Uncertainty and Ambiguity in Terahertz Parameter Extraction and Data Analysis
— Source link

Matthias Krüger, Stefan Funkner, Erik Bründermann, Martina Havenith

Institutions: Ruhr University Bochum

Related papers:

- Material parameter estimation with terahertz time-domain spectroscopy

- Highly precise determination of optical constants and sample thickness in terahertz time-domain spectroscopy

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

- Uncertainty in terahertz time-domain spectroscopy measurement

- A reliable method for extraction of material parameters in terahertz time-domain spectroscopy

Uncertainty and Ambiguity in Terahertz Parameter Extraction and Data Analysis

Matthias Krüger, Stefan Funkner, Erik Bründermann, Martina Havenith*

Ruhr-Universität Bochum, Physikalische Chemie II,

NC7/72, D-44780 Bochum, Germany

Abstract

Phase sensitive measurement techniques, such as THz time-domain spectroscopy or dispersive Fourier transform spectroscopy, are very useful tools to obtain a complete set of optical material parameters. When recording the electric field as a function of time delay between THz and optical pulse, the absorption coefficient and the index of refraction can be extracted. However, the analysis shows ambiguity. Here, we describe an analysis which yields a complete set of mathematical solutions and show how the physically relevant can be deduced. We present a comprehensive mathematical survey for parameter extraction. We have recorded the THz spectra of anthracene and the fatty acid capric acid as examples for weakly absorbing solid samples, and an ionic liquid as an example for a strongly absorbing liquid sample. Finally, we discuss the uncertainty of the obtained optical parameters using error propagation of the Fourier transformation with a simple model and a rigorous mathematical procedure.

INTRODUCTION

Developed in the 1980s and 1990s [1, 2], THz time-domain spectroscopy (THz TDS) has been established as a general spectroscopic technique to study liquids [3], solids [4], gases [5], biomolecules [6, 7], aqueous salt solutions [8], nanoparticles [9], semiconductors [10] and plasmas [11]. The covered frequency bandwidth ranges from 3 cm^{-1} to 140 cm^{-1} (0.1 THz to 4 THz). THz TDS reaches higher frequencies in comparison to microwave dielectric spectroscopy [12, 13] and is restricted to lower frequencies compared to Fourier-transform infrared spectrometers (FTIR). When combining all available techniques the absorption coefficient can be measured over a bandwidth of up to 8 orders of magnitude in frequencies [14].

THz spectroscopy is able to probe intermolecular interactions and collective network motions on a picosecond time scale [15], thus provides complementary information to infrared, fluorescence and NMR spectroscopy. The latter probes localized chemical bonds and local optical and nuclear properties. In contrast, THz spectroscopy probes weak intermolecular forces, long range interactions and it is sensitive to fast solvent fluctuations and solvent-solute coupling [16, 17].

When recording the phase and amplitude of the electric field in THz TDS, it is possible to deduce the absorption coefficient α and the index of refraction n of the sample. However, the individual steps of data analysis are a challenging task requiring a careful extensive extraction procedure.

In fig. 1, the basic procedure for data analysis is shown. Standard data extraction procedures use an algorithm that searches for roots in the pulse propagation function. Starting at a certain set of initial guess values, the algorithms use a stepwise variation of the complex index of refraction until the measured data are reproduced (fig. 1 left). The sample thickness can either be treated as a constant parameter or can be included as a parameter in a global fit. In the latter case, the total variation is calculated which is a measure for the similarity of adjacent data points, and thus a measure for the smoothness of the spectrum. When minimizing the total variation the sample thickness can be deduced [18]. Other algorithms use multiple echoes of the THz pulse [10] or include the determination of the correct sample thickness in the time-domain [19] and in the frequency-domain [20]. For thin samples, the Fourier transformation of the extracted data into a quasi space enables the determination of

2

the thickness with the smallest oscillations of the extracted optical parameter [21]. Naftaly et al. introduces an algorithm, where the spectral data and oscillation due to Fabry-Pérot effects can be separated by deconvolution using Fourier transformation [22].

In this paper, we describe an alternative method based on Cauchy's principal argument which has the advantage that it is independent of initial guess values(fig. 1 middle). This takes into account the inherent ambiguity of the index of refraction calculated from the measured data.

A second issue addressed in this paper is the determination of the uncertainty of the obtained complex index of refraction of the sample (fig. 1 right). In general, there are many sources for noise and systematic errors in THz TDS, such as intensity fluctuations of the pump laser, the THz antenna, spatial and spot-size fluctuations due to mechanical delays, multiple reflections within plane parallel optical components, noise arising from the sample's inhomogeneity and scattering properties, or electronic background noise of the detector. The spectral noise can be separated into contributions from the emitter, detector and shot noise contributions [23]. Whereas an increase in the number of measurements decreases the statistical error, further repetitions lead to long measurement times and might be accompanied by long-term drifts and instabilities of the signal. On the other hand, with the advent of fast-scanning THz TDS such as vibrating membrane or asynchronous optical sampling spectrometers [24], the measurement speed can be largely increased. We present a procedure that determines the confidence intervals for the complex index of refraction of the sample starting with the uncertainty of the time-domain data. Using implicit function theorem, a rigourous error propagation as well as a simplified model are presented.

This paper focuses on the analysis of TDS spectra in transmission geometry [3], although the conclusions are transferable to other geometries such as reflection [25], ellipsometry [26], interferometry [27] or differential TDS [28].

THEORETICAL BACKGROUND

Basic definitions

In the following, we simplify the notation by writing a tilde over complex variables and define $\tilde{\epsilon} := \tilde{\epsilon}(\nu)$, $\tilde{n} := \tilde{n}(\nu)$, $\alpha := \alpha(\nu)$, keeping in mind that these parameters frequency-

3

dependent. Average or mean values are indicated by a bar, for example, $\bar{n}$. Parameters that are time-dependent are marked by a circumflex, e.g. $\hat{E} := E(t)$. Furthermore, all parameters that refer to the sample and reference measurements will have the subscripts s and r. An asterisk * denotes the complex conjugate of a complex parameter.

THz TDS records the electric field $\hat{E}$ as a function of time. The procedure of data analysis starts with a discrete Fourier transformation of the time-dependent discrete signal $\hat{E}_m$ into a discrete, frequency-dependent Fourier transform $\tilde{E}_k$ [29]

$$\tilde{E}_k = \frac{1}{\sqrt{N}} \left(\sum_{m=0}^{N-1} \hat{E}_m \cos\left(2\pi k \frac{m}{N}\right) - \imath \sum_{m=0}^{N-1} \hat{E}_m \sin\left(2\pi k \frac{m}{N}\right) \right) := \mathrm{FFT}_k\left(\hat{E}\right) \tag{1}$$

with the total number of samples N.

Using THz TDS data analysis it is possible to obtain the linear response from the sample:

$$\tilde{n} = n - \imath\kappa \tag{2}$$

with the real index of refraction n and the imaginary part κ, describing amplitude damping of a wave that is passing through the sample. We introduce the wavevector in vacuum $k_0 = \omega/c = 2\pi\nu/c$. κ can be converted into the absorption coefficient α by

$$\alpha = 2\frac{2\pi\nu}{c}\kappa = 2k_0\kappa \tag{3}$$

Due to interaction between the charge distribution in the medium and the electromagnetic field of the beam, the phase velocity is reduced by propagating through an optical medium. This introduces a phase retardation, expressed in terms of the real part of $\tilde{n}$. n describes the ratio between the phase velocity of light in the sample and the speed of light in vacuum.

The complex index of refraction $\tilde{n}$ can easily be converted into the frequency-dependent complex dielectric function $\tilde{\varepsilon}$

$$\tilde{\varepsilon} = \tilde{n}^2 = \varepsilon' - \imath\varepsilon'' = n^2 - \kappa^2 - \imath 2n\kappa \qquad . \tag{4}$$

The real part of the complex dielectric function ε' describes a phase retardation between the electric field and the dielectric mode in an alternating electric field. The imaginary part ε'' describes the damping of the wave. $\tilde{\varepsilon}$ can be fit to different multi-modal relaxation models, which is a well-known method [3, 30].

Data analysis

An initial analysis of the TDS data can be done by using Lambert-Beer's law to obtain an approximation of the absorption coefficient α_s and the absorption κ_s of a sample with a thickness d

$$\kappa_s = \frac{\alpha_s}{2k_0} = -\frac{1}{2dk_0} \ln(\tilde{\delta}^* \tilde{\delta}) \tag{5}$$

defining the complex ratio of the sample and reference electric fields $\tilde{\delta}$ as

$$\tilde{\delta} = \tilde{E}_s / \tilde{E}_r \qquad . \tag{6}$$

A frequency-independent initial guess $\bar{n}_s$ for n_s, the real part of $\tilde{n}_s$, can be deduced from the measurement of the relative time shift Δt of the maximum of the electric field of the time-domain pulse, which is probing the sample in comparison to the reference. We can use the following estimation [31]:

$$\bar{n}_s \approx \bar{n}_r + \frac{c\Delta t}{d} \tag{7}$$

with d, $\bar{n}_s$, and $\bar{n}_r$ being the sample thickness and the average indices of refraction of the sample and the reference, respectively.

The following phase relation links the phase change of the sample pulse relative to the reference pulse, $\phi_s - \phi_r \equiv \arg(\tilde{\delta}) \mod 2\pi$, to n_s [32]:

$$n_s \equiv n_r - \frac{1}{k_0 d} \arg(\tilde{\delta}) \mod 2\pi \tag{8}$$

An ambiguity arises, because the phase can cycle through many periods of 2π, but the measured phase is restricted to the interval to the interval between $-\pi$ and π. Thus, based upon a single measurement, we are not able to distinguish phase shifts, if they are separated by $n \cdot 2\pi$ with n being an integer. In most cases, the assumption is made that only the first pulse will determine the phase, whereas contributions from further reflections of the pulse, that influence the phase as well, are neglected. However, significant problems arise, in case of high dispersion or large sample thickness. This can be improved when multiple sample thickness d have been probed. In this case, n_s is fitted to the phase $\phi(d)$ and α_s is fitted to $\ln(\tilde{\delta}^* \tilde{\delta})$ [3].

In both approximations (5) and (8), α_s and n_s can be obtained by using either amplitude or phase information. In reality, we have to consider different transmission coefficients at

each optical interface and thus different reflection losses for sample and reference measurements. The transmission coefficients are calculated using the Fresnel equations (see next chapter).

As an example we can choose a solid sample with air ($n_r = 1$) as reference. In this case a correction term can be inserted into Eq. 5and we can use the following approximation for sufficiently small α_s:

$$\alpha_s = -\frac{1}{d}\ln\left(\tilde{\delta}^*\tilde{\delta}\left(\frac{(n_s+1)^2}{4n_s}\right)^2\right)$$ (9)

For samples with high absorption, the complex index of refraction $\tilde{n}_s$ must be used instead of n_s to calculate the Fresnel coefficients.

An alternative method to deduce n_s and κ_s, based upon an initial guess for either n_s or κ_s, is the Kramers-Kronig analysis [33]. It links real and imaginary parts of the complex index of refraction. In the following, we will consider the more common case where discrete values for the initial guess of the absorption κ_s are obtained from the experiments. Then, the integral in the exact equation is replaced by an approximated sum:

$$n_s(\nu_j) \approx n_{\infty,s} + \frac{2}{\pi}\sum_{\substack{p=1\\p\neq j}}^{N}\frac{\kappa_s(\nu_p)\nu_p}{\nu_p{}^2 - \nu_j{}^2}\Delta\nu$$ (10)

with $n_{\infty,s}$ being the index of refraction at infinite frequency and $\Delta\nu$ the discrete sampling rate. $\kappa_s(\nu_p)$ is calculated from $\alpha_s(\nu_p)$ within the measured frequency range. The real part $n_s(\nu_j)$ at the frequency ν_j is determined according to Eq. 10.

Using Eq. 9, α_s can be re-calculated. Now an iterative procedure is used, until the changes in n_s and κ_s are sufficiently small. However, in order to apply the Kramers-Kronig relation, we need information over the entire frequency range. Practically, this is not possible due to the limited bandwidth and the discrete nature of the measurement. The Kramers-Kronig relation is known to be a good approximation as long as α_s is small within the frequency range which has not been investigated.

Pulse propagation

Previously, THz data analysis was based on pulse propagation models [18, 19, 34, 35], i.e. the THz pulse is monitored during propagation. At every interface between two optical media 1 and 2, with an index of refraction n_1 and n_2 respectively, the THz pulse is split

into two pulses, a transmitted pulse and a reflected secondary pulse. At the next interface, the secondary pulse is split into further pulses and so forth. Given an incident angle of 90°, we obtain by the Fresnel equations $r = (n_1 + n_2)/(n_1 - n_2)$ and $t = 2n_2/(n_1 + n_2)$ for the reflection and the transmission coefficients. Within each optical medium, the THz pulse is retarded due to refraction and damped due to absorption. This can be summarized when using the following propagation coefficient:

$$\tilde{P}_{s,r} = \exp\left(\imath k_0 \tilde{n}_{s,r} d\right) \tag{11}$$

The different transmission, reflection and propagation terms of all pulses reaching the detector have to be multiplied yielding the pulse propagation function, a complex, frequency-dependent function that describes the change of the THz pulse by the sample geometry with respect to the reference geometry. Depending on the geometry and number of optical interfaces, many terms have to be considered. However, pulses that are reflected at the first optical interface, pulses that take more time to reach the detector than the maximum time delay of the measurement and pulses that are, due to multiple absorptions, damped such that their amplitude is below the noise level don't have to be considered. Contributions from other pulses have a strong influence on the spectrum. For example, multiple reflections between plane-parallel surfaces (Fabry-Pérot effect), for example the sample, windows of the measurement cell or the detection crystal, generate periodic oscillations in the frequency domain spectrum. To reduce this effect, additional terms, such as series expansions describing multiple reflections, are added to the pulse propagation function. The accuracy of the parameter extraction can be further improved by more complex algorithms, for example the spatially variant moving average filter [36].

Visualization of the pulse propagation function

For a solid sample we obtain the following equation which includes a Fabry-Pérot effect[34]:

$$\tilde{\delta} = \frac{4\tilde{n}_s\tilde{n}_r}{(\tilde{n}_s + \tilde{n}_r)^2}\frac{\tilde{P}_s}{\tilde{P}_r}\left(1 - \left(\frac{\tilde{n}_s - \tilde{n}_r}{\tilde{n}_s + \tilde{n}_r}\right)^2\tilde{P}_s^2\right)^{-1} \tag{12}$$

By rearranging Eq. 12, the following function $F(\tilde{n}_s)$ can be defined:

$$F(\tilde{n}_s) = \frac{4\tilde{n}_s\tilde{n}_r\frac{\tilde{P}_s}{\tilde{P}_r} - \tilde{\delta}(\tilde{n}_s + \tilde{n}_r)^2 + \tilde{\delta}(\tilde{n}_s - \tilde{n}_r)^2\tilde{P}_s^2}{(\tilde{n}_s + \tilde{n}_r)^2 - (\tilde{n}_s - \tilde{n}_r)^2\tilde{P}_s^2} \tag{13}$$

7

Every $\tilde{n}_s$ that solves $F(\tilde{n}_s) = 0$ is a mathematical solution for the complex of index refraction of the sample. F is a frequency-dependent complex function, which is visualized in Fig. 2 using the Hue Saturation and Value (HSV) color scheme. Each complex data point in the considered interval is assigned a specific color and a brightness value. Smaller amplitude corresponds to darker pixel and vice versa, whereas the phase determines the color of the plot point. The roots of the function can be found in the dark areas, whereas the light pixels show singularities of F, i.e. the roots of the denominator.

Search method

For $F(\tilde{n}_s) = 0$, it is necessary that the following equation is valid:

$$F_0(\tilde{n}_s) := 4\tilde{n}_s\tilde{n}_r - \tilde{\delta}(\tilde{n}_s + \tilde{n}_r)^2 \frac{\tilde{P}_r}{\tilde{P}_s} + \tilde{\delta}(\tilde{n}_s - \tilde{n}_r)^2 \tilde{P}_s\tilde{P}_r = 0 \qquad . \tag{14}$$

In order to obtain all roots of $F_0(\tilde{n}_s)$ in a given two-dimensional rectangular interval G_1 with boundary ∂G_1, a special procedure performing a global analysis of all minima is used(Fig. 3). The initial interval for a search should be chosen such that it contains the actual complex index of refraction of the sample. Under the assumption that there is no root directly on ∂G_1 and as singularities have been eliminated beforehand, the total number of roots within G_1 can be calculated by choosing a closed, positively oriented Jordan path γ_1 on the boundary ∂G_1. The Cauchy argument principle can be used to calculate the total number of roots N for $F_0(\tilde{n}_s)$ inside G_1 by

$$N = \frac{1}{2\pi\imath} \int_{\gamma_1} \frac{F_0'(\tilde{n}_s)}{F_0(\tilde{n}_s)} d\tilde{n}_s \tag{15}$$

Eq. (15) can be solved numerically. When the total number of roots in G_1 is calculated, it can be separated into smaller subsets G_2 to G_N, where the procedure is repeated (Fig. 4). All subsets with $N = 0$ can be ignored, whereas all other subsets are re-divided into smaller subsets and so on. By this, all roots in the given interval can successively be identified and calculated with high precision. The procedure can be stopped as soon as the size of a subset is sufficiently small. If the number of roots for this subset is greater than 1, the uncertainty of the measurement is too high to distinguish between two minima.

Error propagation

In order to estimate the statistical experimental uncertainties of the results, typically several measurements are averaged in the frequency domain [37]. Using subsequent Gaussian error propagation of implicit functions, we present here an alternative method which uses directly the recorded time-domain data.

Confidence intervals for $\tilde{\delta}$

We will use the following notation: $\tilde{z} = z_R + \imath z_I$ with z_R and z_I being the real and imaginary part. The partial derivatives of a function $y = f(x_1, x_2, \ldots)$ depending on measured values $x_1, x_2, \ldots$ are denoted as $y_{;x_n} := \partial y / \partial x_n$.

For the pth single measurement, it is possible to approximate the function using a Taylor expansion. Under the assumption that the derivatives are small, averaging commutates with the mapping f, i.e. $f(\bar{x}_1, \bar{x}_2, \ldots) = \bar{y}$. Using the abbreviation $\Delta x_i = x_{i_p} - \bar{x}_i$, we obtain[38]:

$$y_{1_p} - \bar{y} = y_{;x_1} \Delta x_1 + y_{;x_2} \Delta x_2 + \ldots \tag{16}$$

The variance of y is

$$\sigma_y^2 = y_{;x_1} \sigma_{x_1}^2 y_{;x_2} \sigma_{x_2}^2 + \ldots + 2 y_{;x_1} y_{;x_2} \sigma_{x_1 x_2}^2 + \ldots \tag{17}$$

$\sigma_{x_1 x_2}^2$ is the covariance of x_1 and x_2. For a Fourier transformation with frequency index k, the variances of the real and imaginary parts of the amplitude E_k, σ_{kR}^2 and σ_{kI}^2, and their covariance σ_{kRI}^2 are obtained:

$$\sigma_{kR}^2 = \frac{1}{N} \sum_{m=0}^{N-1} \left(\hat{\sigma}_m \cos\left(2\pi k \frac{m}{N}\right) \right)^2 = \frac{1}{2N} \sum_{m=0}^{N-1} \hat{\sigma}_m^2 + \frac{1}{2\sqrt{N}} \Re\left(\mathrm{FFT}_{2k}\left(\hat{\sigma}_m^2\right) \right) \tag{18}$$

$$\sigma_{kI}^2 = \frac{1}{N} \sum_{m=0}^{N-1} \left(\hat{\sigma}_m \sin\left(2\pi k \frac{m}{N}\right) \right)^2 = \frac{1}{2N} \sum_{m=0}^{N-1} \hat{\sigma}_m^2 - \frac{1}{2\sqrt{N}} \Re\left(\mathrm{FFT}_{2k}\left(\hat{\sigma}_m^2\right) \right) \tag{19}$$

$$\sigma_{kRI}^2 = -\frac{1}{2N} \sum_{m=0}^{N-1} \left(\hat{\sigma}_m^2 \sin\left(4\pi k \frac{m}{N}\right) \right) = \frac{1}{2\sqrt{N}} \Im\left(\mathrm{FFT}_{2k}\left(\hat{\sigma}_m^2\right) \right) \tag{20}$$

The last equality in Eqs. (18) to (20) can be processed using the Fast Fourier Algorithm (1) at the second harmonic of k, thus increasing the speed of computation by several orders of

magnitude. For a single frequency we can drop the index k and by using the sample and reference measurements, we obtain the real and imaginary part of $\tilde{\delta}$ according to:

$$\tilde{\delta} = \frac{E_{sR} + \imath E_{sI}}{E_{rR} + \imath E_{rI}} \tag{21}$$

Defining $\vec{\boldsymbol{\Delta}} = (\delta_R, \delta_I)$ and $\vec{\mathbf{E}} = (E_{s_R}, E_{s_I}, E_{r_R}, E_{r_I})$, we calculate the uncertainties σ_{δ_R}, σ_{δ_I} and the covariance $\sigma^2_{\delta_R \delta_J}$ of $\tilde{\delta}$ by

$$\mathbf{COV}(\vec{\boldsymbol{\Delta}}) = \mathbf{J}(\tilde{\delta})\mathbf{COV}(\vec{\mathbf{E}})\mathbf{J}(\tilde{\delta})^{\mathrm{T}} \tag{22}$$

using a higher-dimensional analogon to Eq. 17 with $\mathbf{J}$ being the Jacobian matrix. The covariance matrix $\mathbf{COV}(\vec{\mathbf{X}})$ of a vector $\vec{\mathbf{X}} = (x_1, x_2, \ldots, x_s)$ is defined as

$$\mathbf{COV}(\vec{\mathbf{X}}) = \begin{pmatrix} \sigma^2_{x_1} & \sigma^2_{x_1 x_2} & \cdots & \sigma^2_{x_1 x_s} \\ \sigma^2_{x_2 x_1} & \sigma^2_{x_2} & \cdots & \sigma^2_{x_2 x_s} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{x_s} & \sigma^2_{x_s x_2} & \cdots & \sigma^2_{x_s} \end{pmatrix} \tag{23}$$

$\mathbf{COV}(\vec{\mathbf{E}})$ can be obtained from Eq. 18 to Eq. 20.

Confidence intervals for $\tilde{n}_s$ and $\tilde{n}_r$

In order to deduce the confidence intervals $\sigma^2_{n_s R}$ and $\sigma^2_{n_s I}$ from $\mathbf{COV}(\vec{\boldsymbol{\Delta}})$, the partial derivatives of n_{sR} and n_{sI} with respect to δ_R and δ_I have to be computed. We define:

$$\begin{aligned} h(n_{sR}, n_{sI}, \delta_R, \delta_I) =\ & 4\tilde{n}_s \tilde{n}_r \\ & -(\delta_R + \imath \delta_I)(\tilde{n}_s + \tilde{n}_r)^2 \exp(-\imath k_0 d(\tilde{n}_s - \tilde{n}_r)) \\ & +(\delta_R + \imath \delta_I)(\tilde{n}_s - \tilde{n}_r)^2 \exp(+\imath k_0 d(\tilde{n}_s + \tilde{n}_r)) \end{aligned} \tag{24}$$

with:

$$H : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$((n_{aR}, n_{aI}), (\delta_R, \delta_I)) \mapsto H((n_{aR}, n_{aI}), (\delta_R, \delta_I)) = (h_R, h_I) \tag{25}$$

Let $\bar{n}_s$ be a solution to Eq. (25) with the parameter $\bar{\delta}$, then $H(\bar{n}_s, \bar{\delta}) = 0$. According to the implicit function theorem:

$$\begin{pmatrix} n_{sR;\delta_R}(\bar{\delta}) & n_{sR;\delta_I}(\bar{\delta}) \\ n_{sI;\delta_R}(\bar{\delta}) & n_{sI;\delta_I}(\bar{\delta}) \end{pmatrix} = \begin{pmatrix} H_{R;n_s R}(\bar{n}_s, \bar{\delta}) & H_{R;n_s I}(\bar{n}_s, \bar{\delta}) \\ H_{I;n_s R}(\bar{n}_s, \bar{\delta}) & H_{I;n_s I}(\bar{n}_s, \bar{\delta}) \end{pmatrix}^{-1} \begin{pmatrix} H_{R;\delta_R}(\bar{n}_s, \bar{\delta}) & H_{R;\delta_I}(\bar{n}_s, \bar{\delta}) \\ H_{I;\delta_R}(\bar{n}_s, \bar{\delta}) & H_{I;\delta_I}(\bar{n}_s, \bar{\delta}) \end{pmatrix} \tag{26}$$

10

with the constraint

$$\det \begin{pmatrix} H_{R;n_sR}(\bar{n}_s, \bar{\delta}) & H_{R;n_sI}(\bar{n}_s, \bar{\delta}) \\ H_{I;n_sR}(\bar{n}_s, \bar{\delta}) & H_{I;n_sI}(\bar{n}_s, \bar{\delta}) \end{pmatrix} \neq 0 \quad . \tag{27}$$

Finally, we obtain for the variances with a cross-correlation term

$$\sigma^2_{n_sR} = n^2_{sR;\delta_R}\sigma^2_{\delta_R} + n^2_{sR;\delta_I}\sigma^2_{\delta_I} + 2n_{sR;\delta_R}n_{sR;\delta_I}\sigma^2_{\delta_R\delta_I} \tag{28}$$

$$\sigma^2_{n_sI} = n^2_{sI;\delta_R}\sigma^2_{\delta_R} + n^2_{sI;\delta_I}\sigma^2_{\delta_I} + 2n_{sI;\delta_R}n_{sI;\delta_I}\sigma^2_{\delta_R\delta_I} \tag{29}$$

Simplified error propagation model

In a simplified model, the uncertainty of n_s and κ_s can be calculated by Gaussian error propagation. The errors for the absolute $\sigma_{|\tilde{E}_u|}$ and the phase σ_ϕ of the Fourier transformation for the reference and sample measurements can be calculated as [38]:

$$\sigma^2_{|\tilde{E}_u|} = \frac{1}{|\tilde{E}_u|^2}\left(E^2_{uR}\sigma^2_{uR} + E^2_{uI}\sigma^2_{uI} + 2E_{uR}E_{uI}\sigma^2_{uRI}\right) \qquad u = r, s \tag{30}$$

and for the phase

$$\sigma^2_{\phi_u} = \frac{1}{|\tilde{E}_u|^4}\left(E^2_{uI}\sigma^2_{uR} + E^2_{uR}\sigma^2_{uI} - 2E_{uR}E_{uI}\sigma^2_{uRI}\right) \qquad u = r, s \tag{31}$$

The error bars for the phase $\arg(\tilde{\delta})$ and the absolute $|\tilde{\delta}|$ can be approximated in the following way:

$$\Delta\arg(\tilde{\delta}) = \sqrt{\sigma^2_{\phi_r} + \sigma^2_{\phi_s}} \tag{32}$$

$$\Delta|\tilde{\delta}| = |\tilde{\delta}|\sqrt{\left(\frac{\sigma_{|\tilde{E}_r|}}{|\tilde{E}_r|}\right)^2 + \left(\frac{\sigma_{|\tilde{E}_s|}}{|\tilde{E}_s|}\right)^2} \tag{33}$$

The error bars for the real part Δn_s and the imaginary part $\Delta\kappa_s$ of $\tilde{n}_s$ are obtained by error propagation of Eq. (32) and (33). For thin, as for highly absorbing samples, the uncertainty of the sample thickness must be taken into account. A typical sample thickness for a highly absorbing material is in the range of 25 μm or even thinner, which requires a very accurate thickness measurement to avoid large error bars. We derive

$$\Delta n_s = \frac{\arg(\tilde{\delta})}{k_0 d}\sqrt{\left(\frac{\Delta\arg(\tilde{\delta})}{\arg(\tilde{\delta})}\right)^2 + \left(\frac{\Delta d}{d}\right)^2} \tag{34}$$

$$\Delta\kappa_s = \frac{1}{2k_0}\Delta\alpha = \frac{1}{k_0 d}\sqrt{\left(\frac{\Delta|\tilde{\delta}|}{|\tilde{\delta}|}\right)^2 + \left(\frac{\Delta d}{d}\ln|\tilde{\delta}|\right)^2} \quad (35)$$

Analogous error bars can be derived for the real and imaginary parts of the complex dielectric function:

$$\Delta\varepsilon'_s = 2\sqrt{(\kappa_s\Delta\kappa_s)^2 + (n_s\Delta n_s)^2} \qquad \Delta\varepsilon''_s = 2\sqrt{(n_s\Delta\kappa_s)^2 + (\kappa_s\Delta n_s)^2} \quad (36)$$

Using Fourier Transform Infrared (FTIR) spectroscopy, the sample thickness can be derived when recording the maxima in transmission in the frequency domain:

$$d = \frac{c}{2\bar{n}\nu_e} \quad (37)$$

with $\bar{n}$ being the average index of refraction of the material. Using a commercial FTIR we obtain an accuracy of $\Delta d \leq 1$ μm. Equation (37) is valid as long as n_s is not frequency-dependent, i.e. in case of low sample dispersion. For a precise measurement the choice of sample thickness is extremely important. A thin sample bears the disadvantage that the relative error in the sample thickness is relatively large. A large sample restricts the dynamic range of the spectrometer due to high absorption. Especially for highly absorbing materials such as aqueous solutions or ionic liquids, the absorption is found to increase with frequency, whereas the intensity of the THz source typically decreases at frequencies beyond 1.5 THz. As a consequence, the dynamic range critically drops with increasing sample thickness.

APPLICATION TO EXPERIMENTAL DATA

As an example for the approximation of n_s, we have measured a series of octadecane samples with thicknesses $d = 4.85$ mm, 5.7 mm, 7.1 mm and 9.75 mm, respectively. Figure 4 shows the result of the sample measurements with a scaled x-axis relative to the reference measurement with the optical delay $d = c\Delta t/(\bar{n}_s - \bar{n}_r)$ according to Eq. (7). As expected, the relative optical delay is directly related to the sample thickness d.

Kramers-Kronig relation

One example for which the numeric Kramers-Kronig relation has been applied is displayed in Fig. 5. The index of refraction and absorption coefficient for anthracene were measured

with an asynchronous optical sampling (ASOPS) THz time-domain spectrometer [24]. The sample was pressed into a pill and the empty pill holder was used for the reference measurement. The data were analyzed using a parameter extraction algorithm (cf. 12) retrieving also the index of refraction. For comparison, the iterative numeric Kramers-Kronig analysis according to Eqs. (9) and (10) is shown, which converged within three iterations (Fig. 5).

Both measured and approximated values for the index of refraction agree for frequencies in close proximity to the absorption peak and even for frequencies down to 0.15 THz and below. At high frequencies, i.e. above 2 THz, the measured and calculated curves deviate. This can be attributed to an absorption within the sample which contributions considerable even beyond the measurement bandwidth of 2.5 THz. Both the absorption and dispersion spectra agree qualitatively very well with the spectra given in the literature [39].

Multiple solutions

Figure 6 shows all possible mathematical of Eq. 14 as deduced from a single transmission measurement of a 9.2 mm thick capric acid sample. Solutions with a strongly negative value of the absorption k have been omitted. We find that for a large ratio of sample thickness and wavelength the different solutions are in close proximity (Fig. 6). In general, we expect distinct solutions for different sample thicknesses except for the physically true complex index of refraction. Thus, when using several sample thicknesses, physically irrelevant solutions can be eliminated. Figure 8 shows the resulting dispersion spectrum of capric acid. The error bars were calculated using Eq. 28 and Eq. 29.

If the physically true solution is known for low frequencies, it can be extrapolated to higher frequencies. This can be extrapolated to high-frequency phase-sensitive spectroscopy like dispersive Fourier transform spectroscopy (DFTS).

Many modern microwave spectrometers have a frequency overlap with THz TDS at the low-frequency limit. At the high-frequency limit, FTIR spectrometers with special window materials, beam-splitters and liquid helium cooled detectors can reach down to less than 1 THz. However, only the imaginary part, the absorption coefficient, can be compared, as a FTIR spectrometer does not yield phase information unless it is dispersive (DFTS). In Fig. 7 we show a comparison between the absorption coefficient of the ionic liquid ethylammonium nitrate (EAN) as measured with a THz time-domain, microwave dielectric and far-infrared

FTIR spectrometer [14]. Different solutions for the absorption coefficients for a theoretical sample thickness of 300 μm (blue) and 1 mm (red) are plotted as well as the corresponding indices of refraction. A sample thickness of 300 μm can easily be analyzed, whereas the 1 mm sample shows an ambiguity of the data which can only be solved by extrapolation from the microwave and FTIR frequencies. The physically true solution is shown in black for comparison.

SUMMARY

We presented a method for careful data analysis of THz time-domain data. We show that inherent ambiguities in the data analysis can be eliminated by a combination of different experimental methods or a systematic variation of the investigated sample thickness. This will help to reduce the inherent errors. We also describe how physical meaningful solutions can be separated from mathematical solutions of the problem. The data analysis is applied to THz time-domain spectra of anthracene, ionic liquids and fatty acids.

REFERENCES

––––––––––

∗ Email: `martina.havenith@rub.de,erik.bruendermann@rub.de`

[1] C. Fattinger and D. Grischkowsky, Applied Physics Letters **53**, 1480 (1988).

[2] D. H. Auston, Applied Physics Letters **26**, 101 (1975).

[3] J. T. Kindt and C. A. Schmuttenmaer, The Journal of Physical Chemistry **100**, 10373 (1996), ISSN 0022-3654.

[4] M. Walther, B. Fischer, M. Schall, H. Helm, and P. U. Jepsen, Chemical Physics Letters **332**, 389 (2000), ISSN 0009-2614.

[5] D. Mittleman, R. Jacobsen, R. Neelamani, R. Baraniuk, and M. Nuss, Applied Physics B: Lasers and Optics **67**, 379 (1998).

[6] A. Markelz, S. Whitmire, J. Hillebrecht, and R. Birge, Physics in Medicine and Biology **47**, 3797 (2002).

[7] C. Zhang and S. M. Durbin, The Journal of Physical Chemistry B **110**, 23607 (2006), ISSN 1520-6106.

[8] D. A. Schmidt, O. Birer, S. Funkner, B. P. Born, R. Gnanasekaran, G. W. Schwaab, D. M. Leitner, and M. Havenith, Journal of the American Chemical Society **131**, 18512 (2009), ISSN 0002-7863.

[9] J. Han, W. Zhang, W. Chen, L. Thamizhmani, A. K. Azad, and Z. Zhu, The Journal of Physical Chemistry B **110**, 1989 (2006), ISSN 1520-6106.

[10] J. Dai, J. Zhang, W. Zhang, and D. Grischkowsky, J. Opt. Soc. Am. B **21**, 1379 (2004).

[11] S. Ebbinghaus, K. Schröck, J. C. Schauer, E. Bründermann, M. Heyden, G. Schwaab, M. Böke, J. Winter, M. Tani, and M. Havenith, Plasma Sources Science and Technology **15**, 72 (2006), ISSN 0963-0252.

[12] R. Buchner and G. Hefter, Phys. Chem. Chem. Phys. **11**, 8984 (2009).

[13] H. Weingärtner, A. Knocks, W. Schrader, and U. Kaatze, The Journal of Physical Chemistry A **105**, 8646 (2001), ISSN 1089-5639.

[14] M. Krüger, E. Bründermann, S. Funkner, H. Weingärtner, and M. Havenith, Journal of Chemical Physics (2010).

[15] C. A. Schmuttenmaer, Chemical Reviews **104**, 1759 (2004), ISSN 0009-2665.

[16] D. M. Leitner, M. Gruebele, and M. Havenith, HFSP J. **2**, 314 (2008).

[17] A. G. Markelz, J. R. Knab, J. Y. Chen, and Y. He, Chemical Physics Letters **442**, 413 (2007), ISSN 0009-2614.

[18] T. D. Dorney, R. G. Baraniuk, and D. M. Mittleman, J. Opt. Soc. Am. A **18**, 1562 (2001).

[19] M. Scheller and M. Koch, Journal of Infrared, Millimeter and Terahertz Waves **30**, 762 (2009).

[20] L. Duvillaret, F. Garet, and J.-L. Coutaz, Appl. Opt. **38**, 409 (1999).

[21] M. Scheller, C. Jansen, and M. Koch, Optics Communications **282**, 1304 (2009), ISSN 0030-4018.

[22] M. Naftaly and R. Miles, Optics Communications **280**, 291 (2007), ISSN 0030-4018.

[23] L. Duvillaret, F. Garet, and J.-L. Coutaz, J. Opt. Soc. Am. B **17**, 452 (2000).

[24] A. Bartels, R. Cerna, C. Kistner, A. Thoma, F. Hudert, C. Janke, and T. Dekorsy, Rev. Sci. Instrum. **78**, 035107 (2007).

[25] L. Thrane, R. Jacobsen, P. Uhd Jepsen, and S. Keiding, Chemical Physics Letters **240**, 330 (1995), ISSN 0009-2614.

[26] M. Hangyo, T. Nagashima, and S. Nashima, Measurement Science and Technology **13**, 1727 (2002), ISSN 0957-0233.

[27] S. Krishnamurthy, M. T. Reiten, S. A. Harmon, and R. A. Cheville, Appl. Phys. Lett. **79**, 875 (2001).

[28] Z. Jiang, M. Li, and X.-C. Zhang, Applied Physics Letters **76**, 3221 (2000).

[29] W. Press, *Numerical recipes : the art of scientific computing* (Cambridge University Press, 2007).

[30] A. Stoppa, J. Hunger, R. Buchner, G. Hefter, A. Thoman, and H. Helm, The Journal of Physical Chemistry B **112**, 4854 (2008), ISSN 1520-6106.

[31] J. Chamberlain, J. E. Gibbs, and H. A. Gebbie, Infrared Physics **9**, 185 (1969), ISSN 0020-0891.

[32] M. Naftaly and R. E. Miles, J. Appl. Phys. **102**, 043517 (2007).

[33] R. de L. Kronig, J. Opt. Soc. Am. **12**, 547 (1926).

[34] L. Duvillaret, F. Garet, and J.-L. Coutaz, IEEE Journal of Selected Topics in Quantum Electronics **2**, 739 (1996).

[35] E. Parrott, J. Zeitler, L. Gladden, S. Taraskin, and S. Elliott, Journal of Non-Crystalline Solids **355**, 1824 (2009), ISSN 0022-3093.

[36] I. Pupeza, R. Wilk, and M. Koch, Opt. Express **15**, 4335 (2007).

[37] S. R. Tripathi, M. Aoki, K. Mochizuki, T. Asahi, I. Hosako, and N. Hiromoto, IEICE Electronics Express **6**, 1690 (2009).

[38] J. Fornies-Marquina, J. Letosa, M. Garcia-Gracia, and J. Artacho, Magnetics, IEEE Transactions on **33**, 1456 (1997).

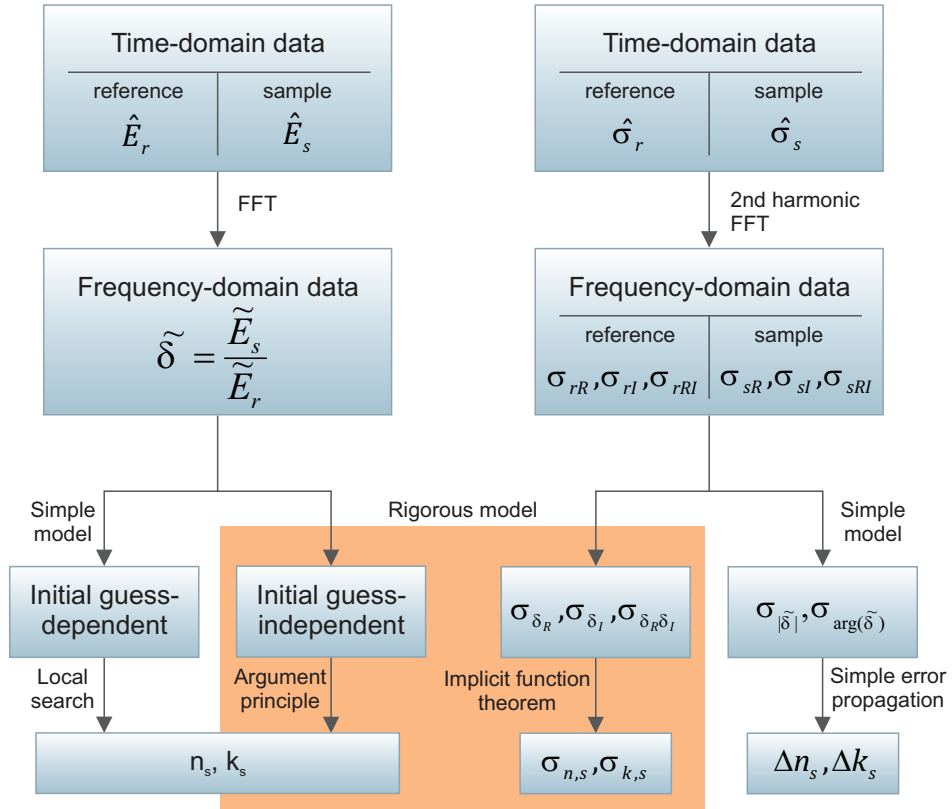[39] J. Han, H. Xu, Z. Zhu, X. Yu, and W. Li, Chemical Physics Letters **392**, 348 (2004), ISSN 0009-2614.

FIG. 1: Flow chart of data analysis starting from measured THz TDS data of the electric field of sample and reference as a function of time and error propagation analysis. The initial guess can be avoided by using a Cauchy argument principle. The error is calculated using the implicit function theorem.
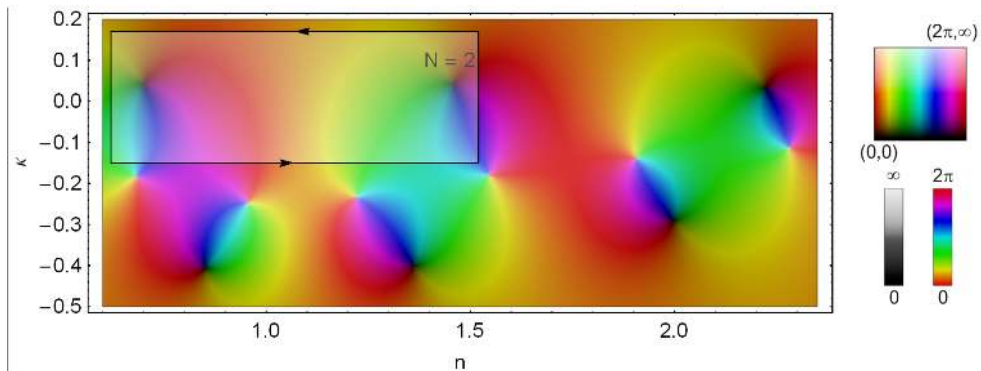
FIG. 2: HSV visualization of $F(\tilde{n}_S)$ for a data set of capric acid in transmission geometry at $\nu = 0.4$ THz. The dark areas a poles, whereas the light areas are roots of $F(\tilde{n}_S)$. In the upper left, one Jordan path is shown containing two solutions. The HSV color is projected on a scale 0 to 2π while the brightness is mapped in such a way that zeros correspond to no brightness and infinity ∞ to maximum brightness.
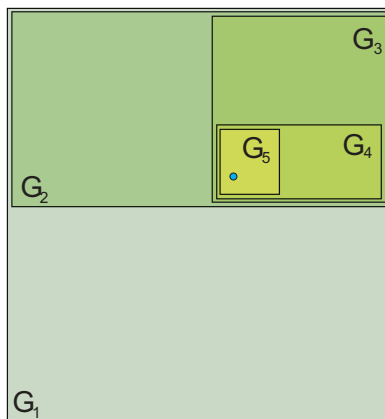
FIG. 3: Scheme of the iterative search method. The initial shape G_1 is subsequently divided into smaller subsets G_2 to G_N. As an example, the blue dot denotes a root of $F(\tilde{n}_S)$..
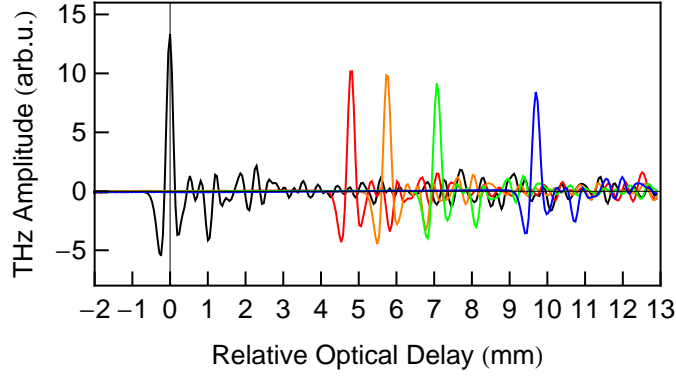
FIG. 4: Illustration of the peak shift (Eqs. (7) and (8)) for measurements of octadecane samples of different thickness. Air was used as a reference (black). The x-axis is scaled with respect to the relative optical delay $d = c\Delta t/(\bar{n}_s - \bar{n}_r)$. The transmitted octadecane pulses are shown in different colors. The corresponding sample thickness was 4.85 mm, 5.7 mm, 7.1 mm, and 9.75 mm.
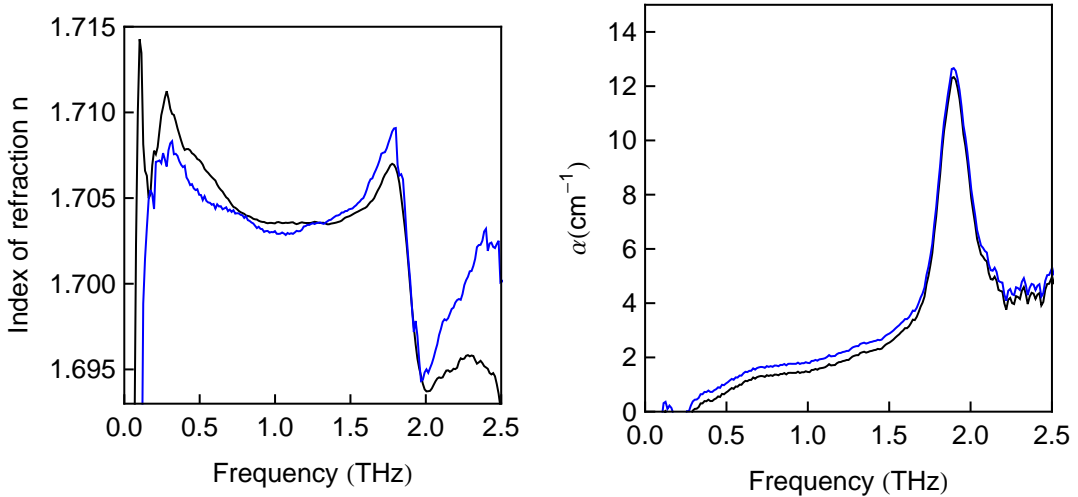

FIG. 5: Left: Real part of the index of refraction spectrum of anthracene a) as analyzed with a parameter extraction algorithm (blue) and b) as obtained from numeric Kramers-Kronig analysis (black). Right: Shown is the THz absorption spectrum of anthracene measured with an asynchronous optical sampling (ASOPS) spectrometer. The two indices of refraction curves coincide very well, but at the low and high-frequency limit, the calculated curve deviates significantly from the measured curve.
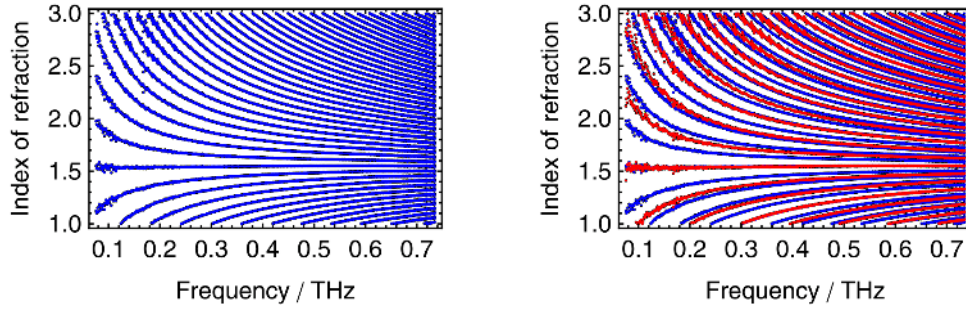
FIG. 6: Left: Index of refraction of capric acid at a thickness of $d = 0.92$ cm (blue) as obtained from an asynchronous optical sampling (ASOPS) THz time-domain spectrometer [24]. Shown are all results of the fitting algorithms in the given interval. The ambiguity of the measurement is especially present at high frequencies. Right: When measuring a second sample of different thickness $d = 0.56$ cm (red). the result shifts. Only the physically true solution will remain unchanged.
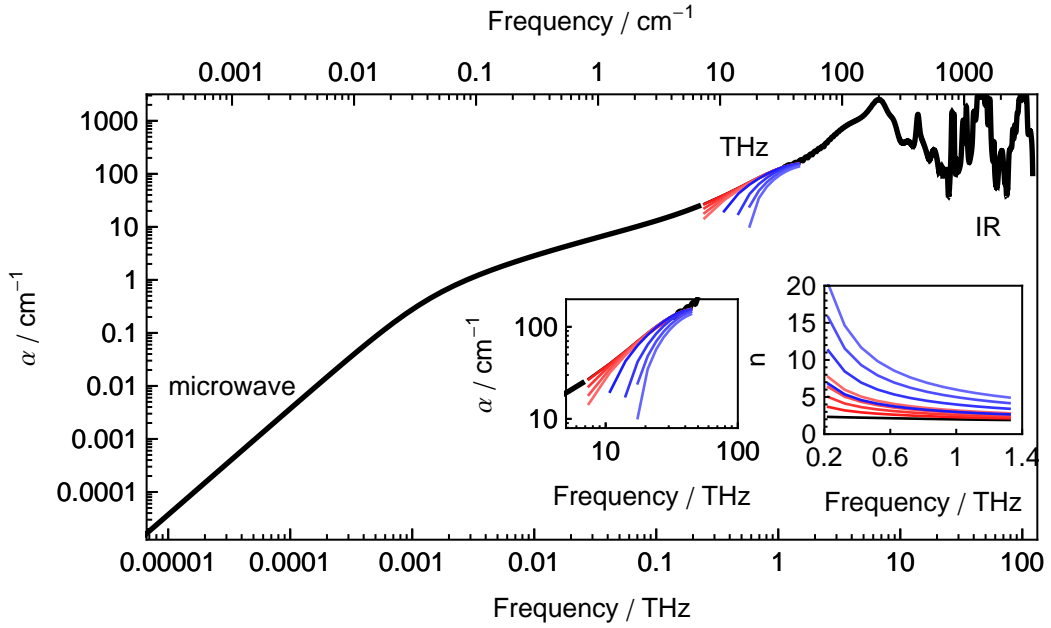
FIG. 7: Experimentally determined absorption coefficient for ethylammonium nitrate (EAN) shown in black. Different experimental setups combining microwave, THz and FTIR data were used (see Ref. [14]). In the small gap between microwave and THz data, the extrapolation of a microwave and THz dielectric fit is plotted [14]. The theoretical absorption coefficient for a 1 mm thick sample is shown in red, the corresponding result for a 300 μm thick sample is plotted in blue. The physically true solution which connects both data is displayed in black. The left inset zooms into the THz range, whereas the right inset shows the corresponding index of refraction n_s. The thin sample has well-separated minima, whereas the thick sample shows an ambiguity that can be solved when we extrapolate to both, the microwave and infrared spectral ranges.
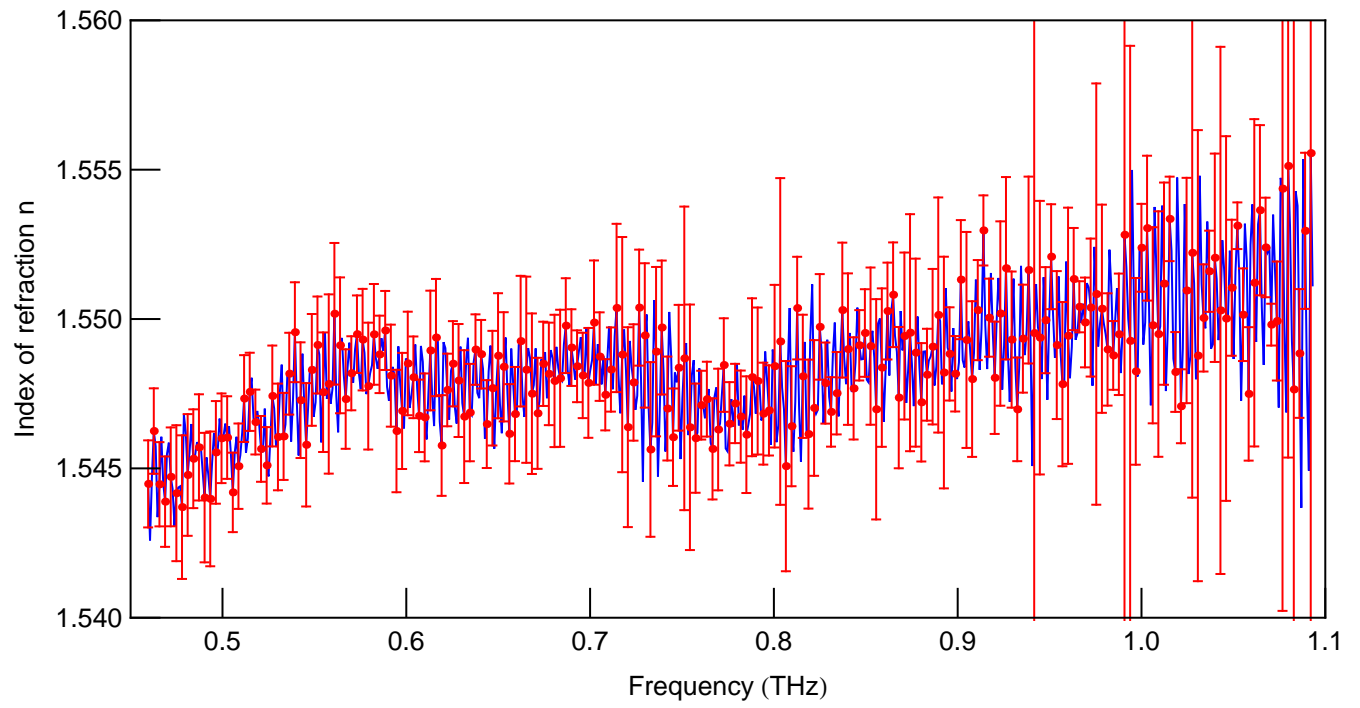
FIG. 8: Frequency-dependent index of refraction with error bars as obtained with Eq. 28 and Eq. 29.