

“Virtual probe design”: Monte Carlo simulation in the design of diagnostic instrumentation

Chih-wen Kan¹, Kort Travis², Daniel Côté³, Konstantin Sokolov^{1,4},
Mia K. Markey^{1,4}

¹ The University of Texas Department of Biomedical Engineering, 1 University Station, Austin, TX, 78712, USA

² Department of Physics, The University of Texas at Austin, 1 University Station, Austin, TX, 78712, USA

³ Université Laval, Département de Physique, Centre de Recherche Université Laval Robert Giffard, 2601 Ch. de la Canardière, F-6500, Québec, Québec, Canada, G1J 2G3

⁴ Department of Imaging Physics, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030

mia.markey@mail.utexas.edu

Abstract: not yet

© 2010 Optical Society of America

OCIS codes: (000.0000) General.

References and links

1. R. Richards-Kortum and E. Sevick-Muraca, “Quantitative optical spectroscopy for tissue diagnosis,” *Annual Review of Physical Chemistry* **47**, 555–606 (1996).
2. K. Sokolov, M. Follen, and R. Richards-Kortum, “Optical spectroscopy for detection of neoplasia,” *Curr. Opin. Chem. Biol* **6**, 651–658 (2002).
3. K. Sokolov, R. Drezek, K. Gossage, and R. Richards-Kortum, “Reflectance spectroscopy with polarized light: is it sensitive to cellular and nuclear morphology,” *Optics Express* **5**, 302–317 (1999).
4. K. Sokolov, L. Nieman, A. Myakov, and A. Gillenwater, “Polarized reflectance spectroscopy for pre-cancer detection,” *Technol Cancer Res Treat* **3**, 1–14 (2004).
5. M. Antonelli, A. Pierangelo, T. Novikova, P. Validire, A. Benali, B. Gayet, and A. De Martino, “Mueller matrix imaging of human colon tissue for cancer diagnostics: how monte carlo modeling can help in the interpretation of experimental data,” *Optics Express* **18**, 10200–10208 (2010).
6. X. Wang, L. V. Wang, C.-W. Sun, and C.-C. Yang, “Polarized light propagation through scattering media: time-resolved monte carlo simulations and experiments,” *Journal of Biomedical Optics* **8**, 608–617 (2003).
7. S. Jacques, J. Roman, and K. Lee, “Imaging superficial tissues with polarized light,” *Lasers in Surgery and Medicine* **26**, 119–129 (2000).
8. J. Mourant, A. Hielscher, A. Eick, T. Johnson, and J. Freyer, “Evidence of intrinsic differences in the light scattering properties of tumorigenic and nontumorigenic cells,” *Cancer* **84**, 366–374 (1998).
9. G. Zonios, L. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M. Feld, “Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo,” *Appl. Opt* **38**, 6628–6637 (1999).
10. A. Hielscher, J. Mourant, and I. Bigio, “Influence of particle size and concentration on the diffuse backscattering of polarized light from tissue phantoms and biological cell suspensions,” *Applied optics* **36**, 125–135 (1997).
11. X. Wang, G. Yao, and L. Wang, “Monte carlo model and single-scattering approximation of the propagation of polarized light in turbid media containing glucose,” *Applied optics* **41**, 792–801 (2002).
12. S. Chandrasekhar, *Radiative Transfer* (Courier Dover Publications, 1960).
13. H. van de Hulst, *Light Scattering by Small Particles* (1957).
14. A. Welch, G. Yoon, and M. van Gemert, “Practical models for light distribution in laser-irradiated tissue,” *Lasers Surg Med* **6**, 488–93 (1987).
15. M. Patterson, B. Wilson, and D. Wyman, “The propagation of optical radiation in tissue i. models of radiation transport and their application,” *Lasers in Medical Science* **6**, 155–168 (1991).
16. L. Wang, S. Jacques, and L. Zheng, “Mcml monte carlo modeling of light transport in multi-layered tissues,” *Computer Methods and Programs in Biomedicine* **47**, 131–146 (1995).

17. G. Yao and L. Wang, "Propagation of polarized light in turbid media: simulated animation sequences," *Optics Express* **7**, 198–203 (2000).
 18. Y. Liu, Y. Kim, X. Li, and V. Backman, "Investigation of depth selectivity of polarization gating for tissue characterization," *Opt. Express* **13**, 601–611 (2005).
 19. D. Côté and I. Vitkin, "Robust concentration determination of optically active molecules in turbid media with validated three-dimensional polarization sensitive monte carlo calculations," *Optics Express* **13**, 148–163 (2005).
 20. F. Jaillon and H. Saint-Jalmes, "Description and time reduction of a monte carlo code to simulate propagation of polarized light through scattering media," *Appl. Opt.* **42**, 3290–3296 (2003).
 21. S. Bartel and A. H. Hielscher, "Monte carlo simulations of the diffuse backscattering mueller matrix for highly scattering media," *Appl. Opt.* **39**, 1580–1588 (2000).
 22. B. Kaplan, G. Ledanois, and B. Villon, "Mueller matrix of dense polystyrene latex sphere suspensions: Measurements and monte carlo simulation," *Appl. Opt.* **40**, 2769–2777 (2001).
 23. X. Guo, M. Wood, and A. Vitkin, "Monte carlo study of pathlength distribution of polarized light in turbid media," *Optics Express* **15**, 1348–1360 (2007).
 24. L. Nieman, M. Jakovljevic, and K. Sokolov, "Compact beveled fiber optic probe design for enhanced depth discrimination in epithelial tissues," *Opt. Express* **17**, 2780–2796 (2009).
 25. C. Bohren and D. Huffman, *Absorption and scattering of light by small particles* (Wiley New York, 1983).
 26. T. Germer, "Scatmech: polarized light scattering c++ class library," .
 27. A. Welch, M. Gemert *et al.*, *Optical-Thermal Response of Laser-Irradiated Tissue* (Springer, 1995).
-

1. Introduction

Probing biological tissue, a scattering medium, with light, has many potential applications in biomedical diagnostics. Biophotonics methods are attractive, because in general they are non-invasive, and this is particularly important for the diagnosis of cancer [1, 2]. Cancerous cells and normal cells have very different morphologies, largely because cancerous cells have fast, generally uncontrolled growth processes. These morphological differences also affect the way these cells interact with light; for example, it was found that the difference (e.g., in scattering/absorption coefficients) between the optical properties of normal and cancerous tissues results in a detectable difference in their reflectance spectra. For these reasons, biophotonics methods provide the promise of early diagnosis of cancer in a relatively inexpensive, and non-invasive fashion.

Polarized light sources, and polarization-sensitive detectors, have been used to provide better separation between shallow and deep signals collected from tissue. This is highly advantageous, as cancer generally originates near the surface layers of organs, and therefore, statistics with more diagnostic relevance will be obtained from these enhanced superficial signals. For example, polarized incident light and polarization-enabled collection fibers have been widely used to probe the structure of superficial tissue. This superficial signal enhancement is due largely to the fact that a significant difference between singly and multiply scattered light is that the latter does not preserve the incident polarization. It has been shown both experimentally and in simulation that polarized light is useful in studying tissue, and that tissue properties can be extracted through changes in the light's polarization state [3, 4, 5, 6, 7, 8, 9, 10, 11].

Radiative transfer theory (RTT) and the associated radiative transfer equation (RTE) [12, 13], is a general theory dealing with the transport of energy through physical media, such as scattering media. The RTT has been the most successful technique applied to the modelling of light propagation in tissue. However, due to its inherent complexity, various approximations have been proposed to simplify the RTE. These approximations concern constraints on the distribution of the sizes of the scatters, constraints on whether scattering or absorption dominates the optical properties, and constraints on whether polarization is or is not included in the model [14, 15]. For example, the diffusion approximation is a low order analytical solution to the RTE, which most commonly assumes multiple scattering events in the tissue and neglects any polarization of the propagating light. Such approximations however, are not always appropriate; and therefore, a numerical solution to the RTE is desirable. The Monte Carlo technique provides a

descriptive and rigorous numerical solution to the RTE. In general, photon Monte Carlo methods are a type of stochastic integration technique, which solve the RTE through the use of a 3D random walk of energy packets, or “photons”, through the tissue volume. Although these methods provide accurate solutions, the drawback of Monte Carlo methods is that an extremely large number of photons must be simulated in order to produce statistically reliable results. Monte Carlo Multi Layer (MCML), implemented by Wang *et al.* [16], is one of the most commonly known implementations of Monte Carlo. Although MCML does not consider polarization, it is considered the standard of the field because it has been tested and validated extensively. More recent Monte Carlo implementations do incorporate polarization [6, 17, 18, 19, 20, 21, 22, 23], although many of these implementations are limited in other respects. For example, most polarized Monte Carlo implementations offer only very limited geometric capability; for example, they do not allow modelling of multiple layers, or of complex objects or of more physically realistic scattering descriptions [6, 17, 18, 11, 19]. For these reasons, only a very few contemporary Monte Carlo implementations are practically useful for instrumentation design and validation.

When light is used for diagnosis, it is important to ensure that it travels through target areas where cancer and pre-cancer are located. Liu *et al.* [18] investigated the depth selectivity of polarization gating for tissue characterization, and concluded that polarization-gated signal is generated primarily by photons emerging from the surface of the medium within a few mean free path lengths from the point of incidence. This study, however, was done considering only one layer of tissue. Guo *et al.* [23] studied the pathlength distribution of polarized light in turbid media, and concluded that the average pathlength of polarization-preserving photons was two to three times smaller than the average pathlength of all collected photons (i.e., where collected photons include all polarizations).

Several fundamental questions must be addressed in order to evaluate the effectiveness of the use of polarization for diagnostic purposes. What are the specific signals collected by parallel and perpendicular polarizers? Is the penetration depth of the light significantly different for multiple layer models? What is the relationship between penetration depth and pathlength? Can signals be discriminated through the use of singly scattered and multiply scattered light? And most importantly, is the instrument actually effectively collecting the signal from the target layer, which is the precancer layer? Answers to these questions will be critical to guide the development of new diagnostic instrumentation, and in addition to validate existing instrumentation designs. Particularly important are considerations of how probe geometry affects the collection of light. For example, in a previous study by our group, a probe was developed to improve the depth resolution by using beveled collection fibers [24], and it would be beneficial to be able to use simulation to validate the signal-to-noise ratio achievable through the use of this type of angled-fiber design in comparison to simpler probe designs using only a flat collection fibers.

In this study, we use a polarization-sensitive Monte Carlo simulation (Pol-MC) to aid in the design of instrumentation for the early detection of epithelial cancer. In contrast to other Monte Carlo implementations, which primarily use the Henyey-Greenstein phase function, the Pol-MC numerical code accurately calculates scattering phase functions based on comprehensive Mie theory calculations, and fully retains polarization information at both scattering and interface transitions. In addition to these advanced features, we have extended the Pol-MC capability to include full multilayer capability, as well as the capability to use scattering mixture-models, which enable a physically realistic treatment for complex scattering media such as biological tissue.

By analysis of simulation results, we compared the use of angled and flat detector geometries in a hypothetical diagnostic instrument. Further, by taking advantage of the stochastic nature of

the Monte Carlo technique itself, we are able to evaluate information not normally accessible through any physical measurement; for example, we are able to evaluate the maximum possible signal-to-background ratio achievable by the instrument through the sub-classification of scattered photon paths, as represented in the simulation, and this information is also applicable to assist in the development of algorithms for the analysis of the instrument's measurements. These promising results provide a proof-of-concept for the application of this type of stochastic computational technique to diagnostic instrument design.

2. Methods

2.1. Polarization sensitive Monte Carlo simulations (Pol-MC)

The numerical simulation of polarization sensitive Monte Carlo is performed using a program, Pol-MC, developed by Dr. Daniel Côté [19]. Several implementation extensions, including multiple layers, geometry checks, detector implementations, and analysis tools, are developed by the authors of this study.

2.1.1. Geometry

The simulation geometry to Pol-MC is specified by a user-supplied input file, in XML format, which may be quite general, and defines the relative positions and optical properties of a collection of geometric objects, such as boxes, infinite layers, and cylinders, at specified locations, along with a number of possible photon source types. In addition to geometric attributes, several optical attributes characterize each object: such as the size of scatters, refractive index for the scatter and its surrounding medium, and absorption coefficient μ_a and scattering coefficient μ_s . Boundaries of geometric objects are represented through the use of primitive surface elements, which then collectively specify the complete geometry. For example, a box has six surface elements, an infinite layer has two, and significantly more complicated “generic” objects are possible which can be represented through the import of any suitable surface mesh. In addition to its location, each surface element retains a set of local basis vectors consisting of two unit vectors (**a** and **b**) in the plane of the surface element and a unit normal vector(\hat{n}), pointing towards the exterior of the object for which the surface element forms part of the boundary. The check for geometric consistency using these surface basis vectors is a critical part of our multilayer implementation, which is further discussed in Section 2.2.2.

2.1.2. Photon propagation

Once the simulation geometry and the illumination sources are fully specified, multiple photons propagating through the tissue geometry are simulated. At each scattering event, interface transition, and or collection or termination event, the program accumulates statistics on current properties of the associated photon. The “photon” structure itself holds the current position, propagation direction, and Stoke's parameters of the form [25]:

$$S = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} E_{\parallel}E_{\parallel}^* + E_{\perp}E_{\perp}^* \\ E_{\parallel}E_{\parallel}^* - E_{\perp}E_{\perp}^* \\ E_{\parallel}E_{\perp}^* + E_{\perp}E_{\parallel}^* \\ i(E_{\parallel}E_{\perp}^* - E_{\perp}E_{\parallel}^*) \end{pmatrix} \quad (1)$$

where E_{\parallel} and E_{\perp} are the complex electric field components. I represents the intensity of the beam, Q and U represent the linear polarization, and V represents the circular polarization. The specific definitions of E_{\parallel} and E_{\perp} are with respect to an arbitrarily selected *plane of reference* including the direction of propagation. The sense is chosen so that $E_{\perp} \times E_{\parallel}$ is in the direction of propagation. Whenever a scattering event takes place, this plane of reference is *defined* so that

it is the plane containing both the incident and the scattered beams[13]. This plane of reference is no longer arbitrary thereafter.

Each photon interaction event occurs statistically, with a distance d determined from the probability distribution $\exp(-\mu_t d)$, where the extinction coefficient μ_t is the sum of the scattering and absorption coefficients $\mu_s + \mu_a$. The photon is then moved by this distance d , while appropriately treating any intervening interfaces [16].

2.1.3. Mie scattering

As described previously, a key difference between an advanced contemporary Monte Carlo implementation such as Pol-MC and a classical Monte Carlo implementation such as MCML is in the treatment of polarization. When a scattering event takes place, both the polarization state and the propagation direction of the photon change. Higher order Mie theory calculations are used to model scattering in Pol-MC; thus, both the polarization state and the propagation direction are updated at each scattering event. As Mie theory in general describes quite complex scattering phase functions, which have no analytic inverse, stochastic sampling algorithms must be applied to generate distribution function samples corresponding to these phase functions. The Mueller matrix of the scattering event is obtained through the use of the SCATMECH library developed by Germer, which is available to the public [26]. Using the Mueller matrix of the scatterer and the current polarization state of the photon, a scattering plane and a scattering angle are statistically sampled. The probability density distribution $P(\theta, \phi)$ of scattering events in a small solid angle around θ and ϕ is $P(\theta, \phi) \equiv \mathbf{I} \cdot \mathbf{S}' \sin \theta$. This probability density distribution itself is sampled by the algorithm presented by Jaillon [20].

2.2. Contributions to Pol-MC

2.2.1. Multiple layers implementation

Côté's [19] original implementation of Pol-MC did not allow the convenient simulation of complex multilayer structures. Prior to the extensions implemented by the authors, all the simulation objects were required to be placed in the interior of the “World” object, which was defined to be vacuum medium; further, object surfaces could not be coincident, that is, no touching layers were possible. In order to correctly treat this latter case, when objects in the simulated space are touching, it is necessary to specially treat this case of coincident surfaces. In our extension, we handle this situation by automatically detecting surface elements that are coincident, then explicitly initializing the surface element structure attributes (`ObjectOutside`) to point the appropriate adjacent object. Figure 1 shows a graphic illustration of how this “touching” of layers is accomplished. Both layer objects have two surface elements: one on top and one on the bottom. When the two surface elements are detected to be at the same location, their `ObjectOutside` are modified so that they correctly point to the adjacent layer.

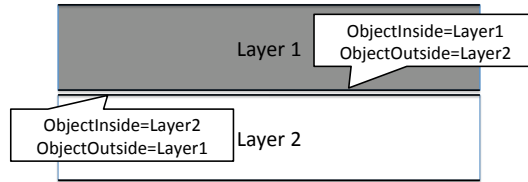


Fig. 1. Illustration of how we handle layers “touching” each other.

2.2.2. Geometry checks

Any computer simulation involves calculation at a finite numerical precision. By implication, simulations containing an extremely large number of geometric computations must correctly deal with surface locations in a manner that is not subject to errors inherent in floating-point computation, such as numerical round-off error. To this end, the manner in which interface locations are tested, and the manner in which a photon makes an interface transition are quite carefully implemented in Pol-MC, and we also added additional refinements to this mechanism. We have incorporated two *safety distances* for geometry checks: PHOTON_SAFETY_DISTANCE is used for securely moving the photon from one side of an interface to another, OBJECT_SAFETY_DISTANCE = 0.5·PHOTON_SAFETY_DISTANCE is used for checking if a certain location is within the interior of an object, and this modification is necessary in order to take correct consideration of surfaces that touch each other, which might otherwise be treated as a geometry initialization error. For example, in Fig. 1, the surface elements that touch each other occupy the same location in space, which is *not* due to geometry setup errors, but is at the explicit request of the end-user.

2.2.3. Detector implementation in Pol-MC simulation

In order to realistically simulate physical instruments, it is necessary to correctly model various detection geometries; for example, detectors in the simulation must behave analogously to collection optical fibers used in experiments. The detector objects used in Pol-MC are fundamentally “box” objects with special settings that allow the collection of photons with no further propagation; that is, a detector is the end-point for any detected simulation photon. Although at first this may seem simple, it should be noted that implicit to every surface element participating in the Pol-MC simulation geometry is the capability of accumulating full angular and polarization statistics on every photon passing through the surface, and this capability is shared by the surface elements defining the detector. Although many real-world detectors have much more complicated geometries, such as the circular surfaces of optical fibers, this simple detector implementation allows for computational simplicity. We use both flat and tilted detectors in the present simulation. In order to tilt a detector, i.e., to place a detector that isn’t perpendicular to the xy-plane, the user explicitly specifies Euler rotation angles (α, β, γ) as part of the detector specification in the input file. The specific angled detector used by our simulation is tilted at 45° from the z-axis.

2.2.4. Mixture model

One of the most significant extensions we made to the Pol-MC program is that of a scattering mixture-model. This model allows the arbitrary mixing of any number of component scatterer types, including even other mixture-model scatterers. For example, to accurately model a distribution of nuclei sizes in scattering tissue, the scatterer mixture would be adjusted to match the size distribution function governing the scatterers. Similarly, a simpler case with perhaps more immediate physical relevance to the present simulation involves the mixture of a small amount of Rayleigh-type scattering associated with the tissue matrix, with the Mie-regime scattering from the cell nuclei in the various model layers. This type of scattering model may be especially important for modeling precancers, since histologically the precancerous basal layer begins with a mixture of scatterers associated with different cell types. Further, the properties of each type of cell may be governed by a statistical distribution. For example, one cell type may scatter analogously to a size distribution of spheres (with optical properties equivalent to normal cell nuclei), and another cell type has a distribution representing abnormal nuclei with appropriate sizes and other properties. The mixture model implemented here is a piece-wise uniform distribution over a user defined number of scatterers, where the probability of the par-

icipation is specified through its associated scattering coefficient μ_s . It will immediately be seen that any probability distribution can be modeled to any desired degree of accuracy using this type of mixture, simply by including enough components. Since $\mu_s = \sum_{i=1}^N \mu_{si}$, where i is the index of scatterer ranging from 1 to N , and μ_{si} is the scattering coefficient of the scatterer i , we can scale the scatterers with the random variable:

$$Pr(X = k) = \frac{\mu_{sk}}{\sum_{i=1}^N \mu_{si}} \quad (2)$$

where X is the uniform random variable representing the probability that the scatterer is of type k . Although our implementation of this extension to Pol-MC is complete, we do not use this model in the simulations shown below, and for this reason further discussion is beyond the scope of this study.

2.2.5. Parallel computing

The main drawback of Monte Carlo methods is that they are extremely computationally intensive. However, it is natural to parallelize photon Monte Carlo since each photon is independent from every other. Thus, photon Monte Carlo can be run on as many separate processes occupying separate cores and separate nodes in a cluster computer, the results of which are subsequently combined. To implement this high degree of scalability, it is important to make the properly initialize the random number generator; otherwise, running many jobs concurrently may cause re-use of the same random number seed, resulting in wasteful duplications. Naive initialization of the random seed is often performed from the system time stamp or from the system process ID number, either of which can be duplicated over a large number of processes running on different nodes. In order to ensure that distinct seeds are used, regardless of the number of independent processes, we modified Pol-MC so that the seed initialization step used the system random device “/dev/urandom”; this choice of initialization effectively reduces the probability that different runs can accidentally use the same random number seed to zero.

2.3. Analysis for the simulation outputs

For this type of simulation, two types of analysis are pertinent. First, one may evaluate the experiments with which simulated variables correspond to physical measurements. We refer to this type of analysis as “outside-the-box”. Second, one can investigate variables that it is impossible to physically measure with current technology. This “inside-the-box” output is possible because complete information about the history of each simulated photon path, including all scattering and interface transitions that the photon has experienced, is available. We will show that this “inside the box” information, when correlated with the model geometry, provides important conclusions with respect to instrument feasibility and can be used to guide the design of instruments and data analysis algorithms.

2.3.1. “Outside-the-box” analyses

For “outside-the-box” analyses, our simulation is configured to deliver variables analogous to those that would be measured by a given instrument. The simulation can provide not only the light intensity collected by the fiber, but also the entire set of Stoke’s parameters, which provide information relevant to polarization sensitive detection. Physical property information such as reflectance or transmittance can also be calculated. In order to obtain full spectral information, the incident wavelength is varied, and the simulation is repeated until the full wavelength range is obtained. Although “outside-the-box” simulations are often less expensive than actually building the instrumentation and conducting experiments, this is not always the case, and the specifics must be judged carefully. The relative costs of running Monte Carlo simulations, which can often be very computationally intensive must be considered.

2.3.2. “Inside-the-box” analyses

A unique advantage of using Monte Carlo simulation is that it produces a complete list of photon traveling paths, i.e., the position, weight, and polarization state of every photon at every step at every move through the simulation geometry. This list of photon traveling record is not easily accessible experimentally, if at all. It should be noted that with respect to effectively using this type of information, it is perhaps even more important than it normally is that the model parameters accurately represent what is being simulated. For example, with respect to simulation of biological tissue specifically, although much information is available, almost as much still remains to be discovered [27].

The signal-to-background ratio (SBR) of an instrument is an important indicator of whether a given design is feasible. Here we give an example of how comprehensive photon path information can be used to estimate the maximum possible s/b achievable by an instrument. With respect to this analysis it must be stressed that whether or not an instrument can actually achieve this s/b depends on the specifics of the physical measurement; that is, whether or not the signal can be physically discriminated from the noise. Here we define the signal as photons detected from a specific layer of tissue via scattering processes. Noise is defined as all detected photons that are not signal. In our analysis of the photon paths recorded, the collected photon packets are classified into the following categories:

- I Photons that passed above the basal layer;
- II Photons that scattered *once* from the basal layer;
- III Photons that scattered *more than once* with at least once from the basal layer; Note that III and II are similar, but II is defined more strictly, and is included in III. Both III and II are possibly considered “signal”.
- IV Photons that passed below the basal layer.

The signal-to-background ratio (SBR) is calculated as

$$SBR_k = \frac{\Phi_k}{I - \Phi_k} \quad (3)$$

where I is the total intensity, k is the index of the photon category, and Φ_k is the surface fluence at the detector surface for all photons from category k , which is defined as:

$$\Phi_k = \sum_{i=1}^N w_i \text{ [Watt]} \quad (4)$$

where w_i is the weight of photon i that hits the surface of the detector.

By definition, fluence is the integral of the radiance over all directions [27]:

$$\phi(\mathbf{r}) = \int_{4\pi} \mathbf{L}(\mathbf{r}, \hat{\mathbf{s}}) d\omega \quad (5)$$

Following this definition, the volume fluence from our simulation can be calculated as:

$$\Phi(x, y, z) = \sum_{i=1}^N w_i(x, y, z) \text{ [Watt]} \quad (6)$$

where $w_i(x, y, z)$ is the weight of the photon (with index i) at position (x, y, z) . Note that volume fluence is different from surface fluence defined in Eq. 4 by their spatial definitions, and with

respect to volume fluence specifically, for reasons of computational expedience, most Monte Carlo simulations accumulate this parameter only at photon *scattering* events, and this preferentially scales the sum in 6 by a factor of μ_t . Keeping this latter in mind, this volume fluence gives us a spatially-resolved volume to show where scattering occurs most often. Combining Eq. 6 and the aforementioned photon categories, we can visualize and compare the different categories of photons detected and see where they were scattered from.

3. Results

3.1. Simulation parameters

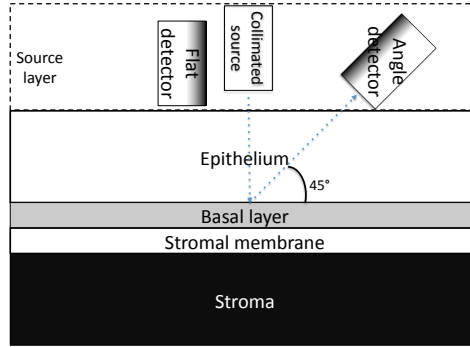


Fig. 2. Illustration of the geometry in our simulation. Photons enter the tissue from a collimated source. The tissue consists of four layers: epithelium, basal/precancer layer, basement membrane, and stroma. These four layers are infinite layers and they touch each other with no gaps in between. Two detectors are placed above the epithelium with very small gaps. The angled detector has a 45° angle with the epithelium, while the flat detector is aligned with the surface of the epithelium. Note that a source layer is placed on top of the epithelium layer both to provide some degree of index-matching and for computational purposes.

In our simulation, the geometry is set up as illustrated in Fig. 2. A four-layer model is simulated: (1) epithelial layer, (2) thin basal layer or precancer layer, (3) basement membrane, implemented as a thin transparent layer, (4) stroma, implemented as a thick layer of scattering material. We also have a collimated light source with photons entering the tissue perpendicular to the surface. Two detectors are placed above the epithelium. The angle detector has a 45° angle with the epithelium, while the flat detector is aligned with the surface of the epithelium.

In the simulation, the optical properties are defined in the input XML files. The wavelength used is 633 nm and all simulation results shown (excepting the wavelength-series) are simulated with 1.28×10^9 photons. Tab. 4 and Tab. 1 show some of the most important optical properties selected in this study. These parameters are chosen carefully to match the values reported in literature, on the conservative side [27]. The properties that we have used for the stromal membrane are selected to be consistent with those of other biological membranes. Although these properties are plausible, we believe that the specific properties of this membrane may be

unknown.

Table 1. Tissue layer thickness used in this study. Four sets of simulation parameters are presented here.

thickness and radius (μm)	Model 1	Model 2	Model 3	Model 4
Source layer	10000	10000	10000	10000
Epithelium	250	250	230	190
Basal/precancer layer	20	20	40	80
Basement membrane	30	30	30	30
Stroma	1000	1000	1000	1000
Basal layer scatterer radius	6 (normal)	8 (abnormal)	8 (abnormal)	8 (abnormal)

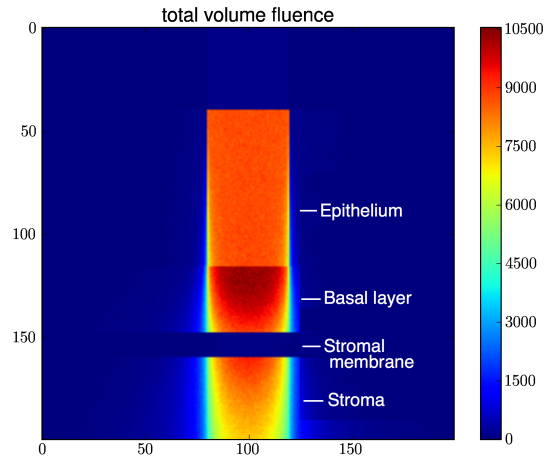
3.2. Scattering event distribution

We generate the total fluence plots (as defined in Eq. 6 to check the scattering distribution spatially. The total scaled fluence for Model 4 is shown in Fig. 3(a). In the figure, we can interpret the intensity as the density of scattering events in each location. The basal layer has a large number of scattering events both because it is closer to the source, and because its scattering coefficient (μ_s) is 120 cm^{-1} , is slightly higher than that of the epithelium which is 100 cm^{-1} . The dark gap between the basal layer and the stroma corresponds to the stromal membrane; with μ_s of zero, the scaled fluence for this region is also zero. Finally, although we would expect the stroma to have high scaled fluence because it is highly scattering ($\mu_s = 150$), it does not show high intensity in the figure; this is mostly because there are less photons that enter the stroma.

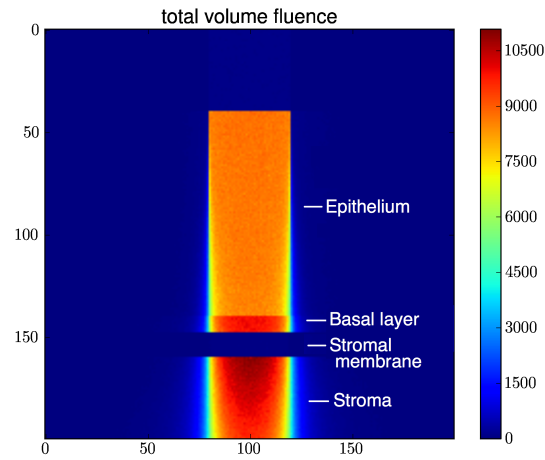
With respect to the scattering intensity in the stroma, this idea is borne out by comparison between the two models; if we compare Model 4 (Fig. 3(a)) with Model 2 (Fig. 3(b)), with all the parameters being the same except the basal layer thickness, we find the scaled fluence in the stroma much higher in Model 2 than in Model 4. This shows that because of the thicker basal layer, and also due to the more complicated scattering geometry, fewer photons are likely to penetrate through the basal layer and enter the stroma in Model 4.

3.3. Studying detector angles

In order to study the specifics of detector geometries, we classified the photon traveling paths and plotted the fluence plots only for photons that are collected by a given detector. This method allows us to specifically evaluate the photons detected, which means it effectively allows us to analyze where scattering events have happened, after the fact of detection, for detected photons. Fig. 4(a) and Fig. 4(b) are two examples of fluence plots from detectors, one from the angled detector and one from the flat detector. We see very clearly from both images that the photons form two beams: one going into the tissue, and one scattered near the bottom of epithelium and in the basal layer. We can also tell from these figures that the fluence is always most significant where the two beams overlap. This means that since the flat detector is placed right next to the source, its distance from the incident beam is smaller, allowing more overlap between the beams. Since the difference between these two fluence plots is largely caused by distances between the detectors and the light source, it suggests that the rotated angle of the detector is not as important a factor as the distance from the source. We find this conclusion to be particularly interesting, as it is counter-intuitive to analysis based strictly on ray-tracing considerations.

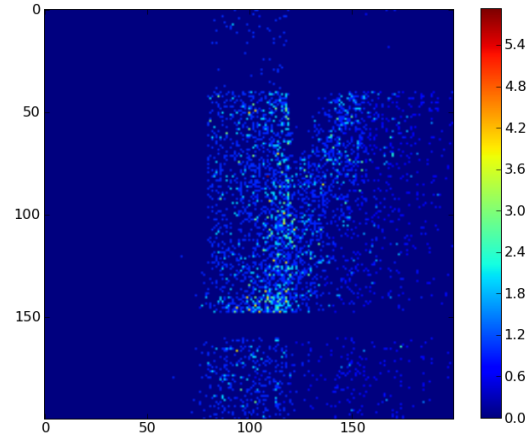


(a)

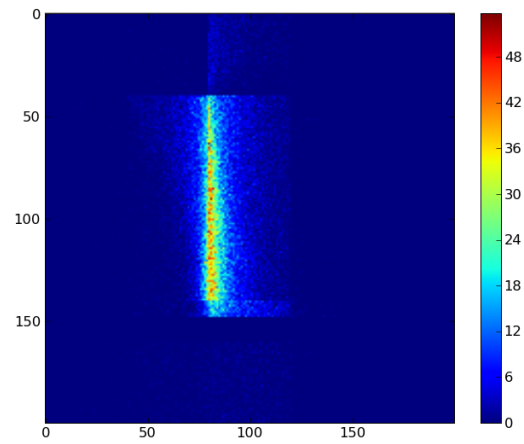


(b)

Fig. 3. Total fluence for Model 4 (Fig. 3(a)) and Model 2 (Fig. 3(b)). Photons enter the tissue from the top. These volume fluence plots represent the total spatial scattering density. The basal layers in both models have high fluence and show a clear boundary between the epithelium and basal layers. Stromal membranes do not scatter, so they have a dark color in these figures. Stroma layers are also highly scattering, but the intensity also depends on how many photons actually enter the layers. Since Model 4 has a thick basal layer, fewer photons enter the stroma and less scattering is indicated.



(a)



(b)

Fig. 4. Fluence for Model 1 with an angle-detector(Fig. 4(a)) and flat-detector (Fig. 4(b)). This image shows the fluence associated with all photons collected by a detector. The photons form two beams: one going into the tissue, and one scattered from the bottom of epithelium and in the basal layer. The fluence is always most significant when the two beams overlap. This plot is for uncollimated detectors, which means all photons arriving at the detector are collected, regardless of their incident angle at the detector surface.

3.4. Detectors with collimation

Collimated light is light traveling in a nearly parallel direction. For a source in a Monte Carlo simulation, which can be seen as a type of ray-optics computation, a collimated source is often perfectly collimated, that is, it is comprised of exactly parallel rays, with no divergence at all. This means that any spread in the beam diameter with propagation is associated with either scattering events, or with interface transitions.

In simulation, we can study the effects of collimation by adjusting the acceptance angle range for photons incident on the detector. In a physical measurement, this selection is automatic due to the *numerical aperture (NA)* of the physical detector. In Fig. 5(a) and Fig. 5(b), we show the fluence plots of photons that scatter from the basal/precancer layer, detected by an angle detector. In Fig. 5(a), we can clearly see many photons that have scattering events right below the detector but not on the straight line path. These scattered photon are eliminated in Fig. 5(b) with a collimated detector. These results further confirm our speculation in Section 3.3 that if detector collimation is not considered the detectors would perform similarly regardless of rotation angles. However, when the detectors are collimated, which here specifically means when we consider only photons with acceptance angles with minimum acceptance cosine = 0.95, fewer photons are collected by the detector, and further, photons that carry information not relevant to the target layer are eliminated by the collimated detectors. This acceptance cosine of 0.95 corresponds to $NA = 0.35$, which is within the range of optical fibers.

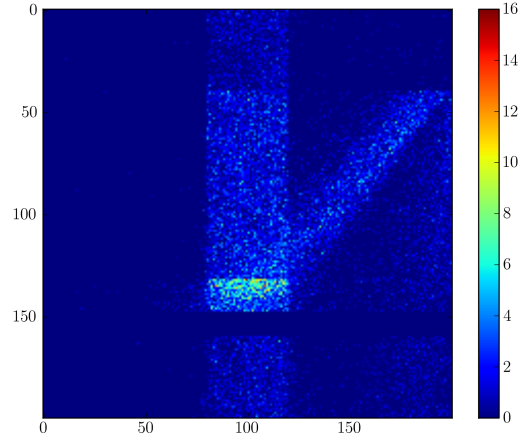
3.5. SBR analysis

The SBR can be calculated by classifying the photon traveling paths in order to determine if the signals collected by detectors were actually from the region of interest. Here we present the SBR from an angle detector, with the signal defined as photons scattering from the basal layer. From Tab. 2 and Tab 3, we can see that SBR for collimated detectors are significantly higher than SBR for uncollimated detectors as much as 12–25X higher. This result agrees with the results in Section 3.4 that collimated detectors remove unwanted noise associated with excessive scattering from outside the region of interest.

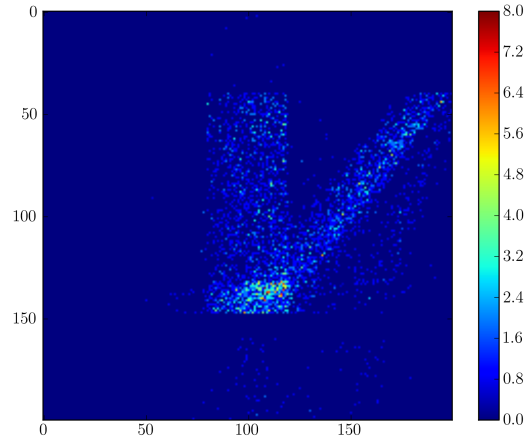
Also, in Tab. 2 and Tab. 3, we see that the thickness of the target layer (in this case, basal layer) dominates SBR. Model 4 has the thickest basal layer of $80 \mu m$, while that of Model 1 and 2 is 20, and Model 3 is 40. An additional factor for SBR may be related to the scatterer size. The difference between Models 1 and 2 is the scatterer size of their basal layers. Model 2 has an abnormal scatterer size of $8 \mu m$, while Model 1 is 6. However, due to the combined interactions of increasing basal layer thickness, and increasing scatterer size, which are present in these simulations, the isolated effect of either factor alone cannot be determined without further simulations.

Table 2. SBR calculation for an angle detector: collimated (minAcceptanceCosine: 0.95). All four models are used here.

	Epithelium	Basal	Stroma	SBR
Model 1	8986.411583	33157.974209	1746.195385	3.089461
Model 2	636.252699	7006.976886	1191.891885	3.832835
Model 3	558.302066	13458.732453	1156.406297	7.848992
Model 4	385.228121	23491.135433	1003.029843	16.921304



(a)



(b)

Fig. 5. Fluence for Model 1 with an uncollimated angle-detector(Fig. 4(a)) and collimated angle-detector (Fig. 4(b)). These images show the fluence caused by all the photons detected by a detector. The collimated detector accepts photons that have acceptance cosine values between 0.95 and 1, whereas the uncollimated detector accepts all photons that hit the detector, regardless of acceptance cosine. These figures show that collimated detectors eliminate many photons that scatter from the stroma and other undesirable locations, and that they enhance the collection of photons scattered from the basal layer. This acceptance cosine of 0.95 corresponds to $NA = 0.35$, which is within the range of optical fibers.

Table 3. SBR calculation for an angle detector: uncollimated (minAcceptanceCosine: 0.0). All four models are used here.

	Epithelium	Basal	Stroma	SBR
Model 1	163053.068867	51621.609665	43415.211736	0.250022
Model 2	111216.721489	21342.653985	25489.689007	0.156120
Model 3	110726.912445	42718.543371	23006.365514	0.319431
Model 4	108527.006232	87430.646328	19167.059612	0.684688

3.6. Wavelength series analysis

By varying the wavelength in the simulation, and running multiple simulation runs, a full spectrum can be calculated. This corresponds to what we refer to as a “physical simulation” in the preceding discussion: that is, it simulates what could actually be measured by an instrument. The simulated spectra are shown in Fig. 3.6 with both flat and angle detectors of Model 1 and Model 3. The wavelength range is 500–700 nm, and 10 spectral points are simulated in this range, which results in a resolution of 20 nm. Although there are “ripples” in these spectra, which may be caused by Mie scattering effects; we find it unlikely that these spectra are representative of physical spectra from biological tissue. It seems likely that this type of ripple effect would usually be “washed out” through the combined action of different sizes of scatterers and multiple scattering effects. For this reason, our implementation of the scattering mixture model may be particularly relevant to future simulation work.

We notice that in the angle-detector spectra, the photon counts are significantly lower than those for flat-detectors. This is predominantly due to the fact that the angle-detectors are further away from the source. The low number of photons used to generate these spectra make us skeptical of the correctness of these plots. However, it may not be practical, at least in this study, to simply increase the number of photons used in the simulation until the spectra are fully resolved. Note that these simulations are extremely intensive in terms of computational time. Each wavelength takes 10^9 photons, which means having 10 points in a spectrum takes 10×10^9 photons, which is very time consuming. At present, it appears that using the Monte Carlo technique to run full spectral simulations with any but the simplest source and detector configurations (i.e. even simpler than those used here) is essentially unfeasible.

4. Discussion and conclusion

Tissue optics has demonstrated its potential in recent years, and various experimental instrument designs have been developed to non-invasively measure signals specific to precancer diagnosis. However, whether the instrument is actually specifically probing the desired signal or not has remained an open question. Researchers indirectly measure instrument parameters such as depth sensitivity, but there is no standard metric to use in instrument design. This fact is particularly evident when considering complex photon path behavior in scattering media, where it is not at all evident that considerations based on simple ray-optics estimates, or analogous measurement, are applicable.

Pol-MC is a tool that provides additional, highly valuable information relating to the history of the photon traveling paths, including polarization. For example, by classifying photon paths into various categories it is possible to understand where the scattering events have the highest density, and the associated concept of the SBR allows us to evaluate whether a given source / detector configuration is collecting from the region of interest. All of this information has great potential to offer investigators a way to understand the full degree of interaction of light with biological tissue, and in addition it offers a standard tool to evaluate multiple objec-

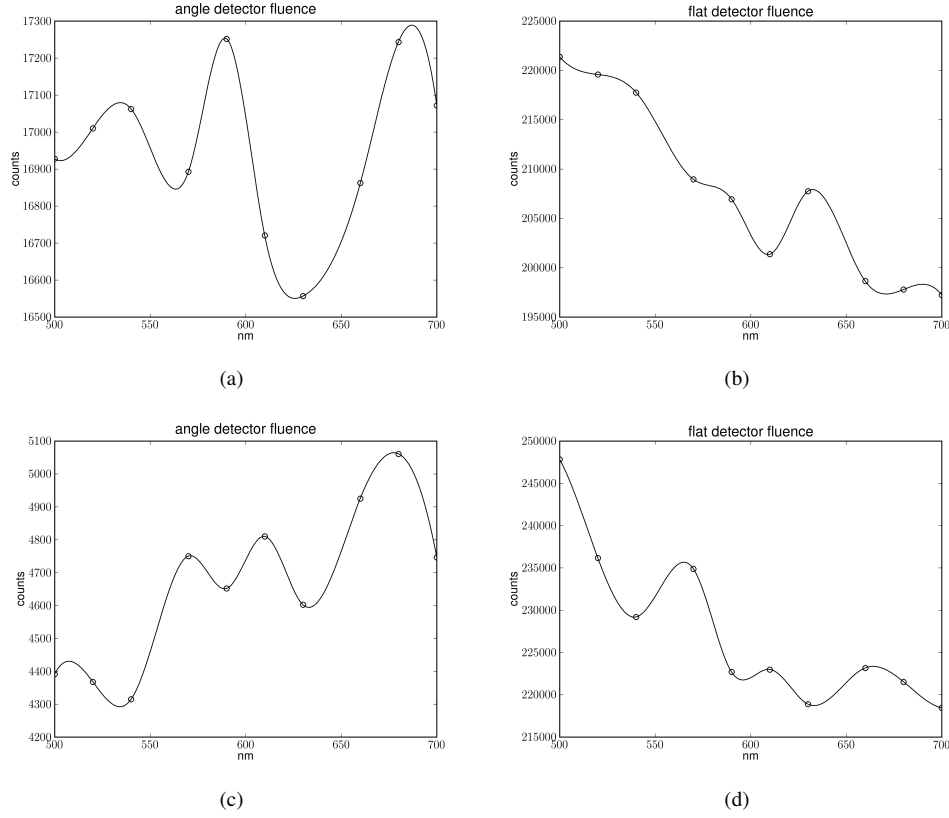


Fig. 6. Simulated spectra for Model 1: angle-(Fig. 6(a)) and flat-detectors(Fig. 6(b)), and Model 3: angle-(Fig. 6(c)) and flat-detectors(Fig. 6(d)). The wavelength range is 500–700 *nm*, and 10 spectral points are simulated in this range, which results in a resolution of 20 *nm*.

tives in instrument design. Furthermore, as most of the widely-used Monte Carlo simulations use the unpolarized Henyey-Greenstein phase function as the scattering angle sampling mechanism, Pol-MC offers a much more realistic model for scattering, including higher order Mie scattering, with full polarization information.

Our study of comparing angle- and flat-detectors shows that for uncollimated detectors, the rotation angle of the detector has little effect on the collection efficiency. However, by adjusting the acceptance angle for photon incidence, a collimated detector can be simulated. Specifically, when the minimum acceptance cosine at the detector has a value of 0.95, the angled-detector is able to eliminate a large fraction of the photons that are scattered from locations outside the region of interest. The results from SBR confirmed this: the collimated detector has a SBR that is much higher than that of the uncollimated detector. This observation is in correspondence to physical intuition, as angled detectors aim at the region overlapping with source illumination [24]; however, there are fine details of the behavior that are counter-intuitive to expectations based solely on a ray-optics analysis.

In our wavelength series analysis, we were able to simulate a series of wavelengths in the range 500–700 *nm*. However, comparing the spectra from flat- and angle-detectors, it seems very likely that the spectra from angle-detectors are exhibiting variations which have much

more to do with the small photon number contributing to the spectra than to actual spectral features which may be present. Since collimated angle-detectors provide very good Max-SNR, this type of detector does eliminate many unwanted photons, and this effect in terms of noise reduction is a reliable result. Nonetheless, this elimination of undesired photons has significant drawbacks in terms of the computational resources actually required to effectively simulate these detectors, in order to achieve the same level of signal strength as uncollimated or flat detectors. From Tab. 3, we can find that approximately half of the photons of uncollimated detectors are eliminated when we collimate the detector. This implies that we may need twice as many photons at the detector, which means more than 10 times in the simulation, for angle-detectors to achieve the same signal strength to get a reasonable spectrum. As the wavelength series analysis is already very computationally expensive, this raises a potential concern for the feasibility of Monte Carlo technique in general to provide spectra with good quality, even for this simple simulation geometry.

There are two random factors that we want to consider in the simulation. One is the variance of scatterers, and the other is inter-patient variance. We have briefly discussed the mixture model we contributed to Pol-MC (Section 2.2.4). This mixture model is not investigated in this paper because it is not in the scope of the preliminary part of this study that we have discussed. However, the mixture model brings in another random variable to model the variance of scatterers. With the use of a model, it will be possible to simulate more than one type of scatterer mixed in an object, which is closer to reality. It will immediately be seen that issues relating to inter-patient variance are significantly more complicated than a simple physical technique such as a mixture model. However, if we extend the stochastic formalism of which the Monte Carlo technique is an example to its full conclusion, it is possible that even a factor such as inter-patient variance could be included in the simulation as simply another random variable. In this case, the simulation would generate a “population” of measurements, which would further facilitate the development of diagnostic algorithms.

The simulation tool we present can be extended to consider various modalities. Considering its ability to provide information such as polarization states of the photon, it can be used to study effects caused by changes of optical properties. Our analyses provide a platform for comparing optical property changes: fluence plots and Max-SNR can be used as metrics to study changes of optical parameters. This “virtual probe design” tool may guide future clinical applications for detection of epithelial cancer.

5. Acknowledgment

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper. URL: <http://www.tacc.utexas.edu>

Table 4. Tissue properties used for this study. *Note that some of the parameters vary between different runs, the layer thickness are specified in Tab. 1. **HG: Henyey-Greenstein.

	thickness (μm)	scattering	g	index(media)	index(scatterer)	μ_a	μ_s	scatterer radius(μm)
Source layer	10000	HG**	0.5	1.335	—	0.1	1	—
Epithelium	varies	Jailon	—	1.369	1.41	1	100	6
Basal layer	varies	Jailon	—	1.369	1.41	1	120	varies
Stromal membrane	30	HG**	0.5	1.41	—	1	0	—
Stroma	1000	Jailon	—	1.369	1.41	1	150	6