



Article

Linearizations of the Spherical Harmonic Discrete Ordinate Method (SHDOM)

Adrian Doicu * and Dmitry S. Efremenko 

Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany; Dmitry.Efremenko@dlr.de

* Correspondence: adrian.doicu@dlr.de

Received: 23 April 2019; Accepted: 22 May 2019; Published: 28 May 2019



Abstract: Linearizations of the spherical harmonic discrete ordinate method (SHDOM) by means of a forward and a forward-adjoint approach are presented. Essentially, SHDOM is specialized for derivative calculations and radiative transfer problems involving the delta-M approximation, the TMS correction, and the adaptive grid splitting, while practical formulas for computing the derivatives in the spherical harmonics space are derived. The accuracies and efficiencies of the proposed methods are analyzed for several test problems.

Keywords: 3D radiative transfer; SHDOM; Jacobian; linearization; adjoint radiative transfer

1. Introduction

Accurate and efficient modeling of radiative transfer is important for the retrieval of earth atmospheric constituent information by means of remote sensing. Real clouds are an inhomogeneous three-dimensional scattering medium. For the new generation of satellite spectrometers with a relative high spatial resolution (such as Sentinel 5 Precursor, Sentinel 4 and Sentinel 5 [1]), it is important to account for the sub-pixel cloud inhomogeneities, or at least, to assess their effect on the radiances at the top of the atmosphere, and in particular, on the trace gas and cloud retrieval results. So far, the majority of operational retrieval algorithms are based on the one-dimensional (1D) radiative transfer model, which neglects all horizontal variability of the atmosphere. However, 1D approximation can lead to the significant errors in the calculated radiances and retrieved parameters [2–6]. In this regard, three-dimensional (3D) radiative transfer models are required.

The SHDOM method developed by Evans is a widely used deterministic method to solve the 3D radiative transfer equation [7]. The radiative transfer equation is solved iteratively by using the spherical harmonic and the discrete ordinate representations of the radiance field. Each iteration consists of four steps: (1) the transformation of the source function from spherical harmonics to discrete ordinates; (2) the integration of the source function along discrete ordinate directions to compute the radiance field; (3) the transformation of the discrete ordinate radiances to spherical harmonics; and (4) the calculation of the source function from the radiance in spherical harmonics space. An important feature of SHDOM is the adaptive grid technique. This technique improves the solution accuracy by increasing the spatial resolution in regions where the source function is changing more rapidly. The difference scheme employed in SHDOM is a characteristic scheme. In the so-called long characteristic method, the integration is performed along the whole characteristic of the differential operator, while in the short characteristic method, the integration is performed along a segment of the characteristic located inside a cell. SHDOM has been modified for use on multiple processors [8], and to include polarization [9–11].

In atmospheric remote sensing and data assimilation, it is necessary to compute the partial derivatives of the outgoing radiances with respect to some atmospheric parameters of interest. For this

purpose, a linearized forward approach, relying on an analytical computation of the derivatives, or a linearized forward-adjoint approach, relying on the application of the adjoint radiative transfer theory, can be used. For a plane-parallel geometry, scalar and vector linearized forward approaches were described in Refs. [12,13], while linearized forward-adjoint approaches were proposed in [14–21]. The forward-adjoint method is extremely efficient because only two radiative transfer calculations are required for derivatives calculations.

A theoretical framework for three-dimensional remote sensing of the atmosphere with adjoint methods was described by Martin et al. [22]. In a subsequent paper, Martin and Hasekamp [23] used the linearized-forward approach to retrieve cloud extinction fields from multi-angle measurements. Although, the analysis was restricted to two-dimensional problems (a two-dimensional domain and a set of directions confined to the unit circle and characterized by one angular variable), the efficiency of adjoint methods for multi-dimensional remote sensing of the atmosphere and surface was demonstrated.

In this paper, linearizations of SHDOM by using a forward and a forward-adjoint approach are presented. After formulating the problem in Section 2, we discuss in Section 3 the fundamentals of the linearized SHDOM with a forward-adjoint approach. The accuracy and efficiency of the linearized models is analyzed in Section 4, while the final section of our paper contains a few concluding remarks.

2. Problem Formulation

We consider the solar radiative transfer in a rectangular cuboid of lengths L_x , L_y and L_z as shown in Figure 1. The top and the bottom faces of the prism are denoted by S_t and S_b , respectively, while the lateral faces are denoted by S_{1x} ($x = 0$), S_{2x} ($x = L_x$), S_{1y} ($y = 0$), and S_{2y} ($y = L_y$). The boundary-value problem for the total radiance at point \mathbf{r} in direction $\mathbf{\Omega}$ consists of the inhomogeneous differential equation (with the thermal emission term neglected)

$$\frac{dI}{ds}(\mathbf{r}, \mathbf{\Omega}) = -\sigma_{\text{ext}}(\mathbf{r})I(\mathbf{r}, \mathbf{\Omega}) + \frac{\sigma_{\text{sct}}(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \mathbf{\Omega}, \mathbf{\Omega}')I(\mathbf{r}, \mathbf{\Omega}') d\mathbf{\Omega}', \tag{1}$$

the top-of-atmosphere boundary condition

$$I(\mathbf{r}_t, \mathbf{\Omega}^-) = \frac{F_0}{|\mu_0|} \delta(\mathbf{\Omega}^- - \mathbf{\Omega}_0), \mathbf{r}_t \in S_t, \tag{2}$$

and the surface boundary condition

$$I(\mathbf{r}_b, \mathbf{\Omega}^+) = \frac{1}{\pi} \int_{\Omega^-} \rho(\mathbf{r}_b, \mathbf{\Omega}^+, \mathbf{\Omega}^-) |\mu^-| I(\mathbf{r}_b, \mathbf{\Omega}^-) d\mathbf{\Omega}^-, \mathbf{r}_b \in S_b. \tag{3}$$

At the horizontal boundaries we assume periodic boundary conditions, i.e.,

$$\begin{aligned} I(x = 0, y, z, \mathbf{\Omega}) &= I(x = L_x, y, z, \mathbf{\Omega}), \\ I(x, y = 0, z, \mathbf{\Omega}) &= I(x, y = L_y, z, \mathbf{\Omega}). \end{aligned} \tag{4}$$

Here, σ_{ext} and $\sigma_{\text{sct}} = \omega\sigma_{\text{ext}}$ are the extinction and scattering coefficients, respectively, ω is the single-scattering albedo, P is the phase function, ρ is the bidirectional surface reflection function, $\mathbf{\Omega}_0 = (\mu_0, \varphi_0)$ with $\mu_0 < 0$, is the solar direction, F_0 is the solar flux, $\mathbf{\Omega}^+$ and $\mathbf{\Omega}^-$ denote an upward and a downward direction, respectively, Ω is the unit sphere, and Ω^+ and Ω^- stand for the upper and lower unit hemispheres, respectively.

Let us define the forward transport operator \mathcal{L} by

$$\begin{aligned}
 (\mathcal{L}I)(\mathbf{r}, \Omega) = & \frac{dI}{ds}(\mathbf{r}, \Omega) + \sigma_{\text{ext}}(\mathbf{r})I(\mathbf{r}, \Omega) \\
 & - \frac{\sigma_{\text{sct}}(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega, \Omega') I(\mathbf{r}, \Omega') d\Omega' \\
 & - \frac{1}{\pi} \delta(z) H(\mu) \mu \int_{\Omega} \rho(\mathbf{r}, \Omega, \Omega') H(-\mu') |\mu'| I(\mathbf{r}, \Omega') d\Omega',
 \end{aligned}
 \tag{5}$$

and the forward source function $Q(\mathbf{r}, \Omega)$ by

$$Q(\mathbf{r}, \Omega) = F_0 \delta(z - L_z) \delta(\Omega - \Omega_0),
 \tag{6}$$

where δ is the Dirac delta function and H is the Heaviside step function. It can be shown that if the radiation field $I(\mathbf{r}, \Omega)$ solves the inhomogeneous operator equation

$$(\mathcal{L}I)(\mathbf{r}, \Omega) = Q(\mathbf{r}, \Omega),
 \tag{7}$$

with the boundary conditions (4),

$$I(\mathbf{r}_t, \Omega^-) = 0, \mathbf{r}_t \in S_t, \text{ and } I(\mathbf{r}_b, \Omega^+) = 0, \mathbf{r}_b \in S_b,
 \tag{8}$$

then $I(\mathbf{r}, \Omega)$ solves the radiative transfer Equation (1) with the boundary conditions (2), (3), and (4). The converse result is also true.

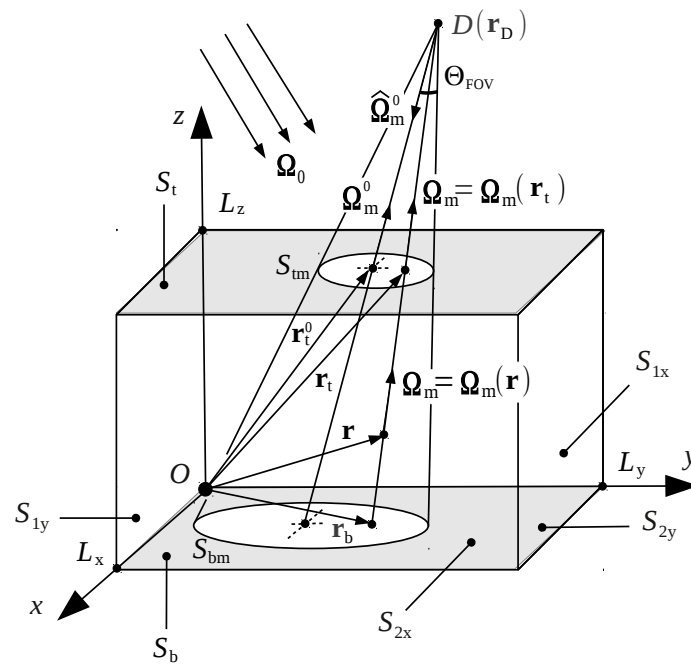


Figure 1. Illustration of the radiative transfer problem.

The adjoint transport operator \mathcal{L}^\dagger is defined through the relation

$$\langle \mathcal{L}I, I^\dagger \rangle = \langle I, \mathcal{L}^\dagger I^\dagger \rangle,
 \tag{9}$$

where the scalar product of the fields I_1 and I_2 is given by

$$\langle I_1, I_2 \rangle = \int_D \int_{\Omega} I_1(\mathbf{r}, \Omega) I_2(\mathbf{r}, \Omega) d\Omega dV
 \tag{10}$$

and D is the domain of the rectangular cuboid. The expression of the adjoint operator, under the assumptions that (i) I satisfies the boundary conditions (4) and (8), and (ii) I^\dagger satisfies the boundary conditions (4),

$$I^\dagger(\mathbf{r}_t, \Omega^+) = 0, \mathbf{r}_t \in S_t, \text{ and } I^\dagger(\mathbf{r}_b, \Omega^-) = 0, \mathbf{r}_b \in S_b, \tag{11}$$

is given by

$$\begin{aligned} (\mathcal{L}^\dagger I^\dagger)(\mathbf{r}, \Omega) &= -\frac{dI^\dagger}{ds}(\mathbf{r}, \Omega) + \sigma_{\text{ext}}(\mathbf{r})I^\dagger(\mathbf{r}, \Omega) \\ &\quad - \frac{\sigma_{\text{sct}}(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega', \Omega) I^\dagger(\mathbf{r}, \Omega') d\Omega' \\ &\quad - \frac{1}{\pi} \delta(z) H(-\mu) |\mu| \int_{\Omega} \rho(\mathbf{r}, \Omega', \Omega) H(\mu') \mu' I^\dagger(\mathbf{r}, \Omega') d\Omega'. \end{aligned} \tag{12}$$

Note that the requirement of homogeneous and periodic boundary conditions for both I and I^\dagger is essential when computing the adjoint of the differential operator d/ds by using the Gauss theorem over the domain D .

In our analysis we consider a situation which is typical for atmospheric remote sensing. Atmospheric radiation can be measured e.g., by a spectrometer, which in the following is called detector. The detector, which measures the radiance at the top of the atmosphere in the direction $\Omega_m^0 = (\mu_m^0, \varphi_m^0)$, $\mu_m^0 > 0$, is represented by a point with position vector $\mathbf{r}_D = (x_D, y_D, z_D)$ and is characterized by a characteristic function of its field of view $\chi_{\text{FOV}} = \chi_{\text{FOV}}(\Omega_m \cdot \Omega_m^0)$. The signal measured by the detector, can be written as

$$\mathcal{I} = \int_{\Omega} \chi_{\text{FOV}}(\Omega_m \cdot \Omega_m^0) I(\mathbf{r}_t(\Omega_m), \Omega_m) d\Omega_m, \tag{13}$$

where $I(\mathbf{r}_t(\Omega_m), \Omega_m)$ is the radiance at the top of the atmosphere in direction Ω_m . For the characteristic function χ_{FOV} , we consider the representation

$$\chi_{\text{FOV}}(\Theta) = \begin{cases} \frac{1}{\Omega_{\text{FOV}}}, & 0 \leq \Theta \leq \Theta_{\text{FOV}} \\ 0, & \text{otherwise} \end{cases}, \tag{14}$$

where Θ_{FOV} and $\Omega_{\text{FOV}} \approx \pi \Theta_{\text{FOV}}^2$ are the collecting polar and solid angles of the detector, respectively, and $\cos \Theta = \Omega_m \cdot \Omega_m^0$. Let $S_{\text{tm}} = S_{\text{tm}}(\mathbf{r}_D, \Omega_m^0)$ be the footprint of the detector on the top face S_t , i.e., an ellipse centered at $\mathbf{r}_t^0 = (x_t^0, y_t^0, L_z)$ with the semi-major and semi-minor axes $b = R_D \Theta_{\text{FOV}} / \mu_m^0$ and $a = R_D \Theta_{\text{FOV}}$, respectively, where $R_D = |\mathbf{r}_D - \mathbf{r}_t^0|$. Using the relation $d\Omega_m = (\mu_m^0 / R_D^2) dS$, where dS is the surface element on the top face S_t , we express the measured signal as

$$\mathcal{I} = \int_{S_t} h(\mathbf{r}_t) I(\mathbf{r}_t, \Omega_m(\mathbf{r}_t)) dS_t, \tag{15}$$

where

$$\Omega_m(\mathbf{r}_t) = \frac{\mathbf{r}_D - \mathbf{r}_t}{|\mathbf{r}_D - \mathbf{r}_t|} = (\mu_m(\mathbf{r}_t), \varphi_m(\mathbf{r}_t)), \tag{16}$$

is the unit vector along the line connecting the points $\mathbf{r}_t \in S_{\text{tm}}$ and \mathbf{r}_D ,

$$h(\mathbf{r}_t) = \begin{cases} \frac{1}{A_{\text{tm}}}, & \mathbf{r}_t \in S_{\text{tm}} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

is the characteristic function associated to the detector footprint, and $A_{tm} = \pi ab = R_D^2 \Omega_{FOV} / \mu_m^0$ is the area of the ellipse.

The radiative transfer in an inhomogeneous three-dimensional medium is usually solved by a numerical method. Considering a spatial discretization of the domain of analysis, the extinction coefficient $\sigma_{ext}(\mathbf{r})$, the single-scattering albedo $\omega(\mathbf{r})$, the Legendre expansion coefficients of the phase function $\chi_n(\mathbf{r})$ (Appendix A), are specified at every grid point of the so-called base grid and trilinearly interpolated within each cell. Furthermore, the bidirectional surface reflection function $\rho(\mathbf{r}_b, \Omega^+, \Omega^-)$ is specified at every grid point on the bottom surface and in all upward and downward discrete ordinate directions.

3. Linearized SHDOM

In atmospheric remote sensing we are interested in the retrieval of say, N_p atmospheric parameters $\xi_p, p = 1, \dots, N_p$, which determine the fields $\sigma_{ext}(\mathbf{r}), \omega(\mathbf{r}), \chi_n(\mathbf{r})$, and $\rho(\mathbf{r}_b, \Omega^+, \Omega^-)$. The retrieval requires the knowledge of the weighting functions, i.e., the partial derivatives of the measured radiance with respect to the atmospheric parameters being retrieved, i.e., $\partial \mathcal{I} / \partial \xi_p$.

3.1. Linearized Forward Approach

The linearized forward approach relies on the analytical differentiation technique and is based on the fact that the radiance can be regarded as the composition of differentiable functions. By applying the chain rule repeatedly to these functions, derivatives of arbitrary order can be computed analytically. After choosing the independent variables with respect to which the derivatives of the radiance are required, the derivatives of each function are computed recursively, i.e., by repeatedly substituting the derivatives of the inner function in the chain rule.

The pertinent equations which are differentiated are given in Appendix A. Essentially, the analytical differentiation technique is applied to all steps of the algorithm including the transformation of the source function from spherical harmonics to discrete ordinates, the integration of the source function along discrete ordinate directions, the transformation of the discrete ordinate radiances to spherical harmonics, the calculation of the source function from the radiance in spherical harmonics space, the adaptive grid procedure, and the acceleration step. The benefit of this method is that no assumptions rather than those of the forward model have to be made. The disadvantage is that not only the radiance and the source function have to be stored at all grid points, but also their derivatives with respect to the atmospheric parameters of interest. As a result, the amount of required memory may exceed the size of RAM (several Gbytes).

3.2. Linearized Forward-Adjoint Approach

In view of Equation (15), the measured signal can be written as

$$\mathcal{I} = \langle Q^\dagger, I \rangle, \tag{18}$$

where

$$Q^\dagger(\mathbf{r}, \Omega) = F_0^\dagger(\mathbf{r}_t) \delta(z - L_z) \delta(\Omega - \Omega_m(\mathbf{r})), \tag{19}$$

with

$$F_0^\dagger(\mathbf{r}_t) = h(\mathbf{r}_t), \tag{20}$$

is the adjoint source function. Here,

$$\Omega_m(\mathbf{r}) = \frac{\mathbf{r}_D - \mathbf{r}}{|\mathbf{r}_D - \mathbf{r}|} = (\mu_m(\mathbf{r}), \varphi_m(\mathbf{r})) \tag{21}$$

is the unit vector along the line joining the points \mathbf{r} and \mathbf{r}_D , and $\mathbf{r}_t = \mathbf{r} + s\mathbf{\Omega}_m(\mathbf{r}) \in S_{tm}$ with $s = (L_z - z)/\mu_m(\mathbf{r})$ and $\mathbf{r} = (x, y, z)$, is the intersection point of this line with the top face S_t (Figure 1). Obviously, because the points \mathbf{r} , \mathbf{r}_t , and \mathbf{r}_D are on the same line, we have $\mathbf{\Omega}_m(\mathbf{r}_t) = \mathbf{\Omega}_m(\mathbf{r})$, and so, $\mu_m(\mathbf{r}_t) = \mu_m(\mathbf{r})$.

The main result of the adjoint radiative transfer theory states that if (i) I solves the forward problem consisting in the operator equation $\mathcal{L}I = Q$ and the boundary conditions (4) and (8), and (ii) I^\dagger solves the adjoint problem consisting in the operator equation $\mathcal{L}^\dagger I^\dagger = Q^\dagger$ and the boundary conditions (4) and (11), then the measured signal \mathcal{I} can be estimated either by the scalar product of the radiance field I (as determined by Q) and the adjoint source function Q^\dagger , or by the scalar product of the adjoint radiance field I^\dagger (as determined by Q^\dagger) and the forward source function Q , i.e.,

$$\mathcal{I} = \langle Q^\dagger, I \rangle = \langle \mathcal{L}^\dagger I^\dagger, I \rangle = \langle I^\dagger, \mathcal{L}I \rangle = \langle I^\dagger, Q \rangle. \tag{22}$$

The partial derivative of the measured signal with respect to some atmospheric parameter ξ_p , can be computed as

$$\frac{\partial \mathcal{I}}{\partial \xi_p} = \left\langle Q^\dagger, \frac{\partial I}{\partial \xi_p} \right\rangle = \left\langle \mathcal{L}^\dagger I^\dagger, \frac{\partial I}{\partial \xi_p} \right\rangle = \left\langle I^\dagger, \mathcal{L} \frac{\partial I}{\partial \xi_p} \right\rangle, \tag{23}$$

where in the first step we assumed that $\partial Q^\dagger / \partial \xi_p = 0$, while in the last step we used the fact that $\partial I / \partial \xi_p$ satisfies the boundary conditions (4) and (8). Taking the partial derivative of the operator equation $\mathcal{L}I = Q$ with respect to ξ_p and assuming $\partial Q / \partial \xi_p = 0$ gives

$$\frac{\partial \mathcal{I}}{\partial \xi_p} = - \left\langle I^\dagger, \frac{\partial \mathcal{L}}{\partial \xi_p} I \right\rangle. \tag{24}$$

It should be pointed out that in general, the linearized forward-adjoint approach can be used to compute the weighted integral of the partial derivative of the radiance field $\partial I / \partial \xi_p$ defined by

$$D_p = \left\langle Q^\dagger, \frac{\partial I}{\partial \xi_p} \right\rangle. \tag{25}$$

In this case, employing the same arguments as in the derivation of Equations (23) and (24), but without assuming that $\partial Q^\dagger / \partial \xi_p = 0$, we get

$$D_p = - \left\langle I^\dagger, \frac{\partial \mathcal{L}}{\partial \xi_p} I \right\rangle. \tag{26}$$

Thus, identifying the adjoint source function from Equation (25) and solving the corresponding adjoint radiative transfer equation $\mathcal{L}^\dagger I^\dagger = Q^\dagger$, we can compute the weighted integral D_p by means of Equation (26).

The forward and adjoint radiative transfer equations $\mathcal{L}I = Q$ and $\mathcal{L}^\dagger I^\dagger = Q^\dagger$, respectively, are related to each other. Replacing $\mathbf{\Omega}$ by $-\mathbf{\Omega}$ in the expression of the adjoint operator \mathcal{L}^\dagger , gives $\mathcal{L}^\dagger(-\mathbf{\Omega})I^\dagger(\mathbf{r}, -\mathbf{\Omega}) = Q^\dagger(\mathbf{r}, -\mathbf{\Omega})$. Defining the pseudo-forward field \hat{I}^\dagger by the relation $\hat{I}^\dagger(\mathbf{r}, \mathbf{\Omega}) = I^\dagger(\mathbf{r}, -\mathbf{\Omega})$ and using the symmetry properties of the phase function $P(\mathbf{r}, -\mathbf{\Omega}, -\mathbf{\Omega}') = P(\mathbf{r}, \mathbf{\Omega}', \mathbf{\Omega})$ and of the normalized reflection function $\rho(\mathbf{r}_b, -\mathbf{\Omega}, -\mathbf{\Omega}') = \rho(\mathbf{r}_b, \mathbf{\Omega}', \mathbf{\Omega})$, we find the identity $\mathcal{L}^\dagger(-\mathbf{\Omega})I^\dagger(\mathbf{r}, -\mathbf{\Omega}) = \mathcal{L}(\mathbf{\Omega})\hat{I}^\dagger(\mathbf{r}, \mathbf{\Omega})$. Thus, the pseudo-forward field \hat{I}^\dagger solves the same type of radiative transfer equation as the radiance field I , i.e., $\mathcal{L}\hat{I}^\dagger = \hat{Q}^\dagger$, where the pseudo-forward source function, defined by $\hat{Q}^\dagger(\mathbf{r}, \mathbf{\Omega}) = Q^\dagger(\mathbf{r}, -\mathbf{\Omega})$, is

$$\hat{Q}^\dagger(\mathbf{r}, \mathbf{\Omega}) = \hat{F}_0^\dagger(\mathbf{r}_t)\delta(z - L_z)\delta(\mathbf{\Omega} - \hat{\mathbf{\Omega}}_m(\mathbf{r})), \tag{27}$$

with $\hat{\mathbf{\Omega}}_m(\mathbf{r}) = -\mathbf{\Omega}_m(\mathbf{r})$, $\hat{\mu}_m(\mathbf{r}) = -\mu_m(\mathbf{r})$, and

$$\hat{F}_0^\dagger(\mathbf{r}_t) = F_0^\dagger(\mathbf{r}_t) = h(\mathbf{r}_t). \tag{28}$$

In other words, $\hat{I}^\dagger(\mathbf{r}, \Omega)$ solves the inhomogeneous differential equation

$$\frac{d\hat{I}^\dagger}{ds}(\mathbf{r}, \Omega) = -\sigma_{\text{ext}}(\mathbf{r})\hat{I}^\dagger(\mathbf{r}, \Omega) + \frac{\sigma_{\text{sct}}(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega, \Omega')\hat{I}^\dagger(\mathbf{r}, \Omega') d\Omega' \quad (29)$$

with the top-of-atmosphere boundary condition

$$\hat{I}^\dagger(\mathbf{r}_t, \Omega^-) = \frac{\hat{F}_0^\dagger(\mathbf{r}_t)}{|\hat{\mu}_m(\mathbf{r}_t)|} \delta(\Omega^- - \hat{\Omega}_m(\mathbf{r}_t)), \quad \mathbf{r}_t \in S_t, \quad (30)$$

and the surface boundary condition

$$\hat{I}^\dagger(\mathbf{r}_b, \Omega^+) = \frac{1}{\pi} \int_{\Omega^-} \rho(\mathbf{r}_b, \Omega^+, \Omega^-) |\mu^-| \hat{I}^\dagger(\mathbf{r}_b, \Omega^-) d\Omega^-, \quad \mathbf{r}_b \in S_b. \quad (31)$$

We are now concern with the computation of the partial derivative $\partial\mathcal{I}/\partial\zeta_p$ by means of Equation (24). Taking the variation of the forward transport operator \mathcal{L} under the assumption that the phase function does not depend on the atmospheric parameter ζ_p , yields (cf. Equation (5))

$$\begin{aligned} (\delta\mathcal{L}I)(\mathbf{r}, \Omega) &= \delta\sigma_{\text{ext}}(\mathbf{r})I(\mathbf{r}, \Omega) \\ &\quad - \delta\sigma_{\text{ext}}(\mathbf{r}) \frac{\omega(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega, \Omega')I(\mathbf{r}, \Omega') d\Omega' \\ &\quad - \sigma_{\text{ext}}(\mathbf{r}) \frac{\delta\omega(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega, \Omega')I(\mathbf{r}, \Omega') d\Omega' \\ &\quad - \frac{1}{\pi} \delta(z)H(\mu)\mu \int_{\Omega} \delta\rho(\mathbf{r}, \Omega, \Omega')H(-\mu')|\mu'|I(\mathbf{r}, \Omega') d\Omega'. \end{aligned} \quad (32)$$

In this section we describe the computation of the partial derivative of the measured signal with respect to a parameter that determine the extinction field; the partial derivatives with respect to parameters that specify the single-scattering albedo and the bidirectional reflection function are given in Appendix B.

To compute the partial derivatives of the forward operator we need to find appropriate interpolation schemes for the extinction coefficient, source function, and radiance field within a grid cell. A major problem arises from the fact that SHDOM uses the integral form of the radiative transfer equation for radiance calculation, and in particular, the assumption that the extinction/source function product varies linearly along the characteristic. Unfortunately, when dealing with the differential form of the radiative transfer equation, this assumption can not be met; instead, we may assume that the extinction/source function product varies linearly within a cell. Thus, different interpolation schemes are used for radiance and derivative calculations.

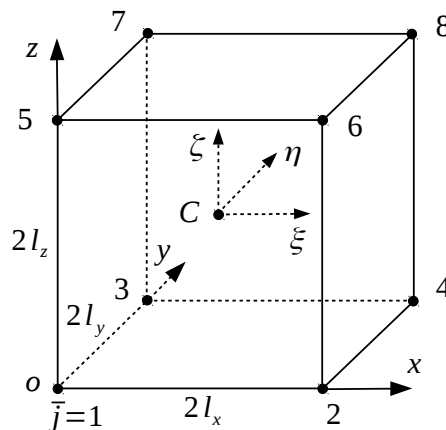


Figure 2. Rectangular cuboid element with lengths $2l_x$, $2l_y$, and $2l_z$. C is the center point of the cell, and $\bar{j} = 1, \dots, 8$ is the local index of a grid point.

In the following, we assume that the partial derivatives $\partial\sigma_{\text{ext}}(\mathbf{r}_i)/\partial\zeta_p$ at all grid points \mathbf{r}_i are known. Let c be the global index of a cell containing a grid point with global index j , and let \bar{j} be the local index of this point within cell c (Figure 2); we define the map g as $(\bar{j}, c) \xrightarrow{g} j$, or equivalently, $j = g(\bar{j}, c)$. For $\mathbf{r} \in D_c$, where D_c is the domain of cell c , we assume that the extinction and extinction/source function product vary linearly within the cell, i.e.,

$$\sigma_{\text{ext}}(\mathbf{r}) = \sum_{\bar{i}} L_{\bar{i}}(\mathbf{R})\sigma_{\text{ext}}(\mathbf{r}_{g(\bar{i},c)}) \tag{33}$$

and

$$\sigma_{\text{ext}}(\mathbf{r})J(\mathbf{r}, \Omega) = \sum_{\bar{i}} L_{\bar{i}}(\mathbf{R})\sigma_{\text{ext}}(\mathbf{r}_{g(\bar{i},c)})J(\mathbf{r}_{g(\bar{i},c)}, \Omega), \tag{34}$$

respectively. Here,

$$J(\mathbf{r}, \Omega) = \frac{\omega(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \Omega, \Omega')I(\mathbf{r}, \Omega') d\Omega' \tag{35}$$

is the source function, $L_{\bar{i}}$ are the first-order interpolation basis functions for a rectangular cuboid element (Appendix C), $\mathbf{R} = \mathbf{r} - \rho_c$ is the position vector of a point in a local coordinate system attached to cell c , $\rho_c = (1/8) \sum_{\bar{i}} \mathbf{r}_{g(\bar{i},c)}$ is the position vector of the center point of cell c , $\mathbf{r}_{g(\bar{i},c)} = \mathbf{r}_i$ is the position vector of the grid point with global index $i = g(\bar{i}, c)$, and the sum $\sum_{\bar{i}}$ is taken over all grid points of a cell. Setting $\sigma_{\text{ext}g(\bar{i},c)} = \sigma_{\text{ext}}(\mathbf{r}_{g(\bar{i},c)})$, we obtain

$$\left(\frac{\partial\mathcal{L}}{\partial\zeta_p} I\right)(\mathbf{r}, \Omega) = \sum_{\bar{i}} L_{\bar{i}}(\mathbf{R})[I(\mathbf{r}, \Omega) - J(\mathbf{r}_{g(\bar{i},c)}, \Omega)] \frac{\partial\sigma_{\text{ext}g(\bar{i},c)}}{\partial\zeta_p}, \tag{36}$$

implying

$$\frac{\partial\mathcal{L}}{\partial\zeta_p} = - \sum_c \sum_{\bar{i}} \int_{\Omega} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \hat{I}^+(\mathbf{r}, -\Omega) [I(\mathbf{r}, \Omega) - J(\mathbf{r}_{g(\bar{i},c)}, \Omega)] \frac{\partial\sigma_{\text{ext}g(\bar{i},c)}}{\partial\zeta_p} dV d\Omega. \tag{37}$$

In the next step, we split the total radiance $I(\mathbf{r}, \Omega)$ into the diffuse and direct radiances, i.e.,

$$I(\mathbf{r}, \Omega) = I_d(\mathbf{r}, \Omega) + \delta(\Omega - \Omega_0)T_0(\mathbf{r}), \tag{38}$$

where

$$T_0(\mathbf{r}) = \frac{F_0}{|\mu_0|} e^{-\tau_{\text{ext}}(\mathbf{r}, \mathbf{r}_0|\Omega_0)} \tag{39}$$

is the transmission along the solar direction Ω_0 from the starting point $\mathbf{r}_0 = \mathbf{r} - s\Omega_0 \in S_t$ to the end point \mathbf{r} (the forward direct beam), and

$$\tau_{\text{ext}}(\mathbf{r}, \mathbf{r}_0|\Omega_0) = \int_0^s \sigma_{\text{ext}}(\mathbf{r}_0 + s'\Omega_0) ds' \tag{40}$$

is the optical depth. Similarly, we split the pseudo-forward radiance $\hat{I}^+(\mathbf{r}, \Omega)$ into its diffuse and direct components, i.e.,

$$\hat{I}^+(\mathbf{r}, \Omega) = \hat{I}_d^+(\mathbf{r}, \Omega) + \delta(\Omega - \hat{\Omega}_m(\mathbf{r}))\hat{T}_m^+(\mathbf{r}), \tag{41}$$

where

$$\hat{T}_m^+(\mathbf{r}) = \frac{\hat{F}_0^+(\mathbf{r}_t)}{|\hat{\mu}_m(\mathbf{r}_t)|} e^{-\tau_{\text{ext}}(\mathbf{r}, \mathbf{r}_t|\hat{\Omega}_m(\mathbf{r}))}, \tag{42}$$

is the transmission along the detector direction $\widehat{\Omega}_m(\mathbf{r})$ from the starting point $\mathbf{r}_t = \mathbf{r} - s\widehat{\Omega}_m(\mathbf{r}) \in S_{tm}$ to the end point \mathbf{r} (the pseudo-forward direct beam). Approximating

$$\Omega_m(\mathbf{r}) = -\widehat{\Omega}_m(\mathbf{r}) \approx \Omega_m(\rho_c) := \Omega_{mc}, \quad \mathbf{r} \in D_c, \tag{43}$$

we are led to the following representation for $\partial\mathcal{I}/\partial\zeta_p$ in terms of diffuse radiances:

$$\frac{\partial\mathcal{I}}{\partial\zeta_p} = T_A + T_B + T_C + T_D + T_E, \tag{44}$$

where

$$T_A = -\sum_c \sum_{\bar{i}} \int_{\Omega} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \widehat{I}_d^{\dagger}(\mathbf{r}, -\Omega) I_d(\mathbf{r}, \Omega) \frac{\partial\sigma_{\text{extg}(\bar{i},c)}}{\partial\zeta_p} dV d\Omega, \tag{45}$$

$$T_B = \sum_c \sum_{\bar{i}} \int_{\Omega} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \widehat{I}_d^{\dagger}(\mathbf{r}, -\Omega) J(\mathbf{r}_{g(\bar{i},c)}, \Omega) \frac{\partial\sigma_{\text{extg}(\bar{i},c)}}{\partial\zeta_p} dV d\Omega, \tag{46}$$

and

$$T_C = -\sum_c \sum_{\bar{i}} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \widehat{I}_d^{\dagger}(\mathbf{r}, -\Omega_0) T_0(\mathbf{r}) \frac{\partial\sigma_{\text{extg}(\bar{i},c)}}{\partial\zeta_p} dV, \tag{47}$$

$$T_D = -\sum_c \sum_{\bar{i}} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \widehat{T}_m^{\dagger}(\mathbf{r}) I_d(\mathbf{r}, \Omega_{mc}) \frac{\partial\sigma_{\text{extg}(\bar{i},c)}}{\partial\zeta_p} dV, \tag{48}$$

$$T_E = \sum_c \sum_{\bar{i}} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \widehat{T}_m^{\dagger}(\mathbf{r}) J(\mathbf{r}_{g(\bar{i},c)}, \Omega_{mc}) \frac{\partial\sigma_{\text{extg}(\bar{i},c)}}{\partial\zeta_p} dV. \tag{49}$$

The above integrals are computed in the spherical harmonics space by assuming the expansions

$$I_d(\mathbf{r}, \Omega) = \sum_{mn} I_{mn}(\mathbf{r}) Y_{mn}(\Omega), \tag{50}$$

$$\widehat{I}_d^{\dagger}(\mathbf{r}, \Omega) = \sum_{mn} \widehat{I}_{mn}^{\dagger}(\mathbf{r}) Y_{mn}(\Omega), \tag{51}$$

$$J(\mathbf{r}, \Omega) = \sum_{mn} J_{mn}(\mathbf{r}) Y_{mn}(\Omega), \tag{52}$$

where $Y_{mn}(\Omega)$ are the orthonormal real-valued spherical harmonic functions (Appendix A), and the sum \sum_{mn} should be understood as

$$\sum_{mn} = \sum_{m=-M}^M \sum_{n=|m|}^N,$$

with N and M being the maximum expansion order and number of azimuthal modes, respectively. The linear approximation

$$f(\mathbf{r}) = \sum_{\bar{j}} L_{\bar{j}}(\mathbf{R}) f(\mathbf{r}_{g(\bar{j},c)}) = \sum_{\bar{j}} L_{\bar{j}}(\mathbf{R}) f_{g(\bar{j},c)}, \tag{53}$$

where $f(\mathbf{r})$ stands for $I_{mn}(\mathbf{r})$, $\hat{I}_{mn}^\dagger(\mathbf{r})$, $T_0(\mathbf{r})$, and $\hat{T}_m^\dagger(\mathbf{r})$, together with the identity $Y_{mn}(-\mathbf{\Omega}) = (-1)^n Y_{mn}(\mathbf{\Omega})$, then yield

$$T_A = - \sum_c \sum_{mn} (-1)^n \sum_{\bar{i}, \bar{j}, \bar{k}} \hat{I}_{mn, g(\bar{j}, c)}^\dagger I_{mn, g(\bar{k}, c)} \frac{\partial \sigma_{\text{extg}(\bar{i}, c)}}{\partial \xi_p} L_{\bar{i}, \bar{j}, \bar{k}}, \tag{54}$$

$$T_B = \sum_c \sum_{mn} (-1)^n \sum_{\bar{i}, \bar{j}} \hat{I}_{mn, g(\bar{j}, c)}^\dagger J_{mn, g(\bar{i}, c)} \frac{\partial \sigma_{\text{extg}(\bar{i}, c)}}{\partial \xi_p} L_{\bar{i}, \bar{j}}, \tag{55}$$

and

$$T_C = - \sum_c \sum_{\bar{i}, \bar{j}, \bar{k}} \hat{I}_d^\dagger(\mathbf{r}_{g(\bar{j}, c)}, -\mathbf{\Omega}_0) T_{0g(\bar{k}, c)} \frac{\partial \sigma_{\text{extg}(\bar{i}, c)}}{\partial \xi_p} L_{\bar{i}, \bar{j}, \bar{k}}, \tag{56}$$

$$T_D = - \sum_c \sum_{\bar{i}, \bar{j}, \bar{k}} I_d(\mathbf{r}_{g(\bar{j}, c)}, \mathbf{\Omega}_{mc}) \hat{T}_{mg(\bar{k}, c)}^\dagger \frac{\partial \sigma_{\text{extg}(\bar{i}, c)}}{\partial \xi_p} L_{\bar{i}, \bar{j}, \bar{k}}, \tag{57}$$

$$T_E = \sum_c \sum_{mn} Y_{mn}(\mathbf{\Omega}_{mc}) \sum_{\bar{i}, \bar{j}} \hat{T}_{mg(\bar{j}, c)}^\dagger J_{mn, g(\bar{i}, c)} \frac{\partial \sigma_{\text{extg}(\bar{i}, c)}}{\partial \xi_p} L_{\bar{i}, \bar{j}}. \tag{58}$$

Here, the interpolations coefficients $L_{\bar{i}, \bar{j}, \bar{k}}$ and $L_{\bar{i}, \bar{j}}$ are given, respectively, by

$$L_{\bar{i}, \bar{j}, \bar{k}} = \int_{D_c} L_{\bar{i}}(\mathbf{R}) L_{\bar{j}}(\mathbf{R}) L_{\bar{k}}(\mathbf{R}) dV, \tag{59}$$

$$L_{\bar{i}, \bar{j}} = \int_{D_c} L_{\bar{i}}(\mathbf{R}) L_{\bar{j}}(\mathbf{R}) dV, \tag{60}$$

while for $i = g(\bar{i}, c)$ and $\omega_i = \omega(\mathbf{r}_i)$, the expansion coefficients of the source function in Equations (55) and (58) are computed as

$$J_{mn, i} = \omega_i \frac{\chi_{n, i}}{2n + 1} I_{mn, i} + \omega_i \frac{\chi_{n, i}}{2n + 1} Y_{mn}(\mathbf{\Omega}_0) T_{0i}, \tag{61}$$

where $\chi_{n, i} = \chi_n(\mathbf{r}_i)$ are the Legendre expansion coefficients of the phase function (Appendix A), i.e.,

$$P(\mathbf{r}, \mathbf{\Omega}, \mathbf{\Omega}') = 4\pi \sum_{mn} \frac{\chi_n(\mathbf{r})}{2n + 1} Y_{mn}(\mathbf{\Omega}) Y_{mn}(\mathbf{\Omega}'). \tag{62}$$

Further comments and algorithm details are given below:

1. The input parameters of the numerical algorithm are the partial derivatives $\partial \sigma_{\text{ext}}(\mathbf{r}_i) / \partial \xi_p$ and $\partial \omega(\mathbf{r}_i) / \partial \xi_p$ at all grid points \mathbf{r}_i , and eventually, the partial derivatives $\partial \rho(\mathbf{r}_{bi}, \mathbf{\Omega}_{ij}^+, \mathbf{\Omega}_{pq}^-) / \partial \xi_p$ at all grid points on the bottom surface \mathbf{r}_{bi} and in all upward and downward discrete ordinate directions $\mathbf{\Omega}_{ij}^+$ and $\mathbf{\Omega}_{pq}^-$, respectively.
2. The interpolation coefficients $L_{\bar{i}, \bar{j}, \bar{k}}$ and $L_{\bar{i}, \bar{j}}$ are computed analytically in a local coordinate system attached to the grid cell.
3. The radiances $\hat{I}_d^\dagger(\mathbf{r}_{g(\bar{j}, c)}, -\mathbf{\Omega}_0)$ and $I_d(\mathbf{r}_{g(\bar{j}, c)}, \mathbf{\Omega}_{mc})$ in the expressions of T_C and T_D , respectively, are computed by the source integration method.
4. For solar problems with the delta-M method [24], the radiative transfer equation is expressed in terms of the scaled quantities

$$\bar{\sigma}_{\text{ext}} = (1 - f\omega)\sigma_{\text{ext}}, \quad \bar{\omega} = \frac{1 - f}{1 - f\omega}\omega, \tag{63}$$

where f is the truncation fraction. At each grid point, we switch from the partial derivatives $\partial\sigma_{\text{ext}}/\partial\zeta_p$ and $\partial\omega/\partial\zeta_p$ to $\partial\bar{\sigma}_{\text{ext}}/\partial\zeta_p$ and $\partial\bar{\omega}/\partial\zeta_p$, respectively, whereby the latter two are computed from Equation (63) by using the chain rule.

5. The partial derivatives of the signal correction in the TMS method of Nakajima and Tanaka [25] are computed analytically (Appendix A).
6. During the adaptive grid procedure, a base grid cell, also called a “parent cell”, is split into “child cells” to achieve higher spatial resolution. In this regard, the sums in Equations (54)–(58) are taken over all child cells.
7. According to the main result of the adjoint radiative transfer theory, the measured signal can be computed by using Equation (15) or (22); in the latter case, the computational formula is

$$\mathcal{I} = F_0 \int_{S_t} \hat{I}_d^\dagger(x_t, y_t, L_z, -\mathbf{\Omega}_0) dS_t. \tag{64}$$

8. The code uses two alternative interpolation schemes for the source function: a linear variation of the extinction/source function product within a cell, i.e., Equation (34), and a linear variation of the source function within a cell, i.e.,

$$J(\mathbf{r}, \mathbf{\Omega}) = \sum_{\bar{i}} L_{\bar{i}}(\mathbf{R}) J(\mathbf{r}_{g(\bar{i},c)}, \mathbf{\Omega}), \tag{65}$$

for $\mathbf{r} \in D_c$. In the second case, the terms T_B and T_E are computed, respectively, as

$$T_B = \sum_c \sum_{mn} (-1)^n \sum_{\bar{i}, \bar{j}, \bar{k}} \hat{I}_{mn,g(\bar{j},c)}^\dagger J_{mn,g(\bar{k},c)} \frac{\partial \sigma_{\text{ext}g(\bar{i},c)}}{\partial \zeta_p} L_{\bar{i}, \bar{j}, \bar{k}}, \tag{66}$$

$$T_E = \sum_c \sum_{mn} Y_{mn}(\mathbf{\Omega}_{mc}) \sum_{\bar{i}, \bar{j}, \bar{k}} \hat{I}_{mg(\bar{j},c)}^\dagger J_{mn,g(\bar{k},c)} \frac{\partial \sigma_{\text{ext}g(\bar{i},c)}}{\partial \zeta_p} L_{\bar{i}, \bar{j}, \bar{k}}. \tag{67}$$

9. In the case of satellite remote sensing we may assume that (i) the distance from the top of the atmosphere to the detector R_D is sufficiently large, so that we can approximate $I(\mathbf{r}, \mathbf{\Omega}_m(\mathbf{r})) \approx I(\mathbf{r}, \mathbf{\Omega}_m^0)$, and (ii) the footprint of the detector is a rectangle of lengths $2a$ and $2b$ centered at \mathbf{r}_t^0 . The second assumption is made because the domain is discretized in rectangular cuboid elements. In this context, the measured signal is computed as

$$\mathcal{I} = \frac{1}{4ab} \int_{S_t} \bar{h}(x_t, y_t) I(x_t, y_t, L_z, \mathbf{\Omega}_m^0) dS_t, \tag{68}$$

where the normalized characteristic function $\bar{h}(x_t, y_t)$ takes the value 1 inside the footprint of the detector and 0 otherwise.

10. The solution of the adjoint problem is a challenging task due to the spatial discontinuity of the pseudo-forward direct beam. In SHDOM, this type of problem is handled by the adaptive grid procedure. Essentially, the adaptive grid supplies extra spatial resolution along the boundaries of the pseudo-forward direct beam, so that there are not large discontinuities in the grid. In order to reduce the number of adaptive grid cells, the step characteristic function of the detector can be replaced by a smooth function, e.g., trapezoidal, Gaussian, cosine [23]. In the first case, the normalized characteristic function reads as

$$\bar{h}(x_t, y_t) = \begin{cases} \bar{h}_0(x_t, \Delta h_x) \bar{h}_0(y_t, \Delta h_y), & (x_t, y_t) \in S_{tm} \\ 0, & \text{otherwise} \end{cases}, \tag{69}$$

where, for example, $\bar{h}_0(x_t, \Delta h_x)$ is given by

$$\bar{h}_0(x_t, \Delta h_x) = c_x \times \begin{cases} \frac{x_t - x_{t\min}^0}{\Delta h_x}, & x_{t\min}^0 \leq x_t < x_{t\min}^0 + \Delta h_x \\ 1, & x_{t\min}^0 + \Delta h_x \leq x_t \leq x_{t\max}^0 - \Delta h_x \\ \frac{x_{t\max}^0 - x_t}{\Delta h_x}, & x_{t\max}^0 - \Delta h_x < x_t \leq x_{t\max}^0 \end{cases} \quad (70)$$

with $c_x = 2a / (2a - \Delta h_x)$, $x_{t\min}^0 = x_t^0 - a$, and $x_{t\max}^0 = x_t^0 + a$.

11. In the linearized forward-adjoint approach, the forward and adjoint problems are solved successively on the same grid; in this way, the interpolation between different grids is avoided. Actually, in the first step, the adjoint problem is solved by using the adaptive grid procedure with a prescribed splitting accuracy ϵ_{split} (Appendix A), and in the second step, the forward problem is solved on the resulting grid without splitting. In the linearized forward approach, the forward problem is solved first by using the adaptive grid procedure, and then, the partial derivatives are computed on the same grid; in this way, a less amount of memory for storing the derivatives of the radiance and the source function with respect to the atmospheric parameters of interest is required.

4. Numerical Simulations

In this section we analyze the accuracy and efficiency of the linearized versions of SHDOM.

4.1. Example 1

We consider a base grid with $N_x = N_y = 20$ and $N_z = 11$ points, and the discretization steps $\Delta x = \Delta y = \Delta z = 0.1$ km. For periodic boundary conditions, when the vertical boundaries in x and y have additional planes of grid points that duplicate the grid points in the planes $x = 0$ and $y = 0$, respectively, the domain of analysis has the sizes $L_x = N_x \Delta x = 2$ km, $L_y = N_y \Delta y = 2$ km, and $L_z = (N_z - 1) \Delta z = 1$ km. The extinction field is Gaussian, i.e.,

$$\sigma_{\text{ext}}(x, y, z) = \sigma_{\text{ext}}^{\max} f(x, y, z), \quad (71)$$

$$f(x, y, z) = f_G(x, x_0, s_x) f_G(y, y_0, s_y) f_G(z, z_0, s_z), \quad (72)$$

$$f_G(u, u_0, s_u) = \exp\left[-\frac{(u - u_0)^2}{2s_u^2}\right], \quad u = x, y, z, \quad (73)$$

with $x_0 = L_x/2$, $y_0 = L_y/2$, $z_0 = L_z/2$, $s_x = L_x/4$, $s_y = L_y/4$, and $s_z = L_z/4$. The parameters $\sigma_{\text{ext}}^{\max}$ is 6 km^{-1} , in which case, the peak optical depth is 3.6. The extinction field in the plane $y = L_y/2$ is illustrated in Figure 3. The number of discrete zenith and azimuthal angles are $N_\mu = 16$ and $N_\varphi = 2N_\mu$, respectively, the single-scattering albedo is $\omega = 0.9$, and a Henyey–Greenstein phase function [26] with the asymmetry parameter $g = 0.8$ is considered. In all our simulations, the solar zenith angle is $\theta_0 = 120^\circ$, and a Lambertian reflecting surface with the surface albedo $A = 0.2$ is chosen. The coordinates of the center of the detector footprint are $x_t^0 = L_x/2$, $y_t^0 = L_y/2$, and $z_t^0 = L_z$, the detector azimuthal angle is $\varphi_m^0 = 0^\circ$, the lengths of the rectangular footprint are $a = 5\Delta x$, and $b = [a / (\Delta y \mu_m^0)] \Delta y$, where $[x]$ is the integer part of x , and the lengths specifying the slope of the characteristic function in Equation (70) are $\Delta h_x = 0.5\Delta x$ and $\Delta h_y = 0.5\Delta y$.

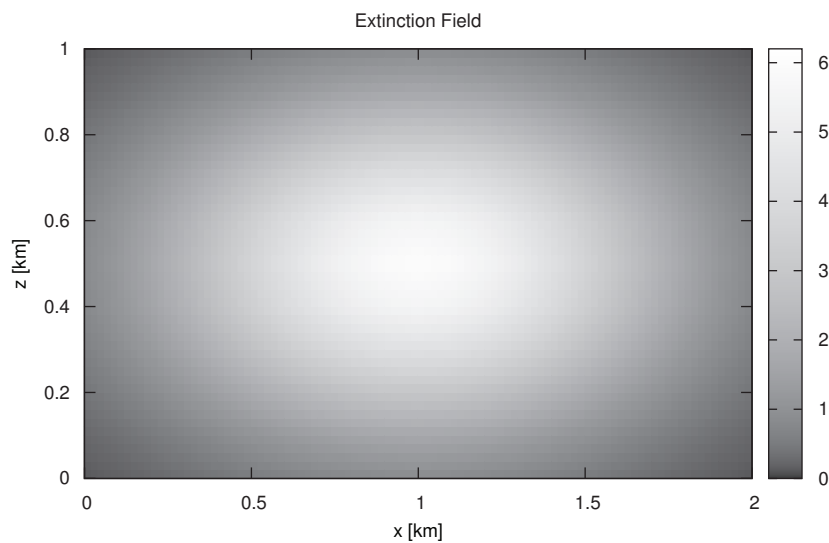


Figure 3. Extinction field in the plane $y = L_y/2$.

In the following, we compute the partial derivatives of the measured signal with respect to the extinction coefficients σ_{extp} at $N_p = 11$ base grid points ($x = x_t^0, y = y_t^0, z = \bar{p}\Delta z$), $\bar{p} = 1, \dots, N_p$. Obviously, if p is the global index of the grid point with the local index \bar{p} in the derivative list, we have

$$\frac{\partial \sigma_{exti}}{\partial \sigma_{extp}} = \delta_{ip}, \quad \frac{\partial \omega_i}{\partial \sigma_{extp}} = 0. \tag{74}$$

First, we compare the accuracies of the measured signal and of its derivatives. The accuracy of SHDOM radiances is determined by the angular and spatial resolution, i.e., the number of discrete ordinates in zenith angle N_μ and, for a specified base grid, the splitting accuracy ϵ_{split} . In Figure 4, we plot the partial derivatives computed with the linearized forward-adjoint approach and the linearized approach for different values of the splitting accuracy ϵ_{split} . Note that the influence of N_μ on the results accuracy is rather low. Taking the values corresponding to $\epsilon_{split}^0 = 5 \times 10^{-4}$ as a reference, we illustrate in Table 1, the relative error in the partial derivatives, expressed as an rms difference normalized by the mean of the quantity over the N_p comparison points, i.e.,

$$\epsilon_{\partial \mathcal{I}} = \frac{\sqrt{\sum_{\bar{p}} \left[\left(\frac{\partial \mathcal{I}}{\partial \sigma_{extp}} \right)_{\epsilon_{split}} - \left(\frac{\partial \mathcal{I}}{\partial \sigma_{extp}} \right)_{\epsilon_{split}^0} \right]^2}}{\sqrt{\sum_{\bar{p}} \left(\frac{\partial \mathcal{I}}{\partial \sigma_{extp}} \right)_{\epsilon_{split}^0}^2}}, \tag{75}$$

and the relative error in the measured signal, i.e.,

$$\epsilon_{\mathcal{I}} = \frac{|\mathcal{I}_{\epsilon_{split}} - \mathcal{I}_{\epsilon_{split}^0}|}{\mathcal{I}_{\epsilon_{split}^0}}. \tag{76}$$

The results show that for the same ϵ_{split} ,

1. the accuracy of the measured signal is higher than that of its derivatives, and
2. the accuracy of the linearized forward-adjoint approach is higher than that of the linearized forward approach.

However, in the next calculations, to avoid a large CPU time and computer memory, we take $\epsilon_{split} = 10^{-3}$. For this splitting accuracy, we show in Figure 5, the adaptive grids in the plane $y = L_y/2$. From these plots we infer that

1. for the linearized forward-adjoint approach, the number of adaptive grid cells is higher in the domain “seen” by the detector along the direction θ_m^0 and in particular, along the boundaries of the pseudo-forward direct beam,
2. for the linearized forward approach, the number of adaptive grid cells is higher in the entire domain along the solar direction $\theta_0 = 120^\circ$ and when large discontinuities in the source function are present, and
3. in general, the number of adaptive grid cells for the linearized forward approach is grater than that for the linearized forward-adjoint approach.

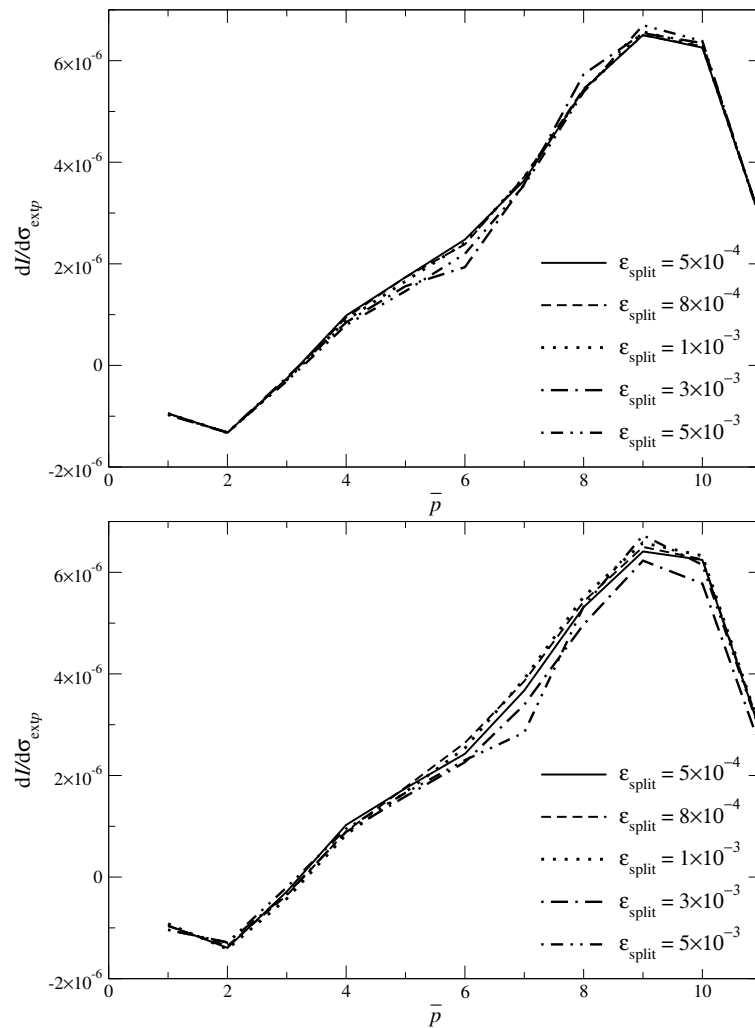


Figure 4. Partial derivatives $\partial I / \partial \sigma_{exp}$ at the base grid points ($x = x_t^0, y = y_t^0, z = \bar{p} \Delta z$), $\bar{p} = 1, \dots, 11$, computed with the linearized forward-adjoint approach (upper panel) and the linearized forward approach (lower panel) for different values of the splitting accuracy ϵ_{split} . The detector zenith angle is $\theta_m^0 = 0^\circ$, and the solar azimuthal angle is $\varphi_0 = 0^\circ$.

Table 1. Relative errors $\varepsilon_{\partial\mathcal{I}}$ and $\varepsilon_{\mathcal{I}}$ for different values of the splitting accuracy $\varepsilon_{\text{split}}$. The results correspond to the linearized forward-adjoint approach (FAA) and the linearized forward approach (FA).

$\varepsilon_{\text{split}}$	FAA		FA	
	$\varepsilon_{\partial\mathcal{I}}$	$\varepsilon_{\mathcal{I}}$	$\varepsilon_{\partial\mathcal{I}}$	$\varepsilon_{\mathcal{I}}$
8×10^{-4}	1.06×10^{-2}	1.16×10^{-4}	3.00×10^{-2}	8.70×10^{-5}
1×10^{-3}	1.42×10^{-2}	1.49×10^{-4}	3.83×10^{-2}	2.07×10^{-4}
3×10^{-3}	4.22×10^{-2}	8.36×10^{-4}	6.62×10^{-2}	7.05×10^{-4}
5×10^{-3}	5.58×10^{-2}	1.04×10^{-3}	7.75×10^{-2}	1.42×10^{-3}

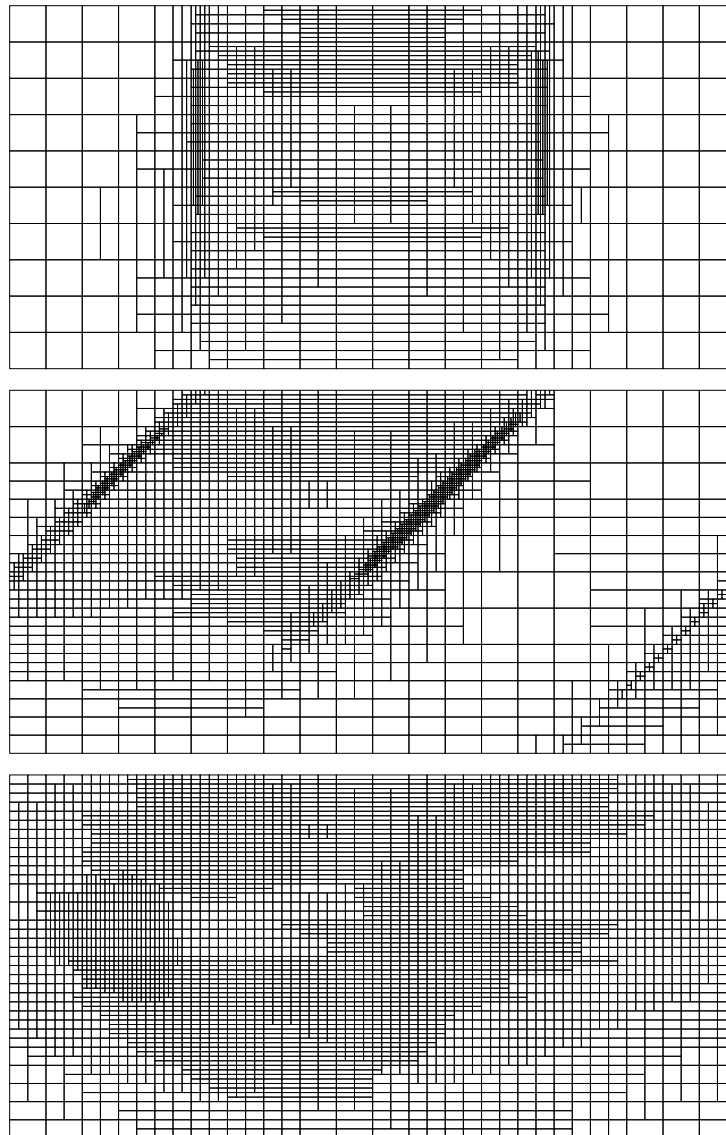


Figure 5. Adaptive grids in the plane $y = L_y/2$ corresponding to the linearized forward-adjoint approach with $\theta_m^0 = 0^\circ$ (upper panel) and $\theta_m^0 = 45^\circ$ (middle panel), and the linearized forward approach (lower panel). The solar azimuthal angle is $\varphi_0 = 0^\circ$. The horizontal axis is the x -axis, while the vertical axis is the z -axis.

In Figure 6 we plot the partial derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ for different values of the detector zenith angle θ_m^0 . The results are computed with the linearized forward-adjoint approach (FAA), the linearized forward approach (FA), and the finite-difference approach (FDA). In the latter case, a centered finite difference scheme is used for derivative calculations. The partial derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ at $N_p = 19$

base grid points ($x = \bar{p}\Delta x, y = y_t^0, z = L_z/2$), $\bar{p} = 1, \dots, N_p$, for different values of the solar azimuthal angle φ_0 are plotted in Figure 7. Taking the values computed by the finite-difference approach as a reference, we illustrate in Table 2 the corresponding relative errors in the partial derivatives, defined, for example, in the case of the linearized forward-adjoint approach as

$$\varepsilon_{\frac{\partial \mathcal{I}}{\partial \sigma}}^{\text{FAA}} = \frac{\sqrt{\sum_{\bar{p}} \left[\left(\frac{\partial \mathcal{I}}{\partial \sigma_{\text{ext}p}} \right)_{\text{FAA}} - \left(\frac{\partial \mathcal{I}}{\partial \sigma_{\text{ext}p}} \right)_{\text{FDA}} \right]^2}}{\sqrt{\sum_{\bar{p}} \left(\frac{\partial \mathcal{I}}{\partial \sigma_{\text{ext}p}} \right)_{\text{FDA}}^2}}. \tag{77}$$

The relative errors corresponding to the linearized forward-adjoint approach are smaller than 2.5×10^{-2} , while the relative errors corresponding to the linearized forward approach are smaller than 4.3×10^{-3} . The relative large errors of the linearized forward-adjoint approach can be explained as follows.

1. As it can be inferred from Figure 5, the linearized forward-adjoint approach uses a rougher discretization grid (depending on the detector zenith angle θ_m^0) as compared to the forward and finite-difference approaches (which use the same grid).
2. In the linearized forward-adjoint approach, the extinction/source function product is assumed to vary linearly within the cell, while in the forward and finite-difference approaches, the extinction/source function product is assumed to vary linearly along the characteristic.

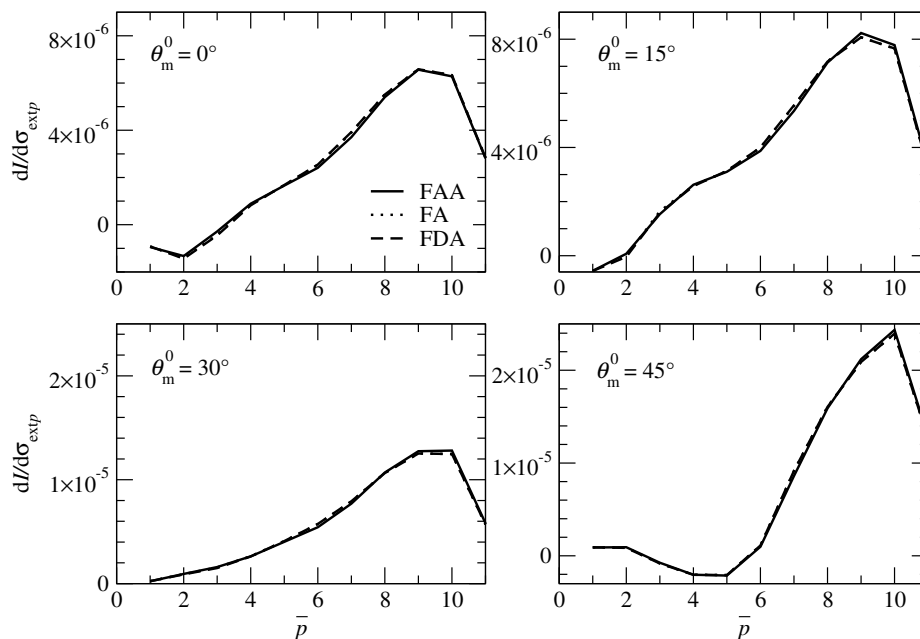


Figure 6. Partial derivatives $\partial \mathcal{I} / \partial \sigma_{\text{ext}p}$ at the base grid points ($x = x_t^0, y = y_t^0, z = \bar{p}\Delta z$), $\bar{p} = 1, \dots, 11$, for different values of the detector zenith angle θ_m^0 . The results correspond to the solar azimuthal angle $\varphi_0 = 0^\circ$ and are computed with the linearized forward-adjoint approach (FAA), the linearized forward approach (FA), and the finite-difference approach (FDA).

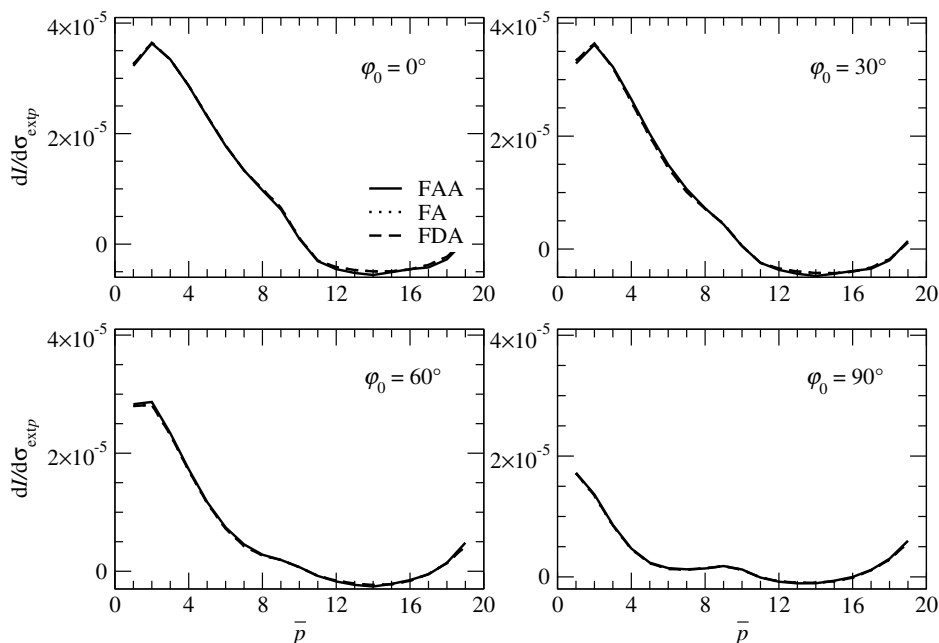


Figure 7. Partial derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ at the base grid points ($x = \bar{p}\Delta x, y = y_t^0, z = L_z/2$), $\bar{p} = 1, \dots, 19$, for different values of the solar azimuthal angle φ_0 . The results correspond to the detector zenith angle $\theta_m^0 = 45^\circ$ and are computed with the linearized forward-adjoint approach (FAA), the linearized forward approach (FA), and the finite-difference approach (FDA).

Table 2. The first part of the table shows the relative errors $\varepsilon_{\partial\mathcal{I}}$ in the partial derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ at the base grid points ($x = x_t^0, y = y_t^0, z = \bar{p}\Delta z$), $\bar{p} = 1, \dots, 11$, for different values of the detector zenith angle θ_m^0 , while the second part shows the relative errors $\varepsilon_{\partial\mathcal{I}}$ in the partial derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ at the base grid points ($x = \bar{p}\Delta x, y = y_t^0, z = L_z/2$), $\bar{p} = 1, \dots, 19$, for different values of the solar azimuthal angle φ_0 . The results are computed with the linearized forward-adjoint approach (FAA) and the linearized forward approach (FA).

θ_m^0	φ_0	$\varepsilon_{\partial\mathcal{I}}$	
		FAA	FA
0°	0°	2.50×10^{-2}	4.02×10^{-3}
15°	0°	2.23×10^{-2}	4.28×10^{-3}
30°	0°	2.42×10^{-2}	3.22×10^{-3}
45°	0°	2.05×10^{-2}	1.29×10^{-3}
45°	0°	7.35×10^{-3}	9.74×10^{-4}
45°	30°	1.64×10^{-2}	6.52×10^{-4}
45°	60°	1.56×10^{-2}	3.97×10^{-4}
45°	90°	1.21×10^{-2}	3.10×10^{-3}

4.2. Example 2

In the second example we consider in addition to the scattering and absorption by a cloud, molecular Rayleigh scattering, and the absorption by NO₂ at the wavelength $\lambda = 443$ nm. The difference to the previous geometry is that now we consider a base grid with $N_z = 16$ points along the z axis, implying that the height of the domain of analysis is $L_z = (N_z - 1)\Delta z = 1.5$ km. The cloud extinction field is given by Equations (71)–(73), while the cloud single-scattering albedo and the phase function are computed by Mie theory [27] for a water-cloud model with a Gamma size distribution

$$P(a) \propto a^\alpha \exp\left[-\alpha\left(\frac{a}{a_{\text{mod}}}\right)\right] \tag{78}$$

of parameters $a_{\text{eff}} = 10 \mu\text{m}$, $a_{\text{mod}} = 2a_{\text{eff}}/3$, and $\alpha = 6$. Here, a is the particle radius, and the droplet size ranges between 0.02 and 50.0 μm . The computed value for the cloud single-scattering albedo is $\omega_{\text{cld}} = 0.865$, the Rayleigh cross-section and depolarization ratios are calculated as in [28], while the pressure and temperature profiles correspond to the US standard model atmosphere [29]. The number density and absorption cross section profiles for NO_2 , $n(z)$ and $C_{\text{abs}}(z)$, respectively, are shown in Figure 8, and the absorption coefficient is calculated as $\sigma_{\text{abs}}(z) = n(z)C_{\text{abs}}(z)$. The peak optical depth is 5.4.

In the following, we compute the partial derivative of the measured signal with respect to maximum value of the extinction field $\sigma_{\text{ext}}^{\text{max}}$ in Equation (71); in this case, we approximate $P \approx P_{\text{cld}}$ and use

$$\frac{\partial \sigma_{\text{exti}}}{\partial \sigma_{\text{ext}}^{\text{max}}} = f(x_i, y_i, z_i) \text{ and } \frac{\partial \omega_i}{\partial \sigma_{\text{ext}}^{\text{max}}} = \frac{\omega_{\text{cld}} - \omega_i}{\sigma_{\text{exti}}} \frac{\partial \sigma_{\text{exti}}}{\partial \sigma_{\text{ext}}^{\text{max}}}. \quad (79)$$

We also compute the partial derivative of the measured signal with respect to the total column of NO_2 , $X = (k_B T_0 / p_0) \int_0^{L_z} n(z) dz$, where k_B is the Boltzmann constant, and p_0 and T_0 are the standard pressure and temperature. Assuming that the number density profile for NO_2 is the a priori profile $n_a(z)$ scaled by a constant, i.e., $n(z) = \alpha n_a(z)$, where α is the scaling factor, we get

$$\frac{\partial \sigma_{\text{exti}}}{\partial X} = \frac{1}{X_a} \sigma_{\text{abs}}(z_i) \text{ and } \frac{\partial \omega_i}{\partial X} = -\frac{\omega_i}{\sigma_{\text{exti}}} \frac{\partial \sigma_{\text{exti}}}{\partial X}, \quad (80)$$

implying $\partial \sigma_{\text{scti}} / \partial X = 0$.

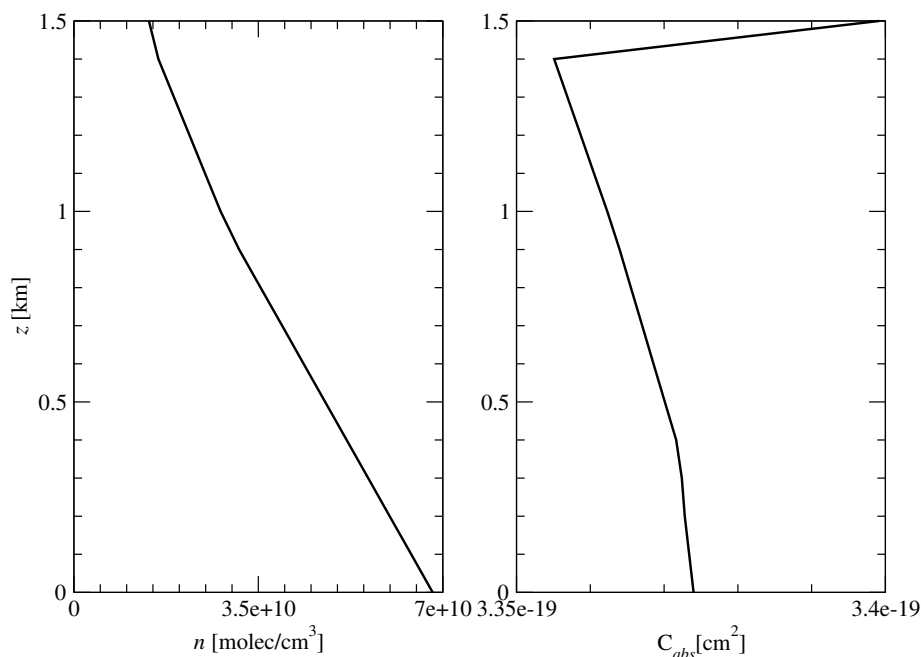


Figure 8. Number density (left) and absorption cross section (right) profiles for NO_2 .

In Tables 3 and 4, we illustrate the partial derivatives $\partial \mathcal{I} / \partial \sigma_{\text{ext}}^{\text{max}}$ and $\partial \mathcal{I} / \partial X$, and the corresponding relative errors using the values computed by the finite-difference approach as a reference. Also shown are the relative errors of the measured signal computed by Equation (22) (or, equivalently, by Equation (64)) with respect to the measured signal computed by Equation (15). In these simulations, the number of discrete zenith angles is $N_\mu = 20$, and the splitting accuracy is $\epsilon_{\text{split}} = 3 \times 10^{-3}$. The relative errors corresponding to the linearized forward-adjoint approach are now smaller than 2.3×10^{-2} , while the relative errors corresponding to the linearized forward approach are

smaller than 3.8×10^{-3} . It should be pointed out that the small errors in the measured signal certify that the adjoint problem is correctly solved.

Table 3. Partial derivative $\partial\mathcal{I}/\partial\sigma_{\text{ext}}^{\text{max}}$, and the relative errors $\varepsilon_{\mathcal{I}}$ and $\varepsilon_{\partial\mathcal{I}}$ for different values of the detector zenith angle θ_m^0 . The results correspond to the solar azimuthal angle $\varphi_0 = 0^\circ$ and are computed with the linearized forward-adjoint approach (FAA) and the linearized forward approach (FA).

θ_m^0	$\partial\mathcal{I}/\partial\sigma_{\text{ext}}^{\text{max}}$	$\varepsilon_{\mathcal{I}}$	$\varepsilon_{\partial\mathcal{I}}$	
			FAA	FA
0°	5.776×10^{-4}	4.20×10^{-4}	3.46×10^{-3}	2.03×10^{-3}
15°	1.439×10^{-3}	5.53×10^{-4}	4.86×10^{-3}	2.76×10^{-4}
30°	2.930×10^{-3}	2.30×10^{-5}	3.40×10^{-3}	2.53×10^{-4}
45°	5.499×10^{-3}	8.19×10^{-5}	1.81×10^{-3}	1.52×10^{-4}

Table 4. Partial derivative $\partial\mathcal{I}/\partial X$, and the relative errors $\varepsilon_{\mathcal{I}}$ and $\varepsilon_{\partial\mathcal{I}}$ for different values of the detector zenith angle θ_m^0 . The results correspond to the solar azimuthal angle $\varphi_0 = 0^\circ$ and are computed with the linearized forward-adjoint approach (FAA) and the linearized forward approach (FA).

θ_m^0	$\partial\mathcal{I}/\partial X$	$\varepsilon_{\mathcal{I}}$	$\varepsilon_{\partial\mathcal{I}}$	
			FAA	FA
0°	-1.184×10^{-3}	4.20×10^{-4}	1.65×10^{-2}	3.42×10^{-3}
15°	-1.284×10^{-3}	5.53×10^{-4}	2.27×10^{-2}	3.70×10^{-3}
30°	-1.428×10^{-3}	2.30×10^{-5}	2.26×10^{-2}	2.82×10^{-3}
45°	-2.529×10^{-3}	8.19×10^{-5}	2.20×10^{-2}	3.06×10^{-3}

In Table 5 we compare the computational efficiency of all linearization methods. Note that the code is parallelized with the OpenMP interface. The simulations are performed on a server Intel(R) Xeon(R) CPU E5-2695 v3 @ 2.30 GHz with 56 cores, while the CPU times are given for the parallelized implementation. Clearly, the most efficient method is the linearized forward-adjoint approach, while the most time-consuming method is the finite-difference approach.

Finally, in Table 6, we show the numbers of adaptive cells and grid points for the test problems considered in this study. Because in the linearized forward-adjoint approach, the numbers of cells and points are smaller than in the linearized forward approach, we infer that the first method is not only faster but also less memory demanding than the second one.

Table 5. CPU time in min:sec for the test problems considered in this study and corresponding to the linearized forward-adjoint approach (FAA), the linearized forward approach (FA), and the finite-difference approach (FDA). The detector zenith angle is $\theta_m^0 = 45^\circ$ and the solar azimuthal angle is $\varphi_0 = 0^\circ$.

Partial Derivatives	FAA	FA	FDA
$\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ (Figure 6)	1:16	27:12	138:19
$\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ (Figure 7)	1:47	32:34	153:26
$\partial\mathcal{I}/\partial\sigma_{\text{ext}}^{\text{max}}$	0:32	2:21	2:24
$\partial\mathcal{I}/\partial X$	0:31	2:20	2:20

Table 6. The first part of the table shows the number of adaptive cells N_{cells} and the final number of grid points N_{points} for computing the derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}p}$ with the splitting accuracy $\epsilon_{\text{split}} = 10^{-3}$, while the second part of the table shows N_{cells} and N_{points} for computing the derivatives $\partial\mathcal{I}/\partial\sigma_{\text{ext}}^{\text{max}}$ and $\partial\mathcal{I}/\partial X$ with the splitting accuracy $\epsilon_{\text{split}} = 3 \times 10^{-3}$. The results correspond to the linearized forward-adjoint approach (FAA) and the linearized forward approach (FA).

ϵ_{split}	θ_m^0	FAA		FA	
		N_{cells}	N_{points}	N_{cells}	N_{points}
10^{-3}	0°	66,652	40,727	293,492	162,279
10^{-3}	15°	81,174	50,381	293,492	162,279
10^{-3}	30°	95,152	59,595	293,492	162,279
10^{-3}	45°	115,186	73,332	293,492	162,279
3×10^{-3}	0°	17,804	13,196	54,164	33,576
3×10^{-3}	15°	20,094	14,942	54,164	33,576
3×10^{-3}	30°	22,674	16,989	54,164	33,576
3×10^{-3}	45°	25,806	19,607	54,164	33,576

5. Conclusions

Two linearization methods for SHDOM have been discussed. The first method is an analytical linearization approach, while the second method is based on the adjoint radiative transfer theory. In the latter case, practical formulas for computing the derivatives of the measured signal in the spherical harmonics space have been derived. Essentially, SHDOM has been specialized for derivative calculations and radiative transfer problems involving the delta-M approximation, the TMS correction, and the adaptive grid splitting.

Our numerical analysis leads to the following conclusions.

1. The linearized SHDOM with analytical derivatives is an accurate approach. However, the method is time-consuming and demands a large computer memory. The reason is that not only the source function has to be stored as a spherical harmonic series at each grid point, but also its derivatives with respect to the atmospheric parameters of interest.
2. The linearized SHDOM with a forward-adjoint approach requires less storage for derivatives calculation, are much faster, but relatively less accurate. The main reason for this lower accuracy is that different interpolation schemes are used for radiance and derivative calculations.

According to our numerical analysis, the relative errors of the linearized forward-adjoint approach are smaller than 2.5%. In this regard, it should be pointed out that in the framework of the Gauss-Newton method, errors of about 2–3% in the Jacobian do not play a significant role. The effect is somehow similar to the computation of the root of an equation by the secant method instead of the Newton method. However, if the errors in the Jacobian are important for the retrieval, the regularized total least squares, which accounts on the errors in both the Jacobian and the data, can be used.

In our analysis, we considered the case of a fixed detector position \mathbf{r}_D and a single-angle measurement Ω_m^0 . For multi-angle measurements and eventually, several detector positions, the solution of several adjoint problems (corresponding to each pair $(\mathbf{r}_D, \Omega_m^0)$) can be avoided by employing the approach proposed in Ref. [23]. The main ideas of this approach, which can be applied for multi-dimensional remote sensing of cloud extinction fields from airborne radiometer measurements, are discussed in Appendix D. However, it should be pointed out that the computational cost for the retrieval of 3D extinction fields is extremely high. A reduction of the computation time can be achieved if a two-dimensional geometry (a constant extinction field along the y -axis, and a set of directions on the unit sphere) is considered. In this case, the application of the forward-adjoint approach requires that both the solar and detector azimuthal angles are zero (the solar and detector directions are in the xz -plane).

Potential application of the linearized SHDOM include the following.

1. Retrieval of cloud model parameters. For broken clouds [30], the indicator function $f(x, y, z)$ in Equation (71) takes the values 1 inside the cloud and 0 inside the clear sky region. As in [31], the extinction field can be parametrized in terms of $\sigma_{\text{ext}}^{\text{max}}$ (cf. Equation (71)), the cloud top height h_t , and the cloud bottom height h_b . The cloud optical thickness and the cloud geometrical parameters can be retrieved in the oxygen A-band by using the correlated k -distribution method for the broadband integration of the gaseous line absorption. Note that SHDOM is able to compute monochromatic and broadband radiative transfer (with a k distribution).
2. Trace gas retrievals under cloudy conditions. The retrieval of total column of trace gases, e.g., O_3 , NO_2 , etc., can be performed by means of the differential optical absorption spectroscopy (DOAS) technique [32]. This approach requires the knowledge of the air mass factor (AMF), i.e., the partial derivative of the measured signal with respect to the total column of the trace gas at a specific wavelength. The main goal of the analysis is the computation of the air mass factor when clouds are outside the footprint of the instrument (effect of horizontal cloud edges on AMF calculation).

Author Contributions: Methodology, A.D.; Software, A.D. and D.S.E.; Writing-Original Draft Preparation, A.D. and D.S.E.; Writing-Review & Editing, A.D. and D.S.E.

Funding: This research received no external funding

Acknowledgments: The authors are pleased to acknowledge very fruitful and stimulating discussions with Frank Evans during the completion of this work.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In this appendix we summarize the SHDOM algorithm in order to emphasize the main equations which are differentiated in the framework of the linearized forward approach.

SHDOM uses both spherical harmonics and discrete ordinates to represent the radiance field during different parts of the solution algorithm. The spherical harmonics are employed for computing the source function including the scattering integral, while the discrete ordinates are used to integrate the radiative transfer equation spatially.

The set of discrete ordinates are a reduced Gaussian grid. There are N_μ Gaussian quadrature cosine zenith angles, μ_j , and N_φ evenly spaced azimuth angles, φ_k . The discrete ordinate set is reduced by having fewer azimuth angles at larger $|\mu|_j$; thus, N_φ depends on j . The corresponding Gauss-Legendre quadrature weights and azimuthal integration weights (normalized appropriately) are w_{μ_j} and $w_{\varphi_{jk}}$, respectively.

The orthonormal real-value spherical harmonics are defined by

$$Y_{mn}(\mu, \varphi) = P_n^{|m|}(\mu)u_m(\varphi), \tag{A1}$$

where $P_n^{|m|}(\mu)$ are the normalized associated Legendre functions, and

$$u_m(\varphi) = \begin{cases} (1/\sqrt{\pi}) \cos(|m|\varphi) & m > 0 \\ (1/\sqrt{2\pi}) & m = 0 \\ (1/\sqrt{\pi}) \sin(|m|\varphi) & m < 0 \end{cases} \tag{A2}$$

are the Fourier harmonics. The source function

$$J(\mathbf{r}, \boldsymbol{\Omega}) = \frac{\omega(\mathbf{r})}{4\pi} \int_{\Omega} P(\mathbf{r}, \boldsymbol{\Omega}, \boldsymbol{\Omega}') [I_d(\mathbf{r}, \boldsymbol{\Omega}') + \delta(\boldsymbol{\Omega}' - \boldsymbol{\Omega}_0)T_0(\mathbf{r})] d\boldsymbol{\Omega}', \tag{A3}$$

$$T_0(\mathbf{r}) = \frac{F_0}{|\mu_0|} e^{-\tau_{\text{ext}}(\mathbf{r}, \mathbf{r}_0|\boldsymbol{\Omega}_0)}, \tag{A4}$$

where

$$\tau_{\text{ext}}(\mathbf{r}, \mathbf{r}_0 | \Omega_0) = \int_0^s \sigma_{\text{ext}}(\mathbf{r}_0 + s' \Omega_0) ds' \tag{A5}$$

with $s = |\mathbf{r} - \mathbf{r}_0|$, is the optical depth from the point \mathbf{r} to the upper domain boundary point \mathbf{r}_0 along the solar direction Ω_0 , is expressed in the spherical harmonic space as

$$J(\mathbf{r}, \Omega) = \sum_{m=-M}^M \sum_{n=|m|}^N J_{mn}(\mathbf{r}) Y_{mn}(\Omega), \tag{A6}$$

where $N = N_\mu - 1$ is the maximum expansion order and $M = N_\varphi / 2 - 1$ is the maximum number of azimuthal modes.

By means of the addition theorem for the Legendre functions, the phase function, expanded in terms of unnormalized Legendre polynomials

$$\tilde{P}_n(\cos \Theta) = \sqrt{\frac{2}{2n+1}} P_n(\cos \Theta), \tag{A7}$$

as

$$P(\mathbf{r}, \cos \Theta) = \sum_{n=1}^{N_{\text{rank}}} \chi_n(\mathbf{r}) \tilde{P}_n(\cos \Theta), \tag{A8}$$

where $\chi_n(\mathbf{r})$ are the Legendre phase function coefficients and N_{rank} is the maximum expansion order of the phase function, is truncated in the spherical harmonic space as

$$P(\mathbf{r}, \cos \Theta) = 4\pi \sum_{m=-M}^M \sum_{n=|m|}^N \frac{\chi_n(\mathbf{r})}{2n+1} Y_{mn}(\Omega) Y_{mn}(\Omega'), \tag{A9}$$

where $\cos \Theta = \Omega \cdot \Omega'$.

The solution method is based on Picard iterations. At the beginning of each iteration step, the expansion coefficients $J_{mn}(\mathbf{r}_i)$, $m = -M, \dots, M$, $n = |m|, \dots, N$, are assumed to be known at all grid points \mathbf{r}_i . They are updated according to the following four computational steps.

Step 1. The source function is transformed to discrete ordinates by means of the relation

$$J(\mathbf{r}_i, \mu_j, \varphi_k) = \sum_{m=-M}^M u_m(\varphi_k) \sum_{n=|m|}^N J_{mn}(\mathbf{r}_i) P_n^{(|m|)}(\mu_j). \tag{A10}$$

Step 2. The discrete ordinate radiance $I_d(\mathbf{r}_i, \mu_j, \varphi_k)$ is computed from the source function $J(\mathbf{r}_i, \mu_j, \varphi_k)$ by integrating the radiative transfer equation. Essentially, (i) the radiances are computed at all grid points in all downward discrete ordinate directions by integrating the radiative transfer equation with the top boundary condition $I_d(\mathbf{r}_{ti}, \mu_j^-, \varphi_k^-) = 0$, where $(\mu_j^-, \varphi_{\mu j}^-)$ and $(w_{\mu j}^-, w_{\varphi j k}^-)$ are the quadrature nodes and weights in the lower hemisphere, (ii) the radiances at the bottom boundary grid points in all upward discrete ordinate directions (μ_j^+, φ_k^+) are computed from the boundary condition

$$\begin{aligned} I_d(\mathbf{r}_{bi}, \mu_j^+, \varphi_k^+) &= \frac{1}{\pi} \rho(\mathbf{r}_{bi}, \mu_j^+, \varphi_k^+, \mu_0, \varphi_0) |\mu_0| T_0(\mathbf{r}_{bi}) \\ &+ \frac{1}{\pi} \sum_{p=1}^{N_\mu/2} \sum_{q=1}^{N_\varphi p} w_{\mu p}^- w_{\varphi p q}^- \rho(\mathbf{r}_{bi}, \mu_j^+, \varphi_k^+, \mu_p^-, \varphi_q^-) \\ &\times |\mu_p^-| I_d(\mathbf{r}_{bi}, \mu_p^-, \varphi_q^-), \end{aligned} \tag{A11}$$

and (iii) the radiances are computed at all grid points in all upward discrete ordinate directions by integrating the radiative transfer equation with the boundary condition (A11).

Step 3. The discrete ordinate radiance at each grid point is transformed to the spherical harmonic space according to

$$I_{mn}(\mathbf{r}_i) = \sum_{j=1}^{N_\mu} w_{\mu_j} P_n^{|\mu_j|}(\mu_j) \sum_{k=1}^{N_{\varphi_j}} w_{\varphi_{jk}} I_d(\mathbf{r}_i, \mu_j, \varphi_k) u_m(\varphi_k). \tag{A12}$$

Step 4. At each grid point, the source function is computed from the radiance in the spherical harmonic space as

$$J_{mn}(\mathbf{r}_i) = \omega(\mathbf{r}_i) \frac{\chi_n(\mathbf{r}_i)}{2n+1} I_{mn}(\mathbf{r}_i) + \omega(\mathbf{r}_i) \frac{\chi_n(\mathbf{r}_i)}{2n+1} Y_{mn}(\mathbf{\Omega}_0) T_0(\mathbf{r}_i). \tag{A13}$$

Some comments are in order.

1. The radiance and source function are initialized before the solution iterations with an Eddington radiative transfer solution on independent columns of the base grid.
2. The explicit form of the transforms in Equations (A10) and (A12) illustrates how the azimuthal and zenith angle parts partially separate. For more than about 12 azimuthal angles, an FFT is used for the azimuthal Fourier transform.
3. In Step 2, the radiative transfer equation is integrated backward from each grid point to a grid cell face that has known radiances at its bounding grid points. In the short-characteristic method, the integration is across just one cell, while in the long characteristic method, the characteristic $\mathbf{\Omega}_{jk} = (\mu_j, \varphi_k)$ is traced backward until the transmission falls below some minimum specified value. In the latter case, the error from interpolating the radiance at the grid cell face is reduced. According to the SHDOM difference scheme, the radiance at the exiting point *A* along a characteristic $\mathbf{\Omega}_{jk}$, $I_d(\mathbf{r}_A, \mathbf{\Omega}_{jk})$, is computed from the integral form of the radiative transfer equation

$$I_d(\mathbf{r}_A, \mathbf{\Omega}_{jk}) = I_d(\mathbf{r}_B, \mathbf{\Omega}_{jk}) e^{-\int_0^{s_0} \sigma_{\text{ext}}(\mathbf{r}_B + s\mathbf{\Omega}_{jk}) ds} + S(s_0, J), \tag{A14}$$

where $I_d(\mathbf{r}_B, \mathbf{\Omega}_{jk})$ is the radiance at the entering point *B*, $s_0 = |\mathbf{r}_A - \mathbf{r}_B|$ is the distance between the points *A* and *B*, and

$$S(s_0, J) = \int_0^{s_0} \sigma_{\text{ext}}(\mathbf{r}_B + s\mathbf{\Omega}_{jk}) J(\mathbf{r}_B + s\mathbf{\Omega}_{jk}, \mathbf{\Omega}_{jk}) \times \left[e^{-\int_s^{s_0} \sigma_{\text{ext}}(\mathbf{r}_B + s'\mathbf{\Omega}_{jk}) ds'} \right] ds \tag{A15}$$

is the integral of the source function along the characteristic. Setting $\sigma_{\text{ext}A,B} = \sigma_{\text{ext}}(\mathbf{r}_{A,B})$ and $J_{A,B} = J(\mathbf{r}_{A,B}, \mathbf{\Omega}_{jk})$, and assuming that the extinction σ_{ext} and the extinction/source function product $\sigma_{\text{ext}}J$ vary linearly along the characteristic, yields

$$S(s_0, J) = (1 - e^{-\tau_{\text{ext}}}) \frac{1}{\sigma_{\text{ext}A} + \sigma_{\text{ext}B}} \left[\sigma_{\text{ext}A} J_A + \sigma_{\text{ext}B} J_B + \frac{\sigma_{\text{ext}A} \sigma_{\text{ext}B}}{6} (J_A - J_B) s_0 \right], \tag{A16}$$

for $\tau_{\text{ext}} \leq 2$, and

$$S(s_0, J) = (1 - e^{-\tau_{\text{ext}}}) \frac{1}{\sigma_{\text{ext}A} + \sigma_{\text{ext}B}} \left[\sigma_{\text{ext}A} J_A + \sigma_{\text{ext}B} J_B + \frac{2\sigma_{\text{ext}A} \sigma_{\text{ext}B}}{\sigma_{\text{ext}A} + \sigma_{\text{ext}B}} (J_A - J_B) \left(1 - \frac{2}{\tau_{\text{ext}}} + \frac{2e^{-\tau_{\text{ext}}}}{1 - e^{-\tau_{\text{ext}}}} \right) \right], \tag{A17}$$

for $\tau_{\text{ext}} > 2$, where

$$\tau_{\text{ext}} = \tau_{\text{ext}}(\mathbf{r}_A, \mathbf{r}_B | \Omega_{jk}) = \int_0^{s_0} \sigma_{\text{ext}}(\mathbf{r}_B + s\Omega_{jk}) ds \tag{A18}$$

is the optical depth along the characteristic. Note that under the above assumptions, the representation (A16) follows from an expansion of the solution to the first order in the path distance s_0 , while the representation (A17) is a particular case of the general solution derived in [33]. The exiting and the entering values of the extinction and extinction/source function product $\sigma_{\text{ext}A}$, $\sigma_{\text{ext}B}$, $\sigma_{\text{ext}A}J_A$, and $\sigma_{\text{ext}B}J_B$ are computed by bilinear interpolation from the grid point values of the faces pierced by the characteristic, while the radiance is also bilinearly interpolated between the surrounding grid points to give the initial radiance $I_d(\mathbf{r}_B, \Omega_{jk})$ for integration.

4. To increase the convergence rate, an acceleration method based on geometrical convergence is applied. At the iteration step K , the “accelerated” source column vector is computed as

$$J_{\text{acc}}^{(K)} = J^{(K)} + a\Delta J^{(K)} \tag{A19}$$

where $J^{(K)}$ is the column vector encapsulating the expansion coefficients $J_{mn}^{(K)}(\mathbf{r}_i)$, $m = -M, \dots, M$, $n = |m|, \dots, N$, at all grid points \mathbf{r}_i , i.e., $J^{(K)} = [J_{mn}^{(K)}(\mathbf{r}_i)]$,

$$\Delta J^{(K)} = J^{(K)} - J^{(K-1)}, \tag{A20}$$

is the residual source vector at the iteration step K , and

$$a = \frac{1 - r \cos \psi + r^{1+\pi/(2\psi)}}{1 + r^2 - 2r \cos \psi} - 1 \tag{A21}$$

with

$$r = \frac{\|\Delta J^{(N)}\|}{\|\Delta J^{(N-1)}\|}, \tag{A22}$$

$$\cos \psi = \frac{(\Delta J^{(N-1)})^T (\Delta J^{(N)})}{\|\Delta J^{(N-1)}\| \|\Delta J^{(N)}\|}, \tag{A23}$$

is the acceleration parameter. The iterations are stopped when the solution criterion

$$\frac{\|\Delta J^{(N)}\|}{\|J^{(N-1)}\|} \leq \epsilon_{\text{sol}} \tag{A24}$$

is satisfied, where ϵ_{sol} is the solution tolerance.

5. If the delta-M scaling method is applied, then the optical parameters σ_{ext} , ω , and χ_n are scaled before their use. The scaled quantities $\bar{\sigma}_{\text{ext}}$, $\bar{\omega}$, and $\bar{\chi}_n$ are given, respectively, by (the dependency on \mathbf{r} is omitted)

$$\bar{\sigma}_{\text{ext}} = (1 - f\omega)\sigma_{\text{ext}}, \tag{A25}$$

$$\bar{\omega} = \frac{1 - f}{1 - f\omega}\omega, \tag{A26}$$

$$\frac{\bar{\chi}_n}{2n + 1} = \frac{1}{1 - f} \left(\frac{\chi_n}{2n + 1} - f \right), \quad n = 0, \dots, N, \tag{A27}$$

where

$$f = \frac{1}{2N + 3} \chi_{N+1} \tag{A28}$$

is the truncation fraction.

6. The radiance at a specified direction Ω_m and location \mathbf{r} is computed by integrating the source function through the medium, while the spherical harmonic representation of the source function is transformed to the desired viewing direction Ω_m .
7. For solar problems with the delta-M method, the TMS method is used to compute the source function. This method replaces the scaled, truncated Legendre phase function expansion for the singly scattered solar radiation by the full, unscaled phase function expansion, while the multiply scattered contribution still comes from the truncated phase function. The source function at grid point \mathbf{r}_i in direction Ω_m is then computed as

$$J(\mathbf{r}_i, \Omega_m) = \sum_{m=-M}^M \sum_{n=|m|}^N J_{mn}(\mathbf{r}_i) Y_{mn}(\Omega_m) + \Delta J(\mathbf{r}_i, \Omega_m), \tag{A29}$$

$$\begin{aligned} \Delta J(\mathbf{r}_i, \Omega_m) = & \frac{F_0}{|\mu_0|} e^{-\bar{\tau}_{\text{ext}}(\mathbf{r}_i, \mathbf{r}_{i0}|\Omega_0)} \left[\frac{\bar{\omega}(\mathbf{r}_i)}{1-f} \frac{1}{4\pi} \sum_{n=1}^{N_{\text{rank}}} \chi_n(\mathbf{r}_i) \tilde{P}_n(\cos \Theta) \right. \\ & \left. - \bar{\omega}(\mathbf{r}_i) \sum_{m=-M}^M \sum_{n=|m|}^N \frac{\bar{\chi}_n(\mathbf{r}_i)}{2n+1} Y_{mn}(\Omega_0) Y_{mn}(\Omega_m) \right], \end{aligned} \tag{A30}$$

where $\cos \Theta = \Omega_0 \cdot \Omega_m$.

8. The adaptive grid evolves from the base grid by splitting cells where more resolution is judged to be needed. The criterion for splitting cells is based on how much the source function times extinction changes across a cell. A cell may be split in half in either of the three Cartesian directions, depending on whether any of them exceed the splitting criterion. For cell c , the cell-splitting criterion is

$$C_{cu} = |\Delta J_{uj}| (1 - e^{-\tau_{uj}}), \tag{A31}$$

where u is one of the three Cartesian directions $x, y,$ and $z,$

$$\begin{aligned} |\Delta J_{uj}| = & \frac{1}{\bar{\sigma}_{\text{ext}uj}} \sqrt{\frac{1}{4\pi} \int_{\Omega} [\sigma_{\text{ext}uj}^+ J(\mathbf{r}_{uj}^+, \Omega) - \sigma_{\text{ext}uj}^- J(\mathbf{r}_{uj}^-, \Omega)]^2 d\Omega} \\ = & \frac{1}{\bar{\sigma}_{\text{ext}uj}} \sqrt{\frac{1}{4\pi} \sum_{m=-M}^M \sum_{n=|m|}^N [\sigma_{\text{ext}uj}^+ J_{mn}(\mathbf{r}_{uj}^+) - \sigma_{\text{ext}uj}^- J_{mn}(\mathbf{r}_{uj}^-)]^2}, \end{aligned} \tag{A32}$$

$\mathbf{r}_{uj}^+ \in F_u^+$ and $\mathbf{r}_{uj}^- \in F_u^-$, $j = 1, \dots, 4$, are two cell corners along direction u , F_u^+ and F_u^- are the two cell faces crossed by direction u , $\tau_{uj} = \bar{\sigma}_{\text{ext}uj} s_{uj}$, $\bar{\sigma}_{\text{ext}uj} = (\sigma_{\text{ext}uj}^+ + \sigma_{\text{ext}uj}^-)/2$, $\sigma_{\text{ext}uj}^+ = \sigma_{\text{ext}}(\mathbf{r}_{uj}^+)$, $\sigma_{\text{ext}uj}^- = \sigma_{\text{ext}}(\mathbf{r}_{uj}^-)$, and $s_{uj} = |\mathbf{r}_{uj}^+ - \mathbf{r}_{uj}^-|$. The criterion is averaged over the four corners, i.e., $\bar{C}_{cu} = (1/4) \sum_{j=1}^4 C_{cu}$, and the cells are sorted by the maximum over the splitting directions $u = x, y, z$ of the averaged splitting criterion \bar{C}_{cu} . At the iteration step K , those cells c with the highest criterion \bar{C}_{cu} above a certain value $\varepsilon_{\text{split}}^{(K)}$ are split first. The new cells may themselves be divided during one solution iteration. As the solution iterations proceed, the cell splitting accuracy $\varepsilon_{\text{split}}^{(K)}$ is gradually lowered to the desired final cell-splitting accuracy $\varepsilon_{\text{split}}$, and so more grid points are added at each iteration during this process.

Appendix B

The partial derivative of the measured signal with respect to a parameter ξ_p that determine the single-scattering albedo at every grid point $\omega_i = \omega(\mathbf{r}_i)$ is computed in the spherical harmonic space as

follows. For $\mathbf{r} \in D_c$, we assume that the extinction/source function product varies linearly within cell c , i.e.,

$$\begin{aligned} \sigma_{\text{ext}}(\mathbf{r})J(\mathbf{r}, \boldsymbol{\Omega}) &= \sum_{\bar{i}=1} L_{\bar{i}}(\mathbf{R})\sigma_{\text{extg}(\bar{i},c)}J(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}) \\ &= \sum_{\bar{i}=1} L_{\bar{i}}(\mathbf{R})\sigma_{\text{extg}(\bar{i},c)} \left[\frac{\omega_{g(\bar{i},c)}}{4\pi} \int_{\Omega} P(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \right]. \end{aligned} \tag{A33}$$

We obtain

$$\begin{aligned} \left(\frac{\partial \mathcal{L}}{\partial \xi_p} I \right) (\mathbf{r}, \boldsymbol{\Omega}) &= - \sum_{\bar{i}} L_{\bar{i}}(\mathbf{R}) \left[\frac{\sigma_{\text{extg}(\bar{i},c)}{4\pi} \right. \\ &\quad \left. \times \int_{\Omega} P(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \right] \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p}, \end{aligned} \tag{A34}$$

yielding

$$\begin{aligned} \frac{\partial E}{\partial \xi_p} &= \sum_c \sum_{\bar{i}} \int_{\Omega} \int_{D_c} L_{\bar{i}}(\mathbf{R}) \hat{\Gamma}^{\dagger}(\mathbf{r}, -\boldsymbol{\Omega}) \left[\frac{\sigma_{\text{extg}(\bar{i},c)}{4\pi} \right. \\ &\quad \left. \times \int_{\Omega} P(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I(\mathbf{r}_{g(\bar{i},c)}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \right] \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p} dV d\boldsymbol{\Omega}. \end{aligned} \tag{A35}$$

Using

$$\int_{\Omega} P(\mathbf{r}, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I_d(\mathbf{r}, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' = 4\pi \sum_{mn} \frac{\chi_n(\mathbf{r})}{2n+1} I_{mn}(\mathbf{r})Y_{mn}(\boldsymbol{\Omega}) \tag{A36}$$

and

$$P(\mathbf{r}, \boldsymbol{\Omega}, \boldsymbol{\Omega}') = 4\pi \sum_{mn} \frac{\chi_n(\mathbf{r})}{2n+1} Y_{mn}(\boldsymbol{\Omega})Y_{mn}(\boldsymbol{\Omega}'), \tag{A37}$$

we end up with

$$\frac{\partial \mathcal{I}}{\partial \xi_p} = T_A + T_B + T_C + T_D, \tag{A38}$$

where

$$T_A = \sum_c \sum_{mn} (-1)^n \sum_{\bar{i}, \bar{j}} \frac{\chi_{n,g(\bar{i},c)}}{2n+1} \sigma_{\text{extg}(\bar{i},c)} I_{mn,g(\bar{i},c)} \hat{\Gamma}_{mn,g(\bar{j},c)}^{\dagger} \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p} L_{\bar{i},\bar{j}}, \tag{A39}$$

$$T_B = \sum_c \sum_{mn} (-1)^n Y_{mn}(\boldsymbol{\Omega}_0) \sum_{\bar{i}, \bar{j}} \frac{\chi_{n,g(\bar{i},c)}}{2n+1} \sigma_{\text{extg}(\bar{i},c)} T_{0g(\bar{i},c)} \hat{\Gamma}_{mn,g(\bar{j},c)}^{\dagger} \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p} L_{\bar{i},\bar{j}}, \tag{A40}$$

and

$$T_C = \sum_c \sum_{mn} Y_{mn}(\boldsymbol{\Omega}_{mc}) \sum_{\bar{i}, \bar{j}} \frac{\chi_{n,g(\bar{i},c)}}{2n+1} \sigma_{\text{extg}(\bar{i},c)} I_{mn,g(\bar{i},c)} \hat{\Gamma}_{mg(\bar{j},c)}^{\dagger} \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p} L_{\bar{i},\bar{j}}, \tag{A41}$$

$$T_D = \sum_c \sum_{mn} Y_{mn}(\boldsymbol{\Omega}_{mc}) Y_{mn}(\boldsymbol{\Omega}_0) \sum_{\bar{i}, \bar{j}} \frac{\chi_{n,g(\bar{i},c)}}{2n+1} \sigma_{\text{extg}(\bar{i},c)} T_{0g(\bar{i},c)} \hat{\Gamma}_{mg(\bar{j},c)}^{\dagger} \frac{\partial \omega_{g(\bar{i},c)}}{\partial \xi_p} L_{\bar{i},\bar{j}}. \tag{A42}$$

Considering the partial derivative of the measured signal with respect to a parameter ξ_p that determine the bidirectional surface reflection function $\rho(\mathbf{r}_b, \Omega^+, \Omega^-)$, we use Equation (32), to obtain

$$\begin{aligned} \frac{\partial \mathcal{I}}{\partial \xi_p} &= \frac{1}{\pi} |\mu_0| \mu_m(\mathbf{r}_b) T_0(\mathbf{r}_b) \widehat{T}_m^+(\mathbf{r}_b) \frac{\partial \rho}{\partial \xi_p}(\mathbf{r}_b, \Omega_m(\mathbf{r}_b), \Omega_0) \\ &+ \frac{1}{\pi} |\mu_0| T_0(\mathbf{r}_b) \int_{\Omega} H(\mu) \mu \widehat{I}_d^+(\mathbf{r}_b, -\Omega) \frac{\partial \rho}{\partial \xi_p}(\mathbf{r}_b, \Omega, \Omega_0) d\Omega \\ &+ \frac{1}{\pi} \mu_m(\mathbf{r}_b) \widehat{T}_m^+(\mathbf{r}_b) \int_{\Omega} \frac{\partial \rho}{\partial \xi_p}(\mathbf{r}_b, \Omega_m(\mathbf{r}_b), \Omega) \\ &\times H(-\mu') |\mu'| I_d(\mathbf{r}_b, \Omega) d\Omega \\ &+ \frac{1}{\pi} \int_{\Omega} H(\mu) \mu \widehat{I}_d^+(\mathbf{r}_b, -\Omega) \\ &\times \left[\int_{\Omega} \frac{\partial \rho}{\partial \xi_p}(\mathbf{r}_b, \Omega, \Omega') H(-\mu') |\mu'| I_d(\mathbf{r}_b, \Omega') d\Omega' \right] d\Omega. \end{aligned} \tag{A43}$$

The integrals over the unit sphere are computed in the discrete ordinate space under the assumption that the partial derivatives $\partial \rho / \partial \xi_p$ are known at all grid points on the bottom surface, and in all downward and upward discrete ordinate directions. For a Lambertian surface, $\rho(\mathbf{r}_b, \Omega^+, \Omega^-)$ is simply the albedo A , and for $\xi_p = A$, the computational formula is

$$\begin{aligned} \frac{\partial \mathcal{I}}{\partial A} &= \frac{1}{\pi} |\mu_0| \mu_m(\mathbf{r}_{bi}) T_0(\mathbf{r}_{bi}) \widehat{T}_m^+(\mathbf{r}_{bi}) \\ &+ \frac{1}{\pi} |\mu_0| T_0(\mathbf{r}_{bi}) \sum_{j=1}^{N_\mu/2} \sum_{k=1}^{N_{\varphi_j}} w_{\mu_j}^- w_{\varphi_{jk}}^- |\mu_j^-| \widehat{I}_d^+(\mathbf{r}_{bi}, \mu_j^-, \varphi_k^-) \\ &+ \frac{1}{\pi} \mu_m(\mathbf{r}_{bi}) \widehat{T}_m^+(\mathbf{r}_{bi}) \sum_{j=1}^{N_\mu/2} \sum_{k=1}^{N_{\varphi_j}} w_{\mu_j}^- w_{\varphi_{jk}}^- |\mu_j^-| I_d(\mathbf{r}_{bi}, \mu_j^-, \varphi_k^-) \\ &+ \frac{1}{\pi} \sum_{j=1}^{N_\mu/2} \sum_{k=1}^{N_{\varphi_j}} \sum_{p=1}^{N_\mu/2} \sum_{q=1}^{N_{\varphi_p}} w_{\mu_j}^- w_{\mu_p}^- w_{\varphi_{jk}}^- w_{\varphi_{pq}}^- |\mu_j^-| |\mu_p^-| \\ &\times \widehat{I}_d^+(\mathbf{r}_{bi}, \mu_j^-, \varphi_k^-) I_d(\mathbf{r}_{bi}, \mu_p^-, \varphi_q^-). \end{aligned} \tag{A44}$$

Appendix C

Referring to Figure 2, the first-order interpolation basis functions for a rectangular cuboid element $L_{\bar{i}}, \bar{i} = 1, \dots, 8$ at a point $\mathbf{r} = (x, y, z)$ are given by

$$L_1(\xi, \eta, \varsigma) = n_-(\xi) n_-(\eta) n_-(\varsigma), \tag{A45}$$

$$L_2(\xi, \eta, \varsigma) = n_+(\xi) n_-(\eta) n_-(\varsigma), \tag{A46}$$

$$L_3(\xi, \eta, \varsigma) = n_-(\xi) n_+(\eta) n_-(\varsigma), \tag{A47}$$

$$L_4(\xi, \eta, \varsigma) = n_+(\xi) n_+(\eta) n_-(\varsigma), \tag{A48}$$

$$L_5(\xi, \eta, \varsigma) = n_-(\xi) n_-(\eta) n_+(\varsigma), \tag{A49}$$

$$L_6(\xi, \eta, \varsigma) = n_+(\xi) n_-(\eta) n_+(\varsigma), \tag{A50}$$

$$L_7(\xi, \eta, \varsigma) = n_-(\xi) n_+(\eta) n_+(\varsigma), \tag{A51}$$

$$L_8(\xi, \eta, \varsigma) = n_-(\xi) n_+(\eta) n_+(\varsigma), \tag{A52}$$

where

$$n_{\pm}(\xi) = \frac{1}{2}(1 \pm \xi) \tag{A53}$$

are the one-dimensional linear interpolation functions, and (ξ, η, ς) are the normalized coordinates of the point in a local coordinate system attached to the center of the cell $\rho_c = (x_c, y_c, z_c)$, i.e.,

$$\xi = \frac{x - x_c}{l_x}, \quad \eta = \frac{y - y_c}{l_y}, \quad \varsigma = \frac{z - z_c}{l_z}. \tag{A54}$$

Appendix D

We consider the case of a detector with a fixed position \mathbf{r}_D , that measures the radiance in K directions Ω_{mk}^0 , $k = 1, \dots, K$. For the measurement direction Ω_{mk}^0 , let $S_{tm}(\Omega_{mk}^0)$ be the detector footprint centered at \mathbf{r}_{tk}^0 . We aim to compute the state vector ξ by minimizing the sum of squared residuals

$$\mathcal{R}(\xi) = \frac{1}{2} \sum_k [\mathcal{I}(\xi, \Omega_{mk}^0) - \mathcal{I}_{mes}(\Omega_{mk}^0)]^2, \tag{A55}$$

where $\mathcal{I}(\xi, \Omega_{mk}^0)$ and $\mathcal{I}_{mes}(\Omega_{mk}^0)$ are the simulated and measured signals of the detector, respectively. In the steepest-descent method characterized by a linear convergence rate, the objective function is approximated by a linear model and the search direction is taken as the negative of the gradient $\mathbf{g} = \nabla \mathcal{R}$. By taking account of Equation (15), we deduce that the p th component of the gradient vector is given by

$$\begin{aligned} g_p &= \frac{\partial \mathcal{R}}{\partial \xi_p} \\ &= \sum_k [\mathcal{I}(\xi, \Omega_{mk}^0) - \mathcal{I}_{mes}(\Omega_{mk}^0)] \frac{\partial \mathcal{I}}{\partial \xi_p}(\xi, \Omega_{mk}^0) \\ &= \int_{S_t} \sum_k [\mathcal{I}(\xi, \Omega_{mk}^0) - \mathcal{I}_{mes}(\Omega_{mk}^0)] \\ &\quad \times h_k(\mathbf{r}_t) \frac{\partial \mathcal{I}}{\partial \xi_p}(\mathbf{r}_t, \Omega_{mk}^0) dS_t, \end{aligned} \tag{A56}$$

where

$$h_k(\mathbf{r}_t) = \begin{cases} \frac{1}{A_{tmk}}, & \mathbf{r}_t \in S_{tm}(\Omega_{mk}^0) \\ 0, & \text{otherwise} \end{cases}, \tag{A57}$$

$A_{tmk} = R_{Dk}^2 \Omega_{FOV} / \mu_{mk}^0$ and $R_{Dk} = |\mathbf{r}_D - \mathbf{r}_{tk}^0|$. Further on, from Equation (25), we find that the adjoint source function reads as

$$Q^\dagger(\mathbf{r}, \Omega) = \sum_k F_{0k}^\dagger(\mathbf{r}_t) \delta(z - L_z) \delta(\Omega - \Omega_{mk}^0), \tag{A58}$$

where

$$F_{0k}^\dagger(\mathbf{r}_t) = [\mathcal{I}(\xi, \Omega_{mk}^0) - \mathcal{I}_{mes}(\Omega_{mk}^0)] h_k(\mathbf{r}_t). \tag{A59}$$

Thus, in each inversion step, we first solve the forward problem and compute the residual $\mathcal{I}(\xi, \Omega_{mk}^0) - \mathcal{I}_{mes}(\Omega_{mk}^0)$, and then solve the adjoint problem with the adjoint source function (A58) and compute each component of the gradient of the residual function g_p by means of Equation (26). In this approach, the pseudo-forward direct beam is a superposition of direct beams corresponding to each Ω_{mk}^0 . Consequently, if the angular domain covered by the measurements is large, it may happens that the spatial discontinuity of the pseudo-forward direct beam is not present.

References

1. Ingmann, P.; Veihelmann, B.; Langen, J.; Lamarre, D.; Stark, H.; Bazalgette Courrèges-Lacoste, G. Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and -5p. *Remote Sens. Environ. Sentin. Missions Oppor. Sci.* **2012**, *120*, 58–69. doi:10.1016/j.rse.2012.01.023.
2. Kokhanovsky, A. The influence of horizontal inhomogeneity on radiative characteristics of clouds: An asymptotic case study. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 817–825. doi:10.1109/TGRS.2003.811075.
3. Marshak, A.; Platnick, S.; Várnai, T.; Wen, G.; Cahalan, R.F. Impact of three-dimensional radiative effects on satellite retrievals of cloud droplet sizes. *J. Geophys. Res.* **2006**, *111*. doi:10.1029/2005jd006686.
4. Efremenko, D.S.; Schüssler, O.; Doicu, A.; Loyola, D. A stochastic cloud model for cloud and ozone retrievals from UV measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *184*, 167–179. doi:10.1016/j.jqsrt.2016.07.008.
5. Zhuravleva, T.B.; Nasrtdinov, I.M.; Russkova, T.V. Influence of 3D cloud effects on spatial-angular characteristics of the reflected solar radiation field. *Atmos. Ocean. Opt.* **2017**, *30*, 103–110. doi:10.1134/s102485601701016x.
6. Efremenko, D.S.; Doicu, A.; Loyola, D.; Trautmann, T. Fast Stochastic Radiative Transfer Models for Trace Gas and Cloud Property Retrievals Under Cloudy Conditions. In *Springer Series in Light Scattering*; Springer International Publishing: Basel, Switzerland, 2018; pp. 231–277. doi:10.1007/978-3-319-70796-9_3.
7. Evans, K. The spherical harmonic discrete ordinate method for three-dimensional atmospheric radiative transfer. *J. Atmos. Sci.* **1998**, *55*, 429–446. doi:10.1175/1520-0469(1998)055<0429:TSHDOM>2.0.CO;2.
8. Pincus, R.; Evans, K. Computational cost and accuracy in calculating three-dimensional radiative transfer: Results for new implementations of Monte Carlo and SHDOM. *J. Atmos. Sci.* **2009**, *66*, 3131–3146. doi:10.1175/2009JAS3137.1.
9. Spherical Harmonic Discrete Ordinate Method (SHDOM) for Atmospheric Radiative Transfer. Available online: <http://coloradolinux.com/shdom/> (accessed on 20 April 2019).
10. Doicu, A.; Efremenko, D.; Trautmann, T. A multi-dimensional vector spherical harmonics discrete ordinate method for atmospheric radiative transfer. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *118*, 121–131. doi:10.1016/j.jqsrt.2012.12.009.
11. Emde, C.; Barlakas, V.; Cornet, C.; Evans, F.; Korokin, S.; Ota, Y.; Labonnote, L.C.; Lyapustin, A.; Macke, A.; Mayer, B.; et al. IPRT polarized radiative transfer model intercomparison project—Phase A. *J. Quant. Spectrosc. Radiat. Transf.* **2015**, *164*, 8–36. doi:10.1016/j.jqsrt.2015.05.007.
12. Spurr, R.; Kurosu, T.; Chance, K. A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *J. Quant. Spectrosc. Radiat. Transf.* **2001**, *68*, 689–735. doi:10.1016/S0022-4073(00)00055-8.
13. Spurr, R. LIDORT and VLIDORT. Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In *Light Scattering Reviews*; Kokhanovsky, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 3, pp. 229–275. doi:10.1007/978-3-540-48546-9_7.
14. Marchuk, G. Equation for the value of information from weather satellite and formulation of inverse problems. *Cosm. Res* **1964**, *2*, 394–409.
15. Marchuk, G.I. *Adjoint Equations and Analysis of Complex Systems*; Springer: Cham, Switzerland, 1995. doi:10.1007/978-94-017-0621-6.
16. Landgraf, J.; Hasekamp, O.; Box, M.; Trautmann, T. A linearized radiative transfer model for ozone profile retrieval using the analytical forward-adjoint perturbation theory approach. *J. Geophys. Res. Atmos.* **2001**, *106*, 27291–27305. doi:10.1029/2001JD000636.
17. Ustinov, E.A. Adjoint sensitivity analysis of radiative transfer equation: Temperature and gas mixing ratio weighting functions for remote sensing of scattering atmospheres in thermal IR. *J. Quant. Spectrosc. Radiat. Transf.* **2001**, *68*, 195–211. doi:10.1016/s0022-4073(00)00022-4.
18. Box, M. Radiative perturbation theory: A review. *Env. Model. Softw.* **2002**, *17*, 95–106. doi:10.1016/S1364-8152(01)00056-1.
19. Walter, H.; Landgraf, J.; Hasekamp, O. Linearization of a pseudo-spherical vector radiative transfer model. *J. Quant. Spectrosc. Radiat. Transf.* **2004**, *85*, 251–283. doi:10.1016/S0022-4073(03)00228-0.
20. Ustinov, E. Atmospheric weighting functions and surface partial derivatives for remote sensing of scattering planetary atmospheres in thermal spectral region: General adjoint approach. *J. Quant. Spectrosc. Radiat. Transf.* **2005**, *92*, 351–371. doi:10.1016/j.jqsrt.2004.08.003.

21. Rozanov, V.; Rozanov, A. Relationship between different approaches to derive weighting functions related to atmospheric remote sensing problems. *J. Quant. Spectrosc. Radiat. Transf.* **2007**, *105*, 217–242. doi:10.1016/j.jqsrt.2006.12.006.
22. Martin, W.; Cairns, B.; Bal, G. Adjoint methods for adjusting three-dimensional atmosphere and surface properties to fit multi-angle/multi-pixel polarimetric measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *144*, 68–85. doi:10.1016/j.jqsrt.2014.03.030.
23. Martin, W.G.; Hasekamp, O.P. A demonstration of adjoint methods for multi-dimensional remote sensing of the atmosphere and surface. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *204*, 215–231. doi:10.1016/j.jqsrt.2017.09.031.
24. Wiscombe, W. The delta-M method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. *J. Atmos. Sci.* **1977**, *34*, 1408–1422. doi:10.1175/1520-0469(1977)034<1408:TDMRYA>2.0.CO;2.
25. Nakajima, T.; Tanaka, M. Algorithms for radiative intensity calculations in moderately thick atmos using a truncation approximation. *J. Quant. Spectrosc. Radiat. Transf.* **1988**, *40*, 51–69. doi:10.1016/0022-4073(88)90031-3.
26. Henyey, L.C.; Greenstein, J.L. Diffuse radiation in the Galaxy. *Astrophys. J.* **1941**, *93*, 70. doi:10.1086/144246.
27. Mie, G. Beitrage zur Optik trueber Medien, speziell kolloidaler Metalloesungen. *Ann. Phys.* **1908**, *330*, 377–445. doi:10.1002/andp.19083300302.
28. Bodhaine, B.A.; Wood, N.B.; Dutton, E.G.; Slusser, J.R. On rayleigh optical depth calculations. *J. Atmos. Ocean. Technol.* **1999**, *16*, 1854–1861. doi:10.1175/1520-0426(1999)016<1854:ORODC>2.0.CO;2.
29. Anderson, G.; Clough, S.; Kneizys, F.; Chetwynd, J.; Shettle, E. *AFGL Atmospheric Constituent Profiles (0–120 km)*; AFGL-TR-86-0110; Air Force Geophysics Laboratory: Hanscom Air Force Base, MA, USA, 1986.
30. Alexandrov, M.; Marshak, A.; Ackerman, A. Cellular Statistical Models of Broken Cloud Fields. Part I: Theory. *J. Atmos. Sci.* **2010**, *67*, 2125–2151. doi:10.1175/2010JAS3364.1.
31. García, V.M.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Linearized radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVER measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 241–251. doi:10.1016/j.jqsrt.2018.03.008.
32. Platt, U.; Stutz, J. *Differential Optical Absorption Spectroscopy: Principles and Applications*; Springer: Berlin/Heidelberg, Germany, 2008.
33. Doicu, A.; Efremenko, D.; Trautmann, T. An analysis of the short-characteristic method for the spherical harmonic discrete ordinate method (SHDOM). *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *119*, 114–127. doi:10.1016/j.jqsrt.2012.12.024.

