Lookup table-based sampling of the phase function for Monte Carlo simulations of light propagation in turbid media

PETER NAGLIČ,* FRANJO PERNUŠ, BOŠTJAN LIKAR AND MIRAN BÜRMEN

University of Ljubljana, Faculty of Electrical Engineering, Laboratory of Imaging Technologies, Tržaška cesta 25, SI-1000, Ljubljana, Slovenia
peter.naglic@fe.uni-lj.si

Abstract: Analytical expressions for sampling the scattering angle from a phase function in Monte Carlo simulations of light propagation are available only for a limited number of phase functions. Consequently, numerical sampling methods based on tabulated values are often required instead. By using Monte Carlo simulated reflectance, we compare two existing and propose an improved numerical sampling method and show that both the number of the tabulated values and the numerical sampling method significantly influence the accuracy of the simulated reflectance. The provided results and guidelines should serve as a good starting point for conducting computationally efficient Monte Carlo simulations with numerical phase function sampling.

References and links

1. B. C. Wilson and G. Adam, "A Monte Carlo model for the absorption and flux distributions of light in tissue," Med. Phys. **10**(6), 824–830 (1983).
2. L. Wang, S. L. Jacques, and L. Zheng, "MCML--Monte Carlo modeling of light transport in multi-layered tissues," Comput. Methods Programs Biomed. **47**(2), 131–146 (1995).
3. F. van Leeuwen-van Zaane, U. A. Gamm, P. B. A. A. van Driel, T. J. A. Snoeks, H. S. de Bruijn, A. van der Ploeg-van den Heuvel, I. M. Mol, C. W. G. M. Löwik, H. J. C. M. Sterenborg, A. Amelink, and D. J. Robinson, "*In vivo* quantification of the scattering properties of tissue using multi-diameter single fiber reflectance spectroscopy," Biomed. Opt. Express **4**(5), 696–708 (2013).
4. K. W. Calabro and I. J. Bigio, "Influence of the phase function in generalized diffuse reflectance models: review of current formalisms and novel observations," J. Biomed. Opt. **19**(7), 75005 (2014).
5. J. J. Bravo, K. D. Paulsen, D. W. Roberts, and S. C. Kanick, "Sub-diffuse optical biomarkers characterize localized microstructure and function of cortex and malignant tumor," Opt. Lett. **41**(4), 781–784 (2016).
6. S. C. Kanick, D. M. McClatchy 3rd, V. Krishnaswamy, J. T. Elliott, K. D. Paulsen, and B. W. Pogue, "Sub-diffusive scattering parameter maps recovered using wide-field high-frequency structured light imaging," Biomed. Opt. Express **5**(10), 3376–3390 (2014).
7. N. Bodenschatz, P. Krauter, A. Liemert, and A. Kienle, "Quantifying phase function influence in subdiffusively backscattered light," J. Biomed. Opt. **21**(3), 35002 (2016).
8. D. M. McClatchy 3rd, E. J. Rizzo, W. A. Wells, P. P. Cheney, J. C. Hwang, K. D. Paulsen, B. W. Pogue, and S. C. Kanick, "Wide-field quantitative imaging of tissue microstructure using sub-diffuse spatial frequency domain imaging," Optica **3**(6), 613–621 (2016).
9. L. G. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," Astrophys. J. **93**, 70–83 (1941).
10. J. R. Mourant, J. Boyer, A. H. Hielscher, and I. J. Bigio, "Influence of the scattering phase function on light transport measurements in turbid media performed with small source-detector separations," Opt. Lett. **21**(7), 546–548 (1996).
11. F. Bevilacqua and C. Depeursinge, "Monte Carlo study of diffuse reflectance at source-detector separations close to one transport mean free path," J. Opt. Soc. Am. A **16**(12), 2935–2945 (1999).
12. L. O. Reynolds and N. J. McCormick, "Approximate two-parameter phase function for light scattering," J. Opt. Soc. Am. **70**(10), 1206–1212 (1980).
13. A. N. Yaroslavsky, I. V. Yaroslavsky, T. Goldbach, and H.-J. Schwarzmaier, "Influence of the scattering phase function approximation on the optical properties of blood determined from the integrating sphere measurements," J. Biomed. Opt. **4**(1), 47–53 (1999).
14. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, 1983).

15. B. Gélébart, E. Tinet, J. M. Tualle, and S. Avrillier, "Phase function simulation in tissue phantoms: a fractal approach," Pure Appl. Opt. J. Eur. Opt. Soc. Part A **5**(4), 377–388 (1996).
16. D. Toublanc, "Henyey-Greenstein and Mie phase functions in Monte Carlo radiative transfer computations," Appl. Opt. **35**(18), 3270–3274 (1996).
17. J. R Zijp and J. J. Bosch, "Use of tabulated cumulative density functions to generate pseudorandom numbers obeying specific distributions for Monte Carlo simulations," Appl. Opt. **33**(3), 533–534 (1994).
18. A. J. Welch and M. J. Van Gemert, *Optical-Thermal Response of Laser-Irradiated Tissue* (Springer, 2010).
19. P. Krauter, S. Nothelfer, N. Bodenschatz, E. Simon, S. Stocker, F. Foschum, and A. Kienle, "Optical phantoms with adjustable subdiffusive scattering parameters," J. Biomed. Opt. **20**(10), 105008 (2015).
20. J. M. Schmitt and G. Kumar, "Optical scattering properties of soft tissue: a discrete particle model," Appl. Opt. **37**(13), 2788–2797 (1998).
21. R. K. Wang, "Modelling optical properties of soft tissue by fractal distribution of scatterers," J. Mod. Opt. **47**(1), 103–120 (2000).
22. E. Alerstam, T. Svensson, and S. Andersson-Engels, "Parallel computing with graphics processing units for high-speed Monte Carlo simulation of photon migration," J. Biomed. Opt. **13**(6), 060504 (2008).
23. U. Utzinger and R. R. Richards-Kortum, "Fiber optic probes for biomedical optical spectroscopy," J. Biomed. Opt. **8**(1), 121–147 (2003).
24. X. Ma, J. Q. Lu, H. Ding, and X.-H. Hu, "Bulk optical parameters of porcine skin dermis at eight wavelengths from 325 to 1557 nm," Opt. Lett. **30**(4), 412–414 (2005).
25. J. F. Beek, P. Blokland, P. Posthumus, M. Aalders, J. W. Pickering, H. J. Sterenborg, and M. J. C. van Gemert, "In vitro double-integrating-sphere optical properties of tissues between 630 and 1064 nm," Phys. Med. Biol. **42**(11), 2255–2261 (1997).
26. P. Thueler, I. Charvet, F. Bevilacqua, M. St Ghislain, G. Ory, P. Marquet, P. Meda, B. Vermeulen, and C. Depeursinge, "*In vivo* endoscopic tissue diagnostics based on spectroscopic absorption, scattering, and phase function properties," J. Biomed. Opt. **8**(3), 495–503 (2003).
27. A. J. Radosevich, A. Eshein, T.-Q. Nguyen, and V. Backman, "Subdiffusion reflectance spectroscopy to measure tissue ultrastructure and microvasculature: model and inverse algorithm," J. Biomed. Opt. **20**(9), 097002 (2015).
28. R. Hennessy, S. L. Lim, M. K. Markey, and J. W. Tunnell, "Monte Carlo lookup table-based inverse model for extracting optical properties from tissue-simulating phantoms using diffuse reflectance spectroscopy," J. Biomed. Opt. **18**(3), 037003 (2013).
29. S. C. Kanick, V. Krishnaswamy, U. A. Gamm, H. J. Sterenborg, D. J. Robinson, A. Amelink, and B. W. Pogue, "Scattering phase function spectrum makes reflectance spectrum measured from Intralipid phantoms and tissue sensitive to the device detection geometry," Biomed. Opt. Express **3**(5), 1086–1100 (2012).

1. Introduction

One of the central mechanisms of light propagation in turbid media such as biological tissue is scattering. In terms of Monte Carlo (MC) simulations, which are often used for modeling the light propagation [1,2], scattering is described by the scattering coefficient and scattering angle typically drawn from the inverse cumulative distribution of the phase function. In recent studies, phase functions have become increasingly important in modeling of the reflectance obtained by spatially-resolved reflectance spectroscopy [3–5] or spatial-frequency domain reflectance spectroscopy [6–8], in particular for small source-detector separations or high spatial frequencies. The Henyey-Greenstein (HG) phase function [9] is most commonly used in the biomedical community since it offers an analytical inverse of the cumulative distribution function (CDF), which is convenient for sampling the scattering angles by a random number drawn from a uniform distribution. However, the HG phase function is well known to underestimate large-angle backward scattering observed in turbid media such as tissue [10]. Consequently, other phase functions were investigated, such as the modified Henyey-Greenstein (MHG) [11], Gegenbauer kernel (GK) [12,13], Mie [14] and Mie fractal [15] phase functions. Unlike the HG and GK, the MHG, Mie and Mie fractal phase functions do not offer analytical inverse of the CDF. In order to use these types of phase functions in the MC simulations, the phase functions have to be numerically sampled. For this purpose, Toublanc has proposed a method for computing the scattering angle from tabulated evenly spaced scattering angles that through the CDF correspond to a random number drawn from a uniform distribution [16]. Zijp *et al.* have proposed a similar approach, however, tabulating the evenly spaced values of the inverse CDF that correspond to the scattering angle [17]. Upon a drawn random number, the scattering angle can be instantly deduced without additional computation. From our experience, we have noticed that the number of the tabulated CDF values as well as the sampling method applied to the phase function can

significantly influence the accuracy of the MC simulated reflectance. Most of the studies that have used numerical phase function sampling as described above, lack the information on the details of the employed sampling method [4–7,10,11]. Consequently, it is very difficult to properly repeat the experiments conducted in those studies.

In this study, we extensively evaluate the existing implementations of the numerical phase function sampling for MC simulations following the methodology of Toublanc and Zijp *et al.*, and propose an improved sampling scheme utilizing non-linearly spaced tabulated CDF values. We show that the accuracy of the MC simulated reflectance acquired by optical fiber probes at various source-detector separations (SDS) significantly depends on the number of the tabulated CDF values and the numerical phase function sampling method, and can become significantly degraded, if the tabulated CDF is too sparse. Moreover, we show that the most prominent phase function property that governs the required number of the tabulated CDF values is the anisotropy factor. This study should serve as a reminder that all the details of the numerical phase function sampling methodology should be carefully considered when performing MC simulations.

2. Theoretical background on sampling from the phase functions

2.1 Phase functions

In general, a scattering phase function is defined as a probability density function $p(\boldsymbol{s}, \boldsymbol{s}')$ for a photon deflecting from the propagation direction $\boldsymbol{s}'$ in the direction $\boldsymbol{s}$. Although turbid media are sometimes anisotropic, biological tissues are often assumed to be isotropic, where the scattering depends only on the angle θ between $\boldsymbol{s}$ and $\boldsymbol{s}'$. Consequently, the phase function can be written as $p(\varphi, \cos \theta)$, where φ is the azimuthal angle given that the z axis is aligned with the propagation direction $\boldsymbol{s}'$ [18]. Under the assumption that φ and $\cos \theta$ are independent variables, the phase function can be written as $p(\varphi, \cos \theta) = p(\varphi) \, p(\cos \theta)$. In the literature, the term phase function is most commonly used for the probability density of $\cos \theta$, since $p(\varphi) = 1/(2\pi)$ for symmetric scatterers. Consequently, we will use the term phase function exclusively for $p(\cos \theta)$ throughout the remainder of this study.

The most commonly used phase function in the biomedical community is the Henyey-Greenstein (HG) phase function [9] that allows simple and computationally efficient description of scattering in biological tissues:

$$p_{\mathrm{HG}}\left(\cos\theta\right) = \frac{1}{2}\frac{1-g^{2}}{\left(1+g^{2}-2g\cos\theta\right)^{3/2}} \,, \tag{1}$$

where g is the anisotropy factor of the phase function. However, since it has been shown that the HG phase function underestimates large-angle scattering [10], other phase functions such as the modified Henyey-Greenstein (MHG) and Gegenbauer kernel (GK) have been proposed. The MHG phase function is defined as a weighted sum of the HG phase function and Rayleigh scattering [11]:

$$p_{\mathrm{MHG}}(\cos\theta) = \beta \, p_{\mathrm{HG}}(\cos\theta) + (1-\beta)\frac{3}{2}\cos^{2}\theta \,, \tag{2}$$

where β adjusts their relative contribution. The GK phase function is a mathematical extension of the HG phase function with an additional free parameter α_{GK} [12,13,19]:

$$p_{\mathrm{GK}}(\cos\theta) = \frac{2\,\alpha_{\mathrm{GK}}\,g_{\mathrm{GK}}(1-g_{\mathrm{GK}}^{2})^{2\,\alpha_{\mathrm{GK}}}}{\left[(1+g_{\mathrm{GK}})^{2\,\alpha_{\mathrm{GK}}} - (1-g_{\mathrm{GK}})^{2\,\alpha_{\mathrm{GK}}}\right](1+g_{\mathrm{GK}}^{2} - 2\,g_{\mathrm{GK}}\,\cos\theta)^{(1+\alpha_{\mathrm{GK}})}} \,. \tag{3}$$

It should be noted that g_{GK} is in general not equal to the anisotropy factor of the GK phase function.

Another common way of describing phase functions in biological tissues is by utilizing the rigorous Mie theory for spherical particles, where a tissue phase function can be expressed as a weighted sum of phase functions arising from different spherical particles. A popular approach is to sum the phase functions of spherical particles that follow the discrete particle fractal size distribution, which was found to adequately describe structural properties of biological tissues and the resulting refractive index variations that cause light scattering [15,20,21]. The discrete particle fractal size distribution is defined as

$$n(d_i) = A d_i^{-\alpha} , \qquad (4)$$

where $n(d_i)$ is the number density of spherical particles with diameter d_i, α the power of the distribution and A is related to the total number density of spherical particles, which determines the amplitude of the scattering coefficient. The phase function of a single spherical particle of diameter d_i is given by [14]

$$p_{\mathrm{Mie}}(\cos\theta) = \frac{\pi}{C_{\mathrm{sca}}} \frac{\left|S_1(\cos\theta)\right|^2 + \left|S_2(\cos\theta)\right|^2}{k^2} , \qquad (5)$$

where C_{sca} is the scattering cross section of the spherical particle, $k = 2\pi n_{\mathrm{med}}/\lambda$, and $S_1(\cos\theta)$ and $S_2(\cos\theta)$ are the elements of the amplitude scattering matrix. In this case, n_{med} represents the refractive index of the medium in which the sphere is suspended and λ the wavelength of light in vacuum. Note that C_{sca}, $S_1(\cos\theta)$ and $S_2(\cos\theta)$ are all functions of the ratio between the refractive index of the medium n_{med} and the refractive index of the spherical particle n_{par}, and also of a dimensionless parameter $x = \pi n_{\mathrm{med}} d/\lambda$. To obtain the phase function of the spherical particles that follow the discrete particle fractal size distribution, the individual Mie phase functions have to be summed in the following way

$$p_{\mathrm{fractal}}(\cos\theta) = \frac{\sum_i d_i^{-\alpha} C_{\mathrm{sca},i} \, p_{\mathrm{Mie},i}(\cos\theta)}{\sum_i d_i^{-\alpha} C_{\mathrm{sca},i}} , \qquad (6)$$

where the two summations go over all the diameters d_i of spherical particles. Thus the obtained phase function depends on the fractal volume dimension α, the lower and upper bounds of the summation and the diameters of the spherical particles.

2.2 Analytical sampling of the phase function

An essential part of the MC simulations of light propagation in turbid media are the scattering events. At each scattering event, the photon packet undergoes a deflection by a scattering angle, which is related to the phase function of the turbid medium. To sample the φ and $\cos\theta$ in the MC simulations, the CDFs of the probability density functions $p(\varphi)$ and $p(\cos\theta)$ have to be derived

$$CDF_\varphi(\varphi') = \frac{1}{2\pi} \int_0^{\varphi'} \mathrm{d}\varphi ,$$

$$CDF_{\cos\theta}(\cos\theta') = \int_{-1}^{\cos\theta'} p(\cos\theta) \, \mathrm{d}\cos\theta , \qquad (7)$$

where $p(\varphi)$ can be set to $1/(2\pi)$ if all the azimuthal angles are equally likely. The obtained CDFs are then usually equated to uniformly distributed random numbers ξ_1 and ξ_2 from [0, 1]. Subsequently, the sampled φ and $\cos\theta$ can be derived by evaluating the corresponding inverse CDF

$$\varphi = 2\pi\xi_1 \, ,$$
$$\cos\theta = CDF_{\cos\theta}^{-1}\left(\xi_2\right).$$

(8)

The computation of an inverse CDF significantly depends on the form of the phase function. For example, the HG phase function is very convenient since the inverse CDF can be expressed analytically. Similar analytical expression can be also found for the GK phase function [13]. Consequently, $\cos\theta$ can be sampled from a closed form analytical expression. Analytical forms of the inverse CDF allow simple and computationally efficient implementation in the MC model. Such implementation is especially important, if the MC model is intended to run on a graphics processing unit (GPU).

2.3 Numerical phase function sampling

If an analytical inverse of the CDF given by the second term in Eq. (8) is not available or the analytical expression is too extensive, the sampling of the $\cos\theta$ can be conducted numerically. Commonly, the sampling from a probability distribution can be performed with the well-known rejection sampling method. However, this method is computationally very inefficient in particular for probability distribution containing peaks, which is the case with the forward-peaked phase functions of biological tissues [17]. Alternatively, a lookup table-based sampling of $\cos\theta$ can be employed.

In the following we present three lookup table-based methods for sampling the scattering angle from the phase function that can be easily incorporated in the MC simulations.

2.3.1 Forward lookup table (FLT) sampling method

The entries of the forward lookup table (FLT) sampling method are the discrete CDF_j values calculated for each $\cos\theta_j$:

$$CDF_j = \int_{-1}^{\cos\theta j} p\left(\cos\theta\right) \, d\cos\theta \, ,$$

(9)

where j goes from 0 to N, $\cos\theta_0 = -1$ and $\cos\theta_N = 1$. Given a random number ξ drawn from a uniform distribution, the task is to find the value of $\cos\theta_\xi$, such that $CDF(\cos\theta_\xi)$ corresponds to ξ (Fig. 1(a)). This can be accomplished by finding an index J for which [16]

$$CDF_J < \xi \leq CDF_{J+1}.$$

(10)

Once index J is found, the sampled $\cos\theta_\xi$ can be estimated by a linear interpolation between the $\cos\theta_J$ and $\cos\theta_{J+1}$ values corresponding to CDF_J and CDF_{J+1}:

$$\cos\theta_\xi = \left(\frac{\cos\theta_{J+1} - \cos\theta_J}{CDF_{J+1} - CDF_J}\right)\left(\xi - CDF_J\right) + \cos\theta_J \, .$$

(11)

The most computationally intensive part of the FLT sampling method is the process of finding index J. For a given value of ξ this can be accomplished by one of the well-known root-finding methods such as bisection or Newton-Raphson method. However, searching through a lookup table in this way is very time consuming and also depends on the size of the lookup table.

To avoid the pitfalls of the FLT sampling method, we describe two alternative numerical approaches.
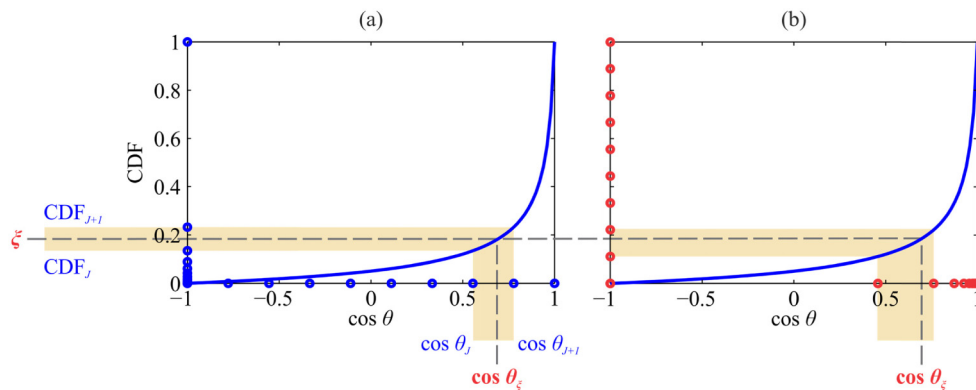
Fig. 1. An example of the CDF for the HG phase function. Drawn random number ξ and the sampled $\cos \theta_\xi$ are given in red. (a) The forward lookup table ($\cos \theta_j$ values and their corresponding CDF_j values) is presented with blue circles. The values of $\cos \theta_j$ in the forward lookup table are evenly spaced. (b) The inverted lookup table ($\cos \theta_j$ values and their corresponding CDF_j values) is presented with red circles. The CDF_j values in the inverted lookup table are evenly spaced.

2.3.2 Inverted lookup table (ILT) sampling method

Zijp et al. have proposed an inverted lookup table (ILT) sampling method where the forward lookup table (evenly spaced $\cos \theta_j$ values and corresponding CDF_j values) is transformed to an inverted lookup table by interpolation (cubic interpolation in our study) to obtain evenly spaced points of the CDF_j values and corresponding $\cos \theta_j$ values [17]. In this way, the index in the inverted lookup table can be computed directly from the drawn ξ without employing iterative root-finding methods, and independently of the lookup table size. Subsequently, a linear interpolation can be used to refine the value of the sampled $\cos \theta_\xi$ (Fig. 1(b)).

2.3.3 Non-linear lookup table (NLT) sampling method

We propose an alternative to the ILT sampling method. The non-linear lookup table is based on an approximate analytical model of the CDF that accounts for much of the information common to the phase functions of biological tissues (highly forward-peaked). The idea behind modeling the CDF is to transform the evenly spaced $\cos \theta j$ values from a forward lookup table using the approximate analytical model and interpolate the obtained values over the exact CDF to get a set of $\cos \theta j^*$ now representing the non-linear lookup table. By drawing a random number ξ, the approximate model instantly yields the index for the evenly spaced $\cos \theta j$. Since the transformation between the domains of $\cos \theta j$ and $\cos \theta j^*$ is bijective, a final linear interpolation between the two domains yields the sampled $\cos \theta_\xi$. The main benefit of the proposed methodology lies in the final linear interpolation that is accomplished between the domains of $\cos \theta_j$ and $\cos \theta_j^*$, which are linked almost linearly, especially if the approximate analytical model fits the CDF well. In contrast to the FLT and ILT sampling methods, where the linear interpolation is performed on the highly non-linear CDF, this should significantly improve the sampling accuracy with respect to the size of the lookup table.

In the following we describe the proposed procedure. Firstly, we compute the forward lookup table of evenly spaced $\cos \theta_j$ values and their corresponding CDF_j values. Next, we fit the selected approximate analytical model $f(\cos \theta)$ to the CDF (Fig. 2, red curve). It should be noted that the selected approximate CDF model f must be analytically invertible and well describe the variability of the employed phase function CDFs. For the purpose of this study, we have used the following model:

$$f(\cos\theta) = \frac{b}{\cos\theta - c} + a . \qquad (12)$$

Given that the model $f(\cos\theta)$ has to satisfy the following conditions to represent a CDF

$$f(-1) = 0 \text{ and } f(1) = 1, \qquad (13)$$

only one free parameter remains in the model f. The value of the free parameter can be found by non-linear least squares minimization of the difference between the model f and the CDF. Once f is known, a non-linear lookup table can be constructed. Each $\cos\theta_j$ value in the forward lookup table is transformed by the model f, see Fig. 2(b). The obtained values are close but not equal to CDF_j. By employing an inverse (cubic) interpolation with the aid of the forward lookup table for each $f(\cos\theta_j)$, the corresponding $\cos\theta_j^*$ can be obtained (red circles in Fig. 2). The connection between the values of $\cos\theta_j^*$ and $f(\cos\theta_j)$ now represents the non-linear lookup table of the CDF.

When a random number ξ is drawn (Fig. 2(b) green diagram), it is first transformed by $f^{-1}(\xi)$ into the domain of evenly spaced $\cos\theta_j$, which instantly yields the index into the lookup table of $\cos\theta_j^*$ values. The $\cos\theta_\xi$ can then be obtained by a final linear interpolation, similar to the FLT and ILT sampling methods. There is one important distinction, however. The NLT sampling method does the final linear interpolation between the evenly spaced $\cos\theta_j$ and $\cos\theta_j^*$ values. The bijection between the two is an identity function, if the model f ideally fits the CDF, which results in an error-free linear interpolation. If the fits are as presented in Fig. 2, the bijection is close to identity function and the final linear interpolation should still work sufficiently well even with small lookup table sizes.
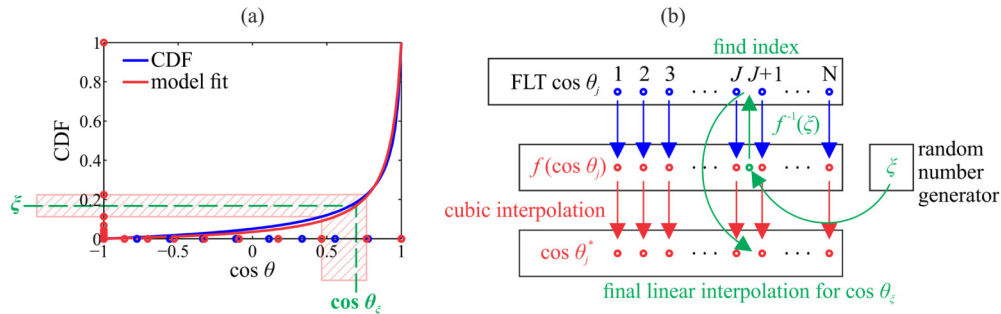

Fig. 2. (a) An example of CDF for the HG phase function. The non-linear lookup table ($\cos\theta_j^*$ values and their corresponding $\mathrm{CDF}_j = f(\cos\theta_j)$ values) is presented with red circles. The evenly spaced points of $\cos\theta_j$ are presented with blue circles. Drawn random number ξ and the sampled $\cos\theta_\xi$ are given in green. (b) Construction of a non-linear lookup table from evenly spaced $\cos\theta_j$. The non-linear lookup table is presented in red. Green diagram highlights the sampling of $\cos\theta_\xi$ from a random number ξ.

3. Materials and methods

3.1 Monte Carlo (MC) simulations

In this study, an MC model (MCML) developed by Wang *et al*. [2] and implemented for CUDA by Alerstam *et al*. [22] was used. The code was modified to account for the source-detector optical fiber probe geometry and to study the reflectance as a function of the source-detector separation (SDS) and numerical phase function sampling. The described setting was adopted since optical fibers probes are frequently used to collect the backscattered light from a turbid medium [23]. In the MC simulations, the optical fiber geometry was treated as a laterally uniform probe-tissue interface, which takes into account only the mismatch between the refractive indices of the turbid medium and optical fibers. The refractive index of the turbid medium was set to $n_m = 1.33$, while the refractive index of the optical fibers was set to

n_{fib} = 1.45. The diameter of optical fibers was set to 200 µm and SDS values of 220, 440, 660, 880, and 1100 µm were considered. The selected range of the SDS should encompass the transition of the reflectance from the sub-diffusive to the diffusion regime [7]. The photon packets were launched uniformly over the source fiber opening and uniformly over the solid angle defined by the numerical aperture of the fiber (NA = 0.22). In order to take advantage of the entire area around the source fibers and thereby reduce the simulation noise, the detection scheme around the source fiber was divided into 5 µm concentric rings. The total weight of the detected photon packets at a particular SDS was then calculated in the post-processing step. Only the photon packets with the output direction within the acceptance cone of the detection fiber were considered (NA = 0.22).

The MC simulations were conducted under the semi-infinite turbid medium geometry. In the simulations, all of the photon packets that drifted more than 0.8 cm laterally or axially from the central point of the source fiber, were terminated. This condition sped up the simulations considerably and was found consistent with the semi-infinite medium geometry.

3.2 Phase functions utilized for evaluating the numerical sampling methods

The most common observables used in the biomedical optics that arise from the phase function are the anisotropy factor g and the parameter γ. The latter is defined as $\gamma = (1-g_2)/(1-g)$, where g_2 is the second Legendre moment of the phase function. We considered several different phase functions to evaluate the three numerical sampling methods across the biological variations of g and γ. The anisotropy factor g can be found to be as low as 0.4 [24] and can reach values of up to 0.98 [25] in the biological tissues. With respect to parameter γ, the values tend to lie between 1.1 and 2.3 [6,26,27]. As a result, we have utilized 15 phase functions listed in Table 1.

Table 1. Phase functions, the corresponding parameters and the resulting observables g and γ used in this study.

Phase Function	Phase Function Parameters		g	γ
GK	$\alpha_{GK} = 0.5$	$g_{GK} = 0.4$	0.4	1.4
	$\alpha_{GK} = 0.82$	$g_{GK} = 0.94$	0.98	2.3
	$\alpha_{GK} = -0.13$	$g_{GK} = 0.93$	0.6	1.4
	$\alpha_{GK} = 5.8$	$g_{GK} = 0.27$	0.8	2.3
MHG	$\beta = 0.90$	$g_{HG} = 0.77$	0.7	1.4
	$\beta = 0.99$	$g_{HG} = 0.98$	0.98	1.9
	$\beta = 0.97$	$g_{HG} = 0.72$	0.7	1.6
	$\beta = 0.96$	$g_{HG} = 0.94$	0.9	1.4
Mie	$n_{med} = 1.33$ $n_{par} = 1.59$ $\lambda = 633$ nm	$d = 0.1$ µm	0.075	0.97
		$d = 1$ µm	0.92	2.2
		$d = 10$ µm	0.90	1.6
		$d = 20$ µm	0.91	1.6
Mie fractal	$n_{med} = 1.352$ $n_{par} = 1.42$ $\lambda = 633$ nm $d_i = [5\text{nm} – 30$ µm$]$ (5 nm steps)	$\alpha = 4$	0.96	2.1
		$\alpha = 4.5$	0.91	1.8
		$\alpha = 5$	0.80	1.5

4. Results

4.1 Influence of the numerical phase function sampling on the Monte Carlo (MC) simulated reflectance

In order to study how the different numerical sampling methods and their corresponding lookup table sizes affect the MC modeling of light propagation in turbid media, we have simulated reflectance maps as a function of the absorption and reduced scattering coefficient at several SDS. For this particular case, a Mie phase function for 10 µm polystyrene microspheres listed in Table 1 was used. Figure 3 presents the relative reflectance error (RRE) maps across different numerical sampling methods for an undersized lookup table

comprising only 100 entries. The RRE maps were obtained by subtracting the ground truth reflectance map from the reflectance map computed by the undersized lookup table and subsequently dividing the difference by the ground truth reflectance map. Since analytical sampling from the Mie phase function is not possible, the ground truth reflectance map was obtained by the NLT sampling method utilizing an oversized lookup table of 5000 entries. Figure 3 shows that the distribution and extent of the RRE significantly depends on the numerical sampling method and also on the SDS. The FLT sampling method seems to perform well at shorter SDS, while the ILT sampling method performs well at longer SDS. The NLT sampling method seems to retain the smallest RRE across all the SDS. For a small lookup table of 100 entries, the RRE can reach up to ± 15%.
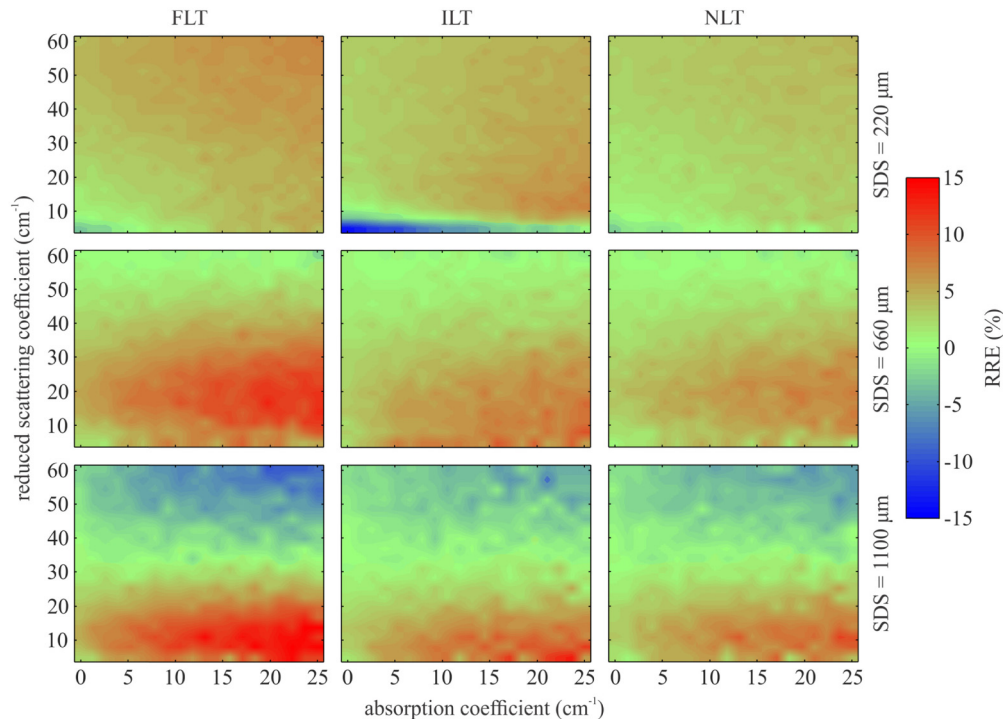

Fig. 3. RRE maps as a function of the absorption and reduced scattering coefficient at three SDS arising from an undersized lookup table employed by the numerical sampling methods (FLT, ILT and NLT).
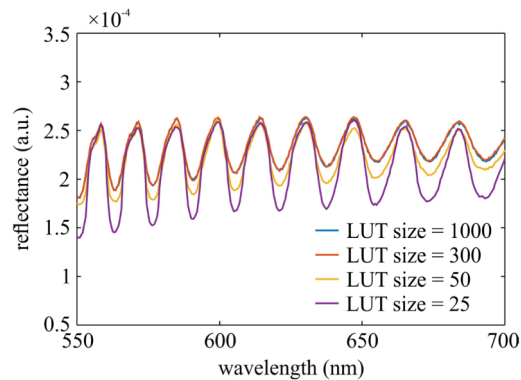

Fig. 4. MC simulated reflectance spectrum of a turbid phantom comprising 5 μm polystyrene spheres computed by varying the lookup table (LUT) size. The phase function was sampled using the ILT sampling method.

Figure 4 shows an example of MC simulated reflectance spectrum from 550 to 700 nm at SDS = 220 μm for a turbid phantom containing 5 μm monodisperse polystyrene microspheres with a wavelength independent absorption coefficient set to 0.1 cm^{-1}. The concentration of the polystyrene microspheres was selected so that the reduced scattering coefficient at 633 nm amounted to 7.61 cm^{-1}. Sampling from the Mie phase function was accomplished with the ILT sampling method. The lobes in the spectrum are a consequence of the wavelength dependency of the Mie phase function of monodisperse spheres. In practice, the lobes are usually somewhat reduced due to the particle size variations. The effect of the lookup table size on the MC simulated reflectance spectrum is apparent. In this particular case, an undersized lookup table leads to an underestimated reflectance.

4.2 Dependence of the relative reflectance error (RRE) on the lookup table size

As shown in Section 4.1, the choice of the numerical sampling method and the underlying lookup table size significantly affect the accuracy of the MC simulations. Consequently, we have further investigated the performance of the numerical sampling methods and the influence of the lookup table size on the accuracy of the MC simulated reflectance. For this purpose, 5 x 5 reflectance maps as a function of the absorption and reduced scattering coefficient were computed for five SDS and all the phase functions defined in Table 1. The absorption coefficient spanned values from 0 to 25 cm^{-1} and the reduced scattering coefficient from 5 to 60 cm^{-1}. The reflectance maps were simulated for each numerical sampling method by progressively increasing the lookup table size from 25 in steps of 25 until the maximum absolute RRE within the map fell under 2%. The ground truth reflectance map was obtained as described in Section 4.1, using the NLT sampling method with a lookup table size of 5000.

Table 2 summarizes the lookup table sizes of the numerical sampling methods that were required to reduce the maximum absolute RRE under 2% at all SDS. It can be observed that the required lookup table size varies dramatically across the employed phase functions. Large lookup tables are required for phase functions with a high anisotropy factor g, in particular for values exceeding 0.9.

Table 2. Minimum lookup table sizes of the numerical sampling methods (FLT, ILT and NLT) that guarantee maximum absolute RRE under 2% for all the utilized phase functions at all SDS.

Phase Function	g	γ	FLT	ILT	NLT
	0.4	1.4	25	50	25
GK	0.98	2.3	850	1450	350
	0.6	1.4	50	50	25
	0.8	2.3	50	475	50
	0.7	1.4	50	75	50
MHG	0.98	1.9	1475	750	800
	0.7	1.6	50	25	25
	0.9	1.4	300	150	175
	0.075	0.97	25	50	25
Mie	0.92	2.2	125	175	125
	0.90	1.7	450	350	250
	0.91	1.6	500	300	275
	0.96	2.1	500	250	300
Mie fractal	0.91	1.8	225	150	125
	0.80	1.5	100	75	75

The performance of the numerical sampling methods was further examined by employing the GK phase function with $g = 0.98$, for which the largest variation of the lookup table size was observed. Figure 5 presents the maximum absolute RRE as a function of the lookup table size for the FLT, ILT, and NLT numerical sampling methods. In addition, the maximum absolute RRE are calculated for each SDS (top row). The corresponding reconstructed GK phase function based on the lookup table size listed in Table 2 is presented in the middle row (blue line). The reconstruction was accomplished by uniformly drawing 10^8 random numbers

from the interval [0,1] and then binning the computed cos θ in a histogram with 1000 bins. In order to overcome the influence of discretization and directly compare the reconstructed and the corresponding true phase function (orange line), the average value of the true phase function was computed within each bin. The relative error of the reconstructed phase function is shown in Fig. 5 (bottom row).

Two important observations can be made from Fig. 5. Firstly, the maximum absolute RRE is notably dependent on the SDS regardless of the numerical sampling method. The FLT sampling method at a given lookup table size exhibits lower maximum absolute RRE for shorter SDS. In contrast, the ILT sampling method performs better at longer SDS. Unlike the FLT and ILT sampling methods, the NLT sampling method seems to perform well across all the SDS.
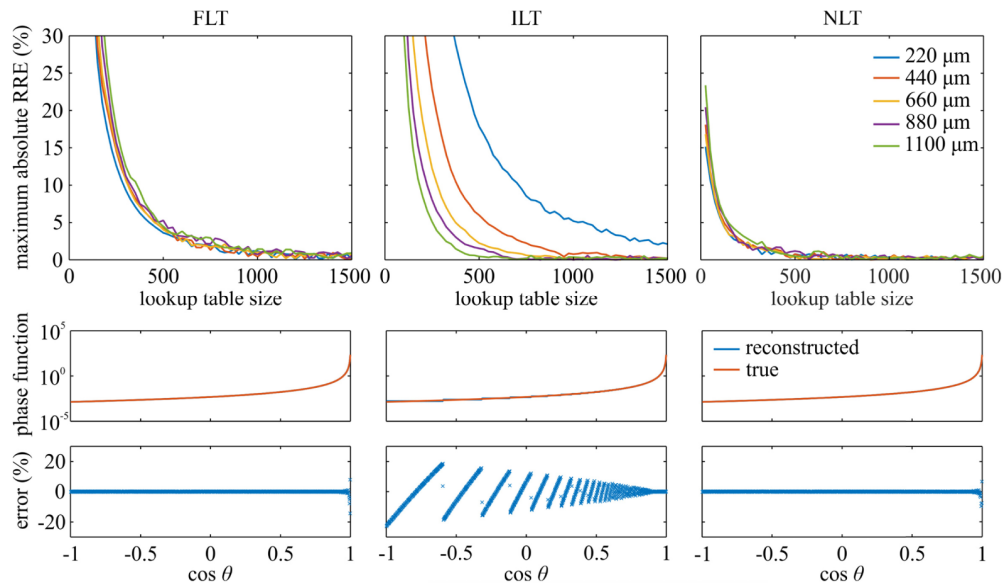

Fig. 5. (top row) Maximum absolute RRE as a function of the lookup table size used by the numerical sampling methods (FLT, ILT and NLT). (middle row) Comparison between the reconstructed and the corresponding true GK phase function obtained by different numerical sampling methods using the lookup table sizes from Table 2. (bottom row) The relative errors of the reconstructed GK phase functions.

Secondly, the maximum absolute RRE can be under 2% even though the reconstructed phase function exhibits relative errors of around ± 20%. This observation suggests that the phase functions do not have to be accurately modeled in the MC simulations, especially at large scattering angles, to yield a satisfactory reflectance prediction.
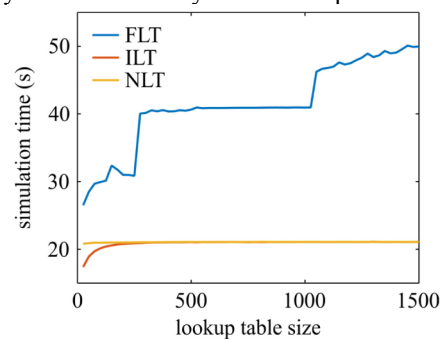

Fig. 6. The average simulation time required to compute a 5 x 5 reflectance map as a function of the lookup table size for each numerical sampling method (FLT, ILT and NLT).

In terms of the simulation time, we compared the numerical sampling methods (FLT, ILT and NLT) for the GK phase function studied above. Figure 6 presents the average simulation time required to compute a 5 x 5 reflectance map normalized by the number of points in the map. Simulations were performed on a personal computer (operating system: Microsoft Windows 7 Enterprise, central processing unit: Intel® Core i7-4770 @ 3.40 Ghz, memory: 16.00 GB RAM, graphics processing unit: Nvidia GeForce GTX 770). A twofold increase in the simulation time is observed for the FLT in comparison to the ILT and NLT sampling methods for a lookup table size larger than 250. For the ILT and NLT sampling methods, the difference is negligible for a lookup table size larger than 250, however, below this point the ILT sampling method seems to perform better.

4.3 Dependence of the required lookup table size on the phase function anisotropy factor

The most prominent dependence of the lookup table size required to keep the maximum absolute RRE under 2% is on the anisotropy factor of the phase function. This is further confirmed by Fig. 7, which shows the RRE maps as a function of the anisotropy factor and reduced scattering coefficient for a MHG phase function with a constant γ of 1.8. The absorption coefficient was set to 2 cm^{-1} and the lookup table size used in all of the numerical sampling methods was set to 200. As the anisotropy factor increases, the RRE rises significantly and can exceed 50% for the FLT method, 25% for the ILT method and 10% for the NLT method. This means that a higher anisotropy factor requires a larger lookup table.
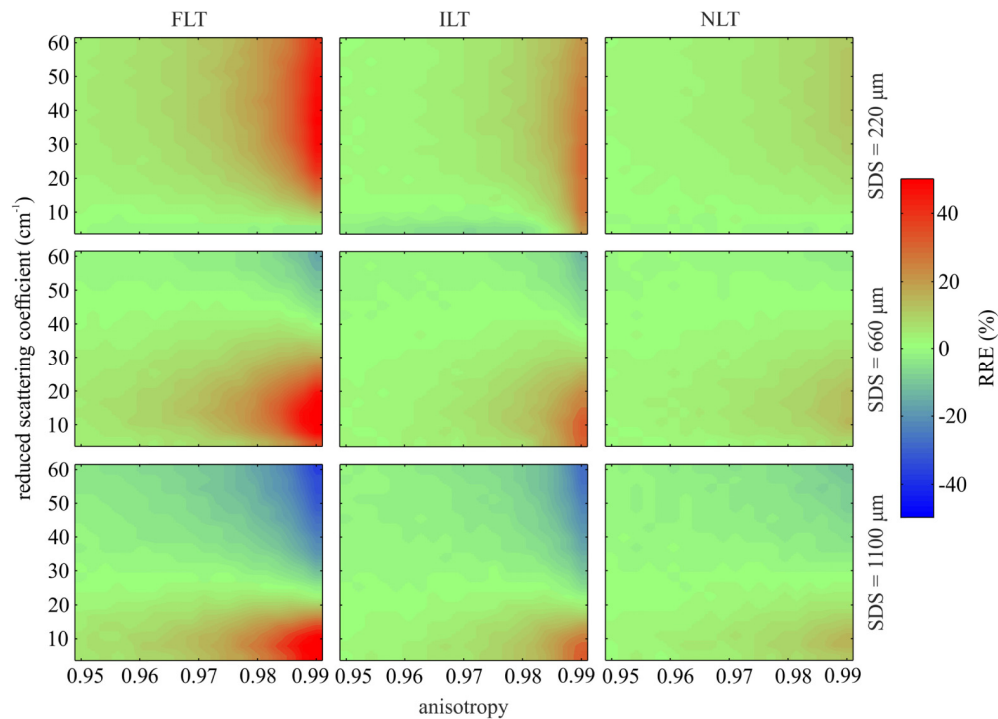

Fig. 7. Relative reflectance error (RRE) maps as a function of the anisotropy factor and reduced scattering coefficient at three SDS arising from an undersized lookup table employed by the numerical sampling methods (FLT, ILT and NLT).

Figure 8 depicts the dependence of the required lookup table size at different SDS obtained for the three numerical sampling methods. It can be clearly observed that regardless of the SDS, the lookup table size significantly increases for anisotropy factors *g* beyond 0.8. Below this value, the required lookup table size is under 100 for all the numerical sampling

methods and all the SDS. We can also observe the difference between the FLT and ILT sampling methods. In particular, for short SDS the largest lookup table is required by the ILT sampling method, while for longer SDS the largest lookup table is clearly required by the FLT sampling method.
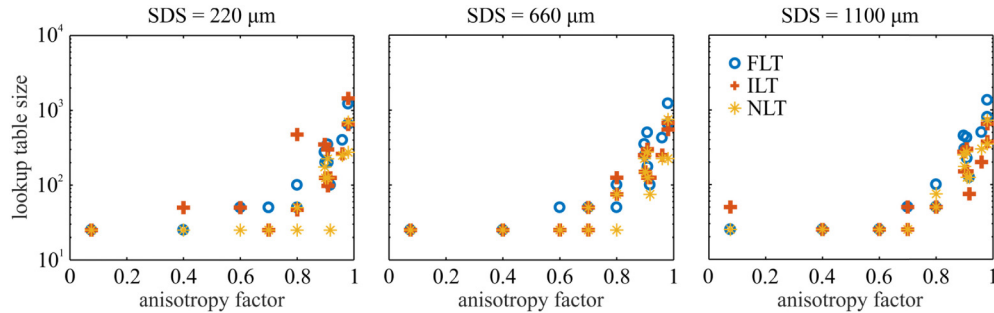

Fig. 8. Minimum lookup table size required to keep the maximum absolute RRE computed by different numerical sampling methods (FLT, ILT and NLT) under 2% as a function of the anisotropy factor for three source-detector separations (SDS).

5. Discussion

Proper sampling of the scattering angle from a phase function is important when simulating light propagation in turbid media. For example, any errors arising from sampling of the scattering angle required at each scattering event in the MC simulations can directly influence the observable quantities such as reflectance. As shown in Figs. 3 and 4, the RRE significantly depends on the numerical sampling method and underlying lookup table size. This suggests that special care should be taken when implementing the phase functions in the MC simulations, and moreover, include the information in research articles as this was not a common practice to date [4–7,10,11].

Especially concerning is the information presented in Fig. 4, since turbid phantoms containing polystyrene microspheres are often used to calibrate MC simulated reflectance acquired with optical fiber probes [28,29]. As such, implementing proper sampling of the scattering angle from the phase function is also crucial for a good calibration that enables comparison between the MC simulations and measurements.

Selection of an optimal lookup table size is especially important when performing the MC simulations on a GPU, where the amount of fast random access memory is limited. Moreover, constructing a large lookup table for phase functions such as the Mie phase function can be computationally intensive. This is especially true for the fractal Mie phase function that requires computation of many Mie phase functions corresponding to monodisperse particles that are integrated over the full range of the employed particle number density.

We have investigated the performance of the numerical sampling methods in the MC simulations for various phase functions with different anisotropy factor g and parameter γ (Table 2). We have observed that the three numerical sampling methods require different sizes of the lookup tables for different phase functions. Furthermore, the results for the GK phase function in Fig. 5 clearly show that the performance of the numerical sampling methods also significantly depends on the SDS. The explanation for this observation can be directly related to the linear interpolation performed during the extraction of $\cos \theta$ from the lookup table (Figs. 1 and 2). The FLT sampling method has evenly spaced $\cos \theta$ values in the lookup table, which implies that for a drawn random number, the CDF will be interpolated equally well for large and small scattering angles. In contrast, the ILT sampling method has evenly spaced CDF values and unevenly spaced corresponding $\cos \theta$ values and thus interpolates the CDF better for small scattering angles. Due to the characteristic shape of the CDF for forward peaked phase functions, the uniformly distributed random numbers evaluate to unevenly

spaced $\cos\theta$ values, where the density of points increases towards the smaller scattering angles. However, according to Bodenschatz *et al.* [7], at short SDS, photons that scatter at larger angles significantly contribute to the reflectance signal, thus making the corresponding part of the phase function important. Consequently, the FLT sampling method is observed to perform better for shorter SDS. This fact is consistent with the observation in the bottom row of Fig. 5. In comparison to the FLT and NLT sampling methods, the ILT sampling method performs significantly worse at larger scattering angles where the density of $\cos\theta$ values in the lookup table drops substantially. Consequently, the ILT sampling method has higher maximum absolute RRE at shorter SDS. At longer SDS, however, the probability of large angle scattering events drops and, due to the forward peaked shape of the phase function, the sampled scattering angles are mostly small. In this case, the ILT sampling method is expected to perform better. Since this cannot be observed from the bottom row of Fig. 5, we have calculated the average relative error within 50 bins of the reconstructed phase function histograms from Fig. 5 that represent the contribution of small scattering angles ($\cos\theta$ between 0.9 and 1). The obtained relative error as a function of the lookup table size is shown in Fig. 9. The average relative error approaches zero faster for the ILT than the FLT sampling method. It should be noted that the best results for all of the SDS shown in Fig. 5 and Fig. 9 are obtained with the NLT sampling method. This is because the linear interpolation in the NLT sampling is accomplished between the indices that correspond to evenly spaced $\cos\theta$ and close to evenly spaced $\cos\theta^*$ (see Sect. 2.3.3 and Fig. 2). The error of linear interpolation depends solely on the quality of the fit between the employed model f and the CDF. For a good fit, the relation between the evenly spaced $\cos\theta$ and $\cos\theta^*$ is close to identity, yielding accurate sampling results for both small and large scattering angles.
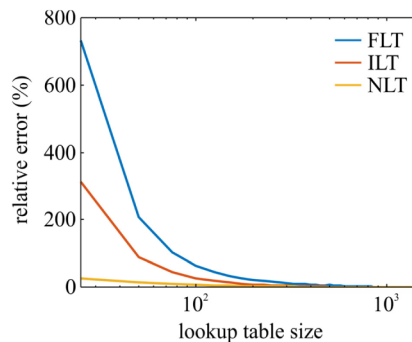

Fig. 9. The average relative error within 50 bins of the reconstructed phase function histograms from Fig. 5 that represent the contribution of small scattering angles ($\cos\theta$ between 0.9 and 1) as a function of the lookup table size for each numerical sampling method (FLT, ILT and NLT).

In Fig. 5 we have made another remarkable observation that high relative errors in the reconstructed phase functions do not necessarily lead to high RRE. This suggests that the phase functions can be implemented in the MC simulations with a low degree of accuracy and still yield satisfactory reflectance estimates. Especially notable is the example for the ILT sampling method that exhibits a $\pm$ 20% error of the reconstructed phase function (large scattering angles) while the simulated reflectance at the shortest SDS still stays within 2% of the ground truth. This observation explains why simple phase functions such as HG can be successfully used in many experimental configurations, for which the phase function has a limited influence on the simulated reflectance. As a consequence, establishing a measure for the accuracy of reflectance simulation should be based on the reflectance itself (as performed in Fig. 5) and not on the relative errors in the reconstructed phase function. Using the latter as a measure could lead to oversized lookup tables, occupying large chunks of memory and requiring long preparation times.

To gain some insight on the computation efficiency of the numerical sampling methods, we have compared the simulation times required to complete a reflectance map. As expected, the simulation time required by the FLT sampling method is the longest and governed by the lookup table size. Since the FLT sampling method does not provide benefits in terms of lookup table size (Table 2), it should be the last choice in terms of MC simulations. The ILT and NLT sampling methods circumvent the use of costly iterative root-finding algorithm through the lookup table, and exhibit generally similar simulation times. However, for lookup tables of size below 250, the ILT sampling method seems to be the fastest. We believe that this is due to the additional computations of the inverse model $f^{-1}(\xi)$ performed by the NLT sampling method. For small lookup tables, the additional computation time becomes dominant over the memory access latency. However, for practical applications where the lookup table size usually exceeds 250 entries, the simulation times for ILT and NLT sampling methods converge.

Finally, the anisotropy factor was proven to be the key observable of the phase functions that determines the lookup table size required for the maximum absolute RRE to drop under 2%. As expected, phase functions with higher anisotropy factor g (i.e., highly forward peaked) require larger lookup tables. For highly forward peaked phase functions, the CDF of the phase function is steep, thus making the linear interpolation in the numerical sampling methods increasingly difficult. Consequently, a large number of $\cos\theta$ values and the corresponding CDF values is required to retain a certain level of accuracy. For phase functions with anisotropy factor below 0.8, a lookup table size of approximately 100 is sufficient. In accordance with the computational efficiency of the numerical sampling methods, our recommendation is to use the ILT sampling method for phase functions with anisotropy factors below 0.8. For phase functions with anisotropy factor above 0.8, we believe that both ILT and NLT sampling methods should work sufficiently well. However, for short SDS, the NLT sampling method exhibits significantly smaller lookup table sizes (in this study shown for the GK phase function) and thus might be preferred over the ILT sampling method.

6. Conclusion

In this study we showed that numerical sampling from the phase function can significantly affect the MC simulated observables such as reflectance. Consequently, the resulting error of simulated reflectance can affect the calibration of MC simulations and thereby comparison with measured reflectance. By direct comparison of the FLT, ILT and NLT sampling methods, we found that their performance significantly depends on the phase function shape (anisotropy factor) and on the SDS when simulating reflectance acquired with optical fiber probes. Phase functions with a higher anisotropy factors require larger lookup tables. In terms of the SDS, at a common lookup table size, the FLT sampling method performs well at shorter SDS, while the ILT sampling method performs well at longer SDS. The NLT sampling method seems to perform sufficiently well at all the studied SDS. Remarkably, a suitable lookup table size should be based on the quantity that is being simulated (e.g., reflectance) and not on the errors that are present in the phase function reconstructed from the lookup table. Since the phase function is sampled many times due to the multiple scattering events, the errors in the reconstructed phase function are generally not a good measure for the quality of the simulated quantity (e.g., reflectance).

Our recommendations on numerical sampling from the phase function in MC simulations of light propagation in turbid media are the following. Firstly, a suitable lookup table size should be determined through a comparison of the simulated observable (e.g., reflectance) to the ground truth. As a first approximation for the suitable lookup table size, the anisotropy factor of the phase function should be considered. Afterwards, the preferred numerical sampling method should be selected according to the simulation time. If the lookup table size is less than approximately 250, the ILT sampling method should be used. For larger lookup

tables, both the ILT and NLT sampling methods are acceptable, although for reflectance acquired close to the source (small SDS), the NLT sampling method was shown to sample the larger scattering angles more accurately than the ILT sampling method.

Funding