Introduction to Geant4

Marc Verderi,
Laboratoire Leprince-Ringuet,
Institut polytechnique de Paris,
Palaiseau,
France

Overview of presentation

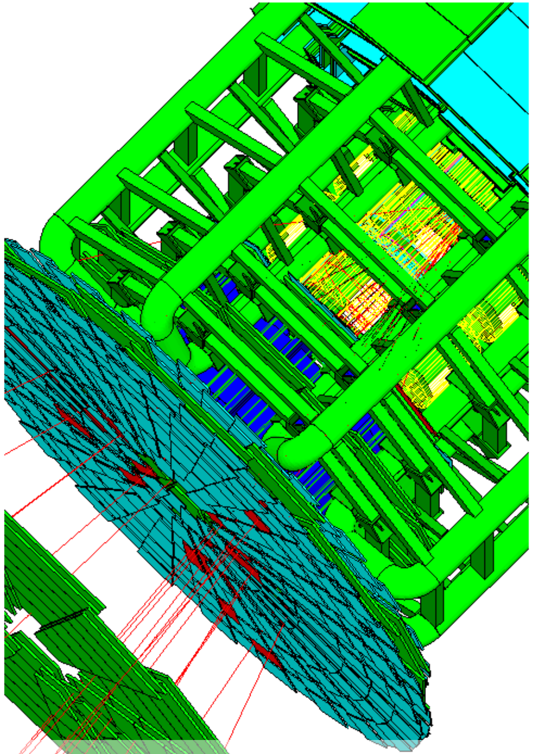
- Geant4 Overview
- Geant4 Application Domains
- Physics Modeling
- Software Aspects
- Coming Challenges

Geant4 Overview

What is Geant4 ?

Geant4 as a Software Toolkit

- **Geant4 is an Object Oriented Monte Carlo particle transport software toolkit for simulating the passage of elementary particles through matter and interacting with it.**
- It started in **1994** as the CERN **RD44** project :
  - **Goal of RD44 : assess the benefit of OO technologies for detector simulation for LHC era** (LHC yet to come at that time !)
    - Medical and space domains requests included since the beginning !
  - **Geant4 v1.0 released in Dec 1998**
    - After alpha release in Apr 1997 and beta one in Jul 1998
- **Key functionalities:**
  - Kernel → to **animate** the system
  - Geometry + navigation & materials → to **describe** the setup
  - Physics processes & tracking → to **generate** the series of **physics** interactions
    - EM (O(100 eV) – PeV), special extensions (O(eV) & O(mK)), hadronic (rest - multi-TeV)
  - Scoring → to **collect data** from the simulation
  - GUI and Visualization drivers → to **pilot** the application and **visualize**
- **"Toolkit"** because users **select** components and **build** their application
  - Not an application like ROOT, or Powerpoint, etc.
- **Users can extend the toolkit !**



The LHC ATLAS Detector

What is Geant4 ?

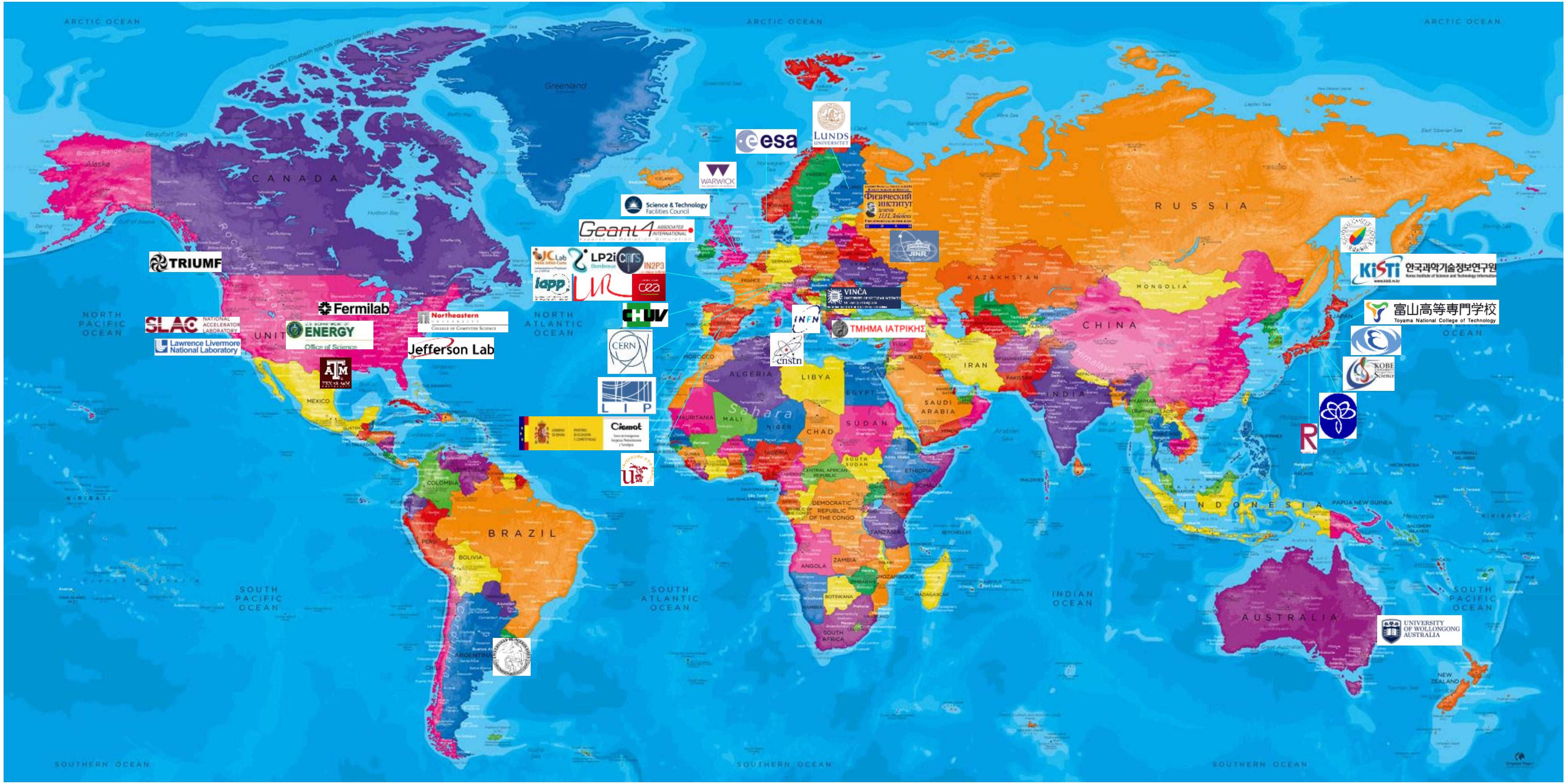
Geant4 as a Collaboration

Some of the members, at last Collaboration Meeting (2022)

- **Geant4 is also the name of the Collaboration maintaining, developing and validating the software**
  - ~130 members + O(10) “contributors” = new light status
  - ~30 FTE
  - ~30 institutes, worldwide
    - (Map of collaborative institutes after)
  - 16 working groups
- **Web site:**
  - http://geant4.cern.ch/
  - **Download area,** documentation, news, announcement of releases, meetings (**Technical Forum**, etc.)

- **Distributed development model:**
  - Based on **GitLab** (geant4-dev repo.)
    - Reserved to members & contributors
  - About 1000 Merge Requests / year
- **Distribution through:**
  - Geant4 Web site
  - **GitHub** instance
    - GitLab mirror for public releases & patches
    - **Open to public for Pull Requests**
  - Special way, CVMFS, for LHC experiments (monthly tag)
- **One public release/year, in December**
  - Latest release: Geant4 v11.1
- **Three general papers:**
  - “Geant4: a simulation toolkit”, S. Agostinelli *et al*., NIM A, vol. 506, no. 3, pp. 250-303, 2003
  - “Geant4 Developments and Applications”, J. Allison *et al.*, IEEE TNS, vol. 53, no. 1, pp. 270-278, 2006
  - “Recent Developments in Geant4”, J. Allison *et al.,* NIM A, vol. 835, pp. 186-225, 2016

Map of Collaborative Institutes

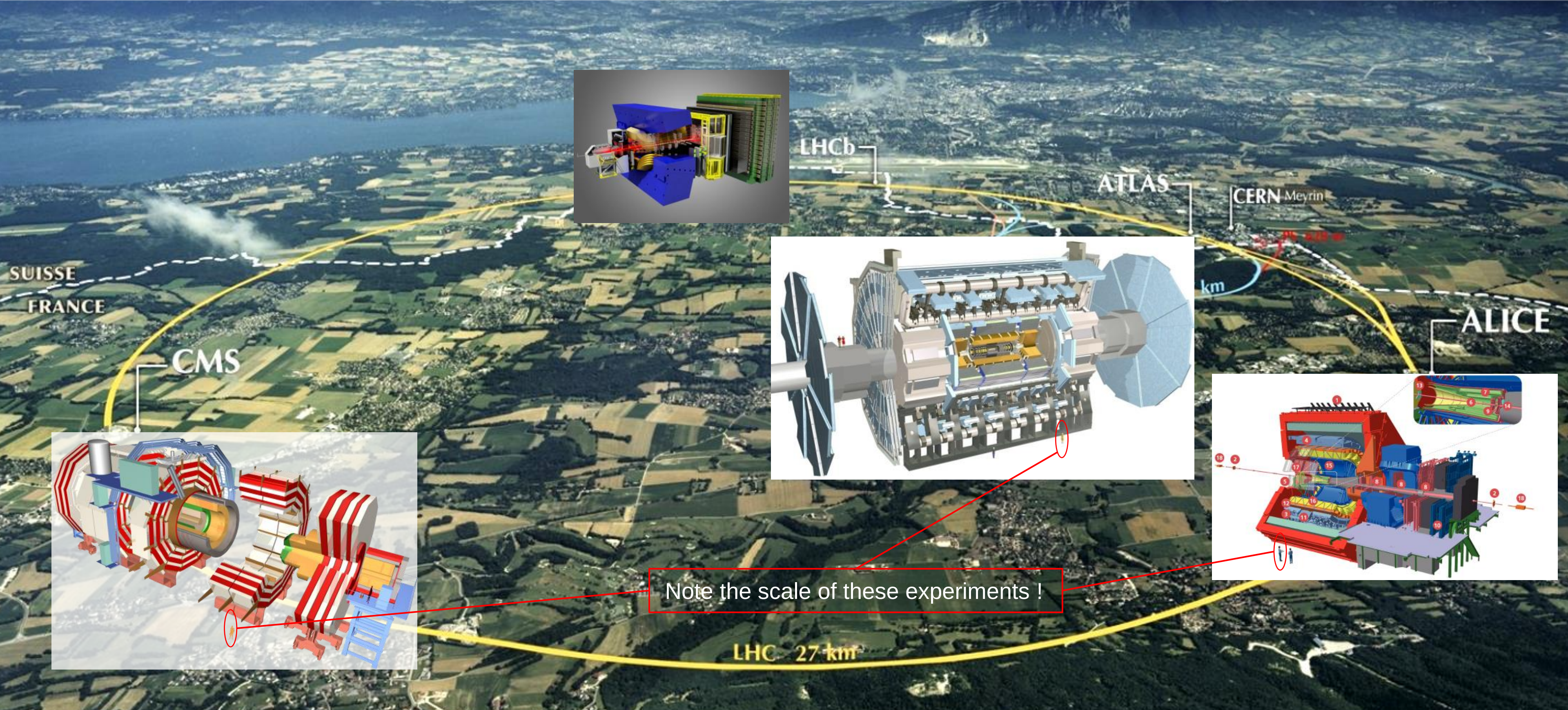


Geant4 Application Domains

Not an exhaustive coverage of domains !

With few examples of application in these selected domains

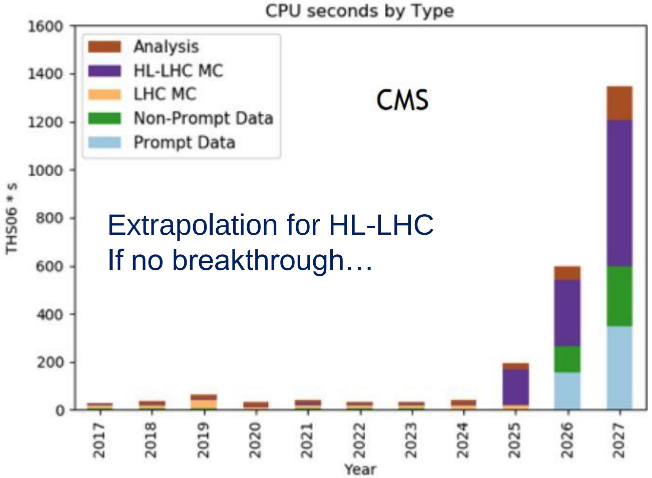
High Energy Physics : LHC @ CERN

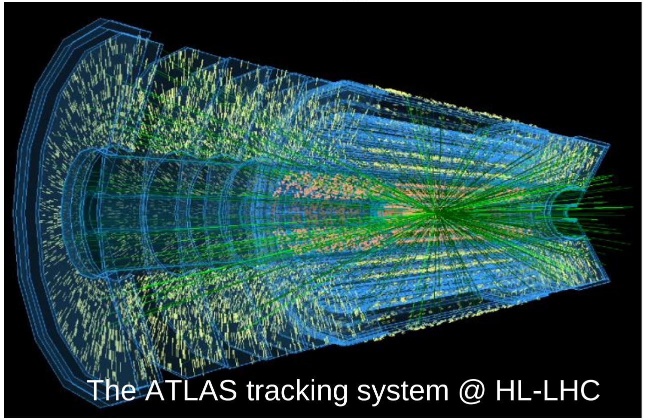


High Energy Physics

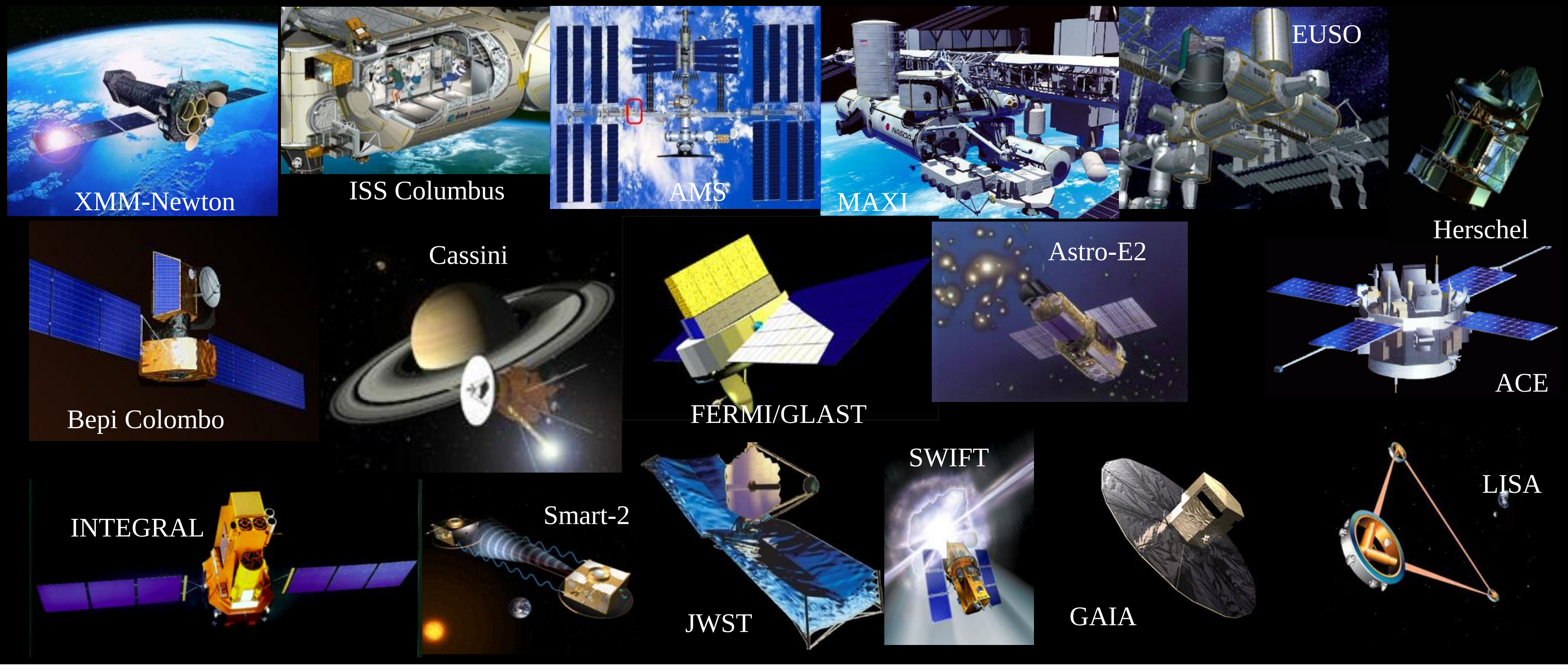
- LHC experiments **very demanding** in terms of simulation
  - **Large detectors :**
    - O(1 – 10) millions of volumes
  - **High energy O(10 TeV) in center of mass:**
    - Generate many tracks, daughters tracks, etc.
    - Up to O(10 k) tracks/events in lead-lead collisions !
  - **Processing time and production volume:**
    - From O(1 s) to O(1 mn) per event
    - O(10^9) events processed / experiment !
    - Processing made on large infrastructure (GRID)
    - → high demand on Geant4 to be **"as fast as possible"**
      - **Each % improvement is a lot of money saved…**

- **New phase HL-LHC (~2027) even more challenging !**
  - **Typical request for O(10) higher throughput !**
    - Together with **better physics** (syst. errors relative to stat. errors)
    - And higher pill-up, higher detector granularity for some part…
  - Triggers quite R&Ds activities:
    - Trying to port **Geant4 on GPU** : **not trivial at all !**
    - Consider part of the simulation with **AI-based fast simulation**



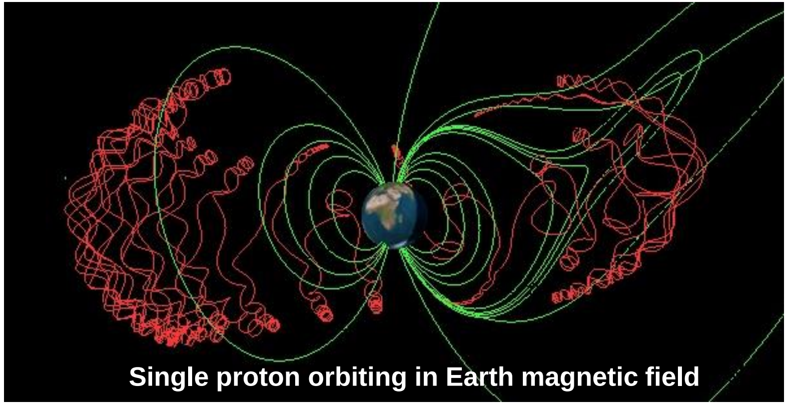


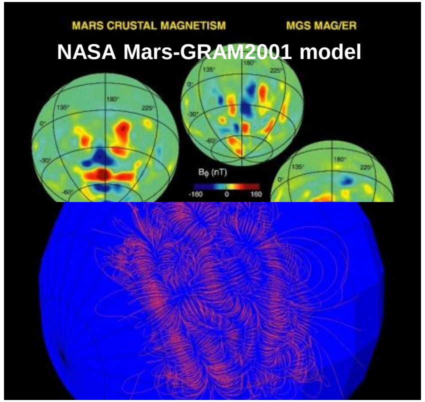
Geant4 in Space



Planetary scale simulation, dosimetry

- **Planetocosmic:**
  - Geant4 simulation of Cosmic Rays in planetary Atmo- /Magneto- spheres
  - Laurent Desorgher *et al.* (Now at ICHUV, Switzerland)

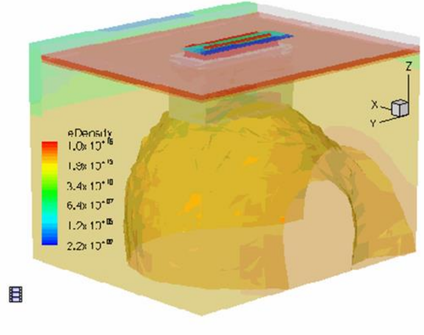

Single proton orbiting in Earth magnetic field

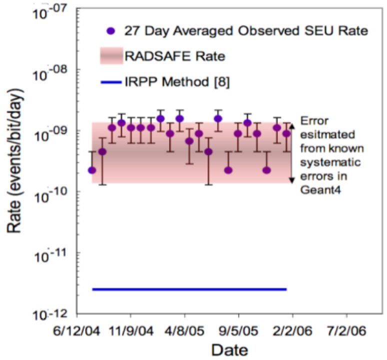


NASA Mars-GRAM2001 model

- **Single event effect rate:**
  - RADSAFE / MRED project
  - Robert A. Weller *et al.* (Vanderbilt University, Nashville, TN, USA)

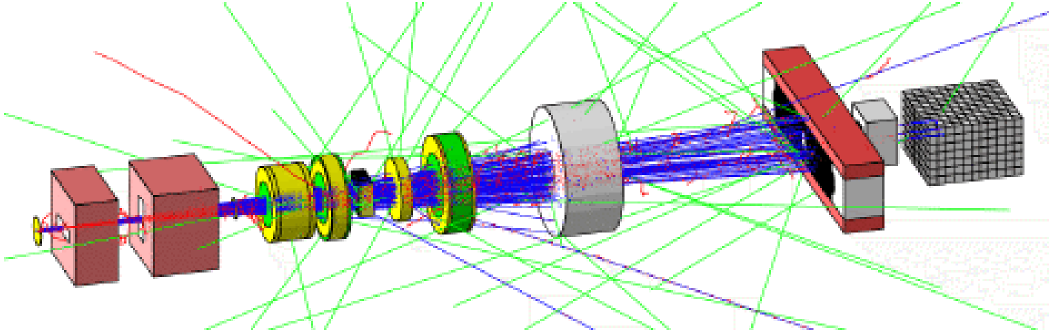

Single Charge Deposition in TCAD: Ne+W Event

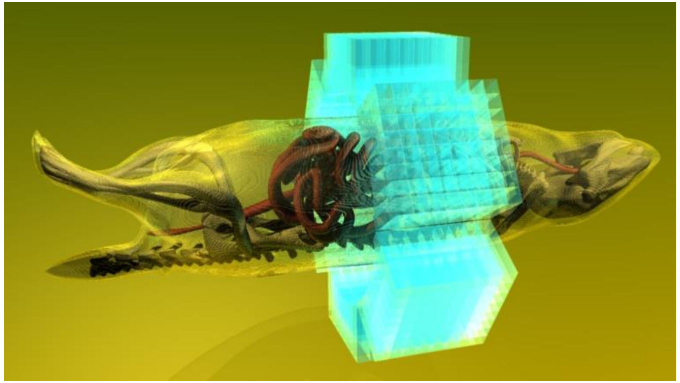


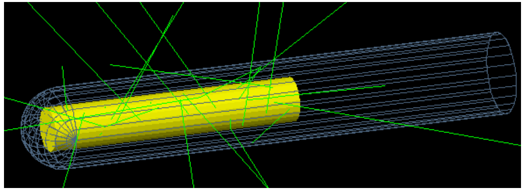


Geant4 in Medical Science

- Main use cases:
  - Beam therapy
  - Brachytherapy
  - Imaging
  - Irradiation study

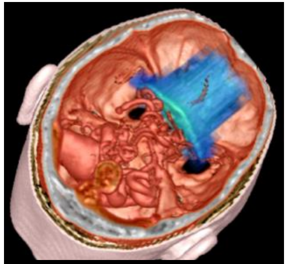




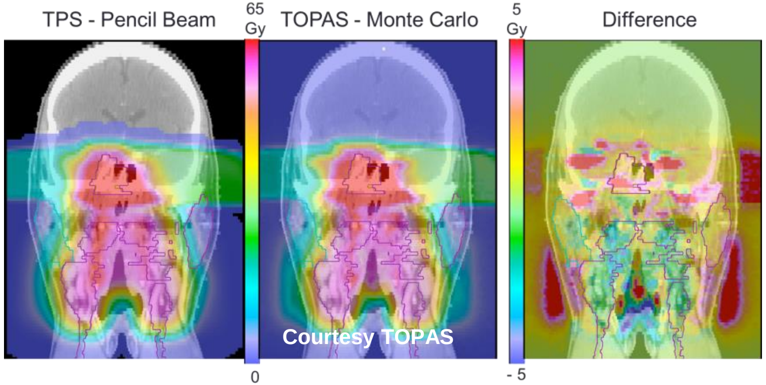


Beam Therapy, Brachytherapy

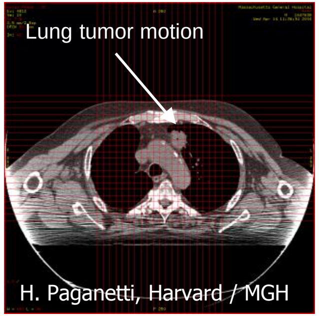
- Beam therapy



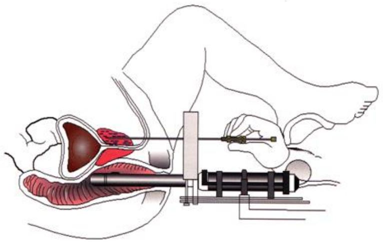
Dose by proton beam
Courtesy of gMocren



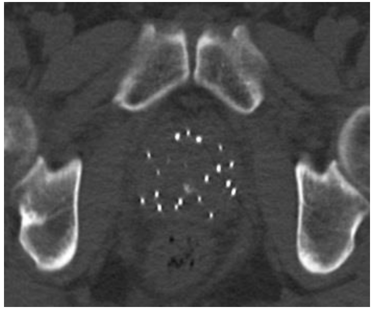
10 beams (6 directions + 4 boost)

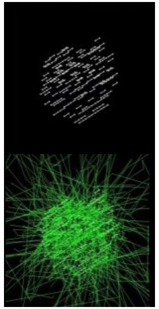


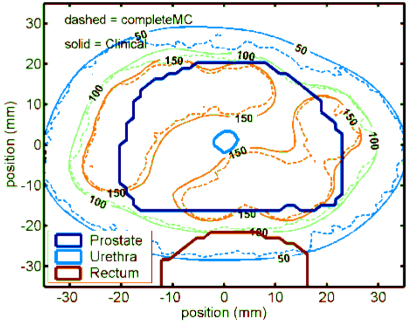
- Brachytherapy



Jean-François Carrier, CHUM

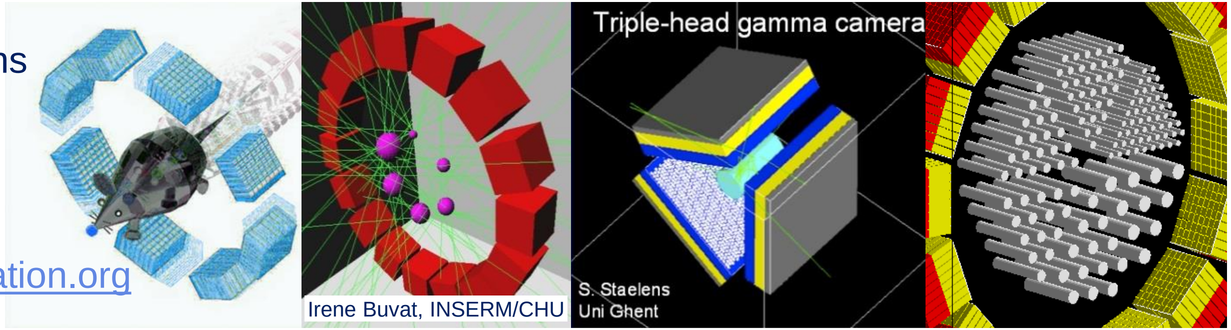




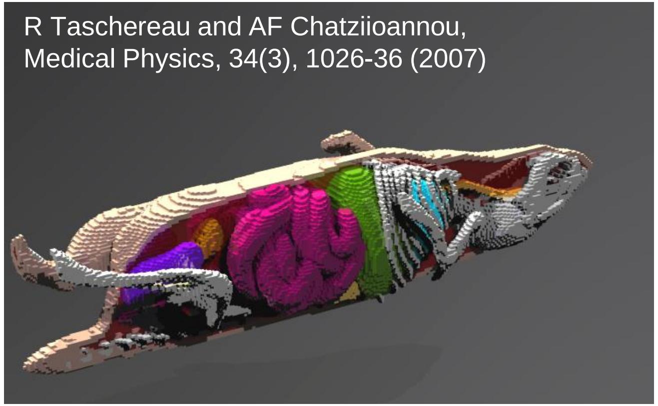


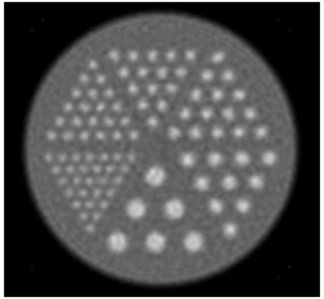
Imaging

- GATE
  - Toolkit for Imaging applications
  - based on the Geant4 toolkit
  - easier to use for Imaging applications
  - http://www.opengatecollaboration.org



- Ex of High resolution phantoms
  - $(400\ \mu m)^3$ voxelized mouse phantom
  - Simulated map of 18-fluorine absorbed dose

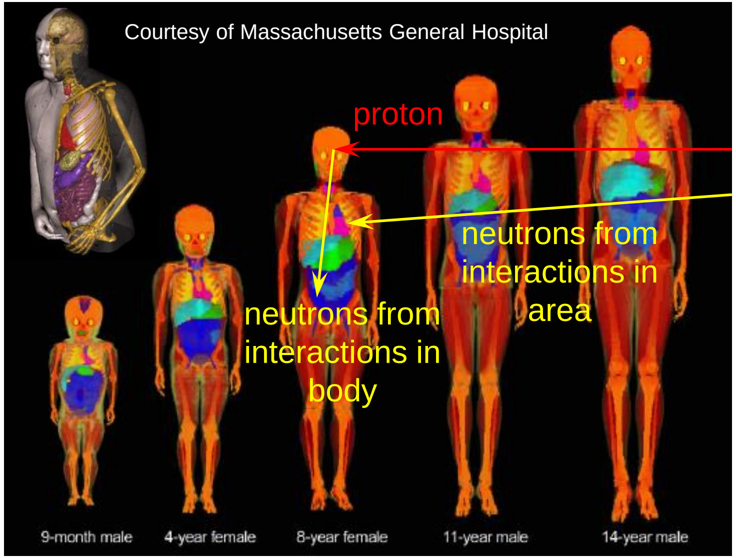




One reconstruction example, extracted from https://doi.org/10.1186/s40658-020-00309-8

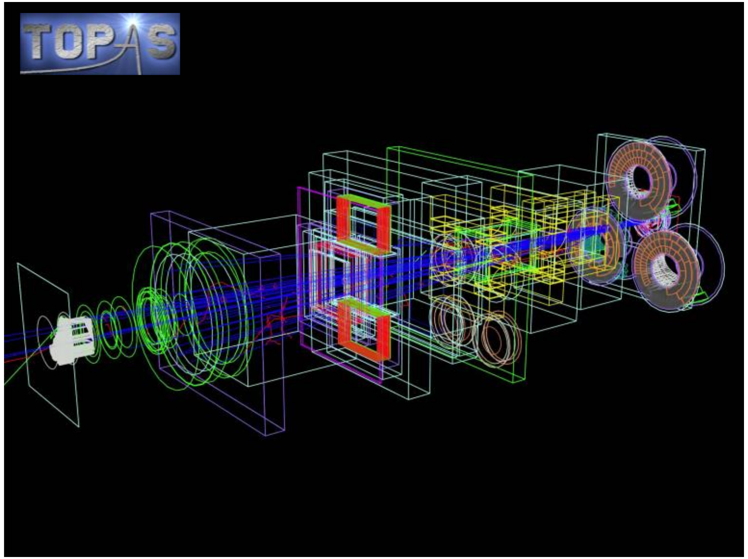
Radiation Studies, Beam Delivery System

- Therapeutic irradiations generate also undesired doses from lost particles or interactions in the body:
  - In particular neutron doses in proton irradiations



  - Ability of handling complex setups useful in estimating these doses

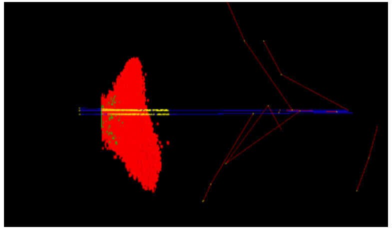
- Time dependent simulation of an IBA double-scattering beam delivery system:



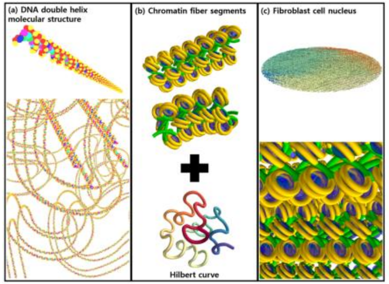
Down to DNA scale : microdosimetry

- Aims at building a **“bottom-up”** understanding of the dose
  - From **damages to the DNA**: single or double-strand break
  - Up to cells → organs → entire body
- The project **Geant4-DNA** deals with this development
  - http://geant4-dna.org/
  - Project initiated O(15) years ago by the **European Space Agency, for manned missions to Mars** !
- **Modeling includes:**
  - **Direct damages to the DNA strands**
  - **Chemistry phase:**
    - **Formation of free radicals + scattering + attack on DNA strands**
    - Chemistry phase important as responsible for **~60% of DNA damages** !
  - Involve processes **down to the O(eV) scale** !
  - **With no “condensed history”** (discussed after), hence time-consuming
- Geant4-DNA has become **very popular in the medical domain**
  - Study of particle irradiation effects, nano-particle enhancement, now used for FLASH, …

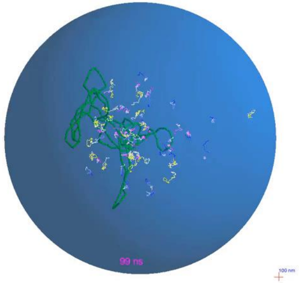
- Same approach used by **“MuElec”** project for microdosimetry in silicon.



Microbeam simulation on a single cell

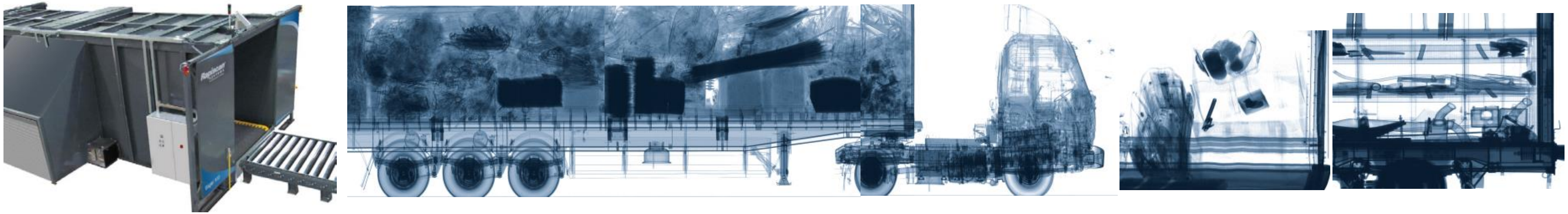


From “human cell” example

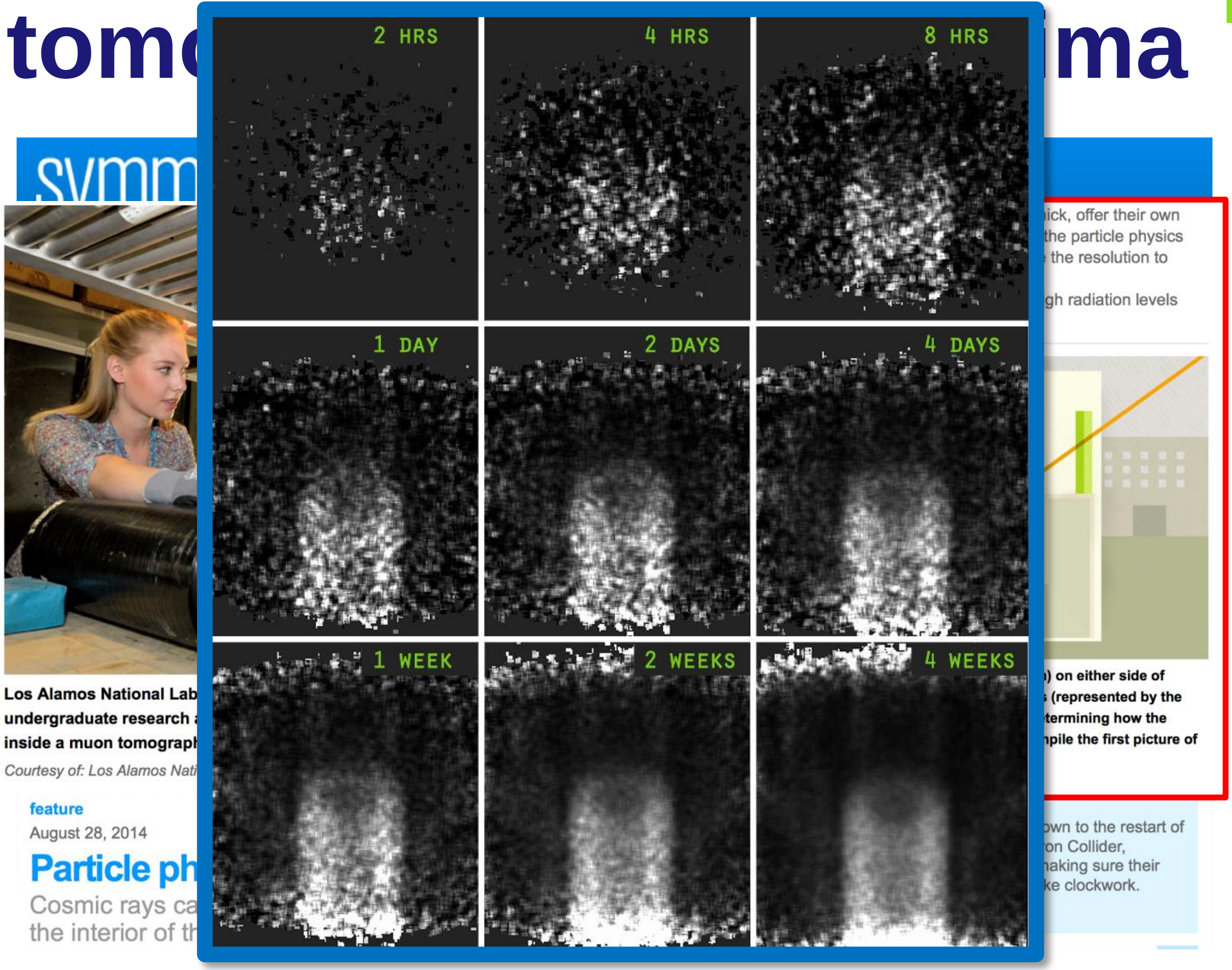


Irradiation of a pBR322 plasmid, including radiolysis - movie courtesy of V. Stepan (NPI-ASCR/LP2iB-CENBG/CNRS/IN2P3/ESA) -

Geant4 in Homeland Security : simulating X-ray cargo radiography

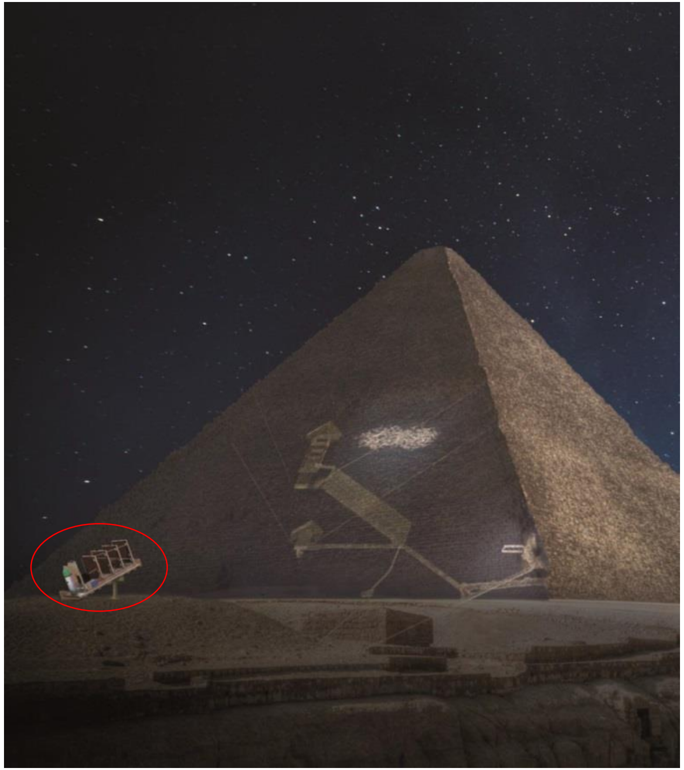


Muon tomo ma


symm
2 HRS
4 HRS
8 HRS
1 DAY
2 DAYS
4 DAYS
1 WEEK
2 WEEKS
4 WEEKS
ick, offer their own
the particle physics
the resolution to
gh radiation levels
) on either side of
(represented by the
termining how the
npile the first picture of
Los Alamos National Lab
undergraduate research
inside a muon tomograph
Courtesy of: Los Alamos Nati
feature
August 28, 2014
Particle ph
Cosmic rays ca
the interior of th
own to the restart of
on Collider,
naking sure their
ke clockwork.

Archeology

- The same **“muongraphy”** technique is used in several areas:
  - **Volcanos**
  - **But also archeology**
- Applied in the discovery of a **big void in the Great Pyramid**
- Geant4 used in the simulation of the muon detection system

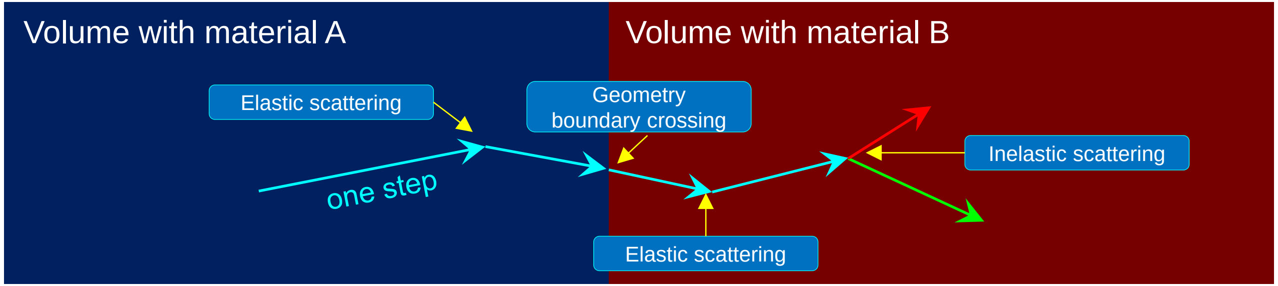


- Images : courtesy of D. Attié & S. Procureur
- Other groups are using this technique in this field:
  - Eg : arXiv:2202.07434v1 [physics.ins-det] 15 Feb 2022

Physics Modeling

Monte-Carlo Particle Transport

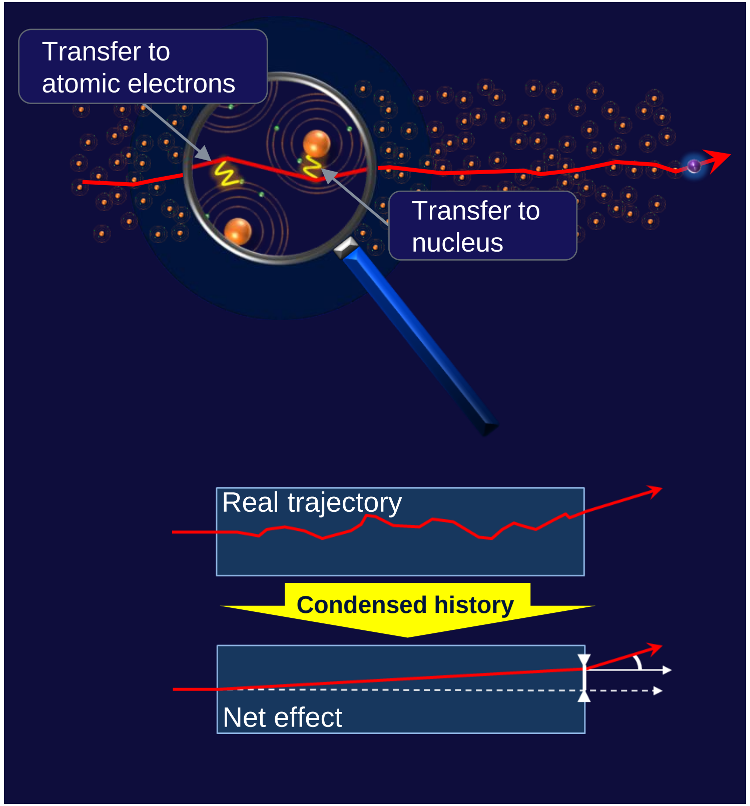
- In a Monte-Carlo particle transport code, particles are transported by **steps**:

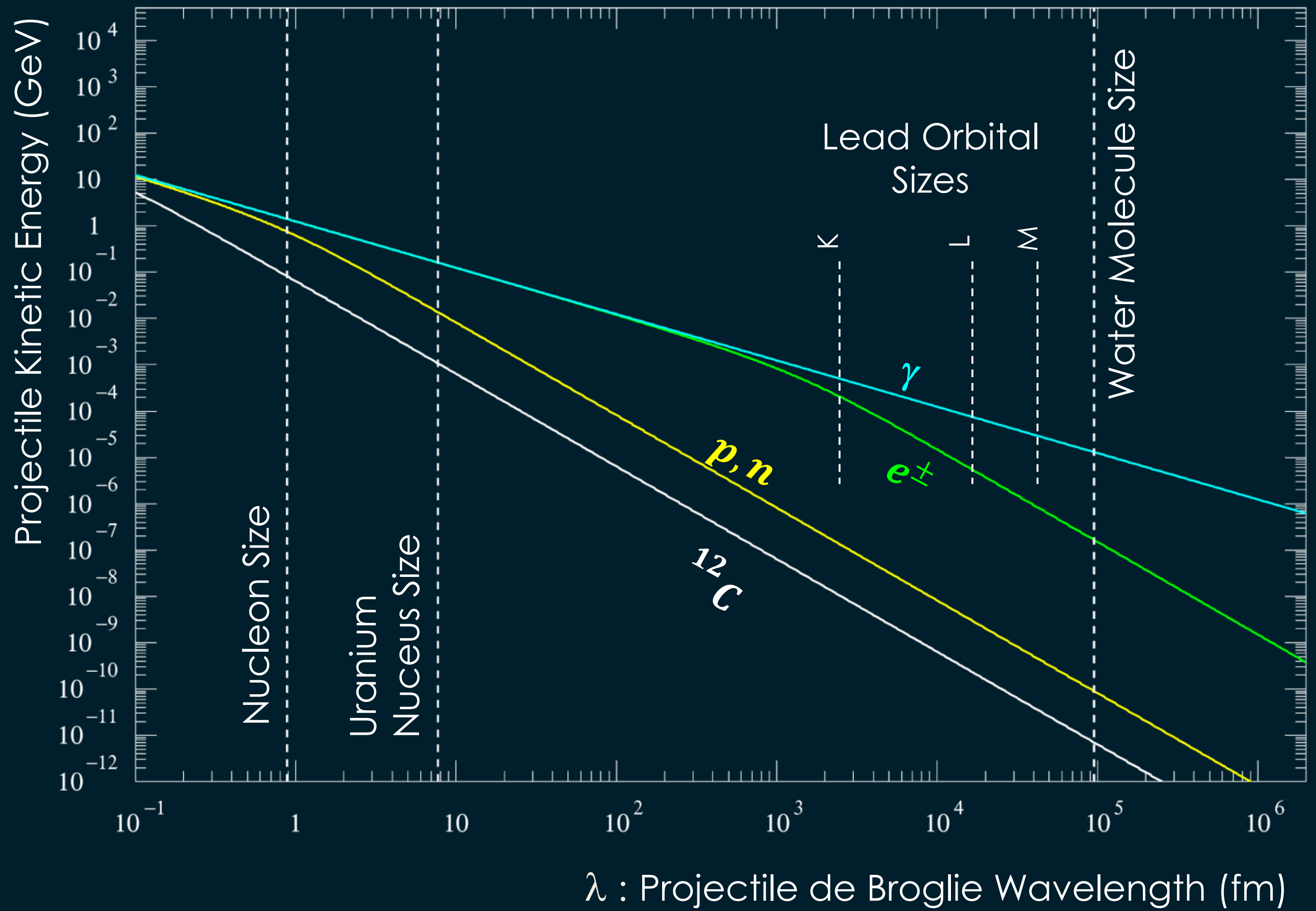


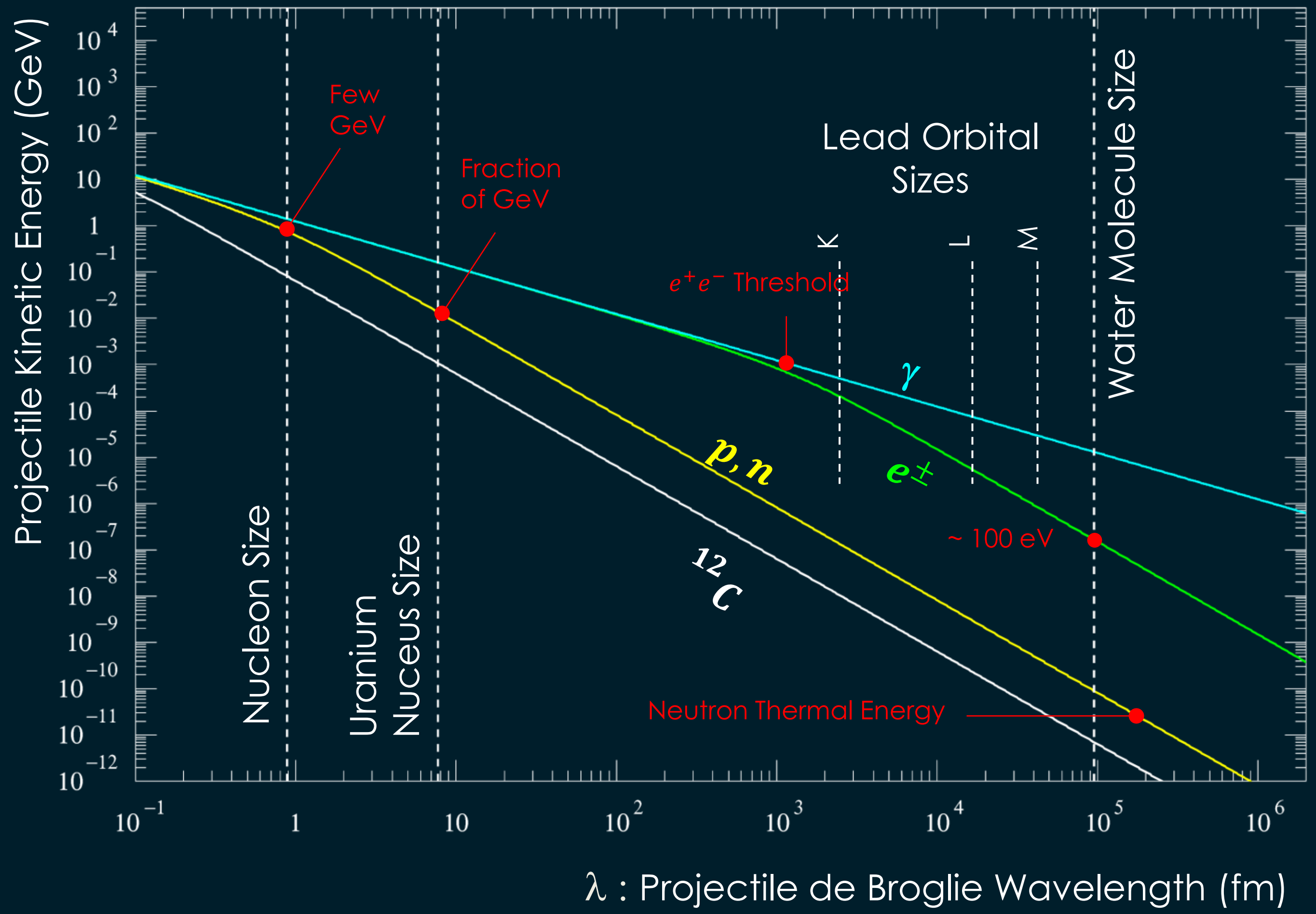
- The particle undergoes interactions by n processes i's
  - Each one being driven by a cross-section σ_i
    - Which depends on the particle type, kinetics & material
  - The cross-section σ_i determines the distance ℓ_i to next interaction for process i :
    - Which is **obtained by sampling** $\boldsymbol{p_i(\ell) = \sigma_i \cdot exp(-\ell \cdot \sigma_i)}$
- **How a step is calculated ?**
  1. Cross-sections σ_i's are calculated at the beginning of the step
  2. Distances ℓ_i are sampled accordingly
     - And the shortest distance is taken, $\ell_{win} = \min(\{\ell_i\}_{i=1\ldots n})$ → this **determines the step length**
  3. The particle is then moved to the end step point, and the **final state of process "win" is applied**
     - If this process produced daughters particles, **they are stacked, and will be tracked too**, later on

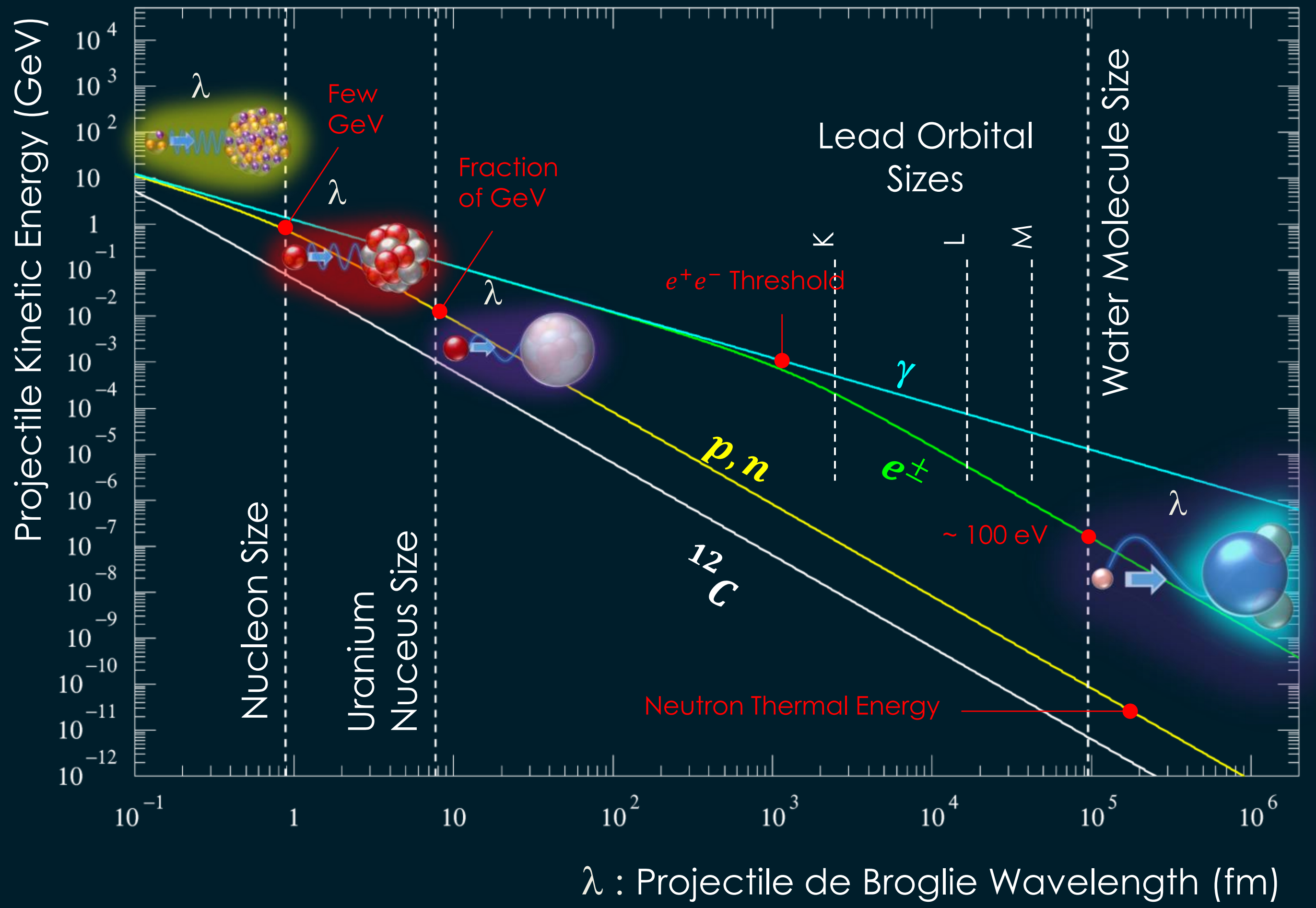
Charged Particle Transport

- The previous algorithm is general in theory
- But is **unaffordable for charged particles:**
  - Ionization generates O(10^6) interactions/mm
  - Bremsstrahlung has infrared divergence
  - → Both generate **very numerous** but **very little interactions**
    - With tiny energy particles, **which almost don't travel**
  - → ***Lot of CPU time*** needed to simulated all these !
- **Condensed History** approach:
  - **Theoretically sum-up** the numerous tiny interactions
    - To generate their **net effect in one step**
    - Tiny energy particles accounted as **"local energy deposit"** and are not created
  - User defines a **threshold** for that : **so-called "cut"**
    - When an ionization e^- –a "δ-ray"-, or a bremsstrahlung γ, is produced above the "cut", it is tracked
    - So the "cut" defines the **limit between continuous and discrete energy loss**




Projectile Kinetic Energy (GeV)
λ : Projectile de Broglie Wavelength (fm)
Nucleon Size
Uranium Nuceus Size
Lead Orbital Sizes
K
L
M
Water Molecule Size
γ
e±
p, n
12C

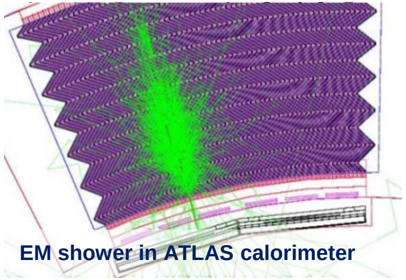

Projectile Kinetic Energy (GeV)
λ : Projectile de Broglie Wavelength (fm)
Few GeV
Fraction of GeV
Lead Orbital Sizes
K
L
M
e^+e^- Threshold
Water Molecule Size
γ
p, n
e±
~ 100 eV
12C
Neutron Thermal Energy
Nucleon Size
Uranium Nuceus Size
10^4
10^3
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-1}
1
10
10^2
10^3
10^4
10^5
10^6



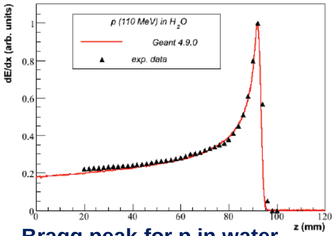
Projectile Kinetic Energy (GeV)
λ : Projectile de Broglie Wavelength (fm)
Few GeV
Fraction of GeV
Lead Orbital Sizes
K
L
M
Water Molecule Size
Nucleon Size
Uranium Nuceus Size
e^+e^- Threshold
γ
p, n
e±
~ 100 eV
^{12}C
Neutron Thermal Energy
λ

Electromagnetic Physics

- **“Standard” Electromagnetic:**
  - Energy range 1 keV – O(100 TeV)
  - Processes for e-, e+, γ
  - Charged hadrons ionization up to 100 TeV
- **Muon, up to PeV**
- **“Low energy” Electromagnetic:**
  - More precise description:
    - PENELOPE 2008 reimplementation
    - LIVERMORE data for cross-sections and final states
    - Energy range down to ~250 eV / ~100 eV
  - Charged hadron ionization
    - ICRU’ 49 & 73 & 90, NIST
  - Material relaxation (PIXE, Auger e-, …)
- **DNA & MuElec:**
  - For microdosimetry studies in DNA and Silicon
  - Processes down to a few eV
  - Chemistry stage for DNA
    - Water radical scattering
- **Optical photon: long wavelength γ** (X-ray, UV, visible)
  - Reflection, refraction, absorption, wavelength shifts, Rayleigh
- **Phonons:**
  - Suited for very low-temperature detectors (tens of mK)

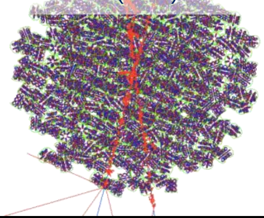


EM shower in ATLAS calorimeter

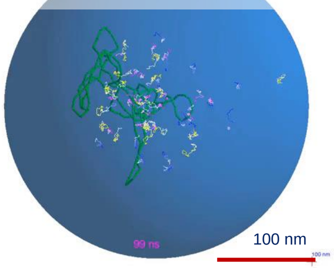


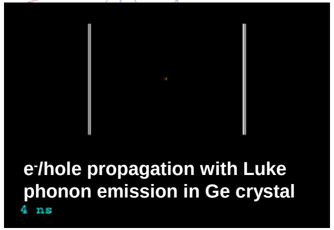
Bragg peak for p in water

Cell nucleus (15 µm diameter) with 6×10^9 base pairs of DNA
NIM B 306 (2013) 158-164

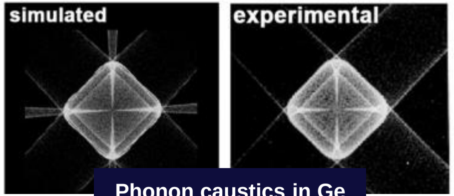


pBR322 plasmid irradiation, including radiolysis



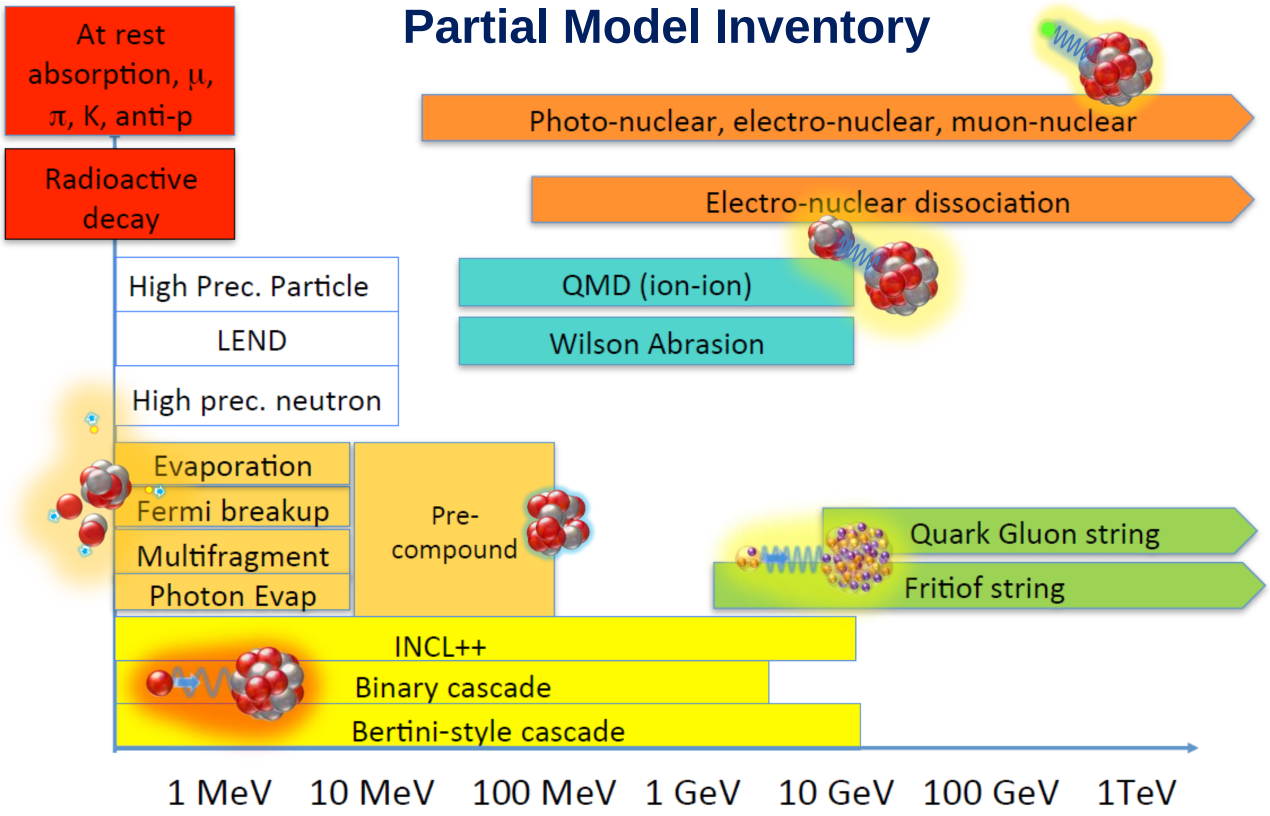


e-/hole propagation with Luke phonon emission in Ge crystal



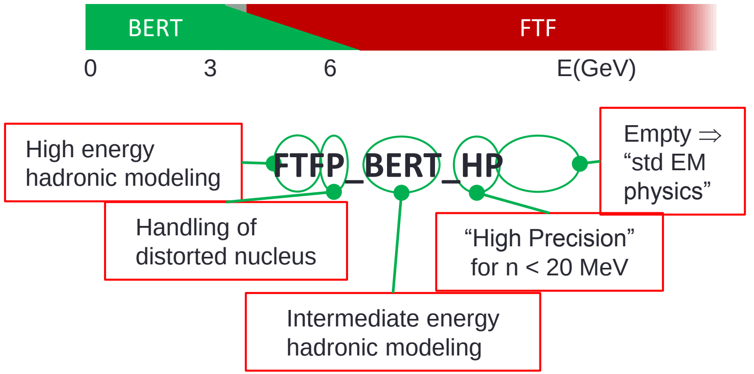
Phonon caustics in Ge

Hadronic Physics



The Physics List Concept

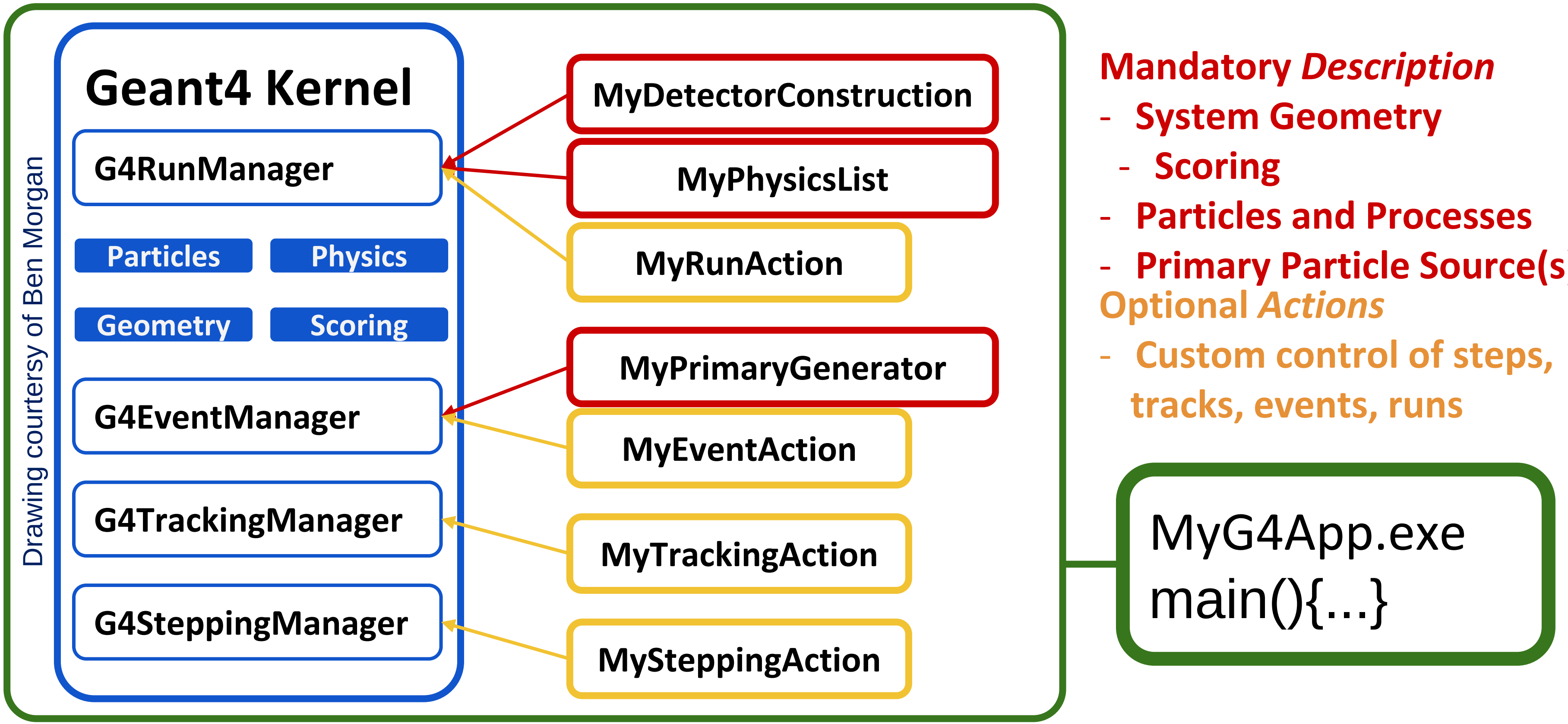
- **There are <u>many</u> physics models in Geant4 !**
  - electromagnetic & hadronic, but also radioactive decay, options for low energy neutrons, low energy electromagnetic, etc.) available in Geant4
  - plus some options like fast simulation, variance reduction (not discussed today)
- Some physics models are:
  - **complementary** (valid on ≠ energy domains)
  - **competitive** (valid on the same energy domain)
- A "physics process" –eg "hadron inelastic" – is **often a composed of several models**
  - **Each model serving one energy domain**
- **The "physics list" concept aims gathering consistent set of "physics processes"**
  - And to configure their underneath "models", parameters, etc
- Geant4 provides **"ready to go"** physics lists, meant to **respond to different use-cases**, eg:
  - **High Energy Physics**
    - With for example LPM effect activated, by no details on atomic structure
  - **Medical**
    - With accurate description of Bragg peak
  - **DNA**
    - With ultra-low energy processes activated, but no precision on high energy side
- **These are continuously monitored**
- They can served as a **basis for more specialized physics**



Example of FTFP_BERT_HP physics list

Software Aspects

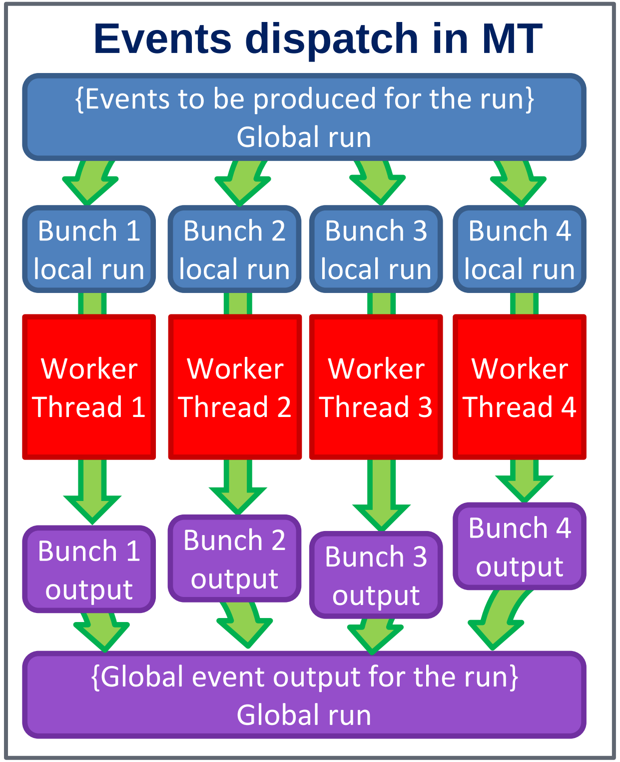
Geant4 Main Components

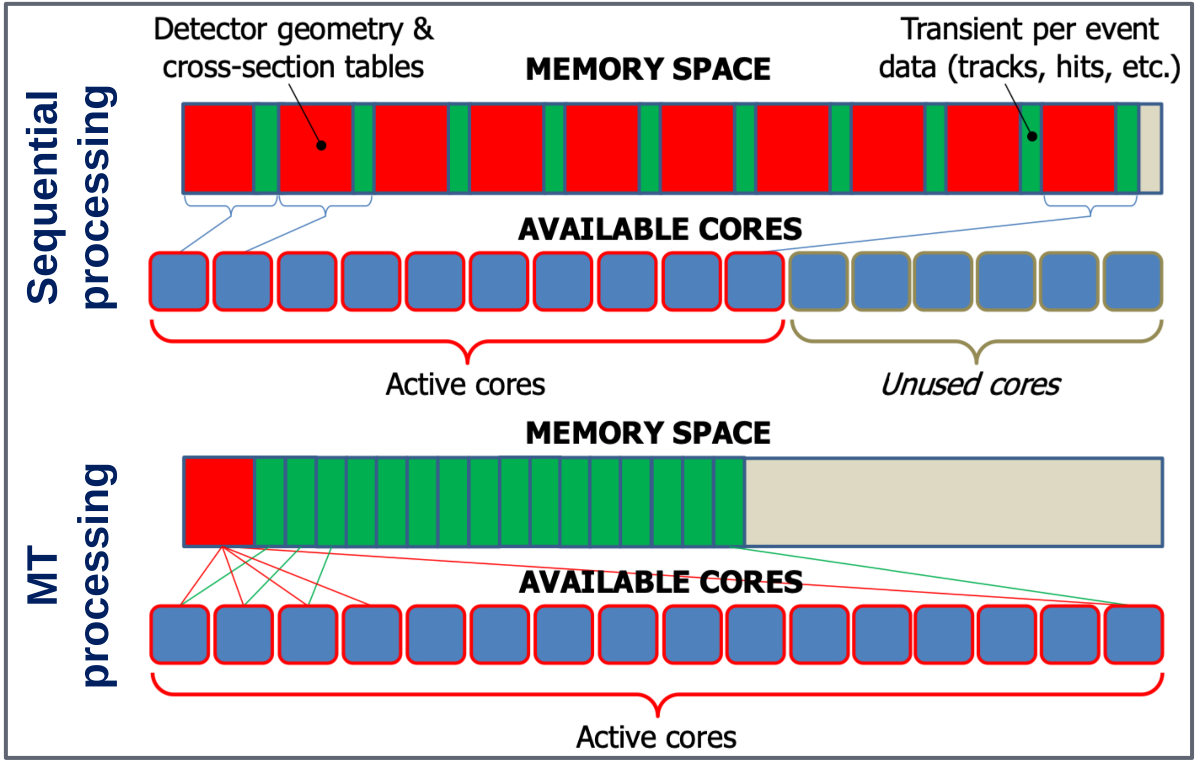


Mandatory *Description*
- System Geometry
  - Scoring
- Particles and Processes
- Primary Particle Source(s)

Optional *Actions*
- Custom control of steps, tracks, events, runs

Multithreading : Efficient Resource Usage



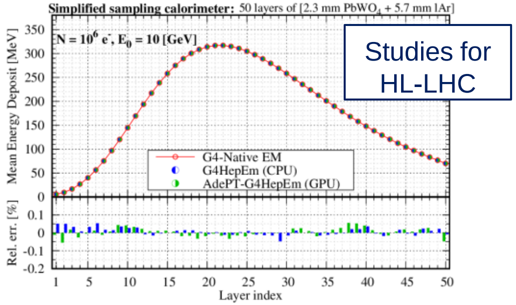


- **MT (introduced in Geant4 10.0, in 2013) resolved the "embarrassingly problem" of Geant4**
- Scheme evolved to "tasking" with 11.0 : more flexible and easier bridge to hybrid computing

Coming Challenges

Human & Technological Challenges

- **Geant4 is next to 30 years old !**
  - Software is still improving…
  - … but "historical" members got older by 30 years !
  - **Not enough new blood** coming & **staying** in Geant4 !
    - And fixed term contract policy is a ***disaster*** for such long terms projects…
  - **New and stable generation of developers is absolutely needed !**
- **Mutation to parallelism(s)**
  - Computing power increasing but **under "parallel" technologies**
    - Many core, SIMD, GPU, etc.
  - **Geant4 resolved the "embarrassingly parallel problem"**
    - With **"event-level parallelism"** (MT and now Tasking)
  - **"Track-level parallelism"** is hoped to improve **compute intensity**
    - By **sharing in one go calculations among several tracks**
      - Using of fine grained parallelism technologies : SIMD, GPU
  - But **high energy Monte-Carlo particle transport code is an evil case** for these technologies !
    - As it leads to **lots of "thread divergences"**
    - That **break the needed "regular" processing flow** to attain high efficiency
  - Ongoing R&Ds to see where we could land…



G4HepEm AdePT

EM shower profile on GPU
NVIDIA card GeForce RTX
Mihaly Novak & Jonas Hahnfeld (CERN)

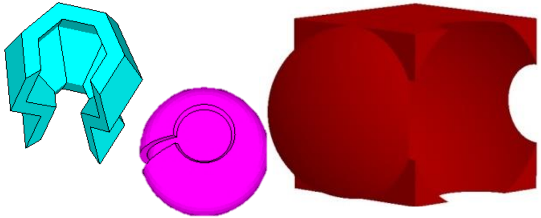
Conclusion

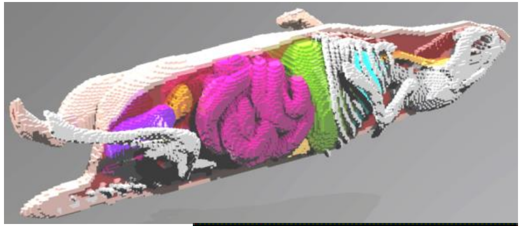
- Geant4 is a **general purpose particle transport Monte Carlo toolkit with wide physics coverage** and numerous application domains
  - Geant4 was initiated in 1994, 29 years ago, and it is still evolving
  - It is a **free** and **Open Software**
  - It is available from https://geant4.web.cern.ch/
  - Its OO structure allowed it to **accommodate many use cases** & hence serve many domains
- Geant4 is also the **name of the Collaboration** in charge of the toolkit
  - About 130 members, for ~30 FTE, distributed worldwide
  - They maintain and develop Geant4 with a **distributed model**
- Despite a long series of successes, Geant4 has **still great challenges to cope with** !
  - In particular, the strong HL-LHC demand poses acute challenges !
  - With **much higher speed requested**, **together with better physics quality**
    - To not increase systematic errors from simulation compared to statistical ones
- But the biggest challenge is a human one : **a new and stable generation of developers is desperately needed !**

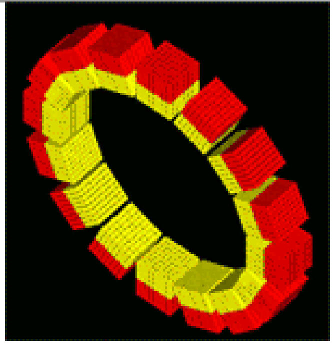
BACKUP

Key geometry capabilities

- Richest collection of shapes
  - CSG (Constructed Solid Geometry), Boolean operation, Tessellated solid, etc.
  - The user can extend
- Describing a setup as hierarchy or ‘flat’ structure
  - Describing setups up to billions of volumes
  - Tools for creating & checking complex structures
  - Interface to CAD
- Navigating fast in complex geometry model
  - Automatic optimization
  - By subdivision of geometry in “voxels” containing a few volumes, with fast navigation between neighbor voxels
- Geometry models can be ‘dynamic’
  - Changing the setup at run-time
    - e.g. “moving objects”





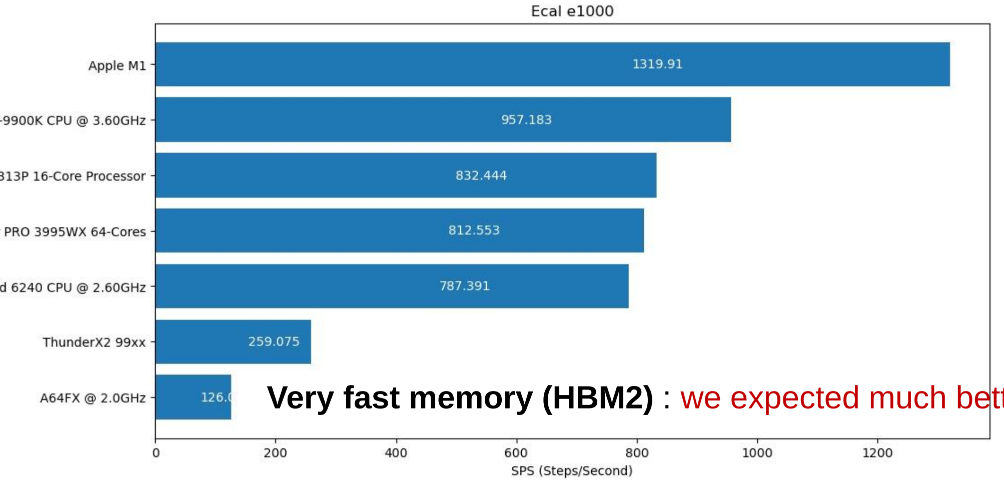


Investigation : hardware benchmark

- GPUs are being largely spread as accelerators in hybrid computing
- But important to **evaluate other hardware solutions**

Apple M1
(not for HPC !)

Chip equipping the Fugaku 415-PFLOPS center



Very fast memory (HBM2) : we expected much better !

Courtesy of Koichi Murakami (KEK)

- **Investigations will continue to better understand the spread in performances !**
- Other hardware will be tested, when available

Can Geant4 run on GPU ?

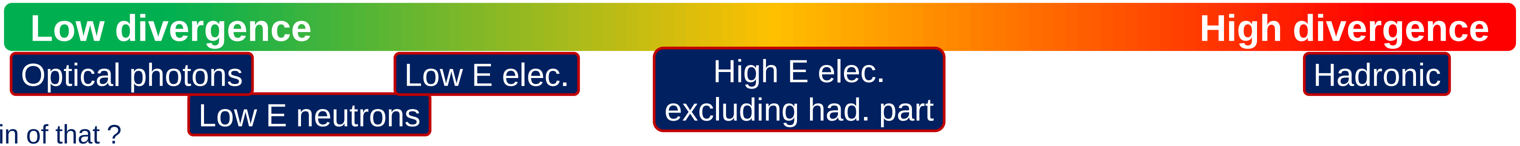
- We have often the question “can Geant4 run on GPU ?”
  - Underlying hope : GPUs are fast, so running Geant4 on GPU would be fast !
  - Not that simple...

- GPU are fantastic to treat **« many very similar things » « behaving almost the same »**
  - Typical example and original motivation : optical photons
  - The treatment can be done in parallel, applying the same calculation to a set of data.
    - And this can be repeated calculation after calculation if the set of data is not destroyed by these calculations.
    - In other words, **no divergences** appear in the data set : the data set remains of « the same nature ».
  - GPUs are designed to make these parallel calculations efficiently, and they are performing nicely !

- But with a Monte Carlo like Geant4:
  - **« many very similar things » → « many very different things » !**
    - Many type of particles !
  - **« behaving almost the same » → « behaving not at all the same » !**
    - Interactions of particles are very different from on type to an other
    - Even particles of same type can undergo very different interactions !

  } → **Source of plenty of divergences !**

- Usage of GPU limited *a priori* to some « sectors », strongly linked to their divergences:



- Net gain of that ?
  - Great for medical applications (demonstrated) : low E elec. in simple geometries.
  - But for HEP and complicated geometries: **ongoing R&D**, first responses expected in a time scale of one year.

Scopus As of 02/2020

Search Sources Lists SciVal

Create account

12,644 documents have cited:

GEANT4 - A simulation toolkit

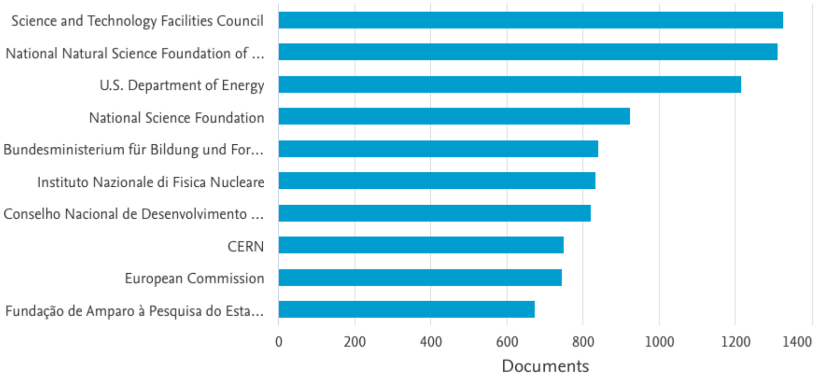
Agostinelli S., Allison J., Amako K., Apostolakis J., Araujo H., Arce P., Asai M., (...), Zschiesche D.

(2003) Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506 (3) , pp. 250-303.

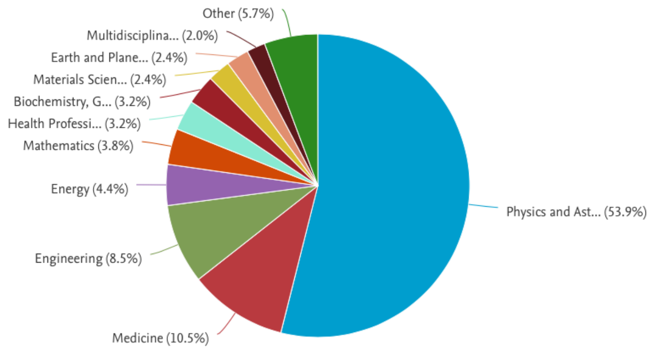
Set feed

Documents by funding sponsor

Compare the document counts for up to 15 funding sponsors.



Documents by subject area

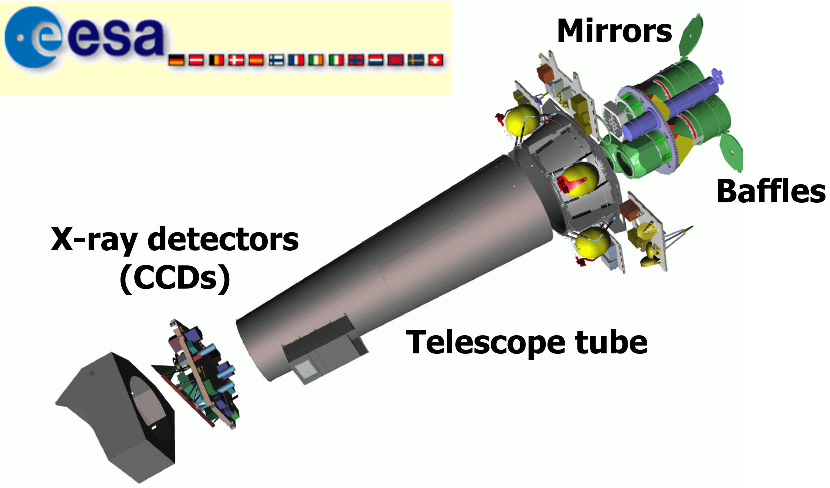


Geant4 - A simulation toolkit

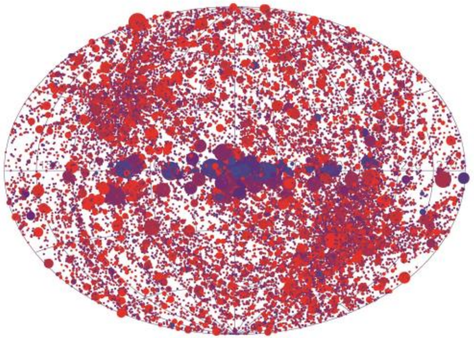
NIM A, vol 506(3), pp250-303, 2003

Significant use across **many** research areas, **considered mission critical for HEP**

XMM Newtown telescope : the start of Geant4 in Space Science



Artist view of XMM

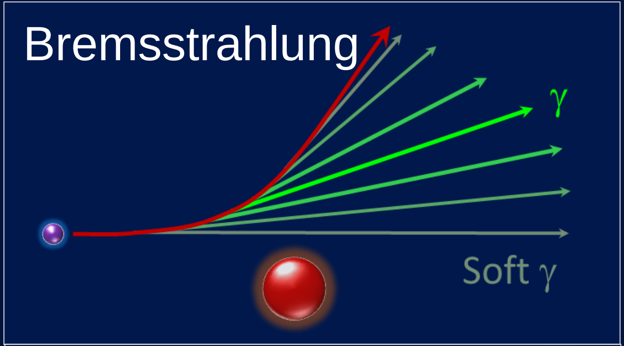


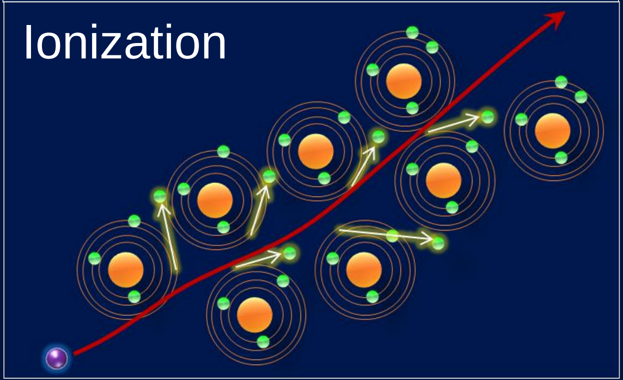
Map of X-ray sources by XMM

- Launch December 1999
  - Expected end of mission December 2025
- Highly elliptic orbit:
  - Perigee 7000 km
  - Apogee 114000 km
- **Flight through the Van Hallen radiation belts**

- Chandra X-ray observatory, with similar orbit, experienced unexpected degradation of CCDs
- Geant4 helped to understand:
  - X-ray telescopes also focus protons very well…
  - which damage the CCD cameras !
- **XMM could take counter-action !**

(Quasi-)Diverging Cross-Sections



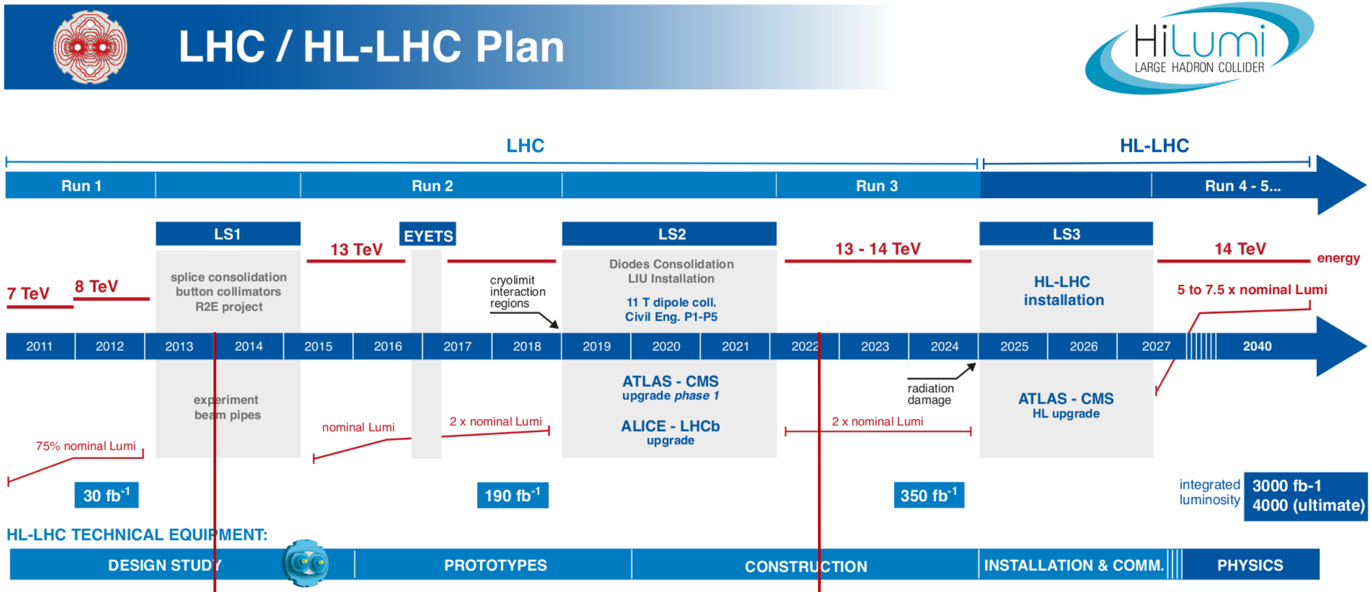


- (Quasi-)diverging cross-sections for:
  - **Bremsstrahlung,** which has a diverging cross-section
    - Produces an infinity of soft γ
      - Which don't fly far away…
    - (Well, some dielectric effect limit the divergence in reality)
  - **Ionization,** which can lead to huge amount of soft e^-
    - Which can't travel far away…

- Under a **"given precision"** these very low energy particles contribute to the **"local energy deposit"**

- Sounds like a specialist issue ?
  - It isn't ! You must specify this "given precision"
  - Which is a "production threshold"
  - Badly named as "cut" in common language
- In Geant4, production thresholds are expressed in range
  - Particles unable to travel beyond "x" distance are not produced
- Typically range cut must be ~ size of smallest volume
  - /run/setCuts 1 mm

HL-LHC & Simulation Time Scales



Geant4 MT, merged to production since 2014

Geant-V R&D, stopped end 2019

TODAY

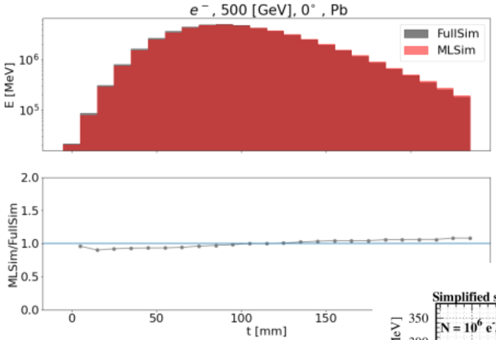
Remaining time for evolution

Three main lines of development

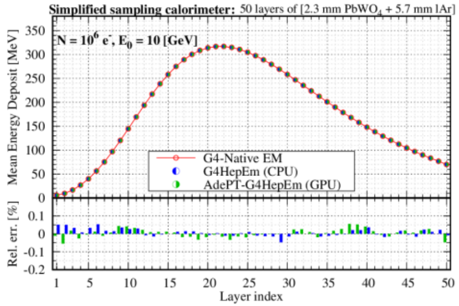
- **Continue with adiabatic improvements of detailed simulation**
  - Technical improvements
    - Code revision, modern standards, more use of acceleration techniques, etc.
  - **<u>And physics precision improvements</u>**
    - If speed-up by factor f, need to improve physics precision by factor $\sqrt{f}$!
  - It is not excluded (but not promised !) that a factor $f \sim 2$ is reachable
    - Generally, few percent gain in speed at each release

- **Improve fast simulation**
  - Classical parametric approaches
    - Well known GFlash for EM showers
  - But also modern techniques based on Machine Learning
    - Might be an approach for hadronic showers and advanced models
  - Important : **fast simulation may be the security net of that era !**
  - And also, explore wider use of biasing techniques

- **R&D accelerators and new architectures**
  - Large spread of GPUs :
    - **High reward @ <u>high risk</u>**
    - Approach : articulate GPU – CPU in an efficient flow
  - **Benchmarks on new architectures**



EM shower AI-based
Dalila Salami & Anna Zaborowska (CERN, SFT)
Variational Auto-Encoders (VAE)



EM shower profile on GPU
NVIDIA card GeForce RTX
Mihaly Novak & Jonas Hahnfeld (CERN, SFT) in the context of the AdePT project

Process	Cross-section → Mean free path	Sample distance to next interaction → min one = **step length**
Process-1	$\sigma_1 \rightarrow \lambda_1 = 1/\sigma_1$	$\ell_1 : \exp(-\ell_1/\lambda_1)/\lambda_1$
Process-2	$\sigma_2 \rightarrow \lambda_2 = 1/\sigma_2$	$\ell_2 : \exp(-\ell_2/\lambda_2)/\lambda_2$
Process-3	$\sigma_3 \rightarrow \lambda_3 = 1/\sigma_3$	$\ell_3 : \exp(-\ell_3/\lambda_3)/\lambda_3$
Geometry	/	$\ell_{geometry}$