Terahertz-dependent identification of simulated hole shapes in oil–gas reservoirs*

Ri-Ma Bao(宝日玛)[1,2,†], Hong-Lei Zhan(詹洪磊)[1,2,†], Xin-Yang Miao(苗昕扬)[2], Kun Zhao(赵昆)[1,2,3,‡], Cheng-Jing Feng(冯程静)[2], Chen Dong(董晨)[2], Yi-Zhang Li(李羿璋)[2], and Li-Zhi Xiao(肖立志)[1]

[1] *State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing* 102249, *China*
[2] *Beijing Key Laboratory of Optical Detection Technology for Oil and Gas, China University of Petroleum, Beijing* 102249, *China*
[3] *Key Laboratory of Oil and Gas Terahertz Spectroscopy and Photoelectric Detection, China Petroleum and Chemical Industry Federation (CPCIF), Beijing* 100723, *China*

(Received 3 February 2016; revised manuscript received 27 May 2016; published online 25 August 2016)

Detecting holes in oil–gas reservoirs is vital to the evaluation of reservoir potential. The main objective of this study is to demonstrate the feasibility of identifying general micro-hole shapes, including triangular, circular, and square shapes, in oil–gas reservoirs by adopting terahertz time-domain spectroscopy (THz-TDS). We evaluate the THz absorption responses of punched silicon (Si) wafers having micro-holes with sizes of 20 μm–500 μm. Principal component analysis (PCA) is used to establish a model between THz absorbance and hole shapes. The positions of samples in three-dimensional spaces for three principal components are used to determine the differences among diverse hole shapes and the homogeneity of similar shapes. In addition, a new Si wafer with the unknown hole shapes, including triangular, circular, and square, can be qualitatively identified by combining THz-TDS and PCA. Therefore, the combination of THz-TDS with mathematical statistical methods can serve as an effective approach to the rapid identification of micro-hole shapes in oil–gas reservoirs.

Keywords: hole shapes, oil–gas reservoirs, THz-TDS, PCA

PACS: 02.70.Rr, 78.20.Ci, 78.30.Jw **DOI:** 10.1088/1674-1056/25/10/100204

1. Introduction

The physical properties of an oil–gas reservoir in the subterranean area clearly influence the analyses of the contribution of the oil and gas reservoir.[1] Rocks with a high pore rate, an influential feature of the earth crust and stratum, have been investigated widely because of the basic theory about reservoirs.[2] The recognition of hole shapes has considerable practical importance in the field of oil–gas exploration, particularly for micro- and nanogaps present in reservoirs. Silicon (Si), a vital substance in geo- and biochemical cycles, is the second most abundant element in the Earth's crust, particularly in the upper continental crust.[3] The physicochemical properties of Si are closely related to the structure and reservoir quality of geological material. Hole shape plays a vital role in the formation of various pore structures, which are used to directly investigate the mechanical properties of an oil–gas reservoir. Therefore, the detection of punch Si wafers with diverse micron-hole shapes is useful in further studying the rock strata.

Because of the developments of ultrashort pulse lasers, semiconductors, and optical detectors, the terahertz time-domain spectroscopy (THz-TDS) has advanced rapidly. THz spectroscopy has been studied and used widely in various fields.[4–22] Several studies of the properties of Si wafers have been conducted by using hole arrays in a THz range. Sharp resonances and enhanced transmission were observed when THz radiation pulses passed through gratings of subwavelength holes.[23] To investigate the resonance characteristics of Si-doped arrays, heavily doped Si dot arrays were measured and multiple plasmon resonances were observed for dot arrays demonstrating diverse dimensions in the THz range.[24] The zero-order THz transmission spectra of an array of subwavelength apertures on ultrathin highly doped Si exhibited highly defined maxima and minima because of the excitations of surface-plasmon polaritons and Wood's anomaly.[25] Therefore, the THz responses of holes are strongly influenced by their shapes and sizes.

In this paper, we present the THz absorption responses of Si wafers, which have been punched with different hole shapes (e.g., triangular, circular, and square) to simulate an oil–gas reservoir. Principal component analysis (PCA) is adopted to cluster the Si wafers punched with the same hole shape. The principal component (PC) distributions clearly indicate successful recognition of the various holes in the evaluated coordinate systems, and the holes demonstrating similar PCs are assembled into an independent space. Therefore, this study

*Project supported by the National Natural Science Foundation of China (Grant No. 61405259), the National Basic Research Program of China (Grant No. 2014CB744302), and the Specially Founded Program on National Key Scientific Instruments and Equipment Development, China (Grant No. 2012YQ140005).

†These authors contribute equally to this work
‡Corresponding author. E-mail: zhk@cup.edu.cn

© 2016 Chinese Physical Society and IOP Publishing Ltd

demonstrates that THz-TDS, combined with statistical methods such as PCA, is an excellent approach to identifying and predicting microholes in oil–gas reservoirs.

2. Experimental methods

The experimental setup included a conventional THz-TDS system with transmission geometry from the Zomega Terahertz Corporation, USA. One of our previous reports comprehensively introduced the relevant parameters and apparatus of the system.[26] In brief, a femtosecond (fs) laser beam was split into a pump beam and a detection beam. The THz pulse was generated by a p-type InAs wafer with ⟨100⟩ orientation pumped by a Ti:sapphire laser with a center wavelength of 800 nm, pulse width of 100 fs, and repetition rate of 80 MHz. A 2.8-mm-thick ⟨100⟩ ZnTe was employed as a sensor, and a standard lock-in technology was used in this setup. To prevent the vapor from being absorbed in the air and enhancing the signal-to-noise ratio (SNR), the setup was covered with dry nitrogen at room temperature.

In this experiment, Si wafers were punched with holes having different shapes, numbers, and sizes. The microholes (triangular, circular, and square holes) were punched onto the Si wafers by using a laser drilling technique. Figure 1 illustrates the overall properties of the three hole shapes, measured using a microscope; the black areas indicate the hole areas. The size of the Si wafer (15 mm (side length) ×15 mm (side length) ×0.4 mm (thickness)) remained constant for all samples, whereas the sizes and number of holes were varied. All the holes were located tightly in an array in the center of a 5-mm-radius circle space, and the pore array region is corresponding to the THz radiation region. Listed in Table 1 are the detailed hole shapes, sizes, and numbers.

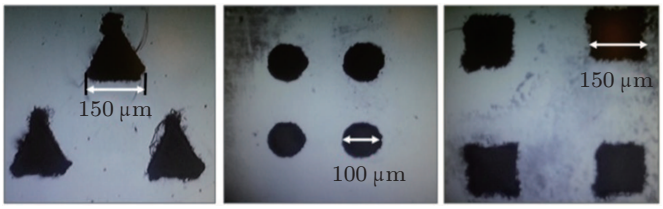


Fig. 1. (color online) Microscope images of the three hole shapes punched on the Si wafers.

Table 1. Samples with various hole shapes, sizes, and quantities.

Triangle hole (33 samples)		Circle hole (32 samples)		Square hole (29 samples)	
Side/μm	Number	Diameter/μm	Number	Side/μm	Number
150	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20	20	20, 50, 100, 150	150	1, 2, 3, 4, 5, 6, 8, 9, 10, 20
160, 180, 200, 250, 300, 350, 400, 450	1, 2	40, 60, 80	1	160, 180	1, 2
500	1, 2, 3, 4, 5, 6	100	1, 2, 3, 4, 5, 6, 7, 8, 9, 20, 30	200	1
		120, 140, 160, 250, 300, 350, 400, 450	1	250, 300, 350, 400, 450	1, 2
		500	1, 2, 3, 4, 5,	500	1, 2, 4, 5 6

In the experiment, the samples and references were initially subjected to THz-TDS. All the samples were tested twice. The relative error of each sample in the two measurements was calculated and did not exceed 10%, relative to the corresponding average spectra. Therefore, the spectral deviations in the two measurements were extremely small because of the stability of the setup, and only the THz data in one of the two measurements were employed in the subsequent calculation and discussion. In addition, selected samples with various hole shapes and sizes were filled with crude oil and tested by THz-TDS to simulate the oil–gas reservoir. Fast Fourier transform (FFT) was used for deriving the THz frequency domain spectra (THz-FDS). According to the derived spectra, THz absorbance spectra were calculated using the relation $-\ln(A_{\mathrm{Sam.}}/A_{\mathrm{Ref.}})$, where $A_{\mathrm{Sam.}}$ and $A_{\mathrm{Ref.}}$ represent the FFT amplitudes of the sample (Si wafers with different hole shapes) and reference (nitrogen) THz pulses, respectively. The effective frequency range (0.1 THz–2.2 THz) of the absorption spectra was determined according to the amplitudes in THz-FDS.

3. Results and discussion

Figure 2 shows the frequency-dependent absorption spectra of the wafers with various hole shapes (i.e., triangular, circular, and square) and that of selected wafers filled with crude oil at a frequency range of 0.1 THz–2.2 THz (i.e., the effective frequency range). Because of the existence of excessive datapoints, the errobars are not shown in Fig. 2 to ensure that the spectra are clearer. Similar spectral values and profiles are found among Si wafers having the same hole shape. The absorbance values of the three hole shapes remained unchanged in a range from 0.5 THz to 1.5 THz; however, they demonstrate the fluctuation in ranges from 0.1 THz to 0.5 THz and from 1.5 THz to 2.2 THz THz. No sharp absorption features are observed in the effective frequency range, and this is highly

consistent with the absorption response of float-zone Si reported in a previous study.[27] In general, the absorption spectra of the Si wafers are highly similar to those of other samples with the same hole shape; furthermore, the deviations between any two samples with the same hole shape are extremely small, indicating that there are few experimental errors and the stability of the setup is verified. Meanwhile, a slight difference is observed in the spectra of wafers associated with the three hole shapes, including wave profiles and values mainly caused by different hole shapes. However, these spectra are still highly overlapped (Fig. 2), and they can be distinguished using other methods, as discussed later. Besides, the spectra of crude oil filled samples resemble each other in a range from 0.1 THz to 1.0 THz and then are separated from each other. In addition, the crude oil filled samples absorb more THz waves than the hole in almost the whole range shown in the figure, owing to the absorptions of oil and water contained in the crude oil.

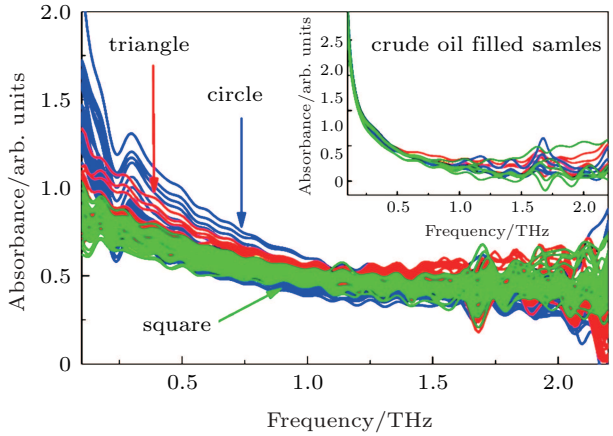


Fig. 2. (color online) THz absorbance spectra of the three groups of punched and crude oil filled samples with triangular (red), circular (blue), and square (olive) holes in a frequency range from 0.1 THz to 2.2 THz, respectively.

In order to detect and distinguish the spectra of holes with various shapes, PCA is employed to cluster the punched Si wafers with similar hole shapes by using THz absorbance spectra in a range of 0.1 THz–2.2 THz as the input, and spectral pretreatments are not performed. As a widely used statistical analysis technique, PCA is a multivariate statistical technique in which the number of dimensions within the data is reduced while retaining the overall variation as much as possible and identifying the potential structure of large spectral data as well as groups within the data sets. Generally, the theory of principal component analysis (PCA) can be described by several steps. The first step is the input of matrix as follows:

$$X = \begin{bmatrix} x_{11}, & x_{12} & \cdots x_{1n} \\ x_{21}, & x_{22} & \cdots x_{2n} \\ \vdots & \vdots & \ddots \\ x_{m1}, & x_{m2} & \cdots x_{mn} \end{bmatrix}, \tag{1}$$

where n is the number of the sample dimensions and m is the number of the samples. Then, the linear transformation of X is calculated and represented with Y as follows:

$$\begin{cases} Y_1 = u_{11}X_1 + u_{12}X_2 + \cdots u_{1n}X_n, \\ Y_2 = u_{21}X_1 + u_{22}X_2 + \cdots u_{2n}X_n, \\ \cdots \\ Y_n = u_{n1}X_1 + u_{n2}X_2 + \cdots u_{nn}X_n. \end{cases} \tag{2}$$

Here,

$$u_{k1}^2 + u_{k2}^2 + \cdots u_{kn}^2 = 1, k = 1, 2, \ldots, n. \tag{3}$$

There is no correlation between Y_i and Y_j ($i \neq j$; i, $j = 1, 2, \ldots, n$). Moreover, Y_1 is the linear combination of X_1, X_2, …, X_n with the largest variance, described as PC1. Y_2 is the linear combination with the largest variance and not related to Y_1, called PC2, etc., all the principal components are calculated. By selecting the results with larger variance, data simplification can be achieved. The method provides a roadmap on how to reduce a complex data set to a lower dimension.

In this study, three groups of spectral absorbance data, each of which follows an ascending order according to porosity, were combined in the order of triangle, circle, and square. Four data sets are then obtained for the calculation below.

PCA calculations result in several variables (PCs) that are related to the original variables and reflect the information about samples. When the scores of the early PCs are plotted against each other, a two- or three-dimensional score space can be obtained; in this space, closely related samples are clustered together and unrelated samples become outliers. However, an individual sample clusters with other samples having related PC properties with it. According to Eqs. (1)–(3), the absorbance spectra in Fig. 2 are used as input X (from X_{1j} to X_{mj}, $m = 94$, $j = 1, 2, \ldots, n$), and the output Y_s are the calculated PCs. The first three PCs are employed here. To determine the degree of separation of the feature vectors associated with the THz absorbance spectra from the three hole shapes, the scores of PCs are used and plotted in Fig. 3, where the position of each sample is reported in a three-dimensional space for the three PCs (PC1, PC2, and PC3).[28]

The plots illustrated in Fig. 3 enable graphically observing the information about both the data samples and variables of a data matrix, where data samples are displayed as points and variables are displayed as vectors.[8,26] The X, Y, and Z axes indicate the scores of PC1, PC2, and PC3 (the contribution rates of three PCs are provided in parentheses). Furthermore, Δ, ◯, and □ indicate the punched Si wafers with triangular, circular, and square holes in the new coordinate system, respectively. Figures 3(a)–3(c) illustrate the classification results of any two groups of holes, including triangle–circle, circle–square, and triangle–square, indicating that the various hole shapes are adequately separated.

The THz absorbance spectra of all samples are then combined and subjected to PCA as shown in Fig. 3(d). The results indicate that the first three PCs (PC1, PC2, and PC3) of the

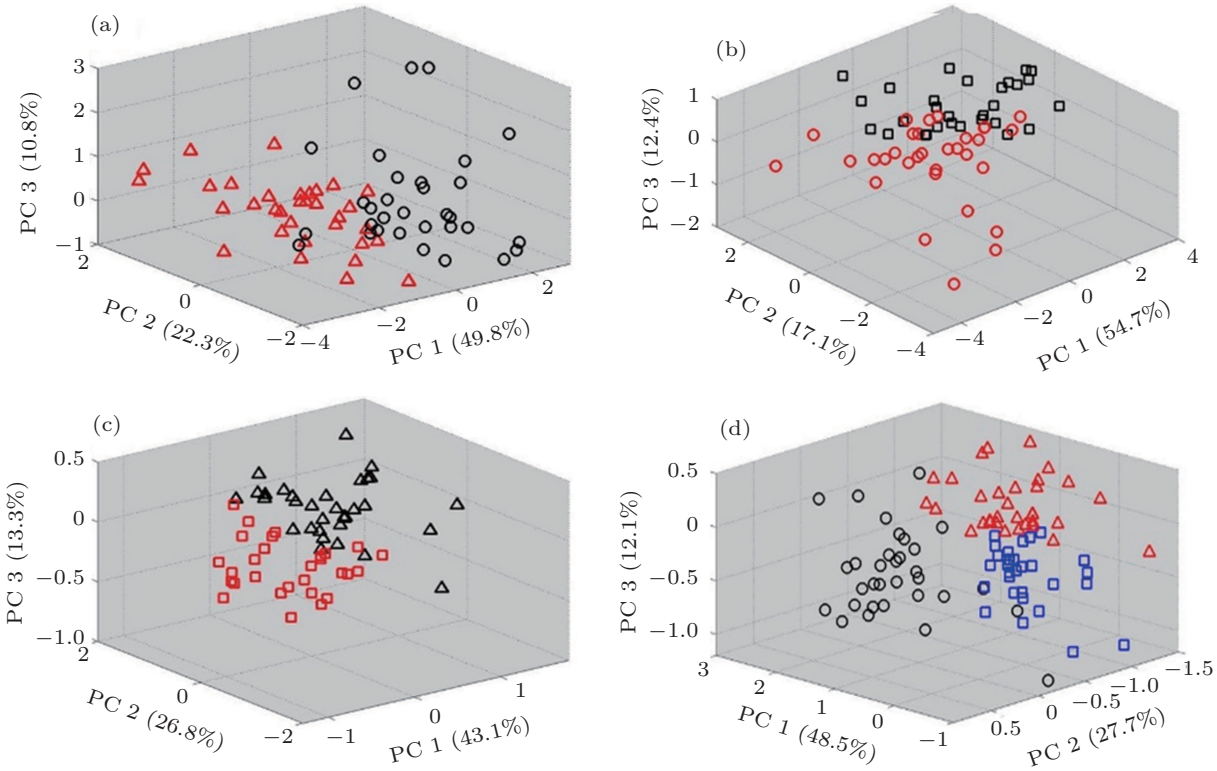


Fig. 3. (color online) Three-dimensional scatter plots of the PCA-THz absorbance spectra for the punched Si wafer with (a) triangular and circular holes, (b) circular and square holes, (c) triangular and square holes, and (d) triangular, circular, and square holes. The percentage values indicate the contribution rates of the corresponding PCs.

data set describe 90.6% of the variance within the data, with the first seven PCs describing 98.0% of the variance with the data. The cluster results show that the three hole shapes demonstrate obvious divergence. However, a common feature observed in the three-dimensional space is that the position points of the triangular and square holes are tightly clustered, but those of the circle holes are scattered. The PCs' deviations of circle are greater than those of triangle and square, and this is possibly due to the greater distance between pores, i.e., bigger porosity. Therefore, PCA is used to identify groups within the data and is performed on the scores of the first three PCs instead of on the original absorbance data to obtain suitable classification; thus, differentiating hole shapes is easier. This study shows that THz-TDS combined with PCA can serve as a promising and effective tool for recognizing microholes in Si wafers and other rock materials.

To verify the accuracy and repeatability of the experimental measurement and calculation procedures, the PC scores of six samples (triangle: 5 holes sized 500 μm and 3 holes sized 150 μm; circle: 5 holes sized 500 μm and 4 holes sized 100 μm; square: 5 holes sized 500 μm and 4 holes sized 150 μm) obtained in another measurement are analyzed and compared with those of the samples measured using PCA. The results indicate that the first three PCs (PC1, PC2, and PC3) of the data set also describe 90.6% of the variance within the data. A two dimensional system is used for identifying the contribution positions of the six new datapoints more clearly. Figure 4 illustrates the X–Z cross-section of the original three-dimensional space, namely the PC1 versus PC3 image, where each datapoint indicates the appropriate area of the hole shape corresponding to that shown in Fig. 3(d). Therefore, we conclude that a new Si wafer with the unknown hole shape, including triangular, circular, and square, can be qualitatively identified by combining the THz technique and PCA, and this combination can be extended to qualitatively identify other hole shapes in rock stratum in the future.

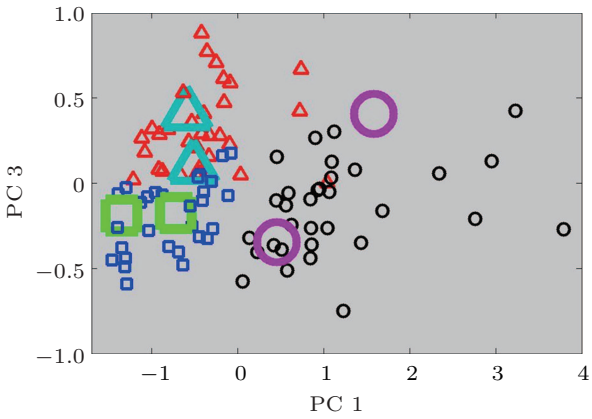


Fig. 4. (color online) PC1 versus PC3 for the data set. The six datapoints indicated by the largest shapes are the new samples in another measurement.

The objective of this study is to realize the identification of microhole shapes by using THz-TDS. The observed changes in the THz absorbance spectra are extremely small, thus posing the question of the reliability of identifying vari-

ous microhole shapes. PCA is employed to demonstrate the accuracy of the process of identifying different shapes by using THz absorbance spectra as the input. Any of the three hole types with varying porosity are clustered into a particular space in a PC system. This combination of THz-TDS and PCA can promote further the identification of hole shapes, in particular on or in rocks storing oil and gas. As a contactless technique, THz-TDS can serve as a supplement to traditional methods in oil–gas detection and exploration fields, which is worth exploring continuously.

4. Conclusions

In this work, we verify qualitatively identifying Si wafers punched with microholes having various shapes, including triangular, circular and square, in a simulated oil–gas reservoir by using THz-TDS. PCA is employed to classify the samples with different hole shapes, and large deviations of PCs are observed in the three-dimensional space. Moreover, the accuracy and repeatability of the experimental measurement are verified by analyzing the PC scores of six samples obtained in another measurement. The results indicate that combining THz-TDS with statistical methods can serve as a contactless and efficient approach to recognizing various hole shapes in oil–gas reservoirs.

Acknowledgment

The authors acknowledge Xin Feng and Tianyu Cai for performing the microscope test shown in Fig. 1.

References

[1] Quintal B 2012 *J. Appl. Geophys.* **82** 119
[2] Wang Y L, Du B Y, Dou X M, Liu J, Shi B Y, Wang D S and Tang H X 2007 *Colloid. Surface A* **307** 16
[3] Savage P S, Georg R B, Williams H M and Halliday A N 2013 *Earth Planet. Sci. Lett.* **365** 221
[4] Horiuchi N and Zhang X C 2010 *Nat. Photon.* **4** 662
[5] Lu X F and Zhang X C 2012 *Phys. Rev. Lett.* **108**
[6] Mittleman D M 2013 *Nat. Photon.* **7** 666
[7] Bao R M, Wu S X, Zhao K, Zheng L J and Xu C H 2013 *Sci. China: Phys. Mech. Astron.* **56** 1603
[8] Zhan H L, Wu S X, Bao R M, Zhao K, Xiao L Z, Ge L N and Shi H J 2015 *Rsc Adv.* **5** 14389
[9] Jiang C, Zhao K, Zhao L J, Jin W J, Yang Y P and Chen S H 2014 *Energ. Fuels* **28** 483
[10] Jin W J, Zhao K, Yang C, Xu C H, Ni H and Chen S H 2013 *Appl. Geophys.* **10** 506
[11] Jin W J, Li T, Zhao K and Zhao H 2013 *Chin. Phys. B* **22** 118701
[12] Suzuki H, Hoshina H and Otani C 2014 *Cryst. Growth Des.* **14** 4087
[13] Qin J Y, Xie L J and Ying Y B *Anal. Chem.* **86** 11750
[14] Lloyd-Hughes J 2014 *J. Phys. D: Appl. Phys.* **47** 374006
[15] Inagaki T, Ahmed B, Hartley I D, Tsuchikawa S and Reid M 2014 *J. Infrared Millim. THz Waves* **35** 949
[16] Ge H Y, Jiang Y Y, Lian F Y, Zhang Y and Xia S H 2014 *Spectrosc. Spect. Anal.* **34** 2897
[17] Dong C, Bao R M, Zhao K, Xu C H, Jin W J and Zhong S X 2014 *Chin. Phys. B* **23** 127802
[18] Yan C, Yang B and Yu Z C 2014 *Anal. Methods* **6** 248
[19] Xie L J, Yao Y and Ying Y B 2014 *Appl. Spectrosc. Rev.* **49** 448
[20] Choi K, Hong T, Sim K I, Ha T, Park B C, Chung J H, Cho S G and Kim J H 2014 *J. Appl. Phys.* **115**
[21] Gao R L, Yuan C X, Wang Y, Zhou Z X, Gong D W, Fang Y and Rong X W 2013 *J. Appl. Phys.* **114** 183510
[22] Zhao Y, Hwan L J, Zhu Y, Nazari M, Chen C, Wang H, Bernussi A, Holtz M and Fan Z 2012 *J. Appl. Phys.* **111** 053533
[23] Rivas J G, Schotsch C, Bolivar P H and Kurz H 2003 *Phys. Rev. B* **68** 201306
[24] Shi X H, Cleary A, Khalid A and Cumming D R S 2009 *Microelectron. Eng.* **86** 1111
[25] Azad A K, Y. Zhao Y and Zhang W 2005 *Appl. Phys. Lett.* **86** 141102
[26] Zhan H L, Wu S X, Bao R M, Ge L N and Zhao K 2015 *Fuel* **143** 189
[27] Dai J M, Zhang J Q, Zhang W L and Grischkowsky D 2004 *J. Opt. Soc. Am. B* **21** 1379
[28] Hwang J, Choi N, Park A, Park J Q, Chung J H, Baek S, Cho S G, Baek J S and Choo J 2013 *J. Mol. Struct.* **1039** 130