Parametric 3D Scene Reconstruction from Imaging Radiometry, Part I: The Path-Recycling Forward Monte Carlo Model

1

Parametric 3D Scene Reconstruction from Imaging Radiometry, Part I: The Path-Recycling Forward Monte Carlo Model

Ian Langmore, Anthony B. Davis, *Member, Opt. Soc. Am., Am. Geophys. Un., Am. Meteor. Soc.*,
and Guillaume Bal, *Member, Am. Math. Soc., Soc. Indust. Appl. Math.*

Abstract

The vast majority of physics-based retrieval algorithms used in remote sensing of atmosphere and/or surface properties are multi- or hyper-spectral, some use multi-angle information; recently, polarization diversity has been added to the available input from sensors and accordingly modeled with vector radiative transfer codes. At any rate, a single pixel is processed at a time using a forward radiative transfer model predicated on 1D transport theory. Neighboring pixels are sometimes considered, but almost always to only formulate a statistical constraint in the inversion based on spatial context. We demonstrate here the potential power that could be harnessed by adding bona fide multi-pixel techniques to the mix. To this effect, we use forward and inverse radiative transfer modeling in 2D (sufficient for this demo) of a single-wavelength imaging sensor's response used to infer the position, size and opacity of an absorbing atmospheric plume somewhere in a deep valley in the presence of a partly-known/partly-unknown aerosol assumed to have an exponential profile with altitude. In the present article, we describe the necessary innovation in forward multi-dimensional radiative transfer. In spite of its notorious reputation for inefficiency, we use a Monte Carlo technique. However, the adopted scheme is highly accelerated without loss of accuracy by using efficiently "recycled" Monte Carlo paths, a methodology adapted from ongoing research in biomedical imaging. In Part II, this forward model is put to work in a Bayesian inversion targeting plume properties and the specific amount of background aerosol.

I. INTRODUCTION, CONTEXT, AND OVERVIEW

OPERATIONAL methods in the remote sensing of atmospheric and surface properties using physics-based techniques have, at a minimum, the same limitations as the adopted forward model for the signal. Such a model is required to translate measured radiances into inherent structural, optical, physical, and chemical properties. However,

I. Langmore and G. Bal are with Columbia University's Department of Applied Physics & Applied Mathematics, 500 W. 120th street, New York, NY, 10027, USA, e-mail: gb2030@columbia.edu.

A. B. Davis, corresponding author, is with the Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Mail Stop 233-200, Pasadena, CA, 91109, USA, e-mail: Anthony.B.Davis@jpl.nasa.gov.

its natural form is to predict radiances for given scene properties. The adopted inversion method that reverses this connection will introduce further limitations, as will instrument noise, and so on.

In retrievals using the UV-through-microwave spectrum, one endemic limitation follows from overly simplified physical and geometrical assumptions in the radiative transfer (RT) underlying the sensor signal prediction. This is most problematic in the VNIR region as it is dominated by atmospheric scattering and surface reflection. These radiation transport processes indeed make the prevailing RT highly nonlocal in the physically intuitive as well as mathematical senses of the word.

When it comes to operational remote sensing missions, data throughput volume considerations weigh heavily in favor of straightforward pixel-by-pixel processing, using the multi- or hyper-spectral dimension of the data as best possible. When available, multi-angle information and, more and more frequently, polarimetric diversity will also be brought to bear on the physics-based remote sensing problem. However, once isolated from any spatial context, the natural assumption for the pixel-scale RT model is horizontal uniformity, i.e., the optical medium is assumed to be an infinite plane-parallel slab with, at most, some vertical structure. This assumption leads immediately to the textbook case of one-dimensional (1D) RT. Very conveniently for computational considerations, the azimuthal dependence is then amenable to a Fourier series decomposition with each mode being independent of the others [1]. Nature, however, is 3D and complex ... as demonstrated by almost every remote sensing image captured since the dawn of the technology! This disconnect between the conceptual model for the remote sensing signal and the reality that produces it can be a significant liability for the inverse problem.

There are two basic kinds of 3D RT effect to worry about when using retrieval algorithms grounded in 1D RT [2]. First, there is the effect small-scale (unresolved) variability that dominates the 1D RT model error when the observation scale (pixel size) is large. Second, there is the "pixel-adjacency" effect that dominates the 1D RT model error when the observation scale is small. For the sake of argument, we can assume that the vertical structure of the medium, including the surface, is perfectly known under any given pixel. The adjacency effect results from the near proximity of other pixels where the surface and/or atmosphere have/has a different structure and thus excite net horizontal fluxes, precisely what is neglected in 1D RT. Much research into these topics has been performed primarily in 3D cloudy scenes [2], [3, and references therein]. Scenarios where the aerosol and molecular atmosphere is horizontally uniform while the surface, assumed flat, is not have also been investigated extensively [4]–[18, and references therein]. Topography effects have been studied as well [19, among others]. In all of these situations, a common question is: How far the 3D effects can be sensed? I.e., what is the scale that divides the above mentioned "small" and "large" pixel sizes? It is fair to say that, overall, much less has done about either large- or small-scale 3D RT effects in current remote sensing operations, let alone applications.

We focus here on the later effect: cross-pixel/adjacency transport. In fact, rather than see this as a nuisance from a 1D RT perspective, we take a resolutely 3D RT perspective and indeed exploit the effect in a remote sensing problem. This is a radical departure from mainstream physics-based remote sensing of the environment where retrieval methods are invariably based on multi- or hyper-spectral analyses of pixel-scale data, or maybe at a coarser resolution. At any rate, the inversion scheme uses a forward 1D RT model that optimally reproduces

the purely spectral observations, e.g., MODerate resolution Imaging Spectrometer (MODIS) [20]. In some cases, this leads to the retrieval of a vertical profile in a thermodynamic variable and/or a chemical composition, e.g., Atmospheric Infra-Red Sounder (AIRS) [21]. In other cases, multi-angle information is brought to bear, e.g., NASA's Multi-angle Imaging Spectro-Radiometer (MISR) [22]. Some of those exploit polarization measurements as well, e.g., the POLarization and Directionality of the Earth's Reflectances (POLDER) [23] sensor (sponsored by CNES).

To demonstrate more clearly how new information can be gained from the 3D RT point of view, we simulate here a retrieval of spatial structure using a purely monochromatic but multi-pixel technique. This of course calls for a 3D RT forward model and an inversion methodology capable of driving it. The 3D RT model, a Monte Carlo scheme, uses a specific parametric representation of the scene. Envision a deep valley or gorge with a known profile in altitude and reflectivity (see Fig. 1 below). This terrain is filled and overlaid with a stratified scattering atmosphere representative of a typical background aerosol that is imperfectly known. Somewhere in this canyon, there is a finite uniform gaseous plume that can be assumed purely absorbing at the selected wavelength (we presumably know its spectrum); also the plume geometry is assumed to be that of a cylinder lying horizontally. The challenge is to determine the position, diameter and opacity of the (admittedly idealized) plume, along with a refined estimate of the ambient aerosol load, using a single-channel imaging sensor that captures reflected and scattered sunlight. This is a scenario of significant interest in nuclear facility monitoring.

The paper is outlined as follows. In the following Section II, we describe in precise and general mathematical terms the forward RT problem that, for simplicity, is cast in two spatial dimensions, hence just one angular variable. In Section III, the specific 2D scene geometry is set, optical properties are prescribed, and remote sensing unknowns are defined. In Section IV, the numerical solution using the new "path-recycling" Monte Carlo scheme is described in simple algorithmic terms, as well as in probabilistic terms using measure-theoretical concepts; its performance is compared with the standard Monte Carlo method. We briefly conclude in Section V.

Appendix A covers certain technical details Monte Carlo on path recycling, starting from the measure-theoretical basis of the method. The path-recycling scheme follows three steps: (1) paths are generated in a reference atmosphere, and those that hit the detector are saved; (2) the (increase/decrease of) probability of these paths reaching the detector in the modified atmosphere is calculated, resulting in a set of weights attached to each path; (3) the weights are added resulting in an unbiased estimate of the detected flux in the modified atmosphere. Appendix B formalizes the notion of a discrete hierarchy of forward Monte Carlo models with varying precision, which is key to the Bayesian analysis presented in Part II. Therein means, variances and (pixel-to-pixel) covariances are computed from first principles.

The companion paper [24], referred to hereafter as "Part II," applies this advance in computational RT to the above-mentioned remote sensing challenge, from a Bayesian perspective on the inverse problem at hand.

II. 2D RADIATIVE TRANSFER EQUATION AND ASSOCIATED BOUNDARY CONDITIONS

In steady-state 3D RT, a monochromatic radiance field I depends on three spatial variables $\vec{r} = (x, y, z)^{\mathrm{T}}$, where z is customarily taken to be altitude above some reference point, and an angular variable $\vec{\Omega}$, with $\|\vec{\Omega}\| = 1$, that

is usually defined by two polar coordinates (θ, ϕ). In the present study, we can restrict our considerations to two-dimensional (2D) RT without losing the ability to provide a proof-of-concept in multi-pixel remote sensing that exploits cross-pixel radiation transport. Consequently, we will use just two spatial coordinates $\vec{r} = (x, z)^{\mathrm{T}}$ and a single angular variable $\vec{\Omega}(\theta) = (\sin\theta, \cos\theta)^{\mathrm{T}}$ where $\theta = 0$ is up (increasing z). We should bear in mind that the units for $I(\vec{r}, \vec{\Omega})$ in 2D are [photons/s/m/rad] (cf. [photons/s/m^2/sr] in the usual 3D setting).

That said, the general steady-state monochromatic RT equation for $I(\vec{r}, \vec{\Omega})$ looks the same in 2D as in 3D:

$$\left[\vec{\Omega} \cdot \nabla + \sigma(\vec{r}) \right] I = \sigma_{\mathrm{s}}(\vec{r}) \int p_{\mathrm{v}}(\vec{r}, \vec{\Omega}' \to \vec{\Omega}) I(\vec{r}, \vec{\Omega}') \mathrm{d}\vec{\Omega}' + Q_{\mathrm{v}}(\vec{r}, \vec{\Omega}), \tag{1}$$

for $\vec{r} \in \mathrm{R}$, an open connected subset of $\mathbb{R}^2$. We have introduced here the extinction, $\sigma(\vec{r})$, and scattering, $\sigma_{\mathrm{s}}(\vec{r})$, coefficients (still in m^{-1}), the phase function $p_{\mathrm{v}}(\vec{r}, \vec{\Omega} \to \vec{\Omega}')$, and the volume source term $Q_{\mathrm{v}}(\vec{r}, \vec{\Omega})$ (expressed here in [photons/s/m^2/rad]). The phase function (in [1/rad]) is normalized in such a way that $\int p_{\mathrm{v}}(\vec{r}, \vec{\Omega} \to \vec{\Omega}') \mathrm{d}\vec{\Omega}' = 1$, with $\mathrm{d}\vec{\Omega}' = \mathrm{d}\theta'$.

Boundary conditions (BCs) can similarly be expressed in very general terms. Letting $\partial\mathrm{R}$ denote the closed boundary of R and $\vec{n}_{\vec{r}}$ its outward normal at $\vec{r}$, with $\vec{r}$-dependent components $(\sin\theta_{\vec{n}}, \cos\theta_{\vec{n}})^{\mathrm{T}}$, we have

$$|\vec{\Omega} \cdot \vec{n}(\vec{r})| I(\vec{r}, \vec{\Omega}) = \alpha(\vec{r}) \int\limits_{\vec{\Omega}' \cdot \vec{n}(\vec{r}) > 0} p_{\mathrm{s}}(\vec{r}, \vec{\Omega}' \to \vec{\Omega}) \, I(\vec{r}, \vec{\Omega}') \, \vec{\Omega}' \cdot \vec{n}(\vec{r}) \, \mathrm{d}\vec{\Omega}' + Q_{\mathrm{s}}(\vec{r}, \vec{\Omega}), \tag{2}$$

for $\vec{r} \in \partial\mathrm{R}$ and $\vec{\Omega}(\theta)$, $\theta \in (-\pi, +\pi]$, such that $\vec{\Omega} \cdot \vec{n}(\vec{r}) < 0$. We have introduced here the surface's albedo $\alpha(\vec{r})$ and its phase function $p_{\mathrm{s}}(\vec{r}, \vec{\Omega} \to \vec{\Omega}')$ (in [1/rad]), with $\vec{\Omega} \cdot \vec{n}(\vec{r}) > 0$, to describe bi-directional reflection, as well as the boundary source term $Q_{\mathrm{s}}(\vec{r}, \vec{\Omega})$ (in [photons/s/m/rad]).

We will refer to the union of R and $\partial\mathrm{R}$ as the "optical medium," and to $(\mathrm{R} \cup \partial\mathrm{R}) \times (-\pi, +\pi]$ as the "transport space."

Equations (1)–(2) determine mathematically the forward 2D RT problem, once R and all the coefficients and phase functions ("optical properties") are specified, e.g., as in the next Section. The "flatland" remote sensing problem is to infer quantitative information about the structure or properties of the optical medium, given some (generally quite sparse) sampling of the $I(\vec{r}, \vec{\Omega})$ field in transport space, typically with $\vec{r} \in \partial\mathrm{R}$ and $\vec{\Omega} \cdot \vec{n}(\vec{r}) > 0$. Ideally, the inferred quantities should be accompanied with an estimate of their uncertainty.

III. 2D SCENE GEOMETRY & OPTICAL PROPERTIES

A. Terrain & Illumination

Figure 1 shows a schematic of the optical medium. The lower boundary is represented by

$$z(x) = \frac{H_{\mathrm{sfc}}}{2} \left(1 - \cos^3(2\pi x/L) \right), -L/2 \le x \le +L/2, \tag{3}$$

where we set $L = 2\pi$ km and $H_{\mathrm{sfc}} = 2$ km. The width-to-depth aspect ratio of this terrain model is therefore $L/H_{\mathrm{sfc}} = \pi$. To form $\partial\mathrm{R}$, this lower boundary is complemented by a "sky" defined by

$$\{z = H_{\mathrm{sky}}; -L/2 \le x \le +L/2\}, \tag{4}$$

where we set $H_{\mathrm{sky}} = 4$ km, and by lateral boundaries

$$\{x = \pm L/2; H_{\mathrm{sfc}} < z < H_{\mathrm{sky}}\}, \tag{5}$$

where we apply open/absorbing BCs. We thus set $\alpha(z) = 0$ for $H_{\mathrm{sfc}} < z \leq H_{\mathrm{sky}}$ in (2). This assumption contrasts with the commonly used periodic lateral BCs over the cell $-L/2 < x < +L/2$ associated an absorbing BC only at $z = H_{\mathrm{sky}}$. However, this does not affect the outcome of the present proof-of-concept.

For simplicity, we emulate scene illumination by an overhead sun (solar zenith angle, SZA, $\theta_0 = \pi$ rad) by setting $Q_{\mathrm{v}}(x, z, \theta) \equiv 0$ in (1) and

$$Q_{\mathrm{s}}(x, z, \theta) = \begin{cases} F_0 \delta(\theta - \theta_0), & \text{if } z = H_{\mathrm{sky}}, \text{ and } |x| < 2.5, \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

in (2). The constraint on x (no light emitted from $2.5 < |x| < L/2 = \pi$) reduces the impact of the simplifying assumption of open (rather than periodic) lateral BCs. Here, F_0 denotes the 2D counterpart (measured in [photons/s/m]) of the incoming solar irradiance at the wavelength of interest.
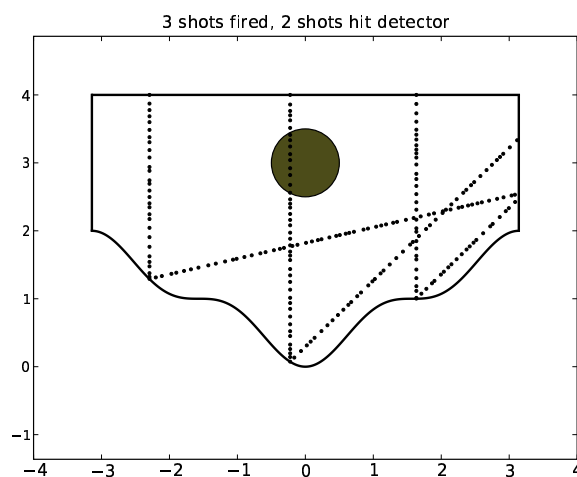

Fig. 1. *Schematic representation of the optical medium.* We show here the terrain with variable height from (3) at the lower boundary, and the upper domain boundaries in (4)–(5). Also, an instance of the absorbing gaseous plume is drawn. For illustration, three solar rays are cast from the overhead direction, one of which intersects the plume region and two of which reach a small hypothetical detector on the right-hand side after one diffuse reflection at the surface.

The surface is assumed everywhere Lambertian, i.e.,

$$p_{\mathrm{s}}(\vec{r}, \vec{\Omega} \to \vec{\Omega}') \equiv |\vec{\Omega}' \cdot \vec{n}(\vec{r})|/2. \tag{7}$$

Note the normalization by 2, instead of π in 3D. It follows directly from the requirement that

$$\int_{\vec{\Omega}' \cdot \vec{n}(\vec{r}) < 0} p_{\mathrm{s}}(\vec{r}, \vec{\Omega} \to \vec{\Omega}') \, \mathrm{d}\vec{\Omega}' = 1$$

where $\mathrm{d}\vec{\Omega}' = \mathrm{d}\theta'$; note that $\vec{\Omega}' \cdot \vec{n}(\vec{r}) = \cos(\theta' - \theta_{\vec{n}(\vec{r})})$ and integration limits are $\theta_{\vec{n}(\vec{r})} \pm \pi/2$. For future (Monte Carlo implementation) use, this means that the new direction of propagation is $\theta' = \theta_{\vec{n}(\vec{r})} + \theta_{\mathrm{s}}(\xi)$ where

$$\theta_{\mathrm{s}}(\xi) = \sin^{-1}(1 - 2\xi), \tag{8}$$

ξ being drawn from a uniform distribution on (0,1). Equivalently, $\theta_{\mathrm{s}}(\xi) = \pm \sin^{-1}\xi$ where the sign is chosen at random. This contrasts with the usual 3D case where $\theta_{\mathrm{s}}(\xi) = \cos^{-1}\sqrt{\xi}$ and the azimuthal angle is chosen at random between $\pm\pi$.

Surface albedo is given by a piece-wise constant function of z:

$$\alpha(\vec{r}) \equiv \alpha(z) = \begin{cases} 0.1, & 0 \le z \le H_{\mathrm{sfc}}/2, \\ 0.5, & H_{\mathrm{sfc}}/2 < z \le H_{\mathrm{sfc}}, \end{cases} \tag{9}$$

as expected, for instance, when vegetation cover changes with altitude. In this case, we go suddenly from sparse to dense vegetation as altitude increases assuming a NIR wavelength (where there are many gaseous absorption features).

B. Aerosols

The optical properties of the atmosphere (in region R) are parameterized as

$$\sigma(\vec{r}) \equiv \sigma(z) \;\; = \;\; \sigma_0 \mathrm{e}^{-cz}, \tag{10}$$

$$\sigma_{\mathrm{s}}(\vec{r}) \equiv \sigma_{\mathrm{s}}(z) \;\; = \;\; \varpi_0 \sigma(z), \tag{11}$$

where

- $\varpi_0 = 0.9$ for the aerosol single scattering albedo (SSA), a typical value, assumed constant with height,
- σ_0 is the extinction coefficient at the low point in the terrain ($x = z = 0$), and $1/c$ is the characteristic scale height of the aerosol.

The "baseline" aerosol is defined by $c = c_0 = 0.5$ km^{-1} ($1/c_0 = 2$ km $= H_{\mathrm{sky}}/2$). In general, the aerosol optical depth (AOD), measured vertically over the low point in the terrain, is given by

$$\tau_{\mathrm{a}}(c) = \frac{\sigma_0}{c} \left(1 - \mathrm{e}^{-cH_{\mathrm{sky}}}\right), \tag{12}$$

and its baseline value is set to unity. This choice yields $\sigma_0 \approx 0.58$ km^{-1}. This scenario corresponds to a relatively heavy aerosol load, in other words, quite hazy observation conditions. However, we will allow for other values further on by varying c.

The scattering phase function is taken to be everywhere equal to a 2D counterpart of the double Henyey–Greenstein (H–G) model [25], namely, [26]

$$p_{\mathrm{v}}(\vec{r}, \theta \to \theta') \equiv f_1 p_{g_1}(\theta_{\mathrm{s}}) + (1 - f_1) p_{g_2}(\theta_{\mathrm{s}}), \tag{13}$$

where $\theta_{\mathrm{s}} = |\theta' - \theta|$ in 2D and [27]

$$p_g(\theta_{\mathrm{s}}) = \left(\frac{1}{2\pi}\right) \frac{1 - g^2}{1 + g^2 - 2g\cos\theta_{\mathrm{s}}}. \tag{14}$$

7

Parameter g is the asymmetry factor of the above 2D H–G phase function model, still defined in 2D as the mean value of $\cos\theta_s$. This leads to $g = f_1 g_1 + (1 - f_1)g_2$ for the double H–G model in (13). We will assume $(f_1, g_1, g_2) = (0.9, +0.8, -0.4)$, hence $g = 0.68$, a representative value for a background aerosol. See Fig. 2 for illustration.
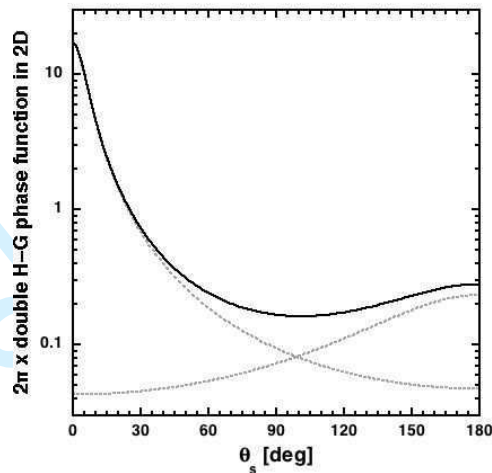

Fig. 2. *Adopted phase function for aerosol scattering.* Expressions for the 2D double H–G model are found in (13)–(14). Parameters are: $f_1 = 0.9$; $g_1 = +0.8$; $g_2 = -0.4$. In gray, we see the two components separately.

Since $\theta_s = \theta' - \theta$ in 2D, the scattering kernel in (1) is simply a convolution of the radiance field with the phase function in direction space. For deterministic solutions of the 2D RT equation, it is therefore important to know the coefficients of the Fourier-cosine series for (14), which are simply g^n (precisely like for the required Legendre series expansion of the 3D H–G model). For a Monte Carlo solution of the 2D RT equation, as is used here, it is best if one knows explicitly the scattering angle $\theta_s(\xi)$ at which the definite integral of (14), starting from $-\pi$, equals $\xi \in [0, 1]$. This turns out to be

$$\theta_s(\xi) = 2\tan^{-1}\left[\left(\frac{1-g}{1+g}\right)\tan\left((\xi - 1/2)\pi\right)\right]. \tag{15}$$

For the double H–G phase function, linear combinations of the Fourier coefficients are computed according to (13). In the Monte Carlo simulations to follow, we draw the scattering angle from either $p_{g_1}(\theta_s)$ or $p_{g_2}(\theta_s)$ based on a simple Bernoulli trial that decides which asymmetry factor to use in (15).

C. Atmospheric Variables

All of the above optical properties are held constant, and assumed known in the remote sensing simulation studies described in later Sections of this report. However, the aerosol stratification parameter c_0 is only a reference value and we consider the aerosol only partially known. The actual atmosphere is given by the same σ_0 in (10)–(11), but with

$$c = c_0 + \delta c \tag{16}$$

where $\delta c.ge. - c_0$ is a perturbation on the baseline inverse scale height c_0. It will automatically impact $\tau_a(c)$ according to (12), hence more (less) total aerosol when $\delta c < (>)0$.

Moreover, the atmosphere contains a uniform plume of purely absorbing gas that we represent parametrically in 2D as a circular region

$$A = \{(x, z)^T \in R; (x - x_p)^2 + (z - z_p)^2 < \rho_p^2\}. \tag{17}$$

Therein the absorption coefficient becomes

$$\sigma_a(\vec{r}) = (1 - \varpi_0)\sigma(z) + k_p, \tag{18}$$

whereas only the first term applies outside of region A. We will vary k_p and, consequently, the optical diameter of the plume $\tau_p = 2\rho_p k_p$. Another interesting property of the plume is its total mass, which is $\propto k_p \rho_p^2$ since k_p is the product of the absorbing particles' volume density—actually *surface* density in 2D—and their (presumably known) cross-section for absorption.

D. Summary

In the simulated retrievals performed in Part II, we will be making inferences about γ, which formally represents the 5-dimensional state vector of remote sensing unknowns, viz.,

$$\gamma = (x_p, z_p, \rho_p, k_p; \delta c). \tag{19}$$

Table I summarizes the properties of the atmosphere-surface system, first as far as the "reference world" is concerned (top 8 rows), then as far as the continuum of "γ worlds" is concerned (bottom 5 rows).

IV. COMPUTATIONAL RADIATIVE TRANSFER MODEL

A. Adopted Monte Carlo Scheme

"Photon" trajectories or histories, as they are commonly but unphysically [28, and references therein] called, are generated in transport space as follows in the plume-free case where $\varpi_0 = \sigma_s/\sigma$ is constant.

1) A starting position and direction $(\vec{r}_0, \vec{\Omega}_0)$ are drawn from the source probability distribution Q_s in (6). In this case, $z_0 = H_{sky}$ and x_0 is drawn randomly from $(-2.5, +2.5)$ and θ_0 is set to π.

2) The photon travels along the path $\vec{r}(t) = \vec{r}_0 + t\vec{\Omega}_0$, $t > 0$, interacting at point $\vec{r}(t_1)$ with a cumulative probability distribution given by

$$P(t) = \Pr\{0 < t_1 < t\} = 1 - \exp\left(-\int_0^t \sigma(\vec{r}_0 + t_1\vec{\Omega}_0)dt_1\right) \tag{20}$$

using (10) with $c = c_0$. In other words, the optical distance would then be defined in terms of t_1: $\tau(t_1) = \int_0^{t_1} \sigma(\vec{r}_0 + s\vec{\Omega}_0)ds$, which is exponentially distributed with unit mean. Such random numders are generated by $-\ln \xi$ where ξ is drawn from a uniform distribution on (0,1).

If the photon does not interact in the volume R, it will always interact with the boundary ∂R. In the present case, the lower surface ("$\cos^3$" terrain) intersects the path just below the starting point, at $(x_0, z(x_0))^T$ using (3), since $\theta_0 = \pi$. This will happen anytime the random number ξ is greater than (20) for $t = z_0 - z(x_0)$.

TABLE I

SUMMARY OF REFERENCE ATMOSPHERE-SURFACE PARAMETERS

Parameter	Symbol	Value	Unit
Minimum height		0	km
Maximum height	H_{sky}	4	km
Terrain height	H_{sfc}	2	km
Terrain width	L	2π	km
Solar zenith angle	θ_0	π	rad
Aerosol SSA	ϖ_0	0.9	[-]
Aerosol asymmetry factor	g	≈ 0.7	[-]
Aerosol scale height	$1/c_0$	2	km
Plume's x-position	x_{p}	varies	km
Plume's z-position	z_{p}	varies	km
Plume's radius	ρ_{p}	varies	km
Plume's absorption coefficient	k_{p}	varies	1/km
Aerosol perturbation	δc	varies	1/km

3) At the interaction point $\vec{r}_1$, the photon is either absorbed or scattered.

- At a surface interaction, $\vec{r}_1 \in \partial R$, the photon will be absorbed with probability $1 - \alpha(\vec{r}_1)$, and the random walk is terminated. If not, it will choose a new direction θ_1 using the probability density function (PDF) $p_{\text{s}}(\vec{r}_1, \theta_0 \to \theta_1)$ in (7), hence using (8).

- At a volume interaction, $\vec{r}_1 \in R$, the photon will be absorbed/terminated with probability $1 - \varpi_0$. If not, it will scatter into a new direction θ_1 using the PDF $p_{\text{v}}(\theta_0 \to \theta_1)$ in (13)–(14), that is, using a Bernouilli trial followed by (15).

4) This continues from Step #2 on, with subscripts incremented by unity, until the photon is absorbed. The main difference is that intersection of the beam with the "cos^3" terrain must now be computed numerically. Exit from the domain is formally accounted for (as an absorption event) by making $\sigma = \infty$ and $\varpi_0 = 0$ outside of R. We note that this will also be true at a detector located on the boundary.

We denote by $n^\star$ the subscript (order of scattering/reflection) at the time of escape or detection.

We note for future reference that this procedure defines a Markov chain and, in Appendix A, Section A-1, a formal probability space is defined for computing means, variances, etc. Also, in view of the "recycling" process described further on, all the positions $\{\vec{r}_0, \vec{r}_1, \vec{r}_2, ..., \vec{r}_{n^\star}, \vec{r}_{n^\star+1}\}$ of all the interactions have to be stored in memory and/or on disk, but only if they end on the detector's surface (at position $\vec{r}_{n^\star+1}$).

Figure 3 illustrates 75 trajectories based on the above algorithm conditioned to end at a small detector on the right-hand side of the medium. More precisely, this detector is defined in transport space by $x = L/2$ and $2.45 < z < 2.55$ along with $0 < \theta < \pi$. When these 75 trajectories—and many more—are tallied and partitioned into 15 equal-sized angular bins from 0^+ to π^-, the resulting radiative fluxes across the lateral boundary at the detector form an "image" of the 2D reference world. Re-centering the angles around $\theta = \pi/2$, the center of the detector's field-of-view, we denote F_i $(i = -7, \ldots, +7)$ where: $i = 0$ is looking horizontally, $i > 0$ is downward (into the variable terrain), and $i < 0$ is upward (toward the sky). Collectively, we will define the image as a formal 15-dimensional vector

$$\mathbf{F}(\gamma) = (F_{-7}, ..., F_{+7}) \tag{21}$$

where, for the moment, $\gamma = (\cdot, \cdot, \cdot, 0; 0)$ since the plume's geometry $(x_\mathrm{p}, z_\mathrm{p}, \rho_\mathrm{p})$ is irrelevant (arbitrary numbers) if $k_\mathrm{p} = 0$ (no absorbers are present). This monochromatic image $\mathbf{F}(\gamma)$ is plotted in Fig. 4 (dashed line). In a typical simulation, 5 to 8 10^5 trajectories ending at the detector were used, starting with as much as $350\times$ more (roughly 2 to 3 10^8 histories). This loss factor is largely dependent on the necessarily finite physical size of the detector.

As expected, radiance from the hazy sky, to the left, is lower than that from the directly and diffusely illuminated reflective surface (with additional path radiance), to the right. Moreover, we recall from (9) that the higher altitudes are assigned a higher (vegetation-type NIR) albedo. Therefore, as the (re-centered) viewing angle increases from 0 to $90°$ and the far end of the high terrain is scanned, high radiance values are found. Then the lower terrain is scanned, resulting lower radiance. The radiance then rebounds when the near side of the higher altitude/reflectivity terrain is reached. Finally, the radiance collapses to near zero values because the nearest terrain is not even illuminated; this artifact goes back to the choice of making the lateral non-terrain boundary "open" (an absorbing BC is applied), and to limit the illumination from $z = H_\mathrm{sky}$ to $|x| < 2.5$ (rather than $L/2 = \pi$). This keeps the ray-tracing simpler, and does not affect the validity of the multi-pixel retrieval demo described in Part II.
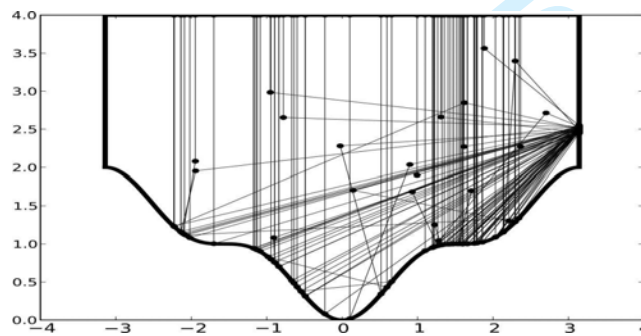
Fig. 3. *Several photon trajectories that all ended at the detector.* 75 trajectories (out of $\approx$26000 total) hit the small detector above the terrain on the right-hand side of the medium. These are the only ones of interest in the following 2D plume characterization by remote sensing.

Note that, although its observations are indexed with the angular coordinate θ in direction space, this is an imaging sensor and we are in the relatively new realm of "multi-pixel" retrieval techniques. (See [29] for an early example

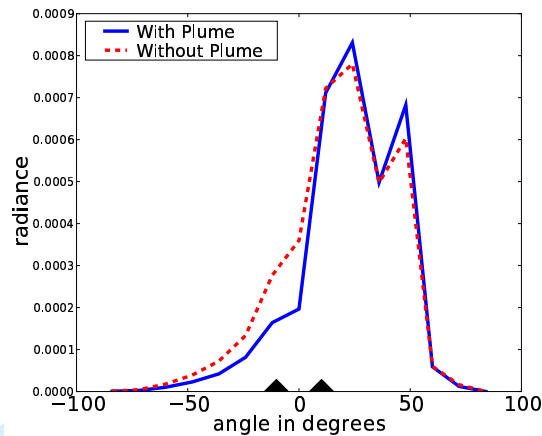Fig. 4.   *Examples of 2D "images" captured by the detector.* Fluxes crossing the right-hand lateral boundary are measured and assigned to the 15 "pixels" in the image. A "fish-eye" optics is assumed (180° field-of-view). Going left to right, they are ordered from zenith ($\theta = 0$) to nadir ($\theta = 180°$), with angles re-centered on the horizontal look ($\theta = 90°$). So, to the left ($-$ve angles), we are looking at sky and, to the right ($+$ve angles), we go from sky into the partially reflective terrain. Two situations are displayed. Dashed line: no absorbing plume and nominal aerosol, $\gamma = (\cdot, \cdot, \cdot, 0; 0)$ in (19), to illustrate the basic Monte Carlo scheme. Solid line: $\gamma = (+0.35 \text{ km}, 2.5 \text{ km}, 0.5 \text{ km}, 0.5 \text{ km}^{-1}; 0.15 \text{ km}^{-1})$, to illustrate the path recycling technique. Direct lines of sight to the plume edges are indicated by the pointers. This plume and aerosol is used as "truth" in the Bayesian inverse problem solution experiments presented in Part II.

of a crude two-pixel/monochromatic/mono-angular technique targeting broken clouds.) The angle-based definition of the pixels used here is entirely traceable to the static sensor's location at close range from the scene. From that vantage, only an angular scan (or a CCD-type focal plane array) can ensure imaging. A space-based counterpart of this simple sensor can collect an image in push-broom mode at a single angle using its orbital motion. Even an imaging focal-plane device at such a large stand-off distance captures light emerging from a small area with quasi-parallel beams (i.e., almost identical θ).

At any rate, the scene reconstruction described further on is inherently multi-pixel in kind, not multi-angle. Truly multi-angle observations would call for multiple sensor positions in the present Monte Carlo simulation. In practice, it can be a single platform moving through space at orbital speeds with multiple push-broom or CCD sensors, e.g., above-mentioned MISR or POLDER missions, respectively.

B. Discussion

The above algorithm is best described as a "brute force" Monte Carlo method. It was coded in Python and could take up to two days to execute on a 4-core 2.6 GHz workstation for 2 to 3 10^8 histories total. With the present detector geometry, only $\approx 1/350$ of these rays contributed to any of the the 15 pixel-scale signals (hence were written to file for future recycling).

Apart from translation from a scripting language to a compiled one and more massive parallelization, the forward Monte Carlo scheme could be accelerated at fixed accuracy in several ways. For instance, in the course of the Monte

Carlo random walk described above, each trajectory could carry, along with 2D coordinates a weight w that accounts for absorption processes, i.e., starting at $w_0 = 1$, it would be reduced by a factor of ϖ_0 at each scattering and by a factor of $\alpha(\vec{r}_1)$ at each reflection; concern over wasting time by pursuing trajectories with negligible weight can be alleviated by using the "Russian roulette" process [30] when a judicious weight threshold is crossed. However, in view of the fact that we will need to recycle all the trajectories that reach the detector, as explained further on, we will need to store not just (x_i, z_i) but also w_i for $i = 0, 1, ..., n^*$. More problematic is that there will be more detection events to revisit and all will carry less than unit weight. So it is not clear that the retrieval method described further on will benefit, only the forward calculation.

The same argument can be made against the "local estimation" technique [3], [30] where a contribution to the observed signal is sent from each and every scattering and reflection event, appropriately weighted by the phase function and the probability of direct transmission to the detector, which can only then be reduced in size to a point. A better idea is to combine backward Monte Carlo, where every simulated path ends at the point detector, with the local estimation technique, this time collecting a contribution from the source at every scattering and reflection along the way through transport space. However, because it is impractical to recycle too many low-weight (i.e., improbable) trajectories, the weight threshold for death by Russian roulette will have to be quite high. The cost of the Russian roulette in increased variance may not be worth the gain in execution time for the forward computation.

In the practical applications that lie beyond the present demonstration, the optimal strategy for the forward computation will depend on how confident we are in the fixed elements of the scene. In this case, we would need to rerun the whole simulation whenever we change the structure of the terrain's shape or height. A change in the position of the source (i.e., time of day) or the detector (e.g., new collection) also mandates a new forward run. At that point, forward or backward local estimation techniques may become more attractive since they can be used quite efficiently to predict signals for multiple detectors or from multiple sources. However, a change in the surface reflectivity map (9) can however be handled by path recycling, to be described next. The same remark applies to a change in the aerosol reference case or its phase function.

Operational circumstances can require a user-directed iteration on the structural properties of the scene (considered fixed in the rest of this study), yet we still need to use a Monte Carlo framework to generate (possibly weighted) paths from source to detector for recycling. In that case, we should seek the most efficient Monte Carlo methods, which are likely to be hybrid ones that use deterministic solutions to achieve radical variance reduction [31, and references therein].

C. Path Recycling

Figure 4 also illustrates simulated observations for one instance of a γ-world (solid line), specifically, when $\gamma = (+0.35 \text{ km}, 2.5 \text{ km}, 0.5 \text{ km}, 0.5 \text{ km}^{-1}; 0.15 \text{ km}^{-1})$ where the key is given in (19). In particular, we see that the absorption optical diameter of the plume, $2\rho_\text{p}k_\text{p}$, is unity. Note the arrow heads on the lower horizontal axis: they indicate the position of the absorbing gaseous plume in the field of view. As expected, the radiance decreases significantly in the pixels that directly image the plume since it reduces the (surface) reflected and (volume) scattered

light that used to stream into them. The radiance increases somewhat in the direction of the more distant terrain ($13°$ and $27°$ bins) as well as the nearby terrain ($54°$ bin). This is because the aerosol element δc in this γ is positive, thus reducing the AOD in (12) from unity (for $c = c_0 \approx 0.58$) to ≈ 0.83 (for $c = c_0 \delta c = 0.65$). The ensuing increased transmittance of direct sunlight to the surface and of surface-leaving radiance to the sensor is therefore overcompensating the reduced path radiance. This is largely because direct sunlight has to be backscattered ($100° \lesssim \theta_s \lesssim 120°$) to reach the detector at these angles, which is unfavorable (cf. Fig. 2).

This second image—in fact a 1D angular scan in 2D space—was computed by recycling the paths that were used to compute the previously discussed reference image (dashed line). Some 334,207 trajectories contributed to these simulated observations (out of a total of $172 \ 10^6$ casted rays), resulting in an average relative Monte Carlo error of 0.44%.

Appendix A describes in full mathematical detail why and how Monte Carlo path recycling works using changes of probability measure. The method has been applied previously in medical imaging applications [32]–[34], but here it is enhanced with new capability. In practice, each path $\vec{r}_i = (x_i, z_i)^{\mathrm{T}}, i = 0, 1, ..., n^* + 1$ that ends at the detector is recalled from memory or disk and assigned a weight of unity. (Here, n^* is necessarily ≥ 1 since only diffuse light can be detected.) It is then reprocessed in two steps.

1) Non-vanishing ρ_p and k_p: First, the weight of the path is changed if any of its n^* segments intersects the circle of radius ρ_p centered at $(x_p, z_p)^{\mathrm{T}}$ that defines the 2D gaseous plume with absorption coefficient k_p. The corresponding multiplicative factor is $\exp(-k_p \ell_{\gamma,i})$ where $\ell_{\gamma,i} \geq 0$ is the length of the intersection of the ith segment with the circle defined in γ. There are highly efficient ways of computing the intersection circles and lines readily adapted to the problem at hand. In practice, the overall factor to use (interpreted in Appendix A, Section A-2, as a change in probability measure) is $\exp(-k_p \ell_\gamma)$ where $\ell_\gamma = \sum_{i=0}^{n^*} \ell_{\gamma,i}$.

2) δc differs from zero: Second, we need to compute the change in the pixel-level signals resulting from redistribution of the background aerosol particle density when δc in (16), the 5th and only non-plume element in γ, is $\neq 0$. It is not obvious that this is possible without casting a whole new set of rays. In Section A-3 of Appendix A it is shown that this is indeed possible and the corresponding change in weight of each detected photon's path is computed explicitly, as in the above for the change in k_p. The computation, however, is more involved than for the effect of the purely absorbing plume that was reduced to straightforward computational geometry. To compute the new signals for $\delta c \neq 0$ (γ world), knowing the ones when $\delta c = 0$ (reference world), we need to look at how δc changes the probability of scattering in one position rather than another along each segment of the broken ray. That part of the photon transport is described mathematically by the cumulative distribution in (20) with the exponential stratification (10) in mind. If we do not want to change the value of c (namely, c_0) in order to leave all the scattering points where they are, that is alright ... as long as the path is properly re-weighted.

The basic question here is whether the new value of c in (16) makes the *given* scattering point, possibly a surface interaction, more or less probable than the what it is in the reference case. In Monte Carlo lingo, this is an application of the "method of dependent paths" [30]. The probability level of the realized scattering point is given by $\mathrm{dP}/\mathrm{d}t$ in (20) using $c = c_0$ while the alternative $\mathrm{dP}_\gamma/\mathrm{d}t$ uses $c = c_0 + \delta c$. Assuming that it ends with a volume

scattering, the weight correction factor we need the ratio

$$\left|\frac{\mathrm{dP}_\gamma}{\mathrm{dP}}\right|_i = \frac{|\mathrm{dP}_\gamma/\mathrm{dt}|_i}{|\mathrm{dP}/\mathrm{dt}|_i} \tag{22}$$

at the ith step where

$$\left|\frac{\mathrm{dP}_\gamma}{\mathrm{dt}}\right|_i = \sigma(\vec{r}_i + t\vec{\Omega}_i)\exp\left(-\int_0^t \sigma(\vec{r}_i + s\vec{\Omega}_i)\mathrm{d}s\right) \tag{23}$$

with $\delta c \neq 0$ in the numerator, and the same with $\delta c = 0$ in the denominator, in expression (10) for $\sigma(\vec{r})$. For a step ending in a boundary (terrain or detector) interaction, we have only the above exponential term to worry about when computing the weight correction ratio. For the whole path, we need to evaluate $|\mathrm{dP}_\gamma/\mathrm{dP}| = \prod_0^{n^*} |\mathrm{dP}_\gamma/\mathrm{dP}|_i$. See §A-3 for the computational procedure using only data from the existing path.

In the end, we need two factors that result from the two terms in (23). On the one hand, we have the product of all the ratios of σ-values, namely, $\exp(-\delta c z_i)$ for all of the *volume* scattering points. On the other hand, we have the product of all the ratios of the exponentials of the reference and perturbed optical distances for segments between *all* (volume and/or surface) scattering points. The latter term amounts to the exponential of σ_0 times a sum of $n^* + 1$ terms that are each a difference of the form $[\mathrm{e}^{-cz_i}(1 - \mathrm{e}^{-ct_i\mu_i})/c\mu_i]_{c=c_0}^{c=c_0+\delta c}$ where t_i is the physical length of the ith segment in the path and $\mu_i = \cos\theta_i$.

3) Summary: The complete change in weight uses the two factors for $\delta c \neq 0$ times the absorbing plume effect, namely, $\mathrm{e}^{-k_{\mathrm{p}}\ell_\gamma}$. That finalizes the algorithmic description of the forward 2D RT model for the observation of interest. In Part II, we turn to the Bayesian inference of the parameterized plume's properties. At that point, only partial knowledge of the aerosol background is assumed, namely, ϖ_0, σ_0, and $p_{\mathrm{v}}(\theta_{\mathrm{s}})$, while c is derived (via δc); the associated AOD $\tau_{\mathrm{a}}(c)$ then follows from (12).

D. Numerical verification/performance of forward model

The accuracy of our forward 2D RT model for the parameterized scene was rigorously verified. In particular, we conducted a variety of tests where

(i) A standard Monte Carlo simulation of transport in an atmosphere parameterized by some value of γ was run until the relative mean square error was less than 0.3%. The mean fluxes are stored as $\mathbf{M}(\gamma) = \{M_i(\gamma),\, i = -7, \ldots, +7\}$.

(ii) Another Monte Carlo simulation was run in an atmosphere parameterized by $\gamma_0 = (0, 1, 0, 0; 0)$ (no plume and baseline aerosol atmosphere). The paths to the detector are stored.

(iii) The paths from (ii) are used in the path-recycling forward model to compute $\mathbf{F}(\gamma)$.

The forward model passes (is deemed verified) if:

- $|F_i(\gamma) - M_i(\gamma)|/M_i(\gamma) \leq 0.01, \forall\, i = -7, \ldots, +7$;
- as γ' becomes sufficiently different from γ, the relative error $|F_i(\gamma') - M_i(\gamma)|/M_i(\gamma), \forall\, i = -7, \ldots, +7$, becomes much worse than 1%.

The forward model was seen to pass for a wide variety of γ values.

The performance increase is dramatic. For example, it took 11,727 minutes to generate approximately 231 million paths (of which $\approx 1/348$ hit the detector). These paths can be recycled in only 30.9 seconds (22,770 times quicker), while one might have expected a speed-up of only about 348 times if one had a Maxwell's daemon that could discriminate at the source rays that would hit the sensor from those that would not. The dramatic difference is due mostly to the fact that the original paths were cast using complicated Python code that explicitly stepped the photons through their path, while the much simpler recycling could be done using optimized code. In any case, recycling paths only involves computing a ratio of weights, and in many cases is generally much quicker than sending the original paths.

V. CONCLUSIONS

We have described an innovative computational approach to multi-dimensional radiative transfer, namely, path-recycling Monte Carlo for simulating remote sensing signals generated by a scene that is spatially variable in two or more dimensions. It is key to the new methodology that the aspects of the spatial structure that are to be varied in the corresponding remote sensing problem are defined parametrically, and hence will have idealized geometry.

The terrain has variable height and varying (height-dependent) albedo and is considered a given throughout the numerical simulations. The adopted terrain model mimics a deep gorge, which would not be unusual for a scenario where one is monitoring a region in search of environmental signatures of nuclear (or other) proliferation activity.

The overlaying atmosphere is composed of a partially known aerosol, with an exponentially decaying density with altitude, and a uniform plume of absorbing gas of unknown location, size, and density (defined via specific absorptivity) that would, for instance, be the observable signature of some covert activity. The plume is parameterized as a circular area and, for simplicity, the radiative transfer unfolds completely in two spatial dimensions $(x, z)^{\mathrm{T}}$. The aerosol's scattering properties (single-scattering albedo and phase function) are assumed known, as is its concentration at the lowest point in the scene; its column-integrated amount (optical depth) is varied by changing the scale-height of the parameterized (exponential) profile. In all, there are 4 plume-related parameters and one aerosol counterpart to retrieve.

In Part II, we demonstrate that atmospheric structure, thus represented parametrically, can be reconstructed reasonably well using a simple imaging sensor at close range when the data is processed using an inherently *multi-pixel* algorithm.

Although they are favored by many in computational 3D radiative transfer, Monte Carlo radiative transfer techniques are notoriously slow to converge in any number of spatial dimensions. Normally, this would make Monte Carlo an impractical way of solving inverse problems in remote sensing, or almost any other application. However, our path-recycling technique reduces the execution time for the limited forward 2D radiative transfer used here to just a few seconds. There is no obvious reason why this procedure could not be implemented in sophisticated 3D scene simulation frameworks that use ray tracing, such as the I3RC community Monte Carlo model [35] or DIRSIG [36].

APPENDIX A

COMPUTATION OF THE CHANGE OF MEASURE

We discuss here the underlying theory and practical implementation of the path-recycling forward Monte Carlo model.

1) Path Measures: The algorithm described in §IV-A induces a measure on the space of finite-length paths, namely,

$$\Omega := \{\omega = (\vec{r}_0, \ldots, \vec{r}_{n^\star+1}) : \vec{r}_j \in \mathrm{R} \cup \partial\mathrm{R}\}$$

where we recall that $n^\star$ is the last order of scattering, including surface reflections, and we take $\vec{r}_{n^\star+1} \in \partial\mathrm{R}$ as the final point. Note that, under reasonable conditions, the (discrete) *stopping time* $n^\star + 1$ of the above Markov chain is $< \infty$. We therefore have a probability measure P_γ. In the special case where $\gamma = \gamma_0 := (0, 1, 0, 0; 0)$, corresponding to no plume and nominal background aerosol, we have our *reference measure* P.

This allows us to define a differential measure $d\mathrm{P}$ and expectation $\mathbb{E}_\mathrm{P}\{\cdot\}$ by

$$\mathrm{P}[A] := \mathbb{E}_\mathrm{P}\{\mathbf{1}_A\} = \int_\Omega \mathbf{1}_A(\omega)\, d\mathrm{P}(\omega) = \int_A d\mathrm{P}(\omega), \tag{A.1}$$

where for $A \subset \Omega$, the *indicator function*

$$\mathbf{1}_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases},$$

and similarly for $d\mathrm{P}_\gamma$.

As a highly relevant example of a subset of paths, consider those that hit (and are necessarily absorbed by) the detector; see Fig. 3 for a few samples. Denote these by a disjoint union

$$\mathcal{D} := D_1 \cup \cdots \cup D_m,$$

meaning that if $\omega \in \mathcal{D}$ then the path ω ended up in the detector, and if $\omega \in D_\nu$ then ω hit the detector with incoming angle θ in the interval $[(\nu - 1)\pi/m, \nu\pi/m)$. Let

$$\mathbf{D} := D_1 \times \cdots \times D_m, \qquad \mathbf{1_D} := (\mathbf{1}_{D_1}, \ldots, \mathbf{1}_{D_m}),$$

and thus our measurement is

$$\mathrm{P}_\gamma[\mathbf{D}] = \mathbb{E}_\gamma\{\mathbf{1_D}\} := (\mathrm{P}_\gamma[D_1], \ldots, \mathrm{P}_\gamma[D_m]).$$

One can similarly define $\mathbb{E}_\mathrm{P}\{\mathbf{1_D}\} = \mathrm{P}[\mathbf{D}]$ in the absence of an absorbing plume. In the main text, we set $m = 15$ and denoted the flux through the νth angular bin (or "pixel") as F_i where $i = 8 - \nu$ (so that $i = 0$ is assigned to a horizontal look to the left in Fig. 1).

2) Restoration of plume absorption k_p: Here we recover only the effect of k_p and, by extension, those of plume geometry contained in the parameter trio $(x_\mathrm{p}, z_\mathrm{p}, \rho_\mathrm{p})$. We assume all other quantities are known and, in particular, that $\delta c = 0$.

Based on (20), the cumulative probability of absorption at the detector in the reference measure P is given by

$$E_\sigma(\vec{r}_0, \ldots, \vec{r}_{n^\star+1}) := \exp\left[-\int_0^{T(\omega)} \sigma\left(\vec{R}(t,\omega)\right) \mathrm{d}t\right], \tag{A.2}$$

a random variable where $\vec{R}(t,\omega) = (x(t,\omega), z(t,\omega))^\mathrm{T}$ is the position of chain ω at time $t \in [0, T(\omega)]$ (in units where the velocity of light is unity), with $T(\omega)$ being the instant of detection. For P_γ, it is therefore given by $E_\sigma(\vec{r}_0, \ldots, \vec{r}_{n^\star+1})\mathrm{e}^{-\alpha\ell_\gamma(\omega)}$, where the random variable $\ell_\gamma(\omega)$ is the total length of intersection of the path ω with the plume parameterized by (the first 4 elements of) γ.

The intersection of an infinite line through a disk can be computed very efficiently via known techniques. From there, it is a simple extension to compute the intersection of a line segment $(\vec{r}_j, \vec{r}_{j+1})$ with a disk, and hence $\ell_{\gamma,j}(\omega)$ is obtained.

Specifically, one first computes

$$\Delta_j = \rho_\mathrm{p}^2 - \left(\left(\vec{r}_\mathrm{p} - \vec{r}_j\right) \times \vec{\Omega}_j\right)^2,$$

where we recall that $\vec{\Omega}_j = (\vec{r}_{j+1} - \vec{r}_j)/\|\vec{r}_{j+1} - \vec{r}_j\|$. If $\Delta_j \leq 0$, the intersection is empty. Otherwise, define $s_{j\pm} = (\vec{r}_\mathrm{p} - \vec{r}_j) \cdot \vec{\Omega}_j \pm \sqrt{\Delta_j}$ and compute s_{j+}. If $s_{j+} \leq 0$, the intersection is empty (the disk is "upwind" from $\vec{r}_j$). Otherwise, compute $\rho_j^2 = (\vec{r}_\mathrm{p} - \vec{r}_j)^2$ and the same for $j+1$. There are than just three possibilities to consider:

- if ρ_j^2 and ρ_{j+1}^2 are both $> \rho_\mathrm{p}^2$, then $\ell_{\gamma,j}(\omega) = 2\sqrt{\Delta_j}$;
- or else, if $\rho_j^2 < \rho_\mathrm{p}^2$ and $\rho_{j+1}^2 > \rho_\mathrm{p}^2$, then $\ell_{\gamma,j}(\omega) = s_{j+}$;
- or else, if $\rho_j^2 > \rho_\mathrm{p}^2$ and $\rho_{j+1}^2 < \rho_\mathrm{p}^2$, then compute $s_{j-} = s_{j+} - 2\sqrt{\Delta_j}$ and set $\ell_{\gamma,j}(\omega) = \|\vec{r}_{j+1} - \vec{r}_j\| - s_{j-}$.

In summary, we have

$$\left|\frac{\mathrm{dP}_\gamma}{\mathrm{dP}}\right|(\omega) = \mathrm{e}^{-k_\mathrm{p}\ell_\gamma(\omega)}, \text{ where } \ell_\gamma(\omega) = \sum_{j=0}^{n^\star} \ell_{\gamma,j}(\omega). \tag{A.3}$$

3) Restoration of γ, including background aerosol perturbation δc: Here we recover the background given by (10)–(11) with (16). Notice that the background absorption/scattering depends only on the height z.

We begin by defining $z_1(\omega), \ldots, z_{n^\star}(\omega)$, the random scattering/reflection heights. Our goal is to compute

$$\mathbb{E}_\gamma\{\mathbf{1}_\mathbf{D}\} = \int_\Omega \mathbf{1}_\mathbf{D}(\omega)\,\mathrm{dP}_\gamma(\omega) = (P_\gamma[D_1], \ldots, P_\gamma[D_m]).$$

We start with the case with no absorbing plume where γ is reduced to $(0, 1, 0, 0; \delta c)$. We will have to differentiate the scattering heights that occur in the volume from those that happen in the volume (due to aerosols) *or* at the lower boundary (due to surface reflection). Let the indices of the former be denoted by $j_1, \ldots, j_{n_\mathrm{s}}$ where n_s is total volume interactions (i.e., bona fide scatterings by aerosol particles).

Here again, we need to compute the Radon–Nikodym derivative [37] $|\,\mathrm{dP}_\gamma/\,\mathrm{dP}|(\omega)$. Note that P_γ differs from P in two ways. First, the integrated extinction coefficient (total cross-section per unit of volume) is $E_{\sigma\gamma}$ rather than

E_σ. Second, the scattering coefficient and phase function in the non-baseline volume are multiplied by a factor $\exp(-\delta c z)$. Therefore, using (10)–(11) and (A.2), but leaving the dependence on ω implicit, we have

$$
\left| \frac{\mathrm{dP}_\gamma}{\mathrm{dP}} \right| = \frac{E_{\sigma^\gamma}(\vec{r}_0, \ldots, \vec{r}_{n^\star+1})}{E_\sigma(\vec{r}_0, \ldots, \vec{r}_{n^\star+1})} \times \frac{\sigma_{\mathrm{s}}^\gamma(\vec{r}_{j_1}) \cdots \sigma_{\mathrm{s}}^\gamma(\vec{r}_{j_{n_{\mathrm{s}}}})}{\sigma_{\mathrm{s}}(\vec{r}_{j_1}) \cdots \sigma_{\mathrm{s}}(\vec{r}_{j_{n_{\mathrm{s}}}})}
$$

$$
= \frac{E_{\sigma^\gamma}(\vec{r}_0, \ldots, \vec{r}_{n^\star+1})}{E_\sigma(\vec{r}_0, \ldots, \vec{r}_{n^\star+1})} \, \mathrm{e}^{-\delta c(z_{j_1} + \cdots + z_{j_{n_{\mathrm{s}}}})}
$$

$$
= E_{\sigma^\gamma - \sigma}(\vec{r}_0, \ldots, \vec{r}_{n^\star+1}) \, \mathrm{e}^{-\delta c(z_{j_1} + \cdots + z_{j_{n_{\mathrm{s}}}})}
$$

$$
= \exp\left(-\int_0^T \left[\sigma^\gamma(\vec{R}(s)) - \sigma(\vec{R}(s)) \right] \mathrm{d}s \right) \times \mathrm{e}^{-\delta c(z_{j_1} + \cdots + z_{j_{n_{\mathrm{s}}}})}
$$

$$
= \exp\left(-\int_0^T \sigma(\vec{R}(s)) \left[\mathrm{e}^{-\delta c z(s)} - 1 \right] \mathrm{d}s \right) \times \mathrm{e}^{-\delta c(z_{j_1} + \cdots + z_{j_{n_{\mathrm{s}}}})}. \tag{A.4}
$$

The 2$^{\mathrm{nd}}$ term, $\exp\left\{ -\delta c(z_{j_1} + \cdots + z_{j_{n_{\mathrm{s}}}}) \right\}$, depends only on the z-coordinate of an identified sub-set of the scattering points $(\vec{r}_1, \ldots, \vec{r}_{n^\star})$. Therefore, like for the plume-related term in (A.3), it can be computed *without casting new rays*: we only need to store the scattering points. We now show that the 1$^{\mathrm{st}}$ term enjoys that feature as well.

Define the scattering/reflection times $T_1, \ldots, T_{n^\star}$, such that $\vec{R}(T_j) := \vec{r}_j$, along with $T_0 = 0$ and $T_{n^\star} = T$. Note that

$$
T_j = |\vec{r}_1 - \vec{r}_0| + \cdots + |\vec{r}_j - \vec{r}_{j-1}|.
$$

When $T_j < s < T_{j+1}$ the Monte Carlo particle is traveling in a straight line given by

$$
x(s) = x_j \pm (s - T_j)\sqrt{1 - \mu_j^2}, \quad z(s) = z_j + (s - T_j)\mu_j,
$$

where μ_j is the vertical direction-cosine of θ_j, the angle between the direction of travel and the upwards unit vector $(0, 1)$; the positive sign is taken if the photon is traveling to the right. We have

$$
\exp\left(-\int_0^T \sigma(\vec{R}(s)) \left[\mathrm{e}^{-\delta c z(s)} - 1 \right] \mathrm{d}s \right)
$$

$$
= \exp\left(-\sigma_0 \int_0^T \mathrm{e}^{-c_0 z(s)} \left[\mathrm{e}^{-\delta c z(s)} - 1 \right] \mathrm{d}s \right\}
$$

$$
= \exp\left(-\sigma_0 \sum_{i=0}^{n^\star} \int_{T_j}^{T_{j+1}} \left[\mathrm{e}^{-(c_0 + \delta c) z(s)} - \mathrm{e}^{-c_0 z(s)} \right] \mathrm{d}s \right). \tag{A.5}
$$

To evaluate this, note that (with $h = c_0$ or $h = c_0 + c$, and assuming $h > 0$),

$$
\int_{T_j}^{T_{j+1}} \mathrm{e}^{-h z(s)} \, \mathrm{d}s = \mathrm{e}^{-h(z_j - T_j \mu_j)} \int_{T_j}^{T_{j+1}} \mathrm{e}^{-h s \mu_j} \, \mathrm{d}s \tag{A.6}
$$

$$
= \frac{\mathrm{e}^{-h z_j}}{h \mu_j} \left[1 - \mathrm{e}^{-h(T_{j+1} - T_j)\mu_j} \right] \tag{A.7}
$$

$$
= \frac{\mathrm{e}^{-h z_j}}{h \mu_j} \left[1 - \mathrm{e}^{-h|\vec{R}_{j+1} - \vec{R}_j|\mu_j} \right]. \tag{A.8}
$$

Each term in (A.5) is evaluated using (A.8). Since this requires only knowledge of the points $\vec{R}_j$, there is no need to re-cast the rays. Finally, (A.4) and (A.5), once evaluated using (A.8), are combined to yield $|\,\mathrm{dP}_\gamma/\,\mathrm{dP}|(\omega)$.

To restore the full effect of γ on $|\,\mathrm{dP}_\gamma/\,\mathrm{dP}|(\omega)$, the term $\mathrm{e}^{-k_\mathrm{P}\ell_\gamma(\omega)}$ computed in Section A-2 (requiring knowledge only of the scattering and reflection points) is factored in to update $|\,\mathrm{dP}_\gamma/\,\mathrm{dP}|(\omega)$. In the path-recycling Monte Carlo code, this Radon–Nikodym derivative is used as the new weight assigned to the random photon path (the RT-related Markov chain) while re-tallying the detector responses.

APPENDIX B

MONTE CARLO MODELS WITH VARYING PRECISION

Adopting the notations and definitions introduced in §A-1, we are in a position to describe our forward 2D Monte Carlo RT model and obtain probabilistic estimates of means and variances (hence errors on the mean). We summarize in this appendix the main results of Bal, Langmore, and Marzouk [38] that are key to the Bayesian approach to the inverse problem presented in Part II.

The model uses importance sampling to compute $\mathrm{P}_\gamma[\mathbf{D}]$ from one fixed set of reference paths. This technique is an advancement over "perturbation Monte Carlo" schemes developed in the context of medical imaging; see, e.g., [32]–[34].

Choosing $N \in \mathbb{N}$, we generate N paths $\{\omega_1, \ldots, \omega_N\}$. Now, for any random variable X,

$$\frac{1}{N} \sum_{j=1}^{N} \mathbf{1}_{\mathbf{D}}(\omega_j) X(\omega_j) \xrightarrow{a.s.} \mathbb{E}_{\mathrm{P}} \{\mathbf{1}_{\mathbf{D}} X\}, \text{ as } N \to \infty,$$

where "*a.s.*" stands for "almost surely". For example, we could generate paths from measure P_γ, and then $N^{-1} \sum_{j=1}^{N} \mathbf{1}_{\mathbf{D}}(\omega_j) \xrightarrow{a.s.} \mathbb{E}_\gamma \{\mathbf{1}_{\mathbf{D}}\}$.

It is important to realize that since we only intend to estimate expectations involving detector hits (e.g., $\mathbb{E}_\gamma \{\mathbf{1}_{\mathbf{D}} X\}$), we only need to store paths that hit the detector. The expected number of detector hits is exactly $N\mathrm{P}[\mathcal{D}] \ll N$.

For every new γ, we could generate a new set of paths and repeat the above procedure. This would be costly since path generation involves complicated steps. Instead, consider fixing one set of reference paths $\{\omega_j, j = 1, \ldots, N\}$ (in practice storing only those that hit the detector) generated by the reference measure P and then set

$$\mathbf{f}_N(\gamma) := \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_{\mathbf{D}}(\omega_k) \left| \frac{\mathrm{dP}_\gamma}{\mathrm{dP}} \right|(\omega_k) \approx \int_\Omega \mathbf{1}_{\mathbf{D}} \left| \frac{\mathrm{dP}_\gamma}{\mathrm{dP}} \right| \mathrm{dP} = \int_\Omega \mathbf{1}_{\mathbf{D}} \, \mathrm{dP}_\gamma = \mathbb{E}_\gamma \{\mathbf{1}_{\mathbf{D}}\}. \tag{B.1}$$

Computation of $\mathbf{f}_N$ requires computing the Radon–Nikodym derivative for the $\approx N\mathrm{P}[\mathcal{D}]$ paths that hit the detector, as described in App. A. This is significantly faster than generating N new paths.

Although already fast, $\mathbf{f}_N$ can be significantly improved by using (for relatively small N) information from a simulation that used a very large N. This is where we depart from the above-mentioned "perturbation Monte Carlo" schemes.

We first generate $N_{\max}$ paths using the reference measure P. Denote by $H^\nu_{\max}$ the collection of paths $\omega_k \in D_\nu$. That is,

$$H^\nu_{\max} := \{\omega_1, \ldots, \omega_{N_{\max}}\} \cap D_\nu.$$

For $\nu = 1, \ldots, m$, the number of observations (in our case, pixels), let

$$H_1^\nu \subset H_2^\nu \subset \cdots \subset H_{\max}^\nu,$$

be nested subsets of $H_{\max}^\nu$ of (fixed, predetermined) size $N_j = |H_j^\nu|$. Note that H_j^ν and $H_{\max}^\nu$ consist of i.i.d. draws from $P[\cdot \,|\, D_\nu]$. Since $|H_{\max}^\nu| = N_{\max}^{-1} \sum_{k=1}^{N_{\max}} \mathbf{1}_{D_\nu}(\omega_k)$, we have

$$\mathrm{Cov}_P\left(|H_{\max}^\nu|, |H_{\max}^\mu|\right) = \frac{1}{N_{\max}} \begin{cases} P[D_\nu] - P[D_\nu]^2, & \nu = \mu \\ -P[D_\nu]P[D_\mu], & \nu \neq \mu \end{cases},$$

where the above $\mathrm{Cov}_P(X, Y)$ is defined as

$$\mathrm{Cov}_P(X, Y) := \mathbb{E}_P\left\{(X - \mathbb{E}_P\{X\})(Y - \mathbb{E}_P\{Y\})^{\mathrm{T}}\right\},$$

with subscript "P" making it clear that expectations are with respect to the probability measure P.

Proof:

$$|H_{\max}^\nu| = \frac{1}{N_{\max}} \sum_{k=1}^{N_{\max}} \mathbf{1}_{D_\nu}(\omega^j), \quad \omega^j \sim P.$$

In other words, it is the sum of $N_{\max}$ i.i.d. random variables ($\mathbf{1}_{D_\nu}/N_{\max}$). The expectation of each random variable is $P[D_\nu]/N_{\max}$. Therefore

$$\mathrm{Cov}_P\left(|H_{\max}^\nu|, |H_{\max}^\mu|\right) = \mathbb{E}_P\left\{ \left(\sum_{k=1}^{N_{\max}} \frac{\mathbf{1}_{D_\nu}(\omega^k)}{N_{\max}} - \frac{P[D_\nu]}{N_{\max}} \right) \times \left(\sum_{\ell=1}^{N_{\max}} \frac{\mathbf{1}_{D_\mu}(\omega^\ell)}{N_{\max}} - \frac{P[D_\mu]}{N_{\max}} \right) \right\}.$$

Since ω^k and ω^ℓ are uncorrelated for $k \neq \ell$ the cross terms are zero, and we are left with a sum of $N_{\max}$ expectations,

$$\frac{1}{N_{\max}^2} \sum_{k=1}^{N_{\max}} \mathbb{E}_P\left\{ \left(\mathbf{1}_{D_\nu}(\omega^k) - P[D_\nu]\right) \left(\mathbf{1}_{D_\mu}(\omega^k) - P[D_\mu]\right) \right\}$$

$$= \frac{1}{N_{\max}} \mathbb{E}_P\left\{ \left(\mathbf{1}_{D_\nu} - P[D_\nu]\right) \left(\mathbf{1}_{D_\mu} - P[D_\mu]\right) \right\}$$

$$= \frac{1}{N_{\max}} \mathbb{E}_P\left\{ \mathbf{1}_{D_\nu} \mathbf{1}_{D_\mu} \right\} - P[D_\nu]P[D_\mu].$$

If $\nu = \mu$, then $\mathbf{1}_{D_\nu} \mathbf{1}_{D_\mu} = \mathbf{1}_{D_\nu}^2 = \mathbf{1}_{D_\nu}$ since $1^2 = 1$ and $0^2 = 0$. On the other hand, if $\nu \neq \mu$ then $\mathbf{1}_{D_\nu} \mathbf{1}_{D_\mu} = 0$ since a photon cannot enter both detectors at once. The result follows. ∎

Although $\{|H_{\max}^\nu|, \nu = 1, \ldots, m\}$ are negatively correlated, so long as $|H_j^\nu|$ may be selected independently of $H_{\max}^\nu$, the sets H_j^ν are independent. We will always ensure this condition holds.

Our improvement on $\mathbf{f}_j$ in (B.1) is $\mathbf{F}_j = (F_j^1, \ldots, F_j^m)$ where

$$F_j^\nu(\gamma) := \frac{|H_{\max}^\nu|}{N_{\max}} \frac{1}{|H_j^\nu|} \sum_{\omega_k \in H_j^\nu} \left| \frac{\mathrm{dP}_\gamma}{\mathrm{dP}} \right| (\omega_k) \tag{B.2}$$

Notice that, if $P = P_\gamma$, then F_j^ν sums $|H_j^\nu|$ i.i.d. draws from $P[\cdot \,|\, D_\nu]$, and each of them scores a hit $|H_{\max}^\nu|/N_{\max}$. In other words, up to the approximations $P_\gamma \approx P$, and $|H_{\max}^\nu|/N_{\max} \approx P[D_\nu]$, $F_j^\nu(\gamma)$ sums $|H_j^\nu|$ random variables, each one recording the exact solution. Hence, up to these approximations, $F_j^\nu(\gamma)$ computes $P[D_\nu]$ with zero variance.

On the practical side, $H_j^\nu \subset H_{j'}^\nu$, for $j < j'$, and therefore the computation of $\mathbf{F}_{j'}$ is quicker after computation of $\mathbf{F}_j$ is done.

The next theorem shows that the estimates $\mathbf{F}_j$ are unbiased. See [38] for a proof.

Theorem B.1.

$$\mathbb{E}_\mathrm{P}\left\{\mathrm{F}_j(\gamma)\right\} = \mathbb{E}_\gamma\left\{\mathbf{1_D}\right\} = \mathrm{P}_\gamma[\mathbf{D}].$$

The following theorem shows that, in the limit where $\mathrm{dP}_\gamma \to \mathrm{dP}$ and $|H_{\max}^\nu| \to \infty$, the $\mathbf{F}_j^\nu(\gamma)$ are uncorrelated zero-variance estimates of $\mathrm{P}_\gamma[D_\nu]$. Again, see [38] for a proof.

Theorem B.2. *As $N_{\max} \to \infty$,*

$$\mathrm{Cov}_\mathrm{P}(\mathrm{F}_j^\nu(\gamma), \mathrm{F}_j^\mu(\gamma)) \to \delta_{\mu\nu} \frac{\mathrm{P}[D_\nu]}{|H_j^\nu|} \times \int_{D_\nu} \left(\left|\frac{\mathrm{dP}_\gamma}{\mathrm{dP}}\right|(\omega) - \frac{\mathrm{P}_\gamma[D_\nu]}{\mathrm{P}[D_\nu]}\right)^2 \mathrm{dP}(\omega).$$

Remark B.1. A similar calculation shows that

$$\mathrm{Cov}_\mathrm{P}(\mathbf{f}_{N_j}^\nu(\gamma), \mathbf{f}_{N_j}^\mu(\gamma)) = \frac{1}{N_j}\begin{cases}\int_{D_\nu}\left(\left|\frac{\mathrm{dP}_\gamma}{\mathrm{dP}}\right| - \mathrm{P}_\gamma[D_\nu]\right)\mathrm{dP}_\gamma, & \nu = \mu \\ -\mathrm{P}_\gamma[D_\nu]\mathrm{P}_\gamma[D_\mu], & \nu \neq \mu\end{cases}.$$

In the expression for $\mathrm{Cov}_\mathrm{P}(\mathrm{F}_j^\nu(\gamma_0), \mathrm{F}_j^\mu(\gamma_0))$, one can replace $|H_j^\nu|$ with $N_j\mathrm{P}[D^\nu]$ and see that, if $\mathrm{dP} \approx \mathrm{dP}_\gamma$, then $\mathrm{Cov}_\mathrm{P}(\mathrm{F}_j^\nu(\gamma), \mathrm{F}_j^\mu(\gamma)) \ll \mathrm{Cov}_\mathrm{P}(\mathrm{F}_j^\nu(\gamma_0), \mathrm{F}_j^\mu(\gamma_0))$. In other words, the variance of our unbiased estimator $\mathbf{F}_j$ is significantly smaller than the estimator $\mathbf{f}_{N_j}$ typically used in aforementioned perturbation Monte Carlo schemes.

REFERENCES

[1] S. Chandrasekhar, *Radiative Transfer*.   Oxford (UK): Oxford University Press, 1950, [reprinted by Dover Publications, New York (NY), 1960].

[2] A. B. Davis and A. Marshak, "Solar radiation transport in the cloudy atmosphere: A 3D perspective on observations and climate impacts," *Reports on Progress in Physics*, vol. 73, no. 2, p. 026801, 2010.

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

[4] J. Otterman and R. S. Fraser, "Adjacency effects on imaging by surface reflection and atmospheric scattering: Cross radiance to zenith," *Appl. Opt.*, vol. 18, pp. 2852–2860, 1979.

[5] Y. Mekler and Y. J. Kaufman, "The effect of Earth's atmosphere on contrast reduction for a nonuniform surface albedo and "two-halves" field," *J. Geophys. Res.*, vol. 85, pp. 4067–4083, 1980.

[6] J. Otterman, S. Ungar, Y. J. Kaufman, and M. Podolak, "Atmospheric effects on radiometric imaging from satellites under low optical thickness conditions," *Remote Sens. Environ.*, vol. 9, pp. 115–129, 1980.

[7] D. Tanré, M. Herman, and P.-Y. Deschamps, "Influence of the background contribution upon space measurements of ground reflectance," *Appl. Opt.*, vol. 20, pp. 3676–3684, 1981.

[8] Y. J. Kaufman, "Solution of the equation of radiative-transfer for remote-sensing over nonuniform surface reflectivity," *J. Geophys. Res.*, vol. 87, pp. 4137–4147, 1982.

[9] D. J. Diner and J. V. Martonchik, "Atmospheric transfer of radiation above an inhomogeneous non-Lambertian ground: 1 - Theory," *J. Quant. Spectrosc. Radiat. Transfer*, vol. 31, pp. 97–125, 1984.

[10] ——, "Atmospheric transfer of radiation above an inhomogeneous non-Lambertian ground: 2 - Computational considerations and results," *J. Quant. Spectrosc. Radiat. Transfer*, vol. 31, pp. 279–304, 1984.

[11] A. Royer, A. Davis, and N. O'Niell, "Analyse des effets atmosphériques dans les images HRV de SPOT," *Can. J. Remote Sens.*, vol. 14, pp. 80–91, 1989.

[12] T. Takashima and K. Masuda, "Simulation of atmospheric effects on the emergent radiation over a checkerboard type of terrain," *Astrophys. and Space Sci.*, vol. 198, pp. 253–263, 1992.

[13] P. N. Reinersman and K. L. Carder, "Monte Carlo simulation of the atmospheric point-spread function with an application to correction for the adjacency effect," *Appl. Opt.*, vol. 34, pp. 4453–4471, 1995.

[14] A. I. Lyapustin, "Three-dimensional effects in the remote sensing of surface albedo," *IEEE Trans. Geosc. and Remote Sens.*, vol. 39, pp. 254–263, 2001.

[15] A. I. Lyapustin and Y. J. Kaufman, "Role of adjacency effect in the remote sensing of aerosol," *J. Geophys. Res.*, vol. 106, pp. 11,909–11,916, 2001.

[16] A. I. Lyapustin and Y. Knyazikhin, "Green's function method for the radiative transfer problem. II. Spatially heterogeneous anisotropic surface," *Appl. Opt.*, vol. 41, pp. 5600–5606, 2002.

[17] A. I. Lyapustin, T. Muldashev, and Y. Wang, "Code SHARM: Fast and accurate radiative transfer over spatially variable anisotropic surfaces," in *Light Scattering Reviews*, A. A. Kokhanovsky, Ed. Heidelberg (Germany): Springer Praxis, 2010, vol. 5, pp. 205–247.

[18] A. A. Semenov, A. V. Moshkov, V. N. Pozhidayev, A. Barducci, P. Marcoionni, and I. Pippi, "Estimation of normalized atmospheric point spread function and restoration of remotely sensed images," *IEEE Trans. Geosc. and Remote Sens.*, vol. 49, pp. 2623–2634, 2011.

[19] R. Richter and D. Schlaepfer, "Geo-atmospheric processing of airborne imaging spectrometry data, Part 2: Atmospheric/topographic correction," *Int. J. of Remote Sens.*, vol. 23, pp. 2631–2649, 2002.

[20] P. E. Ardanuy, D. Han, and V. V. Salomonson, "The MODerate resolution Imaging Spectrometer (MODIS) science and data system requirements," *IEEE Trans. Geosc. and Remote Sens.*, vol. 29, pp. 75–88, 1991.

[21] H. H. Aumann, M. T. Chahine, C. Gautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkranz, W. L. Smith, D. H. Staelin, L. L. Strow, and J. Susskind, "AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, pp. 253–264, 2003.

[22] D. J. Diner, J. C. Beckert, T. H. Reilly, C. J. Bruegge, J. E. Conel, R. A. Kahn, J. V. Martonchik, T. P. Ackerman, R. Davies, S. A. W. Gerstl, H. R. Gordon, J.-P. Muller, R. B. Myneni, P. J. Sellers, B. Pinty, and M. Verstraete, "Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, pp. 1072–1087, 1998.

[23] P.-Y. Deschamps, F.-M. Bréon, M. Leroy, A. Podaire, A. Bricaud, J.-C. Buriez, and G. Sèze, "The POLDER mission: Instrument characteristics and scientific objectives," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, pp. 598–615, 1994.

[24] I. Langmore, G. Bal, and A. B. Davis, "Parametric 3D scene reconstruction from imaging radiometry, Part II: The Bayesian multi-pixel inversion algorithm," *IEEE Trans. Geosc. and Remote Sens.*, submitted.

[25] L. C. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," *Astrophys. J.*, vol. 93, pp. 70–83, 1941.

[26] R. L. White, "Polarization in reflection nebulae. I - Scattering properties of interstellar grains," *Astrophys. J.*, vol. 229, pp. 954–961, 1979.

[27] A. Davis, P. Gabriel, S. Lovejoy, and D. Schertzer, "Scaling laws for asymptotically thick clouds, dimensional dependence - phase function independence," in *IRS'88: Current Problems in Atmospheric Radiation*, J. Lenoble and J.-F. Geleyn, Eds. Hampton (Va): Deepak Publ., 1989, pp. 103–106.

[28] M. I. Mishchenko, "Multiple scattering, radiative transfer, and weak localization in discrete random media: Unified microphysical approach," *Rev. Geophys.*, vol. 46, p. RG2003, 2008.

[29] A. B. Davis, "Cloud remote sensing with sideways-looks: Theory and first results using Multispectral Thermal Imager (MTI) data," *SPIE Proc.*, vol. 4725, pp. 397–405, 2002.

[30] G. Marchuk, G. Mikhailov, M. Nazaraliev, R. Darbinjan, B. Kargin, and B. Elepov, *The Monte Carlo Methods in Atmospheric Optics*. New York (NY): Springer-Verlag, 1980.

[31] G. Bal, A. B. Davis, and I. Langmore, "A hybrid (Monte-Carlo/deterministic) approach for multi-dimensional radiation transport," *Journal of Computational Physics*, vol. 230, pp. 7723–7735, 2011.

[32] C. K. Hayakawa, J. Spanier, F. Bevilacqua, A. K. Dunn, J. S. You, B. J. Tromberg, and V. Venugopalan, "Perturbation Monte Carlo methods to solve inverse photon migration problems in heterogeneous tissues," *Opt. Lett.*, vol. 26, pp. 1335–1337, 2001.

[33] C. K. Hayakawa and J. Spanier, "Perturbation Monte Carlo methods for the solution of inverse problems," in *Monte Carlo and Quasi Monte Carlo Methods 2002*, H. Niederreiter, Ed.   New York (NY): Springer, 2004, pp. 227–242.

[34] J. Chen and X. Intes, "Time-gated perturbation Monte Carlo for whole body functional imaging in small animals," *Opt. Express*, vol. 17, pp. 19,566–19,579, 2009.

[35] http://i3rc.gsfc.nasa.gov/I3RC_community_model_new.htm/.

[36] http://dirsig.org/.

[37] G. E. Shilov and B. L. Gurevich, *Integral, Measure, and Derivative: A Unified Approach*.   Dover Publications, 1978.

[38] G. Bal, I. Langmore, and Y. Marzouk, "Bayesian inverse problems with Monte Carlo forward models," 2011, to be submitted.

Author biographies are provided in Part II.