Metrology of complex refractive index for solids in the terahertz regime using frequency domain spectroscopy

S R Chick, B N Murdin, G Matmon and M Naftaly

1 Advanced Technology Institute, University of Surrey, Guildford GU27XH, United Kingdom
2 London Centre for Nanotechnology and Department of Physics and Astronomy, University College London, London WC1H 0AH, United Kingdom
3 National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, United Kingdom

E-mail: s.chick@surrey.ac.uk

Received 3 May 2018, revised 19 September 2018
Accepted for publication 20 September 2018
Published 19 October 2018

Abstract
Frequency domain spectroscopy allows an experimenter to establish optical properties of solids in a wide frequency band including the technically challenging 10 THz region, and in other bands enabling metrological comparison between competing techniques. We advance a method for extracting the optical properties of high-index solids using only transmission-mode frequency domain spectroscopy of plane-parallel Fabry–Perot optical flats. We show that different data processing techniques yield different kinds of systematic error, and that some commonly used techniques have inherent systematic errors which are underappreciated. We use model datasets to cross-compare algorithms in isolation from experimental errors, and propose a new algorithm which has qualitatively different systematic errors to its competitors. We show that our proposal is more robust to experimental non-idealities such as noise or apodization, and extract the complex refractive index spectrum of crystalline silicon as a practical example. Finally, we advance the idea that algorithms are complementary rather than competitive, and should be used together as part of a toolbox for better metrology.

Keywords: refractive index, optical properties, dielectric, frequency domain spectroscopy, data processing

(Some figures may appear in colour only in the online journal)

Introduction
As the terahertz region above 3 THz becomes more in demand for scientific and industrial applications, with new light sources [1, 2], non-linear media [3], applications in both ultrafast science [4] and spectroscopy [5], and quantum technologies [6] the metrology of optical constants will be of elevated significance. Different experimental equipment can address different spectral regions, with vector- or scalar-network analysis appropriate for the ~100 GHz range [7], time domain spectroscopy (TDS) for the ~3 THz range [5], and ellipsometry across the spectrum (especially close to ~100 THz [8, 9]) all being independently studied, where differing considerations and terminologies apply. In addition, recent experiments in terahertz travelling standards and instrumental intercomparisons [10] have highlighted the utility of Fourier transform infrared spectroscopy (FTIR) techniques as a calibration cross-check with sufficient bandwidth to be an effective comparison for all other methodologies. Continual improvement of the metrology surrounding FTIR techniques is therefore of direct relevance to the general effort of optical parameter characterization in the terahertz regime.

FTIR, as a means to perform c.w. frequency domain spectroscopy (FDS), retains significance due to its ubiquity and relative cost efficiency, as well as decades of instrumental improvement, but especially because it can reach the

4 Author to whom any correspondence should be addressed.
10–30 THz region which all other techniques find challenging [11–13]. Experimental FDS strategies for obtaining refractive index (RI) of solids must take into account whether the sample is plane-parallel, in which case Fabry Perot (FP) interference modifies the sample’s optical properties. While some techniques achieve high accuracy using multiple spectra of the sample—in reflection and transmission, for example [14]—to obtain sufficient information about the material, others leverage the form of FP interference to minimize the demands upon experimental data. The ubiquity of the apparatus lends itself to simpler methods which require only one measurement (plus a reference) to be made, so we focus in this paper upon techniques which can be implemented using a commercial transmission-mode FTIR machine with a broad-bandwidth source.

Historical literature for extracting optical constants from the Fabry Perot interference in FDS (FP-FDS) is plentiful [15–24], but has ignored many of the systematic errors implicit in extraction techniques. Indeed, even comprehensive recent studies [10, 22] leave the sensitivity of their methods uncharacterized. Additionally, the reliance upon physical artefacts for intercomparison purposes leaves us comparatively ignorant about deviations of our measurements from the actual physical constants, either due to noise or systematic errors. Improved metrology for optical constants obtained through FTIR measurements is therefore an essential target for study and a relatively open field at the present.

This paper makes two advances: first we outline an improved methodology for extracting the real and imaginary components of the refractive index of a sample from its FP-FDS transmission function; and second we utilize well known models of that transmission function to characterize the algorithm. By using a known model of the transmission function, we obtain precisely the systematic errors implicit in the data analysis and compare quantitatively the sensitivity of different methodologies to simulated experimental non-idealities such as signal noise and frequency resolution. We show that our proposed algorithm overall outperforms an optimized version of more well-known methods, especially under experimental noise or limited resolution. We then implement a measurement of the complex refractive index of silicon in the band 2–19.5 THz and use it to show that our algorithm is very robust to sources of incoherence in the measurement such as surface roughness, where simpler popular techniques are not.

Theory

Under linearly polarized illumination normal to plane-parallel surfaces, an etalon with complex refractive index \( n = n - \text{i}k \) and thickness \( d \) has the FP-FDS transmission function \( T(f) \) [18, 22, 25]:

\[
T(f) = \frac{1 - \gamma e^{-2\alpha d} e^{-2\gamma k}}{1 + e^{-2\alpha d} e^{-2\gamma k}} \left( 1 - R \right)^2 e^{-\alpha d} \frac{1}{1 + 2\gamma R e^{-2\alpha d} - 2\gamma R e^{-2\alpha d} \cos[\Theta]}.
\]

where \( c \) is the speed of light and \( \alpha = 4k\pi / \lambda \) is the absorption coefficient. The phase \( \Theta = 2\phi + 2nf \frac{\pi}{\lambda} d \) gives rise to the FP interference fringes. The ‘coherence fraction’ \( \gamma \) describes the relative coherence between successive internal reflections in the substrate [22]; \( \gamma = 1 \) represents total coherence, \( \gamma = 0 \) an entirely incoherent process. Unless otherwise stated, we assume in this paper that \( \gamma = 1 \). The surface reflectance of the substrate, \( R \), and the phase shift on internal reflection due to absorption, \( \phi \), are functions of \( n \) and defined by the Fresnel equations [25].

We expect that an experimenter measures \( T(f) \) and then wishes to obtain \( \hat{n}(f) \) with maximum accuracy and using as little a priori information as possible. It is expected that standard data processing methods are available to the experimenter—fast Fourier transforms and curve fitting—but we would like to minimize the computational load if possible, thus prejudicing against techniques which rely very heavily upon local curve fitting to extract \( \hat{n} \) explicitly.

Finally, we assume throughout that the experimental data are sufficiently well resolved that the fringes of equation (1) are visible in the data. When samples are optically thick, or the resolution of the spectrometer is too low, the fringes are insufficiently resolved to apply these algorithms. Some algorithms require thick samples and high resolution to measure accurately and with small datapoint spacing, thus they are more difficult to apply to a wide variety of different samples in a common lab setting.

Divorcing experimental and computational errors

Typically in RI extraction using FP-FDS, an algorithm is introduced and studied theoretically and then applied to a measurement of a relatively well-known sample such as silicon or quartz in frequency bands excluding sharp or strong absorption features. Occasional cross-comparisons between results are found [22], but it is often ambiguous whether any limitations might be fundamental to the algorithm or a feature of the experiment. Better metrology can be enabled by deliberately divorcing these sources of error so that the data analysis algorithms can be compared independently of data.

We compare the different algorithms by applying them to an artificial dataset based on equation (1) and a model \( \hat{n} \). Our model data use equation (1) to calculate a model transmission function \( T_{\text{model}}(f) \) from an input \( \hat{n}(f) \). The model \( n \) & \( k \) are generated by defining \( k_{\text{model}}(f) \) over the frequency interval of interest, and using a published Kramers–Kronig (KK) transform library [26] to produce the corresponding variation in refractive index \( \Delta n_{\text{model}} \). We then add a constant \( n_{\text{add}} \) (since \( n_{\text{add}} \) is not defined by the KK transform of \( k(f) \)) so that we have a physically consistent complex refractive index of magnitude relevant to solid state physics, similar to that of silicon in the 2–20 THz range. We chose \( k_{\text{model}}(f) \) to be a sum of Gaussians, since they decay to zero rapidly which improves the accuracy of the numerical recipe for the KK transform, with the resulting \( n_{\text{model}}(f) \) shown in figure 1. To allow reasonable cross-comparison between theoretical and experimental work, we chose two Gaussians of different amplitudes and widths to approximate the observed absorption function in real data of Si etalons (see results section of this work).
Extracting \( k(f) \) is also essential. The effect on the transmission function of equation (1) of \( k(f) \) is that absorption causes a reduced fringe amplitude. There are several approaches to extracting this quantity using fringe maxima/minima values \( T_\pm \). The simplest approximation for the absorbance \( \alpha d = -\ln T_+ \) is incorrect in all cases [16, 17]. In fact, the error is most significant at high \( n \) and low \( k \) because the high-order internal reflections become more important to the total transmission of the sample—these multiple internal reflections undergo many passes of absorption; thus, the fringe amplitudes are more strongly attenuated than one would expect from the single-pass case.

It is possible to correctly extract the absorption coefficient from \( T_b \), or the ‘median’ of the transmission, \( T_b \), by substituting \( \cos \Theta = b = +1, -1, \) or 0 into equation (1) with corresponding transmission \( T_b \), which produces [16]:

\[
e^{-\alpha d} = -\left( b \cdot 2T_bR - (1 - R)^2 \right) \pm \sqrt{\left( b \cdot 2T_bR - (1 - R)^2 \right)^2 - 4T_b^2R^2}.
\]

We have dropped terms in \( k/n \), which we assume to be negligible since for typical samples in the terahertz region, \( k \ll n \). Extracting the absorbance from data using equation (3) is then possible as a simple extension to the fringe-finding method, as long as the values \( T_b \) can be found from the data.

A few features are common to all variations of the fringe method. Obviously they rely upon accurately obtaining the values \( T_b \) and \( f_b \) from the data. A simple and fast way to obtain \( T_b \) and \( f_b \) uses a peak-finding algorithm, but this is strongly limited by noise. This can be mitigated by refining the measured values using either interpolation or a local fitting.
procedure around each fringe. Figure 1(b) compares a peak-
finding algorithm without refinement to one with refinement
via fitting, and we see that systematic errors due to discre-
tization dominate peak-find algorithms, which refinement
removes at the cost of significant computation time.

Notably, fringe methods give the experimenter informa-
sion solely about the refractive index at the specific points \(f_0\); optically thick samples are desirable for their closely spaced
maxima, but necessitate a very high resolution FTIR instru-
ment. Thick samples with high absorption coefficients tend to
result in lower signal to noise ratio, exacerbating errors due to noise.

Fourier methods and hybrid peak-find techniques

The periodic nature of FP oscillations allows us to inspect the
data in the Fourier space, which we show in this section can
improve the accuracy of fringe analyses. We go on to use the
same principles to derive our new algorithm in the next sec-
tion. In an FTIR experiment, the transmitted light through the
sample as a function of time delay, \(s(\tau)\), and the reference
interferogram, \(q(\tau)\), are related by a convolution with the
Fourier transform of the transmission function, \(T(f)\):

\[
s(\tau) = t(\tau) \otimes q(\tau) \leftrightarrow S(f) = T(f)Q(f),
\]

where \(\otimes\) denotes a convolution. Upper-case symbols denote
frequency domain functions, lower-case symbols denote their
inverse Fourier transforms, and the right-hand side of equa-
tion (4) follows directly from the Fourier convolution theorem.

The periodic features of equation (1), \(T(f)\), have corre-
sponding sharp features in the time domain, \(t(\tau)\), appearing
as a harmonic series leading away from the main centerburst
as illustrated in figure 2. We label these peaks with index \(j\) so
that \(j = 0\) indicates the centreburst, \(j = 1\) indicates the first
order feature, etc. The precise shape of each feature encodes
information about the variations in periodicity of \(T(f)\) over
the measured range.

Supposing that the harmonics in \(t(\tau)\) are well resolved
and \(\hat{n}\) is sufficiently slowly varying that the high-frequency
components are negligible, a smooth Gaussian window func-
tion in the time domain, \(g(\tau; \sigma, \tau_1)\), can be used to isolate
one periodic component. Here \(\tau_1\) is the location in the time
domain of the 1st order feature in \(s(\tau)\), measured relative to
the centreburst at \(\tau_0\). The width parameter \(\sigma\) is set such that
the FWHM \(\Delta \tau < 2\sigma \sqrt{\ln 2} < \tau_0\), where \(\Delta \tau\) is the width of
the feature. The new function \(T'(\tau; \sigma, \tau_1) = g(\tau; \sigma, \tau_1)t(\tau)\)
is Fourier transformed into the complex-valued function

\[
T'(f; \sigma, \tau_1) = G(f; \sigma, \tau_1) \odot T(f).
\]

Taking \(T(f)\) to be an even function makes \(s(\tau)\) real and even
and allows us to take \(g(\tau; \sigma, \tau_1)\) to be real and even, so that
\(G(f; \sigma, \tau_1)\) is a real and even cosinusoidally oscillating func-
tion, and the inequality for \(\sigma\) above sets its width in frequency
between the fringe period and the frequency scale over which
the fringe amplitude varies. \(T'(f)\) retains the FP fringes of
\(T(f)\) and their variation in period, but eliminates their char-
acteristic asymmetrical shape, along with the average back-
ground. The Gaussian windowing process has an additional
advantage of rejecting much of the noise in the data, thus
making \(T'(f)\) more amenable to peak-finding methods than the
raw data. This means it is easier to find \(m\) and hence infer
the real part of the refractive index.

Since the amplitude of \(T'(f)\) is related in quite a compli-
cated way to \(k\), it is not accurate to use the processed fringe
amplitudes to infer \(k\) using equation (3). Instead, returning
to the raw data \(T(f)\), the fringe amplitudes of which may be
refined with fewer fitting variables now that \(f_0\) are known
from \(T'(f)\). Figure 3(a) compares \(T(f)\) to \(T'(f)\) and the
extracted \(R\) for each using the fringe-derived method with
refinement by fitting; the simpler function \(T'(f)\) increases
the contrast between maxima and minima, thus improving
the accuracy of fringe location refinement algorithms. The
resulting increase in accuracy of the recovered \(n\) is
demonstrated in figure 3(b). Notably, all Fourier filtering
methods should be robust to undersampling in the frequency domain as long as the fringes are resolved above the Nyquist limit.

While processing in the time domain has been studied previously [19, 23, 24], previous techniques were limited by computational power when the methods were first derived. The mathematics of smooth apodising functions used in this work is widely appreciated, and peak refinement techniques have been applied separately [10]. The algorithm shown in this section therefore represents a synthesis of best-practice from the literature, to which we will compare our new algorithm proposed in the next section.

**Phase methods**

Let us now assume that our measured dataset has unknown variations in the refractive index, so an extrapolation method might be vulnerable to systematic errors. We would like to have an alternative tool available that may be used as a cross-comparison and that would suffer from different systematic errors. A more sophisticated analysis also might hope to obtain higher datapoint densities, perform better using thinner samples, and still be robust to noise. Here we make use of the Fourier analysis of the last section to introduce a fundamentally different family of methods which meet these expectations.

A general approach can be defined by writing out the derivative of $\Theta$ with respect to frequency, now explicitly acknowledging that $n$ is a function of $f$. Written as a numerical derivative;

$$\frac{\Delta \Theta}{\Delta f} = 2 \frac{\Delta \phi}{\Delta f} + 2 \frac{2\pi}{c} n d + 2 \frac{2\pi}{c} f d \frac{\Delta n}{\Delta f}. \tag{6}$$

To allow for the change in refractive index, $\Delta n$, we require a relationship between the change in phase of the transmission function $\Delta \Theta$ over any given frequency interval $\Delta f$, and the refractive index. Let us assume that the phase change upon reflection is entirely negligible, which we justify in appendix C:

$$\Delta \Theta = 2 \frac{2\pi}{c} d (n \cdot \Delta f + f \cdot \Delta n). \tag{7}$$

Extracting $\Delta \Theta$ may thus help us infer $\Delta n$. Several algorithms have been demonstrated that extract information from the Fourier space [19, 20, 23, 24], which obtain the phase change through manipulation of $T(t)$. Such ‘Fourier phase’ algorithms yield datapoint densities similar to that of the input data, thus allowing the rate of change of the refractive index to be high and measured at every datapoint in the source data. Fourier methods therefore do not require thicker samples to obtain dense sampling, and so thinner samples may be measured—this in turn allows materials with higher absorption indexes to be analysed. Most notably, they allow FTIR instruments with lower frequency resolution to be used.

Present methods still require additional information so that equation (7) may be solved, which has caused these implementations to resort to extrapolation of $\Theta$ to $f = 0$. We thus encounter the same basic problem as fringe methods, and require a new approach. We propose that the problem be addressed using a single global free parameter, by adjustment of which we may obtain the entire spectrum of $n(f)$. If equation (7) is discretised such that a sample of index $i$ has a measured change in phase $\Delta \Theta_i(f_i)$, then we may write a recurrence relation;

$$\Delta \Theta_i = 2 \frac{2\pi}{c} d (n_{i-1} \cdot \Delta f_i + f_i \cdot (n_i - n_{i-1})). \tag{8}$$

Here, $\Delta f$ is the point spacing. Now if the refractive index at any fixed frequency along the interval, $n_0 (f_0)$, is known and $\Delta \Theta_i$ can be experimentally determined, then the entire spectrum of $n_i$ may be found from the relative phase change from point to point, without needing to know the absolute phase. A judicious choice of $f_0$ & $n_0$ is therefore required, but this can be estimated from the spacing of the features in $T(t)$ and further refined by studying the systematic errors caused by...
Figure 4. Systematic errors in the phase method. (a) Extracted phase change $\delta \theta$ from example data as a function of frequency $f$, showing how the convolution effect dramatically suppresses the phase change near the limits of the dataset. Inset: views of the extracted phase change on a magnified frequency axis, with the cutoff regions shown in red and denoted by dashed vertical lines. (b) Systematic errors in extracted RI caused by incorrect initial guess $n_0$; (black) correct guess replicating the input very well, (blue) incorrect guess overestimating $n_0$ by 0.001, (red) incorrect guess underestimating $n_0$ by 0.001. Flat regions near $f = 0$ are regions shown inset in (a).

Incorrect guesses for $n_0$. By avoiding extrapolation we have found a substantially different way to express the problem.

The remaining step in the algorithm is to obtain $\Delta \Theta(f_i)$ from the experimentally measured $T(f_i)$ by using almost the same windowing method as previously equation (5), except that we do not make $g(\tau; \sigma, \tau_1)$ even, we use a single Gaussian only at positive $\tau_1$, so that $G(f; \sigma, \tau_1)$ is the same Gaussian envelope but modulated with $e^{-i\phi(\tau - \tau_0)}$ instead of a real cosine function. This preserves the phase information in the argument of the result, which may be written:

$$T'(f) = A(f) e^{i(\Theta_0 + \delta \theta)},$$

where $\delta \Theta$ is the relative phase, which is a function of frequency due to variation in $n$, and $\Theta_0$ is an irrelevant constant that is determined by both the (unknown) absolute phase and the (arbitrary) phase choice for $g(\tau; \sigma, \tau_1)$. $A(f)$ is an amplitude that is real and slowly varying so long as $R, k, \phi$ are slowly varying. Effectively, the windowing has demodulated the quickly varying FP fringes by a factor of $i_1$ to recover only a slow variation due to the change in refractive index from its baseline. We therefore obtain $\delta \Theta(f)$ by finding the argument of the complex valued $T'(f)$ equations (5) and (9). Once discontinuities in the phase have been removed [28], $\Delta \Theta(f)$ is found numerically. Potential systematic errors in estimating $\tau_0$ may be removed by repeating the procedure to find $\delta \Theta(f)$ using $g(\tau; \sigma, \tau_1)$ and taking the mean.

Systematic errors inherent to the algorithm stem from two primary features. First, since the algorithm is a recurrance relation, any error in $n_{i-1}$ extends to some degree to the extracted $n_i$. Assuming an error $\varepsilon$ is introduced at index $i^{(c)}$, its influence is propagated to following indices in a manner which converges back to the correct solution as $i > i^{(c)}$, as supported theoretically in appendix A. Consequently, an error in the estimate of $n_0$ has a strong influence upon systematic errors in the extracted RI, as shown in figure 4. Errors in $n_{i-1}$ may occur due to an extracted $\delta \Theta$ not representing the actual change in RI, for example if $R, k, \phi$ have strong periodic components through some part of the dataset. Second, since the windowing procedure is equivalent to a convolution, the measured $\delta \Theta$ at each end of the dataset will be smoothly pinned to zero as shown in figure 4(a). This error may be mitigated by discarding the relevant datapoints between $f_{\min}$ & $f_{\min} + 1/(\sigma \sqrt{2})$ prior to applying the recurrence relationship. Major deviations from the expected transmission function can also cause this kind of error, for example where background calibrations become inaccurate due to low source intensity or beam splitter transmission edges etc.

We have now a complete process for obtaining $n$, but the extraction of $k$ is not so straightforwardly obtained. The local average of the function $T(f)$ over one period of the oscillation, which we will denote $T_A(f)$, is analogous to a standard integral:

$$T_A(f) = \frac{c_3}{2\pi} \int_{c_i f' = 2\pi(m-1)}^{c_i f' = 2\pi m} \frac{c_0 df'}{c_1 + c_2 \cos(c_i f')} = \frac{c_0}{\sqrt{c_1^2 - c_2^2}}.$$  

(10)

where the $c_{0,1,2,3}$ are defined by setting the integrand equal to equation (1) and neglecting variations in $c_{0,1,2,3}$ over each period. Hence;

$$T_A(f) = \frac{(1 - R)^2 e^{-\zeta d}}{1 - R^2 e^{-2\zeta d}}.$$  

(11)

We have, as before, dropped terms in $k/n$ and assumed that the terms in $R$ are sufficiently slowly varying for an adequate approximations. Equation (11) has been obtained in the literature by incoherent summation of intensities [17, 29], and can also be obtained by setting $\gamma = 0$ in equation (1), whereas the presented analysis gives a straightforward argument for the same result based on the coherent properties of equation (1) with $\gamma = 1$. Since the local average of $T(f)$ is defined by the centerburst of the interferogram, we may find $T_A(f)$ experimentally by applying the FT window $g(\tau; \sigma, \tau_0)$. The measurement from this approach is entirely independent of coherence of the experiment. Since $\gamma$ generally has its own
spectrum independent of the parameters we are interested in, we have proposed a method which may entirely avoid systematic errors due to an imperfect experiment or sample in a way only achievable previously using laborious fitting techniques [22]. Although extracting $\gamma (f)$ is outside the scope of this work, our result implies that a high resolution measurement of the coherence fraction can be obtained and used to study volume inhomogeneity or surface roughness. The total algorithm is shown diagrammatically in appendix B.

We now use our model data to demonstrate that our algorithm can recover known input data from a model transmission function. We find that the recovered $\hat{n}$ is sufficiently accurate to be essentially indistinguishable from the model inputs when plotted together, so we show the relative error in the extracted refractive index $\Delta n/\hat{n}_{\text{model}}$ expressed in parts per million (ppm) in figure 5(a). Systematic errors in the refractive index are consistently small across the spectrum, and we compare the errors in extraction with errors in the algorithm’s estimate of $T_A (f)$ (left axis), which quantifies the error due to the approximation that the terms in $R$ of equation (10) are slowly varying. Errors due to this approximation are small in this case, but are systematic and will grow as the refractive index varies more rapidly. In consequence, the algorithm will be most accurate for slowly varying refractive indices caused by Debye or Drude interactions [30, 31], and less accurate for resonant Lorentz [32] type interactions—the same is true for all methods presented in this paper.

Figure 5. Model-based testing of refractive index extraction algorithms. (a) (Left axis) Relative error in extracted $n$, in units of parts per million; (right axis) relative weight of the neglected term of equation (6) due to phase change upon reflection, compared to the term kept for use in the phase method, expressed as a percentage. (b) (Left axis) Absolute error in measurement of the incoherent transmission function $T_A (f)$ by the phase method normalized to the maximum value of the incoherent transmission function, expressed as a percentage; (right axis) absolute error in extracted $k$ using the phase method, normalized to the maximum value of $k$, expressed as a percentage.

Results

Canonical studies focus on the practical extraction of refractive index from a real dataset, but algorithms’ inherent flaws should also be studied systematically. Metrological studies generally require a controlled test which is sufficiently independent of experimental limitations, especially here where the precise values of $\hat{n}$ are not already well studied. To demonstrate this principle, we compare our new algorithm to a robust peak-find method which uses Fourier windowing, refinement by fitting, and extrapolation techniques discussed in the Theory section. The two algorithms and their input parameters were optimized in isolation on ideal $T_{\text{model}}$ datasets. We then deliberately simulate experimental error sources and compare the robustness of the different algorithms. By doing so, we gain an insight into the real limitations on refractive index extraction in contemporary settings comparable to travelling standard analysis [10].

We choose the integrated root mean square (RMS) error as a figure of merit. Better performing algorithms will show less overall deviation from the correct result, and this should be reflected in the RMS error—lower is better. Random noise and low resolution are simulated and added to $T_{\text{model}}$ separately. For simplicity, we use a multiplicative factor of $(1 + r)$, where $r$ is a random number drawn from a zero-mean normal distribution of controlled width. Errors due to low resolution sometimes termed ‘slit-width’ errors [16], where
the spectrometer’s resolution effectively smooths $T_{\text{model}}$, are simulated by applying a linear averaging function of varying width. Similar effects can be achieved by collecting a shorter interferogram and zero-padding [28].

Figure 6 shows that our Fourier phase method outperforms our implementation of the fringe method under almost all circumstances, apart from cases where the input data are essentially ideal. The RMS error in $n$ from the phase method is always lower in our example, although this will always be limited by whichever method happens to have the smaller instances of systematic error in this particular dataset. A difference in how the RMS error scales with noise would be the most persuasive argument for the superiority of either algorithm. Figure 6 shows that the fringe method’s RMS error in $k$ increases dramatically as the experimental non-idealities scale up, whereas the phase method’s RMS error in $k$ remains constant. This is true both in the case of random noise, figure 6(a), and in the case of resolution errors, figure 6(b). These kinds of non-idealities represent a range of experimental situations which can be difficult to minimize—detector sensitivity and noise floor, step size and repeatability, and so on. The Fourier phase method is insensitive to these errors since it deliberately averages $F$ as part of the analysis, which is a major reason to choose it over a fringe method.

**Application to silicon optical flats**

We demonstrate the Fourier algorithm’s practical application by measuring the transmission function of a sample of high resistivity float-zone silicon at room temperature. The wafers were purchased commercially with a thickness of $d = (1.0 \pm 0.1)$ mm, and measured to be of thickness $d = (1.070 \pm 0.005)$ mm using profilometry. The transmission spectrum was measured using a Bruker IFS125HR with a datapoint separation of 0.004 cm$^{-1}$ (0.1 GHz) and a nominal resolution of 0.01 cm$^{-1}$ (0.3 GHz) over the region $f = 30–660$ cm$^{-1}$ (0.9–19.8 THz). Our measurement of the transmission function $T_{\text{Si}}(f)$ is shown in figure 7(a), which is qualitatively similar to the modelled transmission function $T_{\text{model}}$. The Fourier phase algorithm was applied with an initial guess $n_0 = 3.4157$, and we show the extracted $\hat{n}$ in figure 7(b). Figure 7(b) also shows the same analysis using our implemented fringe method discussed above. We find an absorption in the region of the transverse-optical (TO) phonon c. 18.5 THz [15, 22] and a corresponding rapid variation in the refractive index. Similarly, a long-baseline variation in the absorption due to the phonon band is matched by a consistent variation in the real part of the refractive index. Both algorithms’ extracted $\hat{n}$ compare well to historical studies of Si [15, 17, 22], with improved datapoint density, but historical comparisons are of limited value where the same test artefact is not used. Both methods’ extracted parameters agree around the extrema of the dataset, particularly at the TO phonon absorption peak. However, there is a major disagreement around 12 THz where the Fourier phase method extracts only a negligible absorbance and the fringe method extracts a significant value.

If the extracted $\hat{n}$ data are substituted into equation (1) assuming $\gamma = 1$ (negligible incoherent scattering), the fringe method results replicate the observed $T_{\text{Si}}$ better than those from the phase method. In reality $\gamma \neq 1$ because there is some incoherent scattering in the sample. The fringe method extraction of $\hat{n}$ is sensitive to $\gamma$ (see appendix D) but the phase method is not (equation (11)). The effect of $\gamma \neq 1$ in the fringe method thus produces a complicated systematic error, especially important when $k$ is small, whereas the phase method produces a systematically better measurement of $k$.

Our experimental observations therefore underline the ability of the new phase method to discriminate absorption effects from other loss processes such as incoherent scattering or low resolution, without requiring both transmission and reflection measurements. Although extracting $\gamma(f)$ is beyond the scope of this paper, we anticipate applications of this feature of the algorithm to study imperfections in samples over large surfaces or volumes. All of our data and analysis techniques are published for free under the GNU general license [27].
Conclusion

The metrological applications of FTIR, and its relevance to the challenging regions of the terahertz spectrum, motivate us to explore new metrological techniques that take advantage of a common and well understood instrument. We have reviewed some methodologies for extracting refractive indices from FTIR measurements of optically flat samples and suggested a new technique that makes the most of the computational ability to filter in the Fourier space. We showed that it is possible to quantify the systematic errors in the algorithm by applying it to model data, and in doing so have demonstrated a thorough metrological analysis of these algorithms. This kind of methodology can allow future intercomparisons between algorithms using a standardised test which allows for comparison to exact analytic results. To ensure that fair comparisons can be drawn and our conclusions replicated, our analysis tools, data, and code for each technique discussed in the paper have been published under the GNU general license [27].

Our proposed algorithm yields the complex refractive index with a very fine datapoint spacing on the order of that of the input data, a major benefit over more commonly utilized fringe methods. We have also discussed how our method prefers to analyse optically thin samples, whereas fringe methods would prefer thick samples—this enables us to measure the optical constants of the material accurately in the presence of higher absorption and with lower frequency resolution. The robustness of the method to incoherence is significant for metrology, since surface roughness and volume inhomogeneities in physical artefacts can be ignored as long as a good measurement of the mean thickness of the samples is obtained.

By the same token, our analysis has also for the first time dedicated significant effort to analysing the limitations of the algorithm, which is a crucial contribution for advancing metrology in the terahertz region. Our technique has managed to avoid the inherent difficulties of KK transformation as part of the extraction technique, so that we avoid some sources of systematic error in other analyses [17]. Some of the limitations we have discussed can be mitigated by using the same principles we have outlined to form an iterative improvement procedure, much like that of King et al [17].

While our algorithm has significant advantages over others, we have established the different systematic errors in each of the different algorithms, and so none can be considered universally optimal. A judicious choice must be made based on the shape of the transmission function, the resolution of the spectrometer, and the approximate optical thickness of the sample. Our new algorithm is therefore part of a toolbox of different techniques which would allow a metrologist to select the best analysis for a given dataset. In cases where particular rigour is desirable, an intercomparison between techniques is enabled using our paradigm of measuring the RMS error in the extraction of a known model dataset, which can be tailored to be similar to the measured data. Automation of this procedure could even allow the choice of the best analysis algorithm without user intervention, thus making the best measurement of complex refractive index straightforward even for a non-expert user.

Acknowledgments

We acknowledge financial support from the UK Engineering and Physical Sciences Research Council (ADRFSS, Grant No. EP/M009564/1).

Appendix A. Propagation of errors in the Fourier phase algorithm

Starting from the recurrence relation:

\[ \Delta \Theta_i = 2\frac{\pi}{c}d \left( n_{i-1} \cdot \Delta f + f_i \cdot (n_i - n_{i-1}) \right). \]

We rearrange to find \( n_i \):

\[ n_i = \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_i + n_{i-1} \frac{f_{i-1}}{f_i} \frac{1}{f_i}. \]

By iterative substitution we can find any \( n_i \) in terms of its previous iterand:

\[ n_{i-1} = \frac{1}{f_{i-1}} \frac{c}{4\pi d} \cdot \Delta \Theta_{i-1} + n_{i-2} \frac{f_{i-2}}{f_{i-1}} \frac{1}{f_{i-1}}. \]

Therefore the \( i \)th iteration is in terms of the \((i - 2)\)th iteration:

\[ n_i = \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i} + \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i-1} + n_{i-2} \frac{f_{i-2}}{f_i} \frac{1}{f_i}. \]

If we deliberately introduce an error by substituting \( n_{i-2} \rightarrow n_{i-2} + \epsilon_{i-2} \):

\[ n_i + \epsilon_i = \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i} + \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i-1} + n_{i-2} \frac{f_{i-2}}{f_i} + \epsilon_{i-2} \frac{f_{i-2}}{f_i}. \]

The final term is the error in \( n_i \) which we can generalise to any error introduced \( I \) iterations previously:

\[ \epsilon_i = f_{i-1} \epsilon_{i-1}. \]

Any error introduced will always reduce in contribution to the total error, by a scaling rule that looks like \( 1/f \) as shown experimentally in the main text. What if the error is instead in the phase, which we argue in the paper is the cause of most of our systematic errors? Recall:

\[ n_i = \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i} + \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i-1} + n_{i-2} \frac{f_{i-2}}{f_i} \frac{1}{f_i}. \]

If we make the replacement \( \Delta \Theta_{i-1} \rightarrow \Delta \Theta_{i-1} + \epsilon_{i-1} \):

\[ n_i + \epsilon_i = \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i} + \frac{1}{f_i} \frac{c}{4\pi d} \cdot \Delta \Theta_{i-1} + n_{i-2} \frac{f_{i-2}}{f_i} + \frac{1}{f_i} \frac{c}{4\pi d} \epsilon_{i-1}. \]

This is reminiscent of the case above, so we can infer that for any phase error \( \epsilon_{i-1} \):

\[ \epsilon_i = \frac{c}{4\pi d f_i} \epsilon_{i-1}. \]

Phase errors therefore drop in significance in a similar manner to errors in \( n \).
Appendix B. Implementing the Fourier phase algorithm

Here we represent the phase algorithm for extracting the complex refractive index in a flow diagram, showing the main calculation and decision steps:

\[ T(f) \rightarrow \text{IFFT} \rightarrow T(\tau) \rightarrow \text{window } \tau_1 \& \text{FFT} \rightarrow e^{-i\tau f} \rightarrow \arg( ) \rightarrow \Delta \Theta(f_i) \rightarrow \text{numerical derivative} \rightarrow \text{assume } n_0 \]

**Appendix C. Systematic errors due to dropped terms**

In our main analysis, we make the assumption that the rate of change of the phase shift upon reflection is negligible. To study mathematically whether this is the case and how robust our algorithm is if this assumption is dropped, we split the derivative into two components:

\[ \frac{d\Theta}{df} = \frac{d}{df} \left[ 2\phi + 2n f \frac{2\pi}{c} d \right] = A + B. \]

We assumed \( A = 0 \) to derive the algorithms in the phase method. Since:

\[ \tan \phi = \frac{2k}{n^2 + k^2 - 1}, \]

we will naturally find that \( A \) is quite a complicated quantity. We can make the assumption that \( k \ll n \) and apply an approximation to find:

\[ A(k \ll n) = \frac{4}{(n^2 - 1)^2} \left( (n^2 - 1) \frac{dk}{df} + 2kn \frac{dn}{df} \right). \]

Hence by substitution:

\[ \frac{d\Theta}{df} = \frac{4}{(n^2 - 1)^2} \left[ (n^2 - 1) \frac{dk}{df} + 2kn \frac{dn}{df} \right] + 2 \frac{2\pi}{c} d \left[ n + f \frac{dn}{df} \right]. \]

In the regime where \( f \sim 10^{12} \) Hz then it is reasonable to conclude that:

\[ \frac{4}{(n^2 - 1)^2} 2kn \frac{dn}{df} \ll 2 \frac{2\pi}{c} d \frac{dn}{df}. \]

So we can drop the term on the LHS above and arrive at the estimate:

\[ \frac{d\Theta}{df} = \frac{4}{(n^2 - 1)^2} \frac{dk}{df} + 2 \frac{2\pi}{c} d \left[ n + f \frac{dn}{df} \right]. \]

The algorithm will tend to be most wrong near stationary points in \( n \), which happen to correlate with peaks in \( \frac{df}{dn} \). If \( \frac{df}{dn} \) dominates the term in square brackets above, then the relative error in \( \Delta \Theta \) will be of opposite sign around each maximum, as observed in the Theory section in the main work. Since the sign of the phase error changes around the maximum, error propagation through the spectrum is minimized by compensation through the mechanism discussed in appendix A.

**Appendix D. Systematic errors due to the coherence factor**

Here we study the effect of the coherence fraction upon the peak extraction method. In the perfectly coherent case \( \gamma = 1 \):

\[ e^{-\alpha d} = \frac{(b \cdot 2 T \pm \alpha) (1 - R^2)}{2 T_\pm R^2}. \]

However, in the general case return to equation (1) of the main work that:

\[ \frac{1 - \gamma^2 e^{-2\alpha d} R^2}{1 - \gamma^2 R^2 e^{-2\alpha d} - 2\gamma R e^{-\alpha d} \cos[\Theta]} \cdot \]

By making the same replacement for \( \cos[\Theta] \):

\[ T_{\pm b} = \left[ 1 - \gamma^2 e^{-2\alpha d} R^2 \right] (1 - R)^2 e^{-\alpha d} \]

\[ 1 + \gamma^2 R^2 e^{-2\alpha d} - 2\gamma R e^{-\alpha d} \]

If we let:

\[ x = e^{-\alpha d}, \]

then the equation can be expressed as:

\[ T_{\pm b} (1 - x^2 R^2) (1 + \gamma^2 R^2 x^2 - 2\gamma R x) = (1 - \gamma^2 x R^2) (1 - R)^2 x. \]

Again, we have assumed that \( n \gg k \). At this point it is clear that we must solve a quartic equation which does not have the
convenience of becoming a simple quadratic in $Y$ under the substitution $Y = x^2$. While the equation must have solutions for $x$ in terms of $\gamma, R, b$ it should not be considered to reduce to the quadratic solution given in the main work, hence the peak method fails. The quartic equation does have a general solution, which is vastly more complicated than we consider within the scope of this paper. Either way, it is reasonable to suppose that an independent estimate for $\gamma$ is essential to extract $\alpha$, which is not acceptable from the position of requiring minimal a priori assumptions about the sample’s properties.

**ORCIDs**

S R Chick [https://orcid.org/0000-0001-8599-6609](https://orcid.org/0000-0001-8599-6609)

**References**

[1] Vicario C et al 2015 High field broadband THz generation in organic materials J. Mod. Opt. 62 1480–5

[2] Hafez H A et al 2016 Intense terahertz radiation and their applications J. Opt. 18 093004

[3] Hafez H A et al 2015 Nonlinear terahertz field-induced carrier dynamics in photoexcited epitaxial monolayer graphene Phys. Rev. B 91 035422

[4] Zhang X C et al 2017 Extreme terahertz science Nat. Photon. 11 16–8

[5] Hwang H Y et al 2015 A review of non-linear terahertz spectroscopy with ultrashort tabletop-laser pulses J. Mod. Opt. 62 1447–79

[6] Chick S et al 2017 Coherent superpositions of three states for phosphorous donors in silicon prepared using THz radiation Nat. Commun. 8 16038

[7] Yang B B et al 2012 A high-Q terahertz resonator for the measurement of electronic properties of conductors and low-loss dielectrics IEEE Trans. Terahertz Sci. Technol. 2 449–59

[8] Galuza A A et al 2016 Developments in THz-range ellipsometry J. Infrared Mill. Terahertz Waves 34 682–708

[9] Galuza A A et al 2016 Developments in THz-range ellipsometry: quasi-optical ellipsometer IEEE Trans. Terahertz Sci. Technol. 6 183–90

[10] Oberto L et al 2017 Measurement comparison among time-domain, FTIR, and VNA-based spectrometers in the THz frequency range Metrologia 54 77–84

[11] Dhillon S S et al 2017 The 2017 terahertz science and technology roadmap J. Phys. D: Appl. Phys. 50 043001

[12] Bell R 1972 Introductory Fourier Transform Spectroscopy (New York: Academic)

[13] Davis S P et al 2001 Fourier Transform Spectrometry (San Diego: Academic)

[14] Sun J and Lucyszyn S 2018 Extracting complex dielectric properties from reflection-transmission mode spectroscopy IEEE Access 6 8302–21

[15] Randall C M and Rawcliffe R D 1967 Refractive indices of germanium, silicon, and fused quartz in the far infrared Appl. Opt. 6 1889–95

[16] Swanepoel R 1983 Determination of the thickness and optical constants of amorphous silicon J. Phys. E. Sci. Instrum. 16 1214–22

[17] King S W and Milosevic M 2012 A method to extract absorption coefficient of thin films from transmission spectra of the films on thick substrates J. Appl. Phys. 111 073109

[18] Hernandez G 1985 Fabry–Perot with an absorbing etalon cavity Appl. Opt. 24 3062–7

[19] Chamberlain J et al 1969 The determination of refractive index spectra by fourier spectrometry Infrared Phys. 9 185–209

[20] Bell E E 1966 Measurement of the far infrared optical properties of solids with a michelson interferometer used in the asymmetric mode: part 1, mathematical formulations Infrared Phys. 6 57–74

[21] Wang M R et al 2016 Extracting dielectric parameter based on multiple beam interference principle and FTIR system in the terahertz range Proc. of the 41st Int. Conf. on Infrared, Millimetre, and Terahertz Waves (Copenhagen) pp 1–2

[22] Grossman E N and McDonald D G 1995 Partially coherent transmittance of dielectric lamellae Opt. Eng. 34 1289–95

[23] Lowenstein E V and Smith D R 1971 Optical constants of far infrared materials. I: analysis of channelled spectra and application to mylar Appl. Opt. 10 577–83

[24] Russell E E and Bell E E 1967 Measurement of optical constants of crystal quartz in the far infrared with the asymmetric Fourier-transform method J. Opt. Soc. Am. 57 341–8

[25] Born M and Wolf E 1980 Principles of Optics 6th edn (Oxford: Pergamon) pp 628–32

[26] Lucarini V et al 2005 Kramers–Kronig Relations in Optical Materials Research (Heidelberg: Springer)

[27] Chick S 2018 Supporting Data for ‘Metrology of Complex Refractive Index for Solids in the Terahertz Regime Using Frequency Domain Spectroscopy’ (Online: Zenodo) (https://doi.org/10.5281/zenodo.1237819) (Dataset)

[28] Withayachumnankul W and Nafaty M 2014 Fundamentals of measurement in terhertz time-domain spectroscopy J. Infrared Millimetre Terahertz Waves 35 610–37

[29] Harbecke B 1986 Coherent and incoherent reflection and transmission of multi layer structures Appl. Phys. B 39 165–70

[30] Zaytsev K I et al 2014 Accuracy of sample material parameters reconstruction using terahertz pulsed spectroscopy J. Appl. Phys. 115 193105

[31] Yurchenko S O and Zaytsev K I 2014 Spectroscopy of nafion in terahertz frequency range J. Appl. Phys. 116 113508

[32] Kuzmenko A M et al 2016 Terahertz spectroscopy of crystal-field transitions in magnetoelectric TmAl12(BO3)3 Phys. Rev. B 94 174419