Proceedings Paper:
Greenall, NR, Linfield, EH, Davies, AG orcid.org/0000-0002-1987-4846 et al. (3 more authors) (2016) Estimation of Spectroscopic Uncertainty and Correlation in Terahertz Time Domain Spectroscopy. In: 41st International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz), 2016. IRMMW-THz 2016, 25-30 Sep 2016, Cophenhagen, Denmark. IEEE . ISBN 978-1-4673-8485-8

https://doi.org/10.1109/IRMMW-THz.2016.7758762

Estimation of Spectroscopic Uncertainty and Correlation in Terahertz Time Domain Spectroscopy

N.R. Greenall[1,3], E.H. Linfield[1], A.G. Davies[1], L.H. Li[1], J.E. Cunningham[1] and A.D. Burnett[1,2].
[1]Department of Electronic and Electrical Engineering, University of Leeds, U.K.
[2]Department of Chemistry, University of Leeds, U.K.
[3] el09ng@leeds.ac.uk

***Abstract*—We present a method of calculating the measurement variance-covariance matrix of a spectroscopic sample's complex refractive index from time-domain statistics in order to estimate uncertainty of a measurement. We compare this method to a numerical analysis and previously derived methodology, and show that our time-based estimate is both accurate and adaptable to complex extraction models.**

I. Introduction

Estimating the measurement uncertainty in spectroscopic measurements allows us to quantify our confidence in the resulting spectra. Statistical sample estimates can often be formed numerically from multiple observations of the sample's complex refractive index, ñ. To form these observations, multiple numerical fittings of a sample's ñ to measurements are required — this is often computationally intractable. A variance-covariance (VC) matrix is a combined measurement of uncertainty and correlation between samples (in either the time or frequency-domain), from which both can be calculated. We have developed a method for estimating the VC matrices of ñ from the time-domain covariance, and showed that this method is accurate compared to a statistical sample estimate. Our results demonstrate that data points across a THz time-domain trace cannot be assumed mutually independent, as in [1], when estimating measurement uncertainty.

II. Propagation of Covariance Matrices

In this paper we model spectroscopic data as a vector of complex multivariate variables, $\boldsymbol{x}$, (in this case samples in either time or frequency) sampled from a normal distribution N with mean $\mu_{\boldsymbol{x}}$ and VC matrix $\Sigma_{\boldsymbol{x}}$, and relation matrix $\Gamma_{\boldsymbol{x}}$ [2]. $\Sigma_{\boldsymbol{x}}$ can be viewed as a sum of the underlying real and imaginary VC matrices, while $\Gamma_{\boldsymbol{x}}$ can be viewed as the difference. In the case where $\boldsymbol{x}$ is real, these matrices are equal. Therefore $\Gamma_{\boldsymbol{x}}$ is required for a complete recovery of the VC matrices of and cross covariance matrices between, the underlying real and imaginary parts of $\boldsymbol{x}$ [2]:

$$\Sigma_{R(x)} = \frac{1}{2}R(\Sigma_{\boldsymbol{x}} + \Gamma_{\boldsymbol{x}}) \quad (1)$$

$$\Sigma_{I(x)} = \frac{1}{2}R(\Sigma_{\boldsymbol{x}} - \Gamma_{\boldsymbol{x}}) \quad (2)$$

$$\Sigma_{R(x)I(x)} = \frac{1}{2}I(\Sigma_{\boldsymbol{x}} + \Gamma_{\boldsymbol{x}}) \quad (3)$$

$$\Sigma_{I(x)R(x)} = \frac{1}{2}I(\Gamma_{\boldsymbol{x}} - \Sigma_{\boldsymbol{x}}) \quad (4)$$

We note that using a singular complex variable is different to the approach taken in Ref. [1], which treats the real and imaginary parts of a variable separately. By treating complex variables as singular, we do not have to consider the internal relation between real and imaginary parts, making our analysis simpler. For variables, represented as $\boldsymbol{y}$, which are a function of $\boldsymbol{x}$, the relations for both $\Sigma_{\boldsymbol{y}}$ and $\Gamma_{\boldsymbol{y}}$ are [3]:

$$\Sigma_{\boldsymbol{y}} = A\Sigma_{\boldsymbol{x}}A^{\dagger} \quad (5)$$

$$\Gamma_{\boldsymbol{y}} = A\Gamma_{\boldsymbol{x}}A' \quad (6)$$

Where A represents a linear operation matrix or, in the case of non-linear operation, a Jacobian matrix which relates $\boldsymbol{y}$ to $\boldsymbol{x}$, and the operations denoted by $'$ and $\dagger$ are the transpose and conjugate transpose respectively.

We thus derive the appropriate matrix for each of the processing functions: the Discrete Fourier Transform (DFT), calculation of the transfer function and fitting of a model to the transfer function. The DFT can be represented by matrix multiplication, the operation matrix can be calculated by the DFT of an identity matrix, I. This gives a matrix, $A_F(= DFT(I))$, which can be used in Eq. 5 and 6:

To calculate the transfer function of a sample, we divide the sample DFT by the reference DFT. When calculating the uncertainty in the transfer function, we use a summation of the sample (S) and reference (R) contributions, calculated using different derivations of A:

$$\Sigma_H = A_S\Sigma_S A_S^{\dagger} + A_R\Sigma_R A_R^{\dagger} \quad (7)$$

$$\Gamma_H = A_S\Gamma_S A_S' + A_R\Gamma_R A_R' \quad (8)$$

The transfer function is a linear function of the sample DFT, and thus A_S is a diagonal matrix, with the inverse of the mean reference DFT along its diagonal. To calculate A_R, we take the derivative of the transfer function with respect to the reference DFT, to obtain a diagonal Jacobian matrix.

$$A_s = diag\left(\frac{1}{F_R}\right) \quad (9)$$

$$A_R = diag\left(\frac{-F_S}{F_R^2}\right) \quad (10)$$

To calculate ñ, we numerically fit a non-linear model to the transfer function, H. This model cannot be easily rewritten so that ñ is a function of H, which makes calculating the Jacobian matrix of ñ with respect to H intractable. Instead, we use the inverse of the Jacobian matrix of H with respect to ñ:

$$A_{\tilde{n}} = diag\left(\frac{dH}{d\tilde{n}}\right)^{-1} \quad (11)$$

We consider two different models, the first a non-resonant model and the second, a model with a finite number of Fabry–Pérot resonances within the sample. In the first case, we use the model [4]:

$$H_0 = \frac{4\,\tilde{n}}{(\tilde{n}+1)^2}e^{-j\frac{\omega l}{c}(\tilde{n}-1)} \quad (12)$$

With the differential being:

$$\frac{dH_0}{d\tilde{n}} = \frac{4e^{-j\frac{\omega l}{c}(\tilde{n}-1)}}{c(\tilde{n}+1)^3}\left(c(1-\tilde{n}) - jl\tilde{n}\omega(\tilde{n}+1)\right) \quad (13)$$

In the second case, we add a term for M reflections within the sample [4]:

$$H_M = H_0 \sum_{\delta=0}^{\delta=M} \left(\frac{1-\tilde{n}}{\tilde{n}+1} e^{-j\frac{\omega l \tilde{n}}{c}}\right)^{2\delta} \quad (14)$$

With the differential being:

$$\frac{dH_M}{d\tilde{n}} = \frac{dH_0}{d\tilde{n}} \sum_{\delta=0}^{\delta=M} \left(\frac{1-\tilde{n}}{\tilde{n}+1} e^{-j\frac{\omega l \tilde{n}}{c}}\right)^{2\delta} + H_0 \left(\frac{2}{(\tilde{n}^2-1)} - \frac{j\omega l}{c}\right) \sum_{\delta=0}^{\delta=M} 2\delta \left(\frac{1-\tilde{n}}{\tilde{n}+1} e^{-j\frac{\omega l \tilde{n}}{c}}\right)^{2\delta} \quad (15)$$

Where ω, l and c are the angular frequency, thickness and free space velocity of light respectively. As both of these are diagonal matrices, the inverses are trivial to calculate.

III. Results

We experimentally verified our estimation method using a 0.5-mm-thick z-cut quartz sample. To calculate ñ of the sample we took 20 reference (air) and 20 sample measurements. We then extracted ñ under two different conditions; the first using the averages of the time measurements, which gave one mean value of ñ, and the second considering the different permutations of reference and sample measurements, to give 400 different calculations of ñ. The first condition was used when estimating the VC matrix from the time-domain, while the second was used to numerically calculate the statistical samples VC matrix post extraction. In Fig. 1 we have given estimates for the uncertainty of the real refractive index, n, and extinction coefficient, κ, which were both derived from $\Sigma_{\tilde{n}}$. These estimates are based on fitting of a transfer function model that does not account for Fabry–Pérot resonances within the sample, this leads to etalons present in n and κ. For comparison, we have included an estimate based on ref [1], which treats time samples as independent measurements.

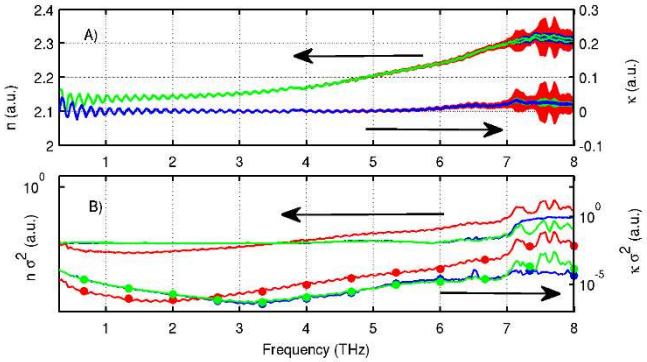

Fig. 1: A) Measurements of n (top of A) and κ (bottom of A), overlapping shaded areas indicates standard error B) Estimates of variance of n (top of B) and κ (bottom of B, dotted): red – estimate of uncertainty based on [1], green – estimate based on time domain, blue – numerical estimate of error.

We find that our estimate from the time-domain data accurately tracks the numerical estimate across a wide bandwidth. By comparison, the method presented in ref [1] is less accurate, primarily because time samples are assumed independent. Our measurement has very low uncertainty until we reach the dynamic range of the instrument at ~7 THz; this can be seen in Fig. 1A, where the shaded error diverges, and in Fig. 1B, where the variance distinctly increases. This is also the point at which our approximation fails and the time estimate diverges from the numerical estimate.

Etalons are present in ñ from resonances within the sample, and these can be resolved by fitting an appropriate transfer model, which will dramatically alter the uncertainty.

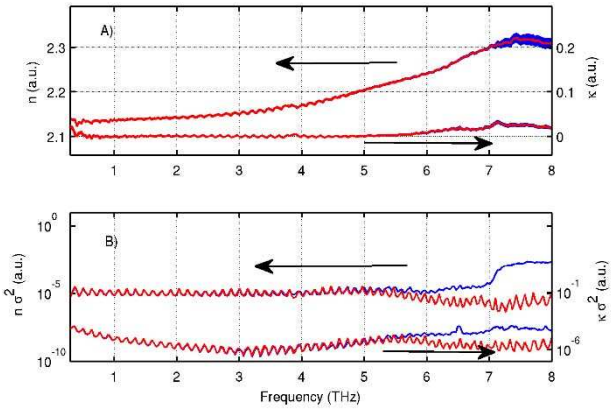

Fig 2: A) Measurements of n (top of A) and κ (bottom of A), overlapping shaded areas indicates standard error, calculated using a resonant model B) Estimates of variance of n (top of B) and κ (bottom of B), calculated using a resonant model: red – estimate based on time domain, blue – numerical estimate of error.

In Fig. 2 we have estimated the uncertainty using a finite resonant model [4]. We find that our method accurately tracks the numerical estimate to about 6 THz, where the numerical estimate tends to converge to the non-resonant estimate in Fig. 1, this is owing to the sample resonance being weak at these frequencies. Our approximation does not include this, and likely fails at higher frequencies because of this. Etalons appear in the uncertainty of ñ owing to time uncertainty. Time uncertainty will create a similar effect to varying the thickness of the sample, producing etalon variation in the extracted ñ as described in ref [4].

IV. Summary

We have demonstrated a rigorous method to estimate the uncertainty in refractive index ñ from THz time domain measurements. This estimate can be adapted to finite resonance models and is significantly more accurate than methods that neglect sample correlation.

V. Acknowledgements

This work was supported by funding from the EPSRC (UK) (EP/I026657/1), the DTRA (US) (HDTRA1-14-C-0013), the Royal Society, and the Wolfson Foundation.

References

[1] W. Withayachumnankul, B.M. Fischer, H. Lin, D. Abbott, "Uncertainty in terahertz time-domain spectroscopy measurement", *Journal of the Optical Society of America B*, vol 25, issue 6, pp. 1059, 2008.

[2] B. Picinbono, "Second-order complex random vectors and normal distributions", *IEEE Transactions on Signal Processing*, vol 44, issue 10, pp. 2637-2640, 1996.

[3] B. Hall, "On the propagation of uncertainty in complex-valued quantities", *Metrologia*, vol 41, issue 3, pp. 173-177, 2004.

[4] I. Pupeza, R. Wilk, M. Koch, "Highly accurate optical material parameter determination with THz time-domain spectroscopy", *Optics Express*, vol 15, issue 7, pp. 4335-4350, 2007.