Principal-Ordinates Propagation for Real-Time Rendering of Participating Media

Oskar Elek\textsuperscript{a,b}, Tobias Ritschel\textsuperscript{a,b}, Carsten Dachsbacher\textsuperscript{c}, Hans-Peter Seidel\textsuperscript{a,b}

\textsuperscript{a}Max Planck Institut Informatik
\textsuperscript{b}MMCI Cluster of Excellence, Saarland University
\textsuperscript{c}Karlsruhe Institute of Technology

Abstract

Efficient light transport simulation in participating media is challenging in general, but especially if the medium is heterogeneous and exhibits significant multiple anisotropic scattering. We present Principal-Ordinates Propagation, a novel finite-element method that achieves real-time rendering speeds on modern GPUs without imposing any significant restrictions on the rendered participated medium. We achieve this by dynamically decomposing all illumination into directional and point light sources, and propagating the light from these virtual sources in independent discrete propagation domains. These are individually aligned with approximate principal directions of light propagation from the respective light sources. Such decomposition allows us to use a very simple and computationally efficient unimodal basis for representing the propagated radiance, instead of using a general basis such as spherical harmonics. The resulting approach is biased but physically plausible, and largely reduces the rendering artifacts inherent to existing finite-element methods. At the same time it allows for virtually arbitrary scattering anisotropy, albedo, and other properties of the simulated medium, without requiring any precomputation.

Keywords: participating media; light scattering; natural phenomena; real-time rendering; physically-based rendering; finite elements

1. Introduction

Scattering, or translucency, greatly contributes to the appearance of many natural substances and objects in our surrounding. Although the problem can be easily formulated as the radiance transfer equation \cite{3,23}, computing a solution can be very costly. Consequently, many existing approaches simplify the problem, e.g. by assuming isotropic scattering or homogeneity of the material, to achieve interactive performance.

In this work we propose a novel algorithm for plausible real-time rendering of heterogeneous participating media with arbitrary anisotropy. The core of our approach is to propagate light in propagation volumes oriented along the principal ordinates of the source illumination. For this we typically use multiple rectangular grids to propagate environmental (distant) lighting, and spherical grids to account for point light sources. In both cases, one dimension of the grids is aligned with the prominent directional part of the source radiance for which the grid has been created. In contrast to previous methods (e.g. \cite{15,11}), discretizing the illumination into directional and point light sources enables us to approximately describe the anisotropy (directionality) of light transport by a single scalar value per grid cell. Specifically, this anisotropy value corresponds to a unimodal function implicitly aligned with the respective principal ordinate. In addition to exploiting data locality and the parallelism of GPUs, the benefit of these decisions is a significant reduction of the false scattering (numerical dissipation) and ray effect (misalignment errors) artifacts arising in many finite-element methods as a consequence of representing the propagated radiance by, e.g. spherical harmonics or piecewise-constant functions. Our main contributions can be summarized as follows:

\begin{itemize}
  \item We propose the concept of Principal-Ordinates Propagation (POP), a deterministic finite-element scheme suitable for real-time simulation of anisotropic light transport in heterogeneous participating media (Sec. 3).
  \item The theory of iterative light propagation in a uniform Euclidean grid using a minimal unimodal propagation basis and explicit alignment with the illumination direction to minimize propagation artifacts and maintain light directionality (Sec. 4).
  \item An extension of the propagation scheme to handle environmental illumination by decomposing it in a set of discrete directions. This includes several new steps, namely specialized prefiltering, importance propagation, and a separate propagation of isotropic residual energy (Sec. 5).
  \item An extension to local light sources via spherical grids, enabling the integration of instant radiosity to simulate light interaction between solid objects and the medium (Sec. 6).
  \item Finally we analyse our approach in a number of diverse scenarios, demonstrating its versatility (Sec. 7).
\end{itemize}

2. Related work

\textbf{Offline methods.} A range of different approaches has been presented to compute solutions to the radiance transport equation for participating environments \cite{3,23}. However, none of the classic techniques provides a satisfying combination of generality, robustness, and, most importantly in our context, speed. Unbiased Monte-Carlo methods, such as bidirectional path trac-
We extend the work of Elek et al. [5], building primarily on the concept of DOM [3] and the more recent light propagation volumes [15, 1]. These approaches are attractive for interactive applications as their grid-based local propagation schemes allow for easy parallel implementation on contemporary GPUs. Our work also shares similarities with the finite-difference time domain method [25], however we only consider the radiance amplitude and in general concentrate on efficiency.

Figure 1: Dense smoke exhibiting strong multiple anisotropic scattering produced by a steam locomotive under complex environment illumination. Our approach renders it dynamically without any precomputations at 25 Hz (Nvidia GeForce GTX 770).

Finite-element methods. Finite-element methods, including volume radiosity [33], the discrete ordinates method (DOM) [3], light diffusion [36], and lattice-Boltzmann transport [10] handle highly multiple scattering well. However, in practice they allow only isotropic or moderately anisotropic scattering, and usually suffer from false scattering (smoothing of sharp light beams) and ray effects (selective exaggeration of scattered light due to discretized directions). Light propagation maps [9] significantly reduce the artifacts, but are still limited to rather low scattering anisotropy.

It can therefore be seen that strong scattering anisotropy is one of the main limiting factors for existing methods. This is unfortunate, as most real-world media exhibit relatively high anisotropy (Heney-Greenstein [11] coefficient $g \approx 0.9$ or more [26]). Although isotropic approximations are acceptable in some cases, this is generally not a valid assumption and one of the primary motivations for our work.

Interactive rendering. Numerous works focus on individual optical phenomena to achieve interactive or real-time performance. These phenomena include light shafts [32, 7], volume caustics [19, 21], shadows [22, 34], and clouds [2]. Various approaches can also be found in visualization literature, e.g. half-angle slicing [17] which empirically computes forward scattering for volume visualization. Sometimes precomputation is used to speed up the rendering of heterogeneous translucent objects [35, 37] or smoke using compensated ray marching [39]. In contrast, we target general multiple scattering in participating media without any precomputation or focus on a particular phenomenon.

We extend the work of Elek et al. [5], building primarily on the concept of DOM [3] and the more recent light propagation volumes [15, 1]. These approaches are attractive for interactive applications as their grid-based local propagation schemes allow for easy parallel implementation on contemporary GPUs. Our work also shares similarities with the finite-difference time domain method [25], however we only consider the radiance amplitude and in general concentrate on efficiency.

Virtually all existing variants or extensions of DOM use a single scene-aligned propagation grid, where every cell stores a representation of the directional radiance function using spherical harmonics (SH) or piecewise-constant functions. This representation is then used to iteratively calculate energy transfer between nearby cells, typically within a local 18- or 26-neighbourhood. However, this representation is only suited for moderately anisotropic scattering at best – especially for anisotropic media under complex (high-frequency) illumination. Such approach causes prominent ray effects and false scattering artifacts (see [9]). We take a different approach and propose to identify the most important light propagation directions (principal ordinates) in the scene and then use multiple propagation grids aligned with these directions, instead of a single one. This enables using a unimodal representation of the angular energy distribution around the principal direction in each grid cell.

3. Principal-Ordinates Propagation

The core idea of our method is to reduce the main drawbacks of previous grid-based iterative methods, namely false scattering and ray effects. These problems stem from the fact that the propagation domain is generally not aligned with the prominent light transport directions. We propose to remedy these issues by using propagation volumes where the propagation domain is explicitly aligned with approximate principal directions of light transport.

Furthermore, we use only a single scalar value per grid cell to describe the local anisotropy of the directional light distribution. In our scheme, we use the well-known Heney-Greenstein (HG) [11] distribution; the aforementioned value, called the anisotropy coefficient, is used to parametrize this distribution. Using principal directions implies that for more complex lighting scenarios we have to use multiple grids that sufficiently
well approximate their directionality; for local light sources we propose to use spherical grids centred around them.

These choices inherently assume that the principal directions can be derived from the initial radiance distribution and do not change strongly when light travels through the medium. However, such variation might occur if the density of the simulated medium changes abruptly. Still we deem this to be a necessary compromise if speed is the priority, and as we discuss in Sec. 4.5, violating this assumption does not cause our algorithm to fail, but only leads to a gradual decrease of accuracy.

In the following sections we first detail our concept of Principal-Ordinates Propagation for a single directional source (Sec. 4). Then we describe how to extend this scheme to environment illumination (Sec. 5) and local light sources (Sec. 6) by using multiple, importance-sampled, rectilinear and spherical propagation volumes respectively. The propagation scheme is explained using radiance as the radiometric quantity; we assume all other quantities (such as irradiance from environment maps or intensity from point lights) to be converted accordingly. All frequently-used notation is summarized in Table 1.

4. Rectilinear grids for directional light

The concept as well as the theory behind our propagation scheme can be best explained for parallel (distant) light travelling along a direction \( \mathbf{d} \) through a region in space (Fig. 2, top). For this case we discretize the space into a uniform rectilinear grid similar to DOM; however, we make sure that one of its dimensions is aligned with \( \mathbf{d} \). For every grid cell \( i \), we store the directional distribution of light and its magnitude \( L_i \) (all computations are performed independently per-wavelength, which is omitted here for brevity). The main difference to DOM is that we represent both the directional distribution of light and the phase function using the HG distribution implicitly aligned with \( \mathbf{d} \). To distinguish radiance and energy, we define the "attenuated radiance" as the radiance distribution that has passed though each cell during the propagation, and the "accumulated energy" as the total number of propagation iterations, as \( L^{\text{acc}} \). The second, accumulation grid \( L^{\text{acc}} \), is needed to accumulate the energy transported through the medium over the course of the computation. Two options are available for implementing \( L^{\text{acc}} \): We could either store the overall radiance distribution that has passed though each cell during the propagation, or alternatively store only the observer-dependent out-scattered radiance at each iteration. We opted for the second approach, because storing the entire directional distribution at each cell is much more expensive than just accumulating the outgoing radiance (which is essentially a single scalar value). Although this of course requires recomputing the solution on every observer position change, it is in agreement with our premise of a fully dynamic algorithm without relying on precomputations.

4.1. Grid initialization

At the beginning each propagation grid—which is scaled to span the entire medium (Fig. 2, top)—needs to be initialized by the incident radiance at each cell. As no scattering has been accounted for yet, the anisotropy is set to an HG coefficient of \( a_i = 1 \), an equivalent to the Dirac function in the direction \( \mathbf{d} \) (Fig. 3). That is, for every cell, we compute the transmittance \( T_i \) (from the point where light enters the medium, travelling along \( \mathbf{d} \) to \( \mathbf{x}_i \)) set to \( L_i = L_i(\mathbf{d}) \cdot T_i \). Note that this can be efficiently computed using ray marching: as our grid is aligned with \( \mathbf{d} \) we can compute the transmittance incrementally along individual "slices" of the grid in a single sweep along \( \mathbf{d} \), accessing each cell only once.
In this section, we describe how to iteratively update the grid to simulate the propagation of light. We use a propagation stencil where the radiance of each grid cell is propagated to its six direct neighbours in every iteration. Specifically, we perform a more GPU-friendly gathering-type computation of how much radiance flows into each grid cell from its neighbours based on their radiance distributions, and then combine these contributions to yield the new distribution at that cell (Fig. 4, right). In the following we denote the neighbouring source cell with index src, and the target destination cell with dst.

**Radiance magnitude contribution.** We first need to determine the amount of radiant energy that flows from cell src towards dst according to the radiance distribution in src. To this end, we efficiently integrate $L(x_{src}, \omega)$ over the solid angle subtended by dst (denoted as $\Omega_{src\rightarrow dst}$ below) using the closed form of the cumulative HG function $F_{bg}(\mu, g) = \int_{-1}^{1} f_{bg}(\mu', g) \, d\mu'$:

$$F_{bg}(\mu, g) = \frac{1 - g^2}{4\pi g} \left( \frac{1}{1 + g^2 - 2g\mu} \right)^{1/2} = \frac{1}{1 + g}. \quad (1)$$

By this we compute the radiance from src travelling towards dst using the transmittance $T_{src\rightarrow dst}$ as

$$\Delta L_{src\rightarrow dst} = L_{src} \cdot T_{src\rightarrow dst} \cdot |\phi_1 - \phi_2| \cdot (F_{bg}(\cos \theta_{src}, a_{src}) - F_{bg}(\cos \theta_2, a_{src}))$$

(2)

using the following approximate parametrization for the subtended solid angle $\Omega_{src\rightarrow dst}$ (depending on mutual positions of src and dst):

$$|\theta_1 - \theta_2| = \begin{cases} (0, \frac{\pi}{2}, \frac{3\pi}{4}) & \text{dst in front of src} \\ (\frac{\pi}{2}, \frac{3\pi}{4}, \frac{\pi}{2}) & \text{dst next to src} \\ (\frac{3\pi}{4}, \pi, 2\pi) & \text{dst behind src} \end{cases} \quad (3)$$

(see Fig. 4, left for a sample illustration of the second case of Eq. 3). Since the HG distribution is rotationally-symmetric (Fig. 4, middle) only the absolute value of the difference of the azimuthal angles $|\phi_1 - \phi_2|$ is required. Note that here the transmittance $T_{src\rightarrow dst}$ accounts *just for absorption* that affects the radiance propagation on its way from src to dst. This is because our scheme treats scattering as a decrease of anisotropy and not as an extinction process, as we show below. In practice, we take the averaged absorption coefficients $\sigma_i$ at the source and destination cells and the distance between their centres, and apply the Beer-Lambert-Bouguer law. To avoid aliasing if the resolution of the propagation grid is much smaller than the medium volume resolution, the medium parameters need to be sampled from a downscaled version of the volume. On GPU, downsampling is a very fast operation, as it corresponds to building a small number of MIP map levels (depending on the ratio between the grid resolutions, but usually 1 or 2).

**Radiance anisotropy contribution.** Similarly to absorption attenuating the radiant energy flowing between neighbouring cells, the anisotropy of the energy propagated from src to dst will decrease due to scattering. In agreement with the radiance transfer equation, in our case this can be easily computed exploiting the self-convolution property of the HG distribution [24]: in a medium with scattering anisotropy of $g$ the radiance anisotropy reduces to $a' = a \cdot g^{|\phi_1 - \phi_2|}$ after travelling a distance $t$ (assuming a constant $\sigma_i$ along this path). We obtain $\sigma_i$ and $t$ the same way as for computing $T_{src\rightarrow dst}$ above. The change of radiance anisotropy from src to dst is therefore

$$\Delta a_{src\rightarrow dst} = a_{src} \cdot g^{|\phi_1 - \phi_2|} \cdot t. \quad (4)$$

We can easily see that this formula cannot lead to an increase of anisotropy, since $g \in [0,1]$. Additionally, in non-scattering media ($\sigma_i = 0$) the directionality will be preserved perfectly.

**Combining contributions from neighbours.** Updating the radiance distribution at the cell dst entails accumulating the contributions from its six neighbours (indexed by src) as

$$L_{dst} = \sum_{src} \Delta L_{src\rightarrow dst}, \quad (5)$$

$$a_{dst} = \sum_{src} \Delta a_{src\rightarrow dst} \cdot \Delta L_{src\rightarrow dst} \cdot \Delta L_{src\rightarrow dst} \cdot \sum_{src} \Delta L_{src\rightarrow dst}. \quad (6)$$

While the radiant energy contributions simply need to be added up, the anisotropy is a weighted average of its neighbours, since the update has to yield an anisotropy value $a_{dst}$ within the valid range. We discuss implications of Eq. 6 in Sec. 4.5.
4.3. Iterating the solution

The update procedure defined by Eqs. 5 and 6 is performed for every cell of \( L^m \) to yield \( L^{m+1} \) for every iteration \( m \). Implementation-wise, this requires maintaining a second grid identical to the propagation grid and swapping these at each iteration.

Additionally, the results of every propagation iteration need to be accumulated in \( L^\text{acc} \) by evaluating the updated distributions in \( L^{m+1} \):

\[
L_{\text{acc},i}^{m+1} = L_{\text{acc},i}^{m} + L_{i}^{m+1}(\mathbf{x}_i, c - \mathbf{x}_i) \\
= f_{\text{acc},i}^{m} + L_{i}^{m+1} \cdot f_{\text{hg}}(\mu, a_{ij}^{m+1})
\]

for every cell \( i \). Here \( c \) is the observer position and \( \mu \) is therefore the dot product of \( d \) and the view direction.

4.4. Upsampling and rendering

When the solution has converged after a sufficient number of iterations, using it for rendering is relatively straightforward. We employ ray-marching to integrate the incoming radiance for every camera ray using the common front-to-back emission-absorption model \([23]\). In this case the emission term corresponds to the scattered radiance accumulated in \( L^\text{acc} \).

As we discuss in Sec. 7, the typical resolutions used for the propagation grids need to be rather small (in most of our examples \( \leq 20^3 \)) for performance reasons. In order to improve the rendering quality with such low grid resolutions it is desired to upsample them prior to their visualization. We use a 3D version of the joint bilateral upsampling \([18]\) where the density field of the medium (i.e. the spatially varying scattering coefficient) is used as a guidance signal. Typically, the density field is significantly more detailed than the propagation volumes; this detail is “transferred” to the solution by the upsampling. According to our experiments, low-resolution propagation grids are usually sufficient for plausible results.

4.5. Discussion of the propagation scheme

Using the unimodal HG function with a single parameter to represent the directional distributions in light transport obviously means that there are distributions in a cell that cannot be represented well. On the other hand, we compensate for this by using multiple grids (see Sec. 5), which in turn can handle anisotropic phase functions significantly better than previous work thanks to the proposed propagation scheme. In comparison, an extremely large number of SH coefficients is required to represent highly anisotropic distributions, and this still does not prevent false scattering issues if a local propagation scheme is employed.

The most heuristic step of our scheme is the recombinantion of reduced anisotropies from neighbouring cells in Eq. 6 (which we further elaborate on in the supplementary materials). The logic behind this formulation is that the radiance distribution at \( dst \) will result from superposing the neighbouring distributions according to how much energy they contribute to \( dst \). The main limitation of this approach lies in the fact that combining multiple HG distributions with different anisotropy values cannot generally be represented by any single HG distribution. Although we have experimented with fitting the resulting HG distribution to the combination of its neighbours in terms of least square error, we found that the simple weighted arithmetic average produces comparable results while keeping the computational cost of this core operation minimal. In addition, Eq. 6 very well preserves the anisotropy of light transported along the principal direction, thus greatly reducing false scattering effects.

Note that there are cases of very heterogeneous media where our approach might locally become inaccurate (see Fig. 5). If light along the principal direction undergoes strong absorption, while light from other directions does not, the resulting light distribution should possibly become skewed, which cannot be represented within our framework. Although this is obviously a failure case of our representation, occurrences of such strong absorption fluctuations are comparatively rare, and more importantly the resulting radiance magnitude in these cases is typically very small (therefore having little impact on the resulting image). Also note that with multiple propagation volumes we can actually reproduce complex multimodal radiance distributions, despite each grid being composed of unimodal HG distributions.

5. Multiple propagation grids for environment lighting

In the previous section we have described our approach for a single directional light source. In order to account for environmental lighting (typically modelled by an environment map), we need to use multiple grids oriented along different principal directions.

In this section we discuss how to choose these directions and, as every grid accounts for light from a finite solid angle, how to prefilter the respective incident radiance to avoid singularity artifacts (see Fig. 6). We further describe how the multiple propagation grids are combined together for rendering. Finally we present an additional (optional) step in the pipeline of our algorithm, which allows splitting the propagation into two stages,
A straightforward approach is importance-sampling the environment map to obtain $N$ directions, $d_n$, each carrying an energy corresponding to its associated portion of the directional domain $\Omega_n$. We can account for the shape of $\Omega_n$ when determining the initial directional radiance distributions (parameter $a_i$ in Sec. 4.1). Recall that the anisotropy parameter of $f_{\text{bg}}$ represents the average cosine of the distribution. We can therefore approximate the initial $a_{n,i} = f_{\Omega_n} - d_n \cdot \omega \ d\omega/|\Omega_n|$, the average cosine between $d_n$ and the directions in $\Omega_n$ and use this value for the grid initialization. In practice, $a_{n,i}$ can be approximated without the integration over $\Omega_n$ for each ordinate or even knowing the shape of $\Omega_n$. As we importance-sample the environment map, the importance of the ordinate $n$ is proportional and (up to a factor) very similar to the actual solid angle of $\Omega_n$. Therefore, we use a heuristic that maps the importance $w_n \in (0,1)$ to anisotropy as $a_{n,i} = (1 - w_n/N)^P$: important ordinates are denser in the directional domain and will have small solid angles and high anisotropy, while less important ordinates are more sparse, and will have larger solid angles and low anisotropy. The scalar factor $\beta > 0$ defines the proportionality and currently needs to be tuned empirically once for each environment map; from our experience this is a simple and quick task.

5.2. Importance propagation

The described sampling scheme can be further improved by considering how much illumination from different directions actually contributes to the image. To this end, we introduce an additional importance propagation step before sampling the environment map: we use a regular grid (perspective-warped into the camera frustum and oriented along the view direction) and propagate importance from the camera through the medium. Thanks to the duality of light transport this is equivalent to the radiance propagation as described before. The result of this propagation is a directional importance distribution stored in the grid cells. By ray-marching this grid we project the importance into the directional domain and create a directional importance map that aligns with the environment map. We then sample the environment map according to its product with the importance map. We show that in certain situations this step improves the sampling result, mainly when a low number of propagation grids is used (see Sec. 7 and Fig. 7). It is also quite cost-effective, since the directional importance function is typically very smooth and therefore only low resolutions for the propagation grid and the directional map are required (all our examples use the resolutions of $16^3$ and $32 \times 16$ respectively).

5.3. Merging multiple grids

Computing the propagation for each of the $N$ principal ordinates yields a separate, view-dependent accumulation grid $L_{\text{acc},n}$ (Sec. 4). Although it is possible to visualize these directly, this is very inefficient as each grid in the set would have to be accessed at every ray-marching step. Because of this we instead opt to combine all $L_{\text{acc},n}$ into a single medium-aligned grid, prior to upsampling and visualization.
Figure 8: Evaluation of the isotropic residuum (IR) propagation. We show the dragon dataset illuminated by two different environment maps, from two opposing viewpoints for each to demonstrate the view dependence of the resulting light distributions. The used medium (milk [26], \( \sigma_t = \{0.91, 1.07, 1.25\} \) m\(^{-1}\), \( g = 0.95\)) has an albedo over 99.9% across the visible spectrum, making it a difficult material to render because of the high number of iterations necessary to converge. The full propagation (64 ordinates, 16\(^3\) grid each) requires about 100 iterations to converge, taking 120 ms (#1) and 90 ms (#2) respectively. In comparison, our heuristic (Eq. 6) switches to isotropic propagation after \( m' = 20\) full anisotropic iterations (\( \epsilon_a = 0.1, \Delta x = 0.25 m\)). The IR propagation requires additional 20 iterations in a single 32\(^3\) propagation grid. The combined propagation time in this case was 28 ms (25 ms anisotropic and 3 ms isotropic) for illumination #1 and 24 ms (20.5 ms anisotropic and 3.5 ms isotropic) for #2; this is about a 4-fold speed-up compared to the full propagation, with negligible visual difference.

We assume the solution to be converged when all \( L_i \) are below a small threshold \( \epsilon_L \). This can however take a large number of iterations for high-albedo media, a problem inherent to all finite-element transport methods. On the other hand, we can observe that scattering reduces the anisotropy of the radiance distribution and we can treat the propagation as (near-)isotropic as soon as \( |a_{\text{ani}}| < \epsilon_a \) \( \forall i, \forall n \), for a small anisotropy threshold \( \epsilon_a \).

As soon as all propagation grids fulfill this criterion the energy from them can be merged into a single grid aligned with the medium, as there is no directionality present anymore. This is similar to merging the accumulation grids (Sec. 5.3), except that here the propagation grids are merged as well and the propagation process switches to isotropic scattering (i.e. the anisotropies \( a_{\text{ani}} \) need not be maintained anymore). This decreases the propagation costs tremendously, as from this point it is performed just for a single global grid instead of one grid per principal ordinate.

In practice, we determine the iteration \( m' \) when we can switch to the cheaper isotropic propagation based on the maximum radiance anisotropy \( \hat{a} \) (for a directional light \( \hat{a} = 1 \), but prefiltering can lower it, to our benefit), and phase function anisotropy \( g \). Both of these parameters are determined at the initialization.

Making use of the HG self-convolution property (Sec. 4.2), from these values we can approximate the distance \( t \) that light has to travel such that its anisotropy falls below \( \epsilon_a \) with

\[
\hat{a} \cdot g^{m'-1} = \epsilon_a, \tag{9}
\]

where \( \bar{s} \) is the average scattering coefficient in the medium. Furthermore, we know that the travel distance depends on the average grid spacing \( \Delta x \) and the number of iterations, i.e. \( t = m' \cdot \Delta x \), and obtain:

\[
m' = \frac{\ln \epsilon_a - \ln \hat{a}}{\ln g} \frac{1}{\sigma_t \cdot \Delta x}. \tag{10}
\]

It is however usually a good idea to use \( m' \) at least equal to the grid resolution along the propagation direction, to allow for light even from the first row of cells to sufficiently penetrate into the rest of the volume (cf. [15]). The subsequent propagation then operates on the residual isotropic radiance in the merged grid, iterating until the residual energy falls below \( \epsilon_L \). Reasonable values for \( \epsilon_a \) are around 0.1, or even higher. In fact this decision is not unlike the one made in similarity theory [38].
Here, after a certain number of scattering events the propagation switches to isotropic scattering, which is accompanied by a switch to a so-called reduced scattering coefficient. This is usually done on an empirical basis and despite the fact that using the Heney-Greenstein phase function allows us to quantify the decision better (cf. [6]) this approach is still an approximation.

Similarity theory also does not apply well to heterogeneous media. Thanks to the fact that we treat scattering as a gradual decrease of anisotropy we can transit to isotropic propagation in a well controlled manner, without changing the propagation parameters or compromising the solution accuracy (aside from small geometric misalignments caused by the grid merging). We demonstrate this in Fig. 8.

6. Radial grids for local light sources

In order to extend our method to local light sources, we use spherical grids with two angular coordinates and a radial coordinate which is again aligned with the initial principal directions of the point source (Fig. 2, bottom). To obtain more isotropic cell shapes, the spacing of shells along the radial coordinate grows exponentially (in proportion to the radial segment length at a given radius). For parametrizing the spherical domain we use the octahedron parametrization [30] mainly as it is simple, provides reasonably uniform sampling, and above all, it discretizes the domain into a 2D square where every cell has four natural neighbors (plus two along the radial axis), similar to rectilinear grids.

The resulting grid is thus topologically equivalent to rectilinear grid (except for being cyclic in the two angular dimensions) and albeit not being uniform, it allows us to approximately treat the space as locally Euclidean and obtain plausible results again using virtually the same propagation scheme as before. The main difference in the propagation is that we have to account for the quadratic fall-off: although we base our propagation on radiance, we have to explicitly compensate for the varying grid cell sizes resulting from the non-uniform shell spacing. To this end, we scale the radiance when propagating along the principal direction in proportion to the radial coordinate spacing. A sample demonstration of this propagation type for a point light in a simple homogeneous spherical medium is shown in Fig. 10.

7. Results and analysis

All results were computed on a PC with a 3.7 GHz Intel Xeon CPU, 16 GB of RAM and an NVidia GeForce GTX 770 GPU with 2 GB of VRAM. Our implementation is written in C++, using OpenGL and GLSL for the GPU code. In all our measurements we use the framebuffer resolution of 800×600 in order to let the computation time be dominated by the propagation rather than ray-marching. Resolutions of the medium density...
Reference comparisons. We first compare our approach to an unbiased Monte-Carlo reference (light tracing), as well as spherical harmonics (SH) DOM, in Fig. 11. It is apparent that the described artifacts prevent SHDOM from handling anisotropic media correctly, despite being theoretically capable to do so. In contrast, POP, despite being biased, reproduces the qualitative appearance well.

A simpler analysis for a single directional propagation is presented as well. Fig. 12 shows comparisons to the reference for a single ordinate, propagating in a simple cubic medium. Again, despite some differences in the appearance, we can see the directional radiance distributions match well. We note that this is actually a difficult case for our method, because the medium regularity and the high density gradient on the faces parallel to the propagation direction violate the alignment assumption, as discussed in Sec. 4.5. However, observe that the directional distributions produced by POP have more complex shapes than the simple Henyey-Greenstein ellipsoid lobes, since they are constituted by a superposition of such lobes. In addition, Fig. 10 shows a simplified analysis similar to this, for a point light source.

Figure 11: Comparison of our Principal-Ordinates Propagation (POP) to SHDOM and a Monte-Carlo reference (light tracing), for a smoke plume 10 m across with $\sigma_i = (2.9, 3.6, 4.2) m^{-1}$, $\sigma_a = (3.4, 3.5, 3.4) m^{-1}$ and $g = 0.9$ using the “Uffizi” environment map as illumination. For POP we used 64 and 125 principal ordinates, grid resolutions of $20^3$ and $30^2$, 10 and 30 propagation iterations, respectively. For SHDOM we have used 5 and 10 bands to represent the directional radiance distribution in each cell and the same grid resolutions. SHDOM required a strong prefiltering to avoid ringing and due to false scattering it fails to reproduce the high scattering anisotropy. Our method compares well to the reference solution, and even with real-time settings it qualitatively matches the overall appearance.

Figure 12: Detailed analysis of our method (POP) in comparison to Monte-Carlo reference (light tracing). We use a sample single-ordinate scenario, with a $16^3$ propagation domain aligned with the medium (unit homogeneous cube, $\sigma_i = 4 m^{-1}$, 100% albedo), using three representative anisotropy values. The plots compare the converged (incident) radiance distributions within a 2D horizontal slice in the middle of the domain.

Datasets are typically in the order of tens in each dimension (but effectively enhanced by procedural noise). Although the number of propagation iterations needs to be chosen empirically at the moment, in general we found that amounts similar to the propagation grid resolution along the propagation dimension is sufficient (around 10–20 in our examples). Please note that we blur the environment maps only for presentation purposes (so that the medium lighting features can be examined better) – the actual illumination is in fact sampled from the full-resolution maps. Other specific scene details are provided in the caption of each discussed figure.

Propagation behaviour. We examine the convergence of our method in Fig. 13. The setting is identical to the second case and in Fig. 12. Notice that because of an absence of absorption the propagation takes a significant number of iterations, even for the small $16^3$ grid. That is the main motivation for introducing the isotropic residuum propagation (Sec. 5.4).

One of the main shortcomings of the importance propagation is its potential temporal incoherency, mostly manifested by temporal flickering. For this reason we filter the importance map both spatially and temporarily, which, however, is not a fully robust solution to the issue. One of our main targets for future work is therefore improving this by explicitly enforcing temporal coherence when the sampled light sources relocate due to camera movement or illumination changes, similarly to, e.g. [31].
We tested our method for clouds with naturally very high scatter-
well, and that correctly handling anisotropic scattering is a key
features of the background illumination, thanks to its adaptivity
Scroring light transport in heterogeneous participating media exhibit-
ing light scattering of virtually arbitrary anisotropy. The method
does not require any precomputations, which makes it well suit-
able for simulating dynamic and evolving media without extra
considerations. Our representation also adapts to and prefilters
the incident lighting. Radiance is represented by the Heney-
Greenstein distribution, and propagated by our novel scheme in
volumes oriented along estimated principal light directions.
In general the steps of the proposed method are physically-
plausible (please refer to the supplementary materials for further

8. Discussion and conclusion

Scattering anisotropy. The shortcomings of current methods in
handling highly anisotropic scattering were the main motivation
for our work, as by far the majority of both natural and artificial
media exhibit anisotropic scattering (cf. [26]).

We tested our method for clouds with naturally very high scatter-
ing anisotropy in comparison to their isotropic versions (Fig. 15).
It can be seen that our propagation scheme handles both cases
well, and that correctly handling anisotropic scattering is a key
to reproducing such media. The same can be observed in Fig. 1,
since steam has properties similar to clouds. Interestingly, grid
resolutions as well as computation times required to render plau-

Figure 13: Convergence of our propagation scheme for the setting described in Fig. 12, with scattering anisotropy $g = 0.7$. The plots show the respective incident radiance distributions within a 2D horizontal slice in the middle of the domain (marked by the red dashed line). The observed strong forward peaks represent the unscattered energy which did not (yet) interact with the medium.

Figure 14: The smoke dataset with an increasing number of ordinates using the “kitchen” environment map ($g = 0.9, 20^3$ grid resolution, 10 propagation iterations). Accounting for importance improves the results, mainly if low numbers of principal ordinates are used. The typical setting we use (shown in the bottom-centre) takes 5 ms for importance propagation, 2 ms for determining the ordinates, 2 ms for grid initialization, 12 ms for propagation, 2 ms for residuum propagation, 4 ms for grid merging and upsampling and 5 ms for ray-marching.

Prefiltering helps to improve the rendering quality in most scen-
arios and we used it to generate all results throughout the paper. It
is particularly indispensable for media with an optical thickness
insufficient to blur the sampled illumination, e.g. as in Fig. 6,
where singularity-like artifacts would appear otherwise. Our
prefiltering removes these artifacts but still allows perceiving
features of the background illumination, thanks to its adaptivity
(as opposed to a naïve prefiltering of the source illumination).

Sensible participating media are rather insensitive to its anisotropy,
i.e. anisotropic media render roughly as fast as isotropic media.
Although a larger number of ordinates might be required to re-
produce high-anisotropy effects, this additional effort is usually
compensated by a decreased complexity of the spatial radiance
distribution, which enables using coarser propagation grids.
Our dual propagation scheme also efficiently handles optically
thick anisotropic media, as seen in Fig. 8. The initial, full prop-
gagation handles the directionally-dependent portion of radiant
energy, while the remaining isotropic residuum is rapidly propa-
gated in the second stage.

Animation. Thanks to the fully dynamic nature of our approach
we can seamlessly handle animated media without any precom-
putations or performance penalty. Fig. 16 shows several frames
of an animated smoke plume coherently rendered at real-time
framerates. In Fig. 17 we then demonstrate the dynamic interac-
tion of POP with surfaces, as described in Sec. 6.

In general, we believe to have demonstrated the versatility of
our method. Our propagation is capable of computing direct
illumination and low-order scattering effects (light shafts), as
well as arbitrary multiple scattering from directional, local and
environment illumination. Orthogonal to this, POP is capable of
simulating media with wide ranges of optical thickness, albedo
and most importantly, scattering anisotropy.

We propose a novel discrete ordinates method capable of comput-
ing light transport in heterogeneous participating media exhibit-
ing light scattering of virtually arbitrary anisotropy. The method
is particularly indispensable for media with an optical thickness
insufficient to blur the sampled illumination, e.g. as in Fig. 6,
Another characteristic inherent to all finite-element methods is their convergence rate depends not only on the propagation domain resolution but also on the optical thickness of the simulated medium; especially for high-albedo media the number of propagation iterations required for producing a converged solution might be prohibitively high. Our approach deals with this issue by using multiple superimposed, relatively small propagation grids, in which a low number of iterations is sufficient to propagate most of the radiant energy (cf. Fig. 8). Media with higher optical thickness also decrease the anisotropy of the propagated light faster, allowing us to switch to the cheaper isotropic propagation mode earlier (Sec. 5.4). The lighting frequencies resulting from the isotropic transport are by definition low and therefore a lower-resolution propagation domain is sufficient here as well.

As future work, we would like to extend our propagation to work with hierarchical or nested grids to handle higher details in media as well as illumination. In general, we believe that the effect of complex lighting on dynamic participating media is an exciting visual phenomenon that deserves more dedicated research, e.g. to better understand human perception of volumetric light or the artistic practice applied to depict it.

Acknowledgements. We would like to thank Anton Kaplanyan for inspiring discussions, Chuong Nguyen, Peter Vangorp and Yulia Gryaditskaya for reviewing early drafts of the article, and the anonymous reviewers for their valuable feedback.

Since the presented method directly relates to DOM it shares some of its basic limitations, such as handling of (surface) boundaries. In volumes with high density gradients (close to opaque surfaces) the light distribution might not be faithfully reproduced by the HG basis aligned with the initial light direction. Also the resolution of every principal grid is limited and the general limitations of discrete sampling apply: for finer details higher resolutions are required. However, the upsampling (Sec. 4.4) and prefiltering (Sec. 5.1) steps help to defer these issues and for typical volume data sets moderate propagation grid resolutions of $8^3$–$20^3$ have shown to be sufficient to handle a wide range of illumination conditions and medium properties.

Another characteristic inherent to all finite-element methods is that their convergence rate depends not only on the propagation domain resolution but also on the optical thickness of the simulated medium; especially for high-albedo media the number of iterations required for producing a converged solution might be prohibitively high. Our approach deals with this issue by using multiple superimposed, relatively small propagation grids, in which a low number of iterations is sufficient to propagate most of the radiant energy (cf. Fig. 8). Media with higher optical thickness also decrease the anisotropy of the propagated light faster, allowing us to switch to the cheaper isotropic propagation mode earlier (Sec. 5.4). The lighting frequencies resulting from the isotropic transport are by definition low and therefore a lower-resolution propagation domain is sufficient here as well.

As future work, we would like to extend our propagation to work with hierarchical or nested grids to handle higher details in media as well as illumination. In general, we believe that the effect of complex lighting on dynamic participating media is an exciting visual phenomenon that deserves more dedicated research, e.g. to better understand human perception of volumetric light or the artistic practice applied to depict it.

Acknowledgements. We would like to thank Anton Kaplanyan for inspiring discussions, Chuong Nguyen, Peter Vangorp and Yulia Gryaditskaya for reviewing early drafts of the article, and the anonymous reviewers for their valuable feedback.

Since the presented method directly relates to DOM it shares some of its basic limitations, such as handling of (surface) boundaries. In volumes with high density gradients (close to opaque surfaces) the light distribution might not be faithfully reproduced by the HG basis aligned with the initial light direction. Also the resolution of every principal grid is limited and the general limitations of discrete sampling apply: for finer details higher resolutions are required. However, the upsampling (Sec. 4.4) and prefiltering (Sec. 5.1) steps help to defer these issues and for typical volume data sets moderate propagation grid resolutions of $8^3$–$20^3$ have shown to be sufficient to handle a wide range of illumination conditions and medium properties.

Another characteristic inherent to all finite-element methods is that their convergence rate depends not only on the propagation domain resolution but also on the optical thickness of the simulated medium; especially for high-albedo media the number of iterations required for producing a converged solution might be prohibitively high. Our approach deals with this issue by using multiple superimposed, relatively small propagation grids, in which a low number of iterations is sufficient to propagate most of the radiant energy (cf. Fig. 8). Media with higher optical thickness also decrease the anisotropy of the propagated light faster, allowing us to switch to the cheaper isotropic propagation mode earlier (Sec. 5.4). The lighting frequencies resulting from the isotropic transport are by definition low and therefore a lower-resolution propagation domain is sufficient here as well.

As future work, we would like to extend our propagation to work with hierarchical or nested grids to handle higher details in media as well as illumination. In general, we believe that the effect of complex lighting on dynamic participating media is an exciting visual phenomenon that deserves more dedicated research, e.g. to better understand human perception of volumetric light or the artistic practice applied to depict it.

Acknowledgements. We would like to thank Anton Kaplanyan for inspiring discussions, Chuong Nguyen, Peter Vangorp and Yulia Gryaditskaya for reviewing early drafts of the article, and the anonymous reviewers for their valuable feedback.


