A Hybrid (Monte-Carlo/Deterministic) Approach for Multi-Dimensional Radiation Transport

Guillaume Bal[*], Anthony B. Davis[†], and Ian Langmore[‡]

May 7, 2011

Abstract

A novel hybrid Monte Carlo transport scheme is demonstrated in a scene with solar illumination, scattering and absorbing 2D atmosphere, a textured reflecting mountain, and a small detector located in the sky (mounted on a satellite or a airplane). It uses a deterministic approximation of an adjoint transport solution to reduce variance, computed quickly by ignoring atmospheric interactions. This allows significant variance and computational cost reductions when the atmospheric scattering and absorption coefficient are small. When combined with an atmospheric photon-redirection scheme, significant variance reduction (equivalently acceleration) is achieved in the presence of atmospheric interactions.

Keywords: Linear Transport; Monte Carlo; Hybrid Methods; Importance Sampling; Variance Reduction; 3D Rendering; Remote Sensing

[*] Department of Applied Physics and Applied Mathematics, Columbia University, 200 S.W. Mudd Building, 500 W. 120th Street, New York, NY, 10027, USA; +1-212-854-4731, gb2030@columbia.edu

[†] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Mail Stop 169-237, Pasadena, CA, 91109, USA; +1-818-354-0450, Anthony.B.Davis@jpl.nasa.gov

[‡] Corresponding author. Department of Applied and Applied Mathematics, Columbia University, 200 S.W. Mudd Building, 500 W. 120th Street, New York, NY, 10027, USA; +1-415-272-6321, ianlangmore@gmail.com

Contents

1 Introduction

1.1 Motivation and Background

Forward and inverse linear transport models find applications in many areas of science including neutron transport [1, 2, 3], medical imaging and optical tomography [4, 5], radiative transfer in planetary atmospheres [6, 7, 8] and in oceans [9, 10], as well as the propagation of seismic waves in the solid Earth [11]. In this paper, we focus on the solution of the forward transport problem by the Monte Carlo (MC) method with, as our main application, remote sensing (an inverse transport problem) of the atmosphere/surface system [12]. In our demonstration, light is emitted from the Sun and propagates in a complex environment involving absorption and scattering in the atmosphere and reflection at the Earth's surface before (a tiny fraction of) it reaches a narrowband detector, typically mounted on a airplane or a satellite.

The integro-differential transport equation (1) may be solved numerically in a variety of ways. Monte Carlo (MC) simulations model the propagation of individual photons along their path and are well adapted to the complicated geometries encountered in remote sensing. Photons scatter and are absorbed with prescribed probability depending on the underlying medium. The output from the simulation, e.g., the fraction of photons that hit a detector, is the expected value of a well-chosen random variable. These simulations are very easy to code, embarrassingly parallel to run, and suffer (in principle) no discretization error. The drawback is that they can be very slow to converge. MC methods converge at a rate $(variance/N)^{1/2}$ where N is the number of simulations, and the *variance* is that of each photon fired. In remote sensing, the (relative) variance is high in large part because the detector is typically small and thus most photons are not recorded by the detector. In order to be effective, even in a forward simulation, MC methods must be accelerated.

One approach to speedup MC simulations is to use quasi-Monte Carlo methods, which steepen the convergence rate from $\sim N^{-1/2}$ to a more negative exponent. However, most MC speedup efforts focus on reducing the variance of each photon. See [2, 3] or the review of more recent work on neutron transport in [14, 15, 13] and on 3D atmospheric radiative transfer in [16, 17]. See also [18] for a thorough introduction to the MC techniques, including variance reduction, used in computer graphics. In problems with a small detector, this is achieved by directing photons toward that detector, and re-weighting to keep calculations unbiased. When *survival-biasing* is used, photons have their weight decreased rather than being absorbed [2, 3].[1] Often, one uses some heuristic (such as proximity to the detector), or some function to measure the "importance" of each region of phase space. In *splitting methods* [2, 3], the photon is split into two or more photons upon identifying that a photon is in a region of high importance. The weight of each photon is then decreased proportionately. Propagating many photons with a low weight is not desirable, therefore splitting is often accompanied by *Russian roulette*. Here, if a photon enters a region of low enough importance, then the photon is terminated with a certain probability, i.e., high chance of absorption if the weight is low; in the rarer alternative outcome of the Bernoulli trial, the weight is increased to keep the simulation numerically unbiased. So there is typically a slight cost in variance to improve efficiency (by terminating low-weighted trajectories). Typically a *weight window* is used to enforce regions of low/high importance. *Source biasing* techniques change the source distribution in order to more effectively reach the detector. More generally, the absorption and scattering properties at any point can be modified, provided photons are re-weighted correctly.

It has long been recognized that the adjoint transport solution is a natural importance function [19, 2, 3, 20, 21, 22, 14, 23, 15]. One can use approximations of the adjoint solution—typically a coarse deterministic solution—to reduce variance. The result is a *hybrid* method (deterministic & MC). The AVATAR method uses an adjoint approximation to determine weight windows [22]. The CADIS scheme in [14] uses an adjoint approximation in both source biasing and weight-window determination. An adaptive technique that successively refines the solution in "important" regions, using the adjoint to designate such regions, is described in [24, 25]. In [19, 2, 3, 20], a zero-variance technique is outlined that uses the true adjoint solution to launch photons that all reach the detector with the same weight ... which happens to be the correct answer. This method is of course impractical since determining the exact adjoint solution everywhere is harder than determining some specific integral of that solution, which is usually the goal of a MC simulation. The LIFT method [20, 21] therefore uses an approximation of the adjoint solution to approximate this zero-variance method.

We adapt the zero-variance technique to the particular problem we have at hand; see Fig. 1 for the type of geometry considered in this paper. The problem we consider has a fixed, partially-reflective, complex-shaped lower boundary, and relatively large mean-free-path (MFP) in the sense that a large fraction of the photons reaching the

[1] Note the somewhat confusing terminology: On the one hand, a method is statistically biased if the expected outcome is not the intended one. On the other, the practice of re-directing photons in favorable directions and/or reducing the number of scattering events is also called biasing. In the latter case the photon has its weight adjusted so that the simulation is unbiased.

detector have not scattered inside the (optically thin) atmosphere. Calculation of the approximate adjoint solution used to emulate zero-variance techniques is difficult and potentially very costly. What we demonstrate in this paper is that partial, "localized" (in an appropriate sense) knowledge of the adjoint solution still offers very significant variance reductions. More specifically, we calculate adjoint solutions that accurately account for the presence of the boundary but do not account for atmospheric scattering (infinite MFP limit). The computation of the adjoint solution thus becomes a radiosity problem with much reduced dimensionality compared to the full transport problem. This, of course, can only reduce variance in proportion to the number of "ballistic" photons that never interact with the atmosphere. When combined with simple rules for allowing atmospheric scattering and sending some photons directly from the atmosphere to the detector, our hybrid method yields very significant variance reduction at relatively minimal cost. Furthermore, the methodology studied is applicable whenever any method is available to deterministically pre-calculate flux over any subset of paths. For instance, complex propagation of light in clouds and its importance could be pre-calculated locally and incorporated into the MC simulations in a similar fashion. This "modular" approach to the description of the adjoint solution is well-adapted to the geometries of interest in remote sensing and avoids complicated, global (hence expensive) deterministic calculations of adjoint transport solutions. Our treatment of the reflecting boundary described in detail in this paper is a first step toward modular adjoint transport calculations and their variance reduction capabilities in remote sensing.

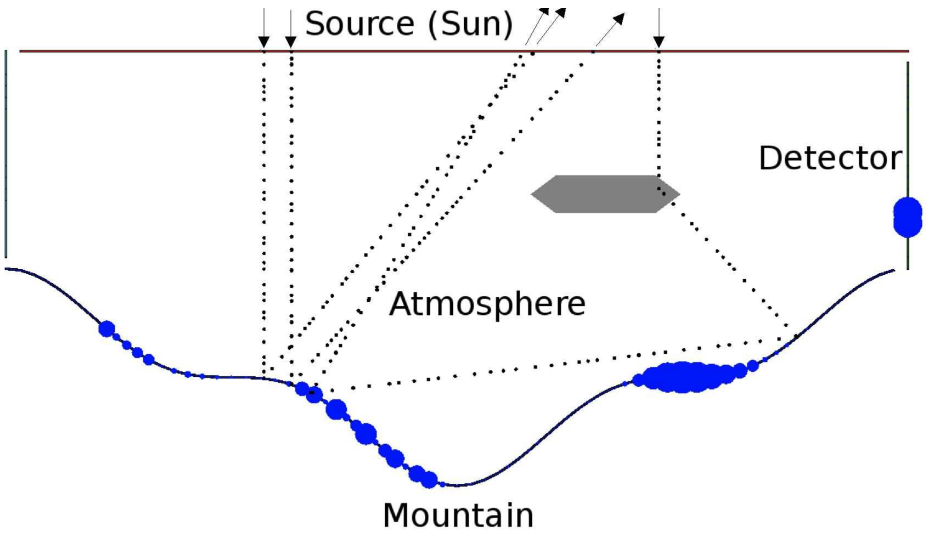

Figure 1: Mountain ($1 - \cos^3 x$ shape), cloud, sky, and detector. Dot size indicates relative adjoint flux strength. Large dots on right-hand-side are the detector (dot size is down-scaled for detector). Dot size on mountain indicates that portions of the mountain are shaded from the detector, and that the surface albedo is varying. See section 3.1 for specifics, as used in the present study.

The rest of the paper is organized as follows. In sections 1.2 and 1.3 respectively, the physical problem and statistical formulation are described. In section 1.4, the analog and survival-biased MC algorithms are presented. In section 2, the surface adjoint importance (SAI) and regularized SAI methods are introduced. These are the hybrid adjoint-based methods at the core of this work. In section 3, numerical estimations of variance reductions and computational speedups are given. For an expanded exposition of techniques and analyses from a mathematical standpoint, we refer the interested reader to [26]. Finally, we summarize our findings in section 4 and conclude with thoughts about potential applications in remote sensing science.

1.2 Problem Setup

Our setup is photon transport in a domain $R \subset \mathbb{R}^d$ ($d = 2, 3$, with $d = 2$ in our present demonstration) described in Fig. 1. The outward normal to the domain boundary ∂R at position r is denoted ν_r. R is the atmosphere, the sky/mountain/sides/detector constitute ∂R. We have one small detector located on the right-side of ∂R, and the goal of simulations is to estimate the photon flux through the detector. Since photons are monokinetic, propagation direction v is a unit vector in the sphere $\mathbb{S}^{d-1}$ embedded in d-dimensional space (unit circle for $d = 2$).

We model radiance (a.k.a. specific intensity or angular photon flux density) $I(r, v)$ in our medium with a boundary source distribution Q. I obeys the following integro-differential transport equation and boundary condition:

$$\begin{aligned} & v \cdot \nabla I(r,v) + \sigma(r) I(r,v) = KI(r,v) \\ & \qquad\qquad I(r,v) = \frac{KI(r,v)}{|\nu_r \cdot v|} + \frac{Q(r,v)}{|\nu_r \cdot v|}, \qquad r \in \partial R, \text{ and } v \cdot \nu_r < 0, \end{aligned} \tag{1}$$

the integral operators being defined by kernels

$$\begin{aligned} Kf(r,v) &= \sigma_s(r) \int_{\mathbb{S}^{d-1}} p(r, v' \to v) f(r, v') \,\mathrm{d}v', && r \in R \\ Kf(r,v) &= \alpha(r) \int_{\nu_r \cdot v' > 0} P(r, v' \to v) |\nu_r \cdot v'| f(r, v') \,\mathrm{d}v' && r \in \partial R, \text{ and } v \cdot \nu_r < 0. \end{aligned} \tag{2}$$

The extinction coefficient, a.k.a. total cross section (per unit of volume) $\sigma(r)$ is the sum of the intrinsic absorption coefficient/cross-section $\sigma_a(r)$ and the scattering coefficient/cross-section $\sigma_s(r)$. For the partially reflecting boundary condition (viewed here as a surface scattering), $\alpha(r)$ is the local value of the albedo. Both volume ($p(r, v' \to v)$) and surface ($P(r, v' \to v)$) phase functions are normalized ($\int p \,\mathrm{d}v' = 1$).

Since the transport problem is linear, we use a normalized boundary source, i.e.,

$$\int_{\partial R} \int_{\nu_r \cdot v < 0} Q(r,v) \,\mathrm{d}\mu(r) \,\mathrm{d}v = 1 \tag{3}$$

where $\mathrm{d}\mu(r)$ is the appropriate measure on the $(d-1)$-dimensional boundary.

Our detector measures photon flux and is described by a "response function," $g(r, v)|\nu_r \cdot v|$, where $g(r, v)$ is zero everywhere except when r is in the physical detector (its aperture or "pupil") and v points out of the boundary. Where $g \neq 0$ it is constant and, furthermore, it is normalized so that $\int g(r, v) \,\mathrm{d}r \,\mathrm{d}v = 1$. The goal of our Monte Carlo method is to compute the detector's signal

$$\int_{\partial R} \int_{\nu_r \cdot v > 0} g(r,v) |\nu_r \cdot v| I(r,v) \,\mathrm{d}\mu(r) \,\mathrm{d}v. \tag{4}$$

In the present study, any v can contribute to the radiometric signal measured at r. To model an imaging detector, direction space would be limited to a finite field-of-view that would in turn be subdivided into individual "pixels."

For future use, we define the function

$$E_\sigma(r, r') := \exp\left\{ -\int_0^{|r-r'|} \sigma(r + t\widehat{r'-r})\,\mathrm{d}t \right\},$$

where $\widehat{r'-r} := (r'-r)/|r'-r|$. Physically, it describes the probability of *direct* transmission of light from point r to point r' (or vice-versa), that is, without suffering any collision.

1.3 Statistical formulation/notation

The measurement defined formally in (4) is approximated in a Monte Carlo simulation by estimating an average

$$S_N := \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_D(\omega_n)$$

where the ω_n are photon paths ($\omega = (r_0, r_1, \ldots, r_k)$) generated by the "analog" chain (meaning analogously to real-life photon travel, cf. Algorithm 1), and the relevant indicator function is $\mathbf{1}_D(\omega) = 1$ if the path hits the detector (hence the subscript D), and $= 0$ otherwise. The paths are random variables and, with $\mathbb{E}\{\cdot\}$ denoting statistical *expectation*, we have

$$\mathbb{E}\{S_N\} = \mathbb{E}\{\mathbf{1}_D\} = \mathrm{P}[D],$$

where the notation $\mathrm{P}[D]$ emphasizes that this is a probability of hitting the detector. We also have $\mathrm{P}[D]$ equal to the desired measurement or signal in (4). For finite N, S_N is not equal to $\mathrm{P}[D]$ exactly. The mismatch is quantified in a statistical sense through the *variance*

$$\mathrm{Var}\{S_N\} := \mathbb{E}\left\{(S_N - \mathrm{P}[D])^2\right\} = \frac{\mathbb{E}\left\{(\mathbf{1}_D - \mathrm{P}[D])^2\right\}}{N} = \frac{\mathrm{Var}\{\mathbf{1}_D\}}{N},$$

since all of the events ω_n contributing to the S_N estimator are independently drawn.

Rather than S_N, one may generate paths according to some modification of real-life photon travel and then estimate

$$\mathrm{P}[D] \approx T_N := \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_D(\omega_n)\frac{\mathrm{dP}^{\mathrm{a}}}{\mathrm{d}\overline{\mathrm{P}}}(\omega_n),$$

where the ratio $\mathrm{dP}^{\mathrm{a}}/\,\mathrm{d}\overline{\mathrm{P}}(\omega)$ is the ratio of the probability density of ω in the analog chain to that in the modified chain. This *importance sampling* technique is widely used in statistics since often times T_N will have lower variance than S_N. Indeed, most of the variance reduction techniques mentioned in the introduction are of this type. In our algorithms we compute this ratio step-by-step and refer to it as a weight (modifier). So, rather than counting photons, we count weighted photons.

For future use we define the following (standard) statistical notations and convention. First, we write $u \sim \mathcal{U}[0,1]$ to indicate that u is a random variable uniformly distributed on the interval $[0,1]$. Second, a probability density such as $\pi(x)$ can be denoted explicitly (e.g., $\pi(x) = (2\pi)^{-1/2} \exp\left\{-x^2/2\right\}$ in the case of the normal distribution with zero mean and unit variance), or it can be given up to a constant (since it must integrate to one). In this last case, we would write $\pi(x) \propto \exp\left\{-x^2/2\right\}$.

1.4 Standard Algorithms

We present here two basic algorithms for Monte Carlo transport. These are well known but we do this in order to demonstrate our notation. Algorithm 1 is often referred to as *analog* since the photons follow a path analogous to photons in the real world.

Algorithm 1 Analog Monte Carlo Transport

```
Choose a starting position/direction $(r_0, v_0)$ according to the sun's source density $Q(r, v)$
Draw $u \sim \mathcal{U}[0,1]$ and cast the photon along the ray $r_0 + tv_0$ until $E_\sigma(r_0 + tv_0) < u$. Call this point $r_1$. If this does not happen before $\partial R$ is reached then set $r_1$ to the boundary point at the intersection with the ray.
if $r_1 \in R$ then
    With probability $\sigma_s(r_1)/\sigma(r_1)$, the photon is not absorbed, and we select $v_1$ using the probability density
        $v_1 \mapsto p(r_1, v_0 \to v_1);$
    otherwise the chain is stopped.
else if $r_1 \in \partial R$ then
    With probability $\alpha(r_1)$ the photon is not absorbed, and we select $v_1$ using the probability density
        $v_1 \mapsto P(r_1, v_0 \to v_1);$
    otherwise the photon is absorbed and we stop the chain.
end if
Continue alternating casts and direction changes until either the photon is absorbed, escapes through the upper boundary ("sky+sides"), or the detector is reached.
```

For use in Algorithm 5 further on, we will need to know the probability density of the analog chain producing a path ω. This is given by

$$\begin{aligned} D_{analog}(r_0, r_1) &= Q(r_0, v_0) E_\sigma(r_0, r_1), \\ D_{analog}(r_0, r_1, r_2) &= D_{analog}(r_0, r_1) K^{analog}(r_1, v_0 \to \widehat{r_2 - r_1}) E_\sigma(r_1, r_2) \end{aligned}$$

and so on. Above K^{analog} is given by

$$K^{analog}(r_1, v_0 \to v_1) := \begin{cases} \sigma_s(r_1) p(r_1, v_0 \to v_1), & r_1 \in R \\ \alpha(r_1) P(r_1, v_0 \to v_1), & r_1 \in \partial R. \end{cases}$$

Algorithm 2 uses a trick known as *survival-biasing* since photons will survive (almost) any interaction with the media. We do this by casting photons while ignoring intrinsic absorption. So, e.g., if a patch of media has σ_a large and $0 < \sigma_s/\sigma \ll 1$ the photon will almost never scatter there. Our weight is then E_σ/E_{σ_s}. When the photon interacts with the surface, then so long as $\alpha > 0$, we do not absorb but multiply the photon weight by α. Another, slightly different but also common, survival-biasing method would cast photons in the same manner as analog, but would eliminate absorption and re-weight by σ_s/σ. So, e.g., if a patch of media has σ_a large and $0 < \sigma_s/\sigma \ll 1$ the photon would likely interact with the media and scatter but not be absorbed there; its weight however would be reduced by a factor of σ_s/σ (known in the radiative transfer literature as the "albedo for single scattering").

Algorithm 2 Survival-Biased Monte Carlo Transport

```
Choose a starting position/direction $(r_0, v_0)$ according to the source density $Q(r, v)$
Draw $u \sim \mathcal{U}[0, 1]$ and cast the photon along the ray $r_0 + tv_0$ until $E_{\sigma_s}(r_0 + tv_0) < u$. Call
  this point $r_1$. If this does not happen before $\partial R$ is reached then $r_1$ is the boundary point
  we have reached. Since we paid no attention to intrinsic absorption during the cast, the
  photon picks up a weight equal to

      $$E_{\sigma_a}(r_0, r_1) = \frac{E_\sigma(r_0, r_1)}{E_{\sigma_s(r_0,r_1)}}$$

if $r_1 \in R$ then
    Select $v_1$ using the probability density

        $$v_1 \mapsto p(r_1, v_0 \to v_1).$$

else if $r_1 \in \partial R$ and $\alpha(r_1) > 0$ then
    Select $v_1$ using the probability density

        $$v_1 \mapsto P(r_1, v_0 \to v_1).$$

    Since we had no chance of boundary absorption, the photon's weight is multiplied by
    $\alpha(r_1)$.
else if $r_1 \in \partial R$ and $\alpha(r_1) = 0$ then
    The photon is absorbed and we stop the chain.
end if
Continue alternating casts and direction changes until either the photon is absorbed,
  escapes, or reaches the detector.
```

2 The Surface Adjoint Importance (SAI) Method

The SAI method uses an approximation to the surface reflection problem to reduce variance coming from surface interactions. It ignores atmospheric effects and therefore, by itself, is statistically biased. In section 2.2 we pair it with other methods to produce an unbiased estimate of the detected flux.

2.1 Pure SAI

Here we ignore atmospheric effects and demonstrate and develop a Monte Carlo method that sends photons from surface point to surface point and then to the detector. If atmospheric effects are not present, and our deterministic solution was perfectly accurate, this method would have zero variance.

The adjoint solution to transport may be developed by considering the L^2 adjoint of the integral solution to transport and reversing the role of the source and detector. Let I^s be the adjoint solution when only surface effects are present. We therefore have

$$I^s(r,v) = \alpha(r) \int_{\nu_r \cdot v' < 0} P(r, v \to v') I^s(r_+(r,v'), v')\, \mathrm{d}v' + g(r,v). \tag{5}$$

This adjoint solution corresponds (in a Monte Carlo viewpoint) to sending photons that start at the detector and travel backwards. Therefore, it will have its maximum at the detector. It will be higher in places that have a clear path to the detector. I^s will be zero at places from which a photon cannot reach the detector. Our numerical solution of (5) is described in the Appendix.

The pure SAI chain is defined by the steps described in the following Algorithm 3.

Algorithm 3 Pure SAI

```
1: Choose a starting position/direction $(r_0, v_0)$ according to the modified source density
       $Q^{sai}(r,v) \propto Q(r,v) I^s(r_+(r,v), v).$
   The photon picks up a weight $Q(r_0,v_0)/Q^{sai}(r_0,v_0)$
2: Cast the photon until it hits the opposing boundary at point/direction $(r_1, v_1) = (r_+(r_0, v_1), v_1)$. The weight is multiplied by
       $\frac{E_\sigma(r_0, r_1)}{1}.$
3: Change direction according to the density
       $K^{sai}(r_1, v_1 \to v_2) = C_{sai}(r_1) P(r_1, v_1 \to v_2) \frac{I^s(r_+(r_1, v_2), v_2)}{I^s(r_1, v_1)}$
   where $C_{sai}(r)$ is a normalization factor depending only on $r \in \partial R$. Since we did not account for boundary absorption, the photon weight is multiplied by $\alpha(r_1)$. The modified direction change must also be taken into account and therefore, in addition, the weight is multiplied by
       $\frac{K^{analog}(r_1, v_1 \to v_2)}{K^{sai}(r_1, v_1 \to v_2)}.$
4: Cast the photon until it hits the opposing boundary. If it hits the detector, stop and record a hit. Else, repeat step 3.
```

For use in Algorithm 5, we will need to compute the probability density of a path $\omega = (r_0, \ldots, r_\tau)$ being generated by pure SAI. This is simply the denominator in the corresponding weights. Denoting this by D_{sai} we have

$$D_{sai}(\omega) = 0, \qquad \text{if } r_j \in R \text{ for any } j,$$

and for paths such that $r_j \in \partial R$ for all j, we define D_{sai} recursively (with $v_j := \widehat{r_{j+1} - r_j}$)

$$\begin{aligned} D_{sai}(r_0, r_1) &= S^{sai}(r_0, v_0) \\ D_{sai}(r_0, r_1, r_2) &= D_{sai}(r_0, r_1) K^{sai}(r_1, v_0 \to v_1), \\ D_{sai}(r_0, \dots, r_k) &= D_{sai}(r_0, \dots, r_{k-1}) K^{sai}(r_{k-1}, v_{k-2} \to v_{k-1}). \end{aligned}$$

Remark 2.1.

- A discretized version of the density Q^{sai} is pre-computed using the (discrete) solution I^s; see section 3. This means that we can pre-compute the normalization factor C_{sai}. The discrete density Q^{sai} will be defined at a number of points (r_i, v_0) where v_0 is the anti-solar direction. We use the density to decide on a center point r_i, and then perturb the starting point by a small (random) amount to eliminate discretization effects in the final solution.
- The direction change pdf is also pre-computed and stored as a discrete pdf over angles. We use the pdf to pick a direction center v_j and then perturb to obtain the new direction.
- Up to numerical error one can see that $C_{sai}(r) = \alpha(r)$. Indeed, dividing (5) through by $I^s(r, v)$ we have

  $$1 = \alpha(r) \int_{\nu_r \cdot v' < 0} P(r, v \to v') \frac{I^s(r_+(r, v'), v')}{I^s(r, v)} \, \mathrm{d}v' + \frac{g(r, v)}{I^s(r, v)}.$$

  So away from the detector $g(r, v) = 0$ and the integral is therefore equal to 1.
- That this method is biased is easy to see: If a region of the atmosphere has non-zero scattering, then it would be possible (in the analog world) to scatter from that point to the detector. This type of interaction is not allowed in a pure SAI world.

This is an implementation of the zero variance adjoint-based chains studied in [2, 3, 20] in the special case where atmospheric effects are not present. Hence (disregarding numerical error), this would be a zero-variance method were atmospheric absorption/scattering absent.

We verify this claim numerically by testing the method in simulations without atmospheric effects. See Fig. 2 where this is tested with both a flat terrain and a curved "$\cos^3$" mountain. The curved mountain increases variance since discretization does not allow the function $r_+(x, v)$ to be implemented perfectly.

2.2 Regularized SAI

Here we use the SAI chain as part of a larger unbiased chain. Since Algorithm 3 does not generate paths following all possible interactions, we must supplement it with an algorithm that does. We then use a number $q_s \in [0, 1]$ to determine the fraction of

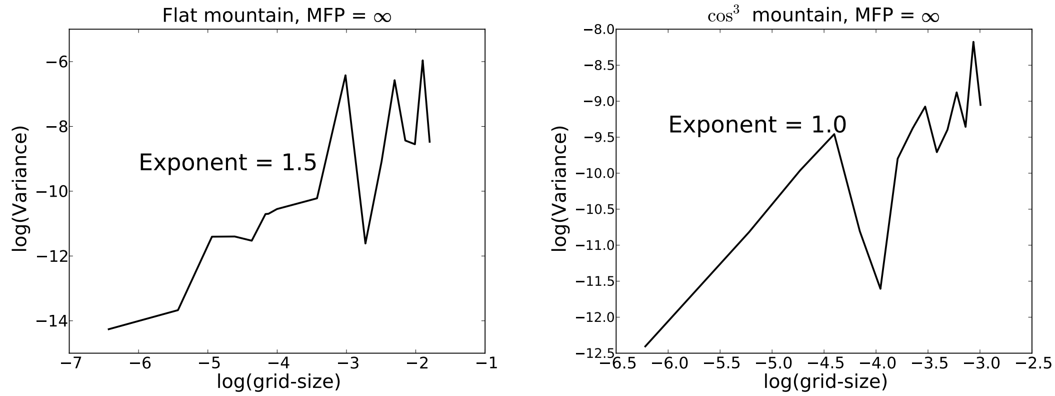

Figure 2: Left: When a flat mountain is used, variance $\sim O(h^{1.5})$ where h is the discretization parameter. Right: On the curved boundary discretization effects are more prevalent and convergence is slower.

photons that travel according to Algorithm 3 (this fraction $= 1 - q_s$), and what fraction according to the supplemental algorithm.

Before describing the regularized SAI algorithm, we present the supplemental Algorithm 4 dubbed "heuristic scattering adjustment." It is a survival-biased algorithm in the sense that no absorption occurs within the atmosphere, or at boundary points (unless the boundary point had $\alpha = 0$, e.g. the sides/sky). It also makes use of a simple scheme to direct a fraction of atmospheric interactions toward the detector. No claim is made to the optimality of this re-direction (it is similar to the technique of *local estimation* [27, 16]). We use Algorithm 4 since it is simple to understand and illustrates the dramatic decrease in variance that can be achieved when two methods (SAI and heuristic) are used together in Algorithm 5 (see also section 3).

Algorithm 4 Heuristic Scattering Adjustment with Parameter $q_v \in [0, 1]$

```
Choose a starting position/direction according to the standard source density $Q(r, v)$
Cast the photon as in Algorithm 2 until it hits the opposing boundary or interacts with
the atmosphere at $r_1$. The photon picks up a weight equal to $E_\sigma(r_0, r_1)/E_{\sigma_s}(r_0, r_1)$
if $r_1 \in R$ then
    With $r_{d_0}$ =“the midpoint of the detector”, compute
        $q_{heu}(r_1, v_0) := 1 - (1 - q_v)\frac{p(r_1, v_0 \to \widehat{r_{d_0} - r_1})}{\|p(r_1, v_1 \to \cdot)\|_{L^\infty}}$
    With probability $1 - q_{heu}$ draw $v_1$ from a uniform distribution of directions pointed
    toward the detector (we call this $f_V(r_1, v_1)$), and with probability $q_{heu}$ draw $v_1$ from
    $p(r_1, v_0 \to \cdot)$. The weight is multiplied by
        $\frac{\sigma_s(r_1)p(r_1, v_0 \to v_1)}{(1 - q_{heu})f_V(r_1, v_1) + q_{heu}p(r_1, v_0 \to v_1)}, \qquad \text{if } r_1 \in R.$
else if $r_1 \in \partial R$ and $\alpha(r_1) > 0$ then
    pick a new direction according to the density $P(r_1, v_0 \to v_1)$. The weight is multiplied
    by $\alpha(r_1)$.
else if If $r_1 \in \partial R$ and $\alpha(r_1) = 0$ then
    the photon is absorbed and we stop.
end if
Continue in this manner until absorption or the detector is reached
```

So at every scattering event, the weight is modified by a ratio of either $\alpha(r)P$ or $\sigma_s(r)p$ to K^{heu} where

$$K^{heu}(r, v \to v') := \begin{cases} (1 - q_{heu})f_V(r, v) + q_{heu}p(r, v \to v'), & r \in R \\ P(r, v \to v'), & r \in \partial R. \end{cases}$$

For use in Algorithm 5 we will need to compute the probability density of a given path generated by Algorithm 4. This is simply the denominator in the corresponding weight. Denote this by $D_{heu}(r_0, r_1, \ldots, r_k)$, which we define recursively by (with $v_j := \widehat{r_{j+1} - r_j}$)

$$\begin{aligned} D_{heu}(r_0, r_1) &= Q(r_0, v_0)E_{\sigma_s}(r_0, r_1), \\ D_{heu}(r_0, r_1, r_2) &= D_{heu}(r_0, r_1)K^{heu}(r_1, v_0 \to v_1)E_{\sigma_s}(r_1, r_2), \\ D_{heu}(r_0, \ldots, r_k) &= D_{heu}(r_0, \ldots, r_{k-1})K^{heu}(r_{k-1}, v_{k-2} \to v_{k-1})E_{\sigma_s}(r_{k-1}, r_k), \end{aligned} \tag{6}$$

and so on.

We now present Algorithm 5, the regularized SAI algorithm that combines pure SAI (Algorithm 3) with the heuristic scattering adjustment (Algorithm 4). Note that any unbiased algorithm may be combined with pure SAI in a similar manner.

Algorithm 5 Regularized SAI with parameters $q_s, q_v \in [0,1]$

```
1: With probability $1-q_s$, generate a path according to Algorithm 3. With probability $q_s$
   generate it according to Algorithm 4.
2: The weight of the path $\omega = (r_1, \ldots, r_\tau)$ is
       $$\frac{D_{analog}(\omega)}{(1-q_s)D_{sai}(\omega) + q_s D_{heu}(\omega)}.$$
```

Algorithm 5 uses SAI to produce paths that interact only with the surface. One could easily devise other algorithms that send paths via the heuristic chain, and once paths interact with the surface they use the SAI chain. This could reduce variance further, but we choose not to study this in order to simplify the presentation.

3 Numerical Results

3.1 Parameter choices in numerical simulations

In the assumed $d = 2$ transport space, we have $r = (x, y)$, where x increases from left to right in Fig. 1 and y increases from bottom to top; $r = (0, 0)$ is the point at the bottom of the valley. For directions, we have $v = v(\phi) = (\cos\phi, \sin\phi)$ where ϕ increases counterclockwise from the $x > 0$ axis.

In the simulations performed with $\sigma = 0$ (no atmospheric interactions), we used both a flat surface (so that our domain was $[-\pi, \pi] \times [2, 4]$) and a "$\cos^3$" surface (Fig. 1). We swept h, with $0.002 < h < 0.2$. We did not use any heuristic scattering adjustment ($q_v = 1.0$). In all cases, we assume an isotropic (Lambertian) redistribution by diffuse surface reflection. This leads to the following surface scattering phase function and assumed surface albedo distribution:

$$P(r, v \to v') \propto \begin{cases} |\nu_r \cdot v'|, & \nu_r \cdot v' < 0 \\ 0, & \text{otherwise} \end{cases}; \qquad \alpha(r) = \begin{cases} 1, & |x| < 2.5 \\ 0, & \text{otherwise} \end{cases}.$$

The cutoff $|x| < 2.5$ was done to simplify the coding (allowed us to use one simple routine for all values of h), and has no theoretical consequence. The source was monodirectional $\phi = -\pi/2$, and given by

$$Q(r, v(-\pi/2)) = \begin{cases} 1/5 & |x| < 2.5,\ y = 4 \\ 0 & \text{otherwise.} \end{cases}$$

In the simulations involving atmospheric interactions ($\sigma > 0$), we used a $\cos^3$ type surface. We compute speedup in a variety of cases. The mean-free-path MFP$= \sigma^{-1}$ was varied as well as q_s, h, and q_v. We swept $0.002 < h < 0.15$. In all cases the atmospheric scattering coefficients were constant with $\sigma_s = 2\sigma_a$ (hence $\sigma_s/\sigma = 2/3$). The atmospheric scattering was given by

$$p(r, v \to v') \propto 1 + (v \cdot v')^2,$$

which mimics a molecular (Rayleigh) in $d = 2$. The other coefficients were chosen to have features (in this case oscillations) on a scale coarser than the fine values of h, and finer than the coarse values.

The surface albedo was chosen to be quite complex (significantly different than the flat surface/constant reflection commonly used). The phase function P was as before (Lambertian), but α is given by

$$\alpha(r) = \begin{cases} 0 & |x| > 2.5, \\ 0.75 + 0.25\sin(2\pi x/0.05) & 1 < x < 2.5, \\ 0.35 + 0.25\sin(2\pi x/0.05) & -2.5 < x < 1, \end{cases}$$

using the same inconsequential cutoff $|x| < 2.5$. Off the mountain there was no scattering (perfectly absorbing boundary).

The source was mono-directional $\phi = -\pi/2$ and given by

$$Q(r, v(-\pi/2)) \propto \begin{cases} 1 + 0.25\sin(2\pi x/0.07) & |x| < 2.5,\ y = 4 \\ 0 & \text{otherwise} \end{cases}.$$

3.2 Speedup (figure of merit)

We start by defining our figure of merit used to compare the different algorithms. We take the viewpoint that each algorithm produces a sequence of paths $\{\omega^n\}_{n=1}^N$ and corresponding random variables $\xi(\omega^n)$ equal to the product of $\mathbf{1}_D(\omega^n)$ times the weight that the photon picked up along the way. To distinguish different methods we write ξ_a for analog, ξ_{sb} for survival-biasing, ξ_{sai} for pure SAI, ξ_{heu} for heuristic scattering, and ξ_q for the regularized SAI method.

For all of these methods, define the approximation after N random draws

$$I_N(\xi) := \frac{1}{N}\sum_{n=1}^{N} \xi(\omega_n).$$

For ξ equal to any of the above methods, $I_N(\xi)$ is an unbiased estimator of $\mathbb{E}\{\xi\} = \mathrm{P}[D]$, i.e., the probability of a detector hit.

The RMS estimation error ε is given by

$$\varepsilon(\xi) := \sqrt{\mathbb{E}\{|I_N(\xi) - \mathrm{P}[D]|^2\}} = \sqrt{\frac{\mathrm{Var}\{\xi\}}{N}}.$$

For a given error level ε, the required number of MC draws is then $N(\varepsilon, \xi) := \mathrm{Var}\{\xi\}/\varepsilon$. The required simulation time $T(\varepsilon, \xi)$ for one estimation of $\mathrm{P}[D]$ is given by

$$T(\varepsilon, \xi) := T_0(\xi) + \tau(\xi)N = T_0(\xi) + \frac{\tau(\xi)\mathrm{Var}\{\xi\}}{\varepsilon^2},$$

where $T_0(\xi)$ is the time needed to compute the deterministic adjoint solution (e.g. at level h when $\xi = \xi_h$), and $\tau(\xi)$ is the expected time for one draw using the appropriate measure for the random variable ξ. We foresee the use of SAI in situations where the boundary remains fixed, but the atmosphere changes (due to, e.g., moving clouds over a fixed surface). We therefore consider the time for m simulations using one boundary,

$$T(\varepsilon, \xi, m) := T_0(\xi) + m\tau(\xi)N = T_0(\xi) + m\frac{\tau(\xi)\mathrm{Var}\{\xi\}}{\varepsilon^2},$$

Schemes may be compared with the ratio

$$\frac{T(\varepsilon, \xi_1, m)}{T(\varepsilon, \xi_2, m)} = \frac{\varepsilon^2 T_0(\xi_1) + m\tau(\xi_1)\mathrm{Var}\{\xi_1\}}{\varepsilon^2 T_0(\xi_2) + m\tau(\xi_2)\mathrm{Var}\{\xi_2\}}.$$

For a deterministic approximation of I^s, we expect $T_0(\xi) \approx C(\xi)h^{-2(d-1)}$. We in fact measure (with $d = 2$) $T_0(\xi_h) \approx 0.017h^{-2}$. Our "benchmark" scheme is survival-biasing. Since ξ_{sb} requires no deterministic solution, the relevant ratio (and our figure of merit) is

$$\mathrm{Speedup}(\xi_q, \varepsilon, m) := \frac{m\tau(\xi_{sb})\mathrm{Var}\{\xi_{sb}\}}{\left(\frac{\varepsilon}{h}\right)^2 C + m\tau(\xi_q)\mathrm{Var}\{\xi_q\}}.$$

We measured speedup when either $m = 10$ or, formally, $m = \infty$ ("Ignoring deterministic solve").

3.3 Variance reduction

Here we analyze the variance of the SAI chain in the presence of atmospheric interactions. Note that even when the error $|\mathrm{P}[D] - \langle I^s, S\rangle|$ is high, we still get good variance reduction. See Fig. 3. This emphasizes the point that the quality of the deterministic solve is not so important in a modular scheme.

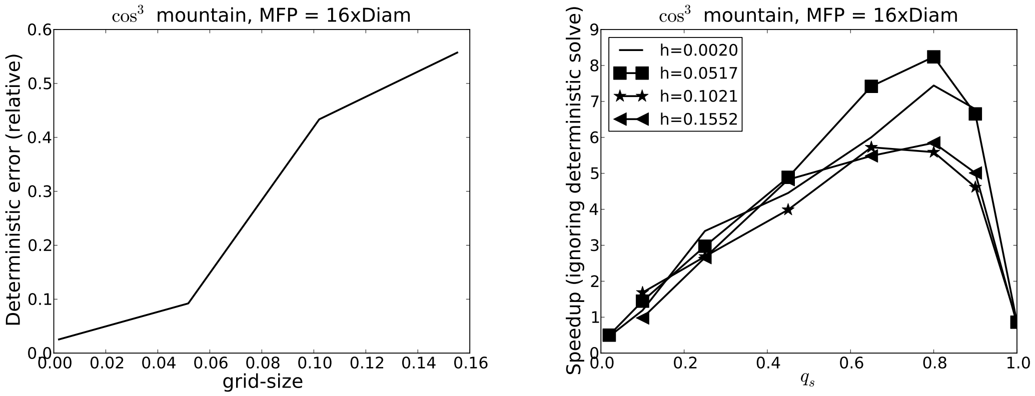

Figure 3: $|\mathrm{P}[D] - \langle I^s, S\rangle|/\mathrm{P}[D]$ is generally lower for smaller h. However, speedup is still very good even for large h. Diam is the maximal diameter of the simulation domain R

Our implementation swept both q_s and q_v. As expected, we see decreasing speedup with increasing atmospheric scattering strength σ. See Fig. 4.

It is important to note that use of adjoint-enhanced surface scattering, and heuristic atmospheric scattering ($q_s < 1$, $q_v < 1$) together is especially helpful. In fact, even with a small MFP = 1.3·Diam (Diam is the maximal diameter of the simulation domain R), we realize good speedup when $q_s = 0.9$, $q_v < 1$. Note that if either $q_s = 1$ or $q_v = 1$ (so no use of either SAI or heuristic scattering adjustment), speedup almost disappears. This is slightly counter-intuitive but may be explained as follows: Each method (SAI or heuristic) significantly increases the number of paths in two significant classes (surface-only and atmosphere-to-detector). Therefore, variance from these path-classes is all

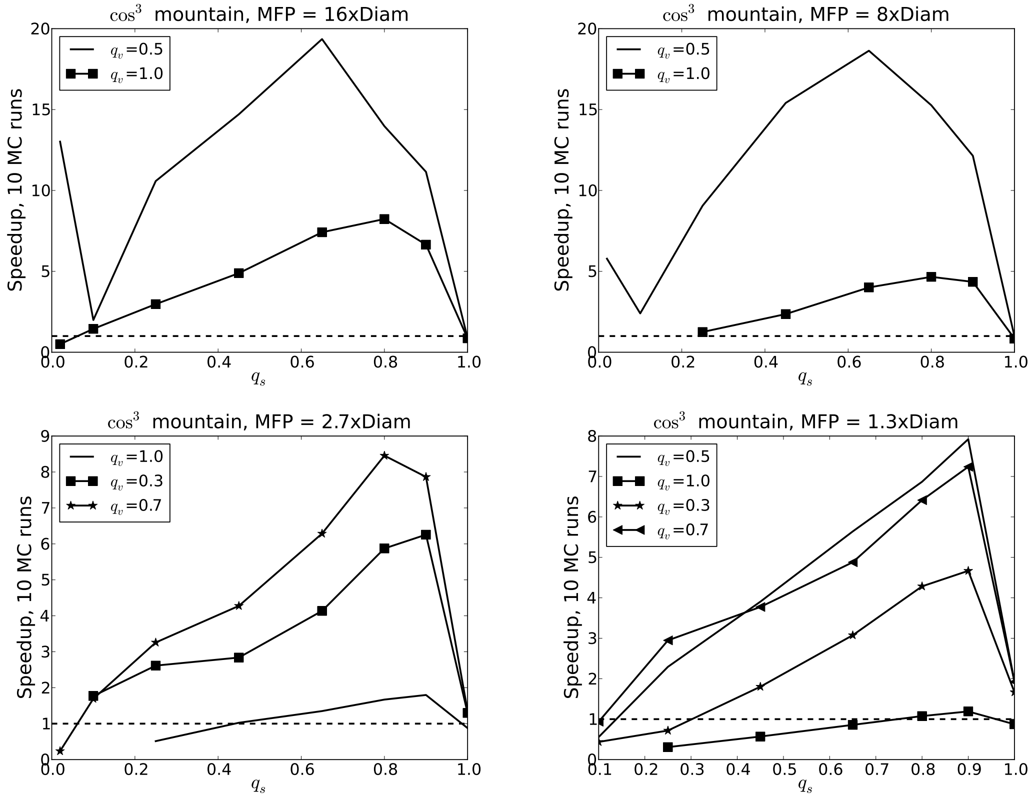

Figure 4: Speedup when using both surface adjoint approximation I^s (with parameter q_s) and heuristic atmospheric scattering (with parameter q_v)

but eliminated. Supposing each of these path-classes accounts for 2/5 of the total paths reaching the detector, by themselves they can only reduce variance by a factor of $1/(1-2/5) = 5/3$. However, together they can reduce variance by a factor of $1/(1-4/5) = 5$.

As one can see, selection of the parameters q_s and q_v makes a significant difference in the resultant variance. We provide some heuristics here and refer the reader to [26] for more details. When $q_s \to 0$ most of the photons will travel on the surface only. The photons that take a route prescribed by the heuristic chain must then carry an additional weight $= 1/q_s$ to compensate for this. For this reason, picking q_s too small results in increased variance. A similar argument holds for q_v. That the optimal q_s is so close to 1 (and greater than the optimal q_v) can also be explained by the fact that paths interacting exclusively with the surface are less likely to occur (in the analog world) than those interacting with the surface and atmosphere.

4 Conclusion and Outlook

A novel method for Monte Carlo transport was presented that uses an approximation of the adjoint (ignoring atmospheric effects) to reduce variance in simulations, equivalently, accelerate convergence to a specified accuracy. This algorithm, the Surface Adjoint Importance (SAI) method, may be combined with any unbiased method to significantly reduce variance coming from surface interactions when the overlaying atmosphere is optically thin. If it is combined with a method that reduces variance coming from atmospheric interactions, significant overall variance reduction is achieved. The implementation is relatively simple, requiring only an approximate adjoint transport solver for the boundary which adds virtually no overhead to the Monte Carlo computation time.

A possible application of this kind of accelerated Monte Carlo modeling in remote sensing is to address "adjacency" effects caused by highly variable terrain, including built environments (urban canyons). The standard adjacency effect is observed when an aerosol layer of moderate optical thickness mixes in an imaging detector's pixel light that has been reflected off surface elements with contrasting albedos in neighboring pixels. This is now a solved problem in the case of a variable-but-flat surface under a uniform atmosphere [28]. However, adjacency effects caused by non-flat terrain are only beginning to be explored, particularly in the thermal IR (where $Q(r, v)$ is determined by temperatures and emissivities).

On a broader scale, our work is an illustration of a modular approach to variance reduction whereby different interactions are handled separately and then pieced together in an unbiased manner. Specifically, these different interactions could be pieced together as in Algorithm 5.

For instance one can envision a "cloud" module where radiation transport inside the cloud (dominated by multiple scattering) is treated off-line in some judicious approximation, and then incorporated into complex scene simulation. In applications driven by surface property retrievals from remote sensing data, efficient modularized Monte Carlo modeling would open the door to advanced atmospheric compensation schemes with broken-cloud capability. This is another wide open frontier recently explored in [29].

Acknowledgments

This work was supported in part by DOE/NNSA Grant No. DE-FG52-08NA28779 and NSF Grants Nos. DMS-0804696 and PHY05-51164, as well as NSF Research Training Grant No. DMS-060DMS-0602235. AD wishes to thank the Kavli Institute for Theoretical Physics at UC Santa Barbara for hospitality and stimulation while finishing this manuscript. Finally, we are grateful to the two anonymous referees whose comments were insightful, and their suggestions lead to a much improved manuscript.

A Appendix: Numerical solution to the adjoint problem

Here, we seek to approximate I^s with I^h, at discretization level h.

Being in $d = 2$ space, this problem in irradiation and radiative exchange between surface elements is not amenable to commonly available radiosity codes, which are naturally designed for productivity in the usual $d = 3$ space. Radiosity is indeed used extensively for heat transport computations in engineering applications [30, among others], as well as in photorealistic computer graphics [31, among others]. At the same time, it is quite easy to encode a 2D radiosity code, as shown below, and there is no need for optimization for our present purposes (demonstrate the SAI speedup). However, beyond this demonstration, there will undoubtedly be applications in 3D radiative transfer per se, and then existing high-performance radiosity codes will become assets.

To simplify computation of our numerical solution we make the assumption

$$P(r, v \to v') = \mathbf{1}_{\nu_r \cdot v > 0}(r, v)\kappa(r, v'),$$

and recall that $g(r, v) = g_0(r) = \mathit{constant}$ whenever $\nu_r \cdot v > 0$ so that $g(r, v) = g_0(r)$. The result is that I^s is then a function of position only. This significantly improves the speed of solving the adjoint problem, as well as the memory requirements for using it. Theoretical results in this paper do not need this assumption, which we make here as a matter of convenience.

We will now discretize the coefficients and approximate the integral operator appearing on the right hand side of (5), denoted now by T. For $r_1 \in \partial R$,

$$TI^s(r_1, v_1) = \alpha(r_1) \int_{\nu_{r_1} \cdot v_2 < 0} K(r_1, v_2) I^s(r_+(r_1, v_2), v_2) \, \mathrm{d}v_2.$$

Notice that Tf is function depending only on r, and in fact only on the boundary values of f. Since g depends only on r, $I^s = \sum_{k=0}^{\infty} T^k g$ will depend only on r and whether or not $\nu_r \cdot v > 0$. We thus define

$$\varphi(r) : = I^s(r, v), \qquad r \in \partial R, \; \nu_r \cdot v > 0.$$

We find that $\varphi : \partial R \to \mathbb{R}$ satisfies the equation

$$\varphi = \mathcal{A}\varphi + g_0, \qquad \mathcal{A}f(r_1) := \alpha(r_1) \int_{\nu_{r_1} \cdot v_2 < 0} K(r_1, v_2) f(r_+(r_1, v_2)) \, \mathrm{d}v_2.$$

In discretizing this operator, and integrals over directions in general, we use the change of variables,

$$\begin{gathered} \int_{\nu_r \cdot v < 0} f(r_+(r, v), v) \, \mathrm{d}v = \int_{\partial R} f(r', v) \partial_\nu N(r, r') \, \mathrm{d}\mu(r'), \\ \partial_\nu N(r, r') := \frac{\nu_r \cdot (r' - r)}{|r' - r|^d}. \end{gathered} \tag{A.1}$$

The term $\partial_\nu N$ is normal derivative (at r) of the free-space Green's function for the Laplacian. One can show (see, e.g., the section on double-layer potentials in [32]) that for $r, r' \in \partial R$, $\nu_r \cdot (r' - r) \lesssim |r' - r|^2$. Therefore it is in fact an integrable function. When $d = 2$ it is moreover bounded.

We now discretize the operator $\mathcal{A}$. First split the boundary into non-overlapping segments $\{\partial R_j\}_{j=0}^{N_p - 1}$ with ∂R_j centered at r_j, with length $|\partial R_j| \le h$. Denote by Rf the (orthogonal) projection of f onto the space of piecewise constant functions (constant on each segment ∂R_j). We also think of Rf as a vector in $\mathbb{R}^{N_p}$ and Rf_j its components.

Then, after the change of variables (A.1) we have (at gridpoint r_i)

$$\begin{aligned}\mathcal{A}f(r_i) &= \alpha(r_i)\int_{\partial R} K(r_i, \widehat{r-r_i})\partial_\nu N(r_i, r) f(r)\,\mathrm{d}\mu(r) \\ &\approx \alpha(r_i) \sum_{\substack{0\le j\le N_p-1 \\ j\ne i}} |\partial R_j| K(r_i, \widehat{r_j - r_i})\partial_\nu N(r_i, r_j) f(r_j) \\ &:= \sum_j A^h_{ij} R f_i.\end{aligned} \tag{A.2}$$

This implicitly defines the matrix A^h.

We now define our discrete approximation to φ as the piecewise constant function (vector) φ^h solving

$$\varphi^h = A^h \varphi^h + Rg. \tag{A.3}$$

We then define approximations $I^h \approx I^s$,

$$I^h(r, v) := \varphi^h(r), \quad r \in \partial R,\ \nu_r \cdot v > 0. \tag{A.4}$$

Note that, in our implementation, we have chosen to represent angular integrals as integrals over the boundary. This works for two reasons. First, as our adjoint solution depends only on position it is convenient to evaluate these sums. Second, if instead a discretization were chosen that was uniform in angle, then (with only finitely many angles) one would often miss the (small) detector in evaluation of the integral.

References

[1] B. Davison and J. B. Sykes. *Neutron Transport Theory.* Oxford University Press, Oxford, 1957.

[2] J. Spanier and E. M. Gelbard. *Monte Carlo Principles and Neutron Transport Problems.* Addison-Wesley, Reading, Mass., 1969.

[3] I. Lux and L. Koblinger. *Monte Carlo Particle Transport Methods: Neutron and Photon Calculations.* CRC Press, Boca Raton, 1991.

[4] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–R93, 1999.

[5] G. Bal. Inverse transport theory and applications. *Inverse Problems*, 25:053001, 2009.

[6] S. Chandrasekhar. *Radiative Transfer.* Dover Publications, New York, 1960.

[7] K. N. Liou. *An Introduction to Atmospheric Radiation.* Academic Press, San Diego, 2002.

[8] A. Marshak and A. B. Davis (Eds.). *3D Radiative Transfer in Cloudy Atmospheres.* Springer, Heidelberg, 2005.

[9] C. D. Mobley et al. Comparison of numerical models for computing underwater light fields. *Appl. Opt.*, 32:7484–7505, 1993.

[10] G. E. Thomas and K. Stamnes. *Radiative Transfer in the Atmosphere and Ocean* Cambridge University Press, Cambridge, 2002.

[11] H. Sato and M. C. Fehler. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth.* AIP Series in Modern Acoustics and Signal Processing. AIP Press, New York, 1998.

[12] J. R. Schott. *Remote Sensing – The Image Chain Approach.* 2nd Edition, Oxford University Press, Oxford, 2007.

[13] C. J. Solomon, A. Sood, T. E. Booth, and J.'K. Shultis. An S_n Approach to predicting Monte Carlo cost with weight-dependent variance reduction. *Trans. Am. Nucl. Soc.*, 103:348–350, 2010.

[14] A. Haghighat and J. C. Wagner. Monte Carlo variance reduction with deterministic importance functions. *Prog. in Nuclear Energy*, 42:25–53, 2003.

[15] J. E. Hoogenboom. Zero-variance Monte Carlo schemes revisited. *Nucl. Sci. & Engr.*, 160:1–22, 2008.

[16] K. F. Evans and A. Marshak. Numerical methods. In *3D Radiative Transfer in Cloudy Atmospheres*, edited by A. Marshak and A. B. Davis, pp. 243-281, Springer, Heidelberg, 2005.

[17] R. Buras and B. Mayer. Efficient unbiased variance reduction techniques for Monte Carlo simulations of radiative transfer in cloudy atmospheres: The solution. *J. Quant. Spectros. & Rad. Trans.*, 112:434-447, 2011.

[18] E. Veach. *Robust Monte Carlo Methods for Light Transport Calculations.* PhD Dissertation, Stanford University, 1997.

[19] M. H. Kalos. Importance sampling in Monte Carlo shielding calculations. *Nucl. Sci. & Eng.*, 16:227–, 1963.

[20] S. A. Turner and E. W. Larsen. Automatic variance reduction for three-dimensional Monte Carlo simulations by the local importance function transform–I: Analysis. *Nucl. Sci. & Eng.*, 127:22–35, 1997.

[21] S. A. Turner and E. W. Larsen. Automatic variance reduction for three-dimensional Monte Carlo simulations by the local importance function transform–II: Numerical results. *Nucl. Sci. & Eng.*, 127:36–53, 1997.

[22] K. A. Van Riper et al. AVATAR – Automatic variance reduction in Monte Carlo calculations. In *Proceedings of the Joint International Conference on Mathematical Methods and Supercomputing in Nuclear Applications*, Saratoga Springs (NY), 6-10 Oct. 1997, Amer. Nucl. Soc., LaGrange, Il, 1997.

[23] J. D. Densmore and E. W. Larsen. Variational variance reduction for particle transport eigenvalue calculations using Monte Carlo adjoint simulation. *J. of Comp. Physics*, 192:387–405, 2003.

[24] M. Ambrose, R. Kong, and J. Spanier. Efficient, automated Monte Carlo methods for radiation transport. *J. of Comp. Physics*, 227:9643–9476, 2008.

[25] R. Kong and J. Spanier. A new proof of geometric convergence for general transport problems based on sequential correlated sampling methods. *J. of Comp. Physics*, 227:9762–9777, 2008.

[26] G. Bal and I. Langmore. Importance sampling and adjoint hybrid methods in Monte Carlo transport with reflecting boundaries. *ArXiv*, arXiv:1104.2550v1, 2011.

[27] G. I. Marchuk et al. *The Monte Carlo Methods in Atmospheric Optics.* Springer-Verlag, New York, 1980.

[28] A. Lyapustin. Radiative transfer code SHARM-3D for radiance simulations over a non-Lambertian nonhomogeneous surface: Intercomparison study. *Appl. Opt.*, 41:5607–5615, 2002.

[29] B. D. Bartlett and J. R. Schott. Atmospheric compensation in the presence of clouds: An adaptive empirical line method (AELM) approach. *J. Appl. Remote Sens.*, 3:1–16, 2009.

[30] R. Siegel and J. R. Howell. *Thermal Radiation Heat Transfer.* 2nd edition, McGraw-Hill, New York, 1981.

[31] P. Brunet and F. W. Jansen (Eds.). *Photorealistic Rendering in Computer Graphics: Proceedings of the 2nd Eurographics Workshop on Rendering (Focus on Computer Graphics).* Springer, New York, 1994.

[32] G. Folland. *Introduction to Partial Differential Equations.* Princeton University Press, Princeton, NJ, 1995.