Pseudo-Marginal Metropolis Light Transport

Joel Kronander\textsuperscript{1} Thomas B. Schön\textsuperscript{2} and Jonas Unger\textsuperscript{3}
\textsuperscript{1} Linköping University, \textsuperscript{2} Uppsala University

1 Introduction

Accurate and efficient simulation of light transport in heterogeneous participating media, such as smoke, clouds and fire, plays a key role in the synthesis of visually interesting renderings for e.g. visual effects, computer games and product visualization. Similar problems also has important applications in other fields such as neutron transport and medical radiation dosimetry. While significant progress has been made towards robust and efficient algorithms for rendering heterogeneous participating media, scenes containing highly anisotropic phase functions, glossy surfaces and complex visibility still remain challenging.

Modern global illumination rendering is based on Monte Carlo methods computing averages of stochastically sampled light paths connecting the sensor with light sources in the scene. The efficiency of these algorithms is directly dependent on the strategy used in the light path sampling. For complex scenes, the Metropolis light transport (MLT), [Veach and Guibas 1997], algorithm provide a powerful framework for sampling the path space. The method is based on using Markov chain Monte Carlo (MCMC) to sample light paths \( \{ \bar{x}^i \}_{i \geq 1} \) according to some target distribution, \( \pi(\bar{x}) \), proportional to the image contribution of the paths.

However, in many applications the target distribution, \( \pi(\bar{x}) \), cannot be evaluated exactly. For example, when rendering scenes with heterogenous participating media, where, in general, computing the transmittance along a path is analytically intractable. This limits direct use of MLT, as it depends on Metropolis-Hastings (MH) acceptance probabilities which cannot be evaluated in closed form. The same problem occurs for other rendering methods relying on MCMC sampling, such as Energy Redistribution path tracing (ERPT) [Cline et al. 2005]. Previous work, [Pauly et al. 2000], proposed to substitute the exact quantity of interest, \( \pi(\bar{x}) \), with a biased approximation, \( \hat{\pi}(\bar{x}) \), to compute the MH probability. However, direct application of this method does not lead to a consistent estimator.

Our main contribution is a new sampling strategy to be used within existing MCMC based rendering methods, e.g. MLT and ERPT, for rendering of scenes with heterogenous participating media. Specifically, we show that any positive and unbiased estimator of the target distribution \( E \left[ \hat{\pi}(\bar{x}) \right] = \pi(\bar{x}) \) can replace the exact quantity to simulate a Markov Chain with a stationary distribution that has a marginal that is the exact target distribution of interest. This enables us to evaluate the transmittance function with recent unbiased estimators, [Novák et al. 2014], leading to significantly shorter rendering times, see e.g. Figure 1. Our approach is based on pseudo-marginal MCMC methods developed for Bayesian inference with intractable likelihoods, [Beaumont 2003; Andrieu and Roberts 2009]. Compared to previous work (relying on biased ray-marching for evaluating transmittance), our method enables simulation of longer Markov chains, a better exploration of the path space, and consequently less image noise.

2 Background and related work

2.1 Path Integral Formulation

Using the path integral formulation of light transport [Veach 1997; Pauly et al. 2000; Wenzel 2013], the \( j \)-th pixel measurement \( I_j \) is expressed as an integral over the space, \( \Omega \), of all valid light transport paths of all lengths:

\[
I_j = \int_\Omega f(\bar{x}) W_j d\mu(\bar{x}), \text{ with } d\mu(\bar{x}) = \prod_{i=0}^{k} d\mu(x_i). \tag{1}
\]

Here a valid path of length \( k \in (0, \infty) \), \( \bar{x} = x_0, \ldots, x_k \) connects a light source to the sensor through a series of scattering vertices, \( x_i \), located on surfaces or in participating media. Here \( W_j = W_j(x_{k-1} \rightarrow x_k) \) denotes the pixel sensitivity/filter at the sensor. The path measure, \( d\mu(\bar{x}) \), is a product measure over the measures of the path vertices, where path vertices on surfaces correspond to area integration, \( d\mu(x_i) = dA(x_i) \), and vertices in a media correspond to volume integration, \( d\mu(x_i) = dV(x_i) \). The measurement contribution function, \( f_j(\bar{x}) \), is defined by

\[
f_j(\bar{x}) = L_e \prod_{i=0}^{k-1} G(x_i, x_{i+1}) T(x_i, x_{i+1}) \prod_{i=2}^{k-1} \rho(x_i). \tag{2}
\]
where \( L_e = L_e(x_0 \rightarrow x_1) \) is the emitted radiance at the light source, \( \rho(x_i) \) denotes scattering functions defined at path vertices, \( G(x_i, x_{i+1}) \) is the geometry term, and \( T(x_i, x_{i+1}) \) is the transmittance defined on the segments between path vertices. For a detailed description of these terms see e.g. [Wenzel 2013].

**Figure 2:** Illustration of the notation for describing the light transport in the scene using the path integral formulation.

### 2.2 Evaluating transmittance

The transmittance between two path vertices, \( x_i \) and \( x_{i+1} \), separated by a distance \( d \) is given by:

\[
T(x_i, x_{i+1}) = \exp \left( - \int_0^d \sigma(x + t\omega_{i,i+1}) dt \right) \tag{3}
\]

where \( \sigma(x) \) denotes the extinction coefficient describing the loss of light due to absorption and out-scattering per unit distance, and \( \omega_{i,i+1} \) denotes a unit length vector pointing from \( x_i \) towards \( x_{i+1} \). A challenge is that the transmittance only can be evaluated analytically in special cases, and is intractable for general heterogeneous media. Ray-marching methods [Perlin and Hoffert 1989] evaluate \( T \) numerically in special cases, and is intractable for general heterogeneous media. Woodcock tracking or pseudo tracking can be used. This method also includes the inner integral \( \int \sigma(x) dt \) using standard quadrature methods. However, these techniques are often computationally inefficient and the resulting transmittance estimates are biased. Instead, an algorithm related to rejection sampling, known as delta tracking, Woodcock tracking or pseudo tracking can be used. This method was developed independently in neutron transport [Woodcock et al. 1965] and plasma physics [Skulderud 1968] for sampling free-flight distances in heterogeneous media given an upper bound of the extinction coefficient, \( \bar{\sigma}(x) \geq \sigma(x), \forall x \). A detailed study of delta tracking was presented by Coleman [1968]. To evaluate the transmittance, a standard technique is to simply average \( M \) delta tracking samples and evaluate the ratio of samples corresponding to a fixed path length \( L \). Improved transmittance estimators based on de-randomization and control variate techniques have also been proposed in the literature. In this paper we use the ratio tracking estimator proposed by Novak et al. [2014] providing a noisy but unbiased estimate of the transmittance function.

### 2.3 Metropolis light transport

To evaluate the path integral (1) the MLT method, [Veach and Guibas 1997], uses MCMC to sample paths from a target probability distribution, \( \pi(\mathbf{x}) \sim \pi(\mathbf{x}) = \frac{L(\mathbf{x})}{Z}, \) where \( L(\mathbf{x}) \) is a scalar contribution function. This function is usually constructed to be similar to the luminance of the path contribution function \( f(\mathbf{x}) \). The path integral is then estimated by the average, \( \hat{f}_\pi = \frac{1}{N} \sum_{i=1}^N \frac{w_i f(x_i)}{Z} \), where \( Z = \int L(\mathbf{x}) d\mathbf{x} \) is an unknown normalizing constant that can be computed using traditional path tracing techniques. To sample paths from the target distribution, \( \pi(\mathbf{x}) \), a set of seed paths are first generated using traditional path tracing techniques. These initial seed paths are then iteratively updated with the MH algorithm to construct a Markov chain on \( \Omega \) that is ergodic and has \( \pi(\mathbf{x}) \) as its stationary distribution. Given a current path \( \mathbf{x} \), a new path, \( \mathbf{x}^{i+1} \), is first proposed using a conditional proposal density, \( q(\mathbf{x}^{i+1} | \mathbf{x}^i) \), and then accepted with probability

\[
\min \left\{ 1, \frac{L(\mathbf{x}^{i+1}) q(\mathbf{x}^i | \mathbf{x}^{i+1})}{L(\mathbf{x}^i) q(\mathbf{x}^{i+1} | \mathbf{x}^i)} \right\} \tag{4}
\]

A sufficient condition for the resulting Markov chain to be irreducible is that \( q \) is positive everywhere. In practice, this is often met by mixing independent proposals with local exploration of the state space. The ERPT, [Cline et al. 2005], method is an extension to the original MLT algorithm, which first samples a large set of paths with standard path tracing methods, and then runs many short MCMC chains starting form each initial sample. Similarly to the original MLT algorithm, MH updates are used to simulate the Markov chains. The main benefit of EPRT is that a well stratified initial sample set can be preserved, e.g. using one path per pixel.

An extended formulation of MLT for rendering scenes with participating media was presented by Pauly et al. [2000]. For heterogeneous media they used a ray-marching algorithm to evaluate the transmittance along sampled light paths, which often requires extensive computational effort and produces biased estimates. A problem (see discussion in the next section) is that if a biased estimate of the transmission is used in place of the exact quantity, the resulting Markov chain does not necessarily target the desired density \( \pi(\mathbf{x}) \). Kelemen et al. [2002] presented an alternative MCMC method. Instead of sampling directly in path space they propose to apply mutations to random number vectors in primary sample space, which are then used to generate light paths via traditional path tracing methods. We refer to this method as Primary sample space Metropolis light transport (PSSMLT). While the method is easy to implement it is limited to the underlying path tracing techniques restricting the possible exploration of the path space. Raub et al. [2008] introduced delta tracking to computer graphics and argued that it could be used to derive an unbiased path tracking algorithm that could then be used with PSSMLT. In this paper we significantly extend their work by showing that we can construct valid algorithms for the larger family of path space MLT and ERPT methods by employing unbiased estimators of the scalar contribution function.

### 3 Replacing the contribution function with an unbiased estimate

In cases where the contribution function is intractable, e.g. rendering heterogeneous media, the standard MH acceptance probability (4) cannot be evaluated. Instead, we propose to employ a pseudo-marginal construction [Andrieu and Roberts 2009], where a non-negative and unbiased estimator of the contribution function \( L(\mathbf{x}) \) is used in place of its intractable counterpart. \( L(\mathbf{x}) \). We show that this allows us to generate a Markov chain with the exact target distribution \( \frac{L(\mathbf{x})}{Z} \) as its stationary distribution although we only have access to an approximation of the contribution function. By employing this construction inside previous MH based rendering algorithms, e.g. MLT and ERPT, we can not only increase their flexibility and applicability but also their efficiency.

Let us first introduce an auxiliary random variable \( u \) denoting all the random quantities generated to construct the estimator \( \hat{L}(\mathbf{x}) \). This random variable follows a probability distribution \( g(u|\mathbf{x}) \) defined on some space \( \mathcal{U} \). Consider now the joint distribution of the light path \( \mathbf{x} \) and the auxiliary variable \( u \)

\[
p(\mathbf{x}, u) = \frac{L(\mathbf{x})}{Z} g(u|\mathbf{x}). \tag{5}
\]

This distribution has the target distribution as one of its marginals, as \( \int g(u|\mathbf{x}) du = \frac{L(\mathbf{x})}{Z} \). Inspired by this joint distribution we
can construct an extended target distribution

\[ \phi(\bar{x}, u) = \frac{\hat{L}_u(\bar{x})}{Z} g(u|\bar{x}) \] (6)

where we have replaced the exact contribution function with an unbiased estimator, \( \hat{L}_u(\bar{x}) \). This extended target distribution can now be targeted directly with a standard MH algorithm in the extended space \( \Omega \otimes \mathcal{U} \). The resulting algorithm will then in fact generate samples \( \{\bar{x}^i\}_{i \geq 1} \) from the target distribution \( \frac{\hat{L}_u(\bar{x})}{Z} \) even though we substituted the exact contribution function with an approximation in (6). To see why this holds, let us consider the marginal with respect to \( u \) of the extended target distribution

\[ \int \phi(\bar{x}, u) du = \frac{1}{Z} \int \hat{L}_u(\bar{x}) g(u|\bar{x}) du. \] (7)

As the contribution function estimate, \( \hat{L}_u(\bar{x}) \), is unbiased we have

\[ \mathbb{E}_{g(u|\bar{x})} [\hat{L}_u(\bar{x})] = \int \hat{L}_u(\bar{x}) g(u|\bar{x}) du = L(\bar{x}) \] (8)

which inserted into (7) proves that \( \int \phi(\bar{x}, u) du = L(\bar{x}) \) is recovered exactly as the marginal of the extended target distribution (6) despite the fact that we used an approximation of the contribution function. An interpretation of these, so called, exact approximation algorithms is that using \( \hat{L}_u(\bar{x}) \) does change the marginal with respect to \( u \) but it does not change the marginal with respect to \( \bar{x} \).

To use a standard MH algorithm in the extended space, a new light path \( \bar{x}^{i+1} \) and auxiliary variable \( u^{i+1} \) is proposed according to

\[ \bar{x}^{i+1} \sim q(\bar{x}^{i+1}|\bar{x}^i), \quad u^{i+1} \sim g(u^{i+1}|\bar{x}^{i+1}) \] (9)

based on the current sample pair \( [\bar{x}^i, u^i] \). Note that, the simulation of \( u^{i+1} \) simply corresponds to generating a new estimate of the contribution function independently of previous estimates. The probability of accepting the new sample pair is then given by

\[
\min \left\{ 1, \frac{\hat{L}_{u^{i+1}}(\bar{x}^{i+1}) g(u^{i+1}|\bar{x}^{i+1})}{\hat{L}_u(\bar{x}^i) g(u^i|\bar{x}^i)} \right\} = \min \left\{ 1, \frac{\hat{L}_{u^{i+1}}(\bar{x}^{i+1})}{\hat{L}_u(\bar{x}^i)} \right\} \frac{q(\bar{x}^i|x^{i+1}) g(u^i|\bar{x}^i)}{q(\bar{x}^{i+1}|\bar{x}^i) g(u^{i+1}|\bar{x}^{i+1})} \] (10)

As we are only interested in the sampled light paths \( \{\bar{x}^i\}_{i \geq 1} \) and not the auxiliary quantities, we do not need to explicitly keep track of the generated auxiliary random variables \( \{u^i\}_{i \geq 1} \). However, it is necessary to keep track of the estimated contribution of the light path \( \hat{L}_u(\bar{x}) \), and keep this value for all comparisons with other proposed paths. Note that, a key requirement for the marginal subchain \( \{\bar{x}^i\}_{i \geq 1} \) to target the desired distribution, is that the estimator \( \hat{L}(\bar{x}) \) is unbiased. Otherwise, using for example a biased estimate, the marginal of the extended target (5) is not necessarily equal to the desired distribution.

### 3.1 Rendering of heterogeneous media

To evaluate the contribution function for a path, \( \bar{x}^j \), in heterogeneous participating media, we estimate the transmission between all vertices along the path using independent ratio trackings [Novák et al. 2014], providing us with unbiased estimators \( \{\hat{T}_i\} \). By combining these estimators, we can form an unbiased estimator of the total path contribution as

\[ \hat{L}_u(\bar{x}) = L_u \left[ \prod_{i=0}^{k-1} G(x_i, x_{i+1}) \hat{T}_i(x_i, x_{i+1}) \right] \left[ \prod_{i=1}^{k-1} \rho(x_i) \right] \]

where \( X^Y \) is the luminance of the spectrally valued function \( X \). The resulting estimator \( \hat{L}_u(\bar{x}) \) is then unbiased as long as the ratio tracking estimators are unbiased and independent.

To propose new paths we use the perturbation strategies originally proposed by Veach and Guibas [1997] extended to handle participating media. To update the position of vertices in media, we use a proposal function similar to the medium perturbation proposed by Pauly et al. [2000] and later extended in the Mitsuba renderer [Jakob 2010]. Here, the medium vertex, \( x_i \), is perturbed with uniform probability either forward or backward along the current edge, with a distance proportional to \( \gamma \sigma(x) \) for all results in this paper we used a mixture of \( \gamma = 50 \) and \( \gamma = 85 \).

As the contribution function estimate, \( \hat{L}_u(\bar{x}) \), is unbiased, the marginal with respect to \( \bar{x} \) can be targeted directly with a standard MH algorithm in the extended space \( \Omega \otimes \mathcal{U} \). The resulting algorithm will then in fact generate samples \( \{\bar{x}^i\}_{i \geq 1} \) from the target distribution \( \frac{\hat{L}_u(\bar{x})}{Z} \) even though we substituted the exact contribution function with an approximation in (6). To see why this holds, let us consider the marginal with respect to \( u \) of the extended target distribution

\[ \int \phi(\bar{x}, u) du = \frac{1}{Z} \int \hat{L}_u(\bar{x}) g(u|\bar{x}) du. \] (7)

As the contribution function estimate, \( \hat{L}_u(\bar{x}) \), is unbiased we have

\[ \mathbb{E}_{g(u|\bar{x})} [\hat{L}_u(\bar{x})] = \int \hat{L}_u(\bar{x}) g(u|\bar{x}) du = L(\bar{x}) \] (8)

which inserted into (7) proves that \( \int \phi(\bar{x}, u) du = L(\bar{x}) \) is recovered exactly as the marginal of the extended target distribution (6) despite the fact that we used an approximation of the contribution function. An interpretation of these, so called, exact approximation algorithms is that using \( \hat{L}_u(\bar{x}) \) does change the marginal with respect to \( u \) but it does not change the marginal with respect to \( \bar{x} \).

To use a standard MH algorithm in the extended space, a new light path \( \bar{x}^{i+1} \) and auxiliary variable \( u^{i+1} \) is proposed according to

\[ \bar{x}^{i+1} \sim q(\bar{x}^{i+1}|\bar{x}^i), \quad u^{i+1} \sim g(u^{i+1}|\bar{x}^{i+1}) \] (9)

based on the current sample pair \( [\bar{x}^i, u^i] \). Note that, the simulation of \( u^{i+1} \) simply corresponds to generating a new estimate of the contribution function independently of previous estimates. The probability of accepting the new sample pair is then given by

\[
\min \left\{ 1, \frac{\hat{L}_{u^{i+1}}(\bar{x}^{i+1}) g(u^{i+1}|\bar{x}^{i+1})}{\hat{L}_u(\bar{x}^i) g(u^i|\bar{x}^i)} \right\} = \min \left\{ 1, \frac{\hat{L}_{u^{i+1}}(\bar{x}^{i+1})}{\hat{L}_u(\bar{x}^i)} \right\} \frac{q(\bar{x}^i|x^{i+1}) g(u^i|\bar{x}^i)}{q(\bar{x}^{i+1}|\bar{x}^i) g(u^{i+1}|\bar{x}^{i+1})} \] (10)

As we are only interested in the sampled light paths \( \{\bar{x}^i\}_{i \geq 1} \) and not the auxiliary quantities, we do not need to explicitly keep track of the generated auxiliary random variables \( \{u^i\}_{i \geq 1} \). However, it is necessary to keep track of the estimated contribution of the light path \( \hat{L}_u(\bar{x}) \), and keep this value for all comparisons with other proposed paths. Note that, a key requirement for the marginal subchain \( \{\bar{x}^i\}_{i \geq 1} \) to target the desired distribution, is that the estimator \( \hat{L}(\bar{x}) \) is unbiased. Otherwise, using for example a biased estimate, the marginal of the extended target (5) is not necessarily equal to the desired distribution.

### 3.1 Rendering of heterogeneous media

To evaluate the contribution function for a path, \( \bar{x}^j \), in heterogeneous participating media, we estimate the transmission between all vertices along the path using independent ratio trackings [Novák et al. 2014], providing us with unbiased estimators \( \{\hat{T}_i\} \). By combining these estimators, we can form an unbiased estimator of the total path contribution as

\[ \hat{L}_u(\bar{x}) = L_u \left[ \prod_{i=0}^{k-1} G(x_i, x_{i+1}) \hat{T}_i(x_i, x_{i+1}) \right] \left[ \prod_{i=1}^{k-1} \rho(x_i) \right] \]

where \( X^Y \) is the luminance of the spectrally valued function \( X \). The resulting estimator \( \hat{L}_u(\bar{x}) \) is then unbiased as long as the ratio tracking estimators are unbiased and independent.

To propose new paths we use the perturbation strategies originally proposed by Veach and Guibas [1997] extended to handle participating media. To update the position of vertices in media, we use a proposal function similar to the medium perturbation proposed by Pauly et al. [2000] and later extended in the Mitsuba renderer [Jakob 2010]. Here, the medium vertex, \( x_i \), is perturbed with uniform probability either forward or backward along the current edge, with a distance proportional to \( \gamma \sigma(x) \) for all results in this paper we used a mixture of \( \gamma = 50 \) and \( \gamma = 85 \).

### 4 Results

To evaluate the performance our new pseudo-marginal approach, we compare it to representative methods for efficient rendering of anisotropic heterogeneous participating media and glossy transfer. All renderings were generated using the Mitsuba renderer [Jakob 2010].
In Figure 1, we show equal time comparisons for our pseudo-marginal ERPT method using ratio tracking. ERPT using Simpson quadrature ray-marching, and PSSMLT with ratio tracking. Note that, Raab et al. [2008] originally proposed to use PSSMLT with delta tracking to evaluate the transmission, however here we instead use the more recent ratio tracking estimator. For the ERPT algorithms, we use the manifold perturbation by Wenzel and Marschner [2012]. For ray-marching based MLT and ERPT, the default step-size selection implemented in Mitsuba was used. The glass cube contains a diffuse torus shape and an isotropic heterogeneous media. Classical volumetric path tracing algorithms, such as bidirectional path tracing, produce very noisy results for this scene. Compared to ray-marching, the significantly faster ratio tracking enables longer Markov chains to be used in the ERPT algorithm and a better exploration of the path space resulting in less image noise. The PSSMLT algorithm has problems finding good light paths in the scene, as it lacks the ability to use the efficient local path perturbations exploited by the ERPT algorithms.

Parameter calibration An important practical consideration when using the pseudo-marginal MH update (10), is how many samples $M$ should be used to compute the unbiased estimator $L_w(\tilde{x})$. If a lower number of samples is used, the mixing of the marginal subchain, $\{\tilde{x}^i\}_{i \geq 1}$, can be poor. For example, if the the exact quantity is overestimated, the chain might get stuck at the current light path leading to a long series of rejected samples. A theoretically justified rule of thumb is to chose $M$ so that the variance of $\log(\hat{L}_w(\tilde{x}))$ is approximately one, see e.g. Doucet et al. [2012]. However, this only holds when the noise in the estimate is independent with respect to the current light path $\tilde{x}$, i.e. an assumption which often does not hold in practice. For most scenes where the MCMC based sampling methods provide a good alternative, i.e. scenes with complex light paths, we have found that using a small $M$ is beneficial. The reason is that the noise introduced by an unbiased approximation is well compensated for by the faster exploration of the path space using longer Markov chains. A comparison of different settings for the number of averaged ratio tracked, $M$, is shown in Figure 4. For all the other result shown in this paper we used a single ratio tracking, $M = 1$, to evaluate the transmission.

5 Future work

Although the examples in this paper are based on replacing the intractable transmission term with unbiased estimators, the algorithm is general. It allows for any unbiased estimator of the scalar contribution function to be used in place of the corresponding exact quantity. This also opens up the possibility of replacing a known quantity with a more efficient approximated quantity, thus enabling more iterations of the underlying Markov chain in the same computational budget. An interesting venue for future work is to investigate applications of such procedures in rendering problems using e.g. expensive shading models.

Acknowledgements This work was supported by the project Probabilistic modelling of dynamical systems (Contract number: 621-2013-5524) funded by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF) through grant IVS11-0081, Linköping University Center for Industrial Information Technology (CENIIT), and the Swedish Research Council through the Linnaeus Environment CADICS.

References


SUPPLEMENTARY MATERIAL FOR PSEUDO-MARGINAL METROPOLIS LIGHT TRANSPORT

1. Additional results

1.1. Varying $M$. Figure 1 shows equal time renderings of the classical scene by Veach filled with anisotropic scattering media ($\gamma = 0.85$), with varying number of averaged ratio trackings to estimate the transmittance. Note that some natural variations occur in the images due to the random nature of the seeding of the Markov chain.

![Figure 1](image.png)

**Figure 1.** Equal time comparisons with varying number of averaged ratio trackings used to estimate the transmission between vertices along the path. Using a low $M$ generally leads to more efficient estimators. The reference was generated using $20\times$ rendering time with $M = 1$. 

1.2. **Disco scene.** Figure 2 shows equal time renderings of a scene with heterogenous anisotropic media (resolution= $256^3$, $g = 0.85$) and glossy transfer, presenting a very difficult case for previous work.

![Figure 2. Equal time comparisons](image)