

# Highly accurate optical material parameter determination with THz time-domain spectroscopy

Ioachim Pupeza, Rafal Wilk, and Martin Koch

*Joint Optical Metrology Center, Technical University Braunschweig and  
Institut für Hochfrequenztechnik, Technical University Braunschweig  
Schleinitzstr. 22, 38106 Braunschweig, Germany*

[martin.koch@ihf.tu-bs.de](mailto:martin.koch@ihf.tu-bs.de)

**Abstract:** We improve the existing data extraction algorithms for THz time-domain spectroscopy (THz TDS) in two aspects. On the one hand, we merge the up-to-date knowledge of THz TDS signal processing into a single powerful optical material parameter extraction algorithm. On the other hand, we introduce a novel iterative algorithm that further enhances the accuracy of the parameter extraction. In contrast to most of the published experiments, we are able to reliably investigate samples with thicknesses as small as 100µm, samples with low indexes of refraction, i.e. close to 1, as well as samples with sharp peaks in the material parameter curves.

© 2007 Optical Society of America

**OCIS codes:** (300.6340) Spectroscopy, infrared; (999.9999) Spatially Variant Moving Average Filter (SVMAF); (070.6020) Signal processing.

---

## References and links

1. L. DuVillaret, F. Garet, and J.-L. Coutaz, "A reliable method for extraction of material parameters in Terahertz Time-Domain Spectroscopy," *IEEE J. Sel. Top. Quantum Electron.* **2**, 739 - 746 (1996).
2. L. DuVillaret, F. Garet, and J.-L. Coutaz, "Highly precise determination of Optical Constants and sample thickness in Terahertz Time-Domain Spectroscopy," *Appl. Opt.* **38**, 409 - 415 (1999).
3. T. Dorney, R. Baraniuk, and D. Mittleman, "Material parameter estimation with Terahertz Time-Domain Spectroscopy," *J. Opt. Soc. Am. A* **18**, 1562 - 1571 (2001).
4. B. Ferguson and D. Abbott, "Signal processing for T-Ray Biosensor Systems," *Smart Electronics and MEMS II* **4236**, 157 - 169 (2001).
5. P. U. Jepsen and B. M. Fischer, "Dynamic range in Terahertz Time-Domain Transmission and Reflection Spectroscopy," *Opt. Lett.* **30**, 29 - 31 (2005).
6. D. Grischkowsky, S. Keiding, M. van Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *J. Opt. Soc. Am. B* **7**, 2006 - 2015 (1990).
7. B. Ferguson and D. Abbott, "Wavelet de-noising of optical terahertz pulse imaging data," *J. Fluct. Noise Lett.* **1**, L65 - L69 (2001).
8. B. Ferguson and D. Abbott, "De-noising techniques for terahertz responses of biological samples," *Microelectronics Journal (Elsevier)* **32**, 943 - 953 (2001).
9. D. L. Donoho, "De-noising by Soft-Thresholding," *IEEE Trans. Inf. Theory* **41**, 613 - 627 (1995).
10. J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM J. Optim.* **9**, 112 - 147 (1998).
11. S. Mickan, K.-S. Lee, T.-M. Lu, J. Munch, X.-C. Zhang and D. Abbott, "Thin Film Characterization using Terahertz Differential Time-Domain Spectroscopy and Double Modulation," *Electronics and Structures for MEMS II Proc. SPIE* **4591**, 197 - 209 (2001).
12. DIN Deutsches Institut für Normung, "DIN 1319 Fundamentals of metrology - Part 4: Evaluation of measurements; uncertainty of measurement," Beuth Verlag GmbH (1999).
13. M. N. Afsar, "Dielectric Measurements of Millimeter-Wave Materials," *IEEE Trans. Microwave Theory and Techniques* **12**, 1598 - 1609 (1984).

## 1. Introduction

*Terahertz time-domain spectroscopy* (THz TDS) is a technique for measuring the complex refractive index of materials over a wide range of frequencies, usually spanning from a few tens of GHz to several THz. In the last decade several authors presented material parameter extraction algorithms to determine the complex refractive indexes of nearly homogenous solid samples with THz TDS, see Ref. [1, 2, 3, 4, 5]. Commonly, for material parameter extraction, a THz pulse which has propagated through a sample is compared to a THz pulse obtained without the sample in its propagation path.

The sample material acts as a Fabry-Perot resonator for the THz pulse. Thus, the measured signal exhibits multiple reflections of the THz pulse. A simplistic method for approximating the optical material parameters considers a narrow time window for the sample waveform, containing only the first pulse propagated through the sample. However, the isolation of the first pulse may not be possible in the cases of low refractive indexes and/or thin samples. Another major drawback of this method is the necessity of a very accurate a priori knowledge of the sample thickness.

The recent papers Ref. [2] and Ref. [3] present algorithms leading to much more accurate results than the simple method mentioned above. Additionally, they allow for an accurate determination of the sample thickness. However, even though very large, the spectrum of materials suited for investigation with these algorithms is constrained by certain boundaries, such as relatively high refractive indexes and/or relatively thick samples (compare with the limits of the method in Ref. [3]). Furthermore, the accuracy of the existing algorithms is restricted by the uncertainty of the measured signals.

In this paper we present two important advancements to THz TDS. On the one hand we improve the existing state of the art material parameter extraction algorithms. We merge the up-to-date knowledge on THz TDS signal processing presented in Ref. [2, 3, 4] into a single powerful material parameter extraction algorithm. On the other hand we introduce a novel algorithm that enhances the accuracy of material parameter extraction by using the uncertainty of the measurements in a constructive way. We define a confidence interval for the theoretically derived transfer function, which allows for the formulation of additional physical boundary conditions to the investigated sample. We introduce the *Spatially Variant Moving Average Filter* (SVMAF), which constitutes the core of an iterative algorithm for finding the optimal material parameter curves within the allowed confidence intervals considering the additional physical boundaries.

With our algorithm we reach a very high accuracy of the extracted complex indexes of refraction of materials. In contrast to other published results, our algorithm performs very accurately even in the difficult cases of thin samples, with thicknesses as small as 100 $\mu$ m, and samples with low refraction indexes. Hence, it has a wide range of applicability.

## 2. The experimental setup

Our THz TDS setup is very similar to the setup described in Ref. [6]. The heart of our setup is a femtosecond titanium sapphire laser operated by a diode pumped solid state laser. It produces 30fs laser pulses with a repetition rate of 80MHz. The laser pulses are used to excite two photoconductive dipole antennas. One part of the optical beam is focused onto the emitter antenna which consists of GaAs onto which two metallic striplines are deposited. The laser pulses generate free carriers in between the striplines which are accelerated by an applied bias.

The resulting time-dependent photocurrent has a subpicosecond rise time and acts as a source for the subpicosecond THz pulses. These pulses are radiated from the antenna and are then collimated, focused and guided by four off-axis paraboloidal mirrors to form an intermediate focus in which the sample is placed. Another antenna is used to detect the THz radiation. In this antenna carriers are excited by the second part of the optical laser pulse which has passed a variable delay line. The electric field of the THz pulse is then used to bias the detector. The induced photocurrent is proportional to this electric field and the transient carrier population. The time-dependent photocurrent is measured by scanning the delay line. The time resolution of this measurement is limited mainly by the carrier lifetime in the silicon on sapphire (SOS) photoconductive antenna detector. For further details see Ref. [6].

### 3. Analytical model of the transfer function

In this section we derive an analytical formula for the transfer function. This is done in close analogy to Ref. [3]. We include this derivation here since it is of capital importance for the understanding of the subsequent sections. Throughout this section we highlight the aspects that we consider particularly important.

#### 3.1. The inverse electromagnetic problem

The method used by THz TDS to compute material parameters is based on measuring the effect caused on the electric field of a THz wave by the pass through the investigated material. Since the parameters are not computed directly, but from their effect on the electric field, this electromagnetic problem is called an *inverse* problem.

The temporal profile of the THz pulse propagated through a material represents the convolution in time domain of the emitted THz waveform with the pulse response of the material. The technique widely used in THz TDS is the comparison of two signals in frequency domain: a *reference signal*  $E_{ref}(\omega)$  and a *sample signal*  $E_{samp}(\omega)$ , where  $\omega$  denotes the angular frequency. Both electric fields in the frequency domain are directly proportional to the Fourier transforms of the respective photocurrents multiplied by  $j\omega$ . The sample signal is obtained by recording the photocurrent of the reference signal, additionally propagated through the sample material.

We shall denote by  $E_{ref}^{ex}(\omega)$  and  $E_{samp}^{ex}(\omega)$  the recorded reference and sample signals, respectively. The frequency dependent experimentally determined transfer function of the sample material is given by their quotient, which corresponds to the deconvolution in the time domain:

$$H_{experiment}(\omega) = \frac{E_{samp}^{ex}(\omega)}{E_{ref}^{ex}(\omega)}. \quad (1)$$

Let  $H_{theory}(\omega)$  be a theoretically derived transfer function of the sample material, which depends on the material parameters. We will derive an analytical expression for  $H_{theory}(\omega)$  in Section 3.3.

Let  $n_1(\omega)$  be the real refractive index and  $\alpha_1(\omega)$  be the absorption coefficient of a material we indicate with the index 1. As usual, we denote the complex refractive index of the material 1 by  $\tilde{n}_1(\omega) = n_1(\omega) - j\kappa_1(\omega)$ , where  $\kappa_1(\omega)$  is the extinction coefficient indicating the absorption. The relation between the absorption coefficient  $\alpha_1(\omega)$  and the extinction coefficient  $\kappa_1(\omega)$  is  $\kappa_1(\omega) = \alpha_1(\omega)c_0/(2\omega)$ , where  $c_0$  is the speed of light in air.

The inverse electromagnetic problem consists in finding the material parameters of the sample material, such that for every frequency  $\omega$  in the considered frequency range, the value  $H_{theory}(\omega)$  is as close as possible to  $H_{experiment}(\omega)$ . As a measure for how close these two functions are, we define the following error function  $Err$ :

$$M(\omega) = |H_{theory}(\omega) - H_{experiment}(\omega)|, \quad (2)$$

$$A(\omega) = \angle H_{theory}(\omega) - \angle H_{experiment}(\omega), \quad (3)$$

$$Err = \sum_{\omega} (|M(\omega)| + |A(\omega)|) \quad (4)$$

Note that all computations in this work are made in discrete time and frequency domains. This explains the sum over all frequencies in equation (4).

Solving the inverse electromagnetic problem for a material 1 with unknown material parameters is therefore equivalent to minimizing the value of *Err* with respect to the functions  $n_1(\omega)$  and  $\kappa_1(\omega)$ . Since  $\omega$  has discrete values, equation (4) implies that this problem is equivalent to minimizing the expression  $|M(\omega)| + |A(\omega)|$  with respect to  $n_1(\omega)$  and  $\kappa_1(\omega)$  for each frequency  $\omega$  separately.

### 3.2. Assumptions and notation

In the subsequent section we derive an analytical formula for the transfer function. The assumptions made about the experimental setup are very important for the validity of this theoretical model. In the following, we give a list of conditions which we consider fulfilled throughout this paper as well as the standard notation that we use.

- *Homogeneity of the investigated sample*

This is tantamount with its material parameters being spatially and directionally invariant. However, the frequency dependence of the material parameters may be arbitrary. In other words, absorption peaks and other variations of the material parameters are detected without losses of the quality of the results.

- *Flat and parallel surfaces of sample materials and no scattering*

We neglect roughness and curvature of the irradiated materials, as well as scattering effects at the surfaces and inside the material. In our analytical transfer function derivation, we assume that all surfaces are perfectly flat and parallel.

- *Dry atmosphere in the environment of the experiment*

We assume that the experimental setup is purged with dry Nitrogen gas to avoid complications due to water vapor.

- *Orthogonal incidence of the THz rays*

For simplicity we presume that the THz rays impinge orthogonally on the irradiated samples. The sample is placed in the spectrometer with an inclination error smaller than  $5^\circ$ , which leads to deviations of the refractive index and the extinction coefficient smaller than 0.002. Thus, we neglect this error.

- *Fresnel coefficients notation*

With  $R_{12}$  and  $T_{12}$  we denote the Fresnel reflection and transmission coefficients, respectively. They describe the reflection and transmission of a THz wave at the interface of material 1 and material 2 in the time domain:

$$R_{12} = \frac{\tilde{n}_2 - \tilde{n}_1}{\tilde{n}_1 + \tilde{n}_2}, \quad (5)$$

$$T_{12} = \frac{2\tilde{n}_1}{\tilde{n}_1 + \tilde{n}_2}. \quad (6)$$

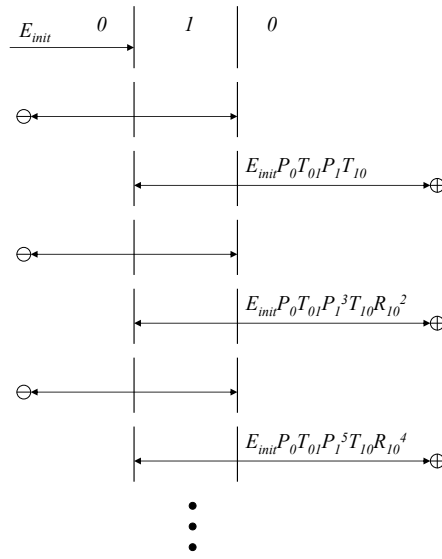


Fig. 1. Ray propagation through a planar, homogenous sample.

Furthermore, the propagation of an electric field  $E_{init}$  entering the material 1 is governed by the following relation:

$$E(z) = E_{init} P_1(z), \quad (7)$$

$$P_1(z) = \exp\left(\frac{-j\omega \tilde{n}_1 z}{c_0}\right), \quad (8)$$

where  $E(z)$  is the electric field after the geometrical propagation length  $z$  and  $P_1(z)$  is the frequency dependent propagation coefficient of the material 1.

- *Standard notation used throughout this work*

For computations, we use the value  $c_{vacuum} = 2.99796 \cdot 10^8$  m/s for the speed of light in vacuum. We approximate the speed of light  $c_0$  in air by  $c_{vacuum}$  and for the complex refractive index of air we use the value  $\tilde{n}_0 = 1.00027 - j0$ . We always indicate the medium air with the index 0. With  $x$  we denote the geometrical length of the optical path between the emitter and receiver antenna and by  $\ell$  the thickness of the investigated sample. The frequency dependent initial electric field leaving the emitter antenna is denoted by  $E_{init}$ .

### 3.3. The analytical formula of the transfer function

Figure 1 shows the ray propagation through the investigated sample in seven different stages. First, the initial electric field  $E_{init}$  impinges on the front (left) surface of the sample. In the second stage, we have depicted the two rays that emerge from  $E_{init}$  being split at the air-sample interface into a transmitted and a reflected portion. The reflected portion ends with the symbol  $\ominus$ , which indicates that this portion of the ray has no influence on  $E_{samp}$  and thus will be discarded. The transmitted portion is equal to  $E_{init} T_{01}$  and propagates through the sample. When

it reaches the second (right) surface of the sample, the electric field of the wave is equal to  $E_{init}T_{01}P_1(\ell)$ . The next sample-air interface causes the beam to split again. This process is depicted in the third stage of the figure. The transmitted portion, which is equal to  $E_{init}T_{01}P_1(\ell)T_{10}$  is a part of  $E_{samp}$ . This fact is indicated by the symbol  $\oplus$  in the third stage of the figure. Subsequently, the reflected part of the beam leads to a periodical reflection pattern, which we refer to as a *Fabry-Perot reflections*. The three dots at the bottom of Fig. 1 indicate the infinite continuation of this reflection pattern. For simplicity, let  $P_1$  denote  $P_1(\ell)$ . Also, the propagation of the electromagnetic wave through the air between the two antennas has to be considered. This implies the multiplication by a factor  $P_0(x - \ell)$ , which we abbreviate by  $P_0$ . Summing up all components of the electric field transmitted through the sample yields the following formula:

$$E_{samp} = E_{init}P_0T_{01}P_1T_{10} \left[ 1 + \sum_{i=1}^{\delta} (R_{10}^2P_1^2)^i \right]. \quad (9)$$

The integer  $\delta$  in equation (9) represents the number of Fabry-Perot reflections taken into account. Theoretically, this number is infinitely large. Practically, this number is upper bounded: on the one hand, the energy of the electric field is attenuated exponentially with every pass through the sample and on the other hand, the time window of our measurement is finite. In Section 4.2 we present an algorithm for the optimal choice of  $\delta$ .

The reference THz pulse propagates through air, thus the following equation holds:

$$E_{ref} = E_{init}P_0(x). \quad (10)$$

Since  $P_0$  denotes  $P_0(x - \ell)$ , the equalities (9) and (10) together with (8) yield the following theoretical transfer function:

$$H_{theory} = \frac{E_{samp}}{E_{ref}} = P_0(-\ell)T_{01}P_1T_{10} \left[ 1 + \sum_{i=1}^{\delta} (R_{10}^2P_1^2)^i \right] \quad (11)$$

#### 4. Material parameter extraction

For a structured presentation of our material parameter extraction algorithms, we have divided this section into four subsections. Each subsection contains a stand-alone step of the extraction process. In the subsections 4.1 to 4.3, we present several improvements of the existing extraction algorithms. In 4.4 we introduce the Spatially Variable Moving Average Filter, which leads to highly accurate results.

##### 4.1. Preliminary signal processing

Before proceeding to the actual material parameter extraction, we perform the following preliminary signal processing:

- *Linear offset compensation of raw time domain data*

We assume the existence of a linear offset in the measured photocurrent, which is an often observed experimental artifact. For its compensation, we subtract the linear offset computed from the mean of several data points before the incidence of the pulse and the mean of the last few data points.

- *Determination of the reliable frequency range*

The reliable frequency range of our experiment is determined as the frequency band, in which the magnitude of the sample signal is greater than or equal to the noise floor of

the experiment. The noise floor is obtained by acquiring data from the spectrometer after blocking the THz beam path, and represents the frequency independent detector noise, as discussed in Ref. [5]. In this fashion, we prevent the misinterpretation of absorption coefficients in regions of the spectrum, where there is a detectable reference signal but the entire radiation is absorbed by the sample. Therefore, our method yields similar results to the analytical technique described in Ref. [5].

- *Noise cancellation techniques*

In Ref. [4, 7, 8], noise cancellation techniques such as wavelet de-noising or Wiener filter are presented for a similar experimental setup. However, such techniques may alter the signal shapes in undesired ways. For example, using different wavelets for de-noising leads to different material parameter curves of the same sample. For a comprehensive description of the wavelet de-noising technique by soft thresholding, consult Ref. [9]. The optimal wavelet for de-noising depends on the THz pulse shape, which may vary for different antennas. Using sub-optimal wavelets may lead to suppression of important peaks in the material parameter curves. Therefore, we discard noise cancellation techniques in our work. Instead, we introduce the novel SVMAF technique described in Section 4.4. While respecting the investigated material's physical properties, our iterative technique suppresses noise on the one hand and statistical errors on the other hand, yielding highly accurate material parameter curves.

#### 4.2. The core of the extraction algorithm

In contrast to Ref. [3], for the minimization of the error  $Err$  expressed in equation (4) with respect to the material parameters  $n_1(\omega)$  and  $\kappa_1(\omega)$  we use the Nelder-Mead algorithm, see Ref. [10]. A similar procedure has been reported in Ref. [11] for the THz differential TDS method. In the following, we derive the initial values which we use for the minimum search algorithm. By measuring the sample thickness with a micrometer screw, we obtain an initial value for the variable  $\ell$ . Furthermore, we assume that the attenuation of the maximum of the magnitude of the THz pulse in the sample is given by the following relation:

$$|E_{s,max}| = |E_{r,max}| \cdot \exp\left(-\frac{\omega}{c_0} \kappa_1 \ell\right), \quad (12)$$

where  $|E_{s,max}|$  and  $|E_{r,max}|$  represent the maximal magnitudes of the sample and reference signals, respectively, in the frequency domain. From equation (12) we obtain the following initial value for  $\kappa_1$ :

$$\kappa_1 = -\frac{1}{\ell} \cdot \frac{c_0}{\omega} \log \frac{|E_{s,max}|}{|E_{r,max}|}. \quad (13)$$

An initial value for the real refractive index is also easily found. While passing the sample, the maximum of the magnitude of the reference signal in the time domain is delayed by the time  $\Delta t$  due to a difference of the real refractive index  $\Delta n = n_1 - n_0$ . This delay can be expressed by the following equation:

$$\Delta t = \frac{\ell \Delta n}{c_0}. \quad (14)$$

From equation (14) we obtain the following initial value for  $n_1$ :

$$n_1 = \frac{c_0 \Delta t}{\ell} + n_0. \quad (15)$$



In order to proceed with the minimum search, we only need to determine  $\delta$  in equation (11). For this purpose, we use the rough approximation of  $n_1$  from equation (15). Let  $t_{max}$  be the length of the time window measured from the incidence of the reference pulse until the end of the measurement. The integer  $\delta$  indicates the number of copies of the main pulse that fit in the considered time window of length  $t_{max}$ . Inside the sample, the wave propagates with the speed  $c_0/n_1$  along a path of length  $\ell$ . Thus, We determine  $\delta$  as the greatest integer, such that the following inequality is satisfied:

$$t_{max} \geq \frac{n_1}{c_0} \ell (1 + 2\delta). \quad (16)$$

Note, that the factor 2 arises from the fact that each reflection at the sample's exit interface traverses the sample back and forth before reaching the detector.

The possibility of setting the termination threshold of the Nelder-Mead search allows us to find values for the material parameters, such that  $H_{theory}$  is arbitrarily close to  $H_{experiment}$ . A typical value we used for this threshold is  $10^{-4}$ .

#### 4.3. Accurate determination of the sample thickness

We employ the total variation technique described in Ref. [3], but, in contrast to Ref. [3], we combine the total variation technique with the Nelder-Mead search. The multiple Fabry-Perot reflections of the ultrashort THz pulse cause an oscillation of the sample signal and, thus, of the transfer function, in the frequency domain. A closer look at equations (5) to (8) and (11) reveals that this oscillation is determined by the sample thickness  $\ell$  and the complex refractive index  $\tilde{n}_1(\omega)$ . Thus, while fitting  $H_{theory}$  to  $H_{experiment}$  by searching the minimum of the error  $Err$ , any uncertainty of  $H_{theory}$  caused by an uncertainty of  $\ell$  has to be compensated by the search variables  $n_1(\omega)$  and  $\kappa_1(\omega)$ . In other words, an uncertainty of  $\ell$  is noticeable in the extracted material parameter curves  $n_1(\omega)$  and  $\kappa_1(\omega)$  as an oscillation due to the Fabry-Perot effect.

We extract material parameter curves for a range of values of the sample thickness  $\ell$ . By implementing a measure of smoothness for these curves, the smoothest curve pair can be chosen from the set of all computed curve pairs. Its corresponding sample thickness indicates the value, which is the closest one to reality among the values in the considered range. In order to define a measure of smoothness we consider the set  $S$  of all evaluation points in the frequency domain and we let the integer  $m$  run through  $S$ . For a given sample thickness  $\ell$ , the total variation of degree one  $TV(\ell)$  is given by the equations:

$$D(m) = |n_1(m-1) - n_1(m)| + |\kappa_1(m-1) - \kappa_1(m)|, \quad (17)$$

$$TV(\ell) = \sum_{m \in S} D(m). \quad (18)$$

In other words, determining the sample thickness is equivalent to choosing  $\ell$  such that  $TV(\ell)$  is minimal. As mentioned above we use the Nelder-Mead algorithm to find the minimum of the error function  $Err$  as described in the previous section. This algorithm provides the material parameters, such that  $H_{theory}$  is arbitrarily close to  $H_{experiment}$ . Therefore, uncertainties of  $H_{theory}$  induced by uncertainties of  $\ell$  have no other possibility to be compensated, but by oscillations of the material parameter curves. For this reason, the total variation technique combined with the Nelder-Mead minimum search yield very accurate results for the sample thickness determination.

#### 4.4. Accuracy improvement with the Spatially Variant Moving Average Filter

##### 4.4.1. The principle

Using the algorithms described in Sections 4.1 to 4.3 for a single pair of reference and sample measurements, we are able to compute a transfer function  $H_{theory}(\omega)$  that lies as close as desired



to  $H_{\text{experiment}}(\omega)$ . However, the values of  $H_{\text{experiment}}(\omega)$  vary slightly for each measurement of the same sample. Hence, one has to accept a confidence interval for the transfer functions. The confidence interval can be modeled on the basis of a set of measurements.

We use this confidence interval for setting additional physical boundary conditions to the extracted material parameter curves. We suppose that a material parameter curve is accurate if it satisfies the additional physical boundaries and if the modeled transfer function  $H_{\text{theory}}(\omega)$  lies within the accepted confidence interval of  $H_{\text{experiment}}(\omega)$ .

#### 4.4.2. The additional physical boundary condition

Several factors, such as neglected scattering processes, the inhomogeneity of the geometry of the sample and additive noise, lead to a rippled shape of the material parameter curves. The additional physical boundary condition that we set to our model is that the material parameters should not vary strongly at consecutive frequency values, i.e., the curves should have a “smooth” shape rather than a rippled one. As we will see in the discussion of the method, this condition does not affect the detection of absorption peaks which investigated samples may exhibit.

#### 4.4.3. The confidence interval

In the following, we assume that several measurements of the reference and sample signals are available. For each considered frequency, we compute an average reference signal  $\overline{E_{\text{ref}}^{\text{ex}}}(\omega)$  having the magnitude and angle equal to the arithmetical means of the magnitudes and of the angles of all reference signals, respectively. Similarly, we obtain the average sample signal  $\overline{E_{\text{samp}}^{\text{ex}}}(\omega)$ . For simplicity, we introduce the following abbreviations:

- $H := H_{\text{experiment}}(\omega)$ , such that:

$$H = \frac{\overline{E_{\text{samp}}^{\text{ex}}}(\omega)}{\overline{E_{\text{ref}}^{\text{ex}}}(\omega)}. \quad (19)$$

- $a := \Re\{\overline{E_{\text{samp}}^{\text{ex}}}(\omega)\}$ ,  $b := \Im\{\overline{E_{\text{samp}}^{\text{ex}}}(\omega)\}$ , such that:

$$\overline{E_{\text{samp}}^{\text{ex}}}(\omega) = a + jb. \quad (20)$$

- $c := \Re\{\overline{E_{\text{ref}}^{\text{ex}}}(\omega)\}$ ,  $d := \Im\{\overline{E_{\text{ref}}^{\text{ex}}}(\omega)\}$ , such that:

$$\overline{E_{\text{ref}}^{\text{ex}}}(\omega) = c + jd. \quad (21)$$

The goal of this subsection is to determine the frequency dependent confidence intervals  $\Delta\Re\{H\}$  and  $\Delta\Im\{H\}$ , such that for each frequency, every pair of values  $\Re\{H_{\text{theory}}(\omega)\}$  and  $\Im\{H_{\text{theory}}(\omega)\}$  of the modeled transfer function is acceptable if it lies in the intervals  $[\Re\{H\} - \Delta\Re\{H\}, \Re\{H\} + \Delta\Re\{H\}]$  and  $[\Im\{H\} - \Delta\Im\{H\}, \Im\{H\} + \Delta\Im\{H\}]$ , respectively. Note, that an equivalent procedure can be derived using the magnitude and angle instead of the real and imaginary parts of the transfer function.

From the equalities (19) to (21) we derive the following frequency dependent equation:

$$H = \frac{a + jb}{c + jd} = \frac{ac + bd + j(bc - ad)}{c^2 + d^2}. \quad (22)$$

The quantities  $a, b, c$  and  $d$  are measured values. They are affected by additive noise. In order to characterize the detector noise of the experimental setup, we performed several measurements without THz radiation, i.e. by blocking the THz beam, which have shown that the noise

level is frequency independent. For this reason, we assume that the measured quantities  $a, b, c$  and  $d$  are affected by additive white noise. Furthermore, we observe variations among different measurements of the same sample materials, that are not correlated to the detector noise. The reason for these variations are fluctuations of the THz signal, as mentioned in Ref. [5]. Laser intensity instabilities or slight changes in the electromagnetic environment lead to such variations. We summarize these effects under the class of statistical errors and we assume that they are independent from the error due to the portion of additive white noise discussed above. In conclusion, we assume that the quantities  $a, b, c$  and  $d$  are affected by two independent error sources: *additive white noise* and a *statistical error*.

In the following, we derive expressions for the intervals  $\Delta x$ , with  $x \in \{a, b, c, d\}$ , considering both error sources. Subsequently, we use these intervals to derive  $\Delta\Re\{H\}$  and  $\Delta\Im\{H\}$ .

Let  $\Delta_N x$ , with  $x \in \{a, b, c, d\}$ , be the confidence interval of the variable  $x$  due to white noise. Note, that  $\Delta_N x$  is frequency independent because of the assumption of white noise. In order to compute  $\Delta_N x$ , we consider a noise array  $N$  that we obtain from the array  $x$  by isolating some last points, such that no signal is detectible in this array. For this purpose, a single pair of reference-sample measurements suffices. Let  $k$  be the length of the array  $N$ . First, we compute the mean of the noise array:

$$m_x = \frac{1}{k} \sum_{i=1}^k N_i. \quad (23)$$

We estimate the confidence interval of the variable  $x$  due to white noise as the standard deviation of the noise array:

$$\Delta_N x = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (N_i - m_x)^2}. \quad (24)$$

Optionally, one may define  $\Delta_N x$  in a different manner. For example, choosing the variance, i.e. the squared standard deviation, may lead to a narrower confidence interval. However, the standard deviation approach delivers a confidence interval  $\Delta_N x$  having the same physical measuring unit as  $x$ .

Let  $\Delta_S x$ , with  $x \in \{a, b, c, d\}$  be the confidence interval of the variable  $x$  due to the statistical error. Let  $u$  be the number of measurements available for the variable  $x$ . If only one reference or sample signal is available, no assertion can be made about the statistical error. In this case, we set  $\Delta_S x = 0$ . If two or more data sets are available, we denote these arrays by  $x^{(1)}, x^{(2)}, \dots, x^{(u)}$ . We compute the mean  $x^{(mean)}$  of all measurements:

$$x^{(mean)} = \frac{1}{u} \sum_{i=1}^u x^{(i)}. \quad (25)$$

We estimate the confidence interval of the variable  $x$  due to the statistical error as the standard deviation of the measurements:

$$\Delta_S x = \sqrt{\frac{1}{u-1} \sum_{i=1}^u (x^{(i)} - x^{(mean)})^2}. \quad (26)$$

Note that the expressions in (25) and (26) have to be computed for every frequency separately. Since the two error sources are statistically independent, we assume that the total confidence interval is equal to:

$$\Delta x = \sqrt{(\Delta_N x)^2 + (\Delta_S x)^2}. \quad (27)$$

As already discussed above,  $\Delta_N x$  is frequency independent in contrast to  $\Delta_S x$ , which is frequency dependent.

Now we can turn our attention to determining the confidence intervals  $\Delta\Re\{H\}$  and  $\Delta\Im\{H\}$ . From equation (22) we get:

$$\Re\{H\} = \frac{ac+bd}{c^2+d^2} =: f(a,b,c,d), \quad (28)$$

$$\Im\{H\} = \frac{bc-ad}{c^2+d^2} =: g(a,b,c,d). \quad (29)$$

Since  $\Delta\Re\{H\}$  and  $\Delta\Im\{H\}$  are functions of  $a, b, c$  and  $d$ , the Gaussian error propagation rule yields:

$$\Delta\Re\{H\} = \sqrt{\left(\Delta a \frac{\partial f}{\partial a}\right)^2 + \left(\Delta b \frac{\partial f}{\partial b}\right)^2 + \left(\Delta c \frac{\partial f}{\partial c}\right)^2 + \left(\Delta d \frac{\partial f}{\partial d}\right)^2}, \quad (30)$$

$$\Delta\Im\{H\} = \sqrt{\left(\Delta a \frac{\partial g}{\partial a}\right)^2 + \left(\Delta b \frac{\partial g}{\partial b}\right)^2 + \left(\Delta c \frac{\partial g}{\partial c}\right)^2 + \left(\Delta d \frac{\partial g}{\partial d}\right)^2}, \quad (31)$$

where  $f$  and  $g$  are the functions defined in equations (28) and (29), respectively. The confidence intervals  $\Delta x$  with  $x \in \{a, b, c, d\}$  are given by equation (27). For simplicity, we discard the covariance between the variables  $a$  and  $b$ , which in principle are correlated. According to Ref. [12], this is equivalent to neglecting the uncertainty of one of them. The same holds for the variables  $c$  and  $d$ . However, this approximation leads to a narrower confidence interval. For this reason, the physical boundary condition which we pose here is stronger than if considering the covariances. The partial derivatives of  $f$  and  $g$  can easily be calculated from equations (28) and (29):

$$\frac{\partial f}{\partial a} = \frac{c}{c^2+d^2}, \quad (32a)$$

$$\frac{\partial f}{\partial b} = \frac{d}{c^2+d^2}, \quad (32b)$$

$$\frac{\partial f}{\partial c} = \frac{ad^2 - ac^2 - 2bcd}{(c^2+d^2)^2}, \quad (32c)$$

$$\frac{\partial f}{\partial d} = \frac{bc^2 - bd^2 - 2acd}{(c^2+d^2)^2}, \quad (32d)$$

$$\frac{\partial g}{\partial a} = -\frac{d}{c^2+d^2}, \quad (32e)$$

$$\frac{\partial g}{\partial b} = \frac{c}{c^2+d^2}, \quad (32f)$$

$$\frac{\partial g}{\partial c} = -\frac{bc^2 - bd^2 - 2acd}{(c^2+d^2)^2}, \quad (32g)$$

$$\frac{\partial g}{\partial d} = \frac{ad^2 - ac^2 - 2bcd}{(c^2+d^2)^2}. \quad (32h)$$

#### 4.4.4. The algorithm

We introduce the Spatially Moving Average Filter (SVMAF) as an extension of the moving average filter, which is also known as adjacent averaging. For material parameter extraction, we have implemented an iterative algorithm using the SVMAF. It provides satisfactory results

usually after up to 5 iterations. After describing in this section the actions performed in each iteration step, we illustrate the mode of operation of the algorithm by means of a graphical model in the subsequent section.

- For every point of the material parameter curves, it replaces the initial value of this point with the mean of the initial values of itself and its two neighbors. This is the reason for the name *moving average*. The first and the last point of the curves keep their initial values.
- It computes the transfer function  $H_{theory}$  with the new values of the material parameters.
- For every point, it checks whether the real and imaginary parts of  $H_{theory}(\omega)$  lie inside the confidence intervals or not. If both parts lie in the respective ranges, the new values of the material parameters are accepted. If not, the material parameters at the respective frequency are reset to the initial values. Since the action of the filter depends on every point, we called it *spatially variant*.
- The new material parameter curves represent the input to the next iteration step.

#### 4.4.5. Visualization

Figure 2 depicts an example of one iteration step using the SVMAF. The upper half of the figure, i.e. the pictures 1 and 3, shows the progression of the real refractive index  $n$  while the lower half, i.e. the pictures 2 and 4, shows the real part of the transfer functions and its confidence interval. For simplicity, we only consider one of the material parameters as well as only the real part of the transfer function in this example. In the first pair of figures we see the initial state: the material parameter curve plotted with a continuous line connecting full dots, represents the initial values of  $n(\omega)$ . The transfer function plotted with a continuous line always represents  $\Re\{H_{experiment}(\omega)\}$ . However, at the beginning of the iteration, the function  $\Re\{H_{theory}(\omega)\}$  is very close to  $\Re\{H_{experiment}(\omega)\}$  since the Nelder-Mead search yields very accurate results. The dashed material parameter curve connecting empty dots, represents the result of the smoothing process in the first step of the algorithm described above and the dashed transfer function represents  $\Re\{H_{theory}(\omega)\}$  computed with the new material parameters. We see that the value of  $\Re\{H_{theory}(\omega)\}$  at the first point lies outside the confidence interval. Thus, the value of  $n$  at this point is reset to its initial value. All other values of  $\Re\{H_{theory}(\omega)\}$  lie inside the confidence interval. Therefore, for these points, the new values are being accepted. The curve consisting of the initial value of the first point and the new values of all other points is the continuous curve in the second plot of  $n$ . It represents the input to the next iteration step.

When computing the new averaged values, depicted dashed on the second  $n$ -plot, we observe that the value of the first point is closer to the initial value than its averaged value on the first plot. This is due to the updated values of its neighboring points. Also, we observe that the new averaged value causes the updated  $\Re\{H_{theory}(\omega)\}$  to entirely lie within the confidence interval. The three dots on the right hand side of the figure indicate that this procedure can be repeated for an arbitrary number of times.

#### 4.4.6. Properties of the SVMAF

Because of the repeated averaging process, the iteratively obtained theoretical material parameter curves are smoother than the initial theoretical curves. Within the experimental error modeled by the confidence intervals, they agree with the experimental curves. The smoothing of peaks in the material parameter curves that physically make sense, modifies the value of  $H_{theory}(\omega)$  at the respective point in an artificial way such that its value exceeds the allowed tolerances for  $\Re\{H_{theory}(\omega)\}$  and/or  $\Im\{H_{theory}(\omega)\}$ . Thus, the SVMAF does not alter such peaks, which is a striking advantage of the SVMAF over common noise cancellation techniques.

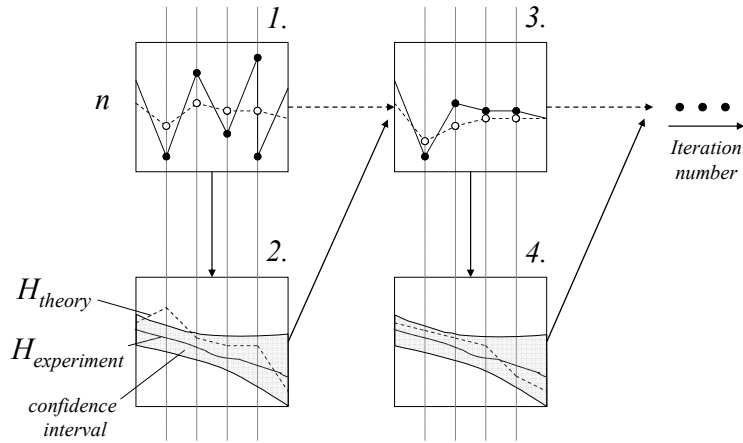


Fig. 2. Illustration of the mode of operation of the SVMAF

## 5. Experimental results

In this section we give some experimental results that exemplify the performance of the material parameter algorithm presented in this paper. For each parameter extraction, we use two reference and two sample measurements. Note that using more measurements leads to a better description of the statistical error and, thus, of the confidence interval. However, the confidence interval is also defined for a single reference and a single sample measurement. In this case, it describes only the additive white noise portion, since no assertion about the statistical error can be made. We use a micrometer screw to measure an initial values for the sample thickness determination algorithm described in Section 4.3. We assume an error of the micrometer screw of  $\pm 10\mu\text{m}$  for thick samples. For this purpose we evaluate the total variation of the material parameter curves for a thickness range around the measured value, in which the difference between two consecutive values is  $1\mu\text{m}$ .

### 5.1. High resistivity silicon wafer

In our first example, we extract the material parameters of a silicon wafer. The thickness measured with the micrometer screw is  $539\mu\text{m}$ . Figure 3(a) shows the temporal profiles of the time domain reference and sample signals. Figure 3(b) depicts the total variation over the sample thickness. The minimum of this curve is located at  $543\mu\text{m}$ , which is the value that we use for the material parameter extraction. The material parameters are depicted in Figures 3(c) and 3(d). The refractive index is in excellent agreement with the one published in Refs. [6] and [13]. Please note that the sample used in Ref. [6] was over 20mm thick, while our sample is nearly 40 times thinner. We observe that the SVMAF leads to much smoother curves than the curves obtained before the SVMAF was applied.

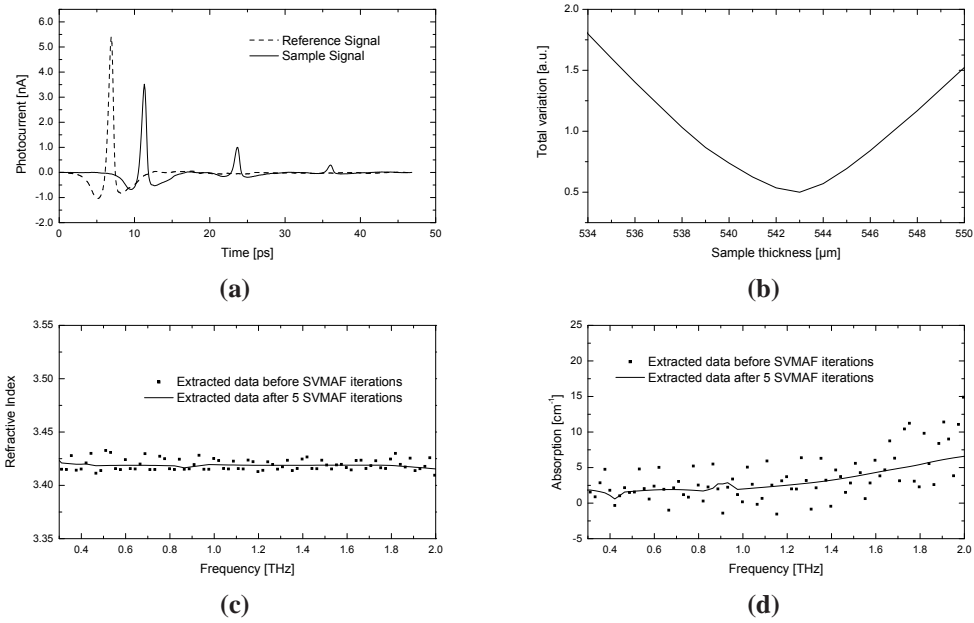


Fig. 3. Material parameter extraction of a 543 $\mu\text{m}$  Si wafer

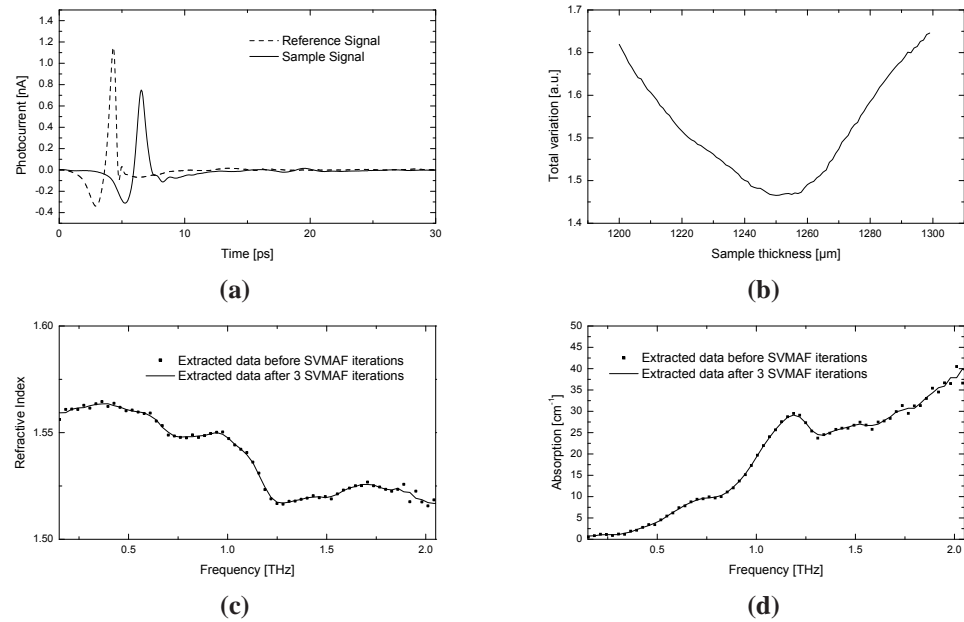


Fig. 4. Material parameter extraction of a 1252 $\mu\text{m}$  leu-gly-gly pellet

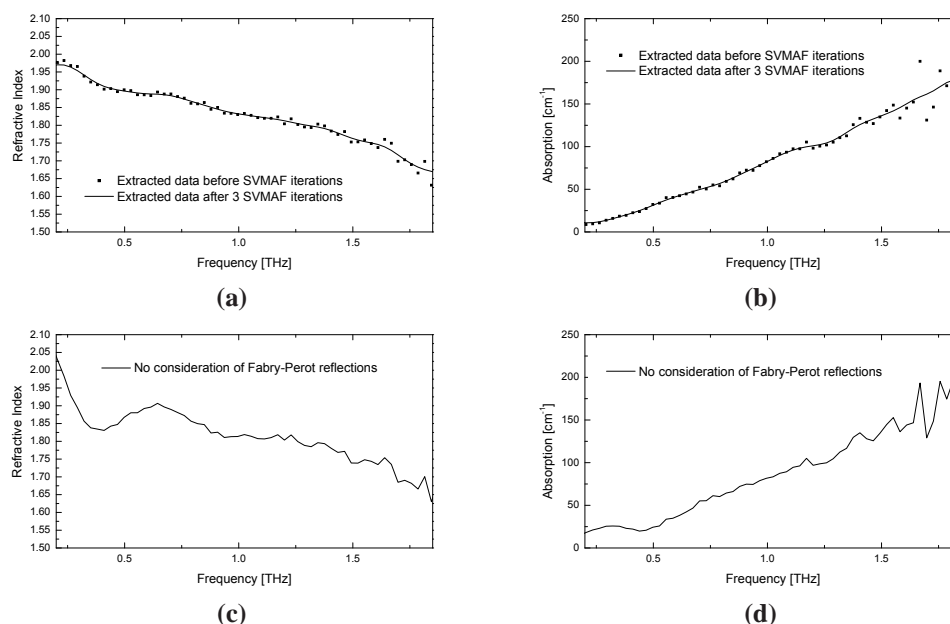


Fig. 5. Material parameter extraction of a 145 $\mu$ m artificial RNA pressing

### 5.2. Pellet consisting of 25mg leu-gly-gly tripeptide and 125mg polyethylene powder

Both material parameter curves of the Si wafer presented in the first example exhibit almost constant shapes. The smoother curves gained due to the SVMAF may lead to the false impression that the SVMAF algorithm only averages neighboring points. Our second example should exclude such an impression by examining a material that shows a clear frequency dependency of the material parameter curves. We investigate a pellet consisting of 25mg leu-gly-gly tripeptide and 125mg polyethylene powder. The material is pressed into a pellet under 380MPa for 10 minutes.

The time domain signals are depicted in Figure 4(a). Even though the Fabry-Perot effect is almost invisible to the eye in the time domain curves, the total variation exhibits a clear minimum at 1252 $\mu$ m, as shown in Figure 4(b). This is due to the very accurate Nelder-Mead algorithm used to determine the error between the modeled and the experimental transfer functions. This fact points out the applicability of our thickness extraction algorithm even in cases with a low Fabry-Perot effect. Figures 4(c) and 4(d) show the material parameters. We observe the strong frequency dependency of both material parameters. Note that the SVMAF algorithm does not affect the peaks of the curves that make sense physically. This is one of the major qualities of the SVMAF technique.

### 5.3. Thin artificial RNA pressing

In the following we investigate a pressing of 22mg of artificial RNA powder. Our goal in this section is to exemplify the performance of our parameter extraction algorithm for very thin samples. The sample under test lies far outside the ranges of experiments documented in several papers, compare to Ref. [3].

With a micrometer screw we measured a thickness of 150 $\mu$ m. However, this measurement may have been inaccurate due to the softness of the the sample. With the total variation tech-



nique we determine the actual sample thickness equal to  $145\mu\text{m}$ .

The computed material parameter curves are shown in Figures 5(a) and 5(b). Obviously, the extraction algorithm works very well for this thin sample with a relatively small real refractive index. Since the sample is very thin, the multiple Fabry-Perot reflections overlap in the time domain. Thus, time windowing the first pulse passing the probe is impossible. In Figures 5(c) and 5(d) we show the material parameters extracted with an algorithm that does not consider the overlapping Fabry-Perot reflections. The experimental data, from which the material parameters were extracted are the same as the data presented in Ref. [14]. We can clearly observe artifacts in the region around and below 0.5THz in the curves caused by the Fabry-Perot effect. Furthermore, Figures 5(c) and 5(d) show that towards higher frequencies, the ripple of both curves increases. This happens due to poor SNR at higher frequencies. Figures 5(a) and 5(b) demonstrate that the SVMFA counters this effect.

## 6. Conclusions

There are two main contributions of this work. First, we have improved the existing material parameter extraction techniques for THz TDS by merging the up-to-date knowledge into a single algorithm. Our method provides highly accurate curves of the complex refractive index of a wide range of samples. It works well for samples as thin as  $100\mu\text{m}$ , where multiple reflections overlap, samples with low indexes of refraction and samples with sharp peaks in the material parameter curves. The second main contribution is the introduction of the Spatially Variant Moving Average Filter which constitutes the core of an iterative algorithm yielding the physically most accurate modeled function out of a set of functions that have to satisfy certain boundaries. In our case, we use THz TDS to determine the material parameters of a homogenous sample by fitting a modeled transfer function to a measured one. Measured transfer functions vary for every measurement. We use these variations to define confidence intervals for the transfer functions which allow for the formulation of supplemental physical boundaries to the material parameter curves. Our iterative algorithm converges to a solution within the confidence intervals. We perceive this solution as the most accurate characterization of the optical parameters of the investigated material to date. However, the idea behind this algorithm may be useful for several other experimental techniques, where a modeled function is fitted to an experimentally determined one.

## Acknowledgments

We thank Radu Cernat from the Institute of High Voltage and EMC, University of Dortmund for the support and useful discussion. Rafal Wilk gratefully acknowledges the financial support from “Gottlieb Daimler and Karl Benz” foundation.