# Montecarlo & GEANT

**Salvatore Fazio**
*Università della Calabria & INFN Cosenza*

**Second European School on the Physics of the Electron-Ion Collider**
**Benicasim, Valencia - June 2025**


# A bit of
# ePIC propaganda

# ePIC- a Collaboration driven by the science!

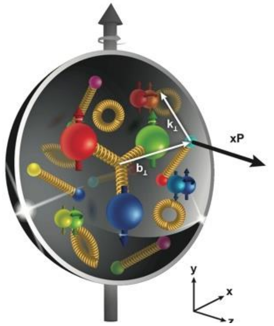
The EIC will unravel the different contribution from the quarks, gluons and orbital angular momentum

SPIN is one of the fundamental properties of matter.
All elementary particles, but the Higgs carry spin.
Spin cannot be explained by a static picture of the proton
It is the interplay between the intrinsic properties and interactions of quarks and gluons

Does the mass of visible matter emerge from quark-gluon interactions?

Atom: Binding/Mass = 0.00000001
Nucleus: Binding/Mass = 0.01
Proton: Binding/Mass = 100

For the proton the EIC will determine an important term contributing to the proton mass, the so-called "QCD trace anomaly



How can we understand their dynamical origin in QCD?
What is the relation to Confinement

How are the quarks and gluon distributed in space and momentum inside the nucleon & nuclei?

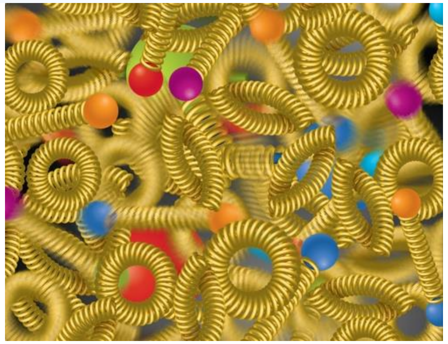
How do the nucleon properties emerge from them and their interactions?

How do the confined hadronic states emerge from quarks and gluons?

Is the structure of a free and bound nucleon the same?

How do quarks and gluons, interact with a nuclear medium?

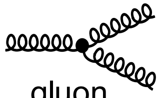
How do the quark-gluon interactions create nuclear binding?



What happens to the gluon density in nuclei? Does it saturate at high energy?

How many gluons can fit in a proton?

How does a dense nuclear environment affect the quarks and gluons, their correlations, and their interactions?



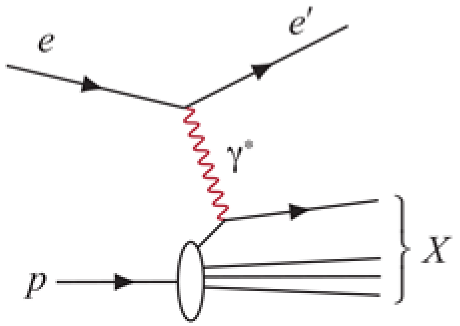
gluon splitting

? =

gluon recombination

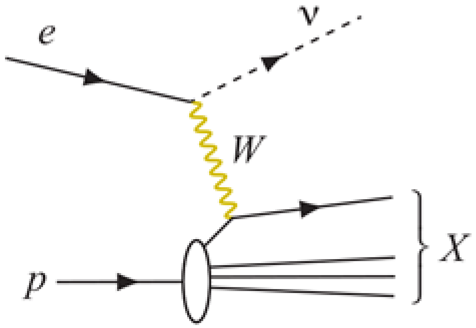
# What process must be measured?

DIS event kinematics - scattered electron or final state particles (CC DIS, low y)



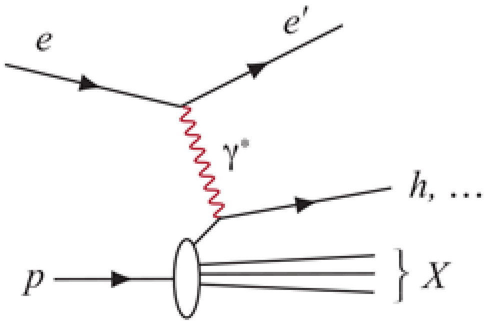

**Neutral Current DIS**

- Detection of scattered electron with high precision - event kinematics



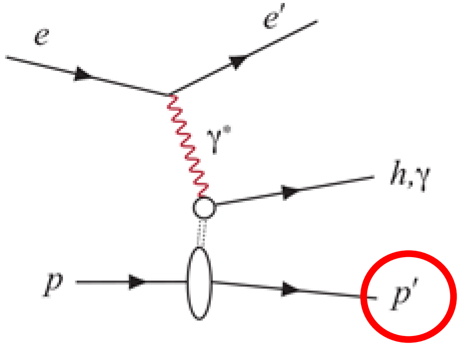

**Charged Current DIS**

- Event kinematics from the final state particles (Jacquet-Blondel method)



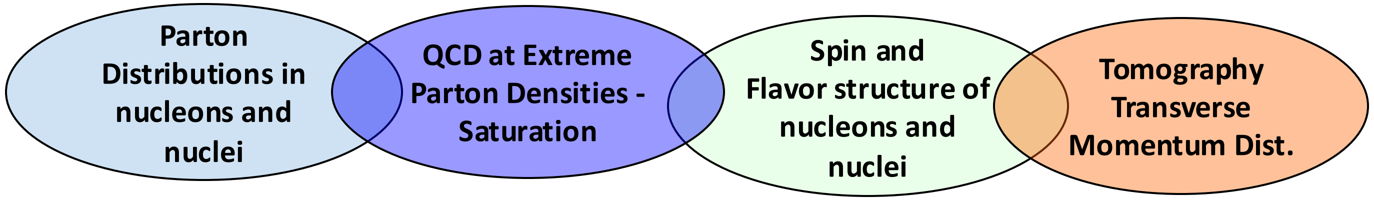

**Semi-Inclusive DIS**

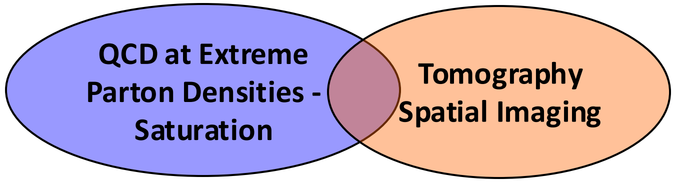
- Precise detection of scattered electron in coincidence with at least 1 hadron




**Exclusive Processes**

- Detection of all particles in event




$\int \mathcal{L}dt$: $\sim 1\ \mathrm{fb}^{-1}$ $\sim 10\ \mathrm{fb}^{-1}$ $\sim 100\ \mathrm{fb}^{-1}$

# Structure of the ePIC Physics Working Groups

**PHYSICS ANALYSIS COORDINATORS**
Salvatore Fazio (Cosenza) - Rachel Montgomery (Glasgow)
Rosi Reed (Lehigh) - deputy

- Each PWG convener is for a two-years term staggered
- Conveners in blue are ending their term in July 2025
- PWGs typically meet by-weekly

**INCLUSIVE PHYSICS**
Tyler Kutz (MIT)
Stephen Maple (Birmingham)

Mailing list: eic-projdet-Inclusive-l@lists.bnl.gov
Indico: https://indico.bnl.gov/category/417/

**SEMI-INCLUSIVE PHYSICS**
Stefan Diehl (UConn)
Ralf Seidl (RIKEN)

Mailing list: eic-projdet-semiincl-l@lists.bnl.gov
Indico: https://indico.bnl.gov/category/418/

**JETS AND HEAVY FLAVOR**
Olga Evdokimov (UIC)
Rongrong Ma (BNL)

Mailing list: eic-projdet-jethf-l@lists.bnl.gov
Indico: https://indico.bnl.gov/category/420/

**EXCLUSIVE, DIFFRACTION AND TAGGING**
Raphael Dupre (Orsay)
Zhoudunming Tu (BNL)

Mailing list: eic-projdet-excldiff-l@lists.bnl.gov
Indico: https://indico.bnl.gov/category/419/

**BSM AND PRECISION EW**
Ciprian Gal (JLab)
Juliette Memmei (Manitoba)

Mailing list: eic-projdet-semiincl-l@lists.bnl.gov
Indico: https://indico.bnl.gov/category/421/

# ePIC Technical Design Report

**pre–TDR (60% design completion) ⟹ December 2025**

**TDR (90% design completion) ⟹ ~ late 2026**

- (pre)TDR are a deliverable of the EIC Project (project manager acts as editor)
  - describe the accelerator + ePIC experiment
  - Chapter 8: (hundreds pages) focus on the ePIC Detector Description, basic performance, Software, and data preservation
  - Chapter 2: (~60 pager) focus on holistic detector performance, physics performance and science reach
    - Holistic detector performance → Technical Coordinator office acts as editor
    - Physics and science reach → Analysis Coordinators act as editors
    - We envision the **publication of a spin—off physics performance paper(s)**

# EIC Early Science Report

- **Previous workshops:**
  - **Sep. 13, 2024:** https://indico.bnl.gov/event/24432/
  - **Jan. 2025, plenary at Coll. Meeting:** https://agenda.infn.it/event/43344/timetable/#20250122.detailed
  - **Apr. 2025, CFNS @ Stony Brook:** https://indico.cfnssbu.physics.sunysb.edu/event/410/overview
- **Goal of this exercise:**
  - **Highlight *meaningful* and *impactful* science within early years of running without undermining the importance of achieving full EIC capabilities**
  - ***Meaningful***: The EIC early science program must engage the collaboration; it must get the collaboration excited about working hard for the future. It must have a balance of *breadth* and *depth*
  - ***Impactful***: The EIC early science program must take the first steps down the path to realizing the EIC science goals
- **Deliverable:**
  - **Report to be published by the ePIC Collaboration (~mid 2026)**

# ePIC science document for the inputs to the EPPSU 2025

March 28, 2025

## Synergies between a U.S.-based Electron-Ion Collider and European Research in Particle Physics

Contact Persons: Stefan Diehl[1], Raphaël Dupré[2], Olga Evdokimov[3], Salvatore Fazio[4], Ciprian Gal[5], Tyler Kutz[6], Rongrong Ma[7], Juliette Mammei[8], Stephen Maple[9], Marco Radici[10], Rosi Reed[11], Ralf Seidl[12], Zhoudunming Tu[13]

On behalf of the ePIC Collaboration and the EIC User Group

**Abstract**

This document is submitted as input to the European Strategy for Particle Physics Update (ESPPU). The U.S.-based Electron-Ion Collider (EIC) aims at understanding how the complex dynamics of confined quarks and gluons makes up nucleons, nuclei and all visible matter, and determines their macroscopic properties. In April 2024, the EIC project received approval for critical-decision 3A (CD-3A) allowing for Long-Lead Procurement, bringing its realization another step closer. The ePIC Collaboration was established in July 2022 around the realization of a general purpose detector at the EIC. The EIC is based in U.S.A. but is characterized as a genuine international project. In fact, a large group of European scientists is already involved in the EIC community: currently, about a quarter of the EIC User Group (consisting of over 1500 scientists) and 29% of the ePIC Collaboration (consisting of ~1000 members) is based in Europe. This European involvement is not only an important driver of the EIC, but can also be beneficial to a number of related ongoing and planned particle physics experiments at CERN. In this document, the connections between the scientific questions addressed at CERN and at the EIC are outlined. The aim is to highlight how the many synergies between the CERN Particle Physics research and the EIC project will foster progress at the forefront of collider physics.

- **The Science document submitted to the European Strategy for Particle Physics panel**
  - arXiv:2504.01236
  - **ZENODO:** https://zenodo.org/records/15102075
  - **Jointly submitted by EICUG and ePIC**
  - **Key Italian contribution!**

# Opportunities!

✓ The EIC provides an unprecedented opportunity for the ultimate understanding of QCD


## New excitement ahead

- Event reconstruction being finalized & novel analysis tools being developed
  - New, more realistic, impact studies
  - TDR has a chapter on physics studies
  - Report on Early Science
- It is NOW the right time to join the efforts and get involved!

  **step 1:** email the conveners of your favorite PWG and **subscribe the mailing list**

  **step 2:** join the **(bi)weeky meetings**

  **step 3:** **actively engage** in studies and efforts


And now…
back to business!

# Overview and credits

- Basics on Monte Carlo methods
- Basics on the Geant4 toolkit
- Hands on exercise: simulate a shower profile in a e.m. calorimeter

- Besides my favourite and most reliable source (wikipedia), these slides and based/inspired by the work of many, including:
  - M. Asai (SLAC, Stanford)
  - A. Dotti (SLAC, Stanford)
  - L. Pandola (INFN-LNS, Catania)
  - H. Jung (DESY, U. of Antwerp)
  - Carlo Mancini Terracciano (U. of Rome La Sapienza)
  - E. Tassi (U. of Calabria & INFN Cosenza)

# Monte Carlo, Generators, GEANT4… simulations

## Discussion…

- What are them?
- Are they the same thing?

**GENERATORS**


**MONTE CARLO**

**GEANT**

# Monte Carlo methods

- A mathematical approach using a sequence of **random numbers** to solve a problem
- Use **stochastic quantities** (random variables) – e.g. the outcome of rolling a dice, the price of a stock, or the time until a radioactive atom decays, average values of mm of rain…
- to solve **Deterministic problems** (definite integral)
- **Generate** N **random "*points*"**, $\vec{x}_i$, in the problem space
- Calculate:
  - $\langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(\vec{x}_i)$
  - $\langle f^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} f^2(\vec{x}_i)$
- The name "Monte Carlo" was coined by Nicholas Metropolis, due to the method's reliance on randomness, like gambling in a casino

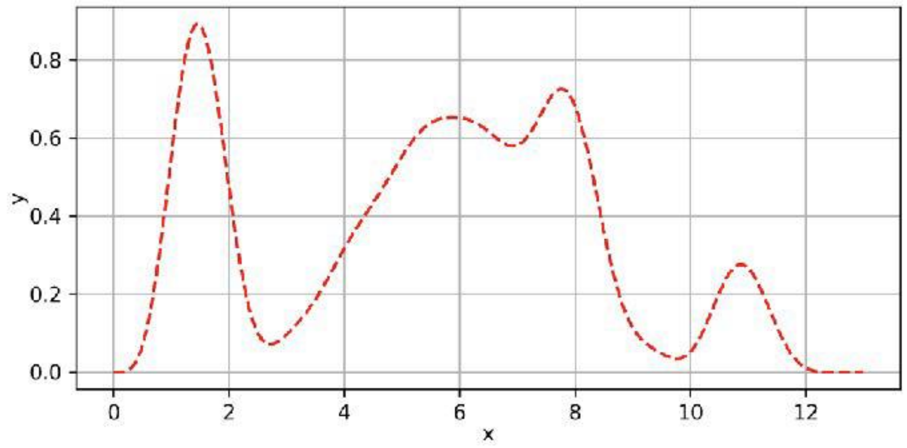
# How to calculate an integral

- Randomly choose couple of numbers $(\vec{x}_i, \vec{y}_i)$ from the range and the domain, respectively, of the function $f$
- The **acceptance-rejection method** proposed by Von Neuman:
  - Accept the fraction of points $\vec{y}_i \leq f(\vec{x}_i)$ in the part of the area below the function
  - This technique is used to generate random numbers for an arbitrary **Probability Density Function** (PDF)

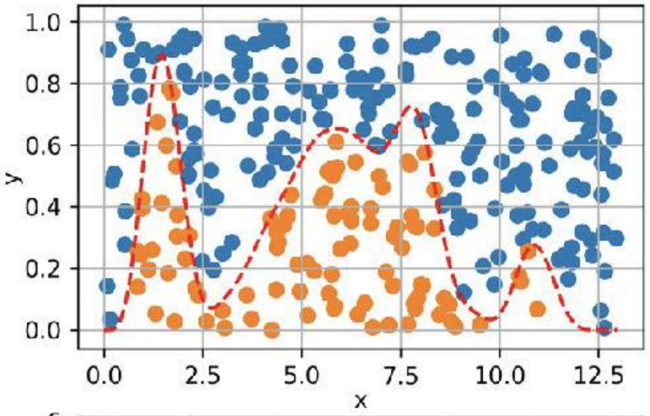
# Example of an integral calculation

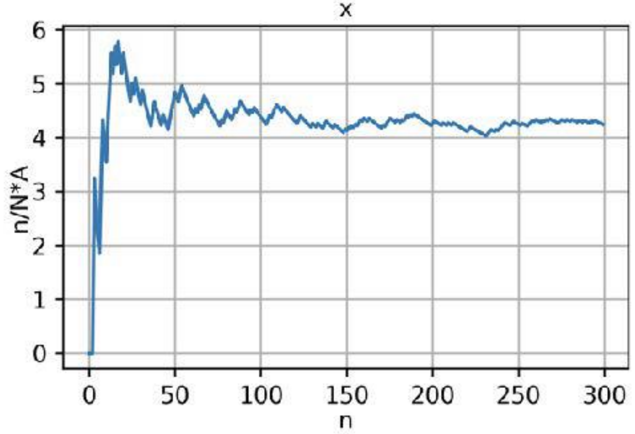
- Imagine you have to calculate the integral of a complicated function like:

$$f(x) = 1.4\left[\sin^4(x)\cos^2\left(\frac{x}{3}\right) + \sin^6\left(\frac{x}{4}\right)\right]e^{-\frac{x}{8}}$$




- The orange points are the accepted = **n**
- The blue are the rejected
- **N** = accepted + rejected



# Example: calculate $\pi$

Let's calculate π

$$f(x) = \sqrt{(1-x^2)}$$

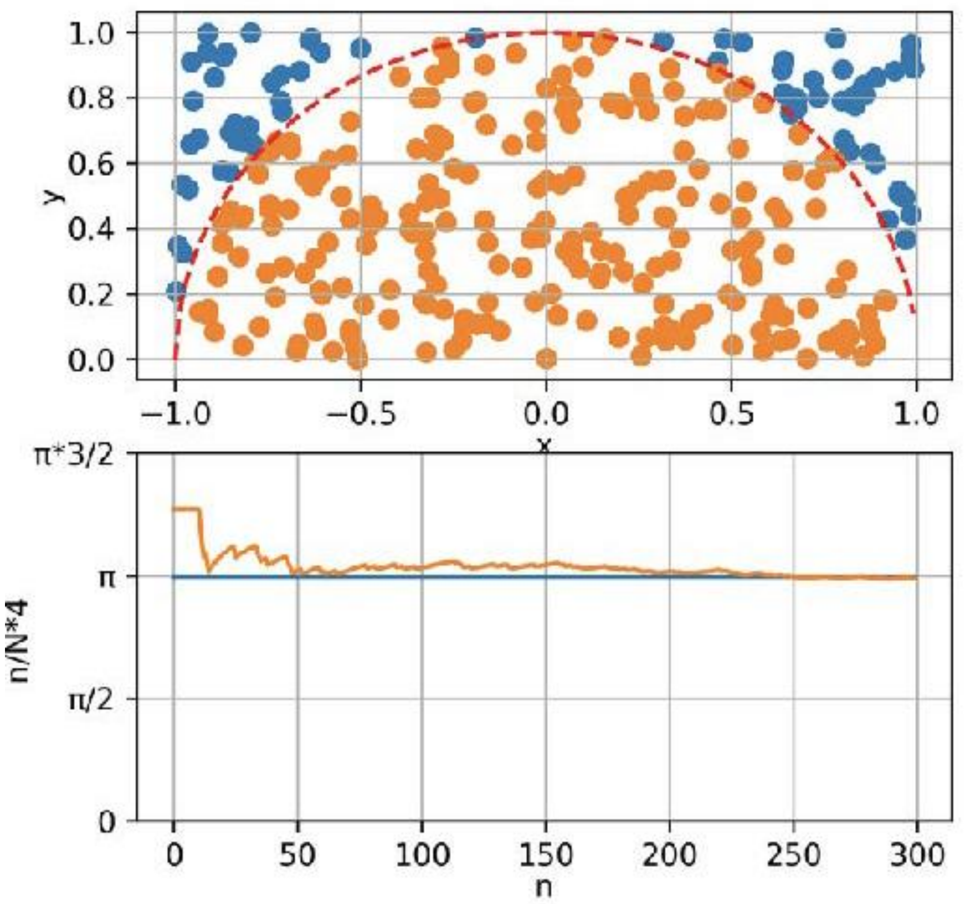
$$A_{circ} = \int_{-1}^{1} f(x)dx = \pi\frac{r^2}{2}$$

$$A_{rect} = \Delta y \cdot \Delta x$$

$$\frac{n}{N} \propto \frac{A_{circ}}{A_{rect}} = \frac{\pi r^2/2}{\Delta x \cdot \Delta y}$$

$$\pi \approx 2\frac{n}{N}\frac{\overbrace{\boxed{\Delta x \cdot \Delta y}}^{=2}}{\underbrace{\boxed{r^2}}_{=1}} = 4\frac{n}{N}$$

- Suppose we don't know the numerical value of $\pi$
- We define it as the ratio of a circle's circumference to its diameter



# Random numbers generators

- At the core of all Monte Carlos there is some mechanism to produce a **long sequence of random numbers** $r_i$ that are uniformly distributed over the interval (0, 1)
  - Digital computers, by design, are incapable of producing random results
- A true random sequence could, in principle, be obtained by coupling to our computer some external device that would produce a truly random signal
  - Such devices rely on physical processes that are inherently unpredictable to generate random numbers
  - **Examples:** Using atmospheric noise, radioactive decay, or quantum phenomena
  - **Advantages:** TRNGs are more secure and produce numbers that are truly random
  - **Use cases:** Cryptography, lottery systems…
  - However, the use of such a random number generator would not always be practical!
- Monte Carlos used in physics usually rely on "pseudo-random number generators"

# Pseudo-random numbers generators

- A **pseudo-random generator** is a deterministic algorithm that, given the previous numbers (usually just the last number) in the sequence, the next number can be efficiently calculated:

$$x_{n+1} = f(x_n, x_{n-1}, \dots, x_0), \text{ where } x_0 \text{ is referred to as the } \textbf{“seed”}$$

- A unique seed returns a unique random number sequence
- Therefore, it is important to use a **new seed every time** that a random selection is initiated
  - A typical error is the use of the same seed for multiple generation, which leads to the generation of the same sample of random numbers
- Because the set of numbers directly representable in the computer is finite, the sequence will necessarily repeat at a certain point. The length of the sequence prior repeating is called **period**
- Many pseudo-random number generators have been proposed and used over the years in a wide variety of Monte Carlo work
  - Designing better random number generators is still an active area of research

# From uniform to other distributions

- **Cumulative distribution** – probability that an event happens between $\infty < x < x_{max}$

$$F(x_{max}) = \int_{\infty}^{x_{max}} f(x)dx$$

- Given a random number in [0,1], find a transformation such that the resulting sequence of $x_i$ is distributed according to $f(x)$

- Inverse transform method: let $x$ be distributed according to $f(x)$, with $F(x)$ in [0,1], then

$$x = F^{-1}(u)$$

  with random number $u = [0,1]$

  - In other words:

    I. If a PDF, $f$, is integrable (called **cumulative**, $F$)

    II. and the cumulative is invertible: $F^{-1}$

    III. It is possible to sample $x$ according to $f$: $x = F^{-1}(u)$, where u is uniformly distributed

# Simple physics application: decay in flight

- Consider a $\pi^+$ with momentum $p$
- The particle's lifetime is a random value with a PDF: $f(t) = \frac{1}{\tau} e^{-\frac{1}{\tau}}$
- Therefore, $t$ can be sampled from the inverse of the cumulative in the interval $r \in [0,1]$

$$f = F^{-1}(r) = -\tau \ln(1 - r)$$

- Pick a decay channel:

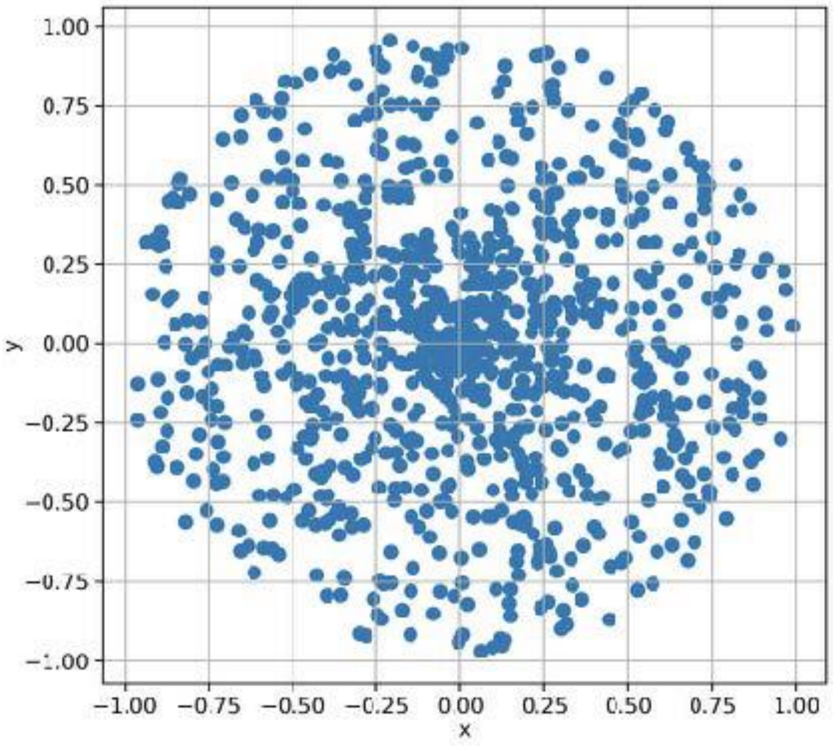
| Mode | | | Fraction $(\Gamma_i / \Gamma)$ |
|---|---|---|---|
| $\Gamma_1$ | $\mu^+ \nu_\mu$ | [1] | $(99.98770 \pm 0.00004)\%$ |
| $\Gamma_2$ | $\mu^+ \nu_\mu \gamma$ | [2] | $(2.00 \pm 0.25) \times 10^{-4}$ |
| $\Gamma_3$ | $e^+ \nu_e$ | [1] | $(1.230 \pm 0.004) \times 10^{-4}$ |
| $\Gamma_4$ | $e^+ \nu_e \gamma$ | [2] | $(7.39 \pm 0.05) \times 10^{-7}$ |
| $\Gamma_5$ | $e^+ \nu_e \pi^0$ | | $(1.036 \pm 0.006) \times 10^{-8}$ |
| $\Gamma_6$ | $e^+ \nu_e e^+ e^-$ | | $(3.2 \pm 0.5) \times 10^{-9}$ |
| $\Gamma_7$ | $e^+ \nu_e \nu \overline{\nu}$ | | $< 5 \times 10^{-6}$ |

Table from PDG

- In the CM frame the decay is isotropic: $\theta \in [0, \pi]$; $\varphi \in [0, 2\pi]$
- Finally, Lorentz-boost in the Lab. frame
- 4 random numbers for one decay!

# Simple physics application: decay in flight

- Why did I sample f in the CM frame?
- What if I sample uniformly $\theta \in [0,2\pi]$; $r \in [0,1]$ ?
  - The extracted points gather at the centre
  - A uniform distribution in polar coordinates is not uniform in orthogonal coordinate system




## ❑ Inverse transform sampling

I. If a PDF, $f$, is integrable (called **cumulative**, $F$)

II. and the cumulative is invertible: $F^{-1}$

III. It is possible to sample $x$ according to $f$:

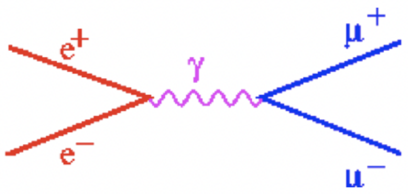
$$x = F^{-1}(u)$$

where u is uniformly distributed

# Now we have a (crude) method… how to simulate physics events?

# Example – building a MC for an $e^+e^-$ process

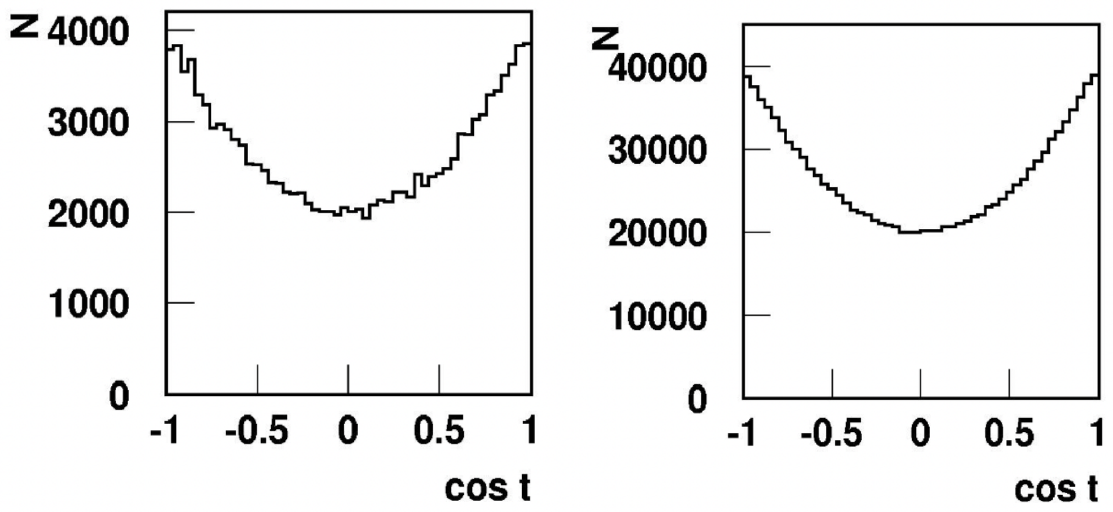
- process: $e^+e^- \rightarrow \mu^+ \mu^-$




- $$\frac{d\sigma}{d\cos\theta\, d\phi} = \frac{\alpha_{em}^2}{4s}\left(1+\cos^2\theta\right)$$

- goal: generate 4-momenta of $\mu$'s, need cm energy s, cos $\theta$, $\phi$

after 100000 events after $10^6$ events




random number R1(0,1): $\phi = 2\pi$ R1
random number R2(0,1): $\cos\theta = -1 + 2$ R2

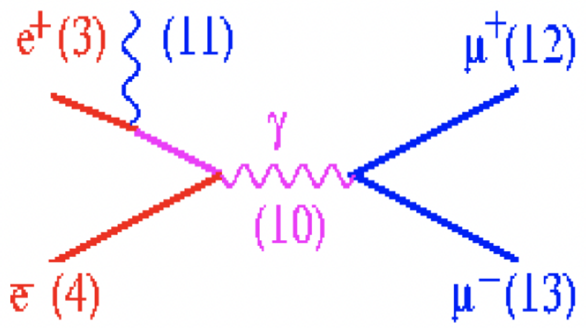
for every R1, R2 use weight with $\frac{d\sigma}{d\cos\theta\, d\phi}$
repeat many times

# MC generators

- example from PYTHIA: Event listing

process: $e^+e^- \rightarrow \mu^+ \mu^-$

| I | particle/jet | KS | KF | orig | p_x | p_y | p_z | E | m |
|---|---|---|---|---|---|---|---|---|---|
| 1 | !e+! | 21 | -11 | 0 | 0.000 | 0.000 | 30.000 | 30.000 | 0.001 |
| 2 | !e-! | 21 | 11 | 0 | 0.000 | 0.000 | -30.000 | 30.000 | 0.001 |
| 3 | !e+! | 21 | -11 | 1 | 0.000 | 0.000 | 30.000 | 30.000 | 0.000 |
| 4 | !e-! | 21 | 11 | 2 | 0.000 | 0.000 | -30.000 | 30.000 | 0.000 |
| 5 | !e+! | 21 | -11 | 3 | 0.143 | 0.040 | 26.460 | 26.460 | 0.000 |
| 6 | !e-! | 21 | 11 | 4 | 0.000 | 0.000 | -29.998 | 29.998 | 0.000 |
| 7 | !Z0! | 21 | 23 | 0 | 0.143 | 0.040 | -3.539 | 56.458 | 56.347 |
| 8 | !mu-! | 21 | 13 | 7 | -9.510 | 1.741 | 24.722 | 26.546 | 0.106 |
| 9 | !mu+! | 21 | -13 | 7 | 9.653 | -1.700 | -28.261 | 29.913 | 0.106 |
| 10 | (Z0) | 11 | 23 | 7 | 0.143 | 0.040 | -3.539 | 56.458 | 56.347 |
| 11 | gamma | 1 | 22 | 3 | -0.143 | -0.040 | 3.539 | 3.542 | 0.000 |
| 12 | mu- | 1 | 13 | 8 | -9.510 | 1.741 | 24.722 | 26.546 | 0.106 |
| 13 | mu+ | 1 | -13 | 9 | 9.653 | -1.700 | -28.261 | 29.913 | 0.106 |
| | | sum: | 0.00 | | 0.000 | 0.000 | 0.000 | 60.000 | 60.000 |




- technicalities/advantages

➔ can work in any frame

➔ Lorentz-boost 4-vectors back and forth

➔ can calculate any kinematic variable

➔ history of event process

# Keeping track of particles

- Particle travel is the most common application of MC in particle physics
- Assume that all the possible interactions are known
- The distance s between two subsequent interactions is distributed as:

$$p(s) = \mu e^{-\mu s}$$

  - where **$\mu$ is a property of the medium** and is proportional to the probability of an interaction per unit length
  - Therefore, $\mu$ is proportional to the **total cross section**:

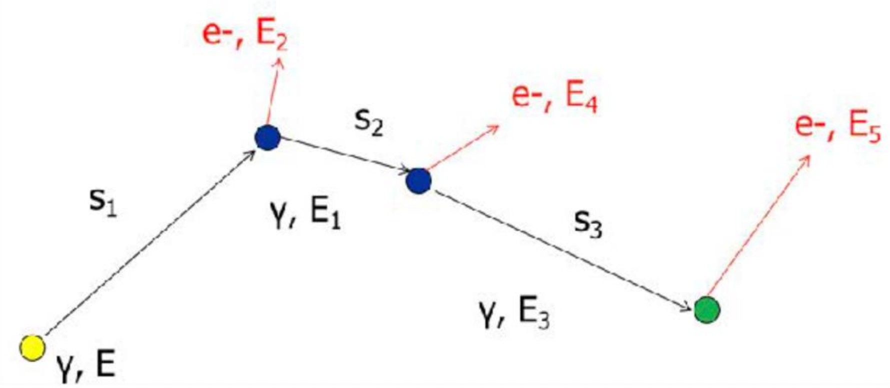
$$\mu = N\sigma = N\sum_i \sigma_i = \sum_i \mu_i$$

  - where $N$ is the number of scattering centers in the medium and $\mu_i$ are the partial cross sections of **all the competing processes**
  - $\mu$ depends on the **density of the material**

# Keeping track of particles

The basic technique works as follows:

1. Divide the particle trajectory in "steps" ($s_i$) of straight free-flight tracks between interactions
   - This could be limited by geometry boundaries of the material
2. Sampling the step length according to $p(s)$
3. Sampling the interaction at the end of the step according to $\mu_i/\mu$
4. Sampling the final state using the physics model of the $i$-th interaction
   - Update the properties of the primary particle
   - Add the possible secondaries produced, that need to be also tracked later
   - Follow (track) all the secondaries, until they are absorbed or leave the geometry

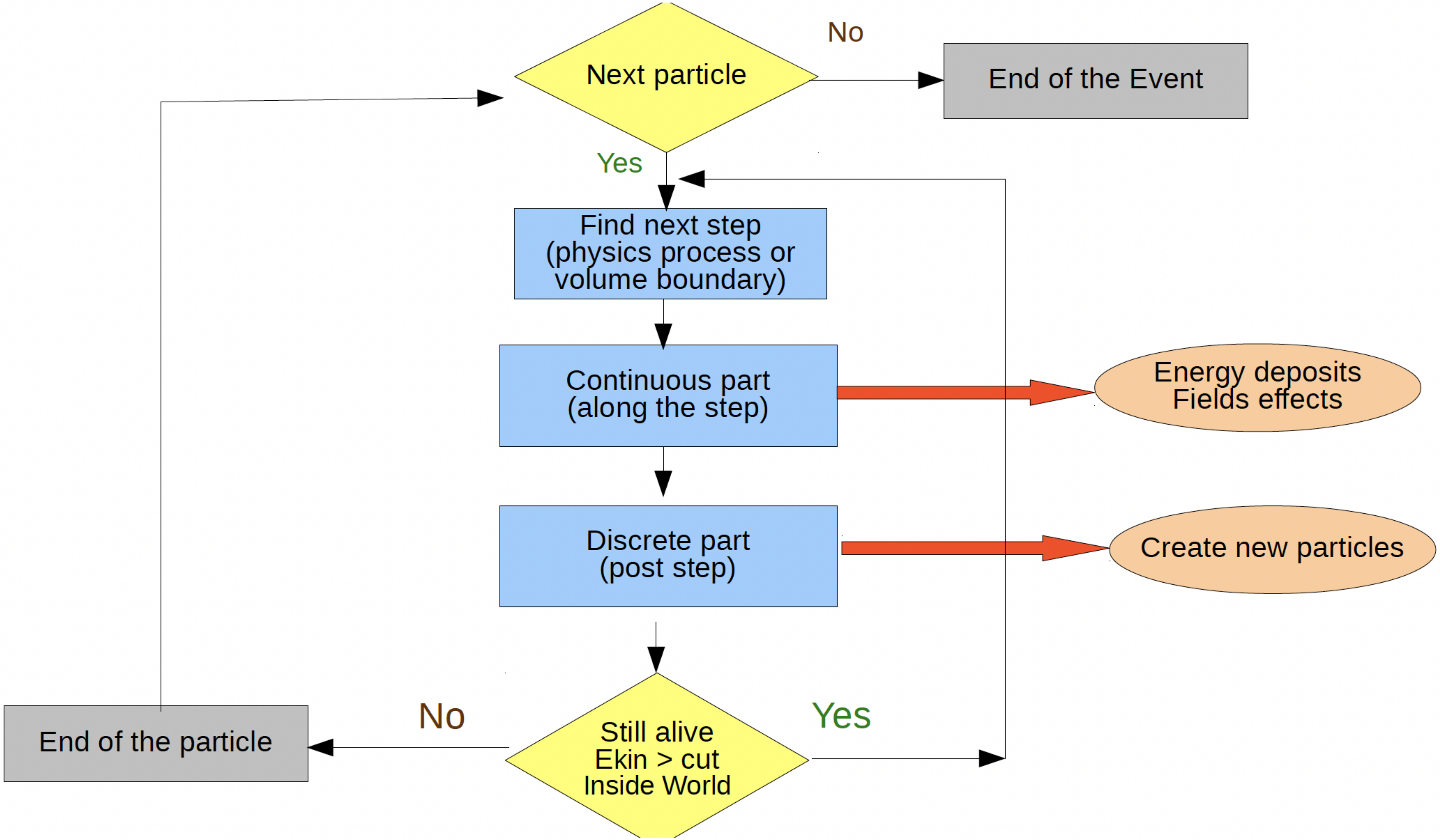


# Keeping track of particles

Describing the travel of particle through matter is not so easy:

- This basic recipe doesn't work well for charged particles
- The **cross section** of some processes (ionisation and bremsstrahlung) is **very high**, so the **steps** would be **very small**
- In each interaction only a **small fraction of energy is lost** and the effects on the particle are small
  - A lot of CPU time is used to simulate many interactions, each having very small effects
- A way of mitigating this issue is:
  1. Explicitly simulate interactions (detailed simulation) only if the energy loss is above a certain threshold $E_0$ (**hard interactions**)
  2. The effects of all sub-threshold **(soft)** interactions are described cumulatively
     - Hard interactions are much less frequent than soft interactions

# A typical flowchart of an event



# Tools to study particle travel through matter

- Luckily, we do not have to reinvent the wheel each time we want to propagate particles generated by a Monte Carlo throughout our experimental setup
- Tools have been put together over decades by many scientists in our community and are continuously being tested and improved
- These tools provide us with reliable ways of simulating travel of particles through matter, and much more!
- The most famous and widely used is GEANT (GEometry ANd Tracking)

GEANT4
A SIMULATION TOOLKIT