4D Cloud Scattering Tomography

Roi Ronen, Yoav Y. Schechner
Viterbi Faculty of Electrical & Computer Eng.
Technion - Israel Institute of Technology
Haifa, Israel
ronen.roi@gmail.com
yoav@ee.technion.ac.il

Eshkol Eytan
Department of Earth & Planetary Sciences
The Weizmann Institute of Science
Rehovot, Israel
eshkol.eytan@weizmann.ac.il

Abstract

We derive computed tomography (CT) of a time-varying volumetric scattering object, using a small number of moving cameras. We focus on passive tomography of dynamic clouds, as clouds have a major effect on the Earth's climate. State of the art scattering CT assumes a static object. Existing 4D CT methods rely on a linear image formation model and often on significant priors. In this paper, the angular and temporal sampling rates needed for a proper recovery are discussed. Spatiotemporal CT is achieved using gradient-based optimization, which accounts for the correlation time of the dynamic object content. We demonstrate this in physics-based simulations and on experimental real-world data.

1. Introduction

Computed tomography (CT) aims to recover the inner structure of three dimensional (3D) volumetric heterogeneous objects [15, 16]. CT has extensive use in many domains. These include medicine [14, 41], sensing of atmospheric pollution [2], geophysics [49] and fluid dynamics [27, 51, 52]. CT requires multi-view imaging [3, 22]. In nearly all CT approaches, the object is considered static during the multi-view acquisition. However, often the object changes while views are acquired sequentially [8, 53]. Thus, effort has been invested to generalize 3D CT to four-dimensional (4D) spatiotemporal CT, particularly in the computer vision and graphics communities [42, 52, 53]. This effort has focused on linear-CT modalities. Linear CT is computationally easier to handle, thus common for decades, mainly in medical imaging [19]. Medical CT often exploits the periodic temporal nature of organ dynamics, to synchronize sequential acquisitions [41].

This paper deals with a more complicated model: scattering CT. It is important to treat this case for scientific, societal and practical reasons. The climate is strongly affected by interaction with clouds [13]. To reduce major errors in climate predictions, this interaction requires a much finer understanding of cloud physics. Current models are based on remote sensing data that is analyzed under the assumption that the atmosphere and clouds are made of very broad and uniform layers. This leads to errors in climate understanding. To overcome this problem, 3D scattering CT has been suggested as a way to study clouds [29, 30, 46].

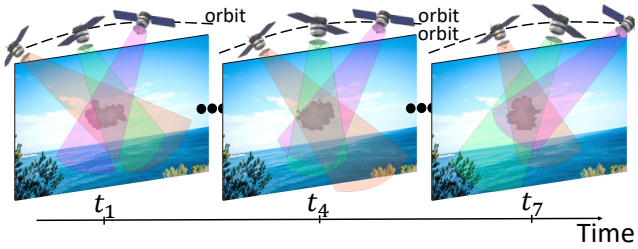

Figure 1. Multiple moving sensors image a time-varying object (cloud) from multiple-views. Tomography seeks the inner content.

Scattering CT of clouds requires high resolution multi-view images from space. There are spaceborne and high-altitude systems that may provide such data, such as AirMSPI [6], MAIA [4], HARP [38], AirHARP [35] and the planned CloudCT formation [43]. These systems are so expensive, that it is unrealistic to deploy them in large numbers to simultaneously image the same clouds from many angles. Therefore, in practice, platforms move above the clouds: a sequence of images is taken, in order to span and sample a wide angular breadth (Fig. 1), but the cloud evolves meanwhile. Hence there are important reasons to derive 4D scattering CT of clouds.

We pose conditions for performing this task. These relate to temporal sampling and angular breadth, in relation to the correlation time of the evolving object. Then, we generalize prior 3D scattering CT, to spatiotemporal recovery using data taken by moving cameras. We present an optimization-based method to reach this. The method is demonstrated both in rigorous simulations and on real data.

2. Theoretical Background

2.1. Cloud Microphysical Parameters

Warm clouds are composed of water droplets. The droplet size distribution is typically parameterized by an effective radius r^{e} and a dimensionless droplet variance v^{e} (see [17] for details). These parameters vary with spatial location $\mathbf{x}$, and assumed here to be uniform in a voxel around $\mathbf{x}$. An additional characterization is the *liquid water content* (LWC), $\mathcal{L}$, which increases linearly with the number-density and volume of droplets in a voxel [17, 30].

A common approximation in convective cloud models is that $\mathcal{L}$ and r^{e} tend to increase with altitude inside a cloud [32, 47, 54]. Let z_0 be the cloud base. For adiabatic convection, at $z > z_0$

$$r^{\mathrm{e}} \approx \xi_{\mathrm{r}}(z - z_0)^{\frac{1}{3}} + r_0^{\mathrm{e}} \ , \quad \mathcal{L} \approx \xi_{\mathcal{L}}(z - z_0) \ , \tag{1}$$

where r_0^{e} is the effective radius at the cloud base and $\xi_{\mathrm{r}}, \xi_{\mathcal{L}} > 0$. The error of this approximation grows as the cloud mixes with its surroundings air at the cloud shell, mainly at its top. The values of $r^{\mathrm{e}}, v^{\mathrm{e}}$ tend to be rather uniform per altitude, while $\mathcal{L}$ can change significantly in 3D [21].

Overall, the vector $\nu_t = [\mathcal{L}_t, r_t^{\mathrm{e}}, v_t^{\mathrm{e}}]$ characterizes a voxel at time t. Concatenating these parameters across all spatial voxels results in a vector $\boldsymbol{\nu}_t$, which expresses the cloud structure at time t.

2.2. The Forward Model

The interaction of radiation with a scattering volumetric object is modelled by 3D *radiative transfer*, which includes multiple scattering. Define $\mathcal{M}_\lambda[\cdot]$ as a differentiable operator that maps microphysical parameters to macroscopic optical parameters corresponding to wavelength λ, using Mie theory [12, 30]. Let

$$L(\mathbf{x}, \boldsymbol{\Omega}, t) = \mathrm{RT}(\mathcal{M}_\lambda[\boldsymbol{\nu}_t]) \tag{2}$$

be the radiance resulting from radiative transfer at each spatial location $\mathbf{x}$ and each direction $\boldsymbol{\Omega}$. There are various algorithms to implement $\mathrm{RT}(\mathcal{M}_\lambda[\boldsymbol{\nu}_t])$, including Monte-Carlo [32, 34] and the spherical harmonic discrete ordinate method (SHDOM). We use the latter, as it is considered trustworthy by the scientific community [9] and has an open-source[1] online code [28] .

A camera observes the scene from a specific location, while each pixel in a camera samples a viewing direction $\boldsymbol{\Omega}$. Hence, imaging (forward model) amounts to sampling the output of 3D radiative transfer at the camera locations and the lines of sight of the pixels. This sampling integrates over the camera exposure time and spectral bands. Camera sampling is denoted by a projection operator $P_{\mathbf{x},\boldsymbol{\Omega}}$.

The forward image formation model $\mathcal{F}(\boldsymbol{\nu}_t)$ yields the expected graylevel at a pixel, at time t:

$$I_{\mathbf{x},\boldsymbol{\Omega},t} = \mathcal{F}(\boldsymbol{\nu}_t) \approx \gamma^{\mathrm{cam}} P_{\mathbf{x},\boldsymbol{\Omega}} \left\{ \mathrm{RT}(\mathcal{M}_\lambda[\boldsymbol{\nu}_t]) \right\} \ . \tag{3}$$

Here γ^{cam} expresses camera properties, including the lens aperture area, exposure time, spectral band, quantum efficiency and lens transmissivity. Eq. (3) assumes that the exposure time is sufficiently short, such that within this time, the scene and the camera pose vary insignificantly.

Empirical measurements include random noise [2, 6, 48]. The noise mainly originates from the discrete nature of photons and electric charges, which yields a Poisson process. There are additional noise sources, and their parameters can be extracted from the sensor specifications. Denote incorporation of noise into the expected signal by the operator $\mathcal{N}$. Then, a raw measurement is

$$y_{\mathbf{x},\boldsymbol{\Omega},t} = \mathcal{N}\left\{ I_{\mathbf{x},\boldsymbol{\Omega},t} \right\} \ . \tag{4}$$

Per t, all multi-view, multi-pixel measurements are concatenated into a vector $\boldsymbol{y}_t$. Concatenating $\boldsymbol{y}_t$ over all t yields the vector $\boldsymbol{y}$.

2.3. Scattering 3D Tomography of Clouds

This section expresses the state of the art in 3D scattering tomography [32, 44, 47], in which $\boldsymbol{\nu}_t$ is assumed to be invariant to t. Hence, t is generally dropped from the derivations here. Estimation of $\boldsymbol{\nu}$ is done by minimization of a cost $\mathcal{E}$, which penalizes the discrepancy between $\boldsymbol{y}$ and the forward model,

$$\hat{\boldsymbol{\nu}} = \underset{\boldsymbol{\nu}}{\mathrm{argmin}} \, \mathcal{E}\left[\boldsymbol{y}, \mathcal{F}(\boldsymbol{\nu})\right] \ . \tag{5}$$

Eq. (5) can be solved efficiently by gradient-based methods. By setting

$$\mathcal{E}\left[\boldsymbol{y}, \mathcal{F}(\boldsymbol{\nu})\right] = \frac{1}{2} \| \boldsymbol{y} - \mathcal{F}(\boldsymbol{\nu}) \|_2^2 \ , \tag{6}$$

the gradient of Eq. (5) with respect to $\boldsymbol{\nu}$ is

$$\mathbf{g}(\boldsymbol{\nu}) \equiv \frac{\partial \mathcal{E}\left[\boldsymbol{y}, \mathcal{F}(\boldsymbol{\nu})\right]}{\partial \boldsymbol{\nu}} = \frac{\partial \mathcal{F}(\boldsymbol{\nu})}{\partial \boldsymbol{\nu}} \left[\mathcal{F}(\boldsymbol{\nu}) - \boldsymbol{y}\right] \ . \tag{7}$$

Gradient-based optimization performs per iteration k

$$\boldsymbol{\nu}(k+1) = \boldsymbol{\nu}(k) - \eta \mathbf{g}[\boldsymbol{\nu}(k)] \tag{8}$$

where η is a step size. In scattering CT, computing the Jacobian $\partial \mathcal{F}(\boldsymbol{\nu}) / \partial \boldsymbol{\nu}$ is complex. However, there are approximations to the Jacobian of 3D RT, which can be computed efficiently [29, 31, 32], making recovery tractable.

[1] Eqs. (3,6,7,16) use the forward model $\mathcal{F}$ and Jacobian $\partial \mathcal{F} / \partial \nu$ as black boxes, agnostic to a specific implementation. A differential RT solver as Mitsuba2 [39] can be used, as any other solver. Currently, however, Mitsuba2 supports neither heterogeneous media having a mixture of spatially varying materials (both air molecules and water droplets with varying densities), nor mixtures of Mie phase functions. To comply with atmospheric science standards, we use SHDOM as our renderer.

We stress that traditional CT methods cannot apply to cloud recovery. To see this, let L^{sun} be the solar irradiance. Let $a^{\text{ground}} \approx 0.05$ be the ground albedo. The range of the optical depth $\mathtt{T}^{\text{cloud}}$ of warm clouds is typically 10-100. The albedo of warm clouds is $a^{\text{cloud}} \approx 0.5$. The radiance directly transmitted from the ground through a cloud to a sensor above is

$$D \sim L^{\text{sun}} a^{\text{ground}} \exp(-\mathtt{T}^{\text{cloud}}) \ . \tag{9}$$

Sunlight reflected above by a cloud has radiance

$$S \sim L^{\text{sun}} a^{\text{cloud}} \ . \tag{10}$$

From these orders of magnitude, $D \ll S$. The measured signal is dominated by S. Suppose a naive approach following traditional CT, associating a measured signal with direct transmission i.e., $S \approx D$. Then from Eqs. (9,10), CT estimates $\hat{\mathtt{T}}^{\text{cloud}} \approx \log(a^{\text{ground}}/a^{\text{cloud}})$. From these orders of magnitude, $\hat{\mathtt{T}}^{\text{cloud}} < 0$, which is not-physical. Thus, traditional CT is irrelevant for cloud tomography in visible light.

2.4. Temporal Sampling of a Random Object

A temporal sample indexed l corresponds to continuous time t'_l. The time interval between consecutive samples is $T = |t'_{l+1} - t'_l|$. Consider a continuously varying object $\boldsymbol{\beta}_t$. A temporal sample is denoted $\boldsymbol{\beta}^{\text{sample}}_{t'}$. The Nyquist sampling theorem [40] relates to objects whose time-spectrum is limited to temporal frequencies ω satisfying $|\omega| < B$, where B is a cutoff frequency. Then, time domain samples satisfying $T \leq (2B)^{-1}$ can yield reconstruction of $\boldsymbol{\beta}_t$ using a linear superposition:

$$\boldsymbol{\beta}_t(\mathbf{x}) \sim \sum_{t'} w_t(t'|T) \boldsymbol{\beta}^{\text{sample}}_{t'}(\mathbf{x}) \ . \tag{11}$$

There, the superposition uses $w_t(t'|T) = \text{sinc}[(t - t')/T]$.

There is a generalization, however, to cases where the object $\boldsymbol{\beta}_t$ is random and not strictly band-limited. The temporal auto-correlation of $\boldsymbol{\beta}_t$ is

$$\alpha(\tau|\sigma) = \langle \boldsymbol{\beta}_t(\mathbf{x}), \boldsymbol{\beta}_{t-\tau}(\mathbf{x}) \rangle_{t,\mathbf{x}} \ . \tag{12}$$

The function $\alpha(\tau|\sigma)$ generally decays with $|\tau|$, where σ is the effective decay time of α, termed *correlation time*. Two limiting cases are illustrative. For $\sigma \to \infty$, we have $\alpha(\tau|\sigma) \to$ constant. This means that the object $\boldsymbol{\beta}$ is effectively static. In contrast, for $\sigma \to 0$, we have $\alpha(\tau|\sigma) \to \delta(t - t')$, i.e., a Dirac delta function. This means that the object $\boldsymbol{\beta}$ varies so fast, that at any time t its state is uncorrelated to the state at other times.

Once again, $\boldsymbol{\beta}_t$ can be linearly reconstructed from temporal samples using Eq. (11), but $w_t(t'|T)$ can be general. Any sampling rate and reconstruction kernel can be used,

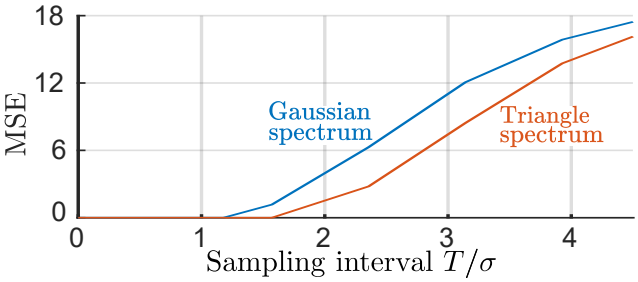

Figure 2. Illustration of Eq. (13) for Gaussian or triangular spectra.

but then there is a reconstruction error. The reconstruction mean squared error (MSE) has a *bound* [36]. Let $\Lambda(\omega)$ be the time-spectrum of the random object, i.e., the Fourier transform of $\alpha(\tau|\sigma)$. The MSE bound [36] is then

$$\text{MSE}(T) = \frac{T}{2\pi} \int_{-\pi}^{\pi} \left[\Lambda(\omega) - \frac{\sum\limits_{q \in \mathbb{Z}} \Lambda^2[\omega - (2\pi/T)q]}{\sum\limits_{q \in \mathbb{Z}} \Lambda[\omega - (2\pi/T)q]} \right] d\omega \ . \tag{13}$$

For illustration, Fig. 2 plots $\text{MSE}(T)$ for objects that have Gaussian or triangular spectra. When $T < \sigma$, the error is negligible, but error accumulates significantly as the sampling interval T increases beyond σ. Hence, to keep reconstruction error small, an efficient temporal sampling interval should satisfy $T \approx \sigma$.

3. Clouds: Correlation Time and Sampling

Warm convective clouds are governed by air turbulence of decameter scale. In these scales [13], the *correlation time* of content in a voxel is about 20 to 50 seconds. This indicates that *4D spatiotemporal clouds can be recovered well using 4D spatiotemporal samples*, if the temporal samples are about 25 seconds apart. The lifetime of a warm convective cloud is typically measured in minutes.

Consider a cloud simulation, described in detail in Sec. 6. The cloud evolves for about 10 minutes. For each cloud voxel, we calculated the temporal auto-correlation of $\mathcal{L}_t$. Similarly, temporal auto-correlations were derived for horizontally-averaged r^{e}_t and v^{e}_t. The auto-correlation functions of $\mathcal{L}_t$ and r^{e}_t are plotted in Fig. 3a (Eq. 12). The auto-correlation function of r^{e}_t behaves similarly to that of v^{e}_t. Clearly, the correlation times of r^{e}_t and of v^{e}_t are very long, comparable to the lifetime of a cloud, and longer than the typical time it takes to acquire multi-view data of clouds. Hence, when recovering microphysical parameters r^{e} and v^{e}, we neglect temporal variations.

On the other hand, $\mathcal{L}_t$ has a short correlation time: $\sigma \sim 25\,\text{sec}$. Hence, 4D recovery is necessary for $\mathcal{L}_t$, if data is sparsely sampled in time. Following the conclusion of Sec. 2.4, it is advisable to sample warm convective clouds at temporal sampling interval of $T \approx \sigma \approx 25\,\text{sec}$.

From Sec. 2.4, at an arbitrary t, reconstruction of $\mathcal{L}_t$

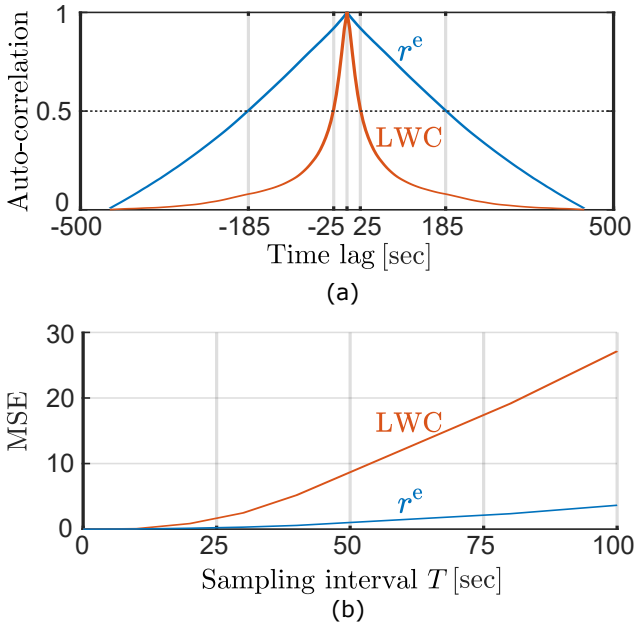

Figure 3. (a) Auto-correlation of cloud field microphysics. The auto-correlation of LWC and r^e decrease to 0.5 after 25 sec and 185 sec, respectively. Typical cloud multi view sensing is about a minute, during which the LWC auto-correlation decreases to 0.26, while the auto-correlation of r^e is 0.82. (b) MSE (Eq. 13) of LWC and r^e.

from samples would use a temporal kernel whose effective width is σ. In CT, however, we do not have direct access to the object samples. We only have noisy projections of radiance scattered by the object. Thus, reconstruction does not involve direct application of the optimal kernel [36], which achieves the bound in Eq. (13). Reconstruction involves a tomographic process, which we describe in Sec. 5.

4. Tomographic Angular Extent

Section 2.4 dealt with sampling of an object, as if 4D measurements are done in-situ. However, in CT, we have no direct access to $\boldsymbol{\nu}_t$: we only measure projections $\boldsymbol{y}_t$. As we discuss now, projections must have a *wide angular breadth, while object evolution is small*.

Consider an extreme case. Let a cloud be temporally constant and reside only in a single voxel, over the ocean. Viewed from space by two cameras simultaneously, cloud recovery here amounts to triangulation. In triangulation, the best cloud-localization resolution is obtained if the angular range between the two cameras is 90°. At small baselines, localization decreases linearly with a decreasing angular extent. When more than two cameras operate, the trend is similar. Consider an error measure that has been used in 3D cloud scattering CT [20, 29, 30, 32],

$$\varepsilon = \frac{\|\boldsymbol{\mathcal{L}}^{\text{true}} - \hat{\boldsymbol{\mathcal{L}}}\|_1}{\|\boldsymbol{\mathcal{L}}^{\text{true}}\|_1} \,. \tag{14}$$

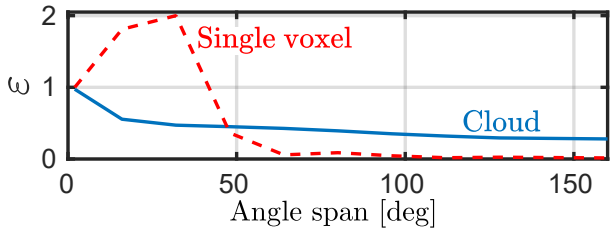

Figure 4. A static heterogeneous cloud and a single-voxel "cloud" (having size 20 m × 20 m × 20 m) are recovered from nine viewpoints using [30]. The plots are of errors defined in Eq. (14).

Here $\hat{\boldsymbol{\mathcal{L}}}$ is the estimated 3D cloud LWC field. Fig. 4 plots ε when CT attempts to recover a single-voxel of a *static* cloud (extending 20 m), when 9 cameras surround it from 500 km away, while the true r^e, v^e are used. Above $\approx 60°$ total angular extent, recovery reaches a limiting excellent quality, but quality is very poor at narrow angle spans.

In general, objects have multiple voxels. In linear-CT (as in medical X-ray CT), information loss due to limited-angle imaging is known as the *missing cone of frequencies* [1, 33]. In scattering CT, with the exception of very sparse objects, the missing cone linear theory does not apply. While a pure theory for nonlinear CT does not exist, the implication of angular span can be assessed numerically. Let us consider a *static* cloud in a single state, simulated as in Sec. 6. There are no dynamics. We can then see how angular sampling alone affects the quality of recovery. Results are shown in Fig. 4. There is a marked degradation of quality if the angular extent is narrow.

So far, this section dealt with static clouds. Clouds are considered nearly static between times t, t' if $|t - t'| < \sigma$. The viewing angular extent covered in those times (and in intermediate times) is denoted $\Theta(t, t')$, in radians. So, within time span approximately equal to σ, good recovery can be achieved only if $2\Theta(t, t')/\pi$ is large. If it is low, then spatial (altitude) resolution in CT recovery is lost. Most CT systems cover wide angular extent, eventually. So, quality is set by the angular *rate*. Define a dimensionless figure

$$\rho = \frac{2\Theta(t, t')}{\pi} \frac{\sigma}{|t - t'|} \,. \tag{15}$$

Good 4D recovery requires $\rho \gtrsim 1$, while $T \lesssim \sigma$ (Sec. 3). The more these conditions are violated, the worse 4D CT is expected to perform.

5. 4D Scattering Tomography

5.1. Estimation of the Liquid Water Content

We now generalize Eq. (8) to 4D CT. Data is captured sequentially at the time set $\mathcal{T} = \{t_1, t_2, \ldots, t_{N^{\text{state}}}\}$, while the object evolves. At each $t' \in \mathcal{T}$, the object is viewed simultaneously from a set of viewpoints $\mathcal{C}_{t'}$, yielding a concatenated data vector $\boldsymbol{y}_{t'}$. At that time, the modelled LWC

is the vector $\boldsymbol{\mathcal{L}}_{t'}$. The set of LWC values in all sampled times is denoted $\mathcal{B} = \{\boldsymbol{\mathcal{L}}_{t'}\}_{t' \in \mathcal{T}}$. At optimization iteration k, the set of all modelled LWC values is $\mathcal{B}(k)$.

Let $w_t(t'|\sigma)$ be a normalized weighting function. Consider the vector

$$\mathbf{g}_t(\mathcal{B}) = \sum_{t' \in \mathcal{T}} w_t(t'|\sigma) \frac{\partial \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial \boldsymbol{\mathcal{L}}_{t'}} \left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] \ , \quad (16)$$

and an iteration move to assess $\boldsymbol{\mathcal{L}}_t$ at arbitrary time t

$$\boldsymbol{\mathcal{L}}_t(k+1) = \boldsymbol{\mathcal{L}}_t(k) - \eta \mathbf{g}_t[\mathcal{B}(k)] \ . \quad (17)$$

We use Eqs. (16,17) iteratively for 4D scattering tomography. We use $w_t(t'|\sigma)$ which decays in effective time σ.

Sec. 2.4 serves as a guideline for the kernel properties. However, as said in Sec. 3, we do not have access to the LWC, but to noisy images. Hence we approximate the optimal kernel using a cropped Gaussian,

$$w_t(t'|\sigma) = s \exp\left(-\frac{|t-t'|^2}{2\sigma^2}\right) \ . \quad (18)$$

Here s is a normalization factor, set so $\sum_{t' \in \mathcal{T}} w_t(t'|\sigma) = 1$ while $w_t(t' \notin \mathcal{T}|\sigma) = 0$. In the limiting case where $w_{t'}(t|\sigma) \to 1/N^{\text{state}}$, the cloud is considered static (correlation time is very long). Then, Eqs. (16,17) degenerate to Eqs. (7,8). That is, recovery of $\boldsymbol{\mathcal{L}}_t$ is tightly related to $\{\boldsymbol{\mathcal{L}}_{t'}\}_{t' \in \mathcal{T}}$. More generally, the correlation time σ has a finite value. Thus, when estimating $\boldsymbol{\mathcal{L}}_t$, there is gradually lower information carried by $\boldsymbol{\mathcal{L}}_{t'}$, as $|t'-t|$ increases, and particularly as $|t'-t| > \sigma$. Thus, an iteration to refine an estimate of $\boldsymbol{\mathcal{L}}_t$ should give less weight to $\boldsymbol{\mathcal{L}}_{t'}$. Eqs. (16,17) provide this capacity in a natural way.

Eq. (16) is equivalent to a gradient of a cost function. This interpretation is detailed in the Supplementary material. The complexity of Eqs. (16,17) is similar to static 3D CT (5), as discussed in the Supplementary material. We performed iterations using L-BFGS-B [55]. Following [30], prior to iterations, the set of voxels to estimate is bounded using space-carving [25]. Space-carving bounds a 3D shape by back-projecting multi-view images. A voxel is labeled as belonging to the object, if the number of back-projected rays that intersect this voxel is greater than a threshold. We adapt this bounding to dynamic scenarios using two ways: (i) by setting a coarse spatial grid for carving and (ii) using a low threshold for labeling voxels as potentially being part of a cloud.

5.2. Estimation of the Effective Radius

We exploit simplifications with regards to r^{e} and v^{e}. As discussed in Sec. 3, we may treat r^{e} and v^{e} as time-invariant per voxel, in the scale of minutes. Furthermore, by Sec. 2.1, r^{e} can be approximated as laterally uniform (varying vertically) [21], using a parametric model (1) [47]. Moreover, it can be often assumed that $v^{\text{e}} = 0.1$ [30].

Hence, we focus here on time-invariant estimation of r^{e}, using the model in Eq. (1), namely, estimating the global parameters $\xi_{\text{r}}, r_0^{\text{e}}, z_0$. This is done intermittently, among optimization of the LWC (Sec. 5.1). Overall we estimate both LWC and r^{e} using the following steps:

{i} Perform exhaustive search on $\xi_{\text{r}}, \xi_{\mathcal{L}}, r_0^{\text{e}}, z_0$, in a coarse grid of values, to minimize Eq. (6), assuming time invariance. This sets the initial values of $r^{\text{e}}, \{\boldsymbol{\mathcal{L}}_t\}_{t \in \mathcal{T}}$.

{ii} Hold r^{e} temporally fixed. Run the gradient-based optimization described in Sec. 5.1, to estimate $\{\boldsymbol{\mathcal{L}}_t\}_{t \in \mathcal{T}}$ in 4D. Here 10 iterations are run.

{iii} Hold $\{\boldsymbol{\mathcal{L}}_t\}_{t \in \mathcal{T}}$ fixed. Perform exhaustive line-search on ξ_{r}, in a fine grid of values, to minimize Eq. (6).

{iv} Return to step {ii}.

6. Simulations

We now test the feasibility of 4D cloud scattering tomography. The tests demonstrate the effect of varying σ, the kernel parameter in Eq. (18), and the importance of the angular breadth. The evolving concentration of cloud water droplets is the main unknown we sense and seek.[2]

For realistic complexity, we use a rigorous simulation based on cloud physics. Clouds are simulated using the System of Atmospheric Modeling (SAM) [24], which is a non-hydrostatic, inelastic large eddy simulator (LES) [18, 37, 50]. It describes the turbulent atmosphere using equations of momentum, temperature, water mass balance and continuity. We couple SAM to a spectral (bin) microphysical model (HUJI SBM) [11, 23] of the droplets' size. It propagates the evolution of the droplet size distribution, by solving the equations for nucleation, diffusional growth, collision-coalescence and break-up. This is done on a logarithmic grid of 33 size bins in the domain $[2\,\mu\text{m}, 3.2\,\text{mm}]$.

The simulation runs according to the BOMEX case [45] of trade wind cumulus clouds near Barbados. Humidity and potential temperature profiles are used as initial conditions, while the surface fluxes and large-scale forcing are constant. The mean horizontal background wind is zero. The horizontal boundary condition is cyclic. The domain is $5.12\,\text{km}$ long (cloud diameter is $\approx 800\,\text{m}$) at $10\,\text{m}$ resolution. The vertical resolution is $10\,\text{m}$ from sea level to $3\,\text{km}$, coarsening to $50\,\text{m}$ above. Cloud tops reach $2\,\text{km}$. The simulation expresses an hour, of which 30 minutes includes the cloud's lifetime. The temporal resolution is $0.5\,\text{sec}$.

We present results using two different time-varying clouds: *Cloud (i)* has size $43 \times 30 \times 45$ voxels (see Fig. 5). *Cloud (ii)* has size $60 \times 40 \times 45$ (see Supplementary material). A voxel size is $10\,\text{m} \times 10\,\text{m} \times 10\,\text{m}$.

[2] Scatter by droplets is usually more dominant and spatiotemporally variable than aerosols. Molecular density changes mainly vertically and is usually known using non-imaging sensors. Molecules then scatter according to the known Rayleigh theory. Thus, we focus on droplets.

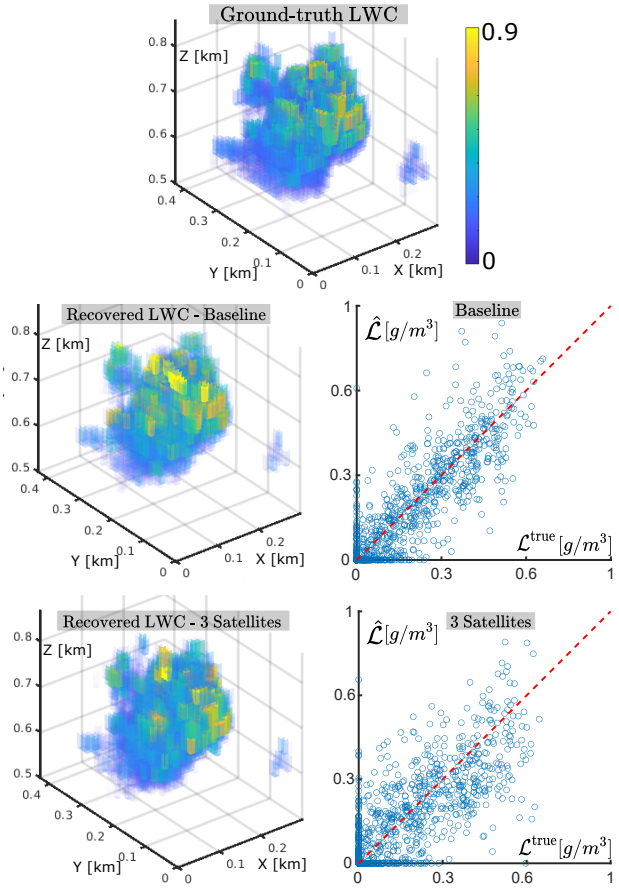

Figure 5. *Cloud (i).* Results of recovery by the `Baseline` and `Setup A` are compared to the ground-truth by a 3D presentation and scatter plots that use 20% of the data points, randomly selected for display clarity. The `Baseline` and `Setup A` scatter plot correlations are 0.9 and 0.86, respectively.

6.1. Rendered Measurements

The scene is irradiated by the sun, whose illumination angle changes in time, relative to the Earth's coordinates, while cameras overfly the evolving cloud. The solar trajectory in Earth coordinates corresponds to Feb/03/2013 at 13:54:30 - 14:01:00 local time, around 38N 123W. We tested several types of imaging setups :

`Setup A`: Three satellites orbit at $500\,\text{km}$ altitude, one after the other. Their velocity is $7.35\,\text{km/s}$. The orbital arc-length between nearest-neighboring satellites is $500\,\text{km}$. At mid-time of the simulation, $t = (t_1 + t_{N^{\text{state}}})/2$, the setup is symmetric around the nadir direction. Then, the setup spans an angular range of $114°$. Each satellite carries a perspective camera. The camera resolution is such that at nadir view, a pixel corresponds to 10m at sea level. Images are taken every $10\,\text{sec}$, during $60\,\text{sec}$, i.e., $N^{\text{state}} = 7$. This setup is illustrated in Fig. 1.

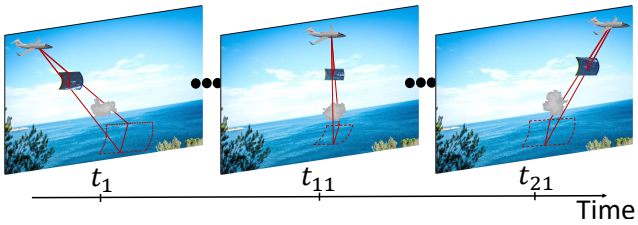

Figure 6. Illustration of `Setup C`. A domain is viewed at 21 push-broom angles, sequentially.

`Baseline`: The baseline uses all the accumulated 21 viewpoints of `Setup A`. However, all viewpoints here have perspective cameras that *simultaneously* acquire the cloud. In other words, this baseline is not prone to errors that stem from temporal sampling. The baseline is used for recovery only at time $t = (t_1 + t_{N^{\text{state}}})/2$.

`Setup B`: This setup is similar to `Setup A`, but it uses only two satellites. Thus, at mid-time of the simulation, the setup spans a $57°$ angular range.

`Setup C`: A single camera, similar to the Multi-angle Spectro-Polarimeter Imager (AirMSPI) [6], is mounted on an aircraft flying $154°$ relative to North at $20\,\text{km}$ altitude. Imaging has a pushbroom scan geometry, having $10\,\text{m}$ spatial resolution at Nadir view. AirMSPI scans view angles in a step-and-stare mode [6]. Based on AirMSPI PODEX campaign [5], we set 21 viewing angles along-track: $\pm 65°$, $\pm 62°$, $\pm 58°$, $\pm 54°$, $\pm 50°$, $\pm 44°$, $\pm 38°$, $\pm 30°$, $\pm\ 21°, \pm 11°$ off-nadir and $0°$ (nadir). For example, three sample angles are illustrated in Fig. 6. It takes $\approx 1\,\text{sec}$ to scan a cloud domain in any single view angle, during which the cloud and solar directions are assumed constant. Dynamics are noticeable *between* view angles.

A spherical harmonic discrete ordinate method (SHDOM) code [10] provides the numerical forward model $\mathcal{F}$. Simulated measurements $\{\boldsymbol{y}_t\}_{t\in\mathcal{T}}$ include noise. The noise model follows the AirMSPI sensors parameters [6, 48]. There, the sensor full-well depth is 200,000 photo-electrons, readout noise has a standard deviation of 20 electrons, and the overall readout is quantized to 9 bits.

6.2. Results of 4D Tomography of the LWC

The rendered and noisy images, in the spectral band of $\lambda = 660\,\text{nm}$, served as input to 4D tomographic reconstruction. The voxel size in the recovery was set to $10\,\text{m} \times 10\,\text{m}$ horizontal, $25\,\text{m}$ vertical and $10\,\text{sec}$ resolution. For parallelization, optimization ran on a computer cluster, where each computer core was dedicated to rendering a modelled image from a distinct angle.

In this section, we recover only the cloud LWC. Hence, we set $r^{\text{e}} = 10\,\mu\text{m}$ and $v^{\text{e}} = 0.1$ to be uniform constants during optimization (though they were not uniform in the

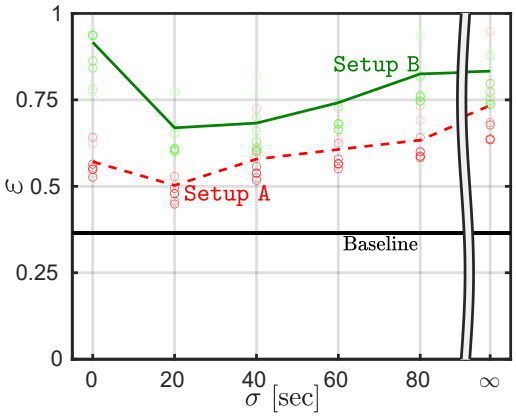

Figure 7. *Cloud (i).* The errors ε_t are marked by colored circles, whose saturation decays the farther the sampling time is from $(t_1 + t_{N^{\text{state}}})/2$. The measure ε is marked by solid or dashed lines, with corresponding colors. The setting $\sigma = \infty$ refers to the solution by the state of the art, i.e. 3D static scattering tomography.

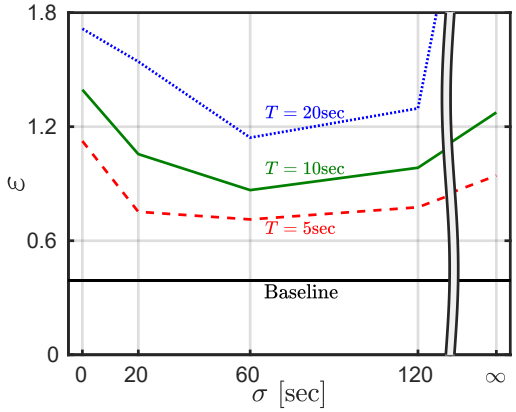

Figure 8. `Setup C`. The error ε of *Cloud (i)*, for different acquisition inter-angular temporal intervals, T. The setting $\sigma = \infty$ refers to the solution by the state of the art, i.e. 3D static scattering tomography.

simulated data). The LWC optimization was initialized by $\{\boldsymbol{\mathcal{L}}_t\}_{\tau \in \mathcal{T}} = 0.01\,\text{g/m}^3$. Convergence was reached in several dozen iterations. Depending on the number of input images, it took between minutes to a couple of hours to converge.

For result assessment, we generalize Eq. (14) to the whole sample set $t \in \mathcal{T}$ by

$$\varepsilon = \frac{1}{N^{\text{state}}} \sum_{t \in \mathcal{T}} \varepsilon_t \ , \ \text{where,} \ \varepsilon_t = \frac{\|\boldsymbol{\mathcal{L}}_t{}^{\text{true}} - \hat{\boldsymbol{\mathcal{L}}}_t\|_1}{\|\boldsymbol{\mathcal{L}}_t{}^{\text{true}}\|_1} \ . \quad (19)$$

From Sec. 2.4, we assess that a value $\sigma \sim 20\,\text{sec}$ is natural. Indeed, this is supported numerically in the plots of $\varepsilon_t, \varepsilon$ for *Cloud (i)* (Fig 7). A naive solution may only use measurements captured at each sampled time t, to solve $\mathcal{L}_t$, *independently* of other times. This solution is reached by $\sigma = 0$ and presented in Figs. 7 and 8. The 3D tomographic results of *Cloud (i)* at $t = (t_1 + t_{N^{\text{state}}})/2$ using `Setup A` are shown in Fig. 5. Recovery used $\sigma = 20\,\text{sec}$. In the Supplementary material we present analogous plots for *Cloud (ii)* and additional results using `Setup B`.

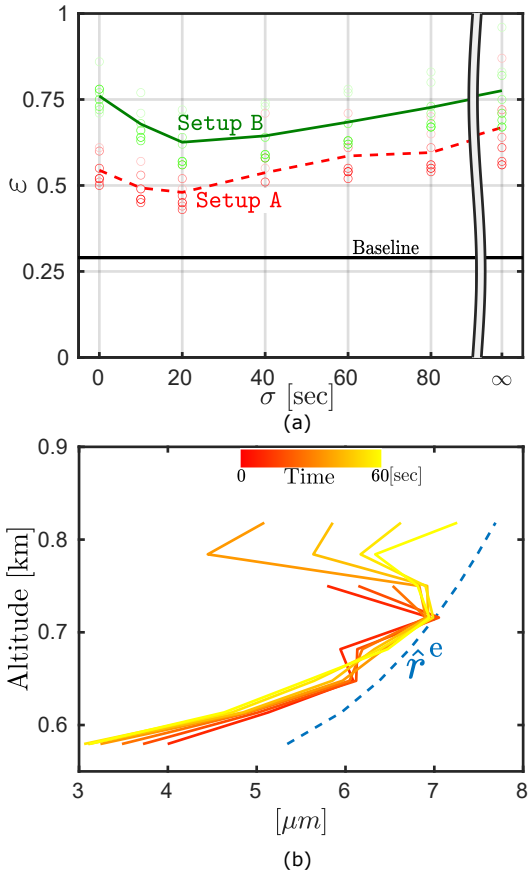

Figure 9. *Cloud (i)*, joint recovery of $r^{\text{e}}, \mathcal{L}$. (a) The errors ε_t of the LWC are marked by colored circles, whose saturation decays the farther the sampling time is from $(t_1 + t_{N^{\text{state}}})/2$. The measure ε is marked by solid or dashed lines, with corresponding colors. The setting $\sigma = \infty$ refers to solution by the state of the art, i.e. 3D static scattering tomography. (b) The estimated r^{e} is presented in blue dashed line. The true horizontally-averaged r_t^{e} are presented in solid lines, for different time samples.

`Setup C` uses a single platform, which is challenging. Results depend significantly on how fast the aircraft flies, i.e., how long it takes to capture the cloud from a variety of angles (up to 21 angles). Fig. 8 compares the results for inter-angle time interval of 5 sec, 10 sec and 20 sec. As expected, quality (ε) improves with velocity. Moreover, if the camera moves slowly (long time interval between angular samples), results improve by using a longer temporal support, observing the cloud from a wider angular range, despite its dynamics.

6.3. Microphysics Estimation

In this section, we recover both LWC and r^{e} as described in Sec. 5.2. We use `Setups A,B` and `Baseline`, as described in Sec. 6.1, with an additional spectral band at $\lambda = 865\,\text{nm}$. Fig. 9 shows the results for *Cloud (i)*.

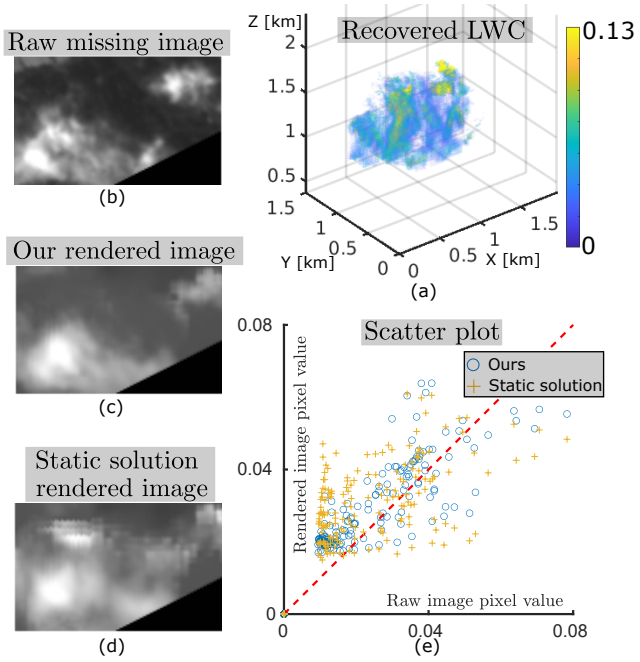

Figure 10. (a) Recovered 3D LWC field using real data. (b) A raw AirMSPI nadir image. Corresponding rendered views of a cloud, that was estimated using data that had excluded the nadir, either by our 4D CT approach (c) or current static 3D CT (d). Gamma correction was applied on (b,c,d) for display clarity. (e) A scatter plot of rendered vs. raw AirMSPI images at nadir. The scatter plot correlations of our solution and the static solution are 0.862 and 0.656, respectively.

7. Experiment: Real World AirMSPI Data

We follow the experimental approach of [29], and use real-world data acquired by JPL's AirMSPI, which flies on board NASA's ER-2. The geometry is exactly as described in `Setup C` in Sec. 6.1, including location and time. An atmospheric domain of size $1.5\,\mathrm{km} \times 2\,\mathrm{km} \times 2\,\mathrm{km}$ in the East-North-Up coordinates is examined. We discretized the domain to $80 \times 80 \times 80$ voxels. Because $N^{\mathrm{states}} = 21$, the total number of unknowns is 10,752,000.

The inter-angle time interval in this experiment is around 20 sec. Based on Fig. 8, we set here $\sigma = 60\,\mathrm{sec}$ in Eq. (18). We want to focus on dynamic tomography of the evolving cloud, and not on global motion due to wind in the cloud field. Hence, we used the pre-processing approach of [29] to align the cloud images. Additionally, the ground albedo is estimated to be 0.04. The pre-processing and albedo estimation are described in the Supplementary material.

A recovered volumetric reconstruction for one time instant is displayed in Fig. 10. We have no ground-truth for the cloud content in this case. Hence we check for consistency using cross-validation. For this, we excluded the nadir image (Fig. 10b) from the recovery process. Thus tomography used 20 out of the 21 raw views. Afterward, we placed the recovered cloud in SHDOM physics-based renderer [10], to generate the missing nadir view. The result is then compared to the ground-truth missing view. Fig. 10 compares the result of this process for two solutions: our 4D tomographic solution, and the state-of-the-art, i.e., 3D static scattering tomography.

	$+54°$ view	nadir view	$-54°$ view
Static solution	1.73	0.94	0.61
Ours	**0.96**	**0.38**	**0.24**

Table 1. Analysis of empirical data in different view angles. Quantitative fit (6) of our 4D result to the data, as compared to the error of state-of-the-art static 3D CT.

The same cross-validation process was repeated for the $\pm 54°$ view angles. Quantitatively, we measure the fitting error using Eq. (6). The results are summarized in Table 1.

8. Discussion

We derive a framework for 4D CT of dynamic objects that scatter, using moving cameras. The natural temporal evolution of an object indicates the temporal and angular sampling needed for a good reconstruction. Given these conditions, 4D CT recovery can be done, even with a small number of cameras. Fig. 7 and specifically `Setup A` may indicate that 4D CT may be achieved using $\sigma = 0$ and strong priors. This possibility should be a welcome topic for further computer vision research. In a sense, our work also uses a cloud prior, which is temporal correlation in clouds (Sec. 3): the correlation is analyzed using signal processing tools, and implemented by gradient weights w_t.

The model introduced in Sec. 5 to account for temporal correlations is independent of the differentiable renderer. Eqs. (16-18) can be combined with any differentiable renderer that can compute the derivative term $\partial \mathcal{F} / \partial \boldsymbol{\mathcal{L}}$ in Eq. (16). This can be either a differentiable renderer based on SHDOM, or a Monte Carlo differentiable renderer like those of [15, 39]. Additionally, some of our findings may be helpful in other fields. Bio-medical CT [7] and flow imaging [26] already have tools exploiting controlled illumination. Elements of this work may add to that toolkit.

Acknowledgements: We thank Aviad Levis and Tali Treibitz for their advice, and 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 project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (CloudCT, grant agreement No. 810370).

References

[1] David A Agard, Yasushi Hiraoka, Peter Shaw, and John W Sedat. Fluorescence microscopy in three dimensions. *Methods in Cell Biology*, 30:353–377, 1989. 4

[2] Amit Aides, Aviad Levis, Vadim Holodovsky, Yoav Y Schechner, Dietrich Althausen, and Adi Vainiger. Distributed sky imaging radiometry and tomography. In *Proc. IEEE ICCP*, pages 1–12, 2020. 1, 2

[3] Rushil Anirudh, Hyojin Kim, Jayaraman J Thiagarajan, K Aditya Mohan, Kyle Champley, and Timo Bremer. Lose the views: Limited angle ct reconstruction via implicit sinogram completion. In *Proc. IEEE CVPR*, pages 6343–6352, 2018. 1

[4] Stacey W Boland, David J Diner, John C Pearson, and Kevin A Burke. NASA's Multi-Angle Imager for Aerosols (MAIA) earth venture instrument investigation. *AGU Fall Meeting*, 2018:GH41C–1443, 2018. 1

[5] David J Diner, Michael J Garay, Olga V Kalashnikova, Brian E Rheingans, Sven Geier, Michael A Bull, Veljko M Jovanovic, Feng Xu, Carol J Bruegge, Anthony B Davis, et al. Airborne multiangle spectropolarimetric imager (AirMSPI) observations over California during NASA's polarimeter definition experiment (PODEX). In *Polarization Science and Remote Sensing VI*, volume 8873, page 88730B. SPIE, 2013. 6

[6] David J Diner, Feng Xu, Michael J Garay, John V Martonchik, Brian E Rheingans, Sven Geier, Anthony B Davis, BR Hancock, Michael A Jovanovic, Veljko M andBull, et al. The Airborne Multiangle Spectropolarimetric Imager (AirMSPI): a new tool for aerosol and cloud remote sensing. *Atmos. Meas. Tech.*, 6(8):2007, 2013. 1, 2, 6

[7] Turgut Durduran, Regine Choe, Wesley B Baker, and Arjun G Yodh. Diffuse optics for tissue monitoring and tomography. *Reports on Progress in Physics*, 73(7):076701, 2010. 8

[8] Marie L Eckert, Wolfgang Heidrich, and Nils Thuerey. Coupled fluid density and motion from single views. In *Computer Graphics Forum*, volume 37, pages 47–58. Wiley Online Library, 2018. 1

[9] K Franklin Evans. The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *Journal of the Atmospheric Sciences*, 55(3):429–446, 1998. 2

[10] K Franklin Evans and J Warren Wiscombe. Improvements to the SHDOM radiative transfer modeling package. In *Proc. 13th ARM Sci. Team Meeting*, 2003. 6, 8

[11] Jiwen Fan, Mikhail Ovtchinnikov, Jennifer M Comstock, Sally A McFarlane, and Alexander Khain. Ice formation in Arctic mixed-phase clouds: Insights from a 3-D cloud-resolving model with size-resolved aerosol and cloud microphysics. *JGR: Atmospheres*, 114(D4), 2009. 5

[12] Jeppe Revall Frisvad, Niels Jørgen Christensen, and Henrik Wann Jensen. Computing the scattering properties of participating media using Lorenz-Mie theory. In *ACM TOG*, pages 60–es. 2007. 2

[13] Ted T Fujita. Mesoscale classifications: their history and their application to forecasting. In *Mesoscale Meteorology and Forecasting*, pages 18–35. Springer, 1986. 1, 3

[14] Adam Geva, Yoav Y Schechner, Yonatan Chernyak, and Rajiv Gupta. X-ray computed tomography through scatter. In *Proc. ECCV*, pages 34–50, 2018. 1

[15] Ioannis Gkioulekas, Anat Levin, and Todd Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. In *Proc. ECCV*, pages 685–701. Springer, 2016. 1, 8

[16] James Gregson, Michael Krimerman, Matthias B Hullin, and Wolfgang Heidrich. Stochastic tomography and its applications in 3D imaging of mixing fluids. *ACM TOG*, 31(4):1–10, 2012. 1

[17] James E Hansen and Larry D Travis. Light scattering in planetary atmospheres. *Space Science Reviews*, 16(4):527–610, 1974. 2

[18] Thijs Heus, Harm JJ Jonker, Harry EA Van den Akker, Eric J Griffith, Michal Koutek, and Frits H Post. A statistical approach to the life cycle analysis of cumulus clouds selected in a virtual reality environment. *JGR: Atmospheres*, 114(D6), 2009. 5

[19] Harish P Hiriyannaiah. X-ray computed tomography for medical imaging. *IEEE Signal Processing Magazine*, 14(2):42–59, 1997. 1

[20] Vadim Holodovsky, Yoav Y Schechner, Anat Levin, Aviad Levis, and Amit Aides. In-situ multi-view multi-scattering stochastic tomography. In *Proc. IEEE ICCP*, pages 1–12, 2016. 4

[21] Wei-Chun Hsieh, Athanasios Nenes, Richard C Flagan, John H Seinfeld, G Buzorius, and H Jonsson. Parameterization of cloud droplet size distributions: Comparison with parcel models and observations. *Journal of Geophysical Research: Atmospheres*, 114(D11), 2009. 2, 5

[22] Anders P Kaestner, Beat Munch, and Pavel Trtik. Spatiotemporal computed tomography of dynamic processes. *Optical Engineering*, 50(12):123201, 2011. 1

[23] Alexander Khain, Andrei Pokrovsky, Mark Pinsky, Axel Seifert, and Vaughan Phillips. Simulation of effects of atmospheric aerosols on deep turbulent convective clouds using a spectral microphysics mixed-phase cumulus cloud model. Part I: Model description and possible applications. *JAS*, 61(24):2963–2982, 2004. 5

[24] Marat F Khairoutdinov and David A Randall. Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *JAS*, 60(4):607–625, 2003. 5

[25] Kiriakos N Kutulakos and Steven M Seitz. A theory of shape by space carving. *IJCV*, 38(3):199–218, 2000. 5

[26] Jonghwan Lee, Weicheng Wu, James Y Jiang, Bo Zhu, and David A Boas. Dynamic light scattering optical coherence tomography. *Optics Express*, 20(20):22262–22277, 2012. 8

[27] Aviad Levis, Daeyoung Lee, A. Joel Tropp, F. Charles Gammie, and L. Bouman Katherine. Inference of black hole fluid-dynamics from sparse interferometric measurements. In *Proc. IEEE ICCV*, 2021. 1

[28] Aviad Levis, Jesse Loveridge, and Amit Aides. *Pyshdom. 2020. Available online*. https://github.com/aviadlevis/pyshdom. 2

[29] Aviad Levis, Yoav Y Schechner, Amit Aides, and Anthony B Davis. Airborne three-dimensional cloud tomography. In *Proc. IEEE ICCV*, pages 3379–3387, 2015. 1, 2, 4, 8

[30] Aviad Levis, Yoav Y Schechner, and Anthony B Davis. Multiple-scattering microphysics tomography. In *Proc. IEEE CVPR*, pages 6740–6749, 2017. 1, 2, 4, 5

[31] Aviad Levis, Yoav Y Schechner, Anthony B Davis, and Jesse Loveridge. Multi-view polarimetric scattering cloud tomography and retrieval of droplet size. *Remote Sensing*, 12(17):2831, 2020. 2

[32] Tamar Loeub, Aviad Levis, Vadim Holodovsky, and Yoav Y Schechner. Monotonicity prior for cloud tomography. In *Proc. ECCV*, pages 24–29, 2020. 2, 4

[33] Fernando Macias-Garza, Kenneth R Diller, and Alan C Bovik. Missing cone of frequencies and low-pass distortion in three-dimensional microscopic images. *Optical Engineering*, 27(6):276461, 1988. 4

[34] Bernhard Mayer. Radiative transfer in the cloudy atmosphere. In *ERCA 2008 - From the Human Dimensions of Global Environmental Change to the Observation of the Earth from Space*, volume 1, pages 75–99. EDP Sciences, 2009. 2

[35] Brent A McBride, J Vanderlei Martins, Henrique MJ Barbosa, William Birmingham, and Lorraine A Remer. Spatial distribution of cloud droplet size properties from Airborne Hyper-Angular Rainbow Polarimeter (AirHARP) measurements. *AMT*, 13(4):1777–1796, 2020. 1

[36] Tomer Michaeli and Yonina C Eldar. High-rate interpolation of random signals from nonideal samples. *IEEE Transactions on Signal Processing*, 57(3):977–992, 2008. 3, 4

[37] Roel A J Neggers, Harm J J Jonker, and Pier Siebesma. Size statistics of cumulus cloud populations in large-eddy simulations. *JAS*, 60(8):1060–1074, 2003. 5

[38] Tim L Neilsen, Jose-Vanderlei Martins, RA Fernandez Borda, Cameron Weston, Crystal Frazier, Dominik Cieslak, and Kevin Townsend. The Hyper-Angular Rainbow Polarimeter (HARP) CubeSat observatory and the characterization of cloud properties. In *AGU Fall Meeting Abstracts*, volume 2015, pages A43A–0237, 2015. 1

[39] Merlin Nimier-David, Delio Vicini, Tizian Zeltner, and Wenzel Jakob. Mitsuba 2: A retargetable forward and inverse renderer. *ACM TOG*, 38(6):1–17, 2019. 2, 8

[40] Alan V Oppenheim. *Discrete-Time Signal Processing*. Pearson Education India, 1999. 3

[41] Tinsu Pan, Ting-Yim Lee, Eike Rietzel, and George TY Chen. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Medical Physics*, 31(2):333–340, 2004. 1

[42] Yiming Qian, Minglun Gong, and Yee-Hong Yang. Stereo-based 3D reconstruction of dynamic fluid surfaces by global optimization. In *Proc. IEEE CVPR*, pages 1269–1278, 2017. 1

[43] Klaus Schilling, Yoav Y Schechner, and Ilan Koren. CloudCT - computed tomography of clouds by a small satellite formation. In *Proc. IAA Symposium on Small Satellites for Earth Observation*, 2019. 1

[44] Yael Sde-Chen, Yoav Y Schechner, Vadim Holodovsky, and Eshkol Eytan. 3DeepCT: Learning volumetric scattering tomography of clouds. In *Proc. IEEE ICCV*, 2021. 2

[45] A Pier Siebesma, Christopher S Bretherton, Andrew Brown, Andreas Chlond, Joan Cuxart, Peter G Duynkerke, Hongli Jiang, Marat Khairoutdinov, David Lewellen, Chin-Hoh Moeng, et al. A large eddy simulation intercomparison study of shallow cumulus convection. *JAS*, 60(10):1201–1219, 2003. 5

[46] Masada Tzabari, Vadim Holodovsky, Omer Shubi, Eshkol Eitan, Orit Altaratz, Ilan Koren, Anna Aumann, Klaus Schilling, and Yoav Y Schechner. CloudCT 3D volumetric tomography: considerations for imager preference, comparing visible light, short-wave infrared, and polarized imagers. In *Polarization Science and Remote Sensing X*, volume 11833, page 1183304. International Society for Optics and Photonics, 2021. 1

[47] Masada Tzabari, Vadim Holodovsky, Omer Shubi, Eshkol Eytan, and Yoav Y Schechner. Advances in 3D scattering tomography of cloud micro-physics. *arXiv preprint arXiv:2012.03223*, 2021. 2, 5

[48] Gerard Van Harten, David J Diner, Brian JS Daugherty, Brian E Rheingans, Michael A Bull, Felix C Seidel, Russell A Chipman, Brian Cairns, Andrzej P Wasilewski, and Kirk D Knobelspiesse. Calibration and validation of airborne multiangle spectropolarimetric imager (AirMSPI) polarization measurements. *Applied Optics*, 57(16):4499–4513, 2018. 2, 6

[49] Tommy E Wright, Mike Burton, David M Pyle, and Tommaso Caltabiano. Scanning tomography of SO2 distribution in a volcanic gas plume. *Geophysical Research Letters*, 35(17), 2008. 1

[50] Huiwen Xue and Graham Feingold. Large-eddy simulations of trade wind cumuli: Investigation of aerosol indirect effects. *JAS*, 63(6):1605–1622, 2006. 5

[51] Guangming Zang, Ramzi Idoughi, Ran Tao, Gilles Lubineau, Peter Wonka, and Wolfgang Heidrich. Space-time tomography for continuously deforming objects. *ACM TOG*, 37, 2018. 1

[52] Guangming Zang, Ramzi Idoughi, Ran Tao, Gilles Lubineau, Peter Wonka, and Wolfgang Heidrich. Warp-and-project tomography for rapidly deforming objects. *ACM TOG*, 38(4):1–13, 2019. 1

[53] Guangming Zang, Ramzi Idoughi, Congli Wang, Anthony Bennett, Jianguo Du, Scott Skeen, William L Roberts, Peter Wonka, and Wolfgang Heidrich. Tomofluid: Reconstructing dynamic fluid from sparse view videos. In *Proc. IEEE CVPR*, pages 1870–1879, 2020. 1

[54] S Zhang, H Xue, and G Feingold. Vertical profiles of droplet effective radius in shallow convective clouds. *Atmospheric Chemistry and Physics*, 11(10):4633–4644, 2011. 2

[55] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM TOMS*, 23(4):550–560, 1997. 5

4D Cloud Scattering Tomography
Supplementary Material

Roi Ronen, Yoav Y. Schechner
Viterbi Faculty of Electrical & Computer Eng.
Technion - Israel Institute of Technology
Haifa, Israel
ronen.roi@gmail.com
yoav@ee.technion.ac.il

Eshkol Eytan
Department of Earth & Planetary Sciences
The Weizmann Institute of Science
Rehovot, Israel
eshkol.eytan@weizmann.ac.il

Abstract

This is a supplementary document to the main manuscript. Here we provide more numerical results. Moreover, this document details pre-processing of real world data, which is presented in section 7 of the main manuscript. Additionally, we elaborate about the computational complexity of the method, give evaluation for the Gaussian weights that were used in the main manuscript and an interpretation using a hidden field for the iterative procedure described in section 5.1 of the main manuscript.

1. Outline

This supplementary material contains five parts. The first part (Sec. 2) elaborates on pre-processing which is applied to real world measurements, presented in Sec. 7 of the main manuscript. This data was collected by the AirMSPI instrument. The second part, Sec. 3, provides an additional example of the temporal auto-correlation of cloud microphysics and more simulation results which were not included in the main manuscript, for space limits. The third part (Sec. 4) analyzes the computational complexity of our proposed method. In Sec. 5, we give an interpretation for the iterative procedure described in Sec. 5.1 of the main manuscript using a hidden field representation. Sec. 6 provides an evaluation for the Gaussian weights that were used in Sec. 5.1 of the main manuscript.

2. Pre-processing Real World Data

The main manuscript presents results using real world measurements. The data were acquired by the AirMSPI instrument. As explained in Sec. 7 of the main manuscript, while AirMSPI flies, clouds move due to wide-scale wind at their altitude. The geometry of AirMSPI's path and the cloud drift during the experiment is presented in Fig. 1 above. In order to eliminate the influence of wide-scale

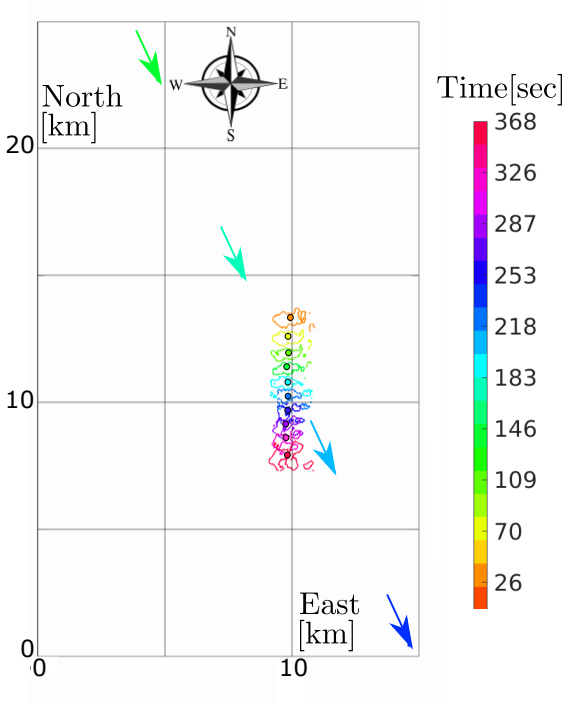

Figure 1. Geometry of the AirMSPI real world setup which led to the data presented in Sec. 7 of the main manuscript. The color represents the locations of the cloud and the AirMSPI instrument in the different time states. The cloud's outer contour and its corresponding center of mass, marked in a circle, are presented per state. The AirMSPI location and velocity are marked by arrows. The arrows point to the AirMSPI flight direction azimuth of 154° relative to the North. Due to the domain size, not all AirMSPI locations are illustrated here. Due to wind, the cloud moves at $57\,\mathrm{km/h}$ in azimuth 182° relative to the North.

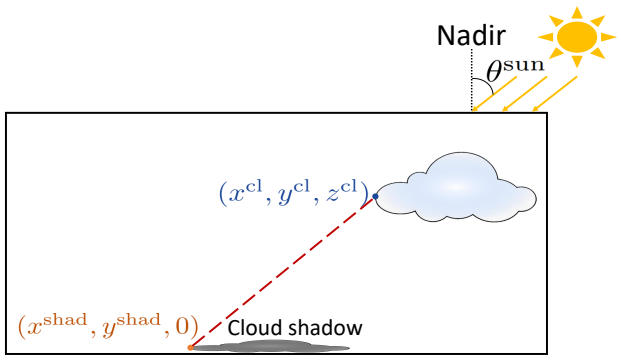

Figure 2. Illustration of estimation of the cloud altitude using a shadow.

wind, a registration process of the cloud images is done. Moreover, for tomographic recovery, we need to have an assessment of the Earth surface albedo, under the clouds. This section describes how pre-processing estimates the wind and albedo.

2.1. Wind Estimation

Clouds are segmented from the surface automatically [12]. Cloudy pixels are then used to estimate the cloud center of mass in each image [7]. A registration of these centers of mass can be done by triangulation. However, triangulation of images of a moving object using a translating camera has an inherent ambiguity. This ambiguity can be solved if the cloud height is known. In this work, we assess this altitude of a cloud by its shadow [1, 6, 8]. Let $(x^{\mathrm{cl}}, y^{\mathrm{cl}}, z^{\mathrm{cl}})$ and $(x^{\mathrm{shad}}, y^{\mathrm{shad}}, 0)$ be a point on a cloud and its corresponding shadow point on the earth surface, respectively (see Fig. 2 above). Let $\varrho^{\mathrm{shad}} = \sqrt{(x^{\mathrm{cl}} - x^{\mathrm{shad}})^2 + (y^{\mathrm{cl}} - y^{\mathrm{shad}})^2}$. We obtain $x^{\mathrm{cl}}, y^{\mathrm{cl}}, x^{\mathrm{shad}}$ and y^{shad} from the AirMSPI images. Given the solar zenith angle relative to the nadir θ^{sun}, the altitude z^{cl} satisfies

$$z^{\mathrm{cl}} = \frac{\varrho^{\mathrm{shad}}}{\tan(\theta^{\mathrm{sun}})} \ . \tag{1}$$

For the example shown in Sec. 7 of the main manuscript, we estimated the cloud base height as $\approx 500\,\mathrm{m}$ and its top at $\approx 1100\,\mathrm{m}$. Indeed taking MODIS/AQUA [2] retrievals of cloud top heights, indicate that the clouds' top in the region[1] does not exceed 1000 m, which makes our approximation reasonable.

We approximate the cloud horizontal velocity by back-projecting the images from the locations of the cameras to the altitude of z^{cl}. From the center of mass of these back-projections, we assess the velocity. We register the camera locations so the projections of the center of mass of all images intersect at the same point at altitude of z^{cl}. The images and the new locations of the camera are the input for the 4D tomographic recovery.

[1]This data applies over the coast of California, 38N 122W,on Feb/03/2013 at 13:30 local time.

2.2. Surface Albedo Estimation

3D radiative transfer calculations require the surface albedo. We use non-cloudy pixels to estimate the albedo. Let $\mathcal{Y}$ be a set of non-cloudy pixels. We estimate the surface albedo a^{ground} as,

$$\hat{a}^{\mathrm{ground}} = \underset{a}{\mathrm{argmin}} \sum_{y \in \mathcal{Y}} ||y - \mathcal{F}(\boldsymbol{\beta}^{\mathrm{air}}; a)||_2^2 \ , \tag{2}$$

where $\boldsymbol{\beta}^{\mathrm{air}}$ represents the extinction coefficient of air in 3D with no clouds. Here $\mathcal{F}(\boldsymbol{\beta}^{\mathrm{air}}; a)$ is a rendering (forward) model, where the surface albedo is set to be a. That is, sunlight interacts only with the air and the surface. Scattering by air is assumed to be known [5, 13]. The optimization problem is solved by the Brent minimization method [3], implemented by the SciPy package [11]. For the example shown in Sec. 7 of the main manuscript, the surface albedo is estimated to be 0.04.

3. Additional Simulations

3.1. Cloud Temporal Spectrum

Sec. 3 of the main manuscript indicates that the correlation time of a convective cloud at 10 sec resolution is about 20 to 50 seconds. Thus, a temporal sampling period of 30 sec or shorter is required. We assess this in an additional cloud simulation. We conducted a single cloud simulation in high resolution, using small changes, relative to the simulation described in Sec. 6 of the main manuscript. The simulation parameters and setting are similar. However, the perturbation that initiates the convection and turbulent flow has a smaller horizontal size. This creates a smaller cloud with a horizontal width of ≈400 m. This cloud is more sensitive to mixing and evaporation than the cloud in the main manuscript whose width is ≈800 m. Because mixing with the environment is more intense here, the clouds' growth is inhibited. It cannot exceed a height of 1400 m, compared to a 2000 m ceiling of the cloud in the main manuscript.

Using the same process described in Sec. 3 of the main manuscript, the temporal auto-correlation functions of $\mathcal{L}_t$ and horizontally-averaged r_t^{e} are presented in Fig. 3[Top] in the next page. The auto-correlation function of r_t^{e} behaves similarly to that of v_t^{e}. Here the required temporal sampling period is more tolerable compared to the presented temporal sampling period in the main manuscript.

3.2. Additional Tomography Results

Recall that our method is demonstrated on two simulated clouds, *Cloud (i)* and *Cloud (ii)*, using several types of imaging setups: `Setup A`, `Setup B` and `Baseline`.

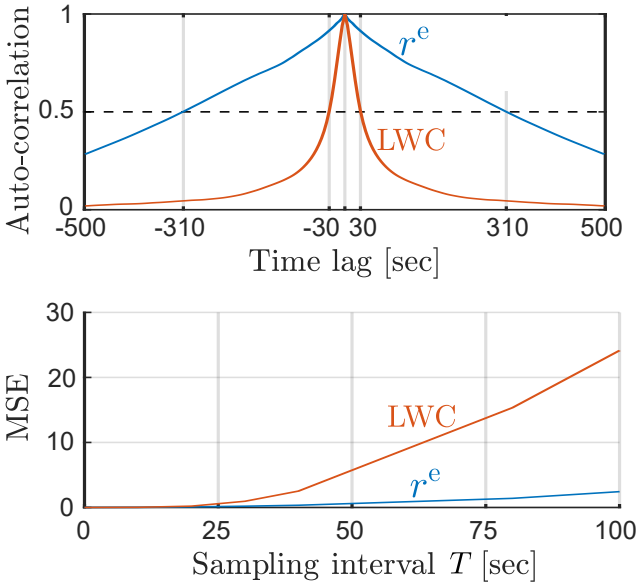

Figure 3. [Top] Auto-correlation of cloud field microphysics. The auto-correlation of LWC and r^e decreases to 0.5 after 30 sec and 310 sec, respectively. [Bottom] MSE (Eq. 13 in the main manuscript) of LWC and r^e.

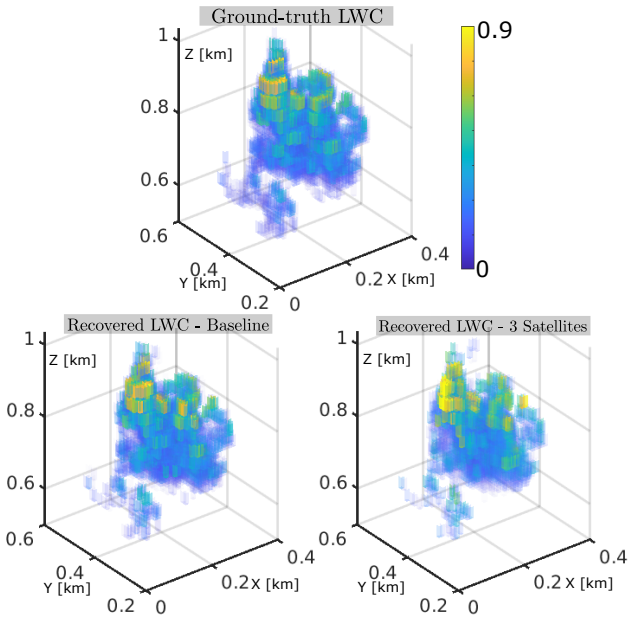

Figure 4. *Cloud (ii)*. Results of recovery by the `Baseline` and `Setup A` are compared to the ground-truth.

Fig. 4 above presents the 3D tomographic results of *Cloud (ii)* at $t = (t_1 + t_{N^{\text{state}}})/2$ using `Setup A`. The recovery used $\sigma = 20$ sec. Moreover, recall the error measures as Eq. (19) defined in the main manuscript. Fig. 5 above presents $\varepsilon_t, \varepsilon$ for *Cloud (ii)*. It reinforces the assessment that a value $\sigma \sim 20$ sec is natural, as explained in Sec. 3 of the main manuscript.

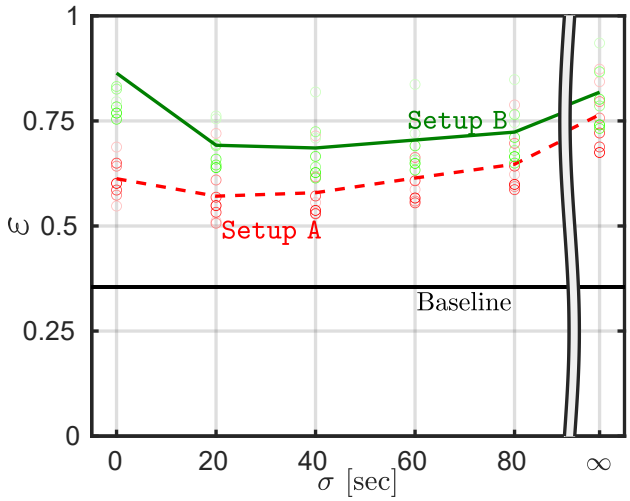

Figure 5. *Cloud (ii)*. The errors ε_t in the main manuscript are marked by colored circles, whose saturation decays the farther the sampling time is from $(t_1 + t_{N^{\text{state}}})/2$. The measure ε in the main manuscript is marked by solid or dashed lines, with corresponding colors. The setting $\sigma = \infty$ refers to the solution by the state of the art, i.e. 3D static scattering tomography.

	$\sigma = 20$ sec	$\sigma = 60$ sec	$\sigma = \infty$
`Setup A`	**0.45**	0.55	0.64
`Setup B`	**0.6**	0.66	0.74
`Setup C`	0.79	**0.74**	0.9

Table 1. The error ε_t in the main manuscript at $t = (t_1 + t_{N^{\text{state}}})/2$ of *Cloud (i)*, when $T = 10$ sec. The state of the art 3D static CT complies with $\sigma = \infty$.

Figs. 6 and 7 in the next page respectively visualize the results of *Cloud (i)* and *Cloud (ii)*. The 3D cut-sections of the error $|\boldsymbol{\mathcal{L}}_t^{\text{true}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{L}}}_t(\mathbf{x})|$ at $t = (t_1 + t_{N^{\text{state}}})/2$ are presented for `Setup A`, `Setup B` and `Baseline` in Figs. 6 and 7[Top]. Fig. 7[Bottom] uses scatter plots to compare the ground-truth to the results obtained by either the `Baseline`, `Setup A` or `Setup B`. Also, we compare the recovery results at $t = (t_1 + t_{N^{\text{state}}})/2$ quantitatively for the three setups in Table 1 above.

4. Computational Complexity

Sec. 5.1 of the main manuscript introduces an iterative procedure for 4D CT estimation of cloud LWC

$$\boldsymbol{\mathcal{L}}_t(k+1) = \boldsymbol{\mathcal{L}}_t(k) - \eta \mathbf{g}_t[\mathcal{B}(k)] \,, \tag{3}$$

where

$$\mathbf{g}_t(\mathcal{B}) = \sum_{t' \in \mathcal{T}} w_t(t'|\sigma) \frac{\partial \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial \boldsymbol{\mathcal{L}}_{t'}} \left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] \,. \tag{4}$$

The time complexity for solving Eqs. (3,4) above is governed by the gradient calculation $\mathbf{g}_t(\mathcal{B})$. Computing the Ja-

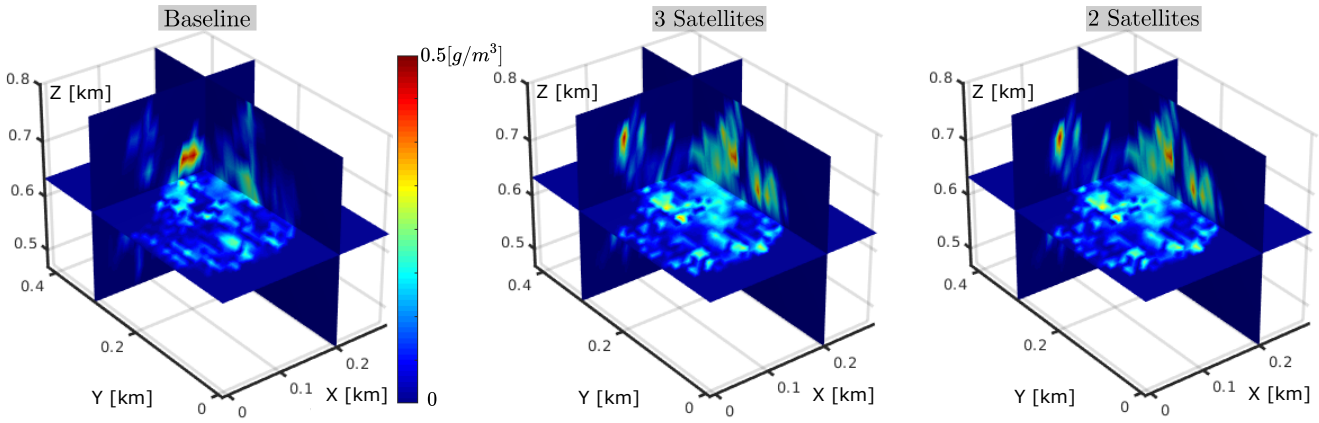

Figure 6. *Cloud (i).* 3D cut-sections of the error $|\boldsymbol{\mathcal{L}}_t^{\text{true}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{L}}}_t(\mathbf{x})|$ at $t = (t_1 + t_{N^{\text{state}}})/2$ for `Baseline`, `Setup A` and `Setup B`.

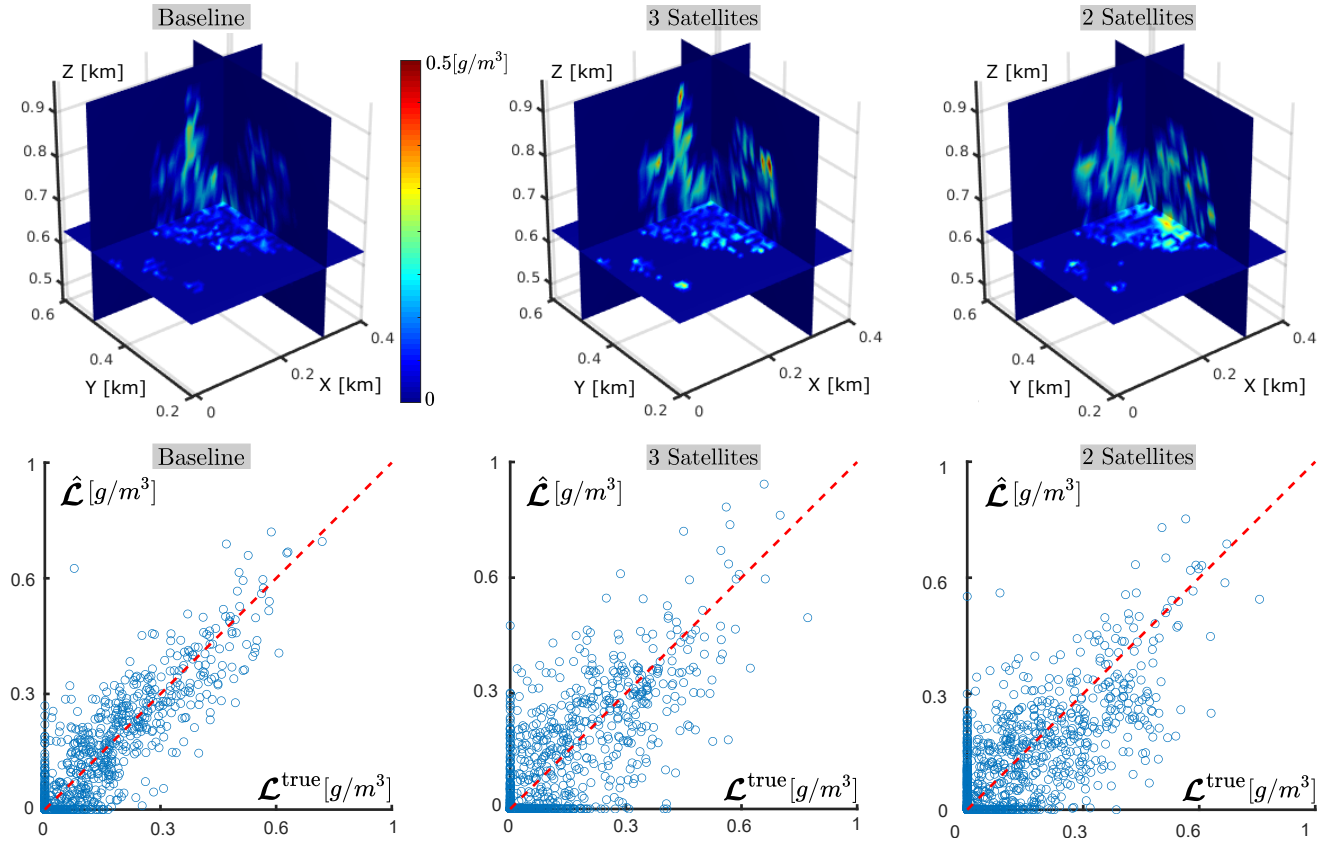

Figure 7. *Cloud (ii)* comparison for `Baseline`, `Setup A` and `Setup B`. [Top] 3D cut-sections of the error $|\boldsymbol{\mathcal{L}}_t^{\text{true}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{L}}}_t(\mathbf{x})|$ at $t = (t_1 + t_{N^{\text{state}}})/2$. [Bottom] Scatter plots that use randomly selected 20% of the data points, for display clarity. The `Baseline`, `Setup A` and `Setup B` scatter plot correlations are 0.92, 0.82 and 0.78, respectively.

cobian $\partial \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) / \partial \boldsymbol{\mathcal{L}}_{t'}$ in Eq. (4) is complex, thus it is established numerically by a surrogate function that evolves through iterations [7, 9]. Calculating the gradient includes two dominant time-consuming processes that are executed in alternation. The first process calculates the forward model for the N^{state} cloud states $\{\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)\}_{t' \in \mathcal{T}}$. The second process sums over the entire set of measurements, which does not depend on the number of cloud states that we seek to recover.

A spherical harmonic discrete ordinate method (SHDOM) code is used for computing the numerical forward model $\mathcal{F}(\cdot)$ and the Jacobian. SHDOM iteratively

updates the estimation of 3D radiation fields until convergence. Calculating the forward model for the N^{state} cloud states can be done in parallel. Thus, the time complexity is governed by the temporal state, for which the SHDOM forward model code takes the longest time to compute. By calculating the forward model for all cloud states in parallel, the time complexity of gradient calculation is insensitive to the number of cloud states N^{state}. Algorithm 1 is a pseudo-code of our algorithm. The source code is publicly available at [10].

Algorithm 1 4D Cloud Scattering Tomography

Require: $\{\boldsymbol{y}_{t'}\}_{t'\in\mathcal{T}}$ and $\sigma \geq 0$
  $\mathcal{B}(0) = \{\boldsymbol{\mathcal{L}}_{t'} = 0.01\}_{t'\in\mathcal{T}}$
  $k = 0$
  repeat
    Calculate $\{\mathcal{F}\left[\boldsymbol{\mathcal{L}}_{t'}(k)\right]\}_{t'\in\mathcal{T}}$ ▷ In parallel
    Approximate $\left\{\frac{\partial\mathcal{F}}{\partial\boldsymbol{\mathcal{L}}_{t'}}\left[\boldsymbol{\mathcal{L}}_{t'}(k)\right]\right\}_{t'\in\mathcal{T}}$ ▷ Ref. [7]
    Calculate $\{\mathbf{g}_{t'}[\mathcal{B}(k)]\}_{t'\in\mathcal{T}}$ ▷ Eq. (4)
    $\forall t \in \mathcal{T}$ update $\boldsymbol{\mathcal{L}}_t(k+1)$ ▷ Eq. (3)
    $k = k + 1$
  until converge $\{\boldsymbol{y}_{t'} \approx \mathcal{F}\left[\boldsymbol{\mathcal{L}}_{t'}(k)\right]\}_{t'\in\mathcal{T}}$

As a numerical example, we used 20 iterations of the L-BFGS-B optimization. Using measurements of *Cloud (i)* acquired by `Setup A`, the run-time of the solution by our method was 501 sec. The static solution took 301 sec. In both, the computer was Intel® Xeon® Gold 6240 CPU @ 2.60GHz with 72 cores. Although our method recovers $N^{\text{state}} = 7$ times more voxels, the run-time is less than twice that of the static solution. The time difference is caused by overheads of saving and loading larger data with our method, and nonoptimal task division for the cores.

5. Cost on a Hidden Field

In this section, we present an interpretation for the iterative procedure of solving the problem of 4D CT estimation of cloud LWC (Eqs. 3,4 above). Recall that Eq. (18) in the main manuscript defines $w_t(t'|\sigma)$ as Gaussian weights

$$w_t(t'|\sigma) = s\exp\left(-\frac{|t-t'|^2}{2\sigma^2}\right) , \tag{5}$$

where s is a normalization factor. Let $w_t(t'|\tilde{\sigma})$ be Gaussian weights with variance $\tilde{\sigma}^2$. Suppose the cloud LWC can be represented by

$$\boldsymbol{\mathcal{L}}_t\left[\mathcal{B}^{\text{hidden}}\right] = \sum_{t'\in\mathcal{T}} w_t(t'|\tilde{\sigma})\boldsymbol{\mathcal{L}}_{t'}^{\text{hidden}} , \tag{6}$$

where $\boldsymbol{\mathcal{L}}_t^{\text{hidden}}$ is a hidden representation at time t and $\mathcal{B}^{\text{hidden}} = \{\boldsymbol{\mathcal{L}}_{t'}^{\text{hidden}}\}_{t'\in\mathcal{T}}$. The set $\mathcal{B}^{\text{hidden}}$ is equivalent to the set $\mathcal{B} = \{\boldsymbol{\mathcal{L}}_t\}_{\forall t}$ through a linear transformation having Gaussian weights. Let us formulate 4D CT using the hidden field representation

$$\hat{\mathcal{B}}^{\text{hidden}} = \underset{\mathcal{B}^{\text{hidden}}}{\text{argmin}} \sum_{t\in\mathcal{T}} \mathcal{E}\left[\boldsymbol{y}_t, \mathcal{F}\left(\mathcal{B}^{\text{hidden}}\right)\right] . \tag{7}$$

Recall that $\boldsymbol{y}_t$, the measurements acquired at time t, depends explicitly only on the cloud state at this time, $\boldsymbol{\mathcal{L}}_t$. Thus,

$$\mathcal{E}\left[\boldsymbol{y}_t, \mathcal{F}\left(\mathcal{B}^{\text{hidden}}\right)\right] = \frac{1}{2}\|\boldsymbol{y}_t - \mathcal{F}\left(\boldsymbol{\mathcal{L}}_t[\mathcal{B}^{\text{hidden}}]\right)\|_2^2 . \tag{8}$$

Eq. (7) above can be solved efficiently by gradient-based methods. The gradient of Eq. (7) above is

$$\frac{\partial}{\partial\boldsymbol{\mathcal{L}}_t^{\text{hidden}}}\sum_{t'\in\mathcal{T}}\mathcal{E}\left[\boldsymbol{y}_t, \mathcal{F}\left(\mathcal{B}^{\text{hidden}}\right)\right] = \sum_{t'\in\mathcal{T}}\frac{\partial\mathcal{E}\left[\boldsymbol{y}_{t'}, \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)\right]}{\partial\boldsymbol{\mathcal{L}}_{t'}}\frac{\partial\boldsymbol{\mathcal{L}}_{t'}}{\partial\boldsymbol{\mathcal{L}}_t^{\text{hidden}}} . \tag{9}$$

From Eq. (8) above,

$$\frac{\partial\mathcal{E}\left[\boldsymbol{y}_{t'}, \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)\right]}{\partial\boldsymbol{\mathcal{L}}_{t'}} = \frac{\partial\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial\boldsymbol{\mathcal{L}}_{t'}}\left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] , \tag{10}$$

while from Eq. (6) above,

$$\frac{\partial\boldsymbol{\mathcal{L}}_{t'}}{\partial\boldsymbol{\mathcal{L}}_t^{\text{hidden}}} = w_t(t'|\tilde{\sigma}) . \tag{11}$$

From Eqs. (9,10,11) above, for optimizing problem (7) above, the gradient is

$$\mathbf{g}_t^{\text{hidden}}(\mathcal{B}) = \sum_{t'\in\mathcal{T}} w_t(t'|\tilde{\sigma})\frac{\partial\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial\boldsymbol{\mathcal{L}}_{t'}}\left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] . \tag{12}$$

A gradient-based approach then performs per iteration k:

$$\boldsymbol{\mathcal{L}}_t^{\text{hidden}}(k+1) = \boldsymbol{\mathcal{L}}_t^{\text{hidden}}(k) - \eta\mathbf{g}_t^{\text{hidden}}[\mathcal{B}(k)] \tag{13}$$

where η is a step size. Every iteration, $\mathcal{B}(k)$ is updated by Eq. (6) above,

$$\boldsymbol{\mathcal{L}}_t(k+1) = \sum_{t'\in\mathcal{T}} w_t(t'|\tilde{\sigma})\boldsymbol{\mathcal{L}}_{t'}^{\text{hidden}}(k+1) . \tag{14}$$

Substitute Eq. (13) above into Eq. (14) above,

$$\boldsymbol{\mathcal{L}}_t(k+1) = \boldsymbol{\mathcal{L}}_t(k) - \eta\sum_{t'\in\mathcal{T}} w_t(t'|\tilde{\sigma})\mathbf{g}_{t'}^{\text{hidden}}[\mathcal{B}(k)] . \tag{15}$$

Now, we use the approximation

$$w_t(t'|\sigma) \approx \sum_{t''\in\mathcal{T}} w_t(t''|\tilde{\sigma})w_{t'}(t''|\tilde{\sigma}) , \tag{16}$$

where

$$\sigma = \sqrt{2}\tilde{\sigma} \ . \tag{17}$$

We explain the approximation in Eq. (16) above using properties of continuous Gaussian PDFs, as we now explain. Let $f_\mu(\xi|\tilde{\sigma})$ and $f_{\mu'}(\xi|\tilde{\sigma})$ be two Gaussian PDFs with variance $\tilde{\sigma}^2$ and respective expectations μ and μ'. Being a PDF,

$$\int_{-\infty}^{\infty} f_{(\mu+\mu')/2}(\xi|\tilde{\sigma}/\sqrt{2})\, d\xi = 1 \ . \tag{18}$$

From [4],

$$f_\mu(\xi|\tilde{\sigma}) f_{\mu'}(\xi|\tilde{\sigma}) = f_\mu(\mu'|\sqrt{2}\tilde{\sigma}) f_{(\mu+\mu')/2}(\xi|\tilde{\sigma}/\sqrt{2}) \ . \tag{19}$$

From Eq. (18) above,

$$\begin{aligned} f_\mu(\mu'|\sqrt{2}\tilde{\sigma}) &= f_\mu(\mu'|\sqrt{2}\tilde{\sigma}) \int_{-\infty}^{\infty} f_{(\mu+\mu')/2}(\xi|\tilde{\sigma}/\sqrt{2})\, d\xi \\ &= \int_{-\infty}^{\infty} f_\mu(\mu'|\sqrt{2}\tilde{\sigma}) f_{(\mu+\mu')/2}(\xi|\tilde{\sigma}/\sqrt{2})\, d\xi \ . \end{aligned} \tag{20}$$

From Eqs. (19,20) above,

$$f_\mu(\mu'|\sqrt{2}\tilde{\sigma}) = \int_{-\infty}^{\infty} f_\mu(\xi|\tilde{\sigma}) f_{\mu'}(\xi|\tilde{\sigma})\, d\xi \ . \tag{21}$$

Discretizing the integral of Eq. (21) above, an approximate finite sum yields Eq. (16) above.

Substitute Eq. (16) above into Eq. (4) above

$$\mathbf{g}_t(\mathcal{B}) \approx \sum_{t' \in \mathcal{T}} \sum_{t'' \in \mathcal{T}} w_t(t''|\tilde{\sigma}) w_{t'}(t''|\tilde{\sigma}) \frac{\partial \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial \boldsymbol{\mathcal{L}}_{t'}} \left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] \ . \tag{22}$$

Swap the summation order of Eq. (22) above:

$$\mathbf{g}_t(\mathcal{B}) \approx \sum_{t'' \in \mathcal{T}} w_t(t''|\tilde{\sigma}) \sum_{t' \in \mathcal{T}} w_{t'}(t''|\tilde{\sigma}) \frac{\partial \mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right)}{\partial \boldsymbol{\mathcal{L}}_{t'}} \left[\mathcal{F}\left(\boldsymbol{\mathcal{L}}_{t'}\right) - \boldsymbol{y}_{t'}\right] \ . \tag{23}$$

Substitute Eq. (12) above into Eq. (23) above and use the property that $w_{t'}(t''|\tilde{\sigma}) = w_{t''}(t'|\tilde{\sigma})$. This yields,

$$\mathbf{g}_t(\mathcal{B}) \approx \sum_{t'' \in \mathcal{T}} w_t(t''|\tilde{\sigma}) \mathbf{g}_{t''}^{\text{hidden}}(\mathcal{B}) \ . \tag{24}$$

Substituting Eq. (24) above into Eq. (15) above yields the iteration move in Eq. (3) above

$$\begin{aligned} \boldsymbol{\mathcal{L}}_t(k+1) &\approx \boldsymbol{\mathcal{L}}_t(k) - \eta \sum_{t' \in \mathcal{T}} w_t(t'|\tilde{\sigma}) \mathbf{g}_{t'}^{\text{hidden}}[\mathcal{B}(k)] \\ &= \boldsymbol{\mathcal{L}}_t(k) - \eta \mathbf{g}_t[\mathcal{B}(k)] \ . \end{aligned} \tag{25}$$

Thus, the iterative procedure for 4D CT estimation in the main manuscript (Eqs. 3,4 above) can be interpreted as solving Eq. (7) above.

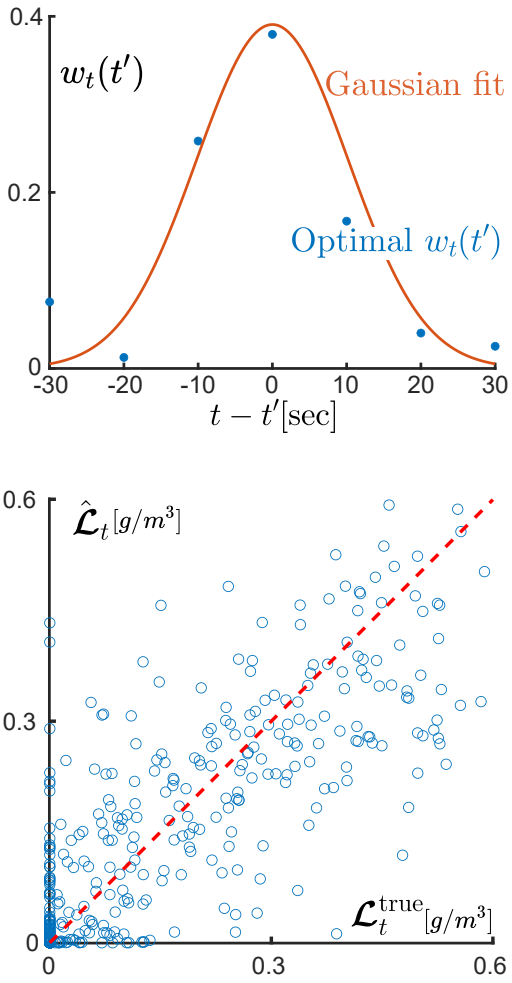

Figure 8. [Top] Optimal $w_t(t')$ are presented in blue circles and their Gaussian fit in red line. [Bottom] Comparison of true $\boldsymbol{\mathcal{L}}_t$ and $\hat{\boldsymbol{\mathcal{L}}}_t$ by a scatter plot.

6. Kernel Assessment for Cloud Tomography

Recall Eq. (11) in the main manuscript,

$$\boldsymbol{\beta}_t \sim \sum_{t'} w_t(t'|T) \boldsymbol{\beta}_{t'}^{\text{sample}} \ . \tag{26}$$

We now assess the approximation of using cropped Gaussian as the kernel $w_t(t'|T)$ for recovering cloud LWC. In tomography, we do not have direct sampling of the $\boldsymbol{\mathcal{L}}_{t'}^{\text{sample}}$ at time $t' \in \mathcal{T}$. We only have projected images. Let $\tilde{\boldsymbol{\mathcal{L}}}_{t'}$ be the tomographic recovery of the LWC using only measurements acquired at time t'. We approximate $\boldsymbol{\mathcal{L}}_{t'}^{\text{sample}} \approx \tilde{\boldsymbol{\mathcal{L}}}_{t'}$. Hence instead of Eq. (26) above, suppose we approximate $\boldsymbol{\mathcal{L}}_t$ as

$$\hat{\boldsymbol{\mathcal{L}}}_t(\{w_t(t')\}_{t' \in \mathcal{T}}) \approx \sum_{t' \in \mathcal{T}} w_t(t') \tilde{\boldsymbol{\mathcal{L}}}_{t'} \ . \tag{27}$$

Let us seek the optimal set of weights $\{w_t(t')\}_{t' \in \mathcal{T}}$ by

$$\{\hat{w}_t(t')\}_{t' \in \mathcal{T}} = \underset{\{w_t(t')\}_{t' \in \mathcal{T}}}{\operatorname{argmin}} ||\boldsymbol{\mathcal{L}}_t - \tilde{\boldsymbol{\mathcal{L}}}_t(\{w_t(t')\}_{t' \in \mathcal{T}})||_2^2 \ . \quad (28)$$

Fig. 8[Top] herein shows the optimal $w_t(t')$. The plot shows that the weights are approximately Gaussian. Fig. 8[Bottom] herein shows a scatter plot of $\hat{\boldsymbol{\mathcal{L}}}_t$ vs. true values $\boldsymbol{\mathcal{L}}_t^{\text{true}}$ for *Cloud (i)*, in `Setup A`. Fig. 8[Bottom] indicates that $\hat{\boldsymbol{\mathcal{L}}}_t(\{w_t\}_{t \in \mathcal{T}})$ based on this Gaussian-weight set yields a good approximation of $\boldsymbol{\mathcal{L}}_t$. These results support the use of Gaussian weights for 4D CT of cloud LWC (Eqs. 3,4 above).

References

[1] Austin Abrams, Kylia Miskell, and Robert Pless. The episolar constraint: Monocular shape from shadow correspondence. In *Proc. IEEE CVPR*, pages 1407–1414, 2013. 2

[2] NASA Federal agency. *NASA Earth Data Site*. https://worldview.earthdata.nasa.gov/. 2

[3] Richard P Brent. *Algorithms for minimization without derivatives*. Courier, 2013. 2

[4] Paul Bromiley. Products and convolutions of Gaussian probability density functions. *Tina-Vision Memo*, 3(4):1, 2003. 6

[5] Howard R Gordon, James W Brown, and Robert H Evans. Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone color scanner. *Applied Optics*, 27(5):862–871, 1988. 2

[6] Michael Hatzitheodorou and John R Kender. An optimal algorithm for the derivation of shape from shadows. In *Proc. IEEE CVPR*, pages 486–487, 1988. 2

[7] Aviad Levis, Yoav Y Schechner, Amit Aides, and Anthony B Davis. Airborne three-dimensional cloud tomography. In *Proc. IEEE ICCV*, pages 3379–3387, 2015. 2, 4, 5

[8] Gregoris Liasis and Stavros Stavrou. Satellite images analysis for shadow detection and building height estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119:437–450, 2016. 2

[9] Tamar Loeub, Aviad Levis, Vadim Holodovsky, and Yoav Y Schechner. Monotonicity prior for cloud tomography. In *Proc. ECCV*, pages 24–29, 2020. 4

[10] Roi Ronen. *4D Cloud Scattering Tomography. 2021. Available online*. https://github.com/ronenroi/4D-Cloud-Scattering-Tomography. 5

[11] The SciPy community. *SciPy. 2020. Available online*. https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize_scalar.html. 2

[12] Flavio R Velasco. Thresholding using the ISODATA clustering algorithm. Technical report, Univ. of Maryland College Park, Computer Science Center, 1979. 2

[13] Menghua Wang. A refinement for the Rayleigh radiance computation with variation of the atmospheric pressure. *International Journal of Remote Sensing*, 26(24):5651–5663, 2005. 2