Optimization problems in the estimation of parameters of thin films and the elimination of the influence of the substrate

Ernesto G. Birgin * Ivan E. Chambouleyron † José Mario Martínez ‡

Date of receipt: October 30, 2001
Date of receipt of the revised version: May 8, 2002

Abstract

In a recent paper, the authors introduced a method to estimate optical parameters of thin films using transmission data. The associated model assumes that the film is deposited on a completely transparent substrate. It has been observed, however, that small absorption of substrates affect in a nonnegligible way the transmitted energy. The question arises of the reliability of the estimation method to retrieve optical parameters in the presence of substrates of different thicknesses and absorption degrees. In this paper, transmission spectra of thin films deposited on non-transparent substrates are generated and, as a first approximation, the method based on transparent substrates is used to estimate the optical parameters. As expected, the method is good when the absorption of the substrate is very small, but fails when one deals with less transparent substrates. To overcome this drawback, an iterative procedure is introduced, that allows one to approximate the transmittance with transparent substrate, given the transmittance with absorbent substrate. The updated method turns out to be almost as efficient in the case of absorbent substrates as it was in the case of transparent ones.

Key Words: Methods of steepest descent type, Physical optics, Nonlinear programming, Mathematical programming, Applications of mathematical programming.

AMS Subject Classification: 49M10, 49M37, 65K05, 78A10, 90C90.

Aditional Key Words and Phrases: Optimization, Estimation of parameters, Unconstrained minimization, Constrained optimization, Thin films, Optical constants, Transmittance.

---

*Department of Computer Sciences IME-USP, University of São Paulo, Rua do Matão 1010, Cidade Universitária, 05508-090, São Paulo SP, Brazil. This author was supported by PRONEX-Optimization 76.79.1008-00, FAPESP (Grants 01/04597-4 and 02/00094-0) and CNPq (Grant 300151/00-4). Corresponding author. FAX: +55(11)3091-6134. e-mail: egbirgin@ime.usp.br

†Department of Applied Physics, Institute of Physics, University of Campinas, CP 6065, 13083-970 Campinas SP, Brazil. This author was supported by FAPESP and CNPq. e-mail: ivanch@ifi.unicamp.br

‡Department of Applied Mathematics IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil. This author was supported by PRONEX-Optimization 76.79.1008-00, FAPESP (Grant 01/04597-4), CNPq and FAEP-UNICAMP. e-mail: martinez@ime.unicamp.br

1 Introduction

In a recent paper [1] we introduced a pointwise unconstrained minimization method (PUMA) for estimating the thickness and the optical constants of thin films using transmission data. We used a transformation of a box-constrained optimization problem that came, after a suitable change of variables, from a convex-constrained optimization problem defined in [4, 5]. For solving the optimization problems we used spectral projected gradient techniques [2, 3]. In [13] applications to the estimation of parameters of real synthesized films are shown. This method assumes that the transmittance of a thin film deposited on a thick substrate obeys a model given in [14] (formula A1). See, also, [7, 12]. In this model, the transmittance is a function of the wavelength λ, the refractive index of the substrate s, the thickness of the film d, the refractive index of the film $n(\lambda)$ and the attenuation coefficient of the film $\kappa(\lambda)$. The inverse problem addressed in [1] consists of recovering the above parameters using transmission data.

The formulation [14] for computing transmissions does not use the thickness of the substrate and assumes that the substrate is transparent. However, real substrates are not completely transparent and it has been observed that this affects in a nonnegligible way the amounts of transmitted and reflected energy.

Practical measurements are also affected by the fact that pure waves of a single wavelength are not observed but transmission takes place with respect to a beam of waves of different length, according to a slit of the order of 1 nanometer.

The question addressed in this paper is: How do the thickness and absorption of the substrate and the size of the slit affect the estimates produced by PUMA? To answer this question we simulated the transmission through the 5 films considered in [1] with different conditions of substrate thickness and absorption and different slits. In this way, for each film we obtain 8 different spectra. We use these data to estimate thickness, refraction and absorption of the film using PUMA. The evaluation of results lead us to introduce an iterative procedure that eliminates the influence of the substrate absorption producing an estimation of the transmittance with transparent substrate using, as input, the transmittance with absorbent substrate.

This paper is organized as follows. In Section 2 we describe the way in which we did the simulations. In Section 3 we show the results of these simulations and we compare the spectra obtained under different conditions of the substrate and the slit. In Section 4 we apply the method [1] to the estimation of thickness, film refraction and absorption using all the spectra generated in the simulations. In Section 5 we introduce the iterative procedure that eliminates substrate absorption and we apply it to the films generated before. Conclusions are given in Section 6.

2 Description of the direct problem with transparent substrate

Suppose that we have a multilayer system of m layers, defined by their (complex) refractive indices $\tilde{n}_0, \ldots, \tilde{n}_{m-1}$. We write, for $\nu = 0, 1, \ldots, m-1$,

$$\tilde{n}_\nu = n_\nu - i\kappa_\nu.$$

The real part n_ν is the refraction coefficient and κ_ν is called the attenuation coefficient. We assume that the first and the last (semi-infinite) layers are transparent, so that $\tilde{n}_0 = n_0$ and $\tilde{n}_{m-1} = n_{m-1}$. Usually, the first layer is air, so that $\tilde{n}_0 = n_0 = 1$. In our application, $\tilde{n}_{m-1} = 1$ too. The interfaces between layers (assumed to be perfectly horizontal) are given by $x = L_1, \ldots, x = L_{m-1}$. In the first semi-infinite transparent layer ($x < L_1$), an incident wave is defined, given by

$$u(x,t) = E_T^0 \exp\left[i(wt - kx)\right],$$

where t represents time. This wave generates transmitted and reflected waves in all the layers. (In the last layer the reflected wave is null.)

In layer ν, for $\nu = 0, 1, \ldots, m-1$, the transmitted and reflected waves are given by

$$u_T^\nu(x,t) = E_T^\nu \exp\left[i(wt - k_\nu x)\right]$$

and

$$u_R^\nu(x,t) = E_R^\nu \exp\left[i(wt + k_\nu x)\right].$$

The first can be interpreted as a summation of infinitely many "transmitted small waves" and the second as a summation of "reflected small waves". The coefficient k is related to the wavelength λ by

$$k = \frac{2\pi}{\lambda}.$$

Moreover,

$$k_0 = k \quad \text{and} \quad k_\nu = \frac{k\tilde{n}_\nu}{\tilde{n}_0}$$

for $\nu = 1, \ldots, m-1$. Since there are no reflected waves in the last semi-infinite layer, we have that

$$E_R^{m-1} = 0.$$

Using the continuity of the waves and their derivatives with respect to x at the interfaces $L_1, \ldots, L_{m-1}$, we get, for $\nu = 1, 2, \ldots, m-1$:

$$E_T^{\nu-1} \exp\left(-ik_{\nu-1}L_\nu\right) + E_R^{\nu-1} \exp\left(ik_{\nu-1}L_\nu\right) = E_T^\nu \exp\left(-ik_\nu L_\nu\right) + E_R^\nu \exp\left(ik_\nu L_\nu\right)$$

and

$$-k_{\nu-1}E_T^{\nu-1} \exp\left(-ik_{\nu-1}L_\nu\right) + k_{\nu-1}E_R^{\nu-1} \exp\left(ik_{\nu-1}L_\nu\right) = -k_\nu E_T^\nu \exp\left(-ik_\nu L_\nu\right) + k_\nu E_R^\nu \exp\left(ik_\nu L_\nu\right).$$

So, using $k_\nu = k\tilde{n}_\nu/\tilde{n}_0$, we obtain:

$$\begin{pmatrix} 1 & 1 \\ -\tilde{n}_{\nu-1} & \tilde{n}_{\nu-1} \end{pmatrix} \begin{pmatrix} \exp(-ik_{\nu-1}L_\nu) & 0 \\ 0 & \exp(ik_{\nu-1}L_\nu) \end{pmatrix} \begin{pmatrix} E_T^{\nu-1} \\ E_R^{\nu-1} \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 1 \\ -\tilde{n}_\nu & \tilde{n}_\nu \end{pmatrix} \begin{pmatrix} \exp(-ik_\nu L_\nu) & 0 \\ 0 & \exp(ik_\nu L_\nu) \end{pmatrix} \begin{pmatrix} E_T^\nu \\ E_R^\nu \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} E_T^\nu \\ E_R^\nu \end{pmatrix} = \begin{pmatrix} \exp(ik_\nu L_\nu) & 0 \\ 0 & \exp(-ik_\nu L_\nu) \end{pmatrix} \frac{1}{2\tilde{n}_\nu} \begin{pmatrix} \tilde{n}_\nu + \tilde{n}_{\nu-1} & \tilde{n}_\nu - \tilde{n}_{\nu-1} \\ \tilde{n}_\nu - \tilde{n}_{\nu-1} & \tilde{n}_\nu + \tilde{n}_{\nu-1} \end{pmatrix}$$

$$\times \begin{pmatrix} \exp(-ik_{\nu-1}L_\nu) & 0 \\ 0 & \exp(ik_{\nu-1}L_\nu) \end{pmatrix} \begin{pmatrix} E_T^{\nu-1} \\ E_R^{\nu-1} \end{pmatrix}.$$

Let us write, for $\nu = 1, \ldots, m-1$,

$$A_\nu = \frac{1}{2\tilde{n}_\nu} \begin{pmatrix} \tilde{n}_\nu + \tilde{n}_{\nu-1} & \tilde{n}_\nu - \tilde{n}_{\nu-1} \\ \tilde{n}_\nu - \tilde{n}_{\nu-1} & \tilde{n}_\nu + \tilde{n}_{\nu-1} \end{pmatrix}.$$

Then,

$$\begin{pmatrix} E_T^{\nu+1} \\ E_R^{\nu+1} \end{pmatrix} = \begin{pmatrix} \exp(ik_{\nu+1}L_{\nu+1}) & 0 \\ 0 & \exp(-ik_{\nu+1}L_{\nu+1}) \end{pmatrix}$$
$$\times A_{\nu+1} \begin{pmatrix} \exp(-ik_\nu[L_{\nu+1} - L_\nu]) & 0 \\ 0 & \exp(ik_\nu[L_{\nu+1} - L_\nu]) \end{pmatrix}$$
$$\times A_\nu \begin{pmatrix} \exp(-ik_{\nu-1}L_\nu) & 0 \\ 0 & \exp(ik_{\nu-1}L_\nu) \end{pmatrix} \begin{pmatrix} E_T^{\nu-1} \\ E_R^{\nu-1} \end{pmatrix}.$$

Let $d_\nu \equiv L_{\nu+1} - L_\nu$ ($\nu = 1, \ldots, m-2$) be the thickness of layer ν. We define, for $\nu = 1, \ldots, m-2$,

$$M_\nu = A_{\nu+1} \begin{pmatrix} \exp(-ik_\nu d_\nu) & 0 \\ 0 & \exp(ik_\nu d_\nu) \end{pmatrix}.$$

Then, setting for simplicity and without loss of generality, $L_1 = 0$,

$$\begin{pmatrix} E_T^{m-1} \\ E_R^{m-1} \end{pmatrix} = \begin{pmatrix} \exp(ik_{m-1}L_{m-1}) & 0 \\ 0 & \exp(-ik_{m-1}L_{m-1}) \end{pmatrix} M_{m-2} \times \ldots \times M_1 A_1 \begin{pmatrix} E_T^0 \\ E_R^0 \end{pmatrix}.$$

Define, now,

$$M \equiv \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = M_{m-2} \times \ldots \times M_1 A_1.$$

and

$$M' \equiv \begin{pmatrix} M'_{11} & M'_{12} \\ M'_{21} & M'_{22} \end{pmatrix} = \begin{pmatrix} \exp(ik_{m-1}L_{m-1}) & 0 \\ 0 & \exp(ik_{m-1}L_{m-1}) \end{pmatrix} M.$$

Using $E_R^{m-1} = 0$ we obtain:

$$E_R^0 = -\frac{M'_{21}}{M'_{22}} E_T^0 = -\frac{M_{21}}{M_{22}} E_T^0$$

and

$$E_T^{m-1} = (M'_{11} - \frac{M'_{12}M'_{21}}{M'_{22}})E_T^0 = \exp(ik_{m-1}L_{m-1})(M_{11} - \frac{M_{12}M_{21}}{M_{22}})E_T^0.$$

Finally, define

$$r_{0,m-1} = -\frac{M_{21}}{M_{22}}$$

and

$$t_{0,m-1} = \exp(ik_{m-1}L_{m-1})(M_{11} - \frac{M_{12}M_{21}}{M_{22}}). \quad (1)$$

According to (1), the transmitted wave in the final layer is

$$(M_{11} - \frac{M_{12}M_{21}}{M_{22}})E_T^0 \;\; \exp\left[i(wt - k_{m-1}(x - L_{m-1})\right].$$

For energy computations, since $|\exp(ik_{m-1}L_{m-1})| = 1$, the presence of this factor in the computation of $t_{0,m-1}$ is irrelevant. The transmitted energy from layer n_0 to layer n_{m-1} is defined by

$$\text{Transmitted energy} = n_{m-1}|E_T^{m-1}|^2$$

and the reflected energy in layer n_0 is:

$$\text{Reflected energy} = n_0|E_R^0|^2.$$

Consequently, the transmittance $T(\lambda)$ and the reflectance $R(\lambda)$ are:

$$T(\lambda) = \frac{n_{m-1}}{n_0}|M_{11} - \frac{M_{12}M_{21}}{M_{22}}|^2 \tag{2}$$

and

$$R(\lambda) = |\frac{M_{21}}{M_{22}}|^2.$$

In many real situations the transmission of a pure wave of a definite wavelength cannot be measured. Instead, we measure an average of transmissions generated by a beam of waves that pass through a small slit. So, instead of $T(\lambda)$ we may take the view that we observe the integral

$$\text{Average slit transmitance} = \frac{1}{2\Delta\lambda}\int_{\lambda-\Delta\lambda}^{\lambda+\Delta\lambda} T(\lambda)d\lambda,$$

where $\Delta\lambda$ measures the half-ize of the slit.

If the last finite layer (called substrate from here on) is transparent, and we think $T(\lambda)$ as depending only on the thickness of this layer, we obtain a periodic function with period $\lambda/(2n_{m-2})$. This period is, usually, much less than the typical substrate thicknesses considered in applications. In many cases the period is far less than the measuring error of the thickness of the substrate. Therefore, a reasonable approximation of the observed transmission is the average of transmissions over that period. This gives:

$$\text{Average thickness transmittance} = \frac{2n_{m-2}}{\lambda}\int_{\bar{L}_{m-1}}^{\bar{L}_{m-1}+\lambda/(2n_{m-2})} T(\lambda)dL_{m-1}. \tag{3}$$

Clearly, this integral does not depend on L_{m-1}.

The integral (3) (for $m = 4$) can be computed analytically. See [9] (pp. 22-23). Swanepoel [14] (citing [12] and [7]) gave a practical organization of the calculation of this integral in the 4-layer case, where all except the second layer (the film) are transparent. Assume that the first layer is air, the second is a thin absorbing film, the third is a transparent substrate and the fourth is, again, air. We call d the thickness of the film, κ its attenuation coefficient, n its refraction coefficient, s the refraction coefficient of the substrate and λ the wavelength. In this case, (3) reduces to the following calculations:

$$\widehat{T} = \text{ Average thickness transmittance} = \frac{Az}{B - Cz + Dz^2} \tag{4}$$

where

$$A = 16s(n^2 + \kappa^2) \tag{5}$$

$$B = [(n+1)^2 + \kappa^2][(n+1)(n+s^2) + \kappa^2] \tag{6}$$

$$C = [(n^2 - 1 + \kappa^2)(n^2 - s^2 + \kappa^2) - 2\kappa^2(s^2+1)]2\ \cos\varphi$$
$$-\kappa[2(n^2 - s^2 + \kappa^2) + (s^2+1)(n^2 - 1 + \kappa^2)]2\ \sin\varphi \tag{7}$$

$$D = [(n-1)^2 + \kappa^2][(n-1)(n-s^2) + \kappa^2] \tag{8}$$

$$\varphi = 4\pi n d/\lambda, \quad z = exp(-\alpha d), \quad \alpha = 4\pi\kappa/\lambda. \tag{9}$$

3 Direct problem and simulations with general substrate

We want to simulate the spectral response (transmittance) of a film with the following characteristics:

- Thickness: d
- Refractive index: $n(\lambda)$
- Attenuation coefficient: $\kappa(\lambda)$.
- The film is deposited on a substrate given by:
- Thickness: d_S
- Refractive index: $s(\lambda)$
- Attenuation coefficient: $\kappa_S(\lambda)$.

According to (2) we have:

$$T(\lambda) = T(\lambda, d, n(\lambda), \kappa(\lambda), d_S, s(\lambda), \kappa_S(\lambda)).$$

For simulating the effect of a slit of size Δ and the variations of the substrate thickness, we compute:

$$T_{obs}(\lambda) = \int_{\lambda-\Delta}^{\lambda+\Delta} \int_{d_S}^{d_S+\mu/(2s)} T(\mu, d, n(\mu), \kappa(\mu), \delta, s(\mu), \kappa_S(\mu)) d\delta d\mu. \tag{10}$$

The integral (10) must be computed numerically. To ensure high precision we performed several experiments that lead us to a conservative choice of 1000 points for the integration with respect to the slit and 2000 points for integrating with respect to the substrate thickness. Therefore, for a given slit, substrate thickness and substrate attenuation, we have a spectral response that can be compared with the one obtained using the hypothesis of transparent substrate.

We simulated, using (10), the spectral responses of the five films considered in [1]. We also simulated the transmittance using (4), which corresponds to a transparent substrate and is the model used for the recovering procedure PUMA [1].

The five films considered are described below.

Film A: a-Si:H thin film deposited on a glass substrate. True thickness $d^{true} = 100$ nm. Refraction index given by formula (42) of [1] and absorption coefficient given by formula (43) of [1]. Wavelength range : $[540 - 1530]$ nm.

Film B: Similar film to A, with $d^{true} = 600$ nm. Wavelength range : $[620 - 1610]$ nm.

Film C: Hydrogenated amorphous germanium thin film deposited on a crystalline silicon substrate. $d^{true} = 100$ nm. Refraction index given by formula (44) of [1] and absorption coefficient given by formula (45) of [1]. Wavelength range : $[1250 - 2537]$ nm.

Film D: Similar film to C, with $d^{true} = 600$ nm.

Film E: Metal oxide thin film deposited onto glass. True thickness $d^{true} = 80$ nm. Refraction index given by formula (46) of [1] and absorption coefficient given by formula (47) of [1]. Wavelength range : $[360 - 657]$ nm.

The transmissions of these films have been computed using the algorithm given in Section 2 under the following conditions:

Case 1: Thin substrate (thin), Small slit (slit), Small attenuation (att).

Case 2: Thin substrate (thin), Small slit (slit), Large attenuation (ATT).

Case 3: Thin substrate (thin), Large slit (SLIT), Small attenuation (att).

Case 4: Thin substrate (thin), Large slit (SLIT), Large attenuation (ATT).

Case 5: Thick substrate (THICK), Small slit (slit), Small attenuation (att).

Case 6: Thick substrate (THICK), Small slit (slit), Large attenuation (ATT).

Case 7: Thick substrate (THICK), Large slit (SLIT), Small attenuation (att).

Case 8: Thick substrate (THICK), Large slit (SLIT), Large attenuation (ATT).

Case 9: Transparent substrate (computed using (4)).

In the classification above, the quantitative specification of the experiment conditions are:

- Thin substrate (thin): 0.2 millimeters.
- Thick substrate (THICK): 2 millimeters.
- Small slit (slit): 0.5 nanometers.
- Large slit (SLIT): 2 nanometers.
- Small attenuation (att): Constant attenuation coefficient equal to 10^{-7}.
- Large attenuation (ATT): Linearly decreasing attenuation coefficient being 10^{-5} at $\lambda = 500$ nanometers, and 10^{-7} for $\lambda \geq 2000$ nanometers.

Each film is now associated with 9 simulated spectra. We compare these spectra using two tables. In the first table (Features), we give:

- #max: The number of local maxima (excluding extrema of the interval in λ) of the transmittance.
- #min: The number of local minima (excluding extrema of the interval in λ) of the transmittance.
- argmax: The wavelength for which the global maximum of the transmittance occurs.
- max: Global maximum of the transmittance.
- argmin: The wavelength for which the global minimum of the transmittance occurs.
- min: Global minimum of the transmittance.
- average: Average transmittance over the whole spectrum considered.

Substrate information		Film features						
Case	Description	#max	#min	argmax	max	argmin	min	average
1	thin, slit, att	1	1	778	0.908	540	1.02E-4	0.4542
2	thin, slit, ATT	1	1	778	0.884	540	9.72E-5	0.4451
3	thin, SLIT, att	1	1	778	0.908	540	1.04E-4	0.4542
4	thin, SLIT, ATT	1	1	778	0.883	540	9.90E-5	0.4451
5	THICK, slit, att	1	1	778	0.906	540	1.01E-4	0.4531
6	THICK, slit, ATT	1	1	788	0.689	540	6.38E-5	0.3713
7	THICK, SLIT, att	1	1	778	0.906	540	1.03E-4	0.4531
8	THICK, SLIT, ATT	1	1	788	0.689	540	6.50E-5	0.3713
9	Swanepoel	1	1	778	0.909	540	1.02E-4	0.4543

Table 1: Film A - Features

Substrate information		Film features						
Case	Description	#max	#min	argmax	max	argmin	min	average
1	thin, slit, att	5	4	1422	0.916	620	1.02E-4	0.5329
2	thin, slit, ATT	5	4	1422	0.908	620	9.78E-5	0.5244
3	thin, SLIT, att	5	4	1422	0.916	620	1.04E-4	0.5329
4	thin, SLIT, ATT	5	4	1422	0.908	620	1.00E-5	0.5245
5	THICK, slit, att	5	4	1422	0.915	620	1.01E-4	0.5318
6	THICK, slit, ATT	5	4	1422	0.832	620	6.91E-5	0.4548
7	THICK, SLIT, att	5	4	1422	0.915	620	1.04E-4	0.5318
8	THICK, SLIT, ATT	5	4	1422	0.832	620	7.06E-5	0.4548
9	Swanepoel	5	4	1422	0.917	620	1.02E-4	0.5330

Table 2: Film B - Features

In Table 6 we display the distances between each spectrum and the spectrum defined by Case 9. The distance between two spectra is defined as the average value, with respect

Substrate information		Film features						
Case	Description	#max	#min	argmax	max	argmin	min	average
1	thin, slit, att	2	3	2537	0.471	1533	0.441	0.4523
2	thin, slit, ATT	1	2	2537	0.471	1495	0.436	0.4495
3	thin, SLIT, att	0	1	2537	0.471	1533	0.441	0.4523
4	thin, SLIT, ATT	0	1	2537	0.471	1495	0.436	0.4495
5	THICK, slit, att	0	1	2537	0.470	1533	0.440	0.4515
6	THICK, slit, ATT	0	1	2537	0.470	1263	0.386	0.4258
7	THICK, SLIT, att	0	1	2537	0.470	1533	0.440	0.4515
8	THICK, SLIT, ATT	0	1	2537	0.470	1263	0.386	0.4258
9	Swanepoel	0	1	2537	0.471	1533	0.441	0.4524

Table 3: Film C - Features

Substrate information		Film features						
Case	Description	#max	#min	argmax	max	argmin	min	average
1	thin, slit, att	3	2	2473	0.527	1456	0.437	0.4865
2	thin, slit, ATT	3	2	2473	0.527	1456	0.432	0.4836
3	thin, SLIT, att	3	2	2473	0.527	1456	0.437	0.4864
4	thin, SLIT, ATT	3	2	2473	0.527	1456	0.432	0.4837
5	THICK, slit, att	3	2	2473	0.527	1456	0.436	0.4857
6	THICK, slit, ATT	3	2	2498	0.526	1456	0.388	0.4589
7	THICK, SLIT, att	3	2	2473	0.527	1456	0.436	0.4857
8	THICK, SLIT, ATT	3	2	2498	0.526	1456	0.388	0.4589
9	Swanepoel	3	2	2473	0.527	1456	0.437	0.4866

Table 4: Film D - Features

to λ, of $|T_i(\lambda) - T_j(\lambda)|$, where T_i represents the transmittance in Case i and T_j is the transmittance for Case j.

The analysis of Tables 1–6 suggest that, as expected, the transmittances corresponding to Cases 1 and 3 are more similar to the transmittance generated assuming a transparent substrate. On the other hand, the characteristics of the transmittances generated with large and small slit seem to be very similar. To confirm this, we show, in Table 7, the distances between the cases $(\cdot$, slit, $\cdot)$ and $(\cdot$, SLIT, $\cdot)$ for each one of the films considered.

As can be seen in Table 7, the differences of transmissions for different slits are far below the usual measuring errors due to other factors. This seems to indicate that in models that simulate this physical phenomenon the influence of the slit can be disregarded. It must be noted, however, that this happens because one is considering averages with respect to thickness, which have the property of simulating the perturbations due to the slit and, perhaps, other perturbations such as the ones associated of slight deviation from the normal of the incident light.

Substrate information		Film features						
Case	Description	#max	#min	argmax	max	argmin	min	average
1	thin, slit, att	1	1	396	0.879	360	1.45E-3	0.7523
2	thin, slit, ATT	1	1	396	0.823	360	1.34E-3	0.7148
3	thin, SLIT, att	1	1	396	0.879	360	4.96E-3	0.7523
4	thin, SLIT, ATT	1	1	396	0.823	360	4.60E-3	0.7148
5	THICK, slit, att	1	1	396	0.874	360	1.44E-3	0.7488
6	THICK, slit, ATT	1	1	657	0.506	360	6.80E-4	0.4523
7	THICK, SLIT, att	1	1	396	0.874	360	4.93E-3	0.7488
8	THICK, SLIT, ATT	1	1	657	0.506	360	2.35E-3	0.4523
9	Swanepoel	1	1	396	0.880	360	1.24E-3	0.7527

Table 5: Film E - Features

Substrate information		Distances				
Case	Description	Film A	Film B	Film C	Film D	Film E
1	thin, slit, att	1.22E−04	1.27E−04	1.08E−04	1.17E−04	4.08E−04
2	thin, slit, ATT	9.23E−03	8.54E−03	2.83E−03	2.95E−03	3.79E−02
3	thin, SLIT, att	1.37E−04	1.89E−04	8.10E−05	8.49E−05	7.25E−04
4	thin, SLIT, ATT	9.25E−03	8.52E−03	2.82E−03	2.93E−03	3.81E−02
5	THICK, slit, att	1.22E−03	1.24E−03	8.18E−04	8.48E−04	3.89E−03
6	THICK, slit, ATT	8.30E−02	7.82E−02	2.66E−02	2.77E−02	3.00E−01
7	THICK, SLIT, att	1.23E−03	1.23E−03	8.17E−04	8.49E−04	4.18E−03
8	THICK, SLIT, ATT	8.30E−02	7.82E−02	2.66E−02	2.77E−02	3.00E−01

Table 6: Distances between transmittance with and without absorbent substrates

4 Estimation of film thickness and optical constants

In this section, we use PUMA to estimate thickness and optical parameters (attenuation and refraction) of the films generated in the previous section. We proceed exactly in the way described in our paper [1]. As a result, we obtain the estimations of thickness given in Table 8. In all the cases, we used 100 points for our estimation. This seems to be the minimum number of points that allows one to get reliable results.

Table 8, as well as the observation of the estimates of $n(\lambda)$ and $\kappa(\lambda)$ so far obtained, confirms our suspicion that thick and absorbent substrates seriously affect the estimates produced by PUMA. In general, in the cases (THICK, ·, ATT) we obtained poor estimates of the thickness and the optical parameters. On the other hand, with thin and nearly transparent substrates, the estimates are, in general, similar to the ones reported in [1], where the data were generated using [14]. An exception worth mentioning is Film C, where even in the cases (thin, ·, att) the estimates were not good. However, looking at Figure 7 of [1] we see that even when the data were generated using the transparent-substrate assumption, it was difficult to distinguish the quadratic errors for thicknesses between 97 and 105 nm. Therefore, we are in the presence of an essential underdetermination of the film thickness in this case.

	Film A	Film B	Film C	Film D	Film E
Case 1 - Case 3	3.829E−5	1.562E−4	7.893E−5	7.053E−5	3.270E−4
Case 2 - Case 4	3.726E−5	1.526E−4	7.830E−5	6.996E−5	3.042E−4
Case 5 - Case 7	1.957E−5	1.460E−4	6.693E−6	7.139E−6	3.233E−4
Case 6 - Case 8	1.454E−5	1.158E−4	6.218E−6	6.713E−6	1.570E−4

Table 7: Distances between (·, slit, ·) and (·, SLIT, ·)

Substrate information		Estimated thickness (nm)				
Case	Description	Film A	Film B	Film C	Film D	Film E
1	thin, slit, att	100	600	104	600	80
2	thin, slit, ATT	99	598	104	599	75
3	thin, SLIT, att	100	599	99	600	80
4	thin, SLIT, ATT	99	597	96	599	72
5	THICK, slit, att	100	599	100	600	82
6	THICK, slit, ATT	91	562	73	588	50
7	THICK, SLIT, att	100	599	99	600	82
8	THICK, SLIT, ATT	88	561	90	588	40

Table 8: Thickness estimation using PUMA

5 An iterative scheme that eliminates substrate absorption

The experiments of the previous section motivated us to introduce an iterative algorithm that eliminates the substrate absorption. Because of the good behavior of the algorithm based on the transparent model [1] we want to use this algorithm as a subroutine for the general procedure.

In principle, we want to approximate the integral (10). However, since numerical evidence shows that the effect of the slit is not relevant (when we take into account the thickness indetermination!) we can replace (10) by the single integral

$$I(\lambda) = \int_{d_S}^{d_S+\lambda/(2s)} T(\lambda, d, n(\lambda), \kappa(\lambda), \delta, s(\lambda), \kappa_S(\lambda)) d\delta. \quad (11)$$

In [9], pp. 22-23, it has been proved that, when $\kappa_S(\lambda) = 0$, this integral is

$$I(\lambda) = \frac{T_1(\lambda)T_2(\lambda)}{1 - R_1(\lambda)R_2(\lambda)}, \quad (12)$$

where $T_1(\lambda)$ is the transmittance from the first layer to the substrate, $T_2(\lambda)$ is the transmittance from the substrate to the last semi-infinite layer, $R_1(\lambda)$ is the reflectance on the backside of the film corresponding to a wavelength $\lambda/s(\lambda)$ and $R_2(\lambda)$ is the reflectance on the backside of the substrate. Therefore,

$$T_1(\lambda) = \frac{A_1 x}{B_1 - C_1 x + D_1 x^2}, \quad (13)$$

where

$$\begin{aligned}
\alpha &= 4\pi\kappa(\lambda)/\lambda, \\
\varphi &= 4\pi n(\lambda)d/\lambda, \\
x &= \exp(-\alpha d), \\
A_1 &= 16s(\lambda)(n(\lambda)^2+\kappa(\lambda)^2), \\
B_1 &= [(n(\lambda)+1)^2+\kappa(\lambda)^2][(n(\lambda)+s(\lambda))^2+\kappa(\lambda)^2], \\
C_1 &= [(n(\lambda)^2-1+\kappa(\lambda)^2)(n(\lambda)^2+\kappa(\lambda)^2-s(\lambda)^2)-4\kappa(\lambda)^2 s(\lambda)]2\cos(\varphi) \\
&\quad -\kappa(\lambda)[2(n(\lambda)^2-s(\lambda)^2+\kappa(\lambda)^2)+2s(\lambda)(n(\lambda)^2-1+\kappa(\lambda)^2)]2\sin(\varphi), \\
D_1 &= [(n(\lambda)-1)^2+\kappa(\lambda)^2][(n(\lambda)-s(\lambda))^2+\kappa(\lambda)^2],
\end{aligned}$$

$$T_2(\lambda) = \frac{4s(\lambda)}{(s(\lambda)+1)^2}, \tag{14}$$

and

$$R_2(\lambda) = 1 - T_2(\lambda) = \frac{(s(\lambda)-1)^2}{(s(\lambda)+1)^2}. \tag{15}$$

It can be proved that the formulae A1 of [14] exactly represent the expression (12). Knowing $I(\lambda)$, $T_1(\lambda)$, $T_2(\lambda)$ and $R_2(\lambda)$ (12)-(15), $R_1(\lambda)$ can be computed as

$$R_1(\lambda) = \frac{1}{R_2(\lambda)}[1 - T_1(\lambda)T_2(\lambda)/I(\lambda)]. \tag{16}$$

A good approximation for (11) has been given by Cisneros [6] (see, also, Chapter 11 of [8]). In principle, it is possible to adapt the philosophy of [1] to the model given by formulae A1-A10 of [6]. In this case the whole automatic differentiation procedure would be redone, the objective function would be changed and the nice properties associated with formula A1 of [14] might be lost.

Therefore, we preferred to introduce a different procedure that allows us to use PUMA exactly in the way it is coded, in an iterative way. The idea is to eliminate the influence of the substrate absorption from the data by means of a sequence of simple steps that involve a few applications of PUMA.

Let us define $T_1(\lambda), T_2(\lambda), R_1(\lambda), R_2(\lambda)$ as in (12)-(16). Then, in the general case, a good approximation for (11) is

$$J(\lambda) = \frac{T_1(\lambda)T_2(\lambda)\theta(\lambda)}{1 - R_1(\lambda)R_2(\lambda)\theta(\lambda)^2}, \tag{17}$$

where

$$\theta(\lambda) = \exp(-4\pi\kappa_S(\lambda)d_S/\lambda). \tag{18}$$

Assume that $\kappa_S(\lambda)$ (and, hence, $\theta(\lambda)$) is known. Assume that the measured transmittance function is $T_{obs}(\lambda)$. Our procedure begins by using PUMA to estimate d, $\kappa(\lambda)$ and $n(\lambda)$. Using these estimations we compute, for all λ in the grid, values $a(\lambda), b(\lambda)$ such that

$$a(\lambda) = T_1(\lambda)T_2(\lambda), \tag{19}$$

$$b(\lambda) = R_1(\lambda)R_2(\lambda), \quad (20)$$

and

$$a(\lambda)\theta(\lambda) = T_{obs}(\lambda)(1 - b(\lambda)\theta(\lambda)^2), \quad (21)$$

in a way that will be specified below. Using these $a(\lambda)$ and $b(\lambda)$ we compute, for all λ in the grid,

$$\widetilde{T}(\lambda) = \frac{a(\lambda)}{1 - b(\lambda)}. \quad (22)$$

The function $\widetilde{T}(\lambda)$ is a first approximation to the transmittance with elimination of the substrate absorption. Then the procedure restarts using $\widetilde{T}(\lambda)$ as "observed transmittances". The algorithmic description of this procedure is given below.

Algorithm 1:

Assume that the observed transmittances $T_{obs}(\lambda)$ are given for all $\lambda \in G$ (the grid). Set the first approximation $\widetilde{T}_0(\lambda) = T_{obs}(\lambda)$. Suppose that $\theta(\lambda)$ is known for all $\lambda \in G$. Set $k \leftarrow 0$.

Step 1. *Estimate d, $\kappa(\lambda)$, $n(\lambda)$*

Use PUMA to estimate d, $\kappa(\lambda)$, $n(\lambda)$ using $\widetilde{T}_k(\lambda)$ as "observed transmittances".

Step 2. *Define a new transmittance without substrate absorption*

Step 2.1. *Compute $a_k(\lambda), b_k(\lambda)$*

Compute $\widehat{T}(\lambda)$ using formulae (5)-(10) (A1 of [14]). Using d, $\kappa(\lambda)$ and $n(\lambda)$ obtained in the previous step, compute

$$a_k(\lambda) = T_1(\lambda)T_2(\lambda), \ \ b_k(\lambda) = 1 - a_k/\widehat{T}(\lambda). \quad (23)$$

Step 2.2. *Compute approximations of $a(\lambda)$ and $b(\lambda)$*

For all $\lambda \in G$, compute $\tilde{a}_k(\lambda), \tilde{b}_k(\lambda)$ as the solution $a(\lambda), b(\lambda)$ of:

$$\text{Minimize } (a(\lambda) - a_k(\lambda))^2 + (b(\lambda) - b_k(\lambda))^2 \quad (24)$$

subject to

$$a(\lambda)\ \theta(\lambda) = T_{obs}(\lambda)(1 - b(\lambda)\ \theta(\lambda)^2, \quad (25)$$

$$0 \le a(\lambda) \le 1,\ 0 \le b(\lambda) \le 1,\ a(\lambda) + b(\lambda) \le 1. \quad (26)$$

Step 2.3. *Compute an approximation of the transmittance without substrate absorption*

For all $\lambda \in G$, compute

$$\widetilde{T}_{k+1}(\lambda) = \frac{\tilde{a}_k(\lambda)}{1 - \tilde{b}_k(\lambda)}.$$

Step 3. *Stop or begin a new iteration.*

If the distance between $\widetilde{T}_{k+1}(\lambda)$ and $\widetilde{T}_k(\lambda)$ is very small, stop. Else, set $k \rightarrow k + 1$ and go to Step 1.

Clearly, when the stopping criterion, say

$$\frac{1}{\#G} \sum_{\lambda \in G} \frac{|\widetilde{T}_{k+1}(\lambda) - \widetilde{T}_k(\lambda)|}{\widetilde{T}_k(\lambda)} \leq \varepsilon,$$

is satisfied, then, by (23) and (25), the obtained optical parameters $d, n(\lambda), \kappa(\lambda)$ satisfy approximately

$$T_{obs}(\lambda) = \frac{T_1(\lambda)T_2(\lambda)\theta(\lambda)}{1 - \theta(\lambda)^2 R_1(\lambda)R_2(\lambda)}$$

and the phenomenological constraints of [1]. Therefore, they can be accepted as solutions of the estimation problem.

Table 9 shows the distance between the transmittances computed using Algorithm 1 and the transmittances without substrate absorption, and Table 10 shows the estimated thicknesses using these transmittances.

Substrate information		Distances				
Case	Description	Film A	Film B	Film C	Film D	Film E
1	thin, slit, att	2.66E−05	2.66E−05	7.28E−05	6.44E−05	2.54E−05
2	thin, slit, ATT	9.30E−04	1.19E−03	1.31E−03	1.36E−03	4.44E−04
3	thin, SLIT, att	2.87E−05	1.56E−04	1.24E−05	1.21E−05	3.47E−04
4	thin, SLIT, ATT	9.42E−04	1.20E−03	1.29E−03	1.34E−03	7.50E−04
5	THICK, slit, att	9.94E−06	9.74E−06	6.57E−06	5.67E−06	2.83E−05
6	THICK, slit, ATT	9.03E−03	1.23E−02	1.29E−02	1.31E−02	4.89E−03
7	THICK, SLIT, att	2.69E−05	1.55E−04	2.05E−06	3.30E−06	3.49E−04
8	THICK, SLIT, ATT	9.00E−03	1.22E−02	1.29E−02	1.31E−02	4.90E−03

Table 9: Distances beteentransmittance with transparent substrate and Transmittance generated by Algorithm 1

Substrate information		Estimated thickness (nm)				
Case	Description	Film A	Film B	Film C	Film D	Film E
1	thin, slit, att	100	600	112	600	80
2	thin, slit, ATT	100	600	112	599	82
3	thin, SLIT, att	100	600	100	600	82
4	thin, SLIT, ATT	100	600	103	599	80
5	THICK, slit, att	100	600	100	600	82
6	THICK, slit, ATT	98	591	115	587	82
7	THICK, SLIT, att	100	600	100	600	82
8	THICK, SLIT, ATT	98	592	116	587	82

Table 10: Thickness estimation using Algorithm 1

6 Conclusions

We have simulated the transmittance of the 5 films analyzed in [1] with different conditions of substrate thickness and absorption. We have verified that, in the case with large absorption and thickness in the substrate, the theoretical transmittance differs significantly from the transmittance of the film deposited on a completely transparent substrate. This is shown in Table 6 of this paper. Moreover, these differences have an important influence on the estimation of thickness, absorption and refraction of the film. The good news is that the slit has almost no influence, so it may be disregarded when one deals with this type of phenomenon. The reason is that the formulae for measured transmittances (without the slit) are averages over the substrate thicknesses and this average tends to simulate, also, the influence of the slit. The elimination of the slit as an influential factor is important because, with the slit, the transmittance can be computed only using expensive numerical integration procedures.

In order to eliminate, as much as possible, the absorption of the substrate, we have introduced an iterative process, based on a fixed point projection idea, which, in the limit, produces the transmittance on a transparent substrate given the observed transmittance. Applying this process we obtained new "transmittances on transparent substrates" which are compared with the theoretical transmittances with transparent substrate in Table 9. Here we can observe that the new transmittances are much closer to the transparent-substrate ones than the transmittances given by "observations". So, the "filtered" transmittances can be used to estimate the optical parameters of the film. In all cases, the application of PUMA to the filtered transmittances is much more successful than its application to the observed ones. In this paper we assumed that the absorption coefficient of the substrate is known. In future works we plan to estimate this coefficient within the optimization procedure. Moreover, it would be interesting to apply the philosophy of PUMA directly to approximations of the transmittance on non-transparent substrates, like the one introduced in [6].

The most important challenge that is being addressed now is the estimation of optical constants in cases where the constraints do not need to be satisfied exactly. A scheme based on "soft" and "hard" constraints allows one to use nonlinear programming techniques based on the inexact restoration philosophy [10, 11].

Acknowledgements

We are indebted to two anonymous referees for hepful comments and many corrections on the original manuscript.

References

[1] E. G. Birgin, I. Chambouleyron and J. M. Martínez. Estimation of the optical constants and the thickness of thin films using unconstrained optimization. *Journal of Computational Physics* 151, 862–880 (1999).

[2] E. G. Birgin, J. M. Martínez and M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization* 10, 1196-1211 (2000).

[3] E. G. Birgin, J. M. Martínez and M. Raydan. Algorithm 813: SPG – Software for convex-constrained optimization. *ACM Transactions on Mathematical Software* 27, 340–349, (2001).

[4] I. Chambouleyron, J. M. Martínez, A. C. Moretti and M. Mulato. Optical constants of thin films by means of a pointwise constrained optimization approach. *Thin Solid Films* 317, 133–136 (1998).

[5] I. Chambouleyron, J. M. Martínez, A. C. Moretti and M. Mulato. The retrieval of the optical constants and the thickness of thin films from transmission spectra. *Applied Optics* 36, 8238–8247 (1997).

[6] J. I. Cisneros. Optical characterization of dielectric and semiconductor thin films by use of transmission data. *Applied Optics* 37, 5262–5270 (1998).

[7] J. Keradec. *Thesis*. L'Université Scientifique et Médicale de Grenoble (1973).

[8] Z. Knittl. *Optics of thin films*, Wiley, New York (1971).

[9] H. M. Liddell. *Computer-aided techniques for the design of multilayer filters*, Adam Hilger Ltd, Bristol (1981).

[10] J. M. Martínez. Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. *Journal of Optimization Theory and Applications* 111, 39-58 (2001).

[11] J. M. Martínez and E. A. Pilotta. Inexact restoration algorithms for constrained optimization. *Journal of Optimization Theory and Applications* 104, 135-163 (2000).

[12] A. Mini. *Thesis*. L'Université Scientifique et Médicale de Grenoble (1982).

[13] M. Mulato, I. Chambouleyron, E. G. Birgin and J. M. Martínez. Determination of thickness and optical constants of a-Si:H films from transmittance data. *Applied Physics Letters* 77, 2133–2135 (2000).

[14] R. Swanepoel. Determination of the thickness and optical constants of amorphous silicon. *J. Phys. E: Sci. Instrum.* 16, 1214–1222 (1983).