

Fast analytical scatter estimation using graphics processing units

Harry Ingleby^{a,b,c}, Jonas Lippuner^a, Daniel W. Rickey^{a,b,c}, Yue Li^a and Idris Elbakri^{a,b,c,*}

^a*Division of Medical Physics, CancerCare Manitoba, McDermot Avenue, Winnipeg, Manitoba, Canada*

^b*Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba, Canada*

^c*Department of Radiology, University of Manitoba, Winnipeg, Manitoba, Canada*

Received 13 December 2013

Revised 13 January 2015

Accepted 16 January 2015

Abstract.

PURPOSE: To develop a fast patient-specific analytical estimator of first-order Compton and Rayleigh scatter in cone-beam computed tomography, implemented using graphics processing units.

METHODS: The authors developed an analytical estimator for first-order Compton and Rayleigh scatter in a cone-beam computed tomography geometry. The estimator was coded using NVIDIA's CUDA environment for execution on an NVIDIA graphics processing unit. Performance of the analytical estimator was validated by comparison with high-count Monte Carlo simulations for two different numerical phantoms. Monoenergetic analytical simulations were compared with monoenergetic and polyenergetic Monte Carlo simulations. Analytical and Monte Carlo scatter estimates were compared both qualitatively, from visual inspection of images and profiles, and quantitatively, using a scaled root-mean-square difference metric. Reconstruction of simulated cone-beam projection data of an anthropomorphic breast phantom illustrated the potential of this method as a component of a scatter correction algorithm.

RESULTS: The monoenergetic analytical and Monte Carlo scatter estimates showed very good agreement. The monoenergetic analytical estimates showed good agreement for Compton single scatter and reasonable agreement for Rayleigh single scatter when compared with polyenergetic Monte Carlo estimates. For a voxelized phantom with dimensions $128 \times 128 \times 128$ voxels and a detector with 256×256 pixels, the analytical estimator required 669 seconds for a single projection, using a single NVIDIA 9800 GX2 video card. Accounting for first order scatter in cone-beam image reconstruction improves the contrast to noise ratio of the reconstructed images.

CONCLUSION: The analytical scatter estimator, implemented using graphics processing units, provides rapid and accurate estimates of single scatter and with further acceleration and a method to account for multiple scatter may be useful for practical scatter correction schemes.

Keywords: Scatter estimation, compton scatter, rayleigh scatter, graphics processing units

1. Introduction

Cone-beam computed tomography (CBCT) is a relatively new imaging modality with expanding use in many areas, including image-guided radiation therapy and breast imaging [1]. In CBCT, the large

*Corresponding author: Idris Elbakri, Division of Medical Physics, CancerCare Manitoba, 675 McDermot Avenue, Winnipeg, Manitoba, Canada R3E 0V9. Tel.: +1 204 787 2856; Fax: +1 204 775 1684; E-mail: Idris.Elbakri@cancercare.mb.ca.

irradiated volume results in substantial scatter at the detector, which can seriously degrade image quality [2]. Both hardware- and software-based methods for mitigating the effects of scatter have been investigated [3,4]. Software-based or computational scatter estimation and correction have shown promise as a method of reducing image quality degradation due to such scatter [5–7]. One approach is for patient-specific scatter estimates to be subtracted from measured projection data to obtain primary-only data [6]. Alternatively, a scatter estimator could be incorporated within the image formation model of a reconstruction algorithm [7], providing a better match between estimated and measured data and resulting in more accurate reconstructed images. In order to be clinically useful, both approaches require a fast and accurate scatter estimation algorithm that can account for heterogeneous objects and varying imaging geometries and beam characteristics [5,7].

Computational scatter estimators are generally based on either Monte Carlo (MC) simulation or analytical calculation [5,6]. Monte Carlo methods are more accurate but generally time-consuming [7]. A comprehensive MC simulation package, such as EGSnrc, incorporates complex physics models that result in excellent agreement with experimental data [8]. The penalty for obtaining this accuracy is lengthy computation time, which may range from minutes to days depending on the desired level of statistical uncertainty and available computer hardware.

A number of researchers have implemented scatter correction using MC scatter estimates and obtained improvements in image quality with scatter-corrected reconstructions. For example, Jarry et al. subtracted MC scatter estimates from measured kilovoltage cone beam computed tomography projection data and obtained improvements in contrast and artifact reduction in reconstructed images [6]. Zbijewski and Beekman implemented an accelerated MC scatter estimator within an iterative scatter artifact reduction scheme for cone beam micro computed tomography, resulting in a reduction in cupping artifacts from scatter [7]. Mainegra-Hing and Kawrakow developed a fast MC scatter estimator for CBCT and incorporated it in an iterative scatter correction algorithm [9]. The major challenge with these techniques is computation time. The latter two methods implemented variance reduction techniques and also employed denoising or smoothing of noisy (i.e. lower photon count) MC data to accelerate the process [7,9].

Analytical scatter estimators for voxelized data have also been developed. Wells et al. developed a method to analytically estimate first- and second-order Compton scatter for single photon emission computed tomography and later applied it experimentally, showing improved image quality for scatter-corrected reconstructions [5]. Yao and Leszczynski derived an approximate analytical expression for first-order Compton scatter as a function of the primary fluence at the detector [10] and applied it iteratively to estimate primary and scatter fluence at the detector for 6 MV CBCT [11]. These analytical techniques were fast, produced noise-free scatter estimates, and were also more amenable to parallelization.

In addition to scatter correction in conventional clinical imaging modalities and geometries, there has been growing interest in scatter tomography and spectroscopy, where the scatter signal is used to reconstruct an image or provide other relevant information [12–14]. For example, Aviles et al. presented an algorithm for Compton-scatter CT using a ring geometry detector, and used Monte Carlo [12] simulations to generate scatter data from a primary beam. The same group [13] also proposed Compton-scatter CT system using a line of energy differentiating detectors. In the latter work an analytical scatter simulation method was used. Since these works focused on demonstrating feasibility, the computational requirements were not stated.

A recent approach to accelerating computationally intensive simulations through parallel processing is to use graphics processing units (GPUs), which were originally developed for computer gaming applications [15]. In the past few years, GPUs have evolved from specialized graphics processors into sophisticated programmable parallel processors that can be used for general-purpose computation. Modern

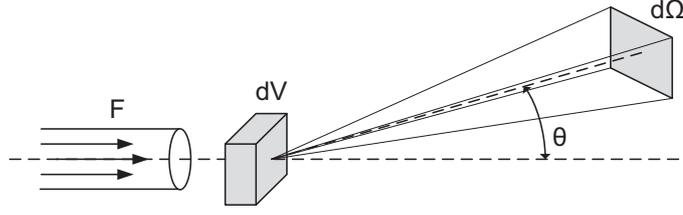


Fig. 1. Geometry for photons scattered by single Compton and Rayleigh interactions from a homogeneous target with volume dV to a detector element subtending solid angle $d\Omega$ at scattering angle θ .

GPUs contain hundreds of cores, which are capable of executing thousands of instances of a single function simultaneously with different input data. Using the Compute Unified Device Architecture (CUDA) (NVIDIA Corp., Santa Clara, CA, USA), C programs can be executed on NVIDIA GPUs to take full advantage of the parallel architecture. As analytical scatter estimation techniques typically require many repetitions of the same calculation with different input parameters, they are very amenable to parallel processing and thus well-suited for implementation using CUDA and execution on a GPU.

In this paper, we present an analytical method to estimate first-order Compton and Rayleigh scatter images for a voxelized heterogeneous object in a CBCT geometry, implemented in the CUDA environment for execution on NVIDIA GPUs. We also present preliminary reconstruction results with numerical phantoms to demonstrate the potential use of this scatter simulation tool scatter artifact reduction. Our long-term goal is to apply this technique for patient-specific scatter compensation in breast computed tomography. While in this work we focus on the CBCT geometry for the purpose of scatter correction, the computation acceleration presented can be useful to other geometries such as those considered for Compton scatter CT [12–14].

2. Methods and materials

2.1. Analytical model for first-order Compton and Rayleigh scatter

Consider a uniform monoenergetic photon beam of energy $h\nu$ incident on a homogeneous target of volume dV as shown in Fig. 1. The number of scattered photons arriving at a detector element at scattering angle θ , subtending a solid angle $d\Omega$, is given by:

$$N(\theta, h\nu) = F \frac{d\sigma}{d\Omega}(\theta, h\nu) \rho dV d\Omega \quad (1)$$

where F is the fluence incident on the target volume, $\frac{d\sigma}{d\Omega}(\theta, h\nu)$ is the differential scattering cross-section for a given scattering process and ρ is the number density of scattering centers in the target volume [16]. The target volume dV is assumed to be small, so that photon attenuation pre- and post-scatter within the volume is negligible.

First-order Compton and Rayleigh scatter processes were modeled as follows. The Klein-Nishina differential cross-section for Compton (incoherent) scattering at scattering angle θ and incident photon energy $h\nu$ is:

$$\frac{d\sigma}{d\Omega}(\theta, h\nu) = \frac{r_0^2}{2} \left(\frac{h\nu'}{h\nu} \right)^2 \left(\frac{h\nu}{h\nu'} + \frac{h\nu'}{h\nu} - \sin^2 \theta \right) \quad (2)$$

where r_0^2 is the classical electron radius and $h\nu'$ is the energy of the scattered photon [17]. The energy of the scattered photon as a function of $h\nu$ and θ is:

$$h\nu' = \frac{h\nu}{1 + \left(\frac{h\nu}{m_0c^2}\right) (1 - \cos \theta)} \quad (3)$$

where m_0c^2 is the rest energy of the electron (511 keV) [17]. The Klein-Nishina approximation assumes unbound (free) electrons, initially at rest, and thus neglects the effects of electron binding energies.

The differential cross-section for Rayleigh (coherent) scattering at scattering angle θ , using the form factor approximation, is:

$$\frac{d\sigma}{d\Omega}(\theta, x) = \frac{r_0^2}{2} (1 + \cos^2 \theta) F_M^2(x) \quad (4)$$

where $F_M^2(x)$ is the molecular coherent scattering form factor as a function of momentum transfer, x [18]. The value of the momentum transfer x , in units of \AA^{-1} is given by

$$x = \frac{h\nu}{12.39852} \sin\left(\frac{\theta}{2}\right) \quad (5)$$

when the incident photon energy, $h\nu$, has units of keV [18]. The molecular form factor is calculated using the free-gas model as:

$$\frac{F_M^2(x)}{W} = \sum \frac{w_i}{M_i} F^2(x, Z_i) \quad (6)$$

where W is the molecular weight of the material, w_i and M_i are the mass fraction and atomic mass of element i , respectively, and $F^2(x, Z_i)$ is the atomic coherent scattering form factor for the element with atomic number Z_i at momentum transfer x [18].

In Compton scattering, the density of scattering centers is given by the electron density of the target. For Rayleigh scattering, the density of scattering centers is given by the density of atoms (for elements) or molecules (for compounds) in the target.

In this work, we restricted our approach to estimation of first-order Compton and Rayleigh scatter. Single scatter could represent a large fraction of total scatter although this depends on photon energy and the size of the irradiated object. For example, Kyriakou et al. found that single scatter represented approximately 70% of total scatter for a 12 cm diameter water phantom at 40 keV [19]. Patient or phantom heterogeneity influences the spatial distribution of single scatter much more strongly than that of multiple scatter, which is more diffuse. Finally, the computational burden increases exponentially for a voxelized geometry as higher-order scatter terms are considered. We plan to address multiple scatter in future work.

2.2. Extension to voxelized geometry

We extended the geometry shown in Fig. 1 to include a point source irradiating a three-dimensional voxelized phantom in cone-beam geometry, with scattered photons detected by a two-dimensional pixelized detector, as shown in Fig. 2. The point source irradiated the front face of the phantom, and the

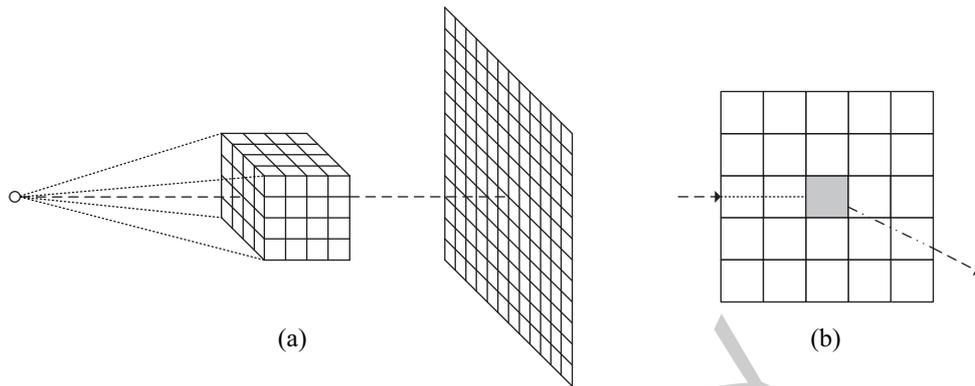


Fig. 2. (a) Cone-beam imaging geometry, with a voxelized phantom irradiated by a point source. Singly scattered photons are detected by a pixelated detector. (b) Detail of ray path through voxels on central axis, showing ray for pre-scatter attenuation (····) and post-scatter attenuation (---) in voxels before and after scattering voxel (shaded).

number of photons scattered from individual phantom voxels to individual detector pixels by single Compton and Rayleigh interactions was calculated as described previously. Both pre- and post-scatter attenuation were accounted for by ray tracing. Each phantom voxel was assigned a known material and was characterized by the following properties: elemental composition (required to calculate molecular coherent scattering form factors), electron density, and linear attenuation coefficient as a function of energy. We assumed an ideal photon-counting detector with 100% detection efficiency.

Each voxel in the phantom was treated as an individual target volume, per Eq. (1). Fluence incident on each voxel in the phantom was corrected for beam divergence and pre-scatter attenuation. Pre-scatter attenuation at each scattering voxel was estimated by ray tracing from the point source to the scattering voxel via intervening voxels with known linear attenuation coefficients using a fast incremental ray tracing algorithm. The number of photons scattered to a given detector pixel from a given scattering voxel in the phantom was calculated using discretized versions of Eq. (1) through Eq. (6). Scattered photon counts were corrected for post-scatter attenuation along the path from the scattering voxel to the detector pixel, again using the incremental ray tracer. As stated previously, attenuation with the volume of the scattering voxel was neglected. The total scattered photon signal at each detector pixel is thus a sum of the contributions from each phantom voxel. This process resulted in two scatter estimates, one for each of Compton and Rayleigh single scatter.

We simplified the scatter estimation process using several assumptions. First, all first-order scatter events for a given phantom voxel were assumed to occur at the voxel center, as opposed to being distributed throughout the voxel volume. Second, all photons scattered from a given phantom voxel to a given detector pixel were assumed to have a common trajectory, which was along a ray connecting the voxel center to the pixel center. Third, post-scatter attenuation of the scattered photons was calculated based on the energy of the incident photons rather than the energy of the scattered photons. Given that any Compton single scatter photons reaching the detector should have scatter angles of much less than 90° for any realistic imaging geometry, this simplification should not have introduced significant errors. Together, these simplifications dramatically reduced the computational burden associated with estimating scatter projections.

Calculating post-scatter attenuation was the most demanding portion of the simulation. For example, for an object consisting of $64 \times 64 \times 64$ voxels and a detector with 256×256 pixels, determining the post-scatter attenuation required tracing approximately 1.7×10^{10} rays. Consequently, we developed a

Table 1
EGSnrc Monte Carlo transport parameters used for all Epp simulations

Transport parameter	Value
Global ECUT	0
Global PCUT	0
Global SMAX	1e10
ESTEPE	0.25
XIMAX	0.5
Boundary crossing algorithm	EXACT
Skin depth for BCA	0
Electron-step algorithm	PRESTA-II
Spin effects	On
Brems angular sampling	Simple
Brems cross sections	BH
Bound compton scattering	Off
Pair angular sampling	Simple
Photoelectron angular sampling	Off
Rayleigh scattering	On
Atomic relaxations	On
Electron impact ionization	On

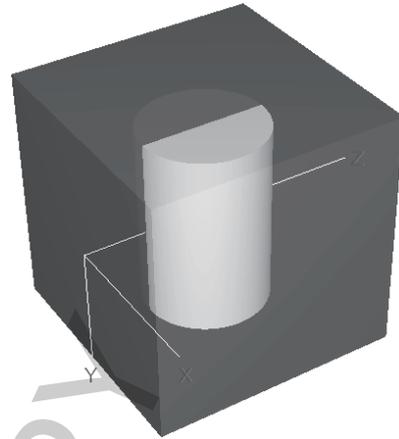


Fig. 3. Half-water (gray), half-aluminum (white) cylindrical phantom showing the simulation coordinate system. The central ray from the source to the detector is along the positive z -axis.

fast incremental ray tracer, which operated by stepping along a specified linear path with a user-defined step size. At each step, the cumulative radiological path length was updated by tri-linear interpolation of the linear attenuation coefficients for neighboring voxels. The ray tracer path did not include the scattering voxel itself.

2.3. Implementation using graphics processing units

Given that the ray tracer performed the same basic calculation repetitively using different subsets of voxel and pixel data as inputs, the ray tracing portion of the simulation was very amenable to GPU implementation. The code that ran on the GPU was a single function, the kernel. The GPU executed thousands of threads in parallel, each of which represented a single instance of the kernel. The number of registers available in the GPU limited the number of threads that could be executed simultaneously. The maximum number of threads was much less than the total number of rays to be traced, so the total simulation was divided into subsections, each corresponding to a certain number of phantom voxels.

Threads were grouped into blocks, and all threads in the same block could access shared memory. In our code, all threads in one block operated on a single voxel but different pixels, with voxel geometry data stored in shared memory. The first thread in every block read the material data for the voxel from global memory, performed some voxel-specific calculations and stored the results in the shared memory where all other threads in that block could access it. Then all subsequent threads calculated the scatter to the individual pixels and traced rays from the voxel center to the pixel centers. Source and detector configuration data were stored in constant memory, along with some phantom data, such as the number of voxels in each dimension. Attenuation data were stored in three-dimensional texture memory, which allowed the ray tracer to exploit hardware-accelerated interpolation.

2.4. Validation against Monte Carlo simulations

Monte Carlo simulations carried out using EGSnrc software have shown excellent agreement with measured data [8]. We thus validated our GPU-based analytical scatter estimates by comparing them to

scatter estimates from EGSnrc MC simulations of an identical imaging geometry. We used Epp, a C++ user code for EGSnrc that provides built-in capabilities for simulating scatter images with a voxelized geometry [20]. The EGSnrc MC simulation parameters used are listed in Table 1.

The EGSnrc simulations were executed on a Linux cluster, with seven nodes and 34 processors, running SUSE Linux 11. The cluster nodes were of varying ages, and thus contained different types of processors. The two newest nodes each had two Intel Xeon quad-core 3.16 GHz processors and 16 GB of RAM. Using the EGS script `run_user_code_batch`, each MC simulation was subdivided into eight smaller parallel processes which were distributed across available processors using the TORQUE scheduler. The analytical simulations were run on a single NVIDIA 9800 GX2 video card containing two 128-core GPUs. The card had 1 GB onboard memory and the clock rates for graphics, processor and memory were 600, 1500 and 1000 MHz, respectively.

We simulated two cylindrical phantoms, 10 cm in diameter and 10 cm in height, surrounded by air in a cube 12.8 cm on a side. One phantom was a homogeneous water cylinder, while in the other was a cylinder longitudinally divided into a water semi-cylinder and an aluminum semi-cylinder, as shown in Fig. 3. This phantom provided a severe test of the algorithm given the asymmetry in its material composition. For the Monte Carlo simulations, the phantoms were generated using the elementary geometry classes available in the `egspp` EGSnrc C++ class library. For the analytical simulations, the phantoms were represented by cubic voxelized volumes. Four versions of the water cylinder phantom were created for the analytical simulations, using grids of 16^3 , 32^3 , 64^3 and 128^3 voxels. The half-water, half-aluminum cylinder phantom used a 128^3 voxel grid. The largest phantom size possible was limited by the GPU card's memory to approximately 192^3 voxels.

The point source uniformly irradiated the front face of the phantom with 10^{11} photons. The number of photons was chosen to give a target mean statistical uncertainty of 1%, based on calculation of the expected mean single scatter photon count per detector pixel. The ideal detector consisted of a 256×256 grid of 2 mm pixels. The source to isocenter distance was 25 cm as was the isocenter to image distance. The compact imaging geometry specified was based on our interest in CBCT breast imaging.

The analytical and MC simulations carried out are listed in Table 2. Each simulation produced two scatter estimates, one for Compton single scatter and one for Rayleigh single scatter. Each estimate represented the counts per pixel of photons arriving at the detector that had undergone one Compton scatter or one Rayleigh scatter while traversing the phantom. The Epp MC code provided a built-in means of generating these estimates by tracking photon interactions.

We calculated the scatter to primary ratio (SPR) and single scatter fraction (SSF) for monoenergetic (38 keV) and polyenergetic (60 kVp) MC simulations using the 10 cm diameter water cylinder to determine if they were consistent with previously reported values. The SPR was calculated as the ratio of the mean of the scatter count to the mean of the primary count over the 16 central detector pixels. The SSF was calculated by taking the ratio between the sum of the single scatter count and the sum of the total scatter count over the whole detector. The SSF was calculated for both Compton and Rayleigh single scatter individually as well as their sum.

We obtained quantitative comparisons between the analytical and MC scatter estimates by calculating the root mean square difference between the analytical and MC projections over the image pixels. The RMS difference was scaled by the mean photon count in the MC images to obtain a percentage difference, ε ,

$$\varepsilon = \frac{100}{N\bar{I}_{MC}} \sqrt{\sum_j (I_{MC}(j) - I_A(j))^2} \quad (7)$$

Table 2
Summary of analytical and MC simulations

Simulation type	Phantom	Size		Source spectrum	
		cm	Voxels		
Analytical	10 cm Ø water cylinder	12.8	16 ³	38 keV monoenergetic	
			32 ³	38 keV monoenergetic	
			64 ³	38 keV monoenergetic	
			128 ³	32 keV monoenergetic	
			128 ³	38 keV monoenergetic	
			128 ³	46 keV monoenergetic	
			128 ³	46 keV monoenergetic	
Analytical	10 cm Ø water/aluminum cylinder	12.8	128 ³	32 keV monoenergetic	
			128 ³	38 keV monoenergetic	
			128 ³	46 keV monoenergetic	
MC	10 cm Ø water cylinder	12.8	n/a	32 keV monoenergetic	
			n/a	38 keV monoenergetic	
			n/a	46 keV monoenergetic	
			n/a	40 kVp spectrum	
			n/a	60 kVp spectrum	
			n/a	80 kVp spectrum	
	MC	10 cm Ø water/aluminum cylinder	12.8	n/a	32 keV monoenergetic
				n/a	38 keV monoenergetic
				n/a	46 keV monoenergetic
				n/a	40 kVp spectrum
				n/a	40 kVp spectrum
				n/a	60 kVp spectrum
				n/a	60 kVp spectrum

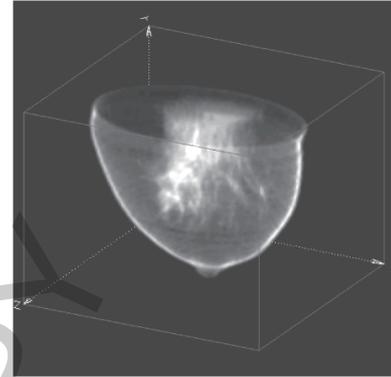


Fig. 4. Anthropomorphic breast phantom used to illustrate scatter artifact reduction in reconstructed images.

where I_{MC} and I_A are the MC and analytical scatter images, respectively, j is an index over the image pixels, N is the number of pixels in each image and \bar{I}_{MC} is the mean of the MC scatter image.

Using fewer voxels to represent the phantom will decrease computation time but is also expected to reduce accuracy. We thus investigated the impact of varying the degree of voxelization of the analytical phantom on the accuracy of the scatter estimates relative to the MC results. The 10 cm diameter water cylinder was modeled using phantoms of 16³, 32³, 64³ and 128³ voxels, corresponding to voxel sizes of 8, 4, 2 and 1 mm. We carried out monoenergetic analytical simulations at 38 keV with these phantoms and compared the scatter estimates with those from monoenergetic MC simulations at the same energy using ϵ values.

For simplicity and speed, the analytical estimator was restricted to monoenergetic incident photons. Scatter estimates for a polyenergetic source would require that the output from multiple simulations at discrete energies be summed using an appropriate spectral weighting. To determine whether monoenergetic scatter estimates differed substantially from polyenergetic estimates, we carried out monoenergetic and polyenergetic MC simulations for both phantoms and compared the results to monoenergetic analytical simulations using ϵ values. The mean energy of the polyenergetic spectra was used for the corresponding monoenergetic simulations.

We used SpekCalc software [21] to generate spectra for 40, 60 and 80 kVp tube potentials with filtration of 0.8 mm beryllium (representing the tube window) and 4 mm aluminum. The software provided the mean photon energy for each of the spectra, these being 32, 38 and 46 keV, respectively.

For the water/aluminum cylinder phantom, we qualitatively compared the 38 keV analytical and 60 kVp MC scatter estimates by displaying them as grayscale images, where the grayscale represents the scattered photon count in each detector element. We calculated a difference image between the two estimates for both Compton and Rayleigh single scatter and also displayed it as a grayscale image. We extracted horizontal profiles from the estimates representing the mean of the 8 central rows of detector pixels and plotted these profiles.

The original implementation of the analytical scatter estimator was a Matlab script which called a C executable optimized for calculating post-scatter attenuation [21]. To obtain an estimate of the reduction in computation time provided by the GPU implementation, we compared times required for the GPU-based estimator with those for the original Matlab/C code and the MC simulations.

We checked the statistical uncertainty of the MC simulations against our target level of 1%. We assumed a Poisson distribution for the number of photons arriving at a given detector pixel, and calculated uncertainty as the standard deviation over the mean converted to a percentage, $100 C^{-1/2}$, where C is the detected photon count.

2.5. Cone-beam reconstruction

To illustrate the potential of this method for scatter correction, we also simulated projection data for a voxelized anthropomorphic breast phantom. The phantom was comprised of $191 \times 150 \times 191$ (1 mm) voxels. The phantom consisted of an outer layer of skin and a mixture of 50% glandular tissue and adipose tissue, embedded in air. Figure 4 shows the phantom. The geometry was similar to that described above, with the breast pendant, and the incident photon energy was 38 keV. We performed a MC simulation of 360 projections of the phantom at intervals of one degree, with 3×10^9 incident photons per projection. The accelerated analytical algorithm was used to generate first order Compton and Rayleigh scatter images for all 360 projection. A modified version of an open-source version of the FDK algorithm was used to reconstruction the cone-beam data [23]. We reconstructed primary projections, total (primary plus scatter) projections and total projections minus the analytically computed first order scatter. We computed the RMS of the analytic scatter estimate relative to the MC estimate and the contrast-to-noise ratio (CNR) between adipose and 50% glandular tissue regions in the reconstructed images. The circular regions in the top left panel of Fig. 9 illustrate the areas used to measure the CNR using the following formula:

$$CNR = \frac{R_1 - R_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad (8)$$

where R_1/σ_1 and R_2/σ_2 are the average pixel values/standard deviations of the 50% glandular region and the adipose region, respectively.

3. Results and discussion

Table 3 gives the SPR and SSF for the monoenergetic and polyenergetic MC simulations of the 10 cm diameter water cylinder. Scatter represents approximately one-third of the total photon count at the detector center, and single scatter represents the majority of the scatter count over the whole detector. These values are consistent with those reported previously [19]. Accurate estimation of the single scatter count should thus prove useful for scatter compensation and correction.

The effects of varying the number of voxels used to represent the analytical 10 cm water cylinder phantom are shown in Table 4. The largest difference between analytical and MC estimates occurs, as expected, for the most coarsely discretized phantom, i.e. 8 mm voxels. The ε values are the same for 2 mm and 1 mm voxels, suggesting that 2 mm voxels are sufficient for acceptably accurate scatter estimation for this phantom. Given the results shown in Table 4 and the memory constraint of this GPU, all subsequent simulations used phantoms with 128^3 voxels. Newer GPU cards offer much higher levels of onboard memory and can support phantoms of 1024^3 voxels or larger.

Table 3

SPR and SSF for monoenergetic (38 keV) and polyenergetic (60 kVp) MC simulations of a 10 cm diameter water cylinder

Simulation	SPR (%)	SSF		
		Compton (%)	Rayleigh (%)	Total (%)
Monoenergetic (38 keV)	53	41	25	66
Polyenergetic (60 kVp)	49	42	24	66

Table 5

Scaled RMS differences, ϵ , between monoenergetic analytical simulations and corresponding monoenergetic and polyenergetic MC simulations of a 10 cm diameter water cylinder

ϵ			ϵ		
Monoenergetic MC			Polyenergetic MC		
Energy (keV)	Compton (%)	Rayleigh (%)	Energy (kVp)	Compton (%)	Rayleigh (%)
30	1.2	3.0	40	2.9	6.8
38	1.1	2.8	60	3.8	17.1
46	1.3	3.0	80	10.7	21.6

Table 4

Effect of varying the number of voxels used for the 10 cm diameter water cylinder phantom on the scaled RMS difference, ϵ , between monoenergetic analytical and MC simulations at 38 keV

Number of voxels	Voxel (mm) size	ϵ	
		Compton (%)	Rayleigh (%)
16^3	8	3.4	8.7
32^3	4	2.3	3.4
64^3	2	1.1	2.8
128^3	1	1.1	2.8

Table 6

Scaled RMS differences, ϵ , between monoenergetic analytical simulations and corresponding monoenergetic and polyenergetic MC simulations of a 10 cm diameter half-water, half-aluminum cylinder

ϵ			ϵ		
Monoenergetic MC			Polyenergetic MC		
Energy (keV)	Compton (%)	Rayleigh (%)	Energy (kVp)	Compton (%)	Rayleigh (%)
30	1.5	3.4	40	3.0	8.4
38	1.4	3.0	60	3.8	18.0
46	1.4	3.1	80	8.8	21.3

The effects of limiting the analytical simulation to a monoenergetic source are shown in Table 5, where monoenergetic analytical simulations are compared with monoenergetic and polyenergetic MC simulations for the 10 cm water cylinder. As expected, agreement between the monoenergetic analytical and MC estimates is within a few percent. Larger differences exist between the monoenergetic analytical and polyenergetic MC estimates, and the difference increased with beam energy. The differences remained, however, generally below 20%.

We expected the efficacy of representing a polyenergetic spectrum with a single energy to degrade with increasing spectral width. The spectral width increased substantially with increasing tube potential, corresponding to the increasing ϵ values. Characteristic radiation in the 80 kVp spectrum also contributed to the increased ϵ when comparing the polyenergetic 80 kVp MC result with the monoenergetic 46 keV analytical estimate. It would be possible in a practical application of this method to narrow the beam spectrum using additional filtration.

Representing a polyenergetic spectrum with a monoenergetic beam for estimating single scatter is a trade-off between accuracy and computation time. If the results of the monoenergetic analytical simulations provide sufficient accuracy for effective scatter correction, then the trade-off is worthwhile. If higher accuracy is required, then some form of polyenergetic simulation will be required. For example, using two representative energies with an appropriate weighting may prove sufficient. Further work is required to determine what single energy or weighted sum of multiple discrete energies adequately represents a polyenergetic spectrum for this application.

The half-water, half-aluminum cylinder is a more demanding simulation than the homogeneous water phantom because both scatter and attenuation are increased in the aluminum portion. Table 6 shows that the scaled RMS differences between analytical and MC scatter estimates for the polyenergetic case are generally only slightly higher than those of the water cylinder (Table 5), despite the strong asymmetry in

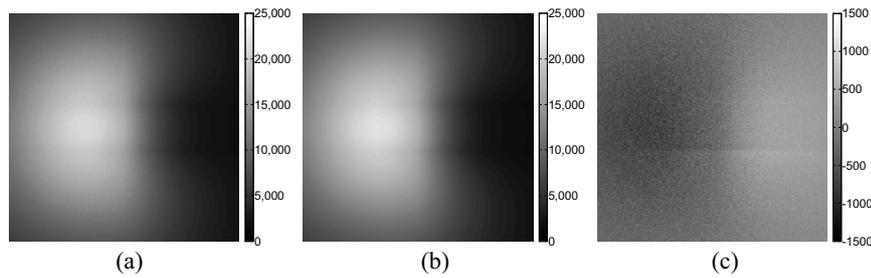


Fig. 5. Compton single scatter estimates as grayscale images for a half-water, half-aluminum cylindrical phantom from (a) the polyenergetic Monte Carlo simulation, (b) the monoenergetic analytical scatter estimate, and (c) the difference between them.

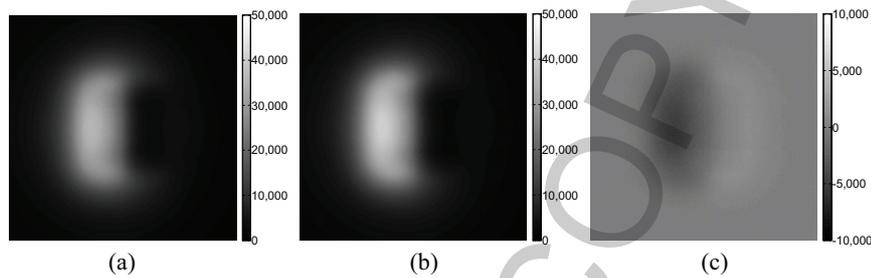


Fig. 6. Rayleigh single scatter estimates as grayscale images for a half-water, half-aluminum cylindrical phantom from (a) the polyenergetic Monte Carlo simulation, (b) the monoenergetic analytical scatter estimate and (c) the difference between them.

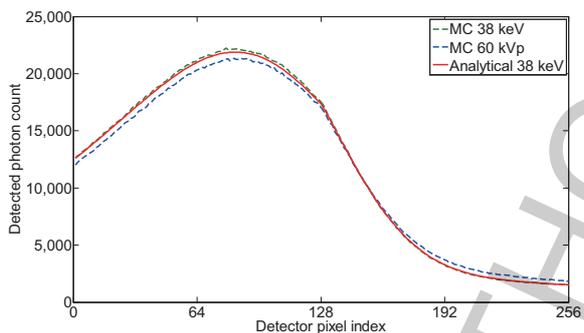


Fig. 7. Profiles representing the mean of the eight central detector rows for analytical and MC Compton single scatter estimates for a half-water, half-aluminum cylindrical phantom. (Colours are visible in the online version of the article; <http://dx.doi.org/10.3233/XST-150475>)

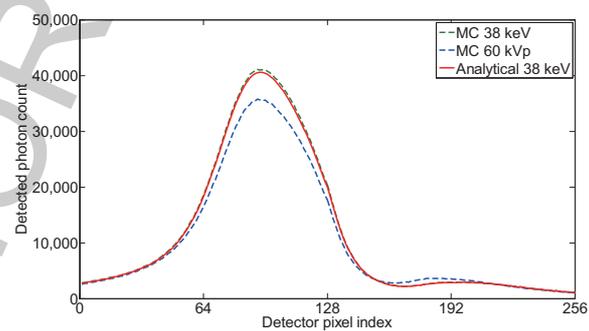


Fig. 8. Profiles representing the mean of the eight central detector rows for analytical and MC Rayleigh single scatter estimates for a half-water, half-aluminum cylindrical phantom. (Colours are visible in the online version of the article; <http://dx.doi.org/10.3233/XST-150475>)

electron density and linear attenuation between aluminum and water. The differences remain generally lower than 20%.

Figures 5 and 6 show the Compton and Rayleigh single scatter estimates for the 38 keV analytical and 60 kVp MC simulations of the water/aluminum cylinder displayed as grayscale images. The horizontal and vertical axes represent detector extent and the grayscale level corresponds to the photon count per detector pixel. As expected, the detected photon count behind the water portion of the phantom is much higher than behind the aluminum portion of the phantom. Additionally, the Compton single scatter dis-

Table 7

Time required for analytical (GPU and MATLAB/C) and MC simulations for scatter estimation for a 10 cm water cylinder irradiated by a 38 keV monoenergetic source

Simulation		Time	Processor
Analytical (GPU)	16 ³ voxels	1.1 s	GPU
	32 ³ voxels	4.2 s	GPU
	64 ³ voxels	48 s	GPU
	128 ³ voxels	670 s	GPU
MC	10 ¹¹ histories	21 h	CPU (8 processors)
Analytical (MATLAB/C)	16 ³ voxels	150 s	CPU (single processor)
	32 ³ voxels	1700 s	CPU (single processor)
	64 ³ voxels	6.2 h	CPU (single processor)
	128 ³ voxels	96 h	CPU (single processor)

Table 8

Scaled RMS differences, ε , between analytical simulations and corresponding MC simulations of the anthropomorphic breast phantom

Projection angle	Type	RMS difference (%)
3 degrees	Compton	1.23
	Rayleigh	1.80
72 degrees	Compton	1.58
	Rayleigh	2.27
186 degrees	Compton	1.24
	Rayleigh	1.81

tribution is much more diffuse than the Rayleigh distribution. The difference images, Figs 5(c) and 6(c), show that the analytical estimator overestimates the photon count behind the water and underestimates it behind the aluminum for both Compton and Rayleigh single scatter.

Figures 7 and 8 show horizontal profiles taken from the single scatter estimates shown in Figs 5 and 6. In general, analytical and MC estimates agree well for Compton single scatter while larger discrepancy is observed between analytical and MC estimates for Rayleigh single scatter.

Over all, the monoenergetic analytical estimate appears to be a reasonable representation of the polyenergetic MC estimate for Compton single scatter for this phantom. However, the agreement is not as good for Rayleigh single scatter. The Rayleigh single scatter distribution is much more sharply peaked than the Compton distribution, and the difference between monoenergetic analytical and polyenergetic MC estimates is greatest in the region of highest scatter count.

There are several possible sources of discrepancy between the analytical and MC estimates. Both the analytical estimator and the EGSnrc MC package used the Klein-Nishina approximation for Compton scatter and the form factor approximation using the free gas model for Rayleigh scatter. We extracted relevant material data, including compositions, densities, linear attenuation coefficients and form factors, directly from the EGSnrc data files and used these in the analytical estimator. The discrepancies between the analytical and monoenergetic MC results would thus have resulted from the geometrical model simplifications, the fact that we neglected the energy change in Compton scattered photons when calculating post-scatter attenuation and possible inaccuracies introduced by the ray tracer. When comparing the analytical and polyenergetic MC estimates, larger differences were expected given the energy dependence of scatter cross-sections and attenuation coefficients. The strongly forward peaked nature of Rayleigh scatter relative to the smoother distribution of Compton scatter would also have amplified the discrepancies.

Table 7 gives the time required for analytical and MC simulations of the 10 cm water cylinder phantom with the 38 keV monoenergetic source. To illustrate the speed-ups possible with GPU processing, we have included timing results for the same analytical simulations using our original MATLAB/C code. Times shown for the analytical simulations were obtained using a single NVIDIA 9800 GX2 dual-GPU video card. The MC simulation was distributed over eight processors on the Linux cluster, so the time given represents roughly one-eighth of the time that would be required using a single processor. The MATLAB/C simulations were also run on the Linux cluster, but using only one Intel Xeon 3.16 GHz processor. Note that we used a single dual-GPU video card. More acceleration could be expected from using additional cards or more recent and powerful ones.

Table 9

CNR between adipose and 50% glandular tissue in the reconstructed images of the anthropomorphic breast phantom

Projection angle	Primary reconstruction	Total projection reconstruction	Reconstruction with 1 st order scatter correction
Central axial slice ($z = 0$)	1.38	0.18	0.57

The time required for the analytical simulations (both GPU and MATLAB/C) increased nonlinearly with increasing phantom voxelization. The time required for both implementations of the analytical estimator (GPU and MATLAB/C) was primarily a function of the number of rays traced and the number of steps along each ray for calculating post-scatter attenuation. As expected, time required for the MC simulations scaled nearly linearly with the number of histories simulated.

For a single NVIDIA 9800 GX2 video card with two primary GPUs, and 128 cores per GPU, the time required to generate Compton and Rayleigh single scatter estimates was 670 seconds. If a slightly coarser level of discretization was considered acceptable for scatter correction, a 64^3 voxel phantom required only 48 seconds. Using multiple video cards of the same type, the processing time should scale as the reciprocal of the number of GPUs available. Assuming the use of multiple GPU cards with faster processors and more onboard memory, it should be possible to obtain reasonably accurate single scatter estimates in a few seconds.

Extending this scenario to CBCT requires scatter estimation for a large number of projections and would drastically increase the computing load and time required. However, given that scatter distributions are in general smooth and slowly varying, it should not be necessary to estimate scatter images for every projection angle. Scatter images could be estimated for a subset of projection angles, perhaps every 10° to 20° , and interpolated to provide data for scatter-corrected reconstructions. Additional GPU cards and computational resources would also enable this computation to be performed in a practical amount of time.

In the MC simulations, our goal for statistical uncertainty in our scatter estimates was 1%, averaged over all detector pixels. The MC simulation with the lowest mean photon count at the detector was the half-water, half-aluminum cylinder. Using 10^{11} incident photons produced mean photon counts at the detector for Compton and Rayleigh single scatter of 9978 and 7229, respectively, for the 38 keV monoenergetic case and 9791 and 6644, respectively, for the 60 kVp polyenergetic case. These photon counts resulted in statistical uncertainties of 1.0%, 1.2%, 1.0% and 1.2%, respectively. The maximum mean statistical uncertainty of the MC estimates was thus roughly 1%, given that all other simulations had higher mean photon counts at the detector.

There was good agreement between MC and analytically estimated scatter for the anthropomorphic phantom. The RMS difference between the MC and analytical first order scatter is listed in Table 8 for 3 projections selected arbitrarily.

Figure 9 illustrates the central axial slice through the original phantom, the primary projection reconstruction, the total projection reconstruction and the image reconstructed from the total projection minus the first-order analytic scatter estimate. The same display scale was used for all the images.

The reconstructed images show depressed attenuation values in the case of total projections and no scatter compensation. These images demonstrate that scatter can cause the CNR to drop significantly. With the first-order scatter correction, the CNR increases by more than a factor of 3 (Table 9). Similar results were obtained in other slices of the phantom but are not shown here for the sake of brevity. The aim of this paper was to present the GPU-accelerated computation of first order scatter, not to present a fully developed scatter correction algorithm. Scatter correction requires estimating multiple scatter and knowledge of the material information necessary to compute first order scatter. It is feasible, for

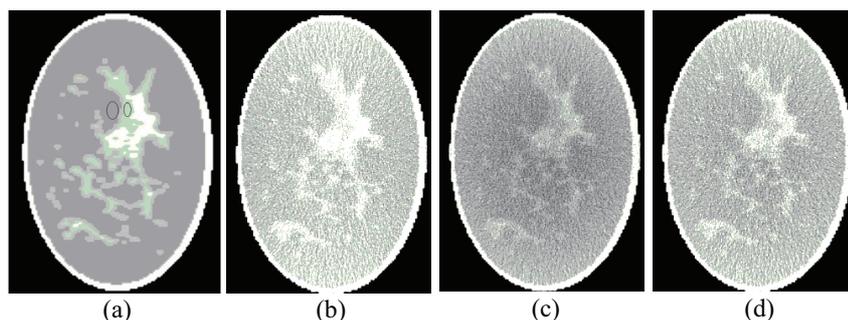


Fig. 9. Central axial slice of anthropomorphic breast phantom. (a) original phantom slice. The elliptical regions show the areas used for CNR calculation. (b) reconstruction of primary projections. (c) reconstruction of total projections. (d) reconstruction of total projections minus analytically estimated first order scatter. (Colours are visible in the online version of the article; <http://dx.doi.org/10.3233/XST-150475>)

example, to combine the analytic first-order scatter simulation with direct scatter measurements obtained in the shadow of the collimator [24]. It may also be possible, in the case of breast imaging, to assume homogenous tissue and use the material information of 50% glandular tissue, for example. These topics are the subject of future work.

4. Conclusions

We have presented an analytical method to estimate first-order Compton and Rayleigh scatter at a detector for a voxelized heterogeneous object in CBCT geometry. Our analytical scatter estimator was implemented in the CUDA environment for execution on an NVIDIA GPU. We validated our analytical estimates by comparing them with MC scatter estimates.

When comparing monoenergetic simulations for two different phantoms at three different energies, we obtained very good agreement ($\varepsilon < 5\%$) for both Compton and Rayleigh single scatter, we obtained useful agreement ($\varepsilon_{\max} \sim 10\%$) between monoenergetic analytical estimates and polyenergetic MC estimates for Compton single scatter. Agreement between monoenergetic analytical and polyenergetic MC estimates for Rayleigh single scatter was reasonable ($\varepsilon_{\max} \sim 20\%$). For both Compton and Rayleigh single scatter, the degree of discrepancy increased with energy. To achieve higher accuracy, polyenergetic analytical simulations using a weighted sum of monoenergetic contributions may be necessary. Further work on the most accurate means of representing a polyenergetic spectrum with a single energy or a small number of discrete energies in this context is required.

The GPU implementation of our analytical single scatter estimator provided a tremendous speed-up relative to both our CPU-based analytical code and to MC simulation using EGSnrc. A GPU-based analytical estimator represents a useful step towards patient-specific scatter estimation and correction. The method could be combined with empirical scatter measurements in the collimator shadows, in a manner similar to that proposed by Siewerdsen et al. [24] to more fully characterize the scatter count at the detector. The ongoing rapid increase in GPU computing power available will further reduce computational time requirements and enable clinical use of such methods in the future.

Acknowledgement

The authors wish to thank the CancerCare Manitoba Foundation for funding support for this research.

References

- [1] B. Chen and R. Ning, Cone-beam volume CT breast imaging: Feasibility study, *Medical Physics* **29** (2002), 755–770.
- [2] J.H. Siewerdsen and D.A. Jaffray, Cone-beam computed tomography with a flat-panel imager: Magnitude and effects of x-ray scatter, *Medical Physics* **28** (2001), 220–231.
- [3] E.P. Ruhrnschopf and K. Klingenberg, A general framework and review of scatter correction methods in x-ray cone-beam computerized tomography, Part 1: Scatter compensation approaches, *Med Phys* **38**, 4296–4311.
- [4] Ruhrnschopf, Ep and K. Klingenberg, A general framework and review of scatter correction methods in cone beam CT, Part 2: Scatter estimation approaches, *Med Phys* **38**, 5186–5199.
- [5] R.G. Wells, A. Celler and R. Harrop, Analytical calculation of photon distributions in SPECT projections, *IEEE Transactions on Nuclear Science* **45** (1998), 3202–3214.
- [6] G. Jarry, S.A. Graham, D.J. Moseley, D.A. Jaffray, J.H. Siewerdsen and F. Verhaegen, Characterization of scattered radiation in kV CBCT images using Monte Carlo simulations, *Medical Physics* **33** (2006), 4320–4329.
- [7] W. Zbijewski and F.J. Beekman, Efficient Monte Carlo based scatter artifact reduction in cone-beam micro-CT, *IEEE Transactions on Medical Imaging* **25** (2006), 817–827.
- [8] D.W.O. Rogers, Fifty years of Monte Carlo simulations for medical physics, *Physics in Medicine and Biology* **51** (2006), R287–301.
- [9] E. Mainegra-Hing and I. Kawrakow, Fast Monte Carlo calculation of scatter corrections for CBCT images, *Journal of Physics: Conference Series* **102** (2008).
- [10] W. Yao and K.W. Leszczynski, An analytical approach to estimating the first order x-ray scatter in heterogeneous medium, *Medical Physics* **36** (2009), 3145–3156.
- [11] W. Yao and K.W. Leszczynski, An analytical approach to estimating the first order scatter in heterogeneous medium, II, A practical application, *Medical Physics* **36** (2009), 3157–3167.
- [12] J.E.A. Aviles, S. Pistorius, I.A. Elbakri, R. Gordon and B. Ahmad, A 1st generation scatter CT algorithm for electron density breast imaging which accounts for bound incoherent, coherent and multiple scatter: A Monte Carlo study, *Journal of X-Ray Science and Technology* **19** (2011), 477–499.
- [13] T. Chighvinadze and S. Pistorius, The effect of detector size and energy resolution on image quality in multi-projection Compton scatter tomography, *Journal of X-Ray Science and Technology* **22**(1) (2014), 113–128.
- [14] W. Cong and G. Wang, X-ray scattering tomography for biological applications, *Journal of X-Ray Science and Technology* **19** (2011), 219–227.
- [15] G. Pratz and L. Xing, GPU computing in medical physics: A review, *Medical Physics* **38**, 2685–2697.
- [16] J. Wang, Z. Chi and Y. Wang, Analytic reconstruction of Compton scattering tomography, *Journal of Applied Physics* **86** (1999), 1693–1698.
- [17] F.H. Attix, Introduction to radiological physics and radiation dosimetry, Wiley-VCH, Weinheim, 2004.
- [18] M.E. Poletti, D. Goncalves and I. Mazzaro, X-ray scattering from human breast tissues and breast-equivalent materials, *Physics in Medicine and Biology* **47** (2002), 47–63.
- [19] Y. Kyriakou, M. Meyer and W.A. Kalender, Technical note: Comparing coherent and incoherent scatter effects for cone-beam CT, *Physics in Medicine and Biology* **53** (2008), 175–85.
- [20] J. Lippuner, I.A. Elbakri, C. Cui and H.R. Ingleby, Epp: A C++ EGSnrc user code for x-ray imaging and scattering simulations, *Medical Physics* **38** (2011), 1705–1708.
- [21] G. Poludniowski, G. Landry, F. DeBlois, P.M. Evans and F. Verhaegen, SpekCalc: A program to calculate photon spectra from tungsten anode x-ray tubes, *Physics in Medicine and Biology* **54** (2009), 433–438.
- [22] H.R. Ingleby, I.A. Elbakri, D.W. Rickey and S. Pistorius, Analytical scatter estimation for cone-beam computed tomography, in: E. Samei and J. Hsieh, eds, SPIE 7258: *Physics of Medical Imaging*, SPIE, Orlando, FL, 2009.
- [23] OSCaR: An open-source cone-beam CT reconstruction tool for imaging research.
- [24] J.H. Siewerdsen, M.J. Daly, B. Bakhtiar, D.J. Moseley, S. Richard, H. Keller and D.A. Jaffray, A simple, direct method for x-ray scatter estimation and correction in digital radiography and cone-beam CT, *Medical Physics* **33** (2006), 187–197.