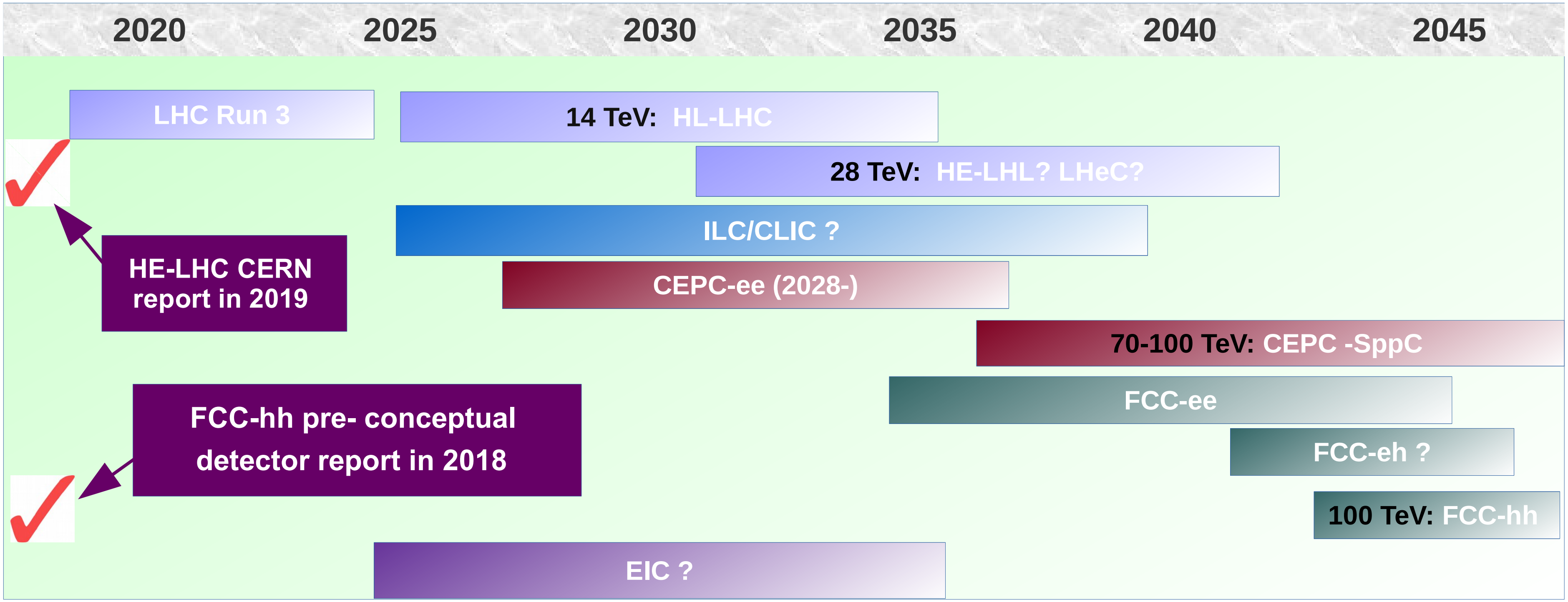
Monte Carlo simulations for future collider studies

S. Chekanov
+ contributors (www link)

HEP/ANL

Oct. 4, 2016

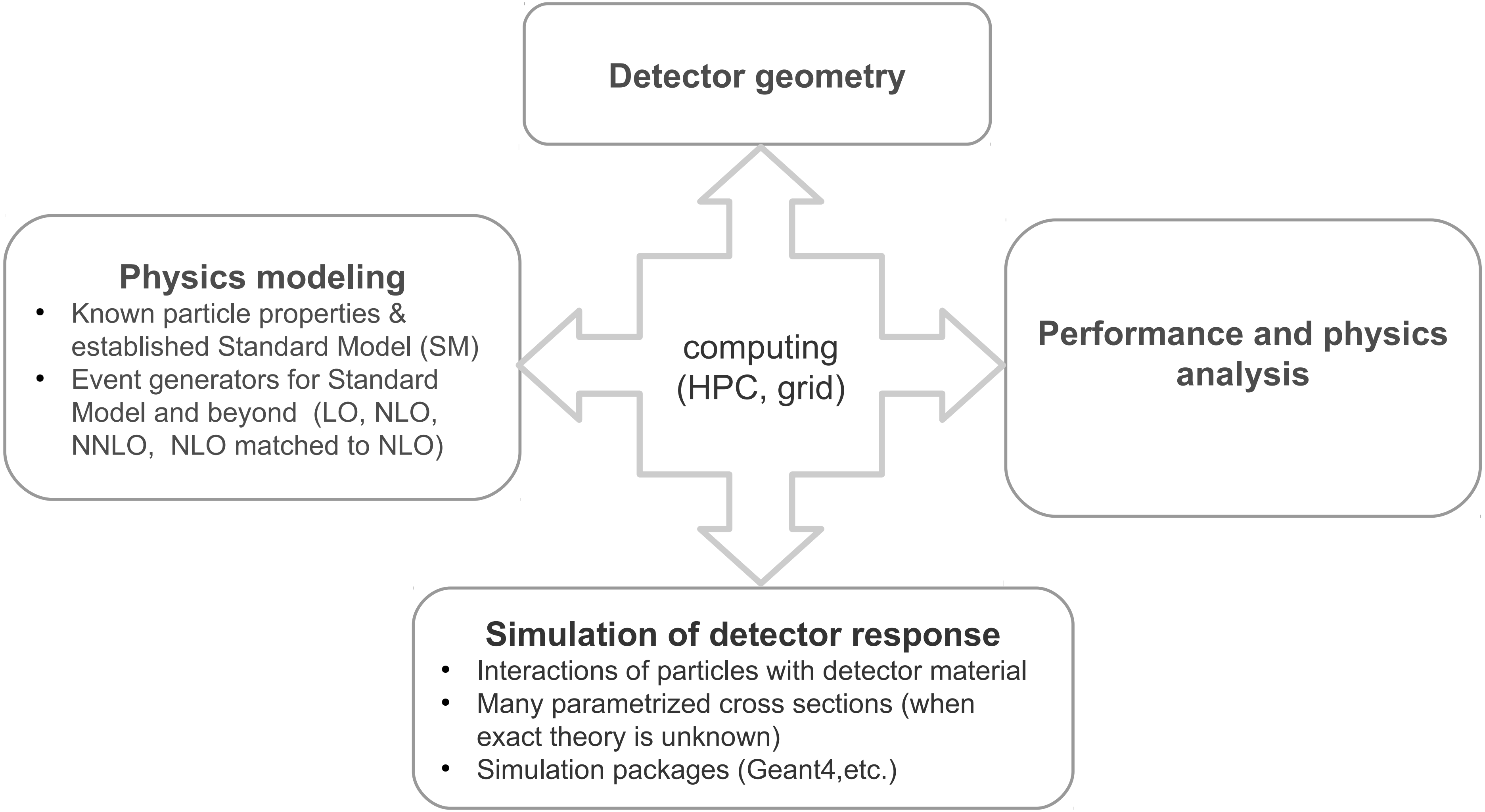
Timeline of particle collision experiments



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

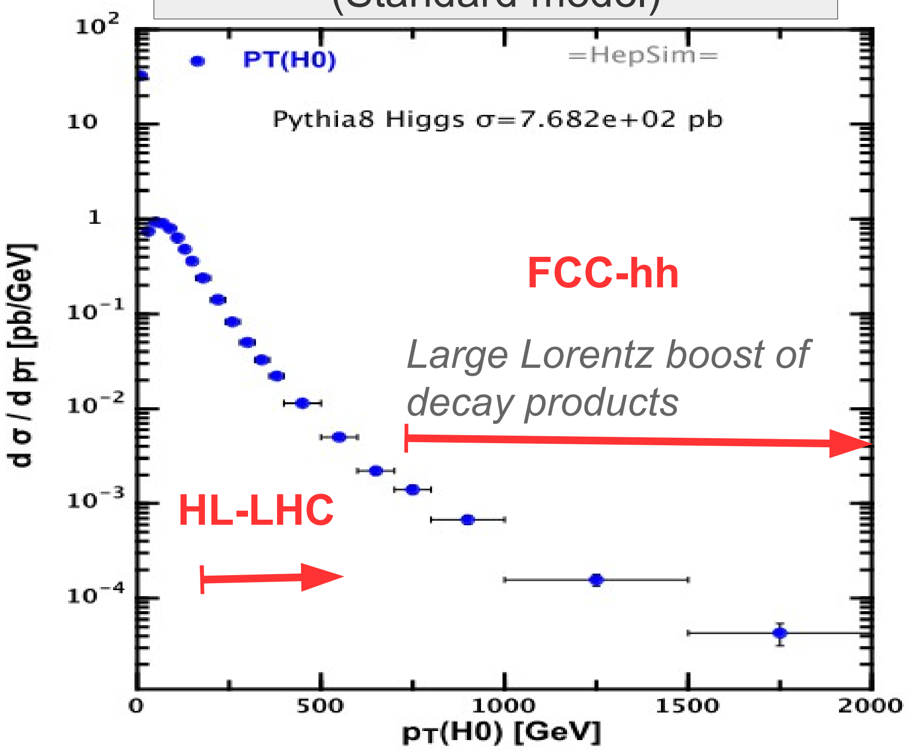
Simulations for particle-collision experiments



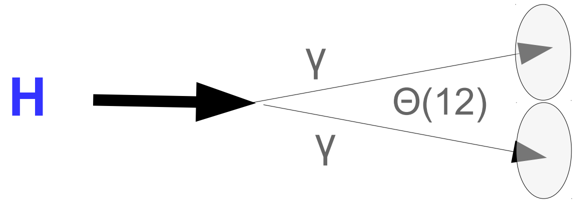
Why do we need simulations? Higgs example

- 100 TeV collider will hunt for M~30 TeV particles decaying to Higgs/W/Z bosons
- Completely new kinematic regime → very challenging for detector designs
- The detector must be optimized to reconstruct Higgs with pT>1 TeV

100 TeV pp
Differential Higgs cross section
(Standard model)



SM predictions: ~100,000 Higgs / ab^{-1} for pT>1 TeV

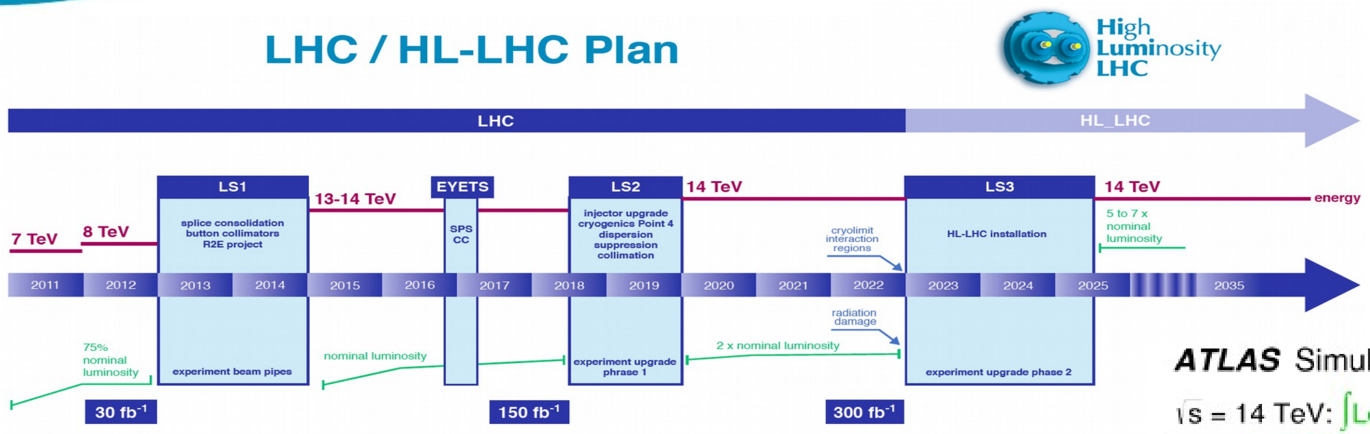


Just kinematics:
pT(H)>2 TeV → **~ 5 deg** between γ's
pT(H)>10 TeV → **~ 1 deg** between γ's

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

The HL-LHC Project

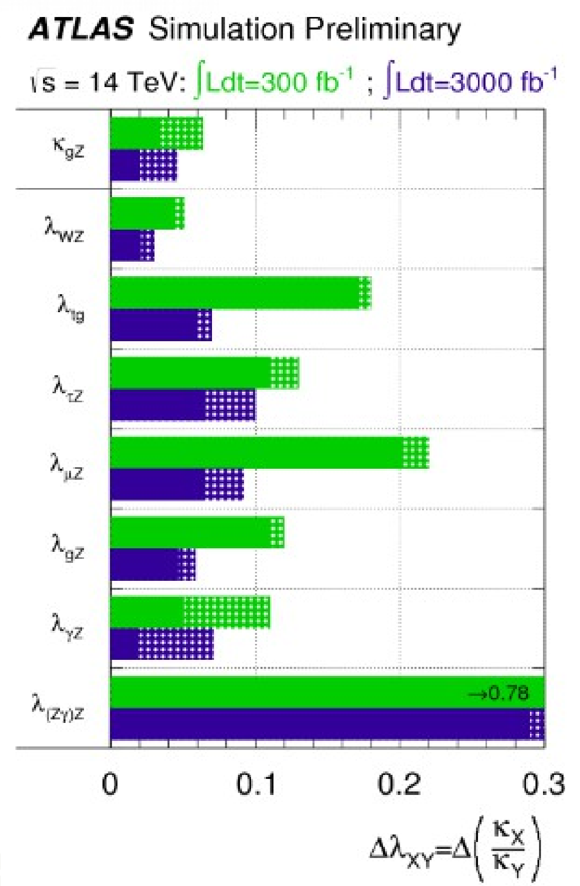
High Luminosity Large Hadron Collider



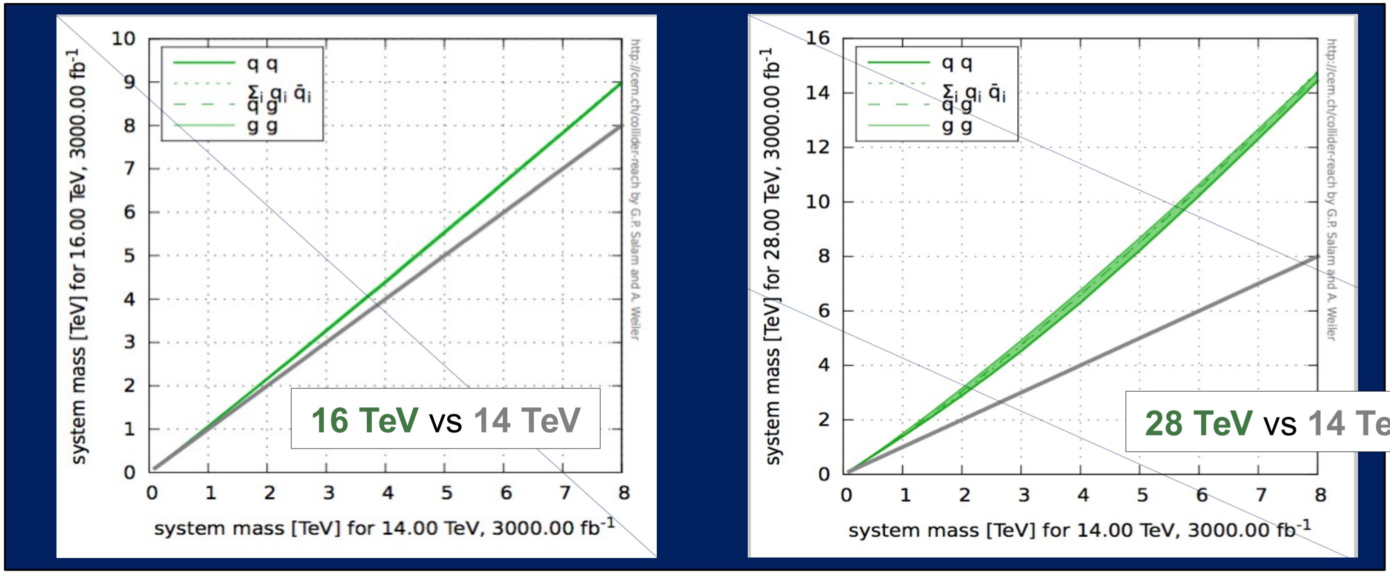
increase luminosity (rate of collisions) by a factor of 10 beyond the original design value of the LHC
(from 300 to 3000 fb^{-1})

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)



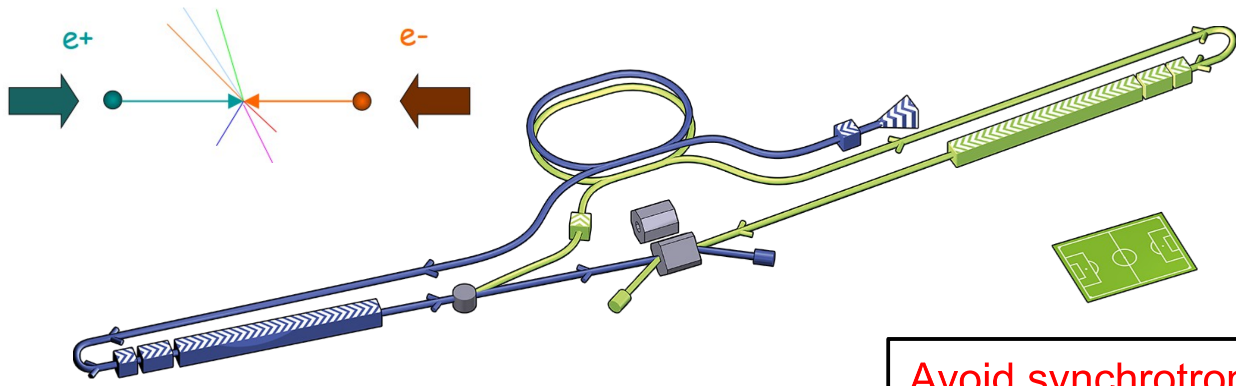
F.Gianotti
+ CERN management

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 ($\sqrt{s}$~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

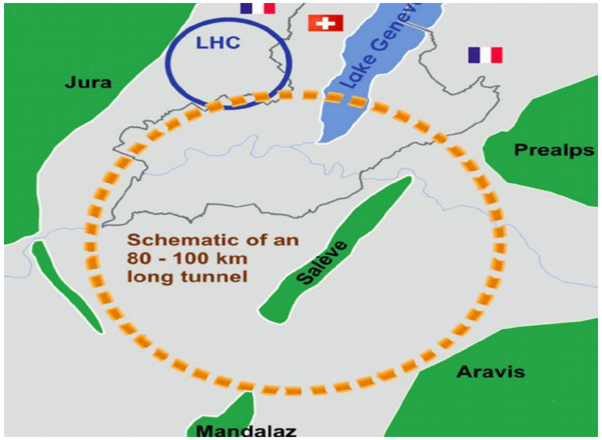
Advantages over proton-proton collisions:
simple initial state (e+e-), variable energies, momentum conservation, democratic production of particles

- 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)

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

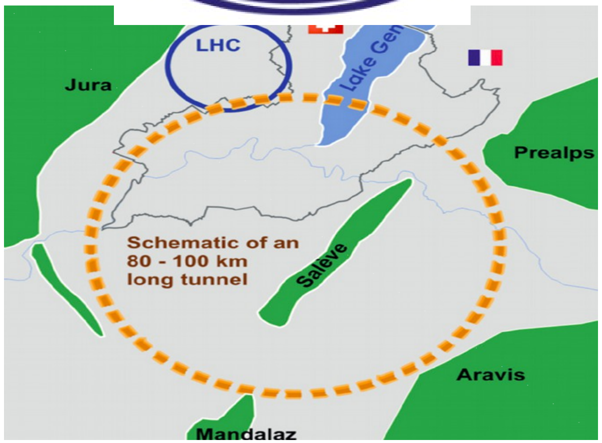


Circular Electron Positron Collider (CEPC)

- 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 <=$30\times10^{34}cm^{-2}s^{-1}$, 25(5) ns, pileup 1020(204)
- Many challenges for the detector!
- Conceptual Design Report (CDR) by 2018

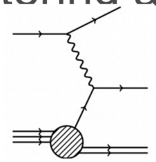
Super Proton-Proton Collider (SppC)

- 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**

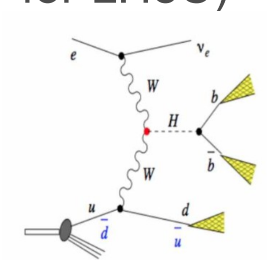
Energy frontier + intensity frontier: LHeC, FCC-ep (CERN), Electron-ion collider (EIC)

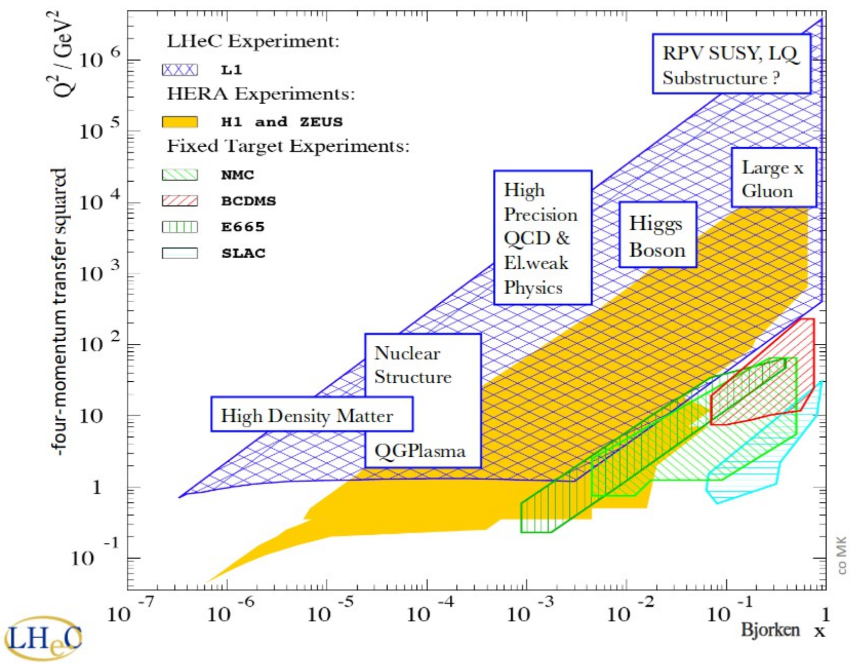
- **LHeC:** 7 GeV proton collided with 20-60 GeV electron ($s^{1/2=}$ = 1.3 TeV)
- **HE-LHC:** 15 TeV proton collided with 60 GeV electrons ($s^{1/2=}$ = 1.9 TeV)
- **FCC-ep:** 50 GeV proton collided with > 20 GeV electron ($s^{1/2=}$ = 3.5 TeV)
- **EIC** – electron-ion collider JLab/BNL: low energy 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 W
→ H (~200 fb^{-1} for LHeC)

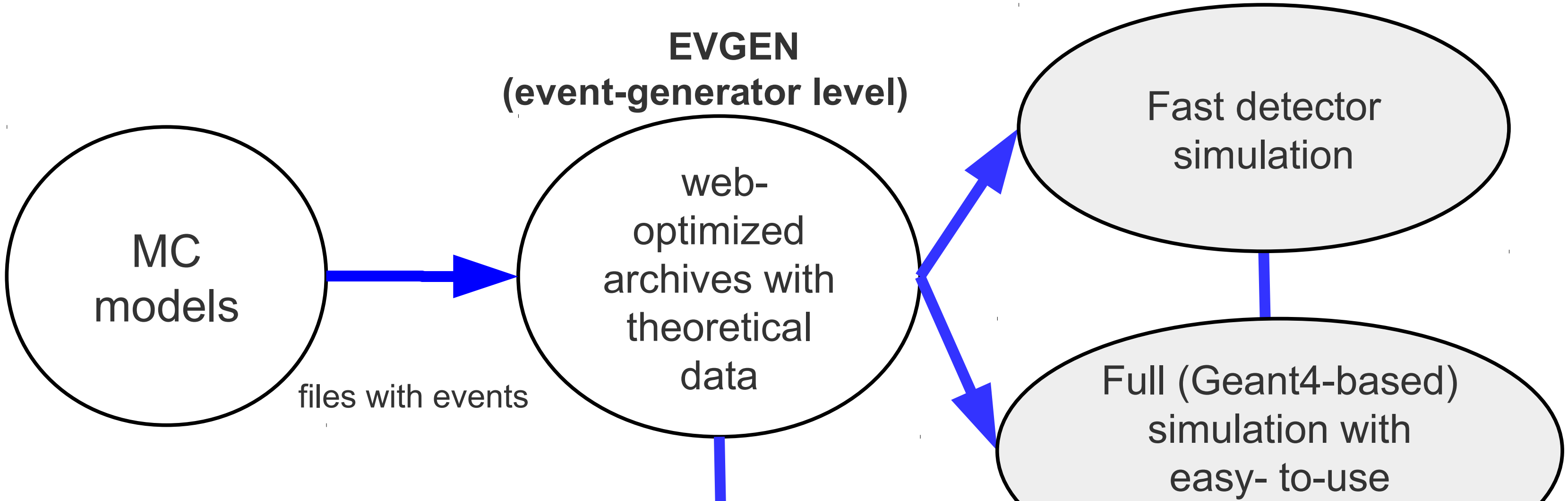




→ Studies of gluon density at large x

Snowmass 2013 and beyond

- **2013:** First large-scale computation for HEP community using Open-Science grid described in "Snowmass Energy Frontier Simulations" (arXiv:1309.1057)
- **2014:** First version of the HepSim repository at ANL for long-term preservation + extension to 100 TeV
- **2015:** Includes "reconstruction tags" with Geant4 simulations



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

S.C. "HepSim: a repository with predictions for high-energy physics experiments", Advances in High Energy Physics, vol. 2015, ID136093, (2015)

OPEN ACCESS **http://**

Long-term availability & preservation

Long-term preservation of theoretical calculations

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

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

ε ~ 0.01-1	- for LO MC
ε << 0.01	- for NLO etc.

- **ε << 1:**
  - Madgraph5 etc. (NLO+PS+hadronisation), ALPGEN
  - Some fast-converging NLO calculations (MCFM, jetPHOX etc)
  - MC with ε ~ 1 but after mixing with pile-up (CPU intensive)
- **ε ~ 1:** Less appropriate approach for:
  - LO simulations (Pythia6, Pythia8)
  - Some NLO programs with slow convergence
    - requires too large data volumes to keep weighted events

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

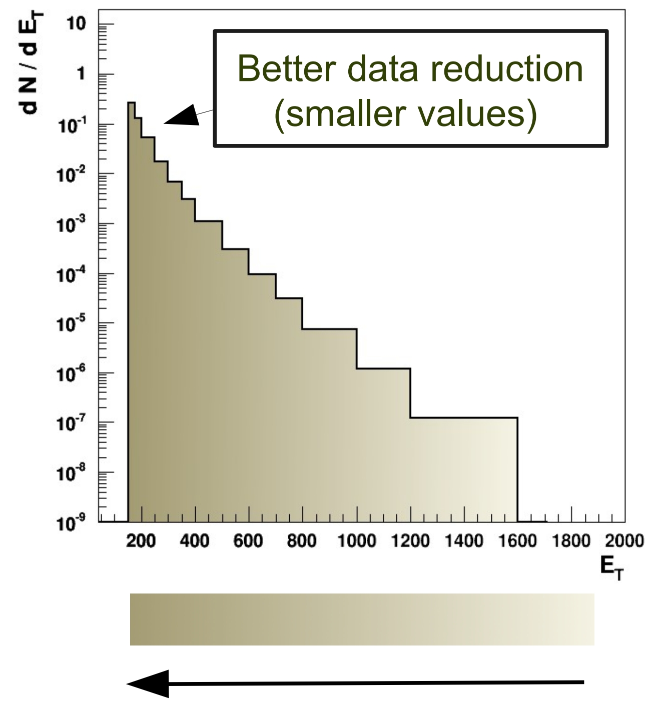
Number of used bytes depends on values.
Small values use small number of bytes

Google's Protocol buffers protobuf
Protocol Buffers - Google's data interchange format

- Effective file-size reduction for pile-up events
  - Particles with small momenta → small nr of bytes used
- Installed on HPC (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/

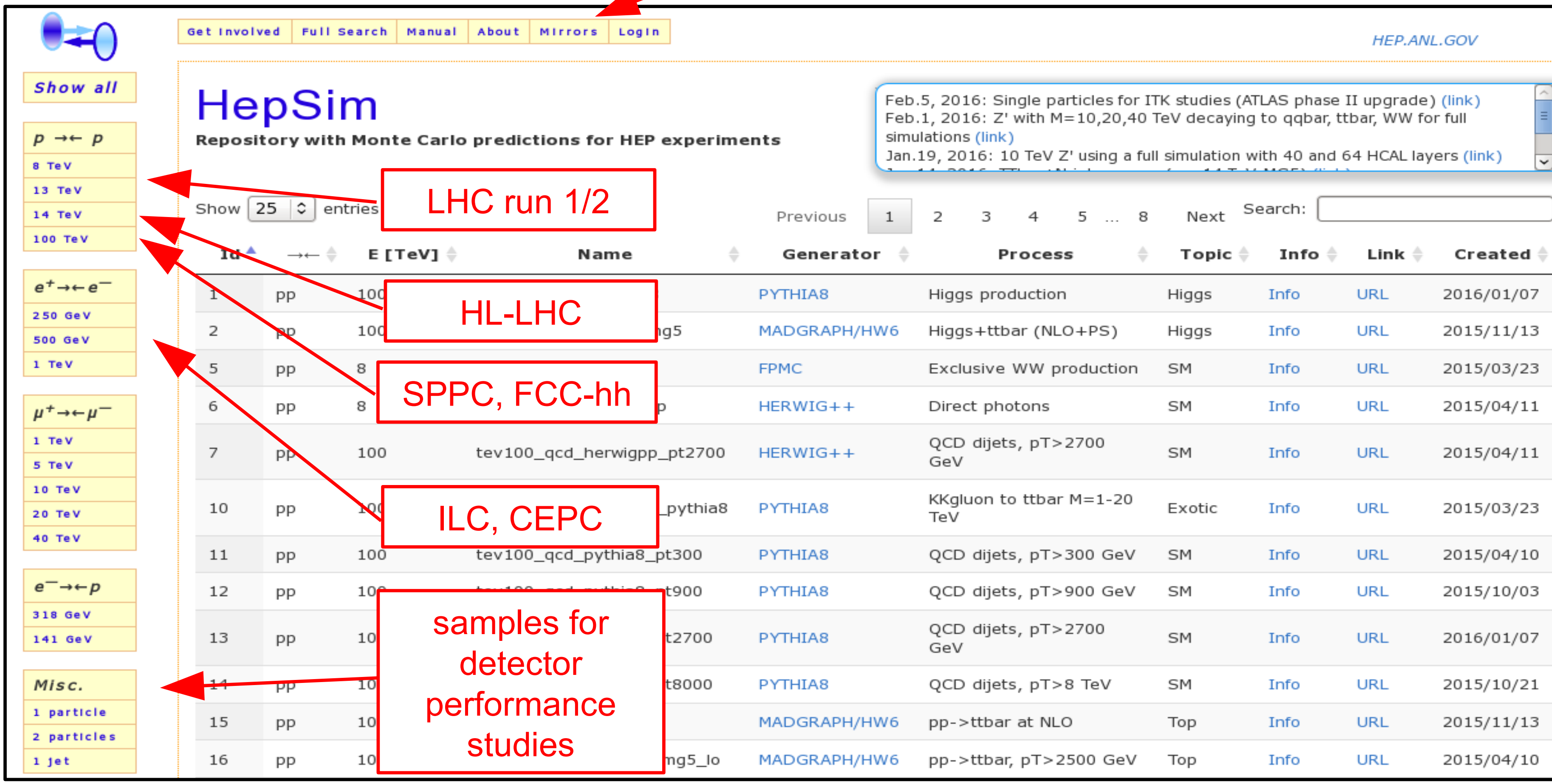
8-bytes → varint



compression strength keeping precision of representation constant

HepSim event simulations

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



Available: EVGEN files (LO,NLO, etc), fast simulations, full Geant4 simulations

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

$p \to\leftarrow p$
8 TeV
13 TeV
14 TeV
100 TeV

$e^+ \to\leftarrow e^-$
250 GeV
500 GeV
1 TeV

$\mu^+ \to\leftarrow \mu^-$
1 TeV
5 TeV
10 TeV
20 TeV
40 TeV

$e^- \to\leftarrow p$
318 GeV
141 GeV

Misc.
1 particle
2 particles
1 jet

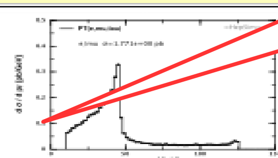
Repository with Monte Carlo predictions for HEP experiments

Sep.15, 2016. Z'(5 TeV) to diffe
Aug.28, 2016: rfull010, rfull01

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
Download URL:	http://mc.hep.anl.gov/asc/hepsim/events/ee/250gev/pythia6_zpole_ee
Status:	Available
Mirrors:	http://portal.nersc.gov/project/m1758/data/events/ee/250gev/pythia6_zpole_ee
EVGEN size:	0.826 GB
Fast simulation:	rfast001 (info) 100 / 1.75 GB 10/13/2015
Full simulation:	rfull002 (info) 16 / 1.27 GB 11/20/2015; rfull001 (info) 15 / 1.18 GB 10/17/2015
Fast/Full size:	4.20 GB
Record slimmed:	No
Events weighted:	No
Submission time:	Tue Oct 13 14:28:55 CDT 2015
Updated on:	
User description:	PYTHIA version 6.4. Z production (Zpole) with decays to e+e-. Other de
File metadata:	**ProMC version:** 4; **Nr events:** 1000; **Varint E:** 1000000; **Varint L:** logfile.txt; **Last modified:** 2015-10-15 20:31:08; **Settings:** PYTHIA-6 mix events; NTOT 0 0 1000 ! Number of events; ECME 0 0 250.0 ! CM e 0 0 839264 ! random seed; MSEL 0 0 0 ! all mixed events; PMAS 6 1 17 91.1876 ! Z boson mass; PMAS 24 1 80.3850 ! W boson mass; PMAS 25 mass; MSUB 1 0 1 ! ffbar to Z; MSTP 43 0 2 ! Z only, no gamma; MDME MDME 175 1 0 ! U U~; MDME 176 1 0 ! S S~; MDME 177 1 0 ! C C~; M B~; MDME 179 1 0 ! T T~; MDME 182 1 1 ! E- E+; MDME 183 1 0 ! NU_E 1 0 ! MU+ MU-; MDME 185 1 0 ! NU_MU+ NU_MU-; MDME 186 1 0 ! TAU- TA 1 0 ! NU_TAU- NU_TAU+; PARJ 71 0 10 ! ctau=10mm; MSTJ 22 0 2 !;

Validation:

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

URL for EVGEN files (download or streaming)

Reconstruction tags 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

Searching for reconstruction tags

Reconstruction tags include fast (Delphes) and full (SLIC/Geant4) datasets for various detector configurations

Example: looking for the tag **rfast005** (Delphes, official FCC detector, v5)

$e^+ \rightarrow\leftarrow e^-$
- 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

	Dataset Name	Generator	EVGEN	Fast simulation	Full simulation
1	tev100_higgs_bbar_pythia8	PYTHIA8	URL	**rfast005** (info) rfast003 (info) rfast002 (info) rfast001 (info)	
2	tev100_higgs_pythia8	PYTHIA8	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
3	tev100_higgs_ttbar_mg5	MADGRAPH/HW6	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
4	tev100_mg5_2HDMexovv	MADGRAPH/PY6	URL	**rfast005** (info)	rfull001 (info)
5	tev100_mg5_ttbar_bjet	MADGRAPH/PY6	URL	**rfast005** (info) rfast002 (info)	
6	tev100_mg5_ttbar_jet	MADGRAPH/HW6	URL	**rfast005** (info) rfast002 (info)	
7	tev100_minbias_a2_pythia8	PYTHIA8	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
8	tev100_minbias_a2_pythia8_l3	PYTHIA8	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
9	tev100_minbias_a2_pythia8_nosl	PYTHIA8	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
10	tev100_pythia6_higgs_zz_4l	PYTHIA6	URL	**rfast005** (info) rfast002 (info)	rfull009 (info) rfull008 (info) rfull006 (info) rfull001 (info)
11	tev100_pythia8_allh2	PYTHIA8	URL	**rfast005** (info) rfast002 (info) rfast001 (info)	
12	tev100_pythia8_higgs_bbar	PYTHIA8	URL	**rfast005** (info) rfast003 (info) rfast002 (info) rfast001 (info)	
13	tev100_pythia8_higgs_zz_4l	PYTHIA8	URL	**rfast005** (info) rfast002 (info)	

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

fast → Delphes fast simulation (ROOT files)

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

ProMC files on several public web servers

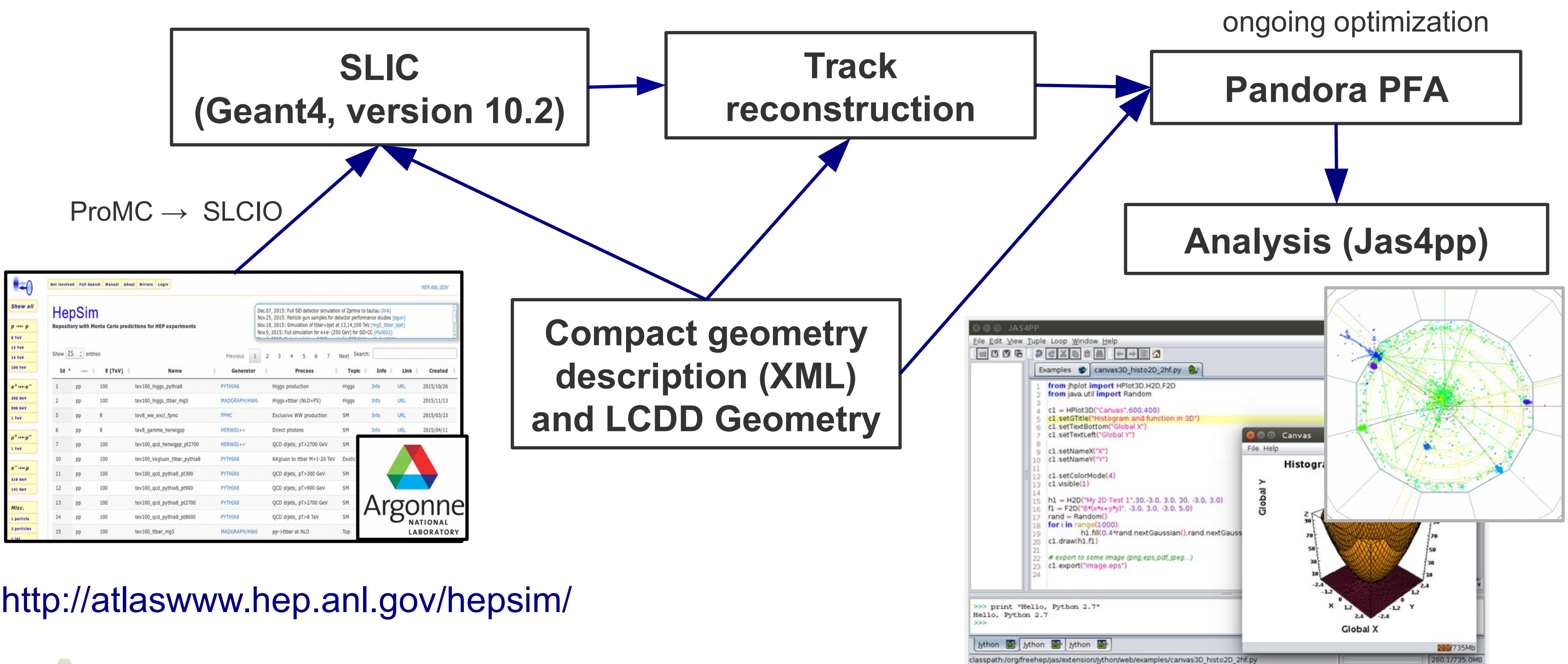
HepSim

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

HepSim software for full simulations

Simulator for the Linear Collider (SLIC) software:

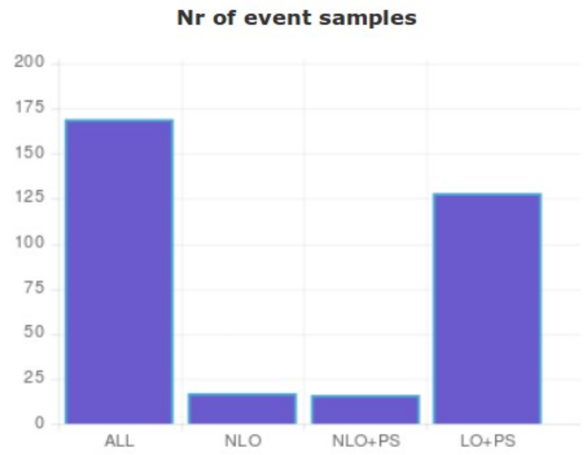
- Optimized for the SiD detector at SLAC (T.Johnson, N.Graf, J.McCormick, J.Strube)
- Re-purposed for future pp collider studies (S.C., A.Kotwal, J.Strube, J.Marshall)
  - *Geant4 10.2, implemented Fast PandoraPFA*
  - *Integrated with HepSim ProMC EVGEN files*
  - *Deployed on Open-Science Grid (OSG)*
  - *Analysis: C++/Root or **Jas4pp** (ANL,S.C,E.May). Based on Jas3 (SLAC)*



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

HepSim event statistics

(excluding fast and Geant4 detector simulations)

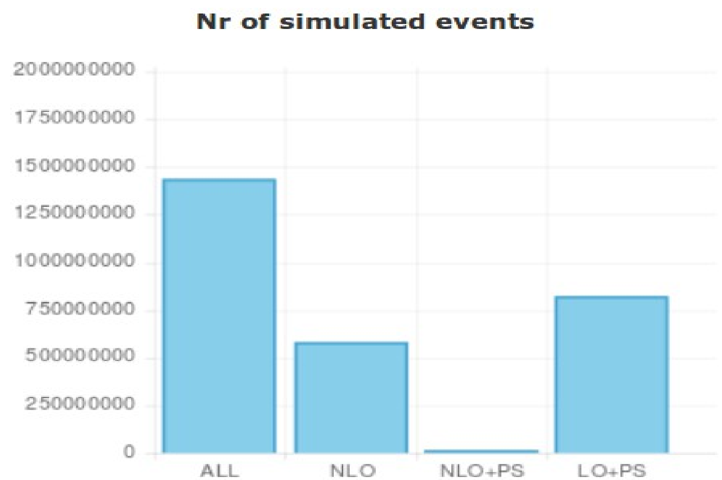


Number of public file servers	3
Number of event samples	208
Number of NLO samples	17
Number of NLO+PS samples	17
Number of LO (+PS) samples	144
Number of events	1560741507
NLO events	583000000
NLO+PS events	32860595
LO (+PS) events	859498212
Total size (GB)	6897.468
NLO size (GB)	238.06
NLO+PS size (GB)	348.693
LO (+PS) size (GB)	6292.482
Number of files	334606

~210 Monte Carlo samples

~1.6 billion EVGEN events

- ~ 10% after fast simulations(Delphes)
- ~ 0.1% after Geant4 simulations



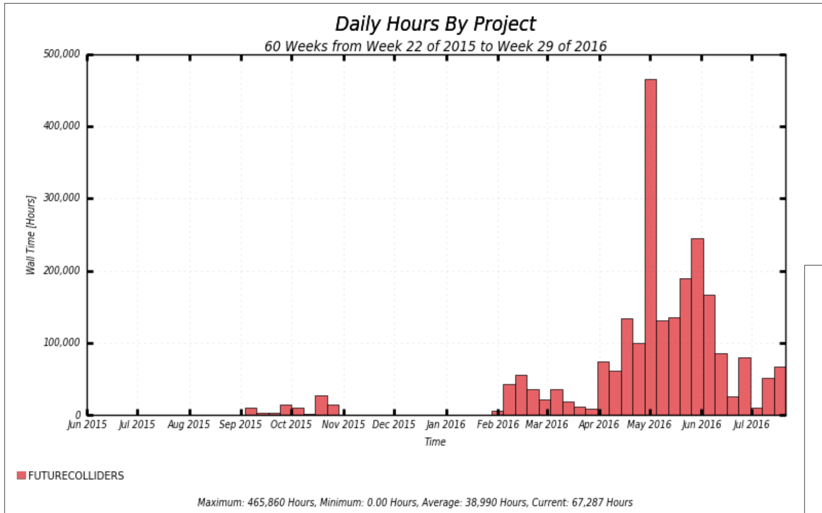
Platforms for event generations (EVGEN)

- 10% → BlueGene/Q (ANL/Mira) (Jetphox, MCFM)
- 50% → HEP-ANL (mainly Madgraph)
- 40% → OSG-CI grid and USATLAS CI (for phase II)

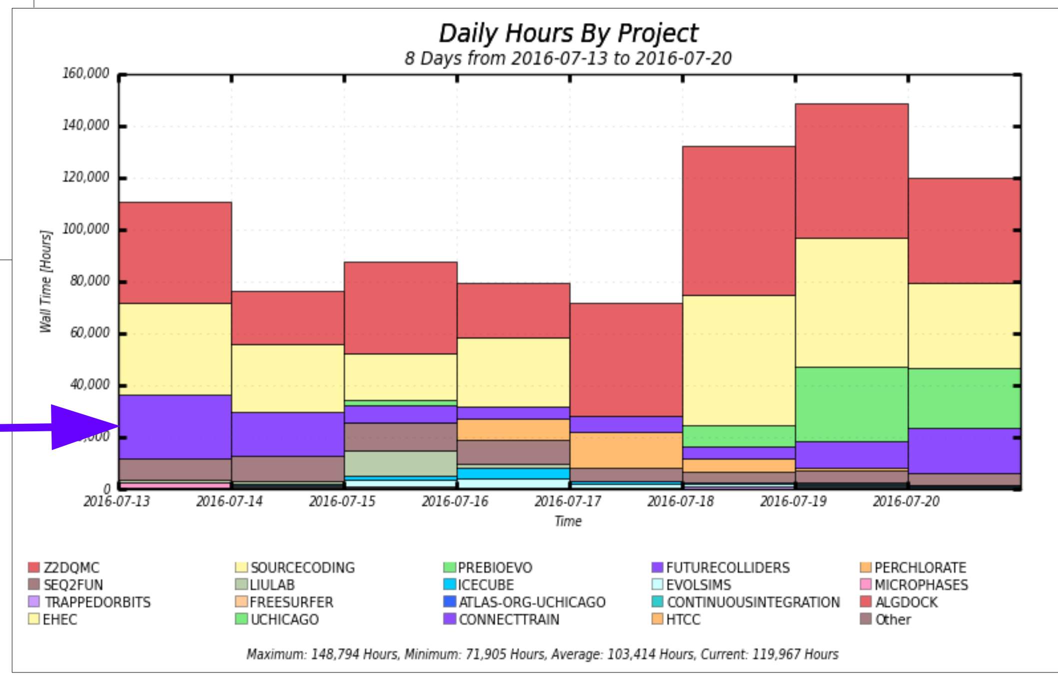
CPU usage for SLIC (Geant4) simulations

osg connect

OSG-Connect "FutureColliders" project for HepSim jobs



CPU*h OSG-Connect



HepSim simulations (SLIC/Geant4) ~ 16 GB RAM per job

2.5 million CPU*h in 2016 using OSG-grid for Geant4 simulations

(equivalent to ~10 million CPU*h on HPC BlueGene/Q Mira)

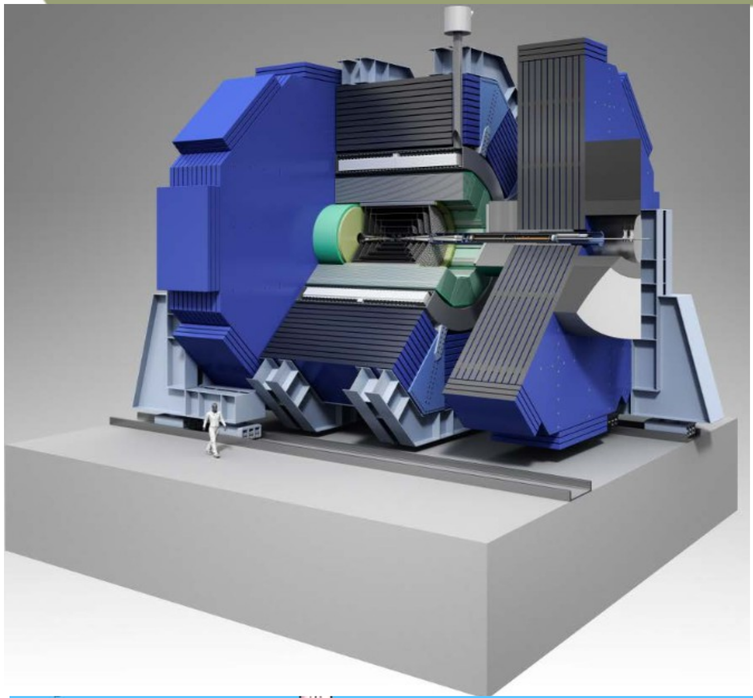
ANL High Performance Computing (HPC)

- ANL is home of Mira, an IBM Blue Gene/Q supercomputer at the Argonne Leadership Computing Facility supported by DOE, is equipped with 786,432 cores for open science projects
- Can run MCFM, Jetphox, Pythia6/8 with basic HEP libraries
  - a few HepSim samples were created on Mira
- Current main activity - ALPGEN Monte Carlo samples (V+jets) for the ATLAS (LHC)
  - J.Childers, T.Uram, T.LeCompte, M.Papka, D.Benjamin http://www.sciencedirect.com/science/article/pii/S0010465516302843
- New: AURORA $\rightarrow$ 2019 target date for user access

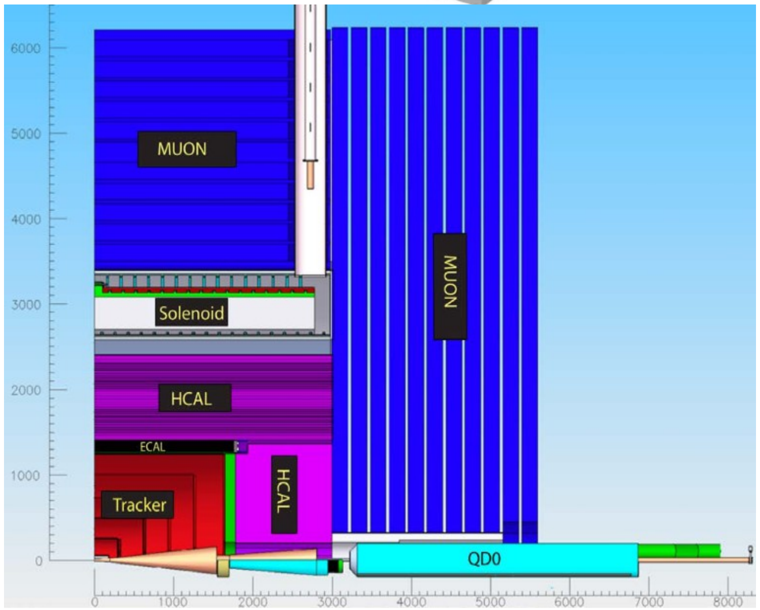
System Features	Mira	Aurora
Compute Nodes	49,152	>50,000
Processor	PowerPC A2 1600 MHz	3rd Generation Intel Xeon Phi
System Memory	768 TB	>7 PB DRAM and persistent memory
System Interconnect	IBM 5D torus interconnect with VCSEL photonics	2nd Generation Intel Omni-Path Architecture with silicon photonics
File System Capacity	26 PB GPFS	>150 PB Lustre
Intel Architecture (x86-64) Compatibility	No	Yes
Peak Power Consumption	4.8 MW	13 MW

http://aurora.alcf.anl.gov/

SiD detector for ILC



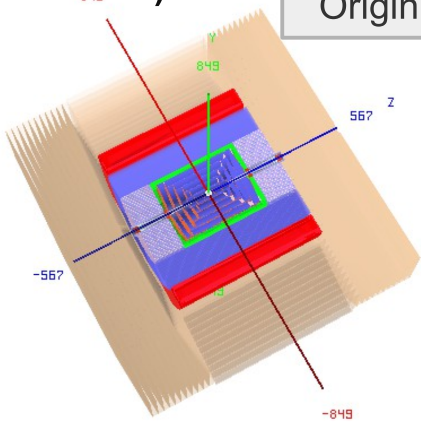
- Multi-purpose detector for the ILC
- Conceived at SLAC (USA LC Physics Group)
- The key characteristics:
  - 5 Tesla solenoid
  - Silicon tracker: 25/50 um readout pitch
  - ECAL: (0.35 cm cell size, W / silicon)
  - HCAL:
    - 1x1 cm cell size (RPC for LOI3*)
    - 40 layers for barrel (HCAL) ~4.5 λ_I
- Optimized for particle-flow algorithms (PFA)
- Fully configurable using SLIC software



‘All-silicon’ design concepts supported in HepSim

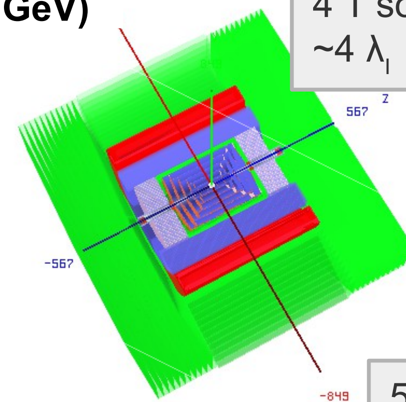
SiD
($e^+ e^-$ up to 1 TeV)

Original SiDLO3

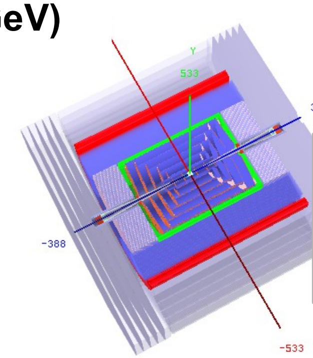


SiCPEC
($e^+ e^-$ 250 GeV)

4 T solenoid
~4 λ_I digital HCAL



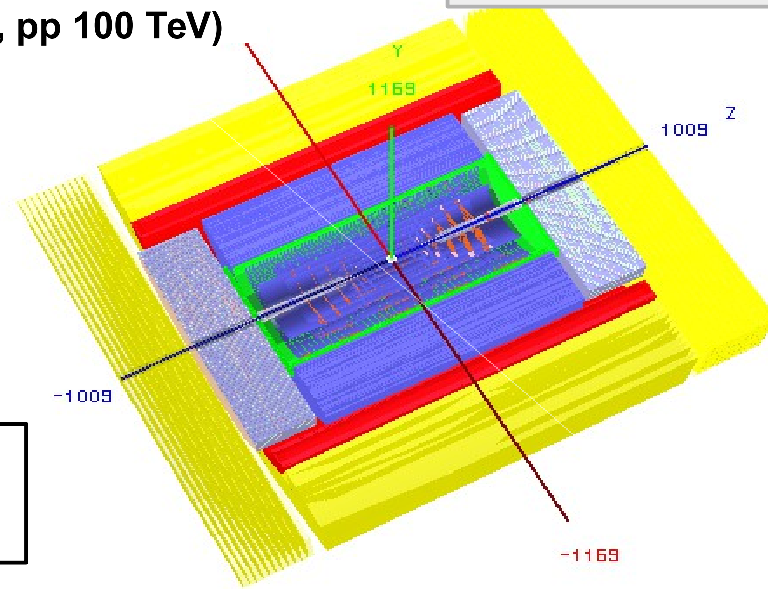
SiEIC
(ep, 141 GeV)



1 T solenoid
~3 λ_I HCAL
Smaller size

SiFCC
(FCC-hh, pp 100 TeV)

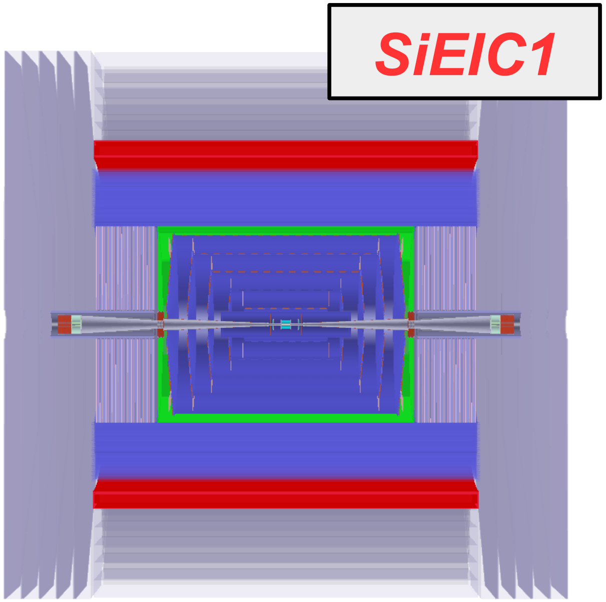
5 T solenoid
~12 λ_I Scint/Fe HCAL
Larger size

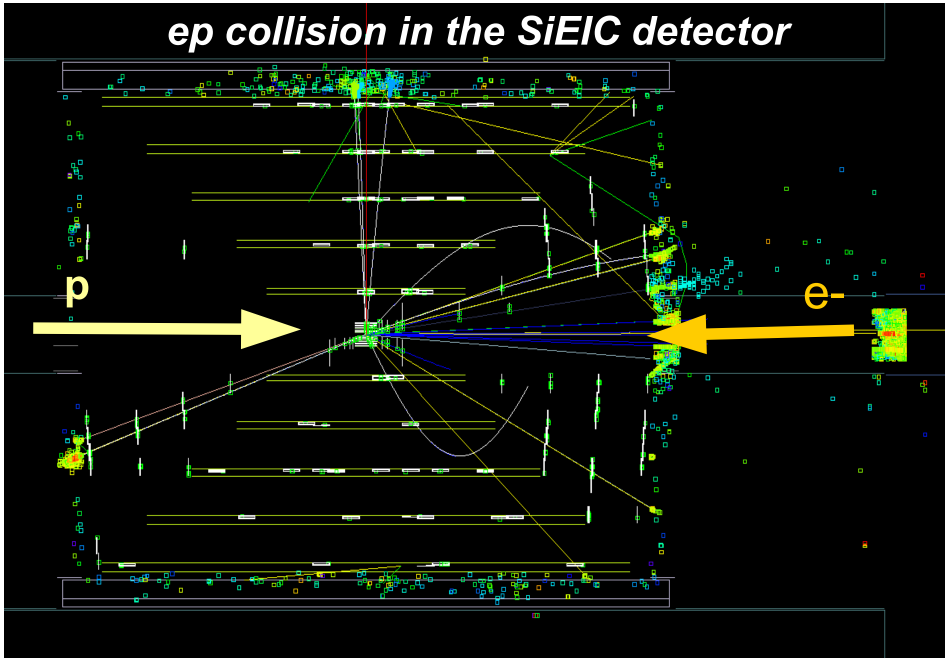


HepSim provides single-particle and physics events after simulation & reconstruction

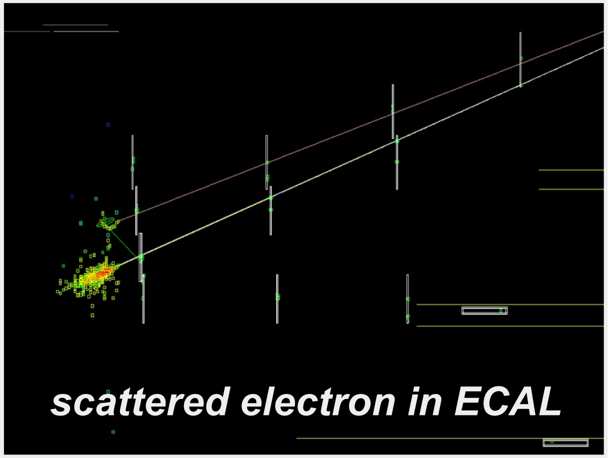
EIC collisions in the SiEIC detector

- Re-purpose SiD for the Electron-Ion Collider (EIC)
- Optimized SiD detector concept for EIC collisions:
  - smaller size, thiner CAL, 1 Tesla solenoid etc..





PFA electron energy: **16.92 GeV**
"EVGEN" truth energy: **16.90 GeV**



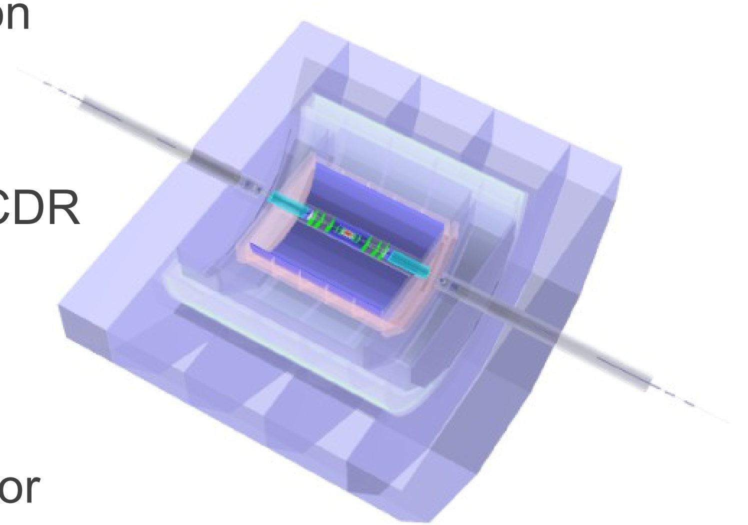
DIS sample (Q^2> 5 GeV^2) → "HEP" like (HERA)
CM energy = 141 GeV ("EIC-like")
Monte Carlo samples available from HapSim

CEPC detector studies

A CEPC detector based on the ILD detector concept

- ILD detector is the baseline of the CEPC simulation group at IHEP (Beijing)
  - M. Ruan, Y. Fang, G. Li, Q. Li, X. Moa etc.
- Ongoing optimization of the detector concept for CDR
- Ongoing Higgs studies using Pythia6 samples
  - see a presentation at ICHEP 2016, Chicago

- A possible second option based on the SiD detector conceived by the USA LC Physics Group?
  - Many similarities in the design choices
  - Similar ILCSOFT software: PFA, LCO format etc..



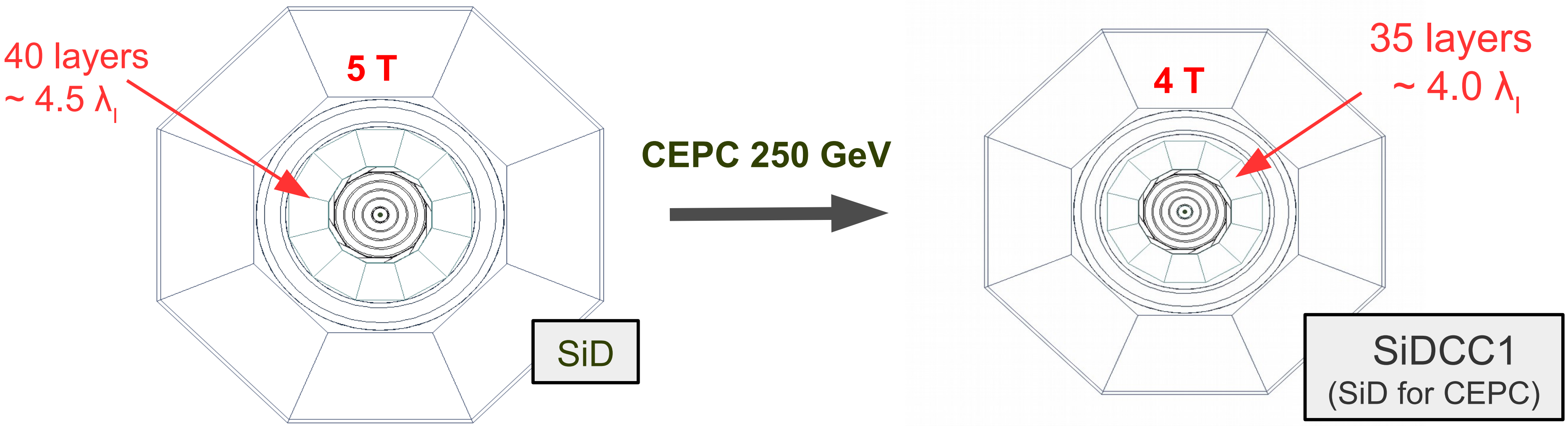
- *3.5 T solenoid*
- *Time Projection Chamber (TPC) for tracks*

Designing a detector for CEPC ($e^+ e^-$ CM E=240-250 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

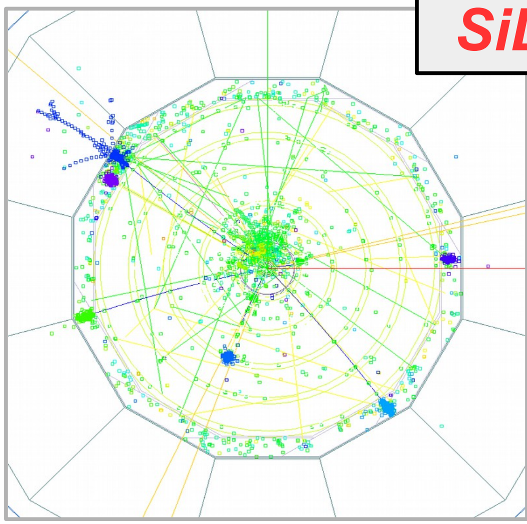


Design a light, cost-optimized version of the SiD detector for CEPC and use several physics processes to benchmark its performance

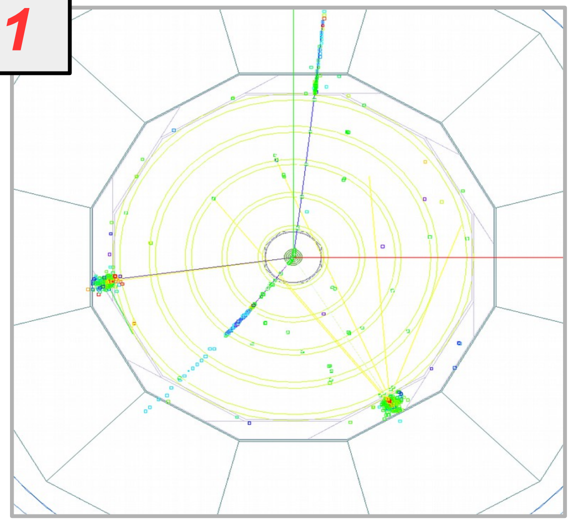
S.C. and M.Demarteau, Conceptual Design Studies for a CEPC Detector.
arXiv:1604.01994. HKUST IAS 2016 proceeding

Example II: Simulations for CEPC (e^+e^- 250 GeV CM energy)

SiDcc1



H(125) → 4 e



H(125) → γγ

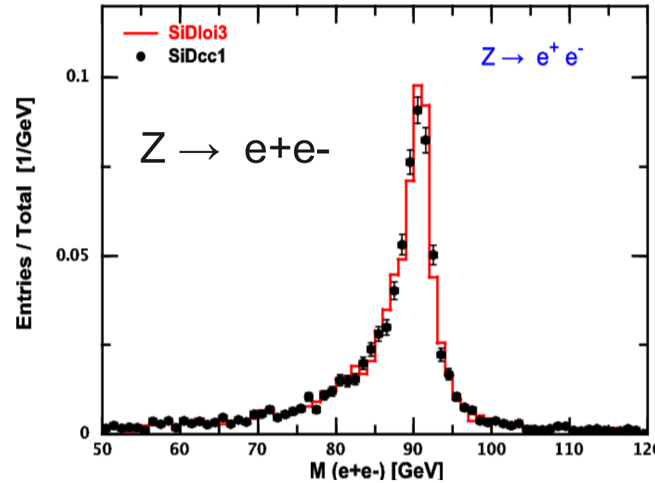
~ 10,000 reconstructed events for each physics channel after PFA (Pandora) reconstruction

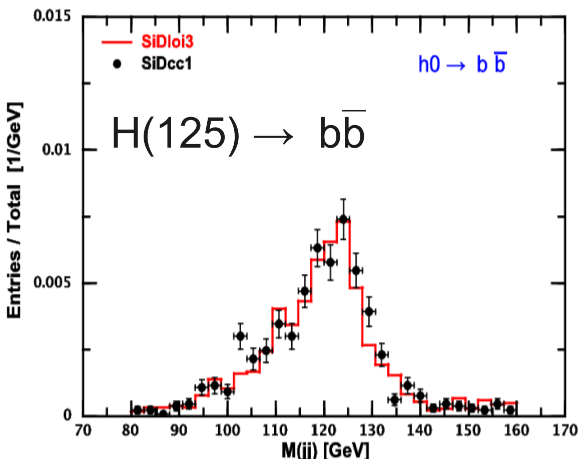
- **Available full simulations for the SiD and SiDCC (for CPC) detectors:**
  - Z → e+e-, Z → tau tau, Z → mu+mu-, Z → $b\bar{b}$
  - H(125) → $b\bar{b}$ H(125) → γγ, H(125) → ZZ* → 4l, H(125) → tau+tau-

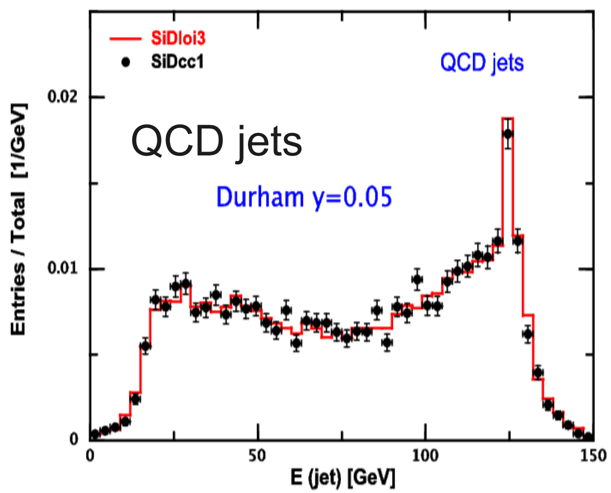
Comparing SiD with SiDCC1

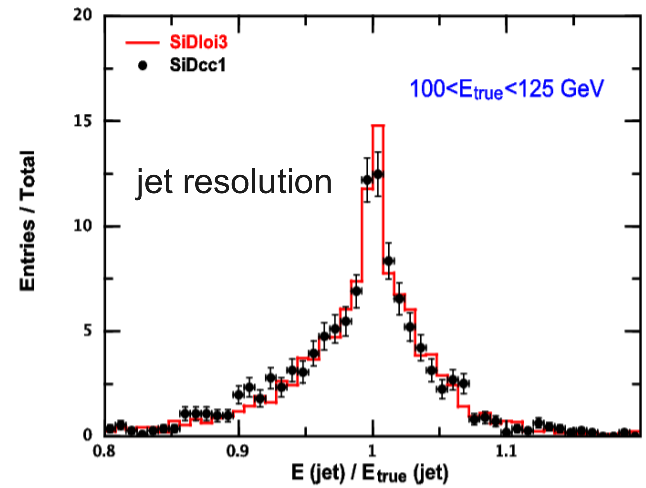
Done with Jas4pp

- 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
- Particle flow objects to reconstruct invariant masses and jet energy resolutions (Durham jets)







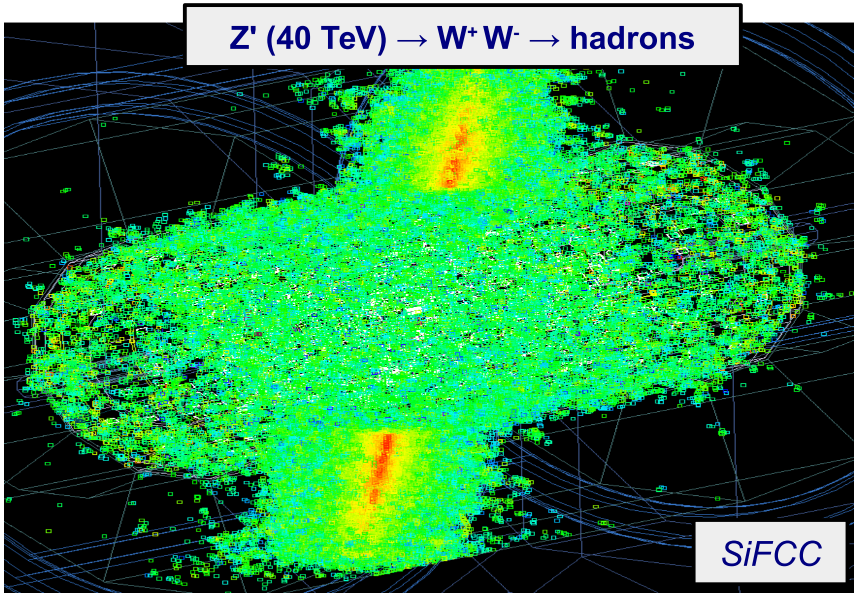


Simplification of the SiD detector does not compromise physics performance

S.C. and M.Demarteau. arXiv:1604.01994. HKUST IAS 2016 proceeding

High-granularity hadronic calorimeter for tens-TeV physics at FCC-hh, SppC and HE-LHC

With contributions from:
M.Beydler (ANL) 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)



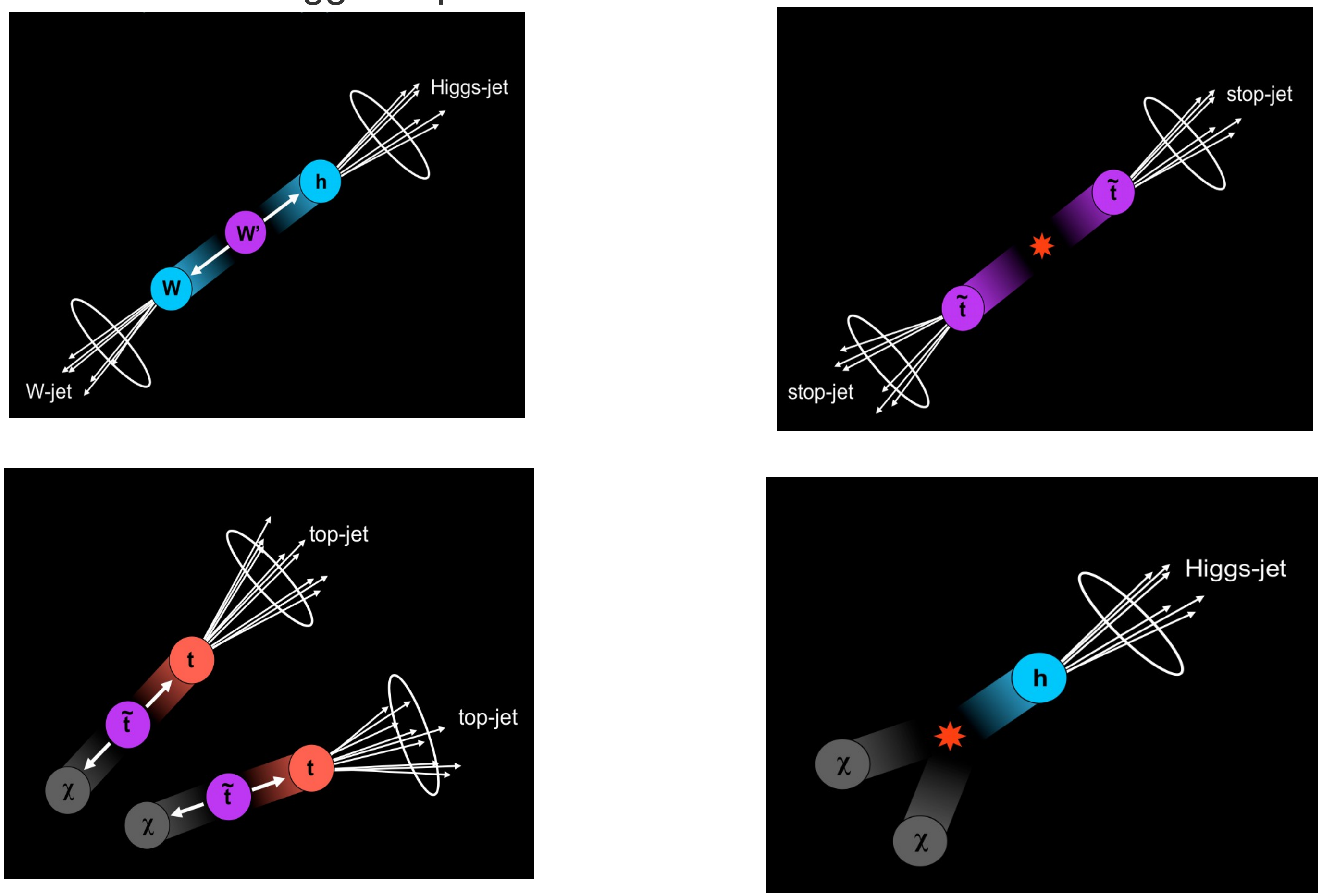
Two 20 TeV jets in ~12 λ_I calorimeter

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

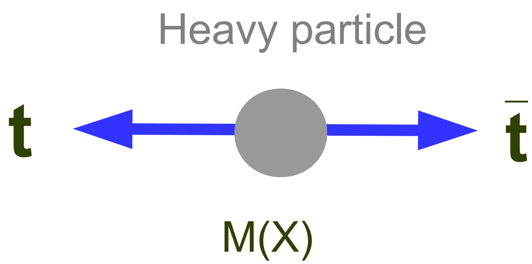


TeV-scale pair-produced

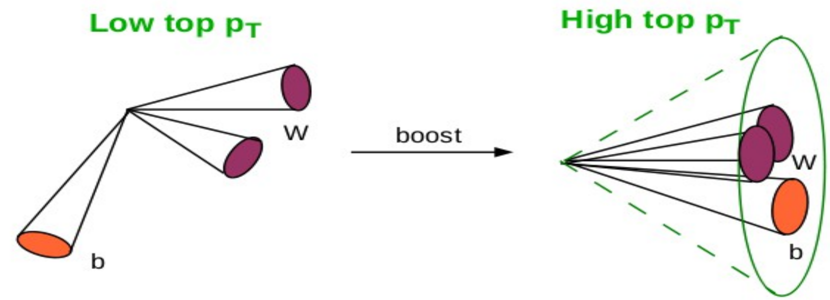
SM + dark matter

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

Boosted top from high-mass particles



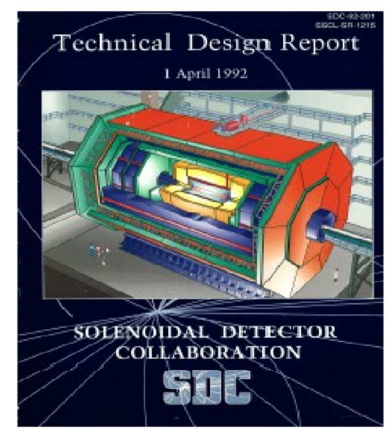
ΔR ~ 2 pT / m(top)



www.quantumdiaries.org

- 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



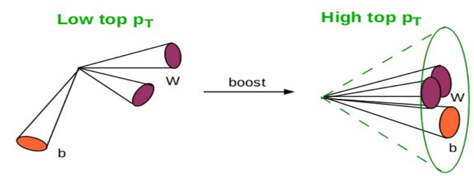
Hadronic calorimeter (HCAL) for next collider experiments

- Physics goals of future colliders - search for particles with masses 10-50 TeV that can decay to Higgs, W, Z, top decays

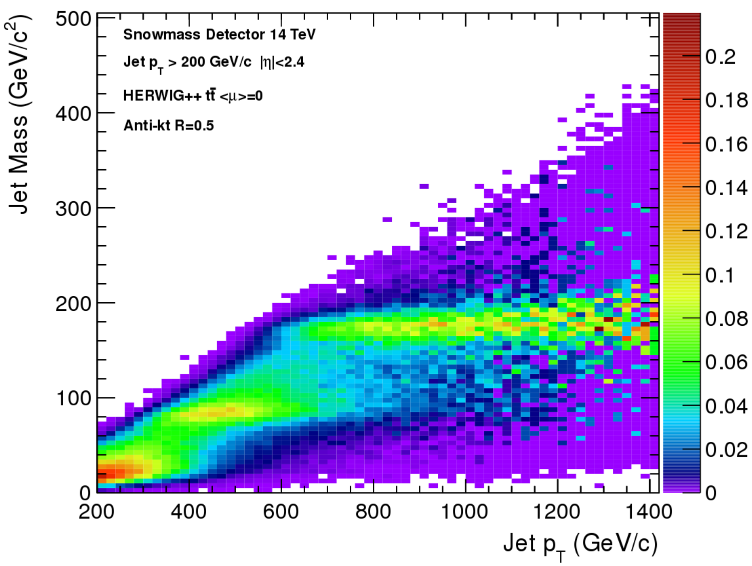
  → narrow jets with pT>5-25 TeV from Higgs, W, Z, top decays

- How to build a HCAL that can:
  - measure jet energy (up to 30 TeV)
  - resolve internal structure of narrow jets

Δ R(qq)

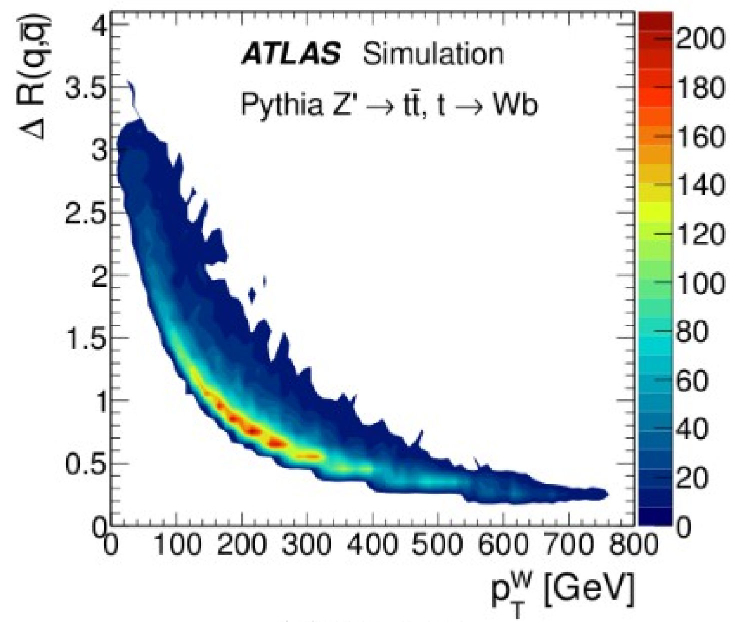


Mass of jets originating from boosted top quark



Typical HCAL cell size for ATLAS & CMS HCAL

distance in pseudorapidity-phi between two quarks from W



(b) $W \to q\bar{q}$

Detector requirements driven by physics at 100 TeV

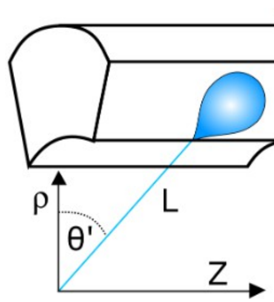
(what we already know)

- **Good containment up to pT(jet)~30 TeV: 12 λ_I for ECAL+HCAL** — done
  - affects jet energy resolution
  - leakage biases, etc.
- **Small constant term for HCAL energy resolution: $c < 3\%$** — done
  - dominates jet resolution for pT>5 TeV
  - important for heavy-mass particles decaying to jets
- **Longitudinal segmentation:** — to be done
  - Not studied → *require realistic Geant4 simulations*
- **Good transverse segmentation for resolving boosted particles:** — **this study**
  - baseline is $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$ from Delphes fast simulations
    - 5x5 cm assuming ~ATLAS-like inner radius (~2.3 m from IP)
  - *Require realistic Geant4 simulations*

See presentations given at the FCC week 2016 (Rome) and CALOR 2016

Resolution for single pions

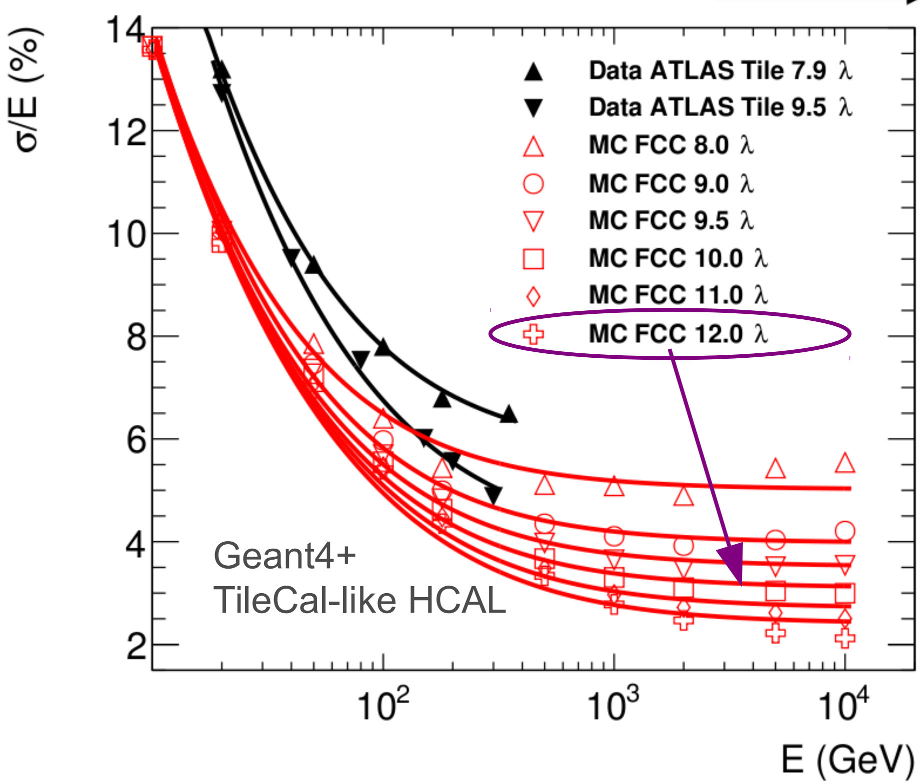
ATLAS-like setup
based on Geant4



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

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
- Stochastic term is close to 45%/√E
- Constant term improves by ~20% with increase of **1λ_I**

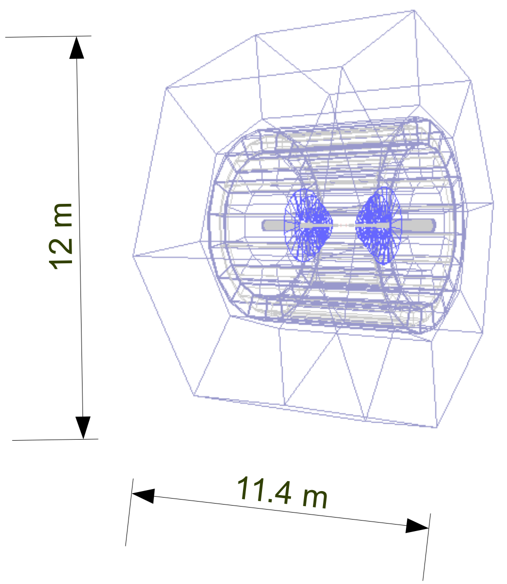
12 λ_I calorimeter:
- no leakage up to 10 TeV
- constant term c~2.5%

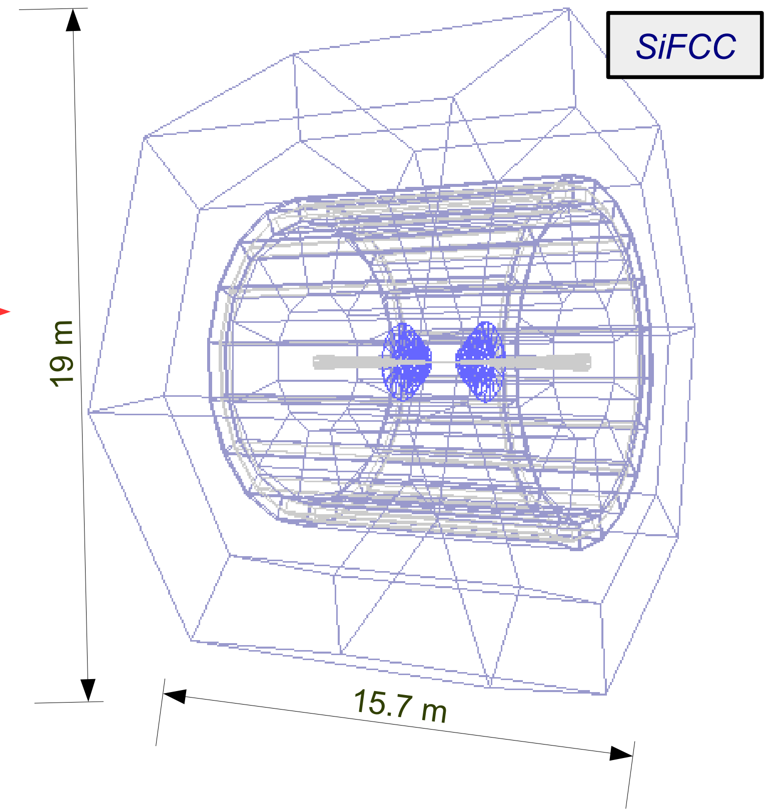
T.Carli, C.Helsens, A.Henriques Correia, C.Solans: 016_JINST_11_P09012. arXiv:1604.01415

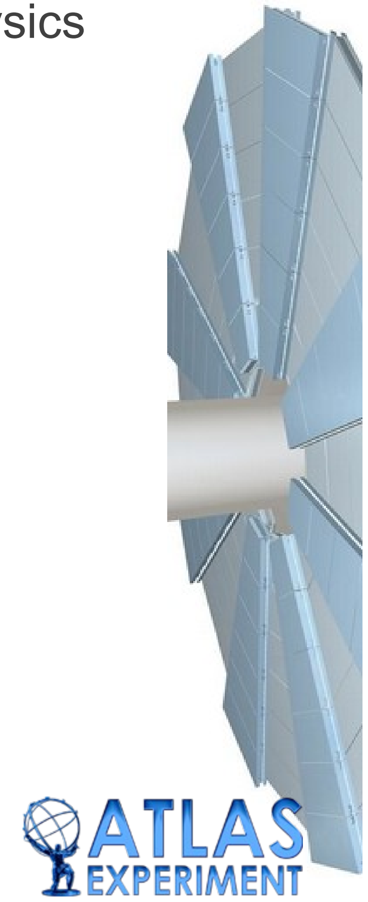
SiFCC detector for performance studies

- Design a FCC-like detector using SiD (ILC) detector software
- Study energy resolution, response and granularity for ~tens TeV physics

SiD e+e- (sidloi3) detector







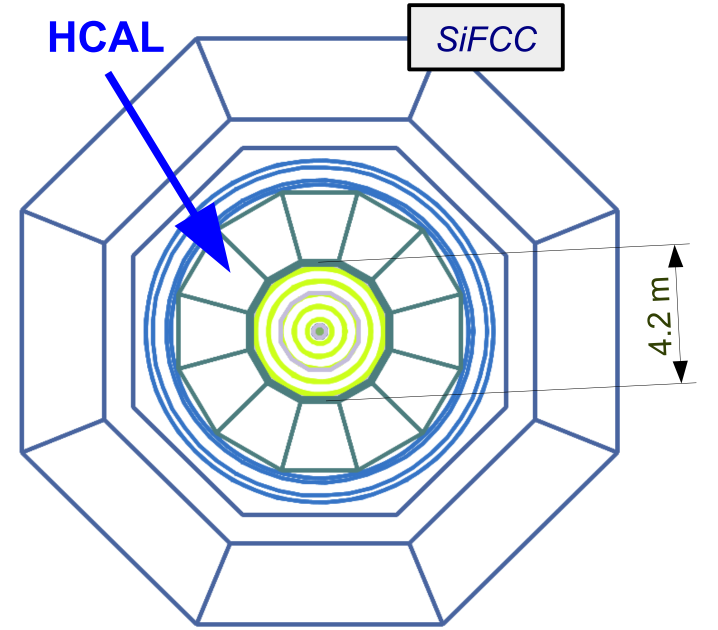
SiFCC (v7) detector:

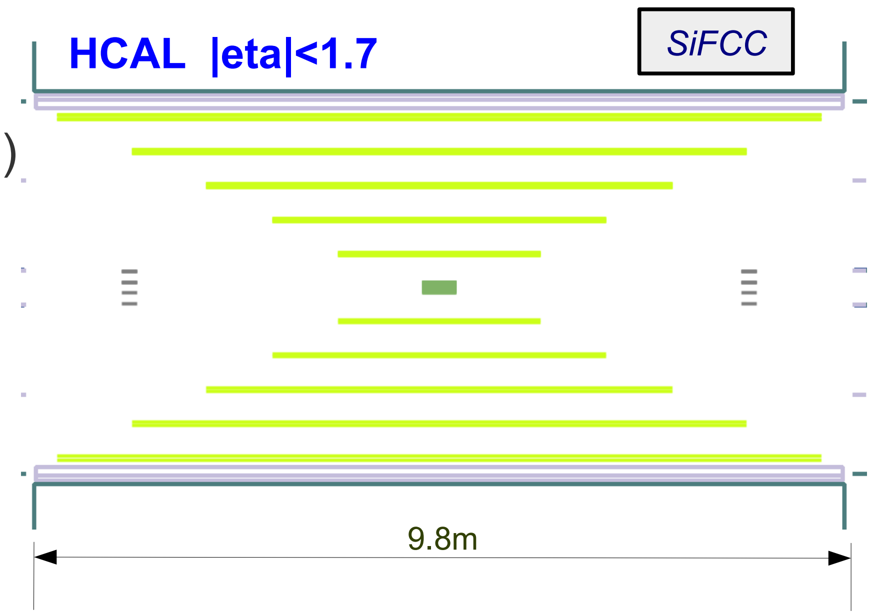
- Multipurpose, high granularity, compact detector
- 30% smaller than ATLAS (R=25 m vs R=19), 30% larger than CMS (R=14.6 m vs R=19 m)

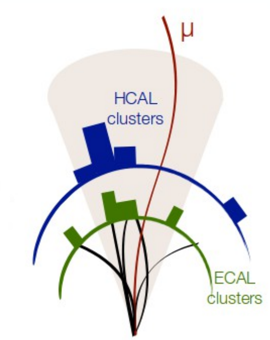
Characteristics of SiFCC (version 7)

http://atlaswww.hep.anl.gov/hepsim/soft/detectors/sifcch7/sifcch7.html

- **5 T solenoid outside HCAL**
- **Si pixel and outer trackers:**
  - 20 um pixel (inner), 50 um (outer)
- **ECAL (Si/W):** 2x2 cm. 32 layers, ~35 X0
- **HCAL (Scint. / Fe)** **~ FCC-hh baseline**
  - 5x5 cm cells: $\Delta\eta \times \Delta\varphi = 0.022 \times 0.022$
  - x4 smaller than for CMS & ATLAS
  - 64 longitudinal layers → 11.3 λ_I
  - 3.1% sampling fraction
- > 150 M non-projective cells (ECAL+HCAL)







Optimized for Particle Flow Algorithms

High granularity HCAL for 100 TeV physics?

- Baseline for past & operational detectors:
  - **transverse cell size is similar or larger than nuclear interaction length: λ_I**
- **Recent high-granularity HCAL: CMS (upgrade), CALICE R&D:**
  - 2x2 or 1x1 cm cell sizes required to reconstruct PFA & separate particles
- **Main question for a 100 TeV collider:**

Can reconstruction of jets and particles at tens-TeV scale benefit from small HCAL cells (<< λ_I) ?

Data with simulations available from HepSim repository: http://atlaswww.hep.anl.gov/hepsim/

SiFCC detector version (Fe/Scin. HCAL)	Transverse size of HCAL cells (cm)	Transverse size of HCAL cells in λ_I	Simulation tag in HepSim
SiFCC-v7 (baseline)	5X5 cm	~ λ_I /4	rfull009
SiFCC-v8 (traditional)	20x20 cm	~ λ_I	rfull010
SiFCC-v9 (as ECAL)	2x2 cm	λ_I /8	rfull011
SiFCC-v10 (fine)	1x1 cm	λ_I /17	rfull012

Energy reconstruction in HCAL (SiFCC)

SLIC (Geant4 10.1)
Inelastic models for pi+/p/n
(validated<400 GeV):
QGSP: 12 GeV - 100 TeV
FTFP: 9.5 GeV - 25 GeV
BertiniCascade: 0 eV - 9.9 GeV
Elastic:
ElasticLHEP/ Gheisha: 0 eV - 100 TeV

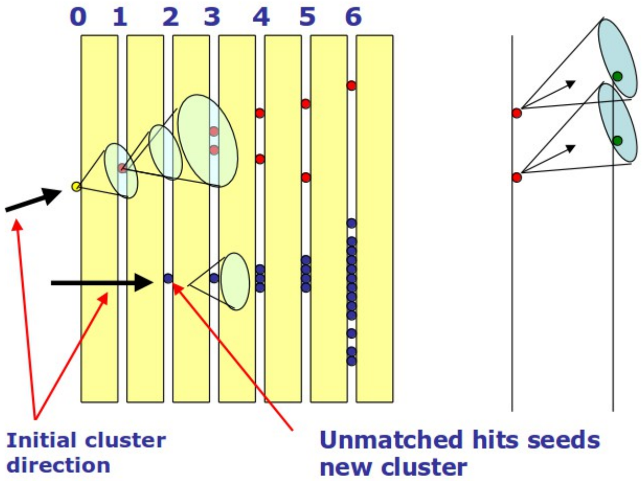
Sampling fractions & digitization
Reconstruction of cell energies

Building topological 3D RecoClusters using PandoraPFA

Building jets from RecoClusters using antiKT R=05 jet algorithms

Data analysis

From M.Thomson



Cone algorithm
Start from inner layer and work outward

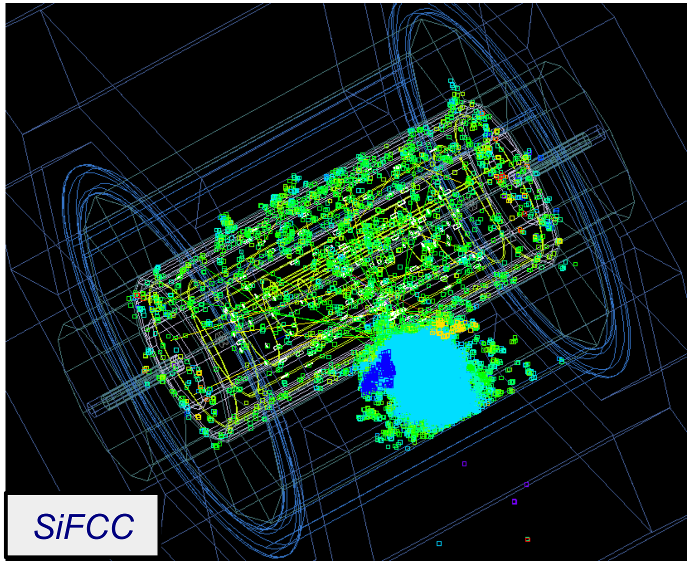
* Pandora PFA objects with track information are used
→ requires optimization

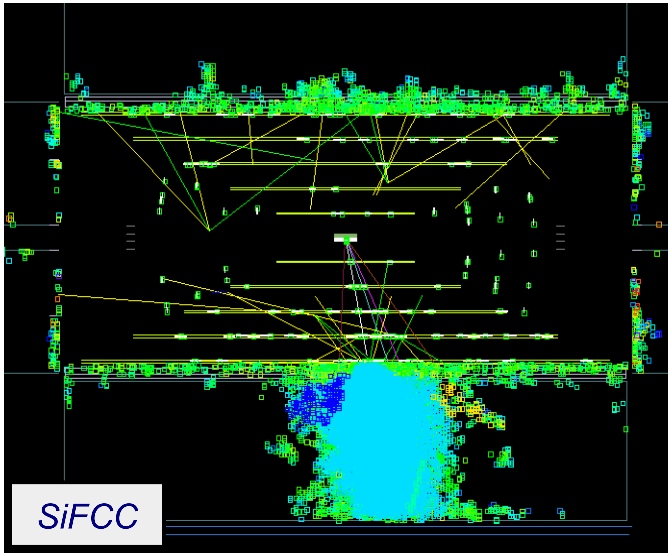
Response to single particles: 8 TeV pions

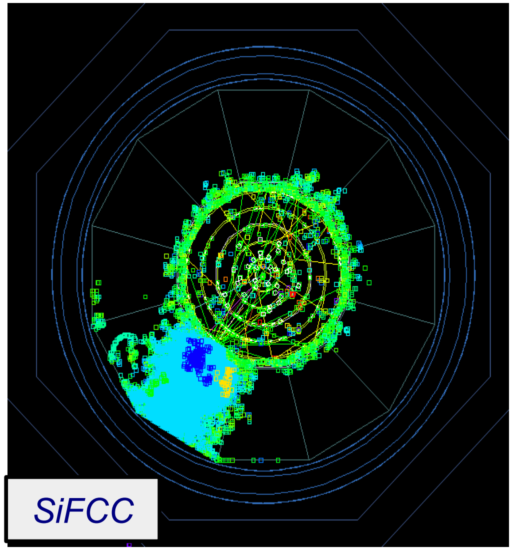
Example: True momentum of π+ : 8.156 TeV

After SiFCC reconstruction (>1.5 M HCAL cells):

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



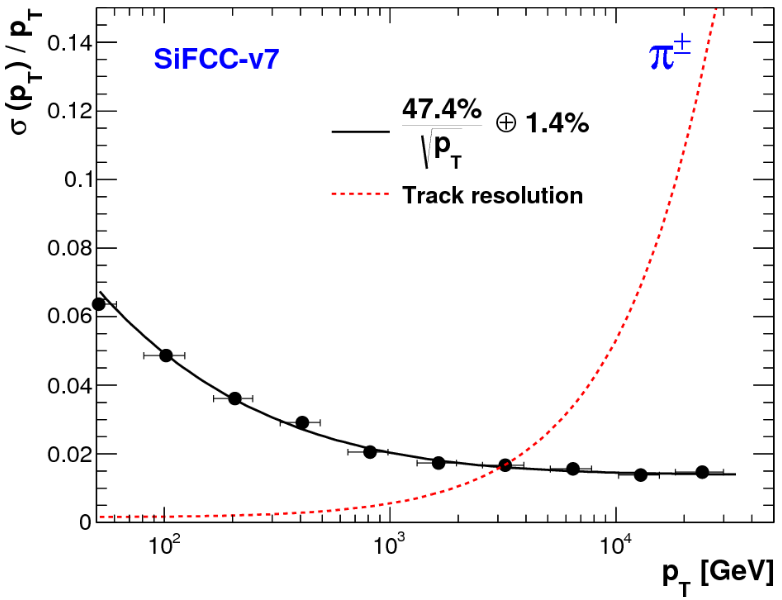


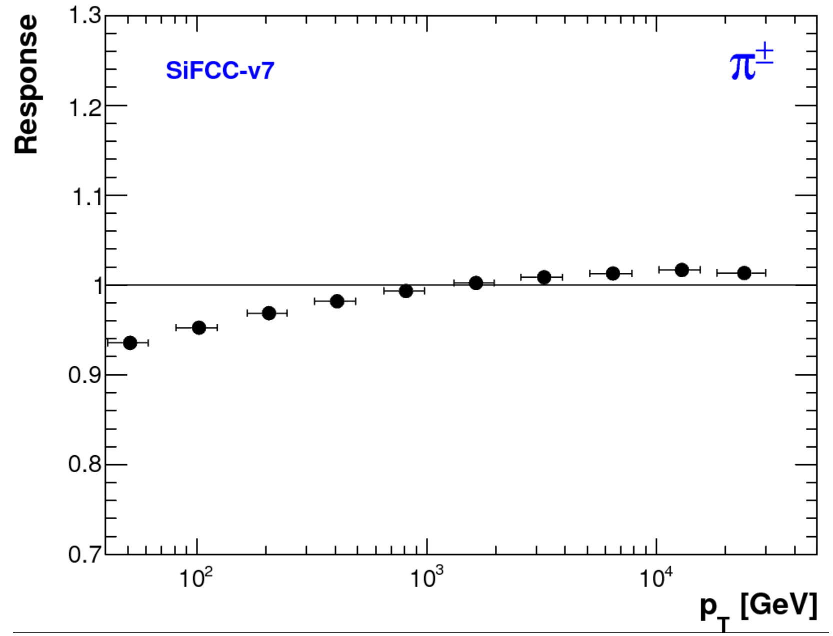


Based on **HepSim**:   http://atlaswww.hep.anl.gov/hepsim/info.php?item=201

Response to hadrons: $\pi^{\pm}$

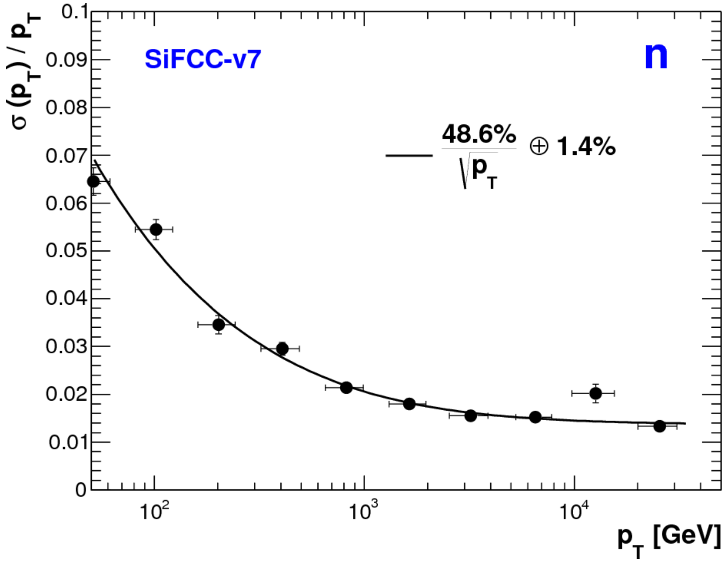
- Single pi+ randomly distributed in Eta & Phi
- pT is reconstructed by collecting energies from all RecoClusters

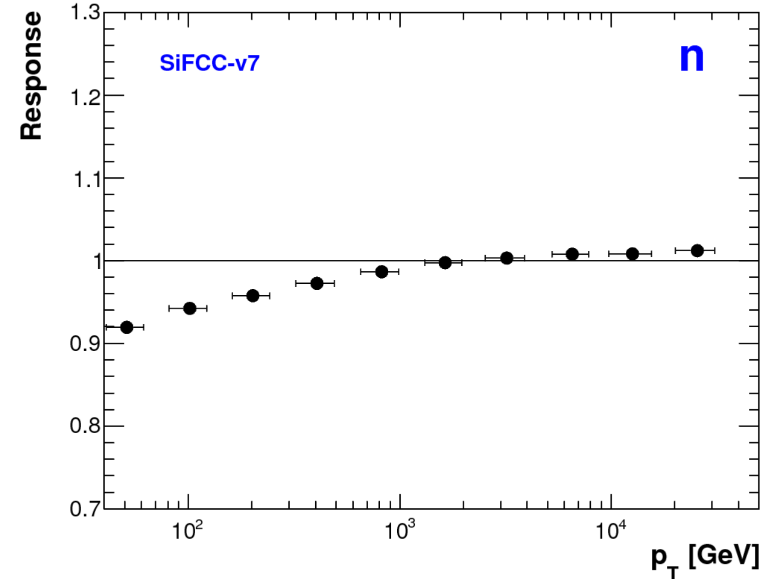


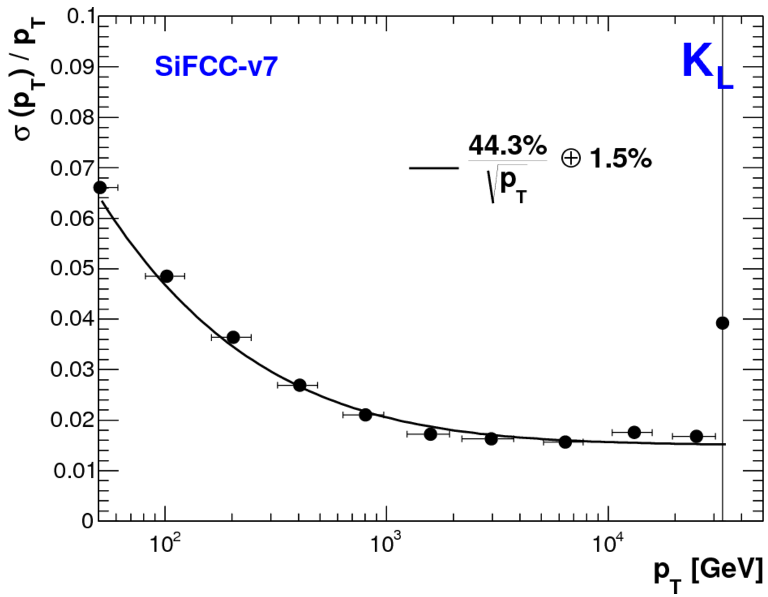


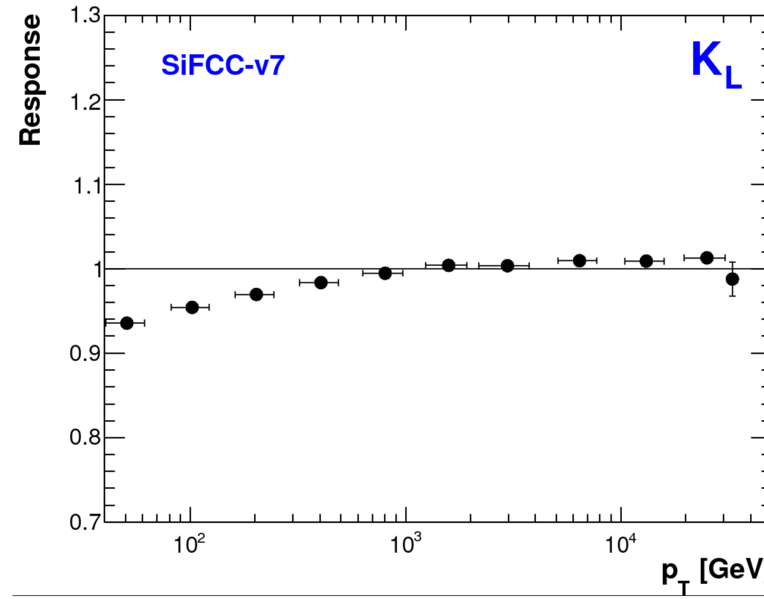
- ~47% sampling term, 1.4% constant term (the noise is small)
  - the sampling term is consistent with ATLAS-like setup (arXiv:1604.01415)
- Calorimeter resolution is better than for SiTracker for pT>3 TeV
  - tracks were studied using single muons
- Calorimeter response is non-linear → should be corrected by MC (e/p, material correction etc.)

Response to neutrons and K_L

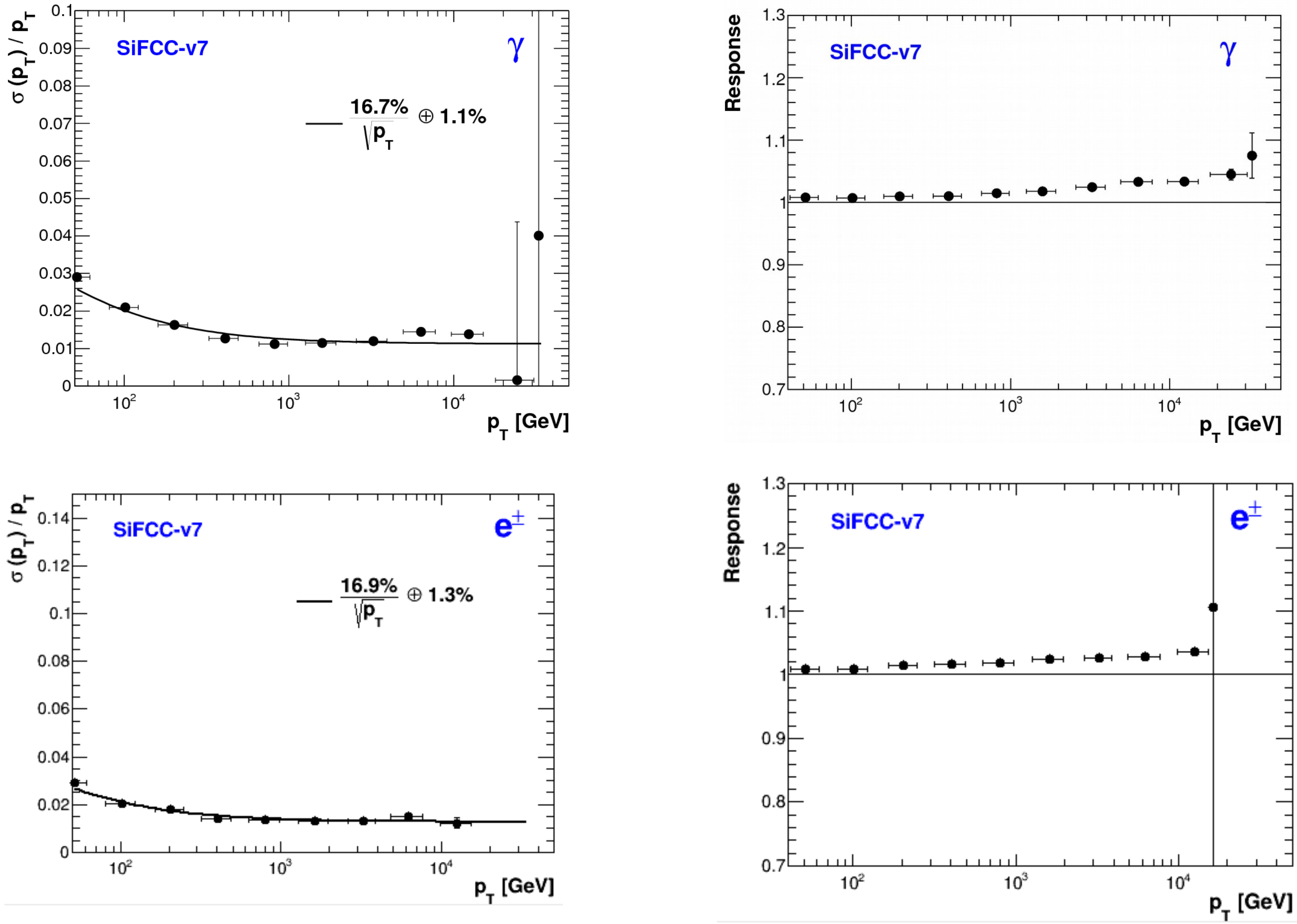








Single particle resolution and response (e/γ/π^0)

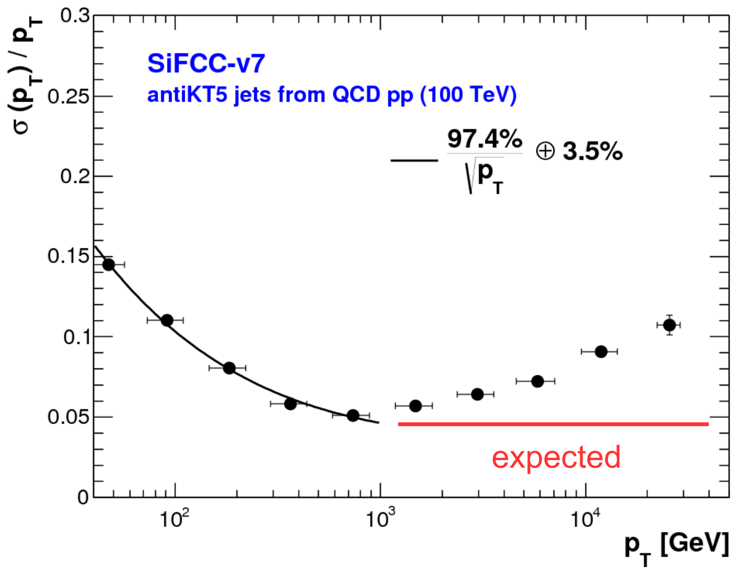


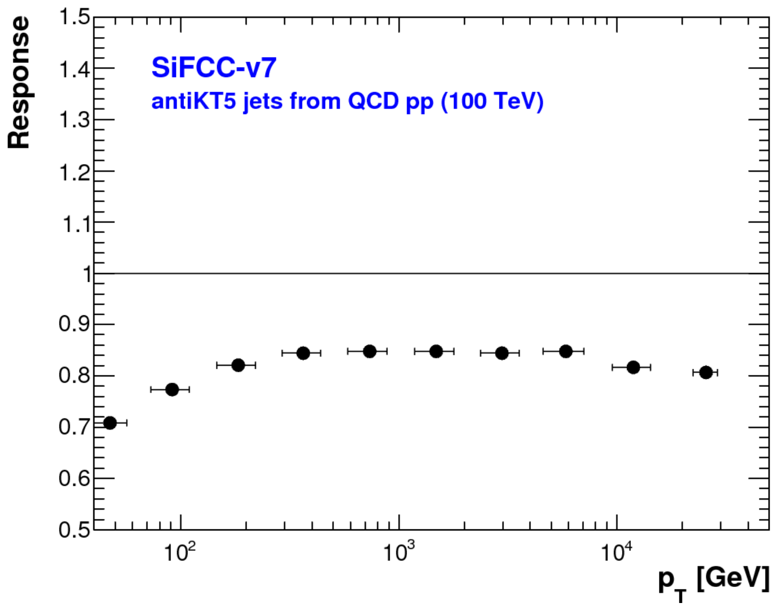
- Reasonable performance of ECAL: ~17% sampling term, 1.3% constant term

Jet energy resolution & response

tev100_qcd_pythia8_ptall
with rfull009 tag

- Jets from 100 TeV pp collisions generated with Pythia8 with different pT(min)
- Use RecoClusters for antiKT jet algorithms with size R=0.5 (not PFA, no tracks)

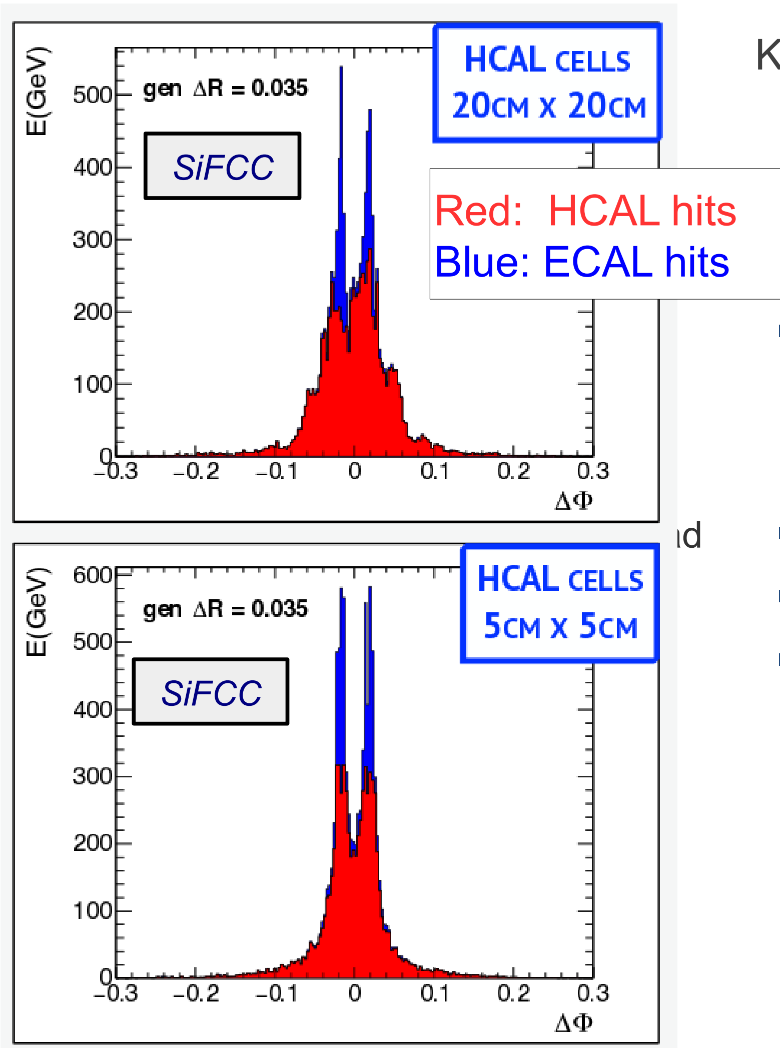


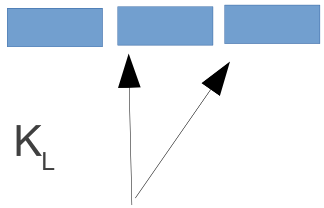


- Jet energy resolution is similar to ATLAS jets before correction ("EM" scale) for pT<2 TeV
- Jet response is lower than for single particles (curved tracks, e/p effect, inactive material, etc).
  - Requires jet energy corrections
- Surprise: resolution increases above 2 TeV and reaches 0.1 at 30 TeV
  - The result is consistent across various similar studies using SiFCC (i.e. Z' events etc.)
  - Searching for explanations (Geant4? Reconstruction problem?)

HCAL segmentation and spacial separation of hadronic showers

truth-level separation 0.035 rad (2 deg)





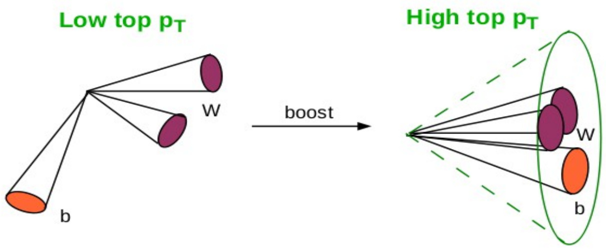
T.Nhan
Presented at Boost2016/ICHEP16

- Generate two K_L (E=100 GeV) particles at η=0.
  - First K_L is always at φ=0
  - Second is shifted by $\Delta\varphi$=1,2,3.. deg
- Simulate and reconstruct with SiFCC
- Calculate energy of hits in Φ with respect to Φ=0
- Repeat for different HCAL cell sizes

Small HCAL cells (~ λ_I /4 size) helps separate hadronic showers produced by two K_L separated by 2 deg

Physics processes for boosted jet studies

- Muon collisions to speed up calculations: no complications due proton beams
- Benchmark process: Z' with masses 10, 20, 30, 40 TeV and ΔΓ(Z')~ 1 MeV:
  - $\mu^+\mu^- \rightarrow Z' \rightarrow W^+W^-$
  - $\mu^+\mu^- \rightarrow Z' \rightarrow q\bar{q}$
  - $\mu^+\mu^- \rightarrow Z' \rightarrow t\bar{t}$
  - $\mu^+\mu^- \rightarrow Z' \rightarrow$ tau+tau-
  - $\mu^+\mu^- \rightarrow Z' \rightarrow b\bar{b}$

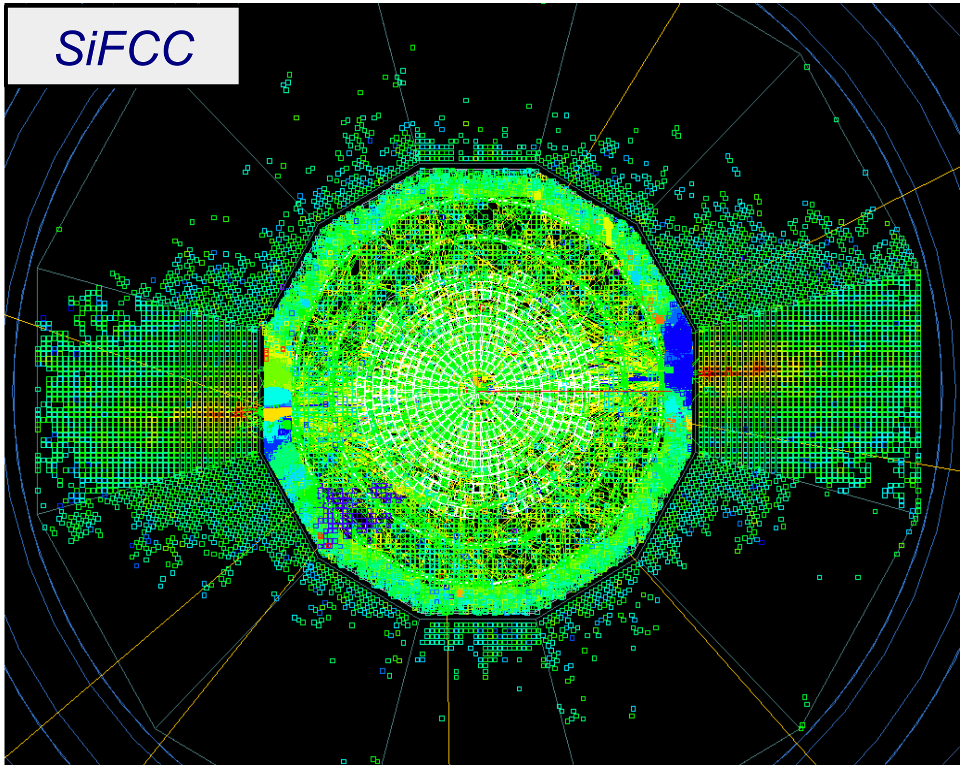


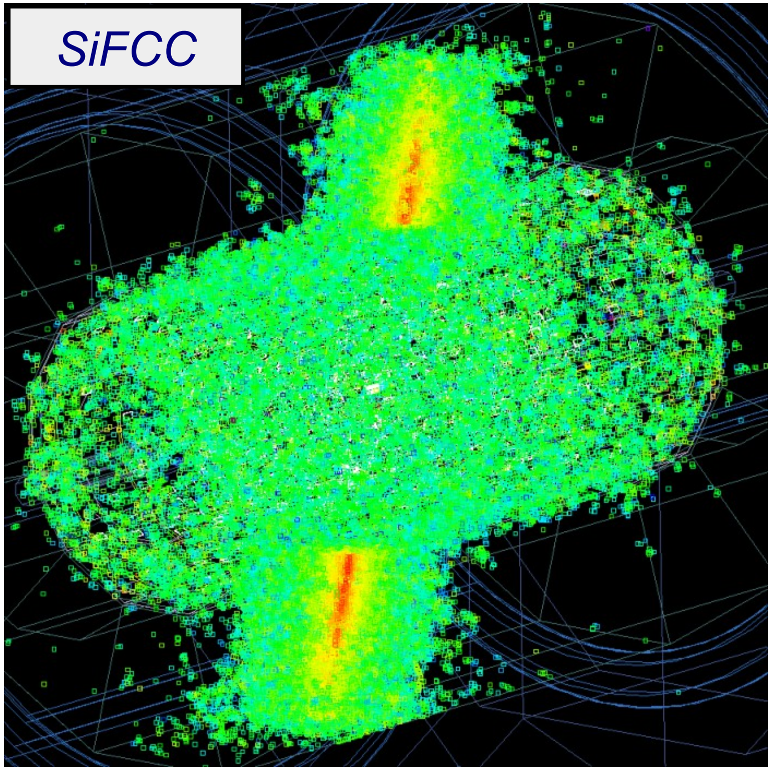
- Use substructure techniques to identify WW, $t\bar{t}$ and compare with $Z' \rightarrow q\bar{q}$

  - about 2000 fully reconstructed events per sample (Tracks, PFA, CaloClusters, HITS)
  - created on Open-Science Grid (UChicago/ANL. ~100,000 CPU*h)

Event display of Z' (40 TeV) → $q\bar{q}$

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



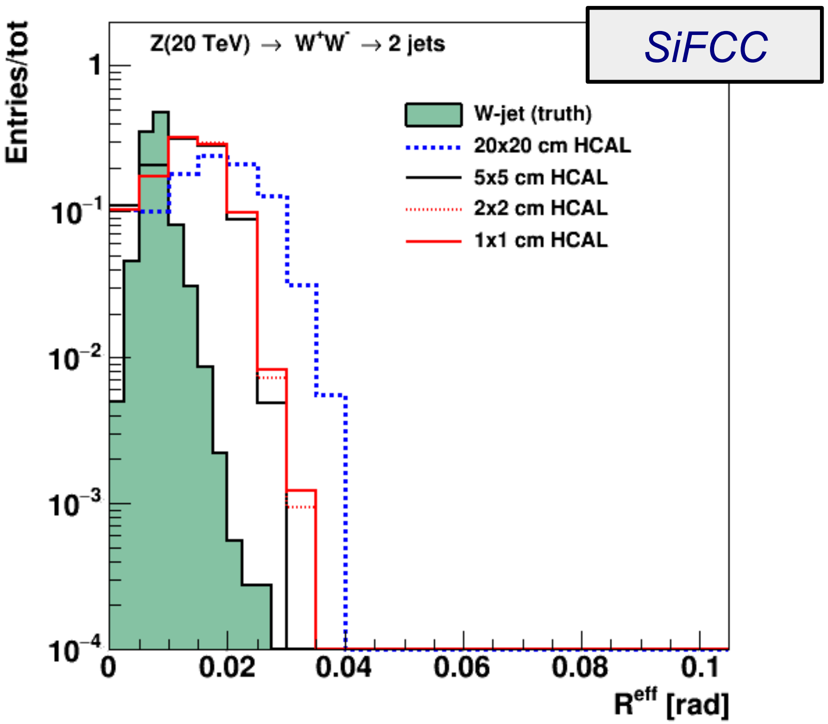


Effective jet radius for antiKT5 jets

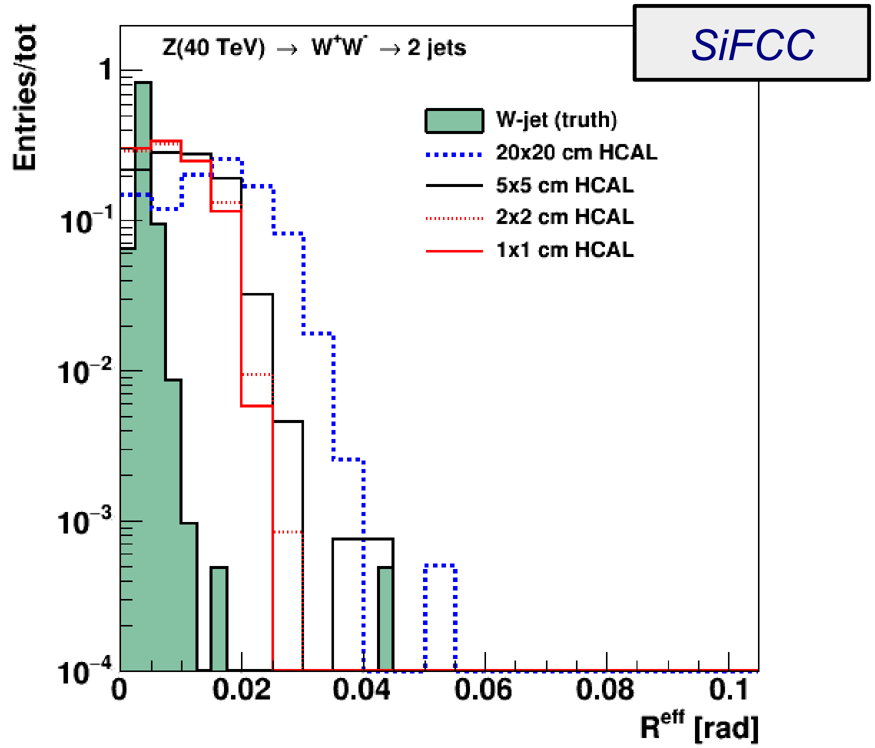
Phys. Rev. D 91 (2015) 034014

Sum over all distances between constituents and jet center, weighted with E(const) / E(jet)

W-jets from Z'(20 TeV)



W-jets from Z'(40 TeV)

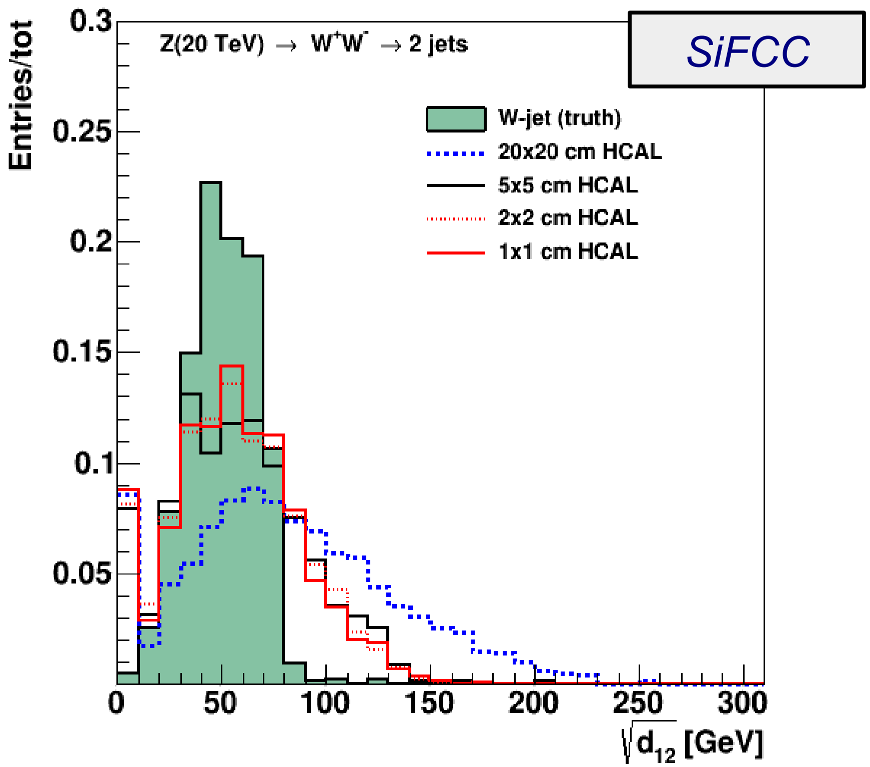


- Jets with pT>10,20 TeV, each from W decays (qqbar)
- Narrow ($\Delta R \sim 2^*$ pT / M(W)) compared to QCD jets (not shown)
- 5x5 cells better reflect true effective jet size compared to ~20x20 cm (ATLAS/CMS)
- Small difference between 2 cm and 1cm cell sizes

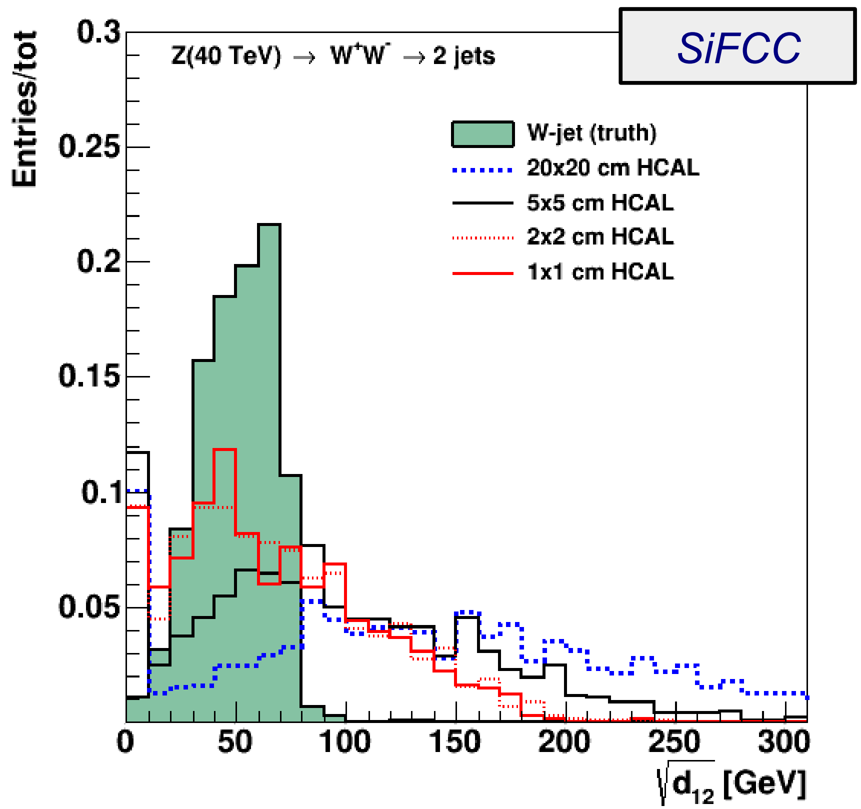
Jet splitting scale: d_{12}

Kt scale at which a jet splits into 2. Used to differentiate QCD jets from 2-body decays (W,H,etc)

W-jets from Z'(20 TeV)



W-jets from Z'(40 TeV)



- Jets with pT>10,20 TeV, each from W decays (qqbar)
- 5x5 cells better reflect true effective jet size compared to 20x20 cm (ATLAS)
- Small difference between 2 cm and 1 cm cell sizes

Summary of HCAL studies for energy frontier

- **First realistic physics processes for boosted topologies have been simulated and reconstructed up to 30 TeV**

- **Overwhelming evidence that we gain useful information from cell sizes smaller than nuclear interaction length for hadronic showers initiated by multi-TeV particles & jets**
  - Optimal size using RecoClusters **~ λ_I /4** or (or **~5x5 cm** for Fe/Sci HCAL of SiFCC)
  - Consistent with previous studies based on fast simulations (CPAD2015, FCC weeks)

- **Cost-effective technology is required to build high granularity calorimeter with large dynamic range of cells (<10$/channel?!)**

Summary

- **First public Monte Carlo repository with fast and full detector simulations**
- **Enable physics & detector-performance studies for current & future colliders + community outreach**
  - 1.6 billion events at the EVENT level for public downloads
  - Significant number of fast and fully reconstructed events for ep, μμ, ee, pp collisions (13-100 TeV) & single-particle samples for detector studies
- **14 articles, ~25 presentations since 2014 (linked to WWW):**
  - Physics reach studies for HL-LHC, HE-LHC, FCC-hh etc.
  - Calorimeter studies (cell granularity)
  - Tracking optimization at multi-TeV scale
  - Software development
- **2.5 million CPU*h from OSG-grid. OSG-Connect support from UChicago**
- **Contributions from 17 students/scientists**
- **Your contribution is welcome!**

How to contribute to HepSim

In addition to physics & detector performance studies, you can contribute to the simulations too!

- Generate EVGEN archive files with physics processes
- Validate using the HEPSIM tools (if you can)
- Contribute to the software tools
- Setup a HepSim mirror:
  - data server with HTTP access
  - can maintain your own EVGEN & full simulation files

Thanks!

People

Contributions to HepSim

Here is a list of people who contributed to the project:

- S. Chekanov (main developer and maintainer)
- E. May - ProMC format development, benchmarks on BlueGene/Q (ANL), Jas4pp debugging
- 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
- H. Gray - Delphes card for FCC geometry
- J. Strube (PNNL) - LCIO/SLIC for full simulation
- A. Kotwal (Duke Univ.) - LCIO/SLIC for full simulation
- J. Adelman (NIU) - debugging post-Snowmass Delphes 3.3 card for 13/14 TeV
- S. Padhi (prototyping Snowmass Delphes3.1 during Snowmass 2013)
- K. Pedersen (alternative b-tagging for rfast003)
- Shin-Shan Yu (heavy higgs MG5 simulations)
- Joel Zuzelski (ANL, SULI 2016) - SLCIO reader, converter promc2slcio, new tracking geometry for SiF
- Boruo Xu (Bono) (xu@hep.phy.cam.ac.uk) - help with moving to new pandora
- John Marshall (marshall@hep.phy.cam.ac.uk) - adaptation of slicPandora for fast Pandora in HepSim

We also acknowledge the computer support by:

- Lincoln Bryant and Bala Desinghu (OSG-Connect)
- David Champion and Rob Gardner (ATLAS-connect / MWT-Tier2)

We apologies in advance if some names are missing.

Using High Performance computing at ANL

- ANL is home of MIRA, an IBM Blue Gene/Q supercomputer at the Argonne Leadership Computing Facility supported by DOE, is equipped with 786,432 cores for open-science projects
- Several MC models ported on Mira: Jetphox, MCFM, Alpgen, Pythia8
- ALPGEN Monte Carlo samples (V+jets) created for the ATLAS (LHC) collaboration
  - See: J.Childers, T.Uram, T.LeCompte, M.Papka, D.Benjamin http://www.sciencedirect.com/science/article/pii/S0010465516302843
- 2019: target date for user access for AURORA

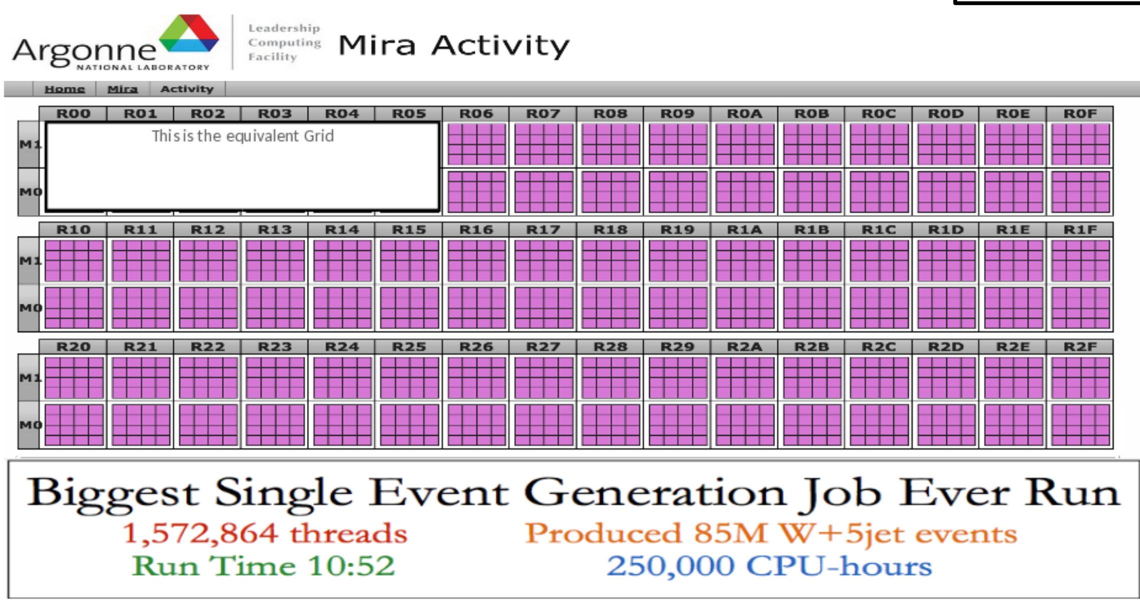
System Features	Mira	Aurora intel CRAY
Compute Nodes	49,152	>50,000
Processor	PowerPC A2 1600 MHz	3rd Generation Intel Xeon Phi
System Memory	768 TB	>7 PB DRAM and persistent memory
System Interconnect	IBM 5D torus interconnect with VCSEL photonics	2nd Generation Intel Omni-Path Architecture with silicon photonics
File System Capacity	26 PB GPFS	>150 PB Lustre
Intel Architecture (x86-64) Compatibility	No	Yes
Peak Power Consumption	4.8 MW	13 MW

http://aurora.alcf.anl.gov/

ALPGEN and Sherpa on Mira

- W/Z+N jets (N>3) takes tens of minutes per event to generate (tree-level matrix element (QCD ME)
- Ran serial AlLPGEN in parallel with minimal MPI additions for random number seeds and file I/O
- Used in ATLAS publications

ANL team: J.Childers, T.Uram, T.LeCompte, M.Papka, D.Benjamin
http://www.sciencedirect.com/science/article/pii/S0010465516302843



Mira activity figure shows ALPGEN job and the white box represents the size of the US ATLAS Grid if it was running the same job. Mira was >6x bigger for this work.

Ongoing work on another popular Monte Carlo simulation - **Sherpa**