Nonlinear Hyperspectral Unmixing With Graphical Models

Rob Heylen[iD], *Member, IEEE*, Vera Andrejchenko[iD], *Student Member, IEEE*, Zohreh Zahiri,
Mario Parente, *Senior Member, IEEE*, and Paul Scheunders, *Senior Member, IEEE*

Abstract—In optical remote sensing, phenomena such as multiple scattering, shadowing, and spatial neighbor effects generate spectral reflectances that are nonlinear mixtures of the reflectances of the surface materials. Using hyperspectral images, the obtained spectral reflectances can be unmixed. We present a general method for creating nonlinear mixing models, based on a ray-based approximation of light and a graph-based description of the optical interactions. This results in a stochastic process which can be used to calculate path probabilities and contributions, and their weighted sum. In many cases, a closed-form equation can be obtained. We illustrate the approach by deriving several existing mixing models, such as linear, bilinear, and multilinear mixing (MLM) models popular in remote sensing, layered models for vegetation canopies, and intimate mineral mixtures. Furthermore, we use the proposed technique to derive a new mixing model, which extends the MLM model with shadowing. Experiments on artificial and real data show the positive traits of this model, which also demonstrates the power of the graphical model approach.

Index Terms—Hyperspectral unmixing, spectral mixing models.

I. INTRODUCTION

HYPERSPECTRAL unmixing [1] concerns the decomposition of the spectra captured in the pixels of a hyperspectral image into their constituent components, or endmembers. In most applications, one assumes that a finite number of elemental pure materials exist, and that each of these endmember materials might be present in the instantaneous field of view (IFOV) of the pixel. Spectral mixing will then take place, where incoming light can undergo complex interactions with different components in the scene, resulting in the observed spectrum. Spectral mixing models attempt to describe these optical interactions as a function of the endmember spectra, their abundances, and possibly other parameters such as geometrical, atmospheric, or physical parameters [2].

One of the most basic mixing models is the linear mixing model (LMM), where one assumes that an observed spectrum is a linear combination of endmember spectra, with the abundances as their linear coefficients, and Gaussian noise present. One often assumes that the abundances cannot be negative, leading to an abundance nonnegativity constraint (ANC). Many authors also consider an abundance sum-to-one constraint (ASC), reasoning that the entire spectral signal has to be decomposed into endmember contributions. This latter constraint is often dropped in recent publications due to several reasons, such as allowing shadows and scaling, to describe nonlinear mixing effects [3], and to allow for L_1 sparsity in the abundance vectors [4].

The physical interpretation behind the LMM is that every incoming photon or light ray will interact once with a specific endmember before reaching the observer. The probability of interaction is proportional to the abundance of the endmember in the pixel's IFOV, leading to the ANC and ASC. This interpretation breaks down when the scene is more complex— geometrical structures such as tree canopies and those found in urban scenarios will cause shadowing, multiple reflections, and secondary illumination. Transmittance through objects, for instance, in canopies or mineral mixtures, can play a large role and is not considered in the LMM. Intimate mixtures, where the scene is more adequately described by a space filled with particles, also show drastically different mixing behaviors compared to linear mixtures [5].

Several more advanced mixing models have been derived that attempt to improve upon this situation, and these can be divided into several classes [6], [7] depending on the complexity. A popular class of nonlinear mixing models in remote sensing is the bilinear models. These models extend the LMM with second-order reflections, which describe photons or light rays interacting with two materials instead of one, in order to describe secondary illumination effects. Many bilinear models have been proposed, with wildly varying assumptions on the free parameters or underlying physics [8]–[17]. Although bilinear models can be useful to address some forms of nonlinearity caused by secondary reflections, they are often not based on sound physical principles, and the mixing equations can allow highly unphysical spectra (e.g., reflectance values far outside the [0, 1] interval).

The second class of nonlinear models is based on simplified geometrical modeling of the scene and the resulting

reflection processes. These techniques often resort to layered structures for modeling certain scenarios, such as tree canopies [18] or intimate mineral mixtures [19]–[21]. Due to the layered structure, an infinite number of reflections between the layers are allowed, and by employing ray-based or radiosity-based analysis, mixing models can often be derived as a series of higher order interactions between objects.

Alternatively, one can attempt to extend bilinear models to all orders of interactions, taking care that the physical interpretation of the mixing model is not lost, and by avoiding the introduction of a large number of free parameters. Recently, such extensions were introduced, e.g., based on a polytope decomposition method [22] or a harmonic description of higher order combinations of endmembers, beyond the classical polynomial combinations [23]. Another recent attempt at such an extension is the multilinear mixing (MLM) model [24], based on a simple probabilistic model of endmember interactions—at each interaction with an endmember, the light ray will have further interactions with a probability P, and go to the observer with probability $(1 - P)$. This way, all orders of interactions between endmembers can be accommodated while introducing only a single additional free parameter on top of the abundances.

Alternatively, kernel-based nonlinear models were defined that include additive nonlinear fluctuations on top of the linear model, defined in reproducing kernel Hilbert spaces [25], [26]. Pure data-driven methods, based on manifold learning [27], [28], or making use of machine learning [29], [30] were developed as well.

Even more advanced are the radiative transfer models, often employed for modeling intimate mixtures of minerals. Typically, the medium is represented as a half-space filled with particles with a given density and distribution of physical attributes. By solving the radiative transfer equations, excellent forward models can be created that predict the resulting reflectance [5], [31]–[33]. These models are often hard to invert, and therefore less fit for use in remote sensing. A notable exception is the Hapke model [5] which is often encountered in the remote sensing literature. However, it must be noted that several severe simplifications have been applied in order to make the model invertible, and that the full Hapke model is much more advanced than the approximations used in remote sensing.

A common issue between these different classes of models is the varying levels of motivation from a physics standpoint. The purely data-driven unmixing techniques do not assume any underlying physics at all. Others employ highly idealized models (e.g., the LMM), or focus on only a single aspect (e.g., second-order interactions in bilinear models) without taking the possible consequences on the mixing equation into account, resulting in mixing equations that allow highly nonphysical behavior such as reflectances that are negative or significantly higher than one. Layered and radiosity-based approaches typically do start from a detailed physical modeling of the allowed interactions, but show significant differences in the interpretation, description, and solving method, which make it hard to compare or unify such models.

In this paper, our goal is to develop a more general framework for deriving mixing models, aimed at the hyperspectral unmixing application. The proposed framework allows the preservation of the simplicity provided by spectral mixture analysis while allowing the introduction of some of the physical effects of the radiative transfer modeling.

The framework is based on graphical models, where the starting point is to describe the allowed optical interactions that a light ray can undergo when it travels from light source to observer. Here, it is important to ensure that physical and mathematical constraints are incorporated at each allowed interaction, such as conservation of energy or sum to one of probabilities. These graphical models then lead to a set of allowed paths, where each path has a given probability and contribution to the total spectral signal. A weighted sum over all these paths leads to the mixing equation of the model.

This framework allows the user to focus on the allowed physical interactions first, avoiding many pitfalls commonly encountered in the construction of mixing models. Several constraints often arise naturally, such as nonnegativity or sum-to-one constraints. Furthermore, several existing mixing models can be derived in this framework, and we will provide some simple and more advanced examples, such as the LMM, bilinear models, the layered model of Borel and Gerstl [18], the Shkuratov model for intimate mineral mixtures [21], and the MLM model [24]. Next, as an example, we will employ the framework to derive a new mixing model, which extends the MLM model with shadows. This shadow model will be illustrated on several artificial and real data sets, where we show the excellent performance in detecting shadows and higher order interactions, further demonstrating the power of the proposed framework.

Note that the forward formulation of illumination, multiple scattering, and other nonlinear effects is common to any optical image not only to hyperspectral images but the latter allows for spectral unmixing.

The simplicity of the framework is, at the same time, its weakness, since more complicated spectral interaction effects such as fluorescence, luminescence, absorption and diffusion, and external effects such as multiple viewing angles and dependence from the digital elevation model cannot be taken easily into account.

This paper is organized as follows. In Section II, the proposed framework is introduced. Section III presents several examples, where we employ this framework to derive the mixing equations of well-known mixing models, such as linear, bilinear, multilinear, and layered models, and one example of an intimate mixing model. Section IV introduces a new mixing model, where the MLM model is extended with the capability of handling shadows. Section V describes the illustrating experiments and Section VI contains the conclusions, followed by the bibliography.

II. Stochastic Process of Optical Reflections

In a hyperspectral image, the signal captured in each pixel is a composite of all photons that have reached the sensor in a certain time frame. Typically, one assumes that these photons

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS

3

are emitted by a light source and interact with the objects in the scene before reaching the sensor. Note that, typically, a pre-processing chain is employed to remove atmospheric effects, perform geometrical correction, pixel resampling, calibration, and optical corrections. We do not consider these in this paper, and assume that a perfect reflectance image of the scene is available.

To predict the observed signal as a function of the scene composition, one can build a statistical model for the possible paths that a single photon can follow. As we are interested in the overall average of photons reaching the sensor, we employ a ray-based approximation of light to model these paths, where light rays are emitted by the source and interact with objects, and each interaction alters the optical properties of the light ray.

We assume that the number of possible interactions is finite, leading to a discrete state space, and that the interaction process is a discrete-time stochastic process. Let the discrete set S contain all possible interactions that a light ray can undergo. A light path is then defined by the random variable $\{X_n\}_{n \geq 0}$ with $\forall n : X_n \in S$. The probability of observing a certain path is given by

$$P(X_0 = s_0, X_1 = s_1, \ldots, X_n = s_n), \quad s_0, \ldots s_n \in S. \quad (1)$$

It is often convenient to assume that this stochastic process obeys the Markov property

$$P(X_{n+1} = s | X_0, \ldots, X_n) = P(X_{n+1} = s | X_n) \quad (2)$$

$$P(X_{n+1} = s_j | X_n = s_i) = p_{ij}. \quad (3)$$

This means that the stochastic process is memoryless. The next state depends only on the current state, and not previous states. Furthermore, the transition probabilities between states are constant, and can be represented by a transition matrix. Such processes can be easily represented by directed graphs, where each node represents a state, and the directed edges represent the transition probabilities between states.

For our purpose, we are only interested in paths reaching the observer, which means that every path eventually ends in the state in S associated with the observer, where it will never leave. This is an example of an absorbing Markov chain. Furthermore, the path always begins at a light source, meaning that X_0 is always associated with a light source state. Note that there can be multiple light sources.

At each state, the optical properties of the light ray will be altered. Let $T(s_i)$ be the operator that acts on the light ray in state $s_i \in S$. The total effect on the light ray when a path $(X_0 = s_0, X_1 = s_1, \ldots, X_n = s_n)$ is followed is then given by $\prod_{i=1}^{n} T(s_i)$. In the context of optical interactions, this alteration typically involves a relative change in the spectrum of the light ray according to the reflectance of the object associated with that state, but more exotic effects can be modeled as well, e.g., fluorescence and luminescence.

To determine the observed signal x, a weighted average over all possible paths needs to be taken. Let L describe the path length, i.e., number of interactions that a light ray undergoes
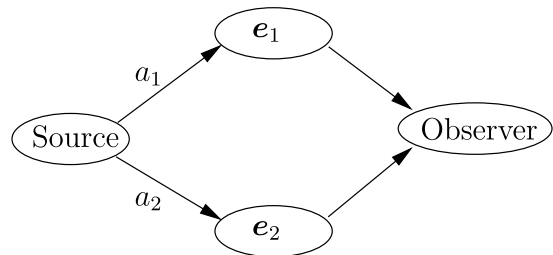

Fig. 1. States and transition probabilities in the LMM, for a two-endmember scenario. Unmarked edges have an implicit probability of one.

between the source state s and the observer state o

$$x = \sum_{L=1}^{\infty} \left(\sum_{s_1 \in S} \cdots \sum_{s_L \in S} \right) P(\text{path}) \prod_{i=1}^{L} T(s_i)$$

$$P(\text{path}) = P(X_0 = s, X_1 = s_1, \ldots, X_L = s_L, X_{L+1} = o). \quad (4)$$

In practice, the spectrum of the light ray is described by a vector in spectral space. The states in S are typically light sources, endmember interactions, the observer state, and possibly other states depending on the model. The operators $T(s)$ that act on the light ray are constant vectors in the case of light sources, componentwise multiplication with a reflectance vector for reflections, or an identity operator for the observer state.

In the next section, several examples will be given where we cast existing mixing models in the proposed stochastic process framework.

III. EXAMPLES

A. Linear Mixing Model

In the LMM, one assumes the presence of p endmember spectra $\{e_i\}_{i=1}^{p}$, and an observation model where the spectrum is represented as a d-dimensional vector, representing the reflectance at specific wavelength intervals. The LMM is then given by

$$x = \sum_{i=1}^{p} a_i e_i, \quad \begin{cases} \sum_i a_i = 1 \\ \forall i : a_i \geq 0 \end{cases} \quad (5)$$

where the abundances $\{a_i\}_{i=1}^{p}$ are positive and sum to one (convexity constraints). In the proposed description, this corresponds to a state space $S = \{s, e_1, e_2, \ldots, e_p, o\}$, with s the light source, e_i endmember i, and o the observer. The operators associated with these states are

1) $T(s)x = 1$.
2) $T(e_i)x = e_i x$.
3) $T(o)x = x$.

The transition probabilities between states p_{ij} can be derived from the graph in Fig. 1, which describes the states and possible transitions in the LMM for a two-endmember scenario. From this graph, it is clear that each path must originate from the source s. The next state will be one of the endmembers e_i with probability a_i. After reaching an endmember, the next state is the observer state o, where the path finishes. There are, hence, p unique paths of path length one (see also Table I), and the weighted summation over all these paths immediately leads to mixing equation (5) and its constraints.

TABLE I

POSSIBLE PATHS IN THE LMM, THEIR PROBABILITIES,
AND THE CONTRIBUTION TO THE TOTAL SPECTRAL
SIGNAL OF EACH PATH

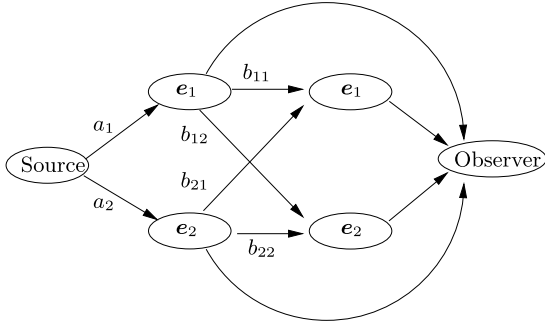| Path	Probability	Contribution
$s - e_i - o$	a_i	e_i

Fig. 2. States and transition probabilities in a subset of bilinear models. Unmarked edges have an implicit probability of one.

TABLE II

SEVERAL POPULAR BILINEAR MODELS, AND THEIR
PARAMETERS FOR THE GRAPH IN FIG. 2

Name	Parameters
Fan	$\forall i \geq j : b_{ij} = 0$ $\forall i < j : b_{ij} = a_j$
GBM	$\forall i \geq j : b_{ij} = 0$ $\forall i < j : b_{ij} = \gamma_{ij} a_j$
PPNM	$\forall i, j : b_{ij} = b a_j$

B. Bilinear Models

In bilinear models, the LMM is extended with second-order interactions, representing secondary illumination. Many bilinear models have been proposed in the recent literature, and several of them can be represented by a single generic graph, shown in Fig. 2 for a two-endmember scenario. Two layers of endmembers have been employed to accommodate both first- and second-order interactions. Light rays coming from the source first go to the endmembers in the first layer, with probabilities equal to their abundances. Next, two paths are followed simultaneously: The light ray goes to the observer with probability one but also goes to the endmembers in the second layer with probabilities dependent on the model. At each interaction with an endmember, the light ray is attenuated by the reflectance value of the endmember.

The values for the generic transition probabilities $\{b_{ij}\}_{ij}$ are given in Table II for several well-known bilinear models. These are the Fan model [11], the generalized bilinear model (GBM) [12], [34], [35], and the polynomial post-nonlinear model (PPNM) [14]. Note that several additional constraints are present, such as the ANC and ASC on the abundances $\{a_i\}_i$, and possible constraints on the additional free parameters.

Note that the sum of outgoing probabilities in the endmember nodes in the first layer is larger than or equal to one, and often significantly so. This causes the modeled spectra to have larger reflectances than those obtained by the LMM. Furthermore, these models often contain asymmetries, such as light rays that are allowed from e_1 to e_2 but not from e_2 to e_1, or missing self-interactions. Several of these issues can

be easily resolved by reformulating the mixing equations with symmetrized equations, or by deriving new models with slight changes, such as allowing self-interactions or inclusion of a sum-to-one constraint on outgoing probabilities.

Furthermore, several other bilinear models exist that cannot be represented by the diagram in Fig. 2, such as the model proposed by Nascimento and Bioucas-Dias [8], Somers *et al.* [9], Halimi *et al.* [12], Chen and Vierling [15], Meganem *et al.* [16], or Qu *et al.* [17]. For each of these models, a graph representation can be created as well.

C. Layered Models

In the past, several layered models have been proposed for hyperspectral mixing [18]. Such models may also be described in the proposed framework. Consider, for instance, the model proposed by Borel and Gerstl [18], which models a vegetation canopy as a single layer with a certain leaf area index (LAI) above soil, and a light source that emits radiation E_0. By using radiosity theory, three fluxes of light are defined as follows.

1) B_1: Upwelling radiation from the leaves in the canopy layer, which consists of reflected downwelling source radiation, and upwelling soil radiation that is transmitted through these leaves.
2) B_2: Downwelling radiation from the leaves, consisting of transmitted downwelling source radiation, and reflected upwelling soil radiation.
3) B_3: Upwelling radiation from the soil, consisting of reflected source radiation and reflected downwelling radiation from the leaves.

This leads to the following set of coupled equations. Let τ be the transmittance of the leaves in the canopy, e_l and e_s be the reflectance of the leaves and soil, respectively, and a be the LAI of the leaves in the canopy layer

$$B_1 = a e_l E_0 + a \tau B_3 \tag{6}$$

$$B_2 = a \tau E_0 + a e_l B_3 \tag{7}$$

$$B_3 = (1-a) e_s E_0 + e_s B_2. \tag{8}$$

By defining an upside and downside of a leaf as L_u and L_d, respectively, and soil as S, this model can be described with a directed graph (see Fig. 3). Note that the sum-to-one constraint on outgoing edge probabilities in the leaf vertices does not apply here. When a leaf is reached, both the transmitted and reflected paths have to be considered, leading to a branching into two paths, both with probability one. All possible paths, and their probabilities and contributions, are given in Table III.

After grouping similar terms together, this leads to the total weighted sum

$$\boldsymbol{x} = a e_l + \sum_{n=0}^{\infty} (a e_s e_l)^n ((1-a) e_s (1 - a + a\tau))$$

$$+ \sum_{n=0}^{\infty} (a e_s e_l)^n a e_s \tau (1 - a + a\tau)$$

$$= a e_l + \frac{(1 - a + a\tau)^2 e_s}{1 - a e_s e_l}. \tag{9}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

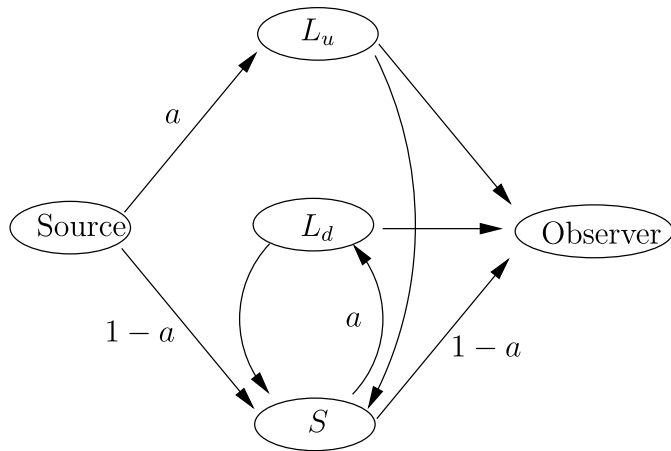HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS

5

Fig. 3. States and transition probabilities in the layered model. Unmarked edges have an implicit probability of one.

TABLE III

POSSIBLE PATHS UP TO LENGTH 4, THEIR PROBABILITIES AND CONTRIBUTIONS IN THE LAYERED MODEL FROM [18]. ALL RAYS ORIGINATE FROM THE SOURCE s AND END IN THE OBSERVER STATE o, THESE ARE OMITTED FOR BREVITY

Path	Probability	Contribution
L_u	a	e_l
S	$(1-a)(1-a)$	e_s
$L_u - S$	$a(1-a)$	τe_s
$S - L_d$	$(1-a)a$	$e_s\tau$
$L_u - S - L_d$	a^2	$\tau e_s\tau$
$S - L_d - S$	$(1-a)a(1-a)$	$e_s e_l e_s$
$L_u - S - L_d - S$	$aa(1-a)$	$\tau e_s e_l e_s$
$S - L_d - S - L_d$	$(1-a)aa$	$e_s e_l e_s\tau$

This solution is identical to the off-hotspot bidirectional reflectance distribution function provided in [18] [(13) and (2)], up to a multiplicative constant π.

D. Intimate Mixing Models

Some intimate mixing models can be cast into the proposed framework as well, and as an example, we consider the Shkuratov model [21]. The derivation of this model consists of several steps. Consider a parallel light beam falling onto an infinite half-space, partly filled with a homogeneously distributed particulate medium. We make a distinction between forward and backward scattering, and consider only the average optical properties of particles (see [21] for detailed calculations of these averages). First, we want to derive the optical properties of a single particle. We define the following.

1) T_e and T_i are the average transmittances of light penetrating into a particle from the outside, and exiting a particle from the inside, respectively.
2) R_i is the average coefficient of internal reflection.
3) R_b and R_f are the average backward and forward reflectance coefficients for light reflecting off the outside of the particle.
4) W_m is the probability of the light beam emerging backward at the mth interaction, and $(1 - W_m)$ is the probability of emerging forward. An interaction is
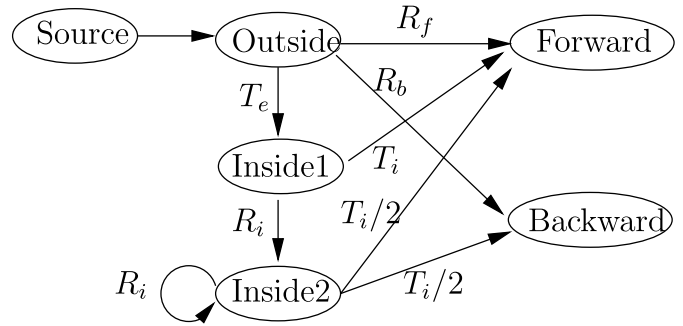

Fig. 4. States and transition probabilities in the Shkuratov model for a single particle.

TABLE IV

POSSIBLE PATHS, THEIR PROBABILITIES, AND CONTRIBUTIONS IN THE SHKURATOV MODEL FOR A SINGLE PARTICLE

Path	Probability	Contribution
OF	R_f	1
OB	R_b	1
O I_1 F	$T_e T_i$	$e^{-\tau}$
O I_1 I_2 F	$T_e R_i T_i/2$	$e^{-2\tau}$
O I_1 I_2 B	$T_e R_i T_i/2$	$e^{-2\tau}$
O I_1 I_2 I_2 F	$T_e R_i^2 T_i/2$	$e^{-3\tau}$
O I_1 I_2 I_2 B	$T_e R_i^2 T_i/2$	$e^{-3\tau}$
O I_1 I_2 I_2 I_2 F	$T_e R_i^3 T_i/2$	$e^{-4\tau}$
O I_1 I_2 I_2 I_2 B	$T_e R_i^3 T_i/2$	$e^{-4\tau}$

defined as reaching a particle interface, either from the inside or the outside.

5) Attenuation of light inside the particle is a volumetric effect and can be described by an average optical density τ.

Some simplifications are considered in [21]. By defining $R_e = R_b + R_f$, we can write $T_e = 1 - R_e$ and $T_i = 1 - R_i$. Specifically, this means that a light beam will either reflect or transmit at a particle boundary, and that there is no absorption or attenuation at this interface. Furthermore, we consider that a light beam forgets its past, and that the probability of emerging backward or forward after internal reflections is equal: $\forall i > 2 : W_i = 1/2$. This changes for W_2, which describes a light beam that enters and exits a particle without internal reflections. Such a light beam has zero probability of emerging backward, hence $W_2 = 0$.

The reflection/transmission process can be described by a directed graph, as shown in Fig. 4. The probabilities are indicated on each vertex and obey the sum-to-one constraint. There are two "Inside" states defined, to accommodate light going straight through a particle (the $W_2 = 0$ exception). Note that this diagram can be simplified significantly if one also considers $W_2 = 1/2$, as the "Inside1" state is then no longer required. The "Outside" state does not alter the optical properties, while each "Inside" state represents an internal traversal of the particle, and will attenuate the light by a factor $e^{-\tau}$, with τ the average optical density of the particles. There are also two observer states, one for forward and one for backward scattering.

The possible paths, and their probabilities and contributions, are given in Table IV. From this table, we can derive that the total reflectance observed by the forward and backward

6

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
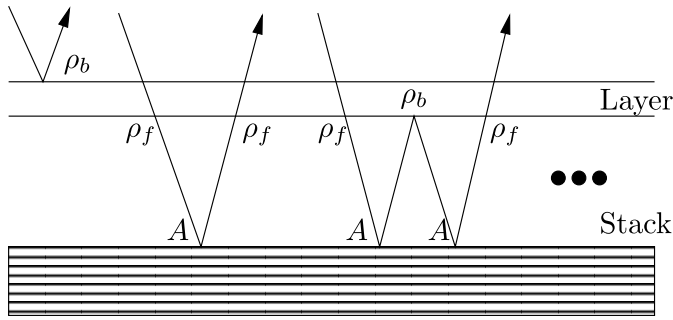

Fig. 5. Different orders of reflections in the Shkuratov model, with a reflection off the upside of the layer, and the first- and second-order reflections between the layer and the stack.

observers is given by

$$r_f = R_f + T_e T_i e^{-\tau} + \frac{T_e R_i T_i e^{-2\tau}}{2}$$

$$+ \frac{T_e R_i^2 T_i e^{-3\tau}}{2} + \frac{T_e R_i^3 T_i e^{-4\tau}}{2} + \cdots$$

$$= R_f + T_e T_i e^{-\tau} + \frac{T_e T_i R_i e^{-2\tau}}{2(1 - R_i e^{-\tau})} \quad (10)$$

$$r_b = R_b + \frac{T_e R_i T_i e^{-2\tau}}{2} + \frac{T_e R_i^2 T_i e^{-3\tau}}{2} + \cdots$$

$$= R_f + \frac{T_e T_i R_i e^{-2\tau}}{2(1 - R_i e^{-\tau})}. \quad (11)$$

These equations describe the forward and backward reflectances of a single particle and are equal to [21, eqs. (9a) and (9b)]. In the next step, one considers a homogeneous distribution of such particles in a half-space, with parallel incoming light. This setup is regarded as an infinite stack of layers of particles, where each layer has a thickness equal to the average particle diameter. If the volume fraction filled by particles is denoted by $q \in [0, 1]$, the probability that a light ray hits a particle when traversing a single layer is also given by q if the packing is random [21]. The forward and backward reflectances of a single layer are then given by

$$\rho_b = q r_b \quad (12)$$
$$\rho_f = q r_f + (1 - q). \quad (13)$$

If one considers a half-infinite stack of such layers, the total reflectance of this stack will not change if one layer is added or removed. This observation allows us to calculate the total reflectance (see Fig. 5). Note that transmission through a single layer corresponds with a factor ρ_f, reflection off a single layer with a factor ρ_b, and reflection off an infinite stack of layers with the reflectance A, which is the property to be determined

$$A = \rho_b + \rho_f^2 A + \rho_f^2 \rho_b A^2 + \rho_f^2 \rho_b^2 A^3 + \cdots \quad (14)$$

$$= \rho_b + \frac{\rho_f^2}{1 - \rho_b A}. \quad (15)$$

This leads to solution (12) found in [21], which we omit for brevity.
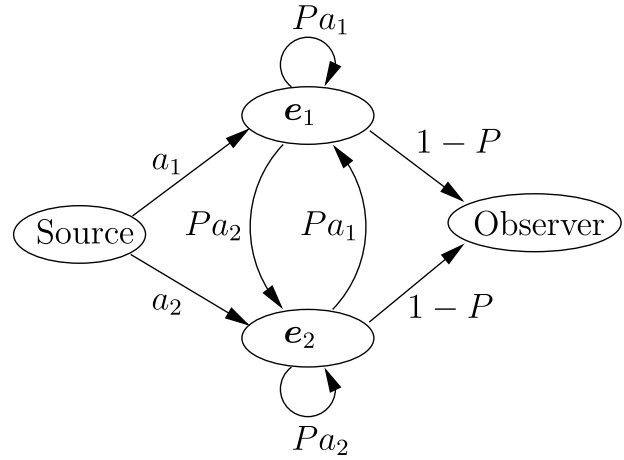

Fig. 6. States and transition probabilities in the MLM model for a two-endmember scenario.

E. Multilinear Mixing Model

Recently, we have proposed the MLM model [24] as an extension of bilinear models to include infinitely many interactions. This model has been derived by employing the proposed framework. We make the following assumptions.

1) A light ray incoming from the source will interact with at least one material.
2) After each interaction with a material, the ray will have a probability P of undergoing further interactions, and a probability $(1 - P)$ of escaping the scene and reaching the observer.
3) The probability of interacting with material i is proportional to its abundance a_i.
4) When a light ray is scattered by material i, its intensity changes according to that material's reflectance e_i.

This leads to the graphical model depicted in Fig. 6 for a two-endmember model, and the possible paths, probabilities and contributions, in Table V.

The probability for the sequence $(e_{i_1}, e_{i_2}, \ldots, e_{i_R})$ before reaching the observer is given by

$$\text{Prob}(e_{i_1}, e_{i_2}, \ldots, e_{i_R}) = (1 - P) P^{R-1} a_{i_1} a_{i_2} \ldots a_{i_R}. \quad (16)$$

The spectral contribution of such a path is $\prod_{j=1}^{R} e_{i_j}$. The resulting mixing equation then becomes

$$x = \sum_{R=1}^{\infty} \left(\sum_{i_1=1}^{p} \cdots \sum_{i_R=1}^{p} \right) (1 - P) P^{R-1} \prod_{k=1}^{R} (a_{i_k} e_{i_k})$$

$$= (1 - P) \sum_{i=1}^{p} a_i e_i + (1 - P) P \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j e_i \odot e_j$$

$$+ (1 - P) P^2 \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{p} a_i a_j a_k e_i \odot e_j \odot e_k \cdots$$

$$= \frac{(1 - P) \sum_{i=1}^{p} a_i e_i}{1 - P \sum_{i=1}^{p} a_i e_i}. \quad (17)$$

It is easy to see that the MLM model accommodates all orders of interactions between end members, and scales these

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS

7

TABLE V
POSSIBLE PATHS UP TO LENGTH 3, THEIR PROBABILITIES AND
CONTRIBUTIONS IN THE MLM MODEL. ALL RAYS ORIGINATE
FROM THE SOURCE s AND END IN THE OBSERVER
STATE o, THESE ARE OMITTED FOR BREVITY

Path	Probability	Contribution
e_i	$a_i(1-P)$	e_i
$e_i \, e_j$	$a_i Pa_j (1-P)$	$e_i e_j$
$e_i \, e_j \, e_k$	$a_i Pa_j Pa_k (1-P)$	$e_i e_j e_k$

TABLE VI
POSSIBLE PATHS UP TO LENGTH 3, THEIR PROBABILITIES
AND CONTRIBUTIONS IN THE EXTENDED MLM MODEL

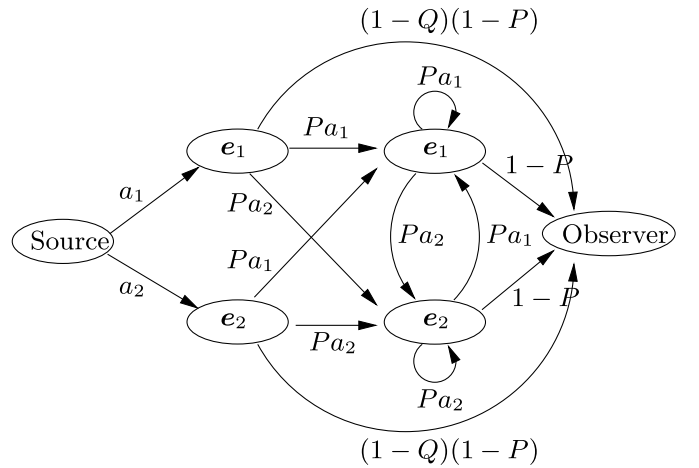| Path	Probability	Contribution
e_i	$(1-Q)a_i(1-P)$	e_i
$e_i \, e_j$	$a_i Pa_j (1-P)$	$e_i e_j$
$e_i \, e_j \, e_k$	$a_i Pa_j Pa_k (1-P)$	$e_i e_j e_k$

Fig. 7. Graphical model describing the SMLM model. By introducing two layers of endmembers, a distinction can be made between direct light rays and ones that underwent multiple reflections.

higher order interactions with a single parameter P with a clear physical interpretation (the probability of interacting again with an endmember). We refer to [24] for more details.

IV. SHADOW MODEL BASED ON GRAPH REPRESENTATION

A. Extending the MLM With Shadows

In this section, we extend the MLM model with the capability of dealing with shadows by employing the proposed framework, resulting in the shadow MLM (SMLM) model. As seen in Section III-E, the MLM model extends the LMM with the probability that a light ray undergoes additional interactions with endmembers after each interaction with a constant probability P (also see Fig. 6). This model has been shown to be able to handle higher order interactions. Also, shadows can play an important role in the composition of the scene and are commonly encountered in many urban and vegetation scenarios. These shadows can be partial, as only part of the IFOV of a pixel can be shadowed, or the geometrical structure which casts the shadow can be smaller than the IFOV (e.g., small trees and bushes). By definition, any object that is shadowed has no direct line of sight to the light source. Disregarding possible atmospheric illumination, this means that any light which is observed from a shadowed area underwent multiple reflections. When we consider the MLM model and its series expansion, this means that only interactions of order two and higher are allowed in a shadowed area, while the first-order interactions can only contribute in directly illuminated parts of the pixel. This intuitive reasoning leads to the path probability illustrated in Table VI, which equals Table V, but with the first-order paths rescaled with a shadow parameter $(1-Q)$, with $Q \in [0, 1]$. The value $Q = 0$ means no shadow, while $Q = 1$ indicates a situation with full shadow. This leads to the following mixing equation for the SMLM model

$$x = \frac{(1-P)\sum_{i=1}^{p} a_i e_i}{1 - P \sum_{i=1}^{p} a_i e_i} - Q(1-P) \sum_{i=1}^{p} a_i e_i. \quad (18)$$

A graphical model supporting these observations can be constructed by considering both the viewpoints of the light source and the observer: while the source cannot observe shadows by definition (this is the hotspot direction), the observer can. Every light ray that leaves the source will have the first interaction with some endmember, proportional to its abundance. After this first interaction, we consider three possibilities: the light ray goes directly to the observer, has another interaction with an endmember, or it is absorbed because the observer is not visible. We assign a variable $Q \in [0, 1]$ to this latter probability. Typically, the distances to source and observer are much larger than the geometrical sizes of the objects in the scene, and one can consider parallel illumination and observation. In this case, the size of the illuminated part of the scene, not visible to the observer due to occlusion, is equal to the size of the shadowed fraction of the scene (this can be easily seen by reversing the roles of source and observer). This indicates that the Q variable plays the role of shadow weight, indicating the size of the shadow component in the scene. The corresponding graphical model is depicted in Fig. 7.

V. EXPERIMENTS

To illustrate the performance of the SMLM model, we have unmixed several artificial and a real-world data set. The results are evaluated quantitatively by employing the reconstruction error (RE), given by the Fröbenius norm between the reconstructed data set $Y = \{Y_j\}_{j=1}^{N}$ and the input data set $x = \{x_j\}_{j=1}^{N}$, and the average absolute abundance error (AE)

$$\text{RE} = \frac{1}{N} \sqrt{\sum_{j=1}^{N} (\mathbf{X}_j - \mathbf{Y}_j)^2} \quad (19)$$

$$\text{AE}(X - Y) = \frac{1}{pn} \sum_{j=1}^{N} \sum_{i=1}^{p} |a_{ji}(\mathbf{X}) - a_{ji}(\mathbf{Y})|. \quad (20)$$

We compared the results with two alternatives.

1) The first algorithm we used for comparison is the standard fully constrained LMM, with the ANC and ASC imposed on the abundances. The mixing equation for the LMM is given by (5).

2) As we observed that the REs from the LMM model are sometimes relatively large, we also considered the LMM with an artificial shadow endmember included [36]. This is accomplished by extending the endmember matrix with a column of zeros, and results in the shadow LMM (SLMM) model. It is easy to show that the SLMM is related to nonnegatively constrained least-squares (NCLS) unmixing, with the exception that all abundances in the SLMM are restricted to the interval [0, 1], while in the NCLS model, the endmember abundances would be nonnegatively constrained, and negative shadow contributions would be allowed. The mixing equation for the SLMM model is equivalent to the artificial addition of a zero-vector shadow endmember to the LMM

$$
\boldsymbol{x} = (1 - Q) \sum_{i=1}^{p} a_i \boldsymbol{e}_i, \quad
\begin{cases}
\sum_i a_i = 1 \\
\forall i : a_i \geq 0 \\
Q \in [0, 1]
\end{cases}
\tag{21}
$$

$$
= \sum_{i=1}^{p} \hat{a}_i \boldsymbol{e}_i + \hat{a}_{p+1} \overline{0}, \quad
\begin{cases}
\sum_{i=1}^{p+1} \hat{a}_i = 1 \\
\forall i : \hat{a}_i \geq 0
\end{cases}
$$

$$
\text{with }
\begin{cases}
\forall i \in 1, \ldots, p : \hat{a}_i = (1 - Q)a_i \\
\hat{a}_{p+1} = Q = 1 - \sum_{i=1}^{p} \hat{a}_i.
\end{cases}
\tag{22}
$$

Equation (21) can be unmixed by employing constrained optimization, while (22) can be unmixed with the fully constrained least-squares unmixing (FCLSU) algorithm [37], and transformed into the equivalent interpretation from (21).

3) The third algorithm we used for comparison is the MLM model of Section III-E. The mixing equation for the MLM model is given by (17).

The mixing equation for the proposed SMLM model is given by (18). Furthermore, note that the $P \to 0$ limit of the SMLM model equals the SLMM model, the $Q \to 0$ limit gives the MLM model, and the $P, Q \to 0$ limit gives the LMM model.

For LMM and SLMM, we employed FCLSU, based on the NCLS algorithm of Lawson and Hanson [38], and for the MLM and SMLM model, we performed constrained optimization of the RE versus the model parameters via sequential quadratic programing.

A. Artificial Data Sets

We have first employed different algorithms on artificial data sets that were generated by the mixing models themselves. This allows us to assess the performance in an ideal scenario, where all informations such as endmember spectra, abundance vectors, and metadata, are known, and can be used to generate quantitative measurements. As all the alternatives for the SMLM model are limiting cases, these experiments also allow us to assess the importance of adding higher order interactions, shadowing, or both, in the LMM.

TABLE VII

RE FOR THE ARTIFICIAL UNMIXING EXPERIMENT, WITH $p = 10$ ENDMEMBERS, $N = 10^4$ PIXELS, AND SNR = 50 (TOP HALF) AND SNR = 200 (BOTTOM HALF)

SNR 50	LMM data	SLMM data	MLM data	SMLM data
LMM algo	3.885	24.844	13.787	29.188
SLMM algo	3.876	3.028	4.081	4.129
MLM algo	3.871	3.069	3.860	3.878
SMLM algo	**3.869**	**3.013**	**3.854**	**3.848**
SNR 200	LMM data	SLMM data	MLM data	SMLM data
LMM algo	0.7945	7.492	4.709	9.232
SLMM algo	0.7916	0.616	1.005	1.055
MLM algo	0.7913	0.633	0.788	0.793
SMLM algo	**0.7905**	**0.613**	**0.786**	**0.783**

TABLE VIII

AVERAGE ABSOLUTE AEs FOR THE ARTIFICIAL UNMIXING EXPERIMENT, WITH $p = 10$ ENDMEMBERS, $N = 10^4$ PIXELS, AND SNR = 50 (TOP HALF) AND SNR = 200 (BOTTOM HALF)

SNR 50	LMM data	SLMM data	MLM data	SMLM data
LMM algo	**0.0355**	0.0888	0.0675	0.1080
SLMM algo	0.0406	**0.0462**	0.0561	0.0645
MLM algo	0.0416	0.0623	**0.0453**	0.0574
SMLM algo	0.0431	0.0508	0.0473	**0.0564**
SNR 200	LMM data	SLMM data	MLM data	SMLM data
LMM algo	**0.0084**	0.0705	0.0495	0.0959
SLMM algo	0.0119	**0.0160**	0.0272	0.0337
MLM algo	0.0122	0.0336	**0.0164**	0.0283
SMLM algo	0.0126	0.0196	0.0179	**0.0250**

To generate the data sets, p endmember spectra were randomly selected from the United States Geological Survey spectral library of minerals.[1] Next, N abundance vectors $\mathbf{a}_j = (a_{j1}, a_{j2}, \ldots, a_{jp})^T (j = 1, \ldots, N)$ were randomly generated, uniformly drawn from a unit simplex and hence obeying the ANC and ASC. The shadow parameter Q and the nonlinearity parameter P were both independently drawn from a half-normal distribution with $\sigma = 0.3$. Values larger than one were set to zero. The motivation here is that one typically does not expect many pixels with very large contributions from higher order interactions or that are fully shadowed.

After parameter selection, the mixing models themselves are used to generate the data sets [(5), (21), (17), and (18) for the LMM, SLMM, MLM, and SMLM models, respectively], and the Gaussian noise with a given signal-to-noise ratio (SNR) is added. These data sets are then unmixed using different mixing models. The RE and the AE can be found in Tables VII and VIII, respectively, for several values of the SNR.

These tables show that the SMLM model always gives the smallest RE, regardless of the mixing model employed. This is to be expected, as the SMLM model has the highest number of free parameters, and the other models are limiting cases of the SMLM model. Furthermore, the REs obtained by the MLM model are close to those obtained by the SMLM model, even on data sets that include shadowing. This suggests that higher order interactions and shadowing are treated the same way by the MLM model, by reducing all orders of interactions instead of only the first order as in the SMLM model.
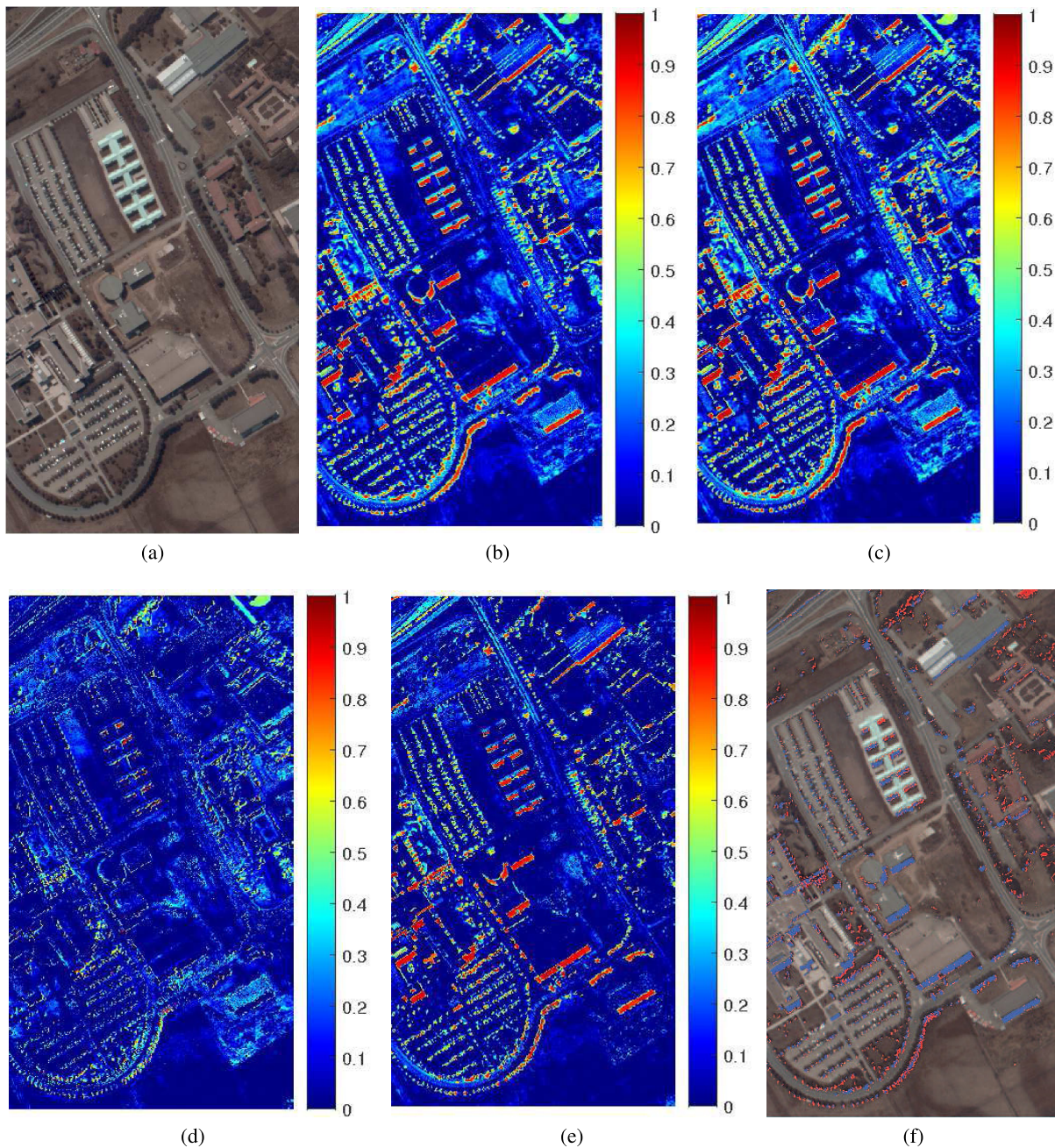
---

[1] https://speclab.cr.usgs.gov/spectral-lib.html

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS

9

Fig. 8. (a) False color RGB image. (b) Shadow abundance values from the SLMM model. (c) Nonlinearity P-values from the MLM and (d) SMLM model. (e) Q-values from the SMLM model, for the Pavia data set (all parameters range from 0 to 1 and are visualized using a colorcode, ranging from dark blue to dark red). (f) Thresholded P (in red) and Q-values (in blue) of the SMLM model overlaid over the RGB image.

From the AE table, one can conclude that each model excels at unmixing its own data set, achieving the smallest AEs. For linearly mixed data, the proposed method performs slightly worse than SLMM and MLM, while in the other cases, it is the best alternative to the model that generated the data. Also, obtaining comparably small REs does not mean that the AEs will also be small.

B. Real Data Set—Pavia Image

The data set that was employed in this experiment was the Pavia University data set. The Pavia University data set is captured by the ROSIS sensor and is a 103-band data set in the range of 430–838 nm [39], with a spectral resolution of 4 nm and 610 by 340 pixels. It depicts part of the campus of Pavia University, Italy. Note that 59 pixels were removed from the Pavia University data set as they contained improper signals. A false color Red Green Blue (RGB) image of the Pavia image is shown in the first column of Fig. 8, by using bands 40, 20, and 10 for the R, G, and B images, respectively.

The following experiment is conducted on the Pavia image. For this image, ground truth pixels are available for eight material classes and a shadow class. In order to evaluate the mixing models for their ability to model shadow, we generate one endmember for each of the eight material classes, as the average spectrum of all training pixels corresponding to that material. No shadow endmember is generated, since we want

to validate to what extent the shadow is treated by the mixing models. In the case of the SLMM model, an additional shadow endmember is defined as the zero vector, in the case of SMLM, the Q-value should take the shadow into account.

After endmember selection, the data sets were unmixed with the respective models. The SLMM, MLM, and SMLM models also produce additional metadata. The MLM and SMLM models contain a P-value for every pixel, indicating the probability of higher order interactions, while the SLMM and SMLM models contain a variable describing the level of shadow in the pixel (this is the Q-value of the SMLM model, and the abundance of the artificial shadow endmember in the SLMM model). These maps are shown in Fig. 8.

The values of the metaparameters range from 0 to 1 and are shown in a colormap ranging from dark blue to dark red. We also include a figure showing for the proposed method, in red, all pixels having P-values larger than 0.5, and in blue, all pixels having Q-values larger than 0.6, overlaid on top of the RGB image.

From visual inspection, the following observations can be made.

1) In general, the shadow abundance map of SLMM is very similar to the P-map of MLM, indicating that both models use their respective parameter to model both shadow and multiple reflections. At the same locations, the SMLM model generates either P- or Q-values or both.

2) The highest values of the parameters are observed in shadowed areas from buildings. In these areas, the shadow abundance values of SLMM and the P-values of MLM are very similar and close to 1. The SMLM model, on the other hand, shows high values of Q but P is zero. These areas are expected to be partially lit by higher order reflections, but the spectral reflectance is very low and all the models seem to overcompensate for this with shadow.

3) However, for partially shadowed pixels at the transitions between these building shadows and fully lit areas, SMLM shows high P-values.

4) High values of P and Q also appear at regions near vegetation. In general, the tree canopies seem to be modeled linearly ($P = Q = 0$), while high Q-values appear in the shadows of these trees. Multiple reflections become visible in partially shadowed pixels at the edges of these shadows.

5) In some vegetation-related regions, the SMLM model seems to show more multiple reflections than shadow. This may indicate low and/or less dense vegetation.

These results are to be expected, as higher order interactions are mainly expected in or near complex 3-D structures such as trees, and cause light to be received from partially shadowed areas. This means that in the MLM model, high P-values can be expected wherever the spectrum is significantly smaller than the endmembers, whether due to higher order interactions or shadowing. The SLMM model also mixes up between shadowing and multiple interactions. The SMLM model however can make this distinction, and

TABLE IX
AVERAGE ABSOLUTE ABUNDANCE DIFFERENCES BETWEEN THE PROPOSED AND THE OTHER MODELS, FOR DIFFERENT GROUPS OF PIXELS

Group	AE (SMLM-MLM)	AE (SMLM-SLMM)	AE (SMLM-LMM)
A	1.5e-05	2.2e-06	1.5e-05
B	1.5e-02	3.5e-02	1.0e-01
C	1.0e-06	2.1e-02	1.1e-01
D	8.5e-03	5.6e-02	1.5e-01
All	4.3e-03	1.8e-02	6.7e-02
Shadow	2.5e-02	1.0e-01	1.7e-01

shows differences between the effects of higher order interactions and shadowing.

For a more quantitative study of the behavior of the metaparameters P and Q from the SMLM model, we subdivided all pixels into four different groups, according to the obtained values for P and Q. The different groups are as follows.

1) *Group A ($P = Q = 0$):* This group contains 77.910 pixels out of a total of 207.400 pixels. It is verified that these are exactly the pixels for which P of MLM equals zero, and almost identical to the group of pixels where the shadow abundance of SLMM equals zero (total of 77.677 pixels). For this group of pixels (let us refer to them as linearly mixed pixels), the obtained abundance values for all models are very similar (AE of 10^{-5}, see Table IX).

2) *Group B ($P = 0$ and $Q \neq 0$):* This is the group that is characterized by our model as (pure or partially) shadowed pixels. This group contains 53.781 pixels. For these pixels $P \neq 0$ in the MLM model, while the obtained Q-values are almost identical to the shadow abundance values in the SLMM model. However, the differences in obtained abundance values with the SLMM model are larger than with the MLM model.

3) *Group C ($P \neq 0$ and $Q = 0$):* This group contains 65.158 pixels. This is the group that is characterized by our model as pixels containing multiple reflections, but without being shadow pixels. For these pixels, the shadow abundances in the SLMM model differ from 0. The obtained P-values are identical to the P-values of the MLM model. For these pixels, the obtained abundances are almost identical to the abundances obtained by the MLM model.

4) *Group D ($P \neq 0$ and $Q \neq 0$):* This group contains 10.551 pixels. This is the group that is regarded by our model as (partially) shadowed pixels containing multiple reflections. For these pixels, we have verified that SLMM compensates for this by larger values of the shadow abundances, while MLM compensates for this by larger values of P. For this group, the obtained abundances by our method are closest to the ones obtained by MLM. To verify whether our model provides more correct abundances, actual ground truth data of fractional abundances of materials under shadowed pixels would be required.

Overall, over all groups of pixels, the differences in abundance values are the smallest between our model and MLM.
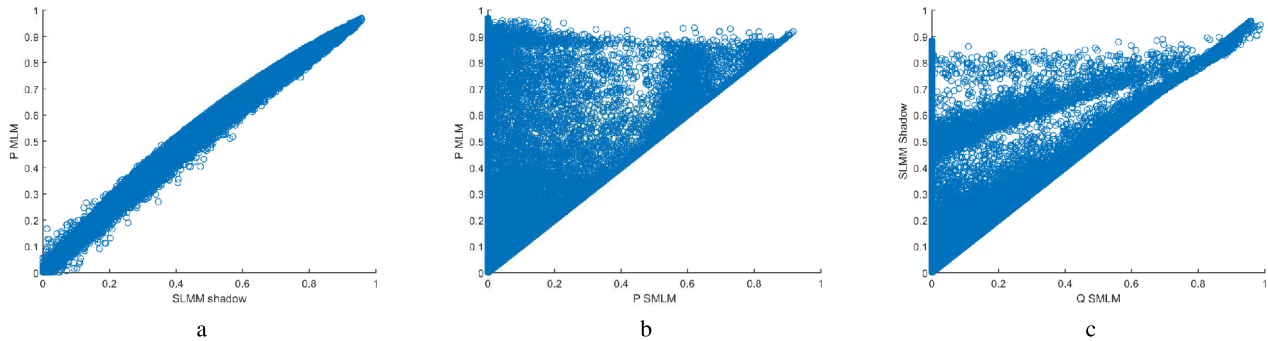
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS

11

Fig. 9. Scatterplots from the Pavia image of (a) shadow abundance of SLMM versus the P-value of MLM, (b) P-value of SMLM versus the P-value of MLM, and (c) Q-value of SMLM versus the shadow abundance of SLMM.

TABLE X

RUNTIME IN SECONDS FOR THE PAVIA DATA SET

Algorithm	time$_{Pavia}$ (s)
LMM	118
SLMM	144
MLM	2790
SMLM	4260

Scatterplots of the parameters, as shown in Fig. 9, confirm these findings. It can be observed from Fig. 9(a) that indeed the values of P of MLM correspond to the shadow abundance values of SLMM. Moreover, Fig. 9(b) shows that the P-values of SMLM are always equal to or lower than the ones of MLM. We verified that all pixels in group C lie exactly on the diagonal. Finally, Fig. 9(c) shows that the Q-values of SMLM are always equal to or lower than the shadow abundance values of SLMM (except for the values close to 1), we also verified that all pixels in group B lie exactly on the diagonal.

A particular group of pixels is formed by the shadow class. We performed a basic classification using an support vector machine classifier based on all training pixels of the Pavia image, including shadow as a class. The classifier has identified 10.593 pixels as shadow pixels. From these pixels, we found that none belong to group A, 7.426 to group B, 898 to group C, and 2.269 to group D. The large majority of the pixels of the shadow class are characterized by our method as shadowed pixels with (D) or without (B) multiple reflections. The observed abundance differences between the models are larger than was the case for any of different groups. All algorithms were developed in MATLAB and run on an Intel Core $i7 - 6700$ K CPU, 4.00 GHz machine with 4 cores. The runtimes of the algorithms on the Pavia image are shown in Table X. As can be seen, the runtimes increase with the increase in model complexity from LMM to SLMM to MLM to SMLM. Note that the runtimes for the multilinear models are relatively high due to the use of sequential quadratic programing and can be significantly improved by employing dedicated optimization schemes, such as, for instance, the alternating direction method of multipliers technique in [40] for unmixing with the MLM model.

VI. CONCLUSION

In this paper, we have introduced a general method of interpreting the optical interactions that lead to spectral mixing models, based on graphical models. By performing a weighted summation over all possible paths in such a graphical representation, closed-form equations for the resulting mixing model can often be obtained. This is illustrated by deriving several popular mixing models in this framework with varying levels of complexity, such as the LMM, the MLM, the layered model by Borel and Gerstl, and the Shkuratov model for intimate mixing. These illustrations show the flexibility of this interpretation. Furthermore, new mixing models can be easily derived within this framework. We give a demonstration where we extend the existing MLM model with shadows, defined as regions in the image that receive no direct illumination from the source. The resulting mixing model is demonstrated on artificial and real data sets and is capable of correctly identifying shadowed areas. Future work concerns the derivation of more powerful and flexible mixing models within this framework. Possible examples are the inclusion of skylight illumination (i.e., diffusive radiation from the atmosphere) in the presented shadow model, radiation arising from multiple reflections between neighboring pixels, point spread functions, or the adjacency effect, and variability in the endmember spectra.

REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[2] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[3] M. A. Veganzones *et al.*, "A new extended linear mixing model to address spectral variability," in *Proc. IEEE Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Lausanne, Switzerland, Jun. 2014, pp. 1–4.

[4] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE Workshop Hyperspectral Image Signal Process. (Whispers)*, Reykjavik, Iceland, Jun. 2010, pp. 1–4.

[5] B. Hapke, "Bidirectional reflectance spectroscopy: 1. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, Apr. 1981.

[6] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.

[7] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.

[8] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," *Proc. SPIE*, vol. 7477, p. 74770I, Sep. 2009.

[9] B. Somers *et al.*, "Nonlinear Hyperspectral Mixture Analysis for tree cover estimates in orchards," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1183–1193, Jun. 2009.

[10] R. B. Singer and T. B. McCord, "Mars—Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," in *Proc. 10th Lunar Planetary Sci. Conf.*, 1979, pp. 1835–1848.

[11] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. J. Remote Sens.*, vol. 30, no. 11, pp. 2951–2962, 2009.

[12] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[13] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a polynomial post nonlinear model for hyperspectral imagery," in *Proc. IEEE ICASSP*, May 2011, pp. 1009–1012.

[14] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3017–3025, Jun. 2012.

[15] X. Chen and L. Vierling, "Spectral mixture analyses of hyperspectral data acquired using a tethered balloon," *Remote Sens. Environ.*, vol. 103, no. 3, pp. 338–350, 2006.

[16] I. Meganem, P. Déliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear–quadratic mixing model for reflectances in urban environments," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 544–558, Jan. 2014.

[17] Q. Qu, N. Nasrabadi, and T. D. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4404–4423, Jul. 2014.

[18] C. C. Borel and S. A. W. Gerstl, "Nonlinear spectral mixing models for vegetative and soil surfaces," *Remote Sens. Environ.*, vol. 47, no. 3, pp. 403–416, 1994.

[19] T. Hiroi and C. M. Pieters, "Effects of grain size and shape in modeling reflectance spectra of mineral mixtures," *Proc. Lunar Planetary Sci.*, vol. 22, 1992, pp. 313–325.

[20] M. Kinoshita and M. Miyamoto, "A model for analysis of the spectral reflectance of mineral mixtures," in *Proc. Symp. Antarctic Meteorites*, vol. 3, 1990, pp. 230–239.

[21] Y. Shkuratov, L. Starukhina, H. Hoffmann, and G. Arnold, "A model of spectral albedo of particulate surfaces: Implications for optical properties of the moon," *Icarus*, vol. 137, no. 2, pp. 235–246, 1999.

[22] A. Marinoni, J. Plaza, A. Plaza, and P. Gamba, "Nonlinear hyperspectral unmixing using nonlinearity order estimation and polytope decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2644–2654, Jun. 2015.

[23] A. Marinoni, A. Plaza, and P. Gamba, "Harmonic mixture modeling for efficient nonlinear hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4247–4256, Sep. 2016.

[24] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 240–251, Jan. 2016.

[25] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, Jan. 2013.

[26] R. Ammanouil, A. Ferrari, C. Richard, and S. Mathieu, "Nonlinear unmixing of hyperspectral data with vector-valued kernel functions," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 340–354, Jan. 2017.

[27] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Nonlinear spectral unmixing of hyperspectral images using Gaussian processes," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2442–2453, May 2013.

[28] R. Heylen and P. Scheunders, "A distance geometric framework for nonlinear hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1879–1888, Jun. 2014.

[29] Z. Mitraka, F. Del Frate, and F. Carbone, "Nonlinear spectral unmixing of Landsat imagery for urban surface cover mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 3340–3350, Jul. 2016.

[30] J. Li, X. Li, B. Huang, and L. Zhao, "Hopfield neural network approach for supervised nonlinear spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 1002–1006, Jul. 2016.

[31] S. Douté and B. Schmitt, "A multilayer bidirectional reflectance model for the analysis of planetary surface hyperspectral images at visible and near-infrared wavelengths," *J. Geophys. Res.*, vol. 103, no. E13, p. 31367–31390, 1998.

[32] M. Shepard and P. Helfenstein, "A laboratory study of the bidirectional reflectance from particulate samples," *Icarus*, vol. 215, no. 2, pp. 526–533, Oct. 2011.

[33] M. Shepard and P. Helfenstein, "A test of the Hapke photometric model," *J. Geophys. Res.*, vol. 112, no. E3, 2007, Art. no. E03001.

[34] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unmixing hyperspectral images using the generalized bilinear model," *Proc. IEEE IGARSS*, Jul. 2011, pp. 1886–1889.

[35] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Bilinear models for nonlinear unmixing of hyperspectral images," in *Proc. IEEE Workshop Hyperspectral Image Signal Process. (Whispers)*, Lisbon, Portugal, Jun. 2011, pp. 1–4.

[36] T. Roper and M. Andrews, "Shadow modelling and correction techniques in hyperspectral imaging," *Electron. Lett.*, vol. 49, no. 7, pp. 458–460, Mar. 2013.

[37] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[38] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1974.

[39] P. Gege, M. Bachmann, and S. Holzwarth, "Airborne remote sensing: Rosis bonn 2008 summary report," Inst. Photogramm., Univ. Stuttgart, Stuttgart, Germany, Tech. Rep. 2008ROSIS-CRT-Vaihingen, 2008. Accessed: Feb. 5, 2018. [Online]. Available: http://www.ifp.uni-stuttgart.de/dgpf/update2711%5C2008ROSIS_CRT_Vaihingen.pdf

[40] Q. Wei, M. Chen, J.-Y. Tourneret, and S. Godsill, "Unsupervised nonlinear spectral unmixing based on a multilinear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4534–4544, Aug. 2017.

Rob Heylen (M'10) received the B.S. degree, the M.S. degree, and the Ph.D. degree in physics, with a focus on statistical mechanics, from Katholieke Universiteit Leuven, Leuven, Belgium, in 2001, 2003, and 2008, respectively.

In 2009, he joined the Vision Laboratory, Department of Physics, University of Antwerp, Antwerp, Belgium, as a Post-Doctoral Researcher. In 2013, he joined the Computer and Information Science and Engineering Department, University of Florida, Gainesville, FL, USA. He was a FWO Post-Doctoral Fellowship with the University of Antwerp. He is currently a Post-Doctoral Researcher with the Medical Imaging Research Center, Department of Nuclear Medicine, Katholieke Universiteit Leuven. His research interests include hyperspectral image processing and computational physics.

Vera Andrejchenko (S'16) received the B.S. degree in computer engineering and information technologies from the University of Ss. Cyril and Methodius, Skopje, Macedonia, in 2009, and the M.S. degree in artificial intelligence from Katholieke Universiteit Leuven, Leuven, Belgium in 2013.

She was a Marie Curie Research Fellow with the Interfaculty Department of Geoinformatics, Z-GIS, University of Salzburg, Salzburg, Austria. In 2015, she joined the Vision Laboratory, Department of Physics, University of Antwerp, Antwerp, Belgium, as a Ph.D. Researcher. Her research interests include machine learning methods and inverse problems for classification and fusion of hyperspectral data.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

HEYLEN *et al.*: NONLINEAR HYPERSPECTRAL UNMIXING WITH GRAPHICAL MODELS
13

Zohreh Zahiri received the bachelor's degree in civil engineering from Shahid Ashrafi Esfahani University, Isfahan, Iran, in 2009, the M.S. degree in architectural conservation from the Isfahan University of Art, Isfahan, in 2013, and the Ph.D. degree in characterization of facade materials using close-range hyperspectral dataž from the Department of Civil Engineering, University College Dublin, Dublin, Ireland.

She is currently a Post-Doctoral Researcher with the Vision Laboratory, Department of Physics, University of Antwerp, Antwerp, Belgium. Her research interests include the application of hyperspectral imaging in civil engineering and related areas well as data analysis methods.

Mario Parente (M'05–SM'13) received the B.S. and M.S. *(summa cum laude)* degrees in telecommunication engineering from the University of Federico II of Naples, Naples, Italy, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA.

He was a Post-Doctoral Associate with Brown University, Providence, RI, USA. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, MA, USA. His research interests include combining physical models and statistical techniques to address issues in remote sensing of Earth and planetary surfaces, identification of ground composition, geomorphological feature detection and imaging spectrometer data modeling, reduction and calibration for NASA missions, the study of physical models of light scattering in particulate media. He developed machine learning algorithms for the representation and processing of hyperspectral data based on statistical, geometrical, and topological models. Furthermore, he has developed solutions involving the integration of color and hyperspectral imaging and robotics for the identification of scientifically significant targets for rover and orbiter-based reconnaissance. He has supported several scientific teams in NASA missions such as the Compact Reconnaissance Imaging Spectrometer for Mars, the Mars Mineralogy Mapper (M3), and the Mars Science Laboratory ChemCam science teams. He is currently a Principal Investigator with the SETI Institute's Carl Sagan Center for Search for Life in the Universe and a member of the NASA Astrobiology Institute.

Dr. Parente serving as an Associate Editor for IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Paul Scheunders (SM'98) received the B.S. degree and the Ph.D. degree in physics, with a focus on statistical mechanics, from the University of Antwerp, Antwerp, Belgium, in 1983 and 1990, respectively.

In 1991, he joined the Vision Laboratory, Department of Physics, University of Antwerp, as a Research Associate, where he is currently a Full Professor. He has authored or co-authored over 200 papers in international journals and proceedings in the field of image processing, pattern recognition, and remote sensing. His research interests include remote sensing and hyperspectral image processing.

Dr. Scheunders is a Senior Member of the IEEE Geoscience and Remote Sensing Society. He served as an Program Committee Member in numerous international conferences. He is an Associate Editor of the IEEE TRANSACTIONS IN GEOSCIENCE AND REMOTE SENSING.