Improved deconvolution of mineral reflectance spectra

Ronan Rialland, Charles Soussen, Rodolphe Marion, Véronique Carrère

Improved deconvolution of mineral reflectance spectra

Ronan Rialland, Charles Soussen, *Member, IEEE,* Rodolphe Marion, and Véronique Carrère

***Abstract*—Reflectance spectroscopy is a widely used technique for mineral identification and characterization. Since modern airborne and satellite-borne sensors yield an increasing number of hyperspectral data, it is crucial to develop unsupervised methods to retrieve relevant spectral features from reflectance spectra. Spectral deconvolution aims to decompose a reflectance spectrum as a sum of a continuum modeling its overall shape and some absorption features. We present a flexible and automatic method able to deal with various minerals. The approach is based on a physical model and allows us to include noise statistics. It consists of three successive steps: *(i)* continuum pre-estimation based on non-linear least-squares; *(ii)* pre-estimation of absorption features using a greedy algorithm; *(iii)* refinement of the continuum and absorption estimates. The procedure is first validated on synthetic spectra, including a sensitivity study to instrumental noise and a comparison to other approaches. Then, it is tested on various laboratory spectra. In most cases, absorption positions are recovered with an accuracy lower than 5 nm, enabling mineral identification. Finally, the proposed method is assessed using hyperspectral images of quarries acquired during a dedicated airborne campaign. Minerals such as calcite and gypsum are accurately identified based on their diagnostic absorption features, including when they are in a mixture. Small changes in the shape of the kaolinite doublet are also detected and could be related to crystallinity or mixtures with other minerals such as gibbsite. The potential of the method to produce mineral maps is also demonstrated.**

***Index Terms*—Mineral reflectance spectra, hyperspectral images, HySpex, EGO model, AGM procedure.**

I. Introduction

During the past thirty years, numerous studies have demonstrated the importance of hyperspectral remote sensing for mapping surface composition and properties, *e.g.*, the structure, pigment composition, dry matter and water content of vegetation, coastal and inland water quality, and snow properties [1]. Modern airborne and spaceborne imaging spectrometers such as AVIRIS [2], [3], HySpex, PRISMA [4] or EnMAP [5] give rise to hyperspectral images of various parts of the world. These sensors provide reflectance spectra in the visible near-infrared (VNIR, $[400 - 1300]$ nm) and the short-wave infrared (SWIR, $[1300 - 2500]$ nm). Mapping the physico-chemical properties of soil and rock from hyperspectral data is of great interest for ore exploration, environmental monitoring, and climate change estimation [6].

R. Rialland and R. Marion are with CEA/DAM/DIF, F-91680 Bruyères-le-Châtel, France (e-mail: {ronan.rialland,rodolphe.marion}@cea.fr).

C. Soussen is with CentraleSupélec at Laboratoire des Signaux et Systèmes (Université Paris-Saclay-CNRS-CentraleSupélec), F-91190 Gif-sur-Yvette, France (e-mail: charles.soussen@centralesupelec.fr).

V. Carrère is with Université de Nantes at Laboratoire de Planétologie et Géodynamique (UMR 6112), F-44300 Nantes, France (e-mail: veronique.carrere@univ-nantes.fr).

Mineral reflectance spectra have been extensively analyzed [7]–[9]. They are composed of diagnostic and secondary absorption bands, the position of which mainly depends on the chemical composition [10]. Absorptions in the SWIR range are generally narrow and deep, whereas VNIR absorptions tend to be broader and weaker. These absorption features are respectively related to vibrational processes in the SWIR (rotation, vibration, stretching, and bending of molecule bonds) versus electronic processes (transitions and charge transfers) in the VNIR [10]. The number of absorption features as well as their position, shape and amplitude depend on mineral composition and concentrations [10], mixture with other components [11], observation scale, and instrumental noise [12]. The overall shape of a mineral reflectance spectrum, called continuum, depends on the surface (*e.g*, grain size, roughness, humidity) and illumination conditions [10].

The paper aims to develop an unsupervised method to retrieve relevant spectral features from mineral reflectance spectra. It is well-known that a given mineral exhibits considerable spectral variability, which can be intrinsic, extrinsic, or due to environmental conditions [13]. In this respect, the identification and characterization of minerals from their reflectance spectra is a challenging problem. More specifically, the number of absorption features is related to the observed minerals, and the absorption shapes depend on many physico-chemical properties [10]. Also, the continuum part of the spectra is induced by environmental conditions. It prevents a direct comparison of an unknown spectrum with established reference spectra, *i.e.*, from spectral libraries.

The advanced spectral processing methods can be categorized into two broad groups: data-driven and knowledge-based approaches [14]. Data-driven approaches are model-free methods aiming to learn how to identify and characterize the minerals from a set of training data. They include classification algorithms [15], [16] based on reflectance spectral libraries, *e.g.*, ASTER and USGS [17], [18]. However, classical spectral libraries only have from 10 to 15 reflectance spectra per mineral, which is a major bottleneck for classification and learning approaches. Feature reduction techniques cope with the limited availability of training samples [16]. However, they may induce a loss of information. Unmixing approaches are alternative data-driven techniques aiming to automatically retrieve endmembers and their abundances from a set of observed spectra [19]. One can use them in a blind context, where no training data are available. However, they may yield inaccurate results when no pure pixels are present in the data. Also, classical unmixing methods are based on linear mixing models and do not account for endmember variability [20],

leading to possible unrealistic results [14]. A recent trend is to resort to nonlinear mixing models [21] and augmented linear models designed to address spectral variability [22], [23].

Contrary to data-driven approaches, knowledge-based approaches incorporate prior knowledge about the spectral information embedded in a single mineral reflectance spectrum. They make use of the characterization of absorption features described for various minerals in the literature [7]–[9], [24]. Several reviews of these methods are available [14], [25]. One can first distinguish direct methods such as band ratioing [25] and spectral indices [26]. The latter use spectrum amplitudes for specific known wavelengths but do not exploit information regarding the shape of the absorption features. On the other hand, spectral feature analyses were designed for mineral identification based on absorption's position, amplitude and shape. They include frameworks such as Tetracorder and extended versions [27], [28] dedicated to local spectral analysis and the spectral deconvolution techniques [29]–[31] aiming to capture complete information regarding the whole spectrum.

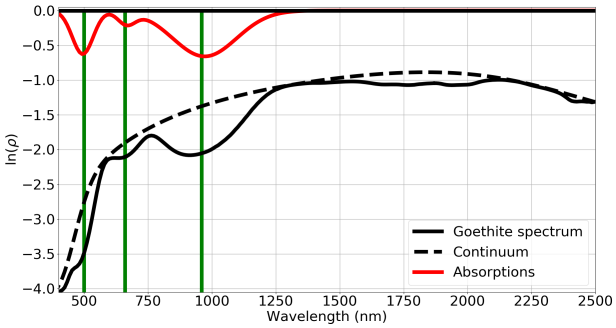
Spectral deconvolution attempts to decompose a given reflectance spectrum as a sum of a continuum and an absorption part. The latter gathers a set of localized parameterized features corresponding to absorptions embedded in the spectrum; see Fig. 1. Several physical models describe the reflectance spectrum, including the Modified Gaussian Model (MGM) [29] and the Exponential Gaussian Optimization (EGO) model [30], [32]. The main difference is that MGM features are symmetric (standard Gaussians), whereas EGO features could be asymmetric. When the asymmetry parameter is 0, the EGO model reduces to the MGM. It is noticeable that absorption features appearing in the VNIR range are symmetric and can be accurately described using the MGM. On the contrary, the EGO model is better suited to absorption features in the SWIR that could be asymmetric [33].

The deconvolution of mineral reflectance spectra has undergone several improvements over the years with increasingly automatic algorithms [31], [34]–[38]. These approaches include the following steps:

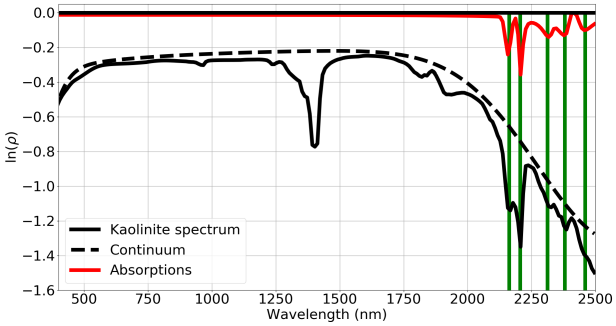
1) continuum removal,
2) MGM or EGO parameter estimation,
3) refinement of the estimated continuum and absorption parameters.

In [37], this three-step approach was dubbed Automatized Gaussian Model (AGM). Although specific algorithmic solutions were put forward in [37] and other references, AGM can be seen as a unifying framework encompassing various contributions. Let us now briefly recall the available solutions for each step and their limitations.

Continuum removal may be carried out using parametric or non-parametric models. The latter usually rely on convex shape assumption, which enables applying geometric methods [34], [39]. This approach is easy to implement but highly sensitive to noise. Indeed, the convex shape is constrained to lie above the data spectrum, which results in artifacts and further false detection of MGM/EGO Gaussians in the case of noisy data. On the other hand, parametric models include second-order polynomials [35] as well as the physical model of [39]. The former are coarse approximate models, whereas the



(a)



(b)

Fig. 1: Logarithm of goethite (a) and kaolinite (b) reflectance spectra as a function of wavelength, derived from the USGS spectral library [18]. The estimated continuums and absorption features are shown with dashed lines and solid red line, respectively. The vertical green lines indicate the positions of the theoretical absorptions used to identify the mineral.

latter is more realistic but exhibits a nonlinear behavior with respect to the physical parameters. This requires to solve a difficult nonlinear least-square problem, with carefully chosen initialization of parameters [30], [35], [37], [40].

The absorption estimation step attempts to determine the number of MGM/EGO Gaussians and estimate the related parameters. Several procedures have been proposed based on spectral derivatives [36], [37] and the continuous wavelet transform [41], [42]. The former retrieve the local minima of the spectrum but are sensitive to noise. Therefore, they may overestimate the number of Gaussians. The latter aim to separate the absorptions from the noise. However, the related threshold parameters are difficult to set. These methods also fall short of estimating the asymmetry parameter.

The refinement step aims to readjust the pre-estimated continuum and absorption features using nonlinear least-squares. It is noticeable that some authors also carry out a post-processing aiming to remove over-estimated Gaussians [37].

The main contribution of the paper is to revisit the AGM procedure for spectral deconvolution. The proposed procedure, named greedy-AGM, includes several substantial improvements. The key novelty is that the number of absorption features is automatically estimated using a greedy strategy. Moreover, the continuum estimation step outperforms classical geometric approaches [39] for noisy data and limits the creation of artifacts. Also, the procedure is designed to avoid

error propagation. Notably, the method is easy to use since no pre-processing of data is required and no expert knowledge is necessary. Note that a preliminary version of this procedure was sketched in the conference paper [43].

The paper is organized as follows. In Section II, we first review the EGO physical model. Then, the proposed greedy-AGM procedure for retrieval of EGO parameters is detailed in section III. Note that greedy-AGM can also be coupled with the MGM physical model with very few adaptations. In section IV, the method is validated on synthetic (noise-free and noisy) data. In section V, it is applied to various experimental spectra related to laboratory measurements and hyperspectral images. This illustrates the potential of greedy-AGM to handle heterogeneous data related to many minerals of interest including calcite, dolomite, goethite, gypsum, kaolinite, and nontronite. Section VI includes a discussion summarizing the main findings and perspectives of this work.

II. EGO MODEL

A. Presentation

The EGO model [30], [32] aims at decomposing the logarithm of a mineral reflectance spectrum ρ as a continuum c and a sum of N asymmetric (modified) Gaussian features:

$$\ln \rho\left(\lambda, \boldsymbol{\theta}\right) = c\left(\lambda, \boldsymbol{\theta}_c\right) - \sum_{i=1}^{N} G\left(\lambda, \boldsymbol{\theta}_{G_i}\right) + n(\lambda) \qquad (1)$$

where λ is the wavelength and $n(\lambda)$ is the noise process. Eq. (1) expresses the dependency upon both continuum ($\boldsymbol{\theta}_c$) and EGO Gaussian ($\boldsymbol{\theta}_G = \{\boldsymbol{\theta}_{G_1}, \ldots, \boldsymbol{\theta}_{G_N}\}$) parameters, respectively. Bold notations refer to a set of parameters.

The continuum part models the overall shape of the spectrum. The drop-off towards the ultraviolet (uv, below 400 nm) is essentially due to iron and towards the SWIR ($water$, beyond 2500 nm) is due to water and OH absorption. The continuum is then expressed as

$$\begin{aligned} c\left(\lambda, \boldsymbol{\theta}_c\right) = & -c_0 - c_1\lambda^{-1} - G_{uv}\left(\lambda, s_{uv}, \mu_{uv}, \sigma_{uv}\right) \\ & - G_{water}\left(\lambda, s_{water}, \mu_{water}, \sigma_{water}\right) \end{aligned} \qquad (2)$$

with $\boldsymbol{\theta}_c = \{c_0, c_1, s_{uv}, \mu_{uv}, \sigma_{uv}, s_{water}, \mu_{water}, \sigma_{water}\}$. G_{uv} and G_{water} are standard Gaussians with amplitudes s_{uv} and s_{water}, positions μ_{uv} and μ_{water} and widths σ_{uv} and σ_{water}, respectively. The offset c_0 and slope c_1 deal with the overall shape.

EGO Gaussians model the wide variety of encountered absorption shapes in the VNIR and SWIR ranges. They are defined as asymmetric functions:

$$G\left(\lambda, \boldsymbol{\theta}_{G_i}\right) = s_i \exp\left(-\frac{1}{2}\frac{\left(\lambda - \mu_i\right)^2}{\left(\sigma_i - k_i\left(\lambda - \mu_i\right)\right)^2}\right) \qquad (3)$$

where s_i is the amplitude, μ_i the position, σ_i the width and k_i the asymmetry of the absorption. Thus, $\boldsymbol{\theta}_{G_i} = \{s_i, \mu_i, \sigma_i, k_i\}$ gathers the parameters of the i-th EGO Gaussian. Note that the EGO model proposed in [30] includes an extra parameter to take the saturation of the absorption bands into account. This complete EGO model is not taken into account here for simplicity. Fig. 1 illustrates the EGO models for goethite and kaolinite spectra with continuum and absorption features estimated by the solution proposed in this paper (note that only the EGO Gaussians corresponding to the mineral diagnostic absorption features are shown).

B. Accuracy and identifiability of the EGO model

The EGO model was successively used to provide information on both the chemical and physical properties of the surface of minerals [33], [34], [37], [44]. Nevertheless, one might encounter the following numerical errors. First, the Gaussians G_{uv} and G_{water} may not be able to adjust the edges of the spectrum perfectly. Also, the VNIR part of the spectrum is modeled by the $c_1\lambda^{-1}$ and $G_{uv}(\lambda)$ which influence the drop-off at the edge of the spectrum. Note also that the range of variation of c_1 is large, from 0 to $\sim 10^3$ nm, since c_1 is weighted by λ^{-1} in Eq. (1).

Moreover, the EGO model may not be identifiable since the decomposition (1) could hold for different sets of EGO parameters $(\boldsymbol{\theta}_c, \boldsymbol{\theta}_G)$. For instance, a sum of EGO Gaussians may be confused with a single EGO Gaussian. Furthermore, large values of the asymmetry parameter k may create unrealistic absorption shapes. Finally, a large number of EGO Gaussians N may lead to overfitting when absorptions are close in position [37].

To avoid identifiability issues, most semi-automatic methods require *(i)* initialization based on prior knowledge of the number of absorption features and/or their center positions [30], [32], [34], [35], [44]; and *(ii)* bounds on the feature parameters [45]. In this paper, we aim at automatically estimating the complete set of EGO parameters N, $\boldsymbol{\theta}_c$, and $\boldsymbol{\theta}_G$.

III. THE GREEDY-AGM PROCEDURE

A. Overview

The input of the greedy-AGM procedure includes a mineral reflectance spectrum $\rho(\lambda)$ and the associated noise covariance matrix $\boldsymbol{\Sigma}$. The noise component is assumed to be zero-mean, Gaussian, and independent between spectral bands. However, the noise variance σ_λ^2 is known to depend on the spectral band λ. The noise covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_\lambda^2)$ is thus diagonal and can be either calculated knowing the sensor characteristics or estimated from the data [46].

Following the AGM methodology proposed in [37], the method is composed of three steps: *(i)* continuum pre-estimation, *(ii)* absorption pre-estimation, and *(iii)* joint refinement. Continuum pre-estimation is carried out using a non-linear least-squares formulation. The resulting pre-estimated continuum is then subtracted from the spectrum:

$$a(\lambda) = c(\lambda, \boldsymbol{\theta}_c^{pre}) - \ln \rho\left(\lambda\right) \qquad (4)$$

where $\boldsymbol{\theta}_c^{pre}$ and $a(\lambda)$ respectively refer to the pre-estimated continuum parameters and the absorption part, that is, the sum of EGO Gaussians in (1). Absorption pre-estimation then automatically retrieves the parameters of the EGO Gaussians, noted $\boldsymbol{\theta}_G^{pre}$ from $a(\lambda)$. The joint refinement step updates the continuum and absorption estimates.

The three steps of greedy-AGM are detailed in the next subsections. The column vector gathering the reflectance values

$\rho(\lambda)$ for every λ is denoted by $\boldsymbol{\rho} \in \mathbb{R}^{N_\lambda}$, with N_λ the number of wavelengths. Similarly, the vectors $\mathbf{c}(\boldsymbol{\theta}_c)$, $\mathbf{a}(\boldsymbol{\theta}_G)$ and $\boldsymbol{\sigma}$ gather the values of $c(\lambda, \theta_c)$, $a(\lambda, \boldsymbol{\theta}_G)$ and σ_λ, respectively, for all λ. The outputs of greedy-AGM will be denoted with hat notations: $\hat{N}$, $\widehat{\boldsymbol{\theta}}_c$, $\widehat{\boldsymbol{\theta}}_G$.

B. Continuum pre-estimation

The following least-squares formulation corresponds to the maximum likelihood estimation of parameters $\boldsymbol{\theta}_c$:

$$\min_{\boldsymbol{\theta}_c} \|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{c} - \ln \boldsymbol{\rho}\left(\boldsymbol{\theta}_c\right)\right)\|^2 \;\text{ s.t. }\; \begin{cases} \mathbf{c}\left(\boldsymbol{\theta}_c\right) - \ln \boldsymbol{\rho} \geq \alpha \boldsymbol{\sigma} \\ \boldsymbol{\theta}_c^{min} \leq \boldsymbol{\theta}_c \leq \boldsymbol{\theta}_c^{max} \end{cases} \tag{5}$$

where $\|\,.\,\|$ refers to the Euclidean norm, and the compact writing $\mathbf{c}(\boldsymbol{\theta}_c) - \ln \boldsymbol{\rho} \geq \alpha \boldsymbol{\sigma}$ refers to N_λ constraints, one per wavelength. In a nutshell, the latter constraints aim to impose that the absorption $a(\lambda)$ defined in (4) is non-negative valued, that is, the amplitude of the EGO Gaussians are positive. The coefficient α is set empirically to adapt the tolerance to the noise standard deviation. The $\boldsymbol{\theta}_c^{min}$ and $\boldsymbol{\theta}_c^{max}$ bounds are defined in subsection III-E.

The least-squares problem (5) is solved using the COBYLA algorithm [47]. Since (5) is a non-linear least-squares problem, the numerical solution $\boldsymbol{\theta}_c^{pre}$ is likely to be a local minimizer, which depends on the initial value of $\boldsymbol{\theta}_c$, denoted $\boldsymbol{\theta}_c^{start}$. Fig. 2 illustrates the initialization strategy. The offset c_0 is set to the maximum value of the spectrum, and the continuum slope c_1 is set to 0 [37]. Starting values of the standard Gaussians G_{uv} and G_{water} are computed similarly in the VNIR and SWIR spectral ranges. Their positions are set to predefined values $\mu_{uv} = 200$ nm and $\mu_{water} = 2800$ nm [40]. Then, we draw the line l_1 joining the point of the spectrum of minimal wavelength and the point λ of the VNIR range for which $\rho(\lambda)$ is maximum. The line l_2 is drawn similarly by considering the last point of the spectrum and the point of maximum amplitude in the SWIR range. The amplitudes s_{uv} and s_{water} are set to the values of l_1 and l_2 at μ_{uv} and μ_{water}, with c_0 removed. The widths σ_{uv} and σ_{water} are computed according to the settings of Fig. 2. This strategy is adapted to various continuum shapes of the minerals of interest, as G_{uv} and G_{water} only influence the VNIR and SWIR ranges, respectively.

C. Pre-estimation of absorption characteristics

Once the continuum is pre-estimated, Eq. (4) yields the absorption signal $a(\lambda)$, which rereads

$$a(\lambda) = \sum_{i=1}^{N} G\left(\lambda, \boldsymbol{\theta}_{G_i}\right) - n(\lambda) + e(\lambda) \tag{6}$$

according to (1), where $e(\lambda)$ refers to the estimation error in the continuum pre-estimation step. Absorption estimation consists of decomposing $a(\lambda)$ into a weighted sum of EGO Gaussians with positive weights. This problem can be effectively solved using greedy algorithms [48], which select the EGO Gaussians one by one in a dictionary containing a large number of synthetic EGO Gaussians. This approach is detailed hereafter. Moreover, the number of selected Gaussians will be adaptively estimated using an information-theoretic criterion.

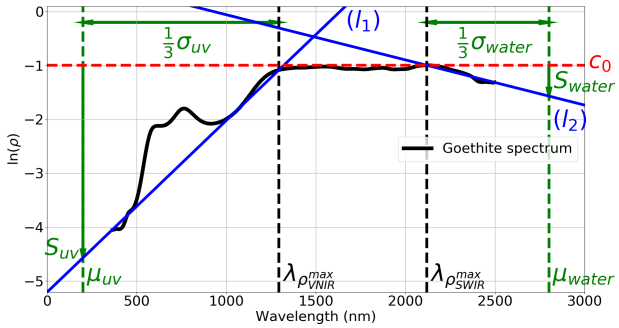


Fig. 2: Computation of the initial continuum parameters $\boldsymbol{\theta}_c^{start}$ for a goethite reflectance spectrum from the USGS spectral library [18]: sample `splib07b_Goethite_GDS134_ASDFRb_AREF`. The lines l_1 and l_2 are drawn in blue, and the maximum of the spectrum c_0 is in dashed red. The parameters of G_{uv} and G_{water} are in green.

1) Estimation of EGO Gaussian parameters: Starting from (6), we adopt an optimization approach akin to (5), which consists of minimizing the negative log-likelihood

$$\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{a} - \sum_{i=1}^{N} G\left(\lambda, \boldsymbol{\theta}_{G_i}\right)\right)\right\|^2 \tag{7}$$

with respect to the EGO Gaussian parameters $\boldsymbol{\theta}_G$. The difficulty of this optimization problem is twofold. First, this is a non-linear least-squares problem. Thus the cost function is likely to be multimodal, *i.e.*, with many local minimizers. Moreover, the number of EGO Gaussians N is unknown.

We propose to solve the latter optimization problem for consecutive values of N, yielding absorption decompositions for many N values. Then, a model order selection rule [49] is applied to select one particular decomposition among all. An attractive feature of greedy algorithms is that they provide solutions to the minimization problems (7) for $N = 1, \ldots, N_{max}$ with a single run. N and N_{max} are the current iteration of the algorithm and the total number of iterations, respectively. The absorption decomposition found for a given N is used as initialization to search the decomposition with $N+1$ Gaussians at the next iteration.

The greedy approach is based on the creation of a very large dictionary $\mathbf{G} \in \mathbb{R}^{N_\lambda \times N_{atom}}$ containing possible spectral features where N_{atom} refers to the number of spectral features in the dictionary (also called atoms). Each atom of $\mathbf{G}$ refers to a specific EGO Gaussian with predefined shape parameters (μ, σ, k), and magnitude $s = 1$. The EGO Gaussian model (3) thus reads as $s_i\,\mathbf{g}_i$ where $\mathbf{g}_i$ refers to a column of $\mathbf{G}$ related to the i-th absorption. In turn, the sum of EGO Gaussians in (7) reads as a matrix-vector product $\mathbf{Gs}$ where $\mathbf{s} \in \mathbb{R}^{N_{atom}}$ are the (non-negative) magnitudes. Thus, the minimization of the residual error (7) rereads:

$$\min_{\mathbf{s}} \|\boldsymbol{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{Gs})\|^2 \quad \text{s.t.} \quad \begin{cases} \mathbf{s} \geq \mathbf{0} \\ \|\mathbf{s}\|_0 \leq N \end{cases} \tag{8}$$

where $\|\mathbf{s}\|_0$ counts the number of non-zero entries in vector $\mathbf{s}$ and N is the number of absorption features embedded in a mineral reflectance spectrum.

Algorithm 1 Solve problem (8) for $N = 1, \dots, N_{max}$. The support S gathers the indices of the nonzero elements in $\mathbf{s}$. $\mathbf{g}_i$ refers to the i-th column of $\mathbf{G}$. $\mathbf{G}_S$ refers to the submatrix gathering the columns of $\mathbf{G}$ indexed by S.

inputs: $\mathbf{a}$, $\mathbf{\Sigma}$, $\mathbf{G}$, N_{max}
outputs: $\mathbf{s}^N$, r^N **for all** N
  $\mathbf{r} \leftarrow \mathbf{\Sigma}^{-1/2}\mathbf{a}$
  $\mathbf{s} \leftarrow \mathbf{0}$
  $S \leftarrow \emptyset$
  for $N = 1$ **to** N_{max} **do**
    $\ell \leftarrow \arg\max_i \{\mathbf{g}_i^T \, \mathbf{\Sigma}^{-1/2}\mathbf{r} \,/\, \|\mathbf{\Sigma}^{-1/2}\mathbf{g}_i\|, i \notin S\}$
    $S \leftarrow S \cup \{\ell\}$
    $\mathbf{s}(S) \leftarrow \arg\min_{\boldsymbol{z}} \{\|\mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}_S\mathbf{z})\|^2, \ \mathbf{z} \geqslant \mathbf{0}\}$
    $\mathbf{s}^N \leftarrow \mathbf{s}$
    $\mathbf{r}^N \leftarrow \mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}\mathbf{s}^N)$
  end for

2) NNOMP, a greedy algorithm: Non-Negative Orthogonal Matching Pursuit (NNOMP) [50], [51] is a greedy search algorithm dedicated to solving problem (8) for consecutive $N = 1, \dots, N_{max}$. As shown in Algorithm 1, NNOMP gradually selects atoms in $\mathbf{G}$ to refine the approximation $\mathbf{a} \approx \mathbf{G}\mathbf{s}$. Initially, no atom is selected, then $\mathbf{s} = \mathbf{0}$. At each iteration, a new atom of $\mathbf{G}$ is added to the current support S, gathering the previously selected atoms. Their amplitudes are then adjusted using a Non-Negative Least-Squares solver [51]. The residual $\mathbf{r} = \mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}\mathbf{s})$ is updated accordingly.

The atom selection rule of NNOMP aims at decreasing at most the norm of the residual $\mathbf{r}$. Thus, in practice, the main absorptions of the spectrum are retrieved in the first iterations, then smaller absorptions are gradually retrieved.

3) Model order selection: Model order selection aims at automatically selecting a solution $\mathbf{s}^N$ among the outputs $\mathbf{s}^N$, $N = 1, \dots, N_{max}$. Many rules such as the Akaike information criterion tend to overestimate the actual number of features because they are designed based on an asymptotic information theory analysis, where the number of observations tends to infinity [49]. Minimum Description Length (MDL) criteria are easily adaptable to greedy algorithms [52]. Moreover, they were adapted to short data records, where the number of parameters is moderately smaller than the data size as in our case [53]. In the case of known noise statistics, they read:

$$\hat{N} = \arg\min_N \left(\ln \|\mathbf{r}^N\| + \frac{\ln N_\lambda \, (N+1)}{N_\lambda - N - 2} \right). \tag{9}$$

This criterion is thus evaluated at each iteration of NNOMP. Then, the NNOMP output $\mathbf{s} = \mathbf{s}^N$ for $N = \hat{N}$ is selected.

4) Dictionary creation and $\boldsymbol{\theta}_G$ pre-estimation: The dictionary $\mathbf{G}$ gathers a large number of synthetic EGO Gaussians computed for a set of position (μ), width (σ), and asymmetry (k) parameters. Each EGO parameter is discretized over a pre-defined grid whose resolution is related to the size of the dictionary. Fine grids induce a greater precision at the cost of increased memory storage and computational time. We emphasize that the dictionary $\mathbf{G}$ is not related to any spectral library of known minerals, *e.g.*, the acquisition of annotated reflectance spectra. The construction of the dictionary is generic and simply relies on the simulation of EGO Gaussians.

The NNOMP algorithm yields the number $\hat{N}$ of selected atoms, their column indices $\mathbf{g}_i$ in the dictionary $\mathbf{G}$ and their magnitudes s_i. Since there is a one-to-one correspondence between a column of $\mathbf{G}$ and a set of parameters (μ, σ, k), the parameters related to the i-th selected atom, *i.e.*, $\boldsymbol{\theta}_{G_i}^{pre} = \{s_i, \mu_i, \sigma_i, k_i\}$ can be retrieved from the knowledge of the location of $\mathbf{g}_i$ in the dictionary $\mathbf{G}$ and from the magnitude s_i.

D. Joint refinement

Once the continuum parameters $\boldsymbol{\theta}_c^{pre}$, the number of EGO Gaussians $\hat{N}$ and the associated parameters $\boldsymbol{\theta}_G^{pre}$ have been pre-estimated, the pre-estimated parameters are refined using a joint optimization process. The non-linear least-squares minimization problem:

$$\min_{\boldsymbol{\theta}_c, \boldsymbol{\theta}_G} \left\| \mathbf{\Sigma}^{-\frac{1}{2}} \left(\ln \boldsymbol{\rho} - \mathbf{c}\left(\boldsymbol{\theta}_c\right) + \sum_{i=1}^{N} \mathbf{G}\left(\boldsymbol{\theta}_{G_i}\right) \right) \right\|^2 \tag{10}$$

is solved for $N = \hat{N}$ using the trust region reflective algorithm [54] with initial parameters $\boldsymbol{\theta}_c^{pre}$ and $\boldsymbol{\theta}_G^{pre}$.

E. Algorithm parametrization

This section summarizes the parameter settings within greedy-AGM.

Bounds on the continuum parameters $\boldsymbol{\theta}_c$ are defined in (5) to ensure that the range of EGO parameters is consistent with their physical interpretation. As $\ln \rho(\lambda)$ is negative valued, the parameters c_0, c_1, s_{uv} and s_{water} in (2) are imposed to be positive. Also, this ensures the drop-offs of the continuum to appear on the edges of the spectrum. The Gaussians G_{uv} and G_{water} are supposed to be centered outside the spectral range of the acquisition sensors with $\mu_{uv} \in \left[0\,,\, \lambda^{min}\right]$ nm and $\mu_{water} \in [\lambda^{max}\,,\, 3000]$ nm. To avoid unrealistic values, lower (0 nm) and upper (3000 nm) bounds are set for both μ_{uv} and μ_{water}. The widths of G_{uv} and G_{water} are unconstrained. The tolerance α in (5) deals with the noise standard deviation. α is set to 3 in the noisy case, and 0 in the noise-free case.

In Algorithm 1, the maximum number of selected atoms N_{max} is set to 20 to limit the possible number of absorptions. Since the estimation of N is adaptive, $\hat{N} = N_{max}$ is rarely reached in practice. Indeed, N often ranges between 1 and 5 for mineral reflectance spectra. To create the dictionary $\mathbf{G}$, we define two grids adapted to the various absorption shapes in the VNIR and SWIR ranges, respectively. The VNIR absorptions tend to be broad, whereas the SWIR absorptions are narrow and occasionally asymmetric. In the VNIR, k_i is set to 0, as VNIR absorptions are generally symmetric. The width σ_i varies between 30 and 380 nm with increments of $\frac{p}{2}$, where p is the wavelength sampling step of the sensor (for AVIRIS, $p \sim 10$ nm). The latter minimum width is set experimentally to avoid selecting atoms of small amplitudes that do not identify with an absorption. The maximum width corresponds to maximum ferric goethite absorption at 920 nm, measured on goethite spectra from the USGS spectral library [18]. The

positions are set on a scale from λ^{min} to 1300 nm in steps of $\frac{p}{2}$. In the SWIR, k_i varies between -0.2 and 0.2 in steps of 0.05 to avoid the creation of unrealistic EGO Gaussians. These values were chosen after several experiments on mineral spectra with asymmetric absorptions in the SWIR. The width varies in the range of 5 to 45 nm in steps of $\frac{p}{2}$. The maximum width is fixed experimentally, based on the application of the greedy-AGM procedure on an alunite spectrum. The positions vary between 1300 nm and λ^{max} in steps of $\frac{p}{10}$.

IV. Validation on synthetic data

The greedy-AGM procedure is first validated on noise-free data. Then, it is compared to classical algorithmic solutions under the AGM framework.

A. Synthetic data description

Three synthetic spectra are generated using the EGO model to represent various continuum and absorption shapes (amplitude, width, asymmetry). They are shown in Fig. 3, and their parameters are given in Tab. I. Spectra 1 and 2 and spectrum 3 have three and four absorption features, respectively. The spectra are convolved to the AVIRIS spectral response function. The objective is to focus on difficult cases with broad and shallow absorptions in the VNIR (like for goethite and hematite), close and narrow absorptions in the SWIR (like for kaolinite), asymmetrical absorptions (like for alunite and calcite). Hereafter, the discrepancy between a synthetic (noise-free) spectrum $\boldsymbol{y}^{syn}$ and its recovered version $\boldsymbol{y}^{est}$ is assessed using the goodness of fit, expressed in dB:

$$r = 10\ \log_{10} \frac{\|\boldsymbol{y}^{syn}\|^2}{\|\boldsymbol{y}^{syn} - \boldsymbol{y}^{est}\|^2}. \tag{11}$$

This indicator is computed for each step of the greedy-AGM procedure, yielding ratios r_1, r_2 and r_3 between the synthetic and estimated continuums $\boldsymbol{c}^{syn}$ and $\boldsymbol{c}^{est}$, absorption spectra $\boldsymbol{a}^{syn}$ and $\boldsymbol{a}^{est}$, and spectra $\ln \boldsymbol{\rho}^{syn}$ and $\ln \boldsymbol{\rho}^{est}$, respectively. Furthermore, the accuracy of recovery of absorption positions μ is assessed using the signed error $\delta\mu = \mu^{est} - \mu^{syn}$.

B. Algorithm validation in the noise-free case

Each synthetic spectrum is deconvolved with the greedy-AGM procedure, see Fig. 3. For this noise-free scenario, the noise covariance $\boldsymbol{\Sigma}$ is set to the identity matrix.

In the continuum pre-estimation step (Fig. 3(d-f)), r_1 rises to 30 dB for all spectra. The errors on the continuum amplitudes are always lower than 10%. These errors are related to the fact that G_{uv} and G_{water} are broad absorptions centered outside the $[400 - 2500]$ nm range, where data are missing. One can distinguish three kinds of artifacts. First, the amplitudes and shapes of broad absorptions can be impacted (*e.g.*, the absorption at 960 nm of spectrum 1). Notably, errors in the VNIR are larger as the continuum is modeled by both and G_{uv} $c_1\lambda^{-1}$, and λ^{-1} is large. Then, broad false absorptions with small amplitude may appear at central wavelengths. Finally, small and narrow false absorptions may appear on the edges of the spectra. Such artifacts are not specific to the proposed continuum removal procedure. For example, the solution in [39], used in the GMEX (Guide for Mineral EXploration) [24] also creates artifacts at both edges of the spectral range.

Fig. 3(g-i) illustrates the absorption pre-estimation step. First, r_2 is close to 20 dB, and the synthetic absorption features are all detected. Isolated absorptions are retrieved with high accuracy, while broad and overlapping absorptions may be slightly shifted from their synthetic positions. The creation of false EGO Gaussians is due to continuum removal artifacts, but their amplitudes are small. For example, in spectrum 1, two false absorptions appear around 1900 and 2500 nm and are fitted by four EGO Gaussians. To further illustrate the impact of continuum artifacts, the absorption pre-estimation step is applied to spectrum 1, where the synthetic continuum has been removed (corresponding to a perfect continuum removal scenario). In Fig. 5, the greedy algorithm retrieved only three absorption features, with no false EGO Gaussian. The first selected Gaussian is the one whose energy is maximal. Then, the isolated Gaussian of largest amplitude and the small overlapping absorption in the VNIR are accurately retrieved.

Finally, in Fig. 3(j-l), r_3 is close to 60 dB, which confirms that the recovered spectrum $\boldsymbol{\rho}^{est}$ perfectly identifies with $\boldsymbol{\rho}^{syn}$. The estimated continuums are consistent with the shape of the three spectra. Absorptions are retrieved with small errors, especially on their positions. The maximum position error is reached for the broad and overlapping absorption at 960 nm of spectrum 1 ($\delta\mu = 40$ nm). We have $\delta\mu < 3$ nm for other absorptions, which is lower than the spectral resolution of airborne sensors such as AVIRIS (~ 10 nm). Also, absorption shape parameters are accurately retrieved such as the broad and overlapping absorption asymmetry at 2165 nm of spectrum 2 ($k = -0.2$). For isolated absorptions, errors on the amplitudes, width, and asymmetry are lower than 1%.

C. Algorithm comparison in the noisy case

The three synthetic spectra are now corrupted with a zero-mean, additive Gaussian noise whose variance is identical on each spectral band. The noise variance is expressed as $\sigma^2 = 10^{-\mathrm{SNR}/10} \ln \overline{\rho}$, where $\overline{\rho}$ is the mean of the synthetic spectrum and SNR is the Signal-to-Noise Ratio, expressed in dB.

The greedy-AGM procedure is compared to classical algorithmic solutions under the AGM framework. The competing procedure will be simply denoted by "AGM". In the latter, continuum removal is carried out in two stages [37], by applying the geometric approach of [39] (convex hull computation), and then adjusting the parametric model (2). Then, the spectral deconvolution algorithm of [34], [36] is applied to retrieve the number of EGO Gaussians and their parameters. The refinement step is identical to the one proposed in subsection III-D. The algorithm outputs are shown in Fig. 4 for $\mathrm{SNR} = 30$ dB.

In noisy cases, greedy-AGM outperforms AGM. In the simulation of Fig. 4, greedy-AGM retrieves the main absorption features (vertical lines). For spectrum 3, the small ones (at 2312 and 2380 nm) cannot be found because their amplitude is lower than the average noise level. On the contrary, some of the main features are not found by AGM, *e.g.*, the one at 2324 nm for spectrum 2. Moreover, a single EGO Gaussian retrieves

	Continuum parameters				Absorption parameters			
	c_0	c_1	G_{uv}	G_{water}	Abs 1	Abs 2	Abs 3	Abs 4
spectrum 1	0.5	500	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 200$ $s = 1.0$	$\mu = 660$ $\sigma = 40$ $s = 0.1$ $k = 0.0$	$\mu = 960$ $\sigma = 125$ $s = 0.25$ $k = 0.0$	$\mu = 2283$ $\sigma = 7$ $s = 0.4$ $k = 0.2$	
spectrum 2	0.5	0.01	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 400$ $s = 0.8$	$\mu = 1760$ $\sigma = 12$ $s = 0.3$ $k = 0.0$	$\mu = 2165$ $\sigma = 45$ $s = 0.4$ $k = -0.25$	$\mu = 2324$ $\sigma = 10$ $s = 0.25$ $k = 0.0$	
spectrum 3	0.2	0.01	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 400$ $s = 1.0$	$\mu = 2162$ $\sigma = 15$ $s = 0.35$ $k = 0.0$	$\mu = 2206$ $\sigma = 17$ $s = 0.45$ $k = 0.0$	$\mu = 2312$ $\sigma = 10$ $s = 0.05$ $k = 0.0$	$\mu = 2380$ $\sigma = 10$ $s = 0.05$ $k = 0.0$

TABLE I: Synthetic EGO parameters.

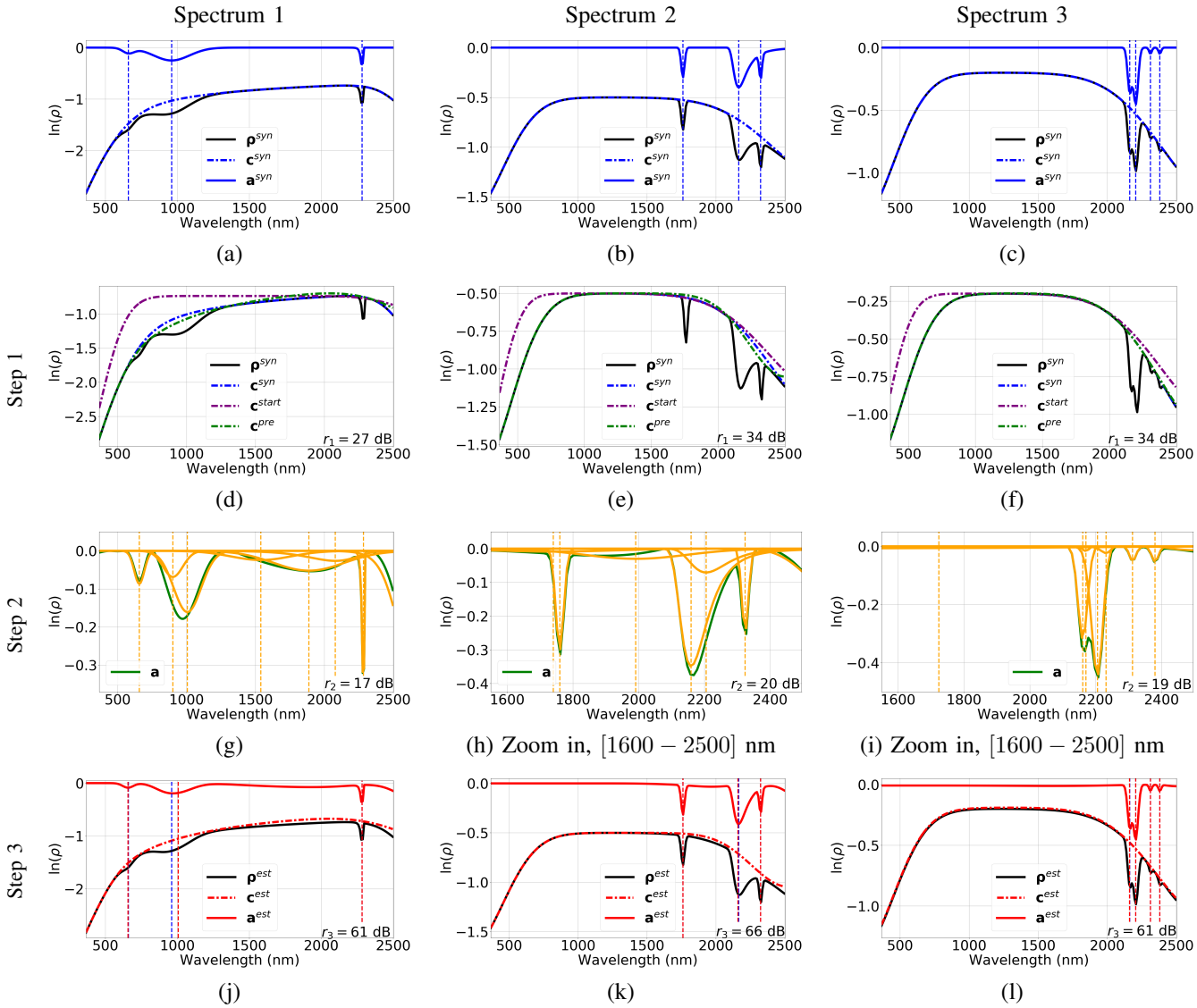
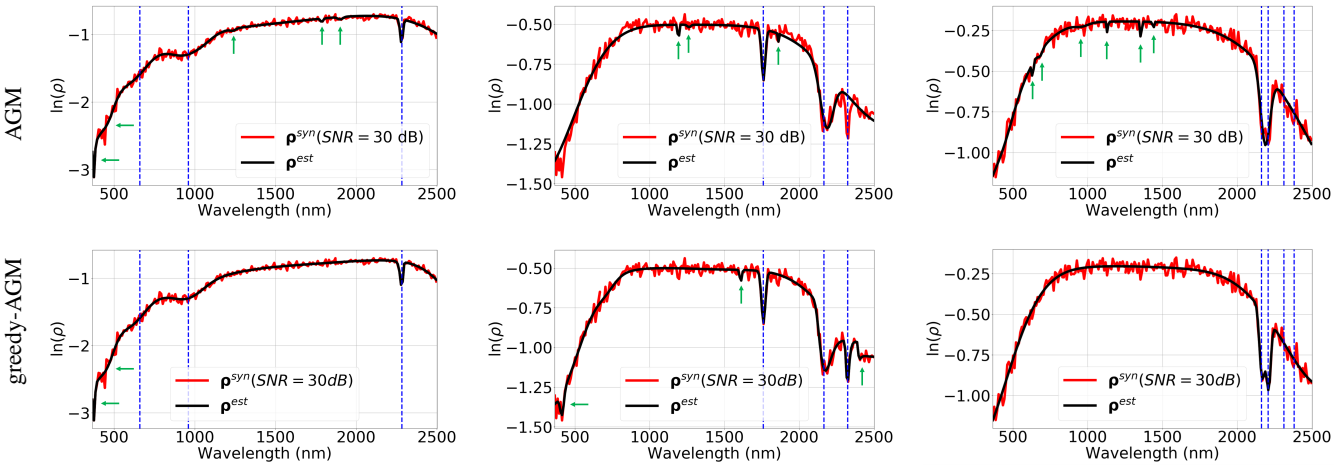


Fig. 3: Results of the greedy-AGM procedure on three synthetic spectra. (a,b,c) Synthetic spectra $\boldsymbol{\rho}^{syn}$ (solid black curve), absorption $\boldsymbol{a}^{syn}$ (solid blue line) and continuum $\boldsymbol{c}^{syn}$ (dashdot blue line). The absorption positions are shown with vertical dashed lines. (d,e,f) Starting $\boldsymbol{c}^{start}$ (purple - dashdot) and pre-estimated continuum $\boldsymbol{c}^{pre}$ (green - dashdot). (g,h,i) Absorption spectra after the continuum pre-estimation step (green) and EGO Gaussians selected by NNOMP (orange). Their positions are shown with vertical dashed lines. (j,k,l) Estimated absorption spectra $\boldsymbol{a}^{est}$ and continuum $\boldsymbol{c}^{est}$ (red). For each step, the goodness of fit (r_1, r_2, r_3) is given in dB.



(a) Spectrum 1 (b) Spectrum 2 (c) Spectrum 3

Fig. 4: Assessment of the AGM (top) and greedy-AGM (bottom) procedures for the three noisy synthetic spectra. Noisy and estimated spectra are shown in red and black, respectively. Synthetic absorption positions are represented as vertical dashed blue lines. Green arrows indicate the location of false EGO Gaussians.

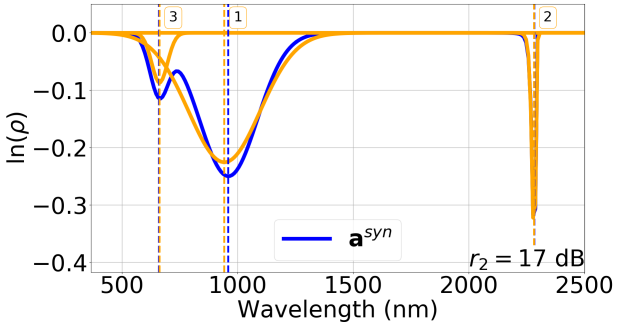


Fig. 5: Synthetic absorption part of spectrum 1 (blue) and EGO Gaussians selected by NNOMP (orange). The order of selection is also indicated. The true absorption positions and their estimates are shown as vertical dashed lines.

the absorption doublet around 2200 nm in spectrum 3. Notably, AGM strongly overestimates the number of EGO Gaussians. Green arrows indicate false EGO Gaussians in Fig. 4. They are due to the creation of artifacts in the continuum pre-estimation step and also to the spectral derivative algorithm for absorption feature retrieval, known to be sensitive to noise.

For SNRs lower than 25 dB, AGM still fails to estimate the number of EGO Gaussians and their parameters without any user supervision. On the contrary, greedy-AGM does not require expert empirical tuning of parameters, save the automatic rules stated in Section III-E. Moreover, the broadest and deepest absorptions are retrieved. The number of estimated EGO Gaussians is underestimated for low SNR, which is a desirable property. Indeed, absorption features of small amplitudes are hidden in the noise, so they cannot be recovered.

Let us remark that the asymmetry parameter k of EGO Gaussians can hardly be recovered from noisy data since a wide range of values of $\boldsymbol{\theta}_G$ yield similar spectral shapes. k can be recovered though for denoised data (SNR > 50 dB).

V. Application to Real Data

A. Greedy-AGM procedure for real data

In hyperspectral imaging, spectral ranges around 1400 and 1900 nm may be affected by the main saturated atmospheric water vapor absorptions. Spectral masks (represented as grey areas in Figs. 6, 8 and 9) are thus applied to remove these spectral bands. The dictionary creation is slightly adapted within greedy-AGM to cope with these missing data, by including dictionary atoms centered in the spectral masks. This allows us to retrieve absorptions inside the spectral masks from available data on the edges of the spectral masks, which avoids the creation of artifacts. Interestingly, the algorithm accuracy is not impacted by the spectral masks. Hereafter, spectral masks are applied to laboratory and image spectra.

B. Results on laboratory data

Here, the goal is to analyze the advantages and limits of the method for minerals often encountered in hyperspectral imaging, that entail spectral features difficult to deconvolve. The ability to identify minerals from their estimated EGO parameters is also discussed.

1) Spectral library description and data preparation: We use the USGS spectral library [18], which gathers laboratory VNIR and SWIR reflectance spectra of various materials. Six reference minerals are considered, for which the electronic and vibrational processes are well-known (Tab. II). The continuum and absorption bands (shape, position of absorptions) differ for each spectrum since the studied minerals belong to different mineralogical classes. To obtain results comparable to those on synthetic data, the spectra are convolved to AVIRIS spectral responses. The estimates yielded by the greedy-AGM will be numerically compared to the theoretical values in Tab. II, seen as ground truth.

Goethite (α-FeO (OH)) is a Fe-hydroxyde mineral. Its reflectance spectrum exhibits two diagnostic absorption features

Mineral - Group	Sample	Diagnostic absorption (nm)	Secondary absorption (nm)
Calcite - Carbonate	splib07b_Calcite_GDS304_75-150um_ASDFRb_AREF	2342	2156
Dolomite - Carbonate	splib07b_Dolomite_HS102.4B_ASDNGb_AREF	2324	2140
Goethite - Fe-hydroxyde	splib07b_Goethite_GDS134_ASDFRb_AREF	660 960	500
Gypsum - Sulphate	splib07b_Gypsum_HS333.4B_(Selenite)_ASDFRa_AREF	1750	1538 2215
Kaolinite - Phyllosilicate	splib07b_Kaolinite_CM9_ASDNGb_AREF	2162 2206	2312 2380
Nontronite - Smectite	splib07b_Nontronite_NG-1.a_ASDNGb_AREF	2283 660 960	2378

TABLE II: Some minerals of interest in the USGS spectral library [18].

at 660 and 960 nm, which are attributed to Fe^{3+} electronic processes that include transition in the iron cations $\left(Fe^{3+}\right)$ from ground state ${}^6A_{1\nu}$ to ${}^4T_{1\nu}$ [55]. The fall-off short of 550 nm is attributed to a conduction band, typical of the trans-opaque iron oxides. We are here evaluating the efficiency of the method to model typically broad absorptions related to ferric iron in the VNIR.

Kaolinite $\left(Al_2Si_2O_5\left(OH\right)_4\right)$ is a phyllosilicate mineral. Its reflectance spectrum exhibits a doublet of diagnostic absorptions at 2162 and 2206 nm. This doublet is related to the Al-OH combination band. It is an indicator of the two-layer dioctahedral structure of kaolinite [56]–[58]. We are here evaluating the ability of the method to estimate a doublet of narrow absorptions with small amplitudes in the SWIR.

Nontronite $\left(Na_{0.33}Fe_2^{3+}\left(Si,Al\right)_4O_{10}\bullet nH_2O\right)$ is a smectite clay. Its reflectance spectrum exhibits three diagnostic absorptions at 660, 960 and 2283 nm. Like goethite, the 660 and 960 nm absorptions are attributed to electronic transition in the Fe^{3+}ion. The 2283 nm absorption is a Fe-OH combination band [58], [59]. The objective here is to test the accuracy of the method in both VNIR and SWIR ranges.

Gypsum $\left(CaSO_4\bullet 2H_2O\right)$ is a sulphate mineral. Its reflectance spectrum exhibits several absorptions including the diagnostic absorption centered at 1750 nm, whose origin is controversial. It is either associated to H_2O combination overtones [60], to $OH-$ or H_2O bending, stretching and rotational overtones and/or $S-O$ bending overtones [61]. It has been partly associated to H_2O because various experiments involving dehydration of gypsum demonstrated an attenuation of this specific absorption with increasing temperature. Also, it becomes narrower and less structured starting at 100°C [60]. The objective here is to evaluate the capacity of the method to detect a diagnostic absorption feature when it is close to the "atmospheric" spectral mask.

Calcite $\left(CaCO_3\right)$ and dolomite $\left(CaMg\left(CO_3\right)_2\right)$ are carbonate minerals with respective diagnostic absorption features at 2342 and 2324 nm, and secondary ones at 2156 and 2140 nm. These absorptions are attributed to overtones and combinations of bending and stretching vibrations of the CO_3^{2-} ion [62]–[64]. The slight shift for dolomite is due to the presence of Mg in addition to Ca associated with the CO_3^{2-} ion. The objective here is to evaluate the capacity of the method to discriminate calcite from dolomite despite their close diagnostic absorption features. The influence of the asymmetry factor on the estimate of absorption characteristics is also evaluated.

Note that a post-processing step on the greedy-AGM results could be applied to remove unrealistic or overfitted EGO Gaussians (represented in blue in the following figures).

2) VNIR absorptions: goethite ferric iron absorptions: Fig. 6 shows the results for goethite. First, the continuum is correctly estimated. It fits the overall shape of the spectrum in the SWIR, where no absorption of mineralogical interest is present. In the VNIR, the position of G_{uv} is close to the edge of the spectrum with a small width ($\mu_{uv} = 402$ nm, $\sigma_{uv} = 65$ nm). Also, c_1 is large ($c_1 = 1110$). Thus, the fit with EGO Gaussians of the absorption at 500 nm is mainly affected by G_{uv} while the ones at 660 and 960 nm depend on c_1. However, the estimated continuum in the VNIR lies above the spectrum, and its shape is consistent with the results on synthetic data; see spectrum 1 in section IV-B.

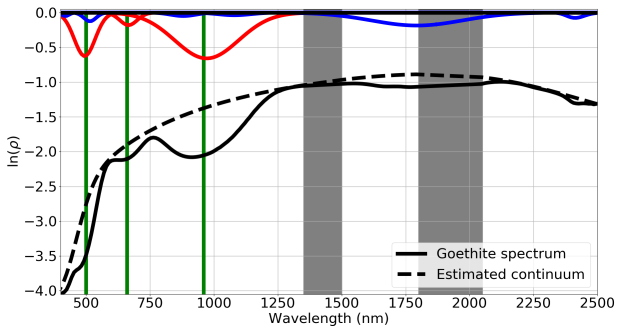


Fig. 6: Goethite spectrum (black), estimated continuum (black-dashed), diagnostic and secondary EGO Gaussians (red), other EGO Gaussians (blue). The true positions of the diagnostic and secondary absorptions are shown as vertical lines.

As expected, the absorption at 960 nm is fitted by a broad EGO Gaussian whose position is slightly shifted ($\delta\mu = 11$ nm). The shift impacts the absorption at 660 nm and tends to adjust both absorptions simultaneously. The absorption at 660 nm is fitted by an EGO Gaussian of reduced amplitude. In addition, NNOMP selects three EGO Gaussians (at 880, 997, and 1079 nm) of low amplitudes to reduce the reconstruction error. An EGO Gaussian accurately adjusts the absorption at 500 nm ($\mu = 494$ nm), together with Gaussians of small amplitudes.

3) SWIR absorptions: kaolinite doublet and secondary absorptions: Fig. 7 shows the results for kaolinite. The estimated continuum is removed to improve readability.

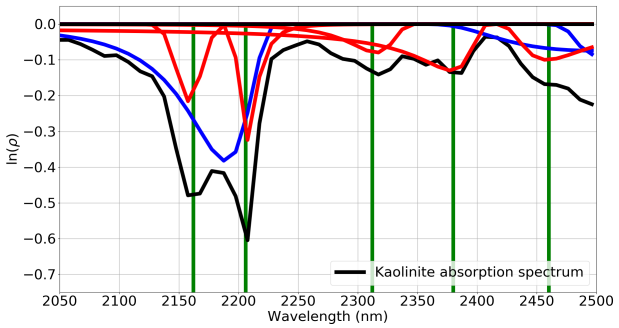


Fig. 7: SWIR kaolinite absorption spectrum (continuum removed, black), diagnostic and secondary EGO Gaussians (red), other Gaussians (blue). The true positions of the diagnostic and secondary absorptions are shown as vertical lines.

Three EGO Gaussians fit the main kaolinite doublet. The broad Gaussian at 2189 nm (in blue in Fig. 7) models the shape of the absorption while the other two (in red), narrow and deep, are located close to the theoretical positions of the doublet ($\delta\mu < 4$ nm). This result is explained by the fact that NNOMP first selects the broad Gaussian to capture the main trends of the spectrum.

Two EGO Gaussians fit the secondary absorptions at 2312 and 2380 nm. Their positions are close to the theoretical ones ($\delta\mu < 3$ nm). The absorptions being close, their amplitudes and shapes are difficult to separate. Thus, the EGO Gaussian centered at 2380 nm is asymmetric ($k = 0.43$), which reduces the amplitude of the one at 2312 nm.

The range at the edge of the spectrum (for the highest wavelengths) is modeled by three EGO Gaussians. The red Gaussian centered at 2458 nm corresponds to an absorption embedded in the spectrum that is not characteristic of kaolinite. Both Gaussians in blue, centered around 2500 nm, compensate for a continuum artifact. These three Gaussians are not helpful for identification purpose since they either fit a non-diagnostic absorption or compensate for artifacts on the edge of the spectrum.

4) VNIR and SWIR absorptions: nontronite continuum and absorptions: Fig. 8 shows the results for nontronite.

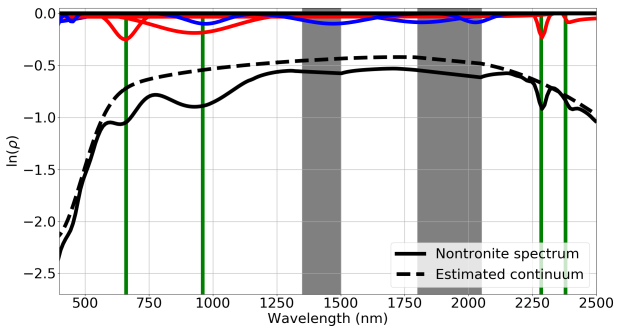


Fig. 8: Nontronite spectrum (black), estimated continuum (black-dashed), diagnostic and secondary EGO Gaussians (red), other Gaussians (blue). The true positions of the diagnostic and secondary absorptions are shown as vertical lines.

The estimated continuum lays above the spectrum and fits the overall shape. Several artifacts appear due to the presence of broad absorptions in the VNIR. Three small and narrow EGO Gaussians are selected in the VNIR edge of the spectrum. Their positions being close to 450 nm, they cannot be interpreted as absorptions. The artifact between 600 and 1250 nm impacts the amplitude of the absorptions at 660 and 960 nm. A smaller EGO Gaussian compensates it at 971 nm. The broad artifact between 1250 and 2100 nm is fitted by three EGO Gaussians centered within the spectral masks (at 1470, 1816, and 2027 nm).

Nevertheless, all the absorptions of interest are retrieved and fitted by unique EGO Gaussians, highlighted in red in Fig. 8. Specifically, one can distinguish two Gaussians in the VNIR, the Gaussian of largest width being slightly shifted from its theoretical position at 960 nm. Two other Gaussians are found in the SWIR part of the spectrum. They correspond to the diagnostic absorption at 2283 nm and the secondary one at 2378 nm. Positions and widths are consistent with theoretical knowledge. However, due to a gap between the continuum and the spectrum, the shape of the secondary absorption is modified. This results in a strong asymmetry of the Gaussian at 2378 nm ($k = -0.73$).

5) Spectral mask: gypsum absorptions: Fig. 9 presents the results for gypsum. The estimated absorptions located outside and within the spectral mask are shown in Fig. 9(a) and Fig. 9(b), respectively. Note that the absorptions at 1538 and 1750 nm are close to the spectral masks. Therefore, they are more difficult to retrieve.

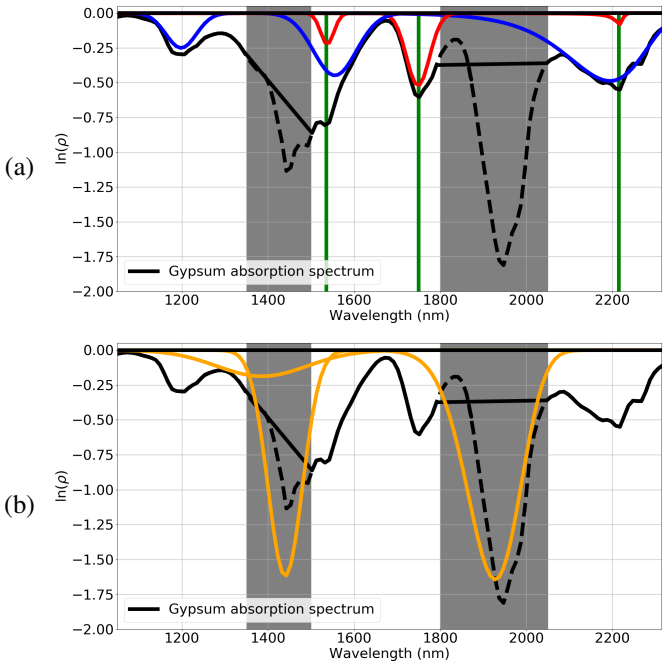


Fig. 9: Gypsum absorption data spectrum including missing data in the spectral masks (plain black; continuum has been removed. The dashed curve indicates the full absorption spectrum). (a). Diagnostic and secondary EGO Gaussians (red), and other EGO Gaussians (blue). The positions of true absorptions are shown as vertical lines. (b) Estimated EGO Gaussians located in the spectral mask (orange).

In Fig. 9(a), the diagnostic absorption at 1750 nm is fitted by an EGO Gaussian whose position and shape are slightly impacted by missing data within the spectral mask ($\delta\mu = 2$ nm). Regarding the secondary absorption at 1538 nm, two Gaussians are retrieved. The first, centered at 1555 nm (in blue), is wide and deep and fits the right edge of the absorption. The second, of lower amplitude and width, gives the exact position of the absorption (in red). At last, the broad absorption at 2215 nm is fitted by a broad EGO Gaussian giving the shape of the absorption and a smaller one, in amplitude and width, positioned at the center of the absorption ($\mu = 2218$ nm). Note that an absorption that is not characteristic of gypsum is retrieved at 1198 nm.

In Fig. 9(b), the edges of each spectral mask centered around 1400 and 1900 nm, respectively, are fitted by unique EGO Gaussians whose positions are close to the center of the masks. Their shape and position are consistent with the spectral mask's theoretical absorptions, shown in the dashed curve. A third (broad) Gaussian is retrieved at 1386 nm, which models the left edge of this spectral mask.

This example shows that it is possible to retrieve absorptions centered close to the spectral masks despite the missing data. The NNOMP algorithm indeed selects EGO Gaussians centered in the spectral mask in order to adjust the edges of the spectrum.

6) Discrimination of close absorptions: calcite and dolomite SWIR absorptions: Fig. 10 gathers the results for calcite and dolomite spectra, with similar spectral shapes. We focus on their close absorptions (diagnostic is separated by 18.5 nm and secondary by 16 nm). Given the AVIRIS spectral resolution, *i.e.*, 10 nm, it is theoretically possible to distinguish the two minerals based on their diagnostic absorption positions. For improved readability, the estimated continuum is removed from the spectrum of Fig. 10.

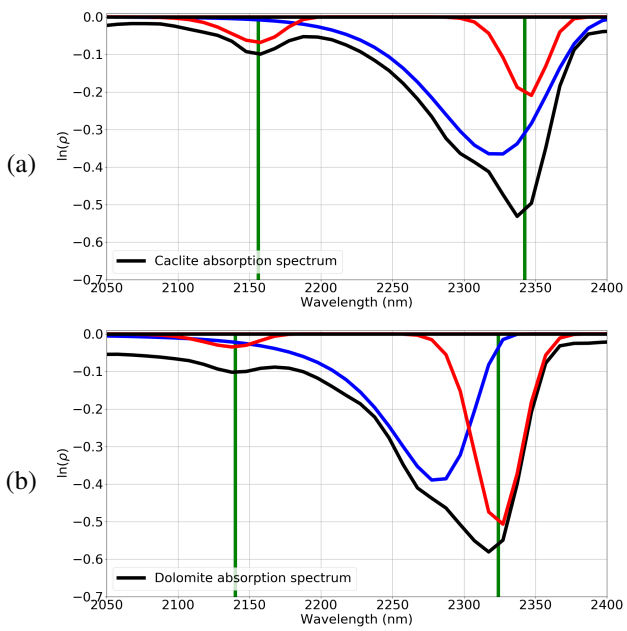


Fig. 10: Calcite (a) and dolomite (b) absorption spectrum (black), diagnostic and secondary absorptions (red) and other absorptions (blue). The positions of the true diagnostic and secondary absorption features are shown as vertical lines.

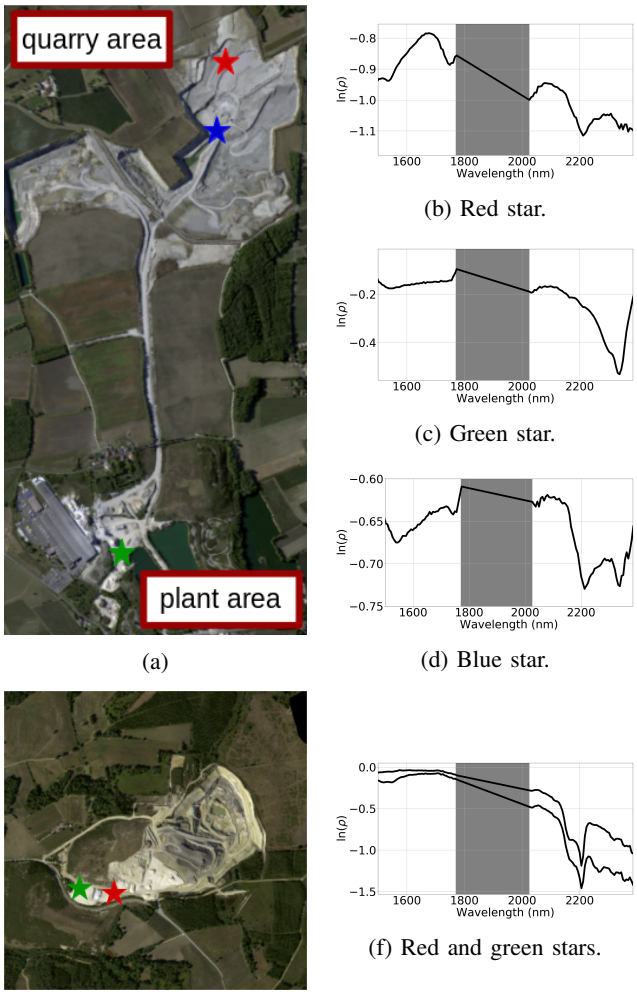
The absorptions of both spectra are retrieved similarly. Two EGO Gaussians fit the main diagnostic absorption, and a separate one is related to the secondary absorption. Similar to kaolinite, the blue Gaussian, broad and asymmetric ($\sigma = 37.5$ nm and $k = 0.13$ for calcite, $\sigma = 27.8$ nm and $k = 0.22$ for dolomite) models the main trend of the diagnostic absorption. The narrower Gaussian in red ($\sigma = 13.7$ nm for calcite and $\sigma = 16.5$ nm for dolomite) is located at the theoretical center of the absorption ($\delta\mu < 2$ nm). Its amplitude varies depending on how the Gaussian shape models the trend. Thus, for calcite, the amplitude of the Gaussian is lower than for dolomite. The secondary absorption is found close to the theoretical position ($\delta\mu < 2$ nm) with a visually coherent shape. However, for dolomite, its amplitude is attenuated by the presence of a wide Gaussian (not represented in Fig. 10(b)), centered before 2050 nm and resulting from a continuum estimation artifact.

This example shows that the diagnostic and secondary absorptions of calcite and dolomite could by successfully retrieved by the method and be used to distinguish calcite from dolomite.

C. Results on hyperspectral images

1) Campaign and data description: The goal is to apply the proposed method to real world image data acquired from airborne sensors. The hyperspectral data of Fig. 11 were acquired during a dedicated campaign over two quarries in France, with HySpex cameras (https://www.hyspex.com). Three minerals of interest are present with specific spectral signatures: gypsum, carbonates (Cherves-Richemont), and kaolinite (Chevanceaux). The minerals entail different physical and chemical characteristics and are used in the production of plasterboard or as aggregates (Cherves-Richemont) and refractory ceramics (Chevanceaux). Images were acquired in September 2019, with 0.5 m and 1.0 m spatial resolution in the VNIR and SWIR ranges, respectively. The number of spectral bands and the spectral resolution in the VNIR and SWIR are 160 and 162 and ~ 4 and 7 nm, respectively. In this study, only SWIR images were used. Images were atmospherically corrected using the ATCOR4 algorithm [65]. Reflectance images were spatially downsampled to 5 m to improve the SNR. Two spectral masks were applied around 1400 and 1900 nm to avoid the effects of saturated atmospheric water vapor absorptions. Representative samples were collected after image acquisition aiming to create a database gathering the possible spectral signatures in the quarries. An ASD FieldSpec®FR3 was used for laboratory measurements.

We analyzed a mixture of gypsum and carbonate (here calcite) and spectra of kaolinite. To apply the method on SWIR spectra, we removed c_1 and G_{uv} from the continuum model (2), and generated a NNOMP dictionary based on the procedure described in section III-C4, without EGO Gaussians centered shorter than 1500 nm. The noise covariance $\mathbf{\Sigma}$ was estimated using HySime [46].



(a) (b) Red star. (c) Green star. (d) Blue star. (e) (f) Red and green stars.

Fig. 11: Hyperspectral images acquired over (a) Cherves-Richemont and (e) Chevanceaux quarries. (b) Gypsum spectrum on the quarry layer. (c) Calcite spectrum from piles used as aggregates. (d) Gypsum-calcite mixture. (f) Kaolinite spectra from two piles.

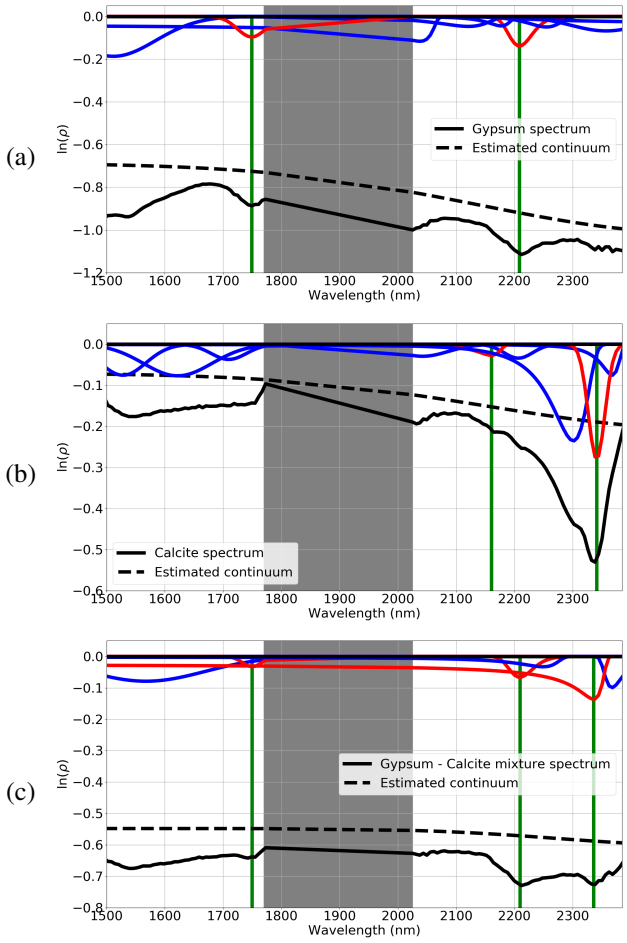


Fig. 12: Deconvolution of gypsum (a) and calcite (b) spectra, and a mixture of both (c). The true positions of the diagnostic and secondary absorptions are shown as vertical lines.

2) Spectral mixture of gypsum and carbonates: In the Cherves-Richemont image, gypsum and calcite can be easily identified from their reflectance spectra (Figs. 11(b,c)). However, for some areas, especially near the roads and the plants, gypsum and calcite constitute mixtures. This results in a spectrum including all the diagnostic and secondary absorptions of gypsum and calcite (Fig. 11(d)). We thus apply the method to three spectra shown in Fig. 12.

The results on gypsum and calcite can be easily compared to the evaluation on laboratory data, see sections V-B5 and V-B6. In Fig. 12(a), the gypsum absorption at 1750 nm is retrieved at $\mu = 1749$ nm with an amplitude of 0.16 (*i.e.*, the sum of all estimated EGO Gaussians at 1749 nm). The broad absorption at 2215 nm is retrieved at $\mu = 2208$ nm using three EGO Gaussians. Also, the continuum parameter c_0 is equal to 0.67. In Fig. 12(b), as expected, the calcite diagnostic absorption is fitted by two EGO Gaussians at ~ 2269 nm, and one for the theoretical position ($\mu = 2341$ nm). The secondary absorption at 2156 nm is retrieved at $\mu = 2160$ nm. The continuum parameter c_0 is equal to 0.07.

The diagnostic absorptions of gypsum and calcite and the secondary absorption of gypsum at 2215 nm are retrieved in the mixture (Fig. 12(c)). First, the absorption at 1750 nm is retrieved at $\mu = 1749$ nm with an amplitude of 0.08 (*i.e.*, the sum of all estimated EGO Gaussians at $\mu = 1749$ nm), which is half the one of pure gypsum. Then, the absorption at 2215 nm is also retrieved at $\mu = 2209$ nm. The calcite absorption at 2340 nm does not match the shape of pure calcite, and only one asymmetric EGO Gaussian is retrieved at 2235 nm. Finally, the continuum parameter c_0 is equal to 0.54, which is between the values for gypsum and calcite.

Based on the estimated EGO Gaussian positions and amplitudes, one can identify the spectrum as a gypsum-calcite mixture (positions) and possibly evaluate the relative concentration of both (amplitudes and c_0). This illustrates the high potential of spectral deconvolution for mineralogical interpretation.

3) Deconvolution of kaolinite spectra: On the Chevanceaux image, kaolinite is easily identifiable from its reflectance spectra (Fig. 11(f)). However, depending on the pile, the kaolinite doublet is more or less well-formed, which may

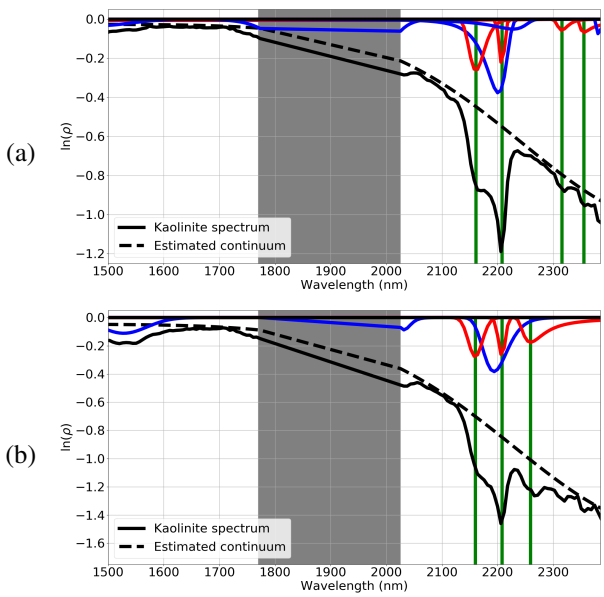


Fig. 13: Deconvolution of kaolinite spectra from two piles. The true positions of the diagnostic and secondary estimated absorptions are shown as vertical lines.

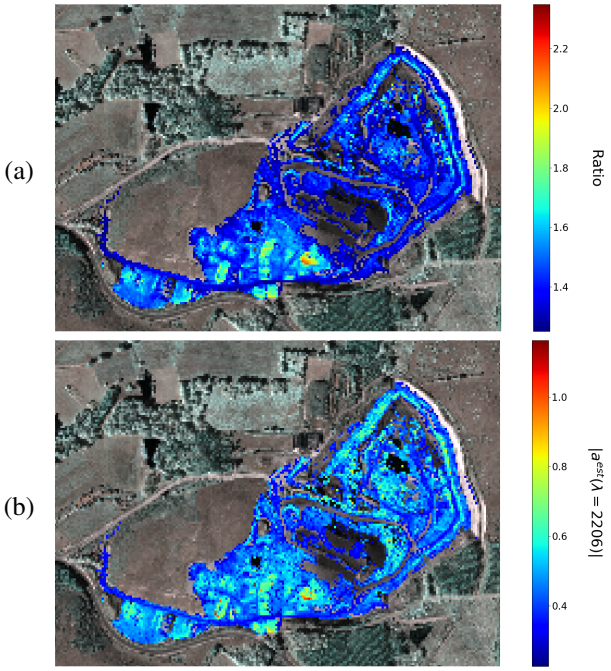


Fig. 14: Kaolinite mapping at Chevanceaux obtained by (a) the band ratioing technique and (b) the greedy-AGM procedure. Piles are clearly visible in the lower left part of the quarry.

be related to well-ordered kaolinite [10]. Results for each spectrum of Fig. 11(f) are shown in Fig. 13. They can be easily compared to those of the laboratory data in paragraph V-B3. Note that an absorption at 2260 nm, probably due to the presence of gibbsite ($AlOOH$), is visible in the spectrum of Fig. 13(b).

For both spectra, the absorption doublet is fitted by a broad EGO Gaussian (in blue) capturing the main shape and two narrow ones (in red). The broad Gaussian is positioned at 2201 and 2193 nm with a width equal to 17.6 and 21.5 nm, respectively. The two narrow EGO Gaussians in red model the absorption doublet and are shifted by less than 1 nm from their true positions. For the spectrum of Fig. 13(a), the small absorptions at 2315 and 2355 nm are well retrieved. On the contrary, in Fig. 13(b), the strong asymmetry of the gibbsite absorption at 2259 nm does not make it possible to detect the small absorptions.

Similar to the gypsum-calcite mixture from the Cherves-Richemont image in paragraph V-C2, the diagnostic absorptions of gibbsite and kaolinite are retrieved, which allows one to identify these two minerals. Also, the kaolinite absorption doublet is fitted similarly for both spectra. So, studying the evolution of its shape over the piles should be possible.

4) Kaolinite mapping: The spatial distribution of kaolinite can be mapped, see Fig. 14. For this purpose, let us first define the following band ratio:

$$R = \frac{0.275\rho(\lambda = 2127\,\text{nm}) + 0.725\rho(\lambda = 2236\,\text{nm})}{\rho(\lambda = 2206\,\text{nm})}. \quad (12)$$

This ratio highlights the presence of kaolinite by quantifying the depth of its absorption band at 2206 nm.

First, a spatial mask is empirically defined to keep only the pixels inside the quarry. The band ratio is calculated for all selected pixels and only those having a value greater than 1.25 are kept (corresponding to 6721 pixels). We thus obtain the map in Fig. 14(a) where the different piles are visible (higher ratios). Then, greedy-AGM is applied to the same pixels. Here, we calculate the sum of the amplitudes of the broad EGO Gaussian and the right narrow EGO Gaussian of the doublet, around 2200 and 2206 nm, respectively (blue and red Gaussians in Fig. 13). Fig. 14(b) shows the resulting map, where the piles are also visible.

The visual comparison of both maps demonstrates the potential of greedy-AGM for mineral mapping. It is worth mentioning that several maps of mineralogical parameters can be generated since one run of greedy-AGM yields the full set of EGO parameters, unlike simpler methods such as band ratioing.

VI. Discussion and Conclusion

A. Highlights of the greedy-AGM approach

In the literature, several methods were proposed to retrieve the parameters of physical models from mineral reflectance spectra for feature reduction and classification purpose [16] or for mineral identification and characterization [29], [34]. These methods suffer from several shortcomings as they need empirical initialization or tuning of algorithm parameters [30], [34], [35]. In particular, estimating the number of absorption features is known to be complicated. Also, propagation of errors often arises with methods that sequentially retrieve the continuum and absorption parts of the spectrum, leading to possible false detection of absorption features [37] and noise sensitivity [34], [37].

We proposed a new automatic approach to retrieve the parameters of physical models. The greedy-AGM procedure

includes several substantial improvements over classical methods. First, the number of absorption features is automatically estimated using a greedy algorithm coupled with an information-theoretic criterion. Second, the continuum estimation step outperforms classical geometric approaches [39] for noisy data and limits the creation of artifacts. Third, the method is easy to use, as no pre-processing of data is required. Automatic rules for setting the algorithmic parameters (number of dictionary atoms, maximum number of absorption features) are given in subsection III-E. We stress that no expert knowledge is necessary for practical use of greedy-AGM.

B. Experimental validation and perspectives

Another significant contribution of the paper is the validation of the greedy-AGM procedure on a rich collection of spectra, including synthetic and laboratory data and real hyperspectral images related to various minerals. This extensive validation illustrates the potential of the method to extract useful spectral information of minerals from noisy reflectance spectra. Greedy-AGM is easily adaptable to absorption features representative of various difficult situations that can be encountered in spectral signatures of minerals: broad and shallow absorptions such as those related to iron in the VNIR, narrow and sharp, close absorptions like the kaolinite doublet, absorptions located close to the atmospheric water vapor absorptions, and asymmetric absorptions. The number of absorptions and their respective parameters were well retrieved. For example, errors on the centers of the absorption features are less than 20 nm and 5 nm for VNIR and SWIR absorptions, respectively. We also demonstrated that the method could deal with complex cases of missing hyperspectral data related to spectral masks.

When applied to hyperspectral images, greedy-AGM was able to handle mixtures of gypsum and calcite, and the kaolinite absorption doublet was well retrieved. We plan to produce maps of diverse geological environments by intensively applying the greedy-AGM procedure in future work. Another critical perspective is to develop an automatic identification procedure from the retrieved EGO parameters. A solution based on a fuzzy logics is currently under study [66].

ACKNOWLEDGMENT

The authors acknowledge the LPG Nantes and GEOFIT EXPERT for the acquisition of the HySpex images. Imerys and Garandeau firms are thanked for their collaboration.

REFERENCES

[1] K. Cawse-Nicholson, P. A. Townsend, D. Schimel, and A. M. A. *et al*, "Nasa's surface biology and geology designated observable: A perspective on surface imaging algorithms," *Remote Sens. Environ.*, vol. 257, p. 112349, 2021.

[2] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 44, no. 2, pp. 127–143, 1993.

[3] L. Hamlin, R. O. Green, P. Mouroulis, M. Eastwood, D. Wilson, M. Dudik, and C. Paine, "Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments," in *Aerospace Conference*, Mar. 2011, pp. 1–7.

[4] R. Loizzo, R. Guarini, F. Longo, T. Scopa, R. Formaro, C. Facchinetti, and G. Varacalli, "Prisma: The Italian hyperspectral mission," in *IEEE Int. Geosc. Rem. Sens. Symp.*, Jul. 2018, pp. 175–178.

[5] L. Guanter, H. Kaufmann, K. Segl, and S. F. *et al*, "The EnMAP spaceborne imaging spectroscopy mission for Earth observation," *Remote Sensing*, vol. 7, no. 7, pp. 8830–8857, 2015.

[6] E. Bedini, "The use of hyperspectral remote sensing for mineral exploration: a review," *J. Hyperspectral Remote Sens.*, vol. 7, no. 4, pp. 189–211, 2017.

[7] J. Adams, "Interpretation of visible and near-infrared diffuse reflectance spectra of pyroxenes and other rock-forming minerals," *Infrared and Raman spectroscopy of lunar and terrestrial minerals*, pp. 91–116, 1975.

[8] G. R. Hunt, "Sprectral signatures of particulate minerals in the visible and near infrared," *Geophys.*, vol. 42, no. 3, pp. 501 – 513, 1977.

[9] R. N. Clark, T. V. V. King, M. Klejwa, G. A. Swayze, and N. Vergo, "High spectral resolution reflectance spectroscopy of minerals," *J. Geophys. Res. Solid Earth*, vol. 95, no. B8, pp. 12 653–12 680, 1990.

[10] R. N. Clark, *Manual of remote sensing, volume 3, remote sensing for the Earth sciences*. John Wiley and Sons, New York, 1999, ch. 1: Spectroscopy of rocks and minerals, and principles of spectroscopy, pp. 3–58.

[11] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, 2013.

[12] F. Kruse, J. Boardman, and J. Huntington, "Comparison of airborne hyperspectral data and eo-1 hyperion for mineral mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1388–1400, 2003.

[13] J. Theiler, A. Ziemann, S. Matteoli, and M. Diani, "Spectral variability of remotely sensed target materials: Causes, models, and strategies for mitigation and robust exploitation," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 8–30, 2019.

[14] S. Asadzadeh and C. R. de Souza Filho, "A review on spectral processing methods for geological remote sensing," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 47, pp. 69 – 90, 2016.

[15] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *Int. J. Appl. Earth Obs.*, vol. 8, no. 1, pp. 3 – 17, 2006.

[16] L. Pompilio, M. Pepe, G. Pedrazzi, and L. Marinangeli, "Informational clustering of hyperspectral data," *IEEE J. Sel. Top. Appl. Earth. Obs. Remote Sens.*, vol. 7, no. 6, pp. 2209–2223, 2014.

[17] A. Baldridge, S. Hook, C. Grove, and G. Rivera, "The ASTER spectral library version 2.0," *Remote Sens. Environ.*, vol. 113, no. 4, pp. 711–715, 2009.

[18] R. Kokaly, R. N. Clark, G. Swayze, K. E. Livo, T. Hoefen, N. Pearson, R. Wise, W. Benzel, H. Lowers, and R. L. Driscoll, "USGS Spectral Library Version 7," U.S. Geological Survey, Report, Apr. 2017.

[19] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "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, 2012.

[20] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1603–1616, Jul. 2011.

[21] 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 Sig. Proc. Mag.*, vol. 31, no. 1, pp. 82–94, 2014.

[22] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, 2019.

[23] T. Uezato, M. Fauvel, and N. Dobigeon, "Hyperspectral unmixing with spectral variability using adaptive bundles and double sparsity," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3980–3992, 2019.

[24] S. Pontual, N. Merry, and P. Gamson, *Spectral interpretation field manual. GMEX spectral analysis guides for mineral exploration*, 3rd ed. Victoria: AusSpec Int. Ltd., 2008, vol. 1.

[25] J. Mustard and J. Sunshine, *Spectral analysis for earth science investigation*, 3rd ed. John Wiley, Mar. 1999, vol. 3, ch. 5.

[26] S. Chabrillat, A. Eisele, S. Guillaso, A. Eisele, C. Rogaß, E. Ben-Dor, and H. Kaufmann, "HYSOMA: An easy-to-use software interface for soil mapping applications of hyperspectral imagery," *Proceedings of the 7th EARSeL SIG Imaging Spectroscopy Workshop*, 2011.

[27] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. M. R. R., and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res. Planets*, vol. 108, no. E12, p. 5131, 2003.

[28] C. Mielke, C. Rogass, N. Boesche, K. Segl, and U. Altenberger, "Engeomap 2.0—automated hyperspectral mineral identification for the german enmap space mission," *Remote Sensing*, vol. 8, no. 2, 2016.

[29] J. M. Sunshine, C. M. Pieters, and S. F. Pratt, "Deconvolution of mineral absorption bands: An improved approach," *J. Geophys. Res. Solid Earth*, vol. 95, no. B5, pp. 6955–6966, 1990.

[30] L. Pompilio, G. Pedrazzi, M. Sgavetti, E. A. Cloutis, M. A. Craig, and T. L. Roush, "Exponential Gaussian approach for spectral modeling: The EGO algorithm I. Band saturation," *Icarus*, vol. 201, no. 2, pp. 781 – 794, 2009.

[31] C. Hecker, F. J. A. van Ruitenbeek, H. M. A. van der Werff, W. H. Bakker, R. D. Hewson, and F. D. van der Meer, "Spectral absorption feature analysis for finding ore: A tutorial on using the method in geological remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 51–71, Jun. 2019.

[32] L. Pompilio, G. Pedrazzi, E. A. Cloutis, M. A. Craig, and T. L. Roush, "Exponential Gaussian approach for spectral modelling: The EGO algorithm II. Band asymmetry," *Icarus*, vol. 208, no. 2, pp. 811 – 823, 2010.

[33] R. Marion and V. Carrère, "Mineral mapping using the Automatized Gaussian Model (AGM) - Application to two industrial French sites at Gardanne and Thann," *Remote Sens.*, vol. 10, no. 1, 2018.

[34] H. D. Makarewicz, M. Parente, and J. L. Bishop, "Deconvolution of VNIR spectra using modified Gaussian modeling (MGM) with automatic parameter initialization (API) applied to CRISM," in *IEEE Workshop on Hyperspectral Image and Signal Processing*, Aug. 2009, pp. 1–5.

[35] H. Clénet, P. Pinet, Y. Daydou, F. Heuripeau, C. Rosemberg, D. Baratoux, and S. Chevrel, "A new systematic approach using the Modified Gaussian Model: Insight for the characterization of chemical composition of olivines, pyroxenes and olivine–pyroxene mixtures," *Icarus*, vol. 213, no. 1, pp. 404 – 422, 2011.

[36] C. Verpoorter, V. Carrère, and J.-P. Combe, "Visible, near-infrared spectrometry for simultaneous assessment of geophysical sediment properties (water and grain size) using the Spectral Derivative–Modified Gaussian Model," *J. Geophys. Res. Earth Surf.*, vol. 119, no. 10, pp. 2098–2122, 2014.

[37] M. Brossard, R. Marion, and V. Carrère, "Deconvolution of SWIR reflectance spectra for automatic mineral identification in hyperspectral imaging," *Remote Sens. Lett.*, vol. 7, no. 6, pp. 581–590, 2016.

[38] M. A. McCraig, G. R. Osinski, E. A. Cloutis, and R. L. F. *et al*, "Fitting the curve in excel®: Systematic curve fitting of laboratory and remotely sensed planetary spectra," *Comp. Geosci.*, vol. 100, pp. 103–114, 2017.

[39] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *J. Geophys. Res. Solid Earth*, vol. 89, no. B7, pp. 6329–6340, 1984.

[40] M. L. Whiting, L. Li, and S. L. Ustin, "Predicting water content using Gaussian model on soil spectra," *Remote Sens. Environ.*, vol. 89, no. 4, pp. 535 – 552, 2004.

[41] P. Du, W. A. Kibbe, and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching," *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, Jul. 2006.

[42] B. Rivard, J. Feng, A. Gallie, and A. Sanchez-Azofeifa, "Continuous wavelets for the improved use of spectral libraries and hyperspectral data," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2850 – 2862, 2008.

[43] R. Rialland, C. Soussen, R. Marion, and V. Carrère, "OMP-based algorithm for mineral reflectance spectra deconvolution from hyperspectral images," in *IEEE Int. Geosc. Rem. Sens. Symp.*, 2020, pp. 4918–4921.

[44] V. L. Mulder, M. Plötze, S. de Bruin, M. E. Schaepman, C. Mavris, R. F. Kokaly, and M. Egli, "Quantifying mineral abundances of complex mixtures by coupling spectral deconvolution of SWIR spectra (2.1–2.4 μm) and regression tree analysis," *Geoderma*, vol. 207-208, pp. 279 – 290, 2013.

[45] J. M. Sunshine and C. M. Pieters, "Estimating modal abundances from the spectra of natural and laboratory pyroxene mixtures using the Modified Gaussian Model," *J. Geophys. Res. Planets*, vol. 98, no. E5, pp. 9075–9087, 1993.

[46] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[47] M. J. D. Powell, "Direct search algorithms for optimization calculations," *Acta Numerica*, vol. 7, p. 287–336, 1998.

[48] M. Elad, *Sparse and redundant representations: From theory to applications in signal and image processing*. New York: Springer, Aug. 2010.

[49] P. Stoica and Y. Selén, "Model-order selection: A review of information criterion rules," *IEEE Sig. Proc. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.

[50] A. M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.

[51] T. T. Nguyen, J. Idier, C. Soussen, and E. Djermoune, "Non-negative orthogonal greedy algorithms," *IEEE Trans. Signal Process.*, vol. 67, no. 21, pp. 5643–5658, Nov. 2019.

[52] C. Soussen, J. Idier, J. Duan, and D. Brie, "Homotopy based algorithms for ℓ_0-regularized least-squares," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3301–3316, Jul. 2015.

[53] F. De Ridder, R. Pintelon, J. Schoukens, and D. P. Gillikin, "Modified AIC and MDL model selection criteria for short data records," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 144–150, Feb. 2005.

[54] M. A. Branch, T. F. Coleman, and Y. Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 1–23, 1999.

[55] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : III. Oxides and hydroxides," *Mod. Geol.*, vol. 2, pp. 195–205, 1971.

[56] G. R. Hunt and J. W. Salisbury, "Visible and near-infrared spectra of minerals ans rocks : I. Silicate minerals," *Mod. Geol.*, vol. 1, pp. 283–300, 1970.

[57] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : VI. Additional silicates," *Mod. Geol.*, vol. 4, pp. 85–106, 1973.

[58] J. L. Bishop, M. D. Lane, M. D. Dyar, and A. J. Brown, "Reflectance and emission spectroscopy study of four groups of phyllosilicates: Smectites, kaolinite-serpentines, chlorites and micas," *Clay Minerals*, vol. 43, no. 1, pp. 35–54, Mar. 2008.

[59] J. Flahaut, M. Martinot, J. L. Bishop, G. R. Davies, and N. J. Potts, "Remote sensing and in situ mineralogic survey of the Chilean salars: An analog to Mars evaporate deposits?" *Icarus*, vol. 282, pp. 152 – 173, 2017.

[60] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : IV. Sulphides and sulphates," *Mod. Geol.*, vol. 3, pp. 1–14, 1971.

[61] E. A. Cloutis, F. C. Hawthorne, S. A. Mertzman, K. Krenn, M. A. Craig, D. Marcino, M. Methot, J. Strong, J. F. Mustard, D. L. Blaney, J. F. Bell, and F. Vilas, "Detection and discrimination of sulfate minerals using reflectance spectroscopy," *Icarus*, vol. 184, no. 1, pp. 121 – 157, 2006.

[62] G. R. Hunt and J. W. Salisbury, "Visible and near-infrared spectra of minerals ans rocks : II. Carbonates," *Mod. Geol.*, vol. 2, pp. 23–30, 1971.

[63] S. J. Gaffey, "Spectral reflectance of carbonates minerals in the visible and near-infrared $(0.35 - 2.55$ microns$)$: Calcite, aragonite and dolomite," *Am. Minerals*, vol. 712, pp. 151–162, 1986.

[64] E. A. Cloutis, S. E. Grasby, W. M. Last, R. Léveillé, G. R. Osinski, and B. L. Sherriff, "Spectral reflectance properties of carbonates from terrestrial analogue environments: Implications for Mars," *Planet. Space Sci.*, vol. 58, no. 4, pp. 522 – 537, 2010.

[65] R. Richter and D. Schlapfer, "Atmospheric/topographic correction for airborne imagery: ATCOR-4 user Guide," *DLR IB*, pp. 565–02, 2012.

[66] R. Rialland, V. Carrère, R. Marion, and C. Soussen, "Identification of minerals from hyperspectral imaging based on a fuzzy logic approach," in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVII*, M. Velez-Reyes and D. W. Messinger, Eds., vol. 11727. SPIE, 2021, pp. 108 – 119.