CHINESE OPTICS LETTERS

Vol. 20, No. 6 | June 2022

Revisiting the relationship between composite multiscale entropy and THz optical parameters with exterior product

Haishun Liu (刘海顺)[1*], Zhenwei Zhang (张振伟)[1**], Meiyan Liang (梁美彦)[2], and Cunlin Zhang (张存林)[1]

[1] Key Laboratory of Terahertz Optoelectronics, Ministry of Education, Department of Physics, Capital Normal University, Beijing 100048, China

[2] Department of Electronics and Information Engineering, Shanxi University, Taiyuan 030006, China

*Corresponding author: phscdream@163.com

**Corresponding author: zhangzw@cnu.edu.cn

Received January 25, 2022 | Accepted March 18, 2022 | Posted Online May 1, 2022

With the framework of exterior product, we investigate the relationship between composite multiscale entropy (CMSE) and refractive index and absorption coefficient by reanalyzing six concentrations of bovine serum albumin aqueous solutions from the published work. Two bivectors are constructed by CMSE and its square by the refractive index and absorption coefficient under vectorization. The desirable linear behaviors can be captured, not only between the defined two bivectors in normalized magnitudes, but also between the normalized magnitude of bivectors pertinent to CMSE and the magnitude of a single vector on the refractive index or absorption coefficient, with the processing of optimum selection. Besides that, the relationship between the coefficients of two bivectors is also considered. The results reveal that plenty of sound linear behaviors can be found and also suggest the scale of 15, 16 and frequency of 0.2, 0.21 THz are prominent for those linear behaviors. This work provides a new insight into the correlation between terahertz (THz) time and frequency domain information.

Keywords: bivector; refractive index; absorption coefficient; terahertz; composite multiscale entropy; linear behavior.

DOI: 10.3788/COL202220.063701

1. Introduction

Terahertz (THz) spectroscopy has made great progress in numerous fields owing to its attractive and exclusive attributes[1–5]. The coherent detection of THz radiation enables the amplitude and phase of temporal signal to be precisely measured in order to conveniently obtain the refractive index and absorption coefficient dependent on frequency. The optical properties of the measured sample are integrated into the THz waveform in the coupling between THz temporal signal and sample, which can be roughly observed by the shape alteration for the sample signal relative to the reference one. This also changes the complexity of the reference waveform, which leads to the introduction of composite multiscale entropy (CMSE)[6,7], sample entropies at different scales, to measure the THz temporal signal. In the case of similar samples with weak variation of signals complexities where it is hard to detect the disparities using CMSE, the variational mode decomposition (VMD) technique is taken into account to decompose the signal into several mode functions with different frequency components because THz signal can be seen as a synthesis composed of sub-signals with different frequency components[8,9]. So, the selected mode functions dependent on the intervals of frequency components can be described by CMSE. From our previous work[9], sample entropy at scale 16 extracted from the selected mode function showed strong pertinence to the refractive index or absorption coefficient at different frequencies as expected, which reveals that the characteristic of sample entropy describing THz mode functions is relevant to optical properties. Nevertheless, the in-depth association needs to be investigated because more valuable information could be captured by a combination of refractive index with absorption coefficients, and the purely numerical information from the refractive index or absorption coefficient sometimes may not fully exhibit the further relationship with sample entropy. Besides that, it is sensible to describe the complexity of signal at multiple scales, thereby considering the sample entropies integration of several scales.

Geometry algebra (GA)[10], as a general and concise computational framework that defines geometric product operation, has been extensively applied to solving the problems in a

mass of domains including signal and image processing[11,12], electrodynamics[13], quantum mechanics[14], relativity[15,16], etc. In the field of THz spectroscopy, the vectorized transfer function of the transmission mode and related properties were given in the GA framework[17]. The method of projective split[18] or conformal split[17,19] was used for substance identification. Additionally, it is also feasible to discriminate samples by measuring the magnitude of the selected vector rejection to the corresponding plane formed by two transfer functions with different sample thickness[20]. Besides that, there are other works on feature extraction or representation by using GA, some of which adopted the strategy of considering high-order structures and interrelations among the features[21–23].

The exterior product, an indispensable component of the geometric product, is involved in some applications. Li *et al.*[24] measured the disparities between THz time domain spectroscopy (TDS) transmission signals by defining the similarity function of exterior products for THz transfer functions. Owing to a simpler operation compared to GA, the exterior product has similar capabilities of acquiring the property of high-order correlation among the original variables, which can be described by a simple example as follows. Assuming **P** and **Q** are two different vectors containing four elements for each, the result of the exterior product can be easily obtained, as shown in Eq. (1). The correlated combinations within all possible elements are considered in the exterior product. So, the exterior product can catch more information and can link with different vectors to show the corresponding geometric meaning subject to different graded subspaces[25]. The novel properties could be found by associating the optical parameters with the exterior product. Additionally, the exterior product and inner product are two different operations. Equation (2) computes the inner product between **P** and **Q**. It can be seen that in Eq. (1) the interactions between **P** and **Q** are more complex than those in Eq. (2). Compared with the exterior product, the scalar result of the inner product can only be achieved by the elements with the same orthogonal basis. Therefore, the inner product has no capability in yielding the interactive results on different dimensional variables.

In this study, we investigate the characteristic of exterior product constructed by the vectorized refractive index and absorption coefficient as a function of frequency and the further relationship between CMSE and absorption coefficient, with the refractive index resorting to the exterior product by reusing the experiment data in the previous work[9]. Compared to our previous work, this study places emphasis on investigating the relationship with the consideration of correlation information on different frequencies or different scales by using the exterior product operation, which is an extension of univariate analysis in previous investigations. Moreover, this work provides an insight into more correlation information between the refractive index and absorption coefficients at different frequencies as well as between CMSE and its square on different scales. It should be note that the parameter, the square of CMSE, was constructed to compute the related exterior product:

$$\begin{aligned}\mathbf{P} &= a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 + a_4\mathbf{e}_4,\\ \mathbf{Q} &= b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3 + b_4\mathbf{e}_4,\\ \mathbf{P}\wedge\mathbf{Q} &= (a_1b_2 - a_2b_1)\mathbf{e}_1\wedge\mathbf{e}_2 + (a_1b_3 - a_3b_1)\mathbf{e}_1\wedge\mathbf{e}_3\\ &\quad + (a_1b_4 - a_4b_1)\mathbf{e}_1\wedge\mathbf{e}_4\\ &\quad + (a_2b_3 - a_3b_2)\mathbf{e}_2\wedge\mathbf{e}_3 + (a_2b_4 - a_4b_2)\mathbf{e}_2\wedge\mathbf{e}_4\\ &\quad + (a_3b_4 - a_4b_3)\mathbf{e}_3\wedge\mathbf{e}_4, \end{aligned} \tag{1}$$

$$\mathbf{P}\cdot\mathbf{Q} = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4. \tag{2}$$

2. Method

It is acknowledged that cross product between vectors is restricted to operate in three dimensions in linear algebra[26]. However, exterior product is a wise candidate on extension to higher dimensions. Exterior product is not only used together with the inner product making up the geometric product that provides a strong mathematical tool in physics and information engineering, but also serves as product operation in exterior algebra[27] that is widely applied in quantum theory[28]. The exterior product between the refractive index vector and absorption coefficient vector, and the magnitude of normalized exterior product can be separately expressed as

$$\mathbf{A} = \mathbf{n}\wedge\boldsymbol{\alpha} = \sum_{i=1}^{k-1}\sum_{j>i}^{k}(n_i\alpha_j - n_j\alpha_i)\mathbf{e}_i\wedge\mathbf{e}_j, \tag{3}$$

$$A_0 = \frac{|\mathbf{n}\wedge\boldsymbol{\alpha}|}{|\mathbf{n}||\boldsymbol{\alpha}|} = \sqrt{\frac{\sum_{i=1}^{k-1}\sum_{j>i}^{k}(n_i\alpha_j - n_j\alpha_i)^2}{\sum_{i=1}^{k}(n_i)^2\sum_{j=1}^{k}(\alpha_j)^2}}, \tag{4}$$

where the two orthonormal basis vectors $\mathbf{e}_i$, $\mathbf{e}_j$ $(i, j = 1, 2, \ldots, k)$ satisfy $\mathbf{e}_i\wedge\mathbf{e}_j = -\mathbf{e}_j\wedge\mathbf{e}_i$ $(i \neq j)$, $\mathbf{e}_i\wedge\mathbf{e}_i = 0$, and the quantity A_0 represents the normalized area spanned by **n** and **α** from geometry. In Eq. (3), quantity **A** is named as bivector (or two-blade denoting a two-dimensional subspace)[21], which extends the concept of vectors. The coefficients of **A** present a new manner in the combination of refractive index with absorption coefficients. The geometric feature of the bivector regarding THz optical parameters could provide a novel perspective for measurement of THz property. In addition, the operation in Eq. (3) leads to dimension increase compared with the original vectors. So, it is necessary to select the dimension of vectors previous to using Eq. (3).

3. Results and Discussion

The bandwidth for absorption coefficients and refractive indices of six concentrations of bovine serum albumin (BSA) solutions in Ref. [9] is from 0.1 to 0.5 THz with the resolution of 0.01 THz. From Eq. (3), taking 5000 μg/mL as an example, the mean profile of **A** can be seen in Fig. 1. It should be noted that the vertical axis

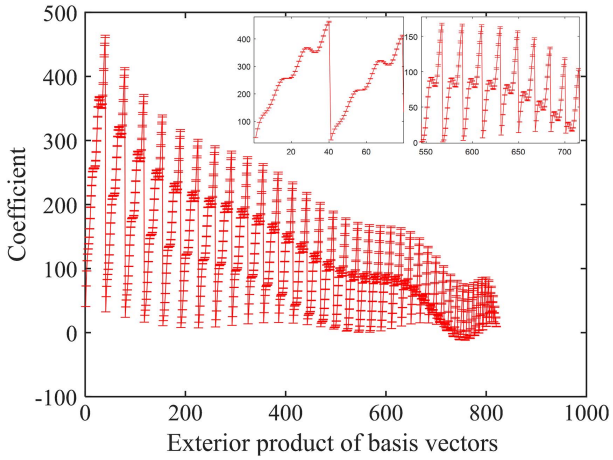

Fig. 1. Mean profile of **A** for 5000 μg/mL BSA aqueous solutions with 83% confidence interval.

represents the coefficients in Eq. (3), and the horizontal axis is the exterior product of basis vectors for all the possible pairs (i.e., basis bivectors), in which the numbers represent $\mathbf{e}_i \wedge \mathbf{e}_j$ $(i = 1, 2, \ldots, k, j = i + 1)$ in order. The profile of the first several basis bivectors is very similar to that of absorption coefficient, and the vector of the absorption coefficient is seemingly associated with **A**. It is accessible that the last several indices represent high frequency components interactions between the refractive index and absorption coefficient, and the profile corresponding to low frequency parts diminishes with the increasing index of the basis bivector. Additionally, for further insight into A_0, we investigated the relationship between A_0 and $|\mathbf{n}|$ as well as between A_0 and $|\boldsymbol{\alpha}|$. It can be seen from Fig. 2 that both of the R-squared reach above 0.9 and the R-squared of linear fitting between A_0 and $|\boldsymbol{\alpha}|$ is higher than those of A_0 and $|\mathbf{n}|$.

In order to further explore the relationship between CMSE and the refractive index or absorption coefficient in the framework of exterior product, similar to the bivector $\mathbf{n} \wedge \boldsymbol{\alpha}$, the bivector on CMSE is defined as $\mathbf{B} = \mathbf{CMSE} \wedge \mathbf{CMSE}^2$, which is composed of CMSE and its square. The quantity B_0, normalized magnitude of bivector **B**, is given as follows:

$$B_0 = \frac{|\mathbf{CMSE} \wedge \mathbf{CMSE}^2|}{|\mathbf{CMSE}||\mathbf{CMSE}^2|}. \tag{5}$$

In compliance with the method of our previous work, VMD was performed for that relationship investigation. In terms of the previous search results[9], CMSE of the second mode function at $\alpha = 200$ and 210 was considered to select the sample entropies of scales with small error bars, and the case of $\alpha = 200$ and 210 shared the highly similar behaviors in the profile of CMSE due to their nearly identical behaviors in respective mode functions. The error bars of the normalized magnitude of **B** are much larger in spite of the linear trend in the optimum solution. The reason could be speculated that some larger error bars exist in the sample entropies of corresponding scales, including some potentially inevitable noise. As the error bar is

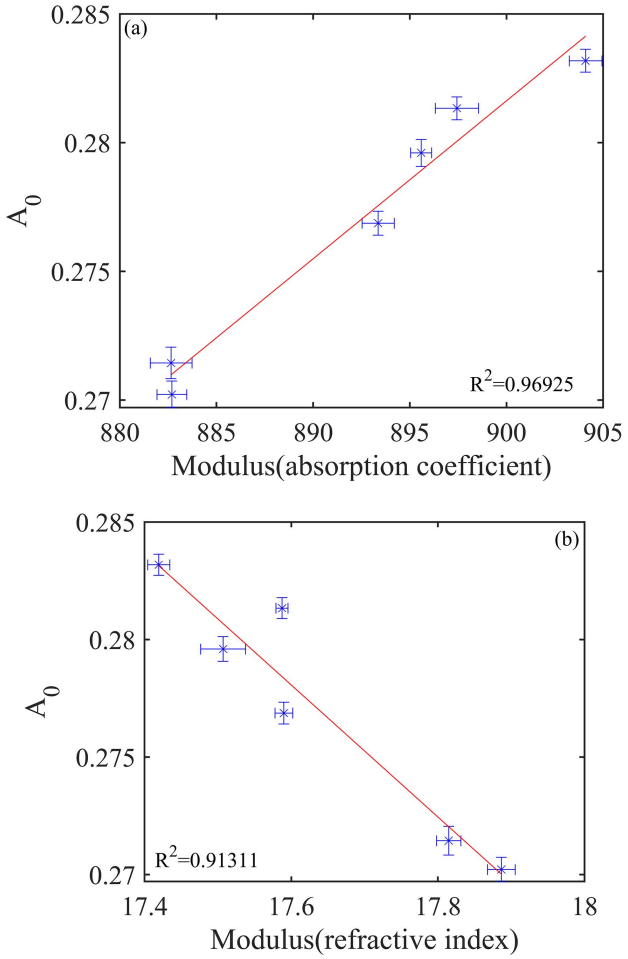

Fig. 2. Linear fitting between A_0 and modulus of the (a) absorption coefficient and (b) refractive index.

heavily dependent on standard deviation, the distribution on standard deviation of CMSE over 20 scales was conducted, as shown in Fig. 3. It seems that the possible sample entropies of more than scale nine could be selected to establish the relationship. In addition, the bandwidth approximate from 0.2 to 0.5 THz is adopted, owing to the frequency spectra of the selected mode functions at $\alpha = 200$ and 210[9]. By trial and error, the scales from 15 to 20 in B_0 were selected to obtain an expected linear dependence on the normalized magnitude A_0 from 0.2 to 0.38 THz [Figs. 4(a) and 4(b)]. More details on the parameter selections can be seen in Supplementary Material. Note that due to the similar behaviors at $\alpha = 200$ and 210, one of them was considered ($\alpha = 210$) in the following discussion. Besides that, the linear fitting between B_0 and $|\mathbf{n}|$ as well as B_0 and $|\boldsymbol{\alpha}|$ can be obtained, which is exhibited in Fig. 5. It is evident that the linear behavior between B_0 and $|\mathbf{n}|$ is much stronger than that between B_0 and $|\boldsymbol{\alpha}|$ or A_0. It also can be compared that, different from the linear behavior of A_0, there is much more linear dependence of B_0 on $|\mathbf{n}|$ and less on $|\boldsymbol{\alpha}|$. Nevertheless, the linear dependence of B_0 on $|\boldsymbol{\alpha}|$ can still be accepted. As a whole, the results indicate that the constructed bivectors on CMSE and

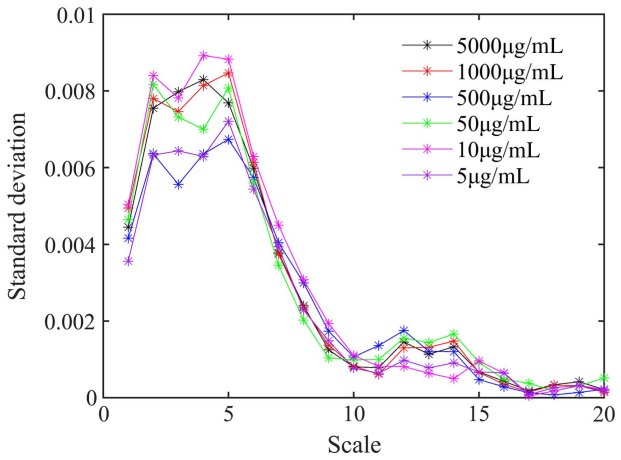

Fig. 3. Distribution on standard deviation of CMSE over 20 scales at $\alpha = 210$.

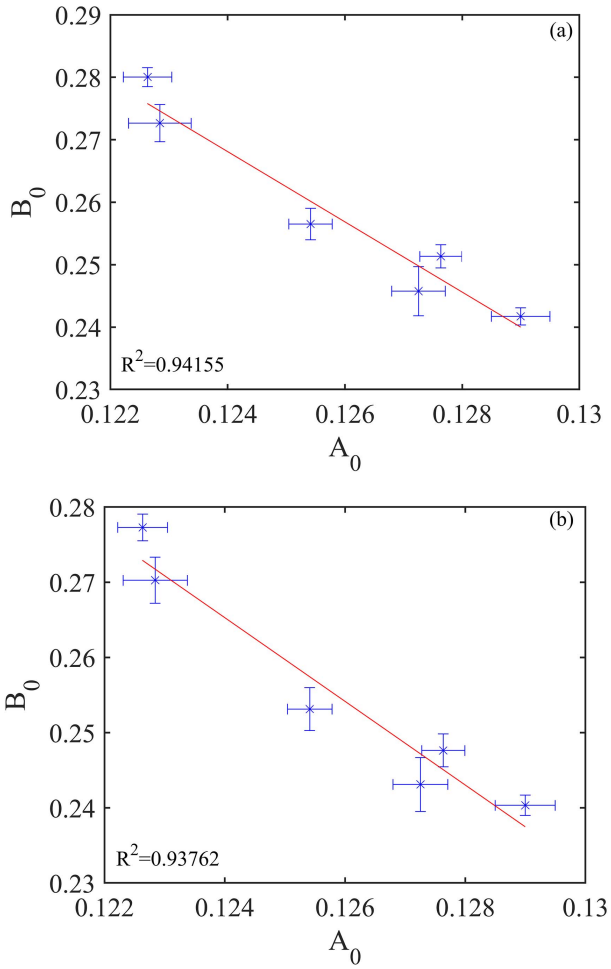

Fig. 4. Linear fitting between B_0 at (a) $\alpha = 200$ and (b) $\alpha = 210$ from scales 15 to 20 and A_0 from 0.2 to 0.38 THz.

optical parameters are feasible to elucidate the highly linear correlation between their respective subspaces. Besides that, the linear relationships between B_0 and $|\mathbf{n}|$, $|\boldsymbol{\alpha}|$ also demonstrate that

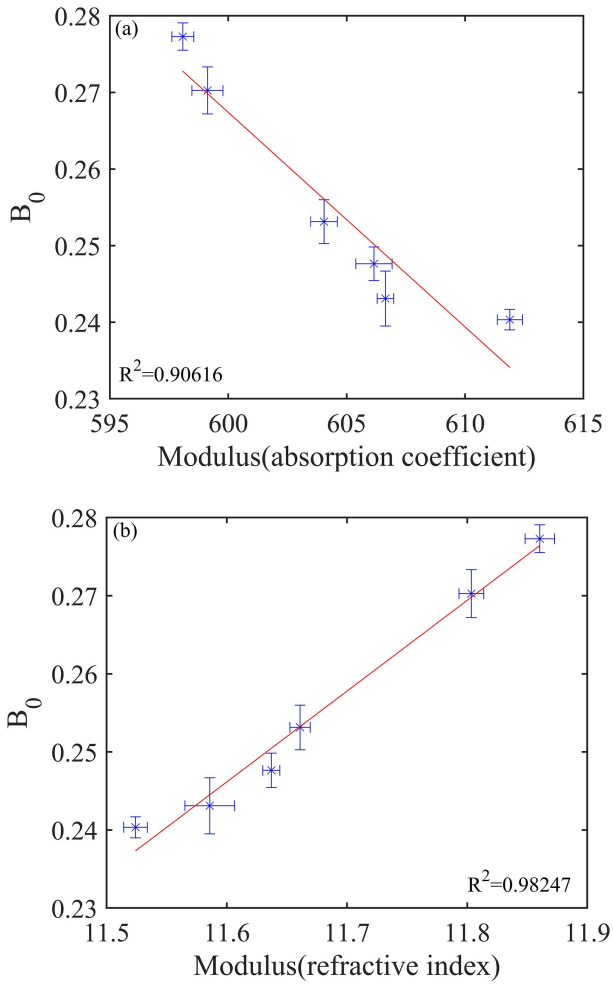

Fig. 5. Linear fitting between B_0 from scales 15 to 20 and modulus of (a) absorption coefficient and (b) refractive index from 0.2 to 0.38 THz.

B_0 is strongly related to the refractive index and absorption coefficient even in the single vector modulus.

Additionally, the correlation on the coefficients of bivectors between **A** and **B** should be further studied because they, respectively, reveal the interrelations between the two correspondingly constituted vectors. The bivector coefficient profiles of **A** from 0.2 to 0.38 THz and **B** from scales 15 to 20 are shown in Figs. 6(a) and 6(b), which separately depict the interactions between the refractive index and absorption coefficient and between CMSE and CMSE^2 with the increasing index of the basis bivector. The behavior in Fig. 6(a) is similar to that in Fig. 1, and some different disparities exist on the coefficients of **B** among these concentrations of samples including index 1 to 9. The linear fitting between them was performed by the exhaustive search that implemented linear fittings for all possible coefficients of **A** (from 0.2 to 0.38 THz) and **B** (from 15 to 20). Some representative cases were chosen to be exhibited in Figs. 7(a) to 7(e), which indicates some specific coefficients of bivectors can successfully describe the linear fitting for some subspaces between **A** and **B**. Furthermore, according to the combinations of basis bivectors $\mathbf{e}_i \wedge \mathbf{e}_j$, Figs. 8(a) and 8(b) show the weight distribution

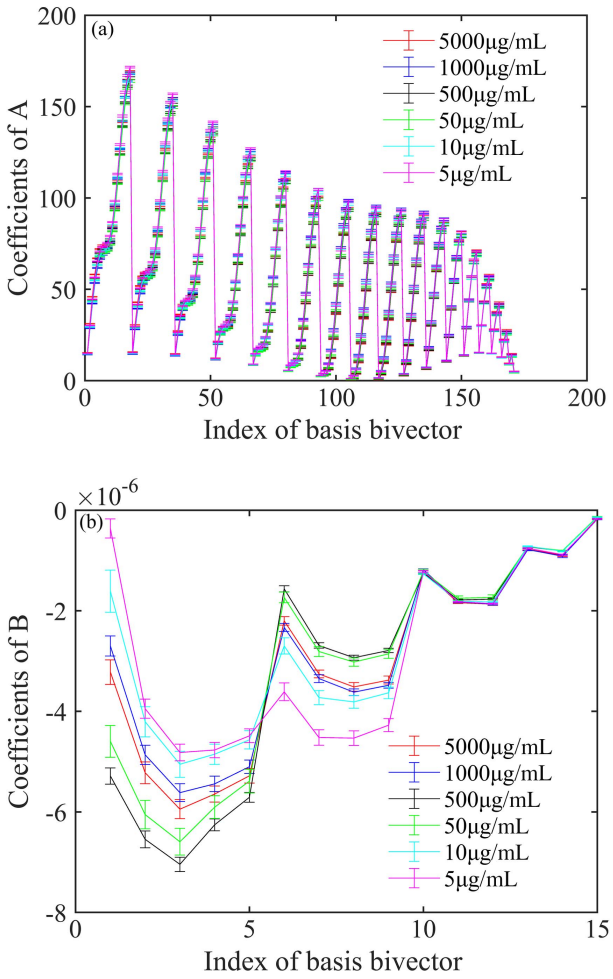

Fig. 6. Profile of the coefficients of (a) **A** from 0.2 to 0.38 THz and (b) **B** from scales 15 to 20.

of the representative frequency pairs and scale pairs in determination linear fitting. The index of the basis bivector was transformed into the upper triangular matrix that is shown as Eq. (6) in order to get a better interpretability. The subscripts $1, 2, \ldots, j$ above correspond to the frequencies 0.2, 0.21, 0.22, ..., 0.38 THz in turn for Fig. 8(a) and the scales 15, 16, ..., 20 in turn for Fig. 8(b). The intensity of the color bar denotes the proportions of highly linear behaviors ($R^2 > 90\%$). It should be noted that the parts in $R^2 < 90\%$ are shown with zeros. It can be seen that the first two lines are noticeable, which demonstrates that the scales 15, 16 and frequencies 0.2, 0.21 THz coupling with the remaining scales or some frequencies can acquire much more sound linear behaviors, thereby implying the importance of the scales 15, 16 and frequencies 0.2, 0.21 THz:

$$\begin{bmatrix} W_{12}, W_{13}, W_{14}, \ldots, W_{1j-3}, W_{1j-2}, W_{1j-1}, W_{1j} \\ W_{23}, W_{24}, W_{25}, \ldots, W_{2j-2}, W_{2j-1}, W_{2j} \\ W_{34}, W_{35}, W_{36}, \ldots, W_{3j-1}, W_{3j} \\ \ldots \\ W_{j-2j-1}, W_{j-2j} \\ W_{j-1j} \end{bmatrix}. \quad (6)$$

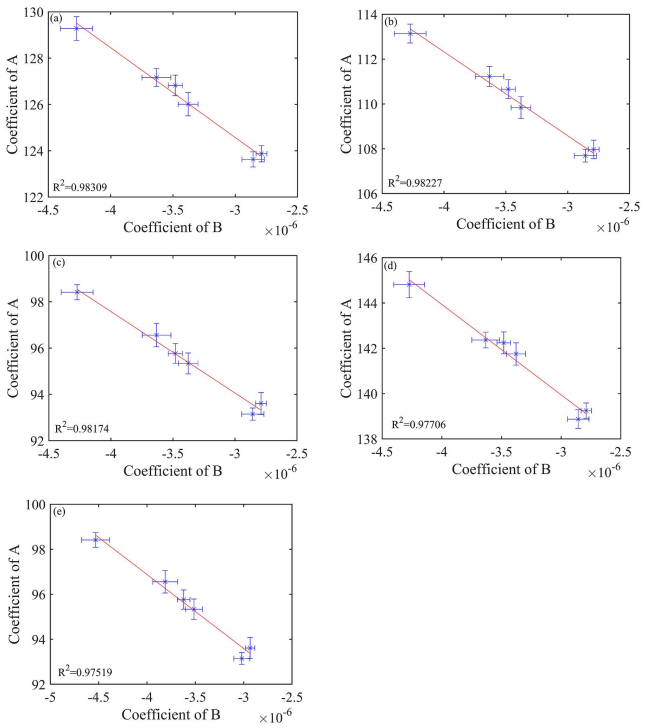

Fig. 7. Linear fitting between coefficients of **A** and **B**, respectively, constructed by basis bivectors at (a) 0.20, 0.34 THz and scales 16, 20, (b) 0.20, 0.33 THz and scales 16, 20, (c) 0.20, 0.32 THz and scales 16, 20, (d) 0.20, 0.35 THz and scales 16, 20, and (e) 0.20, 0.32 THz and scales 16, 19.

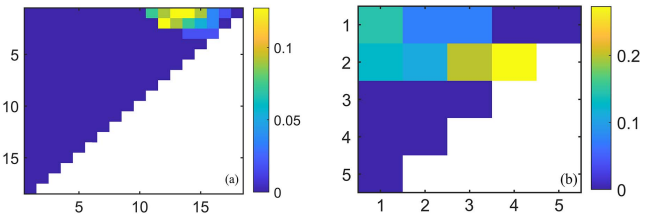

Fig. 8. Weight distribution on coefficients of bivector (a) **A** and (b) **B** in $R^2 > 90\%$.

4. Conclusions

In this work, by reanalyzing the data of previous work, we further study the relationship between CMSE and refractive index or absorption coefficient under the framework of exterior product, which presents a novel insight into the relationship between THz time and frequency domain information. The quantity bivectors $\mathbf{CMSE} \wedge \mathbf{CMSE}^2$ and $\mathbf{n} \wedge \boldsymbol{\alpha}$ were established by vectorizing the corresponding parameters, and their linear behaviors in normalized magnitude and bivector coefficients were given with optimization processing, which provides their correlation on the second graded information (subspaces) and further validates the results of previous work as well. It is feasible to introduce the exterior product to characterize the properties of CMSE, refractive index, and absorption coefficient. Besides that, the coefficients of bivectors have large potential in

sifting important variables depending on the performance of corresponding constructed vectors. The magnitude and coefficients of bivectors established in this work may also fertilize the methodology in substance identification of THz spectroscopy.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Nos. U1837202 and 11804209) and Joint Funds of the Equipment Pre-research and Aerospace Science and Technology (No. 6141B061006).

References

1. Y. Peng, C. Shi, X. Wu, Y. Zhu, and S. Zhuang, "Terahertz imaging and spectroscopy in cancer diagnostics: a technical review," BME Frontiers **2020**, 2547609 (2020).
2. P. Bawuah and J. A. Zeitler, "Advances in terahertz time-domain spectroscopy of pharmaceutical solids: a review," TrAC Trends Anal. Chem. **139**, 116272 (2021).
3. L. Afsah-Hejri, E. Akbari, A. Toudeshki, T. Homayouni, A. Alizadeh, and R. Ehsani, "Terahertz spectroscopy and imaging: a review on agricultural applications," Comput. Electron. Agric. **177**, 105628 (2020).
4. L. Yang, T. Guo, X. Zhang, S. Cao, and X. Ding, " Toxic chemical compound detection by terahertz spectroscopy: a review," Rev. Anal. Chem. **37**, 21 (2018).
5. S. Zhong, " Progress in terahertz nondestructive testing: a review," Front. Mech. Eng. **14**, 273 (2019).
6. S.-D. Wu, C.-W. Wu, S.-G. Lin, C.-C. Wang, and K.-Y. Lee, "Time series analysis using composite multiscale entropy," Entropy **15**, 1069 (2013).
7. R. Zhang, Y. He, K. Liu, L. Zhang, S. Zhang, E. P. MacPherson, Y. Zhao, and C. Zhang, "Composite multiscale entropy analysis of reflective terahertz signals for biological tissues," Opt. Express **25**, 23669 (2017).
8. H. Liu, K. Zhao, X. Liu, Z. Zhang, J. Qian, C. Zhang, and M. Liang, "Diagnosis of hepatocellular carcinoma based on a terahertz signal and VMD-CWSE," Biomed. Opt. Express **11**, 5045 (2020).
9. H. Liu, X. Liu, Z. Zhang, and C. Zhang, "Investigating the relationship between composite multiscale entropy and optical property of bovine serum albumin solutions using THz pulsed system," Opt. Lasers Eng. **137**, 106374 (2021).
10. D. Hestenes and G. Sobczyk, *Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics* (D. Reidel, 1984).
11. R. Wang, M. Shen, T. Wang, and W. Cao, "L1-norm minimization for multidimensional signals based on geometric algebra," Adv. Appl. Clifford Algebras **29**, 33 (2019).
12. Y. Li, W. Liu, X. Li, Q. Huang, and X. Li, "GA-SIFT: a new scale invariant 500 feature transform for multispectral image using geometric algebra," Inf. Sci. **281**, 559 (2014).
13. B. Jancewicz, *Multivectors and Clifford Algebra in Electrodynamics* (World Scientific, 1989).
14. D. Hestenes, "Clifford algebra and the interpretation of quantum mechanics," in *Clifford Algehras and Their Applications in Mathematical Physics* (D. Reidel, 1986), p. 321.
15. F. J. Chinea, "A Clifford algebra approach to general relativity," Gen. Relat. Gravit. **21**, 21 (1989).
16. A. Lasenby, C. Doran, and S. Gull, "Gravity, gauge theories and geometric algebra," Philos. Trans. R. Soc. A **356**, 487 (1998).
17. W. Xie, J. Li, and J. Pei, "THz-TDS signal analysis and substance identification," J. Phys **276**, 012227 (2011).
18. W. Xie, J. Li, and J. Pei, "An analysis of THz-TDS signals using geometric algebra," Proc. SPIE **7277**, 72770E (2008).
19. W. Xie, J. Li, and J. Pei, "THz-TDS signal analysis and substance identification via the conformal split," Sci. China Inf. Sci. **55**, 49 (2012).
20. S. Zhou, D. G. Valchev, A. Dinovitser, J. M. Chappell, A. Iqbal, B. W.-H. Ng, T. W. Kee, and D. Abbott, "Terahertz signal classification based on geometric algebra," IEEE Trans. THz Sci. Technol. **6**, 793 (2016).
21. Y. Xu, W. Hong, Z. Gao, and C. Zheng, "Geometric algebra representation principle of multivariate data dimension-increasing transformation," J. Yanshan Univ. **32**, 393 (2008).
22. Y. Xu, W. Hong, and Z. Gao, "Geometric algebra multi-vector representation method of pattern features," J. Yanshan Univ. **34**, 119 (2010).
23. J. Li and W. Hong, "Feature extraction for breast cancer data based on geometric algebra theory and feature selection using differential evolution," J. Biomed. Eng. **31**, 1218 (2014).
24. J. Li, W. Xie, and J. Pei, "Similarity metrics and their applications of THz-TDS signals," Signal Process. **27**, 1457 (2011).
25. B. Mishra, M. Kochery, P. Wilson, and R. Wilcock, "A novel signal processing coprocessor for *n*-dimensional geometric algebra applications," Circuit. Syst. **5**, 274 (2014).
26. L. Zou, J. Lasenby, Q. Wan, Z. He, and M. Lv, "Geometric algebra in electronics and information engineering: an introduction," Int. J. Elec. Eng. Educ. **53**, 252 (2016).
27. M. El Baz and Y. Hassouni, "Deformed exterior algebra, quons and their coherent states," Int. J. Mod. Phys. A **18**, 3015 (2003).
28. A. Singleton, "Superconformal quantum mechanics and the exterior algebra," J. High Energ. Phys. **2014**, 131 (2014).