

Introducing improved voxel navigation and fictitious interaction tracking in GATE for enhanced efficiency

Niklas S Rehfeld¹, Simon Stute¹, John Apostolakis², Marine Soret³
and Irène Buvat¹

¹ IMNC-UMR 8165 CNRS, Paris 7 University, Paris 11 University, 15 rue Georges Clémenceau, 91406 Orsay Cedex, France

² PH Department, CERN, CH-1211 Geneva 23, Switzerland

³ Hôpital Inter Armées du Val-de-Grâce, 74 Boulevard de Port Royal, 75230 Paris Cedex 05, France

E-mail: rehfeld@imnc.in2p3.fr

Received 15 September 2008, in final form 28 January 2009

Published 17 March 2009

Online at stacks.iop.org/PMB/54/2163

Abstract

Geant4 Application for Emission Tomography (GATE) is a widely used, well-validated and very versatile application for Monte Carlo simulations in emission tomography. However, its computational performance is poor, especially for voxelized phantoms, partly due to the use of a very general particle tracking algorithm. In this work, two methods are proposed to reduce the time spent on particle tracking in the phantom: a newly introduced ‘regular navigation algorithm’ of Geant4 and fictitious interaction tracking (also known as Woodcock tracking) for photons. The speed-up introduced by the two methods was investigated by simulating a PET acquisition with the Allegro/GEMINI GXL PET/CT scanner. The simulation was based on a clinical head-and-neck [¹⁸F]FDG PET/CT scan. The total time spent for the simulation (including initialization, particle tracking and signal processing) was obtained for seven settings corresponding to different tracking options. All seven methods led to very close results with regard to the total number of detected coincidences (less than 0.5% differences), and trues, scatter and random fractions. Acceleration factors of approximately 2.7 ($14 \times 14 \times 9$ voxels) to 27.6 ($378 \times 378 \times 243$ voxels) were obtained in comparison with the fastest available tracking available in GATE 3.1.2.

1. Introduction

Monte Carlo simulations in emission tomography are widely used to investigate new scanner prototypes, and assess image reconstruction and post-processing methods (Buvat and Lazaro 2006). More recently, simulation results have also been used directly in image reconstruction

(Kamphuis *et al* 1998, Beekman *et al* 2002, Rafecas *et al* 2004, Lazaro *et al* 2005). Geant4 Application for Emission Tomography (GATE) (Jan *et al* 2004) is a well validated and very versatile application for Monte Carlo simulations in emission tomography that can be used for highly realistic simulations. It has been used for all aforementioned purposes, but lacks efficiency especially in tracking particles in voxelized phantoms. Clinical whole body (multiple bed) PET simulations with voxelized patient phantoms require typically several tens of thousands of CPU hours of calculation time. At present the simulation of such clinical scans on small computer clusters are therefore prohibitive. Thus, it is desirable to implement improved tracking methods that yield accelerated simulations while preserving the great flexibility of the application. In contrast to recently proposed acceleration methods for GATE single photon emission tomography (SPECT) simulations (De Beenhouwer *et al* 2008), the presented methods are more generally designed to increase the efficiency of tracking in voxelized phantoms. They do not make use of variance reduction techniques and do not assign weights to particles, hence do not affect the statistical properties of the simulated data.

A large portion of the simulation time in GATE is spent to calculate and update paths of particles in the voxelized phantom. In Geant4 (Agnostinelli *et al* 2003, Allison *et al* 2006) objects are modeled using homogeneous *volumes* with constant material properties such as atomic composition and density. Cuboidal volumes are used to model voxels in voxelized phantoms. At each material boundary the linear attenuation coefficient needs to be updated and the remaining path length recalculated. A key operation is therefore to identify the next boundary and the next volume given the present particle position and momentum. In a general purpose Monte Carlo code like Geant4, this operation has to work correctly for volumes of arbitrary shapes and positions and at the same time it should be as fast as possible. For this purpose, in Geant4, additional information concerning the location of the volumes is pre-calculated and stored during the initialization and prior to tracking. This additional information is used to reduce the time needed to find the next boundary/volume during tracking. The time and memory needed for this pre-calculation increase with the number of volumes in the phantom. In GATE 3.1.2 simulations of voxelized phantoms, the amount of pre-calculated information is limited in order to keep reasonable initialization times and a moderate memory overhead. Consequently, the tracking speed is also only moderate.

In voxelized phantoms the positioning and also the shape of the volumes (the voxels) are very regular and therefore a pre-calculation is not necessary, because the next volume can be determined directly and efficiently during tracking. Recently (Arce 2007), a navigation/tracking algorithm for voxelized phantoms with direct voxel localization was introduced into Geant4 that meets the demand for faster tracking in voxelized structures. A key option is that the recalculation of the remaining path length can be omitted when the two neighboring voxels contain the same material. In previous versions of Geant4 voxel boundaries between voxels of the same material always required a recalculation of the path length.

A further acceleration can be achieved for particles with a relatively large mean free path length in comparison to the size of voxels in the phantom by using a sampling technique called *fictitious interaction tracking* (Brown and Martin 2003). This technique is also called *Woodcock tracking* (Woodcock *et al* 1965, Kawrakow and Fippel 2000) or *delta scattering* (Ljungberg *et al* 2005, Tenney *et al* 2004) and can be successfully applied to photon tracking in emission and especially in positron emission tomography (PET), because here the mean free path length is often larger than the voxel dimension.

2. Theory—fictitious interaction tracking

In this method photons of given energy E are tracked as if the total linear attenuation coefficient $\mu_i(E)$ of the voxels i is the same for all voxels. This assumed constant total linear attenuation coefficient is the sum of the real (physically correct) linear attenuation coefficient and a coefficient of an artificial process that does not change the photon state. This process is called fictitious interaction process. In less attenuating voxels fictitious interactions are more probable. Therefore, the process compensates for the wrongly assumed homogeneity of the phantom. This can be done without introducing any bias or changing the photon statistics (Brown and Martin 2003, Woodcock *et al* 1965) by choosing the replacement linear attenuation coefficient $\mu_{\text{repl}}(E)$ such that

$$\forall i \quad \mu_{\text{repl}}(E) \geq \mu_i(E) \quad (1)$$

is true for all voxels. Equation (1) guarantees that the travel distance to the next interaction (real or fictitious) will always be equal or shorter than in reality. In addition, for each interaction in voxel i with material $m = m(i)$ the ratio $R_m(E)$

$$R_m(E) \equiv \frac{\mu_m(E)}{\mu_{\text{repl}}(E)} \quad (2)$$

with $0 < R_m(E) \leq 1$ is calculated. This ratio represents the probability that the interaction is a real interaction. By 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.

The heterogeneity of the phantom is therefore 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.

3. Methods

3.1. Tracking implementation

3.1.1. Parametrized tracking and compressed voxels. These two methods for tracking and volume arrangement in voxelized phantoms have been available in GATE since versions 1.0.0 and 3.0.0, respectively. For these volume arrangements, Geant4 performs the pre-calculation that is mentioned in section 1. In this pre-calculation the space is sliced into one, two or three dimensions. This slicing is not physical but internal in Geant4 for use in locating points and in identifying the next intersection. For each slice (or rather ‘cell’ if the slicing is in three dimensions) it is recorded which volumes (voxels in our case) extend into it or are located within. For a given point in space, the cell in which the point is located can be determined very fast. Then, on average half of the volumes referred to the cell must be searched in order to identify the volume in which the point is located. This procedure works for arbitrarily shaped and positioned volumes. The memory consumption and initialization time increase with the number of cells. Because voxelized phantoms contain a very large number of voxels, the number of cells is reduced in GATE, resulting in a relatively large number of voxels per cell and leading eventually to only moderate tracking speed.

While the standard *parametrized tracking* (Agnostinelli *et al* 2003) uses the voxel grid of the phantom without modification, the method of *compressed voxels* (Taschereau

and Chatziioannou 2008) reduces the number of voxels by combining neighboring voxels which share the same material into larger voxels. This increases the size of the average propagation step and reduces the occupancy of cells. Due to the complexity of a real patient, this reduction is however only moderate and results in acceleration factors of around two. These two methods were not modified but used for comparison as implemented in GATE version 3.1.2.

Another method of creating parametrized volumes in Geant4, using *nested parametrizations* (Aso *et al* 2007, Sarrut and Guigues 2008) and the corresponding navigation, is not utilized in this study or available in GATE.

3.1.2. Regular navigation algorithm. Since Geant4.9.1 a new navigation algorithm, dubbed *regular navigation* (Arce 2007), can be used for the tracking of particles in voxelized volumes. This algorithm is incorporated in our GATE implementation and can be used within the voxelized phantom. The regular navigation algorithm performs fast direct neighboring voxel identification without a large memory overhead. This is the major source of acceleration of the implemented regular navigation algorithm. In addition, boundaries between voxels which share the same material can be ignored. Using this option, the geometry only limits tracking at the boundary between voxels with different materials, providing a significant reduction of the number of propagation steps. The regular navigation algorithm can be used for all types of particles, including photons and electrons.

3.1.3. Fictitious interaction tracking. Fictitious interaction tracking was implemented by using a mechanism of Geant4 called *FastSimulationModel*. This mechanism makes it possible to leave and re-enter Geant4 tracking in a transparent manner. All photons inside the phantom above a specified minimum energy are taken out of Geant4 tracking and are handled by our fictitious interaction tracking. The photons re-enter Geant4 tracking at the border of the phantom, or inside the phantom if the energy drops below the minimum energy (see figure 1). These photons are put on a stack that is then processed by Geant4. Photons of lower energy and electrons inside the phantom are tracked using the regular navigation. If during fictitious interaction tracking the energy of the photon after the interaction is lower than a cut value (E_{kill}) the photon is discarded (see also 3.1.4).

3.1.4. Electron range cut and cut-off energy for photons in the phantom. In addition to the aforementioned tracking and volume arrangement, two other options were implemented. In Geant4 electrons that cannot reach a specified distance (the *range cut*) due to the lack in kinetic energy are not created. It is now possible in GATE for regular tracking and fictitious interaction tracking that electrons in the phantom can have a different range cut than electrons in the detection system. These different electron range cuts were dealt with at the Geant4 level. We used this option to suppress electron simulation in the phantom while simulating electrons in the detection system. If dosimetric questions are not of interest, this does not influence the PET simulation results, but leads to shorter simulation times.

Second, the possibility of discarding photons below a certain energy limit was introduced into GATE (see figure 1). This option is restricted to the phantom and was implemented only for fictitious interaction tracking. This option can also be used to reduce simulation time, but should be applied with caution since it can alter the simulation results. For example, note that photons below the low-energy threshold can still be detected due to the energy resolution, and that even those photons that are not detected can contribute to effects like pile-up and deadtime.

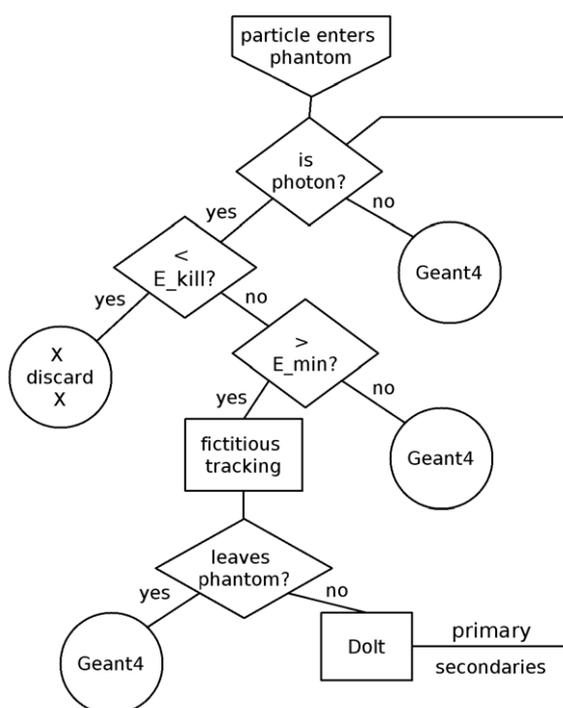


Figure 1. Flow chart showing the integration of fictitious interaction tracking into GATE. The Geant4 *Dolt* function is used to determine the momentum and energy of the primary particle and of the secondary particles after the interaction.

3.1.5. Detection system. Tracking in the rest of the scene (i.e., detection system, ‘world volume’ of Geant4) was not changed. All aforementioned new tracking methods and options can be activated by a few simple commands in the *macro* file that controls the GATE simulation.

3.2. Evaluation

The different tracking methods were compared in terms of simulation outcome and execution time with phantoms that approximate a clinical PET scan using relatively few (1764) up to a large number (3.47×10^7) of voxels.

3.2.1. Phantom. The simulation is based upon data that were acquired in one bed position of a clinical PET/CT [^{18}F]FDG head-and-neck scan. The CT image consisting of $512 \times 512 \times 264$ voxels of dimension $1.172 \times 1.172 \times 1.086 \text{ mm}^3$ was cropped to a smaller region of $378 \times 378 \times 264$ voxels in order to avoid the overlap of the corners of the field of view (=air) with the shielding of the scanner. The resulting image was smoothed with a 3D-Gaussian with $\sigma = (1.172, 1.172, 1.086) \text{ mm}$ and resampled to form images with $378 \times 378 \times 243$, $126 \times 126 \times 81$, $42 \times 42 \times 27$ and $14 \times 14 \times 9$ voxels (see table 1) using the Insight ToolKit ITK (<http://www.itk.org>). These different voxel sizes are considered to investigate the influence of the number of voxels on the simulation time. Samplings with voxel sizes at CT resolution, at PET resolution, and at two lower resolutions are used for comparison.

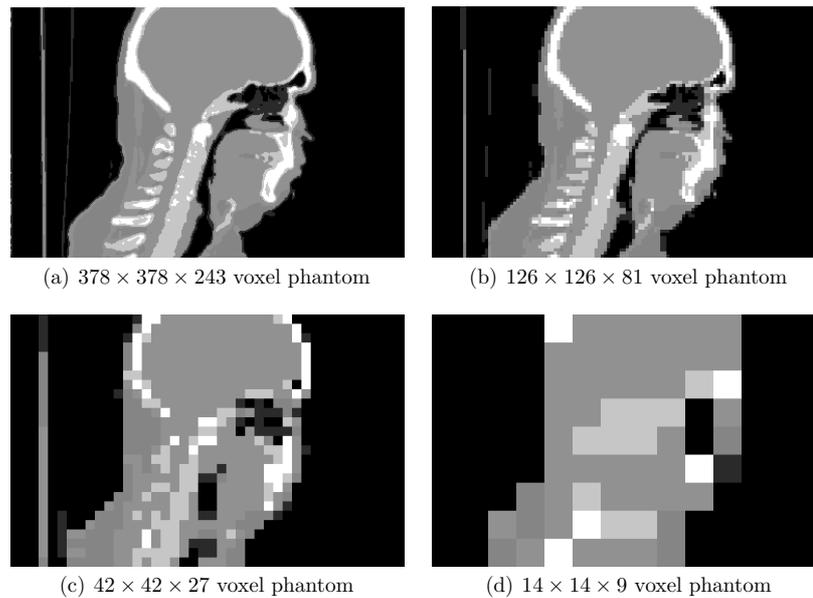


Figure 2. A central sagittal slice of the segmented CT images using various resampling.

Table 1. Voxelization of the simulated phantoms.

Number of voxels	Voxel dimensions
$378 \times 378 \times 243$	$1.172 \times 1.172 \times 1.086 \text{ mm}^3$
$126 \times 126 \times 81$	$3.516 \times 3.516 \times 3.259 \text{ mm}^3$
$42 \times 42 \times 27$	$10.55 \times 10.55 \times 9.78 \text{ mm}^3$
$14 \times 14 \times 9$	$31.64 \times 31.64 \times 29.33 \text{ mm}^3$

Table 2. Hounsfield unit (HU) ranges used to segment the CT images into six different materials.

HU range	Material
-1024 to -291	Air
-290 to -201	Lung
-200 to -31	Adipose
-30 to 109	Brain
110 to 499	Spine bone
500 to 3072	Skull

The resulting images were segmented into six materials (see table 2) providing realistic attenuation maps (see figure 2). The source distribution was taken from the clinical PET scan (see figure 3) and resampled to $126 \times 126 \times 81$ voxels. A total activity of 35 MBq of ^{18}F was used. Activity out of the field of view was not simulated.

In order to further investigate the spatial accuracy of the simulated coincidences, simulations were performed using the same voxelized attenuation phantom (a voxel size of $3.516 \times 3.516 \times 3.259 \text{ mm}^3$), but using only three voxels filled with 20 MBq, 10 MBq

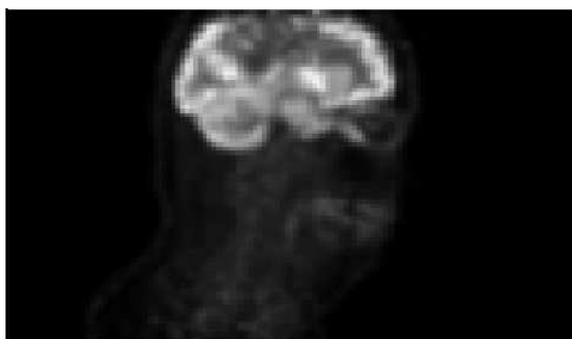


Figure 3. ^{18}F FDG source distribution. In the simulations, the probability of emission is proportional to the intensities of the pixels.

Table 3. Volume arrangement, tracking method, electron range cut and photon cut-off energy of the different settings that were studied.

Setting name	Volume arrangement	Tracking	e^- range cut (phantom, detectors)	γ cut-off energy in phantom
M1	Normal	Parametrized	(0.1 mm, 0.1 mm)	No ^a
M2	Compressed	Parametrized	(0.1 mm, 0.1 mm)	No ^a
M3	Normal	Regular	(0.1 mm, 0.1 mm)	No ^a
M4	Normal	Fictitious	(0.1 mm, 0.1 mm)	No ^a
M5	Normal	Regular	(500 mm, 0.1 mm)	No ^a
M6	Normal	Fictitious	(500 mm, 0.1 mm)	No ^a
M7	Normal	Fictitious	(500 mm, 0.1 mm)	300 keV

^a Geant4 based gamma range cut of 1 mm.

and 5 MBq of ^{18}F FDG (a voxel size of $3.516 \times 3.516 \times 3.259 \text{ mm}^3$). The voxels were placed in the center of the brain, and in the left and right eyes, respectively. This phantom is named three-voxels-source phantom and the corresponding simulation three-voxels-source simulation henceforth.

3.2.2. PET/CT scanner and physics. The simulated scanner was the PET part of the PET/CT scanner Allegro/GEMINI GXL with simulated 8 ns coincidence time window, energy resolution of 15% at 511 keV, 405 keV lower threshold and 665 keV upper threshold. The simulations accounted for photoelectric effect and Compton effect. Both effects were simulated using the ‘standard’ physics modeling of Geant4.9.1p01, since the more precise ‘low-energy’ physics modeling is significantly slower and the ‘standard’ modeling is usually sufficiently accurate for PET simulations. Pairs of 511 keV back-to-back photons were simulated to model the annihilation process, without modeling the actual positron emission and positron range.

3.2.3. Simulations. The simulations were performed using seven different settings for volume arrangement, tracking, electron range cut and photon cut-off energy as listed in table 3. The first two settings are already available in GATE 3.1.2, the other five settings are based on the new tracking algorithms and options.

Setting M1 uses the simplest organization of volumes (*normal*) and the general parametrized tracking (*parametrized*), which does not exploit the regularity of the setup. Setting M2 merges neighboring voxels with the same material (*compressed*), improving the organization of volumes and exploiting better the capabilities of the parametrized tracking.

All other options are based on the new regular navigation algorithm (*regular*) in Geant4.9.1.p01. Settings M4, M6 and M7 in addition use fictitious interaction tracking (*fictitious*) for photons above the energy $E_{\min} = 42.0$ keV, 26.5 keV, 17.5 keV and 12 keV for the $14 \times 14 \times 9$, $42 \times 42 \times 27$, $126 \times 126 \times 81$ and $378 \times 378 \times 243$ phantoms, respectively. In these simulations, the regular navigation is only used for photons of lower energy.

Electrons in both the phantom and the detection system were discarded when the range dropped below 0.1 mm. For regular and fictitious tracking a special electron range cut only for the electrons in the phantom was introduced as an option (used in settings M5, M6 and M7). This electron range cut was set to 500 mm therefore suppressing the production of electrons in the phantom in these simulations. The electron range cut in the detection system was always 0.1 mm.

The photon cut-off energy in setting M7 was set to 300 keV. Photons in the phantom below this threshold were not simulated. This cut-off energy well below the 405 keV lower energy threshold of the simulated PET system assured that practically all photons with energies below 405 keV that could pass the threshold due to the finite energy resolution of the detection system were still simulated.

3.2.4. Accuracy. Simulations were performed to assess the accuracy of the accelerated tracking methods. The new methods used in settings M3–M7 were compared to the standard setting M1. A 10 s PET scan was simulated using the $126 \times 126 \times 81$ voxel phantom (see figures 2 and 3). These simulations were used to determine the total number of coincidences and the fraction of true, scattered and random coincidences. In a second set of simulations the energy of all detected singles was recorded. For this second set of simulations, the lower and upper energy thresholds were removed: the simulated electronics therefore accepted all incoming photons.

In a third set of simulations, the phantom with three voxels filled with activity (see 3.2.1) was used and 10 s acquisitions were simulated using the aforementioned upper and lower thresholds.

3.2.5. Speed. In another set of simulations, a 1 s PET scan was simulated for all seven settings and for all four different phantom samplings (see figure 2). These simulations were used to investigate how the phantom sampling affected the computation time for the different tracking methods.

All simulations were performed using a development version of GATE 4.0.0 together with our code modifications.

3.2.6. Computer hardware. The simulations were performed on a 64 bit Fedora Linux cluster with identical dual processor boards equipped with dual core AMD Opteron CPUs running at 1.6 GHz per core. One front end computer running 32 bit Fedora Linux was used to distribute the simulations and handle the data storage. Each simulation was assigned to a single core and not split. The computer cluster software Condor (Thain *et al* 2005) was used to determine the CPU times.

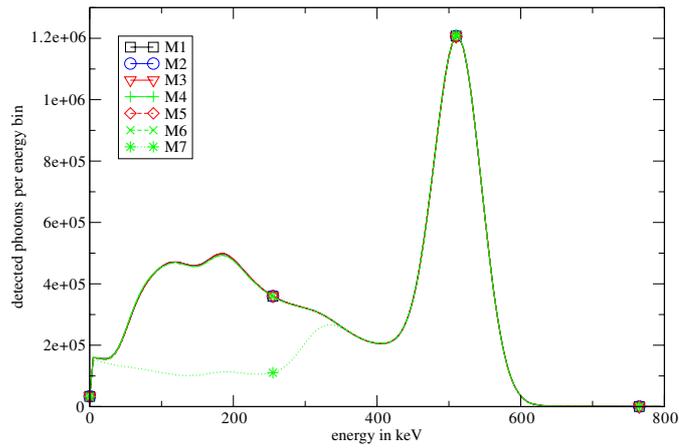


Figure 4. Detected singles in a 10 s simulation without lower or upper threshold using the $126 \times 126 \times 81$ voxel phantom. The detected singles are binned into 5 keV energy intervals. Due to the large number of data points, continuous lines through the data points are drawn and only at four energy positions the data points are shown explicitly by symbols.

Table 4. Simulation results for a 10s scan, $126 \times 126 \times 81$ voxel phantom.

Setting name	Coincidences	Trues	Scatter	Randoms	CPU time
M1	1491 518 ($\equiv 0\%$)	65.01%	25.64%	9.35%	329 255 s
M2	1492 582 (+0.07%)	64.88%	25.70%	9.43%	177 899 s
M3	1493 260 (+0.12%)	64.92%	25.72%	9.35%	46 516 s
M4	1498 574 (+0.47%)	64.78%	25.86%	9.36%	39 478 s
M5	1492 166 (+0.04%)	64.95%	25.68%	9.37%	44 206 s
M6	1496 635 (+0.34%)	64.75%	25.85%	9.39%	18 085 s
M7	1498 545 (+0.47%)	64.70%	25.92%	9.38%	15 394 s

4. Results

4.1. Accuracy

Table 4 shows the total number of detected coincidences and the relative number of true, scattered and random coincidences for the 10 s simulations using the $126 \times 126 \times 81$ voxel phantom (figures 2 and 3). All values agree within 0.5%. The last column shows the total simulation time including initialization, tracking, event and signal processing. The new methods therefore show good agreement with the established ones while accelerating considerably the simulation (settings M3, M4, M5, M6 and M7 are $3.8\times$, $4.5\times$, $4.0\times$, $9.8\times$ and $11.6\times$ faster than setting M2 according to table 4).

Figure 4 shows the spectra of the detected singles for the 7 simulation settings. Clearly, the simulations using settings M1–M6 agree very well. The simulation using setting M7 deviates from the other simulations for energies below 350 keV. This was expected, because in this setting no photons in the phantom below 300 keV were tracked and energy blurring in the detectors was simulated. For settings M2–M7, the relative deviation of the detected singles between 300 keV and 665 keV in comparison with setting M1 can be seen in figure 5. The results agree very well for energies smaller than 600 keV. For energies larger than 600 keV the relative deviation becomes larger for all settings. This is because of the small absolute number

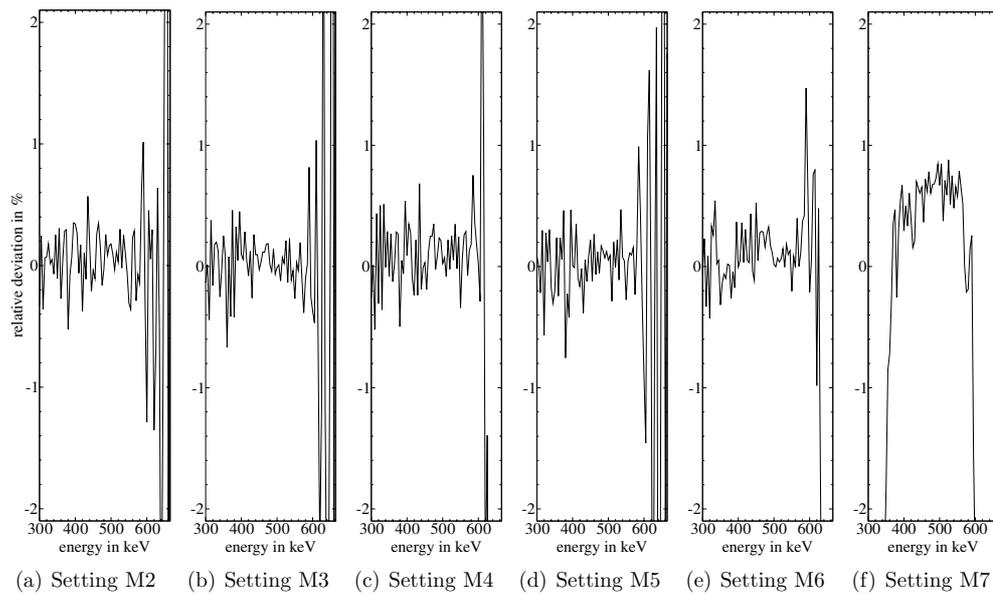


Figure 5. Deviation of the energy spectra (figure 4) of the different tracking settings M2–M7 relative to setting M1.

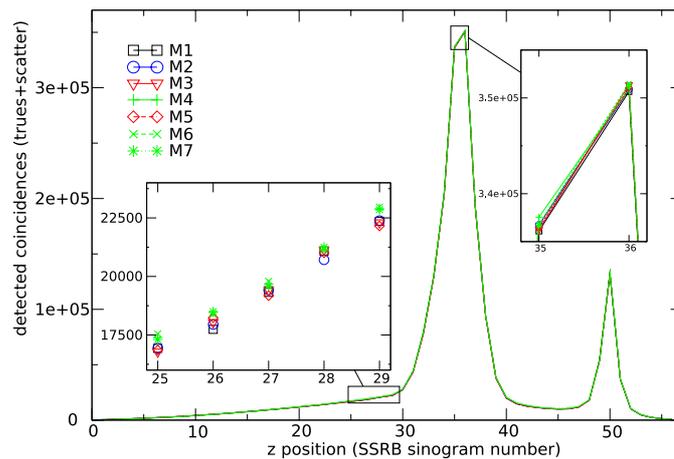


Figure 6. Total detected coincidences (trues+scatter) per SSRB sinogram versus SSRB sinogram number for the three-source-voxels simulation.

of detected singles (see figure 4). In addition, for setting M7 the effect of the photon cut-off energy (deviation for smaller energies) and a general small overestimation of about 0.5% can be observed.

In order to verify the spatial accuracy of the simulated coincidences (trues+scatter) in the three-voxels-source simulation, the detected coincidences were rebinned into 57 sinograms using the single slice rebinning SSRB method (Daube-Witherspoon and Muehlechner 1987). Figure 6 shows the total number of counts per sinogram against axial position (‘axial profile’).

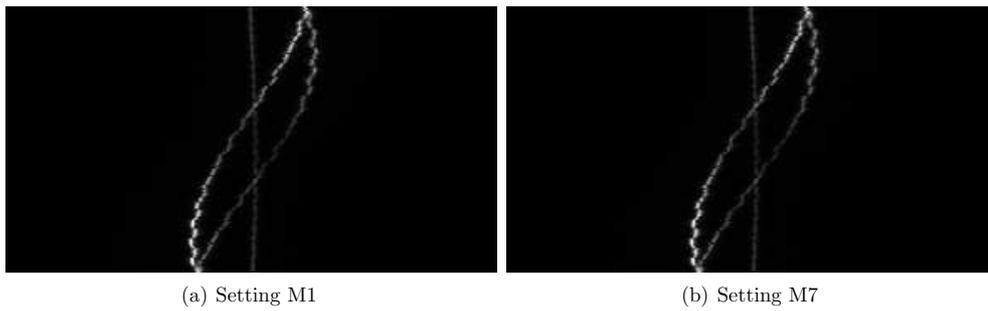


Figure 7. Summed sinograms obtained by the three-source-voxels simulation and settings M1 and M7, respectively.

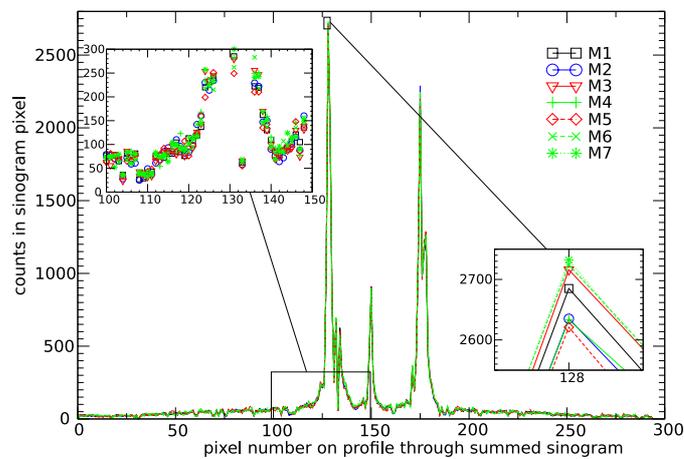


Figure 8. Diagonal profile (lower left to upper right corner) through summed sinograms (figure 7) for settings M1–M7 and the three-source-voxels simulation. Due to the large number of data points, continuous lines through the data points are drawn.

The first peak is due to the two voxels filled with activity that were placed in the two eyes and that were located at the same transaxial plane. The voxel in the brain corresponds to the position of the second peak.

The transaxial spatial accuracy was investigated by summing the obtained SSRB sinograms and taking two diagonal profiles (1. lower left to upper right corner and 2. upper left to lower right corner). The sinograms were summed in order to increase the statistics that were otherwise not sufficient. The summed sinograms for setting M1 and M7 can be seen in figure 7. Figure 8 shows profile 1 for the seven different settings.

The agreement between the results of simulations using setting M1 and the other settings was measured by two tailed two-sample Kolmogorov–Smirnov goodness-of-fit tests. This test was applied to the energy spectra in the range 0–665 keV (figure 4; table 5, left column), the energy spectra in the photo peak range 400–600 keV (figure 4; table 5, second column), the axial profile (figure 6; table 5, third column) and the two diagonal profiles through the summed sinograms (figures 7 and 8; table 5, last column). The p -value and the test statistic D_{\max} are displayed. P -values close to 1 show that the deviations in the results are very likely

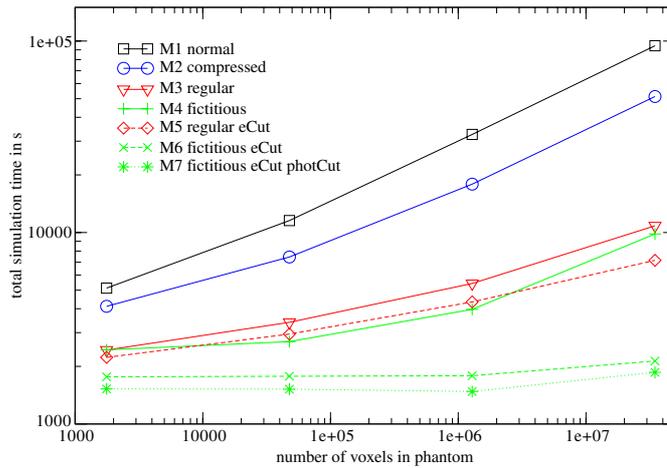


Figure 9. Total CPU time needed for a 1 s simulation of 35 MBq for the seven different tracking approaches including initialization, tracking in phantom and detectors and signal processing. Voxel grid (i.e., different points on the x -axis) were $14 \times 14 \times 9$, $42 \times 42 \times 27$, $126 \times 126 \times 81$, $378 \times 378 \times 243$.

Table 5. P -values obtained by the two-tailed two-sample Kolmogorov–Smirnov goodness-of-fit test applied to the curves M2–M7 compared to M1. The first two columns show the values for the energy spectra (figure 4) in the range (0 keV, 665 keV) and (400 keV, 600 keV), respectively. The third column compares the axial profiles (figure 6). The fourth and fifth columns show the values for the diagonal profiles through the summed sinograms (figure 7). The distance D_{\max} (test statistic) is shown in brackets.

	0–665 keV	400–600 keV	Axial profile	Sinogram profile 1	Sinogram profile 2
M2	1.0 (0.0224)	1.0 (0.0488)	1.0 (0.0175)	0.9644 (0.0407)	0.6990 (0.0576)
M3	1.0 (0.0149)	1.0 (0.0488)	1.0 (0.0351)	0.9952 (0.0339)	0.6990 (0.0576)
M4	1.0 (0.0373)	1.0 (0.0488)	1.0 (0.0351)	0.6990 (0.0576)	0.9990 (0.0305)
M5	1.0 (0.0224)	1.0 (0.0224)	1.0 (0.0351)	0.9952 (0.0339)	0.9848 (0.0373)
M6	1.0 (0.0373)	1.0 (0.0488)	1.0 (0.0526)	0.9952 (0.0339)	0.9999 (0.0271)
M7	4.4×10^{-12} (0.4403)	1.0 (0.0488)	1.0 (0.0702)	0.9999 (0.0271)	0.9644 (0.0407)

due to statistical uncertainties and not due to systematic errors. The small p -value for the test between curves M1 and M7 in figure 4 in the range 0–665 keV (table 5, first column, last row) is the only exception: the deviation between the two curves is systematic since lower energy photons are not simulated in the phantom.

4.2. Speed

Figure 9 shows the dependence of the simulation time on the number of voxels in the phantoms. It can be seen that the execution times of all methods depend on the numbers of voxels. When using fictitious tracking with the suppression of electrons in the phantom (M6 and M7), the execution time stays constant up to phantoms with about 10^6 voxels. The execution time saving for the whole simulation (see figure 10) ranges from almost 2 times (1764 voxels) to around 5–7 times (3.47×10^7 voxels) for the methods M3, M4 M5 compared to the fastest method currently available in GATE (setting M2). Method M6 improves these factors to values of approximately $2.3\times$ to $24\times$, and method M7 to values of $2.7\times$ to $27.6\times$.

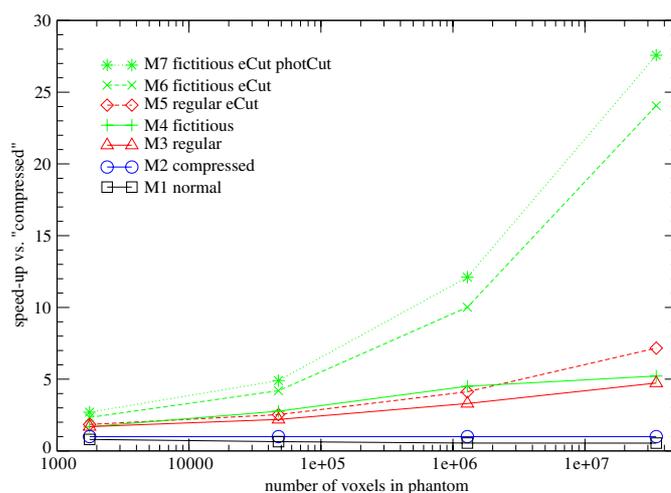


Figure 10. Speed-up relative to the fastest presently available tracking option in GATE 3.1.2 ('compressed').

5. Discussion

For PET coincidence measurements, the accuracy of the new tracking methods/options agree very well with the approved tracking methods of GATE in terms of the number of detected prompts, trues, scatter and randoms. When singles are considered, the 300 keV photon cut-off energy of setting M7 in the phantom led to only minor deviations between the resulting single energy spectrum and the reference energy spectrum (setting M1) above 350 keV. Only a small overestimation of around 0.5% (see figure 5(f)) could be observed which is probably due to reduced dead time of the simulated detectors. Indeed, a fewer low-energy photons are simulated in setting M7. It is possible that some of them would have contributed to dead time and thus the dead time might be reduced if they are not simulated. This effect might increase with higher activities, since dead time becomes more important when activity increases. It is advisable to carefully test the influence of a photon cut-off energy and adjust the cut if necessary.

In dosimetry applications the accuracy of the results is limited by the electron range cut in the phantom. The electron cut must be adjusted to the required accuracy/resolution. Settings M5–M7 are therefore probably not compatible with most dosimetry applications.

The dependence of the simulation time on the number of voxels (see figure 9) varies greatly, from a strong dependence on the size of the voxel grid (setting M1 or M2) to only a weak dependence on the size of the voxel grid (setting M6 or M7). This can be explained as follows.

In simulations with settings M1 and M2, a step is required for each voxel in the setup which is crossed by a particle. In addition, several volumes need to be checked in order to determine the right volume for a given position of the particle (see 3.1.1). Since the average number of voxels per cell is large (in order to reduce typical memory consumption) and the voxels in a cell are searched linearly, this is time consuming.

Settings M3–M6 use the regular navigation algorithm, either exclusively (settings M3 and M5), or as a fallback solution for electrons (setting M4) and low-energy photons (settings M4 and M6). The regular navigation algorithm is faster: the number of steps limited by

the geometry is equal to the number of boundaries encountered between voxels of different materials. As the number of voxels in the phantom increases, the number of boundaries between materials grows, but slower than the total number of voxels. This leads to a (weaker) dependence of the simulation time on the number of voxels in the phantom.

The voxel grid size dependence is stronger for setting M4 since here electrons are also simulated using the regular navigation algorithm. Electron tracking in phantoms with few voxels involves very few voxel boundary crossing due to the short electron range. In phantoms with many voxels such boundary crossings become much more likely. Since photon tracking in setting M4 is faster than in settings M1–M3, this results in a disproportional increase in electron tracking simulation time in comparison to settings M1–M3.

In addition, in simulations with fictitious interaction tracking with regular navigation as fallback solution (settings M4 and M6), particles that are passed to Geant4 are put on a stack. This stack contains no information about the volume/voxel in which the particle resides. Some additional time is required to identify the volume and also to undertake additional bookkeeping, such as creating a new track for each stacked particle.

The initialization time for the voxelized phantoms with up to $126 \times 126 \times 81$ voxels was negligible with around 15 s. The phantom with $378 \times 378 \times 243$ voxels required initialization times of around 2 min. This leads to some additional dependence of the simulation time on the number of voxels, especially for the fast simulations (settings M6 and M7), where the relative increase is larger due the short simulation time.

Finally, although theoretically there should not be any voxel grid size dependence in the fictitious interaction tracking itself, a slight dependence could be introduced by effects due to the limited size of the CPU cache.

The proposed accelerated tracking methods can also be used in single photon emission tomography (SPECT) simulations. While the speed-up in the phantom due to the usage of the regular navigation algorithm should be similar to the presented results, the speed-up due to fictitious interaction tracking is expected to be reduced due to the lower energy of the involved photons. Indeed, the lower energy leads to reduced mean free path hence to an increased number of fictitious interactions per simulated particle.

Preliminary tests using the presented voxelized patient phantoms of different voxel sizes together with a modified version of the SPECT detector used in the SPECT benchmark of GATE (Jan *et al* 2004) (multiplying all dimensions of the detection system by 1.4 to allow the phantom to fit inside it), show comparable and even slightly greater speed-up (almost $3\times$ to $34\times$ for M6/M2, $2\times$ to $9\times$ for M4/M2) than the PET simulations ($2.3\times$ to $24\times$ M6/M2, $1.7\times$ to $5.2\times$ M4/M2).

These surprisingly high accelerations in SPECT can be explained as follows: since the test whether an interaction is real or fictitious is located inside a small efficient loop, the increase in execution time due to the increased number of fictitious interactions is relatively small. In the SPECT benchmark the detection system is simplified and in addition there are no coincidences to be created or sorted like in PET. This leads to shorter simulation times in the detection system and in the signal processing part. This execution time reduction therefore outweighs the increase due to the increased number of fictitious interactions and the net reduction in total execution time due to fictitious interaction tracking is actually greater than in the PET case. This interpretation is supported by the shorter total simulation time of 1322 s for the SPECT M6 simulation versus 1764 s for the PET M6 simulation ($14 \times 14 \times 9$ voxels).

In more realistic SPECT simulations much longer simulation times due to particle tracking in a more detailed and complex detection system are expected and the expected acceleration should be consequently smaller. Irrespective of the total execution time reduction, the

preliminary SPECT results indicate however that 140 keV photon fictitious interaction tracking in the phantom is faster than regular tracking or older tracking implementations of GATE.

6. Conclusions

We presented methods to accelerate photon tracking in GATE without compromising the accuracy, statistics or versatility of GATE. The acceleration is based upon two new tracking algorithms: the regular navigation algorithm of the Geant4 release 9.1 and our implementation of fictitious interaction tracking. In addition, electron range cuts and a photon cut-off energy for the phantom were introduced which yield additional acceleration. The gain in speed for a full PET simulation including tracking, detection and signal processing with voxels at PET resolution was about one order of magnitude in comparison to the fastest tracking available in GATE 3.1.2.

Acknowledgments

We would like to thank Sébastien Jan, Marc Verderi, and Nicolas Grotus for information on GATE, the FastSimulationModel of Geant4, and the Allegro/GEMINI GXL PET/CT scanner, respectively. In addition, we would like to thank Pedro Arce for early access to the regular navigation implementation of Geant4. The work was funded by ANR grant no ANR-06-CIS-004.

References

- Allison J *et al* 2006 Geant4 developments and applications *Trans. Nucl. Sci.* **53** 270–8
- Agnostinelli S *et al* 2003 Geant4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res. A* **506** 250–303
- Arce P 2007 Faster navigation for regular voxel geometries *12th Geant4 Collaboration Workshop (Hebden Bridge, UK, September 2007)* pp 13–9
- Aso T, Kimura A, Yamashita T and Sasaki T 2007 Optimization of patient geometry based on CT data in GEANT4 for medical application *2007 IEEE Nuclear Science Symp. Conf. Record* pp 2576–80
- Beekman F J, de Jong H W A M and van Geloven S 2002 Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation *IEEE Trans. Med. Imaging* **21** 867–77
- Brown F B and Martin W R 2003 Direct sampling of Monte Carlo flight paths with continuously varying cross-sections *Technical Report* (Los Alamos, NM: Los Alamos National Laboratory)
- Buvat I and Lazaro D 2006 Monte Carlo simulations in emission tomography and GATE: an overview *Nucl. Instrum. Methods Phys. Res.* **569** 323–9
- Daube-Witherspoon M E and Muehllehner G 1987 Treatment of axial data in three-dimensional PET *J. Nucl. Med.* **28** 1717–24
- De Beenhouwer J, Staelens S, Vandenberghe S and Lemahieu I 2008 Acceleration of GATE SPECT simulations *Med. Phys.* **35** 1476–85
- Jan S *et al* 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543–61
- Kamphuis C, Beekman F J, van Rijk P P and Viergever M A 1998 Dual matrix ordered subsets reconstruction for accelerated 3D scatter compensation in single-photon emission tomography *Eur. J. Nucl. Med.* **25** 8–18
- Kawrakow I and Fippel M 2000 Investigation of variance reduction techniques for Monte Carlo photon dose calculation using XVMC *Phys. Med. Biol.* **45** 2163–83
- Lazaro D, Bitar Z E, Breton V, Hill D and Buvat I 2005 Fully 3D Monte Carlo reconstruction in SPECT: a feasibility study *Phys. Med. Biol.* **50** 3739–54
- Ljungberg M, Larsson A and Johansson L 2005 A new collimator simulation in SIMIND based on the delta-scattering technique *IEEE Trans. Nucl. Sci.* **52** 1370–5
- Rafecas M, Mosler B, Dietz M, Pogl M, Stamatakis A, McElroy D P and Ziegler S I 2004 Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data *Trans. Nucl. Sci.* **51** 2597–605
- Sarrut D and Guigues L 2008 Region-oriented CT image representation for reducing computing time of Monte Carlo simulations *Med. Phys.* **35** 1452–63

- Taschereau R and Chatziioannou A F 2008 Compressed voxels for high-resolution phantom simulations in GATE *Mol. Imaging Biol.* **10** 40–7
- Tenney C R, Bowsher J E and Jaszczak R J 2004 A delta-scatter Monte Carlo model of a rotating parallel-hole collimator for ^{131}I brain tumor imaging *IEEE Nucl. Sci. Symp. Conf. Rec.* **5** 3084–8
- Thain D, Tannenbaum T and Livny M 2005 Distributed computing in practice: the Condor experience *Concurrency and Computation: Practice and Experience* **17** 323–56
- Woodcock E, Murphy T, Hemmings P and Longworth T 1965 Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry *ANL 7050* (Argonne, IL: Argonne National Laboratory)