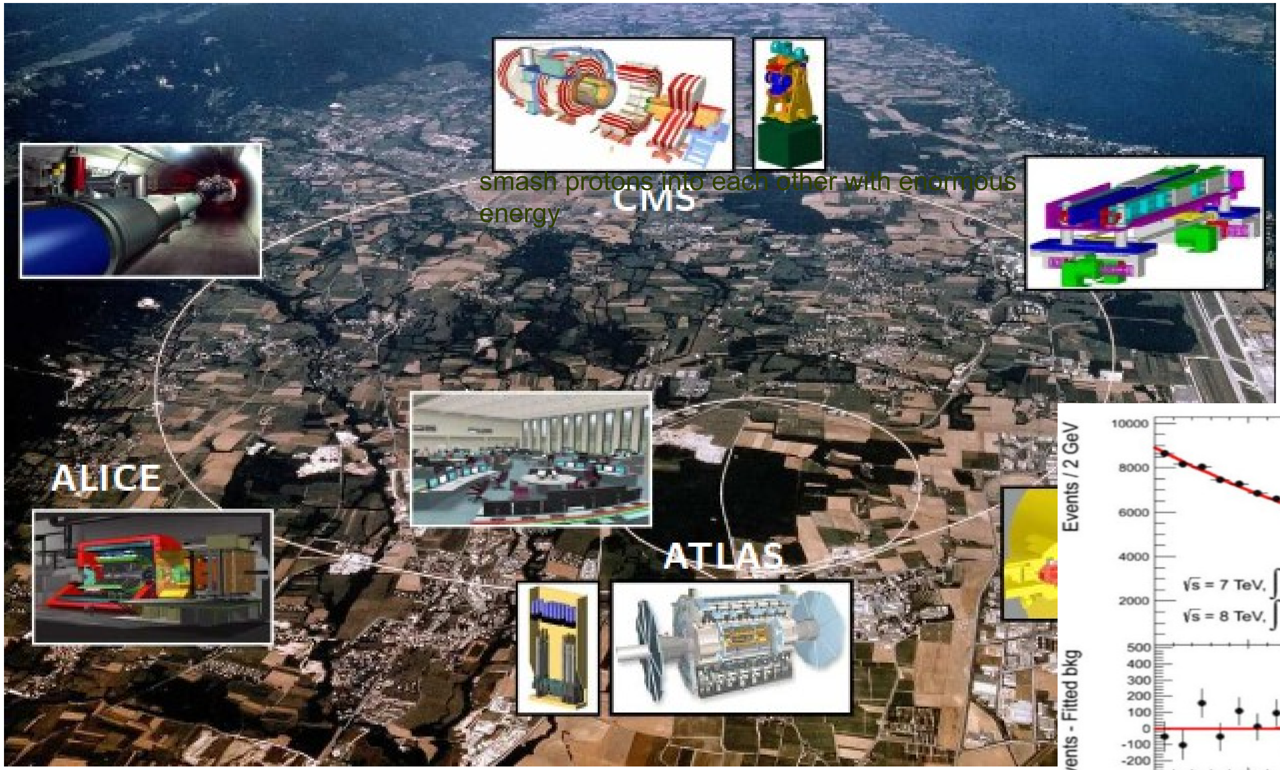
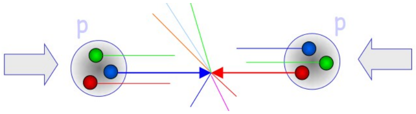
Future high-energy experiments and Monte Carlo simulations for the Energy Frontier

S. Chekanov
HEP/ANL
HEP seminar, Iowa State University
May 4, 2016

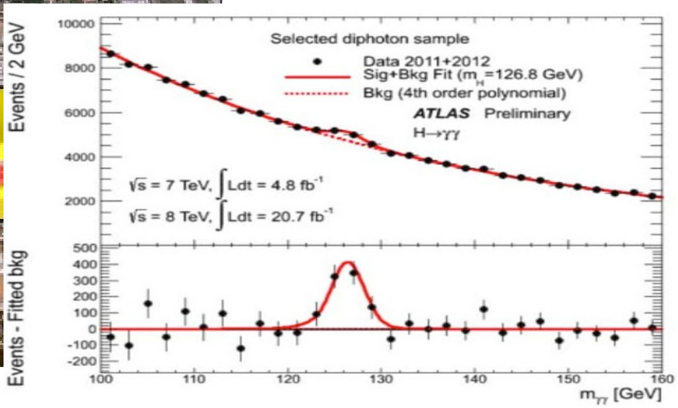
Particle Physics at the LHC

- **Study of the basic elements of matter by smashing subatomic particles at very high energy**
  - Explain the Higgs mechanism proposed in 1960s
  - Find new particles and measure their properties
  - High-precision measurements of the Standard Model





H(125 GeV) → 2 γ



Future of particle collisions

High-Luminosity Large Hadron Collider (HL-LHC)

High-Energy LHC

Large Hadron electron Collider at CERN (LHeC)

ILC (International Linear Collider)

FCC (Future Circular Collider): FCC-ee, FCC-ep, and FCC-hh

CEPC (Circular Electron Positron Collider)

SPPC (Super Proton-Proton Collider)

EIC (Electron Ion Collider)

Future of particle collisions

High-Luminosity Large Hadron Collider (HL-LHC)

High-Energy LHC

Large Hadron electron Collider at CERN (LHeC)

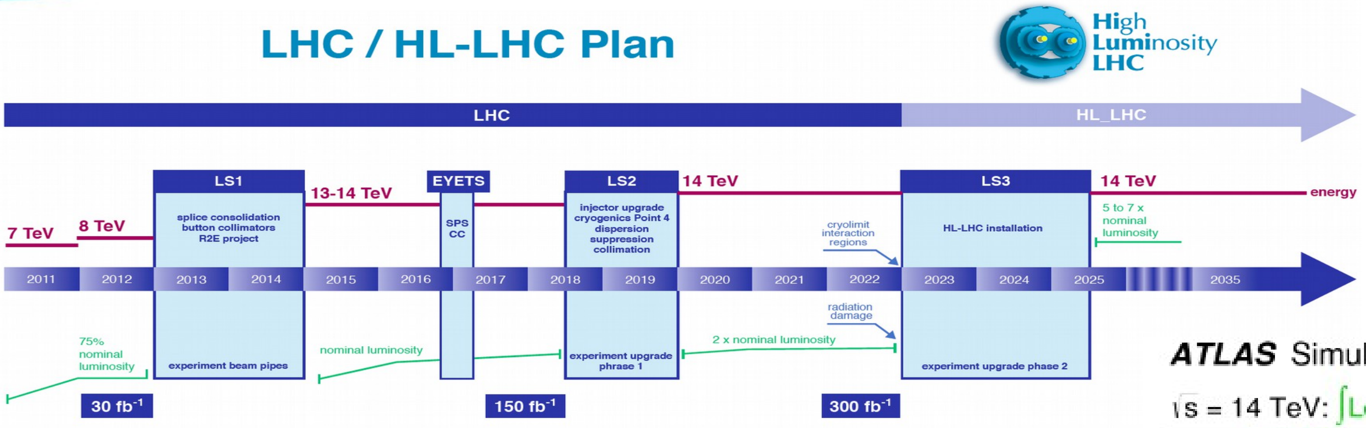
ILC (International Linear Collider)

FCC (Future Circular Collider): FCC-ee, FCC-ep, and FCC-hh

CEPC (Circular Electron Positron Collider)

SPPC (Super Proton-Proton Collider)

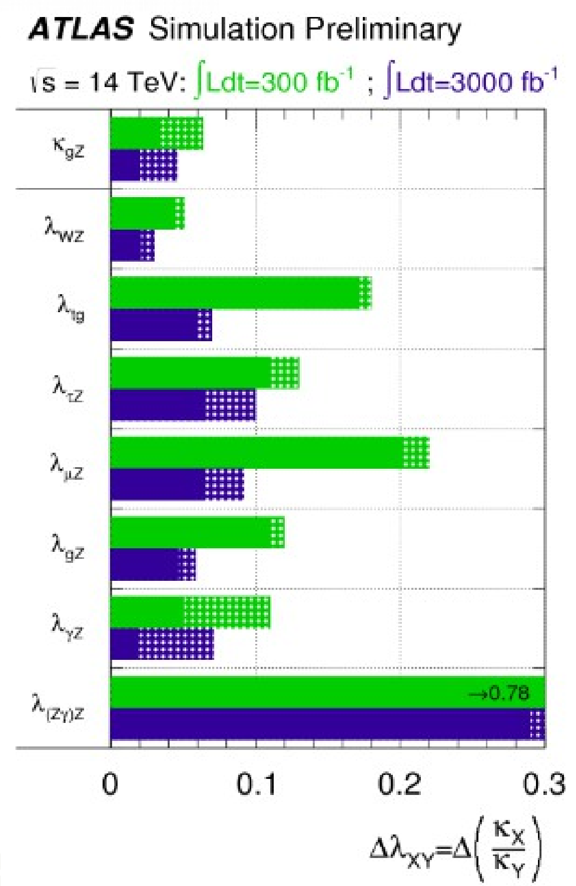
EIC (Electron Ion Collider)



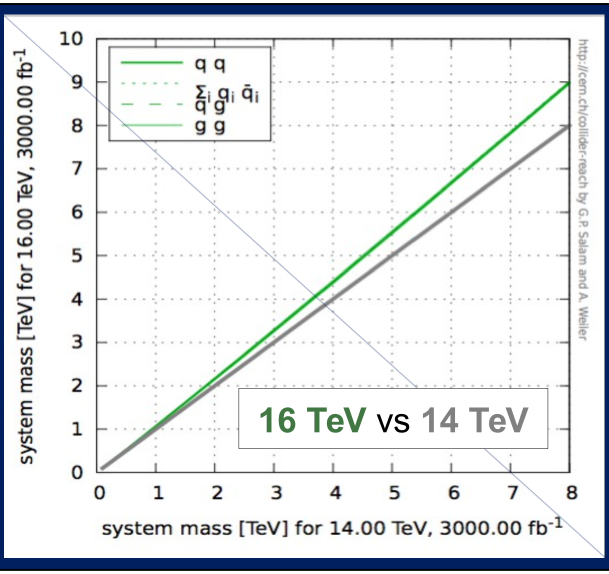
increase luminosity (rate of collisions) by a factor of 10 beyond the original design value of the LHC (from 300 to 3000 fb^{-1})

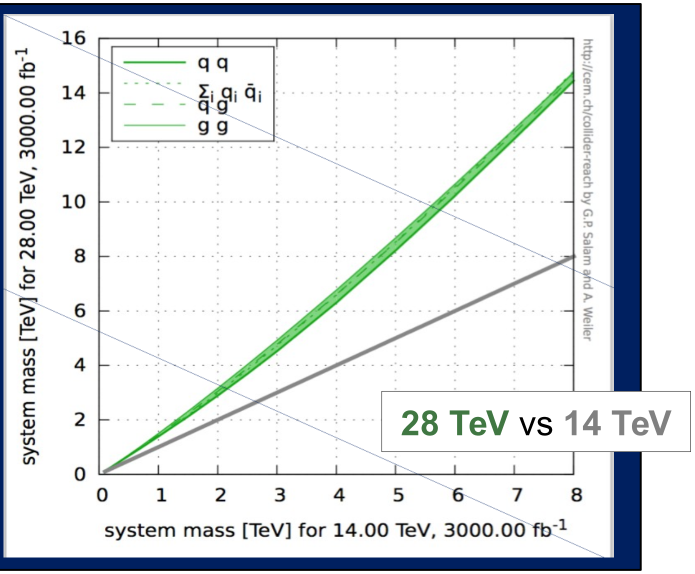
Physics goals:

- Measure existing Higgs decays with better precision
- Rare Higgs decays (μ+μ-, Z-γ, phi), double Higgs production
- Deviations from the SM & high-precision high-pT physics



High-energy LHC (HE-LHC)





F.Gianotti
+ CERN management

Milestone:
update of European Strategy for Particle Physics (~ 2019-2020)

WG set up to explore technical feasibility of pushing LHC energy to:

→ design value: 14 TeV

→ 15 TeV (dipole field of ~9.5 T) beyond (e.g. by replacing dipoles with 11 T Nb_3Sn magnets

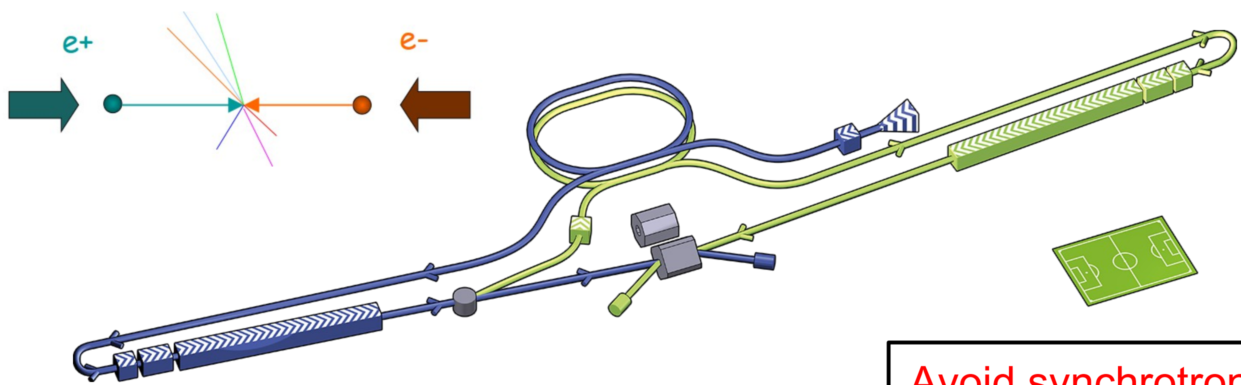
→ Identify open risks, needed tests and technical developments, trade-off between energy and machine efficiency/availability

Report on 1) end 2016, 2) end 2017, 3) end 2018 (in time for ES)

HE-LHC (part of FCC study): ~16 T magnets in LHC tunnel (√s~28 TeV)

- ❑ strong physics case if new physics from LHC/HL-LHC
- ❑ powerful demonstration of the FCC-hh magnet technology
- ❑ uses existing tunnel and infrastructure; can be built at constant budget

International Linear Collider (ILC) and Compact Linear Collider (CLIC)



Avoid synchrotron radiation ~ E^4 / m^3

Advantages over proton-proton collisions:
simple initial state (e+e-), variable energies, momentum conservation, democratic production of particles

- High-precision measurements at e+e-
- Most mature post-LHC era experiment
- ILC: CM energy 500 GeV-1000 GeV
- CLIC: CM energy up to 3000 GeV
- → Interest expressed in Japan in hosting the ILC (~50% contribution)
- → CLIC is considered by CERN (but less advanced)

International Linear Collider (ILC) and Compact Linear Collider (CLIC)

From A.Lankford
(FCC-week, Rome 2016)

A HEPAP subpanel, P5 (Particle Physics Project Prioritization Panel), was responsible for developing a strategic plan, executable over 10 years, in the context of a 20-year global vision, in *realistic* budget scenarios

- USA participation through a decision point within the next 5 years:

  P5-11: Motivated by the strong scientific importance of the ILC and the recent initiative in Japan to host it, the U.S. should engage in modest and appropriate levels of ILC accelerator and detector design in areas where the U.S. can contribute critical expertise. Consider higher levels of collaboration if ILC proceeds.

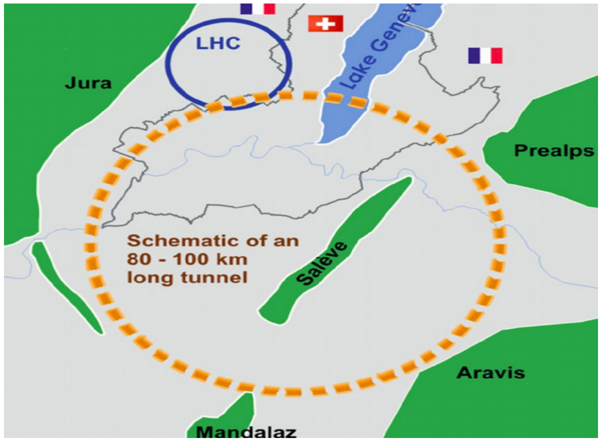
ILC in Scenario C (the 'unconstrained" budget scenario)**:**

Should the ILC go forward, Scenario C would enable the U.S. to play world-leading roles in the detector program as well as provide critical expertise and accelerator components.

Precision electroweak measurements

FCC-ee (formerly known as TLEP)

- e+e- circular collider envisioned in a new 80-100 km tunnel in the Geneva area
- centre-of-mass energy from 90 to 400 GeV
- Key features:
  - ΔM(t) < 10 MeV
  - ΔM(W) < 0.3 MeV
  - ΔM alpha_QED<10^-5
  - ΔM alpha_s (Z) < 0.0001
- Conceptual Design Report (CDR) by 2018



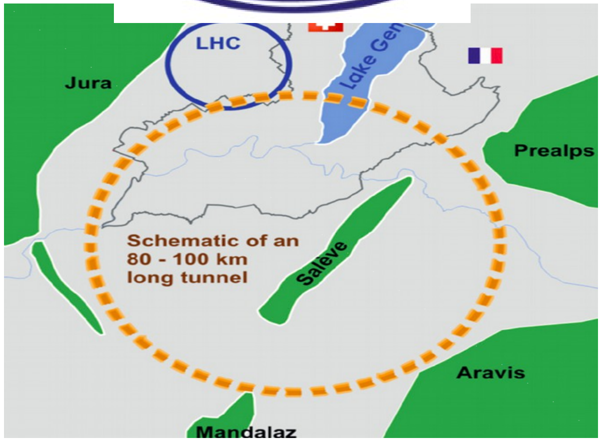
Circular Electron Positron Collider (CEPC)

- e+e- circular “Higgs factory” planned in China
- 240-350 CM energy + higher luminosity (250 fb^{-1} /year)
- Pre-CDR is ready
- Construction: 2021 – 2027. Data talking: 2028-2038

Discovery machines & energy frontier

FCC-hh (CERN) ~ 2040

- Proton-Proton collisions at 100 TeV in the Geneva area
- part of the Future Circular Collider design study (FCC) at CERN
- Physics reach: ~30 TeV for production of new heavy particles
- Peak luminosity <=$30 \times 10^{34} cm^{-2} s^{-1}$, 25(5) ns, pileup 1020(204)
- Many challenges for the detector!
- Conceptual Design Report (CDR) by 2018



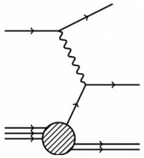
Super Proton-Proton Collider (SppC)

- Proton-proton collisions at 70 TeV in the same tunnel as CEPC
- Physics reach: ~ 25 TeV for masses of new particles
- Construction: **2035-2042**. Data taking: **2042-2055**

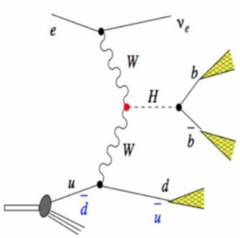
Energy frontier + intensity frontier:
LHeC, FCC-ep (CERN), Electron-ion collider (EIC)

- **LHeC:** 7 GeV proton collided with 20-60 GeV electron ($s^{1/2}$ = = 1.3 TeV)
- **HE-LHC:** 15 TeV proton collided with 60 GeV electrons ($s^{1/2}$ = = 1.9 TeV)
- **FCC-ep:** 50 GeV proton collided with > 20 GeV electron ($s^{1/2}$ = = 3.5 TeV)
- **EIC** – electron-ion collider JLab/BNL: low energy electrons with ions ($s^{1/2}$ <0.14 TeV)
  - tomography with resolution ~1/10 fb, "sweet" spot for reach QCD dynamics

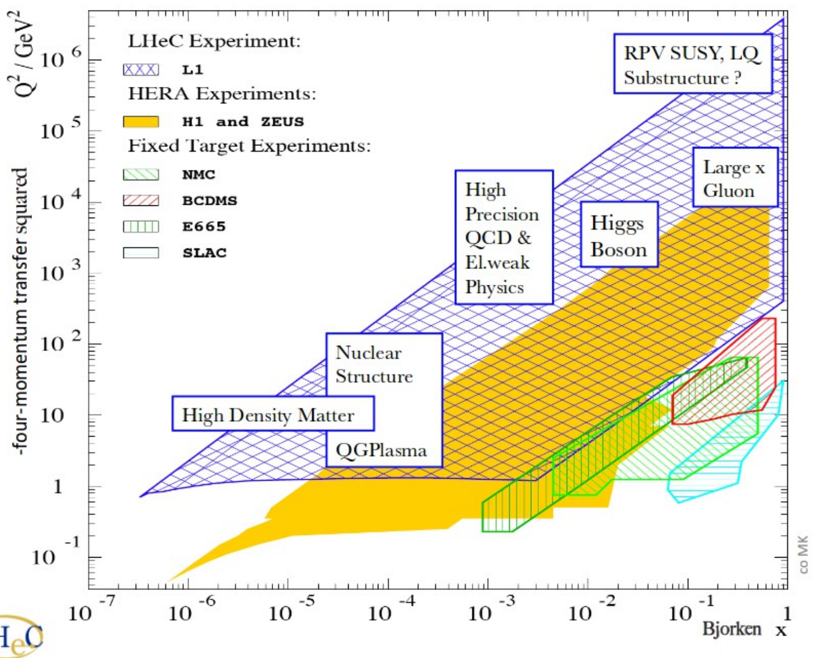
Deep inelastic scattering at the energy frontier



Turn LHC to precision Higgs factory
WW → H (~200 fb^{-1} for LHeC)

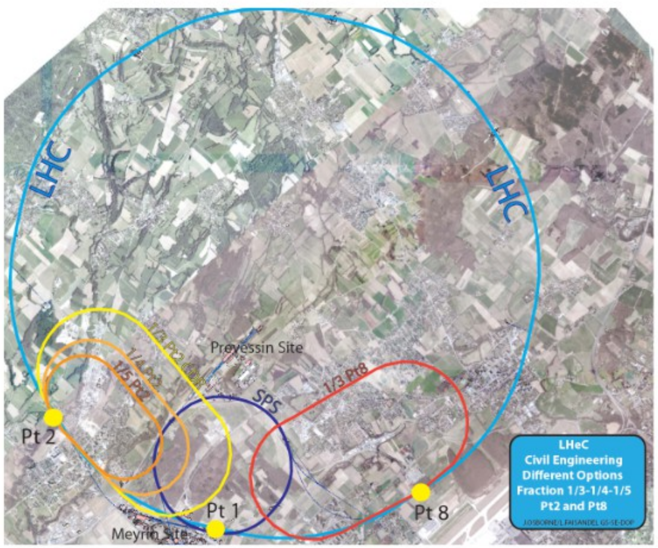


→ Studies of gluon density at large x



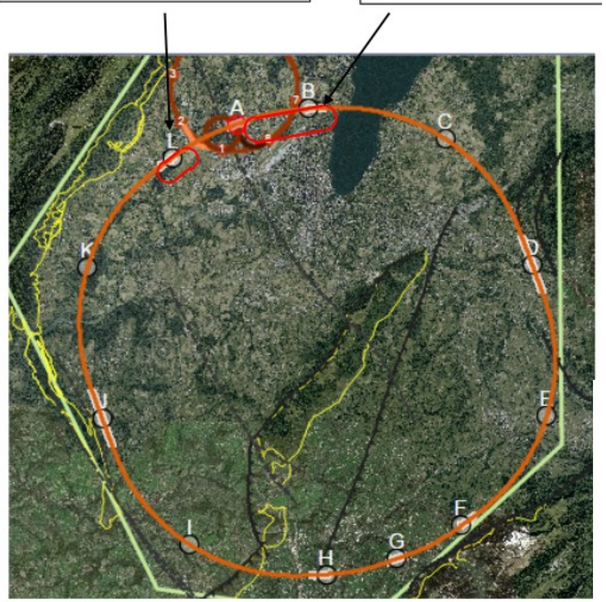
Energy frontier + energy frontier:
LHeC, FCC-ep (CERN), Electron-ion collider (EIC)

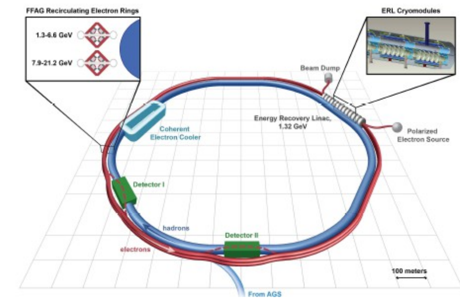
LHeC Machine



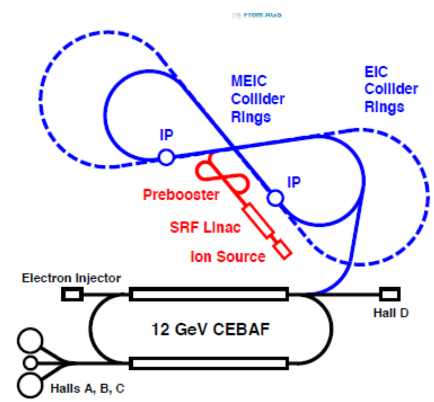
Independent FCC-he
Point L, F, H or B

LHeC / FCC-he
LHC P8 & FCC PB



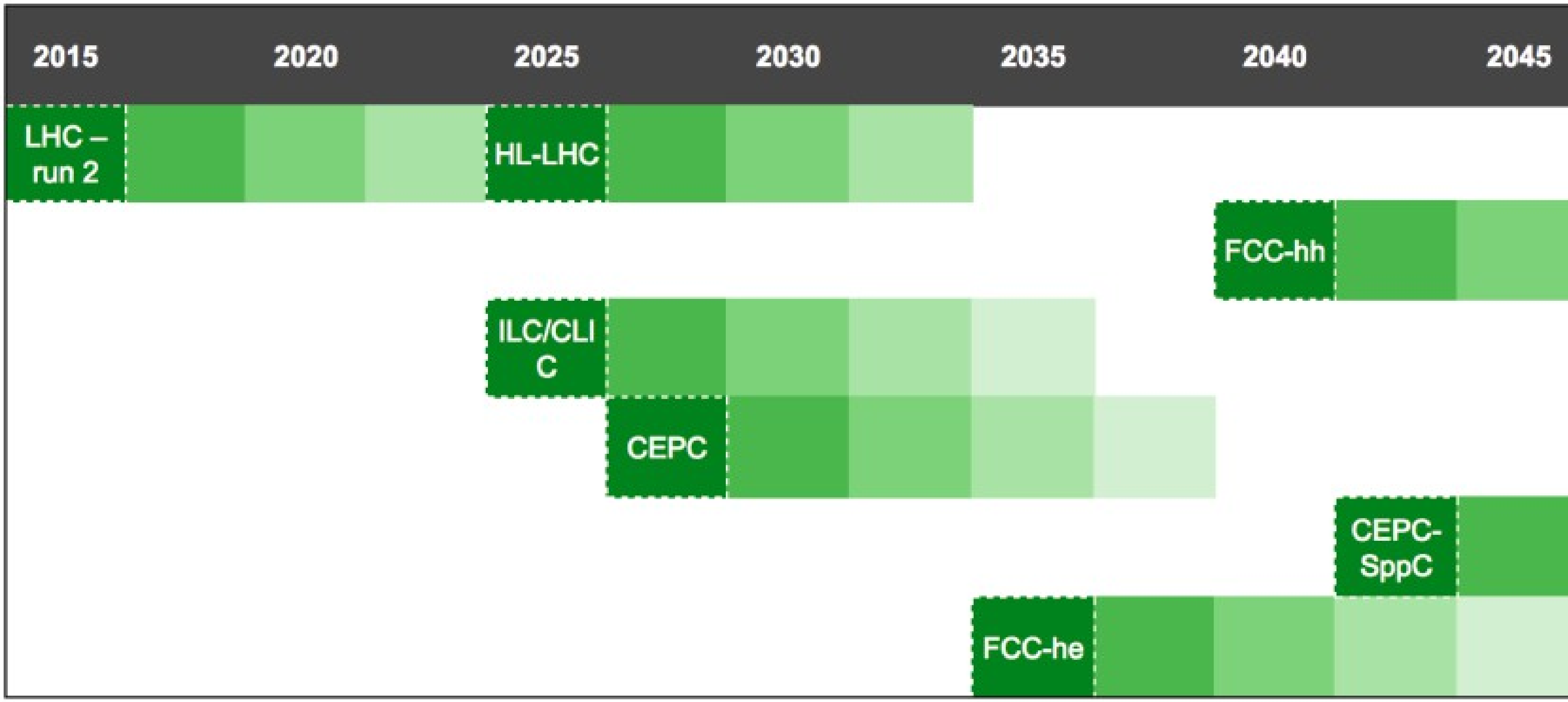


BNL?



JLab?

Timeline



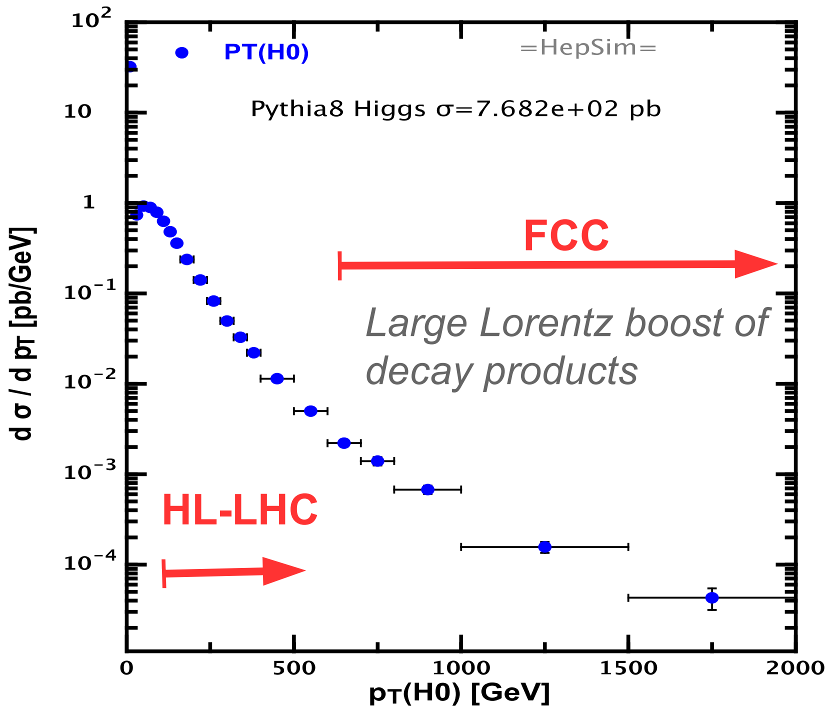
In the next decade we will deal with explorations of physics reach, detector parameters and new technology options for post-LHC era

Requires detailed simulation of physics processes and detector responses

Why do we need simulations? Higgs example

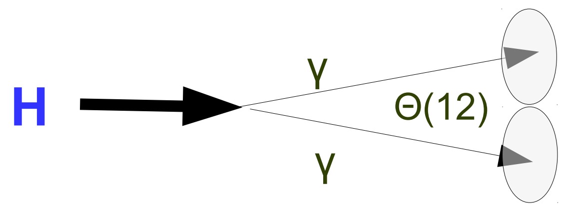
- Completely new kinematic regime → very challenging for detector designs
- 100 TeV collider will hunt for M~30 TeV particles decay to Higgs bosons
- The detector must be optimized to reconstruct Higgs at **pT~0.5-10 TeV**

100 TeV pp. Standard model



Standard model:

~100,000 Higgs / ab^{-1} for pT>1 TeV at LO



Just kinematics:

pT(H)>2 TeV ~ 5 deg separation

pT(H)>10 TeV~ 1 deg separation

Instrumental challenges:

- identify 2 photons separated by 1 degree
- reject $\pi^0 \rightarrow \gamma\gamma$ background at the same time!
- similar problems for electron, b-jets decays

Simulations for particle experiments

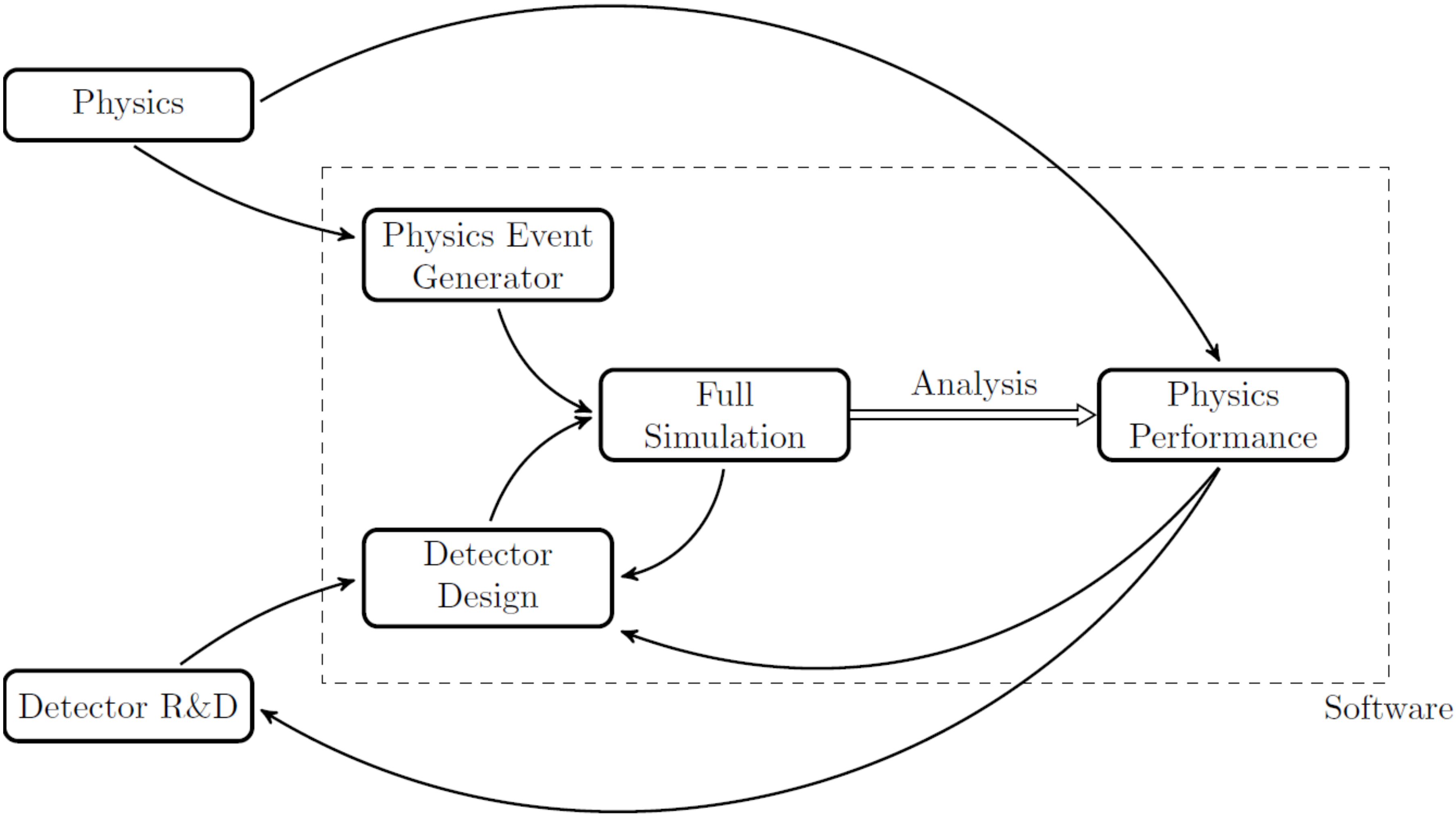


image from W.Armstrong (Physics/ANL)

Simulations for the Energy Frontier

Process modeling

- Known particle properties
- Standard Model (SM) is well established (QCD & QED)
- Event generators at LO, NLO, NNLO, NLO matched to NLO, etc.
- Models beyond the SM with detailed implementation in event generators

Detector response

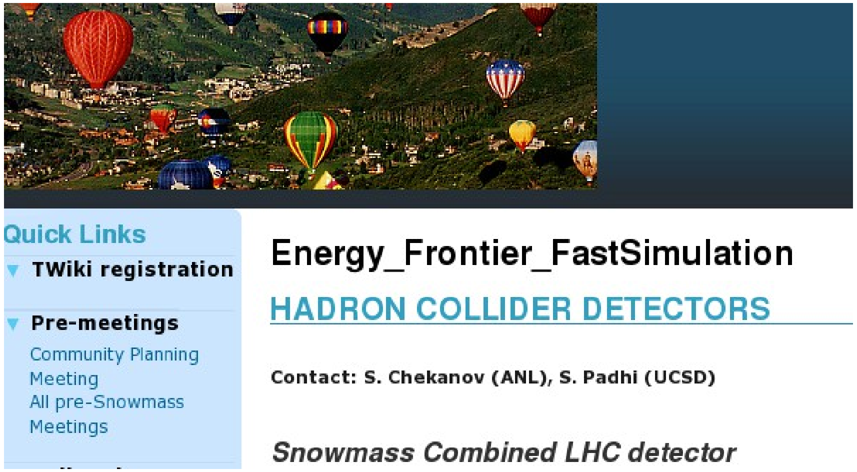
- Interactions of particles with materials
- Many parametrized cross sections (when exact theory is unknown)
- Simulation packages (Geant4,etc.)

Computing

- Fast progress in computer technology
- Open Science HPC and Grid (OSG)

Monte Carlo simulation for DPF (Snowmass 2013)

- First Snowmass meeting with large-scale open-access MC production
  - ~billion events with Delphes fast simulation
  - 140 pileup scenarios for HL-LHC
- Open-science grid (OSG)



Described in the report "Snowmass Energy Frontier Simulations" ***(arXiv:1309.1057)***

Learned Lessons:

- Need to simplify access to data → *use HTTP?*
- Insufficient file storage & large EVGEN event files when using pileup
  - EVGEN files & LOG files removed, ROOT files slimmed

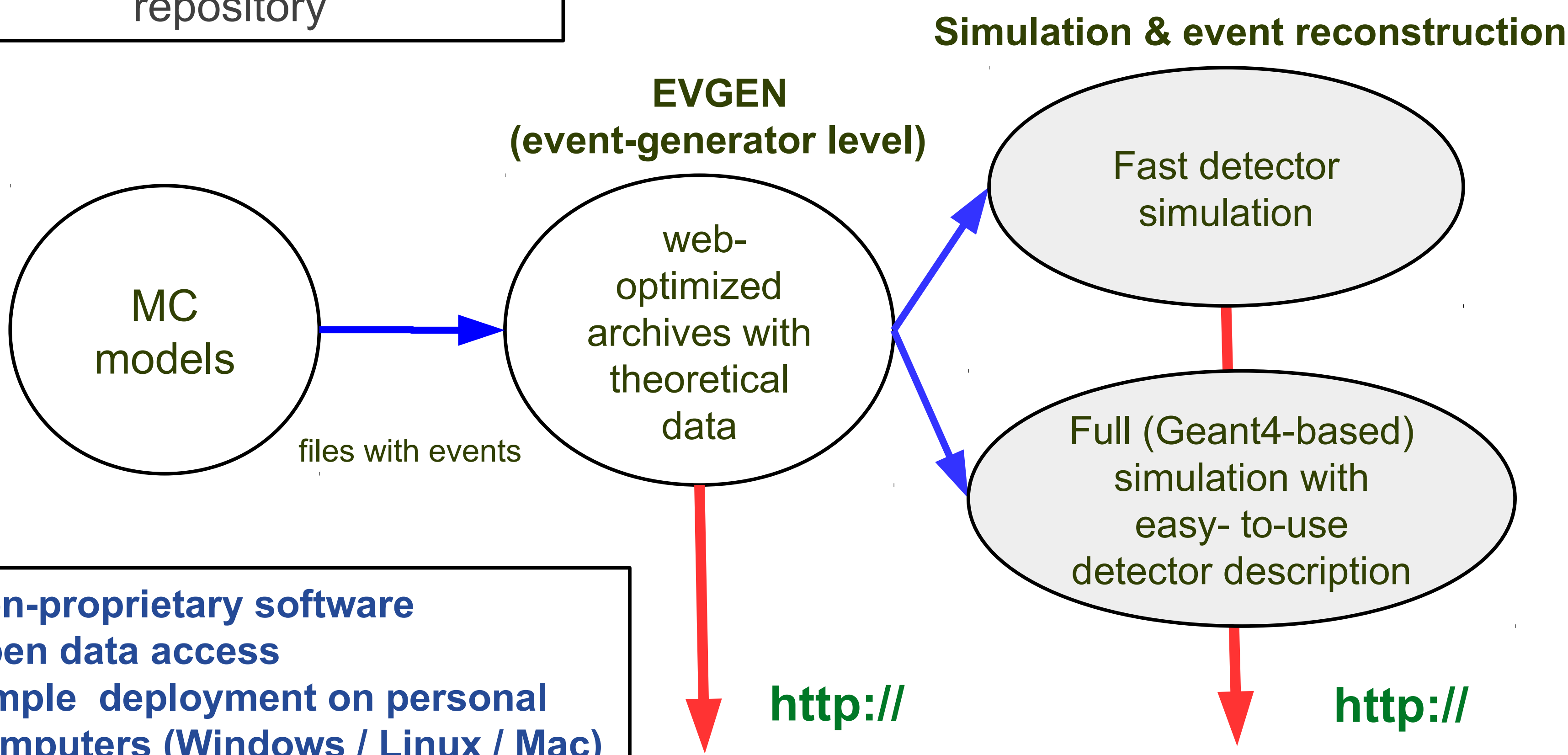
    → *Insufficient information for archiving*
- No sustainable data servers for long-term preservation

  → *Most files cannot be accessed any longer*

Each experiment has its own resources & proprietary tools.
How to share resources using project-specific infrastructure?

Moving forward: Public Repository with Simulations

Learning from Snowmass, building a public Monte Carlo repository



- → **Non-proprietary software**
- → **Open data access**
- → **Simple deployment on personal computers (Windows / Linux / Mac)**

OPEN ACCESS

Long-term availability & preservation

New data format for EVGEN: ProMC

S.C., E.May, K. Strand, P. Van Gemmeren, Comp. Physics Comm. 185 (2014), 2629

- "Archive" self-described format to keep MC events:
  - Event records, NLO,original logfiles, PDG tables etc.
- 30% smaller files than existing formats after compression

Number of used bytes depends on values.
Small values use small number of bytes

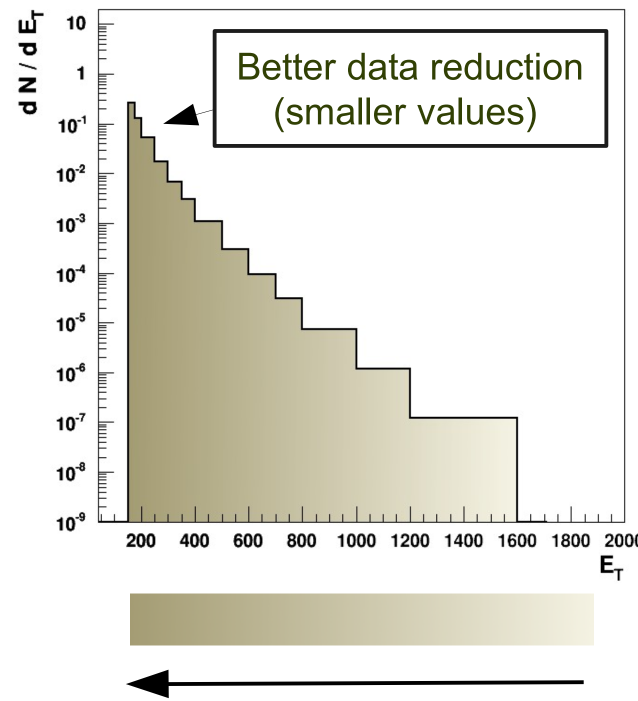
Google's Protocol buffers

protobuf
Protocol Buffers - Google's data interchange format

- Effective file size reduction for pile-up events
  - Particles with small momenta → less bytes used
- Installed on Mira (BlueGene/Q)
- Separate events can be streamed over the Internet:
  - similar to avi frames for web video players

http://atlaswww.hep.anl.gov/asc/promc/

8-bytes → varint



compression strength keeping precision of representation constant

HepSim project http://atlaswww.hep.anl.gov/hepsim/

HEP/ANL

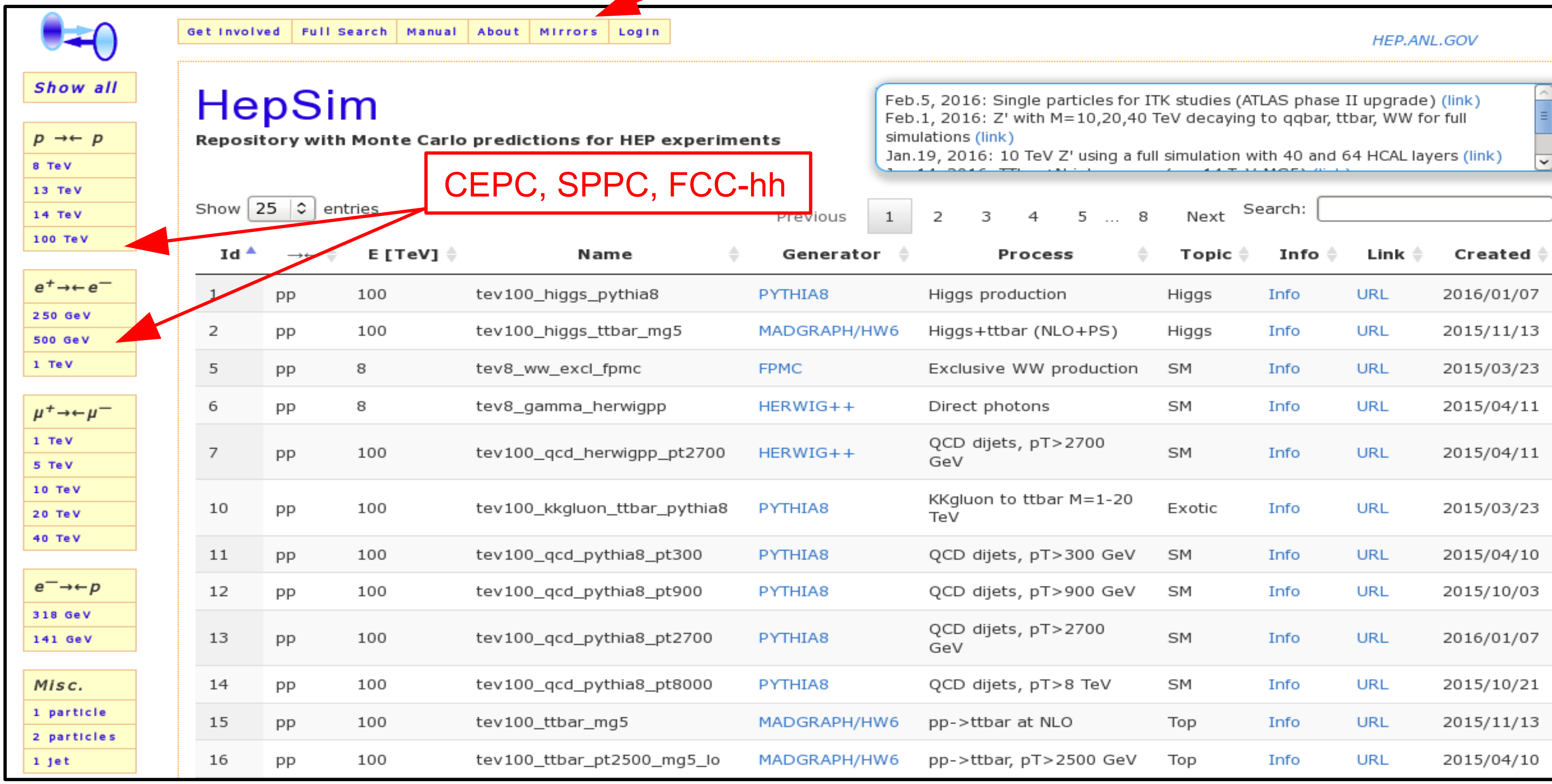
Argonne NATIONAL LABORATORY

- **2013-14: A community project to keep EVGEN files**
- **2015-now: Stores fast and full simulations using “tags”**
- **Used for future circular collider studies (ANL/Fermilab/CERN):**
  - LHC physics
  - Phase-II LHC upgrade
  - HL-LHC (pp 14 TeV 3000 fb-1)
  - FCC-hh studies (100 TeV pp, 3 ab-1)
  - HGCAL for CMS
  - Circular Electron Positron Collider studies
  - EIC
- **Theorists can add their simulations:**
  - .. and analyze events the way experimentalists do!
- **Can be used for outreach too**

HepSim simulation

http://atlaswww.hep.anl.gov/hepsim/

NERSC, CERN mirrors



Get Involved | Full Search | Manual | About | Mirrors | Log In

HEP.ANL.GOV

Show all

$p \rightarrow\leftarrow p$: 8 TeV, 13 TeV, 14 TeV, 100 TeV

$e^+ \rightarrow\leftarrow e^-$: 250 GeV, 500 GeV, 1 TeV

$\mu^+ \rightarrow\leftarrow \mu^-$: 1 TeV, 5 TeV, 10 TeV, 20 TeV, 40 TeV

$e^- \rightarrow\leftarrow p$: 318 GeV, 141 GeV

Misc.: 1 particle, 2 particles, 1 jet

HepSim

Repository with Monte Carlo predictions for HEP experiments

Feb.5, 2016: Single particles for ITK studies (ATLAS phase II upgrade) (link)
Feb.1, 2016: Z' with M=10,20,40 TeV decaying to qqbar, ttbar, WW for full simulations (link)
Jan.19, 2016: 10 TeV Z' using a full simulation with 40 and 64 HCAL layers (link)

CEPC, SPPC, FCC-hh

Show 25 entries

Previous 1 2 3 4 5 … 8 Next

Search:

Id	→←	E [TeV]	Name	Generator	Process	Topic	Info	Link	Created
1	pp	100	tev100_higgs_pythia8	PYTHIA8	Higgs production	Higgs	Info	URL	2016/01/07
2	pp	100	tev100_higgs_ttbar_mg5	MADGRAPH/HW6	Higgs+ttbar (NLO+PS)	Higgs	Info	URL	2015/11/13
5	pp	8	tev8_ww_excl_fpmc	FPMC	Exclusive WW production	SM	Info	URL	2015/03/23
6	pp	8	tev8_gamma_herwigpp	HERWIG++	Direct photons	SM	Info	URL	2015/04/11
7	pp	100	tev100_qcd_herwigpp_pt2700	HERWIG++	QCD dijets, pT>2700 GeV	SM	Info	URL	2015/04/11
10	pp	100	tev100_kkgluon_ttbar_pythia8	PYTHIA8	KKgluon to ttbar M=1-20 TeV	Exotic	Info	URL	2015/03/23
11	pp	100	tev100_qcd_pythia8_pt300	PYTHIA8	QCD dijets, pT>300 GeV	SM	Info	URL	2015/04/10
12	pp	100	tev100_qcd_pythia8_pt900	PYTHIA8	QCD dijets, pT>900 GeV	SM	Info	URL	2015/10/03
13	pp	100	tev100_qcd_pythia8_pt2700	PYTHIA8	QCD dijets, pT>2700 GeV	SM	Info	URL	2016/01/07
14	pp	100	tev100_qcd_pythia8_pt8000	PYTHIA8	QCD dijets, pT>8 TeV	SM	Info	URL	2015/10/21
15	pp	100	tev100_ttbar_mg5	MADGRAPH/HW6	pp->ttbar at NLO	Top	Info	URL	2015/11/13
16	pp	100	tev100_ttbar_pt2500_mg5_lo	MADGRAPH/HW6	pp->ttbar, pT>2500 GeV	Top	Info	URL	2015/04/10

HepSim stores EVGEN files (LO,NLO, etc), fast simulations, full Geant4 simulations

Dataset entry: e+e- collisions (CM energy = 250 GeV). Z → e+e-

- $p \rightarrow\leftarrow p$
  - 8 TeV
  - 13 TeV
  - 14 TeV
  - 100 TeV
- $e^+ \rightarrow\leftarrow e^-$
  - 250 GeV
  - 500 GeV
  - 1 TeV
- $\mu^+ \rightarrow\leftarrow \mu^-$
  - 1 TeV
  - 5 TeV
  - 10 TeV
  - 20 TeV
  - 40 TeV
- $e^- \rightarrow\leftarrow p$
  - 318 GeV
  - 141 GeV
- Misc.
  - 1 particle
  - 2 particles
  - 1 jet

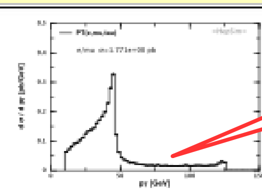
Repository with Monte Carlo predictions for HEP experiments

Feb.1, 2016: Z' with M=10,20,40 TeV decaying t
Jan.19, 2016: 10 TeV Z' using a full simulation w
Jan.14, 2016: TTbar+N jet process (pp, 14 TeV, M

Information about "*gev250ee_pythia6_zpole_ee*" dataset

Name:	*gev250ee_pythia6_zpole_ee*
Collisions:	e+e-
CM Energy:	0.25 TeV
Entry ID:	146
Topic:	SM
Generator:	PYTHIA6
Calculation level:	LO+PS+hadronisation
Process:	Z boson to e+e-
Total events:	2000000
Number of files:	100
Cross section (σ):	1.7765 ± 0.0126 pb
Luminosity (L):	1.126E+06 pb^{-1} *(or)* 1125.7948 fb^{-1} *(or)* 1.1258 ab^{-1}
Format:	ProMC
Submission date:	Tue Oct 13 14:28:55 CDT 2015
Download URL:	http://mc.hep.anl.gov/asc/hepsim/events/ee/250gev/pythia6_zpole_ee
Mirrors:	
MC truth size:	0.826 GB
Fast simulation:	rfast001 (info) |
Full simulation:	rfull002 (info) | rfull001 (info) |
Record slimmed:	No
Events weighted:	No
User description:	PYTHIA version 6.4. Z production (Zpole) with decays to e+e-. Other details in the
File metadata:	**ProMC version:** 4; **Nr events:** 1000; **Varint E:** 1000000; **Varint L:** 10000; **Lo** logfile.txt; **Last modified:** 2015-10-15 20:31:08; **Settings:** PYTHIA-6.4.28; MSE mix events; NTOT 0 0 1000 ! Number of events; ECME 0 0 250.0 ! CM energy (Ge 0 0 839264 ! random seed; MSEL 0 0 0 ! all mixed events; PMAS 6 1 172.5 !; PMA 91.1876 ! Z boson mass; PMAS 24 1 80.3850 ! W boson mass; PMAS 25 1 125. ! Hi mass; MSUB 1 0 1 ! ffbar to Z; MSTP 43 0 2 ! Z only, no gamma; MDME 174 1 0 ! MDME 175 1 0 ! U U~; MDME 176 1 0 ! S S~; MDME 177 1 0 ! C C~; MDME 178 1 B~; MDME 179 1 0 ! T T~; MDME 182 1 1 ! E- E+; MDME 183 1 0 ! NU_E NU_E~; 1 0 ! MU+ MU-; MDME 185 1 0 ! NU_MU+ NU_MU-; MDME 186 1 0 ! TAU- TAU+; 1 0 ! NU_TAU- NU_TAU+; PARJ 71 0 10 ! ctau=10mm; MSTJ 22 0 2 !;

Validation:

Nr	Analysis code	Output plot (SVG)	Output (XML)
1	pythia6_zpole_ee.py Launch Desktop: **hs-ide [URL]**		pythia6_zpole_ee.jdat

URL for EVGEN files (download or streaming)

URL with fast (DELPHES) or Geant4 (SLIC) simulations

Validation distributions created using Python scripts on the Java platform

Run via JavaWeb start by streaming data over the Web

Available Monte Carlo generators

- MG5/PY6 (NLO+PS+hadr): TTbar, Higgs+jj, Higgs+TTbar etc
- MG5/Herwig (NLO+PS+hadr)
- PYHIA8 (many processes)
- FPMC (exclusive WW, Higgs)
- HERWIG++ pp collisions (QCD dijets)
- SuperChic 2 - A Monte Carlo for Central Exclusive Production
- MCFM (NLO):: Higgs -> $\gamma\gamma$, Inclusive gamma, TTbar
- NLOjet++ (NLO) for inclusive jets (bins in pT)
- JETPHOX (NLO) for inclusive photons (bins in pT)
- PYTHIA6 for e+e and mu+mu- collisions
- LEPTO/PYTHIA for ep DIS
- LEPTO/ARIADNE for ep DIS
- Single particle guns (+ pileup)

~20% samples generated on BlueGene/Q (Mira) supercomputer (Jetphox, MCFM)
~40% HEP-ANL (mainly Madgraph)
~40% OSG-CI grid (ANL/UChicago) and USATLAS CI (for phase II)

Long-term preservation of theoretical calculations

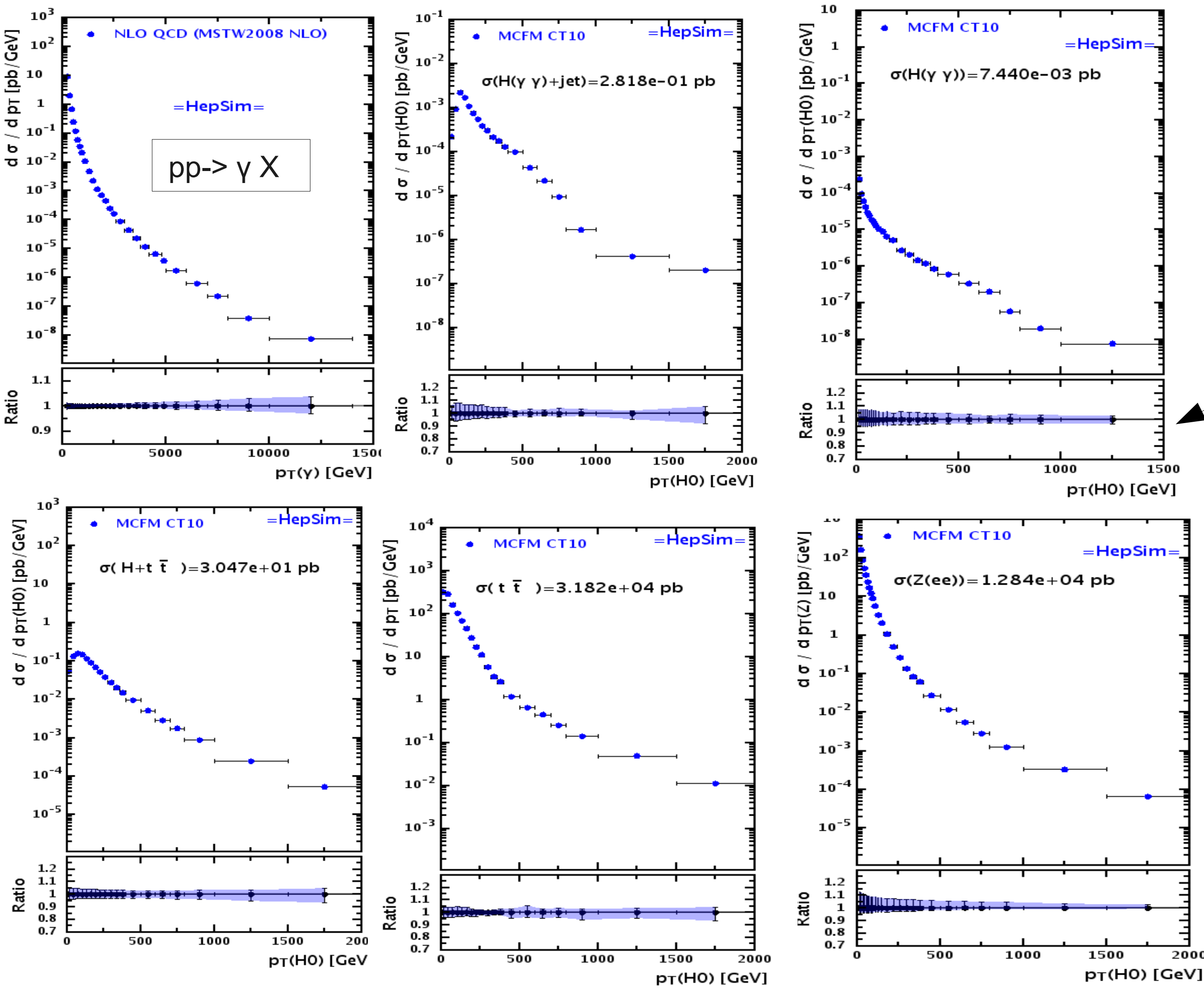
- **Storing Monte Carlo predictions in files makes sense if:**

$$\frac{\text{time to download \& analyze on commodity computer}}{\text{CPU*h needed to create the prediction}} \equiv \varepsilon << 1$$

$\varepsilon \sim 0.01\text{-}1$	- for LO MC
$\varepsilon << 0.01$	- for NLO etc.

- **$\varepsilon << 1$:**
  - Madgraph5 etc. (NLO+PS+hadronisation), ALPGEN
  - Some fast-converging NLO calculations (MCFM, jetPHOX etc)
  - MC with $\varepsilon \sim 1$ but after mixing with pile-up (CPU intensive)
- **$\varepsilon \sim 1$:** Less appropriate approach for:
  - LO simulations (Pythia)
  - Some NLO programs with slow convergence
    - requires too large data volumes to keep weighted events

Examples of differential cross sections for 100 TeV



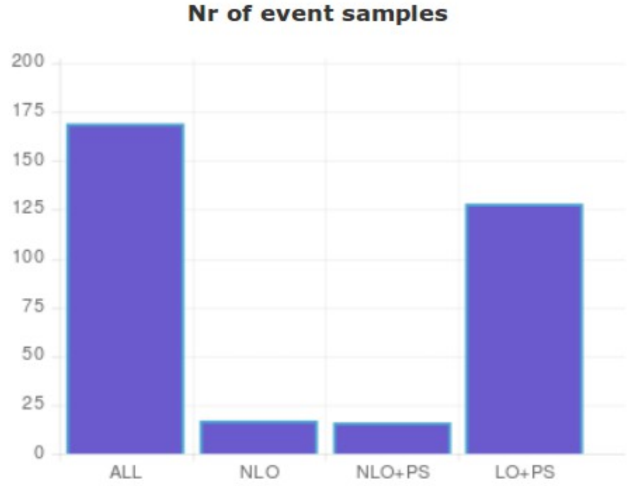
$$\frac{\sqrt{\sum_{i=1}^{N}(\sigma_i-\sigma_0)^2}}{\sigma_0}$$

Data creation (~10GB) takes ~10000 CPU*h

Analysis step takes <30 min

HepSim statistics

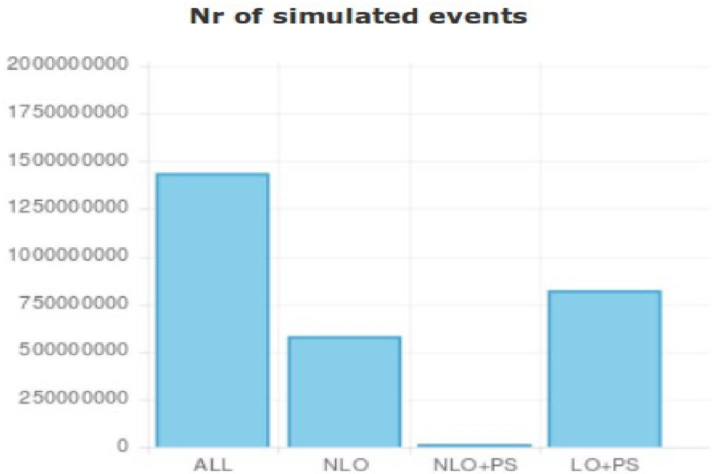
(excluding fast and Geant4 simulations)



Data hosted by:

Nr	Data servers
1	mc.hep.anl.gov
2	raw.stash2.ci-connect.net
3	faxbox.usatlas.org
4	portal.nersc.gov

Number of public file servers	4
Number of event samples	169
Number of NLO samples	17
Number of NLO+PS samples	16
Number of LO (+PS) samples	128
Number of events	1437939816
NLO events	583000000
NLO+PS events	15900595
LO (+PS) events	823536521
Total size (GB)	6486.634
NLO size (GB)	238.06
NLO+PS size (GB)	117.773
LO (+PS) size (GB)	6127.386
Number of files	306046



~ 200 Monte Carlo samples

(some are “compound”, i.e. consists of subsamples)

~1.5 billion events

HepSim repository. How it works

large-scale computing resources

osg connect

ATLAS CONNECT

NERSC

MIRA

Event Generators

- PYTHIA6
- PYTHIA8
- HERWIG++
- Madgraph5
- MCFM
- JetPhox
- FPMC
- NLOjet++
- LEPTO/Ariadne

EVGEN (ProMC)

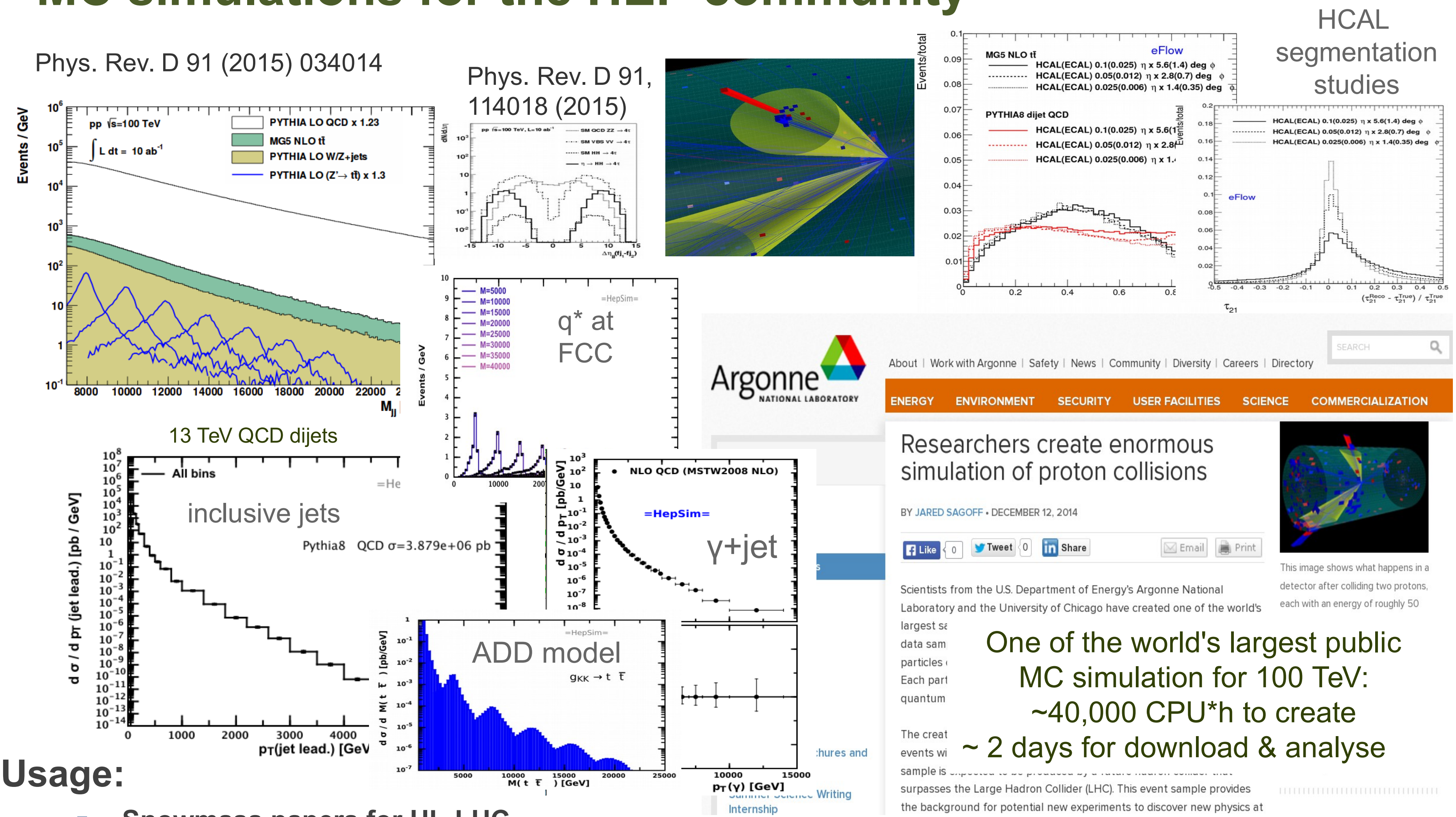
fast → Delphes fast simulation (ROOT)

full → SLIC (Geant4) full simulation and reconstruction software (LCIO)

EVGEN (ProMC) files stored on several public web servers (Apache)

HepSim
- index files
- create metadata
- prepare for batch download
- validate with Jython scripts
- create search database

MC simulations for the HEP community



Usage:

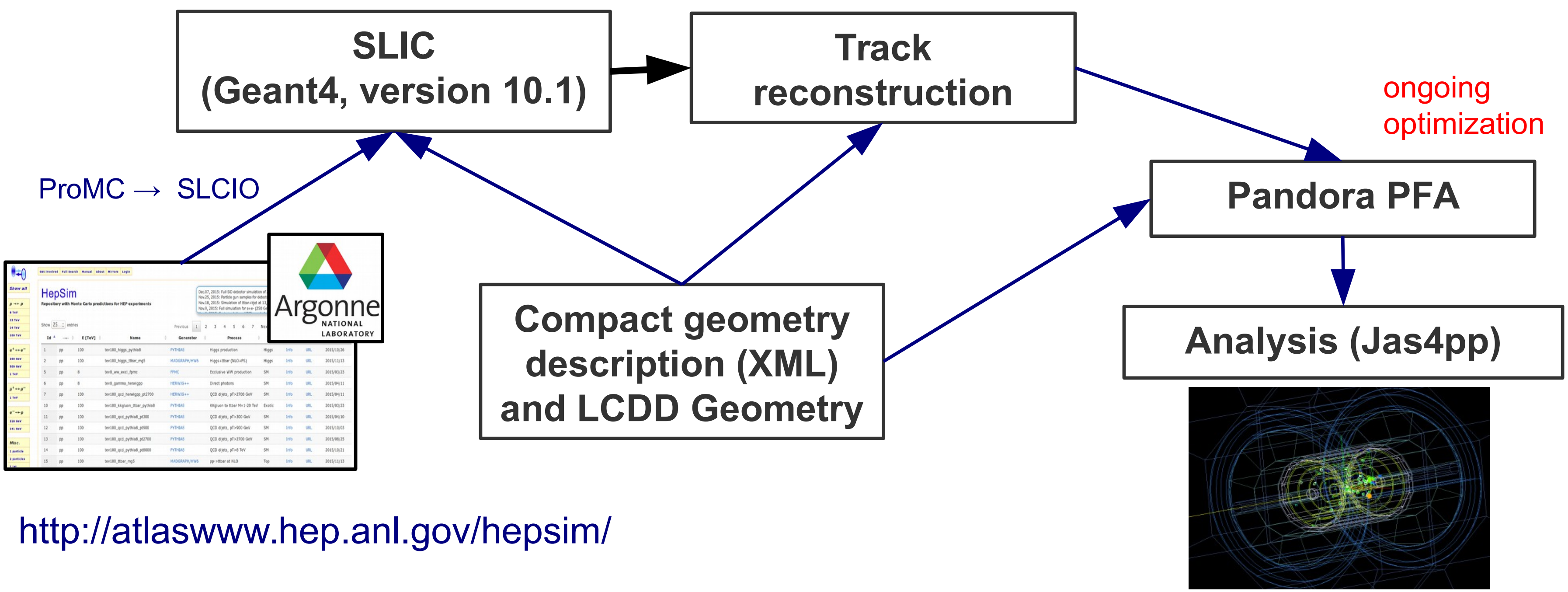
- **Snowmass papers for HL-LHC**
- **ATLAS run I & II analyses:** excl. H^0, excl. WW, direct photons with MCFM NLO, JETPHOX NLO, Long-lived particles, ADD model for gravitons, H → φγ) → validated and shipped to ATLAS
- **FCC physics studies, CPEC (recently)**
- **Detector studies.** List of public talks/papers in **http://atlaswww.hep.anl.gov/hepsim/about.php**

Software for full simulations

Simulator for the Linear Collider (SLIC) software

- Optimized for the SiD detector at SLAC (T.Johnson, N.Graf, J.McCormick, J.Strube)
- Re-purposed for future pp collider studies (S.C., A.Kotwal, J.Strube)
- Integrated with **HepSim.** Deployed on Open-Science Grid (OSG)

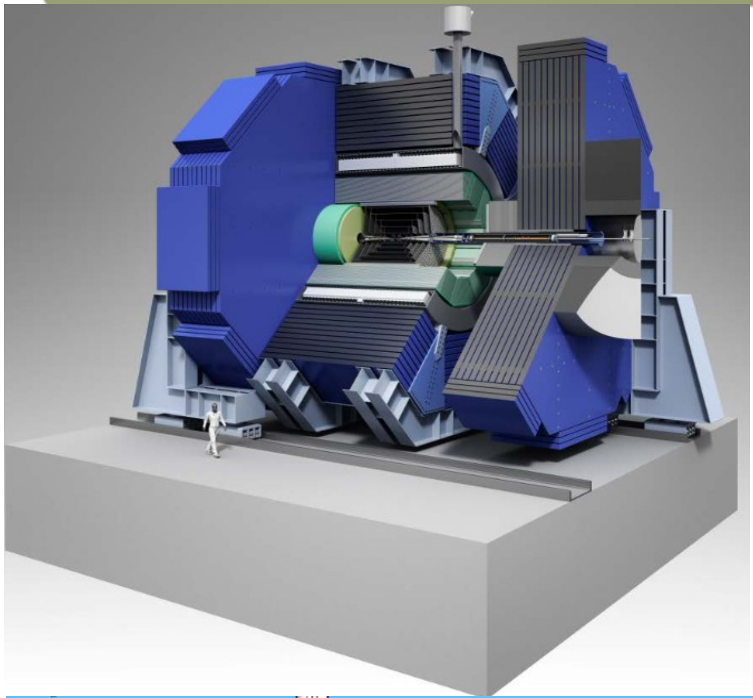
Analysis: C++/Root or **Jas4pp** (ANL,S.C,E.May). Based on Jas3 (SLAC)

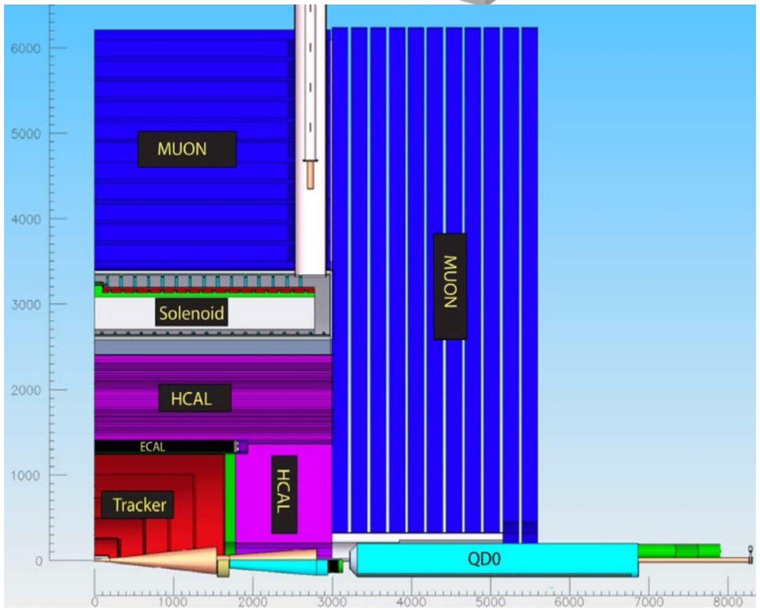


http://atlaswww.hep.anl.gov/hepsim/

SiD detector for ILC

- Multi-purpose detector for the ILC
- The key characteristics of the SiD detector:
  - 5 Tesla solenoid
  - Silicon tracker: 50 um readout pitch
  - ECAL: (0.35 cm cell size, W / silicon)
  - HCAL:
    - 1x1 cm cell size (RPC)
    - 40 layers for barrel (HCAL) ~4.5 λ_I
- Optimized for particle-flow algorithms (PFA)
- Fully configurable using SLIC software





Re-purposing SiD for circular collides

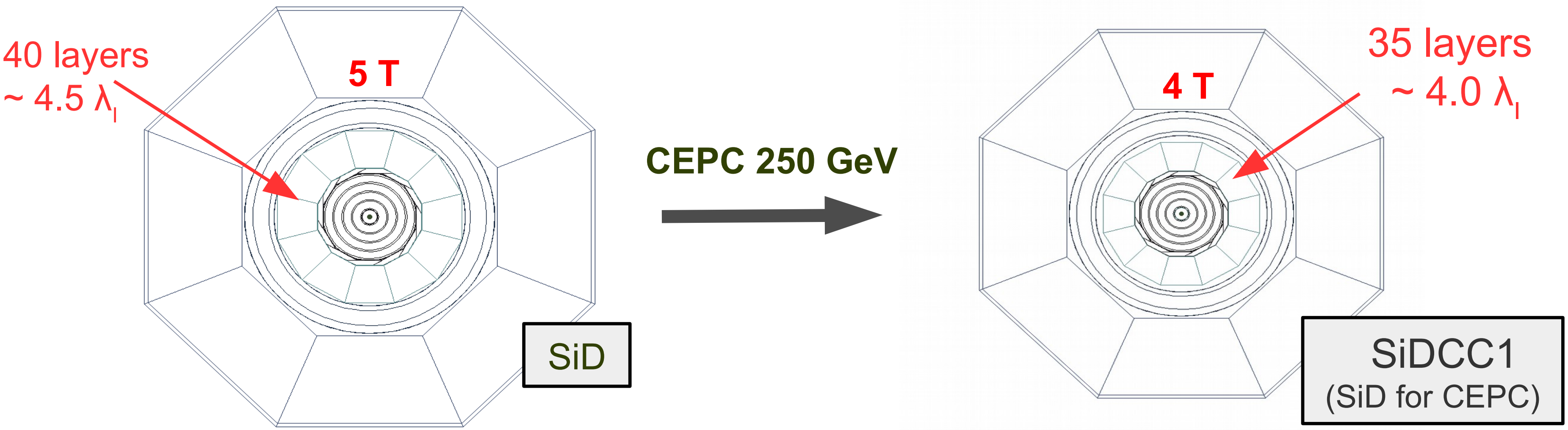
- Re-purpose SiD design and SLIC software for circular colliders:
  - CEPC, EIC, FCC-hh
- Leverage large investments to R&D of the SiD detector, including SLIC software used in the past by the ILC community (SiD+ILD)
- Keep in mind that SiD:
  - is over-designed for CEPC (250 GeV) and expensive ($320M M&S)
  - is too expensive for EIC + requires optimizations
  - requires significant increase in size for FCC-hh (> x2) + more optimizations

Designing a detector for CEPC (e+e- 250 GeV)

SiD detector is designed for ~500 GeV particles/jets (0.5-1 TeV CM energy)
But CEPC will measure particles/jets up to 125 GeV (250 CM energy)

Possible optimizations:

- HCAL:	barrel: 4.5 λ_I (40 layers)	→ 4.0 λ_I (35 layers)
	endcap: 5 λ_I (45 layers)	→ 4.0 λ_I (35 layers)
- Tracking:	5 Tesla	→ 4 Tesla



Design a light, cost-optimized version of the SiD detector for CEPC and use several physics processes to benchmark its performance

S.C. and M.Demarteau, Conceptual Design Studies for a CEPC Detector. arXiv:1604.01994

HepSim samples after full SLIC simulations

- **SiDCC** http://atlaswww.hep.anl.gov/hepsim/list.php?find=gev250%rfull002
- **SiD:** http://atlaswww.hep.anl.gov/hepsim/list.php?find=gev250%rfull001
- Event samples for SiDCC1 (rfull002) and the standard SiD (rfull001):
- Generate Pythia6 processes and process with SLIC:

  - **$Z \rightarrow e^+e^-$**
  - **$Z \rightarrow$ tau tau**
  - **$Z \rightarrow$ mu+mu-**
  - **$Z \rightarrow b\bar{b}$**
  - **$H(125) \rightarrow b\bar{b}$**
  - **$H(125) \rightarrow \gamma\gamma$**
  - **$H(125) \rightarrow ZZ^* \rightarrow 4l$**
  - **$H(125) \rightarrow$ tau tau**

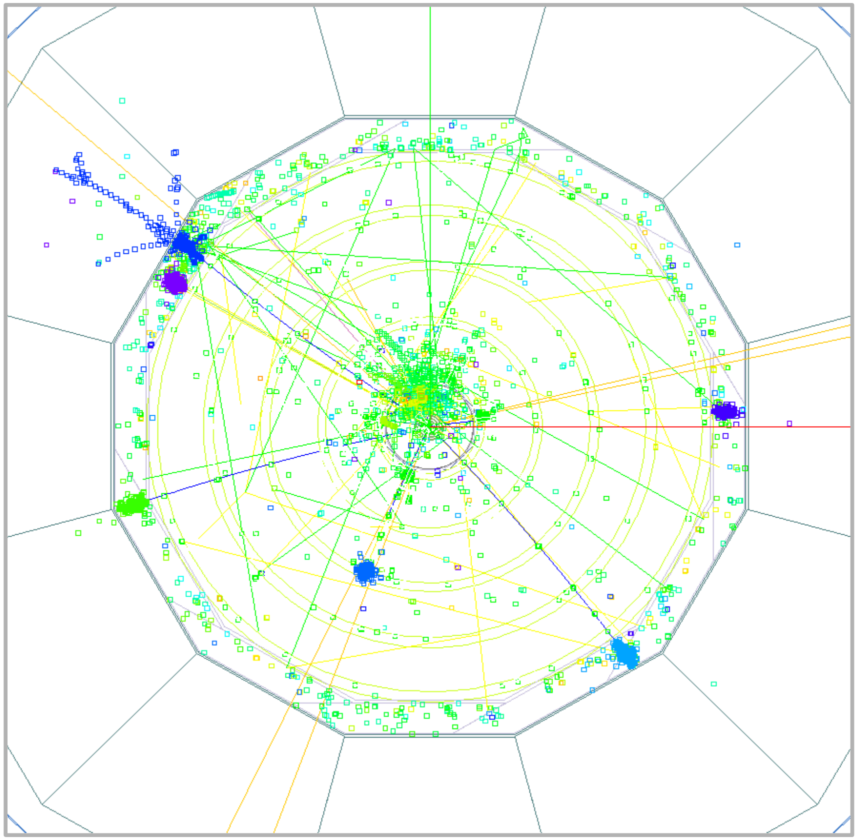
~ 10,000 reconstructed events for each physics channel after PFA (Pandora) reconstruction

URL with manual/examples:

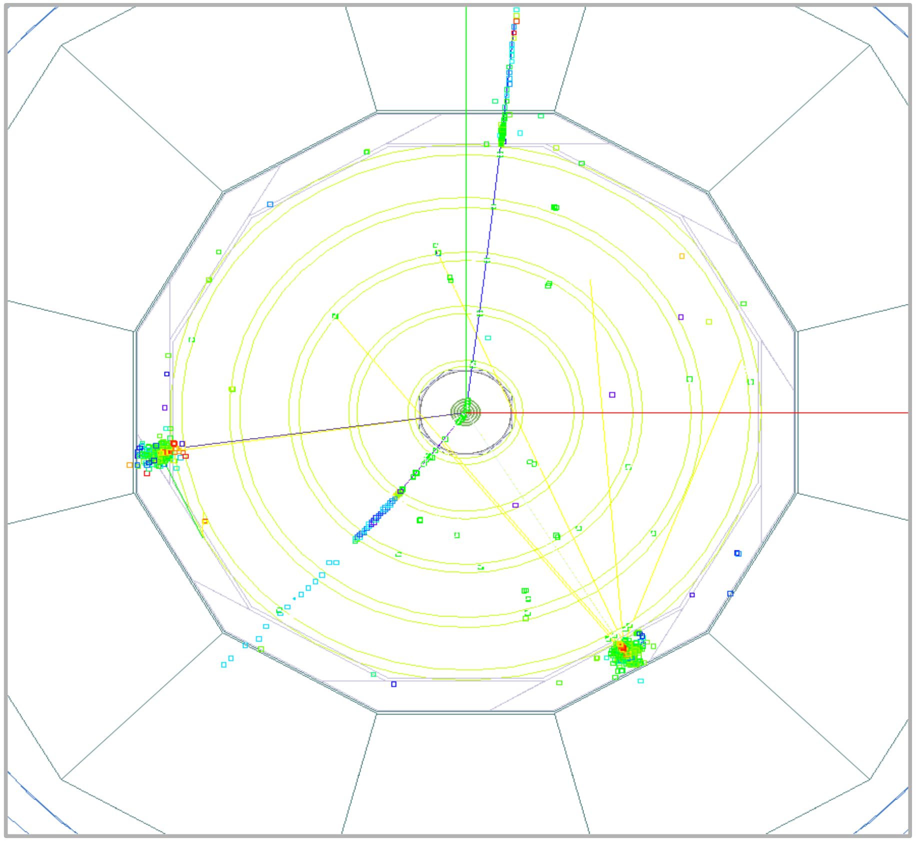
https://atlaswww.hep.anl.gov/asc/wikidoc/doku.php?id=fcs:cepc

Event display (e^+e^- 250 GeV CM energy)

SiDcc1



$H(125) \to 4\,e$

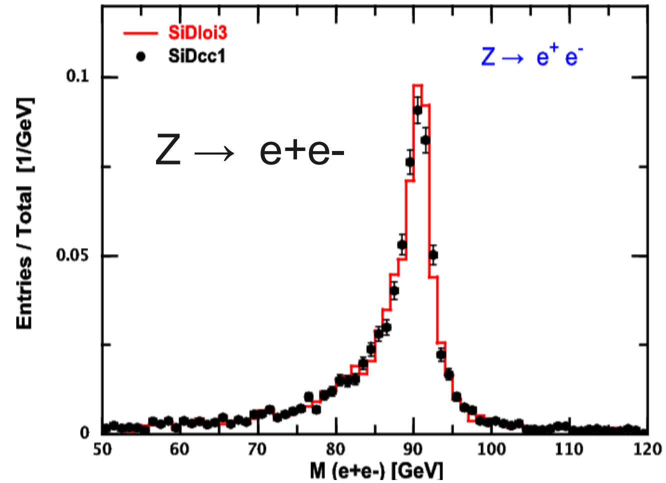


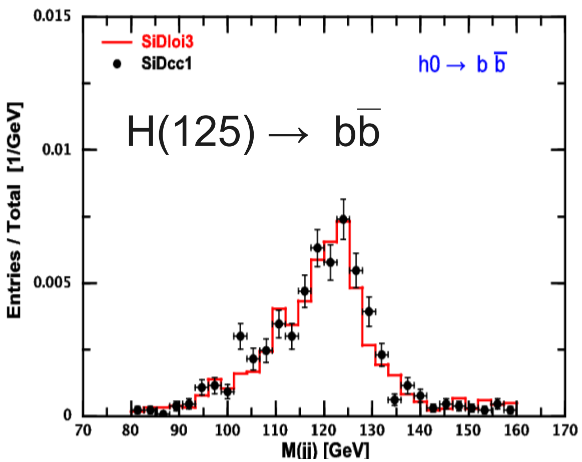
$H(125) \to \gamma\gamma$

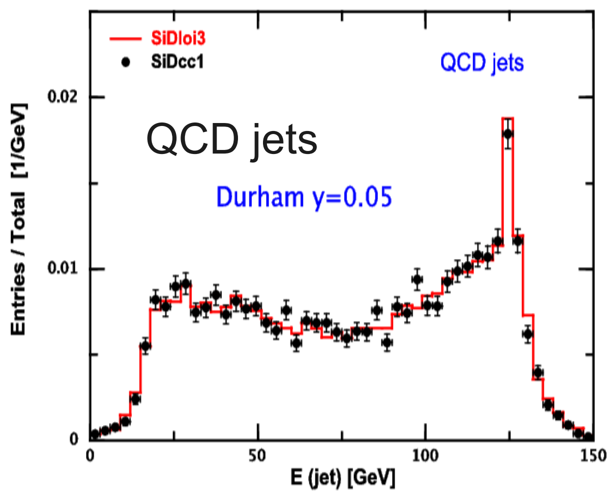
Comparing SiD with SiDCC1

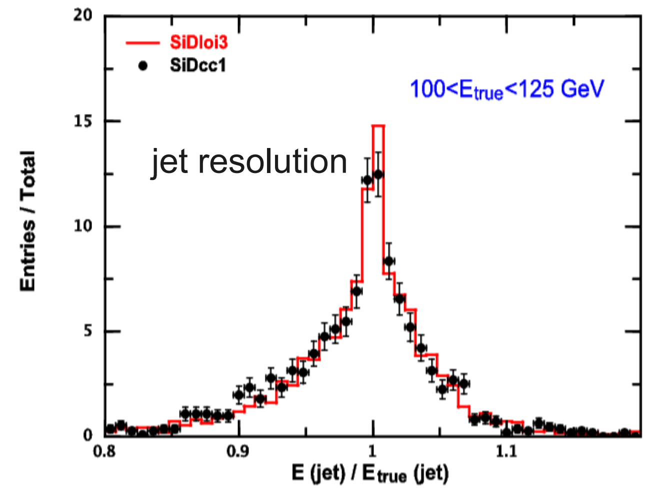
Done with Jas4pp

- Benchmark processes for e+e- (250 GeV)
  - Z → e+e-, Z → $b\bar{b}$, Z → tau+tau-, H → γγ
  - H → 4l, H → $b\bar{b}$, QCD jets
- Use particle flow objects to reconstruct invariant masses and jet energy resolutions using the Durham jet algorithm









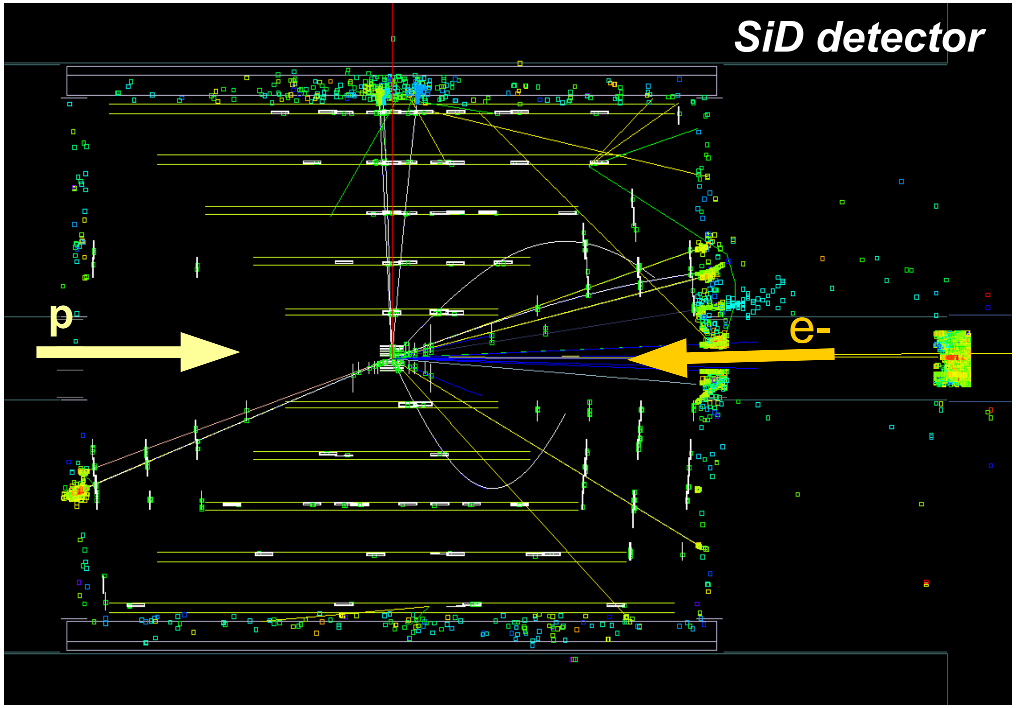
Simplification of the SiD detector does not compromise physics performance

ep collisions in the SiD detector

SiDlo3

(SiD detector)

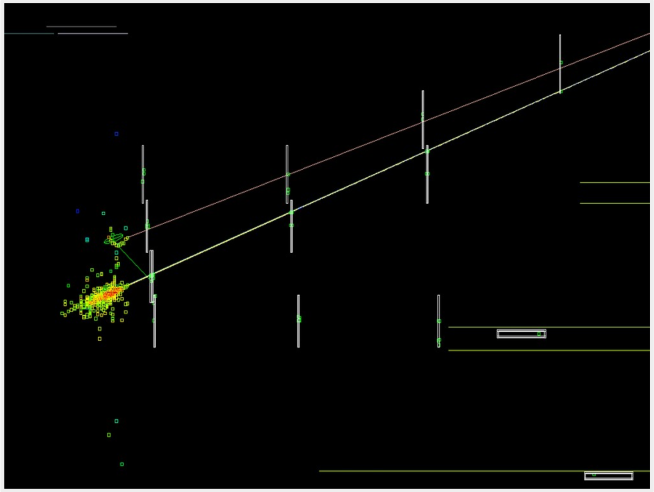
- Re-purpose SiD design for the Electron-Ion Collider (EIC) ?
- Optimize the SiD detector for electron-ion collisions



Reconstructed electron energy from PFA: **E=16.92 GeV**

"EVGEN" energy: **16.90 GeV**

scattered electron in ECAL:



DIS sample (Q^2> 5 GeV^2)
CM energy 141 GeV ("EIC-like")

HepSim Monte Carlo samples:
http://atlaswww.hep.anl.gov/hepsim/info.php?item=159

Converting SiD detector to SiFCC for a 100 TeV pp collider

With contributions from:
A.Kotwal (Fermilab/Duke), L.Gray (Fermilab), J.Strube (PNNL), N.Tran (Fermilab), S. Yu (NCU), S.Sen (Duke), J.Repond (ANL), J.McCormick (SLAC),J.Proudfoot (ANL), A.M.Henriques Correia (CERN), C.Solans (CERN), C.Helsens (CERN)

Requirements for FCC-hh hadronic calorimeter

- **Good containment up to 20 TeV jets**
  - affects jet energy resolution & leakage biases
- **Good longitudinal segmentation**
  - affects jet energy resolution
- **Good transverse segmentation**
  - resolving boosted particles (M~10-40 TeV range)

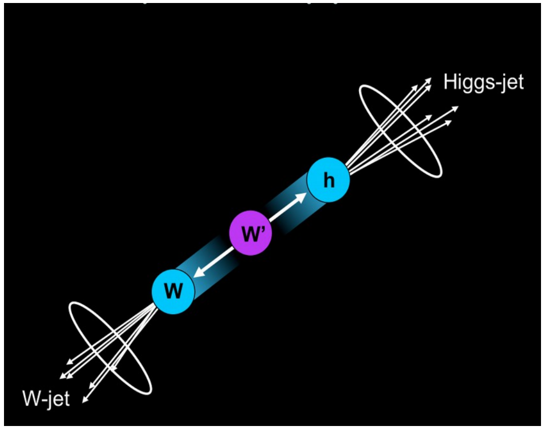
Optimize performance and sensitivity to new physics using appropriate technologies

Require detailed Geant4 simulations

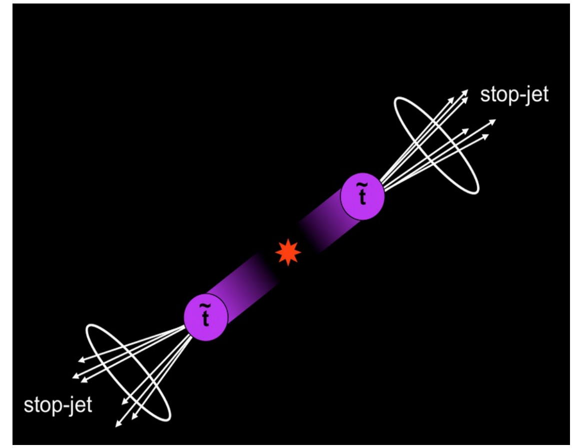
Lateral segmentation. Where does it matter..

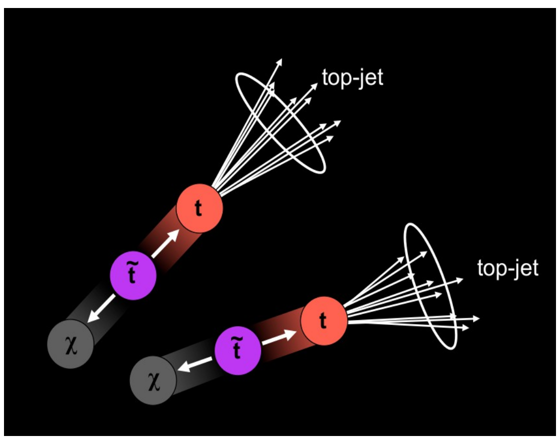
Brock Tweedie. Next steps in the Energy Frontier. LPC@FNAL. Aug. 24, 2014

X → W / Z / Higgs / top

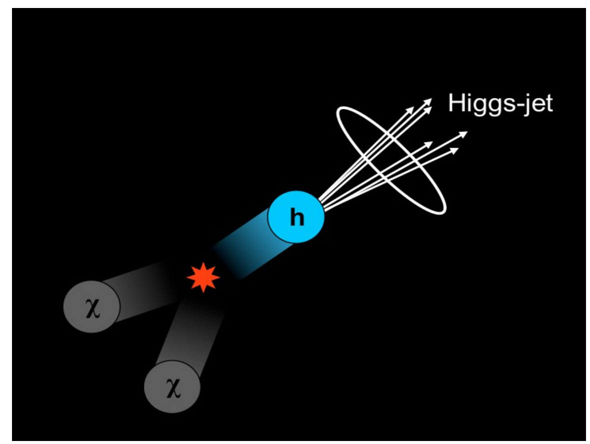


X → quarks/gluons





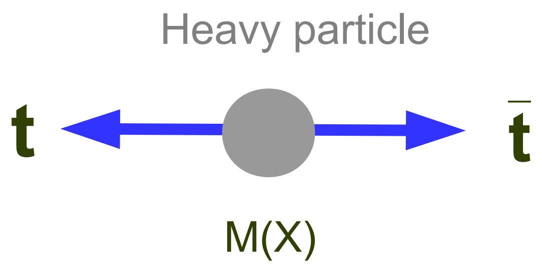
TeV-scale pair-produced



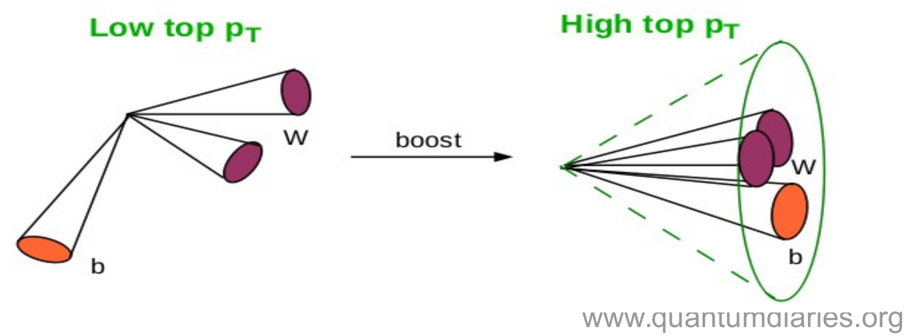
SM + dark matter

Large mass → large Lorentz boost → large collimation of decay products

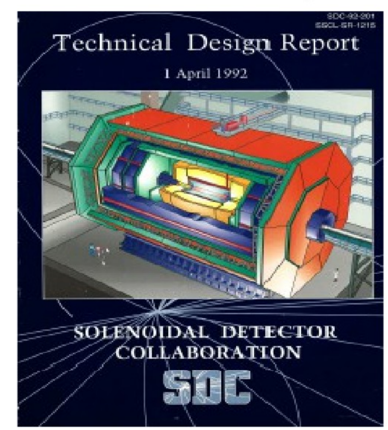
Boosted top from high-mass particles



$\Delta R \sim 2\,pT\,/\,m(top)$



- M(X)~10 TeV → top quarks with pT(top) > 3-5 TeV
- ΔR distance between 2 particles (W,b) from top decay
- SM physics & 10 ab^{-1} for FCC-hh: 5M $t\bar{t}$ events with pT(top)>3 TeV



- **SSC TDR:**
  - **mentions substructure signatures** and large R-jets for boosted Z (SSC-SR-1217 TDR 1992 p 3-26)
- **LHC:**
  - Boosted signatures is **one of the major topics**
- **FCC-hh:**
  - Detector design will be **based on boosted signatures** for top, Z/W, Higgs + modern techniques

Detector requirements driven by physics at 100 TeV

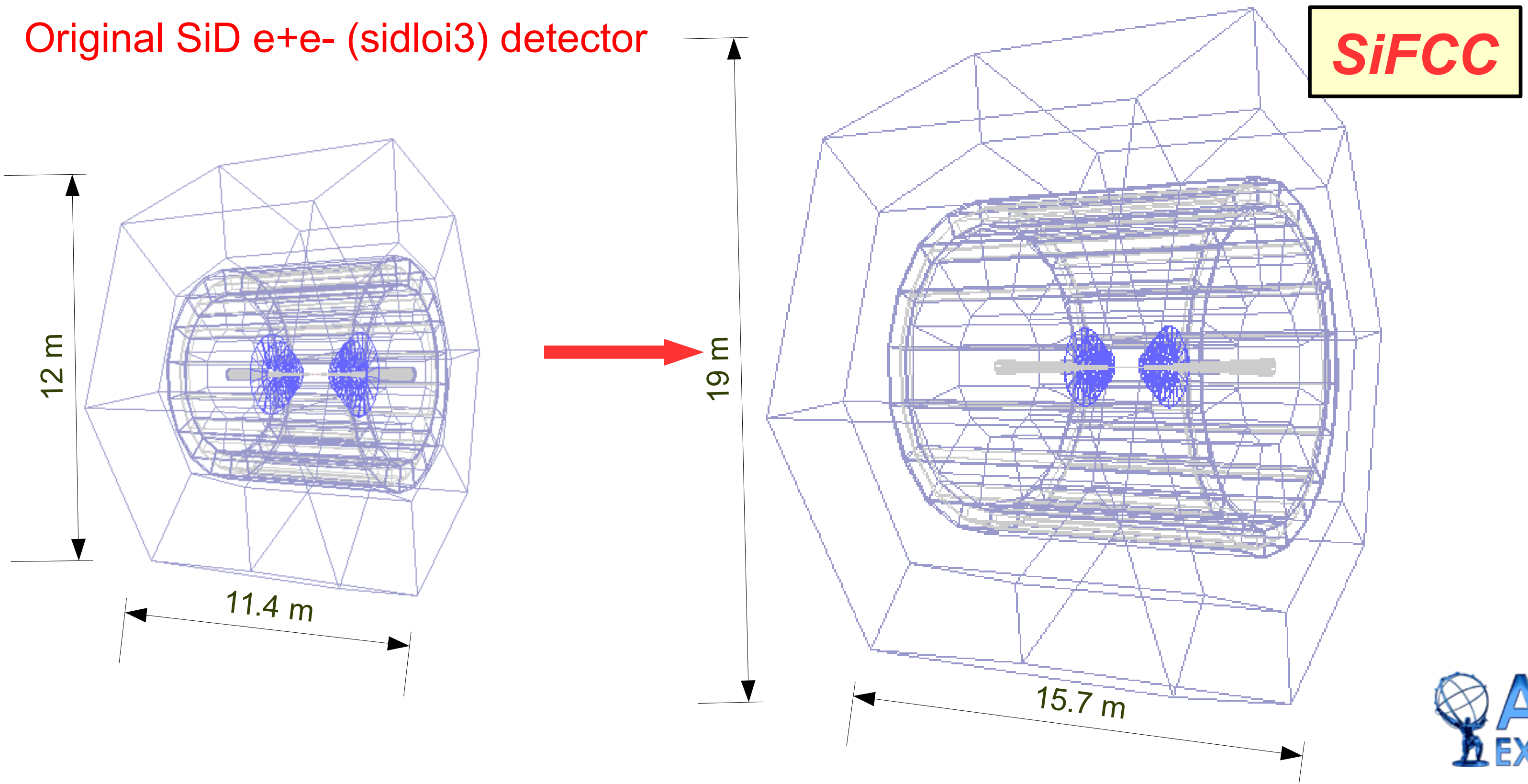
- **Good containment up to pT(jet)~30 TeV: 12 λ_I for ECAL+HCAL**
  - affects jet energy resolution
  - leakage biases, etc.
- **Small constant term for HCAL energy resolution: c < 3%**
  - dominates jet resolution for pT>5 TeV
  - important for heavy-mass particles decaying to jets
- **Longitudinal segmentation:**
  - Not studied
- **Sufficient transverse segmentation for resolving boosted particles:**
  - baseline $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$ from previous Delphes studies
  - 5x5 cm assuming ~ATLAS-like inner radius (~2.3 m from IP)

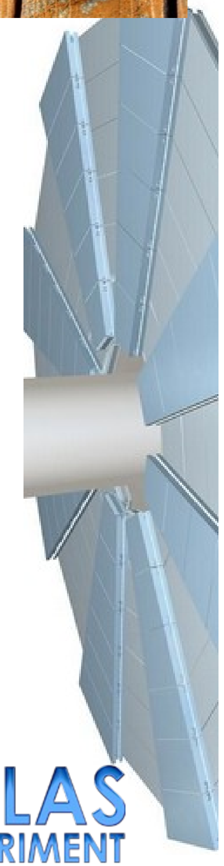
Require:

- detailed Geant4 simulations ..
- realistic reconstruction (including particle flow, i.e. tracks!)

SiFCC detector for performance studies

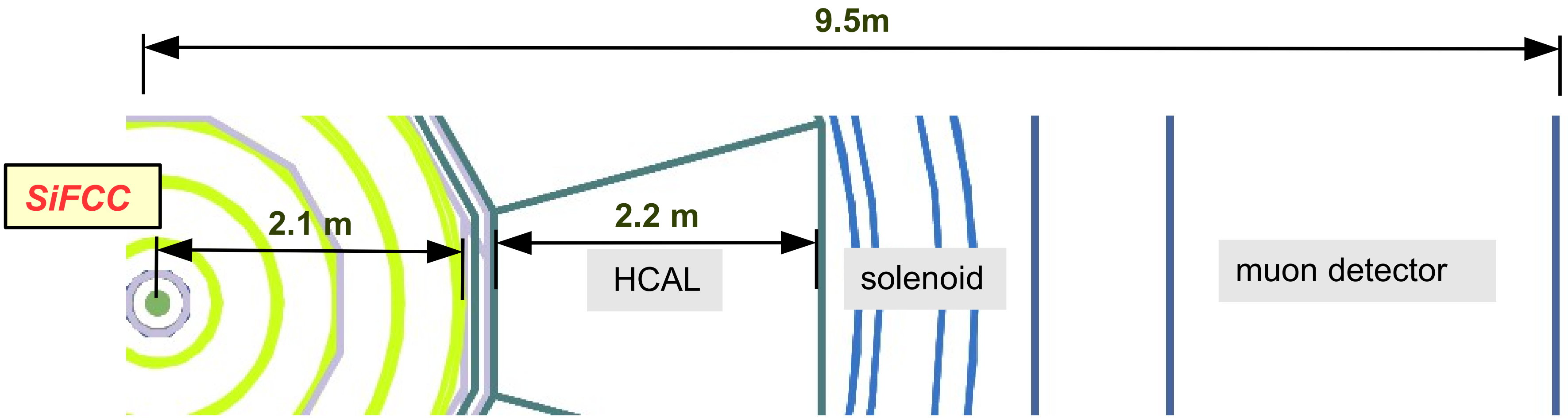
- Re-purpose SiD (ILC) detector and SLIC software
- Leverage large investments to R&D and software designs

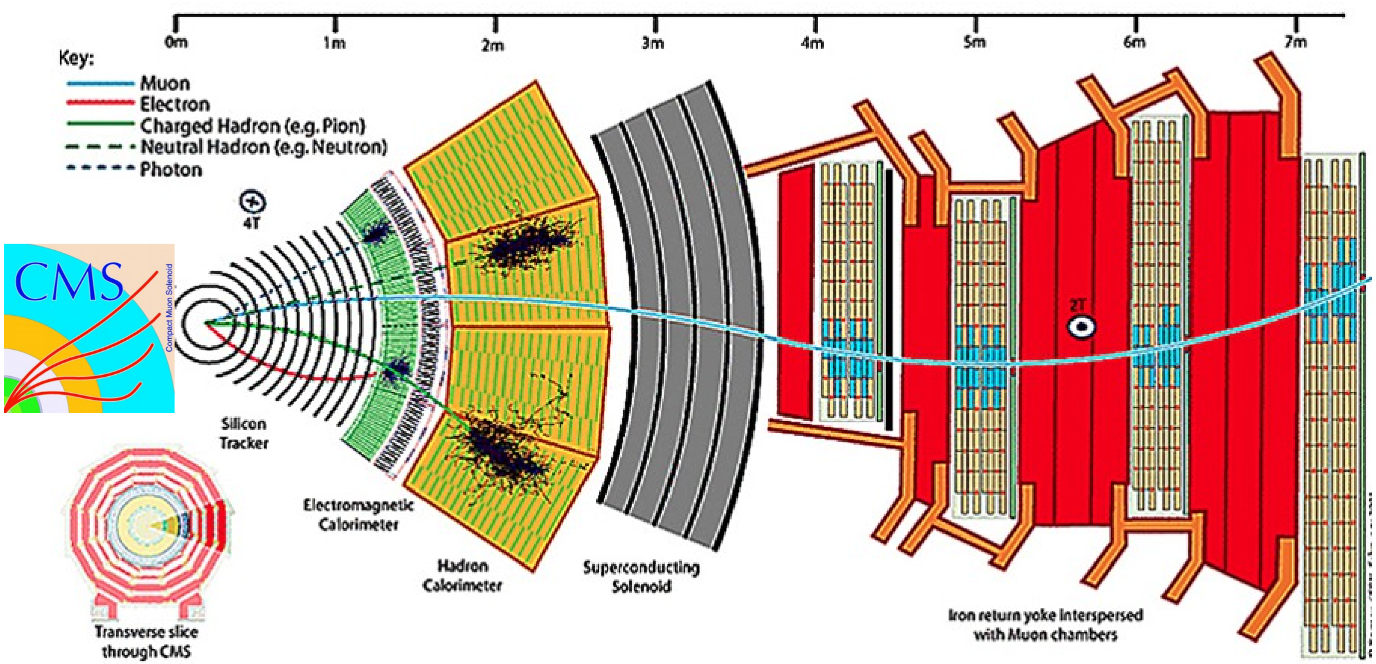




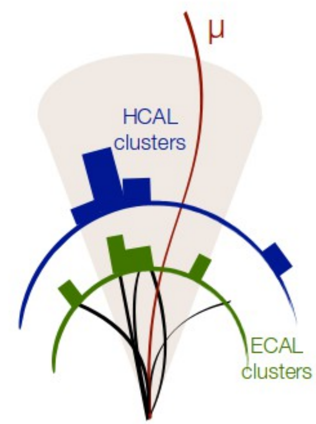
- **SiFCC (v4) detector: Multipurpose, high granularity, compact detector**
  - 30% smaller than ATLAS (R=25 m vs R=19), but with x20 better tracker!
  - 30% larger than CMS (R=14.6 m vs R=19 m)

SiFCC detector vs CMS





Both are optimized for Particle Flow Algorithms

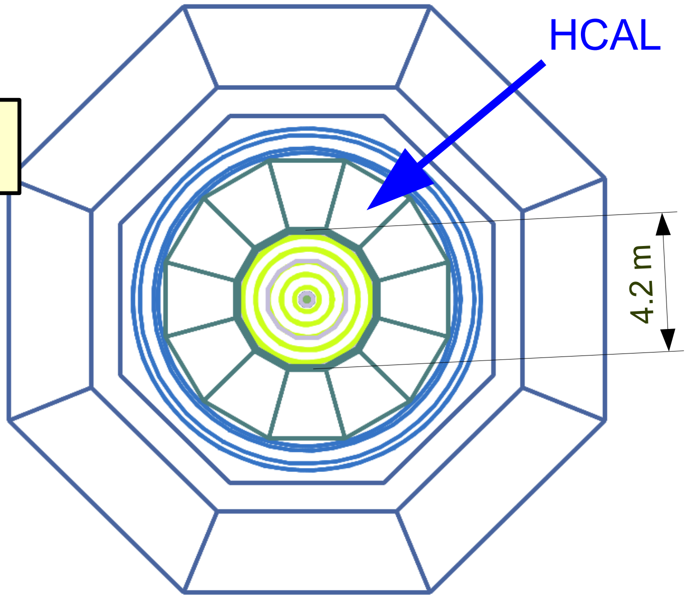


Characteristics of SiFCC (version 4)

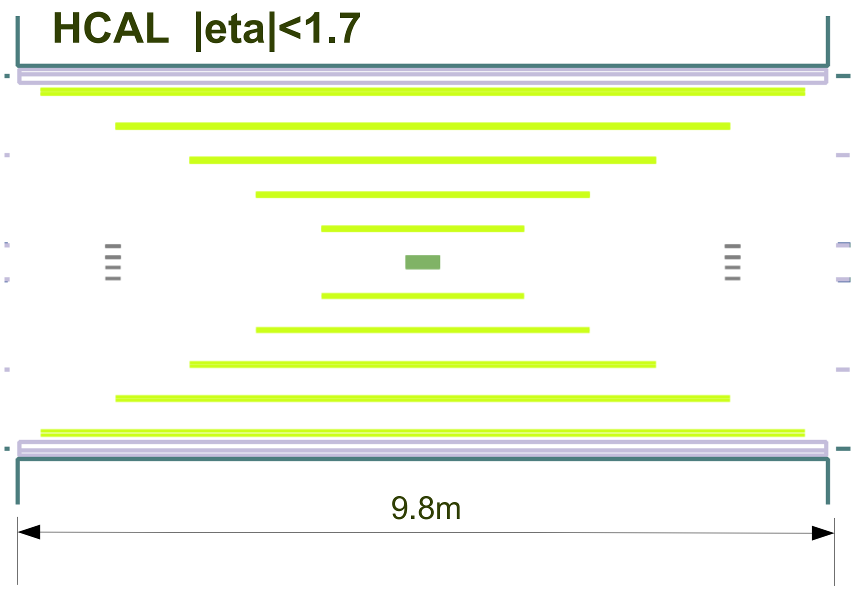
http://atlaswww.hep.anl.gov/hepsim/soft/detectors/sifcch4/sifcch4.html

- **5 T solenoid outside HCAL**
- **Pixel and Outer trackers:**
  - 20 um pixel (inner), 50 um (outer)
- **ECAL (Scint+W):** 2x2 cm. 32 layers, ~35 X0
- **HCAL (Scint+Fe) ~ FCC-hh baseline**
  - 5x5 cm cells: $\Delta\eta \times \Delta\varphi = 0.022 \times 0.022$
    - CMS: $\Delta\eta \times \Delta\varphi = 0.087 \times 0.087$
    - ATLAS: $\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$
  - Longitudinal: 64 layers, 11.3 λ_I
  - 3.1% sampling fraction
  - > 150 million cells, non-projective

SiFCC



trans. cell size: 5 cm ~ λ_I(Fe) / 3



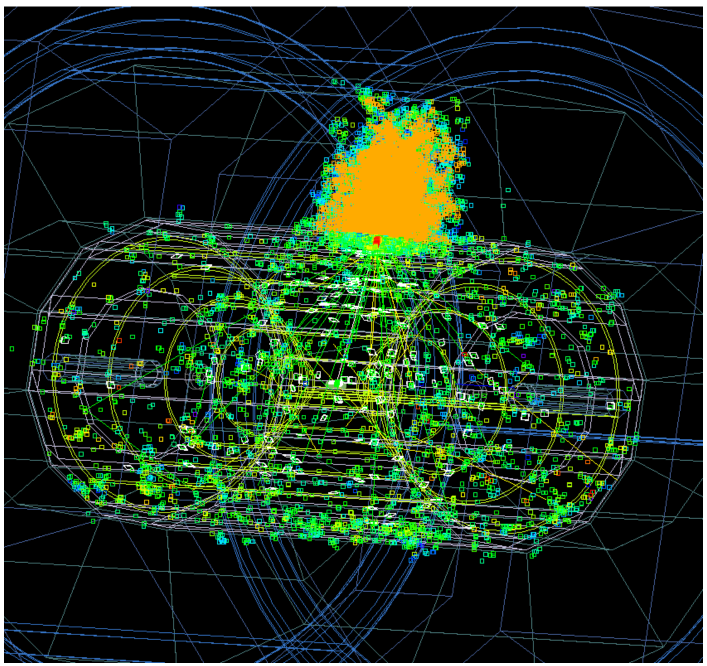
Can reconstruction of TeV-scale objects benefit from small HCAL cells?

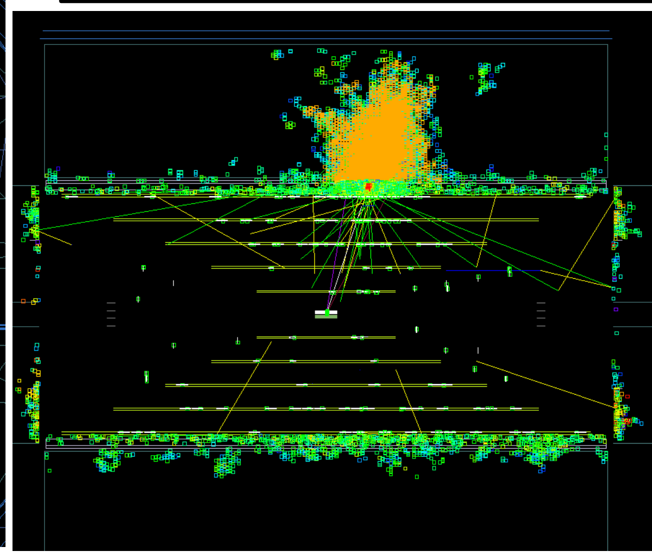
Response to single particles: 1 TeV

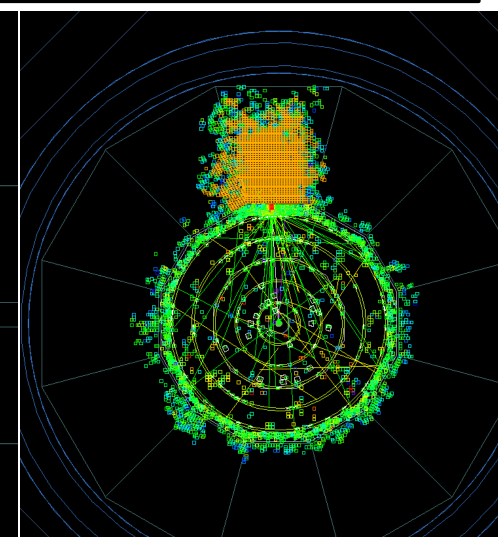
- Use single pions 1 GeV – 10 TeV to study detector performance
- 1 TeV pions are benchmarks used in arXiv:1604.01415 (shown in Washington DC)
  - pT(jet)>30 TeV: ~10% will be carried by 1 TeV hadrons (~9 hadrons/jet)

Example: 1 TeV π+

- 7300 calorimeter hits, 440 SiTracker hits
- 1 reconstructed PFA (pi+) =998 GeV
- 1 reconstructed CaloCluster at 1058 GeV
- Many back-splash interactions







Based on **HepSim**: http://atlaswww.hep.anl.gov/hepsim/info.php?item=182

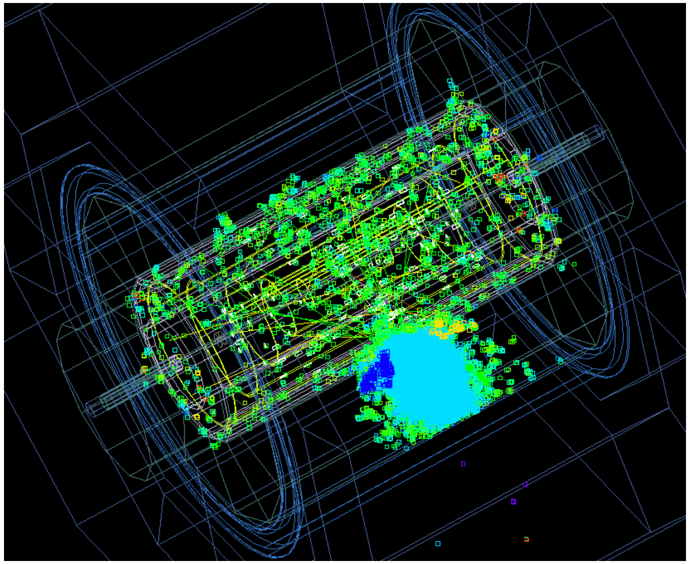
Response to single particles: 8.1TeV pions

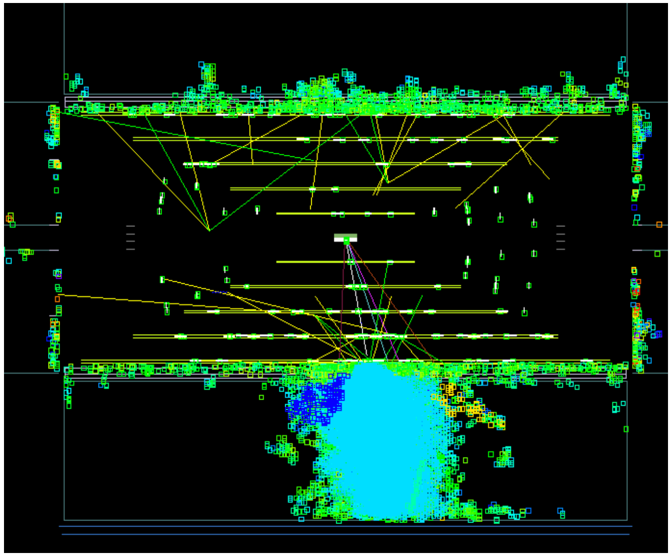
SiFCC

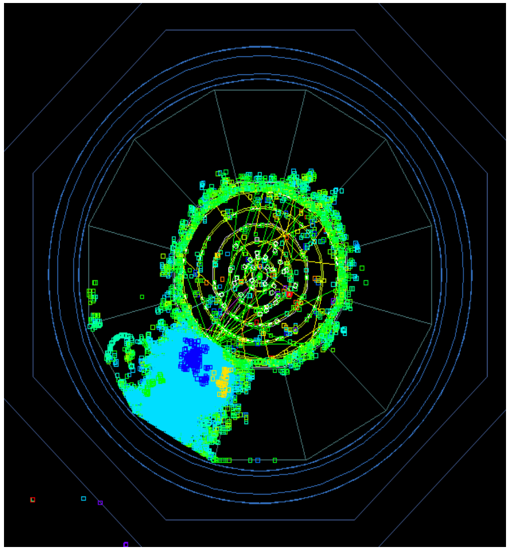
Example: 8.156 TeV π+

Energy leakage outside HCAL?
Energy scale need to be corrected?

- ~30000 calorimeter hits, ~500 SiTracker hits
- 1 reconstructed PFA (pi+)=8.97 TeV
- 1 reconstructed CaloCluster at 8.40 TeV
- Many back-splash interactions



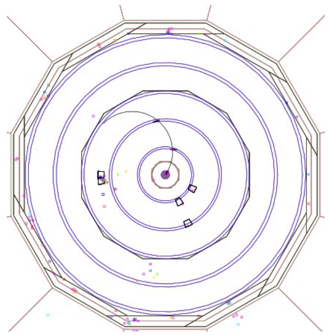


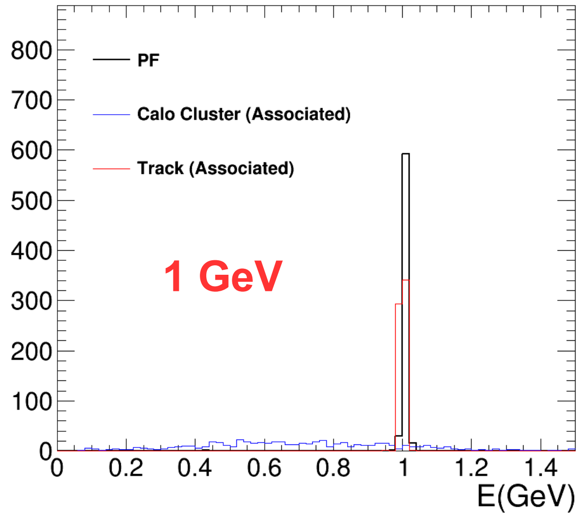


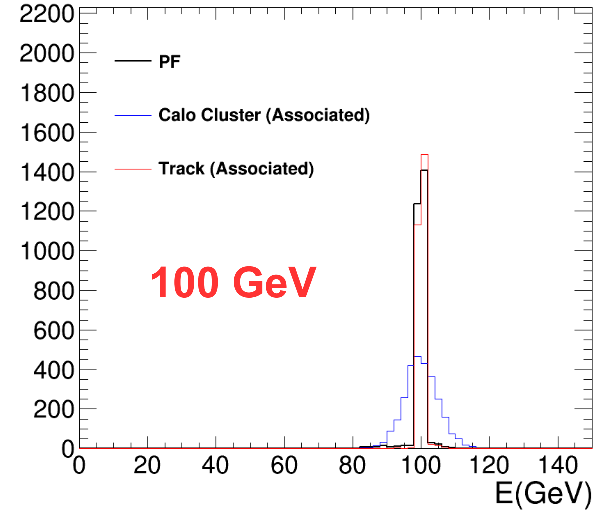
Based on **HepSim**: http://atlaswww.hep.anl.gov/hepsim/info.php?item=201

Detector response to single particles

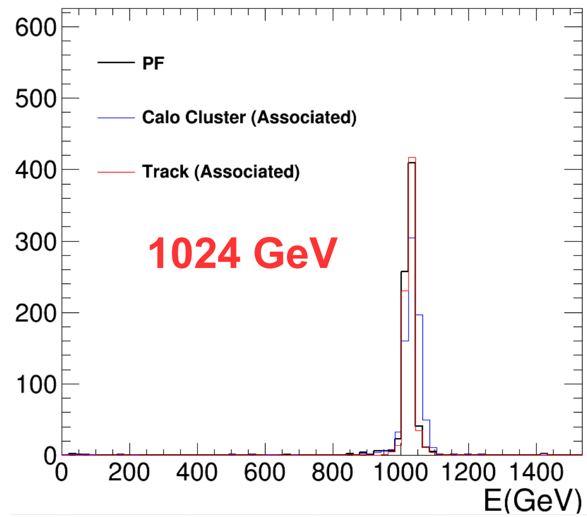
SiFCC

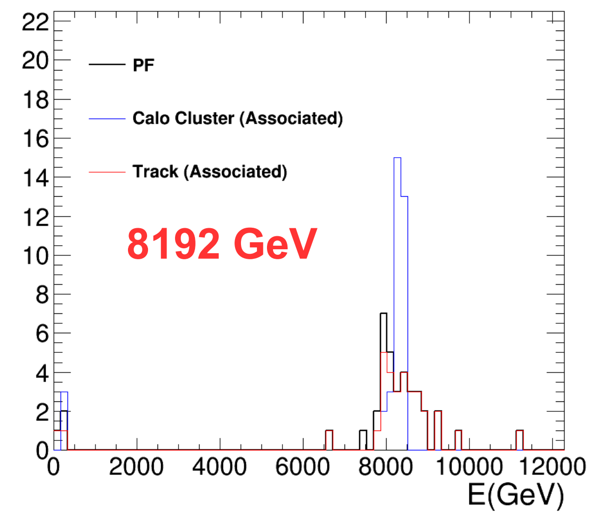






Losses of clusters for < 2 GeV charged particles due to **5 T** field and increased inner radius to 2.1 m



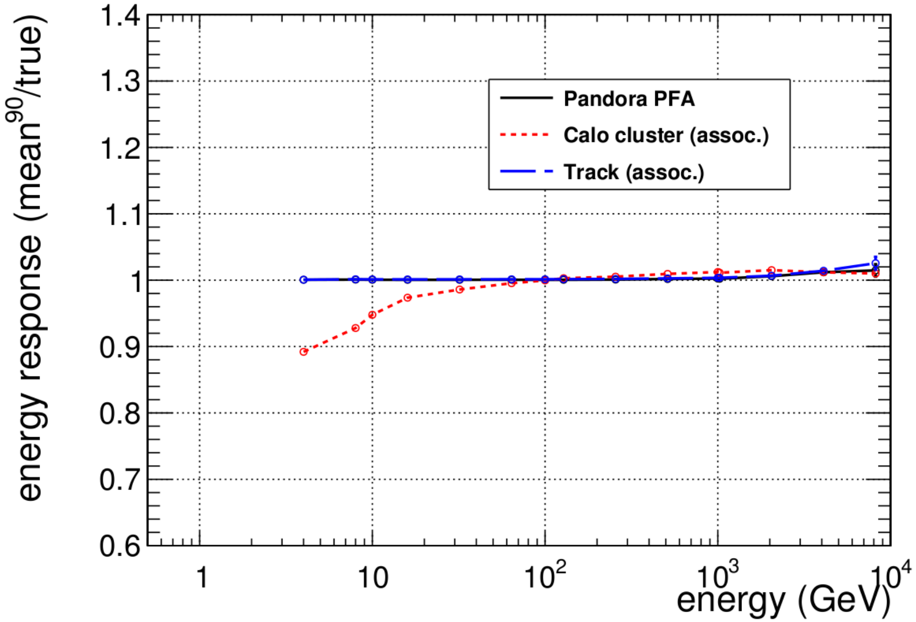


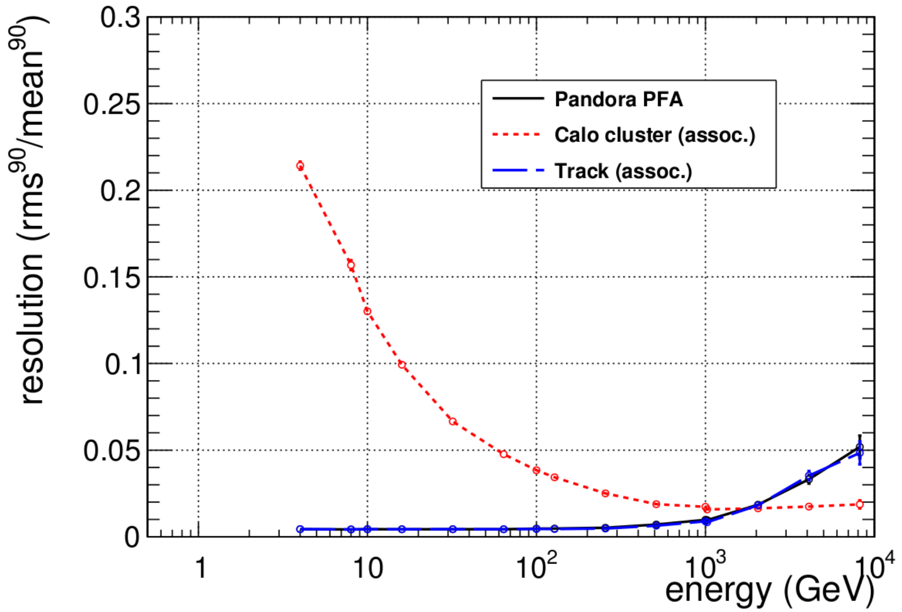
p = 0.3 * B * r

- p – momentum (GeV)
- B - solenoid field (in T)
- r - is the radius (in m)

Single particle response

SiFCC





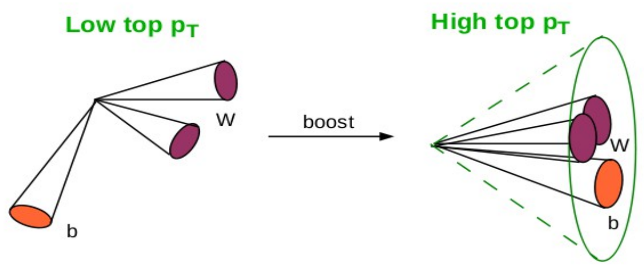
- Losses of clusters with low momentum due to 5 T
- Resolution of tracks & PFA getting worse with energy
- Resolution for CaloClusters is better than PFA/tracks for E> 2 TeV

  ~ 2% for clusters, 5% for tracker near 8 TeV

Estimates based on: dpT/pT= 8 sig*pT/(0.3*B*L2) are more conservative*

Physics processes for boosted jet studies

- Muon collisions to speed up calculations: no complications due proton beams
- Processes for benchmarks:
  - $\mu+\mu- \to Z' \to W+W-$
  - $\mu+\mu- \to Z' \to qq$
  - $\mu+\mu- \to Z' \to t\bar{t}$
  - $\mu+\mu- \to Z' \to$ tau+tau-
  - $\mu+\mu- \to Z' \to$ bbar



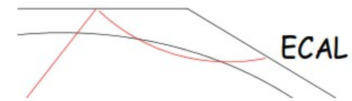
- Reconstructed samples in the LCIO format assuming:
  - $\Delta\Gamma(Z')\sim$ 1 MeV
  - Z'(20 TeV) and Z'(40 TeV)
- Apply favorite substructure techniques to identify WW, $t\bar{t}$ (compare with $Z' \to q\bar{q}$)

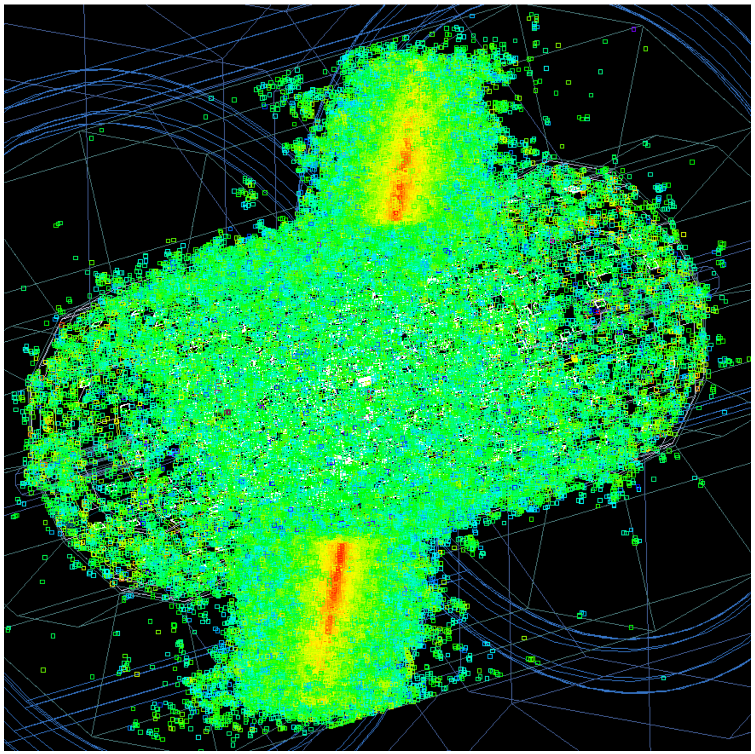
- about 2000 fully reconstructed events per sample (Tracks, PFA, CaloClusters, HITS)
- created on Open-Science Grid (UChicago/ANL. ~100,000 CPU*h)
- Find: http://atlaswww.hep.anl.gov/hepsim/list.php?find=rfull006

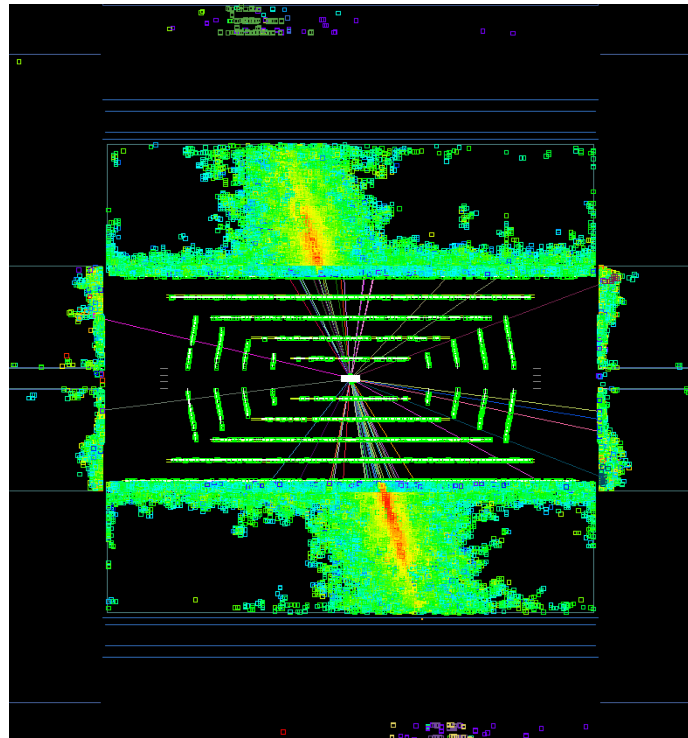
Event display of Z' (40 TeV) → $W^+ W^-$ → hadrons

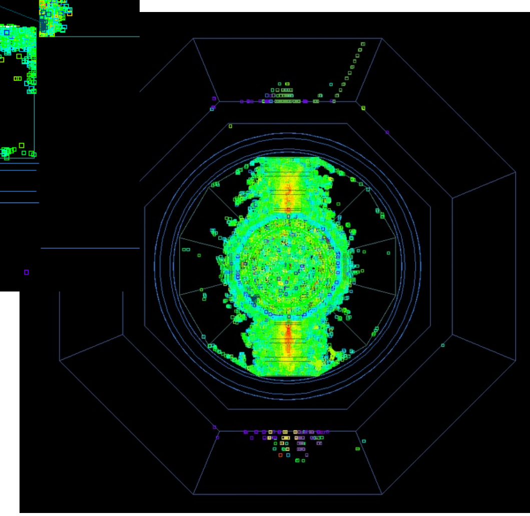
SiFCC

Busy event, large number of back-splash interactions in ECAL/HCAL/Tracker
~4 CPU*h to simulate/reconstruct, 16 GB RAM
→ CPU intensive!









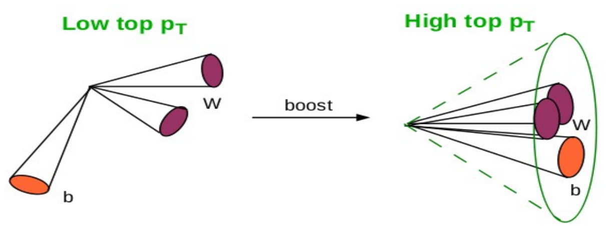
10,000 hits in ECAL
46,000 hits in HCAL
12,000 hits in outer tracker
1,000 hits in the inner tracker

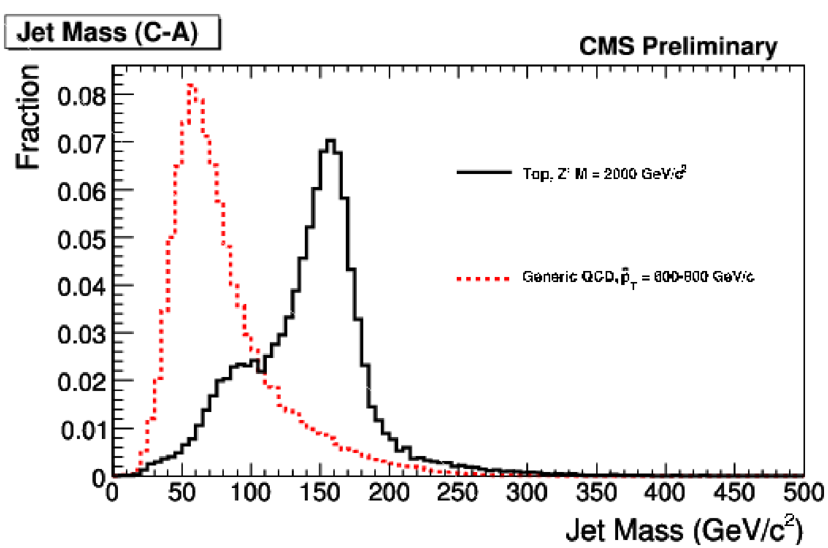
Available for download: Z' → WW, Z' → $t\bar{t}$, Z' → $b\bar{b}$ for different Z' masses

Jet masses for highly boosted jets

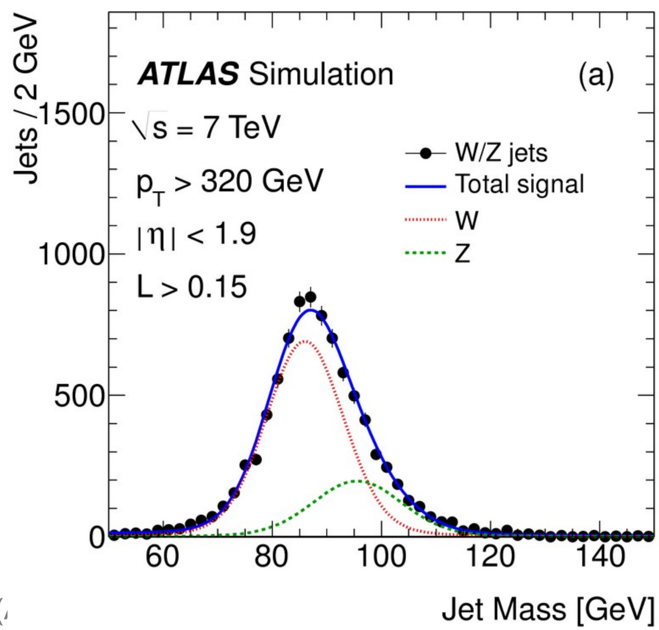
- Simple observable constructed from energies and positions of jet constituents
  - requires high spatial resolution of jet constituents
  - sensitive to calorimeter granularity
- Critical for many searches by ATLAS & CMS
  - signal extraction, background rejection etc: boosted W, top, Higgs etc.

$$m^2(jet) = \sum E_i^2 - \sum \boldsymbol{p}_i^2$$





$W \rightarrow q\bar{q} \rightarrow jet$

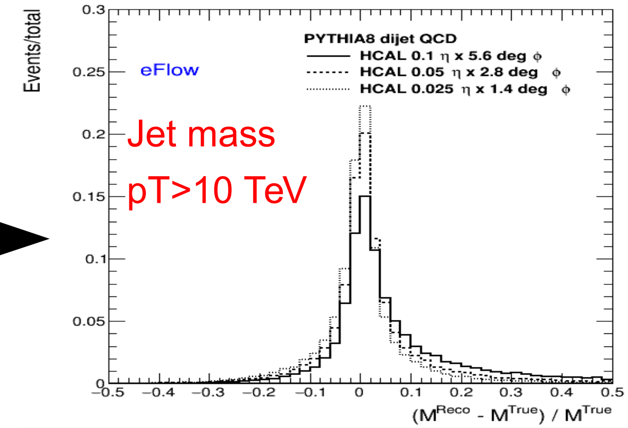


Jet masses for highly boosted jets

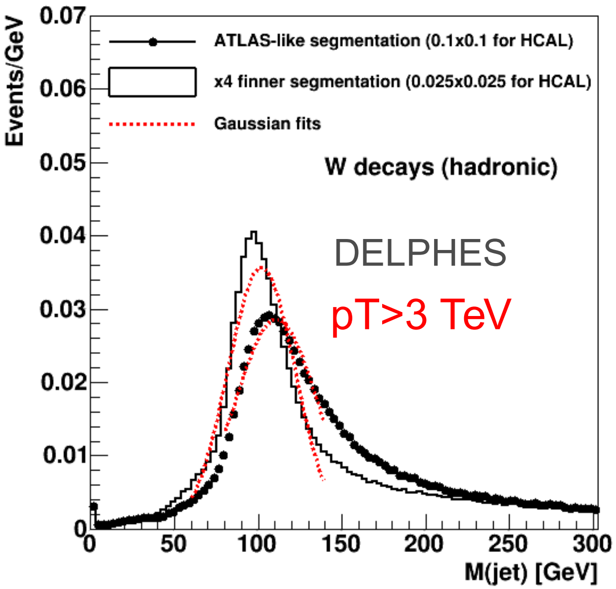
presented at Boost2015 & FCC week in DC

- DELPHES fast simulation shows significant improvement in mass resolution compared to $\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$ cells
  - 80% for $\Delta\eta \times \Delta\varphi = 0.05 \times 0.05$
  - 120% $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$

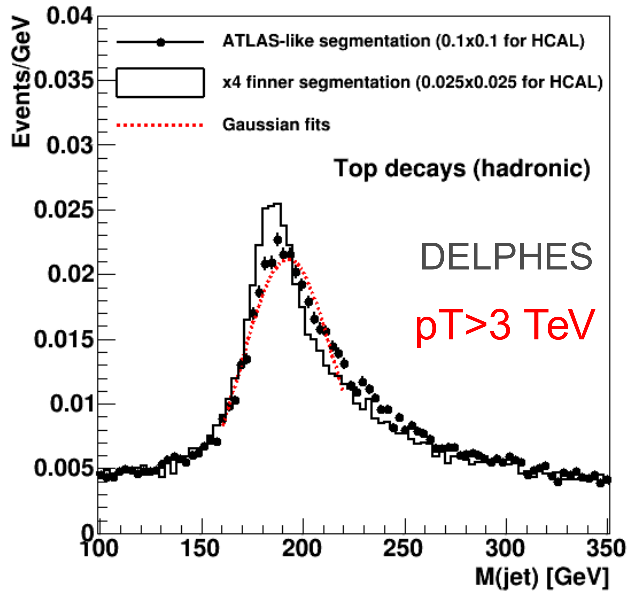
Delphes+HepSim



Boosted W



Boosted top



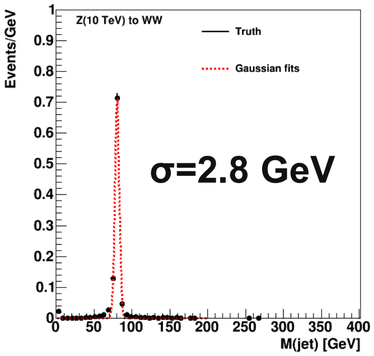
From the Gaussian fits:

W mass:
σ = 23 GeV (0.1x0.1)
σ = 20 GeV (0.025x0.025)

Top mass:
σ = 24 GeV (0.1x0.1)
σ = 21 GeV (0.025x0.025)

Not too realistic:
no longitudinal segmentation, secondary interactions, realistic Geant4 reconstruction, high-pT tracking loses, etc. etc.

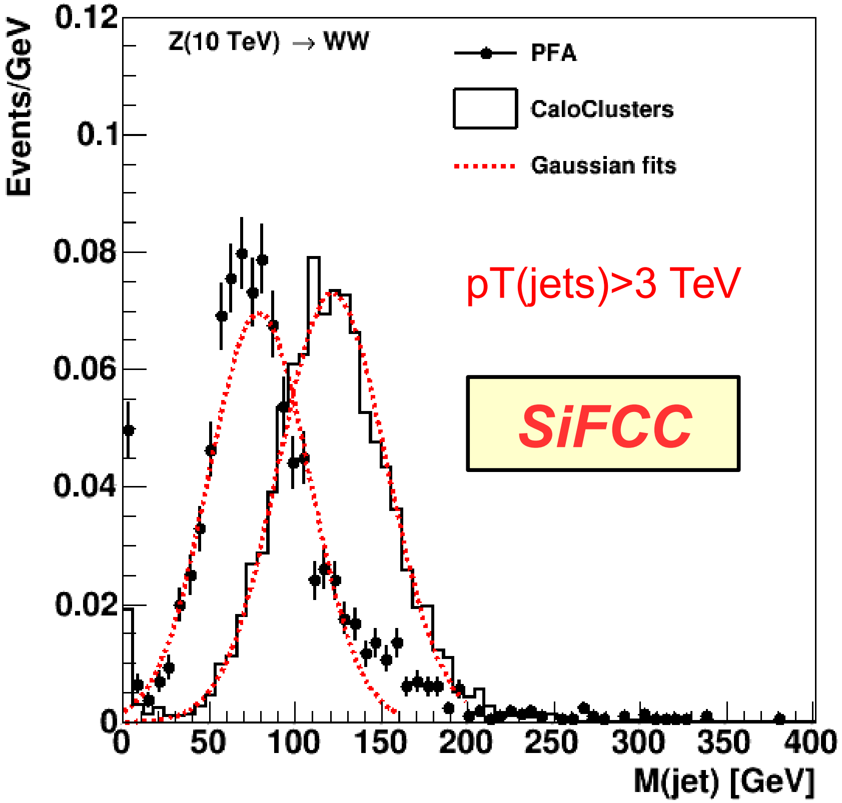
Jet mass for W → q$\bar{q}$ (boosted) in the *SiFCC* detector



Truth-level jets
FastJets R=0.6
pT(jet)> 3 TeV
|η(jet)|<1.2

SiFCC full simulation & reconstruction:

- PFA: σ=29 GeV, peak= 79 GeV
- CaloClusters: σ=31 GeV, peak=121 GeV



- PFA and CaloClusters have similar jet width (dominated by pT~ 5 TeV)
- Shift in jet mass for CaloClusters can be due to:
  - large contribution from secondary interactions & spread of particles in 5T field
  - removing soft constituencies (soft drop) reduces the the jet mass built from clusters
- SiFFC has larger jet width compared to DELPHES (~ 20 GeV)

Contributions to HepSim software

- E. May - ProMC format development, benchmarks on BlueGene/Q (ANL)
- K. Strand (SULI 2014) - ProMC conversion tools
- P. Van Gemmeren - testing ProMC format
- T. Sjöstrand - ProMC integration with Pythia8
- P. Demin - ProMC integration with Delphes
- I. Pogrebnyak - (U.Michigan) software validation toolkit, fastjet in Java
- D. Wilbern (SULI 2015) - Pileup mixing tool based on ProMC
- M. Selvaggi - Delphes card for ILD geometry and “**EIC”-like (requested by S.C.)**
- H. Gray - Delphes card for FCC-hh geometry
- J. Strube (PNNL) - LCIO/SLIC for full simulation
- A. Kotwal (Duke Univ.) - LCIO/SLIC for full simulation
- J. Adelman (NIU) – H+tt sample + post-Snowmass Delphes 3.3 card for 13/14 TeV
- S. Padhi - prototyping Snowmass Delphes 3.1 during Snowmass 2013
- K. Pedersen - alternative b-tagging for rfast003 in HepSim
- Shin-Shan Yu - Heavy Higgs MG5 simulations for HepSim

A lot of help / advise from J.McCormick and N.Graf (SLAC)

Documentation

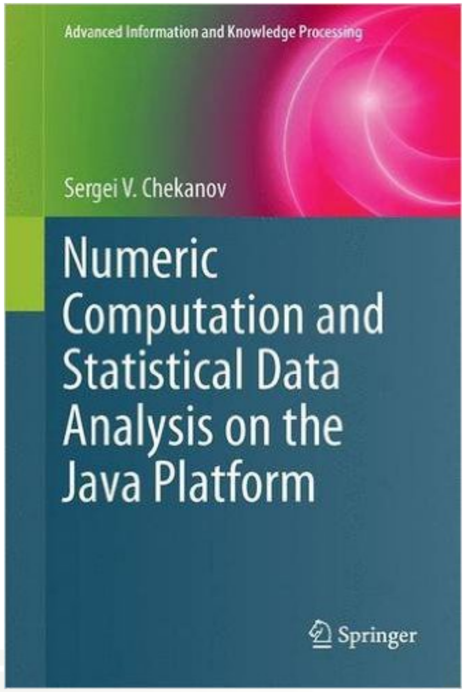
- HepSim short manual: http://atlaswww.hep.anl.gov/hepsim/description.php
- HepSim Wiki:https://atlaswww.hep.anl.gov/asc/wikidoc/doku.php?id=community:hepsim

Physics and detector studies

Here are several links to extending this Wiki for particular detector-performance topics:

- FCC-hh detector studies - explains how to analyses data for FCC-hh detector studies
- SID detector studies - explains how to analyses data for the SiD detector (ILC)
- CEPC detector studies - shows some results with full simulations for CEPC
- EIC detector studies - shows some results with full simulations for EIC
- HCAL studies explains how to analyse ROOT data after fast detector simulations used for FCC studies

- Jas4pp Wiki: https://atlaswww.hep.anl.gov/asc/wikidoc/doku.php?id=asc:jas4pp
- Many examples are coded in Python/Jython and C++
- Look at this book describing much of the Python/Java API →



How to contribute to HepSim

- Generate EVGEN archive files with physics processes
- Validate using the HEPSIM tools (if you can)
- Contribute to the software tools
- Run a data server and maintain your own EVGEN & full simulation files

Support (limited, on a voluntary basis): (contact hepsim@anl.gov)

- HEPSIM integration, deployment, OSG-grid, EVGEN MC, fast sim etc.
  - ANL: S.C.
- Some support for SLIC software (used for ILC)
  - SLAC: N.Graf & J.McCormick, PNNL: J.Strube
- Configure detectors, physics, analysis package for circular colliders
  - ANL/Fermilab: S.C., A.Kotwal

Thanks!

Backup

Benchmarks for EVGEN files

File sizes for 10,000 $t\bar{t}$ events for pp at LHC

ProMC files:
- 12 times smaller than HEPMC
- 30% smaller than ROOT

~30% faster to process (C++/Java)

File format	File Size (MB)	C++ (sec)	CPython (sec)	Java (sec)	Jython (sec)
ProMC	307	15.8	980	11.7 (12.1 +JVM startup)	33.3 (35 +JVM startup)
ROOT	423	20.4	66.7 (PyROOT)	-	-
LHEF	2472	84.7	30.4	9.0 (9.6 +JVM startup)	-
HEPMC	2740	175.1	-	-	-
LHEF (gzip)	712	-	-	-	-
LHEF (bzip2)	552	-	-	-	-
LHEF (lzma)	513	-	-	-	-
HEPMC (gzip)	1021			-	-
HEPMC (bzip2)	837			-	-
HEPMC (lzma)	802	-	-	-	-

ASCII text files
(after compression)

Table 1. Benchmark tests for reading files with 10,000 ttbar events stored in different file formats. For each test, the memory cache on Linux was cleared. In case of C++, the benchmark program reads complete event records using appropriate libraries. CPython code for ProMC file is implemented in pure CPython and does not use C++ binding (unlike PyROOT that uses C++ libraries). In case of LHEF files. JAVA and CPYTHON benchmarks only parse lines and tokenize the strings, without attempting to build an event record, therefore, such benchmarks may not be accurate while comparing with ProMC and ROOT.

https://atlaswww.hep.anl.gov/asc/wikidoc/doku.php?id=asc:promc:introduction

NLO calculations as “ntuples”

Theorists can use it too!

S.C. Adv. High Energy Physics, vol. 2015, 13609

- Several NLO calculations are available (MCFM, JETPHOX, NLOjet++)
- Data structure is different compared to full parton-shower MC
  - “Particle record”: Usually 4-momenta of 3-4 particles per events
  - “Event record”:
    - Event weights (double)
    - Deviations from central weights for different PDF eigenvector sets for calculations of PDF uncertainties

$$w_n = \left[1000 \times \left(1 - \frac{PDF(n)}{PDF(0)}\right)\right]$$

n=1...51 for CT10 PDF

Very large numbers of weighted NLO events can be compactly stored:

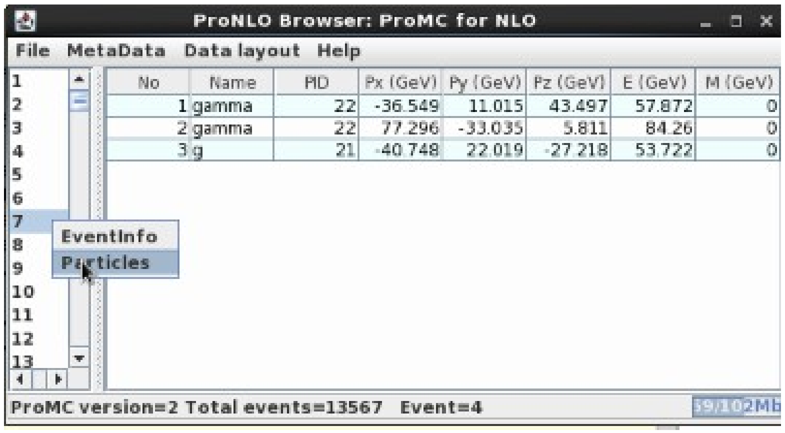
→ *double precision “weights” → int64 varint (deviations) → 2 bytes per weight*

→ *Large deviations are stored using 4 or 8 bytes (rarely)*

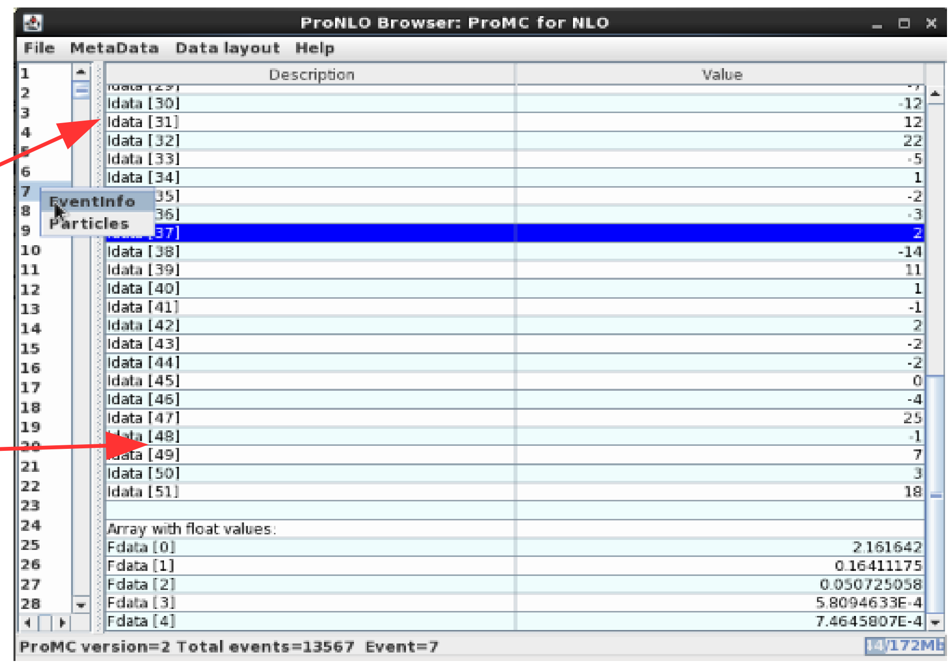
NLO calculations as “ntuples”

MCFM prediction for H(→ γγ)+jet (pp 100 TeV)
“higgsjet_gamgam_mcfm” sample

Some NLO samples using MCFM
have been created on Mira supercomputer
(BlueGene/Q)



4-momenta of particles



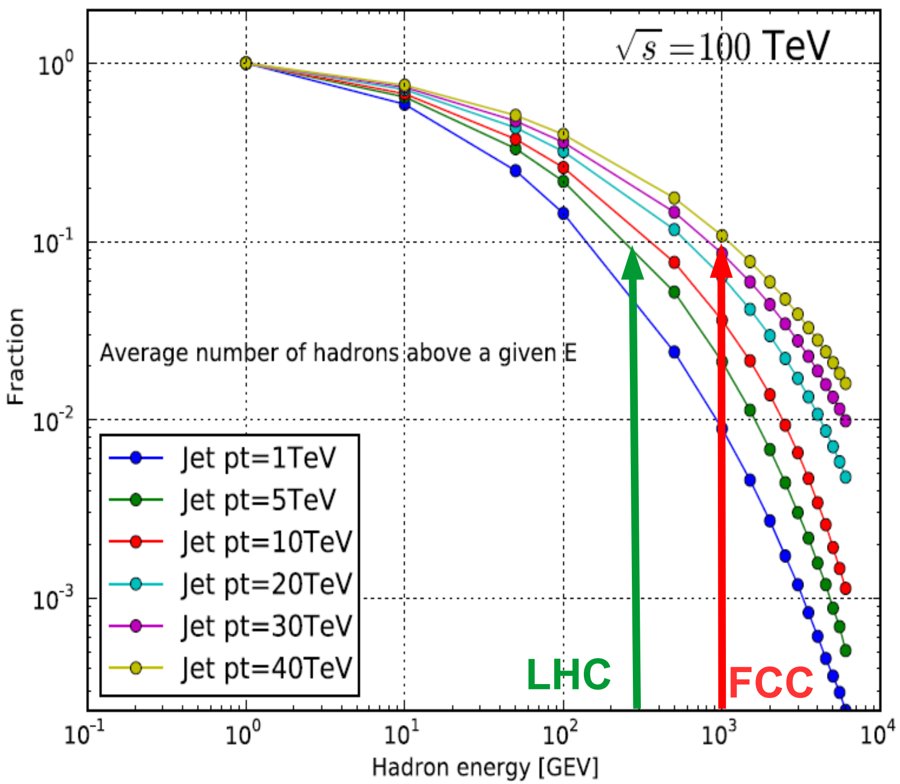
Event weights

PDF variations for CT10 (51)

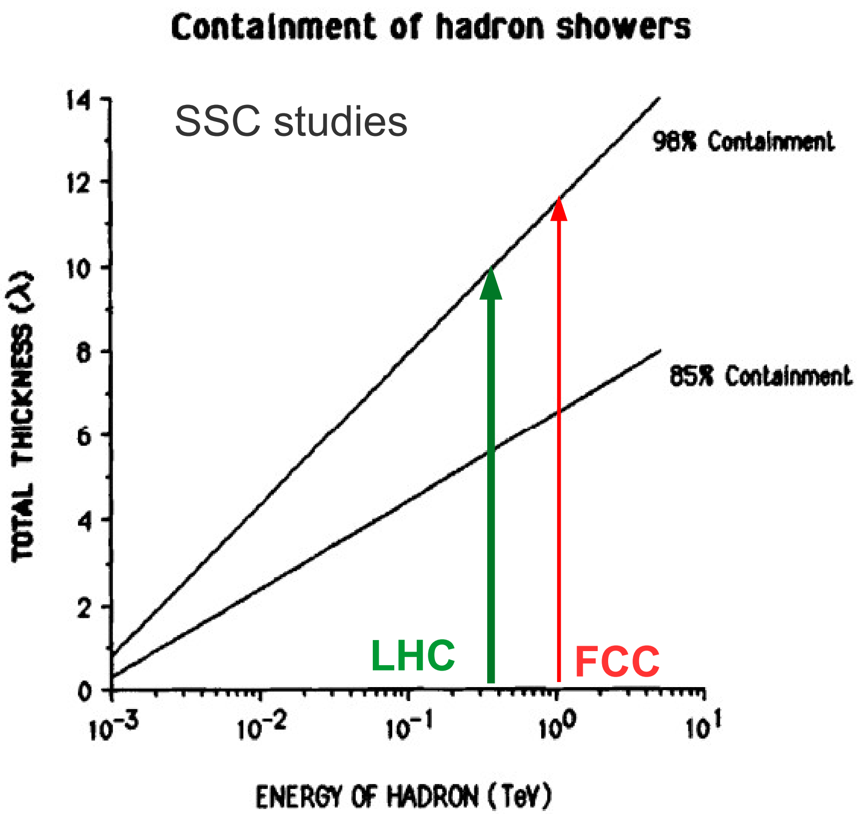
Estimating HCAL depth

Leading particles in high-pT jets

C.Helsens,C.Solans



http://lss.fnal.gov/conf/C860623/p355.pdf



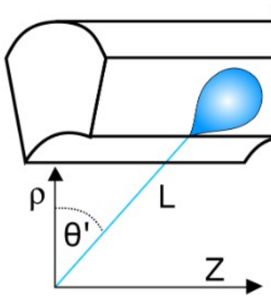
pT(jet)>30 TeV: ~10% will be carried by 1 TeV hadrons (~9 hadrons/jet)

12 λ_I is needed to contain 98% of energy of a 1 TeV hadron

Geant4 simulation agrees with calculations for SSC (.. 1984 Gordon&Grannis. Snowmass)

Resolution for single pions

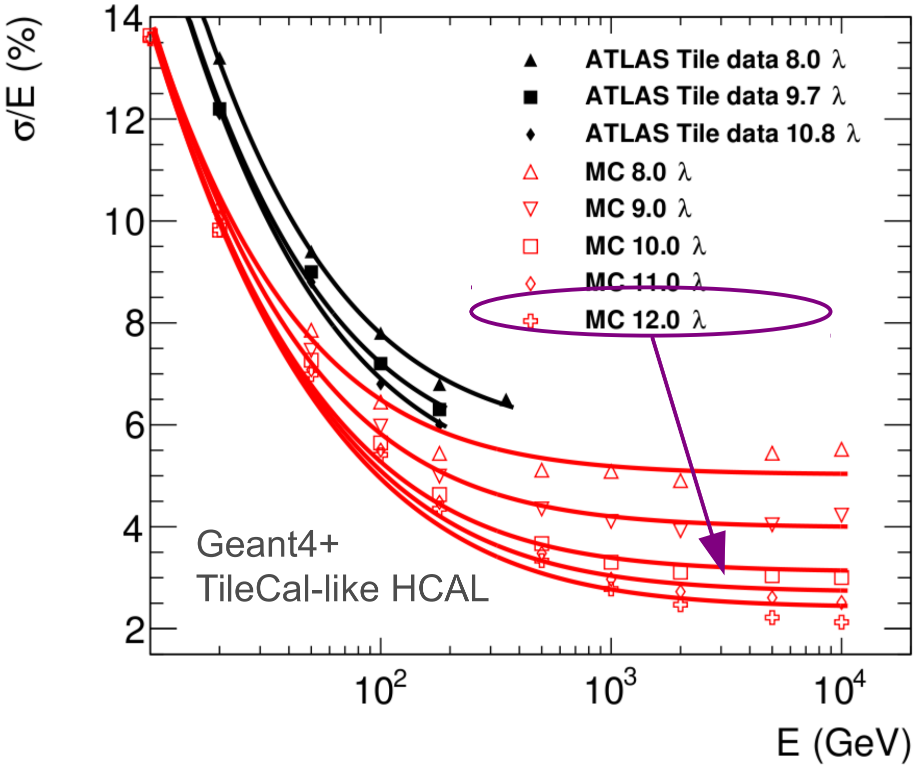
ATLAS-like setup
based on Geant4



$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

a – stochastic/sampling term,
b - electronic noise term
c - constant term

"c" dominates for jet with pT>5 TeV



- Geant4 TileCal inspired simulation based on FTFP_BERT
- Calculate single-particle resolution
- Stochastic term is close to 45%/√E
- Constant term improves by ~20% with increase of **1**$\boldsymbol{\lambda_I}$

Constant term c~2.5% is achievable for 12 $\boldsymbol{\lambda_I}$

T.Carli, C.Helsens, A.Henriques Correia, C.Solans: arXiv:1604.01415