Monte Carlo sampling

Beginners' FLUKA Course

Overview:

General concepts:

- Phase space
- The Boltzmann equation
- Monte Carlo foundations
- Simulation vs. integration

Sampling techniques

- discrete
- by inversion
- by rejection

Results and Errors:

- Statistical errors (single histories, batches)
- Figure of merit

Phase space:

- Phase space: a concept of classical Statistical Mechanics
- Each Phase Space dimension corresponds to a particle degree of freedom
- 3 dimensions correspond to Position in (real) space: x, y, z
- 3 dimensions correspond to Momentum: p_x, p_y, p_z
  (or Energy and direction: E, θ, φ)
- More dimensions may be envisaged, corresponding to other possible degrees of freedom, such as quantum numbers: spin etc
- Each particle is represented by a point in phase space
- Time can also be considered as a coordinate, or it can be considered as an independent variable: the variation of the other phase space coordinates as a function of time constitutes a particle "history"

The Boltzmann equation:

- All particle transport calculations are (explicit or implicit) attempts to solve the Boltzmann Equation
- It is a balance equation in phase space: at any phase-space-point, the increment of particle phase-space-density is equal to the sum of all "production terms" minus a sum of all "destruction terms"
- Production: Sources, "Inscattering", Particle Production, Decay
- Destruction: Absorption, "Outscattering", Decay
- We can look for solutions of different type: at a number of (real or phase) space points, averages over (real or phase) space regions, projected on selected phase space hyperplanes, stationary or time-dependent

Mean of a distribution - 1

In one dimension:

Given a variable *x*, distributed according to ***f(x)***, the mean or average of another function of the same variable A(x) over $[x_{min},x_{max}]$ is given by:

$$\bar{A} = \frac{\int_{x_{min}}^{x_{max}} A(x) f(x) dx}{\int_{x_{min}}^{x_{max}} f(x) dx}$$

Or, introducing the normalized distribution *f'*:

$$f'(x) = \frac{f(x)}{\int_{x_{min}}^{x_{max}} f(x) dx}$$

$$\bar{A} = \int_{x_{min}}^{x_{max}} A(x) f'(x) dx$$

A special case is that of *A(x)=x:*

$$\bar{x} = \int_{x_{min}}^{x_{max}} x f'(x) dx$$

Mean of a distribution - 2

In several dimensions:

- Given n variables, $x,y,z,\ldots$, distributed according to the (normalized) functions $f'(x,y,z,\ldots)$, $g'(x,y,z,\ldots)$, $h'(x,y,z,\ldots)$, the mean or average of a function of those variables $A(x,y,z,\ldots)$ over an n-dimensional domain D, is given by:

$$\bar{A} = \int_x \int_y \int_z \ldots \int A(x,y,z,\ldots) f'(x,y,z,\ldots) g'(x,y,z,\ldots) h'(x,y,z,\ldots) dx\, dy\, dz \ldots$$

- Often impossible to calculate with traditional methods, but we can sample N values of A, by sampling N sets of variables $x_i, y_i, z_i \ldots$ with probability $f' \cdot g' \cdot h' \ldots$ and divide the sum of the sampled values by N:

$$S_N = \frac{\sum_{i=1}^{N} A(x_i, y_i, z_i, \ldots)}{N}$$

- Each term of the sum is distributed like A, integration but also simulation!

Central limit theorem

Central limit theorem:

$$\lim_{N\to\infty} P(S_N) = \frac{1}{\sqrt{2\pi\sigma_A^2/N}} \exp\left[-\frac{(S_N - \bar{A})^2}{2\sigma_A^2/N}\right]$$

- For large values of N, the normalized sum of N independent and identically distributed random variables tends to a normal distribution with mean $\bar{A}$ and variance σ^2_A/N

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \sum_{i=1}^{N} \frac{A(x_i, y_i, z_i, \ldots)}{N} = \bar{A}$$

Monte Carlo* mathematical foundation:

Several possible ways of defining Monte Carlo (MC):

- A mathematical method for Numerical Integration
  - Random sampling techniques
  - Convergence, variance reduction techniques…
- A computer simulation of a Physical Process
  - Physics
  - Tracking
  - Scoring…

Both are valid, depending on the problem one or the other can be more effective (see the examples above)

* Monte Carlo method "inventors": Von Neumann, Ulam, Fermi, Metropolis in the late 40's

Monte Carlo* mathematical foundation:

The central limit theorem is the mathematical foundation of the Monte Carlo method:

In words:

Given any observable A, that can be expressed as the result of a convolution of random processes, the average value of A can be obtained by sampling many values of A according to the probability distributions of the random processes

MC is indeed an INTEGRATION method that allows to solve multi-dimensional integrals by sampling

The accuracy of a MC estimator depends on the number N of samples ($\propto 1/\sqrt{N}$)

Analog Monte Carlo:

In an analog Monte Carlo calculation ("honest" simulation), not only the mean of the contributions converges to the mean of the real distribution, but also the variance and all moments of higher order

$$\mu^{-m} = \int_x \int_y \int_z \cdots \int \left[A(x,y,z,\ldots) - \bar{A} \right]^m f'(x,y,z,\ldots) g'(x,y,z,\ldots) h'(x,y,z,\ldots) dx\, dy\, dz \ldots$$

converge as well:

$$\lim_{N \to \infty} \left[\sum_{i=1}^{N} \left(A_i - \frac{S_n}{N} \right)^m \right]^{\frac{1}{m}} = \mu^{-m}$$

and fluctuations and correlations are faithfully reproduced

Integration efficiency:

- Traditional numerical integration methods (Simpson, etc), converge to the true values as $N^{-1/n}$ where N = number of "points" (interval), and n = number of dimensions
- Monte Carlo converges instead as $1/\sqrt{N}$

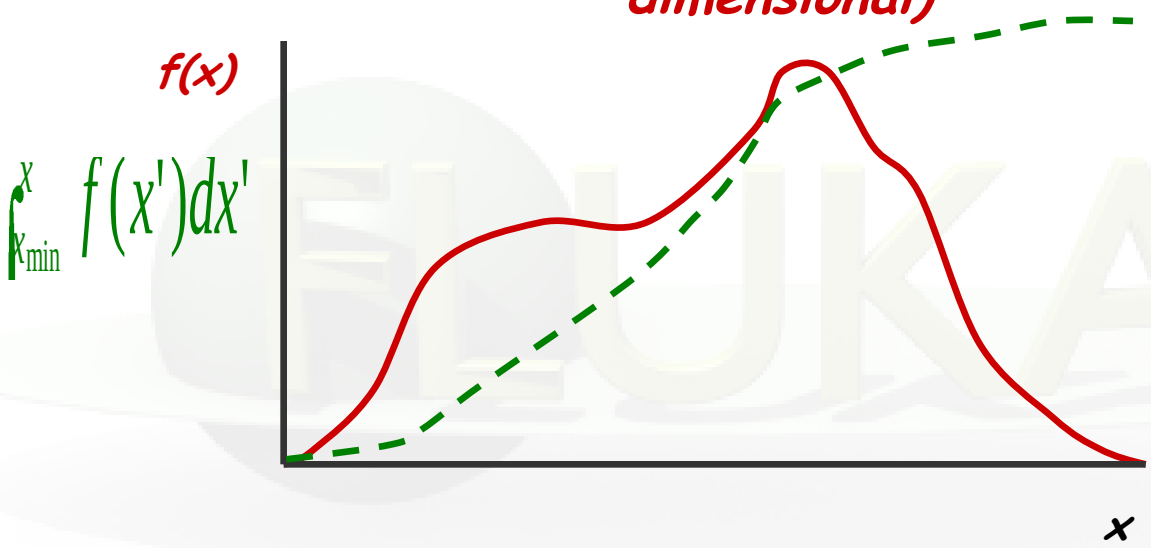
Number of dimensions	Traditional methods	Monte Carlo	Remark
$n = 1$	$1/N$	$1/\sqrt{N}$	MC not convenient
$n = 2$	$1/\sqrt{N}$	$1/\sqrt{N}$	About equivalent
$n > 2$	$1/\sqrt[n]{N}$	$1/\sqrt{N}$	MC converges faster

A typical particle transport Monte Carlo problem is a 7-D problem!
x, y, z, p_x, p_y, p_z and t !!

Random Sampling: the key to Monte Carlo!

The central problem of the Monte Carlo method:

Given a Probability Density Function (pdf), f(x), generate a sequence of x's distributed according to f(x) (x can be multi-dimensional)



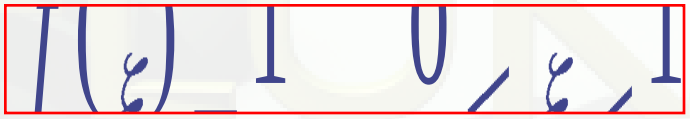
The use of random sampling techniques is the distinctive feature of Monte Carlo

The use of Monte Carlo to solve the integral Boltzmann transport equation consists of:

- Random sampling of the outcome of physical events
- Geometry and material description of the problem

(Pseudo) Random numbers:

- Basis for all Monte Carlo integrations are random numbers, *i.e.* values of a variable distributed according to a pdf (probability distribution function).
- In real world: the random outcome of a physical process
- In computer world: pseudo-random numbers
- The basic pdf is the uniform distribution:



- Pseudo-random numbers are sequences that reproduce the uniform distribution, constructed from mathematical algorithms.
- All computers provide a pseudo-random number generator (or even several of them). In most computer languages (e.g., Fortran 90, C) a PRNG is even available as an intrinsic routine

Sampling from a distribution:

Sampling from a discrete distribution:

- Suppose to have a *discrete* random variable x, that can assume values $x_1, x_2, \ldots, x_n, \ldots$ with probability $p_1, p_2, \ldots, p_n, \ldots$
- Assume $\sum_i p_i = 1$, or normalize it
- Divide the interval [0,1) in n subintervals, with limits

$$y_0 = 0, \; y_1 = p_1, \; y_2 = p_1 + p_2, \ldots$$

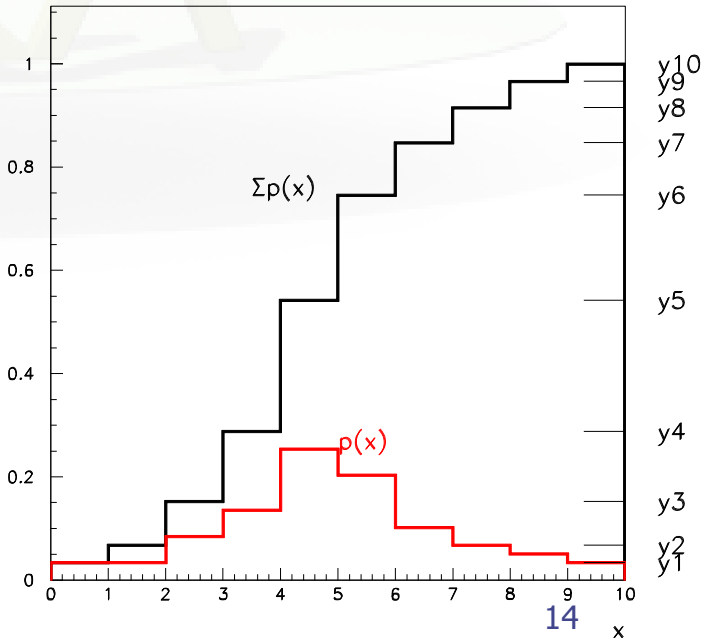
- Generate a uniform pseudo-random number ξ
- Find the interval i^{th} y-interval such that

$$y_{i-1} \leq \xi < y_i$$

- Select $X = x_i$ as the sampled value

Since ξ is uniformly random:

$$P(x_i) = P(y_{i-1} < \xi < y_i) = y_i - y_{i-1} = p_i$$



Sampling from a distribution:

Sampling from a generic continuous distribution:

- Integrate the distribution function ***f(x)***, analytically or numerically, and normalize to 1 to obtain the normalized cumulative distribution

$$F(\xi) = \frac{\int_{x_{min}}^{\xi} f(x)dx}{\int_{x_{min}}^{x_{max}} f(x)dx}$$

- Generate a uniform pseudo-random number ξ
- Get the desired result by finding the inverse value $X = F^{-1}(\xi)$, analytically or most often numerically, i.e. by interpolation (table look-up)

Since ξ is uniformly random:

$$P(a < x < b) = P(F(a) < \xi < F(b)) = F(b) - F(a) = \int_a^b f(x)dx$$

Example:

Take $f(x) = e^{-\frac{x}{\lambda}}$ $x \in [0,-\infty)$

Cumulative distribution:

$$F(t) = \int_0^t e^{-\frac{x}{\lambda}} dx = \lambda \times \left(1 - e^{-\frac{t}{\lambda}}\right)$$

Normalized:

$$F'(t) = \int_0^t \frac{e^{-\frac{x}{\lambda}}}{\lambda} dx = 1 - e^{-\frac{t}{\lambda}}$$

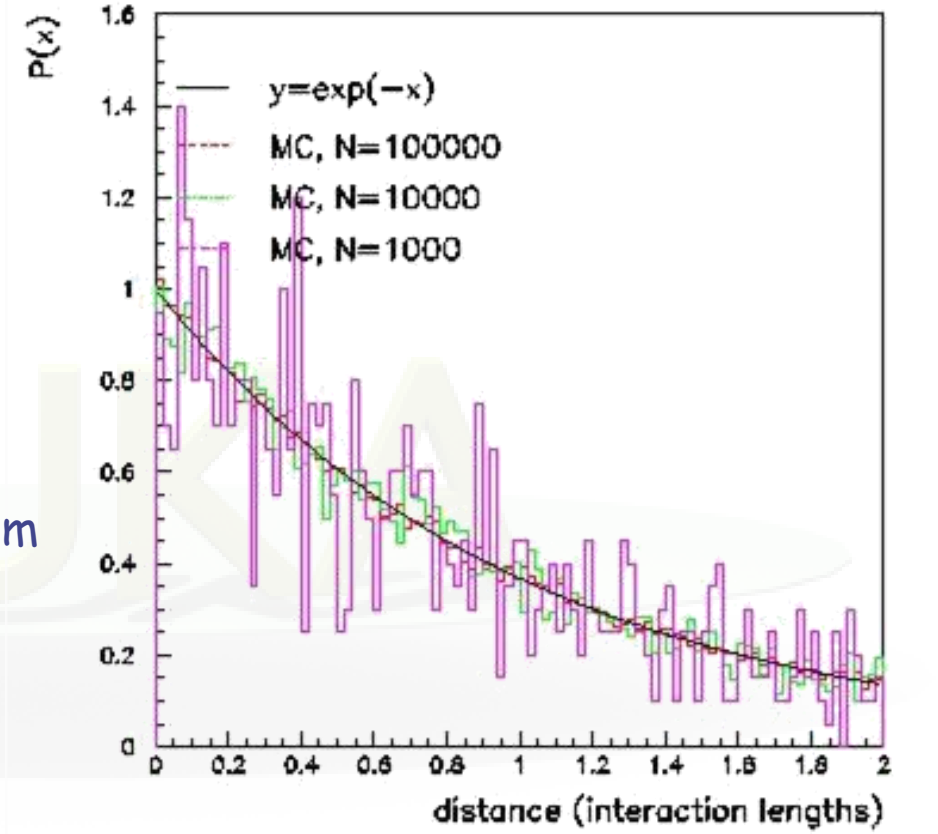
Generate a uniform pseudo-random number $\xi \in [0,1)$

$$1 - e^{-\frac{t}{\lambda}} = \xi$$

Sample t by inverting

$$t = -\lambda \ln(1-\xi)$$

Repeat N times



Practical rule: a distribution can be directly sampled if and only if its pdf can be integrated and the integral inverted

Sampling from a distribution: rejection technique

Rejection procedure:

- Let be $f'(x)$, a normalized distribution function, which cannot be sampled by integration and inversion
- Let be $g'(x)$, a normalized distribution function, which can be sampled, and such that $Cg'(x) \geq f'(x)$, $\forall\ x \in [x_{min}, x_{max}]$
- Sample X from g'(x), and generate a uniform pseudo-random number $\xi \in [0,1)$
- Accept X if $\xi < f'(X)/Cg'(X)$, if not repeat the previous step
- The overall efficiency (accepted/rejected) is given by:

$$R = \int \frac{f'(x)}{Cg'(x)} g'(x)dx = \frac{1}{C}$$

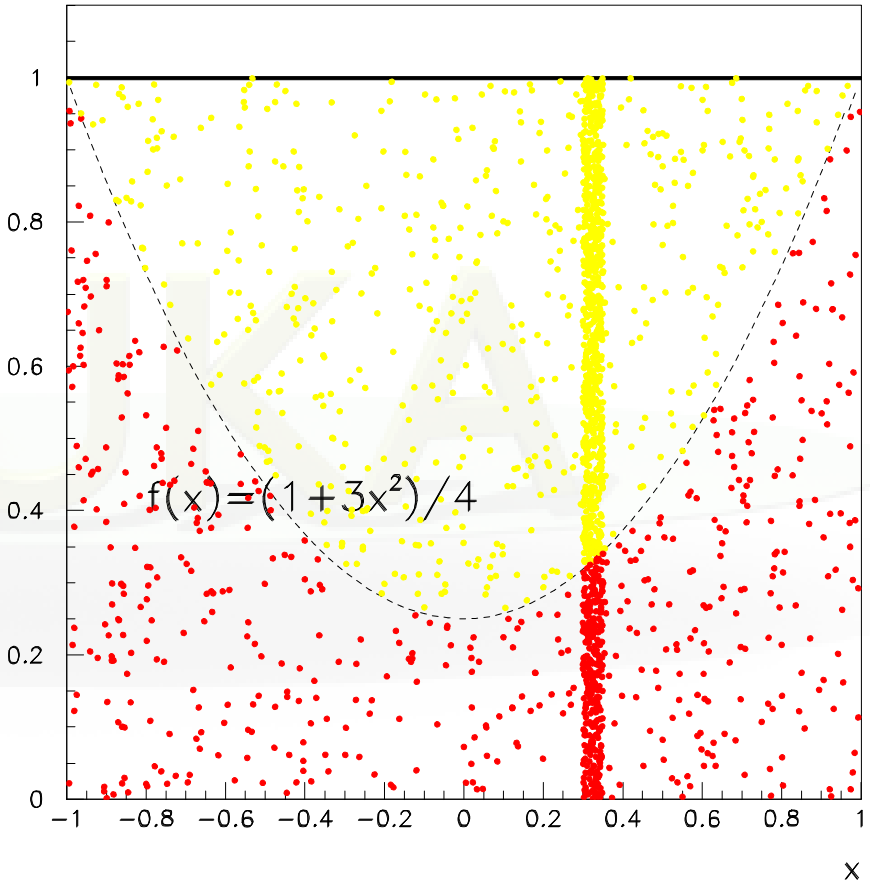
- and the probability that X is accepted is unbiased:

$$P(X)dX = \frac{1}{R} g'(X)dX \times \frac{f'(X)}{Cg'(X)} = f'(X)dX$$

Sampling from a distribution: example

Rejection procedure:

- Let be $f'(x) = (1+3x^2)/4$, $x \in [-1,1]$,
- Take $g'(x)=1/2$, $C=2$
- Generate two uniform pseudo-random numbers ξ_1, $\xi_2 \in [0,1)$
- Accept $X=2\xi_1-1$ if $\xi_2 < (1+3X^2)/4$, if not repeat



FLUKA

FLUKA

Particle transport Monte Carlo:

Assumptions:

- Static, homogeneous, isotropic, and amorphous media (and geometry)
- Markovian process: the fate of a particle depends only on its actual properties, not on previous events or histories
- Particles do not interact with each other
- Particles interact with individual atoms/nuclei/molecules (invalid at low energies)
- Material properties are not affected by particle reactions

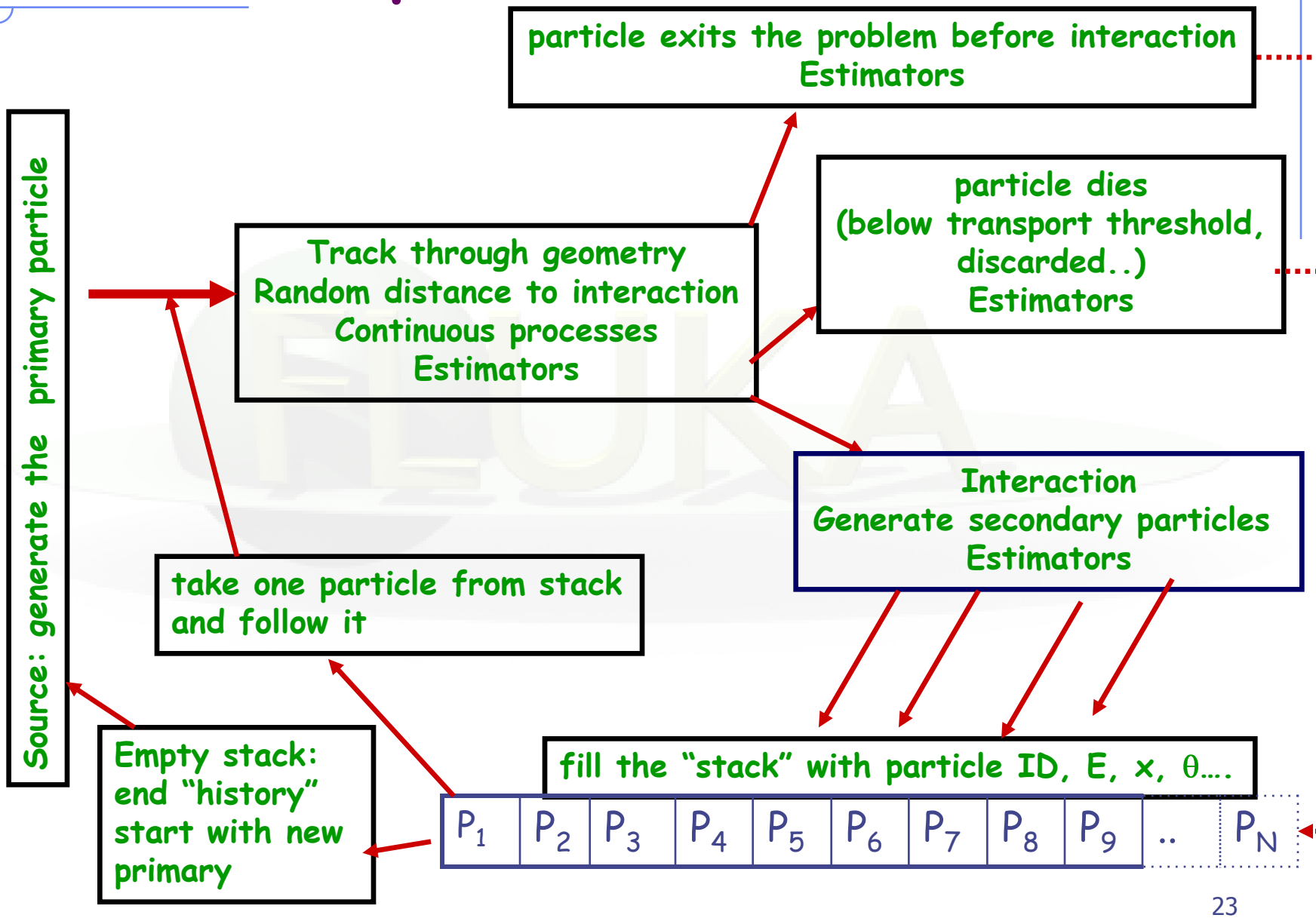
The superposition principle can be used

Particle transport Monte Carlo:

Application of Monte Carlo to particle transport and interaction:

- Each particle is followed on its path through matter.
- At each step the occurrence and outcome of interactions are decided by random selection from the appropriate probability distributions.
- All the secondaries issued from the same primary are transported before a new history is started.
- The accuracy and reliability of a Monte Carlo depends on the models or data on which the pdfs are based
- Statistical accuracy of results depends on the number of "histories"
- Statistical convergence can be accelerated by "biasing" techniques.

Practical implementations



Statistical Errors:

- Can be calculated for single histories, or for batches of several histories
- Distribution of scoring contributions by single histories can be very asymmetric (many histories contribute little or zero)
- Scoring distribution from batches tends to Gaussian for $N \rightarrow \infty$, provided $\sigma^2 \neq \infty$ (thanks to Central Limit Theorem)
- The standard deviation of an estimator calculated from batches or from single histories is an estimate of the standard deviation of the actual distribution ("error of the mean")
- How good is such an estimate depends on the type of estimator and on the particular problem (but it converges to the true value for $N \rightarrow \infty$)

Relative error	Quality of Tally *(from the MCNP Manual)*
50 to 100%	Garbage
20 to 50%	Factor of a few
10 to 20%	Questionable
< 10%	Generally reliable except for point detectors

Statistical Errors (batch statistics)

The variance of the mean of an estimated quantity x (e.g., fluence), calculated out of N batches, is:

$$\sigma^2_{\langle x \rangle} = \frac{1}{N-1}\left[\frac{1}{n}\sum_{i=1}^{N} n_i x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{N} n_i x_i\right)^2\right]$$

where:

- n_i is the number of histories in the i^{th} batch
- $n = \sum n_i$ is the total number of histories in the N batches
- x_i is the average of x calculated in the i^{th} batch: $x_i = \sum_{j=1}^{n_i} x_{ij} / n_i$ where x_{ij} is the contribution to x of the j^{th} history in the i^{th} batch
- In the limit $N=n$, $n_i=1$, the formula applies to single history statistics

Practical tips:

- Use always at least 5-10 batches of comparable size (it is not at all mandatory that they be of equal size)
- Never forget that the variance itself is a stochastic variable subject to fluctuations
- Be careful about the way convergence is achieved: often (particularly with biasing) apparently good statistics with few isolated spikes could point to a lack of sampling of the most relevant phase-space part
- Plot 2D and 3D distributions! In those cases the eye is the best tool in judging the quality of the result