



Multiresolution spectrally-encoded terahertz reflection imaging through a highly diffusive cloak

MAHMOUD E. KHANI,¹  ZACHERY B. HARRIS,¹  MENGKUN LIU,²
AND M. HASSAN ARBAB^{1,*} 

¹Department of Biomedical Engineering, Stony Brook University, 100 Nicolls Rd., Stony Brook, New York 11794, USA

²Department of Physics and Astronomy, Stony Brook University, 100 Nicolls Rd., Stony Brook, New York 11794, USA

*hassan.arbab@stonybrook.edu

Abstract: Turbid media, made of wavelength-scale inhomogeneous particles, can give rise to many significant imaging and spectroscopy challenges. The random variation of the refractive index within such media distorts the spherical wavefronts, resulting in smeared and speckly images. The scattering-induced artifacts can obscure the characteristic spectral fingerprints of the chemicals in a sample. This in turn prevents accurate chemical imaging and characterization of the materials cloaked with a diffusive medium. In this work, we present a novel computational technique for creating spatially- and spectrally-resolved chemical maps through a diffusive cloak using terahertz time-domain spectroscopy. We use the maximal overlap discrete wavelet transform to obtain a multiresolution spectral decomposition of THz extinction coefficients. We define a new spectroscopic concept dubbed the “bimodality coefficient spectrum” using the skewness and kurtosis of the spectral images. We demonstrate that broadband wavelet-based reconstruction of the bimodality coefficient spectrum can resolve the signature resonant frequencies through the scattering layers. Additionally, we show that our approach can achieve spectral images with diffraction-limited resolution. This technique can be used for stand-off characterization of materials and spectral imaging in nondestructive testing and biological applications.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Turbid media, composed of wavelength-scale inhomogeneous particles, can strongly scatter the light, resulting in randomly-distributed rays [1]. Multiple scattering imposes a fundamental limit on the focusing of light due to the optical diffusion [2]. This in turn confines the effective depth of focus to only few multiples of the scattering mean free path [3]. Recent advances in optical imaging, such as wavefront shaping [4] or utilizing the memory effect by angular scanning of the scattered light’s speckle patterns [5], have addressed some of these limitations. However, they require either the measurement of the scattered beam on both sides of the medium [4], or a set of angular-resolved scans of a coherent laser beam [5], thus necessitating long acquisition times and complicated setups. Similarly, optical techniques utilizing only the multiply-scattered photons hold an exciting promise in deep tissue diagnosis [6]. Nonetheless, broadband spectroscopy using multiply-scattered diffuse waves for extraction of intra- or inter-molecular resonance absorption modes has not been demonstrated.

In many crystalline solids and polar gases, the timescales of the weak collective motions and the relaxation phonon modes correspond to the oscillation periods of THz waves [7]. Therefore, THz spectroscopy can provide complementary information to intra-molecular interactions probed by techniques such as infrared spectroscopy [8–12]. However, when a sample is immersed in a turbid medium, although the multiply-scattered waves can still carry the spectral fingerprints,

the scattering artifacts distort or obscure the spectroscopic features [13]. The cross-correlations between diffuse THz waves have been proved to reveal the dynamics of the scatterers [14]. Nevertheless, spectroscopic analysis of dielectric resonant modes using diffuse THz waves has not been possible. Different analytical and numerical approaches, such as dense medium theory [15] or effective medium theory incorporating the Mie scattering amplitudes [16], have been proposed to estimate the THz scattering loss in granular materials and improving the resolution in terahertz time-domain spectroscopy (THz-TDS) imaging [17]. The scattering-induced attenuation in ballistic or quasi-ballistic THz beams in the pressed pellets of α -lactose monohydrate and α -D-glucose has been investigated [18]. A blind demodulation algorithm based on low-rank matrix recovery and alternating minimization has been proposed to remove the sweep distortion originated from densely-layered structures [19]. Signal processing techniques such as wavelet [20] and cepstral analysis [21] have been used to mitigate the scattering artifacts in diffuse THz waves in transmission geometry. Wavelet-based spectral decomposition has been utilized to identify the resonant frequencies of THz reflection coefficients in the presence of severe rough surface scattering [22–24].

Despite many previous attempts, accurate reflection chemical imaging and characterization of materials cloaked with a scattering medium has not been demonstrated. In this work, we present a novel computational technique for creating spatially- and spectrally-resolved chemical maps through a diffusive cloak using THz-TDS. Our approach is based on the bimodality coefficient spectrum, which we define using the skewness and kurtosis of spectral images of a sample. The bimodality coefficient has the unique property of quantifying the modal characteristics of an image histogram distribution. We first decompose the measured reflection spectra using the maximal overlap discrete wavelet transform (MODWT). We reconstruct the spectral images using only a subset of MODWT coefficients such that the bimodality coefficient is maximized at resonant frequencies. To demonstrate the robustness of our technique in reflection imaging of complex samples, we created several spectroscopic Boehler star targets composed of different chemicals with overlapping spectral features. In all the test samples, the scattering effects are designed to be so severe that the raw spectra could not be used to resolve the shape of the Boehler star target or identify the chemicals. However, the experimental results prove that our approach can resolve diffraction-limited spectroscopic images of the sample buried beneath the scattering layer. Importantly, our approach does not rely on hardware modifications to a standard THz spectroscopy setup.

2. Methods

2.1. THz PHASR Scanner

The THz-TDS measurements are obtained using our PHASR (Portable HANdheld Spectral Reflection) Scanner, described in detail previously [25]. In this design, all the optical components shown schematically in Fig. 1(a) are placed inside a 3D-printed fiber-coupled housing (dimensions: $37.3 \times 14 \times 25.1 \text{ cm}^3$), shown in Fig. 1(b). In the PHASR Scanner, the TERA ASOPS (Asynchronous OPTical Sampling) dual-fiber-laser THz spectrometer (Menlo Systems, Inc., Newton, NJ, USA) is incorporated into a handheld, collocated, telecentric imaging system. A THz beam generated by the photoconductive antenna (PCA) in the emitter (E) is collimated using a TPX lens (L1) with 50 mm focal length. The collimated beam is directed towards a gimballed mirror (GM) using a high-resistivity silicon beam splitter (BS). The mirror gimbal is mounted on a two-axis motorized system composed of a goniometer and a rotational stage. It raster scans the collimated beam over the aperture of a custom-made telecentric f - θ lens [26]. Therefore, the focused beam is always in parallel to the optic axis and has a constant focal spot-size of approximately $3.1 \times 2.4 \text{ mm}^2$. The reflected beam retraces the path of the incident beam back to the BS, where it is directed towards the second collimating lens (L2) and is focused on the PCA

inside the detector (D). Previously, we have conducted multiple experiments on the utility of the PHASR Scanner for biomedical sensing applications [27–29].

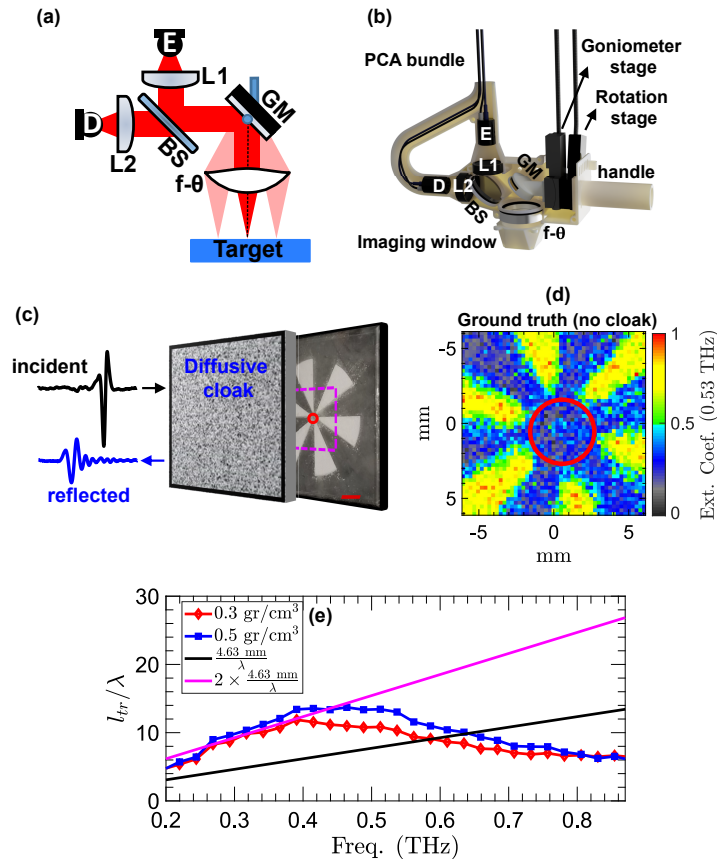


Fig. 1. (a),(b) Schematic and photograph of the optical components inside the PHASR Scanner housing. (c) Boehler star covered by a diffusive scattering cloak with a fixed thickness of about 4.63 mm. Example incident (black) and reflected (blue) THz-TDS waveforms through the scattering layer are shown. (d) Image of the target formed using the extinction coefficients, $\epsilon(f)$, at 0.53 THz without a scattering layer in the beam path (the ground truth). The radius of the red circle in the center is 2.1 mm, corresponding to an 1.11-mm spatial resolution. (e) Wavelength-normalized transport mean free path, l_{tr} , of the scattering layers of 0.3 and 0.5 g/cm^3 volume density. The black and magenta lines show the wavelength-normalized propagation distance of one-way (toward the target) and round-trip (back to the detector) travels through the scattering layer, respectively.

2.2. Imaging target

To assess the imaging performance in spatially resolving the THz resonant modes, the sample chemicals are placed inside the petals of a Boehler star resolution target. Figure 1(c) shows a 3D-printed Boehler star whose petals are filled with α -lactose monohydrate with particle size smaller than $10 \mu\text{m}$. The dashed magenta square in Fig. 1(c) shows a $12 \times 12\text{-mm}^2$ field-of-view. Figure 1(d) illustrates the image of the Boehler star without the presence of any diffusive medium in the beam path. The color axis in Fig. 1(d) is given by the extinction coefficients, $\epsilon(f)$, at 0.53 THz. A THz-TDS measurement is obtained at each $0.25 \times 0.25\text{-mm}^2$ pixel area, resulting in a 48×48 -pixel image of the region enclosed by the magenta square in Fig. 1(c). The petals of

α -lactose monohydrate are clearly discernible in Fig. 1(d). Moreover, the radius of the unresolved circle in the center corresponds to the diffraction-limited imaging resolution at 0.53 THz (focal length = 40 mm, numerical aperture (NA) = 0.40). Figure 1(d) serves as the ground truth for comparison to imaging in the presence of diffusive scattering.

2.3. Turbid scattering medium

The imaging target is cloaked by a scattering layer as shown in Fig. 1(c). The scattering layer is made from low-density polyethylene (LDPE) powders of 180- μm mean particle size. In order to create a scattering layer with known spatial dimensions, we created a 3D-printed housing to be placed on the top of the sample and filled with LDPE. The LDPE powders result in significant Mie scattering over the measurement bandwidth of $f = 0.2\text{--}1.1$ THz. LDPE particles are chosen because they are nonabsorbent in the THz region, and their refractive index is independent of the wavelength. Therefore, the pulse attenuation can only be associated with the scattering. To investigate the robustness of our technique, we study two different cloaks with 0.3 and 0.5 g/cm^3 particle volume density. The thickness of the cloaks, i.e., the beam propagation distance in the turbid medium, is kept the same. In electromagnetic scattering theory, traveling a distance longer than the Boltzmann transport mean free path, l_{tr} , randomizes the beam into ballistic and diffusive fields [30]. Given that the density of LDPE powder is approximately equal to 0.91, we can determine the volume density of the scattering layer by measuring the weight of the particles beforehand. The scattering mean free path, l_{s} , modifies a coherent propagation according to [31],

$$|E_{\text{coh}}(\omega)| = \exp[-z/(2l_{\text{s}}(\omega))]\exp[-\alpha(\omega)z/2]|E_{\text{inc}}(\omega)|, \quad (1)$$

where $|E_{\text{inc}}|$ and $|E_{\text{coh}}|$ are the spectral amplitude of the incident and output beams, respectively, α is the absorption coefficient, and z is the propagation distance. The l_{tr} and l_{s} are related by $l_{\text{s}}/l_{\text{tr}} = 1 - \langle \cos\theta \rangle$, where $\langle \cos\theta \rangle$ is the average cosine of the angles of scattering [32]. The wavelength-normalized l_{tr} of the LDPE cloaks with 0.3 and 0.5 g/cm^3 volume density is given in Fig. 1(e). The l_{tr} at each wavelength is normalized by its corresponding wavelength. The l_{tr} values are measured using a Boehler star target of petals filled with high-density polyethylene (HDPE) and placed underneath the scattering layers. The l_{tr} in both scattering regimes is approximately an order of magnitude larger than the wavelength at the peak wavelength of the THz emissions. The black and magenta lines in Fig. 1(e) respectively show the wavelength-normalized propagation distance of one-way (toward the target) and round-trip (back to the detector) travels through the scattering layer. The 4.63-mm thickness of the scattering layers is chosen such that the round-trip propagation distance is longer than l_{tr} at the resonant frequencies of the several test materials above 0.44 THz.

Figure 2(a) shows the image of the Boehler star in the presence of the cloak with 0.3 g/cm^3 volume density. The color axis represents the $\varepsilon(f)$ at 0.53 THz. It is evident that the petals of α -lactose are not resolved in Fig. 2(a). Here, because a measured ε is composed of the absorption coefficient of the sample, the scattering-induced attenuation, and the Mie scattering-induced spectral artifacts, the petals of α -lactose are obscured. Spatial averaging over disjointed measurements is considered the standard approach to mitigate the scattering effects in THz spectroscopy [21,33]. However, it cannot be used in imaging applications and for characterizing samples made from multiple chemicals when the area covered by each material is not known in advance. The multiresolution spectral analysis approach presented here can mitigate the scattering effects without utilizing spatial averaging. This method relies on decomposition of the measured extinction coefficients using the wavelet transform and reconstruction from wavelet coefficients at specifically-chosen frequency scales. Moreover, unlike the cepstral filtering technique for mitigating the THz scattering artifacts, which requires *a priori* knowledge on the characteristics of the scattering particles to design a low- or band-pass filter [21,34], our technique is robust in that

its implementation is independent of the scattering medium or the properties of the characteristic spectral resonance features of the sample materials.

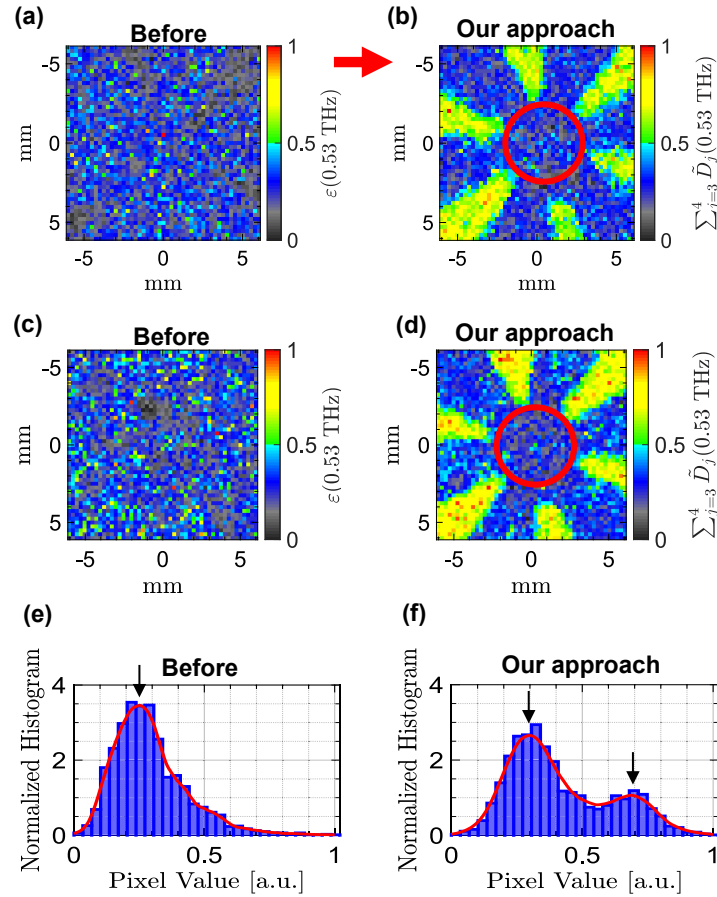


Fig. 2. (a) Image of the α -lactose Boehler star covered by a scattering layer having a 0.3 g/cm^3 volume density. The color axis is given by the measured $\varepsilon(f)$ at 0.53 THz . (b) Image of the same sample formed using $\sum_{j=3}^4 D_j(0.53 \text{ THz})$ as the color map. (c),(d) Similar to (a),(b) for the scattering layer with 0.5 g/cm^3 volume density. (e),(f) Normalized histograms of the images in (a) and (b), yielding the bimodality coefficient values of $\beta = 0.46 < 5/9$ and $\beta = 0.56 > 5/9$, respectively. A probability distribution function (red line) is fitted to each histogram.

2.4. Wavelet multiresolution analysis of extinction coefficients

The discrete wavelet transform (DWT) decomposes a signal into sets of wavelet and scaling coefficients. Each set describes the localized changes in the signal at a specific variation scale. The scale is an interval of signal over which localized, weighted averages of the signal (i.e., the scaling coefficients) and the differences of those averages (i.e., the wavelet coefficients) are calculated. Application of DWT directly to the extinction coefficients results in wavelet scales corresponding to dyadic increments of the frequency sampling interval given by $\sigma_j = 2^{j-1} \delta f$ [35,36]. Here, σ_j is the wavelet scale of the j^{th} level of DWT. The δf is the frequency-domain sampling interval. Its value depends on the sampling rate of measured THz pulses, f_s , in addition to their size, N , and is given by $\delta f = f_s/N$. For example, given a $\delta f \approx 0.02 \text{ THz}$, the 4^{th} and

5th-level wavelet coefficients correspond to $\sigma_4 = 2^3 \times 0.02 = 0.16$ THz and $\sigma_5 = 2^4 \times 0.02 = 0.32$ THz intervals in the extinction coefficients, respectively. The wavelet multiresolution analysis (MRA) of an extinction coefficient measured at a pixel with spatial coordinates of (x, y) , $\varepsilon_{(x,y)}$, using the maximal overlap DWT (MODWT) at J levels of decomposition is given by [35,36],

$$\varepsilon_{(x,y)}(f) = \tilde{S}_J(f) + \sum_{j=1}^J \tilde{D}_j(f). \quad (2)$$

In Eq. (2), \tilde{S}_J represents the smooth vector of level J and is calculated by,

$$\tilde{S}_J(f) = \sum_k \tilde{g}_J(k) \tilde{v}_{J,(x,y)}((f+k) \bmod N). \quad (3)$$

\tilde{D}_j is the details vector of level j and is given by,

$$\tilde{D}_j(f) = \sum_k \tilde{h}_j(k) \tilde{w}_{j,(x,y)}((f+k) \bmod N). \quad (4)$$

In Eq. (3), \tilde{g}_J and $\tilde{v}_{J,(x,y)}$ are the J^{th} -level scaling filter and scaling coefficients of $\varepsilon_{(x,y)}$, respectively. The coefficients of $\tilde{v}_{J,(x,y)}$ correspond to cepstral content of $\varepsilon_{(x,y)}$ in the quefrequency range of $[0 : 1/(2^J \delta f)]$ [36]. The cepstral content refers to the Fourier transformation of the frequency-domain spectra, which is known as the cepstrum (the anagram of spectrum) in the quefrequency (the anagram of frequency) domain [37]. Similarly, \tilde{h}_j and $\tilde{w}_{j,(x,y)}$ in Eq. (4) are the j^{th} -level wavelet filter and wavelet coefficients. The coefficients of $\tilde{w}_{j,(x,y)}$ correspond to the cepstral content of $\varepsilon_{(x,y)}$ in the quefrequency range of $[1/(2^{j+1} \delta f) : 1/(2^j \delta f)]$ [36]. Importantly, reconstructing $\varepsilon_{(x,y)}$ from \tilde{S}_j and \tilde{D}_j at specific scales can separate the characteristic absorption lines from the scattering-induced attenuation and spectral artifacts. Moreover, the absorption lines of different materials can be separated into \tilde{D}_j of various scales.

Figure 2(b) shows the image of the Bohler star formed using MODWT MRA in the presence of the cloak with 0.3 g/cm^3 volume density. The color axis in Fig. 2(b) is $\sum_{j=3}^4 \tilde{D}_j(0.53 \text{ THz})$, which represents the reconstruction of ε from only the third- and fourth-level wavelet coefficients. The coefficients of $\sum_{j=3}^4 \tilde{D}_j(f)$ are associated with the cepstral content of ε in the quefrequency range of $[1/(2^5 \delta f) : 1/(2^3 \delta f)]$. It can be seen that the petals of α -lactose are clearly resolved in Fig. 2(b). In addition, the spatial resolution is comparable to the image of the sample obtained without a cloak in the beam path, shown in Fig. 1(d). In particular, the diameter of the red circle in the center of Fig. 2(b) is found to be 4.2 mm, corresponding to a 1.1-mm spatial resolution using $\frac{4.2 \times \pi}{2 * N_{\text{petals}}} = 1.1$, where N_{petals} is the number of petals in the target [38]. This resolution is larger than the Rayleigh diffraction limit at 0.53 TH because, due to the use of a solid Bohler star mold, the sample powders are not densely-packed under a hydraulic press as is customary when making pellets. Instead, powders of α -lactose are loosely-packed in each petal, which results in additional light scattering by the extra air voids between the particles. Figure 2(c) shows the image of the Bohler star obtained in the presence of the cloak with 0.5 g/cm^3 volume density. Similar to Fig. 2(a), the color axis in Fig. 2(c) represents $\varepsilon(0.53 \text{ THz})$. Figure 2(d) illustrates the result of reconstruction of ε from $\sum_{j=3}^4 \tilde{D}_j$. It can be noticed that the petals of α -lactose are resolved in the image formed using $\sum_{j=3}^4 \tilde{D}_j(0.53 \text{ THz})$. In this paper, MODWT is calculated using the least asymmetric mother wavelet filter with four vanishing moments, i.e., sym4 or LA(8) mother wavelet [36,39], at $J = 6$ levels of decomposition.

2.5. Bimodality coefficient spectrum

In broadband THz spectral imaging of a mixture sample, images formed at resonance absorption frequencies provide a higher visual contrast [10]. We employ this basic concept to find the

MODWT details coefficients that can separate the resonant spectral features from the scattering artifacts. In the examples of Figs. 2(a)-(d), we use the higher-order image statistics to determine a specific combination of MODWT details vectors. This combination is utilized in the MODWT MRA given by Eq. (2) [20]. In particular, we use the skewness and kurtosis to define a new spectroscopic concept named the bimodality coefficient spectrum. Bimodality coefficient can measure the modal characteristics of an image histogram distribution. Skewness (a measure of histogram asymmetry) and kurtosis (a measure of histogram tailedness) represent the third- and fourth-order standardized moments of an image around its mean, respectively [40]. The skewness of an image with n pixels is given by [40],

$$\gamma = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right]^{3/2}}, \quad (5)$$

while kurtosis is calculated by,

$$\kappa = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right]^2}. \quad (6)$$

Here, x_i is the i th pixel value, and \bar{x} is the average. Accordingly, the bimodality coefficient is calculated by [41],

$$\beta(f) = \frac{\gamma^2(f) + 1}{\kappa(f) + \frac{3(n-1)^2}{(n-2)(n-3)}}. \quad (7)$$

The value of β can only vary in the range of 0–1. The β of a uniform distribution is equal to 5/9 [42]. Values greater than 5/9 indicate a bimodal or multimodal distribution. It has been proved that a high level of bimodality in an image histogram corresponds to a better visual contrast [43]. Therefore, it can be deduced that a spectral image with β greater than 5/9 provides a better contrast on the imaging target. For example, Figs. 2(e),(f) show the normalized histogram distributions of the images in Figs. 2(a),(b), respectively. A probability distribution function (the solid red line) is fitted to each histogram. It can be observed that the histogram in Fig. 2(e) has only one peak (i.e., it's unimodal), which is verified by its bimodality coefficient of $\beta = 0.46 < 5/9$. Conversely, the histogram in Fig. 2(f) has two well-separated peaks (i.e., it's bimodal), resulting in a $\beta = 0.56 > 5/9$.

To find a unique combination of MODWT details coefficients, we calculate the β of the images formed at a resonant frequency using all possible combinations of \tilde{D}_j in the MRA reconstruction of Eq. (2). For example, $\sum_{j=2}^4 \tilde{D}_j$ and $\sum_{j=3}^4 \tilde{D}_j$ are two forms of such combinations. At $J = 6$ levels of decomposition, there are $2^{J+1} - 1 = 127$ unique combinations. Subsequently, we select the combination that yields the image with highest β value to reconstruct the ε .

2.6. Principal component analysis

In hyperspectral imaging, the measured spectra are composed of contributions from disparate substances [44]. Therefore, several spectral unmixing techniques have been proposed to distinguish the constituent materials and their proportions in each spectrum measurement [45]. Principal component analysis (PCA) is an unsupervised spectral unmixing approach and a mathematically optimal way to treat similar problems [45]. We can also treat the problem of extracting the resonance absorption signatures from the ε in the presence of scattering as a spectral unmixing case.

PCA identifies a set of orthogonal axes, onto which the projection of the data conveys the highest variance [46]. These orthogonal axes are known as principal components (PC). The PCs along a desired dimension, here the frequency axis, can be obtained using the eigen-decomposition

of the sample covariance matrix given by [44,46],

$$\text{cov}(E) = \frac{1}{N} \sum_{n=1}^N (E(n) - \mu_E(n))(E(n) - \mu_E(n))^T, \quad (8)$$

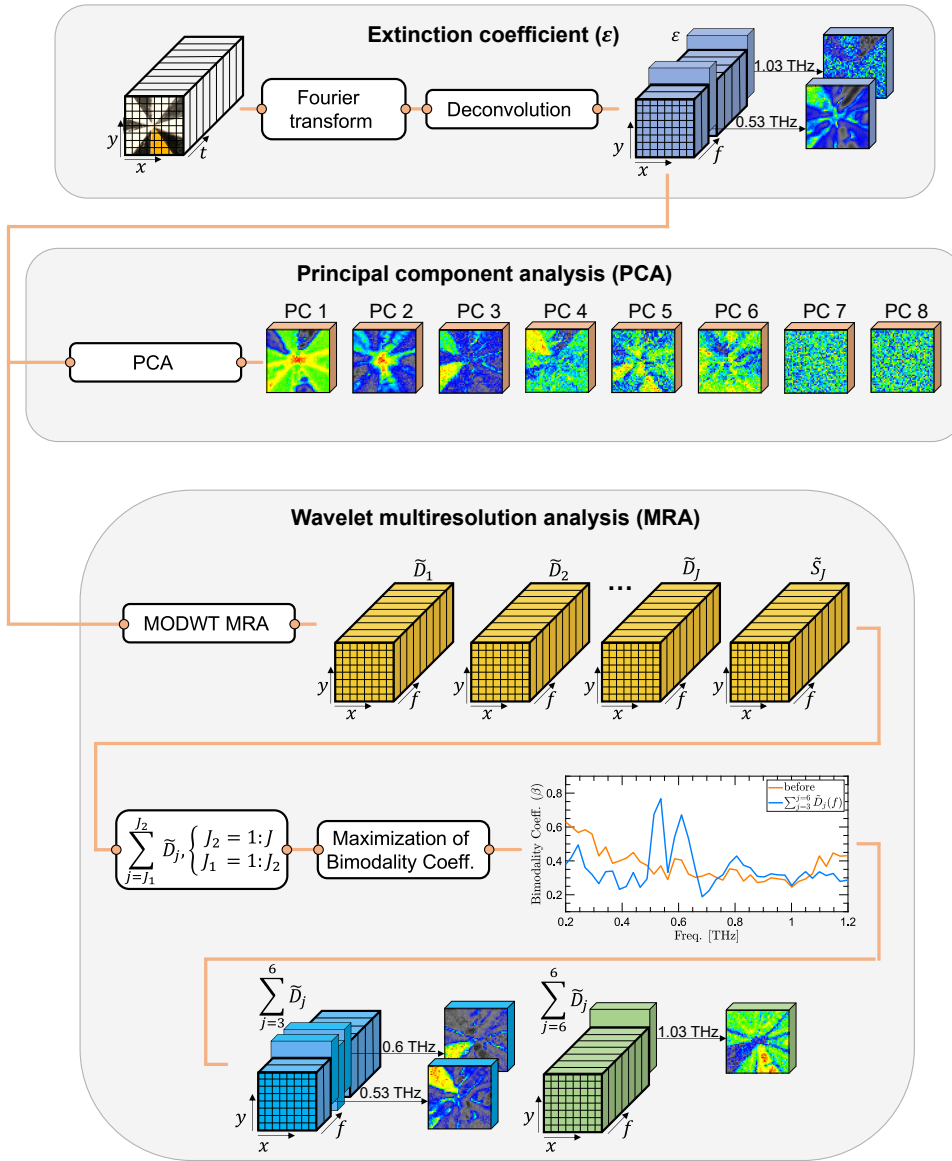


Fig. 3. the computational steps of the proposed MODWT MRA technique and the PCA algorithm. The first row shows that a three-dimensional $\varepsilon_{(x,y)}(f)$ is calculated from the measured time-domain THz pulses. The second row shows that the implementation of PCA on ε fails to preserve the frequency-dependence of ε . The third and fourth rows demonstrate that the frequency information is fully preserved at all stages of the bimodality maximization-based MODWT MRA.

where E is the three-dimensional data cube comprising the images formed using $\varepsilon_{(x,y)}(f_n)$ at N separate spectral bands of f_n . The value of $\mu_E(n)$ represents the average of the image formed at f_n . The eigen-decomposition of the covariance matrix is expressed as,

$$\text{cov}(x) = \mathbf{U}\Sigma\mathbf{U}^T, \quad (9)$$

where \mathbf{U} is composed of the eigen-vectors as its columns. Σ is a diagonal matrix of eigen-values indicating the contribution of the data parallel to their corresponding eigen-vectors in the total variance. Multiplication of E with a subset of \mathbf{U}^T transforms E into a new system of decorrelated (orthogonal) variables oriented along the eigen-vectors. Usually a very small set of PCs capture the majority of the variance.

It should be noted that the characteristic resonances are not necessarily separable using the PCA approach, whereas the MODWT MRA retains the spectral information. Thus, MODWT MRA enables specificity by isolating the petal associated with each resonance into wavelet coefficients at that resonant frequency. The flowchart of Fig. 3 compares the algorithms of MODWT MRA and PCA.

This flowchart emphasizes the unique property of retention of the frequency information by MODWT MRA. The first row shows that a spatio-spectral data cube composed of $\varepsilon_{(x,y)}$ is obtained by Fourier transformation and deconvolution of the measured THz-TDS pulses reflected from the sample. Deconvolution removes the response of the measurement setup to extract the sample response function. The second row illustrates the implementation of the PCA on ε . PCA results in a set of two-dimensional $\text{PC}_{(x,y)}$ which do not preserve the frequency-dependent information. In contrast, the third row shows that the frequency dimension of ε is fully preserved at all stages of the bimodality-based MODWT MRA algorithm. In other words, the MODWT vectors, $\tilde{D}_1, \tilde{D}_2, \dots, \tilde{D}_J$ and \tilde{S}_J , retain the spectral information of the ε in their 3D structure. This property, in turn, enables the identification of specific petals at their resonant frequencies using the bimodality coefficient optimization process.

3. Results

3.1. Boehler star with multiple spectrally-overlapping samples

In this section, we investigate the robustness of the bimodality-based MODWT MRA of THz reflection images in the presence of scattering. Figure 4 shows the obtained experimental results when each of the five petals of a Boehler star is filled by a different chemical powder. Fig. 4(a) shows the sample target made from α -lactose monohydrate, 4-aminobenzoic acid (PABA), 3-hydroxybenzoic acid (3HBA), riboflavin, and high-density polyethylene (HDPE). Figure 4(b) compares the ε of these five chemicals over the frequency range of $f = 0.2 - 1.1$ THz without a scattering layer in the beam path. The distinct characteristic resonant signatures, including α -lactose resonance at 0.53 THz, PABA resonances at 0.6 and 0.8 THz, 3HBA resonance at 0.8 THz, and riboflavin weak resonance at 1.03 THz, are observable in Fig. 4(b). These resonant frequencies are in agreement with the values reported in the literature [47–50]. The extinction coefficients of HDPE (cyan line) and the Boehler star's resin-based building material (red line) do not contain any resonance absorption lines. In addition to the scattering effects described earlier, the overlap between the resonances of PABA and 3HBA with the broadband absorption of the Boehler star resin itself imposes a further computational challenge in resolving the petal associated with each of these chemicals. Furthermore, as the signal-to-noise ratio (SNR) decreases in the higher frequencies, the ε of riboflavin, 3HBA, PABA, α -lactose, and the resin background overlap, thus further obscuring the image of the riboflavin resonance at 1.03 THz. In this example, maximizing the $\beta(f)$ at 0.53 and 0.6 THz over all unique combinations of \tilde{D}_j reveals that $\sum_{j=3}^6 \tilde{D}_j$ yields the highest bimodality value at both resonant frequencies. Figure 4(c) shows the β of the images formed using the ε (red line) in the presence of cloak of 0.3 g/cm^3

volume density, and their reconstruction from $\sum_{j=3}^6 \tilde{D}_j$ (blue line). It can be noticed that in the β of $\sum_{j=3}^6 \tilde{D}_j$, both resonances are identified as the only local maxima with a bimodality value greater than 5/9. The similarity between the shapes of the resonant signatures at 0.53 and 0.6 THz can explain why the same set of \tilde{D}_j is able to extract them at their respective frequencies. The images shown in Figs. 4(d)-(h) are obtained from a 12×12 -mm² area (48×48 -pixels at 0.25-mm pixel size) at the center of the Boehler star. Figure 4(d) shows an image formed by $\sum_{j=3}^6 \tilde{D}_j(f)$ at 0.53 THz with a $\beta = 0.77$. It can be seen that the petal of α -lactose has a higher contrast in comparison to rest of the sample. Figure 4(e) shows an image formed by the same combination at 0.6 THz, yielding a $\beta = 0.75$. This image provides a high contrast by identifying only the petal of PABA. Figures 4(f)-(h) show the images formed by MODWT details vectors, including \tilde{D}_4 (0.8 THz), \tilde{D}_5 (0.8 THz), and \tilde{D}_6 (1.03 THz), respectively. Figure 4(f) demonstrates that $\tilde{D}_4(f)$ can extract and isolate the broader resonant feature of PABA at 0.8 THz, resulting in a higher contrast over the petal of PABA, which is comparable to the petal resolved by $\sum_{j=3}^6 \tilde{D}_j(f)$ at 0.6 THz. In contrast, Fig. 4(e) exhibits that $\tilde{D}_5(f)$ at 0.8 THz yields a higher contrast on the petal of 3HBA. In other words, \tilde{D}_j of different levels can be employed to distinguish between materials of overlapping resonant signatures at the same resonant frequency. It should be noted that because of the shared similarity in shape between the ε of 3HBA and the Boehler star resin material, as shown in Fig. 4(b), the difference between the 3HBA petal and the background material is modest in Fig. 4(g). Finally, Fig. 4(h) shows an image formed by $\tilde{D}_6(f)$ at 1.03 THz, resolving the petal of riboflavin.

The experimental results given by Figs. 4(d)-(h) show that different resonance absorption lines are extracted by MODWT details coefficients of different scales. For example, the image formed using \tilde{D}_4 (0.8 THz) resolves the petal of PABA, whereas the image created by \tilde{D}_6 (1.03 THz) captures the area of riboflavin. This observation can be explained by the difference between the spectral shapes of different resonant signatures. Because the resonance of PABA at 0.8 THz is both sharper and narrower in comparison to the resonance of riboflavin at 1.03 THz, it is extracted by lower-level MODWT details vectors, which correspond to a higher-frequency content in ε . On the contrary, the broader resonant signature of riboflavin has a lower-frequency content. Therefore, it is extracted by MODWT details vector of a higher level.

3.2. Comparison with alternative hyperspectral technique

Figure 5 shows the results of spectral unmixing using the PCA of ε . Fig. 5(a) shows that each PC is formed by a linear, weighted combination of all spectral images along the measurement bandwidth. The weight associated with each frequency in constructing the PCs is determined by the elements of the eigen-vectors of $\text{cov}(E)$. Figure 5(b) gives the eigen-vectors of the first four PCs in the order of capturing a higher variance in the original set. It can be seen that there is no eminent resonant feature in the eigen-vectors of the first (red line) and second (blue line) PCs. However, the eigen-vector of the third PC (green line) has two distinct peaks at 0.53 and 0.6 THz, revealing the resonant frequencies of α -lactose and PABA. Similarly, in the fourth eigen-vector (purple line) the image at 0.53 THz has the greatest value. Therefore, the fourth PC is mainly associated with the resonance of α -lactose.

In comparison to the results obtained using MODWT MRA, our new approach can identify all constituent chemicals in the presence of scattering, albeit with varying levels of success. In contrast, only the third and fourth PCs are able to identify α -lactose and PABA, with loss of specificity between the two. In other words, because each PC is formed by a weighted superposition of all spectral images at different frequencies, it contains contributions from all frequency components. Therefore, the characteristic resonances are not necessarily separable using the PCA approach. In contrast, the wavelet MRA retains the frequency information. Thus, it enables specificity by isolating the petal associated with each resonance into wavelet coefficients at that resonant frequency. This comparison reveals that the scale-frequency decomposition using

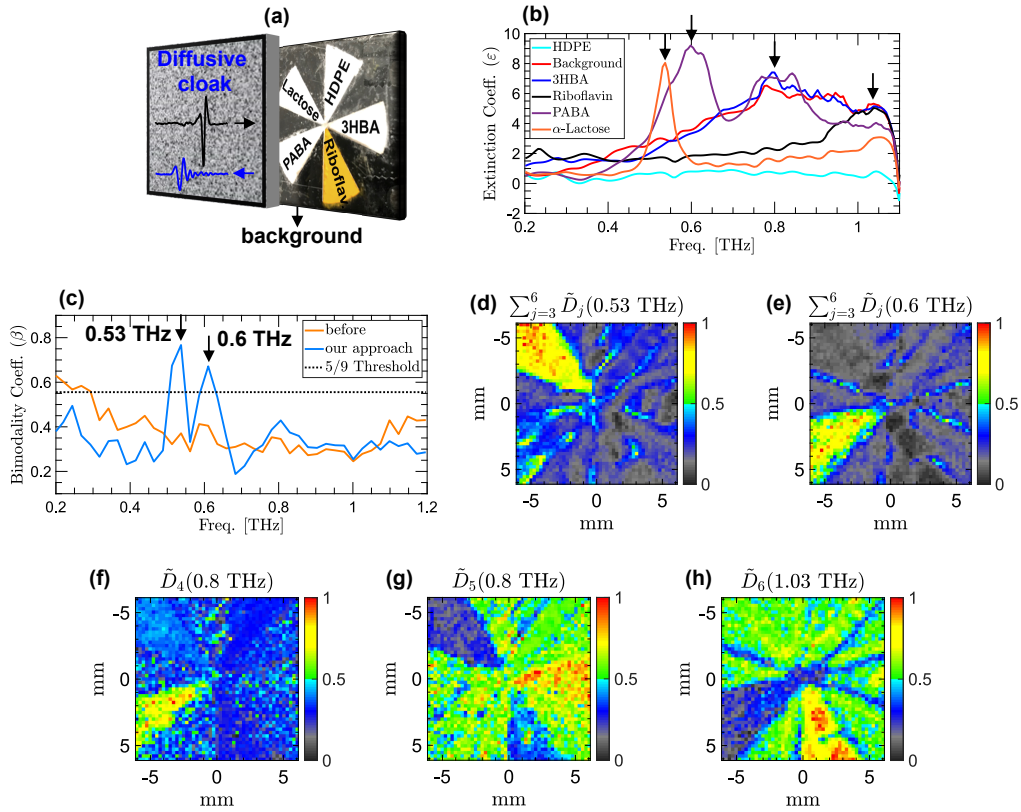


Fig. 4. (a) Boehler star target where each petal is filled by a different chemical powder, including α -lactose, HDPE, 3HBA, riboflavin, and PABA. The Boehler star is covered with a diffusive cloak with 0.3 g/cm^3 volume density. (b) ϵ of the chemicals measured without a scattering layer in the beam path. The black arrows mark the resonant frequencies. (c) β of the images formed using ϵ (orange line) and the images formed using $\sum_{j=3}^6 \tilde{D}_j$ (blue line). (d)-(h) Images formed using different combinations of $\tilde{D}_j(f)$ at the resonant frequencies, including $\sum_{j=3}^6 \tilde{D}_j(0.53 \text{ THz})$ in (d), $\sum_{j=3}^6 \tilde{D}_j(0.6 \text{ THz})$ in (e), $\tilde{D}_4(0.8 \text{ THz})$ in (f), $\tilde{D}_5(0.8 \text{ THz})$ in (g), and $\tilde{D}_6(1.03 \text{ THz})$ in (h).

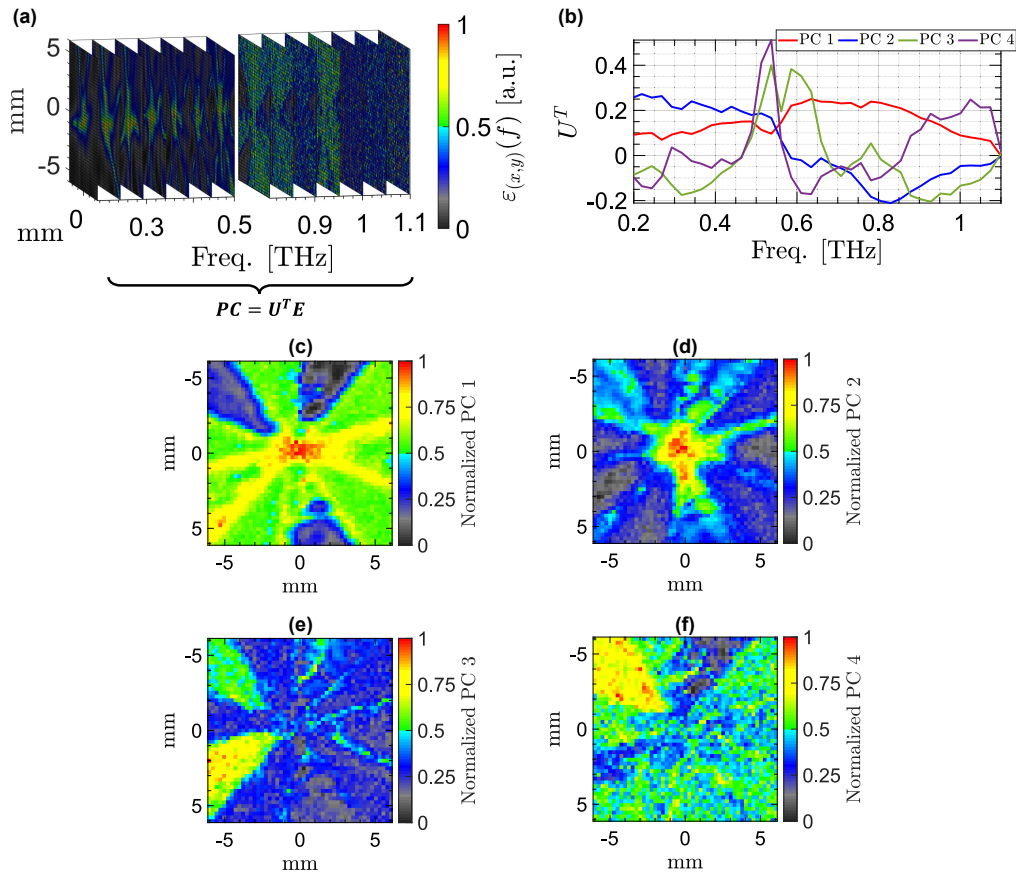


Fig. 5. (a) Forming a PC using a linear, weighted combination of the spectral images along the measurement bandwidth. The weight of each image is determined by the eigen-vectors of $\text{cov}(E)$. (b) Eigen-vectors of the first four PC bands in the order of capturing a higher variance. (c)-(f) Images of the first four PCs. The third and fourth PCs yield the petals of α -lactose and PABA.

wavelet transform is advantageous in studying the characteristic resonant frequencies of mixture samples.

3.3. Performance evaluation

In this section, we describe a quantitative approach for comparison between the images of the mixture Boehler star formed using ε , \tilde{D}_j , and the PCs. We use the image of the Boehler star obtained in the absence of any cloaks in the beam path, shown in Fig. 6(a), as the reference for delineating the boundary pixels of each petal. The color axis in Fig. 6(a) represents the peak-to-peak amplitude of measured THz signals. Figure 6(b) labels the pixels of chemical using a different color and excludes the diffraction-limited resolution circle at the center. The Boehler star background material is set to black, α -lactose to purple, HDPE to blue, 3HBA to green, riboflavin to yellow, and PABA to red. The colored region associated with each material forms the label matrix, as the ground truth, for calculating the mean squared error (MSE) at the resonant frequencies. For example, in the label matrix of α -lactose, all the pixels in the purple region of Fig. 6(b) are equal to one, while every pixel elsewhere equals zero. In order to compare the images formed by ε , \tilde{D}_j , and the PCs to their corresponding label matrix, we first convert

them to a set of binary images. In a binary image, a pixel value is set to one if it is assigned to a given material and is set to zero otherwise. We use an image segmentation approach using a variable threshold over the color axis to form the binary images. The variable threshold is used to create a receiver operating characteristic (ROC) curve for each image [51]. In particular, a new binary image is formed at each cut-off threshold value and is compared to the label matrix for calculating the true positive rate (TPR) and false positive rate (FPR). The ROC curve shows the calculation of TPR or sensitivity on the y-axis versus the FPR or (1-specificity) on the x-axis. A high sensitivity indicates a higher rate of correct assignment of pixels belonging to a material at a specific resonant frequency. A high specificity (1-FPR), on the other hand, implies a higher rate of correct assignment of pixels belonging to the rest of the materials. In each ROC plot in Figs. 6(c)-(g), the diagonal dashed lines show the ROC curve of a random classifier with no

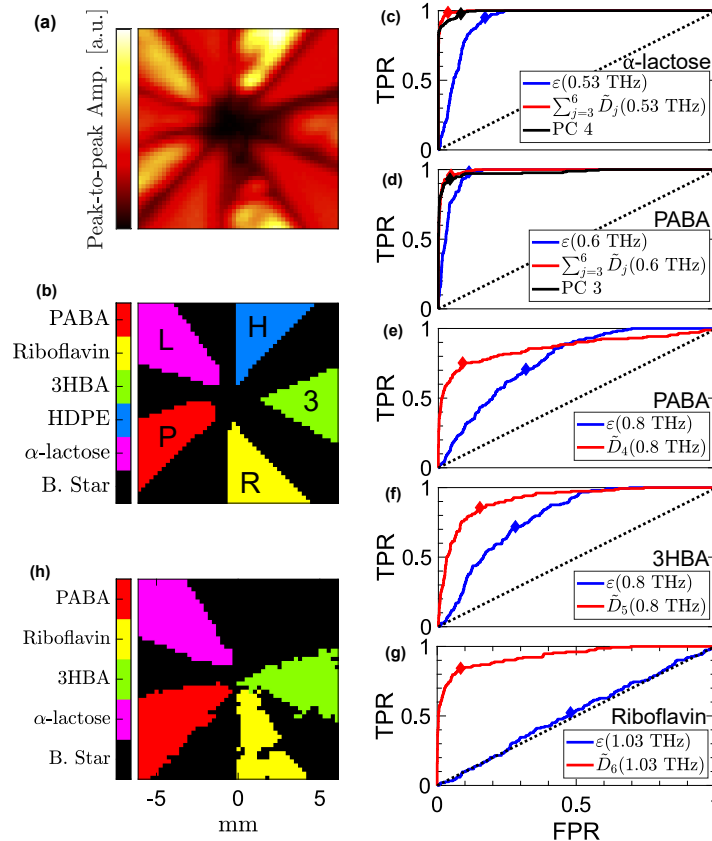


Fig. 6. (a) Image of the mixture sample obtained without the presence of a cloak in the beam path. The image in (a) is formed using the peak-to-peak amplitude of THz signals measured over a $12 \times 12 \text{ mm}^2$ area at the center of the Boehler star. (b) Superposition of the sample label matrices can delineate the pixels associated with each chemical using a different color. (c) ROC curves formed using the images of $\varepsilon(0.53 \text{ THz})$, $\sum_{j=3}^6 \tilde{D}_j(0.53 \text{ THz})$, and PC 4. (d) ROC curves formed using the images of $\varepsilon(0.6 \text{ THz})$, $\sum_{j=3}^6 \tilde{D}_j(0.6 \text{ THz})$, and PC 3. (e)-(g) ROC curves of the images formed using ε and \tilde{D}_j of the materials detected in Figs. 2(d)-(f), respectively. The diamond markers identify the threshold values maximizing the harmonic mean of sensitivity and specificity. (h) Final aggregate chemical map formed by a weighted combination of the thresholded binary images of MODWT details vectors at the resonant frequencies.

predictive power. A higher area under the ROC curve (ROC-AUC) indicates a greater accuracy in differentiation between the pixels associated with a specific resonant mode and the rest of the materials.

Figures 6(c)-(g) show the ROC curves of the images at 0.53 THz (α -lactose), 0.6 THz (PABA), 0.8 THz (PABA), 0.8 THz (3HBA), and 1.03 THz (riboflavin), respectively. The solid blue, red, and black lines show the ROC curves obtained using the images formed by ε , \tilde{D}_j , and PCA, respectively. It should be noted that in the images of PCs, only the petals of α -lactose and PABA are resolved. Therefore, we only include the ROC curves of PC 3 and PC 4 in Figs. 6(c),(d) to compare them with the images obtained by $\varepsilon(f)$ and $\tilde{D}_j(f)$ at 0.53 and 0.6 THz. It can be seen that the ROC-AUC value is always higher for images formed using \tilde{D}_j in comparison to the other approaches. Table 1 gives the MSE values at each resonant frequency for the images formed using ε and \tilde{D}_j . The MSE of the MODWT details vectors is always lower compared to the extinction coefficients, corresponding to an error reduction of 82% at 0.53 THz, 12% at 0.6 THz, 71% at 0.8 THz (PABA), 56% at 0.8 THz (3HBA), and 71 % at 1.03 THz. Because of the inherent trade-off between sensitivity and specificity, to generate a final aggregate chemical map, the optimum threshold values are selected to maximize the harmonic mean of sensitivity and specificity. The resultant cutoff threshold for each chemical are shown using the diamond markers in Figs. 6(c)-(g), In Fig. 6(h), we show the aggregate chemical map formed by a weighted combination of the thresholded binary images formed by \tilde{D}_j . These images include $\sum_{j=3}^6 \tilde{D}_j(0.53$ THz), $\sum_{j=3}^6 \tilde{D}_j(0.6$ THz), $\tilde{D}_5(0.8$ THz), and $\tilde{D}_6(1.03$ THz). Comparing Fig. 6(h) to the assigned labels in Fig. 6(b) reveals that the chemical map resultant of the proposed method provides a higher spatial resolution towards the center of the Boehler star. In other words, areas not previously identifiable using the peak-to-peak amplitude or the raw ε are resolved accurately using our proposed approach. One key limitation of this technique is that materials such as HDPE and the resin background used in the Boehler star structure cannot be differentiated because their dielectric functions do not possess any signature resonant frequencies that can be used to encode the reflected spectra. Future directions of this work include employing time-resolved, polarization-sensitive THz-TDP measurements, which can be used to characterize the scattering events [52]. Furthermore, the use of broadband THz sources, such as air-plasma filaments [53], can enable the implementation of the proposed bimodality coefficient technique at higher frequencies for the identification of even more complex samples.

Table 1. MSE between THz images and the ground truth

Resonant frequency	Extinction coeff.	Proposed method
0.53 THz (α -lactose)	0.17	0.03 (Fig. 4(d))
0.6 THz (PABA)	0.08	0.07 (Fig. 4(e))
0.8 THz (PABA)	0.38	0.11 (Fig. 4(f))
0.8 THz (3HBA)	0.34	0.15 (Fig. 4(g))
1.03 THz (Riboflavin)	0.52	0.15 (Fig. 4(h))

4. Conclusion

We presented experimental results to demonstrate THz spectroscopic reflection imaging through a highly-diffusive cloak. We showed that the scattering effects of a diffusive cloak can obscure the characteristic resonance absorption lines of interest. We presented a multiresolution spectral analysis technique using the wavelet decomposition and reconstruction of THz extinction spectra to mitigate those artifacts. We identified the specific wavelet scales that can extract the obscured resonances of several test materials in a complex sample using the higher-order statistics of the spectral images. In particular, we used the skewness and kurtosis of the spectral images to

define a new spectroscopic concept named the bimodality coefficient spectrum. We showed that by using a unique combination of the wavelet details vectors in reconstructing the extinction coefficients, the resonant frequencies can be identified as the local maxima above a defined 5/9 threshold in the bimodality coefficient spectrum. Additionally, we showed that overlapping signature resonances can be separated into wavelet coefficients of different scales. We evaluated the performance of our technique in the identification and labeling the pixels associated with each chemical. We showed that our approach outperforms the raw extinction spectra and principal component analysis. Finally, we demonstrated that our proposed technique resulted in a chemical map with a higher spatial resolution, despite the presence of a scattering layer, as compared to the images formed using raw THz spectra in the absence of a diffusive cloak.

Funding. Stony Brook University; National Institute of General Medical Sciences (GM112693).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. D. S. Wiersma, "Disordered photonics," *Nat. Photonics* **7**(3), 188–196 (2013).
2. A. P. Mosk, A. Lagendijk, G. Leroose, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," *Nat. Photonics* **6**(5), 283–292 (2012).
3. S. Jeong, Y.-R. Lee, W. Choi, S. Kang, J. H. Hong, J.-S. Park, Y.-S. Lim, H.-G. Park, and W. Choi, "Focusing of light energy inside a scattering medium by controlling the time-gated multiple light scattering," *Nat. Photonics* **12**(5), 277–283 (2018).
4. I. M. Vellekoop, A. Lagendijk, and A. Mosk, "Exploiting disorder for perfect focusing," *Nat. Photonics* **4**(5), 320–322 (2010).
5. J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature* **491**(7423), 232–234 (2012).
6. Z. A. Steelman, D. S. Ho, K. K. Chu, and A. Wax, "Light-scattering methods for tissue diagnosis," *Optica* **6**(4), 479–489 (2019).
7. X. C. Zhang, A. Shkurinov, and Y. Zhang, "Extreme terahertz science," *Nat. Photonics* **11**(1), 16–18 (2017).
8. K. Kawase, Y. Ogawa, Y. Watanabe, and H. Inoue, "Non-destructive terahertz imaging of illicit drugs using spectral fingerprints," *Opt. Express* **11**(20), 2549–2554 (2003).
9. Y. Shen, T. Lo, P. Taday, B. Cole, W. Tribe, and M. Kemp, "Detection and identification of explosives using terahertz pulsed spectroscopic imaging," *Appl. Phys. Lett.* **86**(24), 241116 (2005).
10. H. Zhong, A. Redo-Sanchez, and X.-C. Zhang, "Identification and classification of chemicals using terahertz reflective spectroscopic focal-plane imaging system," *Opt. Express* **14**(20), 9130–9141 (2006).
11. L. Zhang, H. Zhong, C. Deng, C. Zhang, and Y. Zhao, "Terahertz wave reference-free phase imaging for identification of explosives," *Appl. Phys. Lett.* **92**(9), 091117 (2008).
12. Y.-D. Hsieh, S. Nakamura, D. G. Abdelsalam, T. Minamikawa, Y. Mizutani, H. Yamamoto, T. Iwata, F. Hindle, and T. Yasui, "Dynamic terahertz spectroscopy of gas molecules mixed with unwanted aerosol under atmospheric pressure using fibre-based asynchronous-optical-sampling terahertz time-domain spectroscopy," *Sci. Rep.* **6**(1), 28114 (2016).
13. P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas, and P. Gardner, "Resonant mie scattering in infrared spectroscopy of biological materials—understanding the 'dispersion artefact'," *Analyst* **134**(8), 1586–1593 (2009).
14. Z. Jian, J. Pearce, and D. M. Mittleman, "Characterizing individual scattering events by measuring the amplitude and phase of the electric field diffusing through a random medium," *Phys. Rev. Lett.* **91**(3), 033903 (2003).
15. L. M. Zurk, B. Orłowski, D. P. Winebrenner, E. I. Thorsos, M. R. Leahy-Hoppa, and L. M. Hayden, "Terahertz scattering from granular material," *J. Opt. Soc. Am. B* **24**(9), 2238–2243 (2007).
16. M. Kaushik, B. W.-H. Ng, B. M. Fischer, and D. Abbott, "Terahertz scattering by granular composite materials: An effective medium theory," *Appl. Phys. Lett.* **100**(1), 011107 (2012).
17. N. S. Balbekin, M. S. Kulya, A. V. Belashov, A. Gorodetsky, and N. V. Petrov, "Increasing the resolution of the reconstructed image in terahertz pulse time-domain holography," *Sci. Rep.* **9**(1), 180 (2019).
18. M. Kaushik, B. W.-H. Ng, B. M. Fischer, and D. Abbott, "Terahertz fingerprinting in presence of quasi-ballistic scattering," *Appl. Phys. Lett.* **101**(6), 061108 (2012).
19. A. Aghasi, B. Heshmat, A. Redo-Sanchez, J. Romberg, and R. Raskar, "Sweep distortion removal from terahertz images via blind demodulation," *Optica* **3**(7), 754–762 (2016).
20. M. E. Khani, O. B. Osman, and M. H. Arbab, "Diffuse terahertz spectroscopy in turbid media using a wavelet-based bimodality spectral analysis," *Sci. Rep.* **11**(1), 22804 (2021).
21. O. B. Osman and M. H. Arbab, "Mitigating the effects of granular scattering using cepstrum analysis in terahertz time-domain spectral imaging," *PLoS One* **14**(5), e0216952 (2019).

22. M. H. Arbab, D. P. Winebrenner, E. I. Thorsos, and A. Chen, "Retrieval of terahertz spectroscopic signatures in the presence of rough surface scattering using wavelet methods," *Appl. Phys. Lett.* **97**(18), 181903 (2010).
23. M. E. Khani, D. P. Winebrenner, and M. H. Arbab, "Phase function effects on identification of terahertz spectral signatures using the discrete wavelet transform," *IEEE Trans. THz Sci. Technol.* **10**(6), 656–666 (2020).
24. M. E. Khani and M. H. Arbab, "Chemical identification in the specular and off-specular rough-surface scattered terahertz spectra using wavelet shrinkage," *IEEE Access* **9**, 29746–29754 (2021).
25. Z. B. Harris, M. E. Khani, and M. H. Arbab, "Terahertz Portable Handheld Spectral Reflection (PHASR) scanner," *IEEE Access* **8**, 228024–228031 (2020).
26. Z. B. Harris, S. Katletz, M. E. Khani, A. Virk, and M. H. Arbab, "Design and characterization of telecentric f - θ scanning lenses for broadband terahertz frequency systems," *AIP Adv.* **10**(12), 125313 (2020).
27. A. Chen, A. Virk, Z. Harris, A. Abazari, R. Honkanen, and M. H. Arbab, "Non-contact terahertz spectroscopic measurement of the intraocular pressure through corneal hydration mapping," *Biomed. Opt. Express* **12**(6), 3438–3449 (2021).
28. M. E. Khani, Z. B. Harris, O. B. Osman, J. W. Zhou, A. Chen, A. J. Singer, and M. H. Arbab, "Supervised machine learning for automatic classification of in vivo scald and contact burn injuries using the terahertz Portable Handheld Spectral Reflection (PHASR) Scanner," *Sci. Rep.* **12**(1), 5096 (2022).
29. O. B. Osman, Z. B. Harris, J. W. Zhou, M. E. Khani, A. J. Singer, and M. H. Arbab, "In vivo assessment and monitoring of burn wounds using a handheld terahertz hyperspectral scanner," *Adv. Photonics Res.* **3**(5), 2100095 (2022).
30. F. Frezza, F. Mangini, and N. Tedeschi, "Introduction to electromagnetic scattering: tutorial," *J. Opt. Soc. Am. A* **35**(1), 163–173 (2018).
31. J. Pearce and D. M. Mittleman, "Propagation of single-cycle terahertz pulses in random media," *Opt. Lett.* **26**(24), 2002–2004 (2001).
32. A. Ishimaru, *Wave Propagation and Scattering in Random Media*, vol. 2 (Academic Press, 1978).
33. Y. Shen, P. Taday, and M. Pepper, "Elimination of scattering effects in spectral measurement of granulated materials using terahertz pulsed spectroscopy," *Appl. Phys. Lett.* **92**(5), 051103 (2008).
34. S. Schecklman, L. M. Zurk, S. Henry, and G. P. Kniffin, "Terahertz material detection from diffuse surface scattering," *J. Appl. Phys.* **109**(9), 094902 (2011).
35. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.* **11**(7), 674–693 (1989).
36. D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, vol. 4 (Cambridge University, 2000).
37. A. V. Oppenheim and R. W. Schaffer, "From frequency to queffrequency: A history of the cepstrum," *IEEE Signal Process. Mag.* **21**(5), 95–106 (2004).
38. Z. B. Harris, A. Virk, M. E. Khani, and M. H. Arbab, "Terahertz time-domain spectral imaging using telecentric beam steering and an f - θ scanning lens: distortion compensation and determination of resolution limits," *Opt. Express* **28**(18), 26612–26622 (2020).
39. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, 1992).
40. D. N. Joanes and C. A. Gill, "Comparing measures of sample skewness and kurtosis," *J. Royal Statistical Soc. D.* **47**, 183–189 (1998).
41. J. B. Freeman and R. Dale, "Assessing bimodality to detect the presence of a dual cognitive process," *Behav. Res. Methods* **45**(1), 83–97 (2013).
42. R. Pfister, K. A. Schwarz, M. Janczyk, R. Dale, and J. Freeman, "Good things peak in pairs: a note on the bimodality coefficient," *Front. Psychol.* **4**, 700 (2013).
43. D. M. Kim, H. Zhang, H. Zhou, T. Du, Q. Wu, T. C. Mockler, and M. Y. Berezin, "Highly sensitive image-derived indices of water-stressed plants using hyperspectral imaging in swir and histogram analysis," *Sci. Rep.* **5**(1), 15919 (2015).
44. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.* **19**(1), 44–57 (2002).
45. C. Quintano, A. Fernández-Manso, Y. E. Shimabukuro, and G. Pereira, "Spectral unmixing," *Int. J. Remote Sens.* **33**(17), 5307–5340 (2012).
46. T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, vol. 2 (Springer, 2009).
47. E. Brown, J. Bjarnason, A. Fedor, and T. Kortner, "On the strong and narrow absorption signature in lactose at 0.53 thz," *Appl. Phys. Lett.* **90**(6), 061908 (2007).
48. Q. Song, Y. Zhao, R. Zhang, X. Liu, L. Dong, and W. Xu, "Measurement and DFT calculation on terahertz spectroscopy of 4-aminobenzoic acid," *J. Infrared, Millimeter, Terahertz Waves* **31**, 310–318 (2009).
49. Q. Wang, J. Xue, Z. Hong, and Y. Du, "Pharmaceutical cocrystal formation of pyrazinamide with 3-hydroxybenzoic acid: a terahertz and Raman vibrational spectroscopies study," *Molecules* **24**(3), 488 (2019).
50. C. Wang, J. Gong, Q. Xing, Y. Li, F. Liu, X. Zhao, L. Chai, C. Wang, and A. M. Zheltikov, "Application of terahertz time-domain spectroscopy in intracellular metabolite detection," *J. Biophotonics* **3**(10-11), 641–645 (2010).
51. M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine," *Clin. Chem.* **39**(4), 561–577 (1993).

52. K. Xu, E. Bayati, K. Oguchi, S. Watanabe, D. P. Winebrenner, and M. H. Arbab, "Terahertz time-domain polarimetry (THz-TDP) based on the spinning EO sampling technique: determination of precision and calibration," *Opt. Express* **28**(9), 13482–13496 (2020).
53. K. Xu, M. Liu, and M. H. Arbab, "Broadband terahertz time-domain polarimetry based on air plasma filament emissions and spinning electro-optic sampling in GaP," *Appl. Phys. Lett.* **120**(18), 181107 (2022).