Proc. European Conference on Computer Vision (ECCV) 2020

Monotonicity Prior for Cloud Tomography

Tamar Loeub[1], Aviad Levis[2], Vadim Holodovsky[1], and Yoav Y. Schechner[1]

[1] Viterbi Faculty of Electrical Engineering,
Technion - Israel Institute of Technology, Haifa, Israel
[2] Computing and Mathematical Sciences Department,
California Institute of Technology, Pasadena, CA, USA
tamarloeub@gmail.com, aviad.levis@gmail.com,
vholod@ef.technion.ac.il, yoav@ee.technion.ac.il

Abstract. We introduce a differentiable monotonicity prior, useful to express signals of monotonic tendency. An important natural signal of this tendency is the optical extinction coefficient, as a function of altitude in a cloud. Cloud droplets become larger as vapor condenses on them in an updraft. Reconstruction of the volumetric structure of clouds is important for climate research. Data for such reconstruction is multi-view images of each cloud taken simultaneously. This acquisition mode is expected by upcoming future spaceborne imagers. We achieve three-dimensional volumetric reconstruction through stochastic scattering tomography, which is based on optimization of a cost function. Part of the cost is the monotonicity prior, which helps to improve the reconstruction quality. The stochastic tomography is based on Monte-Carlo (MC) radiative transfer. It is formulated and implemented in a coarse-to-fine form, making it scalable to large fields.

Keywords: Scattering, Regularization, Physics-based vision

1 Introduction

Clouds have a key role of the climate system. They are part of the water cycle (the source of freshwater supply), and account for $\approx 2/3$ of the Earth's albedo. Inaccurate models of cloud properties yield the largest errors in climate predictions. To put matters in perspective, an error of $\approx 1\%$ in cloud properties yields errors in current climate predictions, that are comparable to climatic effects of man-made greenhouse gases. To reduce climate prediction uncertainties, it is required to model clouds far better than the state-of-the-art. Models, naturally, need observations to set parameters and statistics. Observations require both relevant raw data and proper ways to extract information from this data, regarding cloud structure.

However, current observations cannot meet this need. Cloud observations on a planetary scale rely on 2D satellite images to provide raw data. Then, analysis of the data assumes the atmosphere is layered. In other words, the prevailing assumption underlying analysis and modeling is that clouds are homogeneous horizontally for infinite extents, having only 1D vertical variations. This is termed

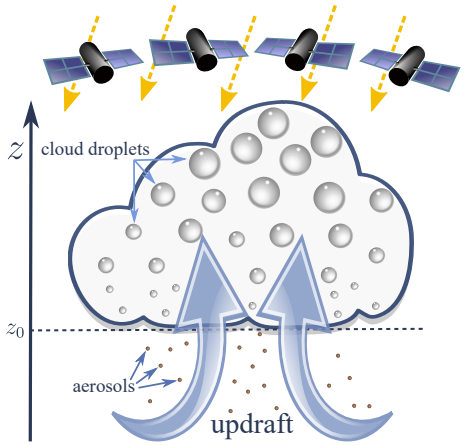

Fig. 1. As air containing water vapor ascends, humidity starts condensing at the condensation level z_0, forming a cloud base. Condensation is initiated by aerosols (condensation nuclei), which are much smaller than water droplets. In the cloud, condensation increases the size (radius) of the droplets, with altitude. The clouds are irradiated by the sun and observed by multiple cameras. From the images, 3D volumetric information is sought.

the *plane parallel* model. Analysis is based on 1D radiative transfer (RT) across layers. However, clouds are naturally 3D volumetric objects having spatial heterogeneity, and light propagates in them according to 3D-RT. This discrepancy yields significant errors and uncertainties the smaller the clouds are. While such clouds are small, they are numerous.

For these reasons, it is important to derive the volumetric structure of small clouds. This requires devising 3D computed tomography (CT) based on multiple scattering [1, 8, 15, 16, 19, 25, 31–33] and passive imaging. Scattering-based CT stands contrary to traditional CT, which relies on linear image formation models [3, 17, 18, 40, 57, 59] and treats scattering as a disturbance. It has recently been shown [14] that scattering-based bio-medical CT can have advantages relative to linear X-ray CT, in terms of radiation dose, chemical analysis and elimination of moving parts.

This paper advances scattering-based tomography of small clouds. To enable imaging over wide scales, the assumed imaging setup is a formation of orbiting satellites, designed to image cloud fields simultaneously from multiple directions (Fig. 1). This is in accordance to a space mission planned to provide raw image data dedicated for tomographic analysis [54]. CT analysis relies on optimized fitting of an image formation model to data, while complying with priors. In this paper, we introduce and use a differentiable monotonicity prior. It is helpful for expressing the nature of convective clouds. Consequently, the prior helps in lowering the recovery error and reaching faster convergence.

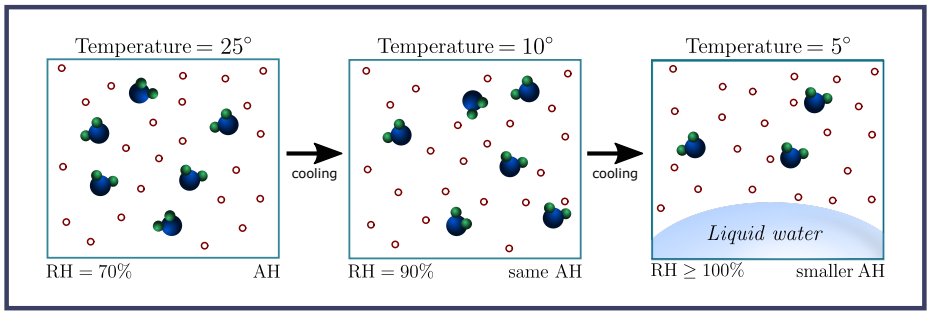

Fig. 2. An illustration of an isolated air parcel, demonstrating cloud-droplet condensation, as described in Sec. 2. The absolute humidity is AH, and the relative humidity is RH. The right-most stage is at saturation or super-saturation

Moreover, our algorithm is scalable to large fields. It is a stochastic approach, based on MC RT, where accuracy is a matter of time to simulate transport of photons. We develop the recovery in a coarse-to-fine (hierarchical) approach, to speed convergence and handling of large fields. This is contrary to prior atmospheric scattering CT, which is based on discrete ordinates [31, 32]. The latter uses high quality RT which is memory-limited: the domain size is severely limited by memory resources, and not designed for grid coarsening.

Proper real raw spaceborne data is expected to become available in a few years [54]. Therefore, we test the methods using simulations. The simulations emulate realistic photon (Poisson) noise, by using realistic photon counts per pixel in the raw images.

2 Droplets in Convective Clouds

To form priors on an object, it is helpful to know its nature. We put to use a cloud trend. First, we note three terms: saturation, absolute humidity and relative humidity. *Absolute humidity* quantifies the density of water vapor in air ($[\mathrm{gr}/m^3]$). *Saturation* is a state at which vapor is at equilibrium with liquid water. *Relative humidity* is the ratio of absolute humidity of an air parcel, to the parcel's absolute humidity at saturation. By definition at saturation, the relative humidity is 100%. For an unsaturated *isolated air parcel*, the absolute humidity is fixed, but relative humidity increases as temperature decreases.

Convective clouds require three ingredients: humidity (water vapor), ascending air current (updraft) and condensation nuclei (aerosols). The ascending air lifts water vapor and aerosols. Ascent leads to cooling of air with altitude z. A lower temperature increases the relative humidity. At some altitude, cooling yields a critical temperature at which relative humidity reaches (or surpasses) 100%. This altitude is termed *condensation level*, or *cloud-base*, z_0. Above this level, the vapor reaches over-saturation and thus starts to condense on the aerosol particles, creating tiny water droplets (Fig. 2). As the air continues to rise,

the *droplets grow in size*. Water vapor that had not condensed into droplets at lower altitude continues to condenses onto the ascending droplets, enlarging them (Fig. 1).

We now formulate the above mathematically. Denote by $\rho_{\rm w} \approx 1 \frac{\rm gr}{\rm cm^3}$ the mass density of liquid water. Denote by r the random radius of a single water droplet. A unit air volume includes many distinct cloud droplets, whose size distribution is $n(r)$. The total number of cloud droplets in the unit volume is $\int_0^\infty n(r)dr$. Aggregating all droplets in the unit volume, the mass of liquid water is termed *Liquid Water Content* (LWC). The LWC and effective radius $r_{\rm eff}$ of these droplets are defined [35] by

$$\text{LWC} = \frac{4}{3}\rho_{\rm w}\int_0^\infty \pi r^3 n(r)dr \quad , \quad r_{\rm eff} = \frac{\int_0^\infty \pi r^3 n(r)dr}{\int_0^\infty \pi r^2 n(r)dr} \ . \tag{1}$$

Intuitively, $r_{\rm eff}$ is a weighted sum of the droplets radiuses, according to the geometric cross section of each droplet. In nature, both $r_{\rm eff}$ and LWC tend to *increases with altitude* inside a cloud [34, 39, 60]. A model for this at $z > z_0$ is

$$r_{\rm eff} \propto (z - z_0)^{\frac{1}{3}} \ , \qquad \text{LWC} \propto (z - z_0) \ . \tag{2}$$

The relations in (2), however, are approximate. They assume, as mentioned in the beginning of this section, that air parcels are *isolated*, exchanging neither material nor heat with surrounding parcels. This assumption is common in atmospheric models, and is termed the *adiabatic* approximation. Real clouds experience mixing with surrounding air parcels, evaporation and in some cases, precipitation. Therefore, we do not impose Eq. (2) as a constraint, but it leads us to regularization that tolerates deviation (Sec. 4).

3 Radiative Transfer and Memory Loads

Many methods in remote sensing of the atmosphere assume a *plane parallel* medium [23, 36, 41], in which the atmosphere is assumed to be layered. There, radiance varies essentially only vertically [10, 27], i.e. the model is 1D-RT. However, we seek recovery of clouds as they are: 3D volumetric and heterogeneous. This is consistent with recent work in computer vision [2, 19, 31]. In a 3D heterogeneous medium, the forward model is 3D-RT. Attempting 3D-RT is computationally complex. Compounding complexity is pursuit of a 3D *inverse problem*: recovering an unknown 3D scattering medium. Solving such an inverse problem requires repeated, iterative calculations: 3D-RT calculations are to be repeated over a hypothetical medium, which varies as iterations progress.

One approach for 3D-RT forward model calculations is deterministic, epitomized by the Spherical Harmonics Discrete Ordinate Method (SHDOM) [11]. As in analysis of reflection in computer vision [4, 50], spherical harmonics help in efficient analysis of angular scattering. SHDOM is fast and accurate for small domain sizes, when many radiometric quantities (i.e. pixels for image synthesis) are sought. However, its memory use is very expensive, as we explain.

SHDOM uses a discrete representation of angles. Let N_{voxels} be the number of voxels and N_{angles} be the number of angles. Typically, $N_{\text{angles}} \approx 500$. To reach accuracy, SHDOM [11] relies on an adaptive subdivision of voxels and a finite number of spherical harmonics. This leads memory requirements to inflate by a factor C. In a typical cloudy scene, $C \approx 75$. Hence, memory load is $\approx CN_{\text{angles}}N_{\text{voxels}} \approx 4 \cdot 10^4 \cdot N_{\text{voxels}}$. This expense inhibits scalability to large domains. Moreover, subdivision is enforced, because SHDOM is based on assumptions of cloud voxels having high internal uniformity and small optical depth. Consequently, SHDOM is not amenable to hierarchical coarse-to-fine analysis.

Contrary to SHDOM, stochastic forward model RT done by MC [38, 58] is not prone to these problems. In MC, simulated photons independently roam the domain and interact randomly (Sec. 5). Angles are not discretized in memory structure per voxel. Hence, memory consumption is $\approx N_{\text{voxels}}$. MC is highly *parallelizable* due to photon independence [48]. Accuracy is controlled (as is runtime) by the number of photons simulated. The variance of the RT result decreases linearly with the number of photons sampled, and does not require adaptive spatial subdivision. It is, therefore, scalable to large domains. Using MC as the forward model along with our hierarchical coarse-to-fine approach, enables solving large scale inverse scattering problems.

4 Inverse Problem

We seek to recover the volumetric optical parameters of a scattering medium, based on two dimensional images, i.e. perform scatter-based tomography. A scattering medium domain is $\mathcal{M} \subset \mathbb{R}^3$. Consider an infinitesimal volume element, voxel v, around location $\boldsymbol{x} \in \mathcal{M}$. The extinction coefficient of the medium at a point $\boldsymbol{x}$ is denoted by $\beta(\boldsymbol{x})$. When multiple particle types exist in a medium, each has its own extinction coefficient. For example, if a volume has air molecules and cloud droplets, then their respective extinction coefficients are $\beta^{\text{a}}(\boldsymbol{x})$ and $\beta^{\text{c}}(\boldsymbol{x})$. The total extinction coefficient at $\boldsymbol{x}$ is

$$\beta(\boldsymbol{x}) = \beta^{\text{a}}(\boldsymbol{x}) + \beta^{\text{c}}(\boldsymbol{x}). \tag{3}$$

The field $\beta^{\text{a}}(\boldsymbol{x})$ is known from auxiliary measurements of the atmosphere, e.g. through radiosondes (weather balloons). Therefore, the computer vision task is to recover the unknown cloud extinction β^{c}_v for each voxel $v \in [1, \ldots, N_{\text{voxels}}]$. The vector of unknowns is $\boldsymbol{\beta} = \left[\beta^{\text{c}}_1, \beta^{\text{c}}_2, \ldots, \beta^{\text{c}}_{N_{\text{voxels}}}\right]$.

Images are obtained from several viewpoints, in different locations and angles. The radiance measurements are indexed by $d \in [1, \ldots, N_{\text{detectors}}]$. They are concatenated into a data vector $\boldsymbol{i}^{\text{measured}}$. We fit to $\boldsymbol{i}^{\text{measured}}$ a corresponding model image set, concatenated to a vector, $\boldsymbol{i}(\boldsymbol{\beta})$. Data fitting is quantified by the difference between measurements $\boldsymbol{i}^{\text{measured}}$ and the corresponding model image set $\boldsymbol{i}(\boldsymbol{\beta})$, using

$$F(\boldsymbol{\beta}) = \frac{1}{2}\|\boldsymbol{i}(\boldsymbol{\beta}) - \boldsymbol{i}^{\text{measured}}\|_2^2 \,. \tag{4}$$

The field $\boldsymbol{\beta}$ is estimated by optimization

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}\geq 0} \left[F(\boldsymbol{\beta}) + R(\boldsymbol{\beta})\right] \ . \tag{5}$$

Here $R(\boldsymbol{\beta})$ is a regularization term, which expresses priors on the medium. The optimization problem in Eq. (5) is solved using stochastic gradient descent (SGD). Sec. 5 describes the image formation model $\boldsymbol{i}^{\text{measured}}$, its implications to $F(\boldsymbol{\beta})$ and the optimization problem in Eq. (5). We now discuss the regularization term.

Regularization

Priors are useful to better constrain results to more physically plausible solutions and aid convergence and quality performance. Recall Sec. 2. For a given droplet size distribution $n(r)$, the extinction coefficient [35] by cloud droplets is

$$\beta^{\text{c}} = \int_0^\infty Q_{\text{eff}}(r)\pi r^2 n(r) dr \ . \tag{6}$$

Here $Q_{\text{eff}}(r)$ is an extinction efficiency factor. For cloud droplets[3], $Q_{\text{eff}}(r) \approx 2$ thus insensitive to r. Combining Eqs. (1,6),

$$\beta^{\text{c}}(\boldsymbol{x}) = \frac{3}{2\rho_{\text{w}}} \frac{\text{LWC}(\boldsymbol{x})}{r_{\text{eff}}(\boldsymbol{x})} \ . \tag{7}$$

Then, considering Eqs. (2,7), for $z > z_0$,

$$\beta^{\text{c}}(\boldsymbol{x}) \propto (z - z_0)^{\frac{2}{3}} \ . \tag{8}$$

As written in Sec. 2, Eq. (2) is an approximation, thus we use Eq. (8) to form a loose prior: $\beta^{\text{c}}(\boldsymbol{x})$ tends to *monotonically increase* with z, within a cloud. The trend of vertical increase in $\beta^{\text{c}}(\boldsymbol{x})$ *only applies* to voxels in which a cloud resides: outside cloud voxels, $\beta^{\text{c}}(\boldsymbol{x}) \equiv 0$. Hence the monotonicity preference is nulled outside a cloud.

For a possible expression for such a prior, define $\mathbb{1}\{\text{statement}\}$ to be a binary column vector. Each element corresponds to a voxel v and satisfies

$$\mathbb{1}\{\text{statement}\}_v = \begin{cases} 1 & \text{statement is true in } v \\ 0 & \text{otherwise} \end{cases} \ . \tag{9}$$

Further, let $\mathbf{D}_z$ be a matrix expressing discrete vertical differentiation, in lieu of continuous differentiation $\partial\beta^{\text{c}}(\boldsymbol{x})/\partial z$. Then, consider this term:

$$R_{\text{mon}} = -\left[\mathbb{1}\left\{\mathbf{D}_z\boldsymbol{\beta}^{\text{c}} > 0\right\}\right]^{\intercal} \mathbb{1}\left\{\boldsymbol{\beta}^{\text{c}} > 0\right\} \ , \tag{10}$$

[3] The factor $Q_{\text{eff}}(r)$ depends on the wavelength λ. For clouds, typically $r \gg \lambda$. Then, it is reasonable to neglect the dependency of Q_{eff} on r and λ [35].

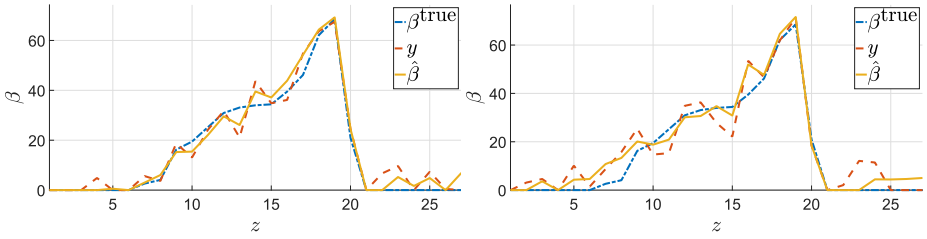

Fig. 3. A 1D example illustrating the denoising capability of the monotonicity prior R_{mon}, when applied to a non-negative signal whose origin tends to be monotonous or null. [Left] The input $\boldsymbol{y}$ noise is $\sim \mathcal{N}(0,5)$. Using Eq. (13), the error in $\hat{\boldsymbol{\beta}}$ is $\epsilon = 15\%$. [Right] The input $\boldsymbol{y}$ noise is $\sim \mathcal{N}(0,10)$. The error in $\hat{\boldsymbol{\beta}}$ is $\epsilon = 20\%$.

where $\intercal$ denotes transposition.

Consider voxels where $\boldsymbol{\beta}^{\text{c}}$ increases vertically. These voxels decrease R_{mon}. This lowers the optimization cost (5), if R_{mon} is used for regularization. Voxels which are not in a cloud or have vertical decrease of the extinction coefficient do not contribute to decreasing of R_{mon}, meaning essentially a higher cost.

Eq. (10) is not differentiable with respect to $\boldsymbol{\beta}^{\text{c}}$. To enable practical optimization based on SGD, we use a softer version. For an arbitrary vector $\mathbf{q}$, let $\tanh(\mathbf{q})$ be a diagonal matrix, where the diagonal is defined by the hyperbolic-tangent of each element of $\mathbf{q}$. Let $\mathbf{1}$ be a column vector, of all whose elements are ones. We use

$$R_{\text{mon}}(\boldsymbol{\beta}^{\text{c}}) = -\mathbf{1}^{\intercal} \tanh\left(c_1 \mathbf{D}_z \boldsymbol{\beta}^{\text{c}}\right) \tanh\left(c_2 \boldsymbol{\beta}^{\text{c}}\right) \mathbb{1}\left\{\boldsymbol{\beta}^{\text{c}} > 0\right\}. \tag{11}$$

The constants $c_1, c_2 \geq 0$ control the slope of the hyperbolic-tangent functions.

In addition, we let the field $\boldsymbol{\beta}^{\text{c}}$ be spatially smooth. This is expressed by penalizing the energy of the Laplacian-filtered field $\boldsymbol{\beta}^{\text{c}}$. The 3D Laplacian operator is represented by matrix $\mathbf{A}$. Overall, the regularization term is

$$R(\boldsymbol{\beta}^{\text{c}}) = \frac{1}{2}\mu_1 \|\mathbf{A}\boldsymbol{\beta}^{\text{c}}\|_2^2 + \mu_2 R_{\text{mon}}(\boldsymbol{\beta}^{\text{c}}) \quad , \tag{12}$$

where $\mu_1, \mu_2 \geq 0$ control the weight of the regularization components. For the gradient of the regularization term, see the supplementary material.

Example of simple use of the monotonicity prior. We illustrate here the effect of the prior on a 1D denoising problem, unrelated to scattering tomography. The original signal $\beta^{\text{true}}(z) \in [0, 70]$, is a representative vertical profile of a cloud, containing 27 samples. The signal $\beta^{\text{true}}(z)$ monotonically increases within part of the domain. The monotonic increase ends with an abrupt fall. The signal has null value in other parts of the domain, as seen in Fig. 3. The noisy input is $y_v = \beta_v^{\text{true}} + n_v$ where $n_v \sim \mathcal{N}(0,5)$ is white and $v \in [1, \ldots, 27]$. The denoised estimate is $\hat{\boldsymbol{\beta}} = \arg\min[\|\boldsymbol{\beta} - \boldsymbol{y}\|_2^2 + 10 R_{\text{mon}}(\boldsymbol{\beta})]$. The result, plotted in Fig. 3, shows that the estimate is less noisy than the data, not blurred, while preserving the main peak. To evaluate the recovery quality we use the following local reconstruction

error measure [2]

$$\epsilon = \frac{\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{\text{true}}\|_1}{\|\boldsymbol{\beta}^{\text{true}}\|_1} \quad . \tag{13}$$

The mean error over $1,000$ realizations of $\boldsymbol{y}$ is $\epsilon = 16\%$, down from 18% without the prior. For a higher noise level, $n_v \sim \mathcal{N}(0, 10)$, $\epsilon = 32\%$ down from 37%.

5 The Forward Model

The image formation model $\boldsymbol{i}(\boldsymbol{\beta})$ is known in the literature [9, 13–15, 19, 38, 48, 49, 52]. We provide it to make the paper self-contained and provide the background for deriving the data-fitting gradient $\partial F(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$. Terms relating to volumetric scattering are often used in dehazing or defogging computer vision literature [5, 22, 28, 42–46, 53, 55, 56].

In an infinitesimal volume, the interaction can be scattering or absorption. Relative to the incoming radiance, the energy scattered in an infinitesimal volume is set by the coefficient $\varpi\beta(\boldsymbol{x})$. Here $\varpi \in [0, 1]$, the *single-scattering albedo*, is the ratio of the scattering coefficient to the total extinction coefficient $\beta(\boldsymbol{x})$. The single-scattering albedo of air molecules and cloud droplets are denote by ϖ^{a}, and ϖ^{c} respectively.

Denote direction by $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathbb{S}^2$, where $\mathbb{S}^2$ is the unit sphere. Scattered light is distributed angularly relative to the radiation incident on point $\boldsymbol{x}$, from direction $\boldsymbol{\omega}$. When scattering particles are oriented randomly and uniformly in a voxel, the probability density for scattering to direction $\boldsymbol{\omega}$ is expressed by a *phase function* $f(\theta)$, where $\theta = \arccos(\boldsymbol{\omega} \cdot \boldsymbol{\omega}')$ is the scattering angle. The phase functions of air molecules and cloud droplets are denoted by $f^{\text{a}}(\theta)$, and $f^{\text{c}}(\theta)$ respectively.

To describe the forward model by MC RT, we draw some definitions and notations from [14, 19]. MC RT aggregates random photon paths. Denote by L_d the set of all possible photon paths via $\mathcal{M}$, connecting a source to a detector d. The source in our case-study is the sun, located at a distant location $\boldsymbol{x}_\odot$. A specific photon path, denoted by $\mathcal{L} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_B)$ is a sequence of B interaction points. One way for photons to contribute a signal to a detector d is for a photon to perform a random walk (via $\mathcal{L}$) from the source to a terminal location $\boldsymbol{x}_B$ at a camera aperture. There, the photon deposits a terminal intensity $I_B(\mathcal{L})$ at a pixel d. A clear signal at detector d is the expectation of a photon to reach d, multiplied by the number of photons (N_{phot}) generated at the source,

$$i_d(\boldsymbol{\beta}) = N_{\text{phot}} \int_{L_d} P(\mathcal{L}) I_B(\mathcal{L}) d\mathcal{L}. \tag{14}$$

Here $P(\mathcal{L})$ is the probability density of a general path $\mathcal{L}$. However, a random walk from the source to a camera pixel is a relatively rare event. For efficient rendering, we sample light paths in the medium using Backward MC.

The Backward MC process, illustrated in Fig. 4, back-propagates mathematically photons from a detector to the source. Denote the direction to the

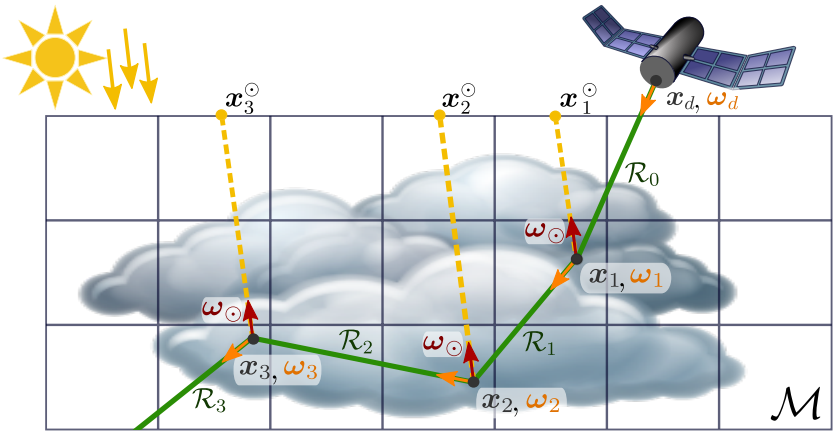

Fig. 4. Backward MC simulated imaging. This illustration demonstrates a photon path from a detector through the medium. The contribution to i_d is the sum of signals along the yellow lines towards the sun (local estimation).

source (sun) by $\boldsymbol{\omega}_{\odot}$. From each camera pixel d, photons are back-projected in a corresponding direction $\boldsymbol{\omega}_d$ and then traced through the medium, as follows:
(i) Initially, a photon has associated intensity I_0. The photon is back projected along a ray denoted $\mathcal{R}_0$, which has a direction $\boldsymbol{\omega}_d$ and starts at detector (pixel) d at location $\boldsymbol{x}_d$, meaning, $\boldsymbol{x}_0 = \boldsymbol{x}_d$.
Per iteration b:
(ii) On ray $\mathcal{R}_b$, a random location $\boldsymbol{x}_{b+1}$ is sampled. The photon propagates to $\boldsymbol{x}_{b+1}$. This random distance, along $\mathcal{R}_b$, is sampled as described in [19]. It is a function of the spatially varying extinction coefficient $\beta(\boldsymbol{x})$.
(iii) If $\boldsymbol{x}_b \notin \mathcal{M}$, the photon is terminated. If $\boldsymbol{x}_b \notin \mathcal{M}$ while $\mathcal{R}_b || \boldsymbol{\omega}_{\odot}$, the photon contributes intensity to pixel d.
(iv) If $\boldsymbol{x}_b \in \mathcal{M}$, then the photon interacts there with a particle. The type of particle (air molecule or cloud droplet) is sampled randomly based on the ratio of extinction coefficients $\beta^{\mathrm{a}}(\boldsymbol{x}_b)$, $\beta^{\mathrm{c}}(\boldsymbol{x}_b)$. If the scattering event at $\boldsymbol{x}_b$ is due to air, the single scattering albedo is $\varpi_b = \varpi^{\mathrm{a}}$. If the scattering event at $\boldsymbol{x}_b$ is due to a cloud droplet, the single scattering albedo is $\varpi_b = \varpi^{\mathrm{c}}$. The intensity carried by the photon is attenuated to $I_b = \varpi_b I_{b-1}$, hence

$$I_b = I_0 \prod_{b'=1}^{b} \varpi_{b'} \ . \tag{15}$$

If I_b is lower than a threshold, the photon can be stochastically terminated [20].
(v) The extinction coefficient at $\boldsymbol{x}_b$ is determined by the type of particle that caused the scattering. Meaning, $\beta(\boldsymbol{x}_b) = \beta^{\mathrm{a}}(\boldsymbol{x}_b)$ if the particle is an air molecule, and $\beta(\boldsymbol{x}_b) = \beta^{\mathrm{c}}(\boldsymbol{x}_b)$ if the particle is a cloud droplet. Denote $\boldsymbol{\omega}_b$ as the unit vector between $\boldsymbol{x}_b$ and $\boldsymbol{x}_{b+1}$. According to the phase function [6, 12, 19], the photon scatters to a new random direction $\boldsymbol{\omega}_{b+1}$. The scattering angle between

$\boldsymbol{\omega}_{b-1}$ and $\boldsymbol{\omega}_b$ is $\theta_{b-1,b} = \arccos(\boldsymbol{\omega}_{b-1} \cdot \boldsymbol{\omega}_b)$. If the photon is scattered by air, the phase function is $f(\theta_{b-1,b}) = f^{\mathrm{a}}(\theta_{b-1,b})$. If scattering is by a cloud droplet, the phase function is $f(\theta_{b-1,b}) = f^{\mathrm{c}}(\theta_{b-1,b})$. Scattering yields a new ray, $\mathcal{R}_{b+1}$ from point $\boldsymbol{x}_{b+1}$ in direction $\boldsymbol{\omega}_{b+1}$, thus a new iteration of propagation (`ii`).

5.1 Efficient Rendering

MC can efficiently estimate i_d using *local estimation* [35]. Index a photon by $p \in [1, \ldots, N_{\mathrm{phot}}]$. Its random path is $\mathcal{L}(d,p)$. Local estimation expresses the probability that a photon scatters from point $\boldsymbol{x}_b$ towards the light source (in our case, the sun at distant location $\boldsymbol{x}_\odot$), without interacting again.

The line segment between $\boldsymbol{x}_{b-1}$ and $\boldsymbol{x}_b$ is denoted $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}$, and contains all points satisfying $\{\boldsymbol{x} = \tilde{\alpha}\boldsymbol{x}_{b-1} + (1-\tilde{\alpha})\boldsymbol{x}_b,\ \forall \tilde{\alpha} \in [0,1]\}$. The transmittance of the medium on this line segment is

$$a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right) = \exp\left\{-\int_{\boldsymbol{x}_{b-1}}^{\boldsymbol{x}_b} \left[\beta^{\mathrm{a}}(\boldsymbol{x}) + \beta^{\mathrm{c}}(\boldsymbol{x})\right] d\boldsymbol{x}\right\} . \tag{16}$$

As illustrated in Fig. 4, the line between $\boldsymbol{x}_b$ and $\boldsymbol{x}_\odot$ intersects with the top of the atmosphere (TOA). This intersection is denoted by $\boldsymbol{x}_b^\odot$. Analogously, define the line segment $\overline{\boldsymbol{x}_b\boldsymbol{x}_b^\odot}$. Along this segment, the transmittance is $a\left(\overline{\boldsymbol{x}_b\boldsymbol{x}_b^\odot}\right)$. The scattering angle to the sun is denoted by $\theta_{b,\odot}$. Using Eq. (15,16), the local estimation contribution of photon p at $\boldsymbol{x}_b$ is

$$I_b^{\mathrm{le}}[\mathcal{L}(d,p)] = I_b f(\theta_{b,\odot})\, a\left(\overline{\boldsymbol{x}_b\boldsymbol{x}_b^\odot}\right) . \tag{17}$$

The MC estimate of $i_d(\boldsymbol{\beta})$ is then

$$i_d(\boldsymbol{\beta}) \approx \sum_{p=1}^{N_{\mathrm{phot}}} \sum_{b=1}^{B} I_b^{\mathrm{le}}\left[\mathcal{L}(d,p)\right] . \tag{18}$$

5.2 Gradient of the Model-Fit

The problem (4,5) is solved iteratively using SGD. This requires, specifically, estimation of the gradient of $F(\boldsymbol{\beta})$. A component of the gradient in voxel v is given by

$$\frac{\partial F(\boldsymbol{\beta})}{\partial \beta_v^{\mathrm{c}}} = \sum_{d=1}^{N_{\mathrm{detectors}}} \left[i_d(\boldsymbol{\beta}) - i_d^{\mathrm{measured}}\right] \frac{\partial i_d(\boldsymbol{\beta})}{\partial \beta_v^{\mathrm{c}}} . \tag{19}$$

We now assess the Jacobian $\partial i_d(\boldsymbol{\beta})/\partial \beta_v^{\mathrm{c}}$, i.e., how image pixels change in response to an infinitesimal perturbation of the medium $\boldsymbol{\beta}$. We describe here the essence of the Jacobian. For more details and mathematical derivations, see the supplementary material.

The domain of voxel v is denoted by $\mathcal{V}_v$. The line segment $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}$ traverses several voxels. Denote by $l_{b-1,v}$ the length of the intersection of $\mathcal{V}_v$ with the line

segment. Similarly, denote by $l_v^{\odot}$ the length of the intersection of $\mathcal{V}_v$ with the line segment and $\overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}}$. Then,

$$\frac{\partial i_d(\boldsymbol{x})}{\partial \beta_v^{\mathrm{c}}} \approx -\sum_{p=1}^{N_{\mathrm{phot}}} \sum_{b=1}^{B} I_b^{\mathrm{le}} \left[\mathcal{L}(d,p)\right] \cdot \left(l_v^{\odot} + l_{b-1,v} - \begin{cases} \frac{1}{\beta_v^{\mathrm{c}}} & \boldsymbol{x}_b \in \mathcal{V}_v, \ \& \text{ scatter} \\ & \text{is by a cloud droplet} \\ 0 & \text{else} \end{cases} \right). \tag{20}$$

Based on Eqs. (19,20) we solve Eq. (5) using SGD.

6 Coarse to Fine

We take several measures to enhance the estimation speed. We use a coarse-to-fine approach, spatially and in RT quality. RT quality is coarsened by using a small number of MC photons when estimating the gradient. The estimate is then noisy (coarse), per iteration. However, this way, SGD iterations are fast, and significantly advance the minimization (5) in its initial stages [51]. For more details and figures regarding specific simulations and performances, see the supplementary material. After a while, however, SGD in the coarse quality yields diminishing returns. Therefore, we use the result based on the coarse quality as initialization for a later stage, where the number of photons increases, leading to finer quality, better accuracy yet slower iterations.

Spatial coarse-to-fine is similar to common analysis in image analysis. Here, $\mathcal{M}$ is first represented in a coarse voxel grid. Thus the vector of unknown extinctions coefficients has a small dimension. Furthermore, the modeled and measured images are represented in a coarse resolution. The use of a small number of pixels and voxels enables significant shortening of runtime while significantly advancing the minimization (5) in its initial stages. We use the result obtained in the coarse voxels as initialization for the next stage: there, both $\mathcal{M}$ and the images are represented by smaller elements, leading to finer spatial resolution, yet slower iterations. This process is repeated in additional stages, where a result obtained in a coarser resolution initializes iterations on a finer-resolution grid. The number of voxels and pixels used in each stage of our examples are specified in the supplementary material.

We used additional speedup measures. Using the data images, we perform space-carving [29] of $\mathcal{M}$. This creates a photo-hull within which clouds potentially reside, pruning many voxels from the optimization. To correspond to the voxel resolution at the intermediate resolution steps described above, the photo-hull map is coarsened by dilation, before subsampling.

Furthermore, Eq. (19) requires two MC simulations: one is for estimating $i_d(\boldsymbol{\beta})$ (forward model); the other is for estimating its Jacobian $\partial i_d(\boldsymbol{\beta})/\partial \boldsymbol{\beta}$. Per iteration we use the same paths to sample both fields, saving about half the runtime. We found that possible bias of the gradient is not a major concern to the inverse problem, as each iteration samples a new random set.

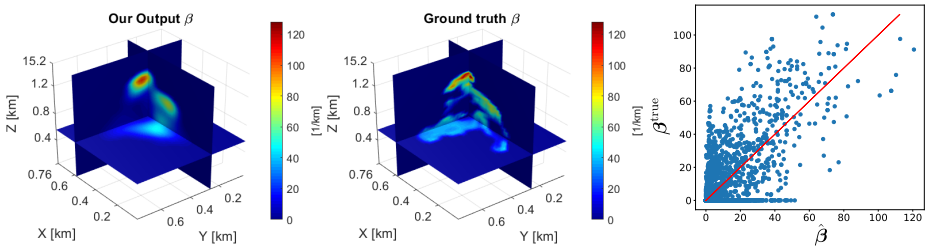

Fig. 5. 3D slices of the medium from scene `C`. [Left] Our recovered medium. [Middle] The true medium. [Right] scatter plot of 25% of the data points, randomly selected. The red line represents ideal reconstruction, where $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{\text{true}}$. Some areas are not well estimated. This is because they are occluded from the viewpoints of the cameras, by higher and denser parts of the cloud.

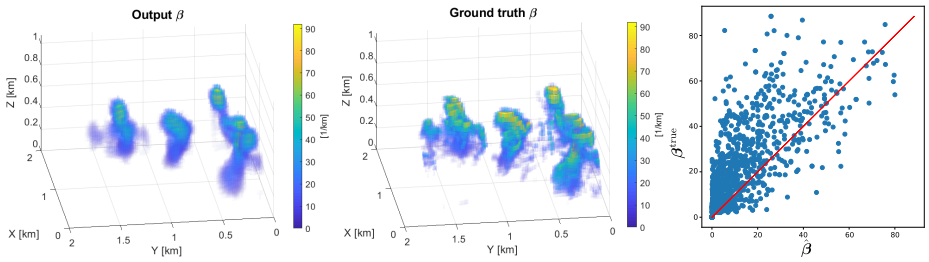

Fig. 6. Scene `F`. [Left] Our recovered medium. [Middle] The true medium. [Right] Scatter plot of 10% of the data points, randomly selected. The red line represents ideal reconstruction, where $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{\text{true}}$.

7 Simulations

Currently, there is no adequate gear for simultaneous multi-view imaging of small clouds from space. However, proper orbiting cameras and data are upcoming in a few years [54]. This paper develops methods in anticipation for the upcoming spaceborne technological advancement. At such a stage, testing is by simulations, as in several recent computer vision papers [15, 19, 31, 32] about scattering-CT. As in [31, 32] a realistically complex scene is created by use a *Large Eddy Simulation (LES)* [7, 37]. This is a tool used by atmospheric scientists as a standard for computationally generating cloud fields from first principles. We deal with two public domain scenes [37], with results shown in Figs. 5,6:

`C` has a single cloud in a 0.76km×0.76km×1.52km domain having 38×38×38 voxels (54, 872 unknowns). Cameras have 76 × 76 pixels, each having 30 m ground-level footprint (nadir view). In rendering of ground-truth data, $N_{\text{phot}} = 8192$ photons. Cameras are at $z = 1.9$km. One camera is 1.8° off the zenith. The other eight reside on a ring $\approx 29°$ around the zenith.

`F` has several clouds in a 1.72km × 1.72km × 1.08km domain having 86 × 86 × 27 voxels (199, 692 unknowns). Cameras have 86 × 86 pixels, each having 20 m

ground-level footprint, $N_{\text{phot}} = 4096$ photons. One camera is 0.4° off the zenith at $z = 2.16$km. The other eight are at $z = 1.86$km on a ring $\approx 33°$ around the zenith.
In both `C` and `F`, the sun is at the zenith, voxel are $20 \times 20 \times 40\,\text{m}^3$ large, and the background surface is Lambertian having albedo 0.05, emulating a dark ocean. Each domain has background $\beta = 0.01$ km^{-1} wherever $\beta^{\text{c}} = 0$. The domain is observed by nine perspective cameras. To reduce the effect of air molecules (which are not our object of interest), sensing focused on a red-colored wavelength 632 nm.

The values of N_{phot} are comparable to the full-well of common machine vision cameras. Due to the randomness of MC rendering, this photon count yields typical Poissonian noise resembling photon noise in cameras. The MC kernel we used is by Mitsuba [21].

After estimation of the photo-hull (Sec. 6), tomographic estimation was initialized by $\beta(\boldsymbol{x}) = 2\,\text{km}^{-1}$. We use the adaptive moment estimation method (Adam) [26], with the following parameters: first moment decay $\eta_1 = 0.9$, second moment decay $\eta_2 = 0.999$, $\varepsilon = 10^{-8}$, and step size is $\alpha = 0.3$ for all scenes.

Reconstruction quality is assessed by two measures [2] which compare the estimated extinction field to the ground-truth. One of them is ϵ from Eq. (13). The other is

$$\delta = \frac{\|\hat{\boldsymbol{\beta}}\|_1 - \|\boldsymbol{\beta}^{\text{true}}\|_1}{\|\boldsymbol{\beta}^{\text{true}}\|_1}, \tag{21}$$

which is a relative bias of the overall extinction.

Results

In the small scene `C`, the results are displayed in Fig. 5. Without using the monotonicity prior, meaning only using a smoothness prior, the results were $\delta = -11.5\%$ and $\epsilon = 87\%$ for a runtime of about 6.5 hours to convergence. Our approach, with the monotonicity prior, yielded $\delta = 5.5\%$ and $\epsilon = 84.5\%$ for the same runtime. Results of the cloud field (scene `F`) are displayed in Fig. 6. Without a monotonicity prior, $\delta = -34.2\%$ and $\epsilon = 54.8\%$, after 127 hours. Afterwards, the error measures increased. The monotonicity prior improved results to $\delta = -28.2\%$ and $\epsilon = 44.7\%$. Results reached convergence after 134 hours.

Comparing our results to previous work, running a code based on [15] for scene `C`, after a fixed 6.5 hours runtime, yielded $\epsilon = 196\%$ and $\delta = 14\%$, and code based on [19] yielded $\epsilon = 174\%$ and $\delta = -9.6\%$, after the same runtime. To compare the SHDOM-based tomography approach [31] (and [32], which uses the same algorithm yet solves for micro-physics), we used an improved code [30] and solved for scene `C`. The deterministic approach resulted in more accurate output, having $\delta = -3.6\%$ and $\epsilon = 72\%$. Moreover, the SHDOM-based tomography code [30] ran for about an hour while our method ran for about 6.5 hours. Recalling Sec. 3, this outcome is not surprising, as the deterministic SHDOM is inherently faster and more accurate than random walks of MC. Yet, this approach has a fundamental challenge to scale to large fields, contrary to MC.

In accordance, for larger scenes, containing a 100s of thousands of voxels, like scene `F`, [31] reported $\epsilon = 70\%$ and $\delta = 30\%$, compared to our results in scene `F`.

Overall, the most significant cost reduction is achieved within minutes, using the coarse scales. The finest scales are much slower, yet enable a significant improvement of the results. See the supplementary for more details.

8 Conclusions

Spaceborne data of simultaneous imaging of cloud fields, in resolution that is required for scattering-CT is not yet accessible by an experiment. However, a future spaceborne mission [54] is planned to provide such data. Hence, priors and algorithms developed for wide field scattering-CT should become useful in several years, enabling scientists to draw new information and understanding of nature. Moreover, advancement of scattering CT should be useful also to some modalities of bio-medical imaging and material engineering.

To enhance speed, good initialization can be obtained in the future by a neural-network. Current differential neural renderers (DNRs) are set for opaque, reflective objects [24, 47]. A DNR based on 3D, volumetric RT in large heterogeneous scattering media requires large training sets that are currently unavailable. Once suitable scattering training data and DNRs become available, it may efficiently initialize our process, which is proper physics-based MC RT analysis. The latter is needed as the results should be trusted by physicists studying nature.

We believe that the corase-to-fine stochastic approach taken in this paper has significant potential. MC can excellently use technologies of parallel computations. The calculations per photon are mutually independent of other photons. Since most of the calculation in MC rendering are similar, it is worthwhile to adapt such a method to GPU, to accelerate the runtime. Encouragingly, the most recent version of Mitsuba [48], states an implementation on GPU. For these reasons, the developments of scattering tomography are timely.

Acknowledgments

We thank Ilan Koren, Eshkol Eytan Liebeskind, and Tom Dror-Schwartz for useful discussions. We thank Johanan Erez, Ina Talmon and Daniel Yagodin for technical support. Yoav Schechner is the Mark and Diane Seiden Chair in Science at the Technion. He is a Landau Fellow - supported by the Taub Foundation. His work was conducted in the Ollendorff Minerva Center. Minvera is funded through the BMBF. This research is funded by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 810370: CloudCT). Aviad Levis is a Zuckerman Postdoctoral Fellow.

References

1. Aides, A., Levis, A., Holodovsky, V., Schechner, Y.Y., Althausen, D., Vainiger, A.: Distributed sky imaging radiometry and tomography. In: ICCP. pp. 1–12. IEEE (2020)
2. Aides, A., Schechner, Y.Y., Holodovsky, V., Garay, M.J., Davis, A.B.: Multi sky-view 3D aerosol distribution recovery. Opt. Express **21**(22), 25820–25833 (2013)
3. Alterman, M., Schechner, Y.Y., Vo, M., Narasimhan, S.G.: Passive tomography of turbulence strength. In: ECCV. pp. 47–60. Springer (2014)
4. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. IEEE Trans. Pattern Anal. Mach. Intell (2), 218–233 (2003)
5. Berman, D., Treibitz, T., Avidan, S.: Air-light estimation using haze-lines. In: ICCP. pp. 1–9. IEEE (2017)
6. Binzoni, T., Leung, T.S., Gandjbakhche, A.H., Ruefenacht, D., Delpy, D.: The use of the Henyey–Greenstein phase function in Monte Carlo simulations in biomedical optics. Phys. Med. Biol **51**(17), N313 (2006)
7. Chung, D., Matheou, G.: Large-eddy simulation of stratified turbulence. part i: A vortex-based subgrid-scale model. J. Atmos. Sci. **71**(5), 1863–1879 (2014)
8. Davis, C., Emde, C., Harwood, R.: A 3-d polarized reversed monte carlo radiative transfer model for millimeter and submillimeter passive remote sensing in cloudy atmospheres. IEEE Trans. Geosci. Remote. Sens. **43**(5), 1096–1101 (2005)
9. Deng, X., Jiao, S., Bitterli, B., Jarosz, W.: Photon surfaces for robust, unbiased volumetric density estimation. ACM Trans. Graph. **38**(4), 46 (2019)
10. Diner, D.J., Martonchik, J.V.: Atmospheric transmittance from spacecraft using multiple view angle imagery. Appl. Opt. **24**(21), 3503–3511 (1985)
11. Evans, K.F.: The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. J. Atmos. Sci. **55**(3), 429–446 (1998)
12. Frisvad, J.R.: Importance sampling the rayleigh phase function. JOSA **28**(12), 2436–2441 (2011)
13. Georgiev, I., Misso, Z., Hachisuka, T., Nowrouzezahrai, D., Křivánek, J., Jarosz, W.: Integral formulations of volumetric transmittance. ACM Trans. Graph. **38**(6), 1–17 (2019)
14. Geva, A., Schechner, Y.Y., Chernyak, Y., Gupta, R.: X-ray computed tomography through scatter. In: ECCV. pp. 34–50. Springer (2018)
15. Gkioulekas, I., Levin, A., Zickler, T.: An evaluation of computational imaging techniques for heterogeneous inverse scattering. In: ECCV. pp. 685–701. Springer (2016)
16. Gkioulekas, I., Zhao, S., Bala, K., Zickler, T., Levin, A.: Inverse volume rendering with material dictionaries. ACM Trans. Graph. **32**(6), 162 (2013)
17. Gorbunov, M.E., Sokolovky, S., Bengtsson, L.: Space refractive tomography of the atmosphere: Modeling of direct and inverse problems (1996)
18. Gregson, J., Krimerman, M., Hullin, M.B., Heidrich, W.: Stochastic tomography and its applications in 3d imaging of mixing fluids. ACM Trans. Graph. **31**(4), 52–1 (2012)
19. Holodovsky, V., Schechner, Y.Y., Levin, A., Levis, A., Aides, A.: In-situ multi-view multi-scattering stochastic tomography. In: ICCP. pp. 1–12. IEEE (2016)
20. Iwabuchi, H.: Efficient monte carlo methods for radiative transfer modeling. J. Atmos. Sci. **63**(9), 2324–2339 (2006)
21. Jakob, W.: Mitsuba renderer (2010), http://www.mitsuba-renderer.org

22. Kaftory, R., Schechner, Y.Y., Zeevi, Y.Y.: Variational distance-dependent image restoration. In: CVPR. pp. 1–8. IEEE (2007)
23. Kalashnikova, O.V., Garay, M.J., Davis, A.B., Diner, D.J., Martonchik, J.V.: Sensitivity of multi-angle photo-polarimetry to vertical layering and mixing of absorbing aerosols: Quantifying measurement uncertainties. J. Quant. Spectrosc. Radiat. Transf. **112**(13), 2149–2163 (2011)
24. Kato, H., Ushiku, Y., Harada, T.: Neural 3d mesh renderer. In: CVPR. pp. 3907–3916. IEEE (2018)
25. Khungurn, P., Schroeder, D., Zhao, S., Bala, K., Marschner, S.: Matching real fabrics with micro-appearance models. ACM Trans. Graph. **35**(1), 1–1 (2015)
26. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
27. Kokhanovsky, A.A.: Light scattering media optics. Springer (2004)
28. Kratz, L., Nishino, K.: Factorizing scene albedo and depth from a single foggy image. In: ICCV. pp. 1701–1708. IEEE (2009)
29. Kutulakos, K.N., Seitz, S.M.: A theory of shape by space carving. Int. J. Comput. Vis. **38**(3), 199–218 (2000)
30. Levis, A., Aides, A.: pyshdom. https://github.com/aviadlevis/pyshdom (2019)
31. Levis, A., Schechner, Y.Y., Aides, A., Davis, A.B.: Airborne three-dimensional cloud tomography. In: ICCV. pp. 3379–3387. IEEE (2015)
32. Levis, A., Schechner, Y.Y., Davis, A.B.: Multiple-scattering microphysics tomography. In: CVPR. pp. 6740–6749. IEEE (2017)
33. Levis, A., Schechner, Y.Y., Davis, A.B., Loveridge, J.: Multi-view polarimetric scattering cloud tomography and retrieval of droplet size. arXiv preprint arXiv:2005.11423 (2020)
34. Lu, M.L., Feingold, G., Jonsson, H.H., Chuang, P.Y., Gates, H., Flagan, R.C., Seinfeld, J.H.: Aerosol-cloud relationships in continental shallow cumulus. JGR: Atmospheres **113**(D15) (2008)
35. Marshak, A., Davis, A.: 3D radiative transfer in cloudy atmospheres. Springer (2005)
36. Martonchik, J.V., Diner, D.J., Kahn, R.A., Ackerman, T.P., Verstraete, M.M., Pinty, B., Gordon, H.R.: Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. IEEE Trans. Geosci. Remote. Sens. **36**(4), 1212–1227 (1998)
37. Matheou, G., Chung, D.: Large-eddy simulation of stratified turbulence. part ii: Application of the stretched-vortex model to the atmospheric boundary layer. J. Atmos. Sci. **71**(12), 4439–4460 (2014)
38. Mayer, B.: Radiative transfer in the cloudy atmosphere. In: EPJ Web of Conferences. vol. 1, pp. 75–99. EDP Sciences (2009)
39. McFarlane, S.A., Grabowski, W.W.: Optical properties of shallow tropical cumuli derived from arm ground-based remote sensing. Geophys. Res. Lett. **34**(6) (2007)
40. Messer, H., Zinevich, A., Alpert, P.: Environmental sensor networks using existing wireless communication systems for rainfall and wind velocity measurements. Instrum. Meas. Mag.e **15**(2), 32–38 (2012)
41. Mobley, C.D.: Light and water: radiative transfer in natural waters. Academic press (1994)
42. Narasimhan, S., Ramamoorthi, R., Nayar, S.: Analytic rendering of multiple scattering in participating media. ACM Trans. Graph. (2004)
43. Narasimhan, S.G., Gupta, M., Donner, C., Ramamoorthi, R., Nayar, S.K., Wann-Jensen, H.: Acquiring scattering properties of participating media by dilution. ACM Trans. Graph. **25**(3), 1003–1012 (2006)

44. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. Int. J. Comput. Vis. **48**(3), 233–254 (2002)
45. Narasimhan, S.G., Nayar, S.K.: Interactive (de) weathering of an image using physical models. In: IEEE Workshop on color and photometric Methods in computer Vision. vol. 6, p. 1. France (2003)
46. Narasimhan, S.G., Nayar, S.K., Sun, B., Koppal, S.J.: Structured light in scattering media. In: ICCV. vol. 1, pp. 420–427. IEEE (2005)
47. Nguyen-Phuoc, T.H., Li, C., Balaban, S., Yang, Y.: Rendernet: A deep convolutional network for differentiable rendering from 3d shapes. In: NeurIPS. pp. 7891–7901 (2018)
48. Nimier-David, M., Vicini, D., Zeltner, T., Jakob, W.: Mitsuba 2: A retargetable forward and inverse renderer. ACM Trans. Graph. **38**(6) (Nov 2019). https://doi.org/10.1145/3355089.3356498
49. Novák, J., Georgiev, I., Hanika, J., Jarosz, W.: Monte carlo methods for volumetric light transport simulation. In: Computer Graphics Forum. vol. 37, pp. 551–576. Wiley Online Library (2018)
50. Okabe, T., Sato, I., Sato, Y.: Spherical harmonics vs. haar wavelets: Basis for recovering illumination from cast shadows. In: CVPR. vol. 1, pp. 50–57. IEEE (2004)
51. Pfeiffer, G.T., Sato, Y.: On stochastic optimization methods for monte carlo least-squares problems. arXiv preprint arXiv:1804.10079 (2018)
52. Pharr, M., Jakob, W., Humphreys, G.: Physically based rendering: From theory to implementation. Morgan Kaufmann (2016)
53. Schaul, L., Fredembach, C., Süsstrunk, S.: Color image dehazing using the near-infrared. In: ICIP. pp. 1629–1632. IEEE (2009)
54. Schilling, K., Schechner, Y.Y., Koren, I.: Cloudct – computed tomography of clouds by a small satellite formation. IAA symposium on Small Satellites for Earth Observation (2019)
55. Sheinin, M., Schechner, Y.Y.: The next best underwater view. In: CVPR. pp. 3764–3773. IEEE (2016)
56. Spier, O., Treibitz, T., Gilboa, G.: In situ target-less calibration of turbid media. In: ICCP. pp. 1–9. IEEE (2017)
57. Vainiger, A., Schechner, Y.Y., Treibitz, T., Avni, A., Timor, D.S.: Optical wide-field tomography of sediment resuspension. Opt. Express **27**(12), A766–A778 (2019)
58. Villefranque, N., Fournier, R., Couvreux, F., Blanco, S., Cornet, C., Eymet, V., Forest, V., Tregan, J.M.: A path-tracing monte carlo library for 3-d radiative transfer in highly resolved cloudy atmospheres. J. Adv. Model. Earth Syst (2019)
59. Wright, T., Burton, M., Pyle, D., Caltabiano, T.: Scanning tomography of so_2 distribution in a volcanic gas plume. Geophys. Res. Lett. **35**(17) (2008)
60. Zhang, S., Xue, H., Feingold, G.: Vertical profiles of droplet effective radius in shallow convective clouds. Atmospheric Chem. Phys **11**(10), 4633–4644 (2011)

Monotonicity Prior for Cloud Tomography
Supplementary Material

Tamar Loeub[1], Aviad Levis[2], Vadim Holodovsky[1], and Yoav Y. Schechner[1]

[1] Viterbi Faculty of Electrical Engineering,
Technion - Israel Institute of Technology, Haifa, Israel
[2] Computing and Mathematical Sciences Department,
California Institute of Technology, Pasadena, CA, USA
tamarloeub@gmail.com, aviad.levis@gmail.com,
vholod@ef.technion.ac.il, yoav@ee.technion.ac.il

Abstract. This is a supplementary document to the main manuscript. Here we provide the mathematical derivation of the gradient of the image formation model and additional details about run times.

1 Outline

This supplementary material contains two parts. The first part (Secs. 2 and 3) details the mathematical derivation of the gradient of the image fitting term, and the prior. The second part, Sec. 4, provides additional information regarding the simulations, i.e., run time, figures, and framework, and configuration details. Moreover, we present an additional test and application: rendering a new viewpoint from the reconstructed medium.

2 Gradient of the Image Fitting Term

We now describe the calculation of the gradient of the image model. Recall Eq. (14) from the main manuscript, which is the analytical formulation of a clear signal at detector d. It is the expectation of a photon to reach d, multiplied by the number of photons (N_{phot}) generated at the source,

$$i_d(\boldsymbol{\beta}) = N_{\text{phot}} \int_{L_d} P(\mathcal{L}) I_B(\mathcal{L}) d\mathcal{L}. \tag{1}$$

Here $P(\mathcal{L})$ is the probability density of a general path $\mathcal{L} \in L_d$. Index a photon by p, and an interaction point by b. The Monte-Carlo (MC) estimation of the signal at detector d is

$$i_d(\boldsymbol{\beta}) \approx \sum_{p=1}^{N_{\text{phot}}} \sum_{b=1}^{B} I_b^{\text{le}} \left[\mathcal{L}(d,p)\right] \ , \tag{2}$$

where $I_b^{\text{le}} \left[\mathcal{L}(d,p)\right]$ is given in Eq. (17) of the main manuscript.

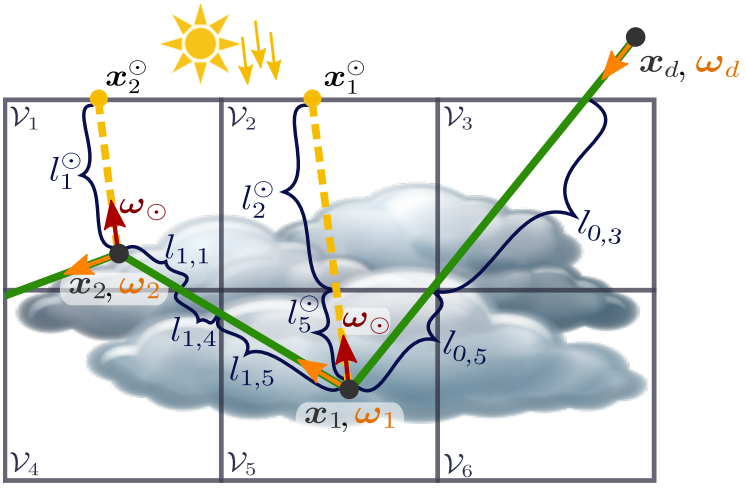

Fig. 1: Illustration of a photon path from a detector through the medium. The gradient is calculated according to the process described in Sec. 2 of this supplementary document, and specifically, Eq. (19) herein.

Note that while $P(\mathcal{L})$ is not needed explicitly for rendering, it is useful to associate a probability to a given path. The usefulness is clear when approaching the inverse problem. Following the definition of the transmittance function $a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right)$, in Eq. (16) of the main manuscript, associate with a scattering event at $\boldsymbol{x}_b$ a probability $a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right) f(\theta_{b-1,b})$. Here, $f(\theta_{b-1,b})$ is the phase function of the scattering angle $\theta_{b-1,b}$. Each scattering event and each extinction sampling is independent. Hence, inspired by [1], the probability P associated with a general path $\mathcal{L}$ is

$$P(\mathcal{L}) = \prod_{b=1}^{B-1} a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right) \beta_b f(\theta_{b-1,b}) \ . \tag{3}$$

Note the different forms of Eq. (1) and Eq. (2) above:

- The probability $P(\mathcal{L})$ in the integral is eliminated during summation, because path probability dictates the MC sampling in the sum.
- Point $\boldsymbol{x}_B$ is not expected to reach the sun, when sampling a small number of photons. Hence the signal in the summation relies on local estimation from all nodes, in lieu of $I(B)$.

We summarize the approximation of MC radiative transfer (RT) using the following operations:

$$I_B \underset{\text{becomes}}{\Longrightarrow} \sum_{b=1}^{B} I_b^{\text{le}}\left[\mathcal{L}(d,p)\right] \quad ; \quad N_{\text{phot}} \int_{L_d} P(\mathcal{L})(\cdot) d\mathcal{L} \underset{\text{becomes}}{\Longrightarrow} \sum_{p=1}^{N_{\text{phot}}} (\cdot) \ . \tag{4}$$

The optimization problem described in the paper is solved iteratively using stochastic gradient descent (SGD). This requires, specifically, estimation of the

gradient of $F(\boldsymbol{\beta})$. A component of the gradient corresponding to voxel v is given by

$$\frac{\partial F(\boldsymbol{\beta})}{\partial \beta_v^{\mathrm{c}}} = \sum_{d=1}^{N_{\mathrm{detectors}}} \left[i_d(\boldsymbol{\beta}) - i_d^{\mathrm{measured}} \right] \frac{\partial i_d(\boldsymbol{\beta})}{\partial \beta_v^{\mathrm{c}}} \ . \tag{5}$$

We now assess $\partial i_d(\boldsymbol{\beta})/\partial \beta_v^{\mathrm{c}}$, i.e., how image pixels change in response to an infinitesimal perturbation of the medium $\boldsymbol{\beta}$. Changes in the medium generally lead to new paths being sampled as described in the paper. However, it is computationally demanding to sample new paths per degree of freedom of $\boldsymbol{\beta}$. To derive the gradient using existing paths, we use Eq. (1) for differentiation:

$$\frac{\partial i_d(\boldsymbol{\beta})}{\partial \beta_v^{\mathrm{c}}} = N_{\mathrm{phot}} \int_{L_d} \frac{\partial [P(\mathcal{L}) I_B(\mathcal{L})]}{\partial \beta_v^{\mathrm{c}}} d\mathcal{L} \ . \tag{6}$$

The term inside the integral of Eq. (6) above can be written

$$\frac{\partial [P(\mathcal{L}) I_B(\mathcal{L})]}{\partial \beta_v^{\mathrm{c}}} = \left[\frac{I_B(\mathcal{L})}{P(\mathcal{L})} \frac{\partial P(\mathcal{L})}{\partial \beta_v^{\mathrm{c}}} + \frac{\partial I_B(\mathcal{L})}{\partial \beta_v^{\mathrm{c}}} \right] P(\mathcal{L}) \ . \tag{7}$$

Then, applying the transformations of Eq. (4) above, the sampled gradient is assessed by

$$\frac{\partial i_d}{\partial \beta_v^{\mathrm{c}}} \approx \sum_{p=1}^{N_{\mathrm{phot}}} \sum_{b=1}^{B} \frac{\partial i_d(p,b)}{\partial \beta_v^{\mathrm{c}}} \ , \tag{8}$$

where

$$\frac{\partial i_d(p,b)}{\partial \beta_v^{\mathrm{c}}} = \frac{I_b^{\mathrm{le}} \left[\mathcal{L}(d,p) \right]}{P(\mathcal{L})} \frac{\partial P(\mathcal{L})}{\partial \beta_v^{\mathrm{c}}} + \frac{\partial I_b^{\mathrm{le}} \left[\mathcal{L}(d,p) \right]}{\partial \beta_v^{\mathrm{c}}} \ . \tag{9}$$

Therefore,

$$\frac{\partial i_d(p,b)}{\partial \beta_v^{\mathrm{c}}} = I_b^{\mathrm{le}} \left[\mathcal{L}(d,p) \right] \frac{\partial \log [P(\mathcal{L})]}{\partial \beta_v^{\mathrm{c}}} + \frac{\partial I_b^{\mathrm{le}} \left[\mathcal{L}(d,p) \right]}{\partial \beta_v^{\mathrm{c}}} \ . \tag{10}$$

We now derive the terms of Eq. (10) above. In calculating I_b^{le}, the only term explicitly dependent on β_v^{c} is $a\left(\overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}}\right)$. Generally, the line segment $\overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}}$ traverses several voxels, as illustrated in Fig. 1 above. Let $\mathcal{V}_v$ be the domain of voxel v. The intersection of voxel v with $\overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}}$ is of length

$$l_v^{\odot} = \left| \overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}} \cap \mathcal{V}_v \right| \ . \tag{11}$$

Notice that if $\overline{\boldsymbol{x}_b, \boldsymbol{x}_b^{\odot}} \cap \mathcal{V}_v = \varnothing$, then $l_v^{\odot} = 0$. Using Eq. (11) above and Eq. (16) of the main manuscript, the transmittance a is

$$a\left(\overline{\boldsymbol{x}_b \boldsymbol{x}_b^{\odot}}\right) = \prod_v \exp\left[-l_v^{\odot} \left(\beta_v^{\mathrm{a}} + \beta_v^{\mathrm{c}} \right) \right] \ . \tag{12}$$

Following Eqs. (10,12) above,

$$\frac{\partial I_b^{\mathrm{le}}\left[\mathcal{L}\left(d,p\right)\right]}{\partial \beta_v^{\mathrm{c}}} = -l_v^{\odot} I_b^{\mathrm{le}}\left[\mathcal{L}\left(d,p\right)\right]. \tag{13}$$

Using Eq. (3) above,

$$\frac{\partial \log\left[P\left(\mathcal{L}\right)\right]}{\partial \beta_v^{\mathrm{c}}} = \sum_{b=1}^{B-1}\left[\frac{\partial \log\left[a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right)\right]}{\partial \beta_v^{\mathrm{c}}} + \frac{\partial \log(\beta_b)}{\partial \beta_v^{\mathrm{c}}}\right]. \tag{14}$$

Here we used the fact that given a scattering particle (air or cloud droplet) at $\boldsymbol{x}_b$, the phase function $f(\theta_{b-1,b})$ is independent of β_v^{c}.

Generally, the line segment denoted $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}$ traverses several voxels, as illustrated in Fig. 1 above. The length of the intersection of a voxel v with the set $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}$ is

$$l_{b-1,v} = \left|\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b} \cap \mathcal{V}_v\right| \ . \tag{15}$$

Similarly to Eq.(11) above, $l_{b-1,v} = 0$ if $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b} \cap \mathcal{V}_v = \varnothing$. Using Eq. (15) above and Eq. (16) of the main manuscript, the transmittance is

$$a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right) = \prod_v \exp\left[-l_{b-1,v}\left(\beta_v^{\mathrm{a}} + \beta_v^{\mathrm{c}}\right)\right] \ , \tag{16}$$

hence

$$\frac{\partial \log\left[a\left(\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}\right)\right]}{\partial \beta_v^{\mathrm{c}}} = -l_{b-1,v} \ . \tag{17}$$

Most line segments $\overline{\boldsymbol{x}_{b-1}\boldsymbol{x}_b}$ cross relatively few voxels. Thus, the output of Eq. (17) above is sparse over the domain.

The term $\partial \log(\beta_b)/\partial \beta_v^{\mathrm{c}}$ in Eq. (14) above is relevant only to the particular voxel at which scattering occurs. Moreover, air density is known irrespective of β_v^{c}. Scattering by an air molecule yields a null value of $\partial \log(\beta_b)/\partial \beta_v^{\mathrm{c}}$. Consequently,

$$\frac{\partial \log(\beta_b)}{\partial \beta_v^{\mathrm{c}}} = \begin{cases} 1/\beta_v^{\mathrm{c}} & \boldsymbol{x}_b \in \mathcal{V}_v \ \ \& \ \ \text{scatter is by a cloud droplet} \\ 0 & \text{else} \end{cases} . \tag{18}$$

Compounding Eqs. (10,13,14,17,18) above,

$$\frac{\partial i_d(\boldsymbol{x})}{\partial \beta_v^{\mathrm{c}}} \approx -\sum_{p=1}^{N_{\mathrm{phot}}}\sum_{b=1}^{B} I_b^{\mathrm{le}}\left[\mathcal{L}(d,p)\right] \cdot \left(l_v^{\odot} + l_{b-1,v} - \begin{cases} \frac{1}{\beta_v^{\mathrm{c}}} & \boldsymbol{x}_b \in \mathcal{V}_v \ \ \& \ \ \text{scattered by a cloud} \\ 0 & \text{else} \end{cases}\right). \tag{19}$$

3 Gradient of The Prior

Recall the formulation of the prior,

$$R(\boldsymbol{\beta}^{\mathrm{c}}) = \frac{1}{2}\mu_1 \|\mathbf{A}\boldsymbol{\beta}^{\mathrm{c}}\|_2^2 + \mu_2 R_{\mathrm{mon}}(\boldsymbol{\beta}^{\mathrm{c}}) \ , \tag{20}$$

where

$$R_{\mathrm{mon}}(\boldsymbol{\beta}^{\mathrm{c}}) = -\mathbf{1}^{\intercal} \tanh\left(c_1 \mathbf{D}_z \boldsymbol{\beta}^{\mathrm{c}}\right) \tanh\left(c_2 \boldsymbol{\beta}^{\mathrm{c}}\right) \mathbb{1}\left\{\boldsymbol{\beta}^{\mathrm{c}} > 0\right\}. \tag{21}$$

The gradient of the regularization term is

$$\frac{\partial R(\boldsymbol{\beta}^{\mathrm{c}})}{\partial \boldsymbol{\beta}^{\mathrm{c}}} = \mu_1 \mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{\beta}^{\mathrm{c}} + \mu_2 \frac{\partial R_{\mathrm{mon}}(\boldsymbol{\beta}^{\mathrm{c}})}{\partial \boldsymbol{\beta}^{\mathrm{c}}} \ . \tag{22}$$

The corresponding gradient of the monotonicity term is

$$\begin{aligned} \frac{\partial R_{\mathrm{mon}}(\boldsymbol{\beta}^{\mathrm{c}})}{\partial \boldsymbol{\beta}^{\mathrm{c}}} = -\Bigg[& c_1 \mathbf{D}_z \cosh^{-2}\left(c_1 \mathbf{D}_z \boldsymbol{\beta}^{\mathrm{c}}\right) \tanh\left(c_2 \boldsymbol{\beta}^{\mathrm{c}}\right) \\ & + c_2 \tanh\left(c_1 \mathbf{D}_z \boldsymbol{\beta}^{\mathrm{c}}\right) \cosh^{-2}\left(c_2 \boldsymbol{\beta}^{\mathrm{c}}\right) \Bigg] \mathbb{1}\left\{\boldsymbol{\beta}^{\mathrm{c}} > 0\right\}. \end{aligned} \tag{23}$$

Here $\cosh^{-2}(\mathbf{q})$ is a diagonal matrix. The diagonal is defined by the squared-reciprocal of the hyperbolic-cosine of each element of the argument vector. Recall that we use SGD to solve the optimization problem in Eq. (5) of the main manuscript,

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta} \geq 0} \left[F(\boldsymbol{\beta}) + R(\boldsymbol{\beta})\right] \ . \tag{24}$$

After each SGD iteration, we clip negative values of $\boldsymbol{\beta}^{\mathrm{c}}$.

4 Additional Details about the Simulations

As described in Sec. 6 of the main manuscript, estimation is done in coarse-to-fine stages. We used six stages in scene `F`, and eight stages in scene `C`. The number of voxels and pixels used in each stage of scene `F`, for example, are plotted in Fig. 2 below.

The cost function in Eq. (24) above is a combination of a data fitting term $F(\boldsymbol{\beta})$, and a regularization term $R(\boldsymbol{\beta})$. Moreover, recall the local reconstruction error measure,

$$\epsilon = \frac{\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{\mathrm{true}}\|_1}{\|\boldsymbol{\beta}^{\mathrm{true}}\|_1} \ . \tag{25}$$

This measure ϵ compares the iteratively-estimated model to the ground truth.

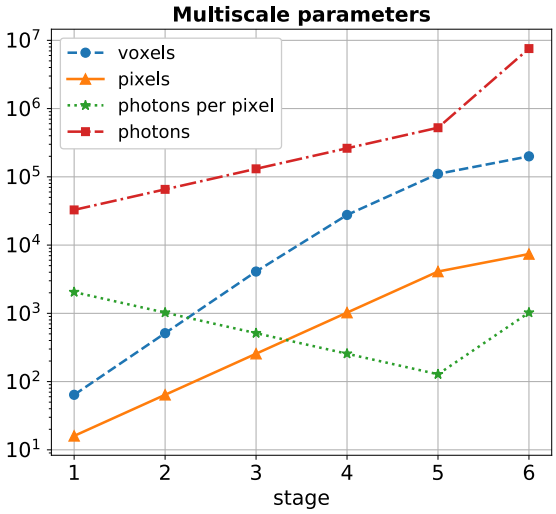

Fig. 2: The coarse-to-fine parameters as a function of the stages of the algorithm. This figure displays the increasing number of voxels, pixels, photons per pixel, and the overall number of photons.

We plot two criteria:

- The cost function from Eq. (24) above, as it evolves during SGD.
- The error measure Eq. (25) above, evolving during SGD.

Both are plotted as a function of time, during runs on our server. Scene `F` ran on a two CPUs Intel® Xeon® Platinum 8175M 2.40GHz with 24 cores each. And scene `C` ran on a two CPUs Intel® Xeon® Processor E5-2680 v4 2.40GHz with 14 cores each.

Overall, the criteria are plotted in Fig. 3 and Fig. 4 herein. Termination of each stage is color-marked in the plots here using a vertical colored line. Run time is displayed in log-scale. Fig. 3 herein corresponds to scene `F`. Fig. 4 herein corresponds to scene `C`.

Fig. 5 herein presents the effect of the monotonicity prior, beyond a smoothness prior (scene `F`). The left plot presents how the cost function Eq. (24) above evolves during the whole SGD process. The right figure presents a zoom-in of the error measure Eq. (25) above in the last stage of the coarse-to-fine solution. The orange lines present the cost Eq. (24) above, and error Eq. (25) above without the monotonicity prior, i.e., $R(\boldsymbol{\beta}^{\mathrm{c}}) = \frac{1}{2}\mu_1\|\mathbf{A}\boldsymbol{\beta}^{\mathrm{c}}\|_2^2$. The blue lines present these respective costs when solving the full optimization problem, as formulated in the main manuscript, including the monotonicity prior. Here the regularization term is $R(\boldsymbol{\beta}^{\mathrm{c}}) = \frac{1}{2}\mu_1\|\mathbf{A}\boldsymbol{\beta}^{\mathrm{c}}\|_2^2 + \mu_2 R_{\mathrm{mon}}(\boldsymbol{\beta}^{\mathrm{c}})$. As stated in the main manuscript, the monotonicity prior improves the quality and convergence.

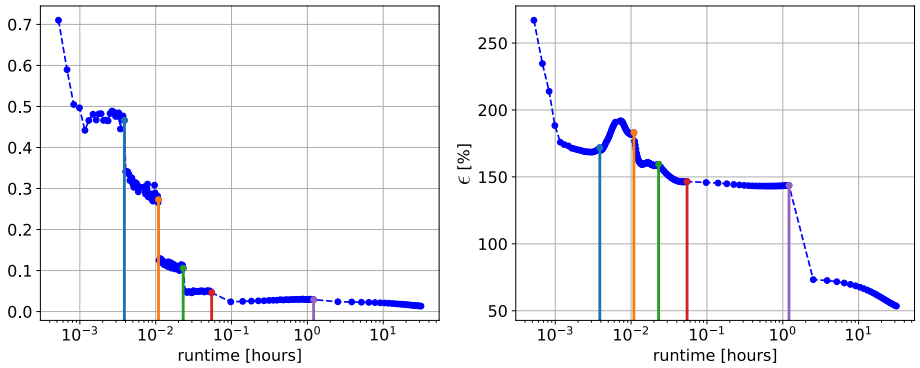

Fig. 3: Recovery of scene `F`. These plots present [Left] the cost function Eq. (24) above, and [Right] the error function ϵ Eq. (25) above, as functions of run time.

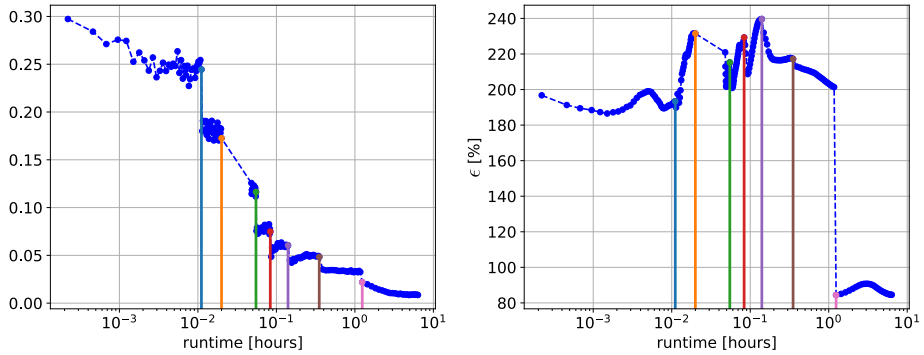

Fig. 4: Recovery of scene `C`. These plots present [Left] the cost function Eq. (24) above, and [Right] the error function ϵ Eq. (25) above, as functions of run time. The hierarchical solution enables a faster convergence to a better solution.

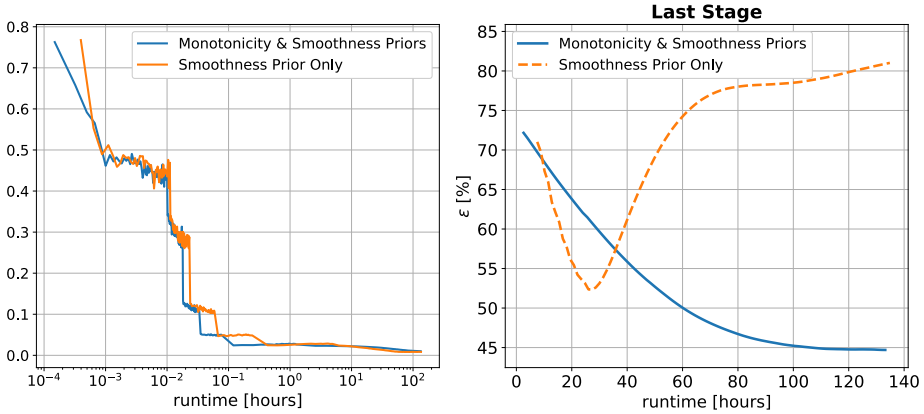

Fig. 5: This figure illustrates the recovery of scene `F`, using exclusively the smoothness prior (orange line), or using monotonicity compounded with smoothness (blue line), as formulated in Eq. (24) above. [Left] The cost function evaluation, as a function of run time. [Right] The error ϵ during the last stage, as a function of run time on our server. The error ϵ significantly improves using the monotonicity prior.

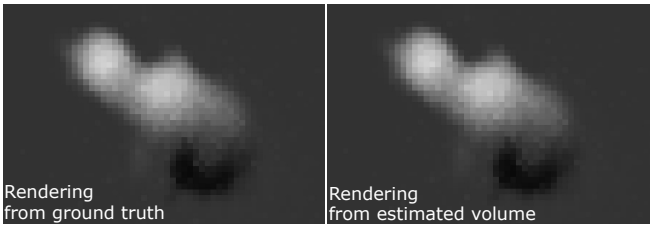

Fig. 6: Rendering a new viewpoint, 50.4° off-nadir. This viewpoint was not used during the reconstruction of scene `C`. This image was not an input to tomographic recovery. [Left] Image rendered from the ground-truth medium. [Right] Image rendered from the estimated medium.

Finally, Fig. 6 presents an application of scattering tomography. Here we render a new viewpoint, that was not included in the input measurements of scene `C`. For comparison, we render an image using the ground truth medium. Then we render the same viewpoint using the estimated medium.

References

1. Gkioulekas, I., Zhao, S., Bala, K., Zickler, T., Levin, A.: Inverse volume rendering with material dictionaries. ACM Trans. Graph. **32**(6), 162 (2013)