This article has been accepted for publication in IEEE Transactions on Pattern Analysis and Machine Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2024.3467913

NeMF: Neural Microphysics Fields

Inbal Kom Betzer, Roi Ronen, Vadim Holodovsky, Yoav Y. Schechner and Ilan Koren

Abstract—Inverse problems in scientific imaging often seek physical characterization of heterogeneous scene materials. The scene is thus represented by physical quantities, such as the density and sizes of particles (microphysics) across a domain. Moreover, the forward image formation model is physical. An important case is that of clouds, where microphysics in three dimensions (3D) dictate the cloud dynamics, lifetime and albedo, with implications to Earth's energy balance, sustainable energy and rainfall. Current methods, however, recover very degenerate representations of microphysics. To enable 3D volumetric recovery of all the required microphysical parameters, we introduce the neural microphysics field (NeMF). It is based on a deep neural network, whose input is multi-view polarization images. NeMF is pre-trained through supervised learning. Training relies on polarized radiative transfer, and noise modeling in polarization-sensitive sensors. The results offer unprecedented recovery, including droplet effective variance. We test NeMF in rigorous simulations and demonstrate it using real-world polarization-image data.

Index Terms—Computational Photography, Inverse Problem, 3D Reconstruction

1 An Important Scientific Problem

Scene reconstruction problems are solved with superior quality and speed thanks to deep neural networks (DNNs). A DNN can represent a continuous spatial field, such as radiance, using a neural field [1]. Based on a neural field, rendering using an ad-hoc or physics-based forward model yields view synthesis. This idea is used in volumetric reconstruction using images [2] and sonar [3], relighting [4], de-scattering [5], [6], [7], [8], non-line-of-sight imaging [9], atomic electron tomography [10], colorization [11], 3D vision using lidar [12] and polarimetric decomposition [13]. The neural-field concept can be used in scientific inverse problems. One example is rendering of moving matter near the event-horizon of a black hole [14], where the forward model is a linear integral (on a curve) through the medium.

In this paper, we define a neural field that is nonlinear in the scene variables, and addresses a different scientific domain: a volumetric map of vectors, where each vector expresses the microphysical properties of scattering particles. While the framework we present is sufficiently general to use in a variety of media, we focus here on microphysics of water droplets in clouds. What is the motivation to focus on clouds, and what is so important about their microphysics? Clouds play critical roles in climate, weather (the essential source of freshwater), and solar-power generation. As we explain below, these roles are determined by the spatial field of microphysics of droplets in a cloud [15], [16]. For example, if a cloud is more reflective, then less sunlight reaches the ground. Noting that clouds account for 2/3 of Earth's albedo, clouds that are more reflective have a cooling effect on Earth (countering global warming), yet they lower generation of solar power, generally. These radiative effects depend on the lifetime of clouds: clouds that take long to dissipate reflect more energy to space rather than transmitting energy to the ground. The lifetime of a cloud is significantly shortened if it precipitates.[1]

1. Clouds, global weather and climate have more complex couplings and feedback effects. We do not detail them here. Their complexity supports the cause of the paper: it is important to resolve the microphysics of clouds, per location in three dimensions.

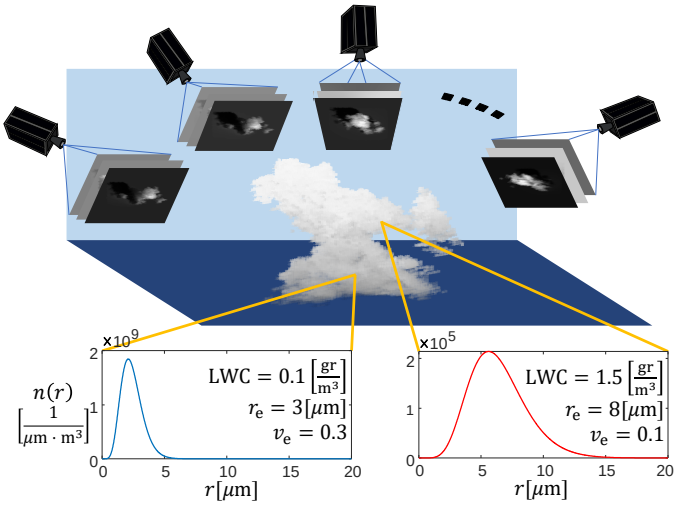


Fig. 1. Simultaneous multi-view polarimetric images of a cloud are acquired from orbit. Based on the images, the cloud's microphysical properties are retrieved volumetrically in 3D. This expresses per location the droplet size distribution $n(r)$, which is parameterized by effective radius r_e, effective variance v_e and liquid water content LWC.

The cloud's albedo, lifetime and ability to precipitate depend on the size distribution of its water droplets, i.e., the cloud microphysics (Fig. 1). Consider several ways in which water vapour (gas) of fixed amount condensates into cloud droplets: few large droplets, many small droplets, or a broad distribution of sizes. A cloud having many small droplets has a higher albedo than a cloud having a few large droplets. This is called the *Twomey effect* [17]. The reason is that with the former, a light ray intersects more air-water interfaces and thus scatters at higher orders. A nice demonstration of this effect is shown in Fig. 2. The Twomey effect means that images can inform about microphysics, as microphysics affect the radiance field.

Precipitation requires droplets to coalesce: this way they can form larger droplets whose weight overcomes vertical air currents, to leave the cloud towards the ground. Suc-

Fig. 2. The jars contain glass spheres and air. In the left jar, the spheres are $4000[\mu m]$ wide. In the right jar, which is significantly brighter, the spheres' width is $40[\mu m]$. Thus, tiny spheres create effectively a much higher albedo of a jar, compared to the effect of the large spheres. This is because, in a volume element of a fixed size, there are many more small spheres, leading to many more scattering events. Image taken and adapted from [17], reproduced with permission from SNCSC.

cess of this process significantly relies [15] on sufficient presence of large cloud droplets in some cloud voxel *and* large variability of sizes in a three dimensional (3D) domain. Thus, mapping cloud microphysics in 3D is a key to many processes [18]. However, there is a huge observational gap. *Neither state of the art nor research level observational methods resolve the effective variance of water droplets.* Moreover, most observational methods estimate the effective radius of droplets just as a cloud-average. Prior research [16] seeks an effective radius that changes only vertically.

We address this challenge by a framework we term *NeMF*: neural microphysics field. The field is a 3D map, where in each point in a cloud there is a vector of microphysical parameters. NeMF is based on a DNN. It trains and infers using *polarimetric multi-view images* of clouds. We use polarization for several reasons. First, polarimetric signals are very sensitive to microphysics [19], [20]. Second, prior art using unpolarized visible light [21], [22], [23], [24], [25], [26] did not yield microphysics.[2] Multi-view image data is required for 3D volumetric recovery, as typical in computed tomography (CT). Here data is acquired by remote sensing.

NeMF is trained in a supervised manner, using an existing labeled dataset of clouds, including their 3D maps of microphysics. For training, these clouds are rendered in multiple views using differentiable vector (polarized) 3D radiative transfer. Our main focus is spaceborne imaging by the geometry of the CloudCT formation, funded by the ERC. We present results also based on real multi-view airborne polarimetric data.

2 Theoretical Background

The theoretical background draws from several sources of prior art. So, we follow and modify notations and definitions used in [20], [24], [28], [29].

2.1 Microphysical Properties of Cloud Droplets

A voxel in a liquid-phase cloud contains an ensemble of microscopic spherical water droplets. A droplet has radius r. In a voxel around 3D location $\boldsymbol{x}$, the *droplets size distribution* is $n_{\boldsymbol{x}}(r)$, per radius and per unit volume $[\frac{1}{\mu\mathrm{m}\cdot\mathrm{m}^3}]$. There are several important parameters. These include, respectively, the *effective radius* $r_{\mathrm{e}}(\boldsymbol{x})$ $[\mu\mathrm{m}]$ of the droplets, a dimensionless *effective variance* $v_{\mathrm{e}}(\boldsymbol{x}) < 1/2$, and the *liquid water content* (LWC) in units of $\left[\mathrm{gr}/\mathrm{m}^3\right]$:

$$r_{\mathrm{e}}(\boldsymbol{x}) = \frac{\int_0^\infty r^3 n_{\boldsymbol{x}}(r)\mathrm{d}r}{\int_0^\infty r^2 n_{\boldsymbol{x}}(r)\mathrm{d}r}, \quad v_{\mathrm{e}}(\boldsymbol{x}) = \frac{\int_0^\infty \left(r-r_{\mathrm{e}}\right)^2 r^2 n_{\boldsymbol{x}}(r)\mathrm{d}r}{r_{\mathrm{e}}^2(\boldsymbol{x})\int_0^\infty r^2 n_{\boldsymbol{x}}(r)\mathrm{d}r},$$
$$\mathrm{LWC}(\boldsymbol{x}) = \frac{4}{3}\pi\rho_{\mathrm{w}}\int_0^\infty r^3 n_{\boldsymbol{x}}(r)\mathrm{d}r. \tag{1}$$

Here ρ_{w} is the density of liquid water in $\left[\mathrm{gr}/\mu\mathrm{m}^3\right]$. The parameter r_{e} is equivalent to a volume-weighted average of the size distribution, divided by the area-weighted average. The weights relate to interaction of a droplet with its surrounding air and aerosols, the cross section for droplet coalescence [30] and cross-section for light extinction [31].

The distribution $n(r)$ can have a general integrable form, or be parameterized as unimodal or bi-modal [32], [33]. Supported by empirical evidence [15], [33], $n(r)$ is often modeled by the unimodal Gamma-distribution [31]

$$n(r) = N\,A\,r^{(v_{\mathrm{e}}^{-1}-3)}\exp[-r/(r_{\mathrm{e}}v_{\mathrm{e}})], \tag{2}$$

where $N = \int_0^\infty n(r)\mathrm{d}r$ is the *total number of droplets* per unit volume, in $[1/\mathrm{m}^3]$. Here $A = (r_{\mathrm{e}}v_{\mathrm{e}})^{(2-v_{\mathrm{e}}^{-1})}/\Gamma(v_{\mathrm{e}}^{-1}-2)$ is a normalization constant and $\Gamma(\cdot)$ is the Gamma function. The microphysics vector of parameters per 3D location $\boldsymbol{x}$ is

$$\boldsymbol{\mu}(\boldsymbol{x}) = [\mathrm{LWC}(\boldsymbol{x}), r_{\mathrm{e}}(\boldsymbol{x}), v_{\mathrm{e}}(\boldsymbol{x})]. \tag{3}$$

For an entire scene, the *microphysics field* is $\mathcal{M} = \{\boldsymbol{\mu}(\boldsymbol{x})\}_{\forall \boldsymbol{x}}$.

The adiabatic model and its limits

Prior art [20], [25] suggests a simplified model for r_{e} and LWC. Let z [km] be the altitude above Earth's surface, and z_0 the cloud-base height. In this simplified model, the microphysics within a cloud vary only vertically, as

$$\mathrm{LWC}(z) = \kappa_{\mathrm{l}}(z - z_0) + \mathrm{LWC}^{\min}, \tag{4}$$

$$r_{\mathrm{e}}(z) = \kappa_{\mathrm{r}}(z - z_0)^{\frac{1}{3}} + r_{\mathrm{e}}^{\min}, \tag{5}$$

where κ_{l}, $\mathrm{LWC}^{\min}$, κ_{r} and $r_{\mathrm{e}}^{\min}$ are parameters. These vertically-monotonous functions assume that a cloud follows the *adiabatic approximation*. This thermodynamic model assumes that a rising air parcel (voxel) is isolated from its surroundings, regarding temperature and materials. As the isolated parcel rises, it expands, thus cools. Cooling leads to condensation of vapour into droplets at the cloud base. Then, the droplets grow in size monotonically with altitude. This approximation loses validity due to mixing of cloud parcels with surrounding air, leading to droplet evaporation, and diffusion of humidity, aerosols and temperature. Such mixing is strongest at the cloud lateral sides and at the cloud top. The adiabatic model (4,5) is thus a guideline for behavior of the cloud *core*, mainly near the cloud base.

2.2 Polarization

Polarization is sensitive to the variables we seek. Polarization is mainly associated with single-scattering, while the single scattering phase function (defined in Sec. 2.3), significantly depends on the microphysics. We thus provide essential definitions relating to polarization.

2. The Nakajima-King [27] method yields a single estimate of the effective radius for a whole cloud, and no effective variance. It relies essentially on short-wave infra-red.

Light that is partially linearly polarized is represented by a 3-element[3] Stokes vector $\mathbf{I} = [I, Q, U]^\top$. Here $I \geq 0$ expresses total (unpolarized) radiance and $\top$ denotes transposition. The real-valued elements Q and U express polarized components in a chosen coordinate system (Sec. 9.1). They satisfy $Q^2 + U^2 \leq I^2$. A Stokes vector can be converted to the Degree of Linear Polarization (DoLP) and Angle of Polarization (AoP) in the coordinate system, using

$$\begin{aligned} \mathrm{DoLP} &= \sqrt{Q^2 + U^2}/I, \\ \tan(2\mathrm{AoP}) &= U/Q, \quad 0 \leq \mathrm{AoP} < 180^\circ, \end{aligned} \tag{6}$$

and vice versa. One coordinate system is defined by the *scattering plane*, which includes the line of sight and solar illumination direction (see Sec. 9.1). In this system,

$$\mathrm{AoP} \approx 90^\circ. \tag{7}$$

Incoherent combination of several light sources (e.g, scattering by different particles) yields addition of Stokes vectors. A linear system that affects light, e.g. Mie or Rayleigh scattering, image projection and optical filtering is expressed by a 3×3 Mueller matrix. A Mueller matrix then acts on an input Stokes vector, to yield the linear system's output Stokes vector [35].

2.3 Polarized 3D Radiative Transfer

The size, shape and material type of particles affect their optical properties [36]. The extinction coefficient at location $\boldsymbol{x}$ in a heterogeneous medium, $\beta(\boldsymbol{x})$, is such a property. In and around clouds, extinction is created by air molecules and cloud water droplets. For clouds, the extinction coefficient [28] is $\beta^{\mathrm{cloud}}(\boldsymbol{x}) = \int n_{\boldsymbol{x}}(r) B(r) \pi r^2 dr$, where $B(r)$ is the extinction efficiency. For liquid-phase clouds observed in the visible spectrum [25], $B \approx 2$. Hence, from Eq. (1),

$$\beta^{\mathrm{cloud}}(\boldsymbol{x}) = \frac{3}{2} \frac{\mathrm{LWC}(\boldsymbol{x})}{r_{\mathrm{e}}(\boldsymbol{x}) \rho_{\mathrm{w}}}. \tag{8}$$

Air density falls exponentially with altitude z [km] above Earth's surface. For optical wavelength λ [μm], the extinction coefficient $[\mathrm{km}^{-1}]$ due to air [37] is

$$\beta^{\mathrm{air}}(\boldsymbol{x}) \approx \beta^{\mathrm{air}}(z) \approx \frac{1.09 \times 10^{-3}}{\lambda^4} \exp\left[-z/H_{\mathrm{air}}\right], \tag{9}$$

where $H_{\mathrm{air}} = 10\mathrm{km}$. Then, $\beta(\boldsymbol{x}) = \beta^{\mathrm{air}}(\boldsymbol{x}) + \beta^{\mathrm{cloud}}(\boldsymbol{x})$.

Consider a domain where light propagates. In our case, sunlight is the source. In a scattering event at $\boldsymbol{x}$, light changes its 3D propagation direction from $\boldsymbol{\omega}'$ to any $\boldsymbol{\omega}$. The normalized angular distribution of scattered radiance is set by a *phase function*. For partially polarized light, the phase function is generalized to a *phase matrix* $\mathbf{P}(\boldsymbol{x}, \boldsymbol{\omega} \cdot \boldsymbol{\omega}')$, which is a Mueller matrix. It converts the Stokes vector of incoming light $\mathbf{I}(\boldsymbol{x}, \boldsymbol{\omega}')$ to a Stokes vector of scattered light, per $\boldsymbol{\omega}$. The matrix $\mathbf{P}(\boldsymbol{x}, \boldsymbol{\omega} \cdot \boldsymbol{\omega}')$ is sensitive to the microphysics $\boldsymbol{\mu}(\boldsymbol{x})$. Scattering by a cloud droplet of radius r is set by a *Mie* phase matrix [28]. A small voxel averages the Mie phase matrix over all droplet sizes, according to their size distribution. Scattering by air molecules is described by the *Rayleigh* phase matrix [28].

3. A 4^{th} vector element expresses partial circular polarization. It is negligible under natural light in the open air [34], allowing work with 3-element Stokes vectors.

The single-scattering albedo, $0 \leq \varpi(\boldsymbol{x}) \leq 1$, is generally determined by $\boldsymbol{\mu}(\boldsymbol{x})$. For example, ϖ is low for spectral wavelengths absorbed by liquid water, and absorption grows with droplet size. In visible light [38], $\varpi \approx 1$.

The set of interactions in propagation, including high-order scattering, is expressed by the vector (polarized) 3D radiative transfer equation (RTE) [39]. As detailed in the supplementary material, the RTE depends on ϖ, $\mathbf{P}(\boldsymbol{x}, \boldsymbol{\omega} \cdot \boldsymbol{\omega}')$, β^{air} and β^{cloud}. Through this dependency, the partially polarized radiance field $\mathbf{I}(\boldsymbol{x}, \boldsymbol{\omega})$ depends on the 3D vector field of microphysical parameters $\mathcal{M}$. Referring to the RTE as an operator, $\mathbf{I}(\boldsymbol{x}, \boldsymbol{\omega}) = \mathrm{RTE}(\mathcal{M})$.

2.4 Neural Fields

A neural field is a DNN-based representation of a spatially continuous variable. A notable example is a neural radiance field (NeRF) [1]. Such a DNN [4], [5], [13] trains to relate multi-view image data to the field. In NeRF, the representation vector stands for volumetric pseudo-optical density (even if the object is opaque in reality) and pseudo-color per spatial location $\boldsymbol{x}$. The vector is used to render images from various viewpoints. NeRF rendering and representation vector are not actually physics-driven. Despite this, NeRF achieves new-view synthesis by pre-training of the DNN, per scene. That is, the trained DNN adapts to the task of view synthesis of a particular scene, despite having a non-physical rendering process. To yield view synthesis, NeRF trains in an unsupervised (self-supervised) manner, without labeled 3D data. It requires a large number of viewpoints. The time-consuming training is done per scene.

Our goal is *not* view synthesis, but scientific output. Thus, the vector inferred by NeMF includes the set of microphysical parameters, per location. As explained below, for meaningful scientific results, learning should be supervised.

3 Why Should Learning Be Supervised?

A question might arise: to estimate microphysics, why not match observed multi-view image data to a differential rendering model (as [16], [21], [28]) or an unsupervised DNN (as NeRF)? Both of these methods do not require prior ground truth labeled scenes. As a result, they have no prior on how microphysics might be distributed in 3D, although they can synthesize new views.

For internal recovery of clouds, however, these approaches can be fundamentally ill-posed, irrespective of the algorithm. Given noisy data, there can be a *set* of spatial and microphysical distributions of a medium, that would fit by radiative transfer. In Fig. 3 and Table 1, seven different atmospheres are examined. `Atmospheres 1-4` have an ellipsoid cloud, 1km wide, 150m thick, with *different* microphysics in each vertical half, as detailed in Table 1. However, all these clouds yield the *same* polarimetric images (up to noise), in *all* view angles of the CloudCT satellite formation. These clouds have a low optical depth, hence the intensity and polarization observed in their images are mainly due to single-scattering. These different 3D scenes are samples from an *equivalence class*: they have equivalent images fitting similarly the rendering model. So, they would also similarly fit unsupervised models based on real images (as NeRF).

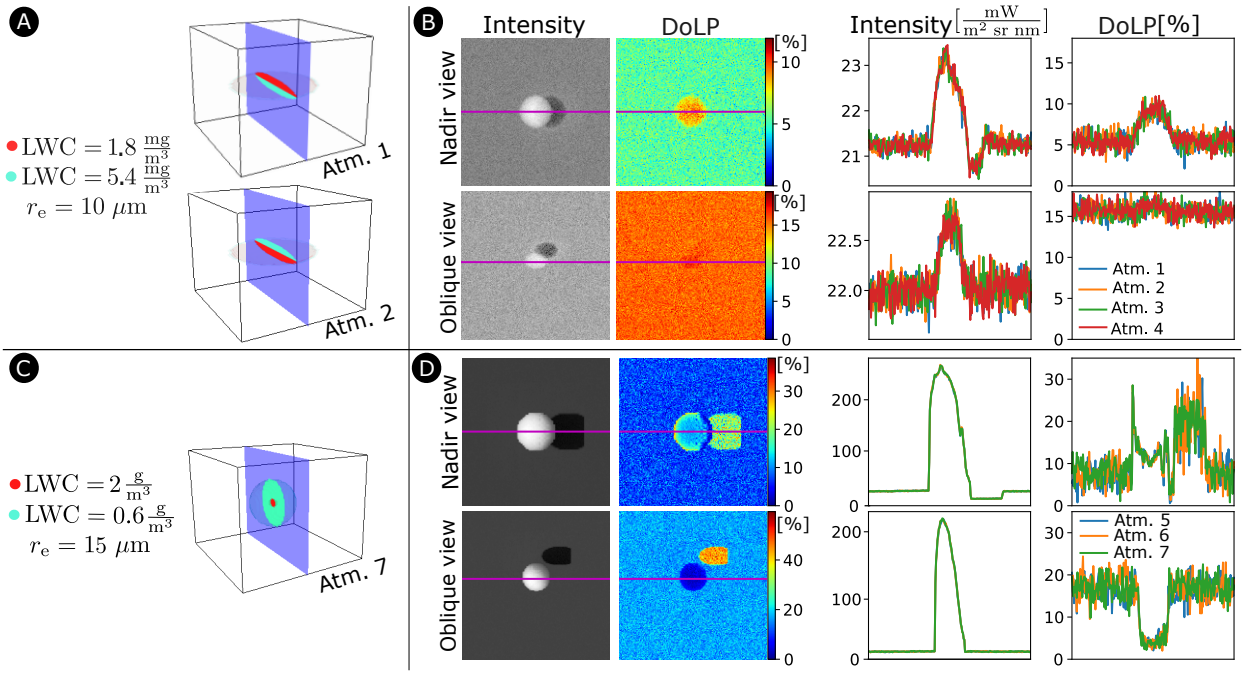


Fig. 3. (A) Cross sections of Atms. 1,2. (B) Intensity and DoLP images (2 view angles), and plots of the center row for Atms. 1-4. (C) Cross section of Atm. 7. (D) Intensity and DoLP images (2 view angles), and center-row plots for Atms. 5-7. Intensity images are contrast-stretched for display.

TABLE 1
Microphysics of the atmospheres (Atms.) presented in Fig. 3. For all, $v_e = 0.1$. Atms. 1-4 have a peak vertical optical depth of 0.1.

Name	Shape	LWC [mg/m^3]	r_e [μm]
Atm. 1	Ellipsoid	top: 5.4, bottom: 1.8	10
Atm. 2	Ellipsoid	top: 1.8, bottom: 5.4	10
Atm. 3	Ellipsoid	top: 4.5, bottom: 3	top: 12, bottom: 8
Atm. 4	Ellipsoid	top: 3.0, bottom: 4.5	top: 8, bottom: 12
Atm. 5	Sphere	core: 0, outer: 600	15
Atm. 6	Sphere	core: 800, outer: 600	15
Atm. 7	Sphere	core: 2000, outer: 600	15

Another equivalence class is sampled in `Atmospheres 5-7`, featuring spherical clouds, 1.2km wide. Their 240m wide core has a variable LWC. Their outer shell has a high optical thickness. As a result, it is not possible to recover the cores' microphysics based on polarimetric multi-view images alone. These clouds appear indistinguishable. Additional examples are in [40].

Data-based ill-posedness or a severe ill-condition is solved by *priors*. By prior observations of *known* 3D clouds, a system learns "how clouds are structured," and what is more likely to be inside them. The system learns this, even if information carried by images is *non-discriminative* or weakly discriminative among possible cloud solutions. To solve ambiguous cases, supervision is (at least partially) needed as a part of the learning process, using labeled ground-truth clouds. This is preferred over purely physics-based or unsupervised methods.

There are additional reasons to favor supervised training. To avoid backlog of *years* of spaceborne data, inference speed is essential. Define σ as the product of spaceborne downlink rate, the per-day communication time with a ground station, and the number of satellites. Define δ as the product of the number of seconds per day, the number of bits per pixel and the number of pixels per km^2. To avoid backlog, the area that has to be processed per second should be larger than σ/δ. For conservative S-band 0.4Mbps downlink, communication at 10 minute/day, 10 satellites and 20m ground resolution, the processing rate should be higher than $1.1\text{km}^2/\text{sec}$.

However, the forward model includes full 3D polarized radiative transfer in a physical volumetric medium that has multiple-scattering. This is significantly more complex than rendering in NeRF models. Physics-based inversion is also complex. Fortunately, in supervised learning, computational resources are invested and used before test image acquisition. This enables fast inference. Moreover, we have only a few viewpoints per object we observe. Therefore, training benefits from many objects (clouds) imaged previously.

4 NeMF

We create a system of neural microphysics fields (NeMF).[4] It learns and infers from polarization images, and reconstructs microphysical properties in a 3D domain. We believe that the architecture can suit other physics problems, that rely on images to recover important underlying characteristics of objects. The input to NeMF is a set of multi-view polarimetric images (denoted $\mathbf{y}$), acquired by N^{cam} cameras, located at $\{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}$. As shown in Fig. 4, per queried *continuous valued* location $\boldsymbol{x}$, NeMF infers a microphysics vector $\hat{\boldsymbol{\mu}}(\boldsymbol{x}) = [\hat{\text{LWC}}(\boldsymbol{x}), \hat{r}_e(\boldsymbol{x}), \hat{v}_e(\boldsymbol{x})]$. A location is queried only if it passes a space carving operation. Space carving [41], [42], based on the acquired intensity images, constrains a domain where cloud droplets may reside.

4. Our code, including NeMF, is available in the public domain at https://github.com/inbalkb/NeMF

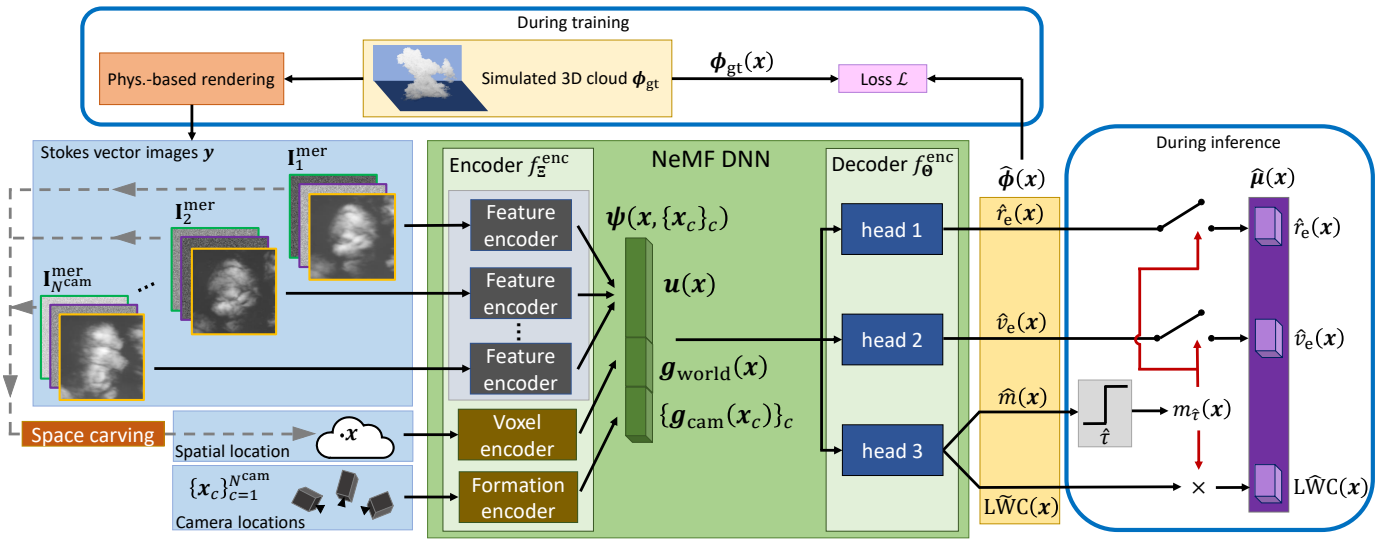


Fig. 4. NeMF workflow. A domain is observed from N^{cam} different viewpoints, resulting in $\mathbf{y} = \{\mathbf{I}_c^{\text{mer}}\}_{c=1}^{N^{\text{cam}}}$, a set of Stokes vector images. Camera locations are encoded as $\mathbf{g}_{\text{cam}}(\boldsymbol{x}_c)$, $\forall c$. A space-carving mask bounds the locations where a cloud might be. In it, a queried 3D location $\boldsymbol{x}$ is encoded to a vector $\mathbf{g}_{\text{world}}(\boldsymbol{x})$. A single convolutional neural network extracts features from all images $\mathbf{y}$. The image features that correspond to 2D projections of $\boldsymbol{x}$ to the cameras form a vector $\mathbf{u}(\boldsymbol{x})$. The vectors $\mathbf{g}_{\text{world}}(\boldsymbol{x})$, $\{\mathbf{g}_{\text{cam}}(\boldsymbol{x}_c)\}_{c=1}^{N^{\text{cam}}}$ and $\mathbf{u}(\boldsymbol{x})$ become a joint input to a three-headed decoder. The decoder outputs a vector $\hat{\boldsymbol{\phi}}(\boldsymbol{x})$ of four parameters, including a cloud existence measure $\hat{m}(\boldsymbol{x})$. The latter is converted to an inferred cloud mask, using a learned threshold $\hat{\tau}$. Microphysical parameters are inferred if $m_{\hat{\tau}}(\boldsymbol{x}) = 1$. Otherwise, the only output at $\boldsymbol{x}$ is $\widehat{\text{LWC}}(\boldsymbol{x}) = 0$. In training, a simulated 3D vector field $\boldsymbol{\phi}_{\text{gt}}$ of cloud microphysics undergoes physics-based rendering, as described in Sec. 4.4. Then, $\hat{\boldsymbol{\phi}}(\boldsymbol{x})$ is compared with $\boldsymbol{\phi}_{\text{gt}}(\boldsymbol{x})$, to optimize the DNN's loss $\mathcal{L}$.

4.1 NeMF Architecture and Inference

NeMF uses an encoder-decoder architecture, illustrated in the green block, titled "NeMF DNN", in Fig. 4. Additional implementation details are in Sec. 4.3, Sec. 6 and the supplementary material. The encoder has three components. One component encodes the set of multi-view polarimetric images, $\mathbf{y}$, using a single *image-feature extractor*, described in Sec. 4.3. This component yields a vector $\mathbf{u}(\boldsymbol{x})$ of image features that correspond to projections of $\boldsymbol{x}$. A second component encodes the 3D location $\boldsymbol{x}$ to a vector $\mathbf{g}_{\text{world}}(\boldsymbol{x})$, using a DNN. The third component is a single DNN, that encodes any camera location $\boldsymbol{x}_c$ to a vector $\mathbf{g}_{\text{cam}}(\boldsymbol{x}_c)$. All these vectors are concatenated to a single vector,

$$\boldsymbol{\psi}\left(\boldsymbol{x}, \{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}\right) = \left[\mathbf{u}(\boldsymbol{x}), \mathbf{g}_{\text{world}}(\boldsymbol{x}), \mathbf{g}_{\text{cam}}(\boldsymbol{x}_{c=1}), ..., \mathbf{g}_{\text{cam}}(\boldsymbol{x}_{c=N^{\text{cam}}})\right] . \tag{10}$$

The three encoder components jointly yield a single high-dimensional function $f_{\boldsymbol{\Xi}}^{\text{enc}}$ of all the inputs. The function is set by a vector of learned parameters $\boldsymbol{\Xi}$,

$$\boldsymbol{\psi}\left(\boldsymbol{x}, \{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}\right) = f_{\boldsymbol{\Xi}}^{\text{enc}}\left[\mathbf{y}, \boldsymbol{x}, \{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}\right] . \tag{11}$$

The vector $\boldsymbol{\psi}\left(\boldsymbol{x}, \{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}\right)$ then serves as input to a *decoder*.

The decoder is a function $f_{\boldsymbol{\Theta}}^{\text{dec}}$ set by a vector of learned parameters $\boldsymbol{\Theta}$:

$$\hat{\boldsymbol{\phi}}(\boldsymbol{x}) = f_{\boldsymbol{\Theta}}^{\text{dec}}\left[\boldsymbol{\psi}\left(\boldsymbol{x}, \{\boldsymbol{x}_c\}_{c=1}^{N^{\text{cam}}}\right)\right] . \tag{12}$$

The decoder outputs a 4-element vector, $\hat{\boldsymbol{\phi}}(\boldsymbol{x}) = [\hat{r}_{\text{e}}(\boldsymbol{x}), \hat{v}_{\text{e}}(\boldsymbol{x}), \widetilde{\text{LWC}}(\boldsymbol{x}), \hat{m}(\boldsymbol{x})]$, as we explain. The decoder has three separate heads. One head infers $\hat{r}_{\text{e}}(\boldsymbol{x})$. The second head infers $\hat{v}_{\text{e}}(\boldsymbol{x})$. The third head infers two values: $\widetilde{\text{LWC}}(\boldsymbol{x})$ and a cloud-existence measure $0 \leq \hat{m}(\boldsymbol{x}) \leq 1$. Define an inferred binary cloud mask using a learned threshold τ:

$$m_\tau(\boldsymbol{x}) = \begin{cases} 1 & \hat{m}(\boldsymbol{x}) > \tau \\ 0 & \text{otherwise} \end{cases} . \tag{13}$$

Consider a queried location $\boldsymbol{x}$, where $m_\tau(\boldsymbol{x}) = 0$. There, the estimated LWC should be nulled. So, the inferred LWC is

$$\widehat{\text{LWC}}(\boldsymbol{x}) = \widetilde{\text{LWC}}(\boldsymbol{x}) \cdot m_\tau(\boldsymbol{x}) . \tag{14}$$

Moreover, if $m_\tau(\boldsymbol{x}) = 0$, values assigned to $\hat{r}_{\text{e}}(\boldsymbol{x}), \hat{v}_{\text{e}}(\boldsymbol{x})$ are meaningless. Therefore, during inference, $\hat{r}_{\text{e}}(\boldsymbol{x}), \hat{v}_{\text{e}}(\boldsymbol{x})$ are reported and used in error calculations (Sec. 6) only where $m_\tau(\boldsymbol{x}) = 1$.

4.2 Training Process

The ground-truth vector corresponding to $\hat{\boldsymbol{\phi}}(\boldsymbol{x})$ is $\boldsymbol{\phi}^{\text{gt}}(\boldsymbol{x}) = [r_{\text{e}}^{\text{gt}}(\boldsymbol{x}), v_{\text{e}}^{\text{gt}}(\boldsymbol{x}), \text{LWC}^{\text{gt}}(\boldsymbol{x}), m_{\text{gt}}(\boldsymbol{x})]$, where the values of $r_{\text{e}}^{\text{gt}}(\boldsymbol{x}), v_{\text{e}}^{\text{gt}}(\boldsymbol{x})$ and $\text{LWC}^{\text{gt}}(\boldsymbol{x})$ are taken from simulated cloud datasets, described in Sec. 7, and

$$m_{\text{gt}}(\boldsymbol{x}) = \begin{cases} 1 & \text{LWC}^{\text{gt}}(\boldsymbol{x}) > 0 \\ 0 & \text{otherwise} \end{cases} . \tag{15}$$

Training the DNN is supervised, minimizing a loss

$$\left[\hat{\boldsymbol{\Theta}}, \hat{\boldsymbol{\Xi}}\right] = \text{argmin}_{\boldsymbol{\Theta},\boldsymbol{\Xi}} \mathcal{L}\left[\{\boldsymbol{\phi}^{\text{gt}}(\boldsymbol{x})\}_{\forall \boldsymbol{x}}, \{\hat{\boldsymbol{\phi}}(\boldsymbol{x})\}_{\forall \boldsymbol{x}}\right] , \tag{16}$$

where $\mathcal{L} = \mathcal{L}_{\text{LWC}} + \mathcal{L}_{\text{mask}} + \mathcal{L}_{\text{droplets}}$, as detailed in Sec. 9.3. After completing Eq. (16), we use a separate set of $N^{\text{val}} = 100$ scenes, to optimize τ of Eq. (13). For this, we calculate the F_1 score [43] of $\{m_\tau(\boldsymbol{x})\}_{\forall \boldsymbol{x}}$ relative to $\{m_{\text{gt}}(\boldsymbol{x})\}_{\forall \boldsymbol{x}}$, and define the error, per scene,

$$\epsilon_{m,\tau} = 1 - F_1\left[\{m_\tau(\boldsymbol{x})\}_{\forall \boldsymbol{x}}, \{m_{\text{gt}}(\boldsymbol{x})\}_{\forall \boldsymbol{x}}\right] . \tag{17}$$

Averaging $\epsilon_{m,\tau}$ over all N^{val} scenes yields a measure $\bar{\epsilon}_{m,\tau}$. We optimize $\hat{\tau} = \arg\min_\tau \bar{\epsilon}_{m,\tau}$.

4.3 Image-Feature Extractor

Feature extraction from any polarization image (a tensor having two spatial dimensions and channels of Stokes elements $[I, Q, U]$) is based on a convolutional neural network (CNN). The convolution kernels couple the polarization channels. The same extractor is used on all $N^{\rm cam}$ images, which are stacked in the batch dimension, in order to parallelize their processing. A detailed illustration and description of the architecture are shown in the supplementary material. Briefly, as in VIP-CT [24], the CNN is based on ResNet [44]. In NeMF, the number of channels increases during propagation through the levels of the CNN, while the size of spatial dimensions decreases.

The CNN extracts features for *discrete-grid* pixels of each image. The 3D location $\boldsymbol{x}$ has continuous-valued coordinates. Also geometric projection of $\boldsymbol{x}$ to camera c falls on a continuous-valued 2D location in the image plane, denoted $\boldsymbol{\chi}_c(\boldsymbol{x})$. For a feature vector of a continuous-valued $\boldsymbol{\chi}_c$, the features on the discrete pixel grid of camera c are interpolated. This process is done $\forall c$. The set of interpolated features corresponding to $\{\boldsymbol{\chi}_c\}_{c=1}^{N^{\rm cam}}$ is concatenated, to form vector $\mathbf{u}(\boldsymbol{x})$.

4.4 Rendering

We train the network using polarization-rendered images. This starts with simulated 3D cloud microphysics. The RTE uses the 3D microphysics to create a light field of Stokes vectors. The light field is projected towards camera c, in direction corresponding to pixel p.

The RTE solver represents Stokes vectors using the *meridian* coordinate system (see Sec. 9.1). The Stokes vectors then transform to the *pixel coordinate system* (Sec. 9.2) using a Mueller rotation matrix $\mathbf{R}\left(\gamma\right)$, where γ is the angle between the meridian and pixel systems. Then, light passes through a pixel-based polarizer, and converts to photo-electrons. The optical system's point spread function is wider than a sensor pixel. So, light spreads over a small area, a *super-pixel*, each having 2×2 sensor pixels. From now on, p indexes a super-pixel. The sensor array has a *mosaic* of tiny polarization filters at angles $[0°, 45°, 90°, 135°]$, so a super-pixel samples each,[5] yielding a 4-element vector of *expected* photo-electron readout $\mathbf{e}[c, p]$. A matrix $\mathbf{W}_{\rm pol}$ expresses how the set of four filters affect the incoming light. The detailed forward model is in Sec. 9.2.

The *actual* number of photo-electrons generated at p is sampled from a Poissonian distribution

$$\mathbf{e}_{\rm meas}[c, p] \sim \text{Poisson}\{\text{mean} = \mathbf{e}[c, p]\} \ . \tag{18}$$

We convert the Poissonian measurements back to a noisy Stokes vector in the meridian system. This is done by a pseudo-inverse of the forward model up to this point:

$$\mathbf{I}^{\rm mer}_{\rm meas}[c, p] = \frac{1}{D}\mathbf{R}^{-1}\left(\gamma\right)\left(\mathbf{W}^{\top}_{\rm pol}\mathbf{W}_{\rm pol}\right)^{-1}\mathbf{W}^{\top}_{\rm pol}\mathbf{e}_{\rm meas}[c, p], \tag{19}$$

where D is a factor accounting for several multiplicative imaging parameters (see Sec. 9.2). Compounding the RTE operator, $\text{RTE}(\mathcal{M})$, the equations in Sec. 9.2 and Eqs. (18,19), a *forward model* relates the cloud's microphysics field $\mathcal{M}$ to a Stokes vector measured in an imager, accounting for 3D radiative transfer, projection, sampling by the camera and noise. The noisy measured data vector is then $\mathbf{y} \equiv \{\mathbf{I}^{\rm mer}_{\rm meas}[c, p]\}_{\forall c,p}$.

5. The sampling, projection and noise model we derive here apply to common wide-field cameras. For special pushbroom cameras, a different model is applicable [16].

5 Assessing the Cloud Core

The cloud core has the least mixing with surrounding air, suffering least from evaporation. Thus, per altitude z, we associate the core with maximal values of LWC and $r_{\rm e}$. For robustness, we average the $N_{\max}$ highest values per z. Let $\{r_{\rm e}^{(i)}(z)\}_{i=1}^{N_{\max}}$ be effective radii of the highest values at altitude z. Let $\{\text{LWC}^{(i)}(z)\}_{i=1}^{N_{\max}}$ be the liquid water content of the highest values at altitude z. Then based on ground-truth microphysics, define

$$r_{\rm e}^{\rm core}(z) = \frac{1}{N_{\max}}\sum_{i=1}^{N_{\max}} r_{\rm e}^{(i)}(z), \tag{20}$$

$$\text{LWC}^{\rm core}(z) = \frac{1}{N_{\max}}\sum_{i=1}^{N_{\max}} \text{LWC}^{(i)}(z). \tag{21}$$

Analogously, define corresponding values based on inferred fields, $\hat{r}_{\rm e}^{\rm core}(z)$ and $\hat{\text{LWC}}^{\rm core}(z)$.

6 Implementation

In the NeMF DNN, the 3D location encoder is a multi-layer perceptron of depth 4 with hidden layers, 64 neurons wide. The camera location DNN encoder is a multi-layer perceptron of the same sizes as the 3D location encoder. NeMF's decoder consists of three heads. Each decoder head has 9 layers: 2048 neurons in the first hidden layer and 512 neurons in each hidden layer besides the first.

In our implementation, NeMF runs on an NVIDIA Tesla V100-DGXS-32GB GPU. Training uses stochastic gradient descent on the loss defined in Sec. 4.2 for $\approx 300,000$ iterations. We use an Adam optimizer with a learning rate of 1e-5 and a weight decay of 1e-5. Each iteration includes 1000 samples of random voxels from the same scene, excluding voxels that fall outside a space-carving region [41]. Space carving is based only on the input intensity images.

7 Simulations

Real multi-view polarimetric data of clouds is currently scarce, and therefore cannot train a DNN yet. So, we use two simulated cloud datasets, *BOMEX* and *CASS*, that are reported in [23], [24]. The domain [45], [46] voxels have height of $40[m]$ and width of $50[m]$. The datasets' specifications are detailed in [24]. Respectively for BOMEX and CASS, training dataset sizes are 5900 and 10808 clouds. Corresponding testing data sizes are 566 and 1000 clouds.

Rendering is implemented using an online open-source physics-based polarimetric radiative transfer solver (vSHDOM) [47]. All cloud datasets and rendered images are accessible at [48]. The sun is set at a northeast azimuth and zenith angle of $25°$. We examine several scenarios, described in [24]. All are based on the CloudCT formation of

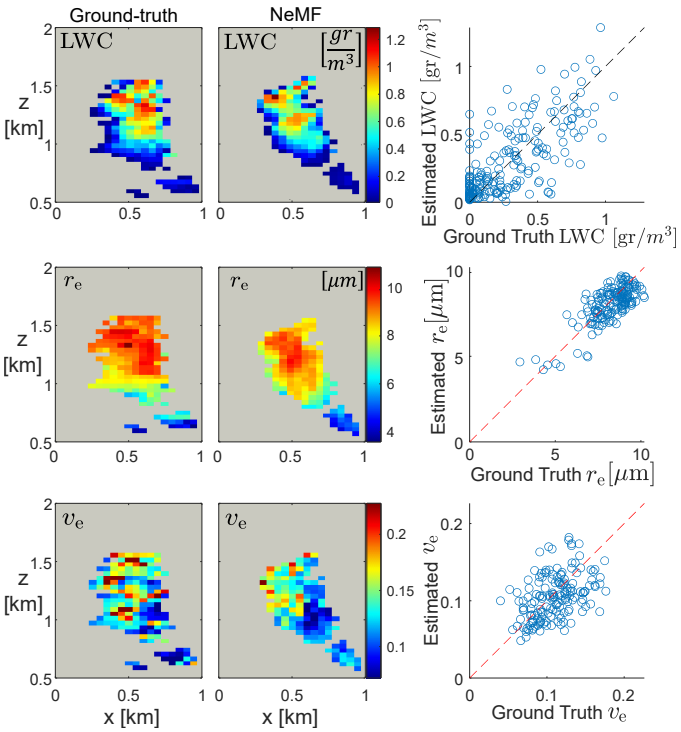


Fig. 5. An example. The 3D scene domain has several disconnected clouds from the BOMEX inference set. The figure shows maximum intensity projections (MIP) and scatter plots, which compare the simulated ground truth with NeMF's estimation.

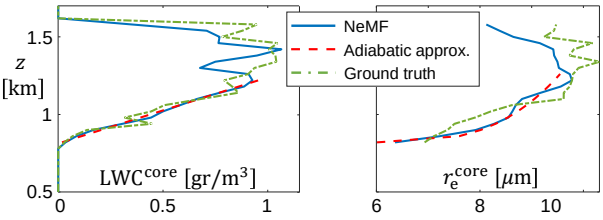


Fig. 6. Microphysical parameters at the assessed cloud core, corresponding to a cloud in Fig. 5. The adiabatic model follows Eqs. (4,5).

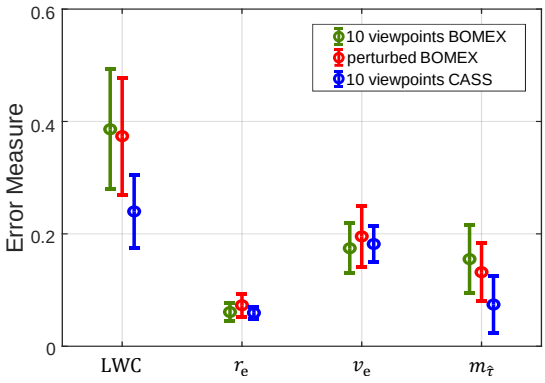


Fig. 7. Simulation results for the test sets. Mean errors of the retrieved parameters, for the three scenarios and test datasets described in Sec. 7 are shown. Bars represent the standard deviation.

10 satellites at 500km altitude, with a distance of $\approx$ 100km between each two neighboring satellites. One scenario has *perturbed* poses of cameras observing BOMEX clouds.

The factor D in Eq. (19) and noise specifications are derived from a SONY IMX250MZR sensor [20], using $620-670[\mathrm{nm}]$ waveband. The maximum image-pixel value corresponds to 90% of the sensor full well (10,500 photo-electrons). Beside Poissonian noise, the readout noise standard-deviation is 2.31 electrons.

7.1 Evaluation Criteria

The microphysics retrieval quality is quantified by relative errors. Let a cloud have a set of cloud voxels $\mathbf{V}$. Then, let

$$\begin{aligned}
\epsilon_{\mathrm{LWC}} &= \frac{\sum_{\boldsymbol{x}\in\mathbf{V}} |\mathrm{LWC}^{\mathrm{gt}}(\boldsymbol{x}) - \hat{\mathrm{LWC}}(\boldsymbol{x})|}{\sum_{\boldsymbol{x}\in\mathbf{V}} |\mathrm{LWC}^{\mathrm{gt}}(\boldsymbol{x})|}, \\
\epsilon_{r_e} &= \frac{\sum_{\boldsymbol{x}\in\mathbf{V}} |m_{\mathrm{gt}}(\boldsymbol{x})\,[r_e^{\mathrm{gt}}(\boldsymbol{x}) - \hat{r}_e(\boldsymbol{x})]\,|}{\sum_{\boldsymbol{x}\in\mathbf{V}} |m_{\mathrm{gt}}(\boldsymbol{x}) r_e^{\mathrm{gt}}(\boldsymbol{x})|}, \\
\epsilon_{v_e} &= \frac{\sum_{\boldsymbol{x}\in\mathbf{V}} |m_{\mathrm{gt}}(\boldsymbol{x})\,[v_e^{\mathrm{gt}}(\boldsymbol{x}) - \hat{v}_e(\boldsymbol{x})]\,|}{\sum_{\boldsymbol{x}\in\mathbf{V}} |m_{\mathrm{gt}}(\boldsymbol{x}) v_e^{\mathrm{gt}}(\boldsymbol{x})|}.
\end{aligned} \tag{22}$$

To quantify errors of $\{m_{\hat{\tau}}(\boldsymbol{x})\}_{\forall \boldsymbol{x}\in\mathbf{V}}$, we use Eq. (17).

Per location, Eq. (8) yields the inferred and ground truth values of the cloud's extinction coefficient, $\hat{\beta}^{\mathrm{cloud}}(\boldsymbol{x})$ and $\beta_{\mathrm{gt}}^{\mathrm{cloud}}(\boldsymbol{x})$, respectively. Then, analogously to Eq. (22), the relative error of the cloud extinction coefficient is

$$\epsilon_{\beta^{\mathrm{cloud}}} = \frac{\sum_{\boldsymbol{x}\in\mathbf{V}} |\beta_{\mathrm{gt}}^{\mathrm{cloud}}(\boldsymbol{x}) - \hat{\beta}^{\mathrm{cloud}}(\boldsymbol{x})|}{\sum_{\boldsymbol{x}\in\mathbf{V}} |\beta_{\mathrm{gt}}^{\mathrm{cloud}}(\boldsymbol{x})|}. \tag{23}$$

Eq. (23) enables comparisons to prior art, including VIP-CT [24], which does not retrieve a microphysics field. This was done by running the VIP-CT code, which resides in [49].

7.2 Results

An example of an inferred BOMEX scene domain is visualized in Fig. 5, for each microphysical parameter. Visualization is via maximum intensity projection (MIP) in the (x, z) plane, where x is a lateral coordinate. A more quantitative assessment is seen in the scatter-plots of Fig. 5.

Fig. 6 examines the largest cloud of Fig 5, using the analysis of Sec. 5 with $N_{\max} = 3$. The functions $\hat{r}_e^{\mathrm{core}}(z)$ and $\hat{\mathrm{LWC}}^{\mathrm{core}}(z)$ derived by NeMF are consistent with the functions $r_e^{\mathrm{core}}(z)$ and $\mathrm{LWC}^{\mathrm{core}}(z)$ derived from ground-truth. Moreover, $\hat{r}_e^{\mathrm{core}}(z)$ and $\hat{\mathrm{LWC}}^{\mathrm{core}}(z)$ fit well the adiabatic model described in Sec. 2.1, from the cloud base, beyond half-way to the cloud top. In this example, $\kappa_{\mathrm{l}} = 2.35[\mathrm{gr}/(\mathrm{km}\cdot\mathrm{m}^3)]$, $\mathrm{LWC}^{\min} = 0.0148[\mathrm{gr}/\mathrm{m}^3]$, $\kappa_{\mathrm{r}} = 4.65[\mu\mathrm{m}/(\mathrm{km})^{1/3}]$ and $r_e^{\min} = 6.01[\mu\mathrm{m}]$.

Analysis was done on the whole inferred test sets. Consequently, the error measures (Eqs. 17,22) are plotted in Fig. 7, for the different scenarios described above. The cloud mask and r_e are inferred well in all three scenarios. Being able to report microphysical parameters that vary in 3D is novel. The ability of NeMF to estimate v_e is significant, as no prior art method in any community achieved this. Using inferred microphysical parameters, Eq. (8) yields $\hat{\boldsymbol{\beta}}^{\mathrm{cloud}}$. Consequently, Eq. (23) yields $\epsilon_{\beta^{\mathrm{cloud}}}$, whose values are comparable to ϵ_{LWC}. They are also comparable to errors resulting from VIP-CT, as seen in Table 2.

Additional statistics are gleaned by estimating the joint probability density functions, $\mathrm{PDF}(r_e^{\mathrm{gt}}, v_e^{\mathrm{gt}})$ and

TABLE 2
Relative errors of the extinction coefficient, $\epsilon_{\beta^{\text{cloud}}}$, comparing NeMF to VIP-CT [24], across three scenarios described in Sec. 7.

Dataset	BOMEX		CASS
Geometry	10 views	perturbed	10 views
VIP-CT	$36 \pm 25\%$	$36 \pm 13\%$	$21 \pm 6\%$
NeMF	$38 \pm 10\%$	$37 \pm 10\%$	$23 \pm 6\%$

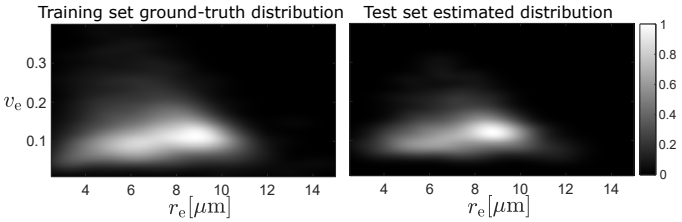


Fig. 8. Joint probability density functions distributions $\text{PDF}(r_e^{\text{gt}}, v_e^{\text{gt}})$ and $\text{PDF}(\hat{r}_e, \hat{v}_e)$. For display clarity only, both are normalized to a maximum value of 1, and then gamma corrected.

$\text{PDF}(\hat{r}_e, \hat{v}_e)$. Respectively, these estimates use random voxels from the training set (3 voxels per cloud) and test set (5 voxels per cloud). Each random sample then spreads a 2D Gaussian in the PDF domain. The PDFs are presented in Fig. 8. They are similar. However, inference tends to avoid rare (very large or very small) values of $\hat{r}_e$ and $\hat{v}_e$. Quantitative comparison of the PDFs uses the Kullback–Leibler (KL) divergence. Because the KL-divergence is defined for discrete-domain distributions, the PDFs are aggregated in discrete domain bins. The result is

$$\frac{\text{KL}\left[\text{PDF}_{\text{gt}}(r_e, v_e) \| \text{PDF}_{\text{est}}(r_e, v_e)\right]}{\log\left(\#\text{ of bins}\right)} = 0.018. \qquad (24)$$

The error measures above are statistics. While most of NeMF's reconstructions agree with the ground-truth, NeMF sometimes fails: a failure case is seen in Fig. 9.

8 Empirical Experiment

Real multi-view polarimetric images of clouds are derived from the Jet Propulsion Laboratory's AirMSPI, a payload polarimetric sensor on the ER-2 aircraft of NASA. AirMSPI has a pushbroom camera, with which it acquires 9 images of a domain along an angular course of $\pm 67°$ relative to the local vertical [50]. For training, we follow Ref. [24] to use a modified BOMEX dataset. Images are projected using AirMSPI's noise model [16] and exploiting the known camera positions $\{\boldsymbol{x}_c\}_{c=1}^{N_{\text{cam}}}$. Training uses five sets of projections from five different flights (not including the inference flight). To assess results, we omit the image taken at $-58.9°$ from the training set. Thus, the system is trained to infer a cloud based on $N^{\text{cam}} = 8$ viewing directions. NeMF infers the cloud microphysics field $\mathcal{M}$ based on the real data. Afterwards, we perform view-synthesis. This synthesis sets the ocean albedo to 0.03 and uses polarized RTE to render the missing polarimetric image.

The rendered polarimetric image is then compared to the true image at $-58.9°$. This rendering-based comparison is also applied on a corresponding physics-based retrieved

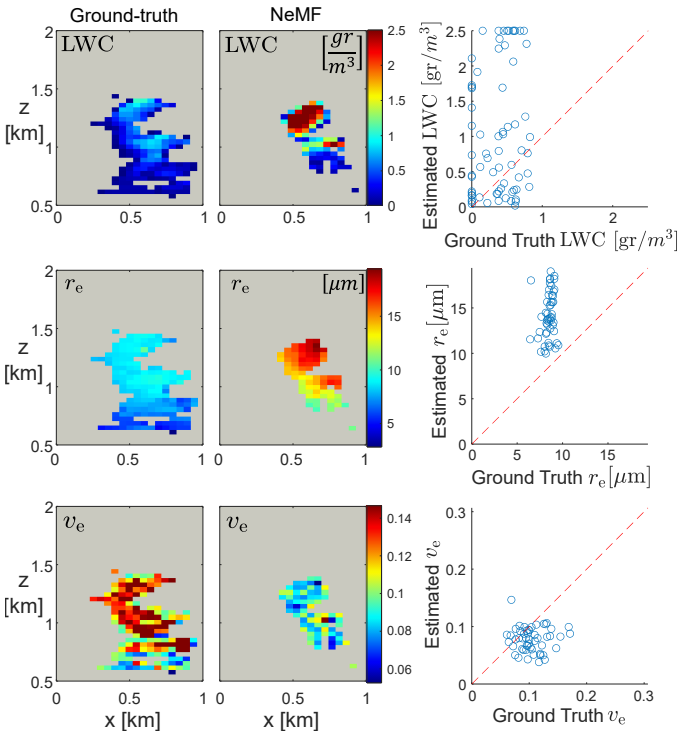


Fig. 9. A failure case. Here the microphysics MIP and scatter plots show significant inconsistencies between inference and ground-truth. This cloud is from the BOMEX test set.

cloud [16], [28]. Towards this, we used results kindly supplied by Aviad Levis [16], [28]. There, only LWC and r_e had been retrieved, while the effective variance had been set to $v_e = 0.1$ for all voxels.

Remote sensing standard uses vectors in the *meridian system*. It is often convenient, however, to use the *scattering coordinate system*, because there, typically, AoP $\approx 90°$. A detailed description of both coordinate systems is found in Sec. 9.1. Therefore, we transfer the polarimetric results (per camera c and pixel p) to the scattering coordinate system. The transformation is detailed in Sec. 9.2. The results of missing-view rendering in Fig. 10 appear in good agreement mutually and with Eq. (7).

Significant benefits of NeMF over prior art are seen in two aspects. First is *runtime*. The recovered surface area in this experiment is $3.6 \times 3.6\text{km}^2$. The physics-based approach took 32 minutes, as reported in [28], in order to converge to a degenerate microphysics field, that does not estimate v_e. NeMF, in contract, needs only 2 seconds on our computer, to infer a full microphysics field that varies in 3D. In other words, NeMF processes $\sim 6.5\text{km}^2/\text{sec}$, while the physics-based inverse polarization rendering analyzes only $0.007\text{km}^2/\text{sec}$. In accordance with the calculation in Sec. 3, NeMF should be able to meet expected remote sensing data rates, whereas the physics-based method would not.

The second significant benefit is the quality of 3D reconstruction. Consider[6] Fig. 11, the NeMF result is shaped much more like a cloud than the "spiky" physics-based retrieval. This is likely because NeMF learns, by training, the spatial characteristics of natural clouds. NeMF retrieves r_e

6. Voxels in which the estimated LWC $\leq 10^{-8}[\text{gr}/\text{m}^3]$ are pruned, being considered numerical artifacts.

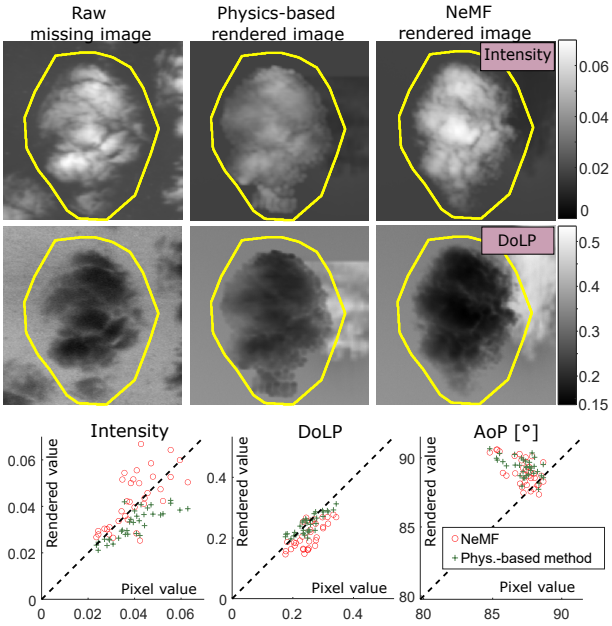


Fig. 10. Real-world experiment results. An AirMSPI polarimetric view was omitted from training. We show it, and images of the same view, rendered from 3D clouds that are estimated by either a physics-based approach, or NeMF. The region surrounded by a yellow line is derived by a space-carving mask. Scatter plots show sampled pixel values, per method. The recovered cloud is visualized in Fig. 11.

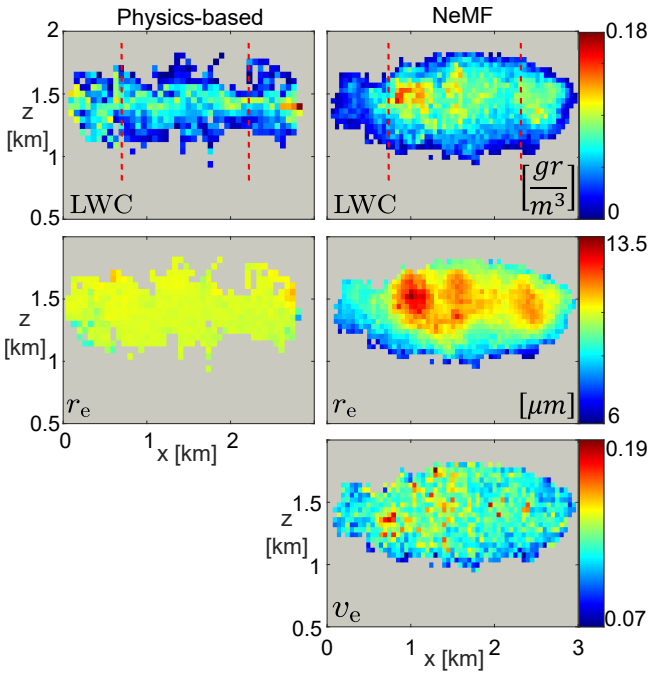


Fig. 11. The recovered cloud microphysics (by MIP), corresponding to the experiment of Fig. 10.

that varies in 3D significantly yet smoothly, as expected from a cloud and the discussion of the adiabatic approximation in Sec. 2.1: at the cloud core, r_e increases monotonically with altitude from the cloud base. Mixing, which undermines the adiabatic process, kicks-in at the cloud sides and top. In contrast, the physics-based retrieval behaves very differently. Moreover, NeMF estimates the v_e field.

Fig. 12 plots the vertical profile of microphysical parameters, at the cloud's core: voxels at the two external quarters of the cloud (marked in red in Fig. 11) in the lateral axes are not being used for the core. The analysis of Sec. 5 then uses $N_{\max} = 10$. Fig. 12 also plots a fit to the adiabatic model. Excluding the cloud top and sides, NeMF inference agrees with Eqs. (4,5), as explained in Sec. 2.1, using $\kappa_l = 0.22[\mathrm{gr}/(\mathrm{km} \cdot \mathrm{m}^3)]$, $\mathrm{LWC}^{\min} = 0.008[\mathrm{gr}/\mathrm{m}^3]$, $\kappa_r = 57.98[\mu\mathrm{m}/(\mathrm{km})^{1/3}]$ and $r_e^{\min} = 5.49[\mu\mathrm{m}]$.

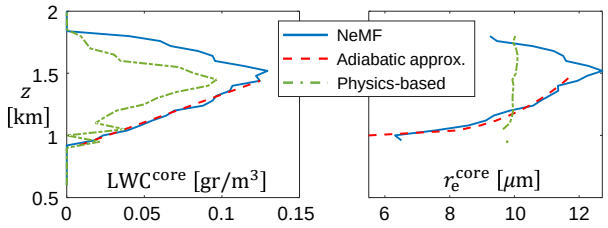


Fig. 12. Microphysics at the assessed cloud core, corresponding to the cloud in Figs. 10,11. The adiabatic model follows Eqs. (4,5).

9 Additional Details

This section explains some technical aspects which help making this paper self-contained.

9.1 Coordinate Systems

A line of sight corresponding to pixel p in camera c, passes through the center of projection of the camera at $\boldsymbol{x}_c$ in a *viewing direction* expressed by unit vector $\boldsymbol{\omega}_{c,p}$. Projection of light involving polarization carries an angular property. For this reason, we describe several relevant 3D coordinate systems, illustrated in Fig. 13.

• The meridian plane is spanned by $\boldsymbol{\omega}_{c,p}$ and the zenith vector $\hat{\mathbf{z}}$ in the observed region. The 3D *meridian system* is thus defined by

$$\hat{\boldsymbol{a}}_1^{\mathrm{mer}} = \frac{\boldsymbol{\omega}_{c,p} \times \hat{\mathbf{z}}}{\|\boldsymbol{\omega}_{c,p} \times \hat{\mathbf{z}}\|}, \quad \hat{\boldsymbol{a}}_2^{\mathrm{mer}} = \boldsymbol{\omega}_{c,p} \times \hat{\boldsymbol{a}}_1^{\mathrm{mer}}, \quad \boldsymbol{\omega}_{c,p}. \tag{25}$$

• The *scattering plane* is particularly important: spanned by the direction of solar radiance vector $\boldsymbol{\omega}_s$ and by $\boldsymbol{\omega}_{c,p}$, this plane directs *single scattering* of sunlight to the camera. Single-scattered sunlight is partially polarized perpendicular to this plane. Hence, relative to this plane, often Eq. (7) is satisfied. The *scattering coordinate system* is defined by

$$\hat{\boldsymbol{a}}_1^{\mathrm{sca}} = \frac{\boldsymbol{\omega}_{c,p} \times \boldsymbol{\omega}_s}{\|\boldsymbol{\omega}_{c,p} \times \boldsymbol{\omega}_s\|}, \quad \hat{\boldsymbol{a}}_2^{\mathrm{sca}} = \boldsymbol{\omega}_{c,p} \times \hat{\boldsymbol{a}}_1^{\mathrm{sca}}, \quad \boldsymbol{\omega}_{c,p}. \tag{26}$$

• Unit 3D vectors $\hat{\mathbf{X}}^{\mathrm{cam}}$ and $\hat{\mathbf{Y}}^{\mathrm{cam}}$ are parallel, respectively, to the rows and columns of the sensor pixel array of camera c. The 3D *pixel system* axes are

$$\hat{\boldsymbol{a}}_2^{\mathrm{pix}} = \frac{\boldsymbol{\omega}_{c,p} \times \hat{\mathbf{Y}}^{\mathrm{cam}}}{\|\boldsymbol{\omega}_{c,p} \times \hat{\mathbf{Y}}^{\mathrm{cam}}\|}, \quad \hat{\boldsymbol{a}}_1^{\mathrm{pix}} = \hat{\boldsymbol{a}}_2^{\mathrm{pix}} \times \boldsymbol{\omega}_{c,p}, \quad \boldsymbol{\omega}_{c,p}. \tag{27}$$

9.2 Forward Model

The RTE creates a light field of Stokes vectors based on a 3D field of microphysics. The light field is projected towards the camera at $\boldsymbol{x}_c$, in direction $\boldsymbol{\omega}_{c,p}$ corresponding to pixel p. An RTE solver represents Stokes vectors using the *meridian*

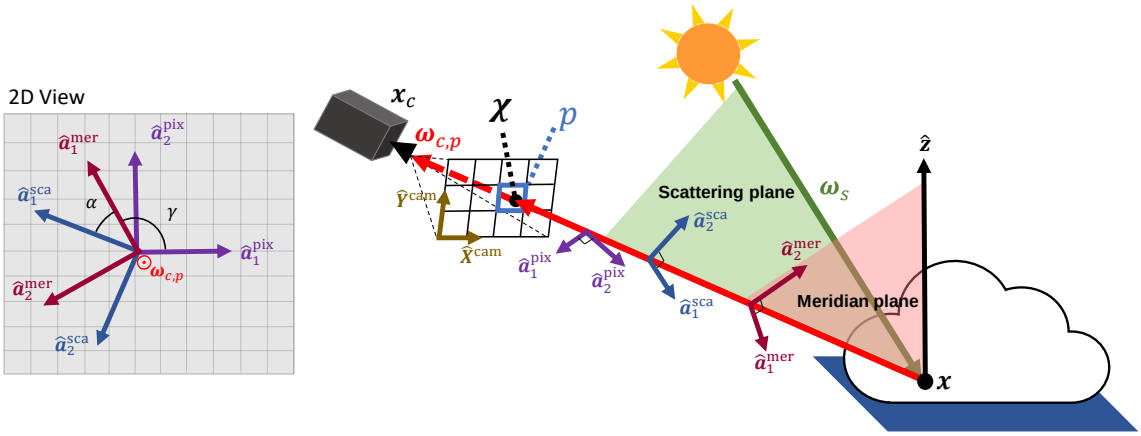


Fig. 13. There are several 3D coordinate systems in this paper, all involving the line of sight $\boldsymbol{\omega}_{c,p}$. In the left-hand-side panel, $\boldsymbol{\omega}_{c,p}$ is directed out of the page. The 3D vectors are thus projected to 2D in the panel.

coordinates (see Sec. 9.1). So, a simulated clean sample is denoted by the Stokes vector $\mathbf{I}^{\text{mer}}\left(\boldsymbol{x}_c, \boldsymbol{\omega}_{c,p}\right)$.

Consider two coordinate systems, rotated by angle ζ relative to each other. Transforming a Stokes vector between these systems is done by a Mueller rotation matrix [51], [52]

$$\mathbf{R}\left(\zeta\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\left(2\zeta\right) & \sin\left(2\zeta\right) \\ 0 & -\sin\left(2\zeta\right) & \cos\left(2\zeta\right) \end{bmatrix} . \tag{28}$$

Specifically, let α be the angle between the *meridian* and *scattering* systems, and γ the angle between the *meridian* and *pixel* systems (Fig. 13). These angles are defined as

$$\alpha = \arccos(\hat{\boldsymbol{a}}_1^{\text{mer}} \cdot \hat{\boldsymbol{a}}_1^{\text{scat}}), \quad \gamma = \arccos(\hat{\boldsymbol{a}}_1^{\text{mer}} \cdot \hat{\boldsymbol{a}}_1^{\text{pix}}). \tag{29}$$

Using Eqs. (28,29), the Stokes vector in the pixel system is

$$\mathbf{I}^{\text{pix}}[c,p] = \mathbf{R}\left(\gamma\right) \mathbf{I}^{\text{mer}}\left(\boldsymbol{x}_c, \boldsymbol{\omega}_{c,p}\right) . \tag{30}$$

Light then passes through a polarizing filter. The filter has angle η relative to the vector $\hat{\mathbf{X}}^{\text{cam}}$, which is aligned with the sensor pixel rows. Then, light is collected by pixel p and converted to photo-electrons. The polarization filter affects the *expected* number e of photo-electrons at pixel p of camera c [53]. This can be modeled by the inner product

$$e[c,p,\eta] = \frac{D}{2}\left[1 \;\; \cos(2\eta) \;\; \sin(2\eta)\right] \mathbf{I}^{\text{pix}}[c,p] . \tag{31}$$

Here the factor D accounts for several multiplicative imaging parameters, including exposure time, lens aperture area, pixel area, transmissivity of the optical system, spectral bandwidth and quantum efficiency.

As mentioned in Sec. 4.4, light spreads over a *super-pixel*, containing 2×2 sensor pixels, having a corresponding mosaic of four polarization filters at angles $[0°, 45°, 90°, 135°]$. We thus let p index a super-pixel. A super-pixel outputs a 4-element vector. Based on Eq. (31), its *expected* photo-electron readout is

$$\begin{aligned} \mathbf{e}[c,p] = \begin{bmatrix} e[c,p,0°] \\ e[c,p,45°] \\ e[c,p,90°] \\ e[c,p,135°] \end{bmatrix} &= \frac{D}{2} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \mathbf{I}^{\text{pix}}[c,p] \\ &\equiv D\mathbf{W}_{\text{pol}}\mathbf{I}^{\text{pix}}[c,p] . \end{aligned} \tag{32}$$

The rest of the imaging model is described in Sec. 4.4.

9.3 Loss Terms

In a training cloud scene, a batch of K voxels has corresponding values of $\hat{r}_{\text{e}}, \hat{v}_{\text{e}}, \text{L}\tilde{\text{W}}\text{C}, \hat{m}$ and $r_{\text{e}}^{\text{gt}}, v_{\text{e}}^{\text{gt}}, \text{LWC}^{\text{gt}}, m_{\text{gt}}$. They form column vectors $\hat{\mathbf{r}}_{\text{e}}, \hat{\mathbf{v}}_{\text{e}}, \mathbf{L}\tilde{\mathbf{W}}\mathbf{C}, \hat{\mathbf{m}}$ and $\mathbf{r}_{\text{e}}^{\text{gt}}, \mathbf{v}_{\text{e}}^{\text{gt}}, \mathbf{LWC}^{\text{gt}}, \mathbf{m}_{\text{gt}}$. Then,

$$\mathcal{L}_{\text{LWC}} = \frac{\|\mathbf{LWC}^{\text{gt}} - \mathbf{L}\tilde{\mathbf{W}}\mathbf{C}\|_2^2}{\|\mathbf{LWC}^{\text{gt}}\|_2} . \tag{33}$$

The values of m, m_{gt} may be interpreted as probabilities that a cloud exists at $\boldsymbol{x}$. Therefore, the mask loss is the binary cross-entropy (BCE),

$$\mathcal{L}_{\text{mask}} = -\frac{1}{K}\left[\mathbf{m}_{\text{gt}}^{\top} \log\left(\hat{\mathbf{m}}\right) + \left(\mathbf{1} - \mathbf{m}_{\text{gt}}\right)^{\top} \log\left(\mathbf{1} - \hat{\mathbf{m}}\right)\right] , \tag{34}$$

where $\mathbf{1}$ is a vector of ones, and $\log(\cdot)$ is element-wise.

In a location that has no cloud droplets at all, i.e., $m_{\text{gt}}(\boldsymbol{x}) = 0$, values assigned to $\hat{r}_{\text{e}}(\boldsymbol{x}), \hat{v}_{\text{e}}(\boldsymbol{x})$ are meaningless. They need to be ignored, and not affect the loss. Training should only rely[7] on voxels for which $m_{\text{gt}} = 1$. Define a diagonal matrix $\mathbf{M}_{\text{gt}} = \text{diag}(\mathbf{m}_{\text{gt}})$, listing $\mathbf{m}_{\text{gt}}$ on the main diagonal. Then,

$$\mathcal{L}_{\text{droplets}} = \frac{\|\mathbf{M}_{\text{gt}}(\mathbf{r}_{\text{e}}^{\text{gt}} - \hat{\mathbf{r}}_{\text{e}})\|_2^2}{\|\mathbf{m}_{\text{gt}}\|_2 \|\mathbf{M}_{\text{gt}}\mathbf{r}_{\text{e}}^{\text{gt}}\|_2} + \frac{\|\mathbf{M}_{\text{gt}}(\mathbf{v}_{\text{e}}^{\text{gt}} - \hat{\mathbf{v}}_{\text{e}})\|_2^2}{\|\mathbf{m}_{\text{gt}}\|_2 \|\mathbf{M}_{\text{gt}}\mathbf{v}_{\text{e}}^{\text{gt}}\|_2} . \tag{35}$$

10 Discussion and Outlook

NeMF is designed to retrieve multiple microphysical parameters in a 3D volumetric heterogeneous domain (a vector field). NeMF also assesses the shape of the object (cloud) in which the microphysical parameters apply. By observing many examples during training, NeMF implicitly learns the nature of the scenes. Hence, it embeds a prior, leading to results whose quality significantly surpasses physics-based methods. In analysis of clouds, NeMF achieves unprecedented retrieval of the effective variance of the droplet size distribution. As the other microphysical parameters, the effective variance is retrieved in 3D.

7. Training loss and inference error calculations only use voxels in which $m_{\text{gt}} = 1$. So, values of r_{e}^{gt} and v_{e}^{gt} where $m_{\text{gt}} = 0$ are irrelevant.

By investing computational effort during supervised training, inference is several orders of magnitude faster than physics-based optimization. As explained, this is important in order to avoid backlog of spaceborne data which accumulates over time.

Generalizations of this work can incorporate non-image inputs. These can include the surface wind at sea level, which affects the reflection of partially polarized light from water. Then, the system can train using different wind conditions, and infer the wind as well as the scatterers. Another generalization is to include as input the solar illumination direction: then, the system can train and infer using a variety of solar angles. The solar angle will be a known input to the system, because it is always known per location and time in scientific contexts.

We believe that NeMF can be adapted to other media. For example, clouds residing above freezing level may be composed of ice crystals, with or without liquid droplets.

Acknowledgments

We thank Aviad Levis, Jesse Loveridge and Eshkol Eytan for the advice. 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), and the Israel Science Foundation (ISF grant 2514/23).

References

[1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *Proc. ECCV*. Springer, 2020, pp. 405–421.

[2] D. Rückert, Y. Wang, R. Li, R. Idoughi, and W. Heidrich, "NeAT: Neural adaptive tomography," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 4, pp. 1–13, 2022.

[3] A. Reed, J. Kim, T. Blanford, A. Pediredla, D. Brown, and S. Jayasuriya, "Neural volumetric reconstruction for coherent synthetic aperture sonar," *ACM Transactions on Graphics (TOG)*, vol. 42, no. 4, pp. 1–20, 2023.

[4] V. Rudnev, M. Elgharib, W. Smith, L. Liu, V. Golyanik, and C. Theobalt, "NeRF for outdoor scene relighting," in *Proc. ECCV*. Springer, 2022, pp. 615–631.

[5] A. V. Sethuraman, M. S. Ramanagopal, and K. A. Skinner, "WaterNeRF: Neural radiance fields for underwater scenes," in *OCEANS - MTS/IEEE U.S. Gulf Coast*, 2023, pp. 1–7.

[6] D. Levy, A. Peleg, N. Pearl, D. Rosenbaum, D. Akkaynak, S. Korman, and T. Treibitz, "SeaThru-NeRF: Neural radiance fields in scattering media," in *Proc. IEEE/CVF CVPR*, June 2023, pp. 56–65.

[7] A. Ramazzina, M. Bijelic, S. Walz, A. Sanvito, D. Scheuble, and F. Heide, "ScatterNeRF: Seeing through fog with physically-based inverse neural rendering," in *Proc. IEEE/CVF ICCV*, 2023, pp. 17 957–17 968.

[8] W.-T. Chen, W. Yifan, S.-Y. Kuo, and G. Wetzstein, "DehazeNeRF: Multi-image haze removal and 3D shape reconstruction using neural radiance fields," in *Proc. IEEE 3DV 2024*. IEEE, 2024, pp. 247–256.

[9] F. Mu, S. Mo, J. Peng, X. Liu, J. H. Nam, S. Raghavan, A. Velten, and Y. Li, "Physics to the rescue: Deep non-line-of-sight reconstruction for high-speed imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[10] T. Chien, C. Ophus, and L. Waller, "Space-time implicit neural representations for atomic electron tomography on dynamic samples," in *NeurIPS 2023 Workshop on Deep Learning and Inverse Problems*, 2023.

[11] Y. Cheng, R. Wan, S. Weng, C. Zhu, Y. Chang, and B. Shi, "Colorizing monochromatic radiance fields," in *Proc. AAAI Conference on Artificial Intelligence*, vol. 38, no. 2, 2024, pp. 1317–1325.

[12] A. Malik, P. Mirdehghan, S. Nousias, K. Kutulakos, and D. Lindell, "Transient neural radiance fields for lidar view synthesis and 3d reconstruction," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[13] A. Dave, Y. Zhao, and A. Veeraraghavan, "PANDORA: Polarization-aided neural decomposition of radiance," in *Proc. ECCV*, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds. Springer, 2022, pp. 538–556.

[14] A. Levis, P. P. Srinivasan, A. A. Chael, R. Ng, and K. L. Bouman, "Gravitationally lensed black hole emission tomography," in *Proc. IEEE/CVF CVPR*, 2022, pp. 19 841–19 850.

[15] V. Pörtge, T. Kölling, A. Weber, L. Volkmer, C. Emde, T. Zinner, L. Forster, and B. Mayer, "High-spatial-resolution retrieval of cloud droplet size distribution from polarized observations of the cloudbow," *Atmospheric Measurement Techniques*, vol. 16, no. 3, pp. 645–667, 2023.

[16] A. Levis, Y. Y. Schechner, A. B. Davis, and J. Loveridge, "Multi-view polarimetric scattering cloud tomography and retrieval of droplet size," *Remote Sensing*, vol. 12, no. 17, p. 2831, 2020.

[17] S. H. Salter, "Solar radiation management, cloud albedo enhancement," in *Geoengineering Responses to Climate Change: Selected Entries from the Encyclopedia of Sustainability Science and Technology*. Springer New York, 2013, pp. 39–51.

[18] O. Boucher, D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann *et al.*, "Clouds and aerosols," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2013, pp. 571–657.

[19] F.-M. Breon and M. Doutriaux-Boucher, "A comparison of cloud droplet radii measured from space," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 8, pp. 1796–1805, 2005.

[20] M. Tzabari, V. Holodovsky, O. Shubi, E. Eytan, I. Koren, and Y. Y. Schechner, "Settings for spaceborne 3-D scattering tomography of liquid-phase clouds by the CloudCT mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.

[21] A. Levis, Y. Y. Schechner, and A. B. Davis, "Multiple-scattering microphysics tomography," in *Proc. IEEE CVPR*, 2017, pp. 6740–6749.

[22] A. Levis, Y. Y. Schechner, A. Aides, and A. B. Davis, "Airborne three-dimensional cloud tomography," in *Proc. IEEE ICCV*, 2015, pp. 3379–3387.

[23] Y. Sde-Chen, Y. Y. Schechner, V. Holodovsky, and E. Eytan, "3DeepCT: Learning volumetric scattering tomography of clouds," in *Proc. IEEE/CVF ICCV*, 2021, pp. 5671–5682.

[24] R. Ronen, V. Holodovsky, and Y. Y. Schechner, "Variable imaging projection cloud scattering tomography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[25] T. Loeub, A. Levis, V. Holodovsky, and Y. Y. Schechner, "Monotonicity prior for cloud tomography," in *Proc. ECCV*. Springer, 2020, pp. 24–29.

[26] J. Lin, M. Farinha, E. Gryspeerdt, and R. Clark, "Volumetric cloud field reconstruction," *arXiv preprint arXiv:2311.17657*, 2023.

[27] T. Nakajima and M. D. King, "Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory," *Journal of the Atmospheric Sciences*, vol. 47, no. 15, pp. 1878–1893, 1990.

[28] A. Levis, A. B. Davis, J. R. Loveridge, and Y. Y. Schechner, "3D cloud tomography and droplet size retrieval from multi-angle polarimetric imaging of scattered sunlight from above," in *Polarization Science and Remote Sensing X*, vol. 11833. SPIE, 2021, pp. 27–41.

[29] O. Avitan, Y. Y. Schechner, and E. Behar, "Using zodiacal light for spaceborne calibration of polarimetric imagers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

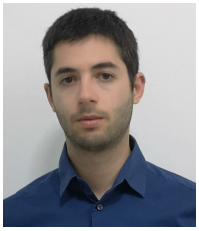
[30] K. Wyser, "The effective radius in large-scale models: impact of aerosols and coalescence," *Atmospheric Research*, vol. 49, no. 3, pp. 213–234, 1998.

[31] J. E. Hansen, "Multiple scattering of polarized light in planetary atmospheres, Part II. Sunlight reflected by terrestrial water clouds," *Journal of the Atmospheric Sciences*, vol. 28, no. 8, pp. 1400–1426, 1971.

[32] M. Klingebiel, A. de Lozar, S. Molleker, R. Weigel, A. Roth, L. Schmidt, J. Meyer, A. Ehrlich, R. Neuber, M. Wendisch *et al.*,

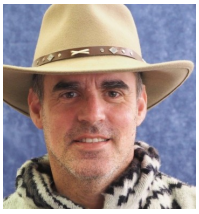
"Arctic low-level boundary layer clouds: in situ measurements and simulations of mono-and bimodal supercooled droplet size distributions at the top layer of liquid phase clouds," *Atmospheric Chemistry and Physics*, vol. 15, no. 2, pp. 617–631, 2015.
[33] M. D. Alexandrov, B. Cairns, K. Sinclair, A. P. Wasilewski, L. Ziemba, E. Crosbie, R. Moore, J. Hair, A. J. Scarino, Y. Hu *et al.*, "Retrievals of cloud droplet size from the research scanning polarimeter data: Validation using in situ measurements," *Remote Sensing of Environment*, vol. 210, pp. 76–95, 2018.
[34] S. Gassó and K. D. Knobelspiesse, "Circular polarization in atmospheric aerosols," *Atmospheric Chemistry and Physics*, vol. 22, no. 20, pp. 13 581–13 605, 2022.
[35] R. Chipman, W. S. T. Lam, and G. Young, *Polarized light and optical systems*. CRC press, 2018.
[36] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *International Journal of Computer Vision*, vol. 48, pp. 233–254, 2002.
[37] R. W. Bergstrom Jr, "Predictions of the spectral absorption and extinction coefficients of an urban air pollution aerosol model," *Atmospheric Environment (1967)*, vol. 6, no. 4, pp. 247–258, 1972.
[38] V. Holodovsky, Y. Y. Schechner, A. Levin, A. Levis, and A. Aides, "In-situ multi-view multi-scattering stochastic tomography," in *Proc. IEEE ICCP*, 2016, pp. 1–12.
[39] B. Mayer, "Radiative transfer in the cloudy atmosphere," in *EPJ web of Conferences*, vol. 1. EDP Sciences, 2009, pp. 75–99.
[40] J. Loveridge, A. Levis, L. Di Girolamo, V. Holodovsky, L. Forster, A. B. Davis, and Y. Y. Schechner, "Retrieving 3d distributions of atmospheric particles using atmospheric tomography with 3d radiative transfer–part 1: Model description and jacobian calculation," *Atmospheric Measurement Techniques*, vol. 16, no. 7, pp. 1803–1847, 2023.
[41] K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," *IJCV*, vol. 38, no. 3, pp. 199–218, 2000.
[42] D. Veikherman, A. Aides, Y. Y. Schechner, and A. Levis, "Clouds in the cloud," in *Proc. ACCV*. Springer, 2015, pp. 659–674.
[43] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and f-score, with implication for evaluation," in *ECIR*. Springer, 2005, pp. 345–359.
[44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, 2016, pp. 770–778.
[45] E. Eytan, A. Khain, M. Pinsky, O. Altaratz, J. Shpund, and I. Koren, "Shallow cumulus properties as captured by adiabatic fraction in high-resolution les simulations," *Journal of the Atmospheric Sciences*, vol. 79, no. 2, pp. 409–428, 2022.
[46] R. H. Heiblum, L. Pinto, O. Altaratz, G. Dagan, and I. Koren, "Core and margin in warm convective clouds–part 1: Core types and evolution during a cloud's lifetime," *Atmospheric Chemistry and Physics*, vol. 19, no. 16, pp. 10 717–10 738, 2019.
[47] A. Levis, J. Loveridge, and A. Aides, *Pyshdom. 2020. Available online*, https://github.com/aviadlevis/pyshdom.
[48] I. Kom Betzer, R. Ronen, V. Holodovsky, Y. Y. Schechner, and I. Koren, *NeMF's cloud datasets and rendered images. 2024. Available online*, https://technionmail-my.sharepoint.com/:f:/g/personal/hybridlab_technion_ac_il/EubYijnY3jBBgGBkeGDRoicBFW2qdi8xfYufLJxkth0w8Q?e=qSnqi7.
[49] R. Ronen, *VIP-CT code. 2022. Available online*, https://github.com/ronenroi/VIPCT.
[50] T. Johnson, M. Tsamados, J.-P. Muller, and J. Stroeve, "Mapping arctic sea-ice surface roughness with multi-angle imaging spectroradiometer," *Remote Sensing*, vol. 14, no. 24, p. 6249, 2022.
[51] M. I. Mishchenko, L. D. Travis, and A. A. Lacis, *Scattering, absorption, and emission of light by small particles*. Cambridge university press, 2002.
[52] S.-H. Baek, D. S. Jeon, X. Tong, and M. H. Kim, "Simultaneous acquisition of polarimetric svbrdf and normals." *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, p. 268, 2018.
[53] R. Wu, A. Jarabo, J. Suo, F. Dai, Y. Zhang, Q. Dai, and D. Gutierrez, "Adaptive polarization-difference transient imaging for depth estimation in scattering media," *Optics Letters*, vol. 43, no. 6, pp. 1299–1302, 2018.

Inbal Kom Betzer is an M.Sc. student at the Electrical and Computer Engineering Faculty, Technion - Israel Institute of Technology. She obtained a B.Sc. degree in Physics in 2017 and a B.Sc. degree in Electrical Engineering in 2019, both from the Technion, Israel. Her research interests primarily involve computer vision and optics, employing both physics-based and data-driven methodologies. Inbal won the Pazy Foundation Award for students in 2020 and was chosen as an Optica Woman Scholar in 2024.



Roi Ronen is a Ph.D. student at the Technion - Israel Institute of Technology and an applied scientist at Amazon Web Services - AI labs. His research explores the interface between machine learning, physics-based imaging and computational photography. He obtained his B.Sc. in 2019 and a M.Sc. in 2021, both at the Electrical and Computer Engineering Department, Technion, Israel. His paper won the JOSA A Emerging Researcher Best Paper Prize for 2021.

Vadim Holodovsky received his M.Sc. degree in Electrical Engineering, Technion, Israel, in 2016. He is currently a Senior Research Associate in the Hybrid-Imaging lab, Technion, Israel. His research interests include computer vision algorithms for 3D retrievals and remote sensing. Vadim's work includes experiment designs, optics, cameras, polarization, and electronic hardware.



Yoav Y. Schechner is the Diane and Mark Seiden Chair in Science. He is a principal investigator and coordinator of the ERC CloudCT project. He was a visiting scientist at Caltech and NASA's Jet Propulsion Laboratory (JPL), a visitor in MIT and research scientist at Columbia University. He graduated from the Technion - Israel Inst. of Technology: B.A. (Physics 1990), M.Sc. (Physics 1994), Ph.D. (EE 2000). He won the Best Student Paper Award at CVPR in 2017, the Best Paper Awards at ICCP in 2013 and 2018, and the Distinguished Lecturer Award, Technion 2020. His interests involve outdoor phenomena and all aspects of imaging.

Ilan Koren explores natural complex systems with a focus on clouds, rain, climate, and ocean atmosphere interactions. He received a Ph.D. degree in atmospheric physics from Tel Aviv University in 2002. He conducted postdoctoral research in NASA Goddard Space Flight Center's Climate and Radiation Branch before joining the Weizmann Institute of Science in 2005.