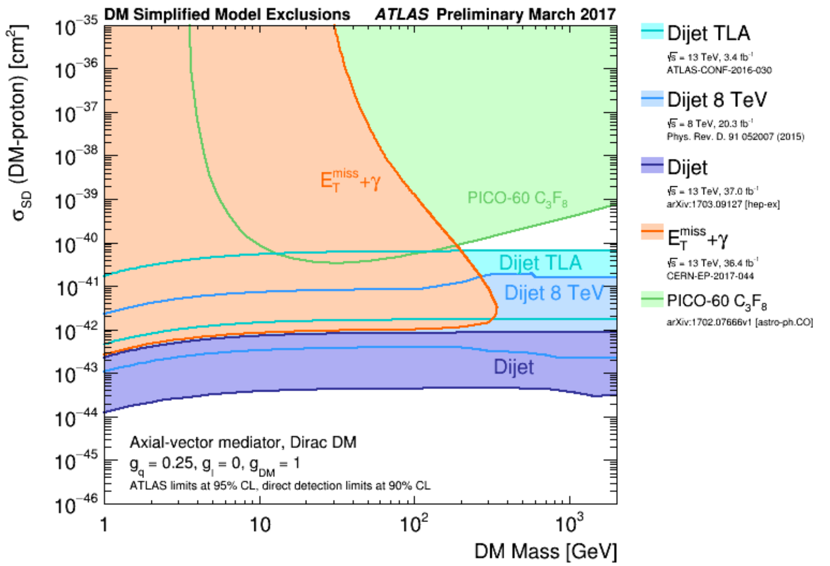
Lepton couplings are different!! Dilepton limit difference not simply due to vector vs axial mediators

DM-Nucleon Scattering

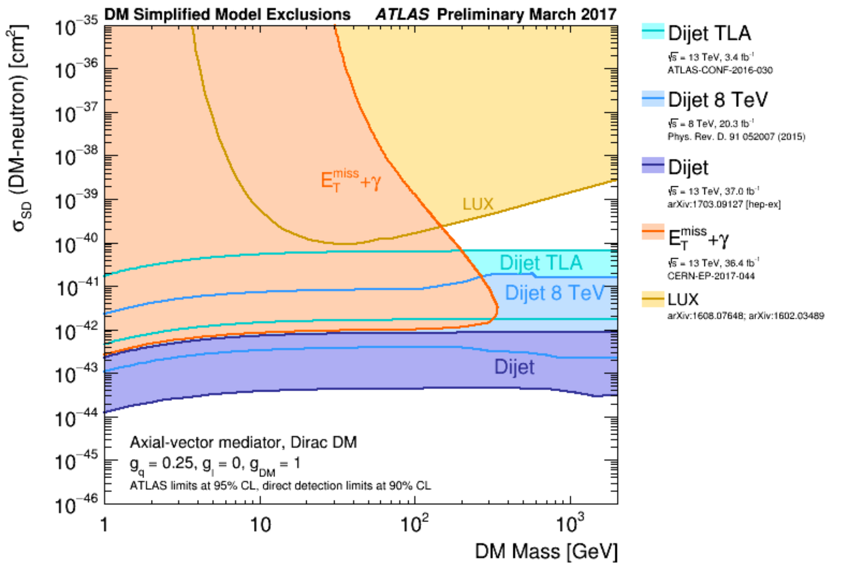
Spin-independent limits



Spin-dependent proton limits



Spin-dependent neutron limits

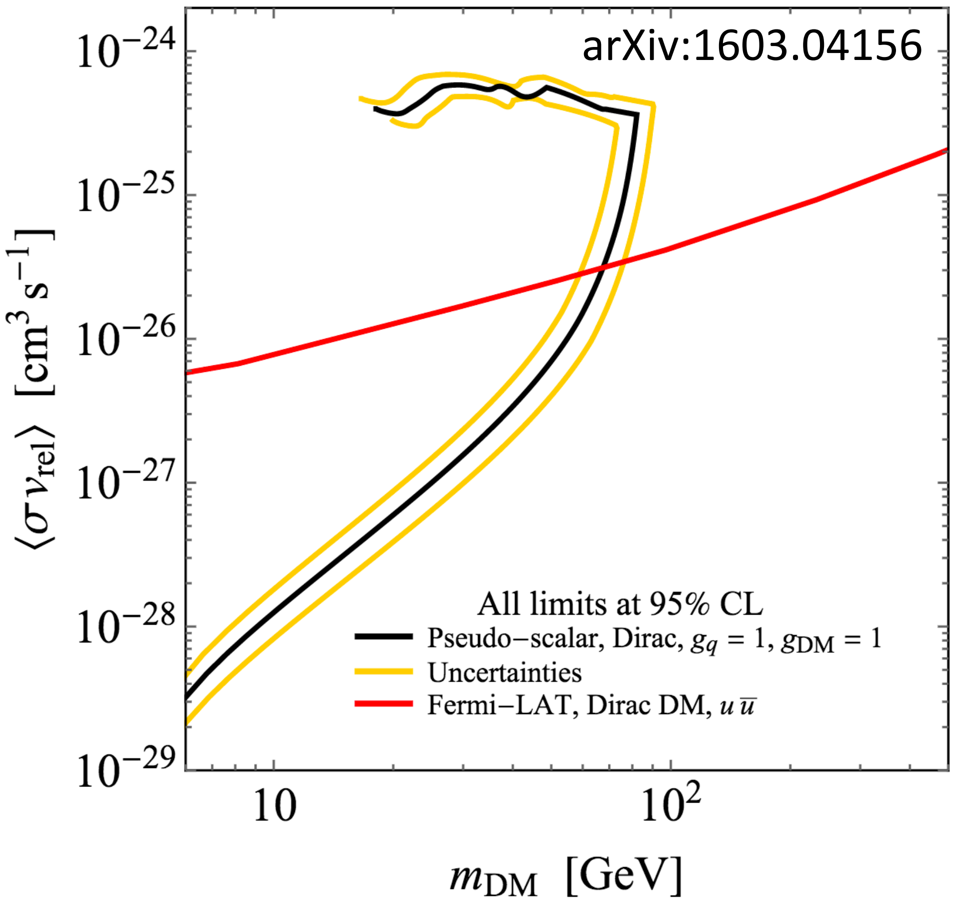


- Direct detection (DD) of DM
  - DD experiments set a limit on the rate of interactions between local DM halo and atomic nuclei
- Translate simplified model collider limits into limits on dm-nucleon effective vertex
  - Constraints on (qq->med->DM DM) or (qq->med->qq) constrain (DM Nucleon -> DM Nucleon)
  - Numerically convert (m_{DM}, M_{med}) limit contours to (m_{DM}, $\sigma_{SI/SD}$):

$$\sigma_{\rm SI} \simeq 6.9 \cdot 10^{-41}\ {\rm cm}^2 \left(\frac{g_q g_{\rm DM}}{0.25}\right)^2 \left(\frac{1\ {\rm TeV}}{M_{\rm med}}\right)^4 \left(\frac{\mu_{n\chi}}{1\ {\rm GeV}}\right)^2$$

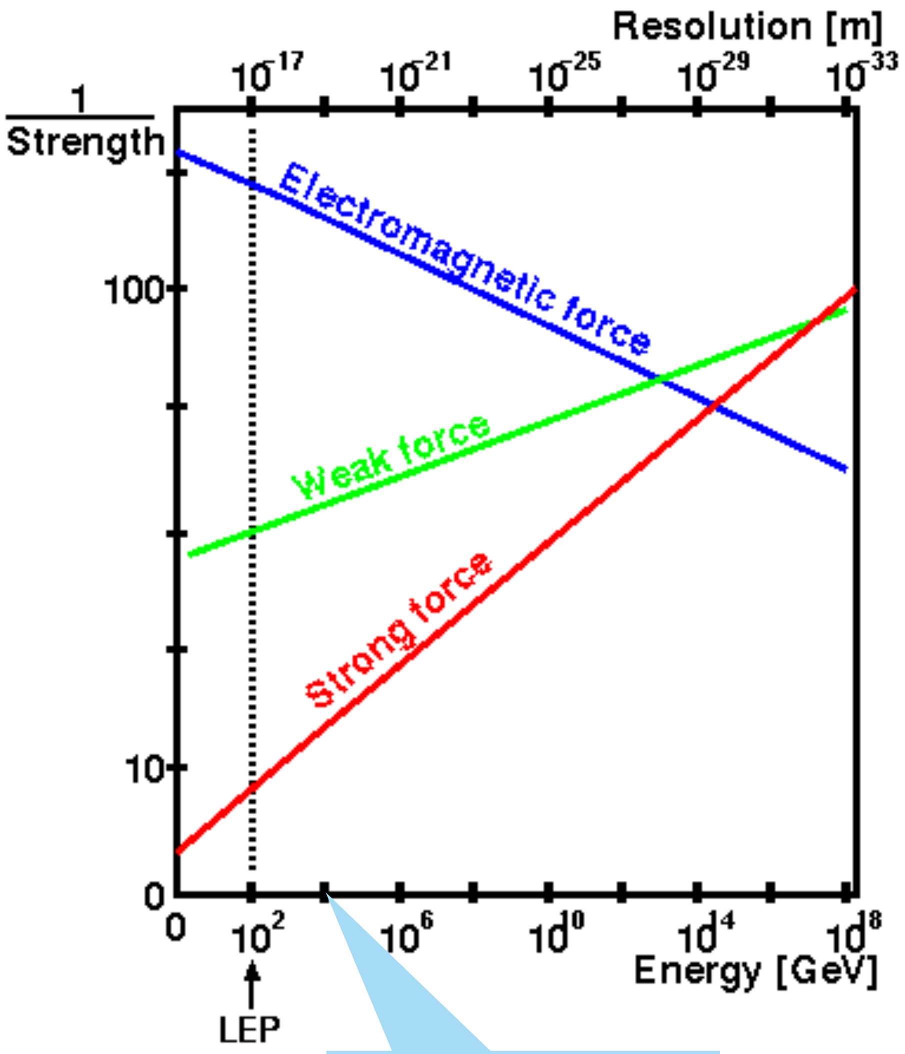
LHC Limits & ID Limits

- Indirect detection (ID) of DM
  - Look for cosmological DM annihilating into SM particles
    - Expect DM to be concentrated in the center of galaxies
    - Can look at dwarf galaxies, the center of the Milky Way, etc
  - Fermi-LAT experiment: gamma ray spectrum
    - Can set limits on DM annihilation into photons
    - Other channels: gamma rays produced by other decay products
- Compare DM production limits with DM annihilation limits
  - LHC mono-X limits should provide complementarity at low DM mass
  - No public LHC result… yet!



Comparison of (fake) LHC limits with Fermi-LAT limits for a model where DM only annihilates into up quarks

Coupling Constants



~LHC energy