Application limits of the scaling relations for Monte Carlo simulations in diffuse optics. Part 1: theory

Caterina Amendola,[1,†] **Giulia Maffeis,**[1,†]
Andrea Farina,[2] **Lorenzo Spinelli,**[2,*]
Alessandro Torricelli,[1,2] **Antonio Pifferi,**[1,2]
Angelo Sassaroli,[3] **Duccio Fanelli,**[4]
Federico Tommasi,[4] **and Fabrizio Martelli**[4]

[1]*Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano 20133, Italy*
[2]*Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche, Piazza Leonardo da Vinci 32, Milano 20133, Italy*
[3]*Department of Biomedical Engineering, Tufts University, 4 Colby Street, Medford, Massachusetts 02155, USA*
[4]*Dipartimento di Fisica e Astronomia, Università degli Studi di Firenze, Via Giovanni Sansone 1, Sesto Fiorentino 50019, Italy*
[†]*These authors contributed equally to this work*
[*]*lorenzo.spinelli@polimi.it*

Abstract: Monte Carlo (MC) is a powerful tool to study photon migration in scattering media, yet quite time-consuming to solve inverse problems. To speed up MC-simulations, scaling relations can be applied to an existing initial MC-simulation to generate a new data-set with different optical properties. We named this approach trajectory-based since it uses the knowledge of the detected photon trajectories of the initial MC-simulation, in opposition to the slower photon-based approach, where a novel MC-simulation is rerun with new optical properties. We investigated the convergence and applicability limits of the scaling relations, both related to the likelihood that the sample of trajectories considered is representative also for the new optical properties. For absorption, the scaling relation contains smoothly converging Lambert-Beer factors, whereas for scattering it is the product of two quickly diverging factors, whose ratio, for NIRS cases, can easily reach ten orders of magnitude. We investigated such instability by studying the probability-distribution for the number of scattering events in trajectories of given length. We propose a convergence test of the scattering scaling relation based on the minimum-maximum number of scattering events in recorded trajectories. We also studied the dependence of MC-simulations on optical properties, most critical in inverse problems, finding that scattering derivatives are ascribed to small deviations in the distribution of scattering events from a Poisson distribution. This paper, which can also serve as a tutorial, helps to understand the physics of the scaling relations with the causes of their limitations and devise new strategies to deal with them.

1. Introduction

The study of photon propagation in highly scattering media is a fascinating field, with a wealth of applications ranging from biomedical optics to astrophysics, from atmospheric optics to the physics of random media, and many others [1]. The most commonly used approach to model photon migration through a highly scattering medium is the radiative transfer equation (RTE) [2], being the Maxwell equations usually unsuitable to solve photon propagation in complex disordered materials with multiple scattering interactions. Monte Carlo (MC) method is a quite

effective and common approach for numerical simulations of photon migration through scattering media based on the same assumptions of RTE – that is describing light composed of particles (photons) governed by energy conservation laws – and simulating many photon trajectories on a statistical basis. In principle, by using MC it is possible to simulate any geometry – *e.g.* the complex structure of the human head – and any combination of optical properties – namely, the absorption coefficient (μ_a), the scattering coefficient (μ_s), the refractive index (n), the scattering phase function [$p(\theta)$]. Yet, even with today's computer and the amazing power offered by graphic processing units (GPUs), the computational time required to provide simulations with acceptable uncertainties can be challenging, in particular when multiple simulations are needed for solving inverse problems.

An attractive perspective to reduce computational time and improve consistency in a set of MC simulations is to apply scaling relations [3], directly derived from fundamental properties of the RTE. Starting from a background MC simulation run with a given set of optical properties [μ_{a0}, μ_{s0}, n_0, and $p_0(\theta)$], it is possible to derive (without re-running a new simulation) a perturbed MC simulation for a new set of optical properties ($\mu_{a0} + \Delta\mu_a$, $\mu_{s0} + \Delta\mu_s$) by applying scaling relations based on photons' pathlengths and number of scattering events in a given trajectory.

This approach was initially introduced in the neutron transport field for sensitivity and accuracy studies in forward problems [4–6]. Later, the same approach was used in the biomedical optics field for forward photon migration problems [3], extending the relations also to perturbation on the phase function [7] and including polarized light [8]. Noticeably the approach was profitably used for inverse problems like estimating tissue optical properties in homogeneous or layered media [9–11]. Further, the scaling relations were also exploited for 3D reconstruction of the optical properties in diffuse [12] or photoacoustic regime [13].

Before being evaluated from the point of view of use in an inverse problem for biomedical applications, a perturbation MC is first of all a relevant support regarding the direct propagation problem and thus could be a very effective tool for training neural networks and artificial intelligence systems that recently are widely used on biological applications [14,15]. Indeed, MC simulations can provide a validated and reliable method for generating robust data sets for training artificial intelligence tools (*e.g.*, machine learning, deep learning) for instance in diffuse optical tomography (DOT) [16]. Therefore, given the extreme real-world complexity of biological tissues, the perturbation approach may allow a rapid generation of data for training networks on situations dedicated to individual patients for which it is reasonable to have prior spatial information obtained by tomographic RX or MRI techniques. This would potentially result in increased speed and effectiveness of the techniques with a significant simplification of the traditional training approach, *i.e.*, by re-running a simulation for different optical properties, involving consideration of the full complexity of the biological tissue being studied. In fact, deep learning training to map the entire case history of possible detected signals requires large multi-type datasets to return robust and accurate reconstructions. It is worth mentioning that the traditional approach to training consists, for instance, of running a large set of MC simulations for different optical properties. Therefore, the use of a perturbation MC forward solver would provide a computationally accurate and efficient light propagation model that is flexible and fast for mapping broad experimental scenarios.

Interestingly, no detailed analysis has been presented to our knowledge on the robustness of the scattering scaling relationship despite its use since many years. Therefore, the aim of the present work is twofold. On one side it is written in the form of a tutorial, to present the foundations of the scaling relations so as to fully appreciate the physics behind and understand the potentialities and limitations of such approach. On the other hand, we add substantial novel material in this field with an in-depth analysis of the convergence properties of the scaling relations and a study of the distribution $p_\ell(k, \mu_s)$ of the number of scattering events k for a trajectory of length ℓ. This distribution is one key component to understand the limits of the scaling relations for a

change of the scattering coefficient. At the same time it is also a source of inspiration for further advancements in dealing with those limits.

The adopted methodology is based on the probabilistic study of the photon trajectories, followed by the analysis of the convergence properties of the scaling relations in different scenarios, encompassing both continuous wave (CW) and time-domain (TD) cases. The general assumptions of the RTE are adopted. In addition, the scattering phase function is considered to be homogeneous all over the medium, though a direct derivation of scaling relations also for non homogeneous scattering phase function is presented in Appendix A, yielding an expression consistent with Ref. [7].

We organized the whole materials in two companion papers. In the present (first) paper, we recall in Section 2.1 the derivation of the scaling relations for a homogeneous medium. The extension of these formulas for non-homogeneous media is reported in Appendix A. Sections 2.2 and 2.3 discuss the convergence of the scaling relations by analysing the structure of the two scaling relations for absorption and scattering. Section 3 investigates the probability function $p_\ell(k, \mu_s)$ for the number of scattering events k undergone by a photon along a path of length ℓ for a given trajectory, which permits to understand the origin of the convergence fragility of the scaling factor for μ_s. Section 4 presents the derivative of the scaling relations with respect to μ_a and μ_s which is at the basis of the inverse problems. The second companion paper [17] is devoted to an extensive analysis of the scaling relations based on numerical results derived by running multiple MC simulations in various scenarios both for CW and TD, for the homogeneous and non-homogeneous geometry, for forward problems and inverse problems. This part will permit to get acquainted with a judicious use of the scaling relations and related methods.

2. Scaling relations

The standard approach to MC simulations is photon-based, that is, it follows the paths of a number of photons, where the direction and distance at which they are scattered or absorbed conform to the probabilistic sampling distributions. In this approach, the simulation evaluates whether a photon is received or not, and a new MC simulation is rerun each time the optical properties are changed, so the simulation is always run with the actual values of the optical properties.

To speed up MC simulations, scaling relations can be applied to an existing MC simulation obtained in the initial state of the medium to generate a new data set with different absorption and scattering properties (new state of the medium). We name this approach *trajectory-based* approach because it is based on knowledge of the paths of photons detected in the initial state. In principle, this approach should be much faster than the classical photon-based approach, in which a new MC simulation is re-run each time the optical properties are changed.

In this paper we follow the trajectory-based approach, which, from a set of possible trajectories in the initial state, calculates the probability of their occurrence also when the optical properties are changed from the initial state. Thus, in the trajectory-based approach the initial set of trajectories are selected as a sample in the space of all the trajectories for the considered case. In practice, the re-emission of photons from the medium is a weighted sum, in accordance with the probability of each trajectory, over the infinity of all possible trajectories. There is a similarity between the path integral method and the trajectory-based approach since both involve the probability for a photon to follow a certain path from an emission point of the source to a certain observation point [18,19].

Both photon-based and trajectory-based approaches converge to an accurate description of photon re-emission probability if an infinite number of photons are considered (first case) or all possible infinite trajectories are included (latter case). In particular, the trajectory-based approach permits to derive a new MC simulation with a different choice of optical properties starting from an existing one by simply changing the weights of every trajectory in accordance with the desired considered variation of the optical properties. Clearly, the newly inferred simulation is accurate

to the extent that the distribution of the original trajectories is sampled with sufficient coverage also for the new situation.

To better highlight the difference between the two methods we notice that, for a change in the optical properties of the medium, in the photon-based method we answer the question whether a photon still propagates through the detector or not, while in the trajectory-based method we ask how much the probability of traveling through the considered trajectory has changed. The answer to the second question, when an enough dense sample of trajectories is considered, permits, for any configuration of the optical properties, to calculate the photon re-emission from the medium. Both approaches lead to the calculation of the Temporal Point Spread Function (*TPSF*), in the first case by re-running a new MC simulation, while in the second case by re-scaling the detection probability of the initial set of trajectories.

2.1. Basic concepts

The scaling relations for the radiation collected by a specific detector can be retrieved from the expression of the fraction of collected radiation that travelled along a specific path in a homogeneous medium. Let us consider a path or trajectory with k scattering events, as the one reported in Fig. 1. We can define a *trajectory* in a medium as a set of points occupied by a travelling photon between a source point and an arbitrary observation point. The photon is passing through the medium and if a detector is placed at the observation point, the photon is annihilated.
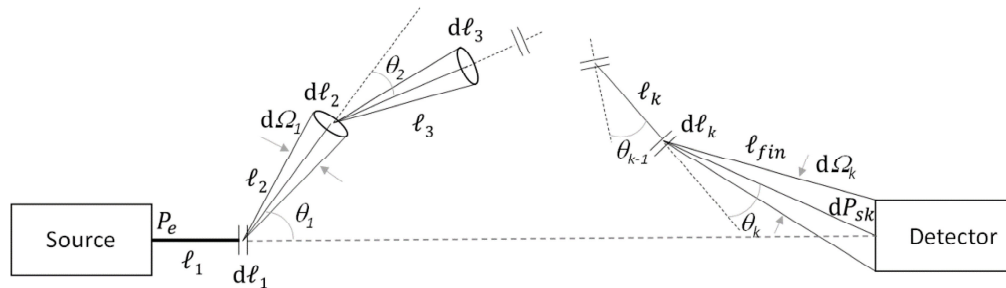
Fig. 1. Example of trajectory from source to detector with k scattering events. In the figure, for simplicity, the azimuth angle is not visible in each highlighted scattering; however, the figure refers to 3D propagation.

In this case, photons travel an ℓ_m path before being scattered at an angle θ_m within a solid angle $d\Omega_m$ and, after k scattering events, are collected by the detector. To derive the expression of the radiation fraction that travels along this specific path, it is necessary to highlight three properties:

- in a medium with absorption coefficient μ_a, and scattering coefficient μ_s, the radiation is attenuated by a factor $e^{-\mu_t \ell}$ when it propagates through a path ℓ, with extinction coefficient $\mu_t = \mu_a + \mu_s$;

- the fraction of radiation scattered in a volume of thickness $d\ell$ is $\mu_s d\ell$;

- the fraction of radiation scattered in a solid angle $d\Omega$ around the scattering angle θ is $p(\theta)d\Omega$.

In the above notation, we have assumed the general case of rotationally symmetric scattering phase functions, *i.e.*, scattering functions which can be represented as $p(\hat{s} \cdot \hat{s}') = p(\theta)$, with $\hat{s}$ direction of the incident radiation and $\hat{s}'$ direction of the scattered radiation.

Given these properties, the expression of the power fraction, $\mathrm{d}P_{sk}$, of radiation traveling along the specific path shown in Fig. 1 is expressed in Eq. (1) as:

$$\mathrm{d}P_{sk} = P_e e^{-\mu_t \ell_{fin}} \prod_{m=1}^{k} e^{-\mu_t \ell_m} \mu_s \mathrm{d}\ell_m \, p(\theta_m)\mathrm{d}\Omega_m =$$

$$= P_e \mu_s^k e^{-\mu_t \left(\sum_{m=1}^{k} \ell_m + \ell_{fin}\right)} \prod_{m=1}^{k} \mathrm{d}\ell_m p(\theta_m)\mathrm{d}\Omega_m \,, \tag{1}$$

where P_e is the power emitted by the source, ℓ_{fin} is the pathlegth travelled by the photons from the last scattering event to the detector, $\mathrm{d}\Omega_k$ is the solid angle within which the radiation is collected by the detector from the last scattering point, and k is the number of scattering interactions along the photon path.

Equation (1) represents the infinitesimal power that starts from the source and reaches the detector following a given trajectory (Fig. 1). It corresponds to the infinitesimal detected power along the same trajectory only for the case of no refractive index mismatch between medium and surroundings. We note that we can generate other trajectories by changing one or more of the azimuthal angles of the original trajectory and keeping all the other parameters unchanged. These trajectories are characterized by different end points inside the medium and the same infinitesimal power transmitted along them. These trajectories are considered equivalent for the application of the scaling relations [Eqs. (2, 3)].

Starting from a MC simulation in a homogeneous medium of optical properties $(\mu_{a0}, \mu_{s0}, p_0)$, Eq. (1) allows to determine the different contributions $\mathrm{d}P_{sk}$ when the optical properties are modified into (μ_a, μ_s, p). The scaling factor relates the new weights $W(\mu_a, \mu_s, p)$ of each trajectory to the ones of the starting MC simulation $W(\mu_{a0}, \mu_{s0}, p_0)$. Thus, it is defined as the ratio between the probability to travel the same path in the two scenarios with different optical properties:

$$W(\mu_a, \mu_s, p) = W(\mu_{a0}, \mu_{s0}, p_0)\frac{\mathrm{d}P_{sk}(\mu_a, \mu_s, p)}{\mathrm{d}P_{sk}(\mu_{a0}, \mu_{s0}, p_0)} =$$

$$= W(\mu_{a0}, \mu_{s0}, p_0)\left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_s - \mu_{s0})\ell} e^{-(\mu_a - \mu_{a0})\ell} \prod_{m=1}^{k} \left(\frac{p(\theta_m)}{p_0(\theta_m)}\right) . \tag{2}$$

The above expression for the scaling formula assumes that changes in the refractive index within the studied medium from initial to final state are not considered. In Eq. (2), ℓ and k represent the total length travelled by the photons in a given trajectory and the number of scattering events, respectively. The weight of the trajectory in the initial state, $W(\mu_{a0}, \mu_{s0}, p_0)$, is 1 if the trajectory is collected by the detector and 0 if it is not collected.

If the propagation is described by the same scattering phase function $p_0(\theta)$ of the ground simulation, the above scaling relation can be simplified. In this case, the scaling factor links the new weights $W(\mu_a, \mu_s)$ of each trajectory to the ones of the starting MC simulation $W(\mu_{a0}, \mu_{s0})$. Thus, it is defined as the ratio between the probability to travel the path of the trajectory in the two scenarios having different absorption and scattering coefficients:

$$W(\mu_a, \mu_s) = W(\mu_{a0}, \mu_{s0})\frac{\mathrm{d}P_{sk}(\mu_a, \mu_s)}{\mathrm{d}P_{sk}(\mu_{a0}, \mu_{s0})} =$$

$$= W(\mu_{a0}, \mu_{s0})\left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_s - \mu_{s0})\ell} e^{-(\mu_a - \mu_{a0})\ell} . \tag{3}$$

We observe that the new weight $W(\mu_a, \mu_s)$ is obtained multiplying the initial weight by the scaling factor $(F_a \, F_s)$ where:

$$F_a = \frac{W(\mu_a, \mu_{s0})}{W(\mu_{a0}, \mu_{s0})} = e^{-(\mu_a - \mu_{a0})\ell} \,, \tag{4}$$

$$F_s = \frac{W(\mu_{a0}, \mu_s)}{W(\mu_{a0}, \mu_{s0})} = \left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_s - \mu_{s0})\ell} . \tag{5}$$

Here, we have explicitly separated the scaling factor due to a change in the absorption coefficient, F_a, from that due to a change in the scattering coefficient, F_s. In the following Sections, the two factors will be discussed for a homogeneous medium, while in Appendix A the heterogeneous case is reported.

One last important observation is that the scaling relations reported in Eqs. (3–5) are valid for equivalent trajectories, as defined after Eq. (1), of the initial medium and need not to refer to the same trajectory. The implication of this observation is that we can gain some knowledge on the scaling relations in a finite medium, by using some well known distributions in the infinite medium geometry which are obtained without fixing a particular end point of a trajectory (see Section 3.1).

2.2. Scaling factor for absorption in a homogeneous medium

The scaling factor F_a reported in Eq. (4) is already widely used in MC simulations to model the effect of changes in the absorption coefficient [11,12].

Let us consider a basic MC simulation for a homogeneous medium with absorption μ_{a0}. Since the absorption coefficient does not vary from point to point, the scaling factor to μ_a depends only on the optical pathlength travelled by the collected photons. In the TD, where the weights of the received photons are classified into a histogram of time-of-flight t, the same scaling factor is applicable to all photons with the same time-of-flight. Then, we can derive a scaling factor for the $TPSF(\mu_a, t)$ resulting from the contributions of all possible trajectories with length $\ell = vt$ from source to detector, being v the speed of light, calculated with respect to a medium with absorption μ_{a0}:

$$F_{TPSF,a}(\mu_{a0} \to \mu_a, t) = \frac{TPSF(\mu_a, t)}{TPSF(\mu_{a0}, t)} = e^{-(\mu_a - \mu_{a0})vt} . \tag{6}$$

Starting from the $TPSF$ obtained for $\mu_{a0} = 0$, we can write:

$$TPSF(\mu_a, t_i) \approx TPSF(\mu_{a0} = 0, t_i)e^{-\mu_a v t_i}, , \tag{7}$$

where t_i represents the average of the i-th time bin, that is assumed to be sufficiently narrow to apply the same attenuation for all photons. We further highlight that Eq. (7) allows one to obtain a new $TPSF(\mu_a, t_i)$ from $TPSF(\mu_{a0} = 0, t_i)$ without resorting to the knowledge of each trajectory. In this case, the relative statistical error on the scaled $TPSF$ will be identical as that on the initial $TPSF$.

If the time window spanned by the $TPSF$, defined as $t_{max} = M\Delta t$, where M is the number of time bins and Δt their width, is large enough to include all the collected photons, $CW(\mu_a)$ can be obtained by integrating the $TPSF$, *i.e.*:

$$CW(\mu_a) = \sum_{i=1}^{M} TPSF(\mu_a, t_i)\Delta t . \tag{8}$$

We want to add a few considerations about the precision and accuracy on the estimated values of $TPSF$ and the CW. For the estimation of the $TPSF(\mu_a, t_i)$ at a given time point, the choice of Δt should be carefully evaluated case by case. By increasing Δt we increase the number of collected photons N_i in $(t_i - \Delta t/2, t_i + \Delta t/2)$, so improving the precision: indeed, the relative error on $TPSF(\mu_a, t_i)$ results $\epsilon_r(t_i, \Delta t) = 1/\sqrt{N_i}$. At the same time, however, we worsen the accuracy, which relies on the approximation that the photon weight $e^{-\mu_a v t_i}$ is constant within the time window. The opposite is true if we decrease Δt. It should lead to a compromise between precision and accuracy.

As for the CW case, we note that, when $\mu_a = 0$, all the received photons N contribute to CW with identical weights ($W_j = 1$), and the relative error on CW is $\epsilon_r = 1/\sqrt{N}$, being in this case $CW(\mu_a = 0) \propto N$ and the precision $\sigma_{CW} \propto \sqrt{N}$. We notice that the precision on the CW is independent of the bin width Δt. For the case $\mu_a \neq 0$ the relative error on CW is larger than for the case $\mu_a = 0$, being the weight of each photon less than 1 ($W_j = e^{-\mu_a \ell_j}$, where ℓ_j is the pathlength travelled by the j-th detected photon). Then, the increase of the relative error with respect to the case $\mu_a = 0$ is given by: $\epsilon_r(\mu_a)/\epsilon_r(\mu_a = 0) = \sqrt{N} \sum_{j=1}^{N} e^{-\mu_a \ell_j}$.

On the other hand, the accuracy of CW is mostly affected by the choice of the bin width, similarly to the *TPSF* case: in fact, at increasingly higher values of μ_a, the approximation that the photon weight is constant within a given time window of width Δt will be a poor one. However, this source of error can be eliminated by choosing a formula for the calculation of CW different form Eq. (8), that bypasses the evaluation of the *TPSF*: $CW = \sum_{j=1}^{N} e^{-\mu_a \ell_j}/(N_e S)$, where N_e is the number of launched photons and S the detector area [compare Eq. (21)].

The use of these expressions therefore becomes particularly risky in geometries where the mean path of the photons is very long, thus making difficult to reconcile a small Δt with a large t_{max}. Two examples are the non-absorbing infinite and semi-infinite media. Analogously, although in principle the scaling relation is applicable even in the presence of *infinite* absorption, the increase of noise due to high μ_a and long time-of-flight could hinder the convergence of the scaling method. In the companion paper [17] of this work, we will show that for configurations of common use (*e.g.*, absorbing slab or semi-infinite medium with $\mu_a < 0.5$ cm^{-1}) the scaling relation for absorption returns accurate results.

2.3. Scaling factor for scattering in a homogeneous medium

The scaling factor F_s for the scattering coefficient reported in Eq. (5) can be divided into two terms: the first, F_{s1}, is a power law factor having as its exponent the number of scattering events, the second, F_{s2}, is an exponential factor that depends on the total pathlength of the photon ℓ:

$$F_{s1} = \left(\frac{\mu_s}{\mu_{s0}}\right)^k, \qquad \text{and} \qquad F_{s2} = e^{-(\mu_s - \mu_{s0})\ell}. \tag{9}$$

These two terms have opposite behaviours: when F_{s1} increases to values >1, F_{s2} decreases to values <1 and *viceversa*. Moreover, they can easily reach huge orders of magnitude, as reported in Table 1. Due to the dependence on ℓ and k of $F_s = F_{s1}F_{s2}$, photons that travelled along paths of identical length can have very different scaling factors based on the number of scattering events.

As discussed for the absorption case, we can try to obtain a scaling factor for the $TPSF(\mu_s, t)$ in a homogeneous medium. However, because of the presence of the number k of scattering interactions, a straightforward calculation as done for the absorption case is not possible. Indeed, this calculation would require the knowledge of the probability distribution $p_\ell(k, \mu_{s0})$ for the number of scattering events k undergone by the received photons with pathlength ℓ. Consequently, the new $TPSF(\mu_s, t)$ can be obtained from the initial $TPSF(\mu_{s0}, t)$ as:

$$TPSF(\mu_s, t) = TPSF(\mu_{s0}, t) \sum_{k=0}^{\infty} p_\ell(k, \mu_{s0}) \left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_s - \mu_{s0})\ell}, \tag{10}$$

where $\ell = vt$. Then, the scaling factor for scattering, resulting from the contributions of all possible trajectories with pathlength ℓ from source to detector, is:

$$F_{TPSF,s}(\mu_{s0} \to \mu_s, \ell = vt) = \frac{TPSF(\mu_s, t)}{TPSF(\mu_{s0}, t)} = e^{-(\mu_s - \mu_{s0})\ell} \sum_{k=0}^{\infty} p_\ell(k, \mu_{s0}) \left(\frac{\mu_s}{\mu_{s0}}\right)^k. \tag{11}$$

Concerning the $CW(\mu_s)$, considerations analogous to the ones reported in the former Section can be done. The time window (t_{max}) can drastically affect CW results, due to the missed

Table 1. Scaling factors F_{s1} and F_{s2} for variations in scattering from μ_{s0} to μ_s for a set of ℓ and k values.

	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 20$ cm^{-1}	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 5$ cm^{-1}	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 11$ cm^{-1}	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 9$ cm^{-1}	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 10.1$ cm^{-1}	$\mu_{s0} = 10$ cm^{-1} $\mu_s = 9.9$ cm^{-1}
k	F_{s1}	F_{s1}	F_{s1}	F_{s1}	F_{s1}	F_{s1}
100	1.27×10^{30}	7.89×10^{-31}	1.38×10^{4}	2.66×10^{-5}	2.70×10^{0}	3.66×10^{-1}
200	1.61×10^{60}	6.22×10^{-61}	1.90×10^{8}	2.66×10^{-10}	7.32×10^{0}	1.34×10^{-1}
300	2.04×10^{90}	4.91×10^{-91}	2.62×10^{12}	1.87×10^{-14}	1.98×10^{1}	4.90×10^{-2}
400	2.58×10^{120}	3.87×10^{-121}	3.61×10^{16}	4.98×10^{-19}	5.35×10^{1}	1.80×10^{-2}
500	3.27×10^{150}	3.05×10^{-151}	4.97×10^{20}	1.32×10^{-23}	1.45×10^{2}	6.57×10^{-3}
600	4.15×10^{180}	2.41×10^{-181}	6.85×10^{24}	3.51×10^{-28}	3.92×10^{2}	2.41×10^{-3}
ℓ [cm]	F_{s2}	F_{s2}	F_{s2}	F_{s2}	F_{s2}	F_{s2}
10	3.72×10^{-44}	5.18×10^{21}	4.54×10^{-5}	2.20×10^{4}	3.68×10^{-1}	2.72×10^{0}
20	1.38×10^{-87}	2.69×10^{43}	2.06×10^{-9}	4.85×10^{8}	1.35×10^{-1}	7.39×10^{0}
30	5.15×10^{-131}	1.39×10^{65}	9.36×10^{-14}	1.07×10^{13}	4.98×10^{-2}	2.01×10^{1}
40	1.92×10^{-174}	7.23×10^{86}	4.25×10^{-18}	2.35×10^{17}	1.83×10^{-2}	5.46×10^{1}
50	7.12×10^{-218}	3.75×10^{108}	1.93×10^{-22}	5.18×10^{21}	6.74×10^{-3}	1.48×10^{2}
60	2.65×10^{-261}	1.98×10^{130}	8.76×10^{-27}	1.14×10^{26}	2.48×10^{-3}	4.03×10^{2}

contribution of large trajectories. In the companion paper [17], we will show how the error increases with the source-detector distance, due to the limited number of simulated trajectories compared to the larger number of possible ones.

Finally, for the reader's convenience, we refer to the Appendix B where the probability of receiving a trajectory when the optical properties are changed from an initial state to a final state is reformulated and justified on the basis of the elementary concepts of optical properties. In this scheme we offer a full justification of Eq. (10) and of the probability distribution $p_\ell(k, \mu_{s0})$. In this way, we provide the reader with all the basic elements necessary to understand the next steps of this study.

3. Probability function $p_\ell(k, \mu_s)$

We saw in the previous section the importance of the probability function $p_\ell(k, \mu_s)$ for the number of scattering events k undergone by a photon along a path of length ℓ regarding the scattering coefficient scaling factor. In this Section, we will calculate the probability function $p_\ell(k, \mu_s)$ starting from the case of an infinite homogeneous medium. In particular, in the infinite medium no fixed end point for a trajectory (*i.e.*, no detector) is considered and the probability functions are calculated based on all emitted photons that have travelled a pathlength ℓ. We also refer to this case as *free propagation*.

3.1. Probability function $p_{\ell,\infty}(k, \mu_s)$ for an infinite non-absorbing homogeneous medium

As a preliminary result, in the Appendix C we derived the expression for the probability density function $f_k(\ell)$ for the length ℓ travelled by a photon at the k-th scattering event in an infinite medium. The $f_k(\ell)$ is somehow related to the $p_\ell(k, \mu_s)$, but they are two different functions. By following a procedure similar to that used in the Appendix C to obtain the function $f_k(\ell)$, one can derive the expression for the probability function $p_{\ell,\infty}(k, \mu_s)$ (discrete, with integer $k = 0, 1, 2, \ldots$) to have k scattering events within a fixed pathlength ℓ in an infinite medium. Therefore, this quantity is not bound to any detector, so that the contribution of all the propagating photons in the medium is considered.

The probability of travelling a length ℓ without undergoing scattering events is:

$$p_{\ell,\infty}(0, \mu_s) = e^{-\mu_s \ell} \ . \tag{12}$$

The probability of travelling a length ℓ undergoing only one scattering event can be obtained by calculating the probability that the first scattering event occurs after a length ℓ_1, $f_1(\ell_1)d\ell_1$ with $0 \le \ell_1 \le \ell$, and that no more scattering events occur till ℓ, $e^{-\mu_s(\ell-\ell_1)}$, and then considering all possible lengths ℓ_1:

$$p_{\ell,\infty}(1, \mu_s) = \int_0^\ell f_1(\ell_1) e^{-\mu_s(\ell-\ell_1)} d\ell_1 = \int_0^\ell \mu_s e^{-\mu_s \ell_1} e^{-\mu_s(\ell-\ell_1)} d\ell_1 =$$
$$= \mu_s e^{-\mu_s \ell} \int_0^\ell d\ell_1 = \mu_s \ell e^{-\mu_s \ell} \ . \tag{13}$$

Then, for a general value k of scattering events, we obtain:

$$p_{\ell,\infty}(k, \mu_s) = \int_0^\ell f_k(\ell') e^{-\mu_s(\ell-\ell')} d\ell' = \int_0^\ell \mu_s \frac{(\mu_s \ell')^{k-1}}{(k-1)!} e^{-\mu_s \ell'} e^{-\mu_s(\ell-\ell')} d\ell' =$$
$$= \mu_s^k \frac{e^{-\mu_s \ell}}{(k-1)!} \int_0^\ell (\ell')^{k-1} d\ell' = \mu_s^k \frac{e^{-\mu_s \ell}}{(k-1)!} \frac{\ell^k}{k} = \frac{(\mu_s \ell)^k}{k!} e^{-\mu_s \ell} \ . \tag{14}$$

The probability function $p_{\ell,\infty}(k, \mu_s)$ is thus a Poisson distribution with mean value $\lambda = \mu_s \ell$. From the properties of the Poisson distribution, we know that its standard deviation σ is equal to the square root of the mean value:

$$\sigma = \sqrt{\lambda} = \sqrt{\mu_s \ell} \ . \tag{15}$$

Furthermore, it is easy to verify, as expected, that the ratio between the probability that k scattering events occur in a trajectory of length ℓ in the infinite non-absorbing medium with scattering coefficient μ_s, and with scattering coefficient μ_{s0}, is:

$$\frac{p_{\ell,\infty}(k, \mu_s)}{p_{\ell,\infty}(k, \mu_{s0})} = \frac{e^{-\mu_s \ell}(\mu_s \ell)^k / k!}{e^{-\mu_{s0} \ell}(\mu_{s0} \ell)^k / k!} = \left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_s - \mu_{s0})\ell} \ , \tag{16}$$

that is identical to the scaling factor for the scattering coefficient for the trajectory received after a path ℓ with k scattering events reported in Eq. (5).

It is now interesting to calculate the scaling factor $F_{TPSF,s}(\mu_{s0} \to \mu_s, \ell)$ reported in Eq. (11) when $p_\ell(k, \mu_{s0})$ is replaced by $p_{\ell,\infty}(k, \mu_{s0})$. Here we stress that $p_{\ell,\infty}(k, \mu_{s0})$ [Eq. (14)] and $p_\ell(k, \mu_{s0})$ [Eq. (10)] are obtained in quite different situations. The former is obtained by following all the photons emitted in an infinite non-absorbing medium and determining the distribution of the number of scattering events (k) when a given pathlength ℓ is travelled. The latter aims to determine the same distribution in a finite medium by considering only detected photons (*i.e.*, usually photons that reach a designated area at the boundary of the medium). Despite the different origin of the two distributions, the use of $p_{\ell,\infty}(k, \mu_{s0})$ will be important (as shown the companion paper [17]) to devise a heuristic test of convergence for the scaled *TPSF*. After quite

straightforward calculations, it results:

$$
\begin{aligned}
F_{\infty,s}(\mu_{s0} \to \mu_s, \ell) &= e^{-(\mu_s - \mu_{s0})\ell} \sum_{k=0}^{\infty} p_{\ell,\infty}(k, \mu_{s0}) \left(\frac{\mu_s}{\mu_{s0}} \right)^k = \\
&= e^{-(\mu_s - \mu_{s0})\ell} \sum_{k=0}^{\infty} \frac{(\mu_{s0}\ell)^k}{k!} e^{-\mu_{s0}\ell} \left(\frac{\mu_s}{\mu_{s0}} \right)^k = \\
&= e^{-\mu_s \ell} \sum_{k=0}^{\infty} \frac{(\mu_{s0}\ell)^k}{k!} \left(\frac{\mu_s}{\mu_{s0}} \right)^k = \\
&= e^{-\mu_s \ell} \sum_{k=0}^{\infty} \frac{(\mu_s \ell)^k}{k!} = 1 \; .
\end{aligned}
\tag{17}
$$

Therefore, we underline that the scaling factor [Eq. (17)] is equal to one whatever the scattering jump is. An intuitive justification of this result can be obtained by noting that the probability distribution of the time-of-flights for the unconstrained propagation through an infinite medium is a constant distribution that does not vary with the scattering coefficient.

Considering that differences between *TPSFs* corresponding to different values of the scattering coefficient are usually relatively small, it can be deduced that $F_{TPSF,s}(\mu_{s0} \to \mu_s, \ell)$ assumes values close to one as well, thus yielding that the probability distribution $p_\ell(k, \mu_{s0})$ relative to a specific geometry does not differ much from $p_{\ell,\infty}(k, \mu_{s0})$. Nevertheless, these deviations, even if small, significantly determine the scaling factor. These statements can be supported by means of MC simulations, as will be shown in the companion paper [17].

We emphasize that the expressions described are valid only for the infinite non-absorbing homogeneous medium, whatever the scattering phase function is. To this extent, we note that the scattering phase function never appears in the presented formulas. As a matter of fact, the scattering phase function determines the shape of the trajectories of fixed length (*i.e.*, angles between the directions before and after each scattering event) and, therefore, their distribution in space, but it has no effect on the distribution of the trajectory lengths, to which all expressions described here refer.

For the sake of the reader in Appendix D the proofs of validity of the distribution $f_k(\ell)$ and of the probability function $p_{\ell,\infty}(k, \mu_{s0})$ are also showed by means of the induction method. Finally, it can be worth noticing that this general property of light propagation for the infinite medium described by the function $p_{\ell,\infty}(k, \mu_s)$ can be also derived from well-known results in renewal theory [20,21].

3.2. Probability function $p_\ell(k, \mu_s)$ for finite geometries

In the previous Section, we showed that in an infinite non-absorbing homogeneous medium the probability function $p_{\ell,\infty}(k, \mu_s)$ for a path of fixed length ℓ with k scattering events is described by a Poisson distribution. In this case, the trajectories can develop unconstrained, with the scattering phase function determining their shape, but not their length.

When instead we consider the trajectories received by a specific receiver at a distance ρ from the source and in a specific geometry, the probability function $p_\ell(k, \mu_s)$ will be different from $p_{\ell,\infty}(k, \mu_s)$. Consider, as an example, the extreme case for lengths ℓ smaller than the source-receiver distance, where there can be no received photons and $p_\ell(k, \mu_s)$ remains undetermined.

In the general case, the probability function $p_\ell(k, \mu_s)$ depends on μ_s, on the scattering phase function (which influences the probability of reception) and on the geometry, *i.e.*, on the volume of the considered medium, on the source-receiver distance ρ, on the features of the detector (area and field of view) and the refractive indices of the internal and external medium.

However, it is reasonable to think that in principle the Poisson distribution can still describe the $p_\ell(k, \mu_s)$ quite well, at least for photons received with pathlengths $\ell \gg \rho$, for which it is expected

that their diffusion-type propagation regime is not very different from the free propagation in the infinite non-absorbing medium. This ansatz, only intuitively introduced here and further investigated in the companion paper [17], is an additional assumption that links $p_\ell(k, \mu_s)$ to $p_{\ell,\infty}(k, \mu_s)$. The correctness of this assumption can be understood by noting that $p_\ell(k, \mu_{s0})$, for the case of no refractive index mismatch between medium and surroundings, does not need to refer to detected photons (*i.e.*, no fixed end point). We can construct this probability function based on equivalent trajectories having different end points inside the medium, linking it to the free propagation case.

On the contrary, when ℓ is slightly greater than ρ it is reasonable to expect significant differences, where the geometry of the problem can play an important role.

In practice, for a general geometry, when only the photons detected by a specific receiver have to be considered, the probability function $p_\ell(k, \mu_s)$ can be reconstructed by exploiting MC simulations, where for each photon received is recorded not only the trajectory length ℓ, but also the number of scattering events k it underwent. Probably, this is the only way to determine $p_\ell(k, \mu_s)$, even if one assumes the simplifying conditions of non-absorbing and homogeneous medium. We will call $p_{\ell,MC}(k, \mu_s)$ the probability function calculated in this way.

According to Eq. (11), the probability function $p_\ell(k, \mu_{s0})$ is necessary for calculating the scaling factor for scattering: therefore, in practical cases, we will be able to apply Eq. (11) only with reconstructed probability functions $p_{\ell,MC}(k, \mu_{s0})$ obtained from MC simulations. This means that the $p_{\ell,MC}(k, \mu_{s0})$, being calculated on a limited number of trajectories, is defined only on a limited range of values of k and is also affected by a statistical error. The uncertainty of the $p_{\ell,MC}(k, \mu_{s0})$ will therefore introduce an error in the scaling factor, which severely limits its applicability.

These considerations highlight the limitations of the practical application of scattering scaling relations, which are indeed not yet in of common use in MC simulations. They are useful only for small scattering variations and, of course, convergence may be even more difficult in non-homogeneous media. However, the computation of the error on the scaled result can always provide a criterion for establishing the level of significance of the outcomes obtained, as will be explained in the companion paper [17].

3.3. Convergence of the scaling factor for scattering

In order to understand the limit of applicability of Eq. (11) when the reconstructed probability function $p_{\ell,MC}(k, \mu_{s0})$ is used and, then, only a limited interval of values of k is available, we consider the following approximation for $p_{\ell,MC}(k, \mu_{s0})$:

$$
p_{\ell,MC}(k, \mu_{s0}) \approx
\begin{cases}
p_{\ell,\infty}(k, \mu_{s0}) = \dfrac{(\mu_{s0}\ell)^k}{k!} e^{-\mu_{s0}\ell} & \text{for } k_{min}(\mu_{s0}) \le k \le k_{max}(\mu_{s0}) \\
0 & \text{elsewhere}
\end{cases}
\tag{18}
$$

where k_{min} and k_{max} are the extremes of the k interval where the $p_{\ell,MC}(k, \mu_{s0})$, calculated by means of the reference MC simulation for μ_{s0}, is defined. In assuming Eq. (18), we took into account that the probability function $p_{\ell,MC}(k, \mu_{s0})$ for $\ell \gg \rho$ (see examples in the companion paper [17]) is always in good agreement with the corresponding $p_{\ell,\infty}(k, \mu_{s0})$.

By adopting Eq. (18), the *TPSF* scaling factor for scattering reported in Eq. (11) becomes:

$$F_{MC,s}(\mu_{s0} \to \mu_s, \ell) = e^{-(\mu_s-\mu_{s0})\ell} \sum_{k=0}^{\infty} p_{\ell,MC}(k, \mu_{s0}) \left(\frac{\mu_s}{\mu_{s0}}\right)^k =$$

$$\approx e^{-(\mu_s-\mu_{s0})\ell} \sum_{k=k_{min}}^{k_{max}} \frac{(\mu_{s0}\ell)^k}{k!} e^{-\mu_{s0}\ell} \left(\frac{\mu_s}{\mu_{s0}}\right)^k = \tag{19}$$

$$= \sum_{k=k_{min}}^{k_{max}} \frac{(\mu_s\ell)^k}{k!} e^{-\mu_s\ell} = \sum_{k=k_{min}}^{k_{max}} p_{\ell,\infty}(k, \mu_s) \,.$$

The *TPSF* scaling factor for scattering is therefore obtained by summing the probability function $p_{\ell,\infty}(k, \mu_s)$ related to the medium with the new scattering coefficient μ_s, between $k_{min}(\mu_{s0})$ and $k_{max}(\mu_{s0})$ calculated for the medium with the initial scattering coefficient μ_{s0}.

The main problems of convergence for the scaling factor $F_{MC}(\mu_{s0} \to \mu_s, \ell)$ reported in Eq. (19) derive from having limited the summation between $k_{min}(\mu_{s0})$ and $k_{max}(\mu_{s0})$, and only secondarily from having approximated $p_{\ell,MC}(k, \mu_{s0})$ with $p_{\ell,\infty}(k, \mu_{s0})$. In particular, the problems will be more relevant when the curves of $p_{\ell,\infty}(k, \mu_{s0})$ and $p_{\ell,\infty}(k, \mu_s)$, relating to the scattering coefficients μ_{s0} and μ_s, respectively, are very spaced apart, since in this case the range of k values where the summation is performed is far from the maximum of $p_{\ell,\infty}(k, \mu_s)$, being $k_{min}(\mu_{s0})$ and $k_{max}(\mu_{s0})$ centered around the maximum of $p_{\ell,\infty}(k, \mu_{s0})$. It is worth noting that the effect of this *truncation* of the $p_{\ell,\infty}(k, \mu_s)$ always involves an underestimation of the scaling factor, whether one scales to larger or smaller μ_s values with respect to μ_{s0}.

In order to make clearer this statement, in Figs. 2 and 3 are compared for $\mu_{s0} = 10$ cm^{-1} and $\mu_{s0} = 50$ cm^{-1} the probability functions $p_{\ell,\infty}(k, \mu_{s0})$ and $p_{\ell,\infty}(k, \mu_s)$ when the ratio μ_s/μ_{s0} is set to 1.1. We can note the progressive separation of $p_{\ell,\infty}(k, \mu_{s0})$ and $p_{\ell,\infty}(k, \mu_s)$ as ℓ increases.

In particular, in Fig. 2, where $\mu_{s0} = 10$ cm^{-1}, we have almost overlapping curves for $\ell = 2.5$ cm (compare Fig. 2(a)), while there is an overlap of less than 50% for $\ell = 80$ cm (see Fig. 2(d)). On the other hand, in Fig. 3, where $\mu_{s0} = 50$ cm^{-1}, a variation of the scattering coefficient by 10%, as considered here, leads to an overlap between $p_{\ell,\infty}(k, \mu_{s0})$ and $p_{\ell,\infty}(k, \mu_s)$ of about 50% already for $\ell = 10$ cm (see Fig. 3(b)), with a rapid decrease for longer pathlengths.

The above figures and Eq. (19) suggest that the real origin of the possible failure of scaling the weights [Eq. (5)] for generating a new *TPSF* due to a change in the scattering coefficient is related to the fact that the trajectories of the initial simulation may probe the investigated medium for a range of scattering events $[k_{min}(\mu_{s0}), k_{max}(\mu_{s0})]$ too off-center with respect to the actual distribution $p_\ell(k, \mu_s)$ of scattering events with the new scattering coefficient. Thus, it is possible to make this statement only on the basis of Eq. (10) and of the ansatz that links the shape of the function $p_{\ell,MC}(k, \mu_s)$ to $p_{\ell,\infty}(k, \mu_s)$. This observation is important in terms of comprehension of the convergence mechanism of the scaling process.

Finally, because scattering variations usually involve *moderate* variations both on the *TPSF* and on the CW, we can consider a strong difference between F_{MC} reported in Eq. (19) and the value 1 as an index of the error due to the truncation of the $p_{\ell,MC}(k, \mu_{s0})$. Thus, the calculation of Eq. (19) offers an heuristic test that can be performed to evaluate whether the estimated $p_{\ell,MC}(k, \mu_{s0})$ has reached a sufficient level of convergence and consequently the calculation of the scattering scaling factor $F_{MC,s}$ reconstructed with MC simulations is reliable. This test can be used to evaluate whether we have a sufficient level of convergence of the scaled MC results. Its failure implies that the range $[k_{min}(\mu_{s0}), k_{max}(\mu_{s0})]$ probed by the selected trajectories in the initial state is too far from the range where $p_\ell(k, \mu_s)$ is significantly different from zero. The results of this test show convergence only for moderate variations of the scattering properties from an initial to a final state for which we have moderate variations on the *TPSF* and on the CW.
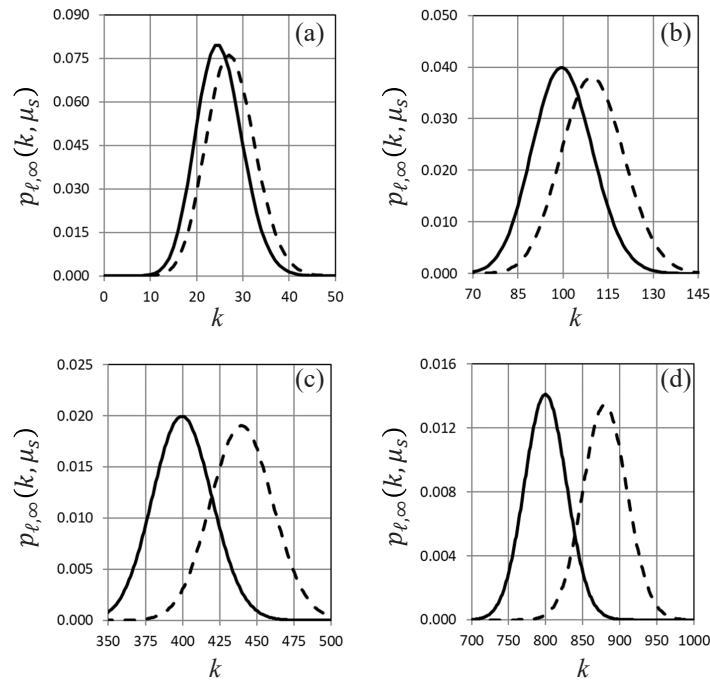
Fig. 2. Plot of $p_{\ell,\infty}(k, \mu_{s0})$ (solid line) and $p_{\ell,\infty}(k, \mu_s)$ (dashed line) for $\mu_{s0} = 10$ cm^{-1} and $\mu_s = 11$ cm^{-1}, respectively, are reported in each panel as a function of the number of scattering events k, for different values of the pathlength: $\ell = 2.5$ cm, panel (a); $\ell = 10$ cm, panel (b); $\ell = 40$ cm, panel (c); $\ell = 80$ cm, panel (d).
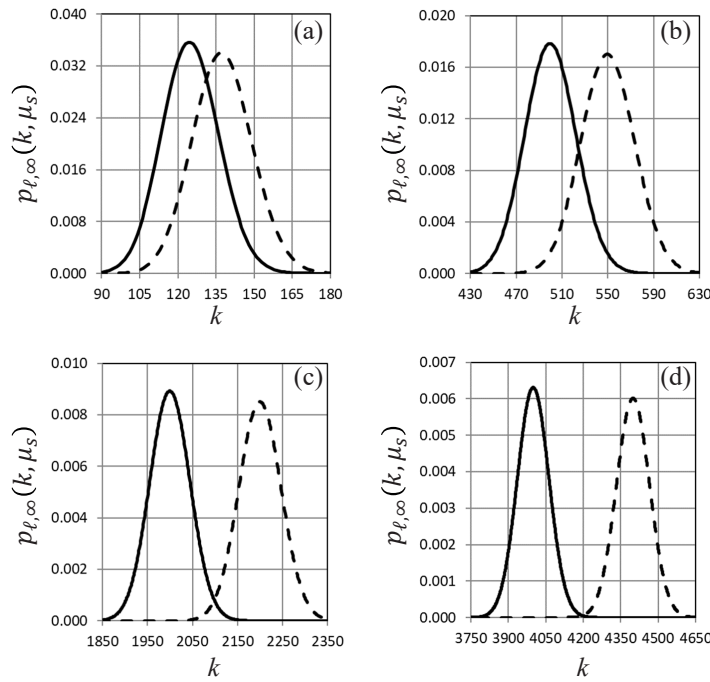
Fig. 3. As in Fig. 2, but for $\mu_{s0} = 50$ cm^{-1} and $\mu_s = 55$ cm^{-1}: $\ell = 2.5$ cm, panel (a); $\ell = 10$ cm, panel (b); $\ell = 40$ cm, panel (c); $\ell = 80$ cm, panel (d).

Details and examples of this test will be provided in the companion paper following this work
[17].

4. Derivatives

The use of an MC forward solver in the inverse problem [9] is subject to the task of evaluating
the derivatives with respect to the absorption and scattering properties of the medium. In this
perspective, it is important to evaluate the convergence properties of the reconstructed derivatives.
In this Section, a general analysis will be made to derive useful formulas for the inverse problems.
The goal will be to provide general guidelines for evaluating the convergence of the derivatives
obtained using the scaling relations introduced in the previous Sections. A validation of the
formulas presented for calculating the derivatives will be provided in the in the companion paper
[17]. By means of MC simulations the calculations with these formulae will be carried out and
verified.

Let us consider a homogeneous medium and, using Eq. (3), the *TPSF* is obtained by classifying
each trajectory according to time-of-flight with the appropriate normalization factor. In particular,
the *TPSF* at the i-th time bin is:

$$TPSF\left(\mu_a, \mu_s, t_i\right) = \frac{\displaystyle\sum_{j \in J_i} W_j\left(\mu_a, \mu_s\right)}{N_e S \Delta t_i} \;, \tag{20}$$

where J_i is the ensemble of photons received in the i-th time bin, N_e is the number of launched
photons, S is the detector area, and Δt_i is the width of the i-th time bin. To the i-th time bin contribute
only photons having run trajectories with length ℓ_j such that $(t_i - \Delta t_i/2)\, v < \ell_j < (t_i + \Delta t_i/2)\, v$, where
v is the speed of light in the medium.

To derive the CW response from these data it is enough to sum all the weights and normalize:

$$CW(\mu_a, \mu_s) = \frac{\displaystyle\sum_{j=1}^{N} W_j\left(\mu_a, \mu_s\right)}{N_e S} \;, \tag{21}$$

where N is the total number of received photons.

To derive the change of the *TPSF* when the absorption or scattering coefficient changes, we
need first to calculate the partial derivatives of Eq. (3) with respect to μ_a and μ_s:

$$\frac{\partial W\left(\partial \mu_a, \mu_s\right)}{\mu_a} = (-\ell)\, W\left(\mu_a, \mu_s\right) \;, \tag{22}$$

$$\frac{\partial W\left(\partial \mu_a, \mu_s\right)}{\mu_s} = \left(\frac{k}{\mu_s} - \ell\right) W\left(\mu_a, \mu_s\right) \;. \tag{23}$$

Considering that, as discussed in Section 2.3, the scaling on μ_s converges for very small
perturbations, Eq. (23) will be useful when the perturbation is calculated for $\mu_s = \mu_{s0}$, where
μ_{s0} is the scattering coefficient of the initial MC simulation. Instead, μ_a can assume any value
because the scaling relation for it always converges. Thus, Eqs. (22, 23) become:

$$\left.\frac{\partial W\left(\partial \mu_a, \mu_s\right)}{\mu_a}\right|_{\mu_s = \mu_{s0}} = (-\ell)\, \partial W\left(\partial \mu_a, \mu_{s0}\right) \;, \tag{24}$$

$$\left.\frac{\partial W\left(\partial \mu_a, \mu_s\right)}{\mu_s}\right|_{\mu_s = \mu_{s0}} = \left(\frac{k}{\mu_{s0}} - \ell\right) \partial W\left(\partial \mu_a, \mu_{s0}\right) \;. \tag{25}$$

We can now derive the expression of the derivative of the *TPSF* with respect to μ_s and μ_a. Starting from Eq. (20) we obtain:

$$\left.\frac{\partial TPSF(\partial\mu_a, \mu_s, t_i)}{\mu_a}\right|_{\mu_s=\mu_{s0}} = \frac{\sum\limits_{j\in J_i} \left.\frac{\partial W_j(\partial\mu_a, \mu_s)}{\mu_a}\right|_{\mu_s=\mu_{s0}}}{N_e S \Delta t_i} =$$

$$= \frac{\sum\limits_{j\in J_i} W_j(\mu_a, \mu_{s0})(-\ell_j)}{N_e S \Delta t_i} \ , \tag{26}$$

$$\left.\frac{\partial TPSF(\partial\mu_a, \mu_s, t_i)}{\mu_s}\right|_{\mu_s=\mu_{s0}} = \frac{\sum\limits_{j\in J_i} \left.\frac{\partial W_j(\partial\mu_a, \mu_s)}{\mu_s}\right|_{\mu_s=\mu_{s0}}}{N_e S \Delta t_i} =$$

$$= \frac{\sum\limits_{j\in J_i} W_j(\mu_a, \mu_{s0})\left(\frac{k_j}{\mu_{s0}} - \ell_j\right)}{N_e S \Delta t_i} \ . \tag{27}$$

We observe that the numerator of Eq. (27) is proportional to the difference between the average number of scattering events multiplied by the mean free path $1/\mu_{s0}$ and the average length of the trajectories of the detected photons falling in the *i*-th temporal bin. Analogously, the numerator of Eq. (26) is proportional to only the average length. If we multiply and divide both equations for the quantity $\sum_{j=1}^{N_i} W_j(\mu_{s0}, \mu_a)$, by using Eq. (20), we obtain:

$$\left.\frac{\partial TPSF(\partial\mu_a, \mu_s, t_i)}{\mu_a}\right|_{\mu_s=\mu_{s0}} = -\langle\ell\rangle_i TPSF(\mu_a, \mu_{s0}, t_i) \ , \tag{28}$$

$$\left.\frac{\partial TPSF}{\partial\mu_s}\right|_{\mu_s=\mu_{sz}} = \langle\frac{k}{\mu_s z} - \ell\rangle_i TPSF(\mu_a, \mu_s 0, t_i) \tag{29}$$

where:

$$\langle\ell\rangle_i = \frac{\sum\limits_{j\in J_i} W_j(\mu_a, \mu_{s0})\ell_j}{\sum\limits_{j\in J_i} W_j(\mu_a, \mu_{s0})} \ , \tag{30}$$

$$\langle\frac{k}{\mu_{s0}} - \ell\rangle_i = \frac{\sum\limits_{j\in J_i} W_j(\mu_a, \mu_{s0})\left(\frac{k_j}{\mu_{s0}} - \ell_j\right)}{\sum\limits_{j\in J_i}^{N_i} W_j(\mu_a, \mu_{s0})} \ . \tag{31}$$

We observe that, for a sufficiently small value of Δt, $\ell_j \approx v t_i$, for $j \in J_i$. Then, $\langle\ell\rangle_i \approx v t_i$, meaning that the calculation of the perturbation with respect to μ_a is quite robust since the corresponding relative error on the mean is practically negligible.

For what concerns the perturbation with respect to μ_s, the calculation of the quantity $\langle k/\mu_{s0}\rangle_i$ can be done by exploiting the probability function $p_\ell(k,\mu_s)$:

$$\langle \frac{k}{\mu_{s0}}\rangle_i = \frac{1}{\mu_{s0}} \sum_{k=0}^{\infty} p_{\ell_i}(k,\mu_{s0})k \,, \tag{32}$$

where we set $\ell_i = vt_i$. In particular, in the special case of an infinite medium and for all the propagating photons, we have to consider the Poisson distribution Eq. (14) for $p_{\ell_i}(k,\mu_{s0})$:

$$\langle \frac{k}{\mu_{s0}}\rangle_i = \frac{1}{\mu_{s0}} \sum_{k=0}^{\infty} p_{\ell_i,\infty}(k,\mu_{s0})k = \frac{\langle k\rangle}{\mu_{s0}} = \frac{\mu_{s0}\ell_i}{\mu_{s0}} = \ell_i \,, \tag{33}$$

for the properties of the Poisson distribution. Then, in the case of infinite medium $\langle k/\mu_{s0} - \ell\rangle_i = \ell_i - \ell_i = 0$ for every time bin and the derivative with respect μ_s is nil. In general, the derivative with respect μ_s depends on the small differences existing between the probability function $p_\ell(k,\mu_s)$ and the correspondent Poisson distribution. As we will see in the companion paper [17], these differences exist only at short pathlengths ℓ, meaning that the scattering properties predominantly affect the short times of the *TPSF*, where the derivative with respect to μ_s is not zero.

In practical cases, for a general geometry, the mean values reported in Eq. (31) are calculated by exploiting the MC simulation for μ_{s0}. However, the smaller is Δt, the smaller is the number of photon trajectories falling in the i-th time bin, thus making the quantity $\langle k/\mu_{s0}\rangle_i$ noisier. Consequently, we expect a slower convergence of $\langle k/\mu_{s0} - \ell\rangle_i$ with respect to $\langle \ell\rangle_i$, but still faster with respect the scaling factor $(\mu_s/\mu_{s0})^k e^{-(\mu_s-\mu_{s0})\ell}$ (see Section 2.3).

The same procedure can be applied to Eq. (21) for the CW, bringing to similar expressions:

$$\frac{\partial CW(\partial\mu_a,\mu_s)}{\mu_a}\bigg|_{\mu_s=\mu_{s0}} = -\langle \ell\rangle CW(\mu_a,\mu_{s0})\,, \tag{34}$$

$$\frac{\partial CW(\partial\mu_a,\mu_s)}{\mu_s}\bigg|_{\mu_s=\mu_{s0}} = \langle \frac{k}{\mu_{s0}} - \ell\rangle CW(\mu_a,\mu_{s0})\,, \tag{35}$$

where the average values are calculated using all the received photons. It is worth noticing that also for the CW, in the special case of an infinite medium, for all the propagating photons, the derivative with respect to μ_s is nil.

In conclusion, this set of formulae provide a direct way to calculate the derivative of the MC forward solver without resorting to additional simulations except the initial background simulation. In the companion paper [17] the results obtained with these formulae will be validated in real MC calculations.

5. Conclusion

In conclusion, we have revised the method of the scaling relations to derive a new MC simulation with a certain set of optical properties from an existing MC simulation obtained with a different set. The method is *trajectory-based* and recalculates the new MC simulation by changing the weights of each trajectory of the original simulation based on the new optical properties. Clearly, the reliability of the new MC simulation depends on how well the original trajectory ensemble probes also the new scenario.

In the first Section, we revised the basic scaling formulas for both scattering and absorption homogeneous media, having extended them to the case of non-homogeneous media in the Appendix A, assuming the same scattering phase function. For absorption, the scaling factor for each trajectory is simply the Lambert-Beer term $e^{-(\mu_{aq}-\mu_{a0q})\ell_q}$ applied to the stretches ℓ_q on the subdomains V_q. This assures a smooth and easy convergence of the scaling relations in

the case of absorption. Conversely, for scattering the weights are the product of two terms: $\left(\mu_{sq}/\mu_{s0q}\right)^{k_q} e^{-(\mu_{sq}-\mu_{s0q})\ell_q}$. The strong divergence of these terms in opposite directions (Table 1) is the ultimate cause of strong convergence instability for the scattering scaling relations.

In Section 3, we dissected the genesis of the instability in the scattering scaling factor by providing original insight into the probability function $p_\ell(k, \mu_s)$ for the number of scattering events k undergone by a photon along a path of length ℓ. For a *free propagation* in an infinite non-absorbing homogeneous medium, $p_\ell(k, \mu_s)$ converges to a Poisson distribution $p_{\ell,\infty}(k, \mu_s)$. By making an ansatz that links $p_\ell(k, \mu_s)$ to $p_{\ell,\infty}(k, \mu_s)$ even for finite geometries, we provides a key to understanding the convergence mechanism of the scattering scaling process. A change in scattering causes a relevant shift of $p_{\ell,\infty}(k, \mu_s)$ (see Fig. 3), leading to a dramatic reduction of the intersection of the two distributions $p_\ell(k, \mu_s)$ and $p_\ell(k, \mu_{s0})$, *i.e.*, in the number of *useful* trajectories, leaving largely under-sampled trajectories in particular for a long pathlength ℓ. This behavior is at the basis of the possible failure of convergence of the scattering scaling relations.

One of the purposes of this paper was to provide new tools for a deeper understanding of the perturbation approach in the MC technique, with the aim of also giving quantitative tools for evaluating its convergence. In this paper we have been primarily focused on providing a comprehensive rereading of the method to improve understanding of the statistical mechanisms that determine its convergence. In doing so, we have highlighted the possible implementation, through Eq. (19), of a heuristic test that can be performed to evaluate whether the calculation of the scattering scaling factor $F_{MC,s}$ reconstructed with MC simulations is reliable. This test exploits the weak deviations existing between the distribution $p_{\ell,\infty}(k, \mu_{s0})$, described by a Poisson function, and $p_{\ell,MC}(k, \mu_{s0})$. The proposed test is to our knowledge original and represents a simple and quick method capable of giving a preliminary assessment of the convergence of scaling rules for scattering used in perturbation Monte Carlo simulations. Details on the implementation of this test will be provided in the companion paper [17].

In the final Section 4, the formulas for a small change in optical properties (*e.g.* to calculate the derivative in an inverse problem) are discussed both for the TD and for the CW case. The scaling factors for the weights of the derivatives are related to the pathlength ℓ for absorption and the difference $(k/\mu_{s0} - \ell)$, between the mean pathlength calculated from the number k of scattering events and the actual length of the trajectory ℓ, for scattering. Again, the probability function for number of scattering events $p_\ell(k, \mu_s)$ plays an important role. However, in this case the calculation of the scattering derivative does not involve changes of μ_s. Thus, there is no shift in the probability distribution $p_\ell(k, \mu_s)$ due to a scattering change and trajectories are well sampled. Therefore, the only contribution that can produce a nonzero derivative with respect to μ_s and a direction for the next iteration in an inverse problem is the slight discrepancy of $p_\ell(k, \mu_s)$ with respect to a Poisson distribution. These differences exist only at short pathlengths ℓ, meaning that the scattering perturbation predominantly affects the short times of the *TPSF*, where the derivative with respect to μ_s is not zero.

The present paper is going to be complemented by a companion paper [17] where practical implementations and convergence properties of the scaling relations are discussed in different cases based on extensive numerical simulations. Overall, this work can be useful to understand the basic physics behind the scaling relations, analyzes the ultimate causes of fragility and validity for applications, and also conceive new strategies to describe and decipher photon migration in scattering media.

Appendix A. Non-homogeneous media

The scaling methods can be easily retrieved also for non-homogeneous media (*e.g.*, layered slab or medium with inclusion). However, in this case, the MC simulation is performed on Q subvolumes V_q with different optical properties $(\mu_{aq}, \mu_{sq}, p_q)$. In this case, the discontinuity surfaces between media with different indices of refraction need to be considered. Following the steps reported in

Section 2.1, the fraction of radiation collected at the detector after k scattering events dP_{sk} [see Eq. (1)], should be multiplied by the reflection $R(\theta_i)$ or transmission $T(\theta_i)$ coefficient. These two coefficients do not depend on the optical properties of the medium, but only on the incidence angle of the photons θ_i. Thus, they cancel out in Eq. (3), and the scaling factors are not affected by discontinuity surface generated by refractive indices mismatch.

The optical properties for the complete set of subvolumes is expressed using the vector notation $(\boldsymbol{\mu_a}, \boldsymbol{\mu_s}, \boldsymbol{p})$, with:

$$\boldsymbol{\mu_a} = \begin{bmatrix} \mu_{a1} & \mu_{a1} & \cdots & \mu_{aQ} \end{bmatrix} , \tag{36}$$

$$\boldsymbol{\mu_s} = \begin{bmatrix} \mu_{s1} & \mu_{s2} & \cdots & \mu_{sQ} \end{bmatrix} , \tag{37}$$

$$\boldsymbol{p} = \begin{bmatrix} p_1 & p_2 & \cdots & p_Q \end{bmatrix} . \tag{38}$$

Thus, if we consider $(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_{s0}}, \boldsymbol{p_0})$ the distribution of optical properties of the starting MC simulation, and $(\boldsymbol{\mu_a}, \boldsymbol{\mu_s}, \boldsymbol{p})$ the ones in the new configuration in each subvolume, the weight of each trajectory is defined as follows:

$$W(\boldsymbol{\mu_a}, \boldsymbol{\mu_s}, \boldsymbol{p}) = W(\boldsymbol{\mu_{az}}, \boldsymbol{\mu_{sz}}, \boldsymbol{p_0}) \times$$

$$\times \prod_{q=1}^{Q} \left\{ \left(\frac{\mu_{sq}}{\mu_{s0q}} \right)^{k_q} e^{-(\mu_{sq}-\mu_{s0q})\ell_q} e^{-(\mu_{aq}-\mu_{a0q})\ell_q} \prod_{m=1}^{k_q} \left(\frac{p_q(\theta_{qm})}{p_{0q}(\theta_{qm})} \right) \right\} , \tag{39}$$

where $W(\boldsymbol{\mu_a}, \boldsymbol{\mu_s}, \boldsymbol{p})$ and $W(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_{s0}}, \boldsymbol{p_0})$ are the weights in the final and initial state, and k_q and ℓ_q are the number of scattering events and the total pathlength travelled within the volume V_q, respectively.

As for the homogeneous medium, a simplified expression can be obtained in case the same scattering function is assumed inside the whole medium. If we consider (μ_{a0q}, μ_{s0q}) the distribution of absorption and scattering coefficients of the starting MC simulation, and (μ_{aq}, μ_{sq}) the ones in the new configuration for each subvolume, the weight of each trajectory is defined as follows:

$$W(\boldsymbol{\mu_a}, \boldsymbol{\mu_s}) = W(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_{s0}}) \prod_{q=1}^{Q} \left\{ \left(\frac{\mu_{sq}}{\mu_{s0q}} \right)^{k_q} e^{-(\mu_{sq}-\mu_{s0q})\ell_q} e^{-(\mu_{aq}-\mu_{a0q})\ell_q} \right\} . \tag{40}$$

Starting from Eq. (40), the scaling factors for absorption and scattering in a non-homogeneous medium can be easily retrieved:

$$F_a = \frac{W(\boldsymbol{\mu_a}, \boldsymbol{\mu_{s0}})}{W(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_{s0}})} = \prod_{q=1}^{Q} e^{-(\mu_{aq}-\mu_{a0q})\ell_q} , \tag{41}$$

$$F_s = \frac{W(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_s})}{W(\boldsymbol{\mu_{a0}}, \boldsymbol{\mu_{s0}})} = \prod_{q=1}^{Q} \left(\frac{\mu_{sq}}{\mu_{s0q}} \right)^{k_q} e^{-(\mu_{sq}-\mu_{s0q})\ell_q} . \tag{42}$$

The scaling factors depend on the attenuation undergone in each subvolume, so that photons with paths of identical length may have very different scaling factors. Even for a fixed time-of-flight, there is a dispersion of weights, leading to an increase of the relative error with respect to the homogeneous case.

Moreover, by starting from Eq. (40), we obtain the expressions for the partial derivatives of $W(\boldsymbol{\mu_a}, \boldsymbol{\mu_s})$ with respect to μ_{aq} and μ_{sq}:

$$\left. \frac{\partial W(\partial\boldsymbol{\mu_a}, \boldsymbol{\mu_s})}{\mu_{aq}} \right|_{\boldsymbol{\mu_s}=\boldsymbol{\mu_{s0}}} = \left(-\ell_q\right) W(\boldsymbol{\mu_a}, \boldsymbol{\mu_{s0}}) \ , \tag{43}$$

$$\left. \frac{\partial W(\partial\boldsymbol{\mu_a}, \boldsymbol{\mu_s})}{\mu_{sq}} \right|_{\boldsymbol{\mu_s}=\boldsymbol{\mu_{s0}}} = \left(\frac{k_q}{\mu_{s0q}} - \ell_q\right) W(\boldsymbol{\mu_a}, \boldsymbol{\mu_{s0}}) \ , \tag{44}$$

where the subscript q refers to the q-th region or voxel.

Appendix B. Probability of detecting a trajectory

In this Appendix we derive the general expression of the probability of detecting a trajectory when the optical properties are changed from an initial state to a final state. We define a trajectory Γ in the medium (see Fig. 1) a set of pairs:

$$\Gamma \equiv (\ell_1, \theta_{in} = 0), (\ell_2, \theta_1), (\ell_3, \theta_2), \ldots, (\ell_{fin}, \theta_k) \ , \tag{45}$$

where ℓ_m is the path travelled by the photon before being scattered at an angle θ_m within a solid angle $d\Omega_m$, with $m = 1, \ldots, k$ being k the number of scattering events of the trajectory, while ℓ_{fin} is the path travelled from the last scattering event to the trajectory end point. We assume that a trajectory starts from a source point and ends either inside the medium or at the boundary when the photon exits the medium. Furthermore, we assume that the injection angle θ_{in} is fixed, setting $\theta_{in} = 0$. A detector D of surface S is a subset of the boundary where photon are received. In a homogeneous medium with fixed μ_a and μ_s we can define an infinitesimal probability to travel along Γ as [compare Eq. (1)]:

$$dP(\Gamma, \mu_a, \mu_s, p) = \mu_s^k e^{-\mu_t \ell} \prod_{m=1}^{k} d\ell_m p(\theta_m) d\Omega_m \ . \tag{46}$$

Strictly speaking Eq. (46) is valid only for $k \geq 1$. In the case the photon goes straight-forwardly from source to detector, *i.e.*, with no scattering events, it results $\ell_{fin} = \ell$ and the probability for such a trajectory is $e^{-\mu_t \ell}$ [compare Eq. (12)]. Then, Eq. (46) can be extended to the case $k = 0$ as:

$$dP(\Gamma, \mu_a, \mu_s, p) = \begin{cases} \mu_s^k e^{-\mu_t \ell} \displaystyle\prod_{m=1}^{k} d\ell_m p(\theta_m) d\Omega_m & \text{for } k \geq 1 \\[2ex] e^{-\mu_t \ell} & \text{for } k = 0 \end{cases} \tag{47}$$

We note that Eq. (47) depends on the optical properties, on the number of scattering events, on the total path length and the set of scattering angles. A different trajectory with the same parameters will be characterized by the same probability. Even here we can introduce the concept of equivalent trajectories (see Section 2.1) that are characterized by the same probability.

Given these premises, we can define the probability to detect a photon within a time range $(t_i - \Delta t/2, t_i + \Delta t/2)$ as:

$$TPSF(t_i, \Delta t, \mu_a, \mu_s) \Delta t S = \int_T dP(\Gamma, \mu_a, \mu_s, p) \ . \tag{48}$$

In Eq. (48) the *TPSF* is the well-known temporal point spread function and $T = T(\Gamma, t_i, \Delta t)$ is the non-numerable infinite ensemble of all the possible trajectories that starts at the source point and ends at the detector in the given time interval. If Δt is small enough, it results that $T \approx T(\Gamma, \ell)$, with $\ell = vt_i$.

We do not address the topic of the probability normalization in the space of all the possible trajectories, giving for granted that is possible to define one. Assume to have two media, one which we define *initial* or *background* medium, with the optical properties μ_{a0}, μ_{s0}, p_0, and the other which is a *final* or *new* medium, characterized by different optical properties μ_a, μ_s, p. The heart of the scaling relation method is to define an estimate of the probability to travel along a given trajectory in the new medium, when an estimate of the same probability is calculated from the background medium. Typically, the estimate in the background medium is calculated with a MC simulation.

If we consider the further simplifying assumption $p(\theta) = p_0(\theta)$, from Eq. (47) it results:

$$\frac{\mathrm{d}P(\Gamma, \mu_a, \mu_s, p)}{\mathrm{d}P(\Gamma, \mu_{a0}, \mu_{s0}, p_0)} = \left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_t - \mu_{t0})\ell} , \tag{49}$$

which is valid for $k \geq 0$. In this derivation we use only the physical and statistical concepts underlying the definition of the optical properties of a medium.

Now, if we substitute the infinitesimal probability to travel along Γ given by Eq. (49) into Eq. (48), we can write:

$$TPSF(t_i, \Delta t, \mu_a, \mu_s)\Delta t S = \int_T \left(\frac{\mu_s}{\mu_{s0}}\right)^k e^{-(\mu_t - \mu_{t0})\ell} \mathrm{d}P(\Gamma, \mu_{a0}, \mu_{s0}, p_0) . \tag{50}$$

In order to manage the number of scattering events k, we can write the trajectory ensemble $T(\Gamma, \ell)$ as the union of disjoint sub-ensembles $T_k(\Gamma, \ell)$ containing trajectories with a defined number k of scattering events:

$$T(\Gamma, \ell) = \bigcup_{k=0}^{\infty} T_k(\Gamma, \ell) . \tag{51}$$

We explicitly note that, since the injection angle is fixed, as we assumed in (45) setting $\theta_{in} = 0$, the sub-ensemble $T_0(\Gamma, \ell)$ contains at most one trajectory, that is the straight-line that connects the source to the detector.

In this framework, we can define the probability function $p_\ell(k, \mu_s)$ for the number of scattering events k undergone by a photon along a path of length ℓ as:

$$p_\ell(k, \mu_s) \equiv \frac{\int_{T_k} \mathrm{d}P(\Gamma, \mu_a, \mu_s, p)}{\int_T \mathrm{d}P(\Gamma, \mu_a, \mu_s, p)} = \frac{\int_{T_k} \mathrm{d}P(\Gamma, \mu_a, \mu_s, p)}{\sum_{k'=0}^{\infty} \int_{T_{k'}} \mathrm{d}P(\Gamma, \mu_a, \mu_s, p)} =$$

$$= \begin{cases} \dfrac{\mu_s^k \int_{T_k} \prod_{m=1}^{k} \mathrm{d}\ell_m p(\theta_m) \mathrm{d}\Omega_m}{\delta_{T_0} + \sum_{k'=1}^{\infty} \mu_s^{k'} \int_{T_{k'}} \prod_{m=1}^{k'} \mathrm{d}\ell_m p(\theta_m) \mathrm{d}\Omega_m} & \text{for } k \geq 1 \\[4mm] \dfrac{\delta_{T_0}}{\delta_{T_0} + \sum_{k'=1}^{\infty} \mu_s^{k'} \int_{T_{k'}} \prod_{m=1}^{k'} \mathrm{d}\ell_m p(\theta_m) \mathrm{d}\Omega_m} & \text{for } k = 0 \end{cases} \tag{52}$$

where δ_{T_0} is equal to 1 if T_0 is not empty and 0 otherwise. We exploited Eq. (47) in the derivation of the last expression for the probability function $p_\ell(k, \mu_s)$ reported in Eq. (52), making explicit the fact that it does not depend on the absorption coefficient μ_a.

Now, we we can develop the calculation of the *TPSF* reported in Eq. (50):

$$
\begin{aligned}
TPSF(t_i, \Delta t, \mu_a, \mu_s)\Delta t S &= \int_T \left(\frac{\mu_s}{\mu_s z}\right)^k e^{-(\mu_t - \mu_{t0})\ell} \, dP(\Gamma, \mu_{a0}, \mu_{s0}, p_0) = \\
&= e^{-(\mu_t - \mu_{t0})\ell} \sum_{k=0}^{\infty} \left(\frac{\mu_s}{\mu_s z}\right)^k \int_{T_k} dP(\Gamma, \mu_{a0}, \mu_{s0}, p_0) = \\
&= e^{-(\mu_t - \mu_{t0})\ell} \sum_{k=0}^{\infty} \left(\frac{\mu_s}{\mu_s z}\right)^k p_\ell(k, \mu_{s0}) \int_T dP(\Gamma, \mu_{a0}, \mu_{s0}, p_0) = \\
&= e^{-(\mu_t - \mu_{t0})\ell} \sum_{k=0}^{\infty} \left(\frac{\mu_s}{\mu_s z}\right)^k p_\ell(k, \mu_{s0}) TPSF(t_i, \Delta t, \mu_a z, \mu_s z)\Delta t S .
\end{aligned}
\tag{53}
$$

Equation (53) fully justifies Eq. (10) and the consequent expression (11) assumed for the scattering scaling factor.

The expression for the probability function $p_\ell(k, \mu_s)$ reported in Eq. (52) can be explicitly calculate in the case of *free propagation*, *i.e.*, when an infinite homogeneous medium and no explicit detector are considered. In this case, we have no constraints on the scattering angles θ_m, while the lengths ℓ_m, travelled between two scattering events, should satisfy:

$$
\sum_{m=1}^{k} \ell_m \leq \ell ,
\tag{54}
$$

because the total length $\sum_{m=1}^{k} \ell_m + \ell_{fin}$ of the trajectory is ℓ and the final path ℓ_{fin} can assume any value between 0 and ℓ.

Now, the integrals present in Eq. (52) can be calculated:

$$
\begin{aligned}
\int_{T_k} \prod_{m=1}^{k} d\ell_m p(\theta_m) d\Omega_m &= \int_{\sum_{m=1}^{k} \ell_m \leq \ell} \prod_{m=1}^{k} d\ell_m \prod_{m=1}^{k} \int_{4\pi} p(\theta_m) d\Omega_m = \\
&= \int_{\sum_{m=1}^{k} \ell_m \leq \ell} \prod_{m=1}^{k} d\ell_m ,
\end{aligned}
\tag{55}
$$

where we exploited the normalization property of the phase function $p(\theta)$ over the whole solid angle.

In order to calculate the last integral in Eq. (55), we set:

$$
I_k(\ell) = \int_{\sum_{m=1}^{k} \ell_m \leq \ell} \prod_{m=1}^{k} d\ell_m \quad \text{for } k = 1, 2, \ldots
\tag{56}
$$

Then, the following recursive relation is valid:

$$
\begin{aligned}
I_k(\ell) &= \int_{\sum_{m=1}^{k} \ell_m \leq \ell} \prod_{m=1}^{k} d\ell_m = \int_{\sum_{m=2}^{k} \ell_m \leq \ell - \ell_1} \prod_{m=1}^{k} d\ell_m = \\
&= \int_0^{\ell} d\ell_1 \int_{\sum_{m=2}^{k} \ell_m \leq \ell - \ell_1} \prod_{m=2}^{k} d\ell_m = \int_0^{\ell} d\ell_1 I_{k-1}(\ell - \ell_1) .
\end{aligned}
\tag{57}
$$

It results:

$$
I_1(\ell) = \int_{\ell_1 \leq \ell} d\ell_1 = \int_0^{\ell} d\ell_1 = \ell ,
\tag{58}
$$

$$I_2(\ell) = \int_0^\ell d\ell_1 I_1(\ell - \ell_1) = \int_0^\ell d\ell_1 (\ell - \ell_1) = \frac{\ell^2}{2} \, , \tag{59}$$

$$\cdots$$

and in general we have:

$$I_k(\ell) = \frac{\ell^k}{k!} \quad \text{for } k \geq 1 \, . \tag{60}$$

It is indeed easy to demonstrate that the last expression satisfies the recursive relation (57).

Now, we can calculate $p_{\ell,\infty}(k, \mu_s)$ starting from Eq. (52) and noting that in this case $\delta_{T_0} = 1$:

$$p_{\ell,\infty}(k, \mu_s) = \frac{\mu_s^k \int_{T_k} \prod_{m=1}^k d\ell_m p(\theta_m) d\Omega_m}{1 + \sum_{k'=1}^\infty \mu_s^{k'} \int_{T_{k'}} \prod_{m=1}^{k'} d\ell_m p(\theta_m) d\Omega_m} =$$

$$= \frac{(\mu_s\ell)^k / k!}{1 + \sum_{k'=1}^\infty (\mu_s\ell)^{k'} / k'!} = \frac{(\mu_s\ell)^k}{k!} e^{-\mu_s\ell} \, , \tag{61}$$

which is valid for $k \geq 0$.

Then, in the case of free propagation the distribution function $p_{\ell,\infty}(k, \mu_s)$ is a Poisson distribution: this result is in agreement with what we found in Sec. 3.1 [see Eq. (14)].

In the general case, *i.e.*, when a detector and/or a finite medium are considered, the analytical calculation of the distribution function $p_\ell(k, \mu_s)$ is not possible, mainly because the integrals present in Eq. (52) cannot be factorized.

Appendix C. Probability density function $f_k(\ell)$

In this Appendix we will derive the expression for the probability density function $f_k(\ell)$, that is the probability that a photon has travelled a length ℓ at the k-th scattering event, in the case of an infinite medium. More precisely, $f_k(\ell)d\ell$ represents the probability that the k-th scattering event occurs between ℓ and $\ell + d\ell$ for the length travelled by a photon.

To obtain the expression for $f_k(\ell)$, we recall that in an infinite non-absorbing homogeneous medium: *i)* the radiation travelling for a length ℓ attenuates by a factor $e^{-\mu_s\ell}$; *ii)* the radiation scattered along a path $d\ell$ is equal to $\mu_s d\ell$ times the incident power. Then, the probability density function that the first scattering event occurs after the photon has travelled a length ℓ will be:

$$f_1(\ell) = \mu_s e^{-\mu_s\ell} \, . \tag{62}$$

In order to obtain $f_2(\ell)$, that is the probability density function that the second scattering event occurs after the photon has travelled a path of length ℓ, we can proceed in the following way. Firstly, we calculate the probability that the first scattering event occurs after the photon has travelled a length ℓ_1 and the second scattering event after a further length $\ell_2 = \ell - \ell_1$. Being these independent events, we have:

$$f_1(\ell_1)f_1(\ell_2) = f_1(\ell_1)f_1(\ell - \ell_1)d\ell_1 d\ell_2 = \mu_s e^{-\mu_s\ell_1} d\ell_1 \mu_s e^{-\mu_s(\ell - \ell_1)} d\ell_2 =$$

$$= \mu_s^2 e^{-\mu_s\ell} d\ell_1 d\ell_2 \, . \tag{63}$$

Then, $f_2(\ell)$ is obtained considering all possible ℓ_1 path with $0 \leq \ell_1 \leq \ell$:

$$
\begin{aligned}
f_2(\ell) &= \int_0^\ell f_1(\ell_1) f_1(\ell_2) d\ell_1 = \int_0^\ell f_1(\ell_1) f_1(\ell - \ell_1) d\ell_1 = \\
&= \int_0^\ell \mu_s^2 e^{-\mu_s \ell} d\ell_1 = \mu_s^2 e^{-\mu_s \ell} \ell = \mu_s (\mu_s \ell) e^{-\mu_s \ell} \, .
\end{aligned}
\tag{64}
$$

Similarly, for $f_3(\ell)$ we have:

$$
\begin{aligned}
f_3(\ell) &= \int_0^\ell f_2(\ell_2) f_1(\ell_3) d\ell_2 = \int_0^\ell f_2(\ell_2) f_1(\ell - \ell_2) d\ell_2 = \\
&= \mu_s^3 e^{-\mu_s \ell} \int_0^\ell \ell_2 d\ell_2 = \mu_s^3 e^{-\mu_s \ell} \frac{\ell^2}{2} = \mu_s \frac{(\mu_s \ell)^2}{2} e^{-\mu_s \ell} \, ,
\end{aligned}
\tag{65}
$$

being ℓ_2 the length travelled before the photon undergoes the second scattering event.

By iterating this procedure, we obtain the general expression for the probability density of having k-th scattering event after travelling a path of total length ℓ:

$$
f_k(\ell) = \mu_s \frac{(\mu_s \ell)^{k-1}}{(k-1)!} e^{-\mu_s \ell}, \text{ for } k = 1, 2, \ldots
\tag{66}
$$

It is easy to verify the correct normalization of the function $f_k(\ell)$:

$$
\begin{aligned}
\int_0^\infty f_k(\ell) d\ell &= \int_0^\infty \mu_s \frac{(\mu_s \ell)^{k-1}}{(k-1)!} e^{-\mu_s \ell} d\ell = \frac{\mu_s}{(k-1)!} \int_0^\infty (\mu_s \ell)^{k-1} e^{-\mu_s \ell} d\ell = \\
&= \frac{1}{(k-1)!} \int_0^\infty x^{k-1} e^{-x} dx = \frac{(k-1)!}{(k-1)!} = 1 \, ,
\end{aligned}
\tag{67}
$$

where set $x = \mu_s \ell$ and used the definition of the gamma function for non negative integers:

$$
\int_0^\infty x^{k-1} e^{-x} dx = \Gamma(k) = (k-1)!, \text{ for } k = 1, 2, \ldots
\tag{68}
$$

By using the expression of $f_k(\ell)$ given in Eq. (66), we can calculate the moments of the distribution:

$$
\begin{aligned}
\langle \ell^m \rangle &= \int_0^\infty \ell^m f_k(\ell) d\ell = \int_0^\infty \ell^m \mu_s \frac{(\mu_s \ell)^{k-1}}{(k-1)!} e^{-\mu_s \ell} d\ell = \\
&= \frac{1}{\mu_s^m (k-1)!} \int_0^\infty (\mu_s \ell)^{m+k-1} e^{-\mu_s \ell} d(\mu_s \ell) = \\
&= \frac{1}{\mu_s^m (k-1)!} \int_0^\infty x^{m+k-1} e^{-x} dx = \\
&= \frac{\Gamma(m+k)}{\mu_s^m (k-1)!} = \frac{1}{\mu_s^m} \frac{(m+k-1)!}{(k-1)!} \, ,
\end{aligned}
\tag{69}
$$

where we set again $x = \mu_s \ell$ and used Eq. (68). In particular, for $m = 1$ we have the mean path travelled by a photon before undergoing the k-th scattering event:

$$
\langle \ell \rangle = \frac{k}{\mu_s} \, .
\tag{70}
$$

It is also possible to calculate the probability that a photon undergoes a scattering event, whatever its number is, in the $d\ell$ after having travelled a path of length ℓ. It can be obtained by

adding the probabilities $f_k(\ell)\mathrm{d}\ell$ that the k-th scattering occurs in $\mathrm{d}\ell$ for any k going from 1 to $+\infty$:

$$
\sum_{k=1}^{+\infty} f_k(\ell)\mathrm{d}\ell = \sum_{k=1}^{+\infty} \mu_s \frac{(\mu_s\ell)^{k-1}}{(k-1)!} e^{-\mu_s\ell}\mathrm{d}\ell = \mu_s e^{-\mu_s\ell} \sum_{k'=0}^{+\infty} \frac{(\mu_s\ell)^{k'}}{k'!}\mathrm{d}\ell =
$$
$$
= \mu_s e^{-\mu_s\ell} e^{\mu_s\ell}\mathrm{d}\ell = \mu_s\mathrm{d}\ell .
$$
(71)

This result can be seen as a confirmation of the correctness of Eq. (66) for $f_k(\ell)$: as a matter of fact, Eq. (71) simply tells us that for a photon propagating in an infinite non-absorbing medium the probability of undergoing scattering is, in accordance with the definition of the scattering coefficient, equal to $\mu_s\mathrm{d}\ell$, regardless of the point where the photon is and the trajectory it is following.

Appendix D. Proof of the validity of $f_k(\ell)$ and $p_{\ell,\infty}(k, \mu_s)$ by the induction method

D.1 Proof by the induction method for $f_k(\ell)$

In this Appendix we will derive the probability density function $f_k(\ell)$, that is the probability density that a photon has travelled a pathlength ℓ when the k-th scattering event occurs, in an infinite non-absorbing medium.

According to Eq. (62), we have:

$$
f_1(\ell) = \mu_s e^{-\mu_s\ell} .
$$
(72)

Then, it results:

$$
f_2(\ell) = f_1(\lambda) \otimes f_1(\lambda)|_{\lambda=\ell} = \int_0^\ell f_1(\lambda)f_1(\ell-\lambda)\mathrm{d}\lambda = \mu_s^2 \int_0^\ell e^{-\mu_s\lambda} e^{-\mu_s(\ell-\lambda)}\mathrm{d}\lambda =
$$
$$
= \mu_s^2 e^{-\mu_s\ell} \int_0^\ell \mathrm{d}\lambda = \mu_s^2 \ell e^{-\mu_s\ell} = \mu_s(\mu_s\ell)e^{-\mu_s\ell} .
$$
(73)

We remind that $\otimes$ is the convolution operator. Here we could derive the expression for $k = 3$, showing the factor $2 = (3-1)!$ at the denominator, but for brevity we skip this step. Assume now that:

$$
f_k(\ell) = f_1(\lambda) \otimes f_1(\lambda) \otimes \cdots \otimes f_1(\lambda)|_{\lambda=\ell} = \mu_s \frac{(\mu_s\ell)^{k-1}}{(k-1)!} e^{-\mu_s\ell} ,
$$
(74)

where the convolution is carried out $k-1$ times. Therefore:

$$
f_{k+1}(\ell) = f_k(\lambda) \otimes f_1(\lambda)|_{\lambda=\ell} = \int_0^\ell f_k(\lambda)f_1(\ell-\lambda)\mathrm{d}\lambda =
$$
$$
= \frac{\mu_s^k}{(k-1)!} \int_0^\ell \lambda^{k-1} e^{-\mu_s\lambda} \mu_s e^{-\mu_s(\ell-\lambda)}\mathrm{d}\lambda =
$$
$$
= \frac{\mu_s^{k+1} e^{-\mu_s\ell}}{(k-1)!} \int_0^\ell \lambda^{(k-1)}\mathrm{d}\lambda = \frac{\mu_s^{k+1} e^{-\mu_s\ell}}{k(k-1)!} \lambda^k .
$$
(75)

This implies that:

$$
f_{k+1}(\ell) = \mu_s \frac{(\mu_s\ell)^k}{k!} e^{-\mu_s\ell} .
$$
(76)

Then, Eq. (74), that is identical to Eq. (66), has been demonstrated by the induction method.

D.2 Proof by the induction method for $p_{\ell,\infty}(k, \mu_s)$

In this Appendix we derive the probability function $p_{\ell,\infty}(k, \mu_s)$, that is, the probability of having k scattering events within a fixed pathlength ℓ in a non-absorbing infinite medium in which the contribution of all propagating photons is considered, so that this quantity is not constrained by any detector. In this case to calculate the desired quantity we need the convolution of two distinct functions.

According to Eq. (12), we have:

$$p_{\ell,\infty}(0, \mu_s) = e^{-\mu_s \ell} . \tag{77}$$

Also, it results:

$$p_{\ell,\infty}(1, \mu_s) = f_1(\lambda) \otimes p_{\lambda,\infty}(0, \mu_s)\big|_{\lambda=\ell} = \mu_s \int_0^\ell e^{-\mu_s \lambda} e^{-\mu_s(\ell-\lambda)} \mathrm{d}\lambda =$$
$$= \mu_s \int_0^\ell e^{-\mu_s \ell} \mathrm{d}\lambda = \mu_s \ell e^{-\mu_s \ell} , \tag{78}$$

$$p_{\ell,\infty}(2, \mu_s) = f_2(\lambda) \otimes p_{\lambda,\infty}(0, \mu_s)\big|_{\lambda=\ell} = \mu_s^2 \int_0^\ell \lambda e^{-\mu_s \lambda} e^{-\mu_s(\ell-\lambda)} \mathrm{d}\lambda =$$
$$= \mu_s^2 e^{-\mu_s \ell} \int_0^\ell \lambda \mathrm{d}\lambda = \frac{(\mu_s \ell)^2 \, e^{-\mu_s \ell}}{2} . \tag{79}$$

Then, we assume:

$$p_{\ell,\infty}(k, \mu_s) = f_k(\lambda) \otimes p_{\lambda,\infty}(0, \mu_s)\big|_{\lambda=\ell} = \frac{(\mu_s \ell)^k}{k!} e^{-\mu_s \ell} . \tag{80}$$

It results:

$$p_{\ell,\infty}(k+1, \mu_s) = f_{k+1}(\lambda) \otimes p_{\lambda,\infty}(0, \mu_s)\big|_{\lambda=\ell} = \mu_s \int_0^\ell \frac{(\mu_s \lambda)^k}{k!} e^{-\mu_s \lambda} e^{-\mu_s(\ell-\lambda)} \mathrm{d}\lambda =$$
$$= \frac{\mu_s^{k+1} e^{-\mu_s \ell}}{k!} \int_0^\ell \lambda^k \mathrm{d}\lambda = \frac{(\mu_s \ell)^{k+1}}{(k+1)!} e^{-\mu_s \ell} . \tag{81}$$

In this way, again exploiting the induction method, we have demonstrated the validity of Eq. (80), that is identical to Eq. (14).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. A. Ishimaru, Introduction (IEEE, New York, NY, 1997), chap. 1, pp. 1–5.
2. S. Chandrasekhar, Radiative Transfer (Dover Publications, Inc., New York, NY, 1960).
3. A. Sassaroli, C. Blumetti, F. Martelli, *et al.*, "Monte Carlo procedure for investigating light propagation and imaging of highly scattering media," Appl. Opt. **37**(31), 7392–7400 (1998).
4. H. Rief, "Generalized Monte Carlo perturbation algorithms for correlated sampling and a second-order Taylor series approach," Ann. Nucl. Energy **11**(9), 455–476 (1984).
5. X-5 Monte Carlo Team, "MCNP-A General Monte Carlo N-Particle Transport Code, Version 5 Volume I: Overview and Theory," Tech. Rep. LA-UR-03-1987, Los Alamos National Laboratory, Los Alamos, NM (2003).
6. B. C. Kiedrowski, "Review of early 21st-century monte carlo perturbation and sensitivity techniques for k-eigenvalue radiation transport calculations," Nucl. Sci. Eng. **185**(3), 426–444 (2017).
7. J. Nguyen, C. K. Hayakawa, J. R. Mourant, *et al.*, "Perturbation Monte Carlo methods for tissue structure alterations," Biomed. Opt. Express **4**(10), 1946–1963 (2013).
8. J. Nguyen, C. K. Hayakawa, J. R. Mourant, *et al.*, "Development of perturbation Monte Carlo methods for polarized light transport in a discrete particle scattering model," Biomed. Opt. Express **7**(5), 2051–2066 (2016).
9. C. K. Hayakawa, J. Spanier, F. Bevilacqua, *et al.*, "Perturbation Monte Carlo methods to solve inverse photon migration problems in heterogeneous tissues," Opt. Lett. **26**(17), 1335–1337 (2001).
10. I. Seo, J. S. You, C. K. Hayakawa, *et al.*, "Perturbation and differential Monte Carlo methods for measurement of optical properties in a layered epithelial tissue model," J. Biomed. Opt. **12**(1), 014030 (2007).
11. C. K. Hayakawa, L. Malenfant, J. Ranasinghesagara, *et al.*, "MCCL: an open-source software application for Monte Carlo simulations of radiative transport," J. Biomed. Opt. **27**(08), 083005 (2022).
12. R. Yao, X. Intes, and Q. Fang, "Direct approach to compute jacobians for diffuse optical tomography using perturbation monte carlo-based photon replay," Biomed. Opt. Express **9**(10), 4588–4603 (2018).
13. A. A. Leino, T. Lunttila, M. Mozumder, *et al.*, "Perturbation monte carlo method for quantitative photoacoustic tomography," IEEE Trans. Med. Imaging **39**(10), 2985–2995 (2020).
14. L. Chicchi, G. Cecchini, I. Adam, *et al.*, "Reconstruction scheme for excitatory and inhibitory dynamics with quenched disorder: application to zebrafish imaging," J. Comput. Neurosci. **49**(2), 159–174 (2021).
15. T. Biancalani, G. Scalia, L. Buffoni, *et al.*, "Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram," Nat. Methods **18**(11), 1352–1362 (2021).
16. N. I. Nizam, M. Ochoa, J. T. Smith, *et al.*, "Monte Carlo-based data generation for efficient deep learning reconstruction of macroscopic diffuse optical tomography and topography applications," J. Biomed. Opt. **27**(08), 083016 (2022).
17. G. Maffeis, C. Amendola, A. Farina, *et al.*, "Application limits of the scaling relations for Monte Carlo simulations in diffuse optics. Part 2: Results," in preparation.
18. L. T. Perelman, J. Wu, I. Itzkan, *et al.*, "Photon migration in turbid media using path integrals," Phys. Rev. Lett. **72**(9), 1341–1344 (1994).
19. R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill, New York, NY, 1965).
20. F. Mainardi, R. Gorenflo, and A. Vivoli, "Beyond the poisson renewal process: A tutorial survey," J. Comput. Appl. Math. **205**(2), 725–735 (2007).
21. S. M. Ross, Introduction to Probability Models (Academic Press, New York, NY, 2014), eleventh ed.