Synthesis inverse mapping: A robust method applicable to atmospheric remote sounding

Didier Fussen, Filip Vanhellemont, and Christine Bingen
Institut d'Aéronomie Spatiale de Belgique, Brussels, Belgium

Received 15 February 2002; revised 15 May 2002; accepted 22 May 2002; published 19 October 2002.

[1] We develop a new inversion method that may be advantageously applied to remote sounding in atmospheric physics. The technique is based on the synthesis of the inversion operator which can be used in noisy measurements where the success of classical regularization and optimal estimation methods strongly depends on the accurate knowledge of the a priori atmospheric state. We outline the benefits of the method: robustness, full control of validity range as well as easy assessment of the smoothing error. An example of spectral inversion is extensively analyzed in the context of aerosol retrievals from limb occultation experiments. *INDEX TERMS:* 0305 Atmospheric Composition and Structure: Aerosols and particles (0345, 4801); 0360 Atmospheric Composition and Structure: Transmission and scattering of radiation; 3260 Mathematical Geophysics: Inverse theory; *KEYWORDS:* remote sensing, inversion method, spectral inversion

Citation: Fussen, D., F. Vanhellemont, and C. Bingen, Synthesis inverse mapping: A robust method applicable to atmospheric remote sounding, *J. Geophys. Res.*, *107*(D20), 4421, doi:10.1029/2002JD002206, 2002.

1. Introduction

[2] Remote sensing of atmospheric gases from spaceborne experiments play an essential role in the monitoring of the Earth's atmospheric evolution. In particular, it allows to determine the vertical distribution of absorbing constituents on a global scale and possibly for long time series. A common characteristic of these optical measurements is the indirect nature of the experiment where the information related to the target physical quantity is conveyed into the measured signal after physical or mathematical transformation. Mostly, the information is integrated along the line-of-sight, loses local details, and it is clear that these can only be discovered by means of a series of measurements focused on the same spatial region.

[3] However, such measurements lead unavoidably to the resolution of an inverse problem: which atmospheric state gave rise to the series of observations? Improvements of efficient inversion algorithms have a direct impact on the quality of the retrieval and the assessment of the error budget. Although inverse problems are universally met in all scientific fields, atmospheric sciences have seen two different schools of thought emerging from various inversion methods that have been published so far.

[4] The first school, the ''Bayesian'', probably played a dominant role in the last thirty years [*Rodgers*, 2000] and is based on the description of experimental errors and measurement in term of probability density functions. These are used to relate prior and posterior information by using Bayes' theorem and the solution is obtained by optimising some property of the a posteriori ensemble of states. Therefore, they are often referred to as ''optimal estimation methods'' (OEM).

[5] The second school, the ''regularization'', has also a long history [*Twomey*, 1985] and is probably more related to the mathematical frame of Fredholm integral equations [*Hansen*, 1992]. Very briefly, it consists of adjudicating a compromise between the agreement with the observations and the smoothness of the solution by selecting an appropriate Lagrange multiplier in the ensemble of possible solutions.

[6] The merits and drawbacks of both approaches can be endlessly discussed and, mostly, historical or pragmatic reasons have conditioned the choice of the selected method. Quite recently, optimal estimation and regularization methods have been cross-compared in the retrieval of ozone profiles from ground-based measurements [*Eriksson*, 2000]. However, the fundamental differences between both techniques are more a matter of philosophical conceptions than actual mathematical divergences as ironically quoted by *Press et al.* [1992].

[7] On the other hand, in our opinion, an interesting point of view in the development of any inversion algorithm is the synthesis approach [*Twomey*, 1985] where the objective is to build an inversion operator that is constrained to possess interesting properties. The most famous technique that makes use of this concept is the Backus-Gilbert method where the smoothing function (the function that describes the successive actions of the forward and inverse models) is requested to be as close as possible to a delta function whereas the magnitude of the operator elements is controlled by means of a ''trade-off'' parameter. However, as expected, error magnification and width of the smoothing function are inversely proportional and the choice of an optimal parameter is somewhat arbitrary.

0148-0227/02/2002JD002206

[8] In this work, we want to exploit the synthesis approach perspective from a more pragmatic point of view: to construct a robust inversion operator that will always work in a given domain of atmospheric states and corresponding measurements, in the manner of a table look-up scheme. Many kinds of inverse problems are encountered in the field of atmospheric remote sounding. The reader can refer to an extensive presentation by *Rodgers* [2000]. The degree of complexity of these problems may vary according to the nonlinearity, the number of measurements during one acquisition, the typical signal to noise ratio and the possible spectral interrelationship between the measured species. Hereafter we will not make any particular assumption about the inverse model complexity. Therefore, we consider our method as quite general and it could be applied for all types of inversion, the only limitation being of course the computational cost of the forward model. The described method is general, for all types of inversion. After deriving the proposed algorithm, we will apply it to a simple although nonlinear atmospheric inverse problem: the spectral inversion of absorbing constituents in an occultation experiment.

2. Synthesis Inverse Mapping (SIM) Method

[9] Occultation experiments can generate different kinds of ill-posed problems. For instance, it is important to realise that successive measurements (during the orbital satellite movement) are very redundant, they carry only a small amount of new information between two signal snapshots. Consequently, any matrix representing the exact forward model or its linear approximation will contain rows that are almost identical or linear combinations. This generates very small eigenvalues which in turn induce large elements in the inverse or pseudo-inverse of the forward model [*Hansen*, 1992]. Fundamental difficulties of inverse problems lie in the unavoidable noise amplification due to these large matrix elements. However, if we know a priori that the signal contains added noise, it is natural to try to remove it or to control it before the inversion itself takes place instead of designing an inversion scheme capable of controlling noise amplification.

[10] Also, many nonlinear inverse problems can only be solved by using iterative algorithms starting from a realistic first guess. Furthermore, they may converge to a local but inadequate solution or to an unstable solution due to the flatness of the merit function. Actually, convergence suggests that there should be only one good solution and that the algorithm could reach it in one step even if there doesn't exist an algebraic way to express it.

[11] This is why the idea of a unique mapping between the measurement space and the solution space is very intuitive and table look-up, its most simple application, is still frequently used in problems where the computing time is prohibitive (e.g. full radiative transfer computations). The unique mapping should always be possible if adjacent domains in the solution space map onto adjacent domains in the measurement space through the action of the forward model and if the solution is not a multivalued function of the measurement. Measurement and solution are supposed to be continuous functions and they could be adequately discretized prior to the mapping. On the other hand, an economical way to represent both functions is to expand them on a suitable basis of orthogonal polynomials and to look for a mapping procedure between the coefficients of the truncated expansions.

[12] Let us consider $f(x)$, the unknown function that we want to retrieve and $g(y)$ the continuous series of measurements identified by the parameter y. For instance, in an occultation experiment, $f(x)$ is the unknown atmospheric absorbing constituent density profile depending on the local altitude x and $g(y)$ is the measured transmittance through the atmosphere when the tangent altitude y of the observed optical path is varying. The forward model, represented by the operator $\mathbf{A}$, expresses the relation between the physical space and the measurement as:

$$\mathbf{A}(f) = g \tag{1}$$

which reduces, in a linear case and after discretization of f and g, to a familiar matrix equation:

$$Af = g \tag{2}$$

[13] Let us call $f_0(x)$ a reasonable reference profile, i.e. an arbitrary profile coming from climatological data, a priori information or scaling requirements, close to $f(x)$ ($f_0(x)$ is not constrained to be the most probable state of the atmosphere but it should lie in the vicinity of the true unknown profile) and $g_0(y)$ is the corresponding generic function for the standard measurement. We can develop the respective deviations with respect to both generic profiles on a basis of orthogonal polynomials $U_i(x)$,$V_j(y)$ up to order n_x,n_y as:

$$F(x) = f(x) - f_0(x) = \sum_{i=0}^{n_x} a_i U_i(x) \tag{3}$$

$$G(y) = g(y) - g_0(y) = \sum_{j=0}^{n_y} b_j V_j(y) \tag{4}$$

[14] The choice of the orthogonal polynomial families depends on the physical range spanned by x and y. However, for a given range, one can select a particular family of polynomials according to a specific weighting function. For example, over a finite interval, Chebyshev are known to approximate the minimax polynomial [*Press et al.*, 1992]. The purpose of the inversion algorithm is to retrieve a good approximation of f from measured g, i.e. to discover an operator B such that:

$$F(x) \simeq \mathbf{B}(G(y)) \tag{5}$$

[15] It is convenient but not mandatory to use a linear transformation, represented by a matrix B, that relates $\vec{a} = (a_0, a_1, \cdots a_{n_x})^T$ and $\vec{b} = (b_0, b_1, \cdots b_{n_x})^T$ by

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n_x} \end{pmatrix} = \begin{pmatrix} B_{00} & B_{01} & \cdots & B_{0n_y} \\ B_{10} & B_{11} & \cdots & B_{1n_y} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n_x0} & B_{n_x1} & \cdots & B_{n_xn_y} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n_y} \end{pmatrix} \tag{6}$$

[16] The practical computation of the b_j coefficients will be explained in the next section. The purpose of a synthesis method is to construct B in order to obtain a solution that

simultaneously possesses good accuracy and robustness against experimental noise. Suppose that, in a realistic atmospheric situation, each a_i is allowed to fluctuate in the range $[\overline{a}_i - \delta a_i, \overline{a}_i + \delta a_i]$. If we simulate K realisations of this situation, this forces the i-th row of B to obey the following system of equations (possibly overdetermined):

$$\begin{pmatrix} b_0^{(1)} & b_1^{(1)} & \cdots & b_{n_y}^{(1)} \\ b_0^{(2)} & b_1^{(2)} & \cdots & b_{n_y}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ b_0^{(K)} & b_1^{(K)} & \cdots & b_{n_y}^{(K)} \end{pmatrix} \begin{pmatrix} B_{i0} \\ B_{i1} \\ \vdots \\ B_{in_y} \end{pmatrix} = \begin{pmatrix} a_i^{(1)} \\ a_i^{(2)} \\ \vdots \\ a_i^{(K)} \end{pmatrix} \tag{7}$$

where the rows of the matrix represent the K forward simulations obtained from equation 4. Equation 7 is more conveniently expressed by

$$D\,\vec{B}_i = \vec{A}_i \tag{8}$$

[17] The least squares solution of equation 8 is easily calculated as

$$\vec{B}_i^{LS} = \left(D^T D\right)^{-1} D^T \vec{A}_i \tag{9}$$

which is naturally unstable due to the lowest eigenvalues of the covariance matrix $D^T D$, i.e. the elements of $\vec{B}_i^{LS}$ tend to be very large and they will overamplify any experimental noise in the right-hand side of equation 6. On the other hand, if we know from our simulation of the K forward model realisations that the solution must lie in the range $[\overline{a}_i - \delta a_i, \overline{a}_i + \delta a_i]$, it is certainly reasonable to impose the following constraint on the size of the $\vec{B}_i$ elements:

$$\vec{B}_i^T \Delta b^2 \vec{B}_i = \vec{B}_i^T \begin{pmatrix} \delta b_0^2 & 0 & \cdots & 0 \\ 0 & \delta b_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta b_{n_y}^2 \end{pmatrix} \vec{B}_i = \delta a_i^2 \tag{10}$$

where the δb_i^2 elements stand for the respective variances of the b_i associated with the K simulations.

[18] The inverse problem reduces now to the constrained minimization of a merit function L_i that combines two criteria: closeness to measured data and regularization compatible with the natural sensitivity produced by the synthetic (and noiseless) simulation. L_i reads:

$$L_i = \| D\vec{B}_i - \vec{A}_i \|^2 + \lambda_i \left(\vec{B}_i^T \Delta b^2 \vec{B}_i - \delta a_i^2 \right) \tag{11}$$

where λ_i is the Lagrange multiplier. Formally, the last equation looks similar to various constrained linear inversion methods. Here, however, the Lagrange multiplier λ_i is not a "trade-off" parameter that can be arbitrarily tuned but it really constrains the solution not to extend outside the bounds that have been used for simulating the mapping between both solution and measurement domains. In other words, there is only one solution for λ_i and this value represents a hard constraint (equation 10). The value of $\vec{B}_i^*$ that minimizes L_i can be algebraically obtained as:

$$\vec{B}_i^* = \vec{B}_i^*(\lambda_i) = \left(D^T D + \lambda_i \Delta b\right)^{-1} D^T \vec{A}_i \tag{12}$$

whereas the value of λ_i can be found by numerically solving the secular [*Golub and Van Loan*, 1996] equation:

$$\vec{B}_i^{T*}(\lambda_i) \Delta b^2 \vec{B}_i^*(\lambda_i) - \delta a_i^2 = 0 \tag{13}$$

[19] The complete construction of B requires $n_x + 1$ resolutions of equations 12–13 but this has to be performed only once. It is clear that any signal for which the expansion coefficients b_i lie in the $\left[\overline{b}_i - \delta b_i, \overline{b}_i + \delta b_i\right]$ range is guaranteed to produce a reasonable solution inside the domain $[\overline{a}_i - \delta a_i, \overline{a}_i + \delta a_i]$ spanned by the K simulations.

3. Error Budget

[20] Let us define $\hat{F}(x)$, the function retrieved by applying the SIM method. According to Equation 3.16 of *Rodgers* [2000], the general error of the inversion $\hat{F}(x) - F(x)$ can be decomposed into two terms, if we accept that the forward model error is negligible in its representation and in the knowledge of its parameters.

[21] The first term is the "smoothing" error and arises from the imperfect restitution of a known profile when converted to a noiseless signal by the forward model and inverted back. The origins of such an error lie in the truncated expansions of equations 3–4, that also imply a finite resolution in the y domain, and the constrained least-squares solution of the mapping (equation 12). As quoted by *Rodgers* [2000], the true state is normally not known and it is impossible to compute the actual smoothing error. Instead, the covariance matrix of a real ensemble of states must be estimated. This is precisely a clear advantage of the SIM method because a natural ensemble to consider is the one that has been used to simulate the K realisations of the forward model. We can compute the covariance matrix of the error on the retrieved $\hat{a}_i$ coefficients:

$$C_a = \mathrm{cov}\left(\widehat{\vec{a}} - \vec{a}\right) \tag{14}$$

[22] The smoothing error variance can be finally computed as:

$$C_s(x) = U C_a U^T \tag{15}$$

where $U(x) = (U_0(x) \ldots U_{nx}(x))$.

[23] The second error term to consider is the retrieval noise, i.e. the spurious contribution that comes from inverted experimental random noise that is added on the signal and that can be described by a covariance $C_G(y)$, mostly diagonal as the noise fluctuations on consecutive measurements are considered independent. How does $C_G(y)$ propagate towards the solution? It is time to explain how to compute the b_j coefficients that represent the signal. If the measured signal is really a continuous function of y, the b_j coefficients are easily obtained by projecting the signal onto the basis functions:

$$b_j = \langle G(y) | V_j(y) \rangle \tag{16}$$

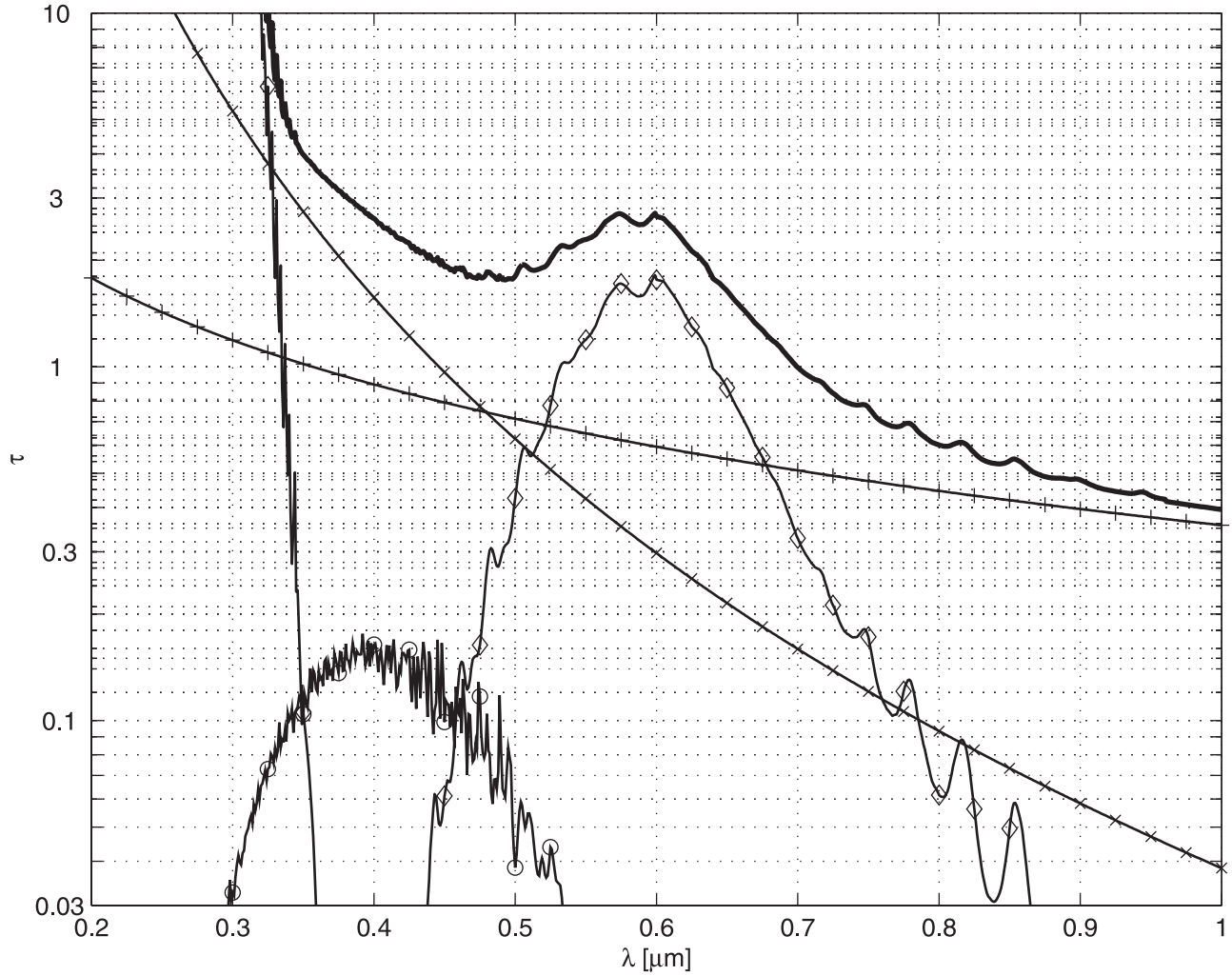

Figure 1. Slant path optical thicknesses at a tangent altitude of 20 km for climatological values. Crosses: air, diamonds: ozone, circles: nitrogen dioxide, plus signs: aerosol, full bold line: total.

If the signal is sampled at discrete values $y_1, y_2, \ldots y_l$, we may write:

$$G = (G(y_1)G(y_2)\ldots G(yl))^T = \vec{b}^T V \tag{17}$$

where V stands for

$$V = \begin{pmatrix} V_0(y_1) & V_0(y_2) & \cdots & V_0(yl) \\ V_1(y_1) & V_1(y_2) & \cdots & V_1(yl) \\ \vdots & \vdots & \ddots & \vdots \\ V_{n_y}(y_1) & V_{n_y}(y_2) & \cdots & V_{n_y}(yl) \end{pmatrix} \tag{18}$$

and $\vec{b}$ can be obtained as

$$\vec{b} = \left(VV^T\right)^{-1} VG^T \tag{19}$$

[24] Consequently, the random error variance is

$$C_r(x) = UB\left(VV^T\right)^{-1} VC_G V^T \left(VV^T\right)^{-1} B^T U^T \tag{20}$$

[25] It is worth noticing that the computation of $\vec{b}$ is equivalent to applying a low pass filter to the noisy signal. This is particulary evident by considering that the integration over y in the inner product of equation 16 smooths high frequency noise out except when the polynomial degree is very high. In other words, the SIM is also less sensitive to experimental noise not only due to the constraint used to build the B matrix but also because it realizes a mapping between the low frequencies of the signal and the low frequencies of the solution profile. High frequencies are simply disregarded in the inversion process.

[26] Assuming that the variances of both error terms can be added, we write for the total error covariance

$$C(x) = C_s(x) + C_r(x) \tag{21}$$

4. A SIM Application: Aerosol Spectral Inversion in Occultations Experiments

[27] We want to illustrate the possible application of the SIM method in a simple example of spectral inversion arising in occultation experiments. The reader is referred to *Fussen et al.* [2002] for a detailed discussion of this inverse problem. Briefly, we consider the atmospheric transmittance

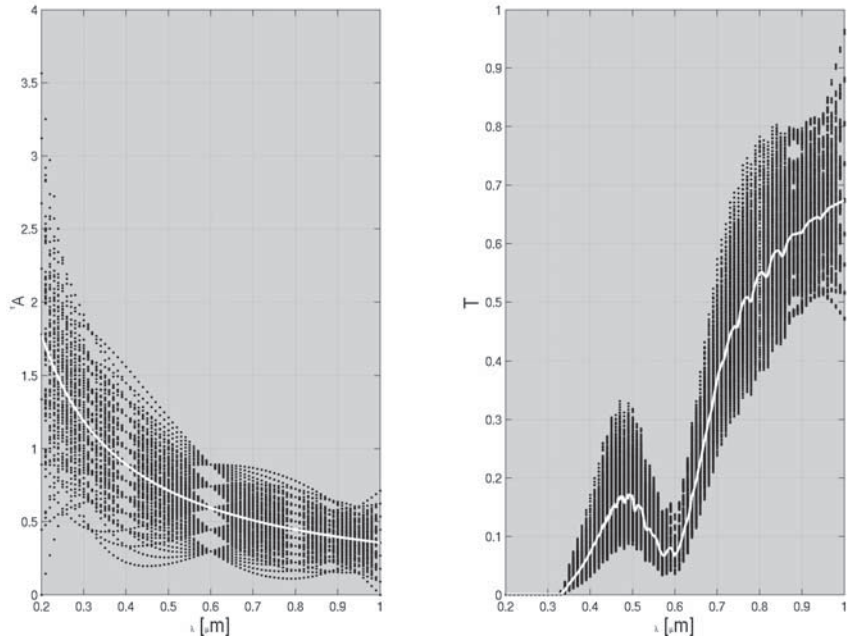

Figure 2. Left: the ensemble of simulated slant path aerosol optical thicknesses (full dots) spread around the reference $\tau_{A0}(\lambda)$ (white full line). Right: the corresponding simulated transmittances around the reference $T_0(\lambda)$.

measured by an ideal monochromatic occultation radiometer working at wavelength λ that can be written:

$$T(h,\lambda) = \exp(-\tau(h,\lambda)) \tag{22}$$

where h is the tangent altitude of the measured ray of light and $\tau(\lambda)$ is the slant path optical thickness due to absorption or scattering.

[28] If we restrict our analysis to the near UV - near IR range ($\lambda \in [0.2, 1.0]$ μm), the true optical thickness (unknown to the experimenter) may be described by:

$$\tau(\lambda) = a_{N_2}\tau_{N_2}(\lambda) + a_{O_3}\tau_{O_3}(\lambda) + a_{NO_2}\tau_{NO_2}(\lambda) + \tau_A(\lambda) \tag{23}$$

where τ_{N_2}, τ_{O_3}, τ_{NO_2} respectively stand for standard atmosphere slant path optical thicknesses of air, ozone and

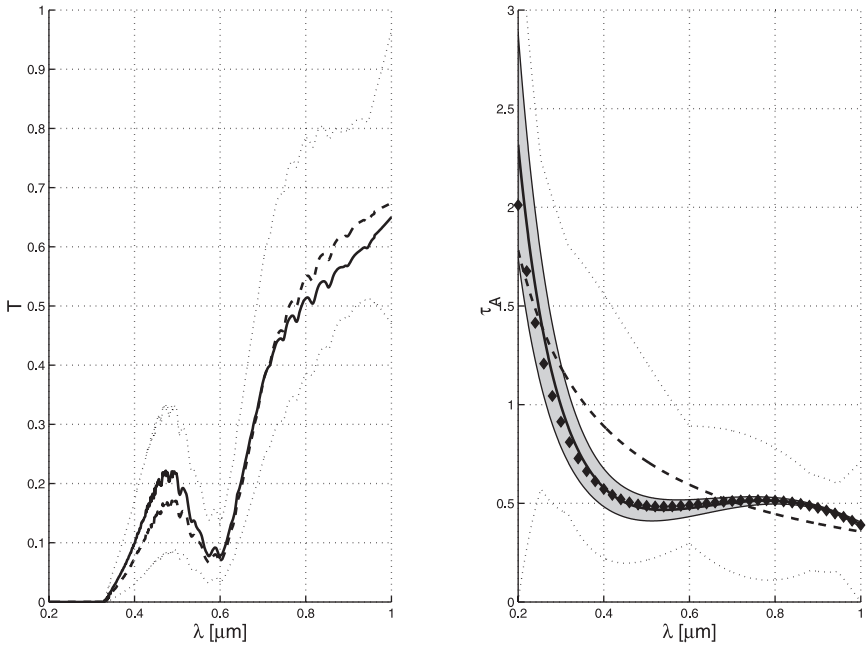

Figure 3. Left: the measured noiseless signal corresponding to the case defined by equation 27 (full line) together with the reference transmittance $T_0(\lambda)$ (dashed line) and the envelope of simulations (dotted lines). Right: Reference $\tau_{A0}(\lambda)$ (dashed line) and envelope (dotted lines). Full line surrounded by the gray band: exact solution and smoothing error. Diamonds: retrieved SIM solution, discretized for the sake of readability.

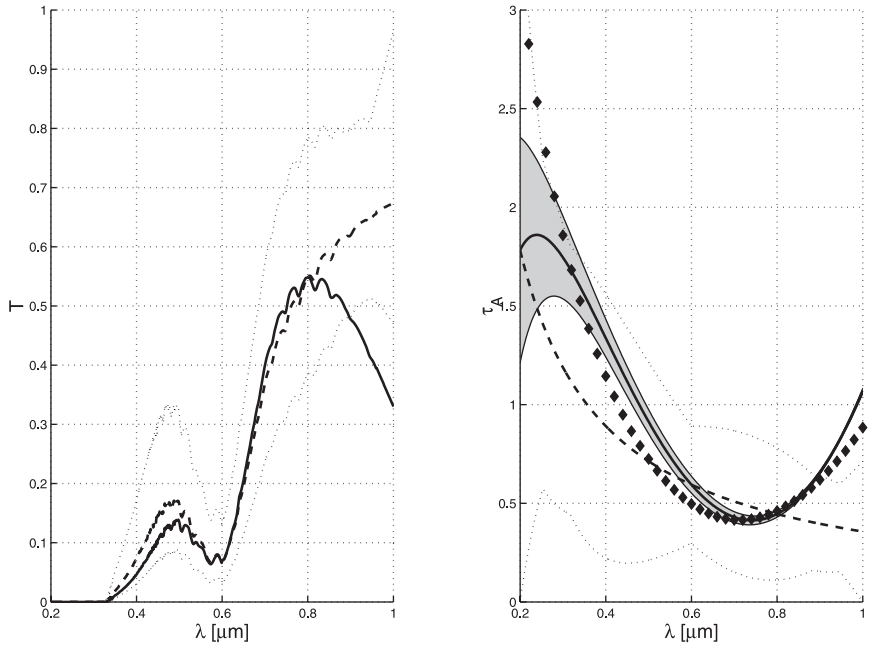

Figure 4. Same as for Figure 3 but for the case defined by equation 28.

nitrogen dioxide. For a standard atmosphere, a_{N_2}, a_{O_3} and a_{NO_2} are equal to 1. $\tau_A(\lambda)$ is the aerosol optical thickness of which the wavelength dependence is always smooth but unknown. The λ^{-1} aerosol wavelength dependence of $\tau_A(\lambda)$ in a moderate volcanic situation [*Fussen et al.*, 2001] is considered as our standard reference ($f_0(x)$ in equation 3) and will be referred to as $\tau_{A0}(\lambda)$ hereafter. In Figure 1, we have represented the partial slant path optical thicknesses for a tangent altitude of 20 km and for climatological values ($a_{N_2} = a_{O_3} = a_{NO_2} = 1$). The general expression for the aerosol optical thickness is chosen as

$$\tau_A(\lambda) = \tau_{A0}(\lambda)(1 + a_0 T_0(\lambda) + a_1 T_1(\lambda) + a_2 T_2(\lambda) + a_3 T_3(\lambda)) \tag{24}$$

where $T_i(\lambda)$ is the Chebyshev polynomial of order i defined on the above-mentioned wavelength range (see *Fussen et al.* [2002]) for a discussion about the maximal order to be used).

[29] In the forward model simulation, 7 parameters have been systematically varied (a_{N_2}, a_{O_3}, a_{NO_2}, a_0, a_1, a_2, a_3) positively and negatively around their mean values by realistic yet arbitrary deviations. We chose ($\delta a_{N_2} = 10\%$, ($\delta a_{O_3} = 20\%$, $\delta a_{NO_2} = 20\%$, $\delta a_{0,1,2,3} = 25\%$) so that a maximal 200% aerosol variability is covered. The full simulation is obtained by selecting all possible combinations of the 7 triplets of values ($-\delta a_X$, 0, $+\delta a_X$), leading to $K = 2187$ realisations of the forward model. Notice that a complete simulation is not indispensable and can be replaced by Monte-Carlo drawing if the forward model computation time does matter. The use of orthogonal polynomials is also an economical way to systematically span all possible aerosol wavelength dependences up to a given order. In Figure 2, we present the ensemble of aerosol realisations and the corresponding ensemble of signals obtained by applying the forward model to the complete ensemble of aerosol, air, ozone and nitrogen dioxide realisations.

[30] The SIM technique will consist in analyzing the difference signal $\delta T(\lambda) = T(\lambda) - T_0(\lambda)$ where $T(\lambda)$ and $T_0(\lambda)$ respectively stand for measured transmittance and reference transmittance (for $T_0(\lambda)$ we use $a_{N_2} = 1$, $a_{O_3} = 1$, $a_{NO_2} = 1$, $a_{0,1,2,3} = 0$). As 7 parameters are varied in the simulation, it is natural to develop $\delta T(\lambda)$ over a basis of 7 Chebyshev polynomials:

$$\delta T(\lambda) = \sum_{j=0}^{6} b_j T_j(\lambda) \tag{25}$$

The inverse mapping reads

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} B_{00} & B_{01} & \cdots & B_{06} \\ B_{10} & B_{11} & \cdots & B_{16} \\ \vdots & \vdots & \ddots & \vdots \\ B_{30} & B_{31} & \cdots & B_{36} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_6 \end{pmatrix} \tag{26}$$

and the B matrix is computed according to the algebra derived in section 2, i.e. row by row. The root of the secular equation 13 is computed by means of standard iterative algorithms. The resulting inverse mapping matrix B turns out to have elements ranging between -18.8 and 16.7, preventing dramatic noise amplification.

[31] A first inversion example is shown in Figure 3 for a hypothetical aerosol spectral behaviour characterized by

$$a_0 = 0, a_1 = 0.1, a_2 = 0.2, a_3 = -0.2 \tag{27}$$

[32] This example could be representative of an underlying bimodal particle size distribution with a fine and a coarse particle mode producing a nonmonotonic wavelength dependence. Such a dependence has been mentioned during the atmospheric relaxation after major volcanic eruptions when very large particles produced by coagulation have been observed [*Russell et al.*, 1996]. The measured signal exhibits clear departure from the reference but lies inside the envelope of the simulation ensemble. The aerosol slant path

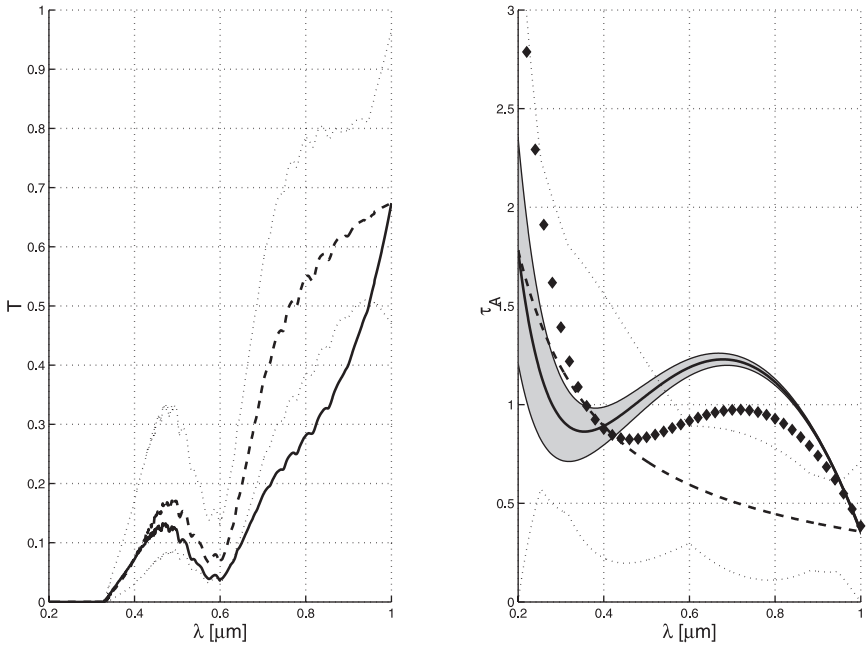

Figure 5. Same as for Figure 3 but for the case defined by equation 29.

optical thickness is well retrieved within the wavelength dependent estimated smoothing error computed by means of equation 15.

[33] The two following examples illustrate the response of the inversion algorithm to a signal produced by an unforeseen aerosol. Of course, the considered aerosol wavelength dependences are rather unrealistic but it is exactly the purpose of the test: can we detect unexpected aerosol extinction with the SIM method? In Figure 4, we selected

$$a_0 = 0.5, a_1 = 0.5, a_2 = 0.5, a_3 = 0.5 \tag{28}$$

which is outside the range of the simulated aerosol parameters but still corresponds to a transmittance mostly bounded by the extremal simulations. The retrieved aerosol is in close agreement with the exact one except at the edges of the wavelength interval. In Figure 5, a more extreme situation

$$a_0 = 0.5, a_1 = 0.5, a_2 = -0.5, a_3 = -0.5 \tag{29}$$

is selected with an unexpected transmittance. The error on the retrieved aerosol is now more pronounced although the

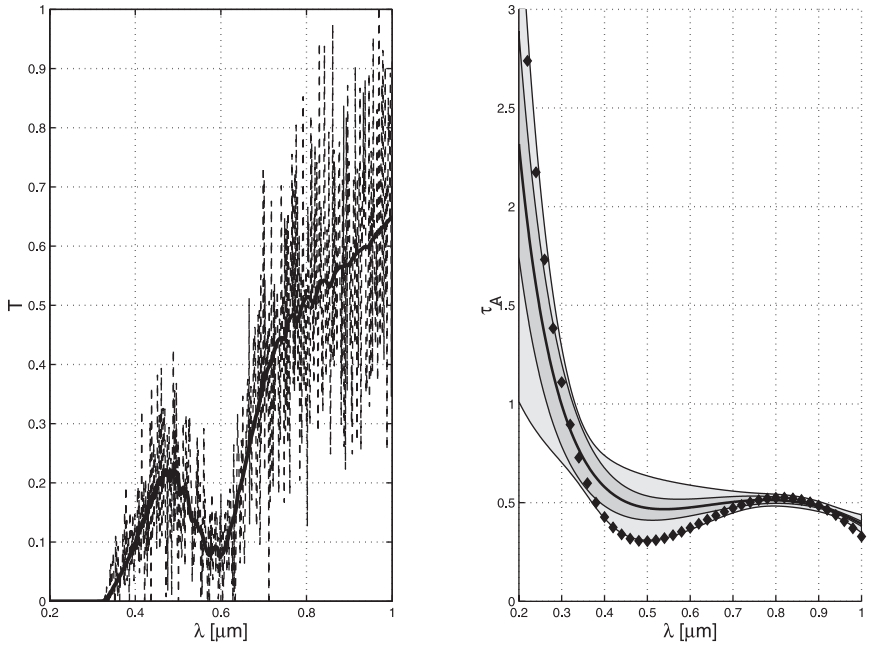

Figure 6. Same as for Figure 3 without references and envelopes. Left: Experimental noise defined by equation 30 and $S = 20$ has been added. Right: inner band: estimated smoothing error; outer band: estimated total error (smoothing + random, see equation 21).

solution is still acceptable as a first approximation of the actual solution. Notice, however, that the smoothing error estimation is not able to predict the retrieval error, as could be expected for a case lying outside the ensemble of simulations. The SIM inversion allows the natural detection of outlying signals just by comparing them to the envelope of simulations or by inspecting their expansion coefficients. It is up to the user to decide whether he should use a more extended simulation range or simply to discard the measurement.

[34] It is essential to check the SIM behavior when the signal is contaminated by experimental (random) noise. For the sake of simplicity we will only consider here shot noise, which is proportional to the square root of the signal for a Poisson distribution. If we call S the detector sensitivity (the dynamic range expressed in photons or electrons per wavelength unit), we may write for the error ΔT on transmittance (equation 22):

$$\Delta T = \frac{\sqrt{T}}{\sqrt{S}} \tag{30}$$

[35] Notice that the noise amplitude changes with the transmittance. In Figure 6, we have simulated the same case as in Figure 3, with important ($S = 20$) noise added. Beside the smoothing error, we have also reported the envelope of the total error estimation as expressed by equation 21. Clearly, the inversion of this very noisy signal leads to a solution with an error larger than the smoothing error at some wavelength regions. However, the retrieval is still acceptable whereas any application of an unregularized inversion method to such a signal would have resulted in a useless (and extremely wavy) solution. The limit of detectability can also be assessed by checking whether the projection of the experimental noise on the basis functions runs over the fluctuation range of simulations.

5. Conclusions

[36] We have developed a robust method based on synthesis inverse mapping. The essence of the method is twofold. Firstly, it promotes the idea of building a synthetic (as opposed to an algebraically derived) operator capable of inverting the mapping associated with the forward model. The operator is constrained to produce only solutions in an arbitrary range (around an arbitrary reference state) chosen by the user but without any reference to a true a priori knowledge or to an optimal vertical resolution. The synthesis of the operator is completely independent of the linear or nonlinear character of the forward model.

[37] Secondly, the SIM method recognizes that there is no reason to invert noise with a simultaneous regularization of its effects. Instead, the constrained mapping is considered to act between the principal components (and the lowest frequencies) of both signal and solution if they are expanded on a suitable basis of orthogonal polynomials. Doing so is equivalent to an optimal filtering of the signal that preserves the most significant part of its information content. In SIM, noise filtering and inversion are almost decoupled.

[38] We hope that the SIM method is a promising inversion technique. In this paper, we have only described the ideas and the general algebra to be used for implementing the method. As a concrete example of the application of SIM, we selected the retrieval of the aerosol extinction from the spectrum measured by an occultation experiment. Although our example is kept synthetic in order to assess the error budget exactly, it is however realistic and very close to what could be measured by well known space instruments like SAGE III on Meteor, GOMOS and SCIAMACHY onboard Envisat.

[39] In the near future we are going to investigate implementations of the method in various cases like optical and spatial inversions, full radiative transfer computations, etc. We will also consider multidimensional generalization.

[40] **Acknowledgments.** This work was partly performed within projects "SADE (Prodex 6)" and "Measurement, Understanding and Climatology of Stratospheric Aerosols (grant MO/35/004)" funded by the SSTC/DWTC service of the Belgian Government.

References

Eriksson, P., Analysis and comparison of two linear regularization methods for passive atmospheric observations, *J. Geophys. Res.*, *105*, 18,157–18,167, 2000.

Fussen, D., F. Vanhellemont, and C. Bingen, Evolution of stratospheric aerosols in the post-Pinatubo period measured by the occultation radiometer experiment ORA, *Atmos. Environ.*, *35*, 5067–5078, 2001.

Fussen, D., F. Vanhellemont, and C. Bingen, Optimal spectral inversion of atmospheric radiometric measurements in the near-UV to near-IR range: A case study, *Opt. Express*, *10*, 70–82, 2002.

Golub, G. H. and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore, Md., 1996.

Hansen, P. C., Numerical tools for analysis and solution of Fredholm integral equations of the first kind, *Inverse Probl.*, *8*, 849–872, 1992.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd ed., Cambridge Univ. Press, New York, 1992.

Rodgers, C. D., Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*, 5587–5595, 1990.

Rodgers, C. D., *Inverse Methods for Atmospheric Sounding*, World Sci., River Edge, N. J., 2000.

Russell, P. B., et al., Global to microscale evolution of the Pinatubo volcanic aerosol, derived from diverse measurements and analyses, *J. Geophys. Res.*, *101*, 18,745–18,763, 1996.

Twomey, S., *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, Elsevier Sci., New York, 1985.

C. Bingen, D. Fussen, and F. Vanhellemont, Institut d'Aéronomie Spatiale de Belgique, 3, avenue Circulaire, B-1180 Brussels, Belgium. (didier.fussen@oma.be)