Unbiased Light Transport Estimators for Inhomogeneous Participating Media

László Szirmay-Kalos¹, Iliyan Georgiev², Milán Magdics¹, Balázs Molnár¹, and Dávid Légrády¹

¹ Budapest University of Technology and Economics  ² Solid Angle

Abstract

This paper presents a new stochastic particle model for efficient and unbiased Monte Carlo rendering of heterogeneous participating media. We randomly add and remove material particles to obtain a density with which free flight sampling and transmittance estimation are simple, while material particle properties are simultaneously modified to maintain the true expectation of the radiance. We show that meeting this requirement may need the introduction of light particles with negative energy and materials with negative extinction, and provide an intuitive interpretation for such phenomena. Unlike previous unbiased methods, the proposed approach does not require a-priori knowledge of the maximum medium density that is typically difficult to obtain for procedural models. However, the method can benefit from an approximate knowledge of the density, which can usually be acquired on-the-fly at little extra cost and can greatly reduce the variance of the proposed estimators. The introduced mechanism can be integrated in participating media renderers where transmittance estimation and free flight sampling are building blocks. We demonstrate its application in a multiple scattering particle tracer, in transmittance computation, and in the estimation of the inhomogeneous air-light integral.

1. Introduction

Physics aims at establishing laws that describe natural phenomena. These laws can be simulated to create realistic behavior. It is also possible to define laws that are different from those of nature, but their simulation still results in the same expected behavior that nature would produce. The rationale for these structurally implausible but behaviorally valid models is the simplification of the simulation process. This paper aims at establishing such laws to render inhomogeneous participating media, which is a challenging problem [CPP⁰⁵, SKSU⁰⁵, Fat⁰⁹]. Analytic solutions are possible only for homogeneous volumes and in special cases [PP⁰⁹]. Numerical approaches are usually based on Monte Carlo quadratures and trace photons or importons (i.e. visibility rays) randomly in the medium [JC⁹⁸, PKK⁰⁰, QXFN⁰⁷]. To obtain estimates, we need to calculate the attenuation of the radiance between two points, called transmittance, and the probability density of the photon’s free flying in the medium. Monte Carlo methods fetch the parameters of the media at discrete points, which can be expensive when the model is procedural. On the other hand, many rays are traced through a pixel for anti-aliasing, motion blur, and complex lighting effects. Thus, there is a need for free flight sampling and transmittance computation methods that process a ray with a small number...
of data fetches, but provide unbiased and preferably low variance estimates that can be averaged into an accurate result.

To attack these problems, we modify the underlying model of the participating media to reduce the number of required fetches. Unlike other methods also modifying the model [ZRB14], our approach aims at the preservation of the expectation in Monte Carlo simulation, which provides additional freedom for model manipulation. The objective of this work is to devise unbiased methods for the computation of scattered and transmitted radiance.

Our main contributions are:

- A mechanism for randomly adding and removing material particles while simultaneously modifying their properties to produce media of arbitrary density where the Monte Carlo light transport provides the true expected values.
- Unbiased particle tracing, free flight sampling, and single scattering estimation methods that do not require an a-priori upper bound for the density.
- A variance analysis of the proposed estimators.
- The application of control variates to scattered radiance computation, which may require negative extinction.

The paper is structured as follows. In Section 2 we define the target problems and survey the previous work on free flight sampling and transmittance estimation. Section 3 presents our mechanism for random medium manipulation and the variance analysis. Among applications, Section 4.1 explains how the medium manipulation can be applied in a multiple scattering particle tracer. Section 4.2 considers the special problem of transmittance estimation. Finally, Section 4.3 addresses the scattering of light emitted by a single point source, which is a building block of efficient global illumination rendering.

2. Problem statement and previous work

In participating media, radiance \( L \) of point \( \vec{e} \) and direction \( \vec{\omega} \) is the contribution gathered by a ray \( \vec{p}(s) = \vec{e} - \vec{o} s \) of length \( S \):

\[
L(\vec{e}, \vec{\omega}) = T_{\vec{e}, \vec{\omega}}(S) L(\vec{p}(S), \vec{\omega}) + \int_0^S T_{\vec{e}, \vec{\omega}}(s) \sigma_t(\vec{p}(s)) \rho(\vec{\omega}, \vec{\omega}') d\vec{\omega}' ds 
\]

where \( \sigma_t(\vec{p}) \) is the extinction coefficient, \( a(\vec{p}) \) is the albedo, \( \rho(\vec{\omega}, \vec{\omega}') \) is the phase function, and

\[
T_{\vec{e}, \vec{\omega}}(s) = \exp \left( - \int_0^s \sigma_t(\vec{p}(t)) dt \right) 
\]

is the transmittance between points \( \vec{e} \) and \( \vec{e} - \vec{o} s \). The integral in the exponent is called the optical thickness.

Radiance estimation according to Equation 1 requires sampling points \( \vec{p}(s) \) along the ray. One possibility is to sample \( s \) with

\[
pdf(s) = \sigma_{\text{samp}}(\vec{p}(s)) \exp \left( - \int_0^s \sigma_{\text{samp}}(\vec{p}(t)) dt \right) 
\]

which is the probability density of the free flight in the medium with extinction coefficient \( \sigma_{\text{samp}}(\vec{p}(s)) \). If \( \sigma_{\text{samp}} \) is equal to real extinction \( \sigma_t \), then the transmittance \( T_{\vec{e}, \vec{\omega}}(s) \) is canceled in the Monte Carlo quadrature. Sampling with this pdf can be done by the inversion method, which requires a random number \( \xi \) uniformly distributed in \([0,1)\) and the solution of the following sampling equation for \( s \):

\[
- \log(1 - \xi) = \int_0^s \sigma_{\text{samp}}(\vec{p}(t)) dt.
\]

In practice, \( \sigma_{\text{samp}} \) is often different from \( \sigma_t \) to allow the analytic solution of the above sampling equation.

Ray marching approximates the optical thickness by stepping along the ray, assuming that the density between the steps is constant. Even with an unbiased estimate of the optical thickness, this algorithm is biased [RSK08], since the transmittance is an exponential, i.e. non-linear function of the optical thickness. To reduce the amount of bias, an excessive number of steps may be required if the medium density has high-frequency variations.

Woodcock tracking [WMHL65, Col68] (also called fictitious interaction tracking or delta-tracking) is an unbiased alternative to ray marching. Woodcock tracking advances in the medium with random-length steps and decides randomly at each visited point whether a real or fictitious collision happens. The expected length of the random steps is determined by the majorant extinction coefficient of the medium. Tight majorants are usually difficult to get especially for procedural media, and non-tight majorants make Woodcock tracking prohibitively inefficient.

Yue et al. [YIC∗10] proposed building a kd-tree to find regions of different majorant extinctions. However, this approach requires the inclusion of a fictitious scattering at each intersected region boundary, where a new random sample needs to be generated and the sample process repeated, which degrades performance when the ray crosses many regions.

In [SKTM11] Woodcock tracking has been generalized to allow arbitrary upper bounding extinction coefficient function by introducing virtual particles that modify the material density but do not change the radiant intensity. That paper addressed free flight sampling as well as transmittance estimation with separating the main part, i.e. the concept of control variates. In this approach, we express the real extinction \( \sigma_t(\vec{p}) \) as a sum of an analytically integrable main or control extinction \( \sigma_{\text{main}}(\vec{p}) \) and a difference extinction:

\[
\sigma_{\text{diff}}(\vec{p}) = \sigma_t(\vec{p}) - \sigma_{\text{main}}(\vec{p}).
\]

As the transmittance is the exponential of the negative extinction’s integral, it will be the product of the transmittance due to the main extinction and the transmittance of the difference extinction.

Novák et al. [NSJ14] considered transmittance estimation only and proposed modifying the weights of light particles instead of randomly terminating them in their ratio tracking method. In physics, Morgan et al. [MK15] investigated the same weighting scheme, and Galtier et al. [GBC∗13] showed the possibility of lifting the requirement of the majorant in their integral formulation. Novák et al. [NSJ14] also combined the weighting scheme with control variates in their residual ratio tracking method.

We build primarily on the results of [SKTM11, NSJ14] and present a probabilistic framework to attack not only the transmittance but also the scattering without requiring majorants, and allowing negative cross sections as well. The presented framework...
subsumes existing methods like Woodcock or residual ratio tracking as special cases, provides intuitive insights into them, and allows for easy proof of unbiasedness as well as variance analysis. Not only is this approach of theoretical interest, but it also has important practical advantages. Majorants of procedurally generated media are difficult or even impossible to get, and our method extends the applicability of unbiased sampling for these media.

3. The method of random medium manipulation

Participating media can be imagined as collections of material particles with which light particles may collide. The probability of collision is proportional to the density of material particles. Upon collision, a light particle can be absorbed or scattered, changing its properties. Light particles have three properties relevant for us: position, direction, and energy. An analog model would state:

- Upon collision, the light particle survives with probability equal to albedo \( a \), its direction is modified randomly according to the phase function, and the energy of the scattered particle \( E_{\text{scat}} \) will be equal to the energy of the incident particle \( E_{\text{in}} \).
- Direction and energy do not change while the light particle travels in free space.
- With probability \( 1 - a \) the light particle is terminated.

Transmittance estimation and free flight sampling depend just on the density of the material particles, so these operations can be made simpler by changing the material particle density. We modify the real extinction \( \sigma_{r} \) to a sampling extinction \( \sigma_{\text{samp}} \) that is appropriate for free flight sampling and transmittance estimation. Simultaneously, the medium parameters and collision laws are also altered to ensure that the expectation of the measured energy is preserved despite the modification of the extinction function. Unbiasedness is guaranteed if the behavior of free-space flight and the expected scaling of the energy at every direction upon a collision are preserved and random medium modifications are statistically independent. We introduce two random operations, one that increases the density and another one that reduces it, allowing to replace the real extinction by an arbitrary sampling extinction.

3.1. Adding virtual material particles

This operation adds virtual particles of albedo 1 and Dirac-delta type phase function (Figure 2). Virtual particles modify neither the energy nor the direction of the light particle during interaction, thus, they can be added safely to the medium without compromising the result of the Monte Carlo simulation. By adding virtual particles the density and consequently the extinction of the material is increased from \( \sigma_{r}(\vec{p}) \) to \( \sigma_{\text{samp}}(\vec{p}) \) by the density of virtual particles.

Let us denote the ratio of the real density and the sampling density by \( r \), which is in \([0,1]\) when virtual material particles are added:

\[
r(\vec{p}) = \frac{\sigma_{r}(\vec{p})}{\sigma_{\text{amp}}(\vec{p})}.
\]

We will omit the dependence of \( r \) on \( \vec{p} \) to simplify the notation.

Whether a collision happens on a real or a virtual particle can be decided randomly, proportionally to the densities. At point \( \vec{p} \), the interacting material particle is real with probability \( \sigma_{r}(\vec{p}) / \sigma_{\text{amp}}(\vec{p}) = r \) and is virtual with probability \( 1 - r \). A real material particle absorbs or scatters the light particle, while a virtual particle transmits it. The rules of interaction specify the energy of the scattered light particle \( E_{\text{scat}} \), the energy of the transmitted particle \( E_{\text{tran}} \), and their probabilities:

\[
E_{\text{scat}} = E_{\text{in}}, \quad E_{\text{tran}} = 0 \quad \text{with prob } ar,
\]

\[
E_{\text{scat}} = 0, \quad E_{\text{tran}} = E_{\text{in}} \quad \text{with prob } 1 - r,
\]

\[
E_{\text{scat}} = 0, \quad E_{\text{tran}} = 0 \quad \text{with prob } (1 - a)r.
\]

A common variance reduction technique is the substitution of a random variable by its mean in an estimator, which is known as Rao-Blackwellisation. In particle tracing this concept replaces random decisions by weights, giving back the expectation of the decisions. We call the resulting scheme the Double Particle Model since at each interaction the incident light particle is broken into a scattered particle and to a transmitted particle:

\[
E_{\text{scat}} = arE_{\text{in}}, \quad E_{\text{tran}} = (1 - r)E_{\text{in}}.
\]

Note that the second equation is similar to ratio tracking [NSJ14], but we handle also scattering and allow arbitrary \( \sigma_{\text{amp}}(\vec{p}) \) functions while that method considered only the transmission with a constant majorant sampling extinction. The two particles pose no problem in case of transmittance estimation when only the transmitted particle is tracked, but can lead to an exponential growth of the particle number if scattered light paths are also simulated.

To keep the number of light particles under control, we randomize the collision process and keep at most one of the two offsprings. The Single Particle Model scores a random decision with three exclusive outcomes: only the scattered light particle is kept with probability \( q_{\text{scat}} \), only the transmitted light particle survives with probability \( q_{\text{tran}} \), and both light particles are terminated with probability \( 1 - q_{\text{scat}} - q_{\text{tran}} \) (Figure 3). The energies should be compensated accordingly:

\[
E_{\text{scat}} = \frac{arE_{\text{in}}}{q_{\text{scat}}}, \quad E_{\text{tran}} = 0 \quad \text{with prob } q_{\text{scat}},
\]

\[
E_{\text{scat}} = 0, \quad E_{\text{tran}} = (1 - r)E_{\text{in}} \quad \text{with prob } q_{\text{tran}},
\]

\[
E_{\text{scat}} = 0, \quad E_{\text{tran}} = 0 \quad \text{with prob } 1 - q_{\text{scat}} - q_{\text{tran}}.
\]
Double Particle Model

\[ R E_{\text{in}} (1-r) E_{\text{tran}} \]

\[ \text{prob } q_{\text{scat}} \]

\[ \text{prob } q_{\text{tran}} \]

Other wise \[ E_{\text{scat}} \]

Only in the terminating version

Single Particle Model

\[ E_{\text{in}} \]

\[ \text{Density: } \sigma \]

\[ \text{Light particles} \]

\[ \text{Material particle} \]

\[ E_{\text{in}} \]

\[ E_{\text{in}} \]

\[ \text{Density: } \sigma_{\text{samp}} \]

\[ \text{Remove particle with prob } 1-r \]

\[ E_{\text{samp}} \]

\[ E_{\text{tran}} \]

\[ E_{\text{in}} \]

\[ E_{\text{in}} \]

\[ \text{Double Particle Model} \]

\[ \text{Single Particle Model} \]

Figure 3: The Single Particle Model is the randomization of the Double Particle Model. Randomization is the inverse of Run-Blacwellization, but can use different probabilities. Only the terminating version allows particles to be absorbed.

We consider two versions for the probabilities. The terminating version sets \( q_{\text{scat}} = ar \) and \( q_{\text{tran}} = 1 - r \), restoring Equation 7 and making the energy of the surviving light particle equal to that of the incident particle. The non-terminating version sets \( q_{\text{scat}} = ar/(ar + 1 - r) \) and \( q_{\text{tran}} = (1 - r)/(ar + 1 - r) \) making sure that exactly one particle survives at each interaction and its energy is independent of whether scattering or transmission happened.

3.2. Randomly removing material particles

If we want to reduce the medium density making \( r \) greater than 1, we keep a material particle with probability \( \sigma_{\text{samp}}(\vec{p})/\sigma_r(\vec{p}) = 1/r \) and remove it with probability \( 1 - 1/r \), modifying the scattering mechanism when the particle is not removed (Figure 4).

As the probability of keeping the real particle is \( 1/r \), the preservation of the expectation of scattered energy \( aE_{\text{in}} \) requires the multiplication by \( r \):

\[
E_{\text{scat}} = arE_{\text{in}}. \tag{10}
\]

On the other hand, if the material particle is removed, which happens with \( 1 - 1/r \) probability, the light particle can go through the empty space resulting in a superfluous light particle having the original energy and direction, which causes \( (1 - 1/r)E_{\text{in}} \) extra expected energy in the transmission direction. To compensate for this, we have to emanate light particles with negative energy \( E_{\text{tran}} \) when the material particle is preserved, which happens with probability \( 1/r \).

The expected energy into the transmission direction should be zero:

\[
\frac{1}{r} E_{\text{in}} + E_{\text{tran}} = 0 \Rightarrow E_{\text{tran}} = (1 - r) E_{\text{in}}. \tag{11}
\]

Note that Equation 8 of the Double Particle Model is equivalent to Equations 10 and 11, so it remains valid also for the case when material particles are randomly removed.

If we want to reduce the medium density making \( r \) greater than 1, we keep a material particle with probability \( \sigma_{\text{samp}}(\vec{p})/\sigma_r(\vec{p}) = 1/r \) and remove it with probability \( 1 - 1/r \), modifying the scattering mechanism when the particle is not removed (Figure 4). The expected energy into the transmission direction should be zero:

\[
\frac{1}{r} E_{\text{in}} + E_{\text{tran}} = 0 \Rightarrow E_{\text{tran}} = (1 - r) E_{\text{in}}. \tag{11}
\]

Figure 4: Randomly removing real particles decreases the density but preserves the expected radiance if interaction rules are simultaneously modified.

The expected energy into the transmission direction should be zero:

\[
\frac{1}{r} E_{\text{in}} + E_{\text{tran}} = 0 \Rightarrow E_{\text{tran}} = (1 - r) E_{\text{in}}. \tag{11}
\]

Note that Equation 8 of the Double Particle Model is equivalent to Equations 10 and 11, so it remains valid also for the case when material particles are randomly removed.

If we want to reduce the medium density making \( r \) greater than 1, we keep a material particle with probability \( \sigma_{\text{samp}}(\vec{p})/\sigma_r(\vec{p}) = 1/r \) and remove it with probability \( 1 - 1/r \), modifying the scattering mechanism when the particle is not removed (Figure 4). The expected energy into the transmission direction should be zero:

\[
\frac{1}{r} E_{\text{in}} + E_{\text{tran}} = 0 \Rightarrow E_{\text{tran}} = (1 - r) E_{\text{in}}. \tag{11}
\]

Note that Equation 8 of the Double Particle Model is equivalent to Equations 10 and 11, so it remains valid also for the case when material particles are randomly removed.

3.3. Negative extinction

If a control variate is used, the difference extinction defined in Equation 5 may also be negative, so the participating medium modification mechanism must be generalized to include this case as well. In a medium of positive extinction light decays, while in material of negative extinction light is amplified similarly to a chain reaction. Establishing a material particle model for this case requires special considerations as it has no direct physical interpretation. First, we consider the Double Particle Model.

Let us realize that if we mix material of positive extinction \( \sigma_r(\vec{p}) \) with material of negative extinction \( -\sigma_r(\vec{p}) \), then we should get the behavior of the empty space when \( \sigma_r = 0 \).

The rules for zero extinction say that the direction and the energy of the light particle are preserved, i.e.

\[
E_{\sigma_r>0}^{\text{scat}} = 0, \quad E_{\sigma_r>0}^{\text{tran}} = E_{\text{in}}. \]

The rules for zero extinction are summarized as:

\[
E_{\sigma_r>0}^{\text{scat}} = a|r|E_{\text{in}}, \quad E_{\sigma_r>0}^{\text{tran}} = (1 - |r|)E_{\text{in}}. \]

The rules for positive extinction can be obtained from the requirement that the total effect of a light particle colliding according to the rule of positive extinction and another light particle with the same energy colliding according to the rule of negative extinction should be equivalent to the total effect of applying the rules of empty space to the two light particles:

\[
E_{\sigma_r>0}^{\text{scat}} + E_{\sigma_r<0}^{\text{scat}} = 2E_{\sigma_r>0}^{\text{scat}}, \quad E_{\sigma_r>0}^{\text{tran}} + E_{\sigma_r<0}^{\text{tran}} = 2E_{\sigma_r>0}^{\text{tran}}. \]

Solving these equations for the negative case, we get:

\[
E_{\sigma_r<0}^{\text{scat}} = -a|r|E_{\text{in}} = arE_{\text{in}}, \]

\[
E_{\sigma_r<0}^{\text{tran}} = 2E_{\text{in}} - (1 - |r|)E_{\text{in}} = (1 - r)E_{\text{in}} \]

since \( \sigma_r \) and \( r \) are negative now.

© 2017 The Author(s)

Computer Graphics Forum © 2017 The Eurographics Association and John Wiley & Sons Ltd.
Note that Double Particle Model equations are valid also for the negative case. However, when the extinction and consequently \( r \) are positive, the radiance intensity is decreased, for negative extinction and \( r \), it is increased.

The Double Particle Model can be randomized to obtain the Single Particle Model also for negative extinction if the non-terminating version is taken:

\[
\frac{a|r|}{a|r| + |1 - r|}; \quad q_{\text{ran}} = \frac{|1 - r|}{a|r| + |1 - r|}. \tag{12}
\]

However, the terminating version is not applicable for negative extinctions since its resulting \( q_{\text{ran}} = |1 - r| \) would be greater than 1, which cannot be a probability.

### 3.4. Variance analysis

Randomly modifying the medium affects not only the cost of simulation but also the variance of the radiance estimates. This section provides formulas for the variance of the transmittance at an arbitrary point along a ray. We consider the change of the expectation and the variance when we make an infinitesimal step along the ray, establishing a differential equation for the unknown variance that has a closed from solution.

Let us consider the expectation \( T(s) = \mathbf{E}[\hat{T}(s)] \) and variance \( V(s) = \mathbf{V}[\hat{T}(s)] \) of sampled transmittance \( \hat{T}(s) \) at distance \( s \). The sampled transmittance is a random variable, which is updated with the following formula when \( s \) is increased by \( ds \):

\[
\hat{T}(s + ds) = \hat{T}(s)W(s, ds)
\]

where \( W(s, ds) \) is a random variable describing the transmittance of differential interval \( [s, s + ds] \) since the transmittance is the probability of survival, so merging two intervals, the resulting transmittance is the product of the respective transmittances. The random variable \( W(s, ds) \) is independent of \( \hat{T}(s) \), thus the expectation and the variance of the transmittance at \( s + ds \) are:

\[
\mathbf{E}[\hat{T}(s + ds)] = \mathbf{E}[\hat{T}(s)]\mathbf{E}[W(s, ds)],
\]

\[
\mathbf{V}[\hat{T}(s + ds)] = \mathbf{V}[\hat{T}(s)]\left(\mathbf{V}[W(s, ds)] + \mathbf{E}^2[W(s, ds)]\right) + \mathbf{V}[W(s, ds)]\mathbf{E}^2[\hat{T}(s)],
\]

where we use that for two independent random variables \( X \) and \( Y \), the identities \( \mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y] \) and \( \mathbf{V}[XY] = \mathbf{V}[X]\mathbf{V}[Y] + \mathbf{V}[X]\mathbf{E}^2[Y] + \mathbf{V}[Y]\mathbf{E}^2[X] \) hold.

Since our models are unbiased, the expected transmittance of \( [s, s + ds] \) is equal to the probability of no interaction:

\[
\mathbf{E}[W(s, ds)] = 1 - \sigma_l(\hat{p}(\tau))ds,
\]

and the expected transmittance is equal to the real transmittance:

\[
\mathbf{E}[\hat{T}(s)] = T(s) = \exp\left(-\int_0^s \sigma_l(\hat{p}(\tau))d\tau\right).
\]

Substituting these and shorthand notations \( T(s) = \mathbf{E}[\hat{T}(s)] \) and \( V(s) = \mathbf{V}[\hat{T}(s)] \) into Equation 13, we obtain:

\[
V(s + ds) = V(s)\left(\mathbf{V}[W(s, ds)] + (1 - \sigma_l(\hat{p}(s))ds)^2\right) + \mathbf{V}[W(s, ds)]T^2(s).
\]

Subtracting \( V(s) \) from both sides, dividing the equation by \( ds \) and taking the \( ds \to 0 \) limit, we can establish the following differential equation for the variance of the transmittance:

\[
\frac{dV}{ds} = V(s)\left(\frac{w(s) - 2\sigma_l(\hat{p}(s))}{w(s)}\right) + w(s)T^2(s),
\]

where

\[
w(s) = \lim_{ds \to 0} \frac{\mathbf{V}[W(s, ds)]}{ds}
\]

is the variance introduction density, which is the only function in this equation that depends on the sampling method, while \( \sigma_l(\hat{p}(s)) \) and \( T(s) \) depend just on the medium. This differential equation can be solved analytically:

\[
V(s) = T^2(s)\left(\exp\left(\int_0^s w(\tau)d\tau\right) - 1\right) = \exp\left(-\int_0^s 2\sigma_l(\hat{p}(\tau))d\tau\right)\left(\exp\left(\int_0^s w(\tau)d\tau\right) - 1\right). \tag{16}
\]

If we use a control variate, then sampled transmittance \( \hat{T}(s) \) is the product of the analytically computed main transmittance \( T_{\text{main}}(s) \) and the sampled transmittance \( T_{\text{diff}}(s) \) of the difference extinction. Thus, the variance of the sampled transmittance is the product of the square of the difference transmittance and the variance obtained by replacing the real extinction with the difference extinction, i.e. \( V(s) = T^2_{\text{main}}(s)V_{\text{diff}}(s) \). However, as \( T^2_{\text{main}}(s)T^2_{\text{diff}}(s) = T^2(s) \), we can come to the conclusion that Equation 16 remains valid also with a control variate, just the real extinction should be replaced by the difference extinction when variance introduction density \( w \) is calculated. If the difference extinction is zero, i.e. the control variate is equal to the real extinction, then the variance introduction density is also zero, which means that the proposed method gives back the zero variance analytic solution.

The variance is said to be stable, i.e. bounded on an arbitrarily long ray if the exponents are negative, i.e. \( w(s) \leq 2\sigma_l(\hat{p}(s)) \). We have to emphasize that it makes sense to use unstable methods as well if the length of the interval where stability does not hold is limited.

**Double Particle Model** In the Double Particle Model, the light particle does not interact with material particles in \([s, s + ds] \), i.e. there is no change of the sampled transmittance \( W_{\text{double}}(s, ds) = 1 \) with probability \( 1 - \sigma_{\text{samp}}ds \), and interaction happens with probability \( \sigma_{\text{amp}}ds \) when the transmittance is scaled with \( W_{\text{double}}(s, ds) = 1 - r \) both for positive and negative difference extinctions. The expectation of the transmittance of \( ds \) in the material of the difference extinction is \( 1 - \sigma_{\text{diff}}ds \). The variance of \( W_{\text{double}}(s, ds) \) is then

\[
V[W_{\text{double}}] = \int_0^1 \left(1 - \sigma_{\text{amp}}ds\right) + \left(1 - \sigma_{\text{diff}}ds\right)^2.
\]

The variance introduction density is the limit of the ratio of the introduced variance and \( ds \):

\[
w_{\text{double}}(s) = \lim_{ds \to 0} \frac{V[W_{\text{double}}]}{ds} = \sigma_{\text{amp}}(r^2 - 2r) + \sigma_{\text{diff}}\sigma_{\text{amp}} = \sigma_{\text{diff}}r
\]

since \( \sigma_{\text{amp}} = \sigma_{\text{diff}} \). When \( \sigma_{\text{amp}} \) is negative, \( r = \sigma_{\text{diff}}/\sigma_{\text{amp}} \) is also negative, making \( w \) always positive. The variance introduction density of the Double Particle Model is symmetric, and can be made
arbitrarily small by reducing $|r|$, i.e. increasing the sampling density and processing proportionally more samples.

**Single Particle Model:** In the Single Particle Model, there is no interaction ($W_{\text{single}}(s,d) = 1$) with probability $1 - \sigma_{\text{samp}} ds$, and interaction happens with probability $\sigma_{\text{samp}} ds$ when the transmittance is scaled by $W_{\text{single}}(s,d) = \frac{1}{q_{\text{tran}}}$. Thus the variance of $W_{\text{single}}(s,d)$ is

$$1^2 (1 - \sigma_{\text{samp}} ds) + \left( \frac{1-r}{q_{\text{tran}}} \right)^2 q_{\text{tran}} \sigma_{\text{samp}} ds - (1 - \sigma_{\text{diff}} ds)^2.$$  

The variance introduction density is then

$$w_{\text{single}}(s) = \lim_{ds \to 0} \frac{V[W_{\text{single}}]}{ds} = \sigma_{\text{diff}} \left( 2 - \frac{1}{r} + \left( \frac{1-r}{q_{\text{tran}}} \right)^2 \right). \quad (18)$$

The terminating version is applicable only for non-negative extinctions ($r \geq 0$) and sets $q_{\text{tran}} = |1 - r|$, which leads to

$$w_{\text{single}} = \begin{cases} \sigma_{\text{diff}} & \text{if } r \leq 1, \\ \sigma_{\text{diff}}(3 - 2/r) & \text{if } r \geq 1. \end{cases} \quad (19)$$

This means that it is not worth increasing the sampling density beyond the threshold of the majorant case, i.e. reducing $r$ below 1, because the variance remains the same, but the sampling cost increases.

The non-terminating version can work both for non-negative and negative extinctions. Substituting its transmission probability form Equation 12 into Equation 18, we obtain

$$w_{\text{single}} = \sigma_{\text{diff}} |1 - r| + a |1 - r| \cdot (20)$$

Note that due to the $|1 - r|$ term the variance introduction density of the Single Particle Model is asymmetric, preferring the positive difference extinctions especially in high albedo media. Unlike in the Double Particle Model, increasing the sampling density can significantly reduce the variance only for low albedo media since the variance introduction density has a lower bound of $[\sigma_{\text{diff}}a$ when $|r|$ goes zero. As the sampling cost increases for lower $|r|$ values, for high albedo media, the optimal choice is $r \approx 1$.

Comparing the terminating (Equation 19) and non-terminating (Equation 20) options, we can see that the terminating option is better for the minorant case ($r > 1$) and the non-terminating for the majorant case ($r < 1$).

**Optimal control variate:** The variance is a monotonically increasing function of the integral of $w(s)$, which is the only factor that depends on the actual strategy. Thus, the optimal choice can be found by minimizing this integral. For the sake of simplicity, let us assume that sampling density $\sigma_{\text{samp}}$ is constant and consider the Double Particle Model:

$$\int_0^s w(t) \, dt = \int_0^s \frac{\sigma_{\text{diff}}(t)}{\sigma_{\text{samp}}} \, dt - \frac{1}{\sigma_{\text{samp}}} \int_0^s (\sigma(t) - \sigma_{\text{main}}(t))^2 \, dt.$$  

This integral is minimal if $\sigma_{\text{main}}(s)$ is an optimal approximation of real extinction $\sigma_r(s)$ in $L_2$ sense. If the control variate is constant, its optimal value is the mean of the real extinction. Note that unlike for conventional application of control variates when a constant main part does not help, here a constant value can also reduce the variance.

4. **Applications**

To demonstrate the application of the proposed model, we render procedurally generated participating media defined by twelve-octave Perlin-noise [EMP'03] and by an analytic function. The albedo is 0.9 and 0.7 in the Perlin-noise model and in the analytic model, respectively, and they scatter according to the Rayleigh phase function. The analytic medium is enclosed in a sphere of radius 10 and center $(0,0,10)$, and its extinction at point $(x,y,z)$ inside the sphere is

$$\sigma_r(x,y,z) = \left( \cos(5(x+y+z)/2) + 1 \right)^5 \cdot \frac{\sin(z/2) + 2}{3}, \quad (21)$$

while it is zero elsewhere.

![Figure 5: The anatomy of the ray going through the center of the Perlin-noise (top) and analytic medium (bottom) used in the experiments. We show one set of explorer points and random interaction points that together define one random sample of the complete transmittance function. Explorer points show where the approximate extinction is evaluated to obtain the control variate. Interaction points are generated when jumping on the ray with free flight distances to estimate the transmittance of the difference extinction. Sampling extinction $\sigma_{\text{samp}}$ is 0.4 for this ray in the analytic medium.](image-url)
even be borrowed from nearby rays. Note that primary rays belonging to a pixel or neighboring pixels cross approximately the same part of the volume, so do shadow rays belonging to a point source. Thus, when a new ray is traced, the points where the extinction has been evaluated in the already processed neighboring rays can be projected to the current ray. The projection involves the substitution of the ray parameter of the borrowed ray into the ray equation of the current ray. By reusing information, the cost of main part definition becomes negligible.

The transmittance along the ray is estimated as the product of the main part transmittance and the weight of a particle transmitted to the current distance. The particle jumps from interaction point to interaction point increasing the distance randomly with the sampled free flight. The free flight distance is obtained by solving Equation 4 using sample density $\sigma_{\text{samp}}$ mimicking $|\sigma_{\text{diff}}|$.

4.1. Particle tracing

The discussed on-the-fly random medium manipulation can be executed during particle tracing, where particles can be either photons emitted by light sources or importons emitted by the camera. To avoid the exponential growth of tracked light particles, we use the Single Particle Model, which randomly keeps either the transmitted or the scattered particle, but not both.

The rules of the Single Particle Model are given by Equation 9. When the next interaction point is to be sampled, we set a desired density function $\sigma_{\text{samp}}(\vec{p})$ and solve Equation 4 to find the interaction point along the ray. Based on the real extinction and the sampling density, the interaction rules are applied, and the process is repeated. The sampling extinction can be a piece-wise constant or piece-wise linear function. The sampling extinction can evolve during rendering, i.e. it can be tuned based on the extinction values of already visited points. The simplest choice is a constant value.

The following algorithm simulates a light particle of properties like position $\vec{p}_{\text{start}}$, direction $\vec{d}$, and energy or importance $E$ until the particle leaves the volume of interest or gets absorbed:

```
SimulateParticle($\vec{p}_{\text{start}}, \vec{d}, E$)
while ($\vec{p}_{\text{start}}$ is inside the volume AND $|E| > 0$) // free flight sampling with $\sigma_{\text{samp}}$
    Solve $-\log(1-rand()) = \int_{0}^{1} \sigma_{\text{samp}}(\vec{p}(\tau))d\tau$ for $s$;
    $\vec{p} = \vec{p}_{\text{start}} + s\vec{d}$; // point of interaction
    $r = \sigma_{\text{albedo}}(\vec{p})/\sigma_{\text{samp}}(\vec{p})$; // fetch real extinction
    $a = \text{Albedo}(\vec{p})$;
    $q_{\text{cat}} = ar$; $q_{\text{ran}} = |1-r|$; // terminating version
    $q_{\text{ran}} = q_{\text{cat}} + q_{\text{ran}}$;
    if ($r < 1$ OR $q_{\text{ran}} > 1$) // adaptive
        $\xi = \text{rand}()$;
        if ($\xi < q_{\text{cat}}$)
            $E = arE/q_{\text{cat}}$;
            $\vec{d} = \text{SamplePhaseFunction}(\vec{d})$;
        else if ($\xi < q_{\text{cat}} + q_{\text{ran}}$)
            $E = (1-r)E/q_{\text{ran}}$; // transmission
        else        // absorption
            $\vec{p}_{\text{start}} = \vec{p}$;
    end
endwhile
```

This program uses the functions $\sigma_{\text{albedo}}(\vec{p})$ returning the real extinction at point $\vec{p}$, Albedo($\vec{p}$) providing the albedo, SamplePhaseFunction($\vec{d}$) that randomly samples a new direction with the phase function of the real material, and rand() that generates a uniformly distributed random number in $[0,1]$. For the sake of simplicity, the pseudo-code assumes that the phase function can be exactly sampled, so the directional pdf cancels it. As no control variate is applied now, $r$ is assumed to be non-negative. The reason of using no control variate is that in particle tracing scattered rays can be very incoherent, thus the cost of main part definition cannot be so simply amortized as in the case of primary rays. According to the variance analysis, the non-terminating version is better for majorants and the terminating version for minorants. Therefore, the above implementation uses an adaptive version that dynamically chooses from the terminating and non-terminating versions at each interaction.

Figure 6 shows the multiple scattering rendering of the medium defined by Equation 21. The medium is illuminated by high dynamic range environment lighting. The rows compare the adaptive SimulateParticle algorithm with a non-terminating and a terminating version. The adaptive version is never worse than the two original versions. As the variance analysis indicates, the variance cannot be further decreased by increasing the sampling extinction beyond the maximum extinction, which is 1 in this medium. On the other hand, the sampling cost expressed by the number of data fetches is inversely proportional to the sampling extinction. To take both the cost and the benefit into account, we use a Figure of Merit (FoM) measure defined as the reciprocal of the product of the Mean Square Error and the number of medium parameter evaluations per pixel. Figure 7 shows the multi-octave Perlin-noise medium.

4.2. Transmittance estimation

If scattering is ignored, the integral of Equation 1 disappears and only the transmittance should be evaluated that attenuates the environment illumination. During transmittance calculation both absorption and out-scattering are just lost in radiance, and out-scattered light particles are not tracked. Thus, it is worth applying the Double Particle Model defined by Equation 8, since it has lower variance and its extra cost of particle splitting does not show up in this application. We also use a control variate since primary rays are coherent, thus main part definition has negligible added cost.

If the main part is appropriately selected, the absolute value of the difference extinction is significantly smaller than the original extinction, which has two important advantages. On the one hand, the variance can be significantly reduced. On the other hand, the expected length of the random steps is the reciprocal of the sampling density mimicking $|\sigma_{\text{diff}}(\vec{p})|$, thus the medium is explored more quickly by taking larger steps.

The transmittance is estimated as the product of the main part transmittance and the importance of a particle transmitted to the end of the considered interval. The particle jumps to interaction points that are generated with incrementally solving the sampling equation for given $\sigma_{\text{samp}}(\vec{p}(s))$ mimicking the absolute difference extinction until the medium is left. The following algorithm computes an unbiased estimate of the transmittance in interval $[0, S]$:

© 2017 The Author(s)
Computer Graphics Forum © 2017 The Eurographics Association and John Wiley & Sons Ltd.
Figure 6: Particle tracing of the medium of Equation 21 with 100 rays per pixel setting the $\sigma_{\text{samp}}$ value differently to demonstrate the effect of the sampling density on the number of volume fetches, i.e. the rendering cost, and on the rendering error. Increasing $\sigma_{\text{samp}}$ reduces the variance until $\sigma_{\text{samp}} < \sigma_t \leq \sigma_{\text{max}} = 1$ on the ray as suggested by the variance analysis of the Single Particle Model in Section 3.4. However, the number of fetches becomes proportional to $\sigma_{\text{samp}}$, thus the sampling cost increases. The FoM efficiency metric is defined as the reciprocal of the product of the mean square error and the cost. Note that we do not require the sampling density to be a majorant, but its appropriate choice is still important as it affects the error-cost trade-off. In this case, the optimum value is $\sigma_{\text{samp}} \approx 0.8$, which is below the maximum extinction, depending on the actual extinction values a given ray can see and also on the variation of the extinction and illumination inside a pixel. If the sampling extinction is too low, the variance grows rapidly, so the estimator becomes poorer.

Figure 7: Particle tracing of a twelve-octave Perlin noise with equal number of medium parameter evaluations, i.e. equal cost, setting sampling density $\sigma_{\text{samp}}$ value differently. The extinction is evaluated 500 times per pixel in each method. The optimal $\sigma_{\text{samp}}$ value is slightly below the extinction of the larger dense regions. In this medium, the maximum of the extinction is 1, but most of the rays meet extinctions less than 0.6.
Transmittance($\bar{\rho}_{\text{start}}, \bar{\omega}, S$)

\[
E = \exp(-\int_0^s \sigma_{\text{main}}(\bar{\tau}(\tau))d\tau);
\]
\[s = 0;\]
\[\text{while}(|E| > 0)\]
\[
\text{Solve}\ \{ -\log(1-\text{rand}(i)) = \int_0^s \sigma_{\text{amp}}(\bar{\tau}(\tau))d\tau\} \text{ for } \Delta s;
\]
\[\bar{\rho} = \bar{\rho}_{\text{start}} + \bar{\omega} \Delta s;\]
\[s = s + \Delta s;\]
\[\text{if } (s \geq S) \text{ break ;}\]
\[
E = E(1 - \sigma_{\text{diff}}(\bar{\rho})/\sigma_{\text{amp}}(\bar{\rho}));
\]
\[\bar{\rho}_{\text{start}} = \bar{\rho};\]
\[\text{endwhile}\]
\[\text{return } E;\]
\[\text{end}\]

Note that this algorithm is identical to the residual ratio tracking [NSJ14] although we have not assumed that $\sigma_{\text{amp}}(\bar{\rho})$ is constant and is a majorant to the absolute value of the difference extinction. Due to the intuitive interpretations of negative energy and negative extinction, we have shown that these hard to meet requirements can be sidestepped. The result of the simulation is still unbiased, with occasionally negative estimates (Figure 8). Such negative estimates are clipped to zero only for visualization, which is not much different from clipping or tone-mapping values outside the dynamic range of our screen.

\[\text{Figure 8: Sample transmittance functions on the ray going through the center of the analytic medium, the real transmittance, and probabilistic lower and upper bounds using the standard deviation that is the square root of the variance obtained with Equation 16. The real extinction and its main part are shown by the lower image of Figure 5, and the sampling extinction is } \sigma_{\text{amp}} = 0.4.\]

In addition to the proposed transmittance estimation algorithm, we also generated images with ray marching and Woodcock tracking setting the sampling density to 1, which is the tightest constant majorant of the extinction and therefore the most optimal choice for Woodcock tracking (Figures 9 and 10). Ray marching worked with a single ray per pixel taking steps of length 0.9, and we set the number of rays per pixel for Woodcock tracking and the new algorithm to make the different methods evaluate the extinction function the same number of times. The new method provides the lowest error and a visually pleasing artifact free image.

4.3. Single scattering of point source lighting

Let us consider a practically important special case, the computation of single scattering of the light of a point source located at $\vec{t}$ (Figure 11). As area lights can be approximated by many point sources, and multiple scattering can be simulated by introducing Virtual Point Lights [ENSD12], this basic operation can be extended to a complete global illumination renderer. If we take the adjoint approach and implement path tracing, variable $E$ will be the visual importance, i.e. the weight of the gathered contribution.

\[\text{Figure 11: Single scattering of point source lighting.}\]

\[\text{In this special case, the radiance gathered by ray } \bar{p}(s) = \vec{e} - \bar{\omega} s\text{ of start } \vec{e}, \text{ direction } \bar{\omega}, \text{ and length } S \text{ is}\]

\[
L(\vec{e}, \bar{\omega}) = T_{\vec{t}, \bar{\omega}}(S)L(\bar{p}(S), \bar{\omega}) + \int_0^S T_{\vec{t}, \bar{\omega}}(s)\sigma_{\text{diff}}(\bar{p}(\bar{p}(s)))\rho(\bar{\omega} - \vec{e})L_{\text{diff}}(\bar{p}(s), \bar{\omega}(s))ds \quad (22)
\]
where $\Phi'(s)$ is the direction from light source $\vec{l}$ to point $\vec{p}(s)$. The incident radiance $L^m$ due to a point source of power $\Phi$ is

$$L^m(s) = \frac{\Phi}{4\pi[D^2 + (s - s_l)^2]T\hat{p}(s)\hat{w}'(s)[|\vec{l} - \vec{p}(s)|]}$$

where $D$ is the distance between the ray and the light source, and $s_l$ is the ray parameter where the ray is the closest to the point source.

To compute the Monte Carlo quadrature of the integral of Equation 22, we need to estimate transmittance $T_{E,\omega}(S)$ and find sample points $\vec{p}(s)$ at which the scattered radiance is evaluated and multiplied by the estimate of transmittance $T_{E,\omega}(s)$. The key observation is that Algorithm Transmittance used to find $T_{E,\omega}(S)$ generates data at interaction points from which a low-variance estimator of the transmittance can be computed at an arbitrary point of the ray. This estimate is not constant between interaction points, but thanks to the control variate it follows an exponential fall off [JNT+11, NNDJ12]. This observation opens the possibility of using arbitrary techniques for sampling scattering points and employing the continuous transmittance function to obtain unbiased low-variance estimates.

One option, called the attenuation-driven sampling, is to directly use the interaction points also for the locations where shadow rays are cast and scattering of the light from the point source is evaluated. The density of interaction points is $\sigma_{\text{sampl}}(\vec{p}(s))$ (see Supplementary Material for the proof). The sampling density of interaction points mimics none of the factors in the integral of Equation 22, thus it provides poor importance sampling and its only advantage is that sample points generated for the transmittance are reused for scattering. The importance sampling aspect can be improved by the concept of Sampling Importance Re-sampling (SIR) [TCEO05]. This rejection sampling scheme increases the variance but saves the expensive shadow ray computation of those interaction points that would have negligible scattered radiance.

Scattering points can be sampled independently of the interaction points, for example, with the Cauchy distribution to mimic the weak singularity of $1/(D^2 + (s - s_l)^2)$ in the scattered radiance [KF12]. We call this source-driven sampling. Alternatively, scattering points can be sampled mimicking $T_{E,\omega}(s)\sigma_{\text{scat}}(\vec{p}(s))$ but using the piece-wise linear approximation of the extinction to allow analytic solution of the sampling equation, which is called scattering-driven sampling. Different approaches can even be combined according to Multiple Importance Sampling (MIS).

The discussed sampling methods have been implemented in a single scattering type ray tracer and we visualized the medium defined by Equation 21 illuminated by environment lighting and also by placing a point source inside the medium. The main part is built from information stored in explored points, which is shared by nearby pixels. The main part of the extinction function is computed by fitting a piece-wise linear function on the explorer points. The sampling density mimicking $|\sigma_{\text{sampl}}(\vec{p}(s))|$ can be set to a constant if the main part absorbs most of the variations. The optimal choice is close to the maximum absolute value of the difference extinction, thus we set it to the maximum obtained during casting the last few neighboring rays.

In case of attenuation-driven sampling, shadow rays are traced from the interaction points. Source-driven sampling obtains scattering points from the Cauchy distribution mimicking $1/(D^2 + (s - s_l)^2)$, and trace shadow rays from these points. Scattering-driven sampling mimics both the free flight until and the albedo at the scattering point. Source-driven and Scattering-driven samples can be generated before sampling the transmittance, and these samples can also be used as explorer points. For shadow rays only the transmittance is calculated.

Rendering results are shown in Figure 12. Different methods fetch the procedural density by the same number of times, which leads to different numbers of rays depending on how many samples per primary ray are needed by a particular method. We can observe that ray marching rendering is noise free but is biased. Woodcock tracking has high variance. The methods incorporating the proposed transmittance estimation are significantly better. If the point source is strong compared to environment lighting, then the combination of source-driven and scattering-driven approaches with MIS is the winner. This option has also been implemented in a production renderer, which was used to compute the image of Figure 13.

Figure 13: Image produced by a production renderer where we integrated the proposed method using the combination of the source-driven and the scattering-driven sampling with MIS.

5. Conclusions
This paper presented a participating media manipulation method that simplifies free flight sampling and transmittance estimation in heterogeneous media while guaranteeing that the original expected radiance values are preserved. The method does not require the knowledge of a majorant extinction coefficient, which extends its applicability to procedural models as well. We have analyzed the variance of the method and have proven that approximate extinction values and especially the control variate can significantly increase the accuracy. Such information can be obtained on-the-fly from rays belonging to the same pixel or going nearby with negligible additional computational cost. Thus, with making the algorithm a little more involved, the variance of the estimators can be decreased.
Acknowledgement

This work has been supported by OTKA K–104476, VKSZ-14 PET/MRI 7T, and ÚNKP-16-4-I projects. The authors are grateful to Marcos Fajardo for the valuable discussions.

References


© 2017 The Author(s) Computer Graphics Forum © 2017 The Eurographics Association and John Wiley & Sons Ltd.

![Figure 12: Equal number of medium parameter evaluations comparison of the proposed single scattering methods to classical single ray per pixel, constant step ray marching and multi-ray per pixel Woodcock tracking. Each method took approximately 300 samples (upper row) and 50 samples (lower row) from the medium per pixel on average to compute all primary and shadow rays. For each method, we show the number of primary rays requiring the same number of medium fetches, and the RMS errors.](image-url)