From physics to medical imaging, through electronics and detectors

Simulation tools

Irène Buvat IMNC CNRS 8165 Orsay

buvat@imnc.in2p3.fr http://www.guillemet.org/irene

June 2012

- Introduction
- The key ingredient: the random number generator
- Monte Carlo simulations in imaging and radiation therapy
 - Modeling a geometry
 - Modeling the sources of particles
 - Modeling the particle-matter interactions
 - Storing the relevant information
- Accelerating the simulations
 - Algorithmic approaches
 - Hardware approaches
- Simulation tools in Emission Tomography, CT and Radiation Therapy
 - General purpose tools
 - Dedicated tools
- Numerical models for patient description
- Applications of MC simulations
- Useful resources

Introduction



• Simulation : a computational model of a physical system.

• A model is a simplified version of reality, in which all parameters (usually hidden in the real system) are perfectly known, i.e. fully under control.

• The role of simulations is to help understand or predict the behavior of a complex system:

- financial markets

- traffic flow
- environmental sciences
- particle physics
- quantum field theory
- astrophysics
- molecular modeling
- light transportations
- radiation therapy
- medical imaging
- ...

• Analytical simulations

- deterministic : yield always the same result with the same input

• Stochastic simulations (= Monte Carlo simulations):

- new paradigm: randomness is an experimental tool !

- it can be used to solve problems !



Basic principle of stochastic simulations

• How can randomness give us solutions to problems?

Basic example: estimation of π



area of the circle : $A_c = \pi$ area of the square : $A_s = 4$

throw dots randomly in the square :

number N_s of dots in the square / number N_c of dots in the circle = $4/\pi$

 $\Rightarrow \pi = 4 N_c / N_s$

• Stochastic simulations (= Monte Carlo simulations):

- Even more puzzling: randomness can be used to solve problems for which no analytical solution can be derived !



• French Comte de Buffon (1777) : estimated π by tossing a needle on a line background.

• Fermi (1930) : random method to calculate the properties of the newly discovered proton

• John von Neumann, Stanislas Ulam and Metropolis (1940's) used simulations to determine the distance that neutrons traveled through radiation shielding

• Rapid growth with the availability of digital computers

• First numerical Monte Carlo simulation of photon transport by Hayward and Hubbell (1954) who generated 67 photon histories using a desk calculator

• Berger (1963) : first coupled electron-pgoton transport code that became known as ETRAN

• Exponential growth in Medical Physics and Medical Sciences



Number of "Monte Carlo" hits in Pubmed

Reasons explaining this growth

• Increase in computing power and availability in the last 5 decades



• Increase availability of powerful software tools











• The legend says that the method was named in honor of Ulam's uncle, who was a gambler, at a suggestion of colleague Metropolis. They both produced the first paper describing the approach:

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 247

SEPTEMBER 1949

Volume 44

THE MONTE CARLO METHOD

NICHOLAS METROPOLIS AND S. ULAM Los Alamos Laboratory

We shall present here the motivation and a general description of a method dealing with a class of problems in mathematical physics. The method is, essentially, a statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of the natural sciences.

A LREADY in the nineteenth century a sharp distinction began to appear between two different mathematical methods of treating physical phenomena. Problems involving only a few particles were studied in classical mechanics, through the study of systems of ordinary differential equations. For the description of systems with very many particles, an entirely different technique was used, namely, the method of statistical mechanics. In this latter approach, one does not concentrate on the individual particles but studies the properties of sets of particles. In pure mathematics an intensive study of the properties of sets of points was the subject of a new field. This is the so-called theory of sets, the basic theory of probabilities prepared the formal apparatus for the use of such models in theoretical physics, i.e., description of properties of aggregates of points rather than of individual points and their coordinates.

Soon after the development of the calculus, the mathematical apparatus of partial differential equations was used for dealing with the problems of the physics of the continuum. Hydrodynamics is the most widely known field formulated in this fashion. A little later came the treatment of the problems of heat conduction and still later the field theories, like the electromagnetic theory of Maxwell. All this is very well known. It is of course important to remember that the study of the

The key ingredient



The random number generator

- It can't.
- Pseudo random numbers are used instead.
- This is actually more practical for debugging, as a sequence of random numbers can be reproduced.

• PRNG = pseudo random number generator (sometimes just called RNG).

• Many PRNG are implemented in mathematical libraries.



- LNG
- Start from a sample r_n.
- $r_{n+1} = (a r_n + c) \mod m$, i.e., remainder of $(a r_n + c)/m$

where a, c and m should be careful chosen (from tables).

- Easy to implement, quick, requires minimal memory.
- Short period: the same number sequence repeats itself after a relatively small number of samples.
- The period is at most equal to m

Source	m	а	с
Numerical Recipes	2 ³²	1664525	1013904223
Borland C/C++	2 ³²	22695477	1
glibc (used by GCC) ^[4]	2 ³¹	1103515245	12345
ANSI C: Watcom, Digital Mars, CodeWarrior, IBM VisualAge C/C++	2 ^{31 [5]}	1103515245	12345
Borland Delphi, Virtual Pascal	2 ³²	134775813	1
Microsoft Visual/Quick C/C++	2 ³²	214013	2531011
Microsoft Visual Basic (6 and earlier) ^[6]	2 ²⁴	1140671485	12820163
RtIUniform from Native API ^[7]	2 ³¹ – 1	2147483629	2147483587
Apple CarbonLib	2 ³¹ – 1	16807	0
MMIX by Donald Knuth	2 ⁶⁴	6364136223846793005	1442695040888963407
VAX's MTH\$RANDOM, ^[8] old versions of glibc	2 ³²	69069	1
Java's java.util.Random	2 ⁴⁸	25214903917	11

- Park and Miller (period = 2^{31} -2)
- Lüscher (period = 10^{171})
- Mersenne Twister invented in 1997 (period = 2^{19937} -1)

Series of tests have been developed to characterize the quality of an RNG. A well known suite of test that can be used is the TestU01 of l'Ecuyer (ACM Trans Math Soft, 2007).

8. SUMMARY AND CONCLUSION

We have described the *TestU01* package for testing random number generators and used it to test several widely used or recently proposed generators. A minority of these generators pass all the tests (in the sense that no *p*-value is outside the interval $[10^{-10}, 1-10^{-10}]$). The default generators of many popular software programs (e.g., Excel, MATLAB, Mathematica, Java standard library, R, etc.) fail several tests miserably. The survivors are the long-period MRGs with good structure, the multiplicative lagged-Fibonacci generators, some non-

Note that in MC simulations for medical physics, an extremely large number of random drawings is performed. The quality of the PRNG is thus of foremost importance to avoid bias. • Most PRNG sample positive integer values between 0 and N.

• A uniform distribution between 0 and 1 is obtained by dividing the value by N.

• Most often, one has to sample from a non-uniform probability density function (PDF).

Several methods can be used to go from a uniform distribution to a non-uniform PDF.

Reminder:

If a PDF is p(x),

the Cumulative Distribution Function CDF is :

$$P(x) = \int_{-\infty}^{x} p(x) dx$$

In the following, we will sample p(x)

- 1. Sample u from the uniform distribution on [0,1]
- 2. Locate u on the y-axis of the CDF P(x) of p(x)
- 3. Find the x value such as P(x)=u



The sampling method can be used whenever the target CDF is invertible

The acceptance - rejection sampling method

- Also called rejection sampling
- Often used when the CDF is not analytically invertible or when the inversion is computationally burdensome
- This is an exact sampling method
- 1. Choose an easy-to-sample PDF g(x) and a constant c so that $c \times g(x) \ge p(x)$ for every x
- 2. Generate a random number y from g(x) (for instance, using the inversion method)
- 3. Generate a random number u distributed uniformly between 0 and 1
- 4. If $u < p(y)/[c \times g(y)]$, then accept y as the random sample
- 5. Otherwise, reject y and start again at step 2

How does that work?

- Sampling y from g gives the wrong distribution for p
- But we reject the sample y in proportion of how far off we are at y



If u < p(y)/ [c x g(y)], then accept y as the random
 sample
Otherwise, reject y and start again</pre>

• Especially when the PDF is derived from experimental data (for example cross-sections)

• The finer the sampling, the lower the bias (interpolation has to be used between samples)

Library functions

- Random sampling functions of many common PDF are available in mathematical code libraries
- The use of these functions is usually recommended, especially for the Gaussian PDF and the Poisson PDF often used in simulations in Medical Physics



ENTERVISION School - Irène Buvat – june 2012 - 21

Monte Carlo simulations in imaging and RT



Basic principles

- Modeling a geometry
- Modeling the sources of particles
 - Nuclear decay in emission imaging
 - Sources of photons for X-ray CT
 - Beams (photons, electrons, protons,
 - Carbon ions) for radiation therapy
- Modeling the interactions between particles and matter
 - Electromagnetic interactions
 - Hadronic interactions
 - within the phantom / patient
 - within the detector in imaging simulations
- Storing relevant information
 - related to the detected events in imaging applications
 - related to dose deposit in radiation therapy applications

 2 types of object / patient / detector models

 analytical description: the phantom/ patient is described using a set of geometric primitives (box, cylinders, spheres, ellipsoids)

- voxelized description: a series of images (i.e. a volume)



from Handbook in Particle Detection and Imaging, Springer, 2012, p 1106

Advantages and drawbacks of the patient description

• Analytical:

- optimizes computational efficiency: navigation through geometric primitives is faster than navigation through voxels

- easy modelling of deformation and motions

- any sampling

- hard to model a realistic patient only with primitives, or the number of geometric primitives gets so high that the benefit of using primitives disappears

• Voxelized:

- can be highly realistic, starting from a CT scan for instance

- when navigating, the simulation must update the object properties at each voxel change: computationally intensive ! • Programming a geometry description is not difficult conceptually, but can be very tedious for complex geometries

⇒ Describing the simulation geometry is reduced to learning the syntax of an input file or learning to operate a Graphical User Interface

- Generate one decay at a time
- Stop through the voxels, or primitives in the patient volume and

- calculate the mean number *n* of decays for the current time frame (activity in Bq x duration of the frame in s)

- sample a Poisson variable with this mean *n* to get the number of decays in this voxel or primitive

- For each decay :

 randomly pick a point within the voxel or primitive for the decay

• randomly pick a time within the time frame from a truncated exponential distribution with λ being the half life of the isotope (or from a uniform distribution if radiactive decay can be ignored) • Single decay mode: the photon is always simulated with the same energy

• Multiple decay mode: photons are produced with different energies and different probabilities, with a time correlation between the various decay products.

If the half life of the intermediate decay products is very short, multiple decays are simulated simultaneously (with the same time stamp)

- Select the direction of travel of the photon given by a unit director vector \vec{v} :
- sample θ from a uniform distribution on [0;2 π)
- sample μ from a uniform distribution on [-1;1]
- Let $\phi = \cos^{-1}(\mu)$
- Set $\vec{v} = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$

• Positron travel before annihilation can be simulated but this is computationally expensive

• Most often, « back-to-back » annihilation photons are simulated instead

• Ideally, if back-to-back simulations, the annihilation location should be chosen away from the positron location (e.g., using the distribution by Palmer and Brownell 1992)

• One photon direction is chosen as in the gamma decay modeling

• The other is chosen either at 180° , or modeling the acolinearity: the angle between the 2 photons is $180^{\circ} + \varepsilon$, where ε is a normal random variable with mean 0° and standard deviation of approximately 5° (Evans 1955).

> Palmer and Brownell 1992 IEEE Trans Med Imaging 11:373-378 Evans 1955 The atomic nucleus McGrax-Hill, New York

Sources of X-rays for transmission imaging

- X-ray sources are polyenergetic and are collimated
- This can be simulated but is computationally expensive
- Most often:
 - a point source is modelled

- an idealized energy distribution is used, sometimes based on experimental measurements or preliminary simulations

- Most often the beam is assumed to be radially symmetric, having a Gaussian angular spread
- Iterative adjustment of the energy and divergence to match experimental profiles



Monte Carlo modelling of a 135 MeV proton beam

Prerequisite: modeling the properties of the object/patient

- Whatever the object / patient / detector description, the material in each component of it should be precisely described, so that particle-matter interaction can be properly simulated.
- Each primitive or voxel is associated with a material, corresponding to a particular chemical composition.

mass fraction

[Mater:	ials]				
Lung:	d=0.26	g/cm3	; n=9		
	+el: name=Hydrogen ;			f=0.103	
	+el: name=Carbon ;			f=0.105	
	+el: na	.me=Ni	f=0.031		
	+el: na	.me=0xy	ygen ;	f=0.749	
	+el: na	.me=So	lium ;	f=0.002	
	+el: na	me=Pho	osphor ;	f=0.002	
	+el: na	.me=Sui	lfur ;	f=0.003	
	+el: na	me=Chi	lorine ;	f=0.003	
	+el: na	.me=Pot	tassium ;	f=0.002	
Hydrogen:	S= H	; Z=	1. ; A=	1.01 g/mole	e
Helium:	S= He	; Z=	2. ; A=	4.003 g/mole	е.
Lithium:	S= Li	; Z=	3. ; A=	6.941 g/mole	е.
Beryllium:	S= Be	; Z=	4. ; A=	9.012 g/mole	е.
Boron:	S= B	; Z=	5. ; A=	10.811 g/mole	е.
Carbon:	S= C	; Z=	6. ; A=	12.01 g/mole	е.

• When a CT is used to describe the patient, one has to convert the Hounsfield Units into materials

• Methods have been described for such a socalled stoichiometric calibration, see for instance:

Schneider et al, Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. Phys. Med. Biol. 45 459–78, 2000 • The most computationally expensive calculation, because it needs to be calculated an extremely high number of times !

• In a uniform medium, the exponential distribution based on the mean-free path of the particle gives the distribution of distances a particle travels before an interaction.

• Using the inversion method, we can sample from this distribution using:

d= -ln (u) / λ

where λ is the mean free path and u follows a uniform distribution on (0;1)

• As we are not tracking through a uniform medium, we calculate:

d=-ln(u)

and we adjust to non-uniform media by travelling until:

$$\sum_{i=1}^{D} \mu_i d_i = d$$

 μ_i is the length of particle path across the ith material segment

 μ_i is the attenuation for i <D

$$\sum_{i=1}^{D} \mu_i d_i = d$$

 μ_i is the length of particle path across the i^{th} material segment

 μ_i is the attenuation for i <D

• If the particle leaves the space defined in our simulation before the sum reaches d, we conclude it has escaped without interacting

• Otherwise, we simulate an interaction at position:

$$(x,y,z) + (\sum_{i=1}^{D} d_i) (u,v,w)$$

where (w,y,z) was the initial position of the particle

and (u,v,w) gave the direction of the particle

- Let's consider the example of a photon undergoing electromagnetic interactions
- 4 types of interactions are possible: photoelectric, Compton, Rayleigh, pair production, so that

$$\mu = \mu_{\rm pe} + \mu_{\rm c} + \mu_{\rm r} + \mu_{\rm pp}$$

- To sample which interaction will take place, we:
- sample u from the uniform distribution on [0,1)
- simulate a photoelectric absorption if $0 \le u < \mu_{pe}/\mu$
- simulate Compton scatter if $\mu_{pe}/\mu \le u < (\mu_{pe} + \mu_c)/\mu$
- simulate Rayleigh if $(\mu_{pe} + \mu_c)/\mu \le u \le (\mu_{pe} + \mu_c + \mu_r)/\mu$
- otherwise simulate pair production

Same principle for other particles, as a function of the interactions they can undergo
- In imaging applications:
 - when a photon is absorbed by photoelectric absorption in tissues, it disappears: end of history for this photon
- In dosimetry applications:
 - when a photon is absorbed by photoelectric absorption, the energy dissipation has to be modelled
- In collimators:
 - the energy might be re-emitted as X-rays characteristics of the collimator component (e.g. lead) and these X-rays can be simulated
- In scintillators:

- the energy is re-emitted as scintillation photons, that can also be simulated (but VERY TIME CONSUMING !) • The photon changes direction and loses energy with:

$$E_{out} = \frac{E_{in}}{[1 + E_{in}(1 - \cos \alpha)/511]}$$

where α is the angle between the incoming and outgoing photon directions.

 α can be sampled from the Klein-Nishina distribution:

$$\frac{E_{out}^{2}}{E_{in}^{2}} \left(\frac{E_{in}}{E_{out}} + \frac{E_{out}}{E_{in}} - 1 + \cos \alpha \right)$$

Kahn developed an acceptance-rejection method to sample the KN distribution.

Beware: this distribution does not account for the binding energy of the electron, which becomes a significant factor for low-energy photons in high Z materials: either use an adjustement (Ljungberg et al 1998) or use Look Up Tables (Cullen et al 2010)

• Cullen et al 2010 http://www-nds.iaea.org/epdl97/

[•] Ljungberg. **1998** Introduction to the Monte Carlo method. In Monte Carlo calculations in nuclear medicine, IOP.

• α determines a cone of possible outgoing directions.



• A uniform distribution on $[0, 2\pi)$ is used to sample that cone and determines the final outgoing direction

• At low photon energy, the likelihood of scattering at an angle α with respect to the initial photon direction is proportional to

 $Z(1 + \cos^2 \alpha)$

in an element with atomic number Z.

• As photon energy and Z increase, atomic form factors should be accounted for.

• Look-up tables should be used at high energy and high Z-materials.

- Only for photons with energy > 1 022 keV
- In medical applications, does not occur in biological tissues, but can occur in high-Z materials (lead) for high energy photons (> 2MeV)
- If pair production, the excess energy above 1022 keV is shared between the electron and the positron as kinetic energy.

• Total cross sections are based on highly sophisticated partial-wave analysis calculations which are known to be accurate to much better than 1%. The details of energy sharing between the electron and positron is not of importance for medical physics applications. • In crystals (PET and SPECT)

- same physics as in the patient/phantom, except the material is different

- the energy deposited in the detector is of interest, as its gets converted into scintillations photons that are ultimately collected by the photomultipliers

- most often, one assumes that the gamma energy is deposited locally: good approximation as electrons have a very short range in detector crystals

scintillation photons can be tracked by general purpose codes, but given they are many, this is not computationally efficient.
For instance, a photon of 140 keV in a NaI (Tl) crystal yields 6000 fluorescence photons

- for most applications, simpler models for the detector response can be used • Inelastic collisions with atomic electrons that lead to ionizations and excitations

- Interactions with energy transfer large compared to the binding energies: Møller (e–) or Bhabha (e+) cross sections

- Bethe-Bloch stopping power theory, excellent agreement with measurements

• Bremsstrahlung in the nuclear and electron fields

- Very well understood at high energies (100+ MeV)

- Well understood at low energies (≤ 2 MeV) in terms of partial-wave analysis calculations

- Interpolation schemes in the intermediate energy range, excellent agreement with measurements

• Elastic collisions with nuclei and atomic electrons: very well understood in terms of partial-wave analysis calculations

• Positrons: annihilation

• Simulation of photon transport: easy as the mean number of event in each photon history is small: the photon is absorbed after single photoelectric interaction or after few Compton interactions (typically ~10)

Much more difficult for electrons and positrons: the average energy loss of an electron at each interaction is very small (~ 0.1 eV).
Many interactions occur before the electron is absorbed in the medium.

- Tractable only for low energy electrons (< 100 keV) or thin targets

• Unlike photons, charged particle undergo a huge number of collisions before being locally absorbed (~ 10⁶ for an electron in a typical radiotherapy application)

⇒ Event-by-event simulation is not practical

Fortunately, most interactions lead to very small changes in energy and/or direction
 ⇒ combine effect of many small-change
 collisions into a single, large-effect, virtual
 interaction ⇒ Condensed History (CH)
 simulation (introduced by Berger in 1963)

• The PDF for these large-effect interactions are obtained from suitable multiple scattering theories

• The CH technique is used by all general purpose MC packages and by fast MC codes specializing in RTP calculations

Modeling hadronic interactions

- Many models:
- Bertini model for p and n
- Cascade exciton model
- Binary cascade model
- JAERI quantum molecular dynamic
- JET AA microscopic transport model
- Boltzmann master equation model
- Shen model
- Tripathi model

- ...

The models should be selected as a function of the intended application (mostly energy and particle involved)

• Fortunately, recommendations are usually published to help the user chose the appropriate model in a given code

Recommendations

- Setting the Physics Parameters for Radiation Therapy Applications
- Setting the Physics Parameters for Proton Therapy Applications
- Setting the Physics Parameters for Carbon Therapy Applications

• All physics relies on the physics models, and in the total and differential cross-section tables

• Photon and electron interactions with atoms and molecules are described by Quantum Electrodynamics

• QED is a very well understood physics theory

• Complications at low energies (as energies and momenta are comparable to the binding energies) or at very high energies (radiative corrections, formation time, possibility to create muons and hadrons, etc)

• Interactions are very simple to model in the energy range of interest for emission tomography, transmission tomography and external beam radiotherapy. Far more challenging for hadrontherapy. Useful resources for physics models

• Photon cross-sections database:

http://www.nist.gov/pml/data/xcom/index.cfm

• X-ray form factors, attenuation and scattering tables

http://physics.nist.gov/PhysRefData/FFast/ html/form.html

• Table of nuclides http://atom.kaeri.re.kr/index.html

Stopping power and range tables
 http://physics.nist.gov/PhysRefData/Star/Text/
 ESTAR.html for electrons
 http://physics.nist.gov/PhysRefData/Star/Text/
 PSTAR.html for protons
 http://physics.nist.gov/PhysRefData/Star/Text/
 ASTAR.html for Helium ions

• X-ray transition energies http://www.nist.gov/pml/data/xraytrans/ index.cfm

• IAEA photon and electron interaction data <u>http://www-nds.iaea.org/epdl97/</u>

- Need for accounting for the imperfect detector response in terms of energy resolution, event positionning, and time resolution
- Energy resolution

 $E_{detected} = E_{deposited} + N(0,\sigma)$

where N is a sample for a centered Gaussian distribution with

$$\sigma = \frac{\text{FWHM} \sqrt{\text{E}_0 \text{E}_{\text{deposited}}}}{2\sqrt{2\ln 2}}$$

with FWHM of the energy response measured at energy E_0 .

When multiple energy deposits occur in the same crystal, they should be summed (before or after adding an error term) to get an estimate of the detected energy.

Models for the detector response

• Spatial resolution

Event positioning is usually performed at the energy-weighted centroid of the interaction location in the crystal

An error term (most often Gaussian) can be added (as for energy determination) to account for the limited spatial resolution of the photomultipliers and associated electronics

• Time resolution

Time management is important in a number of applications, especially PET (for coïncidences, time of flight), dynamic scanning, modelling detector motion, modeling tracer distribution, etc...

For these applications, dedicated models of dead time (paralyzable or non-paralyzable), pile up, random coincidences exist Models of the detector response are almost always based on experimental measurements to get key parameters involved in the models:

- energy resolution at a certain energy
- PMT mispositioning
- pile-up
- dead-time coefficients
- handling of multiple coincidence in PET
- (although the company could tell)
- etc

Storing relevant information out of simulations



• Wording:

The storage of relevant information regarding the simulations can be called:

- histogramming
- scoring
- sorting
- binning
- defining tallies

as a function of the software.

Don't get mixed up by these different words.

• The user can usually define the pieces of information that he wants to store

- Simulation strengths: all parameters related to the history of the particles can theoretically be tracked
- Most relevant information include:
- nature of the detected event (unscattered, scattered)
- initial emission location/energy of the detected event
- detected energy
- number of scatter and type of scatter
- location of interactions within the patient and within the detector
- time of decay
- time of detection
- momentum at emission
- momentum at detection
- particle fluence
- amount of energy deposited in a certain volume

Accelerating the simulations



• The computational burden of MC simulations is still the bottleneck for their wide use in research and clinical environments

- Algorithmic approaches
- Variance reduction techniques
- Fictitious interaction tracking
- Convolution forced detection
- Hybrid simulations
- Hardware approaches

• Variance reduction techniques (also called importance sampling)

- Reduce the variance of the estimates without adding bias

- Principle: simulate more of the particles that have a high likelihood to effectively contribute to the end result (e.g., be detected, deposit energy) than the others

- Bias is avoided by assigning weights to the simulated particles: weights are incremented/ reduced as a function of the biased sampling we use

• Includes stratification sampling and forced detection

• Stratification

Simulate particles in proportion to the probability that they will yield a useful event (i.e. not exit the simulation without yielding any useful signal like event detection or energy deposit)

For instance, simulate more particles towards the detector than in the opposite direction



• Stratification Determination of the weights



• Forced detection

When a photon interacts in a patient:

- force a Compton scatter instead of a photoelectric absorption,

- for the scattered photons to travel towards the detector (e.g., within the acceptance angle of the collimator in SPECT)

In FD, the choice of the photon angle is stochastic, so that each scatter point in the object produces a small noisy cloud on the projection



ENTERVISION School - Irène Buvat – june 2012 - 59

- Also called delta tracking or Woodcock tracking
- Unbiased

• Principle: photons of energy E are tracked as if the total linear attenuation coefficient $\mu_i(E)$ is the same for all voxels i and such that:

 $\mu_{rep}(E) \ge \mu_i(E)$

The travel distance to the next interaction (real or fictitious) will always be equal or shorter than in reality.

For each interaction in voxel i with material m = m(i) $R_m(E) = \mu_m(E) / \mu_{rep}(E)$

represents the probability that the interaction is a real interaction.

Using a uniform random number $r \in [0, 1]$, it is then tested if a real or fictitious interaction takes place.

If $r > R_m(E)$, the particle remains unchanged. This occurs more often in less attenuating materials and reduces the number of real interactions to the correct value.

Fictitious interaction (2)

• The heterogeneity of the phantom is incorporated by introducing fictitious interactions instead of recalculating the remaining path length at each voxel boundary like in conventional tracking.

• When the replacement mean free path length is larger than the voxel size, this leads to fewer real and fictitious interactions than voxel boundaries and thus can lead to reduced simulation time.



Rehfeld et al, Phys Med Biol 2009

• Can be used to estimate the scatter distribution quickly

It is based on an estimate of the detector spatial resolution (and biased), but can be very effective



Photons are forced parallel to the collimator hole axis.

Only the photon weights of these photons are calculated and their sum is stored in voxels. Voxels in layers L(x,y) at different distances z from the collimator are convolved with an appropriate distance dependent collimator response that is pre-stored in a table.

- Combining Monte Carlo simulations with analytical models
- For instance, the contribution of the primary (unscatter) component to the projections is calculated using analytical simulations, while the scatter component is calculated using the MC approach
- Can be pretty accurate for some applications



Zaidi et al, Med Bio Eng Comput 2007

- Optimizing the "cuts" in the simulations
- Cuts are value below which the particles are not tracked anylonger. Can be expressed as an energy, or a distance
- There is not "magic" rule: cuts should always be optimized as a function of the application. This is worth it to save computation time



- Parallelization of the code
- MC simulations well suited for parallelization as particle histories are independent from one another.
- Still, attention should be paid to the initialization of the random seeds not to introduce spurious correlations in the data.
- Apart from the simulation initialization steps, the speed-up factor is close to the number of processors that are used.

- Cluster of workstations
- Arrays of transputers

• Geographically distributed platforms (grids) through a grid middleware (does exist for a number of MC simulation tools)



• And now GPU and clusters of GPU

Simulation tools in Emission Tomography and RT



• Two types of tools :

- General purpose tools for modeling particle-matter interactions :

- Geant4
- MCNPX
- EGS4/EGSnrc
- FLUKA
- PENELOPE

- ...

- Dedicated tools
 - BEAMnrc : Radiation therapy only
 - Simset : PET and SPECT
 - SIMIND : SPECT only
 - SORTEO : PET only
 - GATE : PET, SPECT, CT, and Radiation Therapy

-...

GEANT4: a toolbox



 GEANT4: a toolkit for the simulation of the passage of particles through matter GEANT4 stands for GEometry ANd Tracking 4

• GEANT4 is an Opensource software (C++, objectoriented) widely used in high energy, nuclear and accelerator physics, as well as studies in medical and space science

 GEANT4 is developed by an international collaboration (a total of 38 institutes contributed), organized in 16 working groups <u>https://geant4.cern.ch/collaboration/</u> working_groups.shtml

- First release in 1998, current release is Geant4.9.5
- Two reference papers published :

- Nuclear Instruments and Methods in Physics Research A 506: 250-303, 2003.

- IEEE Trans Nucl Sci 53: 270-278, 2006
- Main resource: <u>http://geant4.cern.ch/</u>

A good introduction to GEANT4 : <u>http://geant4.web.cern.ch/geant4/UserDocumentation/</u> <u>UsersGuides/IntroductionToGeant4/html/index.html</u>



 MCNPX: a general-purpose Monte Carlo radiation transport code for modeling the interaction of radiation with everything.

MCNPX stands for Monte Carlo N-Particle eXtended.

 MCNPX is a free software widely used for nuclear medicine, nuclear safeguards, accelerator applications, homeland security, nuclear criticality, outer space applications, deep underground applications, ...

 MCNPX is developed by the Los Alamos International Laboratory

 First release in 1997 Current release is MCNPX 2.7 MCNPX is written in Fortran 90, runs on PC Windows, Linux, and unix platforms, and is fully parallel (PVM and MPI)

 Main resource: <u>http://mcnpx.lanl.gov/</u>

EGS4/EGSnrc

 EGS4: Monte Carlo code for performing simulations of the transport of electrons and photons in arbitrary geometries

EGS stands for Electron Gamma Shower. Initially developed by the Stanford Linear Accelerator Center (SLAC). Now maintained by the KEK, the japanese High Energy Accelerator Research Organization

 EGSnrc is an extended and improved version of the EGS4 package originally developed at SLAC. EGSnrc is developed by the National Research Council Canada
 significant improvements in the implementation of the condensed history technique for the simulation of charged particle transport and better low energy cross sections

• EGS4 released in 1984, EGSnrc released in 2000 EGS4 is written in MORTRAN, EGSnrc is written in C++ and runs on Unix, Linux, Windows and, to some extent, on Mac OS X.

- Two reference papers published :
- Medical Physics 27: 485-498, 2000
- Medical Physics 27: 499-513, 2000

Main resource: <u>http://irs.inms.nrc.ca/software/egsnrc/</u>



 FLUKA: Fully integrated particle physics Monte Carlo simulation package

Developed by 16 labs around the world.Written in Fortran 77

FLUKA can simulate with high accuracy the interaction and propagation in matter of about 60 different particles, including photons and electrons from 1 keV to thousands of TeV, neutrinos, muons of any energy, hadrons of energies up to 20 TeV and all the corresponding antiparticles, neutrons down to thermal energies and heavy ions. The program can also transport polarised photons (e.g., synchrotron radiation) and optical photons.

"The FLUKA code: Description and benchmarking" Battistoni et al, Proceedings of the Hadronic Shower Simulation Workshop 2006,

Fermilab 6--8 September 2006, M. Albrow, R. Raja eds., AIP Conference Proceeding 896, 31-49, (2007)

"FLUKA: a multi-particle transport code" A. Ferrari, P.R. Sala, A. Fasso`, and J. Ranft, CERN-2005-10 (2005), INFN/TC_05/11, SLAC-R-773 <u>http://www.fluka.org/fluka.php</u>

Main resource:
PENELOPE: a Monte Carlo algorithm and computer code for the simulation of coupled electron-photon transport
 Developed by the Universotry of Barcelona and distributed by the Nuclear Energy Agency

 PENELOPE: acronym for PENetration and Energy LOss of Positrons and Electrons (photon simulations were introduced later)

• First version released in 1996, written in Fortran

 Main resource: <u>http://www.nea.fr/abs/html/nea-1525.html</u> where all papers can be downloaded.

Useful comparison between general purpose codes



- Slightly out of date but still useful:
- Compares the codes in terms of :
 - general information
 - geometry capabilities
 - source
 - physics capabilities
 - scoring and results
 - variance reduction techniques

- Advantages :
 Large community of users
 - Significant manpower involved in code development
 - Many users: extensive testing of the codes
 - Documentation, user support and training



- Limitations:
- Complex as they support many applications
- No necessarily easy to adapt to one's own application

 Monte Carlo simulation system for modelling radiotherapy sources which is based on the <u>EGSnrc</u> code system for modelling coupled electron and photon transport

- Supported by the National Research Council Canada
- BEAMnrc released in 2001, works on Linux, Unix, Windows, and possibly Mac OSX. Fortran F77 and C
- Free for non-commercial users
- Reference paper published :

- D. W. O. Rogers, B. A. Faddegon, G. X. Ding, C.-M. Ma, J. Wei, and T. R. Mackie, BEAM: A Monte Carlo code to simulate radiotherapy treatment units, Medical Physics 22: 503-524, 1995

- complete list of publication on the web site

 Main resource: <u>http://irs.inms.nrc.ca/software/beamnrc/</u> Dedicated to simulations in SPECT and PET



- SimSET stands for SIMulation System for Emission Tomography
- Developed by the University of Washington (Robert Harrison)
- First released in 1993, freely available, written in C, runs on Unix, Linux, Mas OS
- Does not explicitly manage time (so specific functions to model TOF, random coincidences, some limitations regarding the modelling of the scanners
- Main resource: http://depts.washington.edu/simset/html/ simset_main.html

 Dedicated to simulations in planar scintigraphic imaging and SPECT

 Developped by Professor Michael Ljungberg, Lund University, Sweden

first released in

• written in FORTRAN-90, runs on Linux, Mac OS 10, and Windows.

Main resource:

Ljungberg, M. and Strand, S.-E. A Monte Carlo Program Simulating Scintillation Camera Imaging. Comp.Meth.Progr.Biomed. 29, 257-272. 1989 <u>http://www2.msf.lu.se/simind/</u>



 GATE: Geant4 Application for Emission Tomography, Transmission Tomography and Radiotherapy

• GATE is an Opensource software dedicated to the simulation of imaging (SPECT, PET, CT) and radiotherapy, and based on the Geant4 toolbox

• GATE is developed by the international OpenGATE collaboration, including 17 labs

- First release of GATE in May 2004 17 releases since that date Currently: GATE V6.1
- Broad range of applications:
 - Detector design
 - Optimisation of acquisition and processing protocols
 - Assessment of quantification methods
 - Estimation of the system matrix used in tomographic reconstruction
 - Dosimetry
 - Radiation therapy (conventional and hadrontherapy)

- ...

- GATE is based on Geant4: <u>http://www.geant4.org</u>
- GATE is written in C++
- GATE is user-friendly as simulations can be designed and controlled using macros, without any C++ writing

• GATE can simulate SPECT, PET and CT scans and radiotherapy treatments

• GATE is flexible enough to model almost any detector design, including prototypes

• GATE explicitly models time, hence makes it possible to model detector motion, patient motion, radioactive decay, optical photon tracking, dead time, time of flight, tracer kinetics

- GATE can handle analytical or voxelized phantoms
- GATE can run on a cluster architecture and on a grid
- GATE can be freely downloaded, including sources
- GATE can be run on many platforms (Linux, MAC OS, Windows)
- GATE is also distributed as a virtual image

- Online documentation (wiki) about GATE, including FAQ
- Help about the use of GATE can be obtained through the gate-user mailing list (more than 1200 subscribers)

• Archives of the gate-user posts on <u>http://dir.gmane.org/gmane.comp.science.opengate.user</u>

• Many commercial or prototype systems have already been modeled using GATE and most models have been thoroughly validated (list available on <u>http://www.opengatecollaboration.org</u>)

• The GATE project is mostly based on volunteer participation and on the active contribution of GATE developers and users - GATE URL: http://www.opengatecollaboration.org

- GATE user mailing list: gate-users@lists.opengatecollaboration.org

GATE documentation (wiki):
To install GATE:
<u>http://www.opengatecollaboration.org/InstallingGATE</u>
To use GATE:
<u>http://www.opengatecollaboration.org/Documentation</u>

- GATE publications:

http://www.opengatecollaboration.org/Publications

Jan et al, Physics in Medicine and Biology 2004: 4543-4561

INSTITUTE OF PHYSICS PUBLISHING

Phys. Med. Biol. 49 (2004) 4543-4561

PHYSICS IN MEDICINE AND BIOLOGY

PII: S0031-9155(04)80763-2

GATE: a simulation toolkit for PET and SPECT

S Jan¹, G Santin^{2,24}, D Strul^{2,25}, S Staelens³, K Assié⁴, D Autret⁵, S Avner⁶, R Barbier⁷, M Bardiès⁵, P M Bloomfield⁸, D Brasse⁶, V Breton⁹, P Bruyndonckx¹⁰, I Buvat⁴, A F Chatziioannou¹¹, Y Choi¹², Y H Chung¹², C Comtat¹, D Donnarieix^{9,13}, L Ferrer⁵, S J Glick¹⁴, C J Groiselle¹⁴, D Guez¹⁵, P-F Honore¹⁵, S Kerhoas-Cavata¹⁵, A S Kirov¹⁶, V Kohli¹¹, M Koole³, M Krieguer¹⁰, D J van der Laan¹⁷, F Lamare¹⁸, G Largeron⁷, C Lartizien¹⁹, D Lazaro⁹, M C Maas¹⁷, L Maigne⁹, F Mayet²⁰, F Melot²⁰, C Merheb¹⁵, E Pennacchio⁷, J Perez²¹, U Pietrzyk²¹, F R Rannou^{11,22}, M Rey², D R Schaart¹⁷, C R Schmidtlein¹⁶, L Simon^{2,26}, T Y Song¹², J-M Vieira², D Visvikis¹⁸, R Van de Walle³, E Wieërs^{10,23} and C Morel²

Article freely available on http://www.guillemet.org/irene/article/gatepmb.pdf

Jan et al, Physics in Medicine and Biology 2011: 881-901

IOP PUBLISHING

Phys. Med. Biol. 56 (2011) 881-901

PHYSICS IN MEDICINE AND BIOLOGY

doi:10.1088/0031-9155/56/4/001

GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy

S Jan¹, D Benoit², E Becheva¹, T Carlier^{3,4}, F Cassol⁵, P Descourt⁶, T Frisson⁷, L Grevillot⁷, L Guigues⁷, L Maigne⁸, C Morel⁵, Y Perrot⁸, N Rehfeld², D Sarrut⁷, D R Schaart⁹, S Stute², U Pietrzyk¹⁰, D Visvikis⁶, N Zahra⁷ and I Buvat²

Article freely available on http://www.guillemet.org/irene/article/jan2011.pdf

• There are papers comparing code performances, or validation a code against another

• There is no "best" code: the best code depends on the intended application

• Example of code comparison:

. . .

- For PET applications: Buvat et al. Unified description and validation of Monte Carlo simulators in PET. Phys Med Biol 2005

For proton therapy applications: Seravalli et al.
 Monte Carlo calculations of positron emitter yields in proton radiotherapy. Phys Med Biol 2012

- For conventional radiotherapy: Comparison of GATE/GEANT4 with EGSnrc and MCNP for electron dose calculations at energies between 15 keV and 20 MeV. Phys Med Biol 2011

Numerical models for patient description



• Analytical: the phantom/patient is described using a set of geometric primitives (box, cylinders, spheres, ellipsoids)

• Voxelized description: a series of images (i.e. a volume)



from Handbook in Particle Detection and Imaging, Springer, 2012, p 1106

Analytical models

- The MCAT, NCAT, XCAT family
- History

MIRD5 1969 dosimétrie







4D-MCAT 1999-2001 SPECT and PET cardiac and respiratory motions anatomical variability 4D-NCAT 2001 SPECT and PET NURBS representation

NCAT and XCAT: NURBS representation

- NURBS : Non-Uniform Rationale B-Splines
- Primitives allowing for modeling complex curves and 3D surfaces
- Easy to parameterized volumes using (affine) transformation applied to control points



• The NURBS primitives have been parameterized (using the Rhinoceros software) so as to describe the shapes of organs as seen in the « 3D visible human CT »*, after manual segmentation of the different organs

* http://www.nlm.nih.gov/research/visible/visible_human.html

• The XCAT phantom

• Initially, the NCAT was developed for Nuclear Medicine applications. The motivation for the XCAT was to go beyond these low spatial resolution applications

• Combinaison of NURBS and subdivision surfaces (SD)

• Based on the « visible male and female anatomical datasets » of the National Library of Medicine (0.33 x 0.33 x 1 mm³ sampling for the male model, 0.33 x 0.33 x 0.33 mm³ sampling for the female model)

• More than 9000 anatomical structures



Principle of the subdivision surfaces

• To model structure with an arbitrary topological type (brain structure for instance), while NURBS can only model such structures by partitioning the model into a large collection of individual NURBS surfaces (i.e. many parameters).



Initialisation by a coarse polygon mesh



A refinement scheme is used to subdivide and smooth the mesh to produce a smooth surface representation of the object

XCAT: modeling of the cardiac motion

• Model derived from a dynamic CT (100 image per cardiac cycle in the male model, 12 in the female model, using 0.32 x 0.32 x 0.4 mm voxels sampling)

• Includes many anatomical details: coronary vessels, papillary muscles, valves



• This model can be used to simulate heart contraction, torsion, ejection fraction, cardiac twist, heart rates, etc in normal and abnormal hearts.

Modeling of respiratory motion

• Model created based on 30 4D-CT scans, each including 20 respiratory phases.



End-expiration

End-inspiration



ENTERVISION School - Irène Buvat - june 2012 - 93

XCAT: modeling the anatomic variability

- The different structures are parameterized.
- Parameters can be adjusted to model various anatomical models



• Bottleneck: manual parameterization (using the Rhinoceros software), which is extremely tedious and time-consuming: only 30 organs are deformed / 9000 to generate various body habitus.

• Automatic parameterization still has to be implemented: on-going effort using the Large Deformation Diffeomorphic Metric Mapping approach

MCAT family: assigning properties

• Density and composition: values are derived from the CT + anatomical knowledge

• For imaging applications: which tracer uptake in which organ?

- tabulated values in the literature for some tracers (FDG in PET for instance*)

- values estimated from the distributions of values observed in a patient dataset

- values estimated from a single scan that one wishes to reproduce

*

- Ramos et al 2001 FDG-PET standardized uptake values in normal anatomical structures using iterative reconstruction segmented attenuation correction and filtered back-projection *Eur. J. Nucl. Med. Mol. Imaging.* **28** 155–164

- Wang et al 2007 Standardized uptake value atlas: characterization of physiological F18-FDG uptake in normal tissues Mol. Imaging. Biol. **9** 83-90

- Zincirkeser et al, 2007 Standardized uptake values of normal organs on F18-FDG positron emission tomography and computed tomography imaging J. Int. Med. Res. **35** 231-6

MCAT family: advantages and drawbacks



- Fine anatomical modelling (for instance, arterial tree)
- Cardiac and respiratory motions can be easily included
- Based on the continuous NURBS representation, images can then be obtained for any sampling without any approximation









- Hard to fit a given patient scan (feasible, but extremely tedious*)
- The activity / attenuation distribution is piece-wise constant (or a structure has to be subdivided so as to introduce heterogeneities in it)
- Models of lesions have to be developed and combined with the phantoms

* Le Maitre et al **2009** Incorporating patient-specific variability in the simulation of realistic F18-FDG distributions for oncology applications *Proceedings of the IEEE* **97** 2026–2038

Phantoms for small animal imaging

• Following the NCAT strategy for building a phantom, a mouse phantom (MOBY) and a rat phantom (ROBY) have also been developed



• Note that these phantoms are not free (nor those from the NCAT family)

Advanced voxelized models

- Based on real patient scans
- In each voxel:

- density can be assigned based on the Hounsfield Units of a CT scan of the patient

- composition can be assigned by image segmentation (bone, lungs, soft tissues, myocardium, aso)

- activity in each voxel can be obtained for HIGH-RESOLUTION real PET or SPECT scan

- motion can be modelled based on a gated scan (e.g., respiratory gated CT, cardiac gated SPECT)



CT: density and composition



PET and SPECT: activity distribution



• All patients features can be closely reproduced, including heterogeneous uptake within an organ



• PET, SPECT and CT images are not perfect (limited spatial resolution, noisy): errors in the simulation input might propagate during the simulation

Noteworthy: how do the errors in PET and SPECT images used to define a voxelized activity distribution propagate into simulated PET and SPECT images?

Voxelized models: propagation of errors



Voxelized models: propagation of errors



VERY IMPORTANT: when using patient PET or SPECT scans to define the input activity distribution of a simulation, these images should have a very high spatial resolution, regardless of the level of noise, to ensure final spatial resolution in the reconstructed images identical to that used in real patient scans.

Hence, images should be reconstructed with a high number of iterations

•Stute et al **2011** Monte Carlo simulations of clinical PET and SPECT scans: impact of the input data on the simulated images. Phys. Med. Biol. 56: 6441-6457

• Segars PW, Tsui BMW 2009 Evolution of 4-D computerized phantoms for imaging research *Proceedings of the IEEE* **97** 1954-1968

• Le Maitre et al 2009 Incorporating patient-specific variability in the simulation of realistic F18-FDG distributions for oncology applications *Proceedings of the IEEE* **97** 2026–2038

• Zaidi et al 2009 Review of computational anthropomorphic anatomical and physiological models *Proceedings of the IEEE* **97** 1938-1953

• Stute al 2011 Monte Carlo simulations of clinical PET and SPECT scans: impact of the input data on the simulated images. Phys. Med. Biol. 56: 6441-6457

• Stute et al 2012 Realistic and efficient modeling of radiotracer heterogeneity in Monte Carlo simulations of PET images with tumors. IEEE Trans. Nucl. Sci. 59: 113-122

<u>http://www.virtualphantoms.org/index.html</u>

• Zubal anthropomorphic phantom (torso and head) <u>http://noodle.med.yale.edu/zubal/</u>

• NCAT family http://www.bme.unc.edu/~wsegars/phantom.html

 Visible Human project <u>http://www.virtualphantoms.org/index.html</u>

More reading about numerical models

• Book:

-Handbook of anatomical models for radiation dosimetry, Xu and Eckerman eds, CRC Press, 2008



• Studying and optimizing detector design to maximize performance (e.g;, as a function of the detector geometry, time resolution, crystal, aso)

• Identifying the origin of biases in the images and measuring their respective magnitudes (scatter, septal penetration, motion, ...)

• Optimizing acquisition and processing protocols (corrections, post-processing)

• Assisting image production, for instance when used to calculate the system matrix in SPECT or PET reconstruction, or when used for scatter correction

• Determining the accuracy with which specific parameters can be estimated from the images

• Evaluation of the effect of physical, geometrical and design parameters on the scanner performance

• Estimate physical parameters that cannot be measured, like scatter fraction in different detector components

• Validate corrections, such as scatter correction methods, beam hardening correction algorithms

• Generate data for testing reconstruction algorithms in realistic conditions

• Absorbed dose calculations to assess the radiobiological risks in CT scans

• . . .

Applications of MC simulations in radiation therapy

- Validation of Treatment Planning System (that are not full MC)
- Calculation of data for TPS (e.g., dose point kernels)
- Treatment monitoring in hadrontherapy
- Design and commissioning of therapy facility
- Assisting the design of the nozzle
- Simulation of the neutron background
- ...

More reading about MC simulations

• Books:

-Monte Carlo calculations in nuclear medicine Ljungberg, Strand, King eds, IOP, 1998



A new (enhanced) edition of the book is about to be published
Useful references (2)

• Books:

-Therapeutic Applications of Monte Carlo calculations in Nuclear Medicine, Zaidi and Sgouros eds, IOP, 2002



ENTERVISION School - Irène Buvat – june 2012 - 109

Useful references (3)

• Books:

- Handbook of Particle Detection and Medical Imaging, Grupen and Buvat eds, Springer 2012



ENTERVISION School - Irène Buvat – june 2012 - 110