ISSN 0020-4412, Instruments and Experimental Techniques, 2023, Vol. 66, No. 5, pp. 753–760. © Pleiades Publishing, Ltd., 2023.
Russian Text © The Author(s), 2023, published in Pribory i Tekhnika Eksperimenta, 2023, No. 5, pp. 48–55.

GENERAL EXPERIMENTAL TECHNIQUE

A Modification of the Backward Correlation Method for the Brillouin Frequency Shift Accurate Extraction[1]

F. L. Barkov[a],* and Yu. A. Konstantinov[a]

[a] *Perm Federal Research Center, Ural Branch, Russian Academy of Sciences, Perm, 614990 Russia*

**e-mail: fbarkov@pstu.ru*

Received January 23, 2023; revised March 30, 2023; accepted April 2, 2023

Abstract—An improved method for extracting the Brillouin frequency shift in postprocessing of a given Brillouin gain spectrum is presented. Modification of the method made it possible to expand the boundaries of its applicability to the region of noisy spectra with a signal-to-noise ratio (SNR) below 0 dB. The modified method can be successfully used in distributed fiber-optic sensors operating on the Brillouin scattering principle, especially in long-distance sensing lines.

DOI: 10.1134/S0020441223050044

INTRODUCTION

Distributed fiber-optic sensors are increasingly being used in various fields of science and technology. Sensors based on Brillouin scattering play a significant role. Since Brillouin scattering is an inelastic effect of the interaction of photons with acoustic lattice vibrations, phonons, the change in the photon energy is directly related to the phonon energy. The Brillouin frequency shift (BFS) of a photon ν_b is given by the ratio $\nu_b = 2nv/\lambda$, where n is the refractive index, v is the sound velocity of the longitudinal acoustic wave, and λ is the wavelength in vacuum. Both the refractive index and the velocity of sound depend on external influences, such as the temperature and deformation; thus, determining the BFS can yield information about the magnitude of these influences along the fiber.

Spontaneous and forced Brillouin scattering are distinguished. Spontaneous scattering occurs on thermally activated phonons that are present in a fiber at any nonzero temperature. In case of forced scattering, phonons artificially created due to the electrostriction effect (changes in the material density under the influence of a strong electromagnetic field) play a role. To create such phonons, two optical waves of different frequencies are launched into the fiber from different ends.

Brillouin optical time domain reflectometers work on spontaneous scattering, while Brillouin analyzers of the time domain work on forced scattering.

Since forced scattering has a higher power by approximately two orders of magnitude, Brillouin time-domain analyzers are mainly used, especially with a large length of the measured line. A continuous probing optical wave with a frequency ν is launched into the fiber from one end, and a pulsed pumping wave with a higher frequency $\nu + \Delta\nu$ is launched from the other. The pumping-power intensity from the pump wave (and hence the power recorded by the photodetector) depends on the ratio of the frequency difference $\Delta\nu$ and the BFS ν_b. The theoretical dependence of the logarithmic gain on the frequency difference is described by the Lorentzian function: $g(\Delta\nu) = g_B(\Delta/2)^2/((\Delta/2)^2 + (\Delta\nu - \nu_b)^2)$, where g is the logarithmic gain, g_B is the peak gain, and Δ is the line width. The maximum is reached at $\Delta\nu = \nu_b$. By changing the pumping frequency, an experimental spectrum of the Brillouin amplification of a test wave is obtained.

There is another approach [1] in which the pumping wave is continuous and the test wave is pulsed. The principle of operation is completely similar, but spectral scanning is carried out at the frequency of the probe wave, and the result is the Brillouin absorption spectrum of the pump wave. The undoubted advantage of this approach is the large power of the useful response (since most of the power is pumped into the pump wave with this scheme).

For a typical fiber, the typical values are as follows: BFS = 11 GHz, Δ = 30–40 MHz; the BFS sensitivity to temperature is 1 MHz/K and that to deformation is 40–50 MHz/(1000 $\mu\varepsilon$). Thus, the accuracy of extracting the BFS from the Brillouin gain or absorption spectrum directly affects the accuracy of the sensor. As an example, to measure the temperature with an

[1] International Conference "Optical Reflectometry, Metrology, & Sensing," Russia, Perm, May 24–26, 2023.

accuracy of 1 K, the BFS must be determined with an error of no more than 1 MHz.

In general, regardless of the used algorithm, the error in determining the BFS is inversely proportional to the signal-to-noise ratio (SNR) of the available spectrum, proportional to the square root of the frequency scanning step and the square root of the Brillouin spectrum width [2]. Reducing the scanning step, as well as the signal accumulation (obviously leading to an increase in the SNR), leads to an increase in the accuracy of the sensor but reduces its speed at the same time. Therefore, the main trends in the development of time-domain Brillouin reflectometry are to increase the SNR not due to the signal accumulation [3–6] and the search for new algorithms for extracting the BFS.

The used algorithms can be divided into three large groups.

1. Reconstruction of the spectrum [7–10]. The approximation of the spectrum by the Lorentzian function is commonly used. The parameters of the function—the coordinate of the center, the full width at half-magnitude, and the amplitude—are determined based on the original points of the spectrum. For the considered task of searching for the BFS, it is enough to determine only the coordinates of the center. It is worth noting that the reconstruction algorithms are constantly being optimized in terms of both the speed and accuracy of parameter determination. Most commercial reflectometers and analyzers are equipped with such built-in algorithms.

2. Correlation methods [11, 12]. They usually represent the calculation of the cross-correlation function of a given spectrum and some reference spectra. The BFS is determined by which of the reference spectra the maximum value of the cross-correlation function was achieved.

3. Machine learning methods (neural network algorithms) [13–18]. At the training stage, the system looks for connections between some characteristics of the spectra and the BFS, while, at the measurement stage, it determines the BFS by the characteristics of a given spectrum using the found links. Either a connection is immediately sought between the characteristics of the spectrum and the measured physical quantities (e.g., temperature or deformation), and the BFS as an "intermediate link" can be excluded from consideration altogether.

A few years ago, the authors of this study proposed a backward correlation method (BCM) [19]. The essence of this method consists in reflecting the obtained spectrum relative to the central frequency of the scanned range and searching for such a shift of the reflected spectrum relative to the original one, at which the peaks of these two spectra overlap each other as much as possible. For example, if the maximum is exactly at the central frequency in the original spectrum, it will then be the same in the reflected spectrum and the optimal shift is zero. If the maximum in the original spectrum is at 1/4 of the distance from the left border of the spectrum, the maximum in the reflected spectrum will then be already 1/4 of the distance from the right in the reflected spectrum border and the optimal shift will be 1/2 of the spectrum. The method was described in detail in [19]. It occurred that, despite all its simplicity, the BCM yields excellent results at low SNR values and can even surpass the methods of reconstruction of Lorentzian spectra; the gain from using the BCM at low SNRs increases with an increase in the number of points in the spectrum (i.e., with a decrease in the scanning step for a given full range). This was explained by the fact that the value of the convolution of two random (noise) quantities increases proportionally to $\sqrt{N}$ with increasing the fineness of partitioning, where N is the number of partition points, while the value of the signal convolution is proportional to N.

An example of the work of the BCM is shown in Fig. 1. Figure 1a shows the initial spectrum (the spectral power density is expressed in arbitrary units); Fig. 1b shows the reflected (right) and reflected shifted (left) spectra. The magnitude of the shift varies from –400 to +400 MHz since the scanning frequency range is exactly 400 MHz. Figure 1c shows the dependence of the convolution of the original and reflected shifted spectra on the shift magnitude. The maximum convolution is at –170 MHz. Thus, the position of the Brillouin peak maximum determined using the BCM is 10 600 + (–170/2) = 10 515 MHz; here, 10 600 MHz is the central frequency in the scanning range, and –170 MHz must be halved for the reason that the reflected spectrum is shifted to the left (right) by exactly the same amount when the original spectrum is shifted to the right (left). The exact value of the Brillouin peak maximum for the spectrum in Fig. 1, where the spectrum generated with known parameters was used, is 10 513.8 MHz. Thus, in this example, the maximum detection error was 1.2 MHz.

Later, attempts were made to jointly apply the BCM with neural network methods and with spectrum reconstruction methods [20]. The joint application with the methods of spectrum reconstruction gave no significant results, and the consistent application of the BCM and the neural network algorithm based on the generalized linear model allowed the reduction of the error in determining the BFS by 0.4–1.6 MHz depending on the SNR.

According to the above classification, the BCM is logically attributed to correlation methods, although no reference spectra are used in this case.

In this paper, the possibilities of improving the BCM with a view to its potential application in the processing of spectra with even lower SNRs are investigated.

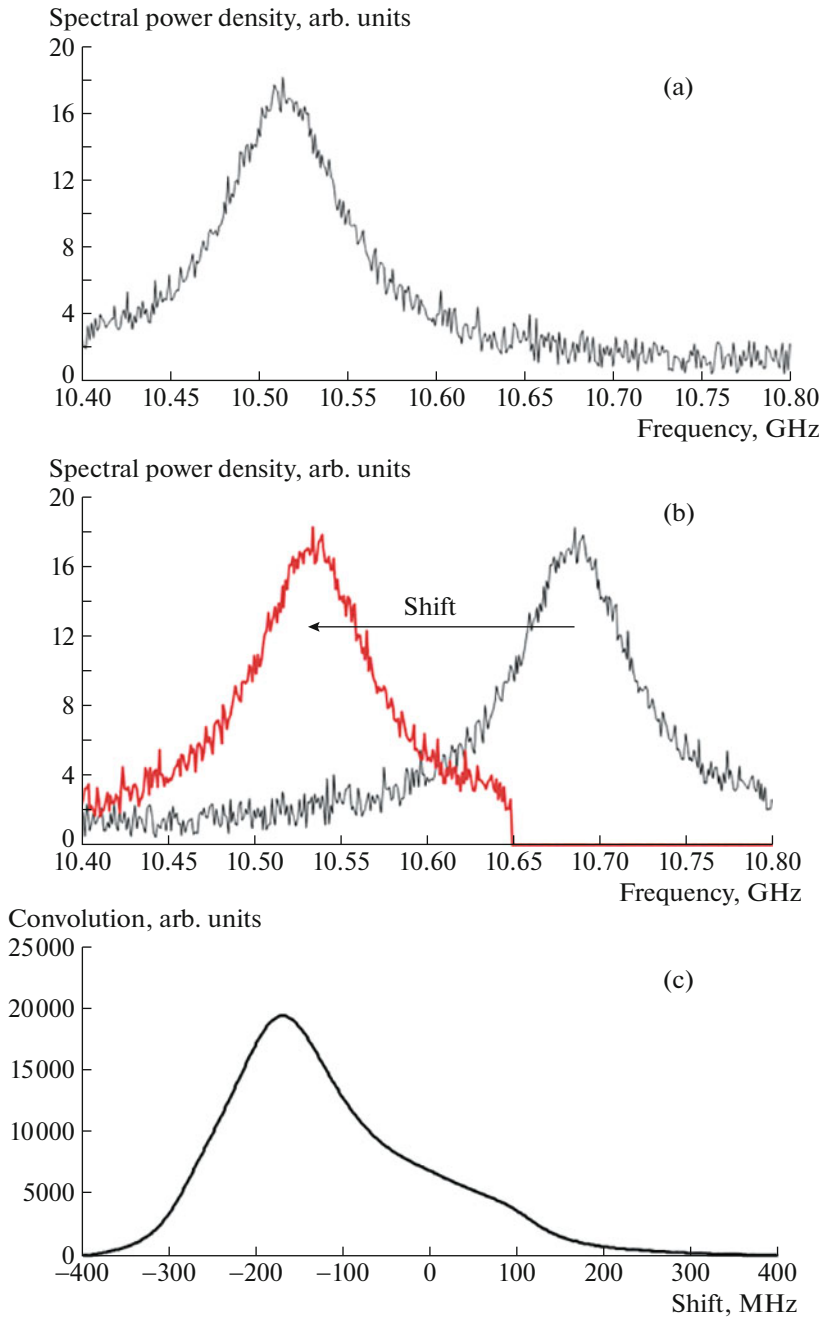

Fig. 1. Illustration of the BCM operation: (a) initial spectrum (the spectral power density is expressed in arbitrary units); (b) reflected nonshifted (right) and reflected shifted (left) spectra; and (c) correlation curve: the dependence of the convolution value on the shift magnitude.

METHODOLOGY

Despite the name the "backward correlation method," to find the optimal shift of the reflected spectrum relative to the original one, the search for the maximum of not correlation but convolution of two signals was used in [19]. In fact, the Pearson correlation coefficient of two discretely set signals F_1 and F_2 is equal to [21]

$$r = \frac{\langle F_1F_2\rangle - \langle F_1\rangle\langle F_2\rangle}{\sigma_{F_1}\sigma_{F_2}},$$

where $\langle\ \rangle$ is the averaging operator and σ are the variances of the corresponding signals.

During the shift, only $\langle F_1F_2\rangle$ changes; therefore, the maximum of the coefficient r must coincide with the maximum of $\langle F_1F_2\rangle$ and, therefore, with the maximum

of $\sum F_1^i F_2^i$, where summing is performed over all points. The expression $\sum F_1^i F_2^i$ is just the convolution of signals.

However, it should be noted that, with a shift, the signal F_2 may go beyond the boundaries of the definition area. As an example, let the signal F_1 consist of 400 values with indexes from 1 to 400. The reflected nonshifted signal consists of the same 400 points in the reverse sequence: $F_2^i = F_1^{400-i}$. With a shift by ten points to the left, F_2^{11} moves to F_2^1, F_2^{12} to F_2^2, etc., up to a transition of F_2^{400} to F_2^{390}. The extreme left points $F_2^1 - F_2^{10}$ do not move anywhere, while there is no correspondence to the points $F_2^{391} - F_2^{400}$ in the nonshifted signal. In order for the dimensions of F_1 and F_2 to coincide, the missing values of the shifted reflected signal (in this example, these are points from F_2^{391} to F_2^{400}) in [19] were filled with zeros.

Strictly speaking, with such a procedure, $\langle F_2 \rangle$ begins to depend on the shift (the greater is the shift, the larger number of nonzero initial elements are replaced by zero), and the above arguments in favor of the fact that the convolution maximum coincides with the correlation maximum lose their force. Hence, the first (I) approach arose: to check how the method behaves when replacing the search for the convolution maximum with the search for the Pearson correlation maximum.

Subsequently, it was additionally found that another problem arises when processing actual spectra with a very low SNR. The authors did not encounter it in [19], because a noise component uniformly distributed in the range of $[-Ns_{\max}, Ns_{\max}]$ was taken when modeling, where $Ns_{\max}$ is the maximum noise value, i.e., the average noise value was zero. In actual spectra, the values of the spectral power density cannot be negative; in addition, a statistical analysis has shown that the noise is distributed according to a normal but not an equiprobable law.

Figure 2 shows the correlation curves of the dependence of the convolution on the shift for various SNRs when processing a signal with nonzero average noise. For a continuous initial spectrum that does not contain noise and is defined in the entire frequency range $[0, \infty)$, the correlation curve should be an ideal Lorentzian function with a width equal to the double width of the initial spectrum (a detailed derivation of this relationship is given, e.g., in [20]). In reality, firstly, there is not the entire spectrum but only a part of it that falls within the frequency scanning range (for this reason, we have to add zero elements at the edges upon a shift); secondly, the spectrum is not continuous but discrete; and thirdly, it contains nonzero noise. All these factors lead to a deviation of the real correlation curve from the ideal Lorentzian function. However, at moderate SNRs, replacing the missing elements with zeros leads only to a curve break and does not affect the search for the maximum in any way, while the picture changes dramatically at very low SNRs (see Fig. 2). The maximum convolution does not now depend at all on the signal position but always corresponds to a zero shift (the curve for the SNR = 1 dB in Fig. 2d). This fact is explained quite simply. The contribution from the products of useful signals is quite small due to the low SNR, and the "parasitic" contribution due to the products of noise components increases with a decrease in the artificially introduced zero signal F_2 values. With a zero shift, such zero values do not exist at all, and the convolution has the maximum value. Accordingly, the BCM in the form that was presented in [19] completely loses its operability.

However, if the replacement of missing elements was performed not by zeros but by the average noise value in the original spectrum, then convolutions from weakly shifted signals would cease to obviously win in value correlative to convolutions from strongly shifted signals. In this case, the impact of the contribution from the product of the missing elements (from F_2^{391} to F_2^{400} in the above-considered example) on the corresponding elements F_1 is approximately the same (accurate within the statistical error) as the contribution from other noise elements. Or, what is absolutely equivalent, it is possible to subtract the average noise from the initial signal and then apply the algorithm used in [19]. This is the second approach (II) (Fig. 3): to work not with the original spectrum but with a preprocessed one, from which the average noise has already been subtracted. Of course, such a signal looks a bit unusual from the point of view of physics: the spectral power density takes negative values at certain points, but the addition of zeros upon a shift does not lead to a strong signal change.

The example shown in Fig. 3 demonstrates the effectiveness of this approach. The maximum of the correlation curve after preprocessing of the spectrum no longer falls on the zero shift; it returned to where it was at high SNRs.

The problem of determining the average noise is proposed to be solved as follows. The width of the Lorentzian peak in the Brillouin scattering spectrum is approximately known; accordingly, the number of points at which the useful signal is of great importance is known. Excluding from consideration, for instance, two or three times the number of points in the original spectrum with the highest values and averaging the value of the spectral power density over the remaining ones, we obtain the average noise. In this case, it does not matter that the remaining points may not form a continuous range (after excluding the largest elements, points from the useful signal area may remain and, conversely, those from the noise area may not

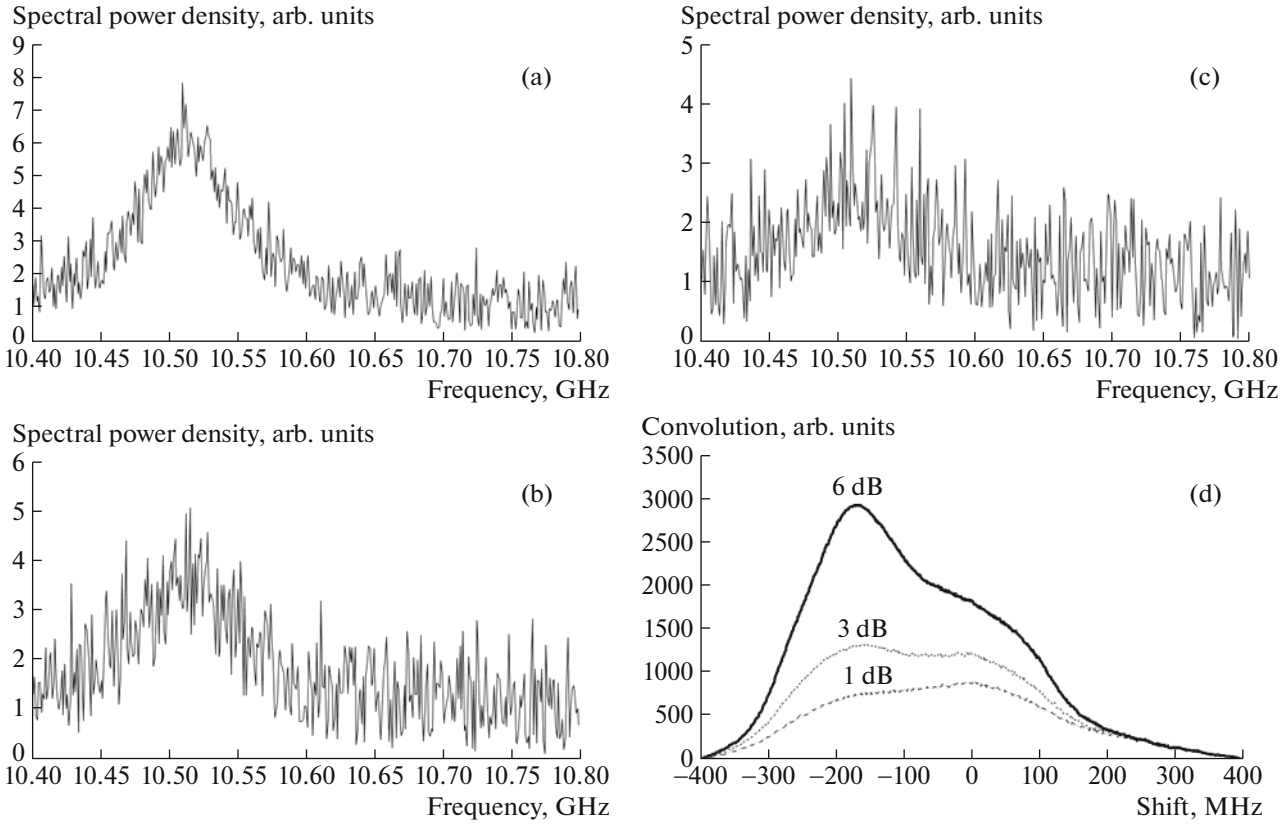

Fig. 2. Evolution of the correlation curve with decreasing the SNR: (a–c) initial spectra with the same maximum position with decreasing SNRs: (a) 6, (b) 3, and (c) 1 dB; (d) corresponding correlation curves.

remain), since it does not matter whether we dropped a point from the noise region or a point with the same value from the signal area. This algorithm was tested on the generated spectra with different SNRs and gave excellent results when excluding both two and three times as many points as the Lorentzian peak width.

Finally, the third approach (III) is to combine the two previous ones: to work with a preprocessed spectrum and look for the maximum correlation.

All three approaches were tested as follows.

1. A certain SNR was set.

2. 1000 spectra were generated with a given SNR and a random position of the Brillouin scattering maximum. The noise component had a nonzero average value and was distributed according to the normal law.

3. For each of the 1000 spectra, all three new methods, as well as the old method from [19], were used to determine the value of the Brillouin peak frequency.

4. The absolute difference between the found and generated Brillouin-peak frequency value was averaged for each of the methods.

5. The procedure was repeated for other SNRs.

All modeling was performed using software specially developed by the authors, which was modified in comparison with the software used when obtaining the results described in [19]. An object-oriented code in the C# language was used.

RESULTS AND DISCUSSION

Figure 4 shows the obtained dependences of the averaged error in determining the BFS on the SNR with the following parameters of the initial spectrum: the scanning frequency range is 10 400–10 800 MHz, the scanning pitch is 4 MHz (i.e., the spectrum consists of 100 points), the half-width of the Brillouin spectrum is 40 MHz, and the Brillouin scattering peak has a random value in the range of 10 450–10 750 MHz.

It can be seen that the usual BCM loses its operability, starting from an SNR of approximately 5 dB, while all three of its modifications at this moment still provide an acceptable accuracy in determining the BFS. The gain in the SNR range is ~5 dB.

There is no significant difference in the results of all three modifications. The authors explain this by the fact that any of the approaches (I or II) eliminates the problem associated with replacing nonzero noise points with zero. This is due to the fact that the signal is not distorted at all during replacement in approach II, while the drop in the convolution of signals when replacing nonzero points with zero is compensated by

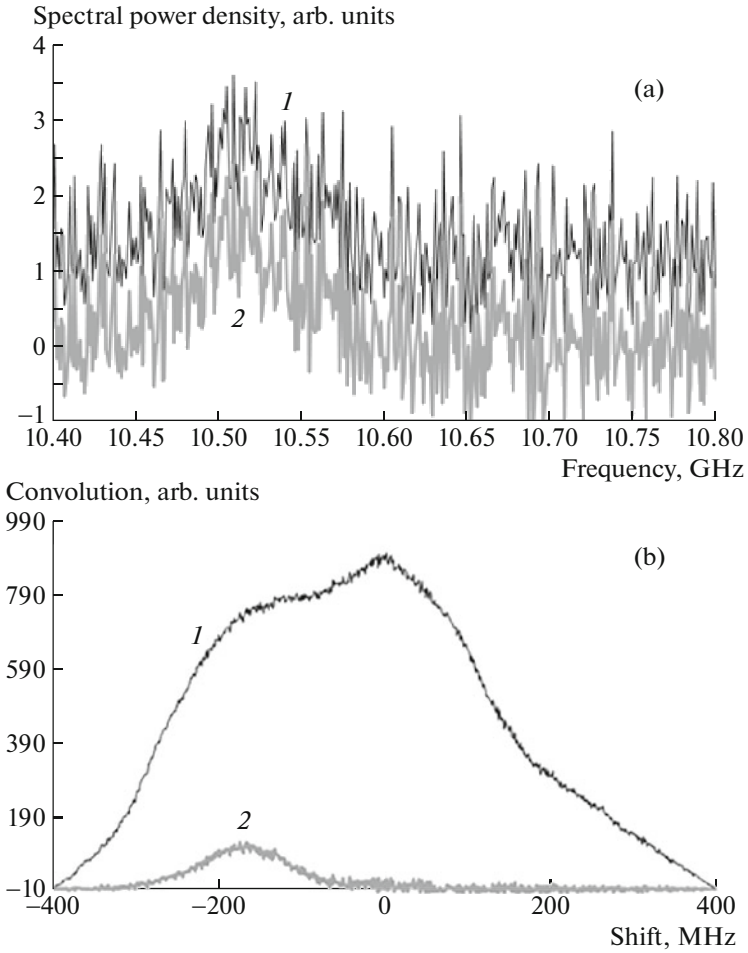

Fig. 3. Illustration of the operation of the approach II-modified BCM: (a) (*1*) initial spectrum and (*2*) spectrum processed by subtracting the average noise; (b) correlation curves obtained when processing the (*1*) initial and (*2*) processed spectra.

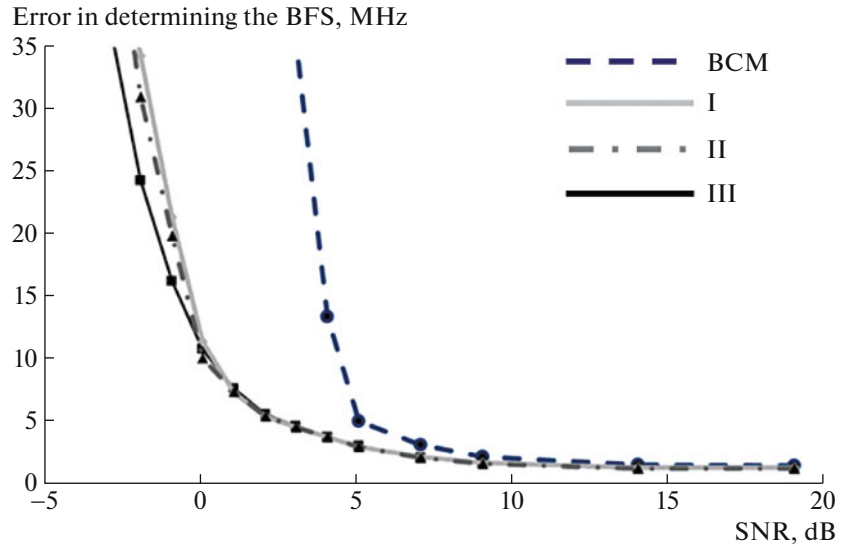

Fig. 4. Dependence of the averaged error in determining the BFS for the scanning range of 400 MHz and a scanning pitch of 4 MHz for the BCM [19] and its modifications I–III.

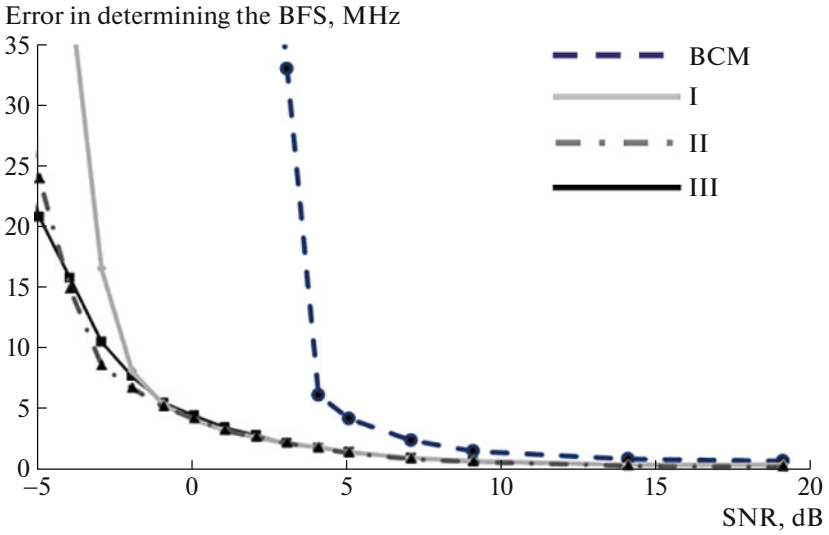

Fig. 5. Dependence of the averaged error in determining the BFS for a scanning range of 400 MHz and a scanning pitch of 0.5 MHz for the BCM [19] and its modification I–III.

a proportional change in $\langle F_2 \rangle$ in approach I. Accordingly, approach III does not provide an additional accuracy gain anymore.

Figure 5 contains similar results but at a scanning pitch of 0.5 MHz (800 points in the spectrum).

The results are similar to the results of the previous experiment, but the gain in the SNR range is much greater; thus, the modified methods continue to yield an acceptable result even in the region of negative SNRs.

It should be noted that there was no additional gain in the SNR compared to the results presented in [19]. However, the results from [19] were obtained only on spectra with a zero average noise value, and the accuracy of conventional BCM decreases significantly when processing actual spectra, while modified methods operate with any noise.

In conclusion, it is worth answering the question: which of the presented modifications of the BCM is the most effective? From the point of view of the number of arithmetic operations, approach II is optimal since fewer operations are required to calculate the convolution of signals than to calculate the Pearson correlation. However, the computational costs of all methods are relatively small (e.g., the entire calculation of data for Fig. 4 takes less than 5 min on a budget office computer); thus, it is logical to use the simplest approach I, especially since, in this case, it does not require calculating the average noise value.

CONCLUSIONS

In this paper, variants of modification of the BCM for processing the spectra obtained in experiments with nonzero average noise were studied. The results obtained with the help of modeling indicate that the modified method continues to be effective in the region of low SNRs, where the usual BCM described in [19] already ceases to work. The gain in the SNR range may be several decibels and depends on the number of points in the spectrum: the more points, the greater is the gain.

Since the correlation curves in the modified method do not contain kinks that are characteristic of conventional BCM [19], it is planned to combine the modified method with the Lorentzian spectrum reconstruction method in the future. Perhaps, this will allow us to correctly extract the BFS from spectra with an even lower SNR: the lower curve in Fig. 4b has a significantly higher SNR than the lower curve in Fig. 4a; therefore, the Lorentzian reconstruction of the shape should give better results. As already noted, attempts to integrate the usual BCM into the chain of processing methods have not brought significant improvements [20].

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

This study was supported by the State Contract no. AAAA-A19-119042590085-2.

ACKNOWLEDGMENTS

We are grateful to A.I. Krivosheev for fruitful discussions and A.R. Davydov for his help in conducting statistical calculations when revealing the law of distribution of the noise spectral component.

REFERENCES

1. Bao, X., Webb, D.J., and Jackson, D.A., *Opt. Lett.,* 1993, vol. 18, p. 1561. https://doi.org/10.1364/OL.18.001561
2. Soto, M.A. and Thévenaz, L., *Opt. Express,* 2013, vol. 21, p. 31347. https://doi.org/10.1364/OE.21.031347
3. Feng, C., Preussler, S., Kadum, J., and Schneider, T., *Sensors,* 2019, vol. 19, p. 2878. https://doi.org/10.3390/s19132878
4. Li, C., Lu, Y., Zhang, X., and Wang, F., *Electron. Lett.,* 2012, vol. 48, no. 18, p. 1139. https://doi.org/10.1049/el.2012.1248
5. Urricelqui, J., Sagues, M., and Loayssa, A., *Opt. Express,* 2014, vol. 22, no. 15, p. 18195. https://doi.org/10.1364/OE.22.018195
6. Zhou, F., Gan, J., Lv, H., and Cui, L., *IOP Conf. Ser.: Earth Environ. Sci.,* 2018, vol. 189, p. 032026. https://doi.org/10.1088/1755-1315/189/3/032026.
7. Feng, C., Lu, X., Preussler, S., and Schneider, T., *J. Lightwave Technol.,* 2019, vol. 37, p. 5231. https://doi.org/10.1109/JLT.2019.2930919
8. Li, C. and Li, Y., *Proc. 2009 5th Int. Conference on Wireless Communications, Networking and Mobile Computing,* Beijing, 2009, p. 24. https://doi.org/10.1109/WICOM.2009.5303692.
9. Yan, Z., Zhong, S., Lin, L., and Cui, Z., *Mathematics,* 2021, vol. 9, p. 2176. https://doi.org/10.3390/math9172176
10. Amini, K. and Rostami, F., *J. Comput. Appl. Math.,* 2015, vol. 288, p. 341. https://doi.org/10.1016/j.cam.2015.04.040
11. Horiguchi, T., Masui, Y., and Zan, M., *Sensors,* 2019, vol. 19, p. 1497. https://doi.org/10.3390/s19071497
12. Farahani, M.A., Castillo-Guerra, E., and Colpitts, B.G., *Opt. Lett.,* 2011, vol. 36, p. 4275. https://doi.org/10.1364/OL.36.004275
13. Ruiz-Lombera, R., Fuentes, A., Rodriguez-Cobo, L., Lopez-Higuera, J.M., and Mirapeix, J., *J. Lightwave Technol.,* 2018, vol. 36, p. 2114. https://doi.org/10.1109/JLT.2018.2805362
14. Lalam, N., Venketeswaran, A., Lu, P., and Buric, M.P., in *Optical Interconnects XXI*, Schröder, H. and Chen, R.T., Eds., Bellingham, WA: SPIE, 2021, vol. 11692, p. 1169213. https://doi.org/10.1117/12.2578509.
15. Wu, H., Wan, Y., Tang, M., Chen, Y., Zhao, C., Liao, R., Chang, Y., Fu, S., Shu, P.P., and Li, D., *J. Lightwave Technol.,* 2019, vol. 37, p. 2648. https://doi.org/10.1109/JLT.2018.2876909
16. Karapanagiotis, C., Wosniok, A., Hicke, K., and Krebber, K., *Sensors,* 2021, vol. 21, p. 2724. https://doi.org/10.3390/s21082724
17. Nordin, N.D., Zan, M.S.D., and Abdullah, F., *Photonics,* 2020, vol. 7, p. 79. https://doi.org/10.3390/photonics7040079
18. Nordin, N.D., Zan, M.S.D., and Abdullah, F., *Opt. Fiber Technol.,* 2020, vol. 58, p. 102298. https://doi.org/10.1016/j.yofte.2020.102298
19. Barkov, F.L., Konstantinov, Y.A., and Krivosheev, A.I., *Fibers,* 2020, vol. 8, p. 60. https://doi.org/10.3390/fib8090060
20. Nordin, N.D., Abdullah, F., Zan, M.S.D., Bakar, A.A., Krivosheev, A.I., Barkov, F.L., and Konstantinov, Y.A., *Sensors,* 2022, vol. 22, p. 2677. https://doi.org/10.3390/s22072677
21. Konstantinov, Yu.A., Kryukov, I.I., Pervadchuk, V.P., and Toroshin, A.Yu., *Quantum Electron.,* 2009, vol. 39, no. 11, p. 1068. https://doi.org/10.1070/QE2009v039n11ABEH014171

Translated by A. Seferov