Journal of Quantitative Spectroscopy & Radiative Transfer 144 (2014) 68–85

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

Adjoint methods for adjusting three-dimensional atmosphere and surface properties to fit multi-angle/multi-pixel polarimetric measurements

William Martin [a,*], Brian Cairns [b], Guillaume Bal [a]

[a] Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA
[b] NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA

ARTICLE INFO

Article history:
Received 10 January 2014
Received in revised form
27 March 2014
Accepted 28 March 2014
Available online 5 April 2014

Keywords:
Adjoint methods
Three-dimensional vector radiative transfer
Linearization
Remote sensing
Parameter derivatives

ABSTRACT

This paper derives an efficient procedure for using the three-dimensional (3D) vector radiative transfer equation (VRTE) to adjust atmosphere and surface properties and improve their fit with multi-angle/multi-pixel radiometric and polarimetric measurements of scattered sunlight. The proposed adjoint method uses the 3D VRTE to compute the measurement misfit function and the adjoint 3D VRTE to compute its gradient with respect to all unknown parameters. In the remote sensing problems of interest, the scalar-valued misfit function quantifies agreement with data as a function of atmosphere and surface properties, and its gradient guides the search through this parameter space. Remote sensing of the atmosphere and surface in a three-dimensional region may require thousands of unknown parameters and millions of data points. Many approaches would require calls to the 3D VRTE solver in proportion to the number of unknown parameters or measurements. To avoid this issue of scale, we focus on computing the gradient of the misfit function as an alternative to the Jacobian of the measurement operator. The resulting adjoint method provides a way to adjust 3D atmosphere and surface properties with only two calls to the 3D VRTE solver for each spectral channel, regardless of the number of retrieval parameters, measurement view angles or pixels. This gives a procedure for adjusting atmosphere and surface parameters that will scale to the large problems of 3D remote sensing. For certain types of multi-angle/multi-pixel polarimetric measurements, this encourages the development of a new class of three-dimensional retrieval algorithms with more flexible parametrizations of spatial heterogeneity, less reliance on data screening procedures, and improved coverage in terms of the resolved physical processes in the Earth's atmosphere.

1. Introduction

More accurate and complete monitoring of cloud and aerosol properties is needed to reduce uncertainties in both the radiative forcing of climate and feedbacks between the radiative forcing and changes in global temperature that are the result of changes to clouds and their properties [1]. While multi-angle polarimetric measurements and plane parallel retrieval methods provide the capabilities necessary for regions that are horizontally homogeneous [2–4], the retrieval of aerosols in broken cloud fields and near cloud edges remains an open challenge which limits the coverage and accuracy of retrievals [5–7]. Using the three-dimensional (3D) vector radiative transfer equation (VRTE) can address this issue by explicitly accounting for the spatial distribution of solar illumination, scattering material, and polarimetric measurements. In contrast to plane-parallel and spherical models for radiative transfer, the 3D VRTE places no default

* Corresponding author. Tel.: +1 212 678 5567.
E-mail address: wgm2111@columbia.edu (W. Martin).

http://dx.doi.org/10.1016/j.jqsrt.2014.03.030

Nomenclature

Vectors are assigned to bold lowercase letters, matrices are assigned to bold uppercase letters, and integral operators are assigned to script uppercase letters.

Remote sensing problem

n, N	index for unknowns and total number
m, M	index for measurements and total number
L	number of wavelengths
$\boldsymbol{a}, a^n$	vector of unknowns and elements
$\hat{\boldsymbol{y}}, \hat{y}^m$	measurements vector and elements
$\boldsymbol{y}(\boldsymbol{a}), y^m(\boldsymbol{a})$	measurement model vector and elements
$\Phi(\boldsymbol{a})$	misfit function

Domain

$\boldsymbol{x}$	position in space
$h(\boldsymbol{x})$	signed distance to boundary
$\nabla h(\boldsymbol{x})$	outward pointing normal
$\boldsymbol{v}$	direction unit vector
$\mathbb{S}^2$	unit sphere
D	spatial domain
∂D	spatial domain boundary
$D \times \mathbb{S}^2$	internal set
Γ_+	outgoing set
Γ_-	incoming set

Single scattering

$\sigma(\boldsymbol{x}; \boldsymbol{a})$	volume extinction coefficient
$\boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}'; \boldsymbol{a})$	volume scattering kernel
$\boldsymbol{R}(\boldsymbol{x}, \boldsymbol{v}_-, \boldsymbol{v}_+; \boldsymbol{a})$	surface reflection kernel

Forward vector radiative transfer

$\boldsymbol{u}$	Stokes vector solution
$\boldsymbol{f}_\odot$	internal light source (internal forward source)
$\boldsymbol{g}_\odot$	solar illumination (incoming forward source)
$\Delta \boldsymbol{f}^n_\odot$	internal source for parameter derivatives
$\Delta \boldsymbol{g}^n_\odot$	incoming source for parameter derivatives
$\mathcal{Z}$	scattering operator
$\mathcal{R}$	reflection operator
$\mathcal{U}_{\boldsymbol{a}}$	forward solution operator

Adjoint vector radiative transfer

$\boldsymbol{w}$	adjoint Stokes vector solution
$\boldsymbol{p}^m_\odot$	polarization analyzer (internal adjoint source)
$\boldsymbol{q}^m_\odot$	polarization analyzer (outgoing adjoint source)
$\Delta \boldsymbol{p}_\odot$	internal source vector for misfit gradient
$\Delta \boldsymbol{q}_\odot$	outgoing source vector for misfit gradient
$\mathcal{Z}^*$	adjoint of scattering operator
$\mathcal{R}^*$	adjoint of reflection operator
$\mathcal{U}^*_{\boldsymbol{a}}$	adjoint solution operator

Forward integral equations

$\mathcal{T}_{00}$	forward streaming operator (internal-to-internal)
$\mathcal{T}_{-0}$	forward streaming operator (incoming-to-internal)
$\mathcal{T}_{0+}$	forward streaming operator (internal-to-outgoing)
$\mathcal{T}_{-+}$	forward streaming operator (incoming-to-outgoing)
$\boldsymbol{f}$	internal solution to the forward integral equations
$\boldsymbol{g}$	incoming solution to the forward integral equations

Adjoint integral equations

$\mathcal{T}^*_{00}$	adjoint streaming operator (internal-to-internal)
$\mathcal{T}^*_{-0}$	adjoint streaming operator (internal-to-incoming)
$\mathcal{T}^*_{0+}$	adjoint streaming operator (outgoing-to-internal)
$\mathcal{T}^*_{-+}$	adjoint streaming operator (outgoing-to-incoming)
$\boldsymbol{p}$	internal solution to the adjoint integral equations
$\boldsymbol{q}$	outgoing solution to the adjoint integral equations

Subscripts

$\odot$	indicates a source vector, i.e. right-hand side
0	defined on or related to the internal set
$+$	defined on or related to the outgoing set
$-$	defined on or related to the incoming set

Superscripts

T	transpose of a vector or matrix
$*$	adjoint of an operator
$'$, $''$	integration variables

restrictions on the spatial variability of the atmosphere and surface [8]. Work to extend coverage with 3D methods has shown promise for determining average cloud optical thickness [9] and cloud top height [10]. However, as a side effect of the increased flexibility, the 3D VRTE leads to retrieval problems with many more unknown parameters and multi-pixel measurement constraints. A significant concern is therefore the extent to which a proposed algorithm scales "gracefully" to large problems. The objective of this work is to formulate an adjoint method for the 3D VRTE which maintains the scalability required for the application to atmospheric remote sensing problems.

Adjoint methods and other linearization procedures can reduce the number of radiative transfer simulations needed over the course of an iterative procedure for fitting data [11,12]. During each iteration, the current estimate of the atmosphere and surface properties is adjusted to improve its fit with measurements. Making the right adjustment requires knowledge of how the fit will change when the adjustment is made. This in turn entails solving the 3D VRTE for each wavelength to evaluate the misfit between model and measurements and its gradient with respect to unknown parameters. It is worth noting that, brute-force numerical differencing could be used to compute the gradient with $\mathcal{O}(LN)$ radiative transfer computations, where N is the number of parameters and L is the number of wavelengths. Adjoint methods provide an alternative route to computing this derivative, and analogous work in the field of medical imaging shows that the number of calls to the radiative transfer solver can be as low as two: one forward solve for evaluating the misfit with data and one adjoint solve for computing the gradient with respect to retrieval parameters [13–15]. Although this result was specific to angularly averaged measurements in the frequency domain with wave length-independent parameters, it is an example of a scalable adjoint method derived for an analogous scalar transport equation. We define a retrieval adjustment procedure as *scalable* if it can be applied to problems with arbitrarily many measurement constraints and unknown parameters without requiring additional calls to a radiative transfer simulation. A scalable method can require $\mathcal{O}(L)$ radiative transfer solutions at each step, but not $\mathcal{O}(LN)$ or $\mathcal{O}(M)$, where M is the number of measurements (including all wavelengths, view angles and pixels).

Previous work on adjoint/linearization methods for remote sensing provides this kind of improved efficiency for plane-parallel [16], spherical [17–19], and pseudo-spherical radiative transfer models [20–22]. Also, adjoint techniques have been used to approximate solutions to the 3D VRTE for atmospheric properties with small deviations from plane-parallel symmetry [23,24]. To our knowledge, the use of adjoint methods to develop a scalable remote sensing methodology that relies on the 3D VRTE has not hitherto been considered. This topic is of current interest due to the advancement of computational techniques, which simulate 3D scalar and vector RT in the Earth's atmosphere [25–29], including the recently released vectorized Spherical Harmonic Discrete Ordinate Method code by Evans [30]. So far, the primary application of these codes has been synthetic studies which assess errors associated with plane-parallel retrievals [31,32]. In our view an equally important direction of research deals with how to incorporate simulations of the 3D VRTE directly into cloud and aerosol retrieval algorithms [9]: for example, using multi-pixel methods to improve single-pixel retrievals by accounting for adjacency effects, or perhaps, using a futuristic 3D parametrization of clouds and aerosols to retrieve their spatial variability in complex scenes with broken cloud cover. The objective of this paper is to provide the necessary theoretical foundation for such endeavors, by extending adjoint methods to allow scalable computations of the misfit function and its gradient using codes that solve the 3D VRTE.

To ensure that the adjoint method derived here meets the needs of the atmospheric remote sensing community, the theoretical description of measurements is consistent with ground-based, air-borne, and space-borne polarimeters, and the parameter-adjustment methods are similar to those used in operational retrieval algorithms [16,33]. The method focuses on minimizing a misfit function for passive measurements of scattered sunlight, but active measurements and measurements at other wavelengths may be included as prior constraints on spatial variability: using high spectral resolution LIDAR to constrain the aerosol scattering coefficient [34] or microwave cloud tomography to constrain cloud-droplet volume concentration [35,36]. Moreover, we formulate the adjoint framework in a manner that is consistent with the complex microphysical parametrizations needed to model single-scattering properties in the Earth's atmosphere [37]. The procedure is outlined using the standard integro-differential form of the 3D VRTE and derived using an equivalent integral formulation, written using concise operator notation for the processes of streaming, scattering and reflection. The integral formulation and associated operators are related to existing numerical solutions, and we describe how to extend such codes to solve adjoint radiative transfer by using the reciprocity principle to write the adjoint Stokes-vector solution in terms of a slight modification to the usual forward solution.

Preliminary definitions are organized in Section 2 with the fundamental adjoint property asserted but left temporarily unproven. The use of adjoint methods in developing a scalable procedure for adjusting parameters as a part of a remote sensing methodology is described in Section 3. Then, the general framework for forward and adjoint 3D VRTE is derived in Section 4 and the fundamental adjoint property is proven in Theorem 1. Supplementary technical results are presented in Appendices A and B.

2. Preliminaries

This section introduces the theoretical framework of the forward and adjoint 3D VRTE as needed for large scale 3D remote sensing of the atmosphere and surface. We generalize adjoint methods to arbitrary boundary conditions, and this entails a mild reformulation of the forward 3D VRTE as a boundary value problem and proof of the *fundamental adjoint property* for the corresponding boundary value problem of adjoint 3D VRTE. The logical progression that we choose to follow is to define the forward and adjoint 3D VRTEs as independent boundary value problems. Then, we prove the fundamental adjoint property which relates them. The first task is to define the domain.

Definition 1 (*Domain*). The spatial region of interest is an open, connected, and bounded set $D \subset \mathbb{R}^3$ with smooth boundary $\partial D \subset \mathbb{R}^3$ — smooth to guarantee that the bounding surface has a continuous outward-pointing normal vector. The region is described implicitly by its *signed-distance-to-boundary* function:

$$h(\boldsymbol{x}) = \begin{cases} -\inf\limits_{\boldsymbol{x}' \in \partial D} \|\boldsymbol{x}-\boldsymbol{x}'\| & \text{for } \boldsymbol{x} \in (D \cup \partial D) \\ \inf\limits_{\boldsymbol{x}' \in \partial D} \|\boldsymbol{x}-\boldsymbol{x}'\| & \text{for } \boldsymbol{x} \in \mathbb{R}^3 \backslash (D \cup \partial D). \end{cases} \quad (1)$$

The useful properties of h are that it is continuous on $\mathbb{R}^3$ and differentiable near the boundary, with gradient $\nabla h(\boldsymbol{x})$ equal to the unit normal vector pointing out of the domain for each $\boldsymbol{x} \in \partial D$. The function value determines if a given point, $\boldsymbol{x} \in \mathbb{R}^3$, is inside the region of interest, $h(\boldsymbol{x}) < 0$; on its boundary, $h(\boldsymbol{x}) = 0$; or outside, $h(\boldsymbol{x}) > 0$.

Taking the direction vector $\boldsymbol{v}$ to be always in the unit sphere, $\mathbb{S}^2 \subset \mathbb{R}^3$; we define the three regions making up the domain of the 3D VRTE: the *internal set*,

$$D \times \mathbb{S}^2 = \{(\boldsymbol{x}, \boldsymbol{v}) : h(\boldsymbol{x}) < 0\}, \tag{2}$$

the *outgoing set*,

$$\Gamma_+ = \{(\boldsymbol{x}, \boldsymbol{v}_+) : h(\boldsymbol{x}) = 0 \text{ and } \boldsymbol{v}_+ \cdot \nabla h(\boldsymbol{x}) > 0\}, \tag{3}$$

and the *incoming set*,

$$\Gamma_- = \{(\boldsymbol{x}, \boldsymbol{v}_-) : h(\boldsymbol{x}) = 0 \text{ and } \boldsymbol{v}_- \cdot \nabla h(\boldsymbol{x}) < 0\}. \tag{4}$$

Inner products of Stokes-vector and source-vector functions are defined for each domain: the *internal inner product*,

$$\langle \boldsymbol{w}, \boldsymbol{u} \rangle_{D\times\mathbb{S}^2} = \int_{h(\boldsymbol{x})<0} \mathrm{d}V_{\boldsymbol{x}} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}}\, \boldsymbol{w}(\boldsymbol{x}, \boldsymbol{v}) \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{v}), \tag{5}$$

the *outgoing inner product*,

$$\langle \boldsymbol{q}, \boldsymbol{u} \rangle_{\Gamma_+} = \int_{h(\boldsymbol{x})=0} \mathrm{d}S_{\boldsymbol{x}} \int_{\boldsymbol{v}_+ \cdot \nabla h(\boldsymbol{x})>0} \mathrm{d}S_{\boldsymbol{v}_+}\, |\boldsymbol{v}_+ \cdot \nabla h(\boldsymbol{x})|\, \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{v}_+) \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{v}_+), \tag{6}$$

and the *incoming inner product*,

$$\langle \boldsymbol{w}, \boldsymbol{g} \rangle_{\Gamma_-} = \int_{h(\boldsymbol{x})=0} \mathrm{d}S_{\boldsymbol{x}} \int_{\boldsymbol{v}_- \cdot \nabla h(\boldsymbol{x})<0} \mathrm{d}S_{\boldsymbol{v}_-}\, |\boldsymbol{v}_- \cdot \nabla h(\boldsymbol{x})|\, \boldsymbol{w}(\boldsymbol{x}, \boldsymbol{v}_-) \cdot \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{v}_-). \tag{7}$$

These elementary inner products appear so often in pairs that is helpful to write the *forward inner product* as

$$\left\langle \begin{Bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{Bmatrix}, \begin{Bmatrix} \boldsymbol{u}|_{D\times\mathbb{S}^2} \\ \boldsymbol{u}|_{\Gamma_+} \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2 \oplus \Gamma_+} = \langle \boldsymbol{p}, \boldsymbol{u} \rangle_{D\times\mathbb{S}^2} + \langle \boldsymbol{q}, \boldsymbol{u} \rangle_{\Gamma_+}, \tag{8}$$

and the *adjoint inner product* as

$$\left\langle \begin{Bmatrix} \boldsymbol{w}|_{D\times\mathbb{S}^2} \\ \boldsymbol{w}|_{\Gamma_-} \end{Bmatrix}, \begin{Bmatrix} \boldsymbol{f} \\ \boldsymbol{g} \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2 \oplus \Gamma_-} = \langle \boldsymbol{w}, \boldsymbol{f} \rangle_{D\times\mathbb{S}^2} + \langle \boldsymbol{w}, \boldsymbol{g} \rangle_{\Gamma_-}. \tag{9}$$

Each of the three inner products in Eqs. (5)–(7) defines a vector space of *square-integrable* functions, for example the internal source vectors, $\boldsymbol{f}$, such that

$$\|\boldsymbol{f}\|^2_{D\times\mathbb{S}^2} = \langle \boldsymbol{f}, \boldsymbol{f} \rangle_{D\times\mathbb{S}^2} < \infty, \tag{10}$$

or the incoming source vectors, $\boldsymbol{g}$, such that

$$\|\boldsymbol{g}\|^2_{\Gamma_-} = \langle \boldsymbol{g}, \boldsymbol{g} \rangle_{\Gamma_-} < \infty. \tag{11}$$

From the operator point of view, the square-integrable functions in Eqs. (10) and (11) are vectors in a linear space, and linear operators will act in much the same way that matrices do. To guarantee this, we give symbolic representation only to linear operators that are bounded. By this convention, a linear operator with symbol, $\mathcal{L}$, will act on a square-integrable vector, $\boldsymbol{f}$, and return another square-integrable vector, $\mathcal{L}[\boldsymbol{f}]$. This follows from the definition of a bounded operator: the linear operator, $\mathcal{L}$, is *bounded* if there exists a value, C, so that

$$\|\mathcal{L}[\boldsymbol{f}]\| \le C \|\boldsymbol{f}\|, \tag{12}$$

for all square-integrable functions, $\boldsymbol{f}$. The smallest such value, C, is called the *operator norm*, $\|\mathcal{L}\|_{\mathrm{op}}$, and it will be needed in Section 4.2 to state the constraints on scattering and reflection that guarantee solve-ability of the 3D VRTE. The adjoint of an operator is defined in the usual sense, as the operator, $\mathcal{L}^*$, which satisfies the adjoint property:

$$\langle \boldsymbol{p}, \mathcal{L}[\boldsymbol{f}] \rangle = \langle \mathcal{L}^*[\boldsymbol{p}], \boldsymbol{f} \rangle. \tag{13}$$

The adjoint, $\mathcal{L}^*$, gives the alternative rule for evaluating the inner product in Eq. (13), so that numerical procedures may use whichever side is more efficient.

In summary, the three distinct subdomains for 3D vector radiative transfer are defined through the utility function, $h(\boldsymbol{x})$: the interior set, $D \times \mathbb{S}^2$; outgoing set, Γ_+; and the incoming set, Γ_-. Each subdomain has an inner product and a set of square-integrable functions. As a convention, we reserve operator notation for bounded linear operators to ensure similarity to matrix algebra. Lastly, the adjoint of a linear operator was defined.

2.1. Forward and adjoint 3D VRTEs

The purpose of this paper is to formulate an efficient procedure for adjusting unknown parameters as a part of an abstract remote sensing problem. The atmosphere and surface properties are described by an unknown parameter vector, $\boldsymbol{a} = (a^n)$ for $0 \le n < N$, from which physical single-scattering properties are derived: extinction, $\sigma(\boldsymbol{x}; \boldsymbol{a})$; scattering kernel, $\boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}'; \boldsymbol{a})$; and reflection kernel, $\boldsymbol{R}(\boldsymbol{x}, \boldsymbol{v}_-, \boldsymbol{v}_+; \boldsymbol{a})$. For a given illumination defined by incoming and internal light sources, the 3D VRTE provides a solution vector $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{v}; \boldsymbol{a})$ to be used in modeling each polarimetric measurement as an inner product over a pair of polarimetric response functions, $\boldsymbol{p}^m_\odot$ and $\boldsymbol{q}^m_\odot$:

$$y^m(\boldsymbol{a}) = \langle \boldsymbol{p}^m_\odot, \boldsymbol{u} \rangle_{D\times\mathbb{S}^2} + \langle \boldsymbol{q}^m_\odot, \boldsymbol{u} \rangle_{\Gamma_+}, \tag{14}$$

$$y^m(\boldsymbol{a}) = \left\langle \begin{Bmatrix} \boldsymbol{p}^m_\odot \\ \boldsymbol{q}^m_\odot \end{Bmatrix}, \begin{Bmatrix} \boldsymbol{u}|_{D\times\mathbb{S}^2} \\ \boldsymbol{u}|_{\Gamma_+} \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2 \oplus \Gamma_+}, \tag{15}$$

with measurement vector $\boldsymbol{y} = (y^m)$ for $0 \le m < M$. This motivates the boundary value problem of the 3D VRTE which defines the Stokes vector solution, $\boldsymbol{u}$, for incoming solar energy.

Definition 2 (*Forward 3D VRTE*). For a fixed parameter vector, $\boldsymbol{a}$, and the corresponding single-scattering properties, σ, $\boldsymbol{Z}$, and $\boldsymbol{R}$; the forward solution $\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{v}; \boldsymbol{a})$ is defined for square-integrable source vectors, $\boldsymbol{f}_\odot$ and $\boldsymbol{g}_\odot$, as the unique solution to the integro-differential equations of the forward 3D VRTE:

$$\boldsymbol{v} \cdot \nabla \boldsymbol{u} + \sigma \boldsymbol{u} - \mathcal{Z}[\boldsymbol{u}] = \boldsymbol{f}_\odot \quad \text{on } D \times \mathbb{S}^2, \tag{16}$$

$$\boldsymbol{u}|_{\Gamma_-} - \mathcal{R}[\boldsymbol{u}|_{\Gamma_+}] = \boldsymbol{g}_\odot \quad \text{on } \Gamma_-. \tag{17}$$

The integral operator for scattering is defined as

$$\mathcal{Z}[\boldsymbol{u}](\boldsymbol{x}, \boldsymbol{v}) = \frac{1}{4\pi} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}'}\, \boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}') \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{v}'), \tag{18}$$

for $(\boldsymbol{x}, \boldsymbol{v}) \in D \times \mathbb{S}^2$, and the integral operator for reflection is defined as

$$\mathcal{R}[\boldsymbol{u}|_{\Gamma_+}](\boldsymbol{x},\boldsymbol{\nu}_-) = \frac{1}{2\pi}\int_{\boldsymbol{\nu}_+\cdot\nabla h(\boldsymbol{x})>0} \mathrm{d}S_{\boldsymbol{\nu}_+}\ |\boldsymbol{\nu}_+\cdot\nabla h(\boldsymbol{x})|\boldsymbol{R}(\boldsymbol{x},\boldsymbol{\nu}_-,\boldsymbol{\nu}_+)\cdot\boldsymbol{u}(\boldsymbol{x},\boldsymbol{\nu}_+), \tag{19}$$

for $(\boldsymbol{x},\boldsymbol{\nu}_-)\in\Gamma_-$. Coupling with boundary conditions is imposed by spatial continuity along outgoing and incoming directions:

$$\boldsymbol{u}|_{\Gamma_+}(\boldsymbol{x},\boldsymbol{\nu}_+) = \lim_{t\searrow 0}\boldsymbol{u}|_{D\times\mathbb{S}^2}(\boldsymbol{x}-t\boldsymbol{\nu}_+,\boldsymbol{\nu}_+), \tag{20}$$

$$\boldsymbol{u}|_{\Gamma_-}(\boldsymbol{x},\boldsymbol{\nu}_-) = \lim_{t\searrow 0}\boldsymbol{u}|_{D\times\mathbb{S}^2}(\boldsymbol{x}+t\boldsymbol{\nu}_-,\boldsymbol{\nu}_-), \tag{21}$$

for $(\boldsymbol{x},\boldsymbol{\nu}_+)\in\Gamma_+$ and $(\boldsymbol{x},\boldsymbol{\nu}_-)\in\Gamma_-$.

Assuming existence and uniqueness for the moment, we define the solution operator for the boundary value problem of the forward 3D VRTE, $\mathcal{U}_{\boldsymbol{a}}\{:\}$:

$$\begin{Bmatrix}\boldsymbol{u}|_{D\times\mathbb{S}^2}\\ \boldsymbol{u}|_{\Gamma_+}\end{Bmatrix} = \mathcal{U}_{\boldsymbol{a}}\begin{Bmatrix}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{Bmatrix}. \tag{22}$$

The forward solution operator, $\mathcal{U}_{\boldsymbol{a}}$, is a 2×2 matrix of integral operators which is parametrized by $\boldsymbol{a}$ and acts on internal and incoming source vectors, $\boldsymbol{f}_\odot$ and $\boldsymbol{g}_\odot$, to give the Stokes vector solution on the internal and outgoing sets, $\boldsymbol{u}|_{D\times\mathbb{S}^2}$ and $\boldsymbol{u}|_{\Gamma_+}$. An explicit formula for the forward solution operator is derived in Section 4.2.1 and given by Eq. (112).

The solution operator, $\mathcal{U}_{\boldsymbol{a}}$, plays the role of a forward solver in the present discussion of an abstract remote sensing problem — one call to a forward solver is equivalent to evaluating the solution operator, $\mathcal{U}_{\boldsymbol{a}}$, for one pair of source vectors, $\boldsymbol{f}_\odot$ and $\boldsymbol{q}_\odot$. Using this operator we write the model for polarimetric measurements in Eq. (15) as the inner product of detector response functions, $\boldsymbol{p}^m_\odot$ and $\boldsymbol{q}^m_\odot$, with the Stokes vector solution for solar illumination:

$$y^m(\boldsymbol{a}) = \left\langle \begin{Bmatrix}\boldsymbol{p}^m_\odot\\ \boldsymbol{q}^m_\odot\end{Bmatrix}, \mathcal{U}_{\boldsymbol{a}}\begin{Bmatrix}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{Bmatrix}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}. \tag{23}$$

This expression allows the computation of all measurements with one call to the solution operator, $\mathcal{U}_{\boldsymbol{a}}$, followed by relatively inexpensive integrations over the polarimetric response function for each detector. Therefore, Eq. (23) is well suited to the task of evaluating many different measurements for fixed internal and incoming source vectors.

However, in remote sensing applications where unknown atmosphere and surface parameters, $\boldsymbol{a}$, are adjusted to fit data, the computation also requires variation of source functions for computing the components of the gradient of the misfit function, $\partial\Phi/\partial a^n$. We will see in Section 3 that the desired quantity has the following form:

$$-\frac{\partial\Phi}{\partial a^n} = \left\langle \begin{Bmatrix}\Delta\boldsymbol{p}_\odot\\ \Delta\boldsymbol{q}_\odot\end{Bmatrix}, \mathcal{U}_{\boldsymbol{a}}\begin{Bmatrix}\Delta\boldsymbol{f}^n_\odot\\ \Delta\boldsymbol{g}^n_\odot\end{Bmatrix}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}, \tag{24}$$

for fixed adjoint source vectors, $\Delta\boldsymbol{p}_\odot$ and $\Delta\boldsymbol{q}_\odot$, and forward source vectors, $\Delta\boldsymbol{f}^n_\odot$ and $\Delta\boldsymbol{g}^n_\odot$, corresponding to derivatives with respect to each parameter, for $0\le n<N$. Forward and adjoint source vectors appearing in Eq. (24) are defined explicitly by Eqs. (42), (43), (46), and (47). To avoid repeated calls to the solver, $\mathcal{U}_{\boldsymbol{a}}$, we seek the adjoint of the forward solution operator. This adjoint operator, $(\mathcal{U}_{\boldsymbol{a}})^*$, is defined to satisfy the adjoint property:

$$\left\langle \begin{Bmatrix}\Delta\boldsymbol{p}_\odot\\ \Delta\boldsymbol{q}_\odot\end{Bmatrix}, \mathcal{U}_{\boldsymbol{a}}\begin{Bmatrix}\Delta\boldsymbol{f}^n_\odot\\ \Delta\boldsymbol{g}^n_\odot\end{Bmatrix}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+} = \left\langle (\mathcal{U}_{\boldsymbol{a}})^*\begin{Bmatrix}\Delta\boldsymbol{p}_\odot\\ \Delta\boldsymbol{q}_\odot\end{Bmatrix}, \begin{Bmatrix}\Delta\boldsymbol{f}^n_\odot\\ \Delta\boldsymbol{g}^n_\odot\end{Bmatrix}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-} \tag{25}$$

Notice that all solver operations now act on the fixed adjoint source vectors, $\Delta\boldsymbol{p}_\odot$ and $\Delta\boldsymbol{q}_\odot$, so that the change in fit can be computed for any unknown parameter by integration. This motivates the definition of the boundary value problem for the adjoint 3D VRTE.

Definition 3 (*Adjoint 3D VRTE*). For fixed parameter, $\boldsymbol{a}$, and single-scattering properties, σ, $\boldsymbol{Z}$, and $\boldsymbol{R}$; we define the adjoint solution, $\boldsymbol{w}$, for square-integrable adjoint source vectors, $\boldsymbol{p}_\odot$ and $\boldsymbol{q}_\odot$, as the unique solution to the adjoint 3D VRTE:

$$-\boldsymbol{\nu}\cdot\nabla\boldsymbol{w}+\sigma\boldsymbol{w}-\mathcal{Z}^*[\boldsymbol{w}] = \boldsymbol{p}_\odot \quad \text{on } D\times\mathbb{S}^2, \tag{26}$$

$$\boldsymbol{w}|_{\Gamma_+}-\mathcal{R}^*[\boldsymbol{w}|_{\Gamma_-}] = \boldsymbol{q}_\odot \quad \text{on } \Gamma_+. \tag{27}$$

The adjoint-scattering operator is defined as

$$\mathcal{Z}^*[\boldsymbol{w}](\boldsymbol{x},\boldsymbol{\nu}) = \frac{1}{4\pi}\int_{\mathbb{S}^2}\mathrm{d}S_{\boldsymbol{\nu}'}\ \boldsymbol{Z}(\boldsymbol{x},\boldsymbol{\nu}',\boldsymbol{\nu})^T\cdot\boldsymbol{w}(\boldsymbol{x},\boldsymbol{\nu}'), \tag{28}$$

for $(\boldsymbol{x},\boldsymbol{\nu})\in D\times\mathbb{S}^2$, and the adjoint-reflection operator as

$$\mathcal{R}^*[\boldsymbol{w}|_{\Gamma_-}](\boldsymbol{x},\boldsymbol{\nu}_+) = \frac{1}{2\pi}\int_{\boldsymbol{\nu}_-\cdot\nabla h(\boldsymbol{x})<0}\mathrm{d}S_{\boldsymbol{\nu}_-}\ |\boldsymbol{\nu}_-\cdot\nabla h(\boldsymbol{x})|\boldsymbol{R}(\boldsymbol{x},\boldsymbol{\nu}_-,\boldsymbol{\nu}_+)^T\cdot\boldsymbol{w}(\boldsymbol{x},\boldsymbol{\nu}_-), \tag{29}$$

for $(\boldsymbol{x},\boldsymbol{\nu}_+)\in\Gamma_+$. Coupling with boundary conditions is imposed by spatial continuity along outgoing and incoming directions:

$$\boldsymbol{w}|_{\Gamma_+}(\boldsymbol{x},\boldsymbol{\nu}_+) = \lim_{t\searrow 0}\boldsymbol{w}|_{D\times\mathbb{S}^2}(\boldsymbol{x}-t\boldsymbol{\nu}_+,\boldsymbol{\nu}_+), \tag{30}$$

$$\boldsymbol{w}|_{\Gamma_-}(\boldsymbol{x},\boldsymbol{\nu}_-) = \lim_{t\searrow 0}\boldsymbol{w}|_{D\times\mathbb{S}^2}(\boldsymbol{x}+t\boldsymbol{\nu}_-,\boldsymbol{\nu}_-), \tag{31}$$

for $(\boldsymbol{x},\boldsymbol{\nu}_+)\in\Gamma_+$ and $(\boldsymbol{x},\boldsymbol{\nu}_-)\in\Gamma_-$.

Assuming existence and uniqueness for the moment, we conclude by defining the adjoint solution operator. For each parameter $\boldsymbol{a}$, the adjoint solution operator, $\mathcal{U}^*_{\boldsymbol{a}}$, maps adjoint source vectors, $\boldsymbol{p}_\odot$ and $\boldsymbol{q}_\odot$, to the adjoint Stokes vector on the internal and incoming sets:

$$\begin{Bmatrix}\boldsymbol{w}|_{D\times\mathbb{S}^2}\\ \boldsymbol{w}|_{\Gamma_-}\end{Bmatrix} = \mathcal{U}^*_{\boldsymbol{a}}\begin{Bmatrix}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{Bmatrix}. \tag{32}$$

The explicit form of this 2×2 matrix of integral operators is derived in Section 4.2.2 and given by Eq. (119).

Since the boundary value problems for the forward and adjoint 3D VRTE are defined independently, the adjoint property, $(\mathcal{U}_{\boldsymbol{a}})^* = \mathcal{U}^*_{\boldsymbol{a}}$, requires proof, and this is done in Theorem 1. In summary the forward 3D VRTE is stated in Definition 2 and can be used to evaluate M radiometric measurements with $\mathcal{O}(L)$ calls to the forward solution operator, $\mathcal{U}_{\boldsymbol{a}}$. We asserted that the adjustment of atmosphere and surface properties would require evaluation of the left hand side of Eq. (25), and noted that this would require $\mathcal{O}(LN)$ calls to $\mathcal{U}_{\boldsymbol{a}}$. In Section 3, we show how the left hand side of Eq. (25) arises naturally as the required quantity in an iterative search, and how evaluating with the adjoint alternative on the right-hand side of Eq. (25)

leads to a scalable procedure for adjusting unknown parameters as a part of a remote sensing algorithm. That is, one which requires only $\mathcal{O}(L)$ evaluations of the solution operators at each step, making the number of calls independent of the size of the problem.

3. Application to remote sensing

In the context of remote sensing of the Earth's 3D atmosphere and surface, the adjoint method provides a means of adjusting arbitrarily many atmosphere and surface parameters to improve their fit with arbitrarily many polarimetric measurements without changing the number of 3D VRTE simulations needed. To provide a concrete example of this, consider the task of retrieving cloud, aerosol, and surface properties for Yellowstone National Park which has a surface area of ten thousand km^2. For measurements, suppose we have access to a hundred satellite images of the park — taken with a single-spectral channel, from different perspectives, and with 1 km resolution. These data provide one million constraints, $M = 1 \times 10^6$. Suppose also that a discretization of the atmosphere and surface is constructed with a total of one-thousand volume and surface elements, and that the volume single-scattering and surface reflection properties are represented by an average of ten parameters per discrete element. These rough assumptions would result in the use of ten thousand parameters to describe the cloud, aerosol, and surface properties, $N = 1 \times 10^4$.

At each step in the retrieval algorithm we must adjust these ten thousand parameters to decrease the collective misfit with one million measurements. We note that if one were to linearize the measurement operator for this problem then the Jacobian matrix, consisting of elements $\partial y^m / \partial a^n$, would have ten billion entries. One of the key strategies of the method outlined here is to avoid the computation and storage of the Jacobian matrix, working instead with the misfit function and its gradient. Since the misfit function is scalar valued, its gradient in this case has only ten thousand elements. Even in this extreme example, the adjoint method described here provides a procedure for adjusting parameters to improve the collective fit with all data using only two calls to the 3D VRTE solver. For multi-wavelength data the required number of calls is $\mathcal{O}(L)$.

Although this example describes the scalability of the adjoint method using a futuristic application involving a full 3D reconstruction of cloud and aerosol properties, it is worth noting that this fits within a hierarchy of methods that start with retrievals assuming a plane-parallel atmosphere and with each pixel being an independent column [16,38,39]. This approach has been extended by Dubovik et al. [33] to include statistical modeling of the co-variation of atmospheric and surface properties in different pixels within the framework of a multi-pixel optimal estimation scheme. A natural addition to their usage of a multi-pixel prior probability distribution would be the usage of a multi-pixel measurement operator, in which the 3D VRTE couples the radiative effects of nearby columns and allows, in the context of clear sky observations, for the proper account of adjacency effects. In this context, the adjoint method would provide a means of adjusting plane-parallel retrievals to correct for 3D, or adjacency effects. Moreover, the scalability result implies that the number of calls to the 3D VRTE solver is independent of the number of columns or pixels so that adjustments can be made to many pixels at once.

The remainder of Section 3 will summarize the methodology which makes this scalability possible. Qualitatively the adjoint method accomplishes this by associating the residual misfit between model and measurements with a single source distribution for the adjoint 3D VRTE. The residual for each individual image pixel is defined as the difference between model and observation, and is specific to the location of the instrument, the field of view of the pixel, and the sensitivity of the polarization analyzer. The weighted sum of these localized and directed residuals over all image pixels gives a single distribution of adjoint sources, and the adjoint 3D VRTE is solved to back-propagate this residual through all orders of multiple scattering. Then, simple integrals can be evaluated to determine the change in fit for all possible adjustments to the unknown parameters. This alternative way of thinking provides a rule for computing the misfit gradient with the desired scalability. The subsequent use of the misfit gradient in numerical optimization routines is discussed in Section 3.3, where each iterative adjustment to cloud, aerosol, and surface properties is written as a solution to an $N \times N$ system of linear equations.

3.1. Model for polarimetric measurements

The chosen setting for these results is the Earth's 3D atmospheric shell bounded between the Earth–atmosphere interface and an arbitrarily large radius out to space. The results extend to any smooth connected sub-region of interest, provided that reasonable horizontal boundary conditions are imposed. In the context of the radiative transfer model of light propagation, incoming solar radiation is scattered in the atmosphere and reflected by the surface, causing measurable radiative effects that vary with location, direction, and polarization.

To setup an abstract remote sensing problem, let $\hat{\boldsymbol{y}} = (\hat{y}^m)$ for $0 \le m < M$ be a vector of single-wavelength multi-angle/multi-pixel polarimetric data taken by a ground, air, or space borne instrument. Let $\boldsymbol{a} = (a^n)$ for $0 \le n < N$ be a vector of N unknown parameters which define a three-dimensional distribution of cloud, aerosol, and surface properties. In practice there will be constraints on the parameter vector, $\boldsymbol{a}$, to guarantee a reasonable physical interpretation (e.g. non-negative particle concentrations). Enforcing such constraints by finitely many linear equalities and convex inequalities is ideal, to give a numerically convenient description of the convex set of all possible states of the atmosphere and surface. We now describe in three steps, how the definitions of Section 2 lead to a useful model for polarimetric measurements of atmospheric radiation as they depend on cloud, aerosol, and surface properties.

First, the values of the volume-extinction coefficient, $\sigma(\boldsymbol{x};\boldsymbol{a})$; the volume-scattering matrix, $\boldsymbol{Z}(\boldsymbol{x},\boldsymbol{\nu},\boldsymbol{\nu}';\boldsymbol{a})$; and the surface-reflection matrix, $\boldsymbol{R}(\boldsymbol{x},\boldsymbol{\nu}_-,\boldsymbol{\nu}_+;\boldsymbol{a})$, must be written explicitly as smooth functions of the vector of parameters, $\boldsymbol{a}$. They must be smooth to guarantee the existence of derivatives, $\partial\sigma/\partial a^n$, $\partial\boldsymbol{Z}/\partial a^n$, and $\partial\boldsymbol{R}/\partial a^n$; and also to guarantee that the integral operators defined in Eqs. (46) and (47) will return square-integrable functions. Furthermore, for all feasible values of the parameter vector, the single-scattering properties must satisfy solve-ability criteria given in Section 4 by Eq. (109). The parametrization of single-scattering properties incorporates both spatial and micro-physical variability. Surface reflection at the Earth–atmosphere boundary is characterized by several spatially dependent parameters. Volume-extinction and scattering properties are modeled as a linear combination of contributions from molecular scattering and various *modes* of airborne-particle. For each mode there are parameters that define loading, size distribution, shape, and complex refractive index. For the present discussion, we assume that there is a well-defined functional relationship between parameters and single-scattering properties, that the functions are smooth with respect to parameters, and that the single-scattering properties satisfy the solve-ability criteria.

Second, the 3D VRTE in Definition 2 is used to solve for the multiple scattering of incoming solar radiation and determine the Stokes vector solution, $\boldsymbol{u}(\boldsymbol{x},\boldsymbol{v};\boldsymbol{a})$. For each feasible parameter vector, $\boldsymbol{a}$, the method requires a solver, $\mathcal{U}_{\boldsymbol{a}}$, that acts on forward source vectors for solar illumination, $\boldsymbol{f}_{\odot}$ and $\boldsymbol{g}_{\odot}$, and returns the Stokes vector solution (restricted to the internal and outgoing sets):

$$\left\{\begin{array}{l}\boldsymbol{u}|_{D\times\mathbb{S}^2}\\ \boldsymbol{u}|_{\Gamma_+}\end{array}\right\}=\mathcal{U}_{\boldsymbol{a}}\left\{\begin{array}{l}\boldsymbol{f}_{\odot}\\ \boldsymbol{g}_{\odot}\end{array}\right\}. \tag{33}$$

The restrictions, $\boldsymbol{u}|_{D\times\mathbb{S}^2}$ and $\boldsymbol{u}|_{\Gamma_+}$, of the full Stokes vector solution, $\boldsymbol{u}$, provide all the information that is necessary to model multi-angle polarimetric measurements.

The third step involves expressing the measurable quantities which correspond to elements of the data vector, $\hat{\boldsymbol{y}}$, as inner products of the solution with detector response functions. Internal measurements are computed as the inner product of the internal Stokes vector, $\boldsymbol{u}|_{D\times\mathbb{S}^2}$, with a polarization analyzer defined on the internal set, $\boldsymbol{p}^m_{\odot}(\boldsymbol{x},\boldsymbol{v})$:

$$y^m(\boldsymbol{a})=\langle\boldsymbol{p}^m_{\odot},\boldsymbol{u}\rangle_{D\times\mathbb{S}^2}\quad\text{for } 0\le m<M_1. \tag{34}$$

Outgoing measurements are computed as the inner product of the outgoing Stokes vector, $\boldsymbol{u}|_{\Gamma_+}$, with a polarization analyzer defined on the outgoing set, $\boldsymbol{q}^m_{\odot}(\boldsymbol{x},\boldsymbol{v}_+)$:

$$y^m(\boldsymbol{a})=\langle\boldsymbol{q}^m_{\odot},\boldsymbol{u}|_{\Gamma_+}\rangle_{\Gamma_+}\quad\text{for } M_1\le m<M. \tag{35}$$

Given the clear apertures of typical Earth observing instruments we note that the polarization analyzers will be effectively Dirac-delta distributions in the location variable with angular integrations being determined by the field of view of the given sensor or pixel. While Dirac distributions are not square-integrable functions, they may be approximated as such to within discretization error. Aircraft measurements taken inside the domain result in a weight that is localized to a point in space:

$$\boldsymbol{p}^m_{\odot}(\boldsymbol{x},\boldsymbol{v})\propto\delta(\boldsymbol{x}^m-\boldsymbol{x}). \tag{36}$$

For ground based instruments, e.g. AERONET [40], a natural route to computing measurements is to localize the position of the instrument to a point on the boundary:

$$\boldsymbol{q}^m_{\odot}(\boldsymbol{x}_+,\boldsymbol{v}_+)\propto\frac{\delta(\boldsymbol{x}^m-\boldsymbol{x}_+)}{|\boldsymbol{v}_+\cdot\nabla h(\boldsymbol{x}^m)|}. \tag{37}$$

Satellite measurements may be taken at a great distance away from the domain, so in this case we suggest projecting the data to the outgoing boundary. This results in a weighting distribution that is singular in direction:

$$\boldsymbol{q}^m_{\odot}(\boldsymbol{x}_+,\boldsymbol{v}_+)\propto\frac{\delta\left(\dfrac{\boldsymbol{x}^m-\boldsymbol{x}_+}{|\boldsymbol{x}^m-\boldsymbol{x}_+|}-\boldsymbol{v}_+\right)}{|\boldsymbol{v}_+\cdot\nabla h(\boldsymbol{x}_+)|}. \tag{38}$$

Again, this singular distribution can be integrated over an actual instrument field of view in practice, with a scaling by the reciprocal, $|\boldsymbol{v}_+\cdot\nabla h(\boldsymbol{x}_+)|^{-1}$, to counteract the weight that appears in the inner product $\langle\cdot,\cdot\rangle_{\Gamma_+}$. Internal measurements require no such scaling. In this way, the formalism used is shown to be consistent with common types of radiometric and polarimetric measurements taken of the atmosphere.

To summarize the process of modeling polarimetric measurements the three steps are as follows: (1) Single-scattering properties are written as smooth functions of N parameters that describe cloud, aerosol, and surface properties. (2) The Stokes vector solving the 3D VRTE is computed as a model for the spatially and directionally dependent field of radiative energy in the atmosphere. (3) Each individual polarimetric measurement is represented as an inner product of the internal or outgoing solution with a polarimetric response function, defined on the same domain. This procedure provides the theoretical connection between observations and the retrieval target of atmospheric composition.

3.2. Data misfit and gradient calculation

In the abstract remote sensing problem, we aim to use multi-angle polarimetric data stored in the M-dimensional vector, $\hat{\boldsymbol{y}}$, to adjust the 3D atmosphere and surface parameters stored in the N-dimensional vector, $\boldsymbol{a}$, and reduce the measurement residual, $\hat{\boldsymbol{y}}-\boldsymbol{y}(\boldsymbol{a})$, to within measurement error. Using the instrument's measurement error covariance matrix, $\boldsymbol{S}_\epsilon$, the misfit of model and data is quantified by the value of the *misfit function*:

$$\Phi(\boldsymbol{a})=\tfrac{1}{2}(\hat{\boldsymbol{y}}-\boldsymbol{y}(\boldsymbol{a}))^T\cdot\boldsymbol{S}_\epsilon^{-1}\cdot(\hat{\boldsymbol{y}}-\boldsymbol{y}(\boldsymbol{a})). \tag{39}$$

To improve the fit we seek to adjust unknown parameters, $\boldsymbol{a}$, to decrease the value of the misfit function. The steepest decrease in Φ is obtained locally in the direction opposite to the gradient, $\nabla\Phi$. The nth element of this N-dimensional vector is given by the following formula:

$$-\frac{\partial\Phi(\boldsymbol{a})}{\partial a^n}=(\hat{\boldsymbol{y}}-\boldsymbol{y}(\boldsymbol{a}))^T\cdot\left(\boldsymbol{S}_\epsilon^{-1}\right)\cdot\frac{\partial\boldsymbol{y}(\boldsymbol{a})}{\partial a^n}. \tag{40}$$

By differentiating Eqs. (34) and (35) with respect to parameter a^n and collecting terms into integration kernels for the internal and outgoing data, we write Eq. (40) as a sum of inner products:

$$-\frac{\partial\Phi}{\partial a^n}=\left\langle\boldsymbol{\Delta p}_{\odot},\frac{\partial\boldsymbol{u}}{\partial a^n}\right\rangle_{D\times\mathbb{S}^2}+\left\langle\boldsymbol{\Delta q}_{\odot},\frac{\partial\boldsymbol{u}}{\partial a^n}\right\rangle_{\Gamma_+}. \tag{41}$$

The adjoint source vectors for differentiating the misfit function, $\boldsymbol{\Delta p}_{\odot}$ and $\boldsymbol{\Delta q}_{\odot}$, are obtained by summing over all detector response functions:

$$\boldsymbol{\Delta p}_{\odot}(\boldsymbol{x},\boldsymbol{v};\boldsymbol{a})\quad=\sum_{0\le m'<M}\sum_{0\le m<M_1}(\hat{y}^{m'}-y^{m'}(\boldsymbol{a}))(\boldsymbol{S}_\epsilon^{-1})_{m'm}\,\boldsymbol{p}^m_{\odot}(\boldsymbol{x},\boldsymbol{v}), \tag{42}$$

and

$$\Delta \boldsymbol{q}_{\odot}(\boldsymbol{x}, \boldsymbol{\nu}_+; \boldsymbol{a}) = \sum_{0 \le m' < MM_1} \sum_{\le m < M} (\hat{y}^{m'} - y^{m'}(\boldsymbol{a}))(\boldsymbol{S}_{\epsilon}^{-1})_{m'm} \boldsymbol{q}_{\odot}^{m}(\boldsymbol{x}, \boldsymbol{\nu}_+). \tag{43}$$

These adjoint source vectors, $\Delta \boldsymbol{p}_{\odot}$ and $\Delta \boldsymbol{q}_{\odot}$, may be visualized as collections of many "search lights" emanating from all measurement pixels at once, with the intensity of each search light equal to a weighted sum of the measurement residuals.

The next key step is to evaluate the derivative of the Stokes vector, $\partial \boldsymbol{u}/\partial a^n$. Although this could be accomplished numerically by finite-difference methods, a better way is to solve the 3D VRTE with modified volume-source and incoming-Stokes vectors. To see how this is possible, we differentiate Eqs. (16) and (17) with respect to parameter a^n and obtain the following 3D VRTE for $\partial \boldsymbol{u}/\partial a^n$:

$$\boldsymbol{\nu} \cdot \nabla \frac{\partial \boldsymbol{u}}{\partial a^n} + \sigma \frac{\partial \boldsymbol{u}}{\partial a^n} - \mathcal{Z}\left[\frac{\partial \boldsymbol{u}}{\partial a^n}\right] = \Delta \boldsymbol{f}_{\odot}^{n}, \tag{44}$$

$$\left.\frac{\partial \boldsymbol{u}}{\partial a^n}\right|_{\Gamma_-} - \mathcal{R}\left[\left.\frac{\partial \boldsymbol{u}}{\partial a^n}\right|_{\Gamma_+}\right] = \Delta \boldsymbol{g}_{\odot}^{n}, \tag{45}$$

with right-hand sides equal to forward source vectors, $\Delta \boldsymbol{f}_{\odot}^{n}$ and $\Delta \boldsymbol{g}_{\odot}^{n}$. The *internal* source vector, $\Delta \boldsymbol{f}_{\odot}^{n}(\boldsymbol{x}, \boldsymbol{\nu})$ for $(\boldsymbol{x}, \boldsymbol{\nu}) \in D \times \mathbb{S}^2$, accounts for the change in extinction and scattering:

$$\Delta \boldsymbol{f}_{\odot}^{n}(\boldsymbol{x}, \boldsymbol{\nu}; \boldsymbol{a}) = -\frac{\partial \sigma}{\partial a^n}(\boldsymbol{x}; \boldsymbol{a}) \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}; \boldsymbol{a}) + \frac{1}{4\pi} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{\nu}'} \frac{\partial \boldsymbol{Z}}{\partial a^n}(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}'; \boldsymbol{a}) \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}'; \boldsymbol{a}). \tag{46}$$

And, the *incoming* source vector, $\Delta \boldsymbol{g}_{\odot}^{n}(\boldsymbol{x}, \boldsymbol{\nu}_-)$ for $(\boldsymbol{x}, \boldsymbol{\nu}_-) \in \Gamma_-$, accounts for the change in reflection:

$$\Delta \boldsymbol{g}_{\odot}^{n}(\boldsymbol{x}, \boldsymbol{\nu}_-; \boldsymbol{a}) = \frac{1}{2\pi} \int_{\boldsymbol{\nu}_+ \cdot \nabla h(\boldsymbol{x}) > 0} \mathrm{d}S_{\boldsymbol{\nu}_+} |\boldsymbol{\nu}_+ \cdot \nabla h(\boldsymbol{x})| \frac{\partial \boldsymbol{R}}{\partial a^n}(\boldsymbol{x}, \boldsymbol{\nu}_-, \boldsymbol{\nu}_+; \boldsymbol{a}) \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}_+; \boldsymbol{a}). \tag{47}$$

Comparing with Definition 2, we see that the left-hand side of the 3D VRTE for the gradient of the Stokes vector solution, $\partial \boldsymbol{u}/\partial a^n$, is identical to the left-hand side of the 3D VRTE of the solution itself, $\boldsymbol{u}$. Therefore the same existence and uniqueness results are applicable. If the forward source vectors for computing parameter derivatives are square-integrable then the solution operator returns the derivative of the Stokes vector with respect to the parameter, a^n:

$$\left\{ \begin{array}{c} \left.\frac{\partial \boldsymbol{u}}{\partial a^n}\right|_{D \times \mathbb{S}^2} \\ \left.\frac{\partial \boldsymbol{u}}{\partial a^n}\right|_{\Gamma_+} \end{array} \right\} = \mathcal{U}_{\boldsymbol{a}} \left\{ \begin{array}{c} \Delta \boldsymbol{f}_{\odot}^{n} \\ \Delta \boldsymbol{g}_{\odot}^{n} \end{array} \right\}. \tag{48}$$

Substituting Eq. (48) into Eq. (41) we can express the equation for the gradient of the misfit function as

$$-\frac{\partial \Phi}{\partial a^n} = \left\langle \left\{ \begin{array}{c} \Delta \boldsymbol{p}_{\odot} \\ \Delta \boldsymbol{q}_{\odot} \end{array} \right\}, \mathcal{U}_{\boldsymbol{a}} \left\{ \begin{array}{c} \Delta \boldsymbol{f}_{\odot}^{n} \\ \Delta \boldsymbol{g}_{\odot}^{n} \end{array} \right\} \right\rangle_{D \times \mathbb{S}^2 \oplus \Gamma_+}, \tag{49}$$

for each a^n with $0 \le n < N$. As written in Eq. (49), computing all elements of the gradient requires N solutions to a 3D VRTE solver at each step of a multi-step iterative procedure.

However, the alternative rule for computing the gradient with the adjoint 3D VRTE is analytically equivalent to Eq. (49), but requires only one additional call to a 3D VRTE simulation, independent of how many parameters are used. The adjoint rule for differentiating the misfit function is as follows:

$$-\frac{\partial \Phi}{\partial a^n} = \left\langle (\mathcal{U}_{\boldsymbol{a}})^* \left\{ \begin{array}{c} \Delta \boldsymbol{p}_{\odot} \\ \Delta \boldsymbol{q}_{\odot} \end{array} \right\}, \left\{ \begin{array}{c} \Delta \boldsymbol{f}_{\odot}^{n} \\ \Delta \boldsymbol{g}_{\odot}^{n} \end{array} \right\} \right\rangle_{D \times \mathbb{S}^2 \oplus \Gamma_-}. \tag{50}$$

Using the fundamental adjoint property, $(\mathcal{U}_{\boldsymbol{a}})^* = \mathcal{U}_{\boldsymbol{a}}^*$, which is proven in Section 4.3 as Theorem 1, along with Eq. (32); the gradient can be written in terms of the adjoint Stokes vector, $\boldsymbol{w}$, solving the adjoint 3D VRTE with adjoint source vectors, $\Delta \boldsymbol{p}_{\odot}$ and $\Delta \boldsymbol{q}_{\odot}$:

$$-\frac{\partial \Phi}{\partial a^n} = \left\langle \mathcal{U}_{\boldsymbol{a}}^* \left\{ \begin{array}{c} \Delta \boldsymbol{p}_{\odot} \\ \Delta \boldsymbol{q}_{\odot} \end{array} \right\}, \left\{ \begin{array}{c} \Delta \boldsymbol{f}_{\odot}^{n} \\ \Delta \boldsymbol{g}_{\odot}^{n} \end{array} \right\} \right\rangle_{D \times \mathbb{S}^2 \oplus \Gamma_-}, \tag{51}$$

$$-\frac{\partial \Phi}{\partial a^n} = \langle \boldsymbol{w}, \Delta \boldsymbol{f}_{\odot}^{n} \rangle_{D \times \mathbb{S}^2} + \langle \boldsymbol{w}, \Delta \boldsymbol{g}_{\odot}^{n} \rangle_{\Gamma_-}. \tag{52}$$

The significance of Eqs. (51) and (52) is that the adjoint solution, $\boldsymbol{w}$, is independent of which a^n appears in the differentiation. Therefore any component of the gradient can be evaluated with the same adjoint solution, $\boldsymbol{w}$, and without further calls to a radiative transfer solver. The procedure involves the comparatively inexpensive operations of weighting the adjoint solution with the source terms from Eqs. (46) and (47) and integrating over the internal and incoming sets. The computation time of these integrations is (at worst) comparable to a single order-of-scattering computation, and could be much faster if one is careful to use a sparse basis for spatial and directional variability.

Using the adjoint method to compute the gradient of the misfit function shifts all the multiple-scattering computations to the residual distribution which depends only on the current atmospheric state and the misfit between observations and the polarized radiance that is generated by the current atmospheric state. The number of solutions to the 3D VRTE required in evaluating Eqs. (51) and (52) is $\mathcal{O}(L)$ and independent of the number of parameters, N. Therefore, rules for adjusting cloud, aerosol, and surface parameters based on the adjoint calculation of the misfit gradient are scalable to retrieval problems with many measurements and unknown parameters.

3.3. Iterative parameter adjustment

To discuss how the adjoint computations of the gradient of the misfit function can be incorporated into a scalable retrieval algorithm, we define a regularized misfit function, Φ_{reg}:

$$\Phi_{\text{reg}}(\boldsymbol{a}) = \Phi(\boldsymbol{a}) + \Phi_{\text{prior}}(\boldsymbol{a}). \tag{53}$$

Prior information is introduced by Φ_{prior}, to give a new maximum-likelihood estimation problem which is less sensitive to measurement noise and to provide a means of imposing additional measurement constraints, for example those from a coordinated LIDAR instrument [41,42]. The retrieval starts with an initial guess, $\boldsymbol{a}_0$, and makes additive adjustments, $\boldsymbol{b}_k$, so that the updated parameter, $\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \boldsymbol{b}_k$, converges to a minimizer of the regularized misfit function, Φ_{reg}.

A common starting place for many optimization methods is the second order Taylor approximation for Φ_{reg}, about any feasible state, $\boldsymbol{a}$:

$$\tilde{\Phi}_{\text{reg}}(\boldsymbol{a}+\boldsymbol{b}) = \Phi_{\text{reg}}(\boldsymbol{a}) + \nabla\Phi_{\text{reg}}(\boldsymbol{a}) \cdot \boldsymbol{b} + \tfrac{1}{2}\boldsymbol{b}^T \cdot \nabla\nabla\Phi_{\text{reg}}(\boldsymbol{a}) \cdot \boldsymbol{b}. \tag{54}$$

With sufficient prior information the Hessian matrix, $\nabla\nabla\Phi_{\text{reg}}(\boldsymbol{a})$, can be made positive definite. In this case the minimum value of the Taylor approximation can be found by solving Newton's equation for step, $\boldsymbol{b}$:

$$(\nabla\nabla\Phi(\boldsymbol{a}) + \nabla\nabla\Phi_{\text{prior}}(\boldsymbol{a})) \cdot \boldsymbol{b} = -(\nabla\Phi(\boldsymbol{a}) + \nabla\Phi_{\text{prior}}(\boldsymbol{a})). \tag{55}$$

A common scenario in setting up Newton's equations is that the Hessian of the prior function is easy to compute and the Hessian of the measurement misfit function is prohibitively expensive. Quasi-Newton methods substitute an approximate Hessian.

In atmospheric remote sensing applications, methods such as Levenberg–Marquardt use a linearized measurement model to approximate the Hessian in terms of the Jacobian matrix [33,39,41]. The Jacobian matrix,

$$\boldsymbol{J}(\boldsymbol{a}_k) = \frac{\partial \boldsymbol{y}(\boldsymbol{a}_k)}{\partial \boldsymbol{a}}, \tag{56}$$

is computed at every step, to approximate the Hessian as follows:

$$\nabla\nabla\Phi(\boldsymbol{a}) \approx \boldsymbol{J}(\boldsymbol{a})^T \cdot \boldsymbol{J}(\boldsymbol{a}). \tag{57}$$

The Jacobian matrix contains the derivatives of measurements with respect to unknown parameters and is, in general, dense with $\mathcal{O}(MN)$ elements. Using the fundamental adjoint property, the inner product for each element may be written in either of the two forms: one using the forward solver,

$$\frac{\partial y^m}{\partial a^n} = \left\langle \begin{Bmatrix} \boldsymbol{p}^m_\odot \\ \boldsymbol{q}^m_\odot \end{Bmatrix}, \mathcal{U}_{\boldsymbol{a}} \begin{Bmatrix} \Delta\boldsymbol{f}^n_\odot \\ \Delta\boldsymbol{g}^n_\odot \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2 \oplus \Gamma_+}, \tag{58}$$

and the other using the adjoint solver,

$$\frac{\partial y^m}{\partial a^n} = \left\langle \mathcal{U}^*_{\boldsymbol{a}} \begin{Bmatrix} \boldsymbol{p}^m_\odot \\ \boldsymbol{q}^m_\odot \end{Bmatrix}, \begin{Bmatrix} \Delta\boldsymbol{f}^n_\odot \\ \Delta\boldsymbol{g}^n_\odot \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2 \oplus \Gamma_-}. \tag{59}$$

In Eq. (58) the forward solution operator, $\mathcal{U}_{\boldsymbol{a}}$, must be evaluated for each different parameter and wavelength and this results in $\mathcal{O}(LN)$ computations. If the number of parameters is small or the computational solver provides Green's functions, then Eq. (58) can be quite efficient, see for example [19,22], but codes which currently solve the 3D VRTE do not compute Green's functions. In the alternative rule given by Eq. (59) the adjoint solution operator, $\mathcal{U}^*_{\boldsymbol{a}}$, must be evaluated for each measurement and this results in $\mathcal{O}(M)$ computations. However, it is worth mentioning that alternative approaches use the single-scattering approximation to formulate an approximate, sparse Jacobian matrix. This idea has shown promise for retrieving the volume-scattering coefficient, σ, using the scalar 3D radiative transfer equation with isotropic and weakly scattering media [43].

To handle the large amount of multiple scattering in clouds, we considered another quasi-Newton method which approximates the Hessian of the measurement misfit function using the gradient of the misfit function, $\nabla\Phi$, at previous iterations. These *gradient-based* methods can take advantage of the scalability of the adjoint rule in Eq. (50). The Hessian of the misfit function is approximated using previous parameter estimates, $\boldsymbol{a}_k$, and previous gradients, $\nabla\Phi(\boldsymbol{a}_k)$. The approximate Hessian, $\boldsymbol{H}_k$, is updated and improved upon at each step:

$$\nabla\nabla\Phi(\boldsymbol{a}_k) \approx \boldsymbol{H}_k(\boldsymbol{a}_0, \ldots, \boldsymbol{a}_k, \nabla\Phi(\boldsymbol{a}_0), \ldots, \nabla\Phi(\boldsymbol{a}_k)). \tag{60}$$

The rule usually guarantees symmetry and non-negative definiteness, and this is the case for the Broyden–Fletcher–Goldfarb-and-Shanno method, which is used in the medical imaging applications [13–15]. With a positive definite prior Hessian matrix $\nabla\nabla\Phi_{\text{prior}}$, the approximate Newton's equation uniquely defines a parameter adjustment, $\boldsymbol{b}_k$, with the following linear system:

$$(\boldsymbol{H}_k + \nabla\nabla\Phi_{\text{prior}}(\boldsymbol{a}_k)) \cdot \boldsymbol{b} = -(\nabla\Phi(\boldsymbol{a}_k) + \nabla\Phi_{\text{prior}}(\boldsymbol{a}_k)). \tag{61}$$

Provided that the problem is well scaled and that the approximate Hessian, $\boldsymbol{H}_k$, is chosen appropriately, the step will result in an improved set of cloud, aerosol and surface properties via the updated parameter $\boldsymbol{a}_{k+1} = \boldsymbol{a}_k + \boldsymbol{b}_k$. Moreover, by using the gradient, $\nabla\Phi$, to set up the local problem, the method can leverage the scalability of the adjoint computation to adjust atmosphere and surface properties with only $\mathcal{O}(L)$ calls to a 3D VRTE solver at each step.

3.4. Pseudo-forward problem

This section describes how to solve the adjoint 3D VRTE using a computer simulation for solving the forward 3D VRTE. To do this we must define two *actions*, α and $\boldsymbol{Q}$, which transform vectors according to the following rules. Action by α changes the sign of the direction argument:

$$\alpha\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\nu}) = \boldsymbol{f}(\boldsymbol{x}, -\boldsymbol{\nu}). \tag{62}$$

Action by $\boldsymbol{Q}$ flips the orientation of circular polarization:

$$\boldsymbol{Q}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\nu}) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \cdot \begin{bmatrix} f_I(\boldsymbol{x}, \boldsymbol{\nu}) \\ f_Q(\boldsymbol{x}, \boldsymbol{\nu}) \\ f_U(\boldsymbol{x}, \boldsymbol{\nu}) \\ f_V(\boldsymbol{x}, \boldsymbol{\nu}) \end{bmatrix}. \tag{63}$$

These actions are their own inverses, $\boldsymbol{Q}^2 = \alpha^2 = \text{Identity}$.

For scattering media which obeys the principles of mirror-symmetry and reciprocity, the kernel for the single-scattering operator satisfies the rule,

$$\boldsymbol{Z}^T(\boldsymbol{x}, \boldsymbol{\nu}', \boldsymbol{\nu}) = \boldsymbol{Q}\boldsymbol{Z}(\boldsymbol{x}, -\boldsymbol{\nu}, -\boldsymbol{\nu}')\boldsymbol{Q}, \tag{64}$$

and the kernel for the reflection operator satisfies the rule,

$$\boldsymbol{R}^T(\boldsymbol{x}, \boldsymbol{\nu}', \boldsymbol{\nu}) = \boldsymbol{Q}\boldsymbol{R}(\boldsymbol{x}, -\boldsymbol{\nu}, -\boldsymbol{\nu}')\boldsymbol{Q}. \tag{65}$$

The adjoint scattering operator can be written in terms of the forward scattering operator,

$$\mathcal{Z}^*[\boldsymbol{w}] = \alpha\boldsymbol{Q}\mathcal{Z}[\alpha\boldsymbol{Q}\boldsymbol{w}], \tag{66}$$

and the adjoint reflection operator can be written in terms of the forward reflection operator,

$$\mathcal{R}^*[\boldsymbol{w}] = \alpha\boldsymbol{Q}\mathcal{R}[\alpha\boldsymbol{Q}\boldsymbol{w}]. \tag{67}$$

By plugging Eqs. (66) and (67) into the adjoint 3D VRTE defined by Eqs. (26) and (27) and acting on both sides with $\alpha\mathbf{Q}$, it is easy to verify that the transformed adjoint Stokes vector, $\alpha\mathbf{Q}\boldsymbol{w}$, solves the forward 3D VRTE with pseudo-forward source vectors, $\alpha\mathbf{Q}\boldsymbol{p}_{\odot}$ and $\alpha\mathbf{Q}\boldsymbol{q}_{\odot}$. Therefore, the adjoint solution operator can be evaluated using the forward solution operator as follows:

$$\mathcal{U}^*_{\boldsymbol{a}}\begin{Bmatrix}\boldsymbol{p}_{\odot}\\ \boldsymbol{q}_{\odot}\end{Bmatrix}=\alpha\,\mathbf{Q}\mathcal{U}_{\boldsymbol{a}}\begin{Bmatrix}\alpha\mathbf{Q}\boldsymbol{p}_{\odot}\\ \alpha\mathbf{Q}\boldsymbol{q}_{\odot}\end{Bmatrix}. \tag{68}$$

This means that a computer code that solves the forward problem can be used to solve the adjoint problem. Provided that it is sufficiently general to accept the transformed source vectors, $\alpha\mathbf{Q}\boldsymbol{p}_{\odot}$ and $\alpha\mathbf{Q}\boldsymbol{q}_{\odot}$, it will output as a solution the transformed adjoint Stokes vector, $\alpha\mathbf{Q}\boldsymbol{w}$. The only additional difficulties in solving for the adjoint solution, $\boldsymbol{w}$, arise in preparing the right-hand side and interpreting the solution.

4. The fundamental adjoint property

This section defines mathematical tools for proving the fundamental adjoint property with the proof given in Section 4.3. The first objective is to define a family of streaming operators. These will enable the formulation of integral equations which are equivalent to the integro-differential equations of 3D VRT, as given by Definitions 2 and 3 in Section 2. The integral equations are presented in Sections 4.2.1 and 4.2.2, along with series expansions for the solution operators, $\mathcal{U}_{\boldsymbol{a}}$ and $\mathcal{U}^*_{\boldsymbol{a}}$. Lastly, in Section 4.3 we state and prove the fundamental adjoint property as Theorem 1. This theorem shows that the adjoint 3D VRTE given in Definition 3 is well defined, that $(\mathcal{U}_{\boldsymbol{a}})^*=\mathcal{U}^*_{\boldsymbol{a}}$, and justifies the use of this property in deriving a scalable procedure for adjusting 3D atmosphere and surface parameters.

4.1. *The streaming operators*

Streaming refers to propagation of radiative information along special line segments called chords.[1] *Chords* are defined to be the open-ended line segments in D whose endpoints lie on the boundary, ∂D. The streaming operators propagate source vectors along these chords: forward streaming operators propagate sources in the positive direction, $\boldsymbol{v}$, and adjoint streaming operators propagate sources in the negative direction, $-\boldsymbol{v}$. In contrast to previous derivations of the ancillary integral equation for 3D radiative transfer, for example that by Davis and Knyazikhin in Chapter 3 of [8], we split the streaming process into four distinct linear operations. This splitting facilitates the treatment of general boundary conditions in both the forward and adjoint 3D VRTE. Moreover, the streaming operators for the adjoint 3D VRTE are, actually, the adjoint operators of the forward streaming operators. A brief summary of the splitting of streaming operators will suffice for readers wishing to move ahead to Section 4.2 and the definition of the integral equations for 3D VRTE.

Each streaming operator acts on, and returns, a function which is defined on one of the three subdomains of vector radiative transfer. For instance, the *internal* streaming operator, $\mathcal{T}_{00}$, acts on an internal source vector, $\boldsymbol{f}$, defined on the internal set, $D\times\mathbb{S}^2$, and it returns a Stokes vector restricted to the internal set, $\boldsymbol{u}|_{D\times\mathbb{S}^2}$. Boundary conditions are managed by the other operators. The *incoming-to-internal* streaming operator, $\mathcal{T}_{-0}$, acts on an incoming source vector, $\boldsymbol{g}$, defined on the incoming set, Γ_-, and it returns a Stokes vector on the internal set, $\boldsymbol{u}|_{D\times\mathbb{S}^2}$. The other two forward streaming operators are named according to their behavior in a similar way. The *internal-to-outgoing* streaming operator, $\mathcal{T}_{0+}$, acts on an internal source vector, $\boldsymbol{f}$, and returns an outgoing Stokes vector, $\boldsymbol{u}|_{\Gamma_+}$; and the *incoming-to-outgoing* streaming operator, $\mathcal{T}_{-+}$, acts on an incoming source vector, $\boldsymbol{g}$, and returns an outgoing Stokes vector, $\boldsymbol{u}|_{\Gamma_+}$.

The adjoints of these streaming operators act on adjoint source vectors and return adjoint Stokes vectors. For instance, the adjoint of the *internal* streaming operator, $\mathcal{T}^*_{00}$, acts on an internal adjoint source vector, $\boldsymbol{p}$, defined on the internal set, $D\times\mathbb{S}^2$, and it returns an adjoint Stokes vector restricted to the internal set, $\boldsymbol{w}|_{D\times\mathbb{S}^2}$. However, the domains of input and output functions are reversed: the adjoint of the *incoming-to-internal* streaming operator, $\mathcal{T}^*_{-0}$, acts on an internal-adjoint source vector, $\boldsymbol{p}$, and returns an incoming-adjoint Stokes vector, $\boldsymbol{w}|_{\Gamma_-}$; the adjoint of the *internal-to-outgoing* streaming operator, $\mathcal{T}^*_{0+}$, acts on an outgoing-adjoint source vector, $\boldsymbol{q}$, and returns an internal-adjoint Stokes vector, $\boldsymbol{w}|_{D\times\mathbb{S}^2}$; and the adjoint of the *incoming-to-outgoing* streaming operator, $\mathcal{T}^*_{-+}$, acts on an outgoing adjoint source vector, $\boldsymbol{q}$, and returns an incoming adjoint Stokes vector, $\boldsymbol{w}|_{\Gamma_-}$.

Adjoint streaming operators are defined to satisfy the following adjoint properties:

$$\langle\boldsymbol{p},\mathcal{T}_{00}[\boldsymbol{f}]\rangle_{D\times\mathbb{S}^2}=\langle\mathcal{T}^*_{00}[\boldsymbol{p}],\boldsymbol{f}\rangle_{D\times\mathbb{S}^2}, \tag{69}$$

$$\langle\boldsymbol{q},\mathcal{T}_{0+}[\boldsymbol{f}]\rangle_{\Gamma_+}=\langle\mathcal{T}^*_{0+}[\boldsymbol{q}],\boldsymbol{f}\rangle_{D\times\mathbb{S}^2}, \tag{70}$$

$$\langle\boldsymbol{p},\mathcal{T}_{-0}[\boldsymbol{g}]\rangle_{D\times\mathbb{S}^2}=\langle\mathcal{T}^*_{-0}[\boldsymbol{p}],\boldsymbol{g}\rangle_{\Gamma_-}, \tag{71}$$

$$\langle\boldsymbol{q},\mathcal{T}_{-+}[\boldsymbol{g}]\rangle_{\Gamma_+}=\langle\mathcal{T}^*_{-+}[\boldsymbol{q}],\boldsymbol{g}\rangle_{\Gamma_-}. \tag{72}$$

The remainder of Section 4.1 is devoted to parametrizing chords for the purpose of defining explicit rules for evaluating each of the forward and adjoint streaming operators. Forward streaming operators are defined in Eqs. (89)–(92), and their adjoints in Eqs. (93)–(96). In Appendix B we prove the adjoint properties of the streaming operators.

4.1.1. *Chords and boundary points*

Because streaming operators propagate information along chords, they are most easily defined with a chord parametrization. The signed distance-to-boundary function, $h(\boldsymbol{x})$, is quite useful for this purpose and enables treatment of non-convex atmospheric regions and surface topography. Using this function we define the unique chord for every internal, incoming, and outgoing point.

For internal points, $(\boldsymbol{x},\boldsymbol{v})\in D\times\mathbb{S}^2$, we define the chord parameters as follows:

$$t=\boldsymbol{x}\cdot\boldsymbol{v}, \tag{73}$$

[1] Also called characteristics.

$$\boldsymbol{x}^{\perp} = \boldsymbol{x} - t\boldsymbol{v}. \tag{74}$$

Extreme values of chord parameter, t, are found by looking along the directions, $\boldsymbol{v}$ and $-\boldsymbol{v}$, to the nearest boundary points:

$$t_{-} = \max\{t_{-} \in \mathbb{R}: t_{-} < t \text{ and } h(\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}) = 0\}, \tag{75}$$

$$t_{+} = \min\{t_{+} \in \mathbb{R}: t_{+} > t \text{ and } h(\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}) = 0\}. \tag{76}$$

For outgoing points, $(\boldsymbol{x}, \boldsymbol{v}_{+}) \in \Gamma_{+}$, the chord parameters are defined as follows:

$$t_{+} = \boldsymbol{x} \cdot \boldsymbol{v}_{+}, \tag{77}$$

$$\boldsymbol{x}^{\perp} = \boldsymbol{x} - t_{+}\boldsymbol{v}_{+}. \tag{78}$$

The opposite extreme is found by looking along, $-\boldsymbol{v}_{+}$, to the nearest boundary point:

$$t_{-} = \max\{t_{-} \in \mathbb{R}: t_{-} < t_{+} \text{ and } h(\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}_{+}) = 0\}. \tag{79}$$

For incoming points, $(\boldsymbol{x}, \boldsymbol{v}_{-}) \in \Gamma_{-}$, the chord parameters are defined as follows:

$$t_{-} = \boldsymbol{x} \cdot \boldsymbol{v}_{-}, \tag{80}$$

$$\boldsymbol{x}^{\perp} = \boldsymbol{x} - t_{-}\boldsymbol{v}_{-}. \tag{81}$$

The opposite extreme is found by looking along, $\boldsymbol{v}_{-}$, to the nearest boundary point:

$$t_{+} = \min\{t_{+} \in \mathbb{R}: t_{+} > t_{-} \text{ and } h(\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}_{-}) = 0\}. \tag{82}$$

In each case, the position-direction pair can be associated with the unique non-empty chord,

$$\boldsymbol{x}'(t'; \boldsymbol{x}^{\perp}, \boldsymbol{v}) = \boldsymbol{x}^{\perp} + t'\boldsymbol{v} \quad \text{for } t_{-} < t' < t_{+}, \tag{83}$$

through the interior, with $h(\boldsymbol{x}'(t')) < 0$. The endpoints of the chord correspond to parameters, t_{-} and t_{+},

$$\boldsymbol{x}_{-} = \boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v} \in \partial D, \tag{84}$$

$$\boldsymbol{x}_{+} = \boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v} \in \partial D, \tag{85}$$

and are on the boundary, since $h(\boldsymbol{x}_{-}) = h(\boldsymbol{x}_{+}) = 0$. This defines the chord representation for all points in the domain of 3D VRTE and provides a useful alternative representation of internal, outgoing, and incoming points:

$$(\boldsymbol{x}, \boldsymbol{v}) = (\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v}), \tag{86}$$

$$(\boldsymbol{x}, \boldsymbol{v}_{+}) = (\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}_{+}, \boldsymbol{v}_{+}), \tag{87}$$

$$(\boldsymbol{x}, \boldsymbol{v}_{-}) = (\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}_{-}, \boldsymbol{v}_{-}). \tag{88}$$

Looking toward future work implementing numerical methods, it is worth noting that this procedure can find chords through complex geometries by determining the zeros of a real-valued, single-variable function: $h(\boldsymbol{x}'(\cdot))$.

There is an important caveat. Although the chord is uniquely defined for each case, the endpoints, $\boldsymbol{x}_{-}$ and $\boldsymbol{x}_{+}$, do not always correspond to elements in the incoming or outgoing sets. The reason for this is that some chords will be tangent to the boundary at one or both end points. This poses a challenge to defining boundary-streaming operators, because there is not necessarily an incoming or outgoing point that corresponds to a location at which we desire to know the value of the streaming operator. However, the set of such points related to boundary-tangent chords will have measure zero in the integrals of interest and can therefore be neglected. We now focus on defining the streaming operators, noting that certain points that correspond to chords that are tangent to the boundary may require alternate definitions.

4.1.2. Rules for evaluating streaming operators

Streaming operators are defined by changing the argument of evaluation from standard representation to chord representation, e.g. from $(\boldsymbol{x}, \boldsymbol{v})$ to $(\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v})$. This is done to isolate the direction, $\boldsymbol{v}$, along which source vectors are integrated. The four forward streaming operators are defined as follows:

$$\begin{aligned}\mathcal{T}_{00}[\boldsymbol{f}](\boldsymbol{x}, \boldsymbol{v}) &= \mathcal{T}_{00}[\boldsymbol{f}](\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v}),\\ &= \int_{t_{-}}^{t} \mathrm{d}t' \left[\exp\left(-\int_{t'}^{t} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{f}(\boldsymbol{x}^{\perp} + t'\boldsymbol{v}, \boldsymbol{v})\right],\end{aligned} \tag{89}$$

$$\begin{aligned}\mathcal{T}_{0+}[\boldsymbol{f}](\boldsymbol{x}_{+}, \boldsymbol{v}_{+}) &= \mathcal{T}_{0+}[\boldsymbol{f}](\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}, \boldsymbol{v}),\\ &= \int_{t_{-}}^{t_{+}} \mathrm{d}t' \left[\exp\left(-\int_{t'}^{t_{+}} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{f}(\boldsymbol{x}^{\perp} + t'\boldsymbol{v}, \boldsymbol{v})\right],\end{aligned} \tag{90}$$

$$\begin{aligned}\mathcal{T}_{-0}[\boldsymbol{g}](\boldsymbol{x}, \boldsymbol{v}) &= \mathcal{T}_{-0}[\boldsymbol{g}](\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v}),\\ &= \exp\left(-\int_{t_{-}}^{t} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{g}(\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}, \boldsymbol{v}),\end{aligned} \tag{91}$$

$$\begin{aligned}\mathcal{T}_{-+}[\boldsymbol{g}](\boldsymbol{x}_{+}, \boldsymbol{v}_{+}) &= \mathcal{T}_{-+}[\boldsymbol{g}](\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}, \boldsymbol{v}),\\ &= \exp\left(-\int_{t_{-}}^{t_{+}} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{g}(\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}, \boldsymbol{v}).\end{aligned} \tag{92}$$

The adjoints of these operators are given by the rules:

$$\begin{aligned}\mathcal{T}^{*}_{00}[\boldsymbol{p}](\boldsymbol{x}, \boldsymbol{v}) &= \mathcal{T}^{*}_{00}[\boldsymbol{p}](\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v}),\\ &= \int_{t}^{t_{+}} \mathrm{d}t' \left[\exp\left(-\int_{t}^{t'} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{p}(\boldsymbol{x}^{\perp} + t'\boldsymbol{v}, \boldsymbol{v})\right],\end{aligned} \tag{93}$$

$$\begin{aligned}\mathcal{T}^{*}_{0+}[\boldsymbol{q}](\boldsymbol{x}, \boldsymbol{v}) &= \mathcal{T}^{*}_{0+}[\boldsymbol{q}](\boldsymbol{x}^{\perp} + t\boldsymbol{v}, \boldsymbol{v}),\\ &= \exp\left(-\int_{t}^{t_{+}} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{q}(\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}, \boldsymbol{v}),\end{aligned} \tag{94}$$

$$\begin{aligned}\mathcal{T}^{*}_{-0}[\boldsymbol{p}](\boldsymbol{x}_{-}, \boldsymbol{v}_{-}) &= \mathcal{T}^{*}_{-0}[\boldsymbol{p}](\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}, \boldsymbol{v}),\\ &= \int_{t_{-}}^{t_{+}} \mathrm{d}t' \left[\exp\left(-\int_{t_{-}}^{t'} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{p}(\boldsymbol{x}^{\perp} + t'\boldsymbol{v}, \boldsymbol{v})\right],\end{aligned} \tag{95}$$

$$\begin{aligned}\mathcal{T}^{*}_{-+}[\boldsymbol{q}](\boldsymbol{x}_{-}, \boldsymbol{v}_{-}) &= \mathcal{T}^{*}_{-+}[\boldsymbol{q}](\boldsymbol{x}^{\perp} + t_{-}\boldsymbol{v}, \boldsymbol{v}),\\ &= \exp\left(-\int_{t_{-}}^{t_{+}} \mathrm{d}t''\, \sigma(\boldsymbol{x}^{\perp} + t''\boldsymbol{v})\right)\boldsymbol{q}(\boldsymbol{x}^{\perp} + t_{+}\boldsymbol{v}, \boldsymbol{v}).\end{aligned} \tag{96}$$

The streaming operators in Eqs. (89), (90), (93), and (95) act on internal source vectors, $\boldsymbol{f}$ or $\boldsymbol{p}$, and require integration over some or all of the chords associated with the point of evaluation. Alternatively, the streaming operators in Eqs. (91), (92), (94), and (96) act on incoming forward source vectors, $\boldsymbol{g}$, or outgoing adjoint source vectors, $\boldsymbol{q}$. These involve only scaling by an attenuation factor. Note that while the operator $\mathcal{T}_{0+}$ in Eq. (90) integrates over a chord from t_{-} to t_{+}, the corresponding adjoint-streaming operator $\mathcal{T}^{*}_{0+}$ in

Eq. (94) has no such integral. The integral over the chord is subsumed in the inner product in Eq. (70).

4.1.3. Properties of streaming operators

Streaming operators are defined to help solve the 3D VRTE and its adjoint by transformation to an equivalent system of integral equations. Under the action of the advective derivative, the forward streaming operators behave as follows:

$$(\boldsymbol{v}\cdot\nabla+\sigma)[\mathcal{T}_{00}[\boldsymbol{f}]]=\boldsymbol{f},\tag{97}$$

$$(\boldsymbol{v}\cdot\nabla+\sigma)[\mathcal{T}_{-0}[\boldsymbol{g}]]=0.\tag{98}$$

Similar properties hold for the adjoint streaming operators:

$$(-\boldsymbol{v}\cdot\nabla+\sigma)[\mathcal{T}^*_{00}[\boldsymbol{p}]]=\boldsymbol{p},\tag{99}$$

$$(-\boldsymbol{v}\cdot\nabla+\sigma)[\mathcal{T}^*_{0+}[\boldsymbol{q}]]=0.\tag{100}$$

These are shown in Appendix A as Theorem 2, and they provide the connection between integral and integro-differential forms of the 3D VRTE.

4.2. Integral equations for 3D VRTE

The integral operators for scattering and reflection, $\mathcal{Z}$ and $\mathcal{R}$, act on Stokes vectors and return source vectors. The integral operators for streaming, $\mathcal{T}_{00}$, $\mathcal{T}_{0+}$, $\mathcal{T}_{-0}$, and $\mathcal{T}_{-+}$, act on source vectors and return Stokes vectors. In an approximate-numerical setting the Stokes vectors and source vectors will be represented by a finite-dimensional vector of basis function coefficients. Furthermore, the integral operators will be approximated by linear-matrix transformations acting on these coefficient vectors. With the discrete analogue of matrix algebra in mind, we use a matrix operator notation to keep track of integral operations. As with the solution operators, $\mathcal{U}_a$ and $\mathcal{U}^*_a$, curly brackets are used, $\{::\}$ and $\{:\}$. The objective is to organize the linear integral operations according to the familiar notation of matrix–vector products from linear algebra:

$$\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}=\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00}[\boldsymbol{f}]+\mathcal{Z}\mathcal{T}_{-0}[\boldsymbol{g}]\\ \mathcal{R}\mathcal{T}_{0+}[\boldsymbol{f}]+\mathcal{R}\mathcal{T}_{-+}[\boldsymbol{g}]\end{Bmatrix}.\tag{101}$$

Normal array operations (such as associativity) behave as expected. For instance, the combined operations of streaming and then scattering can be written in either of the following two ways:

$$\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}=\begin{Bmatrix}\mathcal{Z} & \\ & \mathcal{R}\end{Bmatrix}\begin{Bmatrix}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{Bmatrix}.\tag{102}$$

Empty spaces are assumed to represent a null operator. The analogy with matrices extends to the calculation of adjoints:

$$\begin{Bmatrix}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{Bmatrix}^*=\begin{Bmatrix}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{Bmatrix},\tag{103}$$

where the adjoint of the 2×2 matrix-operator is defined using the joint inner products from Definition 1. This matrix-operator presentation is preferred over the use of indices which are less easily readable.

4.2.1. Forward integral equations

We now present the integral formulation of the 3D VRTE and the so-called successive order of scattering series expansion for its solution. As described in the context of scalar radiative transfer by [25], the Stokes vector, $\boldsymbol{u}$, can be written in terms of the solution vectors, $\boldsymbol{f}$ and $\boldsymbol{g}$, of the forward integral equations:

$$\boldsymbol{u}|_{D\times\mathbb{S}^2}=\mathcal{T}_{00}[\boldsymbol{f}]+\mathcal{T}_{-0}[\boldsymbol{g}],\tag{104}$$

$$\boldsymbol{u}|_{\Gamma_+}=\mathcal{T}_{0+}[\boldsymbol{f}]+\mathcal{T}_{-+}[\boldsymbol{g}],\tag{105}$$

$$\boldsymbol{u}|_{\Gamma_-}=\boldsymbol{g}.\tag{106}$$

The vectors, $\boldsymbol{f}$ and $\boldsymbol{g}$, are called solutions to the forward integral equations because they solve the integral formulation of the forward 3D VRTE:

$$\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}-\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}=\begin{Bmatrix}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{Bmatrix}.\tag{107}$$

This differs from the ancillary integral equations (for diffuse radiation) as they are written in Chapter 3 of [8], in that Eq. (107) treats the internal source vector, $\boldsymbol{f}_\odot$, and incoming source vector, $\boldsymbol{g}_\odot$, as separate entities. This allows us to include direct radiation in the solution, causing the source vectors, $\boldsymbol{f}_\odot$ and $\boldsymbol{g}_\odot$, to be identical to those of Definition 2. In Appendix A, Theorem 3, we show that the set of Eqs. (104)–(107) provide a solution to the 3D VRTE that satisfies the integro-differential formulation in Definition 2.

Eq. (107) is in standard form for a Fredholm integral equation of the second kind. To derive the series expansion for the solution operator, we use a fixed point iteration to obtain the following expression:

$$\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}=\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}^{K+1}\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}+\sum_{k=0}^{K}\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}^{k}\begin{Bmatrix}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{Bmatrix},\tag{108}$$

noting that powers, $\{::\}^k$, indicate repeated application of the integral operator. The solve-ability condition on the 3D VRTE must provide that the first term on the right-hand side of Eq. (108) will decay to zero. This occurs when σ, $\boldsymbol{Z}$, and $\boldsymbol{R}$ are such that the combined operations of streaming and scattering/reflection give an operator with norm less than unity:

$$\left\|\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}\right\|_{\text{op}}<1.\tag{109}$$

If this condition is satisfied, we let $K\to\infty$ in Eq. (108) to obtain the successive order of scattering series solution:

$$\begin{Bmatrix}\boldsymbol{f}\\ \boldsymbol{g}\end{Bmatrix}=\sum_{k=0}^{\infty}\begin{Bmatrix}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{Bmatrix}^{k}\begin{Bmatrix}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{Bmatrix}.\tag{110}$$

One may provide a more formal justification, as in [44], using the completeness of the space of square-integrable functions, as defined in Section 2.

The series in Eq. (110) converges as a square-integrable function and provides the solution vectors, $\boldsymbol{f}$ and $\boldsymbol{g}$, of the forward integral equations and by streaming them according to Eqs. (104) and (105), they provide the Stokes vector solution to the integro-differential form of 3D VRTE:

$$\left\{\begin{array}{l}\boldsymbol{u}|_{D\times\mathbb{S}^2}\\ \boldsymbol{u}|_{\Gamma_+}\end{array}\right\}=\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}. \tag{111}$$

We can thus define the solution operator for the forward 3D VRTE, introduced in Definition 2:

$$\mathcal{U}_{\boldsymbol{a}}\left\{\begin{array}{l}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{array}\right\}=\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\times\sum_{k=0}^{\infty}\left\{\begin{array}{ll}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{array}\right\}^k\left\{\begin{array}{l}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{array}\right\}. \tag{112}$$

4.2.2. Adjoint integral equations

The adjoint 3D VRTE has a completely analogous integral formulation to that of the forward model. Using adjoint streaming operations, we write the adjoint Stokes vector, $\boldsymbol{w}$, in terms of the solution vectors, $\boldsymbol{p}$ and $\boldsymbol{q}$, of the adjoint integral equations:

$$\boldsymbol{w}|_{D\times\mathbb{S}^2}=\mathcal{T}^*_{00}[\boldsymbol{p}]+\mathcal{T}^*_{0+}[\boldsymbol{q}], \tag{113}$$

$$\boldsymbol{w}|_{\Gamma_-}=\mathcal{T}^*_{-0}[\boldsymbol{p}]+\mathcal{T}^*_{-+}[\boldsymbol{q}], \tag{114}$$

$$\boldsymbol{w}|_{\Gamma_+}=\boldsymbol{q}. \tag{115}$$

The vectors, $\boldsymbol{p}$ and $\boldsymbol{q}$, are called solutions to the adjoint integral equations because they solve the integral formulation of the adjoint 3D VRTE:

$$\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\}-\left\{\begin{array}{ll}\mathcal{Z}^*\mathcal{T}^*_{00} & \mathcal{Z}^*\mathcal{T}^*_{0+}\\ \mathcal{R}^*\mathcal{T}^*_{-0} & \mathcal{R}^*\mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\}=\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\}. \tag{116}$$

The proof that the adjoint Stokes vector, $\boldsymbol{w}$, given by these equations satisfies the adjoint 3D VRTE is given in Appendix A as Theorem 4.

The vectors, $\boldsymbol{p}$ and $\boldsymbol{q}$, can be expressed as the infinite successive-order of scattering series:

$$\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\}=\sum_{k=0}^{\infty}\left\{\begin{array}{ll}\mathcal{Z}^*\mathcal{T}^*_{00} & \mathcal{Z}^*\mathcal{T}^*_{0+}\\ \mathcal{R}^*\mathcal{T}^*_{-0} & \mathcal{R}^*\mathcal{T}^*_{-+}\end{array}\right\}^k\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\}. \tag{117}$$

This expression provides the solution vectors, $\boldsymbol{p}$ and $\boldsymbol{q}$, of the adjoint integral equations and by streaming them according to Eqs. (113) and (114), they provide the Stokes vector solution to the integro-differential form of the adjoint 3D VRTE:

$$\left\{\begin{array}{l}\boldsymbol{w}|_{D\times\mathbb{S}^2}\\ \boldsymbol{w}|_{\Gamma_-}\end{array}\right\}=\left\{\begin{array}{ll}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\}. \tag{118}$$

The solution operator for the adjoint 3D VRTE, introduced in Definition 3, can therefore be written as follows:

$$\mathcal{U}^*_{\boldsymbol{a}}\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\}=\left\{\begin{array}{ll}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{array}\right\}\times\sum_{k=0}^{\infty}\left\{\begin{array}{ll}\mathcal{Z}^*\mathcal{T}^*_{00} & \mathcal{Z}^*\mathcal{T}^*_{0+}\\ \mathcal{R}^*\mathcal{T}^*_{-0} & \mathcal{R}^*\mathcal{T}^*_{-+}\end{array}\right\}^k\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\}. \tag{119}$$

4.3. Fundamental adjoint property for the 3D VRTE

For the solution operator of the adjoint problem in Definition 3, we wrote, "$\mathcal{U}^*_{\boldsymbol{a}}$," anticipating that it would be the adjoint of the solution operator for the forward VRTE, which we denoted by "$(\mathcal{U}_{\boldsymbol{a}})^*$". The notation hints that the equation, $(\mathcal{U}_{\boldsymbol{a}})^*=\mathcal{U}^*_{\boldsymbol{a}}$, holds; that the adjoint of the solution operator for the forward 3D VRTE is the solution operator for the adjoint 3D VRTE. In fact, this relation justifies the name *adjoint* 3D VRTE, and must be proven through verification of the fundamental adjoint property:

$$\left\langle\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\},\mathcal{U}_{\boldsymbol{a}}\left\{\begin{array}{l}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}=\left\langle\mathcal{U}^*_{\boldsymbol{a}}\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\},\left\{\begin{array}{l}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-}. \tag{120}$$

An equivalent form of this equation is written in terms of the solutions to the integral equations of 3D VRT, and it is this statement that we prove.

Theorem 1 (*Fundamental adjoint property*). *For any feasible parameter, $\boldsymbol{a}$, and the corresponding single-scattering properties, σ, $\boldsymbol{Z}$, and $\boldsymbol{R}$, satisfying the solve-ability constraint in Eq. (109); the following fundamental adjoint property holds*:

$$\left\langle\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\},\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}=\left\langle\left\{\begin{array}{ll}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\},\left\{\begin{array}{l}\boldsymbol{f}_\odot\\ \boldsymbol{g}_\odot\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-}, \tag{121}$$

where the vectors, $\boldsymbol{f}$ and $\boldsymbol{g}$, solve the forward integral equations with square-integrable source vector, $\boldsymbol{f}_\odot$ and $\boldsymbol{g}_\odot$; and the vectors, $\boldsymbol{p}$ and $\boldsymbol{q}$, solve the adjoint integral equations with square-integrable adjoint source vectors, $\boldsymbol{p}_\odot$ and $\boldsymbol{q}_\odot$.

Proof. We begin by substituting the adjoint integral equation for $\boldsymbol{p}_\odot$ and $\boldsymbol{q}_\odot$, given by Eq. (116), into the left hand side of Eq. (121) and proceed through the following steps:

$$\begin{aligned}&\left\langle\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\},\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}\\ &=\left\langle\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\},\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}\\ &\quad-\left\langle\left\{\begin{array}{ll}\mathcal{Z}^*\mathcal{T}^*_{00} & \mathcal{Z}^*\mathcal{T}^*_{0+}\\ \mathcal{R}^*\mathcal{T}^*_{-0} & \mathcal{R}^*\mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\},\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+};\end{aligned} \tag{122}$$

$$\begin{aligned}&\left\langle\left\{\begin{array}{l}\boldsymbol{p}_\odot\\ \boldsymbol{q}_\odot\end{array}\right\},\left\{\begin{array}{ll}\mathcal{T}_{00} & \mathcal{T}_{-0}\\ \mathcal{T}_{0+} & \mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}\\ &=\left\langle\left\{\begin{array}{ll}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\},\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-}\\ &\quad-\left\langle\left\{\begin{array}{ll}\mathcal{T}^*_{00} & \mathcal{T}^*_{0+}\\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{p}\\ \boldsymbol{q}\end{array}\right\},\left\{\begin{array}{ll}\mathcal{Z}\mathcal{T}_{00} & \mathcal{Z}\mathcal{T}_{-0}\\ \mathcal{R}\mathcal{T}_{0+} & \mathcal{R}\mathcal{T}_{-+}\end{array}\right\}\left\{\begin{array}{l}\boldsymbol{f}\\ \boldsymbol{g}\end{array}\right\}\right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-};\end{aligned} \tag{123}$$

$$\left\langle \begin{Bmatrix} \boldsymbol{p}_\odot \\ \boldsymbol{q}_\odot \end{Bmatrix}, \begin{Bmatrix} \mathcal{T}_{00} & \mathcal{T}_{-0} \\ \mathcal{T}_{0+} & \mathcal{T}_{-+} \end{Bmatrix} \begin{Bmatrix} \boldsymbol{f} \\ \boldsymbol{g} \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_+}$$
$$= \left\langle \begin{Bmatrix} \mathcal{T}^*_{00} & \mathcal{T}^*_{0+} \\ \mathcal{T}^*_{-0} & \mathcal{T}^*_{-+} \end{Bmatrix} \begin{Bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{Bmatrix}, \begin{Bmatrix} \boldsymbol{f}_\odot \\ \boldsymbol{g}_\odot \end{Bmatrix} \right\rangle_{D\times\mathbb{S}^2\oplus\Gamma_-}. \qquad (124)$$

These three steps are justified as follows: Eq. (122) is obtained by substitution of the integral equations of adjoint 3D VRT; Eq. (123) is obtained by using elementary adjoint properties of streaming, scattering and reflection operators; and Eq. (124) is obtained by substitution of the integral equations of forward 3D VRT. The second step can be verified by expanding to a sum of elementary inner products (that is $\langle\cdot,\cdot\rangle_{D\times\mathbb{S}^2}$, $\langle\cdot,\cdot\rangle_{\Gamma_+}$, and $\langle\cdot,\cdot\rangle_{\Gamma_-}$), applying elementary adjoint properties, and collapsing back into the matrix notation. □

5. Conclusion

Adjoint methods can enable the use of 3D VRTE simulations for adjusting 3D atmospheric properties to fit multi-angle, multi-pixel polarimetric measurements of the Earth's atmosphere. This is shown by focusing on computing the misfit function and its gradient, and doing so with only two calls to a 3D VRTE solver for each wavelength. Scalable methods such as the adjoint method presented here will allow the role of the 3D VRTE to transition from a test bed for verifying plane-parallel retrievals to the core engine for performing large-scale retrievals of atmospheric properties for scenes with strongly heterogeneous cloud cover. The primary benefit is that the 3D spatial dependencies of the sampling volume for remote sensing measurements will be explicitly modeled. The lack of default assumptions on cloud horizontal variability will allow for a more flexible parametrization of cloud structure and a more realistic model for measurements of broken cloud fields and the regions near cloud edges. As a near-term application of the adjoint method, Section 3 discussed the use of a multi-pixel measurement operator to correct plane-parallel retrievals that have errors caused by 3D effects, including adjacency effects. Other applications are more ambitious and will require future research into how to parametrize cloud and aerosol properties in 3D. The implementation of the adjoint method using existing codes is a topic of ongoing research, and we welcome collaborations in this effort. The adjoint method provides a way to adjust atmosphere and surface parameters that will scale to large problems — a foundation for a new class of retrieval algorithm, which uses a three-dimensional parametrization of atmosphere and surface properties to simultaneously reconstruct both spatial and microphysical variability.

Acknowledgments

William Martin acknowledges support from NASA Headquarters under the NASA Earth and Space Science Fellowship Program, Grant NNX-10AN85H, and he was partially funded by an IGERT grant from the US National Science Foundation to Columbia University. Also, he is grateful to Alexander Kokhanovsky for organizing a workshop on aerosol remote sensing from space (July 15–19, 2013; International Space Sciences Institute, Bern, CH) and to Thomas Trautmann, Anthony Davis, and the other participants for encouragement and fruitful discussions. Brian Cairns would like to acknowledge support of the work presented here under pre-formulation funding of the Aerosol Cloud Ecosystem (ACE) mission by NASA Headquarters and the ACE program scientists Hal Maring and Paula Bontempi. Guillaume Bal acknowledges support, through JPL Subcontract no. 1497551, from NASA/ESTO's AIST program (3D-TRACE project), and from the NSF under Grant DMS-1108608.

Appendix A. Equivalence of integral and differential equations of 3D VRT

Integral and differential forms of 3D VRT were used interchangeably in the main text of the paper. This appendix justifies such a use by showing that both the forward and adjoint system of integral equations provide a solution to the corresponding differential 3D VRTE. The first theorem states and proves properties from Eqs. (97) to (100). Then these properties will be used in the two theorems that follow: one for forward equivalence and another for adjoint equivalence.

Theorem 2 (*Streaming properties*). *The advective derivative acts on streaming operator output functions according to the following rules*:

$$(\boldsymbol{\nu}\cdot\nabla+\sigma)[\mathcal{T}_{-0}[\boldsymbol{g}]]=0, \qquad (A.1)$$

$$(\boldsymbol{\nu}\cdot\nabla+\sigma)[\mathcal{T}_{00}[\boldsymbol{f}]]=\boldsymbol{f}, \qquad (A.2)$$

$$(-\boldsymbol{\nu}\cdot\nabla+\sigma)[\mathcal{T}^*_{0+}[\boldsymbol{q}]]=0, \qquad (A.3)$$

$$(-\boldsymbol{\nu}\cdot\nabla+\sigma)[\mathcal{T}^*_{00}[\boldsymbol{p}]]=\boldsymbol{p}. \qquad (A.4)$$

Proof. We begin with Eq. (A.1):

$$\begin{aligned}\boldsymbol{\nu}\cdot\nabla\mathcal{T}_{-0}[\boldsymbol{g}](\boldsymbol{x},\boldsymbol{\nu}) &= \frac{\partial}{\partial t}\exp\left(-\int_{t_-}^{t}\mathrm{d}t''\,\sigma(\boldsymbol{x}^\perp+t''\boldsymbol{\nu})\right)\boldsymbol{g}\left(\boldsymbol{x}^\perp+t_-\boldsymbol{\nu},\boldsymbol{\nu}\right),\\ &= -\sigma(\boldsymbol{x})\mathcal{T}_{-0}[\boldsymbol{g}](\boldsymbol{x},\boldsymbol{\nu}).\end{aligned}$$

Next we show Eq. (A.2) using an extension of the Leibniz rule to non-constant limits of integration [45]:

$$\begin{aligned}\boldsymbol{\nu}\cdot\nabla\mathcal{T}_{00}[\boldsymbol{f}](\boldsymbol{x},\boldsymbol{\nu}) &= \frac{\partial}{\partial t}\int_{t_-}^{t}\mathrm{d}t'\,\exp\left(-\int_{t'}^{t}\mathrm{d}t''\,\sigma(\boldsymbol{x}^\perp+t''\boldsymbol{\nu})\right)\boldsymbol{f}\left(\boldsymbol{x}^\perp+t'\boldsymbol{\nu},\boldsymbol{\nu}\right),\\ &= \exp(0)\boldsymbol{f}(\boldsymbol{x}^\perp+t\boldsymbol{\nu},\boldsymbol{\nu})\\ &\quad -\sigma(\boldsymbol{x}^\perp+t\boldsymbol{\nu})\int_{t_-}^{t}\mathrm{d}t'\,\exp\left(-\int_{t'}^{t}\mathrm{d}t''\,\sigma(\boldsymbol{x}^\perp\right.\\ &\quad \left.+t''\boldsymbol{\nu})\right)\boldsymbol{f}(\boldsymbol{x}^\perp+t'\boldsymbol{\nu},\boldsymbol{\nu}),\ =\boldsymbol{f}(\boldsymbol{x},\boldsymbol{\nu})-\sigma(\boldsymbol{x})\mathcal{T}_{00}[\boldsymbol{f}](\boldsymbol{x},\boldsymbol{\nu}).\end{aligned}$$

For Eq. (A.3) we have the following:

$$\begin{aligned}-\boldsymbol{\nu}\cdot\nabla\mathcal{T}^*_{0+}[\boldsymbol{q}](\boldsymbol{x},\boldsymbol{\nu}) &= -\frac{\partial}{\partial t}\exp\left(-\int_{t}^{t_+}\mathrm{d}t''\,\sigma(\boldsymbol{x}^\perp+t''\boldsymbol{\nu})\right)\boldsymbol{q}\left(\boldsymbol{x}^\perp+t_+\boldsymbol{\nu},\boldsymbol{\nu}\right),\\ &= -\sigma(\boldsymbol{x})\mathcal{T}^*_{0+}[\boldsymbol{q}](\boldsymbol{x},\boldsymbol{\nu}).\end{aligned}$$

Lastly, we show Eq. (A.4):

$$-\boldsymbol{\nu}\cdot\nabla\mathcal{T}^*_{00}[\boldsymbol{p}](\boldsymbol{x},\boldsymbol{\nu})$$

$$
\begin{aligned}
&= -\frac{\partial}{\partial t}\int_t^{t_+} \mathrm{d}t' \exp\left(-\int_t^{t'} \mathrm{d}t'' \, \sigma(\boldsymbol{x}^\perp \right.\\
&\qquad \left. + t''\boldsymbol{v})\right) \boldsymbol{p}(\boldsymbol{x}^\perp + t'\boldsymbol{v}, \boldsymbol{v}),\\
&= \exp(0)\boldsymbol{p}(\boldsymbol{x}^\perp + t\boldsymbol{v}, \boldsymbol{v})\\
&\quad - \sigma(\boldsymbol{x}^\perp + t\boldsymbol{v}) \int_t^{t_+} \mathrm{d}t' \exp\left(-\int_t^{t'} \mathrm{d}t'' \, \sigma(\boldsymbol{x}^\perp \right.\\
&\qquad \left. + t''\boldsymbol{v})\right) \boldsymbol{p}(\boldsymbol{x}^\perp + t'\boldsymbol{v}, \boldsymbol{v}), = \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{v}) - \sigma(\boldsymbol{x})\mathcal{T}^*_{00}[\boldsymbol{p}](\boldsymbol{x}, \boldsymbol{v}).
\end{aligned}
$$

The four properties are thus verified. □

Theorem 3 (*Forward 3D VRTE equivalence*). *The Stokes vector $\boldsymbol{u}$ given by the integral formulation of3D VRT in Eqs. (104)–(107) solves the integro-differential 3D VRTE from Definition 2.*

Proof. First, we note that the functions $\boldsymbol{u}|_{\Gamma_+}$ and $\boldsymbol{u}|_{\Gamma_-}$ agree with the limits of $\boldsymbol{u}|_{D\times\mathbb{S}^2}$ along lines approaching the boundary. The boundary conditions are verified by substituting Eqs. (105) and (106) into the left hand side of Eq. (17) and applying Eq. (107) to show the equality with the right-hand side:

$$
\begin{aligned}
&\boldsymbol{u}|_{\Gamma_-} - \mathcal{R}[\boldsymbol{u}|_{\Gamma_+}]\\
&\quad = \boldsymbol{g} - \mathcal{R}[\mathcal{T}_{0+}[\boldsymbol{f}] + \mathcal{T}_{-+}[\boldsymbol{g}]],\\
&\quad = \boldsymbol{g} - \mathcal{R}\mathcal{T}_{0+}[\boldsymbol{f}] + \mathcal{R}\mathcal{T}_{-+}[\boldsymbol{g}],\\
&\quad = \boldsymbol{g}_\odot.
\end{aligned}
$$

The solution on internal points is verified by substituting Eq. (104) into the left hand side of Eq. (16) and applying Theorem 2 and Eq. (107) to show the equality with the right-hand side:

$$
\begin{aligned}
&\boldsymbol{v}\cdot\nabla\boldsymbol{u} + \sigma\boldsymbol{u} - \mathcal{Z}[\boldsymbol{u}]\\
&\quad = (\boldsymbol{v}\cdot\nabla + \sigma)[\mathcal{T}_{00}[\boldsymbol{f}] + \mathcal{T}_{-0}[\boldsymbol{g}]]\\
&\qquad - \mathcal{Z}[\mathcal{T}_{00}[\boldsymbol{f}] + \mathcal{T}_{-0}[\boldsymbol{g}]],\\
&\quad = \boldsymbol{f} - \mathcal{Z}\mathcal{T}_{00}[\boldsymbol{f}] + \mathcal{Z}\mathcal{T}_{-0}[\boldsymbol{g}],\\
&\quad = \boldsymbol{f}_\odot.
\end{aligned}
$$

Thus, the Stokes vector $\boldsymbol{u}$ constructed from solutions $\boldsymbol{f}$ and $\boldsymbol{g}$ of the integral equations solves the differential 3D VRTE. □

Theorem 4 (*Adjoint 3D VRTE equivalence*). *The adjoint Stokes vector, $\boldsymbol{w}$, given by the integral formulation of the adjoint 3D VRTE in Eqs. (113)–(116) solves the integro-differential form of the adjoint 3D VRTE from Definition 3.*

Proof. First, we note that functions $\boldsymbol{w}|_{\Gamma_-}$ and $\boldsymbol{w}|_{\Gamma_+}$ agree with the limits of $\boldsymbol{w}|_{D\times\mathbb{S}^2}$ along lines approaching the boundary. The boundary conditions are verified by substituting Eqs. (114) and (115) into the left hand side of Eq. (27) and applying Eq. (116) to show the equality with the expected right-hand side:

$$
\begin{aligned}
&\boldsymbol{w}|_{\Gamma_+} - \mathcal{R}^*[\boldsymbol{w}|_{\Gamma_-}]\\
&\quad = \boldsymbol{q} - \mathcal{R}^*[\mathcal{T}^*_{-0}[\boldsymbol{p}] + \mathcal{T}^*_{-+}[\boldsymbol{q}]],\\
&\quad = \boldsymbol{q} - \mathcal{R}^*\mathcal{T}^*_{-0}[\boldsymbol{p}] + \mathcal{R}^*\mathcal{T}^*_{-+}[\boldsymbol{q}],\\
&\quad = \boldsymbol{q}_\odot.
\end{aligned}
$$

The solution on internal points is verified by substituting Eq. (113) into the left hand side of Eq. (26) and applying Theorem 2 and Eq. (116) to show the equality with the right-hand side:

$$
\begin{aligned}
&-\boldsymbol{v}\cdot\nabla\boldsymbol{w} + \sigma\boldsymbol{w} - \mathcal{Z}^*[\boldsymbol{w}]\\
&\quad = (-\boldsymbol{v}\cdot\nabla + \sigma)[\mathcal{T}^*_{00}[\boldsymbol{p}] + \mathcal{T}^*_{0+}[\boldsymbol{q}]]\\
&\qquad - \mathcal{Z}^*[\mathcal{T}^*_{00}[\boldsymbol{p}] + \mathcal{T}^*_{0+}[\boldsymbol{q}]],\\
&\quad = \boldsymbol{p} - \mathcal{Z}^*\mathcal{T}^*_{00}[\boldsymbol{p}] + \mathcal{Z}^*\mathcal{T}^*_{0+}[\boldsymbol{q}], \quad = \boldsymbol{p}_\odot.
\end{aligned}
$$

Thus, the adjoint Stokes vector $\boldsymbol{w}$ constructed from the solutions $\boldsymbol{p}$ and $\boldsymbol{q}$ of the integral equations solves the differential 3D VRTE. □

Appendix B. Elementary adjoint property results

Proving the adjoint properties for streaming operators will require us to equate certain multi-variable integrals, and this is facilitated by changing co-ordinates to integrate along chords. After summarizing these coordinate transformations, we will prove the four elementary adjoint properties for streaming operators stated in Eqs. (69)–(72). Following this, we will show the adjoint properties of scattering and reflection operators.

The natural basis to use for streaming integration proofs depends on the direction $\boldsymbol{v}\in\mathbb{S}^2$, and without loss of generality we take the standard basis for $\boldsymbol{v}$ in terms of angles ϑ and φ. This gives the following orthonormal basis for $\mathbb{R}^3$:

$$
\begin{aligned}
\boldsymbol{v} &= [\sin\,\vartheta\,\cos\,\varphi,\;\; \sin\,\vartheta\,\sin\,\varphi,\;\; \cos\,\vartheta]^T,\\
\boldsymbol{\vartheta} &= [\cos\,\vartheta\cos\varphi,\;\; \cos\,\vartheta\,\sin\,\varphi,\;\; -\sin\vartheta]^T,\\
\boldsymbol{\varphi} &= [-\sin\,\varphi,\;\; \cos\,\varphi,\;\; 0]^T.
\end{aligned}
$$

The coordinate transformation for the internal set $D\times\mathbb{S}^2$ is given by the following rules:

$$
\begin{aligned}
t &= \boldsymbol{v}\cdot\boldsymbol{x},\\
y^1 &= \boldsymbol{\vartheta}\cdot\boldsymbol{x},\\
y^2 &= \boldsymbol{\varphi}\cdot\boldsymbol{x},\\
\vartheta &= \vartheta,\\
\varphi &= \varphi.
\end{aligned}
$$

The rules change coordinates so that the chord parameter, t, controls the projection of $\boldsymbol{x}$ along $\boldsymbol{v}$, while variables y^1 and y^2 control location in the 2-dimensional plane orthogonal to $\boldsymbol{v}$. We note that the perpendicular component used in the body of the paper, $\boldsymbol{x}^\perp$, is related to y^1 and y^2:

$$
\boldsymbol{x}^\perp = \boldsymbol{\vartheta}y^1 + \boldsymbol{\varphi}y^2.
$$

The associated surface element $\mathrm{d}y^1\,\mathrm{d}y^2$ is written more compactly as $\mathrm{d}S_{\boldsymbol{x}^\perp}$. For evaluating integrals over the internal set, there is no change in weight associated with the chord transformation and

$$
\mathrm{d}S_{\boldsymbol{v}}\,\mathrm{d}S_{\boldsymbol{x}^\perp}\,\mathrm{d}t = \mathrm{d}S_{\boldsymbol{v}}\,\mathrm{d}x^1\,\mathrm{d}x^2\,\mathrm{d}x^3. \tag{B.1}
$$

For position direction pairs on the outgoing and incoming sets, the change of weight is exactly the cosine of the angle of incidence. For coordinate transformation from the perpendicular component of a chord, $\boldsymbol{x}^\perp$, to the outgoing position, $\boldsymbol{x}_+$, we write

$$
\mathrm{d}S_{\boldsymbol{v}}\,\mathrm{d}S_{\boldsymbol{x}^\perp} = \mathrm{d}S_{\boldsymbol{x}_+}\,\mathrm{d}S_{\boldsymbol{v}_+}\,|\boldsymbol{v}_+\cdot\nabla h(\boldsymbol{x}_+)|, \tag{B.2}
$$

and for coordinate transformations to the incoming position, $\boldsymbol{x}_-$, we write

$$
\mathrm{d}S_{\boldsymbol{v}}\,\mathrm{d}S_{\boldsymbol{x}^\perp} = \mathrm{d}S_{\boldsymbol{x}_-}\,\mathrm{d}S_{\boldsymbol{v}_-}\,|\boldsymbol{v}_-\cdot\nabla h(\boldsymbol{x}_-)|. \tag{B.3}
$$

These transformations are helpful in proving the four elementary adjoint properties for streaming operators.

Theorem 5 (*Internal streaming adjoint property*). *The internal streaming operators $\mathcal{T}_{00}$ and $\mathcal{T}^*_{00}$, defined by Eqs.* (89) *and* (93), *satisfy the adjoint property*:

$$\langle \boldsymbol{p}, \mathcal{T}_{00}[\boldsymbol{f}]\rangle_{D\times\mathbb{S}^2} = \langle \mathcal{T}^*_{00}[\boldsymbol{p}], \boldsymbol{f}\rangle_{D\times\mathbb{S}^2} \tag{B.4}$$

Proof. Beginning with the left hand side of Eq. (B.4), the equality is demonstrated in the following steps:

$$\int_D \mathrm{d}V_{\boldsymbol{x}} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}}\, \boldsymbol{p}(\boldsymbol{x},\boldsymbol{v})^T \cdot \mathcal{T}_{00}[\boldsymbol{f}](\boldsymbol{x},\boldsymbol{v}) = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{(t_-,t_+)} \int_{t_-}^{t_+} \mathrm{d}t \int_{t_-}^{t} \mathrm{d}t'\, \boldsymbol{p}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v})^T \cdot$$
$$\exp\left(-\int_{t'}^{t} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{f}(\boldsymbol{x}^\perp + t'\boldsymbol{v},\boldsymbol{v}), = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{(t_-,t_+)} \int_{t_-}^{t_+} \mathrm{d}t' \left(\int_{t'}^{t_+} \mathrm{d}t\right.$$
$$\left.\exp\left(-\int_{t'}^{t} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{p}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v})\right)^T \cdot \boldsymbol{f}(\boldsymbol{x}^\perp + t'\boldsymbol{v},\boldsymbol{v}), = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{(t_-,t_+)} \int_{t_-}^{t_+} \mathrm{d}t'\, \mathcal{T}^*_{00}[\boldsymbol{p}](\boldsymbol{x}^\perp + t'\boldsymbol{v},\boldsymbol{v})^T$$
$$\cdot \boldsymbol{f}(\boldsymbol{x}^\perp + t'\boldsymbol{v},\boldsymbol{v}), \quad = \int_D \mathrm{d}V_{\boldsymbol{x}} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}}\, \mathcal{T}^*_{00}[\boldsymbol{p}](\boldsymbol{x},\boldsymbol{v})^T \cdot \boldsymbol{f}(\boldsymbol{x},\boldsymbol{v}).$$

This is the right-hand side of Eq. (B.4). We note two key steps. First, we rewrote the integral over space as an integral over chords:

$$\int_D \mathrm{d}V_x \approx \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{(t_-,t_+)} \int_{t_-}^{t_+} \mathrm{d}t.$$

Second, we exchanged the order of integration of $\mathrm{d}t$ and $\mathrm{d}t'$:

$$\int_{t_-}^{t_+} \mathrm{d}t \int_{t_-}^{t} \mathrm{d}t' \approx \int_{t_-}^{t_+} \mathrm{d}t' \int_{t'}^{t_+} \mathrm{d}t.$$

Then, changing back to the original coordinates completed the proof. □

Theorem 6 (*Internal source to outgoing Stokes vector streaming adjoint property*). *The streaming operator $\mathcal{T}_{0+}$ defined by Eq.* (90) *has adjoint $\mathcal{T}^*_{0+}$ defined by Eq.* (94):

$$\langle \boldsymbol{q}, \mathcal{T}_{0+}[\boldsymbol{f}]\rangle_{\Gamma_+} = \langle \mathcal{T}^*_{0+}[\boldsymbol{q}], \boldsymbol{f}\rangle_{D\times\mathbb{S}^2}. \tag{B.5}$$

Proof. Beginning with the definition of the left hand side of Eq. (B.5) we proceed through the following steps:

$$\int_{\partial D} \mathrm{d}S_{\boldsymbol{x}_+} \int_{\boldsymbol{v}_+\cdot\nabla h > 0} \mathrm{d}S_{\boldsymbol{v}_+}\, |\boldsymbol{v}_+ \nabla h(\boldsymbol{x}_+)| \boldsymbol{q}(\boldsymbol{x}_+,\boldsymbol{v}_+)^T \cdot \mathcal{T}_{0+}[\boldsymbol{f}](\boldsymbol{x}_+,\boldsymbol{v}_+) = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{t_+} \boldsymbol{q}(\boldsymbol{x}^\perp + t_+\boldsymbol{v},\boldsymbol{v})^T$$
$$\cdot \int_{t_-}^{t_+} \mathrm{d}t\, \exp\left(-\int_{t}^{t_+} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{f}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v}), = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{t_+} \int_{t_-}^{t_+} \mathrm{d}t \left(\exp\left(-\int_{t}^{t_+} \sigma(\boldsymbol{x}^\perp\right.\right.$$
$$\left.\left. + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{q}(\boldsymbol{x}^\perp + t_+\boldsymbol{v},\boldsymbol{v})\right)^T \cdot \boldsymbol{f}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v}), = \int_D \mathrm{d}V_{\boldsymbol{x}} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}}\, \mathcal{T}^*_{0+}[\boldsymbol{q}](\boldsymbol{x},\boldsymbol{v})^T \cdot \boldsymbol{f}(\boldsymbol{x},\boldsymbol{v}).$$

These steps show the equality. The key was writing each outgoing position $\boldsymbol{x}_+$ as the endpoint $\boldsymbol{x}_+ = \boldsymbol{x}^\perp + t_+\boldsymbol{v}$ of a chord through the domain. Recognizing the definition of $\mathcal{T}^*_{0+}$ and changing to standard coordinates complete the proof. □

Theorem 7 (*Incoming to internal Stokes vector streaming adjoint property*). *The streaming operators $\mathcal{T}_{-0}$ and $\mathcal{T}^*_{-0}$ defined by Eqs.* (91) *and* (95) *are adjoint*:

$$\langle \boldsymbol{p}, \mathcal{T}_{-0}[\boldsymbol{g}]\rangle_{D\times\mathbb{S}^2} = \langle \mathcal{T}^*_{-0}[\boldsymbol{p}], \boldsymbol{g}\rangle_{\Gamma_-}. \tag{B.6}$$

Proof. We begin with the definition of the left hand side of Eq. (B.6) and show the equality with the following steps:

$$\int_D \mathrm{d}V_{\boldsymbol{x}} \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \boldsymbol{p}(\boldsymbol{x},\boldsymbol{v})^T \cdot \mathcal{T}_{-0}[\boldsymbol{g}](\boldsymbol{x},\boldsymbol{v}) = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{t_-} \int_{t_-}^{t_+} \mathrm{d}t\, \boldsymbol{p}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v})^T \cdot$$
$$\exp\left(-\int_{t_-}^{t} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{g}(\boldsymbol{x}^\perp + t_-\boldsymbol{v},\boldsymbol{v}), = \int_{\mathbb{S}^2} \mathrm{d}S_{\boldsymbol{v}} \int_{\boldsymbol{x}^\perp(D)} \mathrm{d}S_{\boldsymbol{x}^\perp} \sum_{t_-} \left(\int_{t_-}^{t_+} \mathrm{d}t\right.$$
$$\left.\exp\left(-\int_{t_-}^{t} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{v})\mathrm{d}t''\right) \boldsymbol{p}(\boldsymbol{x}^\perp + t\boldsymbol{v},\boldsymbol{v})\right)^T \cdot \boldsymbol{g}(\boldsymbol{x}^\perp + t_-\boldsymbol{v},\boldsymbol{v}), = \int_{\partial D} \mathrm{d}S_{\boldsymbol{x}_-} \int_{\boldsymbol{v}_-\cdot\nabla h < 0} \mathrm{d}S_{\boldsymbol{v}_-}\, |\boldsymbol{v}_- \cdot \nabla h(\boldsymbol{x}_-)| \mathcal{T}^*_{-0}[\boldsymbol{p}](\boldsymbol{x}_-,\boldsymbol{v}_-)^T$$
$$\cdot \boldsymbol{g}(\boldsymbol{x}_-,\boldsymbol{v}_-). \qquad \square$$

Theorem 8 (*Incoming to outgoing Stokes vector streaming adjoint property*). *The streaming operators $\mathcal{T}_{-+}$ and $\mathcal{T}^*_{-+}$ defined by Eqs.* (92) *and* (96) *are adjoint*:

$$\langle \boldsymbol{q}, \mathcal{T}_{-+}[\boldsymbol{g}] \rangle_{\Gamma_+} = \langle \mathcal{T}^*_{-+}[\boldsymbol{q}], \boldsymbol{g} \rangle_{\Gamma_-}. \tag{B.7}$$

Proof. We begin with the definition of the left hand side of Eq. (B.7) and show the equality with the following steps:

$$\int_{\partial D} dS_{\boldsymbol{x}_+} \int_{\boldsymbol{\nu}_+ \cdot \nabla h > 0} dS_{\boldsymbol{\nu}_+} |\boldsymbol{\nu}_+ \cdot \nabla h(\boldsymbol{x}_+)| \boldsymbol{q}(\boldsymbol{x}_+, \boldsymbol{\nu}_+)^T \cdot \mathcal{T}_{-+}[\boldsymbol{g}](\boldsymbol{x}_+, \boldsymbol{\nu}_+), = \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}} \int_{\boldsymbol{x}^\perp(D)} dS_{\boldsymbol{x}^\perp} \sum_{t_+} \boldsymbol{q}(\boldsymbol{x}^\perp + t_+ \boldsymbol{\nu}, \boldsymbol{\nu})^T \cdot$$
$$\exp\left(-\int_{t_-}^{t_+} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{\nu}) dt'' \right) \boldsymbol{g}(\boldsymbol{x}^\perp + t_- \boldsymbol{\nu}, \boldsymbol{\nu}), = \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}} \int_{\boldsymbol{x}^\perp(D)} dS_{\boldsymbol{x}^\perp} \sum_{t_-} \left(\exp\left(-\int_{t_-}^{t_+} \sigma(\boldsymbol{x}^\perp + t''\boldsymbol{\nu}) dt'' \right) \right.$$
$$\left. \boldsymbol{q}(\boldsymbol{x}^\perp + t_+ \boldsymbol{\nu}, \boldsymbol{\nu}) \right)^T \cdot \boldsymbol{g}(\boldsymbol{x}^\perp + t_- \boldsymbol{\nu}, \boldsymbol{\nu}), = \int_{\partial D} dS_{\boldsymbol{x}_-} \int_{\boldsymbol{\nu}_- \cdot \nabla h < 0} dS_{\boldsymbol{\nu}_-} |\boldsymbol{\nu}_- \cdot$$
$$\nabla h(\boldsymbol{x}_-)| \mathcal{T}^*_{-+}[\boldsymbol{q}](\boldsymbol{x}_-, \boldsymbol{\nu}_-)^T \cdot \boldsymbol{g}(\boldsymbol{x}_-, \boldsymbol{\nu}_-). \qquad \square$$

Theorem 9 (*Scattering operator adjoint property*). *Scattering operations $\mathcal{Z}$ and $\mathcal{Z}^*$ are adjoint with respect to the internal inner product*:

$$\langle \boldsymbol{p}, \mathcal{Z}[\boldsymbol{u}] \rangle_{D \times \mathbb{S}^2} = \langle \mathcal{Z}^*[\boldsymbol{p}], \boldsymbol{u} \rangle_{D \times \mathbb{S}^2}. \tag{B.8}$$

Proof. The left hand side of Eq. (B.8) is shown to equal the right-hand side by interchanging the order of integration over $\mathbb{S}^2$:

$$\langle \boldsymbol{p}, \mathcal{Z}[\boldsymbol{u}] \rangle_{D \times \mathbb{S}^2} = \int_D dV_{\boldsymbol{x}} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}}\, \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\nu})^T \cdot \frac{1}{4\pi} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}'}\, \boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}') \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}'), = \int_D dV_{\boldsymbol{x}} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}'} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}} \frac{1}{4\pi} \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\nu})^T \cdot \boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}')$$
$$\cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}'), = \int_D dV_{\boldsymbol{x}} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}'} \left(\frac{1}{4\pi} \int_{\mathbb{S}^2} dS_{\boldsymbol{\nu}}\, \boldsymbol{Z}(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}')^T \cdot \boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\nu}) \right)^T \cdot \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\nu}'), = \langle \mathcal{Z}^*[\boldsymbol{p}], \boldsymbol{u} \rangle_{D \times \mathbb{S}^2}. \qquad \square$$

Theorem 10 (*Reflection operator adjoint property*). *The reflection operations $\mathcal{R}$ and $\mathcal{R}^*$ are adjoint*:

$$\langle \boldsymbol{w}|_{\Gamma_-}, \mathcal{R}[\boldsymbol{u}_{\Gamma_+}] \rangle_{\Gamma_-} = \langle \mathcal{R}^*[\boldsymbol{w}|_{\Gamma_-}], \boldsymbol{u}_{\Gamma_+} \rangle_{\Gamma_+} \tag{B.9}$$

Proof. The left hand side of Eq. (B.9) is equated with the right-hand side by interchanging the order of integration:

$$\langle \boldsymbol{w}|_{\Gamma_-}, \mathcal{R}[\boldsymbol{u}|_{\Gamma_+}] \rangle_{\Gamma_-} = \int_{\partial D} dS_{\boldsymbol{x}_-} \int_{\boldsymbol{\nu}_- \cdot \nabla h < 0} dS_{\boldsymbol{\nu}_-} \left| \boldsymbol{\nu}_- \cdot \nabla h \right| \boldsymbol{w}(\boldsymbol{x}_-, \boldsymbol{\nu}_-)^T$$
$$\cdot \frac{1}{2\pi} \int_{\boldsymbol{\nu}_+ \cdot \nabla h > 0} dS_{\boldsymbol{\nu}_+} \left| \boldsymbol{\nu}_+ \cdot \nabla h \right| \boldsymbol{R}(\boldsymbol{x}_-, \boldsymbol{\nu}_-, \boldsymbol{\nu}_+) \cdot \boldsymbol{u}(\boldsymbol{x}_-, \boldsymbol{\nu}_+), = \int_{\partial D} dS_{\boldsymbol{x}_-} \int_{\boldsymbol{\nu}_+ \cdot \nabla h > 0} dS_{\boldsymbol{\nu}_+} |\boldsymbol{\nu}_+ \cdot \nabla h|$$
$$\left(\frac{1}{2\pi} \int_{\boldsymbol{\nu}_- \cdot \nabla h < 0} dS_{\boldsymbol{\nu}_-} |\boldsymbol{\nu}_- \cdot \nabla h|\, \boldsymbol{R}(\boldsymbol{x}_-, \boldsymbol{\nu}_-, \boldsymbol{\nu}_+)^T \cdot \boldsymbol{w}(\boldsymbol{x}_-, \boldsymbol{\nu}_-) \right)^T \cdot \boldsymbol{u}(\boldsymbol{x}_-, \boldsymbol{\nu}_+), = \langle \mathcal{R}^*[\boldsymbol{w}|_{\Gamma_-}], \boldsymbol{u}|_{\Gamma_+} \rangle_{\Gamma_+}. \qquad \square$$

References

[1] C.o.E.S. NRC, A. from Space: A Community Assessment, N. R. C. Strategy for the Future, Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond. The National Academies Press, Washington DC, 2007.

[2] Mishchenko MI, Cairns B, Hansen JE, Travis LD, Burg R, Kaufman YJ, et al. Monitoring of aerosol forcing of climate from space: analysis of measurement requirements. J Quant Spectrosc Radiat Transf 2004;88:149–61.

[3] Waquet F, Riedi J, Labonnote LC, Goloub P, Cairns B, Deuzé J-L, et al. Aerosol remote sensing over clouds using A-train observations. J Atmos Sci 2009:2468–80.

[4] Knobelspiesse K, Cairns B, Mishchenko M, Chowdhary J, Tsigaridis K, van Diedenhoven B, et al. Analysis of fine-mode aerosol retrieval capabilities by different passive remote sensing instrument designs. Opt Express 2012;20:21457–84.

[5] Horváth A, Davies R. Anisotropy of water cloud reflectance: a comparison of measurements and 1D theory. Geophys Res Lett 2004;31:L01102.

[6] Zhang Z, Ackerman AS, Feingold G, Platnick S, Pincus R, Xue H. Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: case studies based on large-eddy simulations. J Geophys Res: Atmos 2012;117:D19208.

[7] Varnai T, Marshak A. Global CALIPSO observations of aerosol changes near clouds. IEEE Geosci Remote Sens Lett 2011;8:19–23.

[8] Marshak A, Davis A, editors. 3D radiative transfer in cloudy atmospheres. Springer; 2005.

[9] Evans KF, Marshak A, Varnai T. The potential for improved boundary layer cloud optical depth retrievals from the multiple directions of MISR. J Atmos Sci 2008;65:3179–96.

[10] Diner DJ, Braswell BH, Davies R, Gobron N, Hu J, Jin Y, et al. The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces. Remote Sens Environ 2005;97:495–518.

[11] Box MA. Radiative perturbation theory: a review. Environ Model Softw 2002;17:95–106.

[12] Rozanov VV, Rozanov AV. Relationship between different approaches to derive weighting functions related to atmospheric remote sensing problems. J Quant Spectrosc Radiat Transf 2007;105:217–42.

[13] Ren K, Bal G, Hielscher AH. Frequency domain optical tomography based on the equation of radiative transfer. SIAM J Sci Comput 2006;28:1463–89.

[14] Balima O, Favennec Y, Boulanger J, Charette A. Optical tomography with the discontinuous Galerkin formulation of the radiative transfer equation in frequency domain. J Quant Spectrosc Radiat Transf 2012;113:805–14.

[15] Saratoon T, Tarvainen T, Cox BT, Arridge SR. A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation. Inverse Probl 2013;29:075006. (19pp).

[16] Hasekamp O, Landgraf J. Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing. J Geophys Res: Atmos 2005;110:D04203.

[17] Ukhinov SA, Yurkov DI. Computation of the parametric derivatives of polarized radiation and the solution of inverse atmospheric optics problems. Russ J Numer Anal Math Model 2002;17:283–303.

[18] Postylyakov OV. Linearized vector radiative transfer model MCC++ for a spherical atmosphere. J Quant Spectrosc Radiat Transf 2004;88: 297–317.
[19] Spurr RJD, Kurosu TP, Chance KV. A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. J Quant Spectrosc Radiat Transf 2001;68:689–735.
[20] Doicu A, Trautmann T. Adjoint problem of radiative transfer for a pseudo-spherical atmosphere and general viewing geometries. J Quant Spectrosc Radiat Transf 2009;110:464–76.
[21] Walter HH, Landgraf J. Linearization of radiative transfer in spherical geometry: an application of the forward-adjoint perturbation theory. In: Kokhanovsky AA, editor. Light scattering reviews. Springer Praxis Books, vol. 5. Berlin, Heidelberg: Springer; 2010. p. 105–45.
[22] Spurr RJD. Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment. J Quant Spectrosc Radiat Transf 2002;75:129–75.
[23] Polonsky IN, Box MA, Davis AB. Radiative transfer through inhomogeneous turbid media: implementation of the adjoint perturbation approach at the first order. J Quant Spectrosc Radiat Transf 2003;78: 85–98.
[24] Box MA, Polonsky IN, Davis AB. Higher order perturbation theory applied to radiative transfer in non-plane-parallel media. J Quant Spectrosc Radiat Transf 2003;78:105–18.
[25] Evans KF. The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. J Atmos Sci 1998;55:429–46.
[26] Pincus R, Evans KF. Computational cost and accuracy in calculating three-dimensional radiative transfer: results for new implementations of Monte Carlo and SHDOM. J Atmos Sci 2009;66:3131–46.
[27] Emde C, Buras R, Mayer B, Blumthaler M. The impact of aerosols on polarized sky radiance: model development, validation, and applications. Atmos Chem Phys 2010;10:383–96.
[28] Cornet C, C-Labonnote L, Szczap F. Three-dimensional polarized monte carlo atmospheric radiative transfer model (3DMCPOL): 3D effects on polarized visible reflectances of a cirrus cloud. J Quant Spectrosc Radiat Transf 2010;111:174–86.
[29] 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;118:121–31.
[30] Evans KF. Spherical harmonic discrete ordinate method for atmospheric radiative transfer, 2014 ⟨http://nit.colorado.edu/shdom.html⟩. [Accessed: March 7, 2014].
[31] Marshak A, Platnick S, Vrnai T, Wen G, Cahalan RF. Impact of three-dimensional radiative effects on satellite retrievals of cloud droplet sizes. J Geophys Res: Atmos 2006;111:D09207.
[32] Zhang Z, Platnick S. An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands. J Geophys Res: Atmos 2011;116:D20215.
[33] Dubovik O, Herman M, Holdak A, Lapyonok T, Tanr D, Deuz JL, et al. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. Atmos Measur Techn 2011;4:975–1018.
[34] Hair JW, Hostetler CA, Cook AL, Harper DB, Ferrare RA, Mack TL, et al. Airborne high spectral resolution lidar for profiling aerosol optical properties. Appl Opt 2008;47:6734–52.
[35] Huang D, Liu Y, Wiscombe W. Determination of cloud liquid water distribution using 3D cloud tomography. J Geophys Res: Atmos 2008;113:D13201.
[36] Huang D, Liu Y, Wiscombe W. Cloud tomography: role of constraints and a new algorithm. J Geophys Res: Atmos 2008;113:D23203.
[37] Mishchenko MI, Travis LD, Lacis AA. Multiple scattering of light by particles: radiative transfer and coherent backscattering. New York: Cambridge University Press; 2006.
[38] Remer LA, Kaufman YJ, Tanre D, Mattoo S, Chu DA, Martins JV, et al. The MODIS aerosol algorithm, products, and validation. J Atmos Sci 2005;62:947–73.
[39] Hasekamp OP, Litvinov P, Butz A. Aerosol properties over the ocean from PARASOL multiangle photopolarimetric measurements. J Geophys Res: Atmos 2011;116:D14204.
[40] Dubovik O, Holben B, Eck TF, Smirnov A, Kaufman YJ, King MD, et al. Variability of absorption and optical properties of key aerosol types observed in worldwide locations. J Atmos Sci 2002;59:590–608.
[41] Rodgers CD. Inverse methods for atmospheric sounding: theory and practice. Series on atmospheric, oceanic and planetary physics, vol. 2. River Edge, NJ: World Scientific Publishing Co. Inc.; 2000.
[42] Tarantola A. Inverse problem theory and methods for model parameter estimation. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM); 2005.
[43] Aides A, Schechner YY, Holodovsky V, Garay MJ, Davis AB. Multi sky-view 3D aerosol distribution recovery. Opt Express 2013;21: 25820–33.
[44] Guenther RB, Lee JW. Partial differential equations of mathematical physics and integral equations. New York: Dover Publications; 1996 [See Chapters 7-3 and 7-4, and Theorem 4-2].
[45] Kaplan W. Advanced calculus. third ed.Reading, MA: Addison-Wesley Publishing Company Advanced Book Program; 1984.