Monte Carlo Simulations: Efficiency Improvement Techniques and Statistical Considerations

Daryoush Sheikh-Bagheri, Ph.D.¹, Iwan Kawrakow, Ph.D.², Blake Walters, M.Sc.², and David W.O. Rogers, Ph.D.³

¹Department of Radiation Oncology, Allegheny General Hospital
Pittsburgh, Pennsylvania

²Ionizing Radiation Standards, National Research Council,
Ottawa, Ontario, Canada

³Physics Department, Carleton University, Ottawa, Ontario, Canada
If linac simulations take too long ...

- Divide the beam into treatment-independent and treatment-dependent components
- Simulate treatment-independent components
  - characterize phase space distribution with a beam model
- Simulate treatment-dependent components and the patient CT together
If linac simulations can be made fast enough ...
Do all at once ...

• Simulate treatment-independent linac components, treatment-dependent components and the patient CT together
Metrics of Efficiency

\[ \epsilon = \frac{1}{\sigma^2 T} \]

**\( T \):** computing time to obtain a variance \( \sigma^2 \)

**\( \sigma^2 \):** variance on the quantity of interest
Q: How can one increase the efficiency?

A: By reducing the computing time that it takes to obtain a sufficiently small variance on the quantity of interest

\[ \epsilon = \frac{1}{\sigma^2T} \]

... easier said than done!
Variance of what?

- Variance of a quantity of interest averaged over a region
- Examples:
    \[
    \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\Delta D_i}{D_i} \right)^2, \quad \text{with } D_i > 0.5 \ D_{\text{max}}
    \]
  - fluence in 1x1 cm² regions in beam
  - dose on central axis or profile, etc.
Statistical Uncertainties

• Without them MC calculated values would be ... useless
• Prerequisite to efficiency estimation
• Central limit theorem
• The batch method
• The history-by-history method
• Pick independent particles ... otherwise correlation
• Only those particles are independent that belong to different histories
• Note particle’s origin when recycling phase-space files
• Latent Variance
Histories ...

- Treatment head (& air)
- Primary
- Patient
- Electron contamination
- Phantom scatter
- Monitor backscatter
- Head scatter
Uncertainties: Computational Considerations

\[ \sigma_X = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N(N-1)}} \]

\[ \sigma_X^2 = \frac{\langle X^2 \rangle - \langle X \rangle^2}{N - 1} \]

\[ \langle X \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i, \]

\[ \langle X^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \]
Making the history-by-history technique computationally feasible

- Trick by Salvat

```plaintext
IF(HIST.NE.LASTHI(K)) THEN
  Q(K)  = Q(K)+QTEMP(K)
  Q2(K) = Q2(K)+QTEMP(K)**2
  QTEMP(K) = DELTAQ
  LASTHI(K) = HIST
ELSE
  QTEMP(K) = QTEMP(K)+DELTAQ
ENDIF

IF(nhist=X_last) THEN
  X_tmp=X_tmp+delta
ELSE
  X=X+X_tmp
  X2=X2+(X_tmp)**2
  X_tmp=delta
  X_last=nhist
ENDIF
```

Sempau et al, Phys Med Biol 46:1163-1186
Latent Variance

• Divide the dose calculation into 2 phases; A, B
  “A” -> the linac simulation resulting in a phase-space
  “B” -> the dose calculation using the phase-space

\[ \sigma^2(\bar{q}) = \frac{1}{N}(A + B) \]

\[ A = \sum_b \langle q_b \rangle^2 \langle n_b^2 \rangle + \sum_{a \neq b} \langle q_a \rangle \langle q_b \rangle \langle n_a n_b \rangle - \langle q \rangle^2 \]

\[ B = \sum_b \sigma^2(q_b) \langle n_b^2 \rangle \]

\[ \sigma_k^2(\bar{q}) = \frac{1}{N}(A + BK^{-1}) \]

Sempau et al, Phys Med Biol 46:1163-1186
History-by-history and batch methods

Walters, Kawrakow and Rogers, Med Phys 29: 2745-2752
Advantage of history by history

3594 voxels in both cases

10 batches

history by history

number of voxels/bin

Dose in a brachytherapy phantom: from Gultekin Yegin
Codes used in radiotherapy

- ITS
- MCNP
- PENELLOPE
- GEANT4
- No VRTs -> EGS and ITS/ETRAN same efficiency
- Other systems slower
- BEAMnrc code significantly more efficient, still not fast enough for routine RTP
BEAMnrc

- a general purpose user-code for simulation of radiotherapy beams
- built on EGSnrc
- freely available for non-commercial use
- lots of built in variance reduction to enhance efficiency, especially for accelerator photon beams
Codes designed to be more efficient

- The Macro Monte Carlo (MMC) code
- The PEREGRINE code
- Voxel Monte Carlo (VMC/xVMC)
- VMC++
- MCDOSE
- The Monte Carlo Vista (MCV) code system
- The Dose Planning Method (DPM)
- and other codes (Keall and Hoban 1996; Wang, Chui, and Lovelock 1998).
Comparative accuracy of dose calculation

18 MV beam
1.5x1.5 cm² beam
## How fast are current codes?

<table>
<thead>
<tr>
<th>Monte Carlo code</th>
<th>Time estimate (minutes)</th>
<th>% max. diff. relative to ESG4/PRESTA/DOSXYZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG4/PRESTA/DOSXYZ</td>
<td>42.9</td>
<td>0, benchmark calculation</td>
</tr>
<tr>
<td>VMC++</td>
<td>0.9</td>
<td>± 1</td>
</tr>
<tr>
<td>MCDOSE (modified ESG4/PRESTA)</td>
<td>1.6</td>
<td>± 1</td>
</tr>
<tr>
<td>MCV (modified ESG4/PRESTA)</td>
<td>21.8</td>
<td>± 1</td>
</tr>
<tr>
<td>RT_DPM (modified DPM)</td>
<td>7.3</td>
<td>± 1</td>
</tr>
<tr>
<td>MCNPX</td>
<td>60.0</td>
<td>max. diff. of 8% at Al/lung interface (on average ± 1% agreement)</td>
</tr>
<tr>
<td>Nomos (PEREGRINE)</td>
<td>43.3*</td>
<td>± 1*</td>
</tr>
<tr>
<td>GEANT 4 (4.6.1)</td>
<td>193.3**</td>
<td>± 1 for homogeneous water and water/air interfaces**</td>
</tr>
</tbody>
</table>

*Note that the timing for the PEREGRINE code also includes the sampling from a correlated-histogram source model and transport through the field-defining collimators. **See Poon and Verhaegen (2005) for further details.

AEIT vs VRT

• Distinguish between a technique that achieves the improved efficiency through the use of approximations

→ approximate efficiency improving technique (AEIT)

• And a technique that does not alter the physics in any way when it increases the efficiency

→ true variance reduction technique (VRT)
AEITs used in the treatment head simulation

- Condensed History Technique (CHT)
- Range Rejection
- Transport Cutoffs
Condensed History Technique (CHT)

In previous talk Iwan talked about this in detail...
Condensed History Technique (CHT)

- $10^6$ elastic and inelastic collisions until locally absorbed
- Berger (1963) introduced the condensed history technique
- “step-size” dependence
- Is an AEIT
- Two main components very strongly influence the simulation speed and accuracy:
  - the “electron-step algorithm”
    ( “transport mechanics” )
  - the boundary-crossing algorithm
Range Rejection

- Discard an electron if its residual range is smaller than the distance to the nearest boundary.
- Region Rejection: Discard more aggressively when “far” away from the region of interest.

- Suggested 1.5 MeV cutoff for 6 MV and up
- By tagging bremsstrahlung photons generated outside target
- Speed up -> a factor of 3
- Negligible (< 0.2%) underestimation of the calculated photon fluence
Transport Cutoffs

• Do not transport further, if the energy drops below a certain threshold (ECUT & PCUT)
• Do not create secondaries if their energy is going to be below a certain threshold (AE & AP)
Splitting and Russian Roulette

- Originally proposed by J. von Neumann and S. Ulam
- The most powerful VRTs used in Treatment Head Simulations
Splitting and Roulette; a schematic

marks termination of an e- or e+ or both by Russian Roulette

incident e- beam

target

collimator
Splitting, Roulette & Particle Weight

\[ 1 \, w_i = 10 \, w_f \]

\[ 10 \, w_i = 1 \, w_f \]

Split! Roulette!
Weight Management for:

Splitting and Russian Roulette

From Rock Mackie

Courtesy of Jinsheng Li, Fox Chase CC
Splitting-based VRTs developed for BEAM/BEAMnrc

- Uniform Bremsstrahlung Splitting (UBS)
- Selective Bremsstrahlung Splitting (SBS)
- Directional Bremsstrahlung Splitting (DBS)
Electrons incident on and transported in the tungsten target ...
No Splitting
And the resulting brem photons ...
Uniform Brems Splitting

Particle weights: \( \frac{1}{N} \)
Uniform Brems Splitting

relative efficiency

russian roulette on

nsplit

no russian roulette

20  100  1000

AUG
Selective Brems Splitting

Particle weights:
Vary between $1/N_{\text{min}}$ and $1/N_{\text{max}}$
Fat Particles

Distribution that has fat particles

No fat particles
Selective Brems Splitting

Sheikh-Bagheri (1999)

Probability of photon emission towards treatment field

\[ P(R_e, E_e, R_f, SSD) = \frac{\int_{\theta_{\text{min}}}^{\theta_{\text{max}}} d\sigma/d\theta_p f(\theta_p) d\theta_p}{\int_{0}^{\pi} d\sigma/d\theta_p d\theta_p} \]

\[ \theta_p/\theta \]

SSD

bremsstrahlung

treatment field

Sheikh-Bagheri (1999)
Selective Brem Splitting (SBS)

![Graph showing relative efficiency vs. nsplit for SBS and UBS with and without Russian roulette.]

- SBS
- Russian roulette on
- No Russian roulette
- UBS
The evolution of splitting routines

No Splitting

SBS
- $e^-$ is aiming off FOI

UBS

DBS
- "Fat" photon surviving Russian Roulette
- Photons emitted (and transported) toward FOI

- $e^-$ is aiming at FOI
**Directional Brem Splitting (DBS)**

- **goal:** all particles in field when reach phase space have same weight

**Procedure**

i) brem from all fat electrons split nsplit times

ii) if photon aimed at field of interest, keep it, otherwise Russian roulette it:

   if it survives, weight is 1 (i.e. fat)

iii) if using only leading term of Koch-Motz angular dist’n for brem: do_smart_brems and similar tricks for other interactions
do_smart_brems

do_smart_brems calculates how many of the nsplit brem photons will head to the field and only generates those photons;
+
samples 1 photon from the entire distribution (if not heading into the field, kept with weight 1)
Probability of photon heading at field

```
  1.000
  0.100
  0.010
  0.001

probability

E / MeV
```

- The graph shows the probability of a photon heading at a specific field as a function of energy (E) in MeV.
- The x-axis represents energy in MeV, ranging from 0 to 6.
- The y-axis represents probability, ranging from 0.001 to 1.000.
- The curve indicates an increasing probability with energy.
DBS (cont)

Play similar tricks for other quantities

• e+ annihilation: \( \text{(uniform\_photons)} \)

• Compton scattering:
  \( \text{(do\_smart\_compton if Klein Nishina)} \)

• Pair production/photo-effect: \( \text{(Russian roulette before sampling)} \)

• Fluorescence: \( \text{(uniform\_photons)} \)
DBS (cont)

Photons
- reaching field have weight 1/nsplit
- outside field are fat

Electrons in the field
- usually fat
- a few have weight 1/nsplit from interactions in the air
Efficiency of fluence calcs

Russian roulette on

relative efficiency

nsplit

20

100

1000

10000

DBS

SBS

UBS
Efficiency of phantom dose calcs

- DBS (nsplit=5000)
- SBS (nsplit=4000)
- UBS (nsplit=40)

russian roulette on

relative efficiency vs. depth in phantom /cm
6 MV, 10x10 cm²

5.5 min CPU on 2GHz AMD 2400+
NBRSPL=1000, split 10
splitting field radius 10 cm

relative dose

scoring radius 2 cm

depth /cm
Electron contamination ...

- Primary
- Phantom scatter
- Head scatter
- Monitor backscatter
- Treatment head (& air)
- Patient
Electron problem

Unlike UBS and SBS, DBS efficiency gain for electrons is only 2

*Basis of the solution*

-electrons are, almost entirely, from flattening filter and below

-major gains are from “taking care” of electrons in primary collimator
Electron solution

Introduce 2 planes

Splitting plane: split weight 1 charged particles nsplit times

(may distribute symmetrically)

Russian roulette turned off below a certain plane and all fat photon interactions split nsplit times
Efficiency increase for $e^-$
Efficiency: total dose

- DBS no fat photons
- DBS - with fat photons
- SBS
- UBS

relative efficiency vs depth in phantom (cm)
DBS summary

DBS improves BEAMnrc’s efficiency by a factor of 800 (10 vs SBS) for photon beams (ignore small dose from photons outside field).

For total dose calculations the efficiency improves by a factor of 150 (5 vs SBS)

SBS is optimized for greater nsplit than previously realized (5000)
Efficiency Improvement Techniques Used in Patient Simulations

- Macro Monte Carlo
- History Repetition
- Boundary-Crossing Algorithms
- Precalculated Interaction Densities
- Woodcock Tracing
- Photon Splitting Combined with Russian Roulette
- Simultaneous Transport of Particle Sets (STOPS)
- Quasi-Random Sequences
- Correlated Sampling
Macro Monte Carlo (MMC)

Kugel
- Incident electron Energy $T_i$
- Emerging electron Energy $T_f$
- electron track

Ellipsoid
- $r, Z, \rho$
- $\alpha$, $\theta$
- $\varepsilon(r, Z, \rho, T_i)$

Centered Kugel
- $r, Z, \rho$
- $\alpha$, $\theta$
Macro Monte Carlo

Kugels  Ellipsoids  Centered Kugels

Electron direction at exit-position
Energy deposition path
Electron Track Repeating

Courtesy of Jinsheng Li, Fox Chase CC
STOPS
(Simultaneous Transport Of Particle Sets)

- Several particles that have the same energy (but not position, direction, weight) form a “Particle Set” and are transported simultaneously.
- This allows material independent quantities such as interpolation indices, azimuthal angles, maximum acceptable step-lengths, etc., to be calculated just once for the set.
- Material dependent quantities such as MS angles and discrete interaction probabilities are sampled separately.
- In particular, if one or more particles in the set undergo a different interaction, set is split into separate sets and each new set transported individually.
Faster Convergence Using Quasi RNs

\[ I = \int_{a}^{b} f(x) \, dx \]

\[ I_m = (b-a) \frac{1}{N} \sum_{i=1}^{N} f(x_i) \]

\[ \lim_{N \to \infty} I_m = I \]
Pseudo-Random vs. Quasi-Random
Pseudo-Random vs. Quasi-Random in 3D
Less Frequently Used VRTs

- Forcing
- CNVR
- Correlated Sampling
- Exponential Transform
Photon interaction forcing

- **Force to interact in a phantom**
  \[ N_\lambda = - \ln \{1 - R[1 - e^{-M_\lambda}]\} \]
  
  \( M_\lambda \) is the thickness of the phantom in number of mean free paths

  The new photon weight: \( W' = W \{1 - e^{-M_\lambda}\} \)

- **Force to interact in a region of a phantom**
  \[ N_\lambda = M_{\lambda 1} - \ln \{1 - R[1 - e^{(M_{\lambda 1} - M_{\lambda 2})}]\} \]
  
  \( M_{\lambda 1} \) is the number of mean free paths to the near boundary of the region and \( M_{\lambda 2} \) to the far boundary of the region.

  The new photon weight: \( W' = W \{e^{-M_{\lambda 1}} - e^{-M_{\lambda 2}}\} \)

Courtesy of Jinsheng Li, Fox Chase CC
Combine Electron Track Repeating with Photon Interaction Forcing and Splitting

Courtesy of Jinsheng Li, Fox Chase CC
CNVR Technique

• Forces the primary photon fluence to be invariant with depth
Correlated Sampling

Main histories

Split histories
Exponential Transform

- Bias the sampling procedure to interact in the regions of interest

\[ N_\lambda = -\beta \ln R \]
\[ \beta = \frac{1}{(1 - C \cos \theta)} \]

C is defined by the user, \( \theta \) is the angle the photon makes with the direction of interest

The new weighting factor: \( W' = W C e^{-N_\lambda \alpha \cos \theta} \)

\( C < 0 \): smaller \( N_\lambda \) for surface problem, shortening
\( 0 < C < 1 \): larger \( N_\lambda \) for shielding problem, stretching

Courtesy of Jinsheng Li, Fox Chase CC
Stretched ($C = 1/2$) and shortened ($C = -1$) distribution compared to an unbiased one ($C = 0$).
(From A.F. Bielajew and D.W.O. Rogers)
Use of Symmetry

Courtesy of Jinsheng Li, Fox Chase CC
PDD for Photon Beams

- **DOSXYZ**
- **MCDOSE, 6MeV**
- **MCDOSE, 15MV**
- **MCDOSE, 2MeV**

- Lung Tissue
- Tissue

<table>
<thead>
<tr>
<th>Depth Z (cm)</th>
<th>Dose (pGy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 5 10 15 20</td>
<td>0 2 4 6 8</td>
</tr>
</tbody>
</table>

- Without VRT
- With VRT

**Courtesy of Jinsheng Li, Fox Chase CC**
To split, or not to split: ... that is the question!

- "Sheikh"speare