A Hybrid Approach to Computational Radiative Transfer in 3D Clouds:

Getting the Best of Monte Carlo and Deterministic Methods

<u>Anthony B. Davis</u>
Jet Propulsion Laboratory
California Institute of Technology
... *and Visiting Scientist at:* Los Alamos National Laboratory
Space & Remote Sensing Group

Guillaume Bal
Columbia University
Department of Applied Physics & Applied Mathematics

JPL COLUMBIA UNIVERSITY

Research Team

Anthony Davis, Laboratory co-PI

- 5 years at NASA – Goddard Space Flight Center, 11 years at LANL, now at JPL/CalTech

Guillaume Bal, University PI

- 2 years Stanford, 2 years U of Chicago, at Columbia for 7 years
- Expertise: wave propagation in random media and time reversal (ONR); imaging in random media using kinetic models (DARPA); inverse transport theory with applications in geophysical and medical imaging (NSF).

Ian Langmore, Columbia University Postdoctoral Fellow

Sponsor: DOE/NNSA/NA-22/SAM

Hybrid Radiation Transport Methods, 1

Present transport codes, for both EM and particulate radiation, are either *all* Monte Carlo or *all* deterministic.

The **deterministic** approach is relatively fast, but in some situations does not adequately reproduce the physics, whereas **Monte Carlo** can more accurately model the physics, but is much slower to converge.

A hybrid RT model would merge the best of Monte Carlo and deterministic methods in a theoretically coherent way. It should be adaptable to problems of very different types, for various sensor types.

Hybrid Radiation Transport Methods, 2

One could imagine use of a deterministic approach wherever it provided adequate results, but would switch over to Monte Carlo when necessary for some components of the computation.

However, in a better approach, the *best of both techniques* would be wholly integrated in a rigorous mathematical framework so that all problems are done more efficiently and accurately.

Hopefully, the deterministic and Monte Carlo approaches are not incompatible and **a conceptually coherent hybrid model can be constructed**.

The resulting numerical models should be able to handle three-dimensional problems.

3D Radiative Transfer Equation, 1

The monochromatic 3D RT equation takes the integro-differential form

$$\mathbf{\Omega} \cdot \nabla I + \sigma(\mathbf{x}) I = \sigma_{\mathsf{S}}(\mathbf{x}) \int_{4\pi} p(\mathbf{x}, \mathbf{\Omega}' \cdot \mathbf{\Omega}) I(\mathbf{x}, \mathbf{\Omega}') \mathrm{d}\mathbf{\Omega}' + S(\mathbf{x}, \mathbf{\Omega}), \quad \text{where}$$

- $I(\mathbf{x}, \mathbf{\Omega})$ the unknown radiance (a.k.a. specific intensity) measured by a sensor at position $\mathbf{x}$ looking into direction $-\mathbf{\Omega}$,
- $\sigma(\mathbf{x})$ is the extinction coefficient,
- $\sigma_{\mathsf{S}}(\mathbf{x})$ is the scattering coefficient,
- $p(\mathbf{x}, \mathbf{\Omega}' \cdot \mathbf{\Omega})$ is scattering phase function describing how radiation is redistributed directionally, and
- $q(\mathbf{x}, \mathbf{\Omega})$ is a source term.

Boundary conditions can express the presence of primary sources or not ("absorbing" boundaries), or partial reflection with or without bi-directional anisotropy.

3D Radiative Transfer Equation, 2

Letting $z = (\mathbf{x}, \mathbf{\Omega})$, we recast the above 3D RT equation *and boundary conditions* in integral form

$$I(z) = KI(z) + Q(z) \quad \text{where} \quad KI(z) := \int k(z', z) I(z') \mathrm{d}z'$$

is the linear transport kernel that can be made explicit in terms of the above coefficients and phase function; $Q(z)$ is the associated source term (uncollided radiance) dependent on $q(z)$ and boundary conditions.

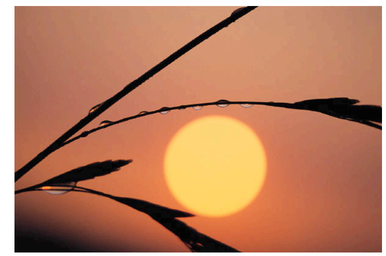
Solution of this RT problem can be expressed as a Neumann series

$$I(z) = (1 - K)^{-1} Q := \sum_{n=0}^{\infty} K^n Q(z),$$

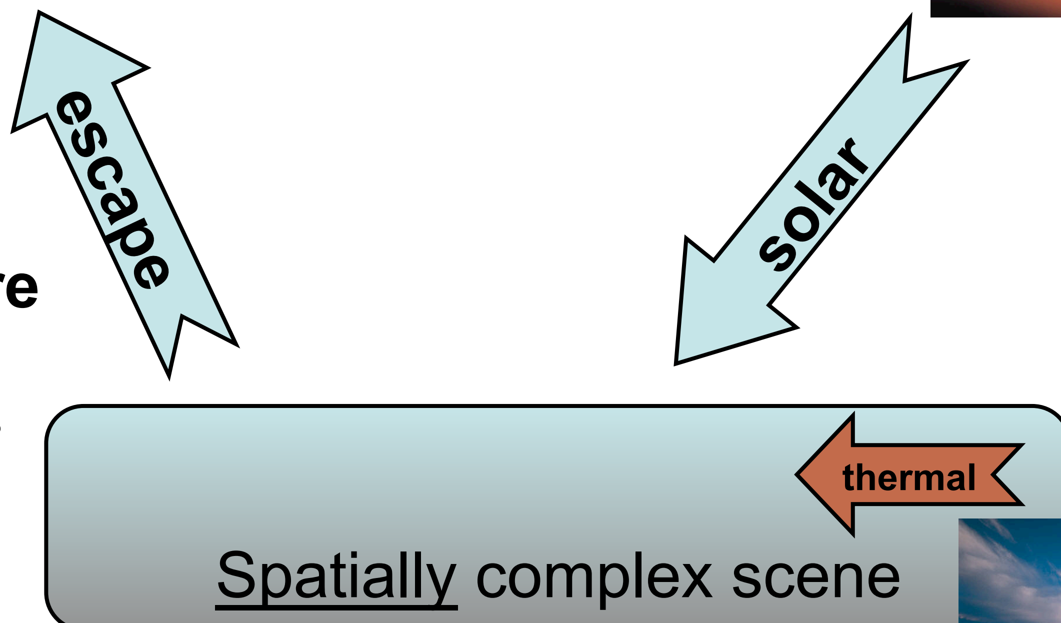
which expresses physically an expansion by orders of scattering. Deterministic computational methods are all based on discretizations of the transport equation in $\mathbf{x}$-space (1D, 2D and 3D meshes) and in $\mathbf{\Omega}$-space (angular quadratures and/or spherical harmonic expansions).

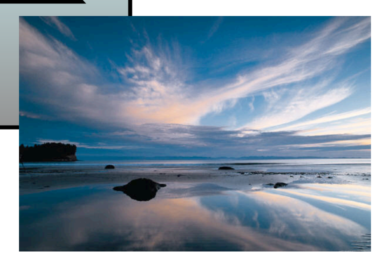
Hybrid MC/deterministic *forward* 3D RT

Physics/chemistry-based analyses



$I(\boldsymbol{x}_{\mathrm{obs}},\Omega_{\mathrm{obs}})$

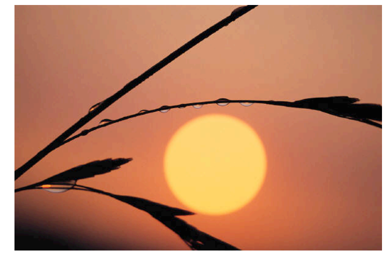


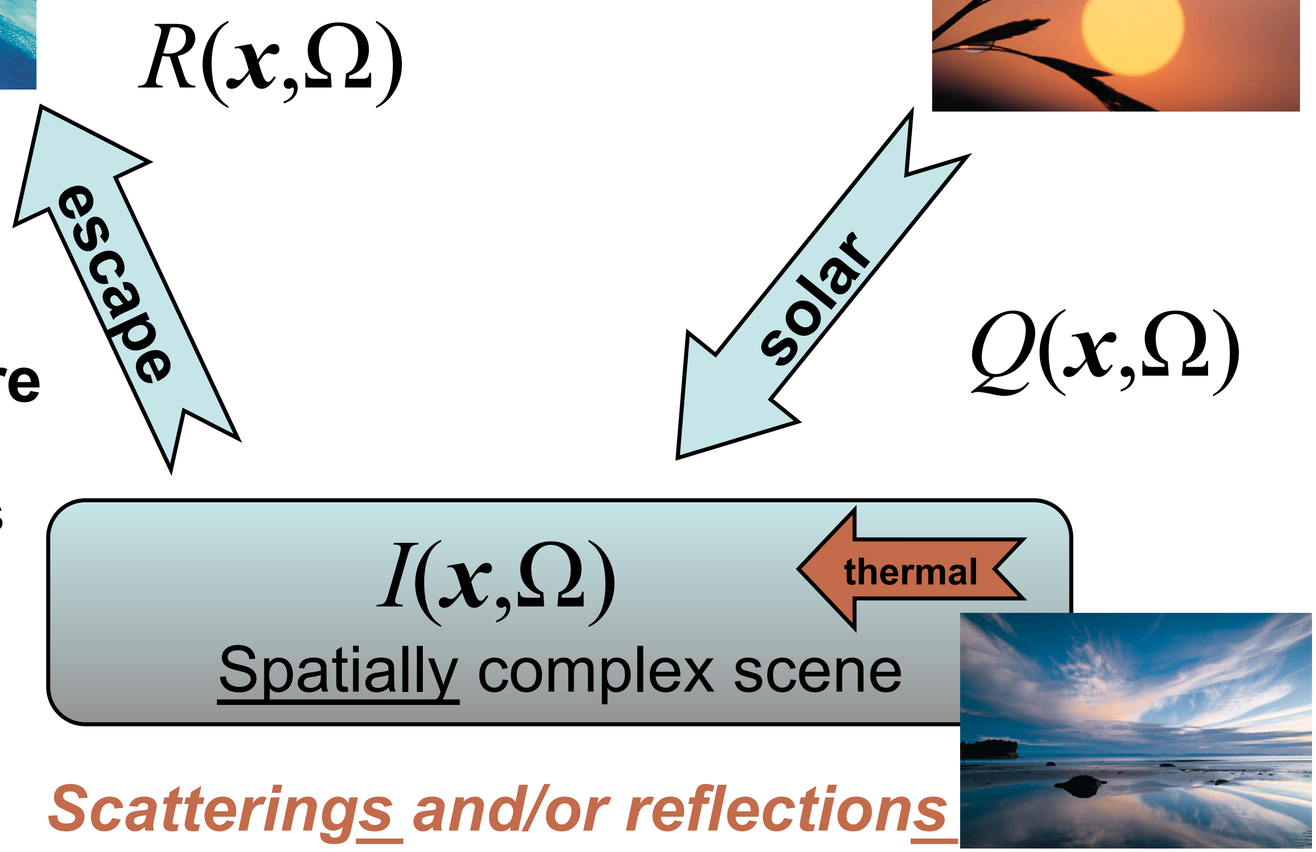


- **3D atmosphere**
  - aerosols
  - clouds/plumes
  - etc.
- **3D surface (roughness)**
  - microscale
  - macroscale

Hybrid MC/deterministic *forward* 3D RT

Physics/chemistry-based analyses





- **3D atmosphere**
  - aerosols
  - clouds/plumes
  - etc.
- **3D surface (roughness)**
  - microscale
  - macroscale

Monte Carlo Methods: Formal Description, 1

Let $R(z)$ be the response function of a detector. We wish to compute

$$\langle R, I\rangle = \int R(z)I(z)\mathrm{d}z = \sum_{n=0}^{\infty} \underbrace{\int \cdots \int}_{(n+1)\times} Q(z_0) \prod_{j=0}^{n-1} k(z_j, z_{j+1})\, R(z_n)\, \mathrm{d}z_0 \cdots \mathrm{d}z_n.$$

This may be estimated with the following random walks $\alpha = \{z_0, \ldots, z_n\}$:

(1) the walk starts at z_0 drawn from $Q^*(z)$ with $\int Q^*(z)\mathrm{d}z = 1$;

(2) the chain concludes at z_j (if so, then $n = j \geq 0$) with probability $p^*(z_j) = 1 - \int k^*(z_j, z')\mathrm{d}z'$ and contributes its current weight$\times R^*(z_j)$ to the tally; if not, then ...

(3) jumps from z_j to z_{j+1} with transition probability $k^*(z_j, z_{j+1})$, and repeats step **(2)** until termination.

Monte Carlo Methods: Formal Description, 2

Define then
$$\xi(\alpha) = \frac{Q(z_0)}{Q^*(z_0)} \left[\prod_{j=0}^{n-1} \frac{k(z_j, z_{j+1})}{k^*(z_j, z_{j+1})}\right] \frac{R(z_n)}{p^*(z_n)},$$
where the various ratios can be interpreted as weights to be assigned to the sequence of steps in the random walk.

We verify that, for random independent identically distributed (i.i.d.) paths α_m drawn according to the above law (rules), we have
$$\langle R, I\rangle = \mathbb{E}\{\xi(\alpha)\} = \sum_{n=0}^{\infty} \int^{n+1} \xi(\alpha) \times Q^*(z_0) \left[\prod_{j=0}^{n-1} k^*(z_j, z_{j+1}) \mathrm{d}z_j\right] p^*(z_n) \mathrm{d}z_n,$$
where the specific choices for $Q^*(z)$, $k^*(z, z')$ and $p^*(z)$ define the specific type of Monte Carlo algorithm. In practice, we compute
$$\mathbb{E}\{\xi(\alpha)\} = \lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} \xi(\alpha_m).$$

Classic Monte Carlo

The classic Monte Carlo method is based on the natural choices

$$Q^*(z) = \frac{Q(z)}{\langle Q, 1 \rangle}, \quad k^*(z, z') = k(z, z'), \quad p^*(z) = 1 - \int k(z, z') \mathrm{d}z'.$$

It can furthermore be shown that

$$p^*(z) \approx \frac{\sigma_{\mathsf{a}}(\mathbf{x})}{\sigma(\mathbf{x})},$$

where $\sigma_{\mathsf{a}}(\mathbf{x}) := \sigma(\mathbf{x}) - \sigma_{\mathsf{s}}(\mathbf{x})$, noting that this approximation breaks down near the boundaries of the medium.

The chain is then easy to simulate, i.e., $k(z, z')$ is easy to sample, and

$$\xi(\alpha) = \langle Q, 1 \rangle \frac{R(z_n)}{p^*(z_n)} \approx \langle Q, 1 \rangle R(\mathbf{x}_n, \boldsymbol{\Omega}_n) \frac{\sigma(\mathbf{x}_n)}{\sigma_{\mathsf{a}}(\mathbf{x}_n)}.$$

The variance may be quite large when $R(z_n)$ varies rapidly (localized detectors) and/or when $\sigma_{\mathsf{a}}(\mathbf{x}_n)$ is small (as in the diffusion regime).

Monte Carlo Variance Estimation

The variance for the Monte Carlo scheme is estimated as

$$\begin{aligned}
\mathbb{E}\{(\xi - \mathbb{E}\{\xi\})^2\} &= \sum_{n=0}^{\infty} \int^{n+1} (\xi(\alpha) - \mathbb{E}\{\xi(\alpha)\})^2 \times \\
&\qquad Q^*(z_0) \left[\prod_{j=0}^{n-1} k^*(z_j, z_{j+1}) \mathrm{d}z_j \right] p^*(z_n) \mathrm{d}z_n = \\
\mathbb{E}\{\xi^2\} - \mathbb{E}\{\xi\}^2 &= \lim_{N \to \infty} \frac{N}{N-1} \left[\frac{1}{N} \sum_{m=1}^{N} \xi(\alpha_m)^2 - \left(\frac{1}{N} \sum_{m=1}^{N} \xi(\alpha_m) \right)^2 \right].
\end{aligned}$$

From there, the central limit theorem gives us

$$\left| \langle R, I \rangle - \frac{1}{N} \sum_{m=1}^{N} \xi(\alpha_m) \right| \lesssim \sqrt{\frac{\mathbb{E}\{\xi^2\} - \mathbb{E}\{\xi\}^2}{N}}.$$

Monte Carlo Variance Reduction

There are only two ways of reducing the RMS error (accelerating the convergence):

- either change the scaling in $N^{-1/2}$ to something faster;
- or minimize the prefactor $(\mathbb{E}\{\xi^2\} - \mathbb{E}\{\xi\}^2)^{1/2}$.

The former route is the goal of *quasi*-Monte Carlo methods, i.e., use only i.d. random numbers (with N^{-1} being the theoretical limit). Our goal here is to follow the second route, as far as possible.

Variance Annihilation

Let us now assume that we can solve the **adjoint RT equation** in integral form

$$I^\dagger(z) = \int k^\dagger(z', z) I^\dagger(z') \mathrm{d}z' + R(z),$$

where $k^\dagger(z', z) = k(z, z')$ is the adjoint kernel to $k(z', z)$. Note that the adjoint source term is the response function of the sensor.
Then choose

$$Q^*(z) = \frac{I^\dagger(z)Q(z)}{\langle I^\dagger, Q\rangle}, \quad k^*(z, z') = \frac{k(z, z')I^\dagger(z')}{I^\dagger(z)}, \quad p^*(z) = \frac{R(z)}{I^\dagger(z)},$$

which generates admissible chains such that $\underline{\xi(\alpha) = \langle I^\dagger, Q\rangle = \langle R, I\rangle}$, i.e., the chain always gives the right solution and its variance vanishes!

However, solving *deterministically* for $I^\dagger$ is at least as hard as solving for I itself. Indeed ...

Adjoint 3D Radiative Transfer Equation

Based on the *defining* requirement that $\langle I^\dagger, Q\rangle = \langle R, I\rangle$ for the large class of valid R and Q functions, it can be shown that the integro-differential form of the adjoint 3D RT equation is

$$-\mathbf{\Omega}\cdot\nabla I^\dagger + \sigma(\mathbf{x})I^\dagger = \sigma_\mathrm{s}(\mathbf{x})\int_{4\pi} p(\mathbf{x},\mathbf{\Omega}\cdot\mathbf{\Omega}')I^\dagger(\mathbf{x},\mathbf{\Omega}')\mathrm{d}\mathbf{\Omega}' + r(\mathbf{x},\mathbf{\Omega}),$$

where we note the reversal of the sign of $\mathbf{\Omega}(\mu,\phi)$ and of the roles of $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ in $P(\cdots)$, and where $r(\mathbf{x},\mathbf{\Omega})$ describes *locally* the sensor response. For instance, $r(\mathbf{x},\mathbf{\Omega}) = \delta(\mathbf{x}-\mathbf{x}_\mathrm{obs})|\mu|\Theta(\pm\mu)$ gives the hemispherical fluxes in the $\pm\hat{\mathbf{z}}$ directions at point $\mathbf{x}_\mathrm{obs}$.

Adjoint boundary conditions are for outgoing radiance. They express the *absence* of adjoint sources if $r(\mathbf{x},\mathbf{\Omega}) \not\equiv 0$, or the *presence* of sensors at the boundaries when $r(\mathbf{x},\mathbf{\Omega}) \equiv 0$. In short, no adjoint radiance (a.k.a. the "importance" field for a given sensor) escapes the medium, unless a sensor is there. If anisotropic reflection, $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ are reversed.

Extreme Variance Reduction?

Assume we have a reasonably good deterministic approximation $I_{\mathsf{D}}^{\dagger}(z)$ of $I^{\dagger}(z)$, such that $I_{\mathsf{D}}^{\dagger}(z) \approx \int k^{\dagger}(z',z) I_{\mathsf{D}}^{\dagger}(z')\mathrm{d}z' + R(z)$. We then choose

$$Q^*(z) = \frac{I_{\mathsf{D}}^{\dagger}(z)Q(z)}{\langle I_{\mathsf{D}}^{\dagger}, Q\rangle}, \; k^*(z,z') = \frac{k(z,z')I_{\mathsf{D}}^{\dagger}(z')}{I_{\mathsf{D}}^{\dagger}(z)}, \; p^*(z) = 1 - \int k^*(z,z')\mathrm{d}z',$$

leading to

$$\xi_{\mathsf{D}}(\alpha) = \frac{R(z_n)}{p^*(z_n)} \times \frac{\langle I_{\mathsf{D}}^{\dagger}, Q\rangle}{I_{\mathsf{D}}^{\dagger}(z_n)}.$$

We can show that variance of $\xi_{\mathsf{D}}(\alpha)$ for this chain is proportional to

$$\langle I_{\mathsf{D}}^{\dagger}, Q\rangle - \langle I^{\dagger}, Q\rangle, \qquad \text{and to} \qquad 1 - R(z_n)/I_{\mathsf{D}}^{\dagger}(z_n)p^*(z_n),$$

both of which vanish when $I_{\mathsf{D}}^{\dagger} = I^{\dagger}$. So, any reasonable approximation $I_{\mathsf{D}}^{\dagger}$ of $I^{\dagger}$ then significantly reduces the variance of $\xi(\alpha)$.

Sampling Issues

The approximation $I^{\dagger}_{\text{D}}$ may be obtained by various means such as, e.g., using a diffusion approximation or a coarse *deterministic* transport simulation. This is how we plan to couple deterministic with Monte Carlo models.

The main difficulty is that the sampling of the modified distribution is far more cumbersome than the original distribution based on $k(z', z)$.

Theoretical results show that in the diffusive regime, always problematic in Monte Carlo, with a non-dimensional mean-free-path of size ε, the variance of the hybridized Monte Carlo method is of order ε^2/N as $\varepsilon \to 0$, for N realizations of the modified chain.

Path Forward, Columbia

We are developing fast sampling methodologies for the modified distribution based on 3D diffusion approximations or coarse deterministic methods.

Currently, a slow MCMC (Monte Carlo Markov Chain) methodology and a rejection methodology have been implemented for simple one dimensional and three dimensional geometries (no boundaries and isotropic scattering).

This effort is being pursued by Guillaume Bal, Ian Langmore and Stan Snelson (an undergraduate who has been working on the problem for about a year). Results are in progress.

Path Forward, JPL

The co-PI will finish verification of variance annihilation theory in "literal" 1D RT, and verify ε^2 scaling for adjoint diffusion approximation in plane-parallel media. These are case where closed-form expressions exist.

He will also explore the idea of accelerating backward Monte Carlo, where it is suspected that the normal determinsitic 3D solution (e.g., from SHDOM) plays the role of the adjoint in forward Monte Carlo.

Path Forward, Team

Develop test cases of interest to the sponsor *and* to the environmental science community, e.g., involving clouds and (or) aerosol (plumes).

We anticipate that it will be necessary to design the hybrid method for the type of scene geometry. Universal hybrid approach seems elusive.


Questions?