JASA ARTICLE

The spectrogram, method of reassignment, and frequency-domain beamforming

Gil Averbuch[a)]
Roy M. Huffington Department of Earth Sciences, Southern Methodist University, Dallas, Texas 75205, USA

ABSTRACT:
A smeared spectrogram is a result of the smoothing kernel in the short-time Fourier-transform (STFT). Besides the smeared energy, time and frequency phase information is also smeared, i.e., spectral components may contain imprecise phase information. The STFT is also used as the basis for more advanced signal processing techniques such as frequency-domain beamforming and cross correlation (CC). Both methods seek the delay time between signals by exploring phase-shifts in the frequency domain. Due to the inexact phase information in some of the time-frequency elements, their phase shifts are incorrect. This study re-introduces the reassigned spectrogram (RS) as a measure to fix the STFT artifacts. Moreover, it is shown that by using the RS, phase shifts can be optimized and improve beamforming and CC results. Synthetic and recorded data are used to show the advantage of using the RS in time-frequency analysis, CC, and beamforming. Results show that, subject to certain constraints, the RS provides exact time-frequency representation of deterministic signals and significantly improve CC and beamforming results. Array analysis of infrasonic signals shows that better results are obtained by either the RS- or STFT-based analysis depending on the signals' spectral components and noise levels. © *2021 Acoustical Society of America.*
https://doi.org/10.1121/10.0003384
(Received 5 October 2020; revised 29 December 2020; accepted 6 January 2021; published online 1 February 2021)
[Editor: Karim G. Sabra] Pages: 747–757

I. INTRODUCTION

The spectrogram is a time-frequency representation (TFR) of non-stationary signals, i.e., frequency- and amplitude-modulated signals. It was developed during the 1940s and, due to its robustness and intuitive formulation, has remained widely used (Bracewell, 2000; Cohen and Loughlin, 2003; Koenig *et al.*, 1946). The spectrogram is defined as the squared modulus of the short-time Fourier transform (STFT), which reads as

$$S(t_0,\omega_0) = \int x(t)h(t_0-t)e^{-i\omega_0 t}dt = e^{-i\omega_0 t_0}\int X(\omega)H(\omega-\omega_0)e^{i\omega t_0}\frac{d\omega}{2\pi}. \quad (1)$$

Here, $x(t)$ is a signal, and $h(t)$ is a smoothing kernel; their frequency-domain forms are $X(\omega)$ and $H(\omega)$, respectively, and the integrals are defined as $\int = \int_{-\infty}^{\infty}$. The squared modulus, $|S(t_0,\omega_0)|^2$, can be understood as a time-frequency energy distribution of the signal, smoothed by the energy distribution of the window $h(t)$. For convenience, TFR values will be presented in terms of absolute values, i.e., $|S(t_0,\omega_0)|$, which later will be combined with phase information and used in the beamforming method (Sec. III). Also, angular frequency, $\omega = 2\pi f$, is used for derivations while the results are presented in terms of frequency, f.

The spectrogram and the STFT are fundamental signal processing tools in many research fields. They are used for time-frequency analysis and as a basis for more advanced analysis methods. For example, in linguistics, the spectrogram is used to characterize the communication and speech of humans and animals (Flanagan, 1972; Mellinger and Clark, 2000; Potter *et al.*, 1994), and in medical imaging, it is used to describe skin vibrations and blood flow (Chen *et al.*, 2017; Jensen, 2006). The spectrogram and STFT-based methods are also used to monitor and characterize sources of anthropogenic noise like aircrafts, sirens, and ships (Angione *et al.*, 2016; Garibbo *et al.*, 2020; Merino-Martínez *et al.*, 2019; Rossi-Santos, 2020; Young, 1973). In underwater acoustics, TFRs are used to detect earthquake signals and describe the nature of their propagation (Evers *et al.*, 2014); TFRs are also used to study wind-induced oceanic waves (Guedes Soares and Cherneva, 2005). Infrasound (low-frequency sound) monitoring is used to detect anthropogenic and geophysical sources of sound in the atmosphere, study their nature, and characterize the atmosphere's state. Such sources can be explosions (Arrowsmith and Bowman, 2017), sprites (Applbaum *et al.*, 2020), ocean waves (Smets and Evers, 2014), and earthquakes (Shani-Kadmiel *et al.*, 2018). In seismology, spectral analysis is used to detect and distinguish between different types of seismic waves (Gal *et al.*, 2016), and in the field of electromagnetism, spectrograms are used to detect radio waves and describe the magnetosphere properties (Smith and Angerami, 1968).

Although the STFT and spectrogram are widely used, some inherent characteristics prevent obtaining a localized TFR. First, in analogy to quantum mechanics' uncertainty

[a)]Electronic mail: gaverbuch@smu.edu, ORCID: 0000-0002-0403-9354.

principle, the Heisenberg-Gabor limit states that it is impossible to have a localized signal in time and frequency (Cohen, 1995; Gabor, 1946); a high spectral resolution requires long time windows, which yield low temporal resolution. In contrast, a high temporal resolution requires short time windows, which yield a low spectral resolution. Second, the spectrogram's time and frequency marginals are not satisfied due to the running window, $h(t)$, which leads to imprecise representations of the signal's bandwidth and duration. However, energy can still be conserved by using a normalized smoothing window (Cohen, 1995; Cohen and Loughlin, 2003).

The simplest example that demonstrates the problem is a spectrogram of the stationary-phase signal, $\exp(i\omega_0 t)$. A Fourier transform of a time window that contains an integer number of cycles yields a delta function at ω_0. However, a spectrogram with the same window size leads to a smeared spectrum. Such spurious spectral distribution can be seen in Fig. 1(b). Throughout this study, a normalized Hanning window is used as a smoothing kernel.

Besides obtaining smeared TFRs, the artifacts caused by the STFT have additional impact on data analysis techniques. Methods like frequency-domain beamforming and cross correlation (CC) are based on the STFT. When beamforming, recorded signals are phase-shifted and summed to maximize the signal power and suppress the noise. It will be shown (Sec. II) that the STFT phase information, i.e., frequency and time, does not always correspond to its TFR location. Therefore, some of the beamforming phase-shifts

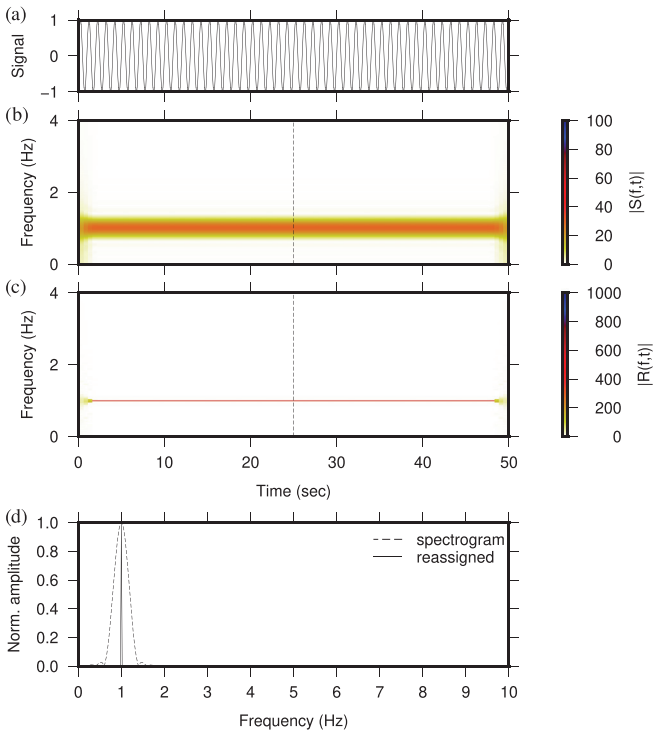

FIG. 1. (Color online) One Hz stationary phase function example. a) Signal. b) Spectrogram. c) Reassigned spectrogram. d) Normalized cross sections (dashed line in b and c) of the spectrogram and reassigned spectrogram.

are inaccurate and prevent maximizing the signal power. Similarly, CC phase-shifts suffer from the same effect. Before continuing, it is important to stress that this **does not** mean that previous work is wrong. The difference between the performed phase-shifts and the one needed is relatively small, and numerous excellent studies have properly used both beamforming and CC techniques.

The method of reassignment corrects for the smeared TFR that is caused by the smoothing kernels of the STFT. In principle, this method uses the time and frequency first moments to calculate the distribution's center of gravity coordinates. Then, the energy from each TFR element is reassigned to the center of gravity, i.e., the reassignment coordinates. Mathematically, the computed reassigned coordinates correspond to the signal's phase information (Kodera *et al.*, 1976). This method repairs the STFT artifacts and sharpens the spectrogram; it yields the reassigned spectrogram (RS). As with the STFT, the RS is also subjected to the Heisenberg-Gabor limit, therefore the resolution, i.e., the ability to resolve two nearby peaks, stays the same. Figure 1(c) shows the RS of the previous stationary-phase signal. As expected, the RS of the signal provides a delta function in the frequency domain [Fig. 1(d)].

The remainder of the text is organized as follows. Section II describes the method of reassignment. Two formulations for computing the reassigned coordinates are presented, followed by numerical examples comparing the STFT-based spectrogram and the RS. Frequency domain beamforming is discussed in Sec. III, and beamforming results, based on the STFT and the RS, are presented. In addition, results of STFT- and RS-based CC are compared. Finally, we discuss the presented results and its implications in Sec. IV.

II. THE REASSIGNED SPECTROGRAM

A. Theory

Our analysis begins with the Wigner-Ville Distribution (WVD). It provides a high TFR resolution for non-stationary signals in the time-frequency plane and will be used to derive the spectrogram and the reassigned coordinates (Bracewell, 2000; Cohen, 1989; Debnath, 2002). The WVD of a signal $s(t)$ is defined as

$$W(s;t_0,\omega_0) = \int s(t_0+\tau/2)s^*(t_0-\tau/2)e^{-i\omega_0\tau}d\tau. \qquad (2)$$

For a specific set of deterministic signals, the WVD provides exact TFRs. For example, the WVD of an impulsive signal $\delta(t-t_0)$ is a vertical line along the frequency axis at t_0. A chirp signal of the form $e^{i(\omega_0 t+\alpha t^2/2)}$ has a WVD that corresponds to $\delta(\omega-\omega_0-\alpha t)$ (Bracewell, 2000). However, a monochromatic cosine signal with angular frequency ω_0 has a distorted WVD with horizontal lines at $\omega=\pm\omega_0$, and a spurious component at $\omega=0$ (Cohen, 1995). Linear combinations of the mentioned signals also yield spurious TFR components. The limited signals for which the WVD

https://doi.org/10.1121/10.0003384

provides accurate TFRs, and the broad range of signals for which it provides inaccurate representations, limit its practical use.

A way to suppress these artifacts is by convolving Eq. (2) with a two-dimensional smoothing kernel, $\Phi(t,\omega)$. The smoothed WVD is defined as

$$W_s(s;t_0,\omega_0) = \int\int \Phi(t,\omega)W(s;t_0-t,\omega_0-\omega)dt\frac{d\omega}{2\pi}. \tag{3}$$

This operation preserves time and frequency invariance while reducing the accuracy of the localized WVD (Auger and Flandrin, 1995; Cohen, 1989). Moreover, by defining the smoothing kernel, $\Phi(t,\omega)$, to be the WVD of a smoothing window, $h(t)$, it can be shown that the smoothed WVD is identical to the well-known spectrogram (Auger and Flandrin, 1995; Frazer and Boashash, 1994; Stankovic, 1994)

$$\begin{aligned} W_{sp}(s;t_0,\omega_0) &= \int\int W(h;t,\omega)W(s;t_0-t,\omega_0-\omega)dt\frac{d\omega}{2\pi} \\ &= |S(t_0,\omega_0)|^2. \end{aligned} \tag{4}$$

A way to overcome the smeared distribution is by the reassignment method. First proposed by Kodera *et al.* (1976) and then furthered developed by Auger and Flandrin (1995), the method shows that the local centers of gravity of the spectrogram correspond to the phase components of the signal. In addition, spectrogram elements in the vicinity of the local centers of gravity may contain phase information with the original signal components. It means that some time-frequency elements contain phase information that corresponds to a different time and frequency, i.e., to the local center of gravity (this concept will become clearer in Sec. II B). The method reassigns energy from time-frequency elements to their local center of gravity. By that, a sharp and accurate TFR of the signal is obtained. Auger and Flandrin (1995) show that reassignment preserves time and frequency invariance and energy conservation, and can localize chirps and impulsive signals.

By computing the time first moment of Eq. (4), the reassigned time coordinate is expressed as

$$\hat{t}(t_0,\omega_0) = t_0 - \frac{\int\int tW(h;t,\omega)W(s;t_0-t,\omega_0-\omega)dtd\omega}{\int\int W(h;t,\omega)W(s;t_0-t,\omega_0-\omega)dtd\omega}. \tag{5}$$

Similarly, the reassigned frequency coordinate is expressed as

$$\hat{\omega}(t_0,\omega_0) = \omega_0 - \frac{\int\int \omega W(h;t,\omega)W(s;t_0-t,\omega_0-\omega)dtd\omega}{\int\int W(h;t,\omega)W(s;t_0-t,\omega_0-\omega)dtd\omega}. \tag{6}$$

Let a signal be composed of N amplitude and frequency modulated components

$$s(t) = \sum_n A_n(t)e^{i(\omega_n t+\phi_n)}, \tag{7}$$

where $n = 1,\ldots,N$, and $\omega_n = \omega_n(t)$. If the components fulfill a separability condition, the STFT yields the analytic signal, i.e., the amplitude and frequency modulated components. It is known that the analytic signal can also be obtained using the Hilbert transform of the waveforms. However, the presented methods to compute the reassigned coordinates are based on the STFT. The separability condition is based on Gabor's uncertainty principle and states that within the vicinity of ω_0 and t_0, there is only one dominant component. Mathematically, the condition is given by

$$\begin{aligned} &|S_n(t_0,\omega_0)|^2 \gg \sum_{l\neq n}|S_l(t,\omega)|^2, \quad |\omega-\omega_0| \ll \epsilon, \\ &|t-t_0| \ll \delta, \end{aligned} \tag{8}$$

where ϵ and δ are small parameters. Under this assumption, the argument of the STFT has the form

$$\arg\{S_n(t,\omega)\} = (\omega_n t + \phi_n), \tag{9}$$

up to a $2\pi m$ factor, where m is an integer. In addition, Kodera *et al.* (1976) show that Eqs. (5) and (6) can be simplified to

$$\hat{t}(t,\omega) = -\frac{\partial}{\partial\omega}\arg\{S(t,\omega)\} \tag{10}$$

and

$$\hat{\omega}(t,\omega) = \omega + \frac{\partial}{\partial t}\arg\{S(t,\omega)\}, \tag{11}$$

respectively. Compared to Eqs. (5) and (6), these compact expressions have a clear physical meaning; Eqs. (10) and (11) show that a local center of gravity is located at the instantaneous time and frequency of the signal. Equally important, these equations imply that some time-frequency elements contain phase information that does not correspond to their coordinates. It will be shown that this characteristic prevents maximizing beamforming results.

Numerically evaluating the phase derivatives introduces challenges in the form of phase-unwrapping, accuracy in the presence of noise, and parting components that violate the separability condition (Fulop and Fitz, 2006; Nelson, 2001). A proposed way to overcome these challenges is by an implicit evaluation of the phase derivatives. Auger and Flandrin (1995) show that Eqs. (5) and (6) can be expressed in terms of STFT with different window functions. The reassigned time coordinate can be formulated as

$$\hat{t}(t,\omega) = t - Re\left\{\frac{S_{\tau h}(t,\omega)\cdot S^*(t,\omega)}{|S(t,\omega)|^2}\right\}, \tag{12}$$

and the reassigned frequency coordinate as

$$\hat{\omega}(t,\omega) = \omega + Im\left\{\frac{S_{dh}(t,\omega)\cdot S^*(t,\omega)}{|S(t,\omega)|^2}\right\}, \tag{13}$$

where $S_{\tau h}$ and S_{dh} correspond to STFTs with smoothing kernels $t\cdot h(t)$ and dh/dt, respectively. These alternative expressions of the phase derivatives are expected to be more reliable (Papandreou-Suppappola, 2003, chapter 5). For the tested signals in Secs. II B and III, both methods provided similar results. A detailed derivation of Eqs. (10)–(13) can be found in the Appendix.

Regardless of the chosen method, the STFT values are reassigned from the original coordinates to the new coordinates as $|S(t,\omega)| \rightarrow R[\hat{t}(t,\omega),\hat{\omega}(t,\omega)]$. The RS can be expressed as

$$R(t',\omega') = \int\int |S(t,\omega)|\delta(t'-\hat{t}(t,\omega)) \times \delta(\omega'-\hat{\omega}(t,\omega))dtd\omega. \tag{14}$$

Numerical recipes of the methods can be found in Nelson (2001), Papandreou-Suppappola (2003, chapter 5), and Fulop and Fitz (2006).

B. Examples

Here, the method of reassignment will be demonstrated on a selection of signals. Hopefully, this section will convince the reader that this method can accurately correct the spectrogram and prove that the smeared spectrum contains phase information that corresponds to the original signal components. Figure 2 shows the reassigned coordinates for a signal composed of two stationary-phase functions. The black line in the bottom frame represents a normalized spectrogram cross section. Peaks at 2 and 4 Hz correspond to the signal's frequencies. However, the smeared spectrum in their vicinity is a consequence of the STFT. One can see that elements of the spectrogram are reassigned to 2 and 4 Hz. According to Eq. (11), the reassigned coordinates prove that frequency bins that do not correspond to the signal's frequencies contain its phase information. For instance, phase information at 1.5 Hz corresponds to 2 Hz. It should be noted that TFR elements with very low amplitudes contain numerical noise, which leads to false reassignment, e.g., 3.5 Hz element is mapped to a frequency higher than 4 Hz. Therefore, a threshold value is needed for the method. Here, the spectrogram RMS is used as a threshold.

For the second example, two superimposed 50 s chirp signals are used. In order to obtain a high time resolution, the spectrogram is computed with a 2-s window and a 90% overlap. Figure 3(b) shows the spectrogram of the signal, and the black lines represent the chirps' phase variation. The RS shows a sharp TFR that mostly agrees with the input phase functions [Fig. 3(c)]; the presence of some spurious components between 20 and 25 s results from spectral components that do not satisfy the separability condition. A

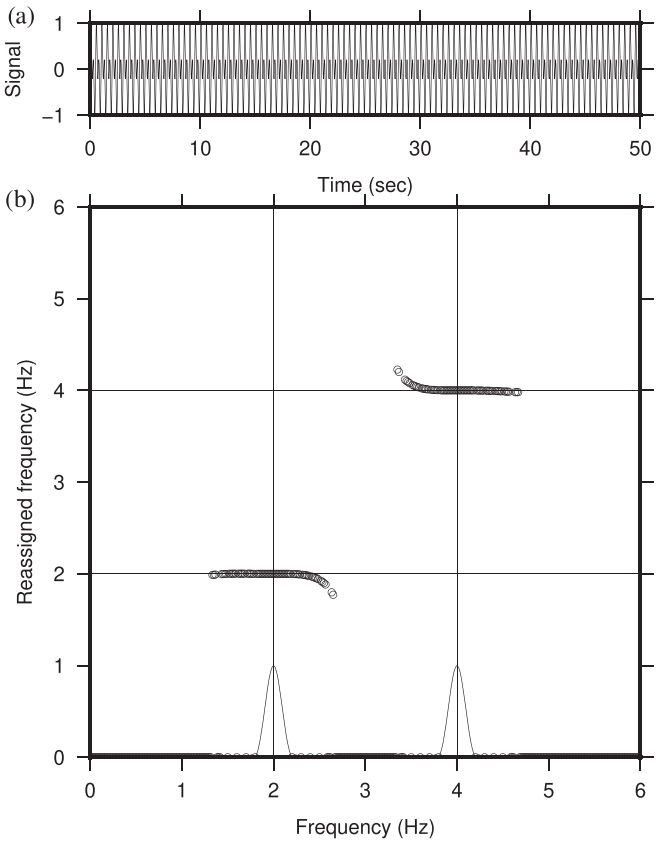

FIG. 2. Reassigned frequency coordinates. a) Superimposed 2 and 4 Hz stationary-phase signals. b) Reassigned frequency coordinates as a function of the spectrogram frequency coordinates. The black line corresponds to a normalized cross section of the signal's spectrogram. Black circles represent the reassigned frequencies' coordinates.

comparison between the TFRs' cross sections shows that the spectrogram misrepresents a broadband signal while it is clear that the signal has distinct narrowband components [Fig. 3(d)]. Also, it shows the RS low-amplitude artifacts that result from not satisfying the separability condition. These artifacts disappear once the separability condition is satisfied.

The third example uses a random signal that is filtered from 1 to 1.5 Hz. As in the previous example, a 2-s window with a 90% overlap is used. Figure 4(b) suggests a frequency range of almost 2 Hz, while Fig. 4(c) shows that the RS is mainly composed of narrowband components. Moreover, in some time elements, the RS appears to bifurcate. This artifact is attributed to either low amplitude components or dissatisfaction of the separability condition.

An essential aspect of data analysis is the choice of the processing parameters; each field with its unique signals has its rule of thumb for defining them. Nevertheless, an optimal choice is gained by experience, and the presented study is no exception. In the previous examples, a short time window was used to gain high temporal resolution at the expense of an initial smeared spectrum in the frequency axes. It was found that accurate, consistent results are obtained by correcting the smeared frequency. Conversely, a long window that provides a higher resolution in frequency with a

https://doi.org/10.1121/10.0003384

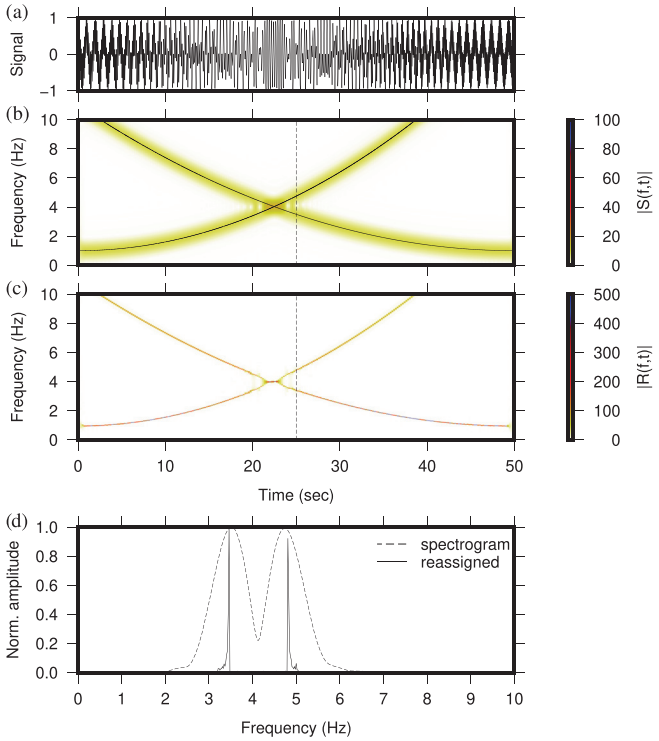

FIG. 3. (Color online) Two chirps example. (a) Signal. (b) Spectrogram. The solid black line corresponds to the true phase function. (c) Reassigned spectrogram. (d) Normalized cross sections (dashed line in b and c) of the spectrogram and reassigned spectrogram.

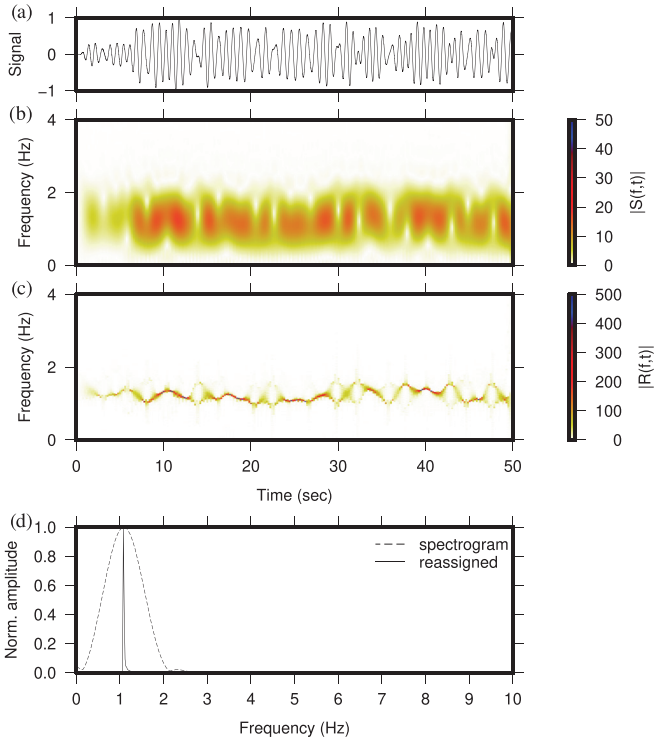

FIG. 4. (Color online) One to 1.5 Hz signal example. (a) Signal. (b) Spectrogram. (c) Reassigned spectrogram. (d) Normalized cross sections (Dashed line in b and c) of the spectrogram and reassigned spectrogram.

smeared spectrum in the time axes could not be rectified appropriately by the reassignment method.

III. APPLICATION TO BEAMFORMING

Frequency-wavenumber (f-k) analysis, or frequency domain beamforming, is an array analysis technique that estimates the recorded waves' parameters. These parameters consist of back azimuth (BA), apparent velocity (AV), and signal-to-noise ratio (SNR). The BA points to the direction of arrival, and AV indicates the plane wave inclination angle. In this study, the algorithm proposed by Smart and Flinn (1971) is used. The algorithm is based on a f-k search (Bartlett, 1948) combined with the Fisher statistic. It is the frequency domain equivalent to the delay-and-sum beamformer (Melton and Bailey, 1957), which provides a statistical framework to estimate the probability of detection.

Assuming a propagating plane wave of the form

$$g(\boldsymbol{x},t) = Ae^{i(\boldsymbol{k}\boldsymbol{x}-\omega t)}, \tag{15}$$

which crosses an N elements array located at $\boldsymbol{x}_n(i = 1, 2, \ldots, N)$. The wave vector

$$\boldsymbol{k} = \begin{bmatrix} k_x \\ k_y \\ k_z \end{bmatrix} = \frac{\omega}{c}\begin{bmatrix} \cos(\theta)\cos(\phi) \\ \cos(\theta)\sin(\phi) \\ \sin(\theta) \end{bmatrix} \tag{16}$$

describes the direction of propagation, and the wavenumber is defined as $k = \omega/c$. The angles θ and ϕ correspond to the grazing angle and azimuth from the north, respectively, and c is the propagation velocity. Defining $S(t, \omega, \boldsymbol{x}_n)$ as the STFT of the recording at location $\boldsymbol{x}_n$, the signal power across the array is computed by

$$P_S(t,\omega,\boldsymbol{k}) = \left|\frac{1}{N}\sum_{n=1}^{N} S(t,\omega,\boldsymbol{x}_n)e^{-i\boldsymbol{k}\cdot\boldsymbol{x}_n}\right|^2. \tag{17}$$

The kernel $e^{-i\boldsymbol{k}\cdot\boldsymbol{x}_n}$, also known as the steering vector, can be read as the phase shift to the center of the array. Phase-shifting and summing the spectral components with the correct wavenumbers will maximize the sum. The total power is defined as

$$P_T(t,\omega) = \frac{1}{N}\sum_{n=1}^{N} |S(t,\omega,\boldsymbol{x}_n)|^2, \tag{18}$$

and the Fisher ratio is

$$F_{\text{ratio}}(t,\omega,\boldsymbol{k}) = \frac{P_S(t,\omega,\boldsymbol{k})}{P_T(t,\omega) - P_S(t,\omega,\boldsymbol{k})}(N-1). \tag{19}$$

The wavenumbers that give the highest Fisher ratio indicate the parameters of the wave that most likely crossed the array. The Fisher ratio is also a measure of the single component SNR *via* $F_{\text{ratio}} = N \cdot \text{SNR}^2 + 1$. Detections can be determined by thresholding on F_{ratio} or SNR (Averbuch *et al.*, 2018).

Other formulations of weighted steering vectors provide different array responses and can improve the results

(Capon, 1969; Neele and Snieder, 1991). Moreover, post-processing algorithms can sharpen the results and detect multiple sources (den Ouden *et al.*, 2020; Schmidt, 1986). However, testing different detectors is not in the scope of this study, and only the presented beamforming formulation is used.

One can see that the phase-shifts in Eq. (17) are wavenumber dependent, and the wavenumbers are frequency-dependent [Eq. (16)]. A basic assumption in this process is that the phase information in $S(t,\omega,\boldsymbol{x}_n)$ corresponds to ω. Thus, the phase shifts and the phase information correspond to the same frequency. However, Sec. II shows that this assumption is not always accurate.

As stated in Sec. I, this characteristic does not imply any flaws in previous studies. The implication of the RS on the frequency domain beamforming is that the phase-shifts can be optimized. In what follows, a comparison between beamforming based on the STFT and RS is presented. The only difference in the beamforming algorithm is that the STFT spectrum, $S(t,\omega,\boldsymbol{x}_n)$, is replaced by $R(t,\omega,\boldsymbol{x}_n)$. However, the phase information needs to be incorporated. In Eq. (14), only the absolute values of the STFT are reassigned to $R(t,\omega)$. Now, the phase information of the assigned time-frequency elements is included, and phase shifting $R(t,\omega,\boldsymbol{x}_n)$ is possible.

Although this section's focus is beamforming, frequency-domain CC is also based on phase-shifts of the recordings. Therefore, the smeared STFT spectrum prevents performing accurate phase-shifts. For clarity, the frequency-domain CC reads

$$(f\star g)(t) = \int f^*(\tau)g(\tau+t)d\tau = \int F^*(\omega)G(\omega)e^{i\omega t}\frac{d\omega}{2\pi}. \tag{20}$$

To find the lag time, t, between time segments of recordings, the arguments $F(\omega)$ and $G(\omega)$ will be time bins of the STFT. Therefore, similarly to Eq. (17), the frequency of the phase shift in Eq. (20) and the functions phase information may differ.

The first beamforming and CC examples are based on a time-shifted signal. Two identical signals are used where one is time-shifted by 0.2 s. Random noise is added to them independently. The processing parameters are set to a five-second window with a 90% overlap. Figure 5(b) shows the average Fisher ratio based on the STFT and RS. Results based on the RS show a significant increase in the computed Fisher values. Moreover, the results are sharper and more precise, i.e., the lobe is smaller. CC results are presented in Fig. 5(c). As expected, the RS-based CC provides higher-amplitude correlation coefficients. However, the RS-based correlation curve is a scaled version of the STFT-based curve, and both curves' maxima point to the correct lag time.

While these results seem very promising, both methods' performance depends on the ability of the RS to provide

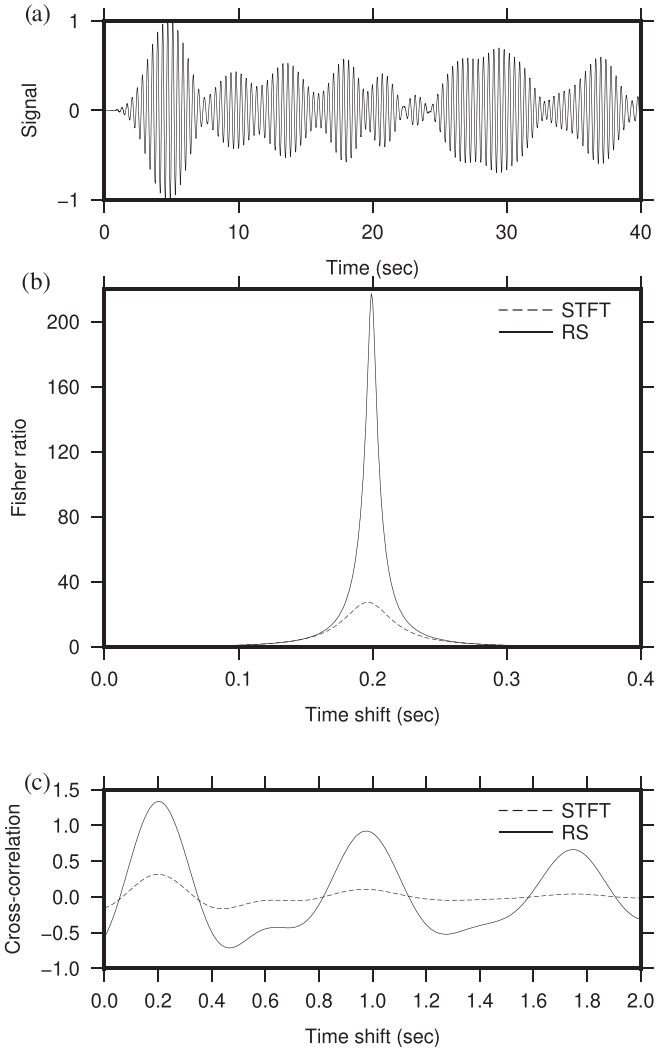

FIG. 5. Fisher ratio and cross correlation examples for two time-shifted signals. (a) Signal. (b) The average Fisher ratio calculated based on the STFT (dashed line) and RS (solid line). (c) The average cross correlation based on the STFT (dashed line) and RS (solid line).

accurate results. For example, in the presence of high noise levels or spectral components that do not fulfill the separability condition, the RS introduces artifacts. Such artifacts may lead to a worse beamforming and CC performance than with the STFT. Hence, the traditional methods that are STFT-based are more forgiving and, therefore, more robust. In addition, the time correction of the RS leads to some time bins that do not contain any signal components. In that case, the STFT-based methods will provide detections, and the RS-based methods will not.

In order to test the method on array data, the eight-element I18DK infrasound array is used. This array is part of the International Monitoring System (IMS), operated by the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). As in the previous example, the method is first tested with synthetic data that consists of a continuous signal in the frequency range of 0.2–0.3 Hz, AV of 340 m/s, and BA of 300 degrees. Due to the array aperture and the signal's frequency content, the processing parameters are set to

https://doi.org/10.1121/10.0003384

a 20-s window with a 90% overlap. Figure 6 shows the beamforming results based on the STFT and RS for two time bins: Figs. 6(a) and 6(c) are based on the RS, and Figs. 6(b) and 6(d) are based on the STFT. Time bins that fulfill the separability condition yield a focused RS and improve the traditional beamforming results [Fig. 6(a)]. However, other time bins yield the opposite results, which can be explained by: (1) Signal components that violate the separability condition lead to incorrect RS. (2) Higher noise levels prevent accurate phase shifts. Due to the higher-amplitude spectral components of the RS, the total power and noise power are higher (than of the spectrogram), which leads to lower SNR values. (3) Time bins without any signals may contain a smeared STFT spectrum that will lead to detections. The results in Figs. 6(c) and 6(d) are probably an outcome of the first two reasons.

After benchmarking the method's performance with synthetic data, recordings from February 10, 2010, are used to test its performance on real data (Fig. 7). The specific day is chosen after verifying that there are coherent microbarom signals present. Microbaroms are low-frequency acoustic perturbations in the atmosphere caused by oscillating standing oceanic waves (Smets and Evers, 2014). Depending on the ocean depth, their frequencies range between 0.2 and 1 Hz (Assink *et al.*, 2014; Averbuch *et al.*, 2019; Smets and Evers, 2014). Here, the recorded signals have dominant microbarom frequencies ranging between 0.2 and 0.3 Hz.

Figure 8 shows the beamforming results, based on the STFT and RS for two time bins. As in Fig. 6, the two methods perform differently at different time bins. Compared to Fig. 8(b), Fig. 8(a) shows a significant improvement in both amplitude and the lobe's width. However, the STFT-based results in Fig. 8(d) are better compared to Fig. 8(c).

IV. DISCUSSION AND CONCLUSIONS

The STFT and the spectrogram are fundamental signal processing tools. Due to the use of a smoothing kernel, the spectrogram is smeared. This leads to an imprecise TFR of the recorded signals, i.e., some time and frequency elements contain spurious energy. Moreover, these elements may contain phase information that does not correspond to its spectrogram coordinates. This study re-introduces the RS. The

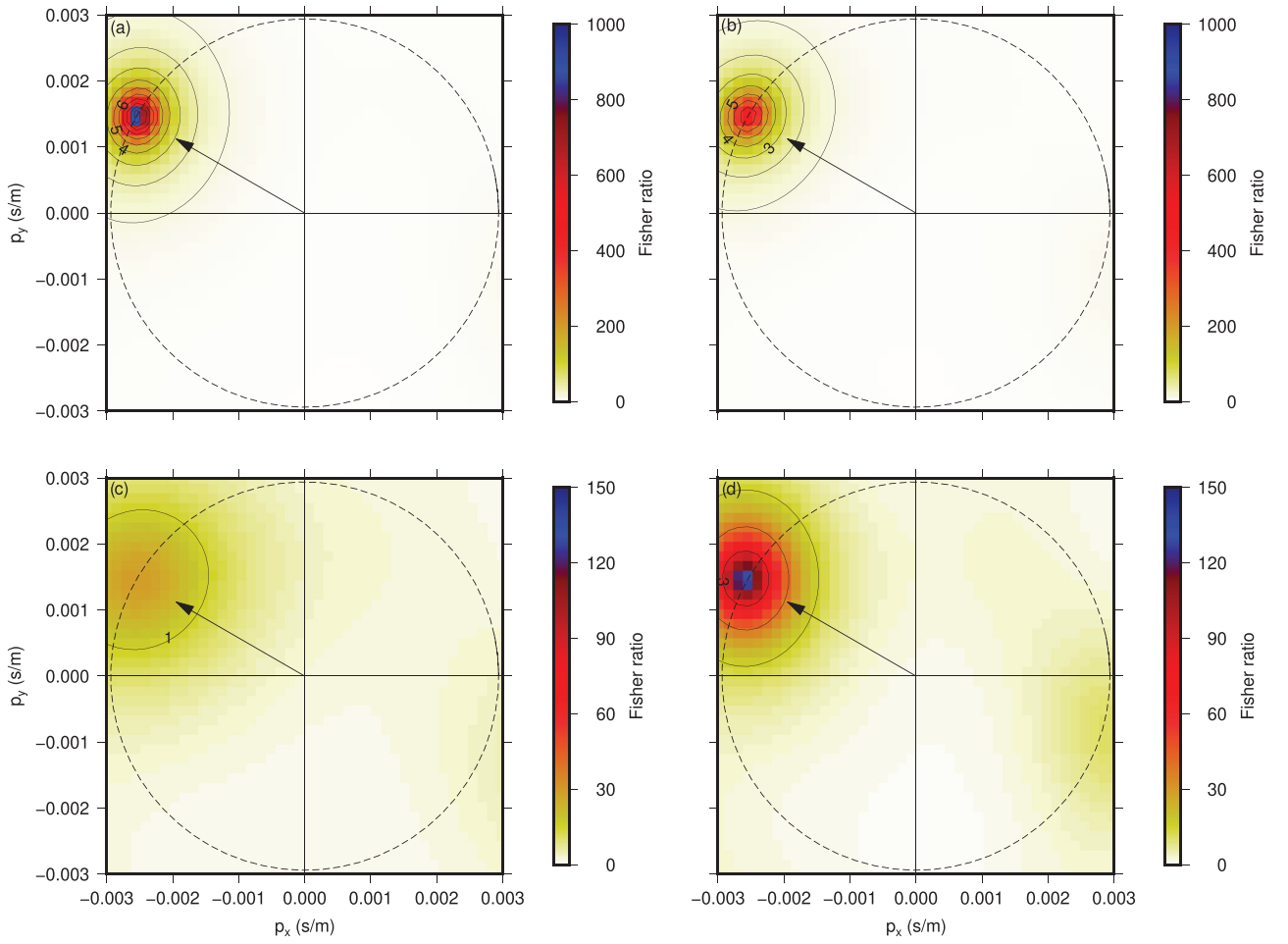

FIG. 6. (Color online) Frequency-wavelength analysis results for a synthetic signal, based on the STFT (b and d) and RS (a and c). Colors correspond to the Fisher ratio, and solid contours represent the SNR. The dashed circle and arrow represent the signal's apparent velocity and BA, respectively. Rows correspond to the same time window.

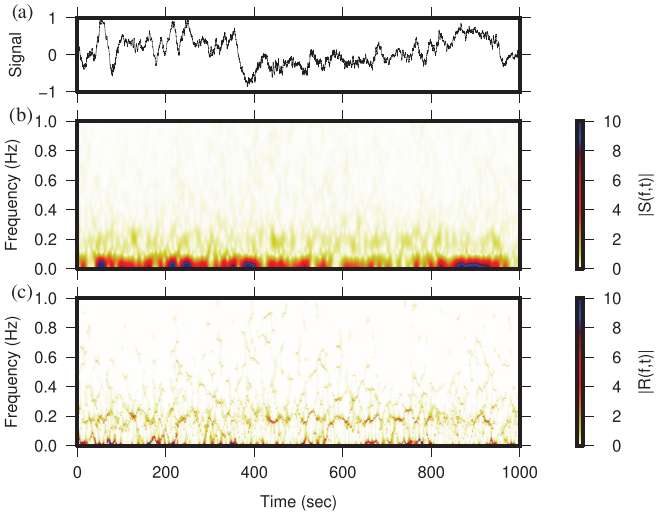

FIG. 7. (Color online) I18DK representative recordings. (a) Signal. (b) Spectrogram. (c) Reassigned spectrogram.

RS fixes the smeared spectrum by reallocating energy to the distribution's center of gravity, which corresponds to the original signal phase information. In addition to time-frequency analysis, STFT is used as the basis for more advanced signal processing techniques such as frequency-domain beamforming and CC. Both methods are based on the phase-shifts of the recordings. Since some TFR elements contain phase information that does not correspond to their coordinates, their phase shifts are inaccurate. It means that by using the RS, phase shifts can be optimized.

Section II B shows three examples that compare the performance of the spectrogram and RS on known signals. Figures 1, 3, and 4 show the stark difference in the obtained TFRs; the spectrograms suggest that all signals are broadband, even though they are actually narrowband. Conversely, the RS-based TFRs provide better representations of the signals. Since the spectrogram is widely used to characterize recorded signals and will probably stay that way, the presented results should be used to highlight the inherent artifacts and limitations of the method.

While the theory and results presented in Sec. II are a reminder of a method that stayed under the radar, to the best of the author's knowledge, this is the first time that the RS is applied to f-k analysis and CC (Sec. III). Figure 5 shows the significant improvement in both the Fisher ratio and CC coefficient for a time-shifted signal. The improvement is based on optimizing the phase-shifts by correcting the

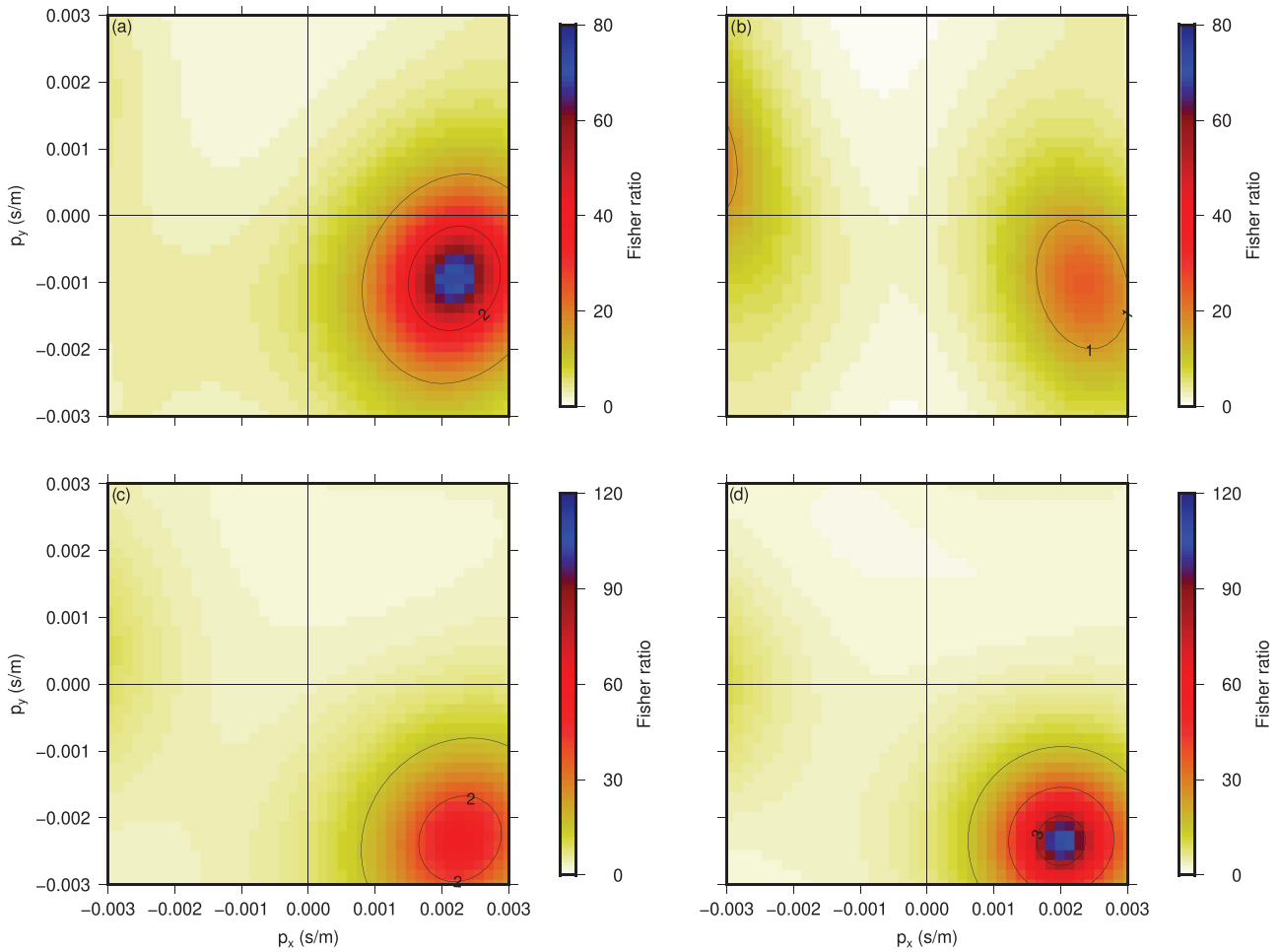

FIG. 8. (Color online) Frequency-wavelength analysis results for I18DK data, based on the STFT (b and d) and RS (a and c). Colors correspond to the Fisher ratio, and solid contours represent the SNR. Rows correspond to the same time window.

https://doi.org/10.1121/10.0003384

smeared spectrum. The Fisher ratio results, based on the RS, show improvement in both amplitude and accuracy. However, the RS-based correlation curve is a scaled version of the STFT-based curve. Therefore, there is no improvement in accuracy although it was expected.

The results from array data beamforming are less conclusive than the synthetic time-shifted signal example. For time bins with signal components that fulfill the separability condition, the RS-based beamforming is superior to the STFT-based analysis. In contrast, the STFT-based beamforming provides better results if the separability condition is violated and/or in the presence of high noise levels and coherent noise. Furthermore, higher amplitudes of the RS lead to an increase in the total power across the array [Eq. (18)]. In that case, if the phase-shifts are not accurate, the signal power and Fisher ratio are significantly low. These facts make the STFT-based beamforming more forgiving in contrast to the RS-base beamforming and, therefore, more robust for infrasound data analysis.

The strict conditions of the RS-based beamforming and its sensitivity to the choice of parameters make it hard to implement operationally. However, the RS can still shed more light on the observations and help choose better processing parameters for the STFT-based beamforming. For example, if one bases her/his analysis on the spectrogram, an initial broad frequency spectrum may be chosen; such a broad spectrum will lead to relatively low SNR values. The low values result from phase shifts that are "too far" from the one needed. By trial and error, excluding some spectral components will increase the SNR since the difference between the performed phase-shifts and those needed is relatively small. Finally, a further decrease in the frequency range will reduce the SNR since the remaining frequency elements contain low amplitudes. Such trial-and-error steps can be reduced by basing the parameters' choice on the RS.

In Sec. III, other beamforming methods were mentioned. Methods like Capon, MUSIC, and Clean, are based on improving the array response and iteratively deconvolving it from the obtained results. While these methods come to improve the beamforming part, the presented study comes to improve the input data. Therefore, the RS can be used by the more advanced beamforming techniques, and it is expected to yield further improvement of the results.

High noise levels, coherent noise, and violation of the separation condition harm the RS performance. For the method to be more applicable to array data, further studies on reducing noise and parting spectral components are needed. It is found that RS-based beamforming performs better on narrowband signals. If a broadband signal is present, performing the array analysis on several narrow frequency bands may be useful. This method may also be applied to measurements with relatively low noise levels, such as medical imaging and aeroacoustic studies.

The given examples in this study are mainly based on continuous narrowband signals. Therefore, one can question the performance of the method in the presence of a broadband transient signal. Auger and Flandrin (1995) show that the RS [Eq. (14)] inherits the WVD property and can localize a delta function in time. However, the theory does not discuss the RS performance on band limited transients, i.e., a filtered delta function. Preliminary numerical tests show the presence of artifacts at the RS of such signals (not shown here). Therefore, in order to apply the RS on such signals (for time-frequency analysis and beamforming) further theoretical and numerical studies on the problem are needed.

To conclude, this study shows the potential benefits of using the RS in time-frequency analysis, frequency-domain cross correlation, and f-k analysis. In addition, the limitations are presented and discussed. Hopefully, this study will encourage further investigation that will lead to improved signal processing capabilities.

ACKNOWLEDGMENTS

I would like to thanks the CTBTO and station operators for the high-quality IMS data and products. IMS data can be accessed through the vDec (see https://www.ctbto.org/specials/vdec/). Also, the author would like to thank François Auger for discussions about the method of reassignment, and Stephen Arrowsmith and the anonymous reviewers for their constructive comments.

APPENDIX: DERIVING THE REASSIGNED COORDINATES

To help the reader, a detailed derivation of the reassigned coordinates [Eqs. (10)–(13)] by Kodera *et al.* (1976) and Auger and Flandrin (1995) is presented. The starting points are the time and frequency first moments of the energy measured by the WVD [(Eq. (4)]. Since the imaginary part does not contribute to the integration, and the reassigned coordinates are real quantities, only the real part of the expression will be considered:

$$\hat{t}(t_0,\omega_0) = t_0 - \frac{\iint tW(h;t,\omega)W(x;t_0-t,\omega-\omega_0)dtd\omega}{\iint W(h;t,\omega)W(x;t_0-t,\omega-\omega_0)dtd\omega} \tag{A1}$$

and

$$\hat{\omega}(t_0,\omega_0) = \omega_0 - \frac{\iint \omega W(h;t,\omega)W(x;t_0-t,\omega-\omega_0)dtd\omega}{\iint W(h;t,\omega)W(x;t_0-t,\omega-\omega_0)dtd\omega}, \tag{A2}$$

which are almost the same as Eqs. (5) and (6). The only difference is that $\omega_0 - \omega$ is now swapped, and the frequency variable elsewhere remains the same (based on the symmetry of the FT).

The WVD can be expressed by the original signal and window functions as

$$\begin{aligned}&\int\int W(h;t,\omega)W(x;t_0-t,\omega-\omega_0)dt\frac{d\omega}{2\pi}\\&=\int\int X^*(\omega-\omega_0)H^*(\omega)e^{-i\omega t_0}\frac{d\omega}{2\pi}e^{i\omega_0 t_0}\\&\quad\times x(t_0-t)h(t)e^{-i\omega_0 t}dt\\&=S^*(t_0,\omega_0)S(t_0,\omega_0).\end{aligned}\tag{A3}$$

Following Eq. (A3), Eq. (A1) can be rewritten as

$$\hat{t}(t_0,\omega_0)=t_0-\frac{e^{i\omega_0 t_0}\int tx(t_0-t)h(t)e^{-i\omega_0 t}dtS^*(t_0,\omega_0)}{S^*(t_0,\omega_0)S(t_0,\omega_0)}.\tag{A4}$$

To obtain the expressions of the reassigned coordinates by Kodera *et al.* (1976), the integral in the numerator can be expressed as

$$\int tx(t_0-t)h(t)e^{-i\omega_0(t-t_0)}dt=-i\frac{\partial S(t_0,\omega_0)}{\partial\omega_0}+t_0S(t_0,\omega_0).\tag{A5}$$

Plugging Eq. (A5) in Eq. (A4) yields

$$\hat{t}(t_0,\omega_0)=\frac{i}{S(t_0,\omega_0)}\frac{\partial S(t_0,\omega_0)}{\partial\omega_0}.\tag{A6}$$

Expressing $S(t_0,\omega_0)$ as its argument and modulus yields

$$\hat{t}(t_0,\omega_0)=-\frac{\partial}{\partial\omega_0}\arg\{S(t_0,\omega_0)\}+\frac{i}{A}\frac{\partial A}{\partial\omega_0}.\tag{A7}$$

Retaining the real part of the expression leaves us with

$$\hat{t}(t_0,\omega_0)=-\frac{\partial}{\partial\omega_0}\arg\{S(t_0,\omega_0)\},\tag{A8}$$

which is equivalent to Eq. (10).

Similarly, Eq. (A2) can be simplified to the form

$$\hat{\omega}(t,\omega)=\omega+\frac{\partial}{\partial t}\arg\{S(t,\omega)\}.\tag{A9}$$

For the expressions proposed by Auger and Flandrin (1995), let us start with Eq. (A4). Now, the integral in the numerator will be expressed as

$$\int tx(t_0-t)h(t)e^{-i\omega_0(t-t_0)}dt=S_{\tau h}(t_0,\omega_0),\tag{A10}$$

where $S_{\tau h}(t_0,\omega_0)$ correspond to the STFT based on the smoothing kernel $t\cdot h(t)$. This representation provides us with Auger's expression for the reassigned time,

$$\hat{t}(t,\omega)=t-Re\left\{\frac{S_{\tau h}(t,\omega)\cdot S^*(t,\omega)}{|S(t,\omega)|^2}\right\}.\tag{A11}$$

Similarly, Eq. (A3) is used to express Eq. (A2) as

$$\hat{t}(t_0,\omega_0)=\omega_0-\frac{e^{i\omega_0 t_0}\int\omega x(t_0-t)h(t)e^{-i\omega_0 t}dtS^*(t_0,\omega_0)}{S^*(t_0,\omega_0)S(t_0,\omega_0)}.\tag{A12}$$

Expressing $\omega h(t)$ as $Im\{dh/dt\}$, the integral in the numerator can be written as

$$\int\omega x(t_0-t)h(t)e^{-i\omega_0 t}dt=Im\{S_{dh}(t_0,\omega_0)\},\tag{A13}$$

where S_{dh} corresponds to the smoothing kernel dh/dt. The reassigned frequency coordinate is expressed as

$$\hat{\omega}(t,\omega)=\omega+Im\left\{\frac{S_{dh}(t,\omega)\cdot S^*(t,\omega)}{|S(t,\omega)|^2}\right\}.\tag{A14}$$

Angione, F., Novak, C., Imeson, C., Lehman, A., Merwin, B., Pagliarella, T., Samardzic, N., D'Angela, P., and Ule, H. (**2016**). "Study of a low frequency emergency siren in comparison to traditional siren technology," Proc. Mtgs. Acoust. **29**, 030008.

Applbaum, D., Averbuch, G., Price, C., Yair, Y., and Ben-Horin, Y. (**2020**). "Infrasound observations of sprites associated with winter thunderstorms in the eastern Mediterranean," Atmos. Res. **235**, 104770.

Arrowsmith, S., and Bowman, D. (**2017**). "Explosion yield estimation from pressure wave template matching," J. Acoust. Soc. Am. **141**(6), EL519–EL525.

Assink, J. D., Waxler, R., Smets, P., and Evers, L. G. (**2014**). "Bidirectional infrasonic ducts associated with sudden stratospheric warming events," J. Geophys. Res. Atmos. **119**(3), 1140–1153, https://doi.org/10.1002/2013JD021062.

Auger, F., and Flandrin, P. (**1995**). "Improving the readability of time-frequency and time-scale representations by the reassignment method," IEEE Trans. Signal Process. **43**(5), 1068–1089.

Averbuch, G., Assink, J. D., Smets, P. S. M., and Evers, L. G. (**2018**). "Extracting low signal-to-noise ratio events with the Hough transform from sparse array data," Geophysics **83**(3), WC43–WC51.

Averbuch, G., Ben-Horin, Y., Smets, P. S. M., and Evers, L. G. (**2019**). "The Mount Meron Infrasound Array: An infrasound array without a noise reduction system," Geophys. J. Int. **219**, 1109–1117.

Bartlett, M. S. (**1948**). "Smoothing periodograms from time-series with continuous spectra," Nature **161**, 686–687.

Bracewell, R. (**2000**). *The Fourier Transform and Its Applications* (McGraw-Hill, New York).

Capon, J. (**1969**). "High-resolution frequency-wavenumber spectrum analysis," Proc. IEEE **57**(8), 1408–1418.

Chen, W. C., Chen, H. Y., Chen, Y. S., Tian, Y., and Lin, I. L. (**2017**). "Self-excited multi-scale skin vibrations probed by optical tracking micro-motions of tracers on arms," J. Appl. Phys. **122**(2), 024701.

Cohen, L. (**1989**). "Time-frequency distributions–A review," Proc. IEEE **77**, 941–981.

Cohen, L. (**1995**). *Time-Frequency Analysis: Theory and Applications* (Prentice-Hall, Upper Saddle River, NJ).

Cohen, L., and Loughlin, P. J. (**2003**). "The marginals and time-frequency distributions," in *Advanced Signal Processing Algorithms, Architectures, and Implementations XIII*, edited by F. T. Luk (SPIE, Bellingham, WA), Vol. 5205, p. 59.

Debnath, L. (**2002**). *Wavelet Transforms and Their Applications* (Birkhäuser Boston, Boston, MA), available at http://link.springer.com/10.1007/978-1-4612-0097-0.

den Ouden, O. F. C., Assink, J. D., Smets, P. S. M., Shani-Kadmiel, S., Averbuch, G., and Evers, L. G. (**2020**). "CLEAN beamforming for the enhanced detection of multiple infrasonic sources," Geophys. J. Int. **221**(1), 305–317.

Evers, L. G., Brown, D., Heaney, K. D., Assink, J. D., Smets, P. S., and Snellen, M. (**2014**). "Evanescent wave coupling in a geophysical system: Airborne acoustic signals from the Mw 8.1 Macquarie Ridge earthquake," Geophys. Res. Lett. **41**(5), 1644–1650, https://doi.org/10.1002/2013GL058801.

Flanagan, J. L. (**1972**). "Speech synthesis," in *Speech Analysis Synthesis and Perception* (Springer, Berlin, Heidelberg), pp. 204–276.

Frazer, G., and Boashash, B. (**1994**). "Multiple window spectrogram and time-frequency distributions," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 4, pp. IV/293–IV/296.

Fulop, S. A., and Fitz, K. (**2006**). "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," J. Acoust. Soc. Am. **119**(1), 360–371.

Gabor, D. (**1946**). "Theory of communication. Part 1: The analysis of information," J. Inst. Electr. Eng. - Part III: Radio Commun. Eng. **93**(26), 429–441.

Gal, M., Reading, A. M., Ellingsen, S. P., Koper, K. D., Burlacu, R., and Gibbons, S. J. (**2016**). "Deconvolution enhanced direction of arrival estimation using one- and three-component seismic arrays applied to ocean induced microseisms," Geophys. J. Int. **206**(1), 345–359.

Garibbo, S., Blondel, P., Heald, G., Heyburn, R., Hunter, A., and Williams, D. (**2020**). "Low-frequency ocean acoustics–Measurement from the Lofoten-Vesterålen Ocean Observatory, Norway," Proc. Mtgs. Acoust. **40**, 040001.

Guedes Soares, C., and Cherneva, Z. (**2005**). "Spectrogram analysis of the time-frequency characteristics of ocean wind waves," Ocean Eng. **32**(14–15), 1643–1663.

Jensen, J. A. (**2006**). "Spectral velocity estimation in ultrasound using sparse data sets," J. Acous. Soc. Am. **120**(1), 211–220.

Kodera, K., De Villedary, C., and Gendrin, R. (**1976**). "A new method for the numerical analysis of non-stationary signals," Phys. Earth Planet. Inter. **12**(2–3), 142–150.

Koenig, W., Dunn, H. K., and Lacy, L. Y. (**1946**). "The sound spectrograph," J. Acoust. Soc. Am. **18**(1), 19–49.

Mellinger, D. K., and Clark, C. W. (**2000**). "Recognizing transient low-frequency whale sounds by spectrogram correlation," J. Acoust. Soc. Am. **107**(6), 3518–3529.

Melton, B. S., and Bailey, L. F. (**1957**). "Multiple signal correlators," Geophysics **22**(3), 565–588.

Merino-Martínez, R., Heblij, S. J., Bergmans, D. H., Snellen, M., and Simons, D. G. (**2019**). "Improving aircraft noise predictions considering fan rotational speed," J. Aircraft **56**(1), 284–294.

Neele, F., and Snieder, R. (**1991**). "Are long-period body wave coda caused by lateral heterogeneity?," Geophys. J. Int. **107**(1), 131–153.

Nelson, D. J. (**2001**). "Cross-spectral methods for processing speech," J. Acoust. Soc. Am. **110**(5), 2575–2592.

Papandreou-Suppappola, A. (**2003**). *Applications in Time-Frequency Signal Processing*, 1st ed. (CRC Press, Boca Raton, FL), p. 432.

Potter, J. R., Mellinger, D. K., and Clark, C. W. (**1994**). "Marine mammal call discrimination using artificial neural networks," J. Acoust. Soc. Am. **96**(3), 1255–1262.

Rossi-Santos, M. R. (**2020**). "Louder than love: Anthropogenic noise overlaps humpback whale songs in coastal soundscapes of Bahia, Brazil," Proc. Mtgs. Acoust. **37**, 040007.

Schmidt, R. O. (**1986**). "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag. **34**(3), 276–280.

Shani-Kadmiel, S., Assink, J. D., Smets, P. S. M., and Evers, L. G. (**2018**). "Seismoacoustic coupled signals from earthquakes in central Italy: Epicentral and secondary sources of infrasound," Geophys. Res. Lett. **45**(1), 427–435, https://doi.org/10.1002/2017GL076125.

Smart, E., and Flinn, E. A. (**1971**). "Fast frequency-wavenumber analysis and Fisher signal detection in real-time infrasonic array data processing," Geophys. J. R. Astronom. Soc. **26**(1–4), 279–284.

Smets, P. S., and Evers, L. G. (**2014**). "The life cycle of a sudden stratospheric warming from infrasonic ambient noise observations," J. Geophys. Res. **119**(21), 12084–12099, https://doi.org/10.1002/2014JD021905.

Smith, R. L., and Angerami, J. J. (**1968**). "Magnetospheric properties deduced from OGO 1 observations of ducted and nonducted whistlers," J. Geophys. Res. **73**(1), 1–20, https://doi.org/10.1029/ja073i001p00001.

Stankovic, L. (**1994**). "Method for time-frequency analysis," IEEE Trans. Signal Process. **42**(1), 225–229.

Young, R. W. (**1973**). "Sound pressure in water from a source in air and vice versa," J. Acoust. Soc. Am. **53**(6), 1708–1716.