



**HAL**  
open science

## **Exponential Gaussian approach for spectral modeling: the EGO algorithm, I. Band saturation**

Loredana Pompilio, Giuseppe Pedrazz, Maria Sgavetti, Edward A. Cloutis, Michael A. Craig, Ted L. Roush

### ► **To cite this version:**

Loredana Pompilio, Giuseppe Pedrazz, Maria Sgavetti, Edward A. Cloutis, Michael A. Craig, et al.. Exponential Gaussian approach for spectral modeling: the EGO algorithm, I. Band saturation. *Icarus*, 2009, 201 (2), pp.781. <10.1016/j.icarus.2009.01.022>. <hal-00533503>

**HAL Id: hal-00533503**

**<https://hal.science/hal-00533503v1>**

Submitted on 7 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## Accepted Manuscript

Exponential Gaussian approach for spectral modeling: the EGO algorithm, I. Band saturation

Loredana Pompilio, Giuseppe Pedrazz, Maria Sgavetti, Edward A. Cloutis, Michael A. Craig, Ted L. Roush

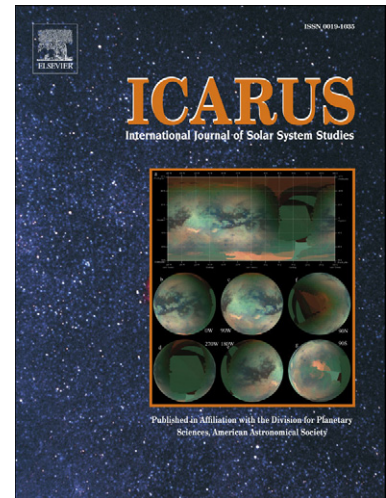
PII: S0019-1035(09)00043-8  
DOI: [10.1016/j.icarus.2009.01.022](https://doi.org/10.1016/j.icarus.2009.01.022)  
Reference: YICAR 8903

To appear in: *Icarus*

Received date: 21 May 2008  
Revised date: 14 January 2009  
Accepted date: 24 January 2009

Please cite this article as: L. Pompilio, G. Pedrazz, M. Sgavetti, E.A. Cloutis, M.A. Craig, T.L. Roush, Exponential Gaussian approach for spectral modeling: the EGO algorithm, I. Band saturation, *Icarus* (2009), doi: 10.1016/j.icarus.2009.01.022

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Exponential Gaussian approach for spectral modeling: the EGO  
algorithm, I. Band saturation

Loredana Pompilio<sup>1,\*</sup>, Giuseppe Pedrazzi<sup>2</sup>, Maria Sgavetti<sup>1</sup>, Edward A. Cloutis<sup>3</sup>,  
Michael A. Craig<sup>3,4</sup>, and Ted L. Roush<sup>5</sup>

<sup>1</sup> Dipartimento di Scienze della Terra, University of Parma, Via Usberti n. 157A,  
Parma, I-43100, Italy

\* Corresponding Author E-mail address: loredana.pompilio@unipr.it

<sup>2</sup> Dipartimento di Sanità Pubblica, Sezione di Fisica, University of Parma, Via  
Voturno 39, Parma, I-43100, Italy

<sup>3</sup> Department of Geography, University of Winnipeg, 515 Portage Ave.,  
Winnipeg, MB, Canada R3B 2E9

<sup>4</sup> Now at the University of Western Ontario

<sup>5</sup> NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035-1000, USA

Pages: 44

Tables: 3

Figures: 9

**Proposed Running Head:** Spectral EGO modeling

**Editorial correspondence to:**

Dr. Loredana Pompilio

Dipartimento Scienze della Terra, Università di Parma

Via Usberti 157/A, 43100 Parma, Italy

Phone: +39 0521905361

Fax: +39 0521905305

E-mail address: [loredana.pompilio@unipr.it](mailto:loredana.pompilio@unipr.it)

**Abstract**

Curve fitting techniques are a widespread approach to spectral modeling in the VNIR range (Burns, 1970; Singer, 1981; Roush and Singer, 1986; Sunshine *et al.* 1990). They have been successfully used to model reflectance spectra of powdered minerals and mixtures, natural rock samples and meteorites, and unknown remote spectra of the Moon, Mars and asteroids. Here, we test a new decomposition algorithm to model VNIR reflectance spectra and call it Exponential Gaussian Optimization (EGO). The EGO algorithm is derived from and complementary to the MGM of Sunshine *et al.* (1990). The general EGO equation has been especially designed to account for absorption bands affected by saturation and asymmetry. Here we present a special case of EGO and address it to model saturated electronic transition bands.

Our main goals are: 1) to recognize and model band saturation in reflectance spectra; 2) to develop a basic approach for decomposition of rock spectra, where effects due to saturation are most prevalent; 3) to reduce the uncertainty related to quantitative estimation when band saturation is occurring.

In order to accomplish these objectives, we simulate flat bands starting from pure Gaussians and test the EGO algorithm on those simulated spectra first. Then we test the EGO algorithm on a number of measurements acquired on powdered pyroxenes having different compositions and average grain size and binary mixtures of orthopyroxenes with barium sulfate. The main results arising from this study are: 1) EGO model is able to numerically account for the occurrence of saturation effects on reflectance spectra of powdered minerals and mixtures; 2) the systematic dilution of a strong absorber using a bright neutral material is not responsible for band deformation.

Further work is still required in order to analyze the behavior of the EGO algorithm with respect to the saturation phenomena using more complex band shapes than pyroxene bands.

**Key Words:** Spectroscopy, Mineralogy

## 1. Introduction

Curve fitting is a widespread approach to model reflectance measurements and infer absorption components within complex bands. Electronic absorption features in reflectance spectra of minerals have usually been considered to be Gaussian-like (Burns, 1970; Singer, 1981; Roush and Singer, 1986), or modified Gaussians (Sunshine *et al.* 1990), when analyzed as log reflectance versus photon energy. Modified Gaussians describe random distribution of average bond lengths ( $r$ ) which is related to the energy via  $1/r$  (Sunshine and Pieters, 1993). However, it is recognized that vibronic coupling, as well as saturation phenomena and the signatures of amorphous materials and impurities may modify absorption band shapes and contribute to flat features in the minimum region and asymmetric profiles in crystal field spectra (Runciman *et al.* 1974; Nolet *et al.* 1979; Burns, 1993; Pieters and Tompkins, 2005; Nicholis *et al.* 2006). To isolate the spectral features from other effects, such as interactions with other absorbing materials, a continuum is used (Singer, 1981; Clark and Roush, 1984; Roush and Singer, 1986; Clark, 2003; Parente and Bishop, 2006). Nevertheless, the physical meaning of the continuum is not completely understood and as a result, its mathematical representation is the subject of active research.

So far, curve fitting techniques have been extensively used in order to decompose complex reflectance spectra of minerals and mixtures. The Modified Gaussian Model (MGM, Sunshine *et al.* 1990) has been successfully applied to model crystal field transitions due to transition elements in various crystal sites (Sunshine and Pieters, 1990; Sunshine and Pieters, 1991; Sunshine and Pieters, 1993; Sunshine *et al.* 1993; Pieters, 1996; Sunshine and Pieters, 1998; Klima *et al.* 2007), as well as both oxygen-metal and intervalence charge transfer bands (e.g., McFadden and Cline 2005). It has been used to

decompose overlapping absorptions in mafic minerals and mixtures (Mustard, 1992; Sunshine and Pieters, 1993), solid solution series (olivines and pyroxenes, Sunshine and Pieters, 1998), natural rock samples and meteorites (Sunshine *et al.* 1993; Schade and Wäsch, 1999; Sugihara *et al.* 2003; McFadden and Cline, 2005), and unknown remote spectra of the Moon and asteroids (Hiroi *et al.* 1995; Hiroi *et al.* 1996; Hiroi and Pieters, 1998; Hiroi *et al.* 2000; Binzel *et al.* 2001; Mayne *et al.* 2006). Recently, it has been implemented in the processing pipeline of OMEGA and CRISM data, for classification purposes (Pinet *et al.* 2007; Baratoux *et al.* 2007; Murchie *et al.*, 2007).

The bulk of samples so far modeled have been natural mixtures of particulate minerals, obtained via crushing the rock samples and sieving the particulate to select a specific grain size range. Few spectroscopic experiments applying the MGM have been performed on rock chips or slices (Schade and Wäsch, 1999; McFadden and Cline, 2005). In addition, the effects due to grain size variations (affecting both band saturation and asymmetry) of particulate mixtures have not been accurately modeled yet. Starting from the work of Craig *et al.* (2007), which showed the effects of the grain size on reflectance spectra, using powdered olivine, orthopyroxene and basalt, we performed a number of experiments using both real and simulated spectra.

In particular, Craig *et al.* (2007) have pointed out that the preferred orientation of non-equidimensional grains, band saturation and contamination due to other phases can affect the overall band shape (and in turn, depth) of the measured spectra. The preferred orientation of grains generally affects tabular or linear crystals, and is not significant for powdered pyroxene. Spectral contributions by accessory and minor phases of the host rock can sometimes be significant, especially for opaque minerals associated with weakly absorbing minerals. Contaminant phases can sometimes strongly modify reflectance

spectra, especially in the case of finely-dispersed Fe and/or Ti oxides which are typically associated with pyroxenes in mafic and ultramafic assemblages. Contamination can be significantly reduced during sample preparation of monomineralic samples. However, contamination will generally be more significant for coarser grain sizes, where the accessory and minor phases cannot be completely separated from the coarse grains.

Band saturation (Clark, 1999) is still an intriguing issue in spectroscopy and it deserves further exploration. The process (or processes) causing band saturation is related to both the amount of the absorbing species within a crystal and the crystal dimensions. When saturation takes place in the VNIR range, the absorption band is reduced in depth and becomes flatter near the minimum. This phenomenon increases the uncertainty in measuring the minimum and makes the band depth a non reliable parameter for quantitative measurements. In addition, the decomposition techniques so far used for spectral modeling purposes could fail to correctly discriminate between saturated absorption bands versus overlapping (Gaussian) absorption bands.

In the present investigation, we test a new decomposition algorithm for VNIR reflectance spectra, designed to model saturated electronic transition bands and the occurrence of neutral substances mixed with the mineral of interest. We call this new algorithm Exponential Gaussian Optimization (EGO) and we test its performances using real measurements on both actual particulate samples and simulations. The main objectives of our efforts are: 1) to test and validate a fitting technique for reflectance spectra which can recognize and model band saturation; 2) to develop a basic approach to decompose complex spectra of rocks, where saturation is most effective; 3) to reduce the uncertainty related to quantitative estimation when band saturation is occurring.

Since pyroxenes are among the most widespread rock-forming minerals within the Solar System, and their spectroscopic behavior is very well known, we used this mineral in attempting to model its spectral features. The strong absorptions of pyroxenes in the VNIR range are diagnostic of their major cation abundances (Cloutis & Gaffey, 1991; Burns, 1993). Pyroxene composition, especially the Fe/Mg ratio, and Ca abundance, as well as the crystallographic structure, can be quite easily estimated using the parameters both directly measured and mathematically derived from the spectral signatures. In addition, spectral decomposition techniques have been successfully used to determine the relative abundance of high- and low-Ca pyroxene in two-pyroxene mixtures (Sunshine *et al.* 1993; Sunshine & Pieters, 1993).

## 2. Modeling strategy

### 2.1 Overview

In the current practice of fitting techniques, band saturation is recognized by the symmetric distribution of residuals with respect to the central wavelength of the modeled band (Sunshine and Pieters, 1993). This criterion prevents introduction of additional Gaussians to the model, although it does not provide any numerical evaluation of the saturation phenomenon, nor does it provide a quantitative evaluation of when the residual distribution indicates that saturation occurs.

In the present investigation we use a multi-target and multi-technique approach aimed at testing the performances of the EGO algorithm as a spectral decomposition technique to be applied in the VNIR range for modeling band saturation. First, we simulate flat absorption bands using the superposition of two modified Gaussians of Sunshine *et al.* (1990) and then fit the resulting spectra with a single EGO profile. This

mathematical analysis aims to evaluate the reliability of the EGO algorithm to model the simulated effects due to band saturation. In addition, an estimate of the model accuracy is provided. Then, we apply the EGO modeling to spectral measurements of pure particulate pyroxenes affected by saturation and mixtures of pyroxenes with barium sulfate, for which the chemical composition and the relative abundances are known. This mathematical analysis aims to establish tolerance limits and variation trends for the different determined parameters. We also qualify the occurrence of saturation by using the results returned from the EGO modeling.

Model results are statistically evaluated and compared to assess the *best fit* models, using: (1) the standard error of the estimate; (2) coefficient of determination  $R^2$ ; (3) adjusted coefficient of determination  $R^2_{\text{adj}}$ ; (4) Akaike's Information Criterion (AIC); (5) correlation matrices; and, (6) analysis of residuals. All the statistics that were used are described in Appendix A. In addition, we evaluate the results according to reliability of the returned parameters.

## 2.2 Exponential Gaussian Optimization (EGO) algorithm

The EGO profile accounts for non-Gaussian behavior of absorption features that can result from nested bands or from the occurrence of saturation effects. Therefore, it is able to model band asymmetry and flattening, respectively. When flattening and asymmetry are negligible, the absorption band is commonly fitted as a single Gaussian or a sum of Gaussians, superimposed on some kind of background. The Gaussian shape is therefore the base function for this mathematical model.

Nested bands are weak absorptions which cannot be resolved neither at the temperature of the spectral measurement nor due to the low chemical abundance of the weaker absorber. When nested bands occur in spectra, they impart slight asymmetry to

the spectral shape. Similarly, when saturation effects become relevant, the band shape is no longer a Gaussian and the fitting process becomes more complex. In the absence of a theoretical model capable of providing an appropriate shape function in order to account for these phenomena, the temptation is to model those changes in shape with additional Gaussians. The higher the number of Gaussians the better is the fit, because more parameters are available to fit the data. Eventually, by using this strategy, we could achieve a very “accurate” fit (mathematically) but questions still remain regarding the physical significance of the additional Gaussians.

Before looking at complicated functions or a distribution of functions, we rely upon a slight modification of the basic Gaussian function by changing the Gaussian profile and symmetry while preserving some relevant parameters (i.e., center). The EGO profile is based on this approach and is summarized in equation 1:

$$EGO(\lambda) = \frac{s}{1 - e^{-\frac{t}{\sigma^2}}} \left[ 1 - e^{-\frac{1}{2} \left( \frac{\lambda - \mu}{\sigma} - \frac{k(\lambda - \mu)}{\sigma + k(\lambda - \mu)} \right)^2} \right] \quad (\text{Eq. 1})$$

where  $s$  is the band intensity,  $\mu$  the center and  $\sigma$  is the width of the EGO profile. The  $t$  parameter is used to model band flattening, and the  $k$  coefficient models asymmetry. The FWHM can be easily calculated via  $FWHM = 2.35482(\sigma)$ . In equation 1, the wavelength  $\lambda$  is expressed as micrometers, which is a common physical unit for spectral measurements within the visible and near infrared ranges. The multiplicative term to the right of the equal sign is a normalization coefficient, to keep the EGO onto the baseline. The EGO defined in this way has the advantage of keeping the same values of intensity, center, and width of the corresponding Gaussian:  $G(s, \mu, \sigma) = s \cdot \exp(-0.5 * [(\lambda - \mu)/\sigma]^2)$ , which is included in (Eq. 1). It can be demonstrated that the Exponential Gaussian profile becomes a Gaussian in wavelength for  $k = 0$  and  $t$  approaching 0 (see Appendix B).

In the following, we will refer to a special case of the general Eq. 1, where  $k = 0$ . Therefore, we will limit our discussion to the band saturation modeling and postpone the analysis of the symmetry of electronic bands to a separate paper. An example of the special EGO behavior is shown in figure 1 where  $t$  varies, while the other parameters are kept constant. Starting from an original Gaussian (where  $t$  is negligible), we progressively increase the effect of band flattening (Fig. 1), by increasing  $t$ . The resulting EGOs retain same center, width and depth as the original Gaussian, while their shapes are changing.

**[Figure 1]**

Starting from the MGM approach and its variants (Sunshine *et al.* 1991; Ueda *et al.* 2002; Parente and Bishop, 2006) and the recommendations on the continuum problem by Clark and Roush (1984), we used a number of EGOs in log reflectance versus wavelength space superimposed onto a continuum linear in wavenumber (Eq. 2).

$$C(\lambda) = c_0 + c_1(\lambda^{-1}) \quad (\text{Eq. 2})$$

Each EGO is intended to model a single absorption feature affected by saturation. Each EGO is described using four parameters (position, intensity, width and flattening coefficient) and the model algorithm is written as follows:

$$\ln R(\lambda) = C(\lambda) + \sum EGO(\lambda) \quad (\text{Eq. 3})$$

where ( $R$ ) is the reflectance spectrum as a function of the wavelength and  $C$  indicates the continuum as a function of wavelength.

In order to achieve high stability of the algorithm, increase speed of computation and to provide solid statistics for improving the level of significance of the fit (Appendix A), the analytical EGO routine has been implemented within the R-environment (R Development Core Team, 2005). R is an open source system for statistical computation

and graphics; it provides a very powerful and freely accessible programming language. In addition, the R code has been found to be very stable and easy to use. In accordance with the open source policy, it is available upon demand.

The initial approximation to the parameters is iterated a number of times via simultaneously optimizing all the parameters of the model function and minimizing the residual function using a Levenberg-Marquardt approach (Moré, 1978; Garbow *et al.* 1980 MINPACK Gnu GSL routines; Press *et al.* 1992 and references therein). We have performed our calculations using various methods other than the Levenberg-Marquardt, such as: BFGS (Broyden, Fletcher, Goldfarb, Shanno), L-BFGS-B [box-constrained, Byrd *et al.* (1994)], stochastic methods [Simulated Annealing (SA) and Genetic Algorithms (GA)] and a combination of stochastic and gradient methods (Rgenoud, R-package, <http://sekhon.berkeley.edu/rgenoud/>). All these methods, if properly used, find nearly identical solutions. In addition, all these methods, as well as the non-linear least squares inversion based on the stochastic technique of Tarantola and Valette (1982) implemented in the MGM original formulation (Sunshine *et al.* 1990) allow a number of constraints to be applied to the parameters. By using the more familiar statistical approach of the Levenberg-Marquardt routine, we do not substantially affect the philosophy of the process but noticeably increase the stability and speed of the computation.

Optimization using the Levenberg-Marquardt algorithm allows the iteration process to terminate when the standard deviation of the estimate is no longer improved beyond an established value, analogous to the original MGM routine. This is accomplished when three convergence parameters are simultaneously satisfied, as discussed in Moré (1978). To prevent the occurrence of a false minimum, we have run

each nonlinear regression several times, each time providing a different set of initial parameters and selecting among the set of regression parameters generated by different fits.

### 3. Spectral Analyses

#### 3.1 Synthetic simulations

The effects of band saturation have been simulated by superimposing two modified Gaussians of Sunshine *et al.* (1990) on a straight horizontal continuum, thereby avoiding the influence of background effects and poorly fitted continua. A Gaussian noise in the order of 1% of band intensity has been added to the resulting band, thus increasing reliability. The initial parameters of the component modified Gaussians have been modified in each simulation. A subset of the resulting spectra is shown in figure 2. We start the simulations with the Gaussian at shorter wavelengths having invariant parameters. We progressively vary the position of the longer wavelength Gaussian relative to the former. We fit the whole dataset using a single EGO (Eq. 1) superimposed onto a straight continuum. All the parameters except the continua coefficients were free to change during the optimization. The statistics attained with the various fitting models are very good. Correlation matrixes reveal poor correlation coefficients between parameters. The occurrence of poor correlation coefficients between couples of parameters retrieved by a fitting routine is the most favorable condition. When parameters display high correlation coefficients, the model is affected by collinearity. Collinearity has two main effects: a) we are no longer able to separate the contribution that each correlated parameter gives to our model. Thus, we cannot reliably establish whether the effects we are intended to model are attributed to a certain parameter or to

the correlated one; b) since the statistical uncertainty related to each parameter in the model mainly depends on those correlation coefficients, the higher the correlation, the stronger is the uncertainty.

Simulations allow assessing the band separation threshold at which band flattening can be recognized and modeled using the  $t$  parameter of Eq. 1. The calibration of  $t$  as a function of band separation returns a description of the flattening in spectra and the maximum and minimum thresholds, which allow the band flattening to be qualified. Since band flattening is diagnostic of saturation effects occurring in spectra, the saturation effect could be inferred from the  $t$  values of the EGO profiles and in turn the occurrence of additional absorptions rejected. In order to relate the band separation, and in turn the  $t$  parameter, to saturation phenomena, we must assess the threshold of distances between component bands which is modeled using a single EGO. That model will return the maximum  $t$  value which signifies band saturation.

As a strategy, we generated three separate datasets resulting from the superposition of two modified Gaussians having constant and equal depths ( $s_1 = s_2 = 0.1$ ) and variable centers and widths (Fig. 2a, b, c). The Gaussian components have the following constant parameters for each simulation: ( $s_1 = s_2 = 0.1$ ;  $\mu_1 = 0.95 \mu\text{m}$ ;  $\sigma_1 = \sigma_2 = 0.05 \mu\text{m}$ ;  $\sigma_1' = \sigma_2' = 0.07 \mu\text{m}$ ;  $\sigma_1'' = \sigma_2'' = 0.08 \mu\text{m}$ ). The longer wavelength Gaussian retains the same depths and widths as the shorter wavelength Gaussian, but is centered 7.5 nm away from the other and progressively moves up to a maximum separation of 150 nm, by steps of 7.5 nm. Depending on the band width used in each simulation, the separation threshold changes accordingly. Narrower Gaussians give rise to an additional detectable absorption in the resulting composite spectra well below the highest threshold

(Fig. 2a). Those occurrences have not been included in the computation. Besides centers, the two Gaussians are identical, thus preventing band asymmetry.

**[Figure 2]**

**[Figure 3]**

The fitting results show that as the separation between the two Gaussians increases, the band width of the composite absorption increases up to 12 nm more than the starting value per each simulation, centers shift toward longer wavelengths, regardless the initial width of the components, and intensities decrease (Fig. 3).

The deformation of band shape due to the presence of the second absorption band is mainly due to flattening around the minimum and thus it is modeled by increasing  $t$ . According to band widths,  $t$  increases following an exponential-like trend as the absorption components move away from one another (Fig. 3). In the simulation with  $\sigma_1 = \sigma_2 = 0.05 \mu\text{m}$  (Fig. 2a),  $t$  starts to increase when the separation of the two components is around 37 nm and reaches a maximum value of 7.0 when band separation reaches 105 nm. Distances up to 40 nm are not uniquely resolved by the  $t$  parameter. Separations larger than 105 nm require an additional band to be included in the model, as shown in figure 2a. As we introduce broader Gaussians in the simulation (Fig. 2b, c),  $t$  starts to increase for higher ranges of band separation (around 70 nm). These results indicate that the composite band used for fitting the spectral simulations is progressively flatter at the bottom.

**[Figure 4]**

If we accept this kind of simulation as representative of saturation effects, where the saturation could be the result of the superposition of several approximately equal Gaussians centered very close to each other, then we would have found a way to

parameterize the effects of saturation on spectra. Within the range of inferred saturation, the width of the composite band varies up to 12 nm and it tends to become higher for larger component Gaussians. This result points out that the saturation does not affect the width of the composite band appreciably, but becomes detectable using the  $t$  parameter of the EGO model. However, composite bands generated with these simulations show a slight increase of width as a function of  $t$ , up to a plateau level (Fig. 5a).  $t$  correlates non-linearly with band width up to 1.5 and 2.0. After that, band width becomes constant while  $t$  increases. Therefore, the plateau level can be used to retrieve the  $t$  values which are diagnostic of saturation effects. All the values above this threshold are highly diagnostic of saturation. The residual distribution can help us understand when the use of additional bands is required. As the distance between the component Gaussians approaches the band separation threshold, the EGO profile is no longer able to resolve separate bands, as shown in figure 4. A distribution of residuals scattered randomly about 0 (see models 1 to 16 in figure 4), suggest that no additional bands are required to fit the data. The residuals here are simply due to the random noise of simulated spectra. Where the residuals display fluctuations larger than the noise (as models 17 to 20 in figure 4), there is evidence of inadequate number of bands in the fit.

**[Figure 5]**

Non linear correlations exist also between  $t$  and the band depth (Fig. 5b). As the two components move away from each other, and in turn  $t$  increases, the composite band becomes weaker. This behavior depicts mathematically what is recognizable by looking at the spectral measurements affected by saturation effects (Fig. 6).

In the following sections, we apply the results of the simulations to laboratory spectra of powdered minerals and mixtures with different grain sizes and end member

proportions. We use natural minerals having different particle size ranges in order to establish the link between the range of  $t$  values indicating band saturation and grain size. The ultimate goal is to estimate the grain sizes from spectra of powdered samples. In order to yield as high as possible accuracy, we need to expand the range of measurements using a closer sampling of grain size splits and model the whole set of measurements. We need further work to accomplish this task.

## 3.2 Laboratory Analyses

### 3.2.1 Data

The data we used for testing the EGO fitting model start with a natural mineral characterized by a simple absorption band in a definite wavelength region, i.e. orthopyroxene. In this specimen, absorption of energy is due largely to  $\text{Fe}^{2+}$  located in the M2 crystallographic site (Burns, 1993). Contributions from  $\text{Fe}^{2+}$  in the M1 site (which causes absorption bands near 0.9 and 1.15  $\mu\text{m}$ ) are approximately an order of magnitude weaker than the M2 absorptions for a given  $\text{Fe}^{2+}$  content (Klima *et al.* 2007). To determine the effect of saturation on spectral data we used 5 powdered splits having different grain size ranges, varying between  $< 45$  and 500-1000  $\mu\text{m}$ . We also used two compositionally different orthopyroxenes to assess the role of differences in the abundance of the absorbing species ( $\text{Fe}^{2+}$  in the M2 site). We then increased the complexity by using binary mixtures made up of one pyroxene and a neutral bright material ( $\text{BaSO}_4$ ), in order to test the behavior of saturation with respect to dilution. Spectra were modeled using a single EGO and up to two modified Gaussians. The dataset and the experimental strategy are described in detail below.

The reflectance spectra used in this investigation have been acquired relative to Spectralon® with an Analytical Spectral Device (ASD) FieldSpec Pro HR field portable spectrophotometer at the University of Winnipeg. The spectral range investigated is between 0.35 and 2.50  $\mu\text{m}$ ; light source is a 50W quartz-tungsten halogen lamp; FOV is  $\sim 5\text{mm}$ ; incidence and emission angles are  $30^\circ$  and  $0^\circ$ , respectively, which produce an oblong illuminated spot  $\sim 7 \times 10\text{ mm}$  across axes; achieved spectral resolution is 2-7 nm with a 1 nm spectral sampling interval; high SNR was achieved by averaging 1000 individual spectra. In the standard operative mode, the FieldSpec calibration software automatically corrects the individual spectra for the dark current.

In order to test the spectral modeling of saturated and unsaturated bands, we used the following laboratory measurements.

Spectra acquired on two different series of powdered splits obtained via crushing and dry sieving of essentially monomineralic pyroxenes (Fig. 6). Pyroxene chemistry is listed in Table 1. Sample PYX023 is enstatite from Mirabel Springs, Mount St. Helens (Lake County, California); sample PYX032 is hypersthene from Egersund (Norway). The grain size splits for both series are  $<45$ , 45-90, 90-250, 250-500 and 500-1000  $\mu\text{m}$ . The spectra of powdered orthopyroxenes show prominent absorptions near 0.9 and 1.9  $\mu\text{m}$  attributable to  $\text{Fe}^{2+}$  in the M2 sites (Burns, 1993). Slight amounts of impurities mainly due to water and/or alteration products are present and account for the narrow absorptions near 1.4, 1.9 and 2.3 for the enstatite. The effects of progressive band broadening and flattening in the minimum region are evident in spectra (Fig. 6) and diagnostic of band saturation (Craig *et al.* 2007).

**[Figure 6]**

The spectral series of mixtures obtained by adding powdered BaSO<sub>4</sub> to the 45-90 μm splits of the same two pyroxenes, using same incidence and emission angles as above is shown in figure 7. The barium sulfate is very fine (on the order of 10 – 15 μm), and was slightly dried before measurements. Mixtures were prepared with the following relative abundances of barium sulfate and pyroxene: 30%/70%, 50%/50%, 70%/30%, and 90%/10%; endmember spectra were also acquired. Barium sulfate provides a powerful diluent of strong absorbers due to its strong scattering and low absorption, and reduces absorption band depths that may arise due to saturation. No additional bands in the 0.90 μm region are introduced by the barium sulfate. Pyroxene band I shows constant width while its depth is progressively reduced as the amount of BaSO<sub>4</sub> in the mixture increases. Band flattening in the minimum region still occurs, arising from the high Fe<sup>2+</sup> content of the endmember pyroxene (see Tab. 1 and Fig. 7b).

**[Table 1]**

Since the long wavelength wing of pyroxene band II (centered near 2.0 μm) is not always included within the spectral range of the measurements (0.35-2.5 μm), we use only pyroxene band I (centered near 1.0 μm) for testing the EGO fitting technique. In order to better constrain the spectroscopic processes causing the absorption of VNIR light, we fit each spectral band individually, isolated by using a straight line continuum tangent to the local reflectance maxima located on either side of the absorption feature of interest, using: 1) One EGO with all the parameters free to change; 2) A single EGO with one or more parameters locked; and 3) MGM modeling, with one or two modified Gaussians and free parameters.

### 3.2.2 Pure particulate pyroxene

We have modeled the spectra acquired on powdered pyroxenes with a single EGO. First, we fit all the parameters. Then we alternatively set  $t = 10^{-8}$ ; and both  $c_0$  and  $c_1$  as to keep the continuum tangent to the spectrum. Finally, we use alternatively one and two modified Gaussians of Sunshine *et al.* (1990) to model the spectra.

Statistically, the *best fit* results were achieved by modeling with two modified Gaussians and alternatively a single EGO, with all the parameters free to change during the optimization process. It is not surprising because those models compute the largest number of parameters to fit the data. However, few correlations result between parameters, especially in the case of hypersthene and in the case of the EGO models, continua cut the spectra at longer wavelengths. Although not desirable, the occurrence of high correlation coefficients between couples of parameters retrieved by a fitting routine is sometimes difficult to avoid. In particular, the coefficients of continua always correlate. Generally, high correlations among the parameters indicate an overestimation of the model results. One of the best ways to deal with multicollinearity is to reduce the number of variables in the model. Since our main goal is to model the absorption behavior other than the continuum, at first glance we fix the continuum to a straight line tangent to the spectra. If correlations do persist, then we fix some additional parameter (e.g., band center, strength...) to some constant value retrieved by previous iterations. Despite the slightly worse statistics, models with continua forced to be tangent to the spectra display no correlations among the parameters and more consistent continua. Based on the assumptions for residuals behind classical regression analysis, one expects them to be roughly normal and approximately independently distributed with a mean of 0 and some constant variance. Such a distribution of residuals as a function of the

wavelength is diagnostic of a good fit. Departures from the previous assumptions usually mean that the residuals contain structure that is not accounted for in the model. Based on statistics, the *best fit* models involve the superposition of a single EGO onto a tangent continuum (Fig. 8). Since our goal is to model the absorption band and not strictly the continuum, we can accept the use of a fixed tangent continuum for the modeling.

**[Figure 8]**

Modeling the spectra of each powdered split with EGOs and tangent continua returns centers between 939 and 973 nm, and between 918 and 930 nm for hypersthene (PYX032) and enstatite (PYX023), respectively (Tab. 2). This compares to the values of 937 - 970 nm and 917 - 925 nm, based on fitting a polynomial to the center of the continuum removed spectra of all the different powdered splits hypersthene and enstatite, respectively (Cloutis and Gaffey, 1991).

**[Table 2]**

Band depths as a function of grain size increase up to a maximum (corresponding to the 45 – 90  $\mu\text{m}$  and 90 – 250  $\mu\text{m}$  splits, for the hypersthene and enstatite, respectively) and then decrease. Therefore, an optimum particle size does exist where band depth is at a maximum and band saturation is not yet occurring. Crown and Pieters (1987) attributed this phenomenon to the differences between the absorption coefficients in the band center and in the continuum and the competing effects of surface and volume scattering. The fact that the strongest bands are correlated with different grain sizes for the hypersthene versus the enstatite is related to their different compositions, specifically the abundance of  $\text{Fe}^{2+}$ . The more Fe-rich hypersthene is a stronger absorber than the enstatite, thus band saturation in hypersthene would occur at a smaller average grain size. This evidence

supports the hypothesis that saturation occurs in Fe-rich pyroxenes at a smaller grain size than Fe-poor pyroxenes.

Band width increases with increasing grain size in the hypersthene EGO results whereas it is quite constant in enstatite EGO results (between 77 and 112 nm and between 72 and 78 nm, respectively). By modeling the pyroxene band in the 1  $\mu\text{m}$  region with a single modified Gaussian, between 10% and 40% broader absorption band width values are returned (between 83 and 138 nm and between 75 and 110 nm, respectively). This is expected because EGO can rely on the  $t$  parameter to model the shape and constrain the band to be narrower.  $t$  varies from 0.46 up to 2.46 for the hypersthene and from 0.06 to 3.71 for the enstatite (Tab. 2). Now the questions are: what is the meaning of the additional parameter; is it related to any spectral effect or simply a mathematical artifact?

The statistical inadequacy of models with 2 modified Gaussians to interpret the measured spectra of powdered pyroxenes is due to the multicollinearity of parameters. Nevertheless, these models are also affected by an objective inadequacy because one absorption component is centered at wavelengths which cannot be attributed to pyroxene (around 850 nm), based on Cloutis and Gaffey (1991). Therefore, assuming that these absorptions in the 1  $\mu\text{m}$  region are attributable to only one pyroxene and based on the results of simulations, we could infer that saturation effects are being detected.  $t$  values higher than 1.5 could be diagnostic of saturation. Therefore, hypersthene produces a saturated band in all the different average grain sizes except the finest ( $< 45 \mu\text{m}$ ). The absorption band in the spectra of enstatite becomes saturated when the particle size is between 90 and 250  $\mu\text{m}$ . The maximum band depth is achieved in those models where  $t$  becomes diagnostic of saturation. This result is in agreement with the evidence of the

optimal particle size which maximizes band depth before saturation affects spectra and reduces intensity.

### 3.2.3 Bright diluent and pyroxene mixtures

The 45 – 90  $\mu\text{m}$  pyroxene splits were used in these mixtures. This size fraction produces the deepest absorption in hypersthene, but not in enstatite spectra, as discussed above. Since the pyroxene modeling described in the previous section revealed that even the 45 – 90  $\mu\text{m}$  splits displayed  $t$  values larger than 0, we aimed to verify if the occurrence of a bright neutral material intimately mixed with the pyroxenes could alter the band shape and produce band broadening or flattening effects, similarly to saturation.

At first glance, spectra do not show band broadening or flattening as the  $\text{BaSO}_4$  amount increases in the mixtures (Fig. 7). The most noticeable effects are: the increase of total reflectance, the suppression of the charge transfer and OH-related bands, the appearance of an absorption feature near 1.9  $\mu\text{m}$  (likely due to hygroscopic water in  $\text{BaSO}_4$ ), the transition from overall red to blue slope, and the decrease of pyroxene band I depth.

The *best fit* results are achieved by a single EGO with free parameters, EGO with tangent continua, and two-modified Gaussians models. However, based on correlation matrices, we reject all the models but those including a single EGO and tangent continua.

EGO models with tangent continua (Fig. 9) return band centers in the ranges 942-949 nm, and 917 - 922 nm, for hypersthene and enstatite, respectively, as listed in Table 3, consistent with the band centers determined by polynomial fitting of continuum-removed absorption bands (937 - 949 nm and 914 - 921 nm, for hypersthene and enstatite mixed with barium sulfate, respectively). The band width results are between 79 - 81 nm and 70 - 73 nm for hypersthene and enstatite, respectively (Tab. 3). The

constancy of band centers and widths indicates that the addition of a neutral bright material to pyroxene does not produce band broadening or shifting, thus confirming our visual observation.

**[Figure 9]**

**[Table 3]**

In both cases, band depths systematically decrease as the amount of BaSO<sub>4</sub> increases in the mixture (Tab. 3). In the enstatite models,  $t$  varies between 1.52 and 2.33; in the hypersthene models  $t$  is in the range 1.56 – 2.42 (Tab. 3). These results mean that the pyroxene band I in both the mixtures retains the shape as the EGO profile diagnostic of pyroxene (section 3.2.2), with some fluctuation of  $t$  with respect to the model results relative to the 45 – 90  $\mu\text{m}$  pyroxene splits (Table 2). Nevertheless, those slight variations are not significant, based on the magnitude of the corresponding values diagnostic of saturation effects (section 3.1) and on the reverse behavior they assume in each mixture. Therefore, the absorption band I in the pyroxene – BaSO<sub>4</sub> mixtures roughly coincides with the EGO used to model the corresponding pyroxene endmembers. Band intensity varies as a function of the relative proportions of the endmembers in the mixtures, as expected.

This implies that the systematic dilution of pyroxene using a bright neutral material imparts a negligible deformation of the pyroxene absorption band, in terms of flattening, while preserving the parameters  $\mu$  (center) and  $\sigma$  (width) throughout the whole set of models. The  $t$  parameter is not dramatically influenced by the dilution with barium sulfate and the pyroxene band I is still diagnostic of saturation (and in turn grain size), as the pyroxene band of the endmembers used in the mixture. The advantage of EGO modeling versus the Modified Gaussian Model is that the returned set of parameters

strongly suggests the influence of some additional effect other than the electronic transition in  $\text{Fe}^{2+}$ . Nevertheless, we still are uncertain as to whether these deviations were due to a single pyroxene or a mixture with a bright neutral material.

#### 4. Conclusions

In the present investigation, we tested a new decomposition approach which is derived from and alternative to the MGM. It has been named Exponential Gaussian Optimization (EGO) and focuses on the use of a sort of exponential Gaussian (Eq. 1) to fit the electronic absorption bands in pyroxenes. The EGO algorithm has been especially designed to account for those absorption bands which are not Gaussian in shape, due to the occurrence of saturation and asymmetry effects. However, here we present the results achieved from modeling band saturation only. After a number of simulations aimed at reproducing flat bands starting from pure Gaussians, the EGO algorithm has been tested on a number of measurements acquired on powdered pyroxenes having different compositions and average grain size and binary mixtures of pyroxenes with barium sulfate.

Our study has yielded the following results:

The  $t$  parameter of equation 1 allows for the occurrence of saturation effects in spectra to be recognized and accounted for numerically. The threshold of  $\sim 1.5$  established for very Fe-rich pyroxenes appears to be diagnostic of saturation;

Saturation indeed occurs in Fe-rich pyroxenes at a smaller grain size than Fe-poor pyroxenes, as expected;

The systematic dilution of a strong absorber using a bright neutral material is not responsible for band deformation. Therefore, flattening of the band at the bottom

imparted by some saturation effects on the spectrum remains detectable and almost invariant, no matter the amount of bright neutral incorporated in the mixture;

The returned set of EGO parameters can suggest the influence of some additional effect other than a single electronic transition responsible for the absorption feature here investigated. In particular, we show here that saturation effects can be found out and parameterized.

Further work is still required in order to analyze the behavior of the EGO algorithm using more complex band shapes (e.g., the complex bands of olivines and glasses and two pyroxene mixtures; vibrational absorptions...) and with respect to the saturation phenomena, which likely affect absorption bands in spectra acquired on rock slabs.

## Appendix A – Statistical analysis

The following is a brief description of the statistics implemented in the R-code for spectral modeling evaluation.

To compare regression models, we calculate the standard error of the estimate (*SEE*) using the following equation:

$$SEE = \sqrt{\frac{SS_{res}}{df}} = \sqrt{\frac{\sum_i (mod - data)^2}{n-p}} \quad (\text{Eq. A1})$$

where  $SS_{res}$  is the residual sum of squares and  $df$  the degrees of freedom of the nonlinear regression,  $mod$  and  $data$  indicate each single value of the model and the spectral measurement, respectively,  $n$  is the total number of data points, and  $p$  is the number of the regression parameters (Glantz and Slinker, 1990).

The coefficient of determination ( $R^2$ ) for multiple regressions is the fraction of the total variance of data that is explained by the model (Glantz and Slinker, 1990).

$$R^2 = 1 - \frac{SS_{reg}}{SS_{tot}} = \frac{SS_{reg}}{SS_{tot}} \quad (\text{Eq. A2})$$

where  $SS_{reg}$  and  $SS_{tot}$  are the sum of squared distances of data points from the regression curve and a horizontal line through the mean of all the data values, respectively.

The adjusted coefficient of determination ( $R^2_{adj}$ ) derives its definition directly from  $R^2$ . Nevertheless, since the  $R^2$  always increases as more variables are added to the model, even if these new variables add little new independent information, the  $R^2_{adj}$  compromises between adding additional variables and losing degrees of freedom. It is defined as follows:

$$R^2_{adj} = 1 - \frac{SS_{reg}/(n-p)}{SS_{tot}/(n-1)} = 1 - \frac{(1-R^2)(n-1)}{(n-p)} \quad (\text{Eq. A3})$$

The model with the highest  $R^2_{\text{adj}}$  will have the smallest prediction errors.

Although being valid, the  $R^2_{\text{adj}}$  values close to 1.0 do not guarantee that the model will not be overspecified and contains serious correlations among parameters.

The computation of the Hessian matrix represents a key point for evaluating the goodness of fit. It is used to improve the approximation of the parameters during the iterative optimization process, as much as to estimate the standard errors of the regression coefficients ( $s_p$ ) by taking the square roots of the diagonal elements of the inverse of the Hessian matrix (Glantz and Slinker, 1990)

$$s_p = \sqrt{SEE^2(\mathbf{D}^T \mathbf{D})^{-1}} \quad (\text{Eq. A4})$$

where  $\mathbf{D}$  is the matrix of partial derivatives of the regression function,  $(\mathbf{D}^T \mathbf{D})$  is the Hessian matrix, and  $SEE$  is the standard error of the estimate (Eq. A1). These standard errors have been used to compute the 95% confidence intervals for the regression coefficients  $p$ , as follows:

$$p_t - t_\alpha s_{p_t} < p_t < p_t + t_\alpha s_{p_t} \quad (\text{Eq. A5})$$

where  $t_\alpha$  is the two-tail critical value for the t-distribution with  $(n-p)$  degrees of freedom.

The Hessian matrix also allows the correlation matrix of the regression parameter estimates to be calculated, which in turn allows the multicollinearity among regression parameters to be diagnosed (Motulsky and Christopoulos, 2003).

The Akaike's Information Criterion (AIC) has been also implemented within the code for comparing nested models. Under the assumption that deviations of points from the regression curve are randomly distributed, the AIC is computed as follows:

$$AIC = n \ln\left(\frac{SS_{\text{res}}}{n}\right) + 2K \quad (\text{Eq. A6})$$

where  $n$  is the number of data points;  $K$  is the number of parameters plus 1. The usefulness of the AIC is in comparing models, so it is the difference between AIC values

to be accounted for. The AIC makes a balance between the variations of goodness of fit as assessed by the residual sum of squares and the increased complexity of the model.

The lower the AIC, the more likely is the model.

### Appendix B – Demonstration of coincidence between EGO and G under certain conditions

Let's define a Gaussian profile as

$$G = s e^{-\frac{t}{2} \left( \frac{x-s}{\sigma} \right)^2} \quad (\text{Eq. B1})$$

and an EGO profile as

$$EGO = -\frac{s}{1 - e^{-\frac{t}{2}}} \left[ 1 - e^{-\frac{t}{2} \left( \frac{x-s}{\sigma} \right)^2} \right] \quad (\text{Eq. B2})$$

Recall that

$$\text{for } x \ll 1 \quad e^{-x} \cong \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = 1 - x + \dots \quad (\text{Eq. B3})$$

if we take a Taylor series around  $t = 0$  and assume  $k = 0$ , then we get:

$$EGO = \frac{s}{1 - \left(1 - \frac{t}{2}\right)} \left[ 1 - \left(1 - \frac{t}{2} e^{-\frac{t}{2} \left( \frac{x-s}{\sigma} \right)^2} \right) \right] = \frac{s}{\frac{t}{2}} \left[ \frac{t}{2} e^{-\frac{t}{2} \left( \frac{x-s}{\sigma} \right)^2} \right] = s e^{-\frac{t}{2} \left( \frac{x-s}{\sigma} \right)^2} = G \quad (\text{Eq. B4})$$

So we demonstrate that for  $t \cong 0$  and  $k = 0$ , the EGO and G become coincident.

### Acknowledgements

We recognize financial support of SIMBIOSYS Experiment from ASI's spatial Program. We wish to thank Jessica Sunshine and an anonymous reviewer for the helpful comments that allowed significant improvements to the manuscript.

**References**

- Baratoux, D., Pinet, P., Gendrin, A., Kanner, L., Mustard, J., Daydou, Y., and Bibring, J. P. 2007. Impact craters mineralogy from OMEGA data: implications on alteration history, ejecta emplacement, and subsurface composition. 7<sup>th</sup> international Conf. on Mars, abstract # 3183.
- Binzel, R. P., Rivkin, A. S., Bus, S. J., Sunshine, J. M., Burbine, T. H. 2001. MUSES-C target asteroid (25143) 1998 SF36: a reddened ordinary chondrite. *Meteorit. Planet. Sci.* 36, 1167-1172.
- Burns, R. G. 1970. Crystal field spectra and evidence of cation ordering in olivine minerals. *Am. Mineral.* 55, 1608-1632.
- Burns, R. G. 1993. *Mineralogical Applications of Crystal Field Theory*, 2nd ed. Cambridge Univ. Press, Cambridge, UK.
- Byrd, R. H., Nocedal, J., Schnabel, R. B. 1994. Representation of quasi-Newton matrices and their use in limited memory methods. *Mathematical programming* 63, 4, 129-156.
- Clark, R. N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: Rencz, A. N. (Ed.), *Remote Sensing for the Earth Sciences, Manual of Remote sensing* 3<sup>rd</sup> ed., John Wiley and Sons, Inc., New York, pp. 3 – 58.

- Clark, R. N. 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.* 108, E12, 5131.
- Clark, R. N., Roush, T. L. 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* 89, 6329-6340.
- Cloutis, E. A., Gaffey, M. J. 1991. Pyroxene spectroscopy revisited: spectral-compositional correlations and relationships to geothermometry. *J. Geophys. Res.* 96, 22809-22826.
- Craig, M. A., Cloutis, E. A., Bailey, D. T. 2007. The effects of grain size, < 45 - 1000  $\mu\text{m}$ , on the reflectance spectrum of planetary analogs from 0.35 - 2.5  $\mu\text{m}$ . *Lun. Planet. Sci.* XXXVIII, abstract #1356.
- Crown, D. A., Pieters, C. M. 1987. Spectral properties of plagioclase and pyroxene mixtures and the interpretation of Lunar soil spectra. *Icarus* 72, 492-506.
- Garbow, B. S., Hillstrom, K. E., Moré, J. J. 1980. Part of the Argonne National Laboratory MINPACK Project, publicly available via the netlib server.
- Glantz, S. A., Slinker, B. K. 1990. *Primer of Applied Regression and Analysis of Variance*. McGraw-Hill Inc., New York.

Hiroi, T., Pieters, C. M. 1998. Modified Gaussian Deconvolution of Reflectance Spectra of Lunar Soils, *Lun. Planet. Sci.* XXIX, abstract #1253.

Hiroi, T., Binzel, R.P., Sunshine, J.M., Pieters, C.M., Takeda, H. 1995. Grain sizes and mineral compositions of surface regoliths of Vesta-like asteroids. *Icarus* 115, 374-386.

Hiroi, T., Pieters, C. M., Noble, S.K. 2000. Improved scheme of Modified Gaussian deconvolution for reflectance spectra of Lunar soils. *Lun. Planet. Sci.* XXXI, abstract #1548.

Hiroi, T., Vilas, F., Sunshine, J. M. 1996. Discovery and analysis of minor absorption bands in S-asteroid visible reflectance spectra. *Icarus* 119, 202-208.

Klima, R. L., Pieters, C. M., Dyar, M. D. 2007. Spectroscopy of synthetic Mg-Fe pyroxenes I: Spin-allowed and spin-forbidden crystal field bands in the visible and near-infrared. *Meteorit. Planet. Sci.* 42, 2, 235-253.

Mayne, R. G., Sunshine, J. M., McCoy, T. J., McSween Jr., H.Y. 2006. Substantial lithologic diversity on 4 Vesta: evidence from the petrology and spectra of Antarctic eucrites. *Lun. Planet. Sci.* XXXVII, abstract #1796.

- McFadden, L. A., Cline, T. P. 2005. Spectral Reflectance of Martian Meteorites: spectral signatures as a template for locating source region on Mars. *Meteorit. Planet. Sci.* 40, 2, 151-172.
- Moré, J. J. 1978. The Levenberg-Marquardt algorithm: implementation and theory. In: Watson, G.A. (Ed.), *Numerical Analysis. Lecture Notes in Mathematics*, Springer-Verlag, Berlin, pp. 105-116.
- Motulsky, H. J. Christopoulos, A. 2003. Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting. GraphPad Software Inc., San Diego CA, [www.graphpad.com](http://www.graphpad.com).
- Murchie, S., et al. 2007. Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on Mars Reconnaissance Orbiter (MRO). *J. Geophys. Res.* 112, E05S03, doi:10.1029/2006JE002682.
- Mustard, J. F. 1992. Chemical analysis of actinolite from reflectance spectra. *Am. Mineral.* 77, 345-358.
- Nicholis, M., Milliken, R. E., Mustard, J. F., Rutherford, M. 2006. VIS-NIR spectral properties of olivine in a basaltic glass: implications for olivine-rich terrains on Mars. *Lun. Planet. Sci.* XXXVII, abstract #2378.

- Nolet, D. A., Burns, R. G., Flamm, S. L., Besancon, J. R. 1979. Spectra of Fe-Ti silicate glasses: implications to remote-sensing of planetary surfaces. *Pro. Lunar Planet. Sci. Conf. 10<sup>th</sup>*, 1775-1786.
- Parente, M., Bishop, J. L. 2006. Deconvolution of reflectance spectra using nonlinear least squares curve fitting: application to Martian meteorites. *Lun. Planet. Sci. XXXVII*, abstract #1535.
- Pieters, C. M. 1996. Plagioclase and maskelynite diagnostic features. *Lun. Planet. Sci. XXVII*, abstract #1031.
- Pieters, C. M., Tompkins, S. 2005. Remote sensing of Lunar mineralogy: the glass Conundrum, *Lun. Planet. Sci. XXXVI*, abstract #1346.
- Pinet, P. C., Heuripeau, F., Clenet, H., Chevrel, S., Daydou, Y., Baratoux, D., Rosemberg, C., Bibring, J. P., Poulet, F., Gondet, B., Mustard, J., leMouélic, S., Bellucci, G., and the OMEGA team. 2007. Mafic mineralogy variations across Syrtis Major shield and surroundings as inferred from visible-near-infrared spectroscopy by OMEGA/Mars Express. 7<sup>th</sup> international Conf. on Mars, abstract #3146.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P. 1992. *Numerical Recipes in FORTRAN. The Art of Scientific Computing 2nd ed.* Cambridge Univ. Press, London (UK).

- R Development Core Team. 2005. R: A language and environment for statistical computing, ISBN 3-900051-07-0, URL: <http://www.R-project.org>, R Foundation for Statistical Computing, Vienna.
- Roush, T. L., Singer, R. B. 1986. Gaussian Analysis of Temperature effects on the Reflectance Spectra of mafic minerals in the 1- $\mu\text{m}$  region. *J. Geophys. Res.* 91, 10301-10308.
- Runciman, W. A., Sengupta, D., Gourley, J. T. 1974. The polarized spectra of iron silicates: II. Olivine: A reply. *Am. Mineral.* 59, 630-631.
- Singer, R. B. 1981. Near-Infrared spectral reflectance of mineral mixtures: systematic combinations of pyroxene, olivine and iron oxides. *J. Geophys. Res.* 86, 7967-7982.
- Schade, U., Wäsch, R. 1999. Near-infrared reflectance spectra from bulk samples of the two Martian meteorites Zagami and Nakhla. *Meteorit. Planet. Sci.* 34, 417-424.
- Sugihara, T., Owada, A., Ohtake, M., Takeda, H. 2003. MGM deconvolution of reflectance spectrum of the Y981031 Lunar Meteorite. *Lun. Planet. Sci.* XXXIV, abstract #1568.
- Sunshine, J. M., Pieters, C. M. 1990. Extraction of compositional information from olivine reflectance spectra: a New Capability for Lunar Exploration. *Lun. Planet. Sci.* XXI, abstract #1223.

- Sunshine, J. M., Pieters, C. M. 1991. Identification of modal abundances in the spectra of natural and laboratory pyroxene mixtures: a key component for remote analysis of Lunar Basalts. *Lun. Planet. Sci.* XXII, abstract #1361.
- Sunshine, J. M., Pieters, C. M. 1993. Estimating modal abundances from the spectra of natural and laboratory pyroxene mixtures using the Modified Gaussian Model. *J. Geophys. Res.* 98, 9075-9087.
- Sunshine, J. M., Pieters, C. M. 1998. Determining the composition of olivine from reflectance spectroscopy. *J. Geophys. Res.* 103, 13675-13688.
- Sunshine, J. M., McFadden, L. A., Pieters, C. M. 1993. Reflectance spectra of the Elephant Moraine A79001 Meteorite: implications for remote sensing of planetary bodies. *Icarus* 105, 79-91.
- Sunshine, J. M., Pieters, C. M., Pratt, S. F. 1990. Deconvolution of mineral absorption bands: an improved approach, *J. Geophys. Res.* 95, 6955-6966.
- Tarantola, A., Valette, B. 1982. Generalized non-linear inverse problems solved using the least-squares criterion, *Rev. Geophys.* 20, 219-232.
- Ueda, Y., Hiroi, T., Pieters, C. M., Miyamoto, M. 2002. Expanding the Modified Gaussian Model to include space weathering effects: estimation of the weathering

degrees of pulse-laser treated olivine samples. Lun. Planet. Sci. XXXIII, abstract  
#1950.

ACCEPTED MANUSCRIPT

**Table 1****Pyroxene chemistry**

<b>Wt. % oxides</b>	<b>Samples</b>	
	PYX023	PYX032
SiO <sub>2</sub>	56.86	50.21
Al <sub>2</sub> O <sub>3</sub>	0.76	1.24
FeO	6.36	23.65
Fe <sub>2</sub> O <sub>3</sub>	0.83	5.11
MgO	34.04	17.57
CaO	0.65	1.59
Na <sub>2</sub> O	0.00	0.00
TiO <sub>2</sub>	0.01	0.19
Cr <sub>2</sub> O <sub>3</sub>	0.45	0.04
V <sub>2</sub> O <sub>5</sub>	tr.	tr.
CoO	0.03	0.06
NiO	0.08	0.01
MnO	0.17	0.53
ZrO <sub>2</sub> *	0.00	0.00
ZnO*	n.d.	n.d.
TOTAL	100.24	100.20
<b>ATOMIC RATIOS</b>		
Mg	89.3	54.9
Fe <sup>2+</sup>	9.4	41.5
Ca	1.3	3.6
<b>NUMBER OF IONS ON THE BASIS OF 6 OXYGENS</b>		
Si	1.965	1.920
Al	0.031	0.056
Al	--	--
V	tr.	tr.
Ti	tr.	0.005
Zr	--	--
Cr	0.012	0.001
Fe <sup>3+</sup>	0.022	0.147
Fe <sup>2+</sup>	0.184	0.756
Mg	1.753	1.002
Ca	0.025	0.065
Co	0.001	0.002
Ni	0.002	tr.
Mn	0.004	0.017
Na	--	--
Zn	--	--
TOTAL	4.000	3.972
Wo	1.21	3.28
En	88.21	50.40
Fs	10.58	46.32

Table 2

EGO fitting results for spectra of the two orthopyroxenes shown in Fig. 8 as a function of grain size.

PYX032 - EGO model - Continuum fixed												
Grain size	c0	uncert.	c1	uncert.	s	uncert.	$\mu$	uncert.	$\sigma$	uncert.	t	uncert.
<b>&lt;45 <math>\mu\text{m}</math></b>	-0.790	--	-0.270	--	0.547	1.68E-03	0.939	2.33E-04	0.077	7.50E-04	0.462	1.41E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.544	0.550	0.938	0.939	0.076	0.079	0.185	0.739
<b>45-90 <math>\mu\text{m}</math></b>	-0.734	--	-0.640	--	0.773	2.51E-03	0.944	2.75E-04	0.080	6.32E-04	2.310	1.39E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.768	0.778	0.944	0.945	0.079	0.081	2.037	2.583
<b>90-250 <math>\mu\text{m}</math></b>	-1.180	--	-0.880	--	0.725	2.36E-03	0.959	3.27E-04	0.094	7.51E-04	2.464	1.42E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.721	0.730	0.958	0.959	0.093	0.096	2.186	2.743
<b>250-500 <math>\mu\text{m}</math></b>	-1.765	--	-0.545	--	0.480	1.64E-03	0.968	3.81E-04	0.108	1.06E-03	1.420	1.54E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.477	0.483	0.968	0.969	0.106	0.111	1.118	1.721
<b>500-1000 <math>\mu\text{m}</math></b>	-2.380	--	-0.340	--	0.401	1.59E-03	0.973	4.55E-04	0.112	1.29E-03	1.302	1.78E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.398	0.404	0.972	0.974	0.110	0.115	0.953	1.652

PYX023 - EGO model - Continuum fixed												
Grain size	c0	uncert.	c1	uncert.	s	uncert.	$\mu$	uncert.	$\sigma$	uncert.	t	uncert.
<b>&lt;45 <math>\mu\text{m}</math></b>	-0.298	--	-0.140	--	0.519	1.63E-03	0.918	2.19E-04	0.072	7.61E-04	0.062	1.48E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.516	0.522	0.918	0.918	0.071	0.074	-0.230	0.353
<b>45-90 <math>\mu\text{m}</math></b>	-0.440	--	-0.200	--	0.847	2.42E-03	0.918	2.09E-04	0.073	5.83E-04	1.072	1.24E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.842	0.851	0.918	0.919	0.071	0.074	0.828	1.316
<b>90-250 <math>\mu\text{m}</math></b>	-0.720	--	-0.300	--	1.061	2.98E-03	0.920	2.18E-04	0.073	4.89E-04	2.365	1.19E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	1.055	1.067	0.920	0.921	0.072	0.074	2.131	2.600
<b>250-500 <math>\mu\text{m}</math></b>	-1.020	--	-0.380	--	0.932	2.74E-03	0.925	2.44E-04	0.075	4.74E-04	3.581	1.30E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.927	0.937	0.924	0.925	0.074	0.076	3.325	3.837
<b>500-1000 <math>\mu\text{m}</math></b>	-1.530	--	-0.350	--	0.752	2.24E-03	0.930	2.58E-04	0.078	4.94E-04	3.706	1.33E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.748	0.757	0.930	0.931	0.077	0.079	3.446	3.967

See Eq. 1 for definition of symbols.  $k = 0$  in all the trials.  
 "Uncert." means uncertainty on the parameter.

Table 3

EGO fitting results for the two BaSO<sub>4</sub> + orthopyroxene mixtures shown in Fig. 9.

PYX032 - BaSO <sub>4</sub> mixture - EGO model - Continuum fixed												
PYX-BaSO <sub>4</sub> ratio	c0	uncert.	c1	uncert.	s	uncert.	μ	uncert.	σ	uncert.	t	uncert.
<b>0% - 100%</b>	-0.713	--	-0.721	--	0.921	2.60E-03	0.942	2.36E-04	0.079	5.32E-04	2.422	1.21E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.916	0.926	0.942	0.943	0.078	0.080	2.185	2.659
<b>30 - 70%</b>	-0.567	--	-0.295	--	0.428	1.21E-03	0.946	2.46E-04	0.082	5.63E-04	2.289	1.21E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.426	0.431	0.946	0.947	0.081	0.083	2.052	2.526
<b>50 - 50%</b>	-0.423	--	-0.080	--	0.227	5.95E-04	0.945	2.23E-04	0.081	5.18E-04	2.155	1.11E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.226	0.229	0.944	0.945	0.080	0.082	1.937	2.373
<b>70 - 30%</b>	-0.257	--	-0.007	--	0.122	3.06E-04	0.945	2.06E-04	0.077	4.56E-04	2.393	1.06E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.122	0.123	0.945	0.946	0.076	0.078	2.186	2.601
<b>90 - 10%</b>	-0.146	--	0.027	--	0.040	1.13E-04	0.949	2.32E-04	0.080	5.87E-04	1.562	1.20E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.040	0.040	0.949	0.950	0.079	0.081	1.327	1.797

PYX023 - BaSO <sub>4</sub> mixture - EGO model - Continuum fixed												
PYX-BaSO <sub>4</sub> ratio	c0	uncert.	c1	uncert.	s	uncert.	μ	uncert.	σ	uncert.	t	uncert.
<b>0% - 100%</b>	-0.495	--	-0.230	--	1.074	2.78E-03	0.917	1.90E-04	0.071	4.87E-04	1.523	1.10E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	1.069	1.080	0.917	0.917	0.071	0.072	1.306	1.739
<b>30 - 70%</b>	-0.485	--	-0.019	--	0.433	1.33E-03	0.917	2.28E-04	0.072	5.74E-04	1.568	1.31E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.431	0.436	0.917	0.918	0.070	0.073	1.311	1.824
<b>50 - 50%</b>	-0.391	--	0.000	--	0.334	1.11E-03	0.918	2.52E-04	0.073	6.15E-04	1.829	1.41E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.332	0.336	0.917	0.918	0.072	0.074	1.553	2.106
<b>70 - 30%</b>	-0.217	--	0.005	--	0.221	6.13E-04	0.918	2.04E-04	0.070	4.90E-04	1.862	1.17E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.220	0.222	0.917	0.918	0.069	0.071	1.632	2.092
<b>90 - 10%</b>	-0.114	--	0.023	--	0.055	1.36E-04	0.922	1.82E-04	0.070	4.06E-04	2.331	1.03E-01
95% confidence intervals	low	high	low	high	low	high	low	high	low	high	low	high
	--	--	--	--	0.055	0.056	0.922	0.923	0.069	0.071	2.128	2.534

See Eq. 1 for definition of symbols.  $k = 0$  in all the trials.  
 "Uncert." means uncertainty on the parameter.

### Figure captions

Figure 1. Mathematical behavior of the EGO profile (solid lines) starting from a pure Gaussian (gray dots) and varying the parameter  $t$  (saturation effect) Intensity, center and width ( $s$ ,  $\mu$ , and  $\sigma$ ) of the initial Gaussian and EGOs are -1, 0, 0.7, respectively. From the inner to the outer curve,  $t$  varies as: 1, 5, 10, 30, 100. As  $t$  increases, the Gaussian becomes flat at the bottom, thus simulating the effect of saturation on absorption bands.

Figure 2. A spectral subset of the simulation dataset. The Gaussian components have the following constant parameters for each simulation:  $s_1 = s_2 = 0.1$ ;  $\mu_1 = 0.95 \mu\text{m}$ . (a)  $\sigma_1 = \sigma_2 = 0.05 \mu\text{m}$ ; (b)  $\sigma_1 = \sigma_2 = 0.07 \mu\text{m}$ ; (c)  $\sigma_1 = \sigma_2 = 0.08 \mu\text{m}$ ). The Gaussian noise has been removed for clarity. The longer wavelength Gaussian retains the same depths and widths as the shorter wavelength Gaussian, but is centered 7.5 nm away from the other and progressively moves up to a maximum separation of 150 nm, by steps of 7.5 nm. According to the width of the component Gaussians, in turn depending on the nature of the electronic transition absorption, as the maximum band separation approaches 60 nm, the resulting composite band shows the occurrence of separate bands (a), evidence of saturation (b), broadening of the absorption (c). The EGO profile is able to model all these occurrences.

Figure 3. The plots show the results of modeling simulated absorption bands of figure 2 using one single EGO superimposed onto a horizontal continuum. The symbols in the legend indicate the simulation runs and make use of the same letters as in figure 2. See the text for discussion.

Figure 4. Residual distributions as a function of the wavelength resulting from modeling a subset of the simulated spectra (i.e., the spectra shown in Fig. 2b), starting from 7.5 up to 150 nm band separation, by steps of 7.5 nm from Model 1 to Model 20. The residual distributions show the capability of the EGO model to accurately fit the simulated bands up to Model 16. Here the residuals are normally distributed around 0 and equal to the random noise added to the simulations (1% of band intensity). The residual distributions resulting from Models 17 to 20 show fluctuations larger than the random noise. These results indicate that the EGO model is no longer able to fit the data when band separation is larger than 120 nm and a further Gaussian is required. Therefore the  $t$  value retrieved by Model 16 is a limit for saturated bands having the set of parameters as the simulation shown in figure 2b. For band separations larger than 150 nm, the resulting composite bands show clear evidence of separate components.

Figure 5. Plots of EGO band width and depth as a function of  $t$  (a and b, respectively), as results from modeling the simulations shown in figures 2a, b and c. The symbols are depicted with the same letters as in figure 2. Band width increases non-linearly with  $t$  up to a plateau level. Threshold  $t$  value is  $\sim 1.5$ . Band depth decays progressively as  $t$  increases. Therefore,  $t$  values higher than 1.5 are diagnostic of saturation, because band width remains constant while depth decreases progressively.

Figure 6. Spectral series of powdered pyroxene splits. a) Enstatite (PYX023); b) Hypersthene (PYX032). From top to bottom in each plot, measurements relative to the following grain size ranges are shown:  $<45 \mu\text{m}$  (solid lines),  $45\text{-}90 \mu\text{m}$  (dashed lines),

90-250  $\mu\text{m}$  (dotted lines), 250-500  $\mu\text{m}$  (short dashed lines), 0.5-1.0 mm (dash-dotted lines).

Figure 7. Spectral series relative to the “dilution experiment”.  $\text{BaSO}_4$  has been added with definite proportions to (a) Enstatite (PYX023) and (b) Hypersthene (PYX032). From top to bottom in each plot, spectra of the  $\text{BaSO}_4$ /pyroxene relative proportions are shown: 90/10 (solid lines); 70/30 (dashed lines); 50/50 (dotted lines); 30/70 (short dashed lines); 0/100 (dash-dotted lines).

Figure 8. Fit results of pyroxenes modeling. Each plot shows the *best fit* results including the residual distribution (top curve in each plot), EGO curve used to model pyroxene band I (black solid lines) superimposed onto the continuum and thus coincident with the model fit, continuum tangent to the absorption to model (dashed lines), the data (gray lines), for the hypersthene PYX032 (a) and the enstatite PYX023 (b). From top to bottom the modeled spectra are displayed as follows:  $<45 \mu\text{m}$ , 45-90  $\mu\text{m}$ , 90-250  $\mu\text{m}$ , 250-500  $\mu\text{m}$ , 0.5-1.0 mm (see Fig. 6 for the reference spectra).

Figure 9. Fit results of  $\text{BaSO}_4$ /pyroxene mixtures modeling. Each plot shows the *best fit* results including the residual distribution (top curve in each plot), EGO curve used to model pyroxene band I (black solid lines) superimposed onto the continuum and thus coincident with the model fit, continuum tangent to the absorption to model (dashed lines), the data (gray lines), for the hypersthene PYX032 (a) and the enstatite PYX023 (b). From top to bottom the modeled spectra are obtained by measuring the mixtures with

the following relative abundances: 0%/100%, 30%/70%, 50%/50%, 70%/30%, and 90%/10% (see Fig. 7 for the reference spectra).

ACCEPTED MANUSCRIPT

