Journal of Infrared, Millimeter, and Terahertz Waves (2022) 43:847–856
https://doi.org/10.1007/s10762-022-00886-y

Algorithm for Determination of Cutoff Frequency of Noise Floor Level for Terahertz Time-Domain Signals

E.S. Reyes-Reyes[1] · R. Carriles-Jaimes[1] · E. Castro-Camus[2]

Received: 27 September 2022 / Accepted: 3 November 2022 / Published online: 26 November 2022

Abstract
The frequency-dependent signal-to-noise ratio of terahertz time-domain signals is a relevant source of uncertainty for parameters measured with it. It also limits the total usable bandwidth of such signals. In the great majority of cases, the processes to establish the limits of this usable bandwidth are determined based on the experience of the user. Therefore, it is desirable to develop a procedure to automate this calculation. In this work, a method to estimate the bandwidth of terahertz time-domain signals is presented. Different spectra were analyzed, showing the potential of the algorithm in the calculation of cutoff frequencies which delimits the usable bandwidth.

Keywords Terahertz · Bandwidth · Cutoff frequency

1 Introduction

THz time-domain spectroscopy (THz-TDS) is currently the most powerful technique for the characterization of the complex dielectric properties of a wide range of materials in the far-infrared part of the spectrum. This has opened interesting opportunities for applications in areas as diverse as industry [1–3], medicine [4–6], art inspection [7, 8], among many others.

✉ E. Castro-Camus
enrique.castrocamus@physik.uni-marburg.de

R. Carriles-Jaimes
ramon@cio.mx

[1] Centro de Investigaciones en Optica, A.C., Loma del Bosque 115, Lomas del Campestre, 37150, Leon, Guanajuato, Mexico

[2] Department of Physics and Material Sciences Center, Philipps-Universität Marburg, Renthof 5, 35032, Marburg, Germany

THz-TDS measures the electric field waveform of a terahertz pulse directly in the time domain and through Fourier transform one can obtain the spectrum. From this information it is possible to calculate the optical parameters of the sample. Typically in most spectra, the amplitude decreases at high frequencies until it is not longer distinguishable from the noise floor [9, 10]. In order to avoid misinterpretation of the spectral features, it is important to estimate the cutoff frequency at which the signal to noise ratio (SNR) approaches one. The position of this cutoff will determine the usable bandwidth for each particular spectra.

Usually, defining the measurable bandwidth limit in THz spectra is not done by a rigorous analysis because, for a trained eye, it is straightforward to distinguish the frequency at which the spectral amplitude decays to the noise floor, so this cutoff frequency is typically calculated "by eye." However, when a large volume of data is analyzed, such as in the analysis of THz images which contain thousands of spectra per image, it is necessary to automate the identification of the cutoff frequency for each spectrum in the data set. In this work, we propose and test an algorithm to estimate the cutoff frequency at which the amplitude of the spectrum decays to the level of the noise floor.

2 Cutoff Frequency Calculation

The function $g_{\mathrm{THz}} = (t - t_0)e^{-(t-t_0)^2/\sigma^2}$ is commonly used to analytically model a single-cycle terahertz pulse in the time domain [11]. In our case, without lose of generality, we take $t_0 = 0$ and $\sigma = 1\,\mathrm{ps}$ in order to keep this analysis as simple as possible. Also, we assume that the pulse contains white noise n_w, which has a constant spectral power density. Therefore, the pulse is expressed as $g(t) = te^{-t^2} + n_w$. Using the Fourier transform, we obtain the THz spectrum

$$G(\omega) = -i\pi^{\frac{3}{2}}\omega e^{-\frac{\omega^2\pi}{2}} + N, \tag{1}$$

where N is the Fourier transform of n_w which is a constant for the case of white noise. In this analysis, N represent the noise floor in the spectrum, which experimentally originates, mainly, from THz detectors and laser fluctuations [12, 13]. At low frequencies, $G(\omega)$ increases proportionally to ω; however, as ω increases, $G(\omega)$ decreases as $e^{-\frac{\omega^2\pi}{2}}$. Eventually, at some cutoff frequency which we call ω_c, the spectral amplitude will have values comparable to N. For frequencies higher than ω_c, the first term of the Eq. 1 will be negligible and, therefore, $G(\omega) \approx N$. Given the above considerations, if we now multiply $G(\omega)$ by ω, the new equation $G_1 = \omega G(\omega)$ will have two different behaviors bounded by ω_c. For $\omega < \omega_c$, the first part of G_1 will have an ω^2 factor multiplying the exponential, resulting in a less abrupt decay of the spectral amplitude; however, after ω_c which correspond to noise floor in the spectrum, G_1 will behave as a linear increasing function ωN. This change of behavior in

G_1 will generate a minimum at the frequency ω_c. The same reasoning can be applied for $G_n = \omega^n G(\omega)$, generating minima at

$$\frac{1}{n} e^{-\omega^2 \pi} \left[\pi^4 \omega^4 - (n+1)\pi^3 \omega^2 \right] = N^2. \tag{2}$$

Therefore, in order to find a reasonable estimate for the cutoff frequency value ω_c, it is sufficient to multiply $G(\omega)$ by ω^n and find the minimum in order to calculate the cutoff frequency ω_c, as shown in Fig. 1 (a). From this figure it is clear that the part of the spectrum corresponding to the noise floor has faster growth rate for higher values of n. As a result, as n increases, the value of ω_c decreases slightly; the difference in ω_c calculated between using G_1 and G_4 was only 160 GHz, which is equivalent to 8.4 % of the total bandwidth in this example. In the Additional Material found in the "Appendix" section at the end of this article, we provide the implementation of the algorithm described above in MATLAB/Octave.

Before analyzing experimental spectra, it is necessary to mention three factors that may affect the performance of our method. The first is the random oscillations in the noise floor, which can generate minima with lower values than the minimum generated in G_n. This problem can be solved by smoothing the spectrum. The second is the noise floor itself, since in some THz systems the noise floor does not remain constant, but has smooth variations with frequency; therefore the generation of the minimum would not be clear. If this is the case, an independent measurement of the instrument noise floor could fix the problem. The third factor is the existence of absorption lines within the "good" signal-to-noise-ratio region, which is also addressed by appropriate smoothing of the spectrum.

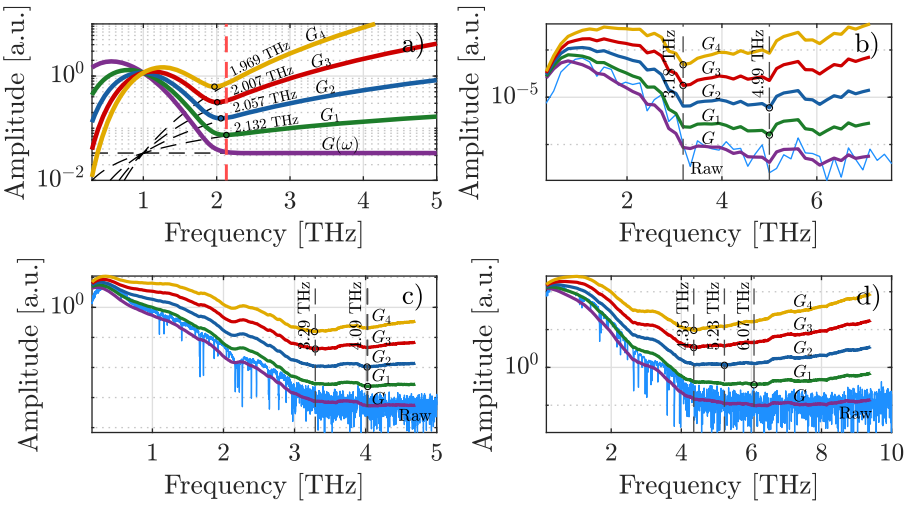

Fig. 1 (a) Calculation of the cutoff frequencies ω_c through $G_n = \omega^n G(\omega)$ for $n = 1, 2, 3, 4$ and $N = 0.0334$. For reference, a red dashed line is shown corresponding to ω_c calculated for G_1. (b)–(d) Calculations of ω_c for $n = 1, 2, 3, 4$ from experimental data

2.1 Results and Discussion

Data to support the considerations described above are shown in Fig. 1 (b–d), where reference spectra of three different THz systems in transmission configuration are shown. For the acquisition of the original time-domain data, various spectrometers were used. In particular

- A home-built spectrometer based on a Ti:sapphire oscillator at a central wavelenght of 800 nm with a pulse duration of 35 fs using a SI-GaAs photoconductive antenna as emitter and a 1 mm [110] ZnTe-crystal-based electrooptic sensor as detector, recording 100 points over a 6.6 ps delay.
- A commercial spectrometer based on an Yb:fiber laser centered at 1064 nm with a pulse duration of 90 fs using a photoconductive anternnas both as emitter and detector, recording 1600 points over 160 ps.
- A commercial spectrometer based on an Yb:fiber laser centered at 1064 nm with a pulse duration of 90 fs using a photoconductive anternnas both as emitter and detector 8000 points over 800 ps.

The spectra in Fig. 1(b–d) correspond to the three spectrometers described above in that same order. It is worth mentioning that we tested our algorithm with spectra of 3 more spectrometers with various characteristics, not shown, and all the results are comparable.

The smoothing was done using the moving average method [14], with a width w_s equal to 3% of the total data recorded in the waveform. The smoothed spectra is denoted as G. As we can see on Fig. 1 (b), there is no difference using G_1 and G_2, calculating the cutoff frequency at 4.99 THz, where it is easy to notice that ω_c must be at lower frequencies. However, increasing $n = 3$, the cutoff frequency is displaced at 3.18 THz, which is more consistent than the previous result. Using G_4, w_c is calculated at the same frequency as G_3. On the other hand, analyzing the spectrum of Fig. 1 (c), the small peak centered at 3.93 THz is an artifact of the THz system which resulted in a miscalculation of ω_c by G_1 and G_2; however, both G_3 and G_4, are not affected by this artifact and agree with the frequency estimated "by eye." In the case of the spectrum shown in Fig. 1 (d), we obtained three different values of ω_c. As in Fig. 1 (b) and (c), both G_3 and G_4 have the same value of ω_c, which suggests that it would be enough to use G_3 to correctly estimate the cutoff frequency in a conventional reference spectrum.

As mentioned in the previous section, since the method is based on the calculation of the minimum generated by multiplying the spectrum by ω^n, any other minimum originated by random noise may result in an incorrect value of ω_c. For this reason, another essential element in our processing is the smoothing, since it facilitates the calculation of the cutoff frequency by omitting the random oscillations in the experimental signal. In order to establish the limits of the degree of smoothing that can be performed to the signals, Fig. 2 shows the spectra and their cutoff frequencies calculated using different values of w_s and multiplying the smoothed spectrum by ω^3.

In the spectrum shown in Fig. 2(a), which corresponds to a reference spectrum recorded from a metallic foil sample, the cutoff frequency calculated using $w_s = 3\%$ of the total number of data in the spectrum was $\omega_c = 2.1$ THz. Although there is a

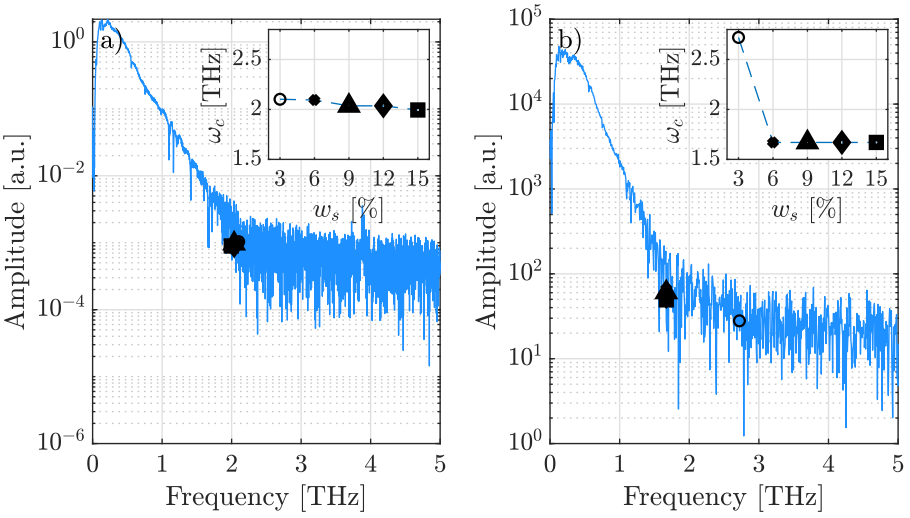

Fig. 2 Spectra and their corresponding cutoff frequencies using $w_s = 3\%, 6\%, 9\%, 12\%$, and 15% for $n = 3$. Insets show ω_c as a function of w_s

change in ω_c as the value of w_s increases, this difference is minimal; by increasing w_s to 15%, the cutoff frequency only decreased by 107 GHz, which is equivalent to 5.1% of the initially calculated bandwidth. This result suggests that increasing the value of w_s in the smoothing for reference spectra does not lead to a significant improvement. It is also worth mentioning that we attempted various smoothing methods which include the Savitzky-Golay, the moving average and the non-centered moving average methods, being the latter, the one that gave best results

On the other hand, Fig. 2 (b) shows the spectrum of a metallic sample with an irregular surface. In this case, the cutoff frequency calculated using $w_s = 3\%$ is $\omega_c = 2.72\,\text{THz}$. However, as w_s increased to 6%, the cutoff frequency decreased to $\omega_c = 1.66\,\text{THz}$, with no further changes for w_s greater than 9%. Based on the results shown in Figs. 1 and 2, as well as many others that we do not show, we can empirically say that it is ideal to use $w_s \geq 6\%$ and $n \geq 3$ for the processing of most spectra, with no significant change in the calculation of ω_c.

As shown, the proposed method analyzes reference signals without major difficulties. However, this method is not limited to calculating the cutoff frequency in reference spectra, but can be used in any spectrum, regardless of the characteristic of the sample or the circumstances in which the signals were recorded. For this purpose, three data sets were analyzed. The first set is shown in Fig. 3(a), which correspond to a high-density polyethylene (HDPE) sample. Waveforms were recorded in reflection using a 3-inch focal length lens. The sample was placed at the focal point of the lens and subsequently moved $\pm 12\,\text{mm}$ along the optical axis, recording the waveforms every 3 mm. By moving the sample away from the focal point, the amplitude spectrum decays but it is not affected evenly across the entire band, which results in a decrease in the effective bandwidth, so this data set will allow the validation of the effectiveness of the algorithm in a very practical manner. Cutoff frequencies were calculated using $w_s = 6\%$ and $n = 3$, and are indicated by circles in each spectrum.

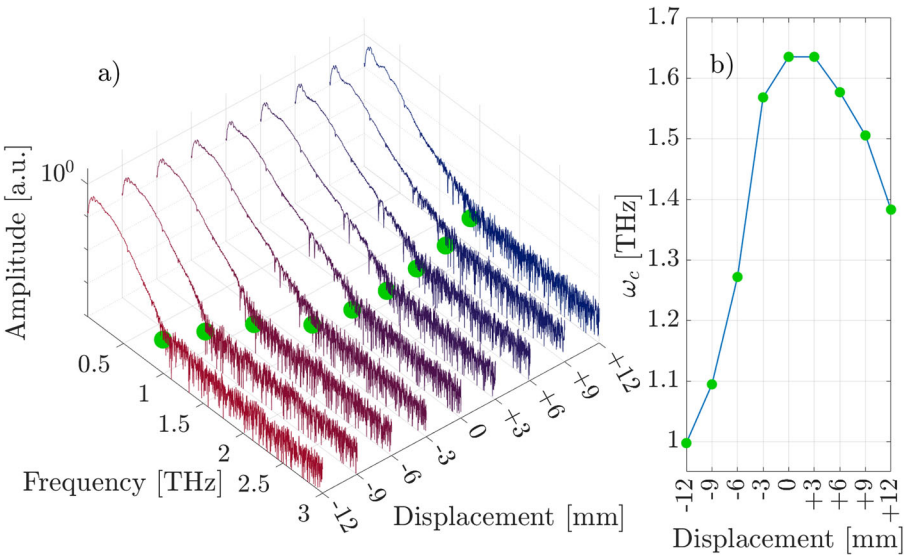

Fig. 3 (a) Spectra of a HDPE sample recorded in reflection geometry. The sample was initially placed at the focal plane of the lens. Subsequently, the sample was displaced ± 12 mm out of the focal length, recording the waveforms every ±3 mm. Circles show the cutoff frequency calculated using $w_s = 6\%$ and $n = 3$. (b) ω_c as a function of sample displacement

As expected, the highest cutoff frequency, which was 1.66 THz, corresponds to the sample placed at the focal point; this frequency was the same when the sample was displaced +3 mm, which is within the Rayleigh length of the focus of the THz beam. In addition, ω_c decreases as the sample moves away from the focal point of the lens, falling to 1 THz for −12-mm displacement. This can be seen in Fig. 3(b), where the cutoff frequencies are shown as a function of sample displacement.

The last two data sets correspond to six signals obtained from a THz image of an artistic painting with two main areas of interest. The first three spectra correspond to an area where mercury sulfide was detected, which has an absorption line centered at 1.12 THz. This feature is important to analyze since the absorption line can be confused with the noise as it is located in a part of the spectrum where the SNR is low. The last 3 spectra were taken from the edges of the image exhibited a very poor measurable bandwidth. These two data sets are of special interest as they are highly complex spectra and will allow us to examine the robustness of the method.

Figure 4 (a–c) shows the spectrum of a measurement where mercury sulfide was present. Although the absorption peak is easily distinguishable, the algorithm correctly calculates the cutoff frequency in (a) and (c). Apparently, the cutoff frequency calculated in (b) should be located at a lower frequency, since clearly the spectral amplitude decays to the noise floor before 2.2 THz. In order to correct the cutoff frequency in this spectrum, the value of w_s was increased to 9%, calculating $\omega_c = 1.5$ THz. For the spectra in Fig. 4(d–f), it can be seen that these signals are extremely noisy, so it is natural to consider increasing the value of w_s. Nevertheless, the calculated cutoff frequencies agree with the "by eye" estimation, this without having increased w_s.

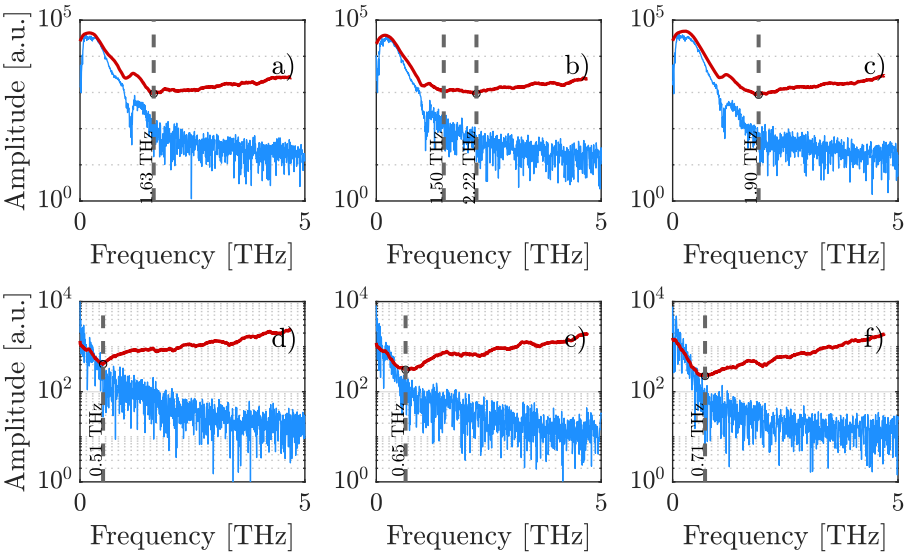

Fig. 4 (a)–(c) Spectra corresponding to a region of interest containing mercury sulfide with its spectral fingerprints at 1.12 THz. (d)–(f) Spectra with narrow measurable bandwidth. Red line indicates G_3, while vertical dashed line indicates the cutoff frequency calculated in each spectrum

Finally, it is important to mention that computation time is a critical factor in this work, since this method is designed to analyze large amounts of data. For this purpose, eight THz time-domain images of 550×460 pixels were analyzed, with a total of more than two million spectra, which were processed in approximately 1840 seconds that corresponds to $\sim 900\,\mu$s per spectrum. The images were processed in a commercial computer (Intel Core i5-11400H with 8 GB memory, using *MATLAB* R2021b).

3 Conclusions

In this work, we present a method that allows the calculation of the measurable bandwidth in THz spectrum through smoothing and subsequent multiplication by ω^n, which generates minima at the frequencies where the spectral amplitude decays to the amplitude of noise floor. By analyzing different spectra, it was shown that for values of $w_s = 6\%$ and $n = 3$ it is possible to calculate the cutoff frequencies in most cases. However, in spectra that present features that propitiate a bad calculation of the cutoff frequencies, it is advisable to use values of w_s higher than 9%. It is important to clarify that these parameters are not fixed and can be adjusted by the user in order for the method to work properly according to the data characteristics and THz system used.

Furthermore, the method can be applied to any signal from any THz-TDS system, since this method analyzes the spectrum itself and does not depend on either the system configuration nor the optical parameters of the sample. This was demonstrated by analyzing spectra of out-of-focus samples, with spectral fingerprints and narrow

measurable bandwidths. Additionally, computational time is not a problem by using a very simple smoothing method and by multiplying and finding minima.

The code, which is the central part of this contribution, is provided below as Additional Material. The data sets used for testing the code can be obtained from the corresponding author upon reasonable request.

Appendix A: Additional material (MATLAB/Octave code)

We provide MATLAB/Octave functions that implement the algorithm presented in this article. The user-provided variables or parameters are as follows: t the time-vector for the time-domain data; Et the electric field waveform for the time-domain data; n_{smooth} the fraction (between 0.0 and 1.0) of the frequency domain points for the width of the smoothing window, and p the power of the frequency factor referred to as n in Section 2.

```
function [w_c,Amp_noisef,w,Ef]= get_noise(t,Et,n_smooth,p)
    w=[1:length(t)]./(t(end)-t(1)); % Natural frequency
    Ef=abs(fft(Et));    % Spectrum

    window=round(length(w) .* n_smooth);    % Window
    [out_f,out_Ef]=SmoothingSpec(w,Ef,window);  % Smoothing
    w0=linspace(1,5,length(out_f));
    Ef_cube=out_Ef.*(w0.^p);     % Smooth spectra times w^p

    [M3,iM3]=max(out_Ef);
    [m3,im3]=min(Ef_cube(iM3+1:end));   % Find min
    Amp_noisef=out_Ef(im3+iM3); % Noise floor amplitude
    w_c=out_f(im3+iM3); % Cuttoff frequency
end
```

```
function [out_w,out_Ef]=SmoothingSpec(w,Ef,window)
    if rem(length(Ef),2)==0 % Case even data size
        sz=size(Ef);
    else
        sz=size(Ef)+1; % Case odd data size
    end

    z=sz(1)/2;  % Keep positive frequencies

    for i=1:z-window
        out_Ef(i)=mean(Ef(i:i+window)); % Smooth
    end

    out_w=w(1:z-window);    % Length cut for w
end
```

Author Contribution E.C.C. proposed the study with support from R.C.J.; R.C.J and E.C.C. supervised the investigation; E.S.R.R. wrote, tested, debugged the various versions of the code with support from E.C.C. and R.C.J.; E.S.R.R. performed the majority of the measurements used and obtained examples from collaborators acquired with a variety of spectrometers; E.S.R.R. wrote the manuscript. All authors reviewed the manuscript and contributed with comments.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors acknowledge the financial support of CONACYT through grants 252939, 255114, 280392, and 294440 and scholarship 931264. ECC also thanks the support from the Alexander von Humboldt Foundation through an Experienced Research Fellowship.

Availability of Data and Materials The code, which is the central part of this contribution, is provided as Additional Material. The data sets used for testing the code can be obtained from the corresponding author upon reasonable request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Sibo Hao, Haochong Huang, Yuanyuan Ma, Shangjian Liu, Zili Zhang, and Zhiyuan Zheng. Characterizations of the calamine tablets by terahertz time-domain spectroscopy. *Optik*, 187:278–284, 2019.
2. Keir N. Murphy, Mira Naftaly, Alison Nordon, and Daniel Markl. Polymer pellet fabrication for accurate thz-tds measurements. *Applied Sciences*, 12(7), 2022.
3. Yu Heng Tao, Anthony J. Fitzgerald, and Vincent P. Wallace. Non-contact, non-destructive testing in various industrial sectors with terahertz technology. *Sensors*, 20(3), 2020.
4. Liu Yu, Liu Hao, Tang Meiqiong, Huang Jiaoqi, Liu Wei, Dong Jinying, Chen Xueping, Fu Weiling, and Zhang Yang. The medical application of terahertz technology in non-invasive detection of cells and tissues: opportunities and challenges. *RSC Adv.*, 9:9354–9363, 2019.
5. Goretti G Hernandez-Cardoso, Lauro F Amador-Medina, Gerardo Gutierrez-Torres, Edgar S Reyes-Reyes, César Augusto Benavides Martínez, Cuitlahuac Cardona Espinoza, José Arce Cruz, Irving Salas-Gutierrez, Blanca O Murillo-Ortíz, and Enrique Castro-Camus. Terahertz imaging demonstrates its diagnostic potential and reveals a relationship between cutaneous dehydration and neuropathy for diabetic foot syndrome patients. *Scientific Reports*, 12(1):1–10, 2022.
6. Yuqi Cao, Pingjie Huang, Xian Li, Weiting Ge, Dibo Hou, and Guangxin Zhang. Terahertz spectral unmixing based method for identifying gastric cancer. *Physics in Medicine & Biology*, 63(3):035016, 2018.
7. FEM Lambert, ES Reyes-Reyes, GG Hernandez-Cardoso, AM Gomez-Sepulveda, and E Castro-Camus. In situ determination of the state of conservation of paint coatings on the kiosk of guadalajara using terahertz time-domain spectroscopy. *Journal of Infrared, Millimeter, and Terahertz Waves*, 41(4):355–364, 2020.
8. Tianhua Meng, Rong Huang, Yuhe Lu, Hongmei Liu, Jianguang Ren, Guozhong Zhao, and Weidong Hu. Highly sensitive terahertz non-destructive testing technology for stone relics deterioration prediction using svm-based machine learning models. *Heritage Science*, 9(1):1–9, 2021.

9. Mira Naftaly. Metrology issues and solutions in thz time-domain spectroscopy: Noise, errors, calibration. *IEEE Sensors Journal*, 13(1):8–17, 2012.
10. Mira Naftaly and Richard Dudley. Methodologies for determining the dynamic ranges and signal-to-noise ratios of terahertz time-domain spectrometers. *Optics letters*, 34(8):1213–1215, 2009.
11. Y.S. Lee. *Principles of Terahertz Science and Technology*. Springer New York, NY, 1 edition, 2009.
12. Lei Hou and Wei Shi. An lt-gaas terahertz photoconductive antenna with high emission power, low noise, and good stability. *IEEE Transactions on Electron Devices*, 60(5):1619–1624, 2013.
13. Peter Uhd Jepsen and Bernd M Fischer. Dynamic range in terahertz time-domain transmission and reflection spectroscopy. *Optics letters*, 30(1):29–31, 2005.
14. Seng Hansun. A new approach of moving average method in time series analysis. In: *2013 Conference on New Media Studies (CoNMedia)*, pages 1–4, 2013.