Impact of Detector Simulation in Particle Physics Collider Experiments

V. Daniel Elvira

UChicago Seminar

February 25th, 2019

Motivation

Throughout my career, I received many requests for material showing concrete examples on how detector simulation helps modern particle physics experiments

As a follow-up of one of these requests, John Harvey, former leader of the Software Group (SFT) at CERN, encouraged me to write a note on the topic

The note found its way to Physics Reports where it was recently published as a review paper:

- "Impact of detector simulation in particle physics collider Experiments", Physics Reports 695 (2017) 1–54

This presentation follows closely the material included in the paper

Outline

Detector simulation is of critical importance to the success of HEP experimental programs, a determinant factor for faster delivery of outstanding physics results

- Introduction
  - History, facts and numbers, modeling tracks and showers, the simulation software chain
- Detector simulation tools
  - Types of simulation, the Geant4 toolkit, physics validation
- Applications of detector simulation to HEP collider experiments
  - Simulation in data analysis, detector design & optimization, software & computing design, testing
- Modeling of particle and event properties and kinematics
  - Geometry and material effects, examples for different final states, the jet cross section story
- Simulation and publication turnaround
- Economic impact and cost of simulation in HEP experiments
- The future

Introduction

History, facts and numbers, simulation software tools and applications

Some history

Accurate computer simulation is essential to design, build, and commission the highly complex detectors in modern HEP experiments, and to analyze & interpret their data

- Old times detector simulation
  - Simple analytic calculations, back-of-the-envelope estimates
- Era of detailed detector simulation started in late 70's early 80's
  - Electron Gamma Shower (EGS), GEometry ANd Tracking (GEANT) software
- GEANT3 software kit to describe complex geometry, propagate particles and model interactions as they traverse different materials and EM fields
  - GEANT3 widely used by CERN, DESY, FNAL experiments. First OPAL (LEP), then L3 and ALEPH, followed by experiments at DESY and FNAL in the 90's
- Other simulation tools are FLUKA and MARS
- Geant4 used by most HEP experiments – limited initially, the norm in 21st century

Why to simulate detectors

- Save time and money, improve the quality and accuracy of physics measurements

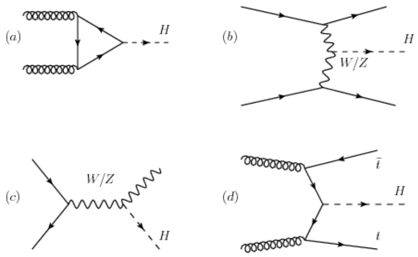
  Design optimal detector, best physics at a given cost, even before fastening the first screw!

- Simulation is not magic

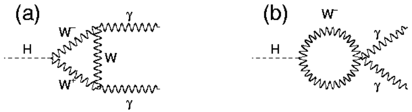
  Particles cannot be “discovered” in a simulated sample which does not model them

- Simulation is essential to HEP experiments

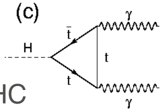
  Teaches physicists what mark a new particle would leave in the detector if it existed



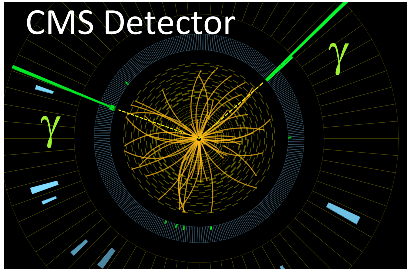
+



Higgs Production and decay to photons at the LHC



and



➔ Higgs discovered in July 2012

SM Higgs prediction:
Higgs is produced at the LHC and decays to two γ’s with given properties for the event and the individual particles

Observation:
two photon events with predicted detector marks are observed

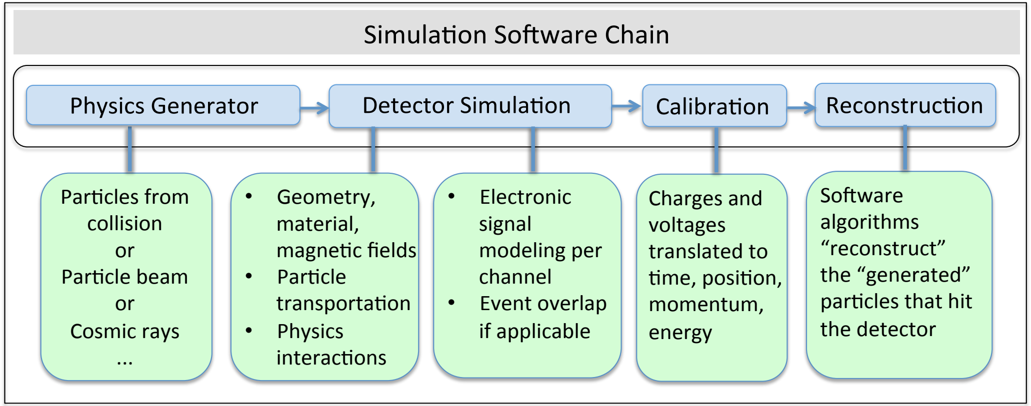
Facts and numbers

The role of detailed detector simulation in HEP experiments has increased during the last three decades to become an essential component

- LHC experiments simulate events at a speed and with physics accuracy never seen before
  - ATLAS/CMS: seconds to minutes per event, tens of billions of events since 2010
  - CDF/D0 (early 1990's): hundreds of thousands of poor quality events, in comparison
- Geant4-based simulation has shortened the time between data-taking and journal submission of increasingly precise physics results at the LHC
  - Other factors being detector and computing technology, a wealth of experience from pre-LHC experiments, better calibration and analysis techniques, communication tools, etc.
- In most experiments, detector simulation takes > 1/2 of all computing resources
- Over the next two decades, detector simulation applications need to deliver orders of magnitude more events with increased physics accuracy and with a flat budget

A daunting challenge for detector simulation tools

Simulation software chain in a typical HEP experiment

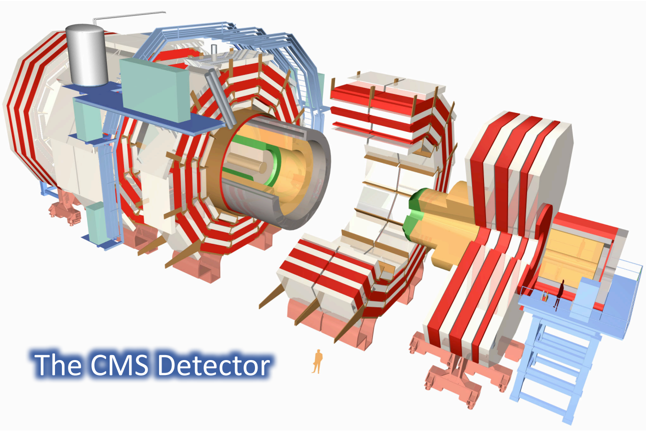


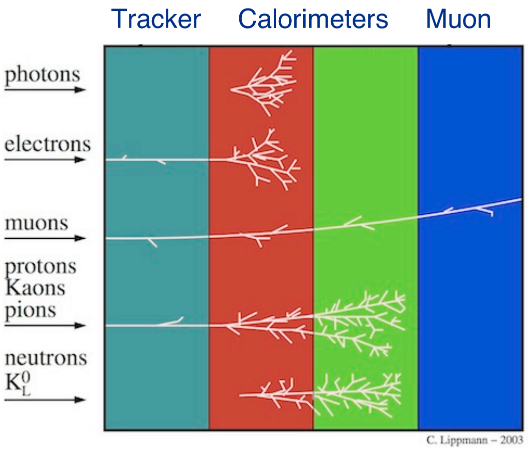
Simulation referred to as "Monte Carlo (MC) simulation"

Simulated events referred to as "MC events, or MC samples"

- Physics generator: provides the final states of the physics process of interest (Pythia, Herwig, Madgraph, Alpgen, etc. in colliders; GENIE, etc. for neutrinos)
- Detector simulation [focus of this presentation]:
  - First stage: passage of generated particles through detector material and magnetic fields
  - Second stage: detector electronics, backgrounds to collision of interest (pileup)
- Calibration: from detector quantities to physics quantities
- Event reconstruction: algorithms, typically the same, applied to real data

Particles through a collider detector: tracks and showers





Tracker measures charged particle trajectories

Calorimeter measures particle energy

Muon system, combined with tracker, measures muon trajectories

(Physics processes: energy loss, multiple scattering,…, etc. "Showers" of secondary particles produced through EM and nuclear interactions)

Hits and energy deposits in millions of detector channels ➔ **x**, **p**, E, time measurements

Particle tracks and particle showers must be modeled accurately

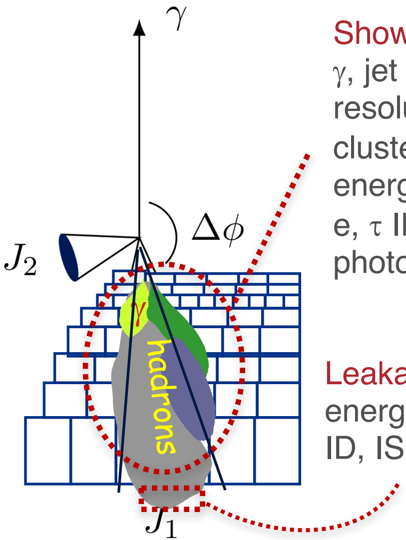
Shower modeling affects physics predictions – examples

The accuracy of the modeling of particle showers in calorimeters

(particle types and multiplicity, E and η, ϕ distribution, response linearity and fluctuations)

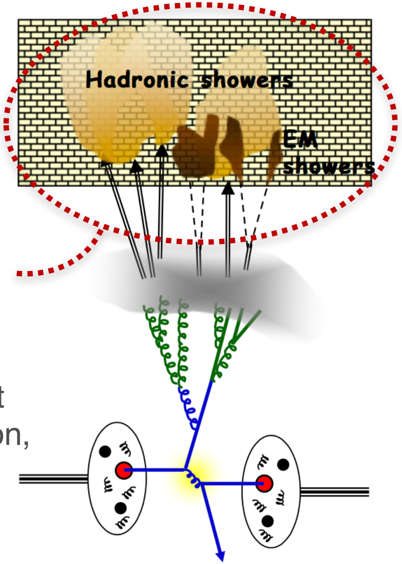
affects the degree of data-to-MC agreement for

– physics object variables, lepton identification (ID) and isolation (ISO) efficiency, etc



Shower mis-modeling affects e, γ, jet energy response and resolutions, jet multiplicity, un-clustered and out-of-cone energy in jet reconstruction, γ, e, τ ID and ISO efficiencies, di-photon and di-lepton separation

Leakage mis-modeling affects jet energy response, μ reconstruction, ID, ISO efficiencies



Impact on physics predictions:

Backgrounds with multi-jets (QCD), and leptons (EWK)

W, Z, top, Higgs mass

QCD cross sections, jet shapes, sensitivity to soft radiation

Detector simulation tools

Types of simulation, the Geant4 toolkit, physics validation

Types of simulation: toy, parametrized, full

- Toy simulation – a few simple analytical equations without a detailed geometry/field description or particle shower development
  - Zeroth order detector or physics studies
    - Output data format may not be the same as real data’s , speed is a small fraction of a second/event
- Parametrized simulation – approximate geometry/field description, parametrized energy response and resolution, shower shapes
  - Computing intensive MC campaigns that would otherwise be prohibitive, i.e. parameter space scanning in BSM signal samples
    - Examples are the CDF QFL simulation (1990’s) and CMS Fast Simulation framework which are tuned to test beam data, single tracks and/or full simulation
    - Output data format is typically identical to real data’s, speed is of the order of a second/event

Types of simulation: toy, parametrized, full

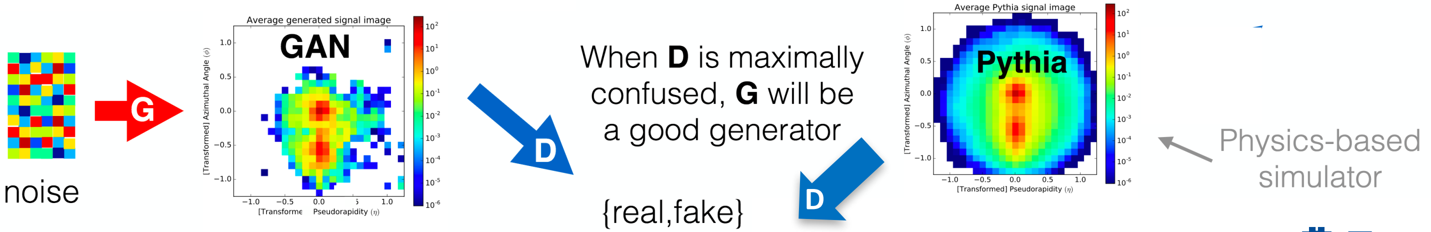
- Full simulation – based on Geant, FLUKA, MARS with detailed geometry/field and shower description, the latter based on individual particle interactions
  - Detector and physics studies where geometry and physics accuracy are important
    - Output format same as real data's, speed is of the order of seconds to minutes per event

Full versus Fast simulation – misleading concept

Experiments are moving towards simulation frameworks with flexibility to incorporate "fast simulation techniques" to a base Geant4 application

Tabulation, shower libraries, parametrization a la GFLASH, Machine Learning

ATLAS ML application to simulation



The Geant4 simulation toolkit

At the core of most full simulation applications at modern collider experiments, i.e. LHC, is the Geant4 toolkit

- International Collaboration of tens of institutions and ~120 physicists and computer professionals, including FNAL, CERN, and SLAC
- Written in OO C++, > 1 million lines of code, > 2000 C++ classes
- Used by almost all HEP experiments (10,000 users), space, and medical applications

20th G4 Collab. meeting at FNAL, USA (2015)

21st G4 Collab. meeting in Ferrara, Italy (2016)

22nd G4 Collab. meeting in Wollongong, Australia (2017)

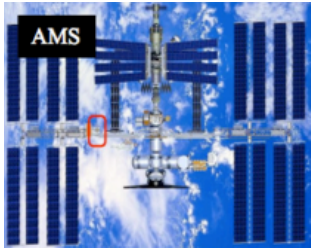
23nd G4 Collab. meeting in Lund, Sweden (2018)

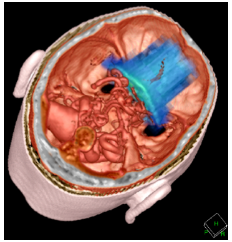
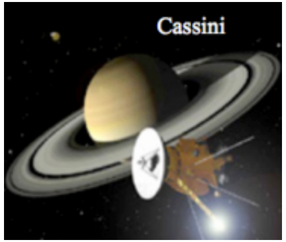
The Geant4 simulation toolkit

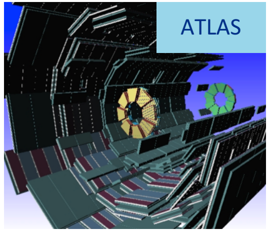
The impressive success of the Geant4-based simulation applications at the LHC experiments is the result of:

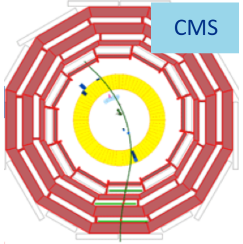
- Many years of hard work, partnership between the experiments and the Geant4 team
- A process to develop, optimize, and validate the many Geant4 physics models
- Different fora served as vehicles of communication, discussion, and information exchange

The use of Geant4 has extended to include high-energy, nuclear and accelerator physics, as well as medical science and treatment, and space exploration.



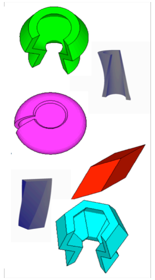




A Geant4-based simulation application

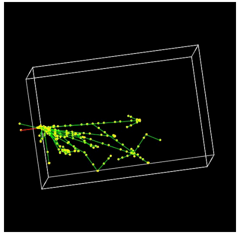
Experiments develop a ”simulation application” (software package) for their detector using Geant4 by assembling each of the following elements:

Detector geometry (shapes and materials)



\+

Particle Propagation through geometry and EM fields



\+

Physics Processes

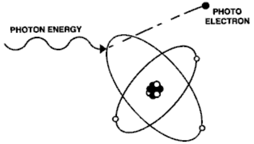
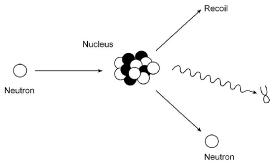


Figure 2-XIV. Gamma Interaction by Photoelectric Effect



The user selects:

- Method of integration of the equation of motion, particle tracking parameters
- “Physics Lists” composed of a subset of the physics models available to describe the interaction of particles with matter for energy between 250 eV and ~100 TeV

Output is a collection of “particle trajectories” and “simulated hits” with position, time, and energy deposited in detector volumes

Validation of the detector simulation physics

A collaborative task involving the Geant4 developers and the experiments

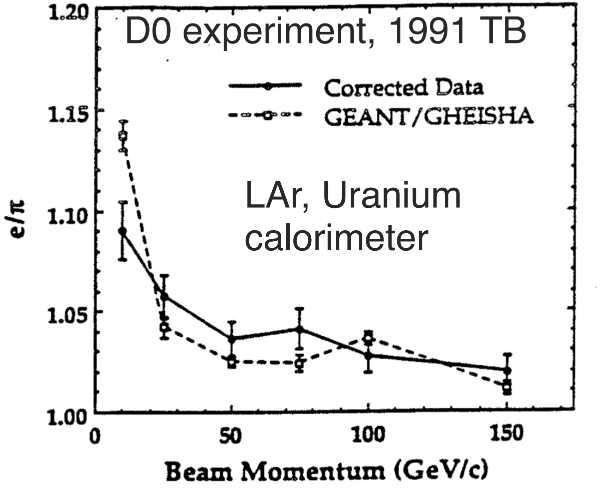
- Thin-target experiments
  - Beams of particles of different types (typically e, π, p) are directed onto thin targets made of materials typically used in HEP experiments (Be, C, Cu, Pb, Fe, etc.)
  - Measure cross sections, angular distributions, particle multiplicities
  - Examples: CALICE, HARP, NA49, NA61
  - Used by the G4 team to validate individual (G4) models at the single-interaction level
- HEP experiments
  - Collider, neutrino, muon experimental data, as well as their associated test beam results are compared to predictions from their Geant4-based simulation applications
  - Quantities are typically energy response functions, shower shapes

These two sets of data are complementary: thin-targets for “first principles” G4 models tuning, HEP experiments for confirmation or small tweaks to the models

Simulation physics validation: HEP experiments – test beams

Collider experiments run test beam (TB) campaigns, used to select among physics lists, guide the G4 team on how to assemble them from individual models

Early times: D0, CDF Experiments (Tevatron, early 90's)



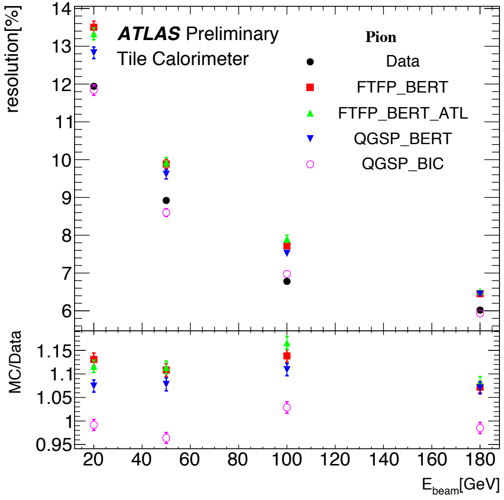
Slow computers – Geant3 full simulation took O(hour/event), limited TB programs, deficient communication technology

- Low statistics MC samples
- Approximations in exchange for time performance
  - D0: simplified geometry, average materials, shower truncation, full Geant3 simulation only for some analyses
  - CDF: use of parametrized simulation (QFL) tuned to minbias/TB data, then Geant3+GFLASH shower parametrizations

e/π response ratio: large statistical uncertainties in GEANT3 prediction (negligible in CMS)
Limited energy range 10-150 GeV, difficult to evaluate high energy region (2-300 GeV in CMS)

Simulation physics validation: HEP experiments – test beams

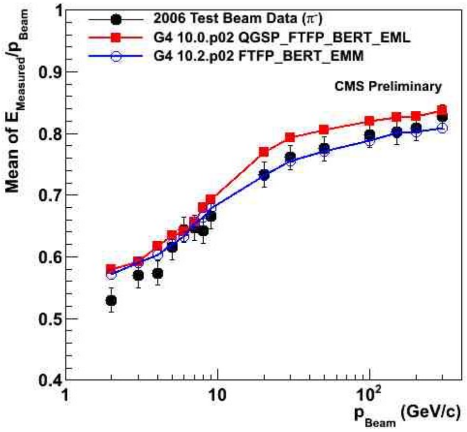
ATLAS 2002–2003 TB



Calorimeter π energy resolution (%) vs. beam energy

- Stat errors only
- MC/data ~ 1.00 -1.15

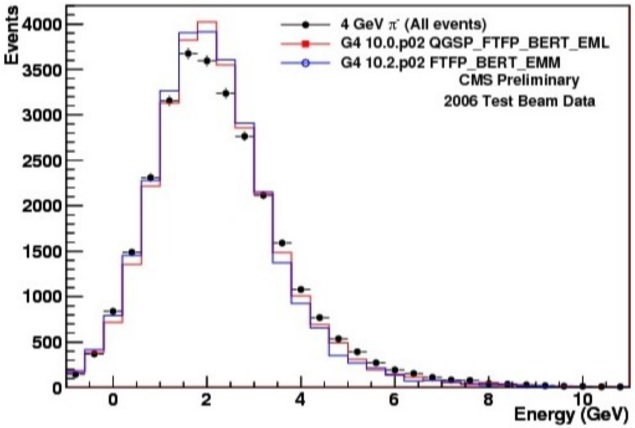
CMS 2006 TB



Calorimeter π energy response vs. beam energy

- Excellent agreement within statistical uncertainties
- MC overestimate trend below ~5 GeV

CMS 2006 TB



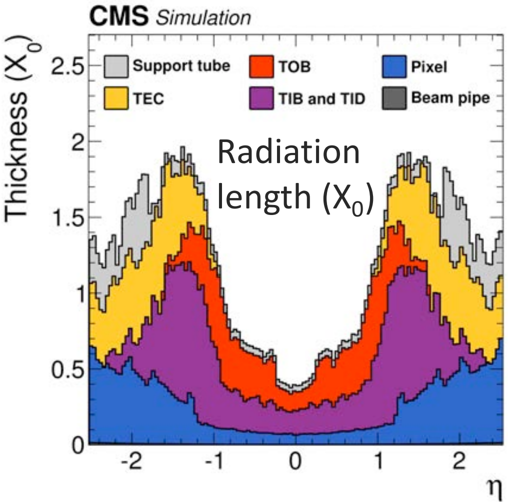
Calorimeter response function

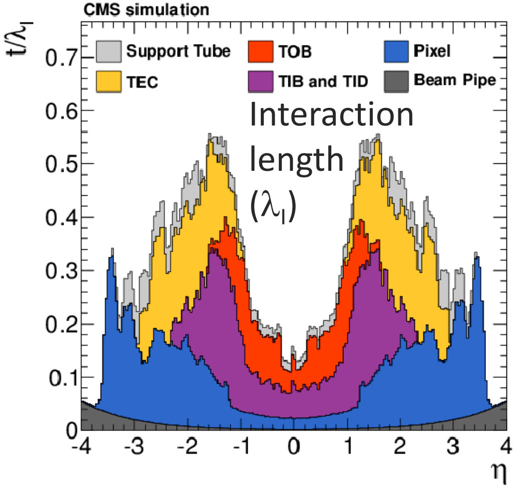
- Good modeling of core and tails critical for jet and E_T^{miss} modeling (jet cross sections and QCD background to BSM measurements)

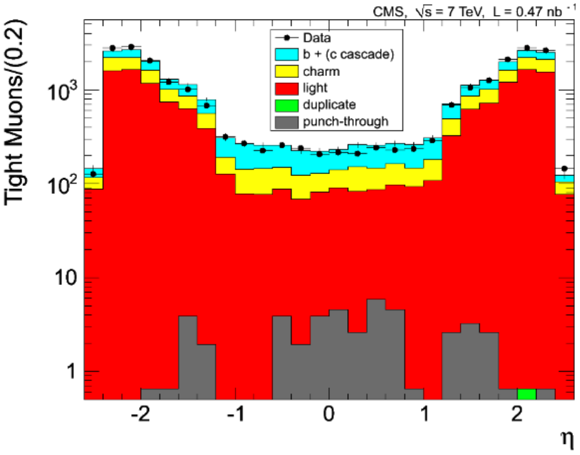
Note small/negligible statistical errors in simulation

Simulation physics validation: HEP experiments – physics runs

Data from collider runs used for final validation of full simulation application







Thickness of CMS silicon tracker from simulation

- Mis-modeling affects energy loss of charged particles, photon conversion (70% in CMS tracker)

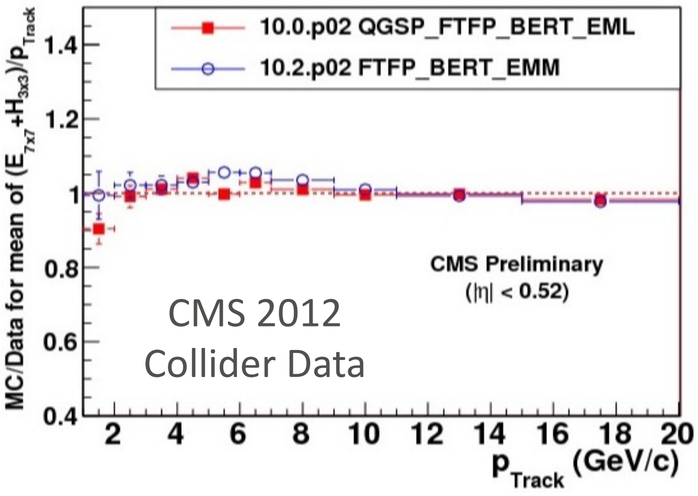
Validated by weighing components of real detector

CMS inclusive μ sample (zero-bias)

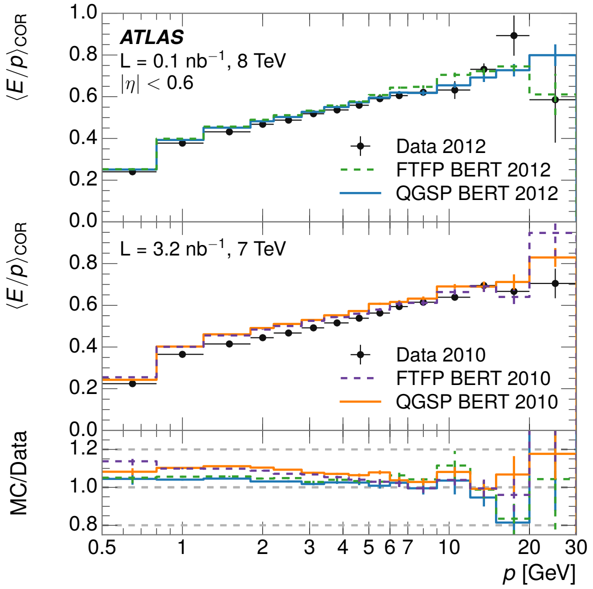
- All sub-detectors used in muon reconstruction (different materials, technologies)

Excellent agreement!

Simulation physics validation: HEP experiments – physics runs



ATLAS 2010/2012 Collider Data →



MC-to-data ratio: calorimeter energy / tracker momentum (single tracks, minbias samples)

- Demonstrates excellent modeling of hadron energy response linearity after calibration, using two independent measurements: calorimeter energy and tracker momentum

MC models data within < 5 % above 0.5 (1) GeV for ATLAS (CMS)

Applications of detector simulation to HEP Collider experiments

Data analysis, detector design and optimization, software & computing design, development and testing

Applications of simulation to data analysis

A few examples of applications to data analysis and interpretation:

- Data-driven methods
  - Techniques applied to real collider data to measure ***physics backgrounds, calibration & alignment factors, resolutions, identification & reconstruction efficiencies, fake rates***, etc
  - Based on detector properties, conservation laws, mathematical tools and analysis
    - Applied to detector-level data and detector-level simulated data as if it were real data
- Closure tests
  - Verify data-driven measurements are correct within the quoted uncertainties
    - Comparing detector level MC measurement with MC truth information

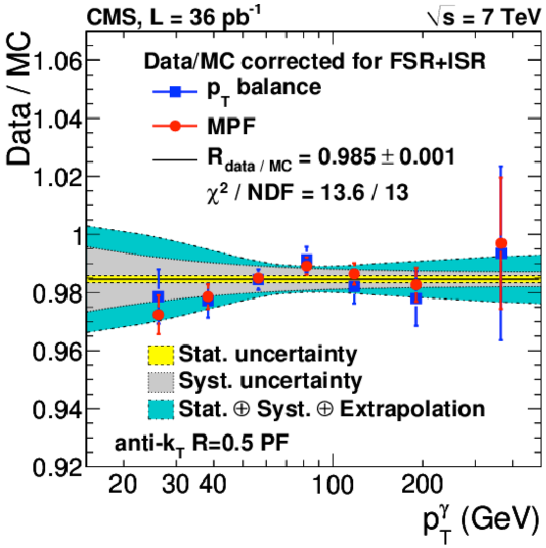
      $T = (MC^{reco\text{-}level} - MC^{truth}) / MC^{truth} \sim 0$ within the uncertainty of the method
- Modeling of signal samples
  - SM precision measurements (i.e. top, W/Z/Higgs), BSM searches
  - Fast simulation to scan large theory parameter space (i.e. SUSY)

Applications of simulation to data analysis – data-driven methods

Corrections in data analysis mostly from MC truth with small "scale factors (SF)"

- SF calculated as ratio of data-driven measurements in detector-level collider data and MC
- The trick is that systematic uncertainties "cancel" in the SF ratio – same method!

- Jet energy response (R_{jet}) or " jet energy scale" (JES)
  - $R_{jet}^{truth\text{-}MC} = p_T^{jet\ reco\text{-}MC}/p_T^{jet\ particle\text{-}level\text{-}MC}$
  - Data-driven methods use di-object p_T balance: multijet, γ+jets, Z+jets samples (conservation laws)
  - $R_{jet} \sim p_T^{jet}/p_T^{\gamma,Z}$ and $SF=R_{jet}^{reco\text{-}data}/R_{jet}^{reco\text{-}MC}$



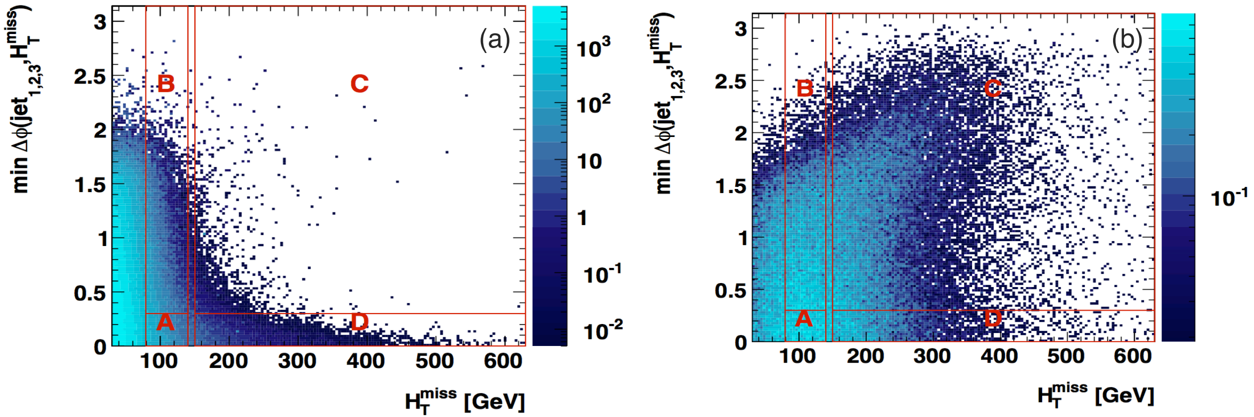
$\boxed{JES = R_{jet}^{truth\text{-}MC} \times SF, \text{ with } SF \sim 0.98 +/- 1\text{-}2\%}$

Accuracy improves as SF→1 within a small uncertainty – excellent MC modeling of the data

Applications of simulation to data analysis – data-driven methods

- Design of control sample/region (CS/CR) and methods for background estimation

Example: QCD background for SUSY searches in many jets + E_T^{miss} final state



CMS Simulation
(a) QCD only sample
(b) SUSY signal sample

Illustration of "factorization" or "ABCD" data-driven method for background estimation

– Simulation used to design CS/CR sample selection (signal depleted)
– Simulation helps develop the data driven method
  - Identify control regions A, B, D and signal region C
  - Decide on function to fit to Δϕ versus E_T^{miss} and study systematics

$$N_C = (N_B/N_A) \times N_D$$

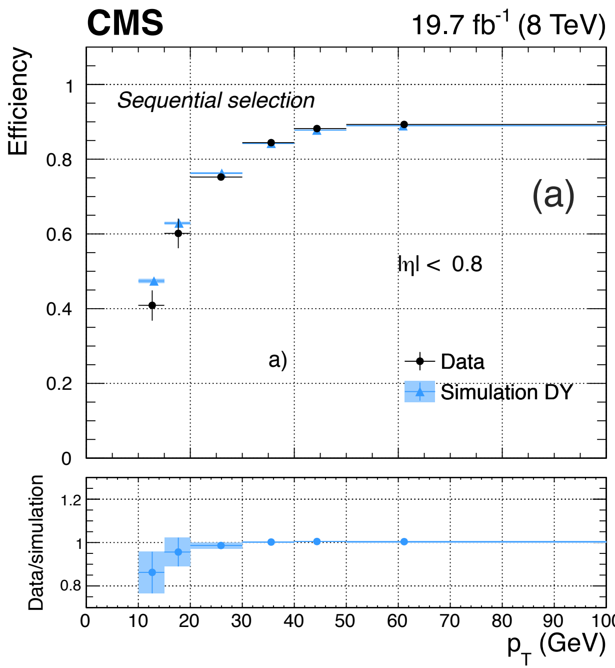
$$N_C = f(H_T^{\text{miss}}) \times N_D$$

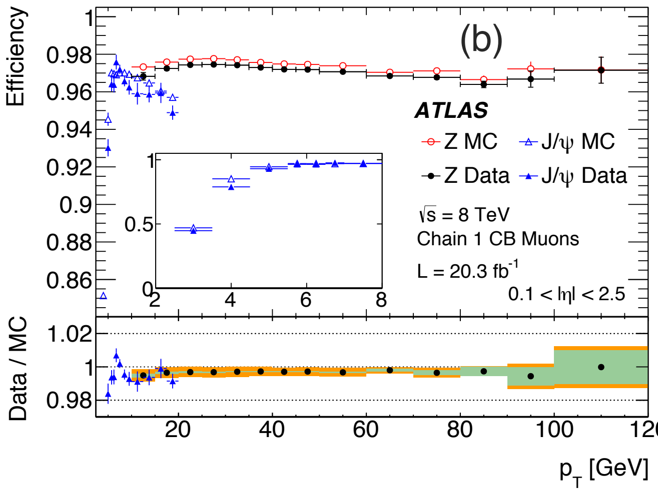
For uncorrelated, or correlated variables

Applications of simulation to data analysis – data-driven methods

- Tag-and-probe (tight-and-loose) method for measurements of efficiencies and fake rates

  Basics: A priori knowledge of an identified reconstructed physics object or "tagged object", then measure fraction of times a software algorithm identifies and reconstructs a "probe object" correctly





(a) CMS electron identification efficiency

(b) ATLAS muon reconstruction and identification efficiency

Method applied to samples:

$Z \to e^+e^-/\mu^+\mu^-,\ J/\psi \to \mu^+\mu^-$

Applications of simulation to data analysis – closure tests

Data-driven methods need to be demonstrated with “closure tests” (T)

- Lack of closure (T$\neq$ 0, outside error band)
  - Indicates the need to go back to the drawing board and understand biases in the procedure excellent MC modeling needed!
- Limitations of simulation at D0 (early 1990’s)
  - **Geant3:** approximate geometry, average material, partial validation of response linearity with data, showers at 95% of total energy deposited (soft contributions, out-of-cone effects missed)
  - **Parametrized “a la CDF” simulation not viable**: no central magnetic field until 2001 $\Longrightarrow$ no single particle response measurement for response tuning

Cause of delay in a number of physics measurements

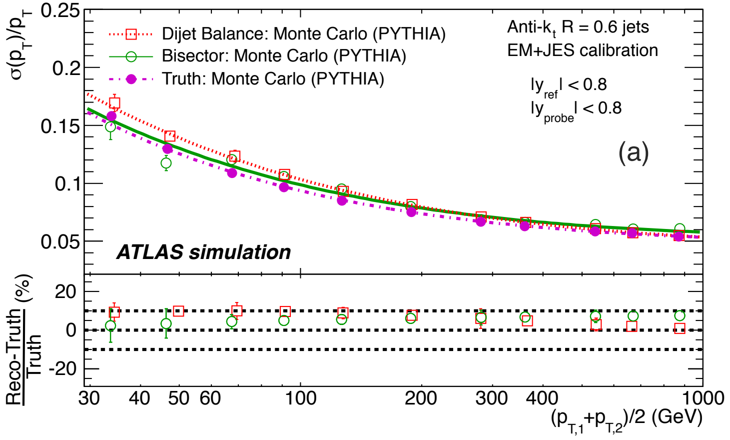
Jet cross sections and other QCD measurements –delayed 1992 $\Rightarrow$ 2000 until JES error $\leq$ 3%

(Lack of large/accurate MC samples to demonstrate data-driven methods by closure for JES)

Applications of simulation to data analysis – closure tests

- Verify data-driven methods are accurate within quoted uncertainties

  T = [(data-driven prediction) – (MC truth value)] / (MC truth value)

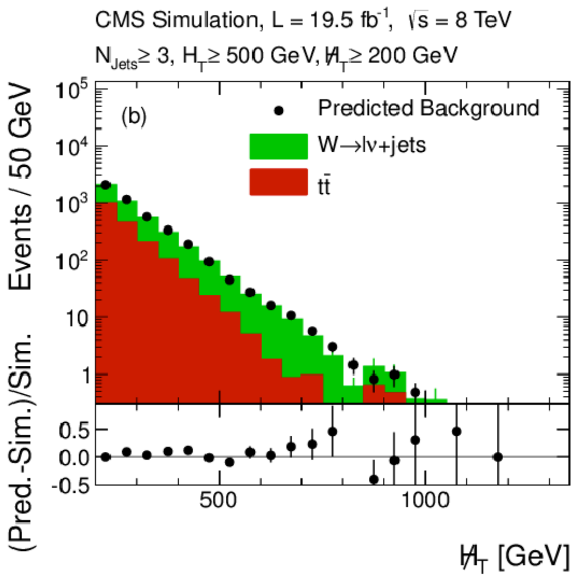


Jet energy Resolution

MC data-driven prediction from dijet asymmetry:

$A = (p_T^{jet\,1} - p_T^{jet\,2}) \,/ =(p_T^{jet\,1} + p_T^{jet\,2})$

Method closes within < 5%



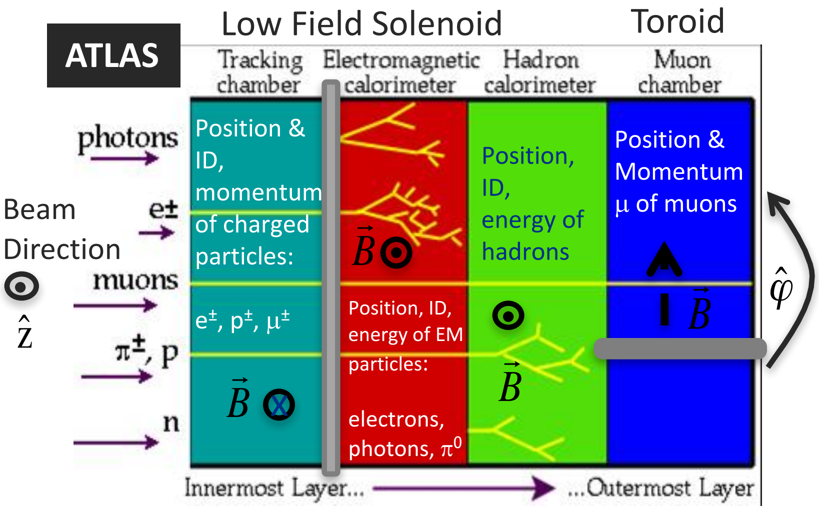
***$W \rightarrow e/\mu\,\nu$* and *$t\bar{t}$* backgrounds to multijet + H_T^{miss} SUSY search**

MC data-driven prediction from inclusive $l\nu$ sample after analysis cuts and lepton efficiency corrections

Closure within stat. errors

Simulation in detector design and optimization

To design a HEP detector, different components, sizes, and are modeled and optimized in simulation for best physics performance



Tracker (in Si detector) optimized varying pixel and strip density, number of layers, angular coverage, amount of material

Calorimeter optimized varying angular coverage and hermeticity, transverse granularity, longitudinal segmentation, materials

Muon system optimized varying wire chamber density, number of layers in the radial direction, angular coverage.

More powerful or weaker magnets allow for more compact (CMS) or larger (ATLAS) detector designs

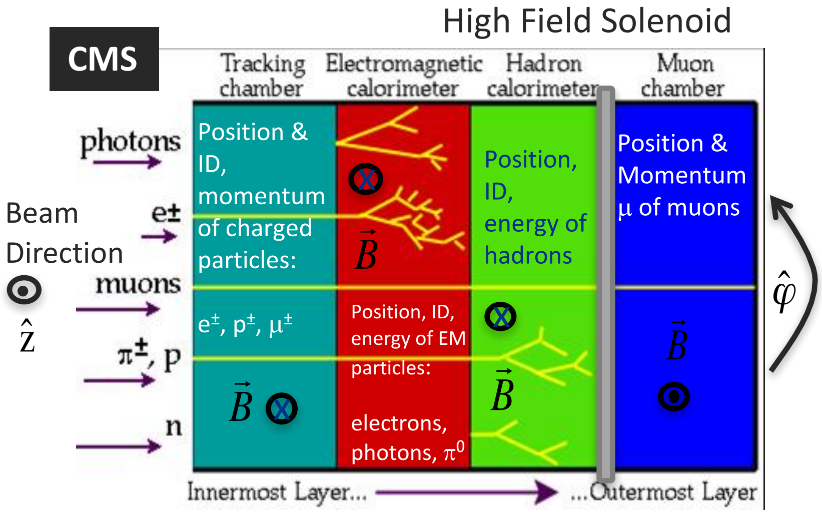
MC campaigns consist of millions of events generated with different detector scenarios

– Make the case for a design, optimize parameters for best physics, impact of de-scoping

(Interesting: detector configurations also adapt to play to the strengths of the Geant4 simulation toolkit)

Simulation in detector design and optimization

To design a HEP detector, different components, sizes, and are modeled and optimized in simulation for best physics performance



Tracker (in Si detector) optimized varying pixel and strip density, number of layers, angular coverage, amount of material

Calorimeter optimized varying angular coverage and hermeticity, transverse granularity, longitudinal segmentation, materials

Muon system optimized varying wire chamber density, number of layers in the radial direction, angular coverage.

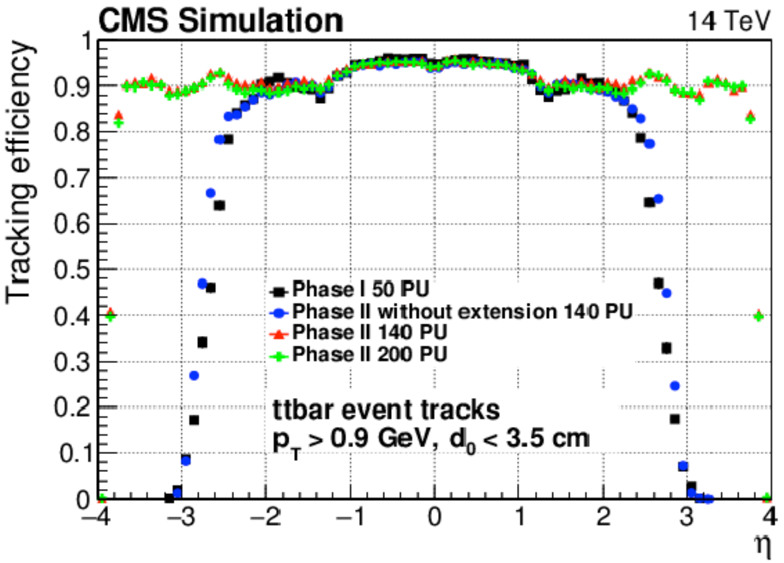
More powerful or weaker magnets allow for more compact (CMS) or larger (ATLAS) detector designs

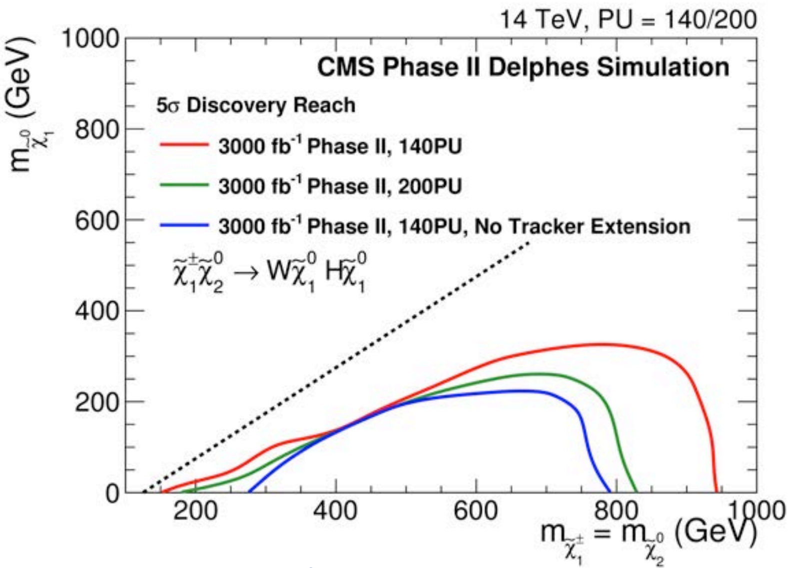
MC campaigns consist of millions of events generated with different detector scenarios

- Make the case for a design, optimize parameters for best physics, impact of de-scoping

(Interesting: detector configurations also adapt to play to the strengths of the Geant4 simulation toolkit)

Simulation in detector design and optimization





CMS tracking efficiency vs. η for various tracker design options and pileup scenarios for HL-LHC

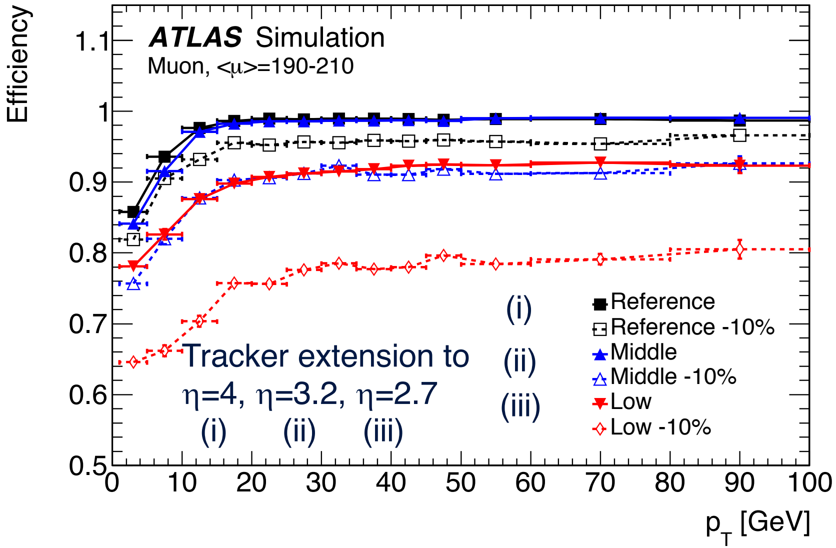
- 2021 detector: 50 (140) pileup events in black (blue)
- 2026 detector: tracker extension to η=3.8 with 140 (200) pileup events in red (green)

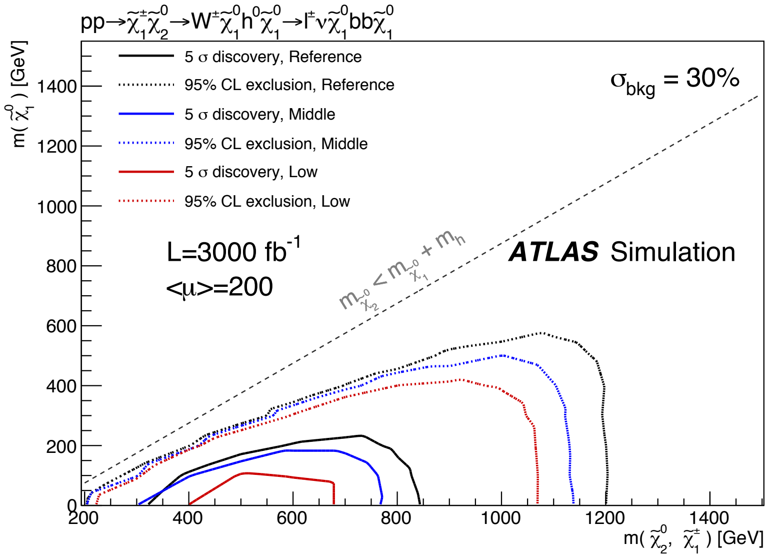
Limit on SUSY $\tilde{\chi}_1^+\tilde{\chi}_2^0$ pair production for various tracker design options and pileup scenarios for HL-LHC

- Mass reach for $\tilde{\chi}_1^+\tilde{\chi}_2^0$ increases from ~750 to ~950 GeV with the tracker extension, reduces back to ~800 GeV when pileup events increase from 140 to 200

Absolute requirement for every HEP experiment seeking approval from funding agencies

Simulation in detector design and optimization





ATLAS muon reco + ID efficiency vs. p_T for various HL-LHC tracker design options (200 pileup events)

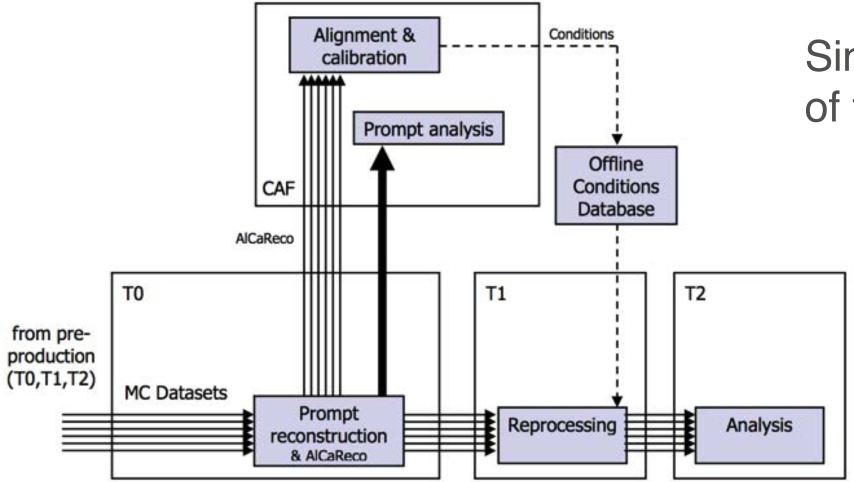
- Reference to Low de-scoping cost 10% in muon efficiency (black squares vs full red triangles)

Limit on SUSY $\tilde{\chi}_1^+\tilde{\chi}_2^0$ pair production for various tracker designs and 200 pileup events for HL-LHC

- $\tilde{\chi}_1^+\tilde{\chi}_2^0$ discovery reach improves from mass ~700 GeV to ~850 GeV for Low to Reference tracker designs

Absolute requirement for every HEP experiment seeking approval from funding agencies

Simulation in software and computing design and testing



Simulation is essential to develop each element of the workflow and data flow for data handling

- Worldwide LHC Computing Grid (WLCG) divided in four tiers: 0, 1, 2, 3
- Each tier performs difference services: acquisition, reconstruction, simulation, storage, data analysis

Combined procedure tested in Computing, Software, and analysis challenges (CSA) in CMS

- System stress tested at 25%, 50%, and 75% capacity in 2006, 2007, and 2008
- 150 million events simulated, trigger rates modeled, and data reconstructed, skimmed, calibrated
- Data transfers between centers, monitoring of event file size, memory and CPU consumption

The realism of these tests resulted in computing systems performing as predicted

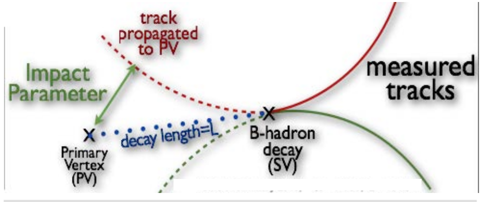
Modeling of particle and event properties and kinematics

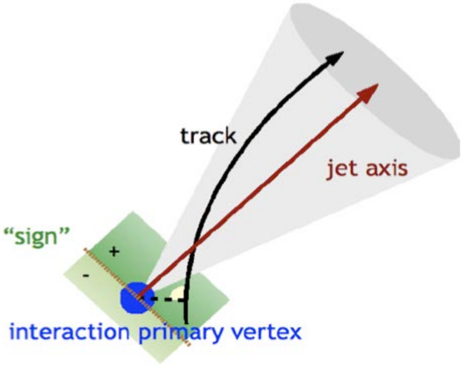
Tagging of heavy quarks, W, Z, and photon event distributions, missing transverse energy distributions

Modeling of particles and event properties: b jets

Modeling of b-jet reconstruction/identification is a critical simulation benchmark

- SM measurements: top decays to b, W and flavor tied to EWSB mechanism
- BSM searches: SUSY and EWSB related through hierarchy problem





b-jet identification (b-tagging) depends on impact parameter of charged-tracks and reconstructed decay vertices in the jet, lepton presence

Impact parameter (IP) is the point of closest approach between the track and the primary vertex

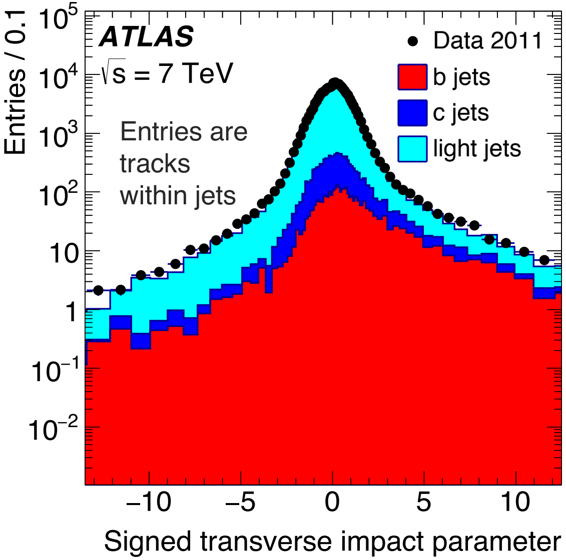
- b-quarks have positive IP while light jets have IP ~ 0
  - Resolution effects give positive and negative values in a real detector

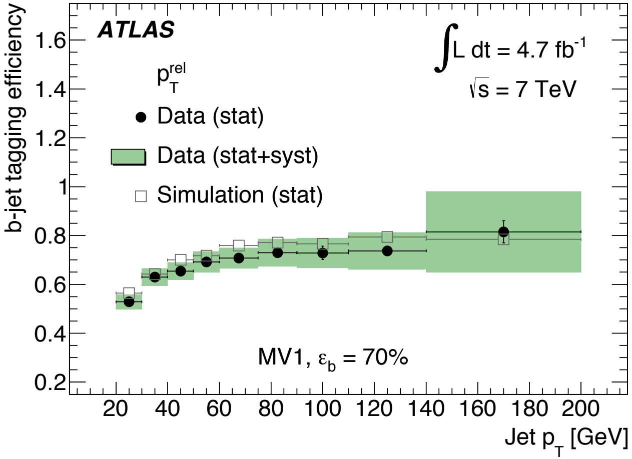
Good simulation of IP variables is necessary for accurate measurement of b-tagging efficiencies and fake rates using data-driven methods

- Derived from data-driven methods applied to samples of jets with a muon

Need excellent modeling of material budget, energy loss, ionization, multiple scattering, noise, pileup mainly in tracker

Modeling of particles and event properties: b jets





Data-to-MC scale factors used to adjust MC truth efficiencies used in data analysis

Mis-tag rates: light jet passing for b-jet (not shown)

- From negative taggers
- Modeling tricky – tracks come from tail of IP distributions
- data-to-MC fake rate ratio deviates from 1.0 by:

  CMS: 20% for 0.01-0.03 mis-tag probability
  ATLAS: factor 2-3 for mis-tag rate in 0.002-0.005 range

ATLAS Signed Transverse IP significance, $S_{d_0} \equiv d_0/\sigma_{d_0}$ for heavy and light quarks in a di-jet sample

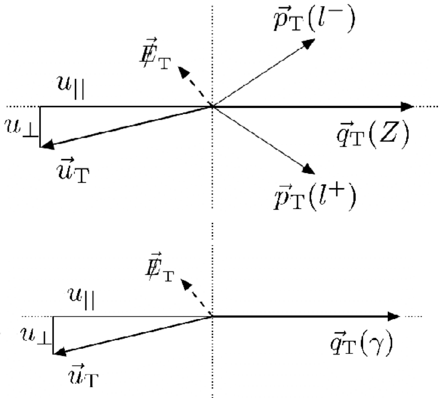
- b, c quark distributions positive and asymmetric, light quark distribution slightly positive and almost symmetric
- Excellent Data/MC agreement except in tails of distribution – resolution smearing more difficult to model
- Simulation models b-tagging efficiency within <5% (absolute)

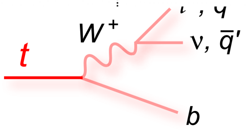
CMS 3D Impact Parameter distribution and efficiencies show similar agreements with MC (not shown)

Modeling of particles and event properties: W, Z, photons

Gauge bosons are at the core of SM measurements (W, Z, top mass and properties) and contribute backgrounds to most BSM searches

- Topologies and kinematics of W/Z/γ + jets events must be modeled with high accuracy
  - Generators are limiting factor for accuracy, particularly in multi-jet events with heavy flavor





W/Z + jets background typically estimated from data in BSM searches

- Simulation is used to study (di-)lepton + jets control samples to design data-driven methods (distribution shapes, variable correlation, etc)

MC truth used to predict sub-dominant SM backgrounds

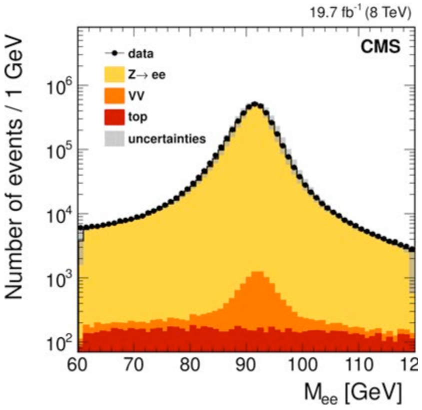
- $VH, t\bar{t}, t\bar{t}Z, t\bar{t}W, t\bar{t}H$

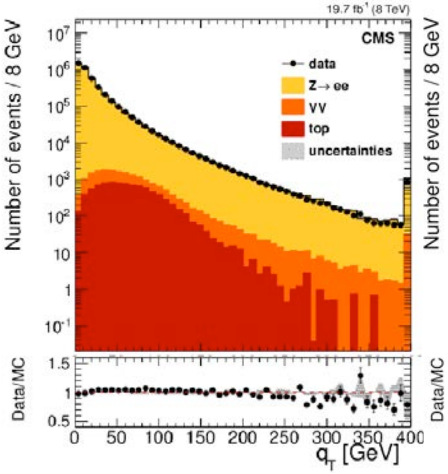
Detector simulation accuracy enters through modeling of γ, e/μ, jets, and E_T^{miss} (E_t^{miss} coming from neutrino in W decay, energy resolution in hadronic recoil)

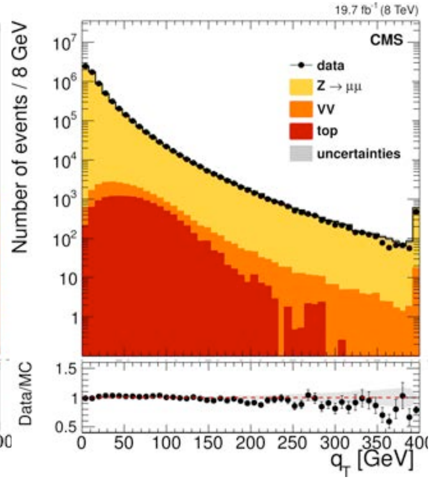
- material budget in tracker, EM and hadron calorimeter showers

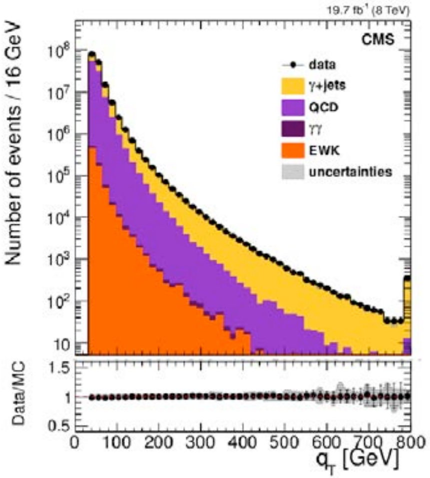
$$M_T^W = \sqrt{2p_T^l p_T^\nu(1-\cos(\phi_l-\phi_\nu))} \qquad M^Z = \sqrt{2p^{l1}p^{l2}(1-\cos(\phi_{l1}-\phi_{l2}))}$$

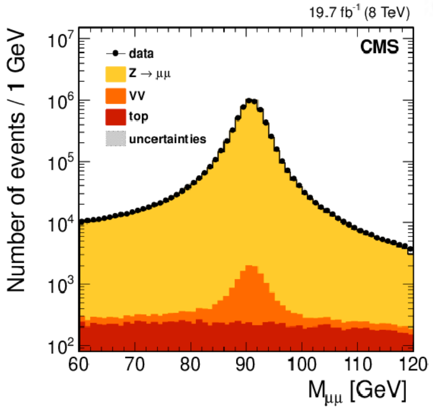
Modeling of particles and event properties: W, Z, photons







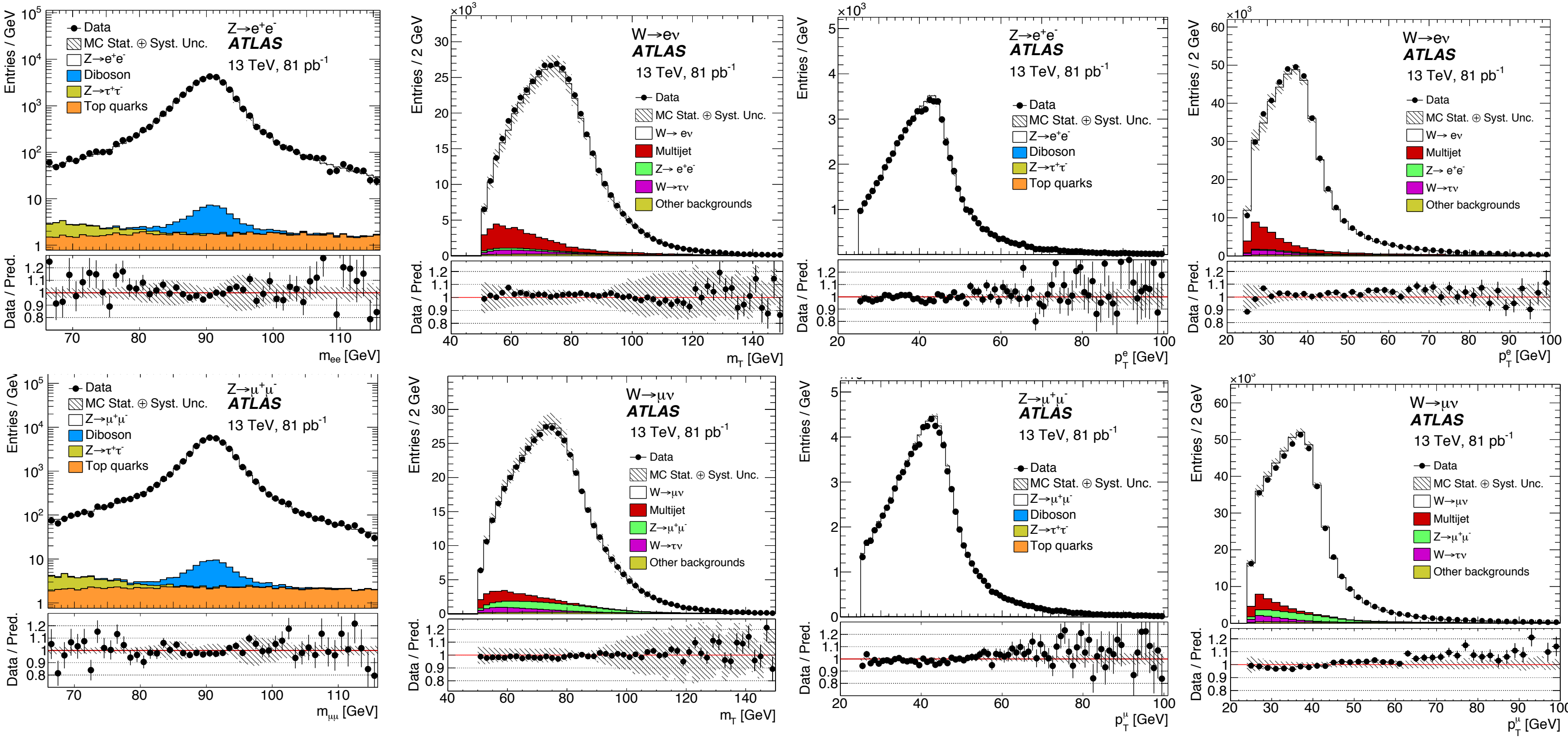




CMS Z mass and Z/γ p_T (q_T in plots) distributions

- Z mass measured from di-lepton p_T's: electrons and muons
- Agreement is excellent within statistical and systematic uncertainties
- In the range of $p_T > 200$ GeV for the Z + jets sample, the simulation overestimates the data by a difference that grows linearly with p_T
  - Most probably coming from generator mis-modeling of event topology and kinematics

Modeling of particles and event properties: W, Z, photons



ATLAS W/Z mass and e/μ p_T distributions in the electron and muon W/Z decay channels

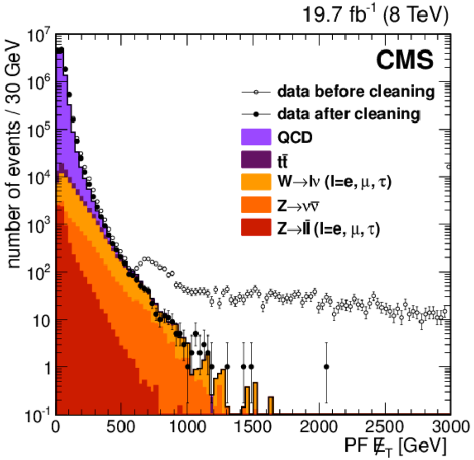
- Impressive agreement within <10% in the domain ranges with good statistics

Modeling of particles and event properties: missing E_T

Event missing transverse energy: E_T^{miss} or $\vec{\not{E}}_T = -\sum_{\mathbf{particles}} (p_x\hat{\mathbf{i}} + p_y\hat{\mathbf{j}})$

Modeling E_T^{miss} is among the most challenging simulation tasks: depends on all types of particles, hadronic showers from jets, and un-clustered energy

- Paramount importance in BSM SUSY, ED, dark matter searches, Higgs characterization
- Intrinsic low-med (high) E_T^{miss} in SM (BSM) searches, or E_T^{miss} from detector mis-measurement

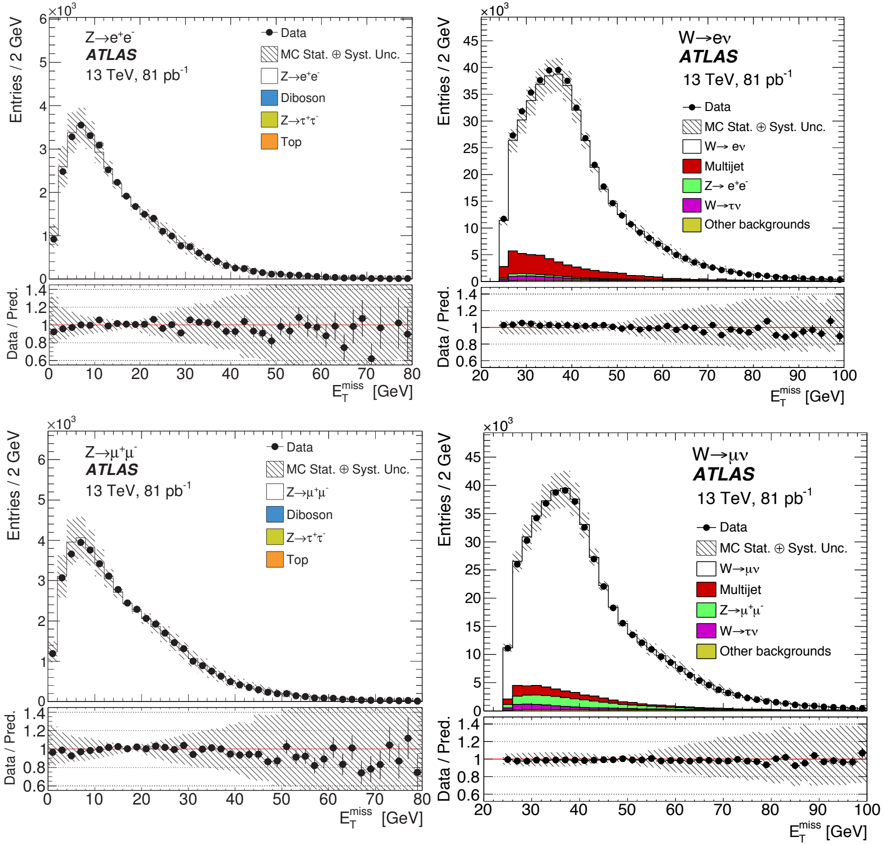


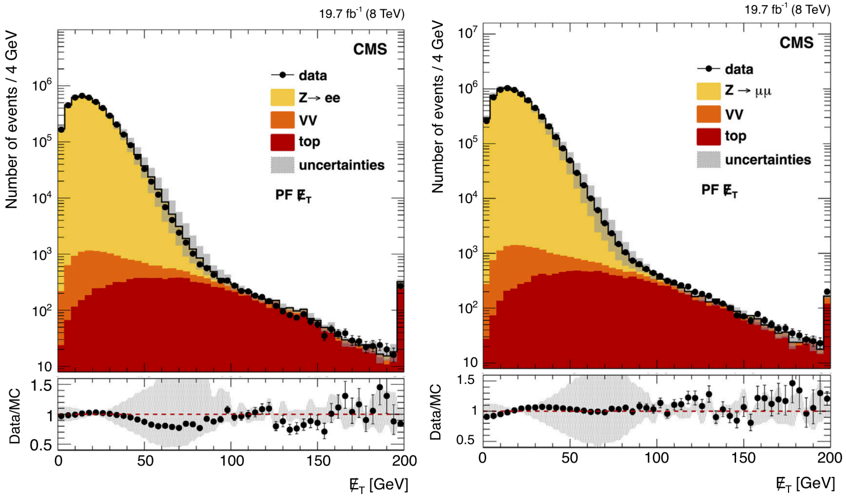
E_T^{miss} distribution for CMS di-jet events before and after applying the software algorithms to remove events with spurious E_T^{miss}

- Agreement is excellent > 500 GeV and worsens below 500 GeV as the QCD contribution increases and becomes dominant
- E_T^{miss} QCD background estimates in SM/BSM analyses typically not taken from MC
  - Shower fluctuations and un-clustered energy not modeled accurately enough
  - Impossible to demonstrate that all sources of spurious events in the tails have been identified and modeled in the MC with the correct rates

Low-med E_T^{miss} from invisible decays (neutrinos) better modeled than high E_T^{miss} tails in multi-jet samples with origin in resolution or detector malfunction

Modeling of particles and event properties: missing E_T

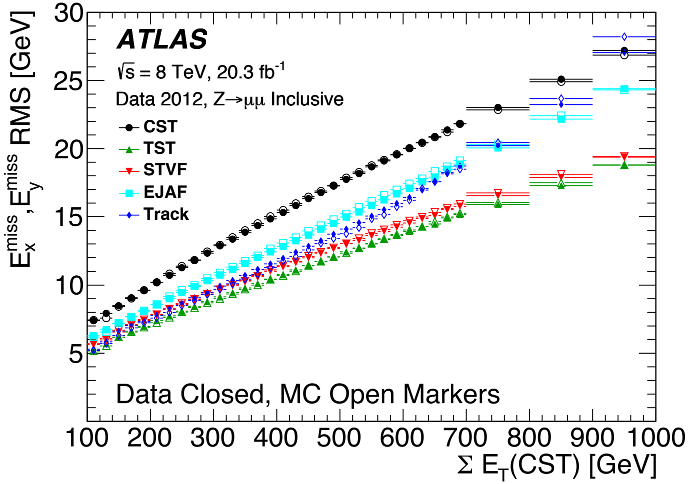




ATLAS and CMS E_T^{miss} In W/Z + jets samples

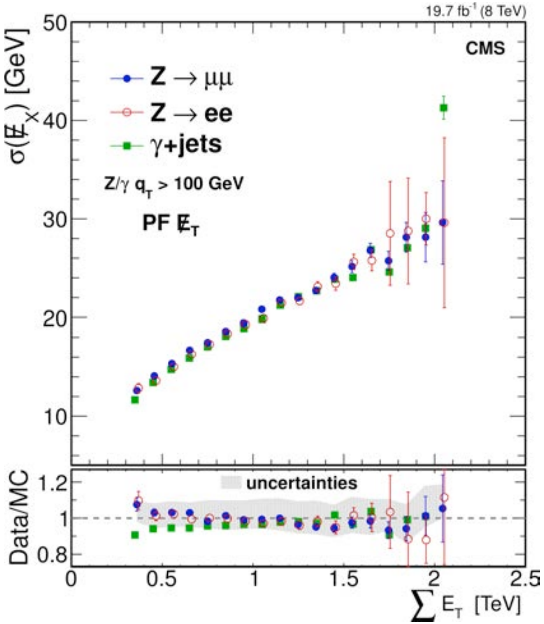
- Real and spurious E_T^{miss} in W + jets
- Spurious E_T^{miss} in Z + jets
- MC models E_T^{miss} in Z + jets to within ~20% in CMS and in W/Z + jets within ~10% in ATLAS
- In both experiments, systematic uncertainties grow above 50% in the range where hadronic shower mis-measurement dominates (~50-90 GeV)

Modeling of particles and event properties: missing E_T



ATLAS RMS distribution from x and y components of E_T^{miss} vs. scalar sum of the E_T of the physics objects in a $Z(\mu\mu)$ + jets sample

- CST, TST, STVF, EJAF refer to different algorithms to reconstruct/calibrate un-clustered energy
- Data-to-MC E_T^{miss} resolution agree within < 5%



CMS RMS E_T^{miss} projections along x and y vs. scalar sum of the E_T of the physics objects in $Z(ee/\mu\mu)$ + jets and γ + jets samples

- Photons and leptons not in the scalar sum
- E_T^{miss} resolutions modeled within a ~10% accuracy

Simulation and jet cross sections

The emblematic example of the jet cross section measurements

Simulation and jet cross sections

Jet cross sections are useful to illustrate the impact of simulation in data measurements

- Dependence on single source of systematic uncertainty: the jet energy scale (JES)
- JES accuracy relies on the accuracy of
  - Data-driven methods, modeling of hadronic response and resolution in parametrized or full simulation
    - 1-10 GeV hadrons difficult to model – affect even high p_T jets because energy of constituents grows slowly, approximately as square root of jet energy

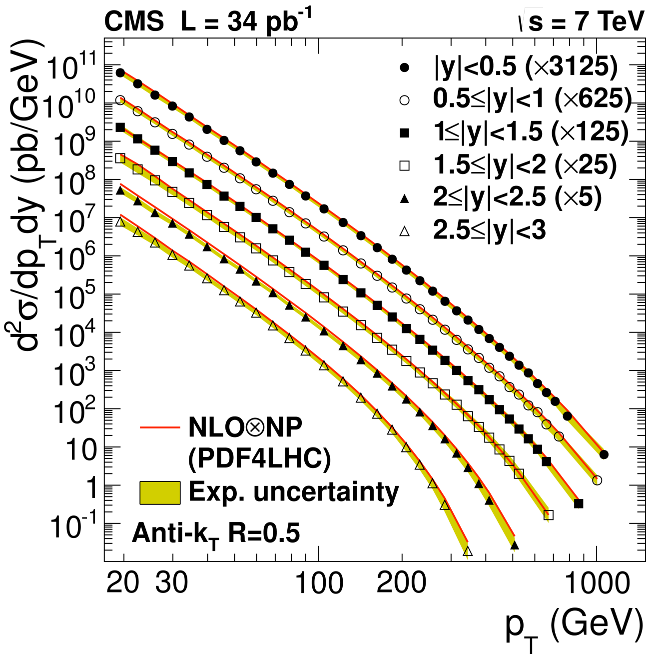
Plots in next slides display "first published measurements", not "latest and most precise" of each experiment: ATLAS, CMS, CDF, D0

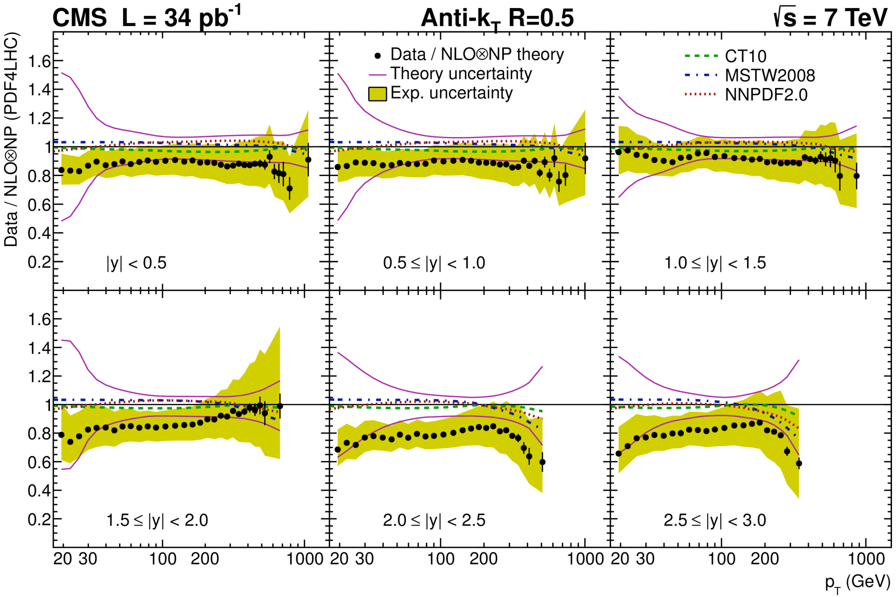
- Comparisons evaluate accuracy of NLO-QCD theoretical predictions and not of generators or detector simulation tools (not MC-to-data ratios like previous plots!)
  - Data corrected for detector effects to "particle level" – equivalent to all orders theoretical predictions including hadronization effects

Highlight – the role of simulation in the relationship between the size of systematic uncertainties and the publication timeline

Jet cross sections at the Tevatron: a story about limited test beam data, tuning of parametrized simulation, the long process to develop data-driven methods with little aid from simulation

Simulation and jet cross sections: CMS

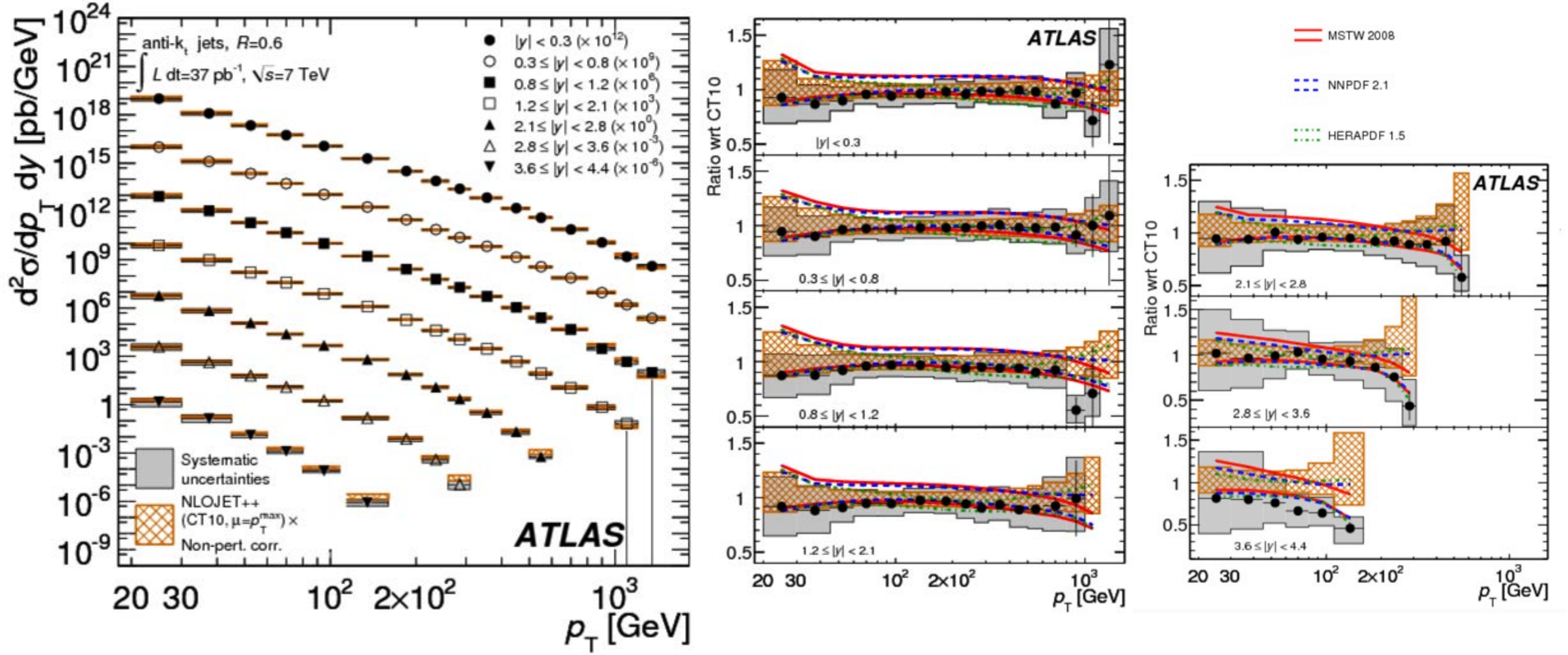




CMS inclusive jet cross sections based on L=34 pb^{-1} from the 2010 run

- Data taking started in 3/2010 and the measurement was published in 6/2011, 7 months after the end of the run
- Extends up to lyl = 3 with 10%–20% uncertainties in the most central and 15%–30% in the most forward regions

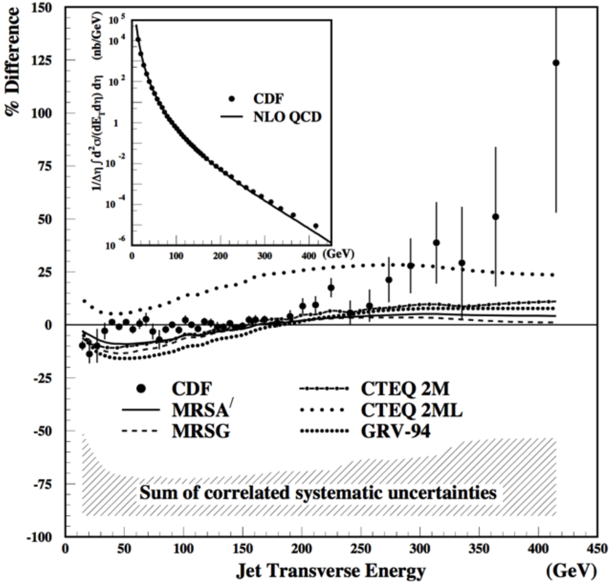
Simulation and jet cross sections: ATLAS

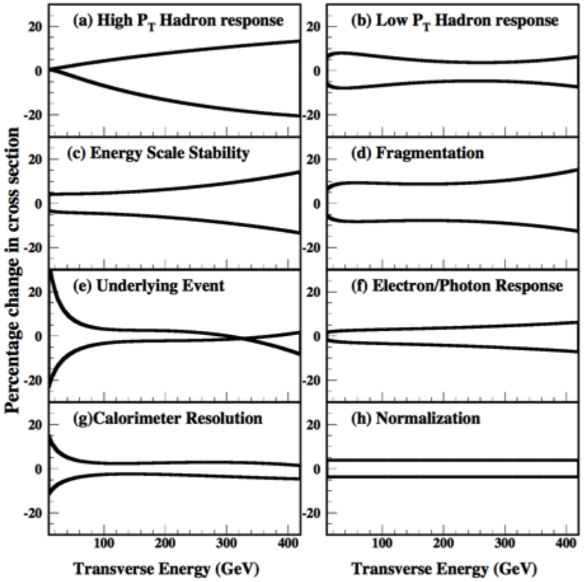


ATLAS inclusive jet cross sections based on L=37 pb^{-1} from the 2010 run

- Full 37 pb-1 dataset published in 4/2012, but intermediate result with half data published in 10/2010
- Up to lyl = 4.4 with uncertainties similar to CMS in the most central and 12-40% in the most forward regions

Simulation and jet cross sections: CDF





CDF extension to high η:

Needed QFL (Parametrized Simulation) tuned in the End Plug Calorimeter

- GFLASH parametrization in 2002–2003.

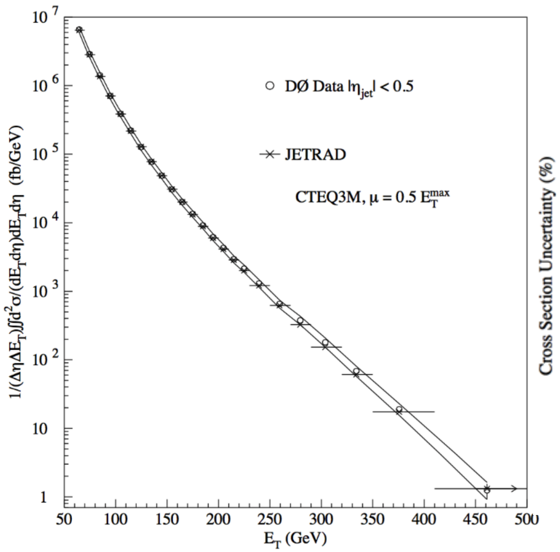
Also, JES in forward regions use di-jet balance techniques, affected by resolution biases

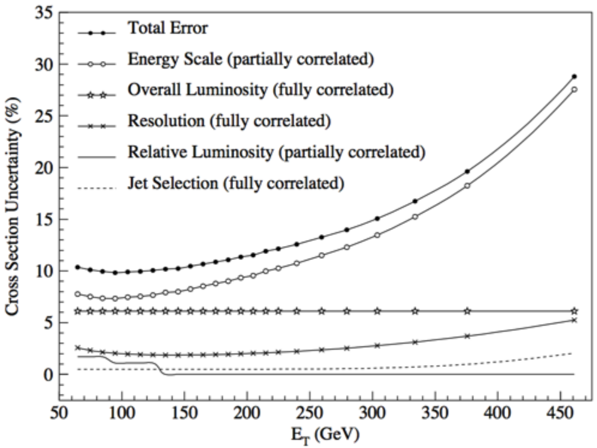
- Understood with large and accurate MC samples

CDF inclusive jet cross sections based on L=19.5 pb^{-1} from run 1a

- Full 19.5 pb-1 run 1a dataset published in 1/1996, five years after the start of the run
- Measurement in the central 0.1<|η|<0.7 region with uncertainties in the 20-35% range
- CDF published the run 2 inclusive jet cross sections in |η|<2.1 (2008)

Simulation and jet cross sections: D0





D0 inclusive jet cross sections based on L=92 pb^{-1} from full run 1

- D0's first inclusive jet cross section publication came out in 1999, eight years after the start of the run
- Restricted to the $|\eta|<0.5$ region, systematic uncertainties in the 10-30% range
- In 2001, D0 extended the measurement to the $|\eta|<3$ region
- D0 published the run 2 inclusive jet cross sections in $|\eta|<2.4$ (2011)

Parametrized Simulation not viable
- No solenoid in the tracker
- Scarce test beam data

In situ calibration based on data-driven methods (only option)
- Developed from scratch
- Lengthy process without the aid of reliable MC samples

D0 did not deliver a result with competitive uncertainties until 1999
- While CDF published with large uncertainties in 1989 (Run 0) and 1992 (early run 1 data)

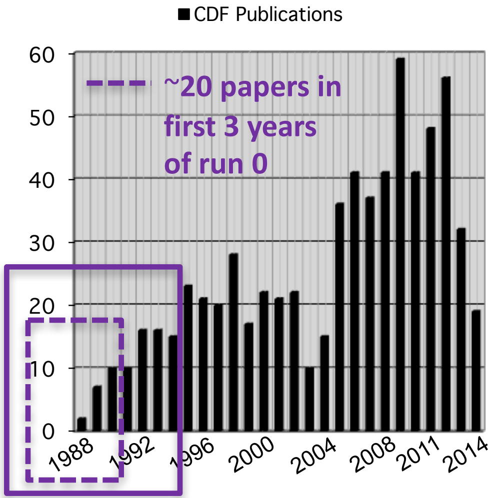
Simulation and publication turnaround

The CDF, D0, ATLAS, and CMS examples

Simulation and publication turnaround

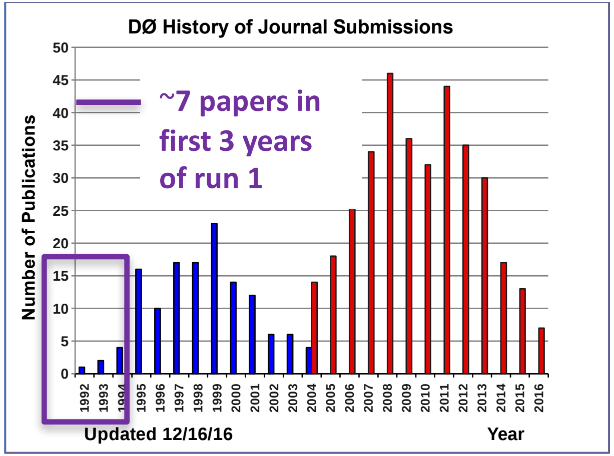
The Tevatron program coincided with the dawn of the era of detector simulation toolkits –1988 to the start of LHC experiments

– Transition between sporadic to systematic use of Geant3-based full simulation by the end of run 2



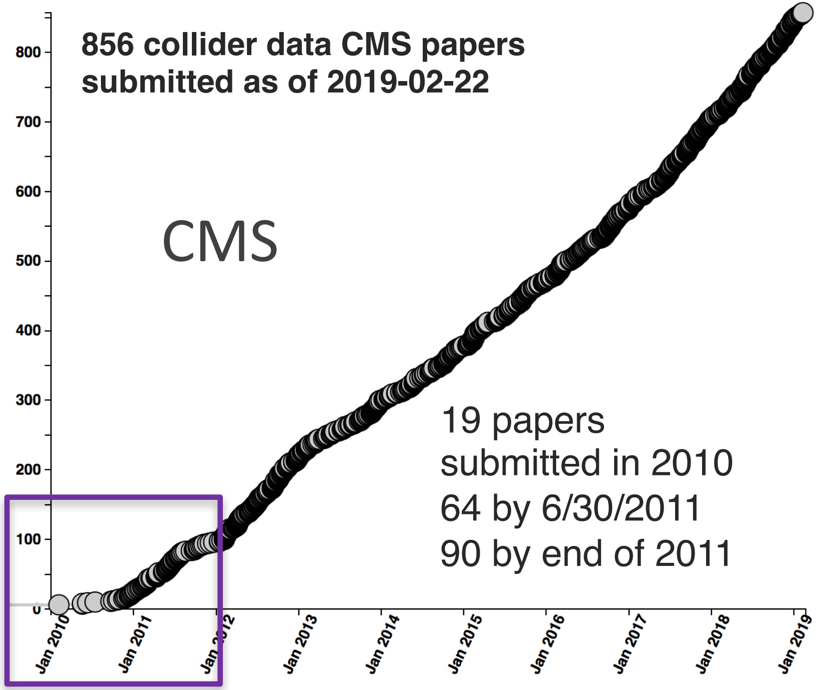
CDF (left) **D0 (right)**

- CDF started to operate in 1988 (run 0), D0 in 1992 (run 1)
- CDF first physics paper in 1988, and jet cross sections in 1989 (one and two years after start)
- D0 first physics paper in 1994, and jet cross sections in 1999 (2 and 7 years after start)



CDF published faster than D0 in their first run (run 0 for CDF and run 1 for D0) – one of the reasons being the presence of a solenoidal field which allowed for calibration and parametrized simulation tuning

Simulation and publication turnaround



Simulation shortened significantly the detector and physics commissioning time

- Computing model, software worked basically as in designed specifications
- Reconstruction software, calibration and analysis data-driven methods performed out-of-the-box

Examples of papers submitted the first year CMS & ATLAS:

- **CMS:** Dijet cross sections, top pair production, W/Z cross sections, J/ψ and direct photon production, BSM searches for gluinos and leptoquarks
- **ATLAS:** Inclusive jets and dijet cross sections, W/Z cross sections, J/ψ and direct photon production, top pair cross sections, jet shapes measurement

Factors for LHC faster than TeV: thousands vs. hundreds of members, detector & computing technology

– But simulation had a direct impact through the effect on calibration, corrections, analysis methods

Economic impact and cost of simulation in HEP collider experiments

The CMS case, Geant4

Economic impact/cost of simulation in HEP collider experiments

We define "simulation chain" physics generation, interaction with matter (G4), readout modeling, reconstruction, analysis

- Took 85% of CPU resources used by CMS, while G4 module took 40% of total (Run 1, 2)
- ATLAS's simulation application is 8-9 times slower than CMS's and uses significantly more resources than CMS in physics generation
- Rest of resources used in reconstruction and analysis of real collider data

CMS in more detail taken from (analysis of 2012/May 2015-May 2016 periods)

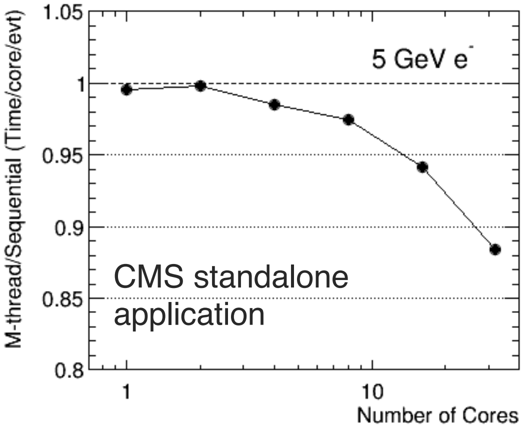
- 540k/860k core months corresponding to 45/70k CPU cores at full capacity (half in G4)
- Purchasing cost is 5/8 million dollars
- Cost of physical hardware including life-cycle, operation, maintenance
  - 0.9 cents/core hour (FNAL), or 1.4 cents/core hour (what FNAL paid industry in 2017)
- Annual cost of simulation in CMS: 3.5-6.2/5.5-10 million dollars
- Improvements of 1%, 10%, 35% in G4 time performance would render 50-80k, 500-800k, 1.8-2.8M dollars savings to CMS

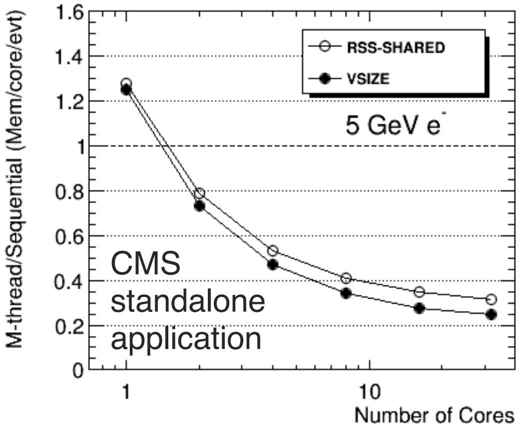
Computing needs of HL-LHC program are 10-100 higher depending on simulation and reconstruction solutions implemented – reconstruction will take a larger fraction (pileup)

Economic impact/cost of simulation in HEP collider experiments

Design, development, validation, operation, support of simulation toolkits, such as Geant4, as well as development of the experiment applications add to the cost

- In 22 years of existence, investment on G4 was ~ 500 person-years or > 100M dollars
- How much more it would have costed to design, optimize, commission, operate detectors, as well as the physics programs without the Geant simulation toolkit?





Geant4 introduced multithreading capabilities in 2013 – event level

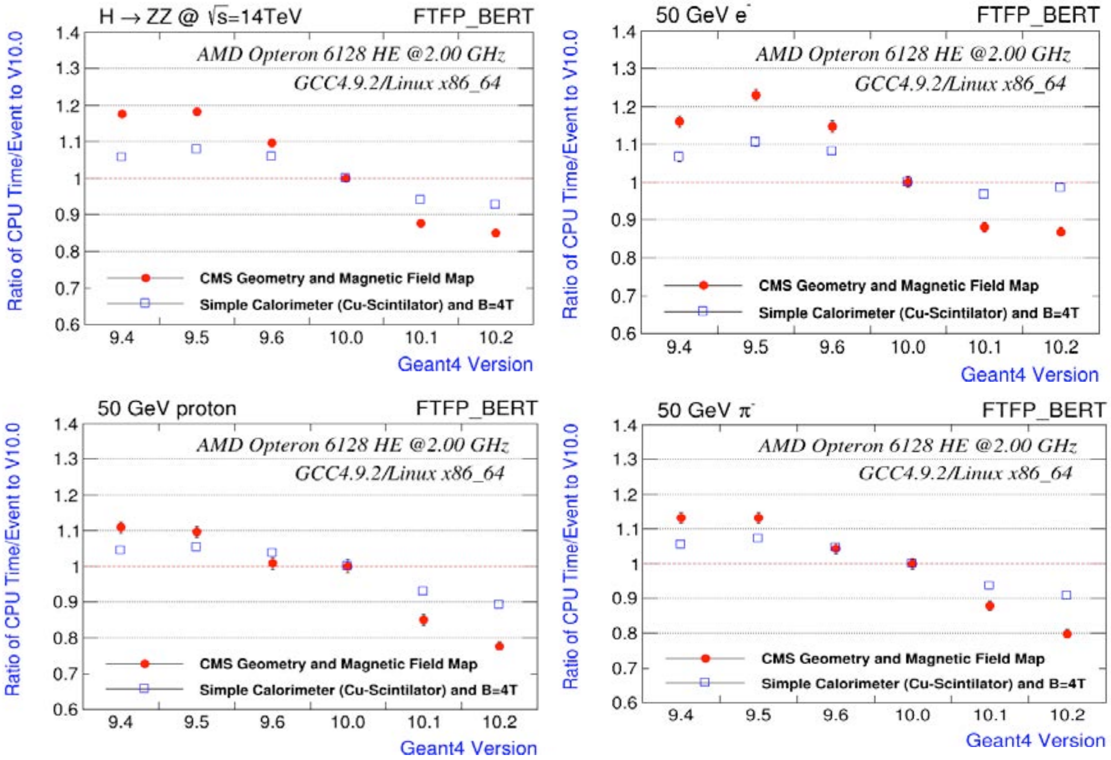
- Time performance does not improve, deviates from perfect scaling: ~10% for 30 cores
- Memory use improves significantly ~170MB in first event, ~30MB by each additional thread

Corollaries: 1- the cost of physics software is a significant fraction of the cost of detectors

2- the cost of simulation and reconstruction should be a factor in detector design

Economic impact/cost of simulation in HEP collider experiments

The G4 Collaboration has gone to great lengths to improve computing performance



During the 2010-2015 period:

- Time performance improvement was of the order of 35% (simple calorimeter & CMS standalone)
- Double digits CPU improvement while physics accuracy also improved

Remember a 35% faster G4 means ~2-3M dollars/year savings in CMS

(or we can do 35% more simulation)

But this is not enough!

The future

Better physics accuracy and increased speed by means of novel programming techniques and modern computing architectures

The future

Next generation HEP experiments will require orders of magnitude more simulated events with improved physics accuracy

- The effort to improve the physics and computing performance of simulation tools (and reconstruction algorithms) require immediate attention
- Transistor density growth is more or less keeping with Moore's law but clock speed has been flat since 2003
  - Leverage core count growth in multicore machines, use new generation coprocessors, re-engineer code using fine grained parallelization for accelerators and HPC systems
  - Use of machine learning techniques to replace the detector simulation step

The simulation community is working hard on improved physics models and software & computing R&D to meet the challenges:

A Roadmap for HEP Software and Computing R&D for the 2020s
(https://arxiv.org/abs/1712.06982)
HEP Software Foundation Community White Paper Working Group - Detector Simulation
(https://arxiv.org/abs/1803.04165)

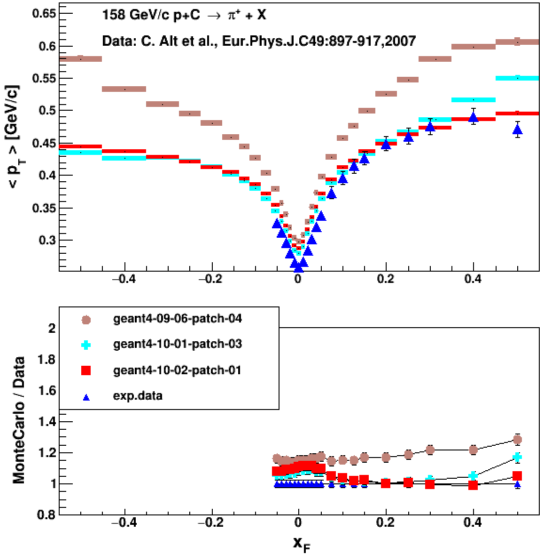
… which is a topic for another seminar

Backup slides

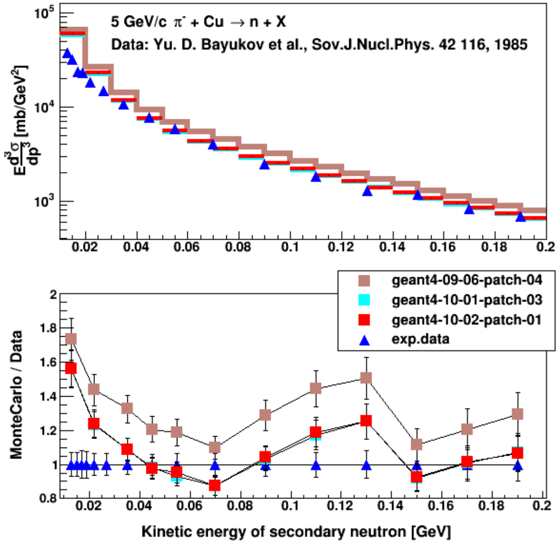
Geant4 physics validation: thin-target experiments

The first step is for the Geant4 team to validate individual physics models with thin-target experiments and tune their associated parameters. Examples:

NA49 experiment



ITEP-771 experiment

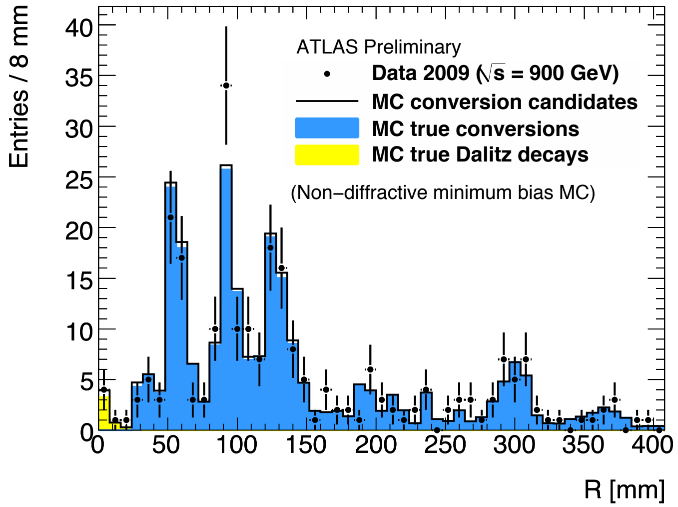


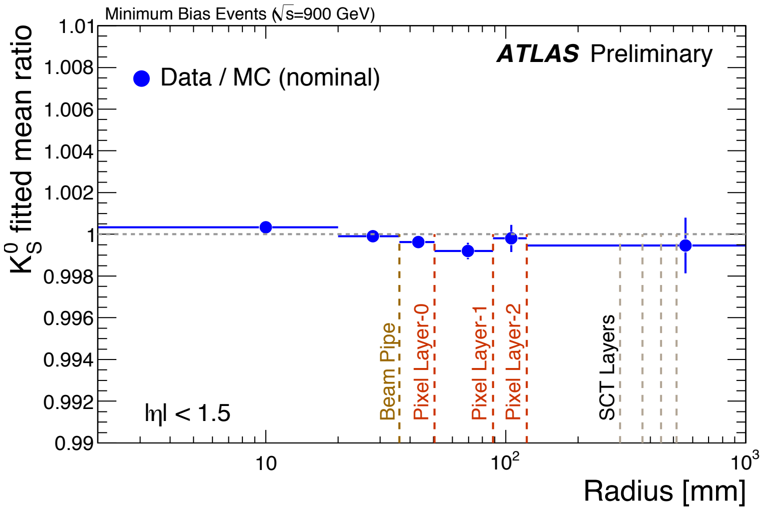
Validation of:

- FRITIOF Pre-compound Model
  - Strings in hadron-nucleon collision, nucleus de-excitation
- Bertini Cascade Model
  - Final states for hadron inelastic scattering (intra-nuclear cascade)

G4 prediction/data ratios have improved over time for each new release

Simulation physics validation: HEP experiments – physics runs





ATLAS: position of γ conversion vertices in the radial direction (data, MC, MC truth)

- γ conversion mis-modeling would affect physics with photons (Higgs, QCD, BSM)

Excellent agreement

ATLAS: K^0_s mass measured from reconstructed tracks versus distance in radial direction

- material modeling affects energy loss and multiple scattering

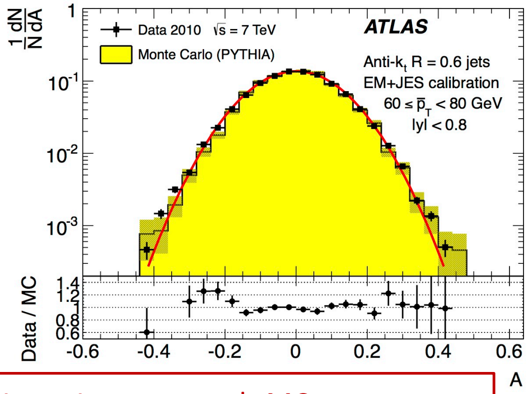
MC models data mass measurement to within < 1%

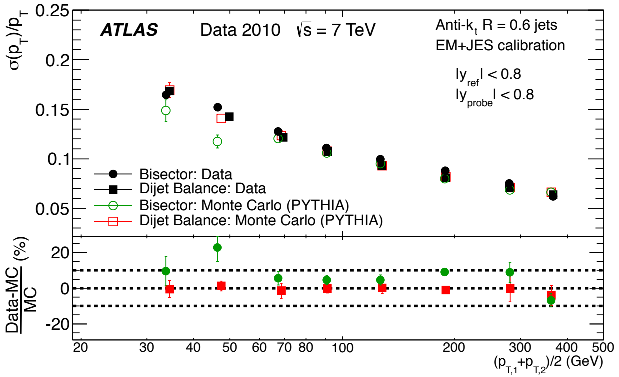
Applications of simulation to data analysis – data-driven methods

Where is the magic ? Most systematic uncertainties are 100% correlated numerator-to-denominator, they cancel in the SF=$X^{reco\text{-}data}/X^{reco\text{-}MC}$ ratio

- X is not measured, but the data/MC ratio, which contains information on MC mis-modeling of X

- Jet energy resolution
  - Asymmetry distribution $A=(p_T^{jet\ i}-p_T^{jet\ j})/ =(p_{Ti}^{jet\ i}+p_{Tj}^{jet\ j})$
  - Relative energy resolution $\sigma(p_T^{jet})/p_T^{jet}= \sqrt{2}\,\sigma(A)$
  - As in the JES case:

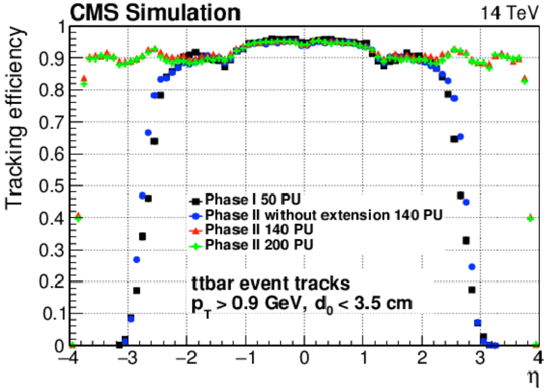




Jet Energy Resolution (JER)= $JER^{truth\text{-}MC}$ x SF
SF ~ 1-1.10 depending on the method

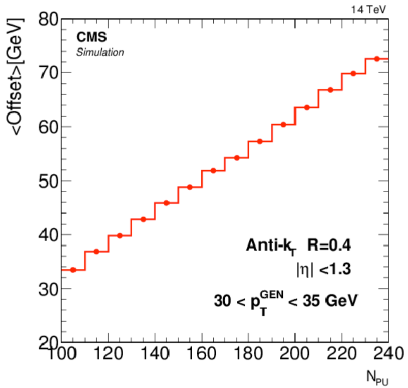
Non-Gaussian tails coming from non-linear energy response are difficult to model

Simulation in detector design and optimization



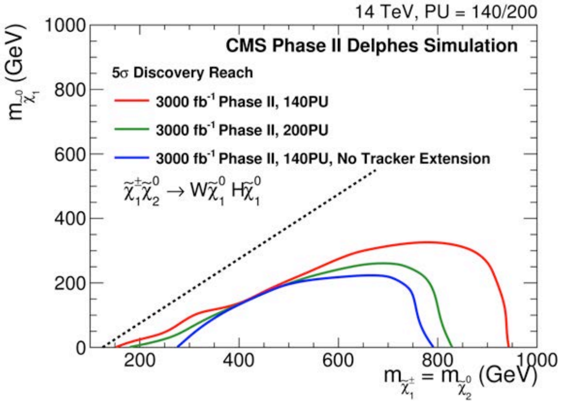
CMS tracking efficiency vs. η for various tracker design options and pileup scenarios for HL-LHC

- 2021 detector with 50 (140) pileup events in black (blue)
- 2026 detector including tracker extension to η=3.8 with 140 (200) pileup events in red (green)



Spurious energy in jets (offset) vs. number of pileup events

- Modifies jet multiplicity of the event, distorts jet energy response, degrades jet energy and missing transverse energy resolutions

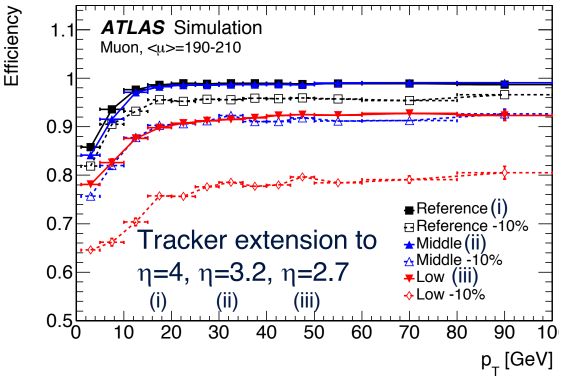


Limit on SUSY chargino-neutralino pair production for various tracker design options and pileup scenarios for HL-LHC

- Chargino, neutralino mass reach increases from ~750 GeV to ~950 GeV with the tracker extension but reduces back to ~800 GeV when the number of pileup events increase from 140 to 200

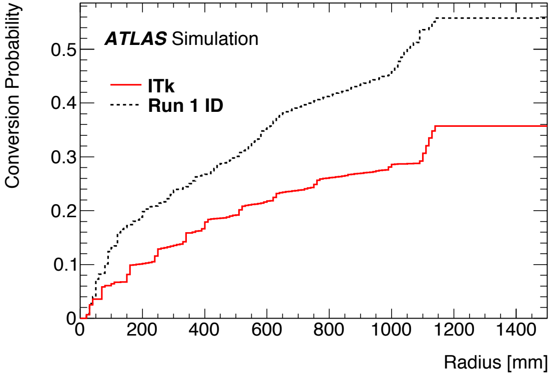
Absolute requirement for every HEP experiment seeking approval from funding agencies

Simulation in detector design and optimization



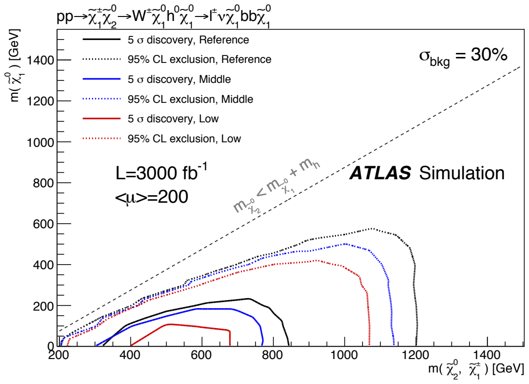
ATLAS muon reconstruction plus dentification efficiency vs. p_T for various HL-LHC tracker design options (200 pileup events)

- De-scoping from Reference to Low scenarios cost 10%in muon efficiency (compare black squares to full red triangles)



Photon conversion cumulative probability vs. the distance from the interaction vertex

- Results for current tracker shown in black and upgraded tracker (iTK) shown in red
- Significant reduction in photon conversion probability is predicted for the iTK



Limit on SUSY chargino-neutralino pair production for various tracker designs and 200 pileup events for HL-LHC

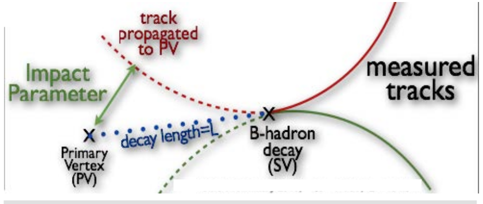
- Discovery reach improves from a chargino-neutralino mass of ~700 GeV to ~850 GeV for the Low and Reference tracker designs respectivelly

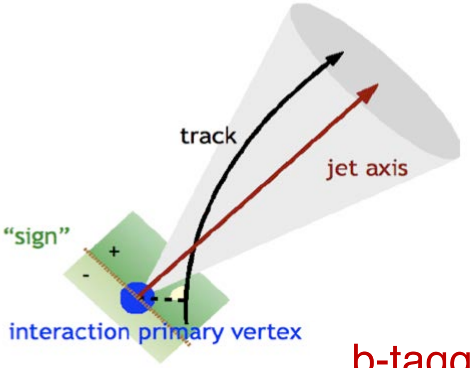
Absolute requirement for every HEP experiment seeking approval from funding agencies

Modeling of particles and event properties: b jets

Modeling of b-jet reconstruction/identification is an important simulation benchmark

- SM measurements: top decays to b, W and flavor tied to EWSB mechanism
- BSM searches: SUSY and EWSB related through hierarchy problem





b-jet identification (b-tagging) depends on impact parameter of charged-tracks in a jet, reconstructed decay vertices in the jet, lepton presence

3-D impact parameter (3-D IP) is the point of closest approach between the track and the primary vertex

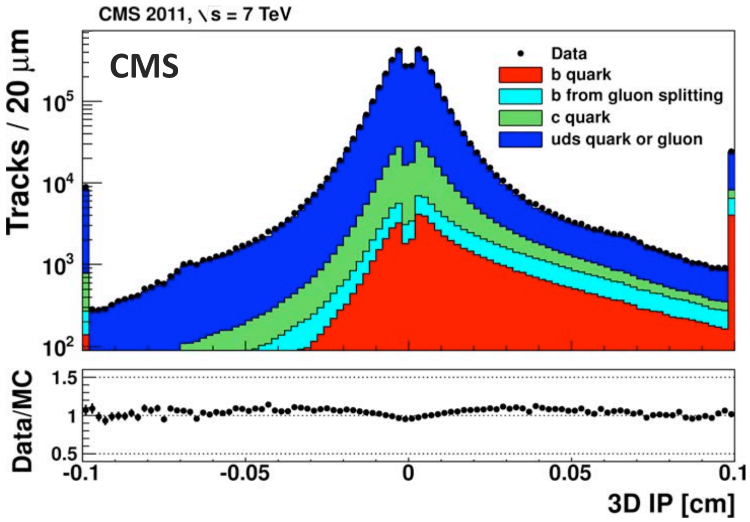
- b-quarks have positive 3-D IP while light jets have close to zero 3-D IP
- In a real detector, resolution effects give positive and negative values

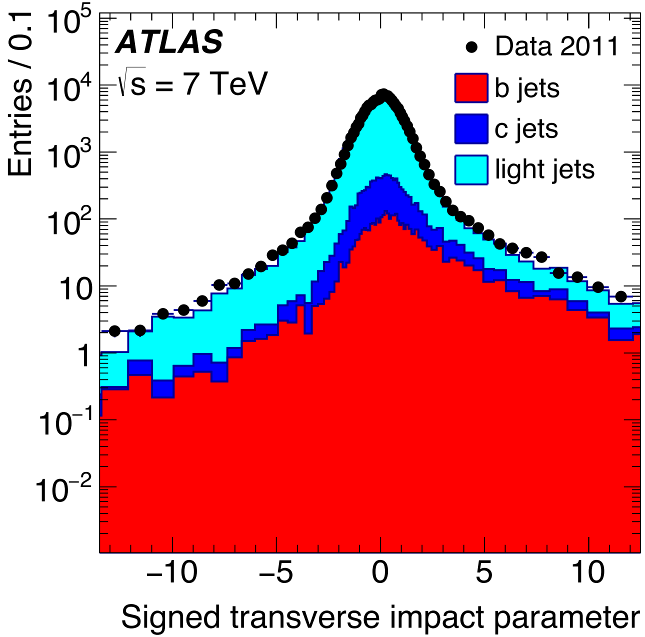
3-D IP distribution asymmetric: + mean for b jets and ~0 for light jets

Good modeling of IP variables necessary for accurate measurement of b-tagging efficiencies and fake rates using data-driven methods

b-tagging simulation depends on modeling of material budget, energy loss, ionization, multiple scattering, noise, pileup mainly in tracker

Modeling of particles and event properties: b jets





CMS 3D Impact Parameter distribution for heavy and light quarks in a di-jet sample

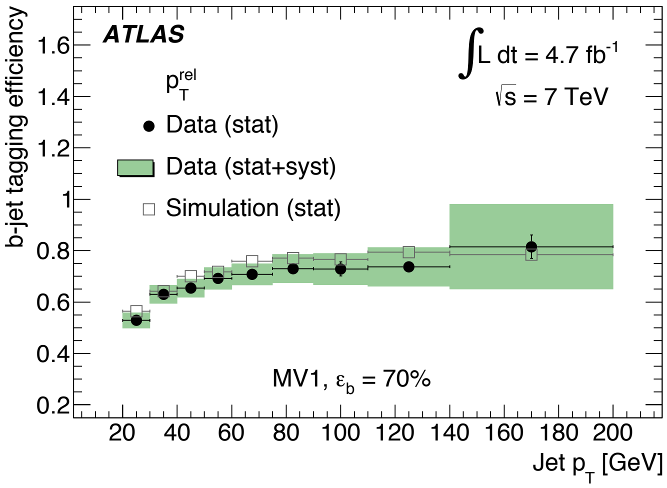
- Excellent data/MC agreement, within 10%
- b and c quark distributions (red, light blue, green) positive and asymmetric
- uds quark distributions slightly positive and symmetric

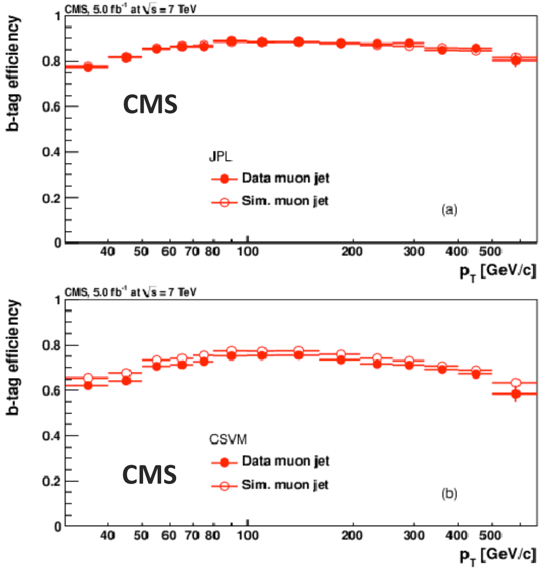
ATLAS Signed Transverse IP significance, $S_{d_0} \equiv d_0/\sigma_{d_0}$

- Same behavior as 3D IP
- Excellent agreement except at tails of distribution – resolution smearing more difficult to model

Modeling of particles and event properties: b jets

b-tagging efficiencies derived from data-driven method from samples of jets with a muon.





Data-to-MC scale factors used to adjust MC truth efficiencies used in physics measurements

Mis-tag rates (not shown) derived from "negative taggers"

- Modeling of mis-tag rates (light jet passing for a b-jet) is tricky because contributing jets come from tail of IP distributions
- p_T dependence of data-to-MC fake rate ratio:

CMS: 20% for mis-tag rate in 0.01-0.03 range
ATLAS: factor 2-3 for mis-tag rate in 0.002-0.005 range

ATLAS and CMS b-tagging efficiencies for MV1, JPL, and CSVM algorithms

- Simulation models b-tagging efficiency within <5% in both experiments
- ATLAS and CMS efficiencies cannot be compared because algorithms are tuned to different efficiency operating point depending on tolerated fake rates

Fat jet substructure

Pure data sample of semi-leptonic $t\bar{t}$ compared with MC prediction

– N-subjetiness and jet softdrop mass evaluated for fat jets