Digital Signal Processing 42 (2015) 1–26

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms ☆

Dmytro Iatsenko, Peter V.E. McClintock, Aneta Stefanovska *

Department of Physics, Lancaster University, Lancaster LA1 4YB, UK

ARTICLE INFO

Article history:
Available online 3 April 2015

Keywords:
Time–frequency analysis
Windowed Fourier transform
Wavelet transform
Synchrosqueezing

ABSTRACT

Time–frequency representations (TFRs) of signals, such as the windowed Fourier transform (WFT), wavelet transform (WT) and their synchrosqueezed versions (SWFT, SWT), provide powerful analysis tools. Here we present a thorough review of these TFRs, summarizing all practically relevant aspects of their use, reconsidering some conventions and introducing new concepts and procedures to advance their applicability and value. Furthermore, a detailed numerical and theoretical study of three specific questions is provided, relevant to the application of these methods, namely: the effects of the window/wavelet parameters on the resultant TFR; the relative performance of different approaches for estimating parameters of the components present in the signal from its TFR; and the advantages/drawbacks of synchrosqueezing. In particular, we show that the higher concentration of the synchrosqueezed transforms does not seem to imply better resolution properties, so that the SWFT and SWT do not appear to provide any significant advantages over the original WFT and WT apart from a more visually appealing pictures. The algorithms and Matlab codes used in this work, e.g. those for calculating (S)WFT and (S)WT, are freely available for download.

1. Introduction

Identification and quantification of the oscillatory components present in a given signal is a classic problem of signal processing, and time–frequency analysis has been one of the most successful approaches for solving it. Thus, it is often useful to study the signal's structure in both time and frequency simultaneously, which can be done by considering a specifically constructed projection of the signal onto the time–frequency plane. Such projections are called time–frequency representations (TFRs) and there are a number of different kinds, depending on the exact way in which the projection is carried out.

Time–frequency analysis is especially useful for signals containing many oscillatory components with time-varying amplitudes and/or frequencies, which is a very common scenario for real-life signals. The paradigmatic example is the heart beat: although the cardiac frequency is usually localized around ∼1 Hz, it varies continuously about this average; these variations are hard to analyze in the time domain, or by using the Fourier transform (FT), but they can easily be traced in the time–frequency plane [1–6]. The TFRs have thus become established as very powerful tools that are routinely applied in almost every area of science, from image processing and finance to geophysics and the life sciences [4,7–18].

General aspects of time–frequency analysis, and the properties of the different existing TFRs, have been thoroughly discussed in a number of excellent books [7,8,19–25] and reviews [9,13,26–28]. In a sense, however, there is too much information available, as well as some aspects remain to be revised. Therefore, instead of briefly reviewing all existing TFRs, we concentrate on an in depth study of only four specific types (to be discussed below), considering all their practically relevant aspects and clarifying various issues related to their use.

In general, there exist two main TFR classes: linear and quadratic. Although each of these has its own advantages, we will consider only linear TFRs, such as the windowed Fourier transform (WFT) and wavelet transform (WT). We consider these types because, first, they are more readily interpretable on account of being additive (the TFR of a sum of signals = the sum of the TFRs for each signal) and, secondly, linear TFRs offer the possibil-

☆ Work supported by the Engineering and Physical Sciences Research Council, UK (Grant No. EP/I00999X/1).

* Corresponding author.
E-mail addresses: dmytro.iatsenko@gmail.com (D. Iatsenko), p.v.e.mcclintock@lancaster.ac.uk (P.V.E. McClintock), aneta@lancaster.ac.uk (A. Stefanovska).

http://dx.doi.org/10.1016/j.dsp.2015.03.004

ity of extracting and reconstructing individual components, which can be problematic for quadratic representations. Additionally, we also consider the synchrosqueezed WFT and WT (abbreviated as SWFT and SWT) [5,6,29–31], which represent a particular nonlinear transformation of the original WFT and WT to increase their concentration. Although not additive, the SWFT and SWT still allow for straightforward extraction and reconstruction of the signal's components.

The purpose of the present work is therefore to: (i) bring together and summarize the minimal but sufficient knowledge needed to understand and apply the WFT and WT effectively; (ii) cover the aspects of their use that are not discussed and clarified adequately in the literature; (iii) consider in detail the synchrosqueezed transforms and their properties; (iv) advance the theory and extend the applicability of the (S)WFT and (S)WT by introducing certain improvements, procedures and concepts; (v) provide well-optimized and user-friendly algorithms and MatLab codes (available at [32]), appropriate for any window or wavelet; (vi) investigate some important related issues.

The paper falls into two parts. The first part provides a tutorial review of the above mentioned TFRs and their properties, supplemented with some additional contributions. The particular emphasis is placed on the TFR-based estimation of the properties of the components present in the signal, and on some underlying caveats. An experienced reader who is familiar with all these topics can leave the first part for reference, and proceed directly to the second one, which considers more advanced issues. In particular, it is shown that synchrosqueezing does not seem to provide significant advantages apart from a more visually appealing picture, while having some drawbacks.

The main results of the work are summarized in the ultimate section. All notation, abbreviations, terminology, assumptions and conventions used throughout the paper are listed in Appendix A. Supplementary Material covers some technical details that are omitted from the main text.

Remark 1.1. Based on different TFRs, many advanced methods have been developed, such as wavelet coherence [16,33,34] and phase coherence [35–37], TFR-based harmonic identification [38], wavelet bispectral analysis [39,40], and others. However, it does not seem feasible to review appropriately all such techniques in one work, given their large number, and we do not consider them here. Rather, we concentrate on the in depth study of the basic time–frequency representations, which provide the foundation for all these advanced techniques.

Part 1. Tutorial review

This part provides a tutorial review of the (S)WFT and (S)WT, considering their main properties and various related issues. We start by discussing the basic concepts, such as the notion of the AM/FM component and the analytic signal approach. The WFT, WT, SWFT and SWT are then defined, and different aspects of their use are discussed in detail.

2. Analytic signal

One of the basic notions of time–frequency analysis is the *AM/FM component* (or simply *component*), which is defined as a function of time t of form

$$x(t) = A(t)\cos\phi(t) \quad \big(\forall t:\ A(t) > 0,\ \ \phi'(t) > 0\big). \tag{2.1}$$

The time-dependent values $A(t)$, $\phi(t)$ and $\nu(t) \equiv \phi'(t)$ are then called the instantaneous amplitude, phase and frequency of the component (2.1) (see [41–43] for a more detailed discussion of their definitions and related issues).

Given that the signal is known to be of the form (2.1), the natural question is how to find its associated $A(t)$, $\phi(t)$ and $\nu(t)$. The most convenient way of doing this is the analytic signal approach. However, before considering it, a few additional notions should be introduced. Thus, for an arbitrary function $f(t)$, its Fourier transform (FT), positive and negative frequency parts, time-average and standard deviation will be denoted as $\hat{f}(\xi)$, $f^+(t)$, $f^-(t)$, $\langle f(t)\rangle$ and $\mathrm{std}[f(t)]$, respectively:

$$\hat{f}(\xi) \equiv \int_{-\infty}^{\infty} f(t)e^{-i\xi t}dt$$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\xi)e^{i\xi t}d\xi = \langle f(t)\rangle + f^+(t) + f^-(t),$$

$$f^+(t) \equiv \frac{1}{2\pi}\int_{0^+}^{\infty} \hat{f}(\xi)e^{i\xi t}d\xi, \quad f^-(t) \equiv \frac{1}{2\pi}\int_{-\infty}^{0^-} \hat{f}(\xi)e^{i\xi t}d\xi,$$

$$\langle f(t)\rangle = \frac{\int f(t)dt}{\int dt}, \quad \mathrm{std}[f(t)] \equiv \sqrt{\langle [f(t)]^2\rangle - [\langle f(t)\rangle]^2}. \tag{2.2}$$

Here and in what follows the integrals are taken over $(-\infty, \infty)$ if unspecified (or, in practice, over the full time duration of $f(t)$).

For a given signal $s(t)$ (which is always assumed to be real in this work), its doubled positive-frequency part is called its *analytic signal* and will be denoted as $s^a(t)$:

$$s^a(t) \equiv 2s^+(t) \quad \big(s(t) = \langle s(t)\rangle + \mathrm{Re}[s^a(t)]\big). \tag{2.3}$$

The analytic signal is complex, so its dynamics can easily be separated into amplitude and phase parts. For signals represented by a single component (2.1), the analytic amplitude and phase $A^a(t), \phi^a(t)$ match closely the true amplitude and phase $A(t), \phi(t)$, thus providing an easy way to estimate them:

$$A(t) \approx A^a(t) \equiv |s^a(t)|, \ \ \phi(t) \approx \phi^a(t) \equiv \arg[s^a(t)]. \tag{2.4}$$

The approximate equality (2.4) will be called the analytic approximation. According to the Bedrosian theorem [44] (see also [41,42]), this approximation is exact when the spectrum of $A(t)$ lies lower than the spectrum of $e^{i\phi(t)}$, and there are no intersections between the two; the amplitude and phase of the component are then uniquely defined. Usually, however, there is a small discrepancy between the true and analytic amplitude/phase (considered in detail in Supplementary Section 2), but related errors are typically much smaller than those of the other methods of amplitude and frequency estimation [45].

However, real-life signals rarely consist of only one component, and they usually also contain noise. In what follows, we assume that the signal can be represented by a sum of AM/FM components, each of which satisfies the analytic approximation (2.4), plus some noise $\eta(t)$:

$$s(t) = \sum_i x_i(t) + \eta(t) = \sum_i A_i(t)\cos\phi_i(t) + \eta(t),$$

$$\forall t, i:\ A_i(t) > 0,\ \phi_i'(t) > 0,\ [A_i(t)\cos\phi_i(t)]^+ \approx A_i(t)e^{i\phi_i(t)}/2, \tag{2.5}$$

In this case, the analytic signal will represent a mix of the amplitude and phase dynamics of all components contained in the signal (additionally corrupted by noise), so their individual parameters cannot be recovered from it. One should therefore employ more sophisticated techniques, able to distinguish the different components within a single time-series.

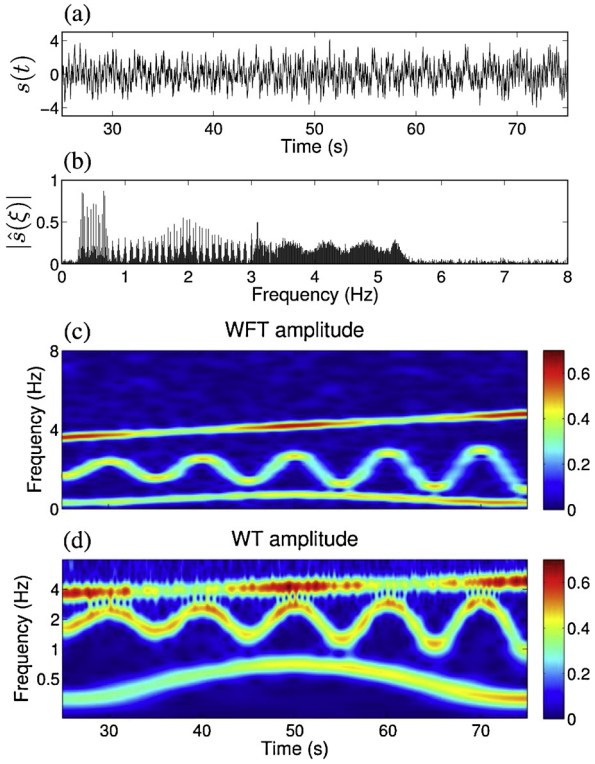

Fig. 1. Different representations of a signal composed of three AM/FM components and corrupted by white Gaussian noise. (a): Signal in the time domain. (b): Signal in the frequency domain, given by its Fourier transform. (c, d): Signal in the time–frequency domain, given by its WFT and WT (see Sections 3.1 and 3.2 below), respectively. Note the logarithmic frequency scaling in (d).

This is illustrated in Fig. 1, which shows an example of a particular signal representation in the time, frequency, and time–frequency domains, with the latter being given by its WFT and WT, to be discussed below. As can be seen, although all representations by definition contain the same amount of information about the signal, in the case of Fig. 1 the most readily interpretable view of this information is provided in the time–frequency domain. Note that, although signal representation (2.5) is not unique, in practice one aims at the sparsest among such representations, i.e. the one characterized by the smallest number of components $x_i(t)$.

3. Time–frequency representations (TFRs)

As illustrated in Fig. 1, instead of studying a signal in either of the (one-dimensional) time ($s(t)$) or frequency ($\hat{s}(\xi)$) domains, it is often more useful to consider its TFR, i.e. to study the signal in a (two-dimensional) time–frequency plane. Such an approach gives the possibility of tracking the evolution of the signal's spectral content in time, which is typically represented by variations of the amplitudes and frequencies of the components from which the signal is composed. Note that in the present section we will sometimes use the notions of time, frequency and time–frequency resolutions of the TFR, which will be further clarified in Section 4 below. We also refer the reader to classical books and reviews (e.g. [7,20,21,25]) for more details on the WFT/WT and their resolution characteristics.

3.1. Windowed Fourier transform (WFT)

The windowed Fourier transform (WFT), also called the short-time Fourier transform or (in a particular form) the Gabor transform [46], is one of the oldest and thus most-investigated linear TFRs. The WFT $G_s(\omega,t)$ of the signal $s(t)$ can be constructed as:

$$G_s(\omega,t) = \int s^+(u)g(u-t)e^{-i\omega(u-t)}du = \frac{1}{2\pi}\int_0^\infty e^{i\xi t}\hat{s}(\xi)\hat{g}(\omega-\xi)d\xi, \tag{3.1}$$

where $g(u)$ is the chosen *window function* and $\hat{g}(\xi)$ is its FT (without loss of generality, we assume $\operatorname{argmax}|\hat{g}(\xi)| = 0$); the use of $s^+(t)$ instead of simple $s(t)$ is needed to remove the interference with negative frequencies (see Supplementary Section 3 for a discussion of this issue).

The WFT is an invertible transform, so that the original signal in both time and frequency domains can be recovered from it as (see Supplementary Section 5 for derivations):

$$s^a(t) = C_g^{-1}\int G_s(\omega,t)d\omega, \quad s(t) = \langle s(t)\rangle + \mathrm{Re}[s^a(t)],$$
$$\hat{s}(\omega>0) = \widetilde{C}_g^{-1}\int G_s(\omega,t)e^{-i\omega t}dt, \quad \hat{s}(-\omega) = \hat{s}^*(\omega),$$
$$C_g \equiv \frac{1}{2}\int \hat{g}(\xi)d\xi = \pi g(0), \quad \widetilde{C}_g \equiv \int g(t)dt = \hat{g}(0). \tag{3.2}$$

In numerical applications, the WFT is calculated via the inverse FFT algorithm applied to the frequency domain form of (3.1). The full procedure including all related issues discussed in this work is summarized in Supplementary Section 2, with the corresponding codes being freely available at [32].

Gaussian window. Unless otherwise specified, all considerations and formulas in this work apply for an arbitrary window $g(t)$, $\hat{g}(\xi)$ (for most common window forms and their properties, see Supplementary Section 7). However, in what follows we pay particular attention to the Gaussian window function

$$g(u) = \frac{1}{\sqrt{2\pi}f_0}e^{-u^2/2f_0^2} \quad \Leftrightarrow \quad \hat{g}(\xi) = e^{-f_0^2\xi^2/2}, \tag{3.3}$$

which we use for simulations. It is commonly used on account of its unique property of maximizing the "classic" joint time–frequency resolution of the transform (see e.g. pp. 43–45 in [7]), while the trade-off between the time and frequency resolutions of this window is controlled by the *resolution parameter* f_0 (by default we use $f_0 = 1$).

3.2. Wavelet transform (WT)

The (continuous) wavelet transform (WT) is the other well-known linear TFR. In contrast to the WFT, it has logarithmic frequency resolution; in other respects the two TFRs are quite similar. The WT $W_s(\omega,t)$ of a signal $s(t)$ for a chosen *wavelet function* $\psi(u)$ can be constructed as

$$W_s(\omega,t) = \int s^+(u)\psi^*\Big(\frac{\omega(u-t)}{\omega_\psi}\Big)\frac{\omega du}{\omega_\psi} = \frac{1}{2\pi}\int_0^\infty e^{i\xi t}\hat{s}(\xi)\hat{\psi}^*\Big(\frac{\omega_\psi\xi}{\omega}\Big)d\xi, \tag{3.4}$$

where

$$\omega_\psi \equiv \operatorname{argmax}|\hat{\psi}(\xi)|, \tag{3.5}$$

is the wavelet peak frequency, and we use the positive-frequency part of the signal $s^+(t)$ (2.2) to avoid interference with negative frequencies (see Supplementary Section 4). Note that the WT is commonly expressed through scales $a(\omega) = \omega_\psi/\omega$, but in (3.4) we have already made transformation to frequencies.

The reconstruction formulas for the case of the WT become (see Supplementary Section 5 for derivations)

$$s^a(t) = C_\psi^{-1} \int_0^\infty W_s(\omega,t)\frac{d\omega}{\omega}, \quad s(t) = \langle s(t)\rangle + \mathrm{Re}[s^a(t)],$$

$$\hat{s}(\omega>0) = \widetilde{C}_\psi^{-1} \int W_s(\omega,t)e^{-i\omega t}dt, \quad \hat{s}(-\omega) = \hat{s}^*(\omega),$$

$$C_\psi \equiv \frac{1}{2}\int_0^\infty \hat{\psi}^*(\xi)\frac{d\xi}{\xi},$$

$$\widetilde{C}_\psi \equiv \left[\int \psi(t)e^{-i\omega_\psi t}dt\right]^* = \hat{\psi}^*(\omega_\psi), \tag{3.6}$$

where, in contrast to the WFT (3.2), the signal is reconstructed from its WT by integrating $W_s(\omega,t)$ over the frequency logarithm $d\omega/\omega = d\log\omega$, which is standard for the WT-based measures. Note that, for a meaningful WT, the wavelet's FT $\hat{\psi}(\xi)$ should vanish at zero frequency:

$$\hat{\psi}(0) = \int \psi(t)dt = 0, \tag{3.7}$$

which is called the *admissibility condition*.

Similarly to the WFT, the WT is computed by applying the inverse FFT algorithm to the frequency domain form of (3.4). All details of the procedure can be found in Supplementary Section 2, and the corresponding codes are available at [32].

Morlet wavelet. Except where otherwise specified, all considerations and formulas in this work apply for an arbitrary wavelet $\psi(t), \hat{\psi}(\xi)$ (for most common window forms and their properties, see Supplementary Section 7). However, the so-called Morlet wavelet [47] is worthy of special consideration. It is constructed in analogy with the Gaussian window and takes the form

$$\psi(u) = \frac{1}{\sqrt{2\pi}}\left(e^{i2\pi f_0 u} - e^{-(2\pi f_0)^2/2}\right)e^{-u^2/2},$$

$$\hat{\psi}(\xi) = e^{-(\xi-2\pi f_0)^2/2}\left(1 - e^{-2\pi f_0 \xi}\right), \tag{3.8}$$

where, in analogy with (3.3), f_0 is the *resolution parameter* determining resolution properties of the wavelet (we use $f_0 = 1$ by default), while the second term in $\psi(u)$ is needed to establish the wavelet admissibility condition (3.7). The fact that the Morlet wavelet is used so commonly for the continuous WT is attributable to the belief that it maximizes the time–frequency resolution, though in what follows (Section 4 and Supplementary Section 7) we will see that this is not really the case.

Lognormal wavelet. Since the WT has logarithmic frequency resolution, it seems more appropriate to construct a wavelet using $\log\xi$ as its argument. Therefore, a more correct WT analogy to the Gaussian window (3.3) would probably be not Morlet, but the lognormal wavelet

$$\hat{\psi}(\xi) = e^{-(2\pi f_0 \log\xi)^2/2}, \ \xi > 0 \tag{3.9}$$

As discussed in Supplementary Section 7, the resolution properties of the wavelet (3.9) are generally better than that of the Morlet, and it has a variety of other useful properties. For example, it is "infinitely admissible", i.e. $\int \xi^{-n}\hat{\psi}(\xi)d\xi/\xi$ is finite for any $n \geq 0$, and one therefore can reconstruct any order time-derivative of the component's amplitude and phase from its WT (see Section 5.2). This makes lognormal wavelet a preferable choice, though we will still employ the Morlet wavelet (3.8) in our simulations just because, apart from being the most widespread choice, it has more in common with the other wavelet forms and thus better demonstrates what one typically gets.

3.3. Synchrosqueezed WFT/WT (SWFT/SWT)

Synchrosqueezing [5,6,29–31] provides a way to construct a more concentrated representation from the WFT or WT. The underlying idea is very simple, namely to join all WFT/WT coefficients corresponding to same phase velocities (the first derivative of the unwrapped WFT/WT phase) into one SWFT/SWT coefficient. In mathematical terms, the definition of the WFT/WT instantaneous frequency $\nu_{G,W}(\omega,t)$ is

$$\nu_G(\omega,t) = \frac{\partial}{\partial t}\arg[G_s(\omega,t)] = \mathrm{Im}\left[G_s^{-1}(\omega,t)\frac{\partial G_s(\omega,t)}{\partial t}\right],$$

$$\nu_W(\omega,t) = \frac{\partial}{\partial t}\arg[W_s(\omega,t)] = \mathrm{Im}\left[W_s^{-1}(\omega,t)\frac{\partial W_s(\omega,t)}{\partial t}\right]. \tag{3.10}$$

The SWFT $V_s(\omega,t)$ [29] and SWT $T_s(\omega,t)$ [6] are then

$$V_s(\omega,t) = C_g^{-1}\int \delta(\omega-\nu_G(\tilde{\omega},t))G_s(\tilde{\omega},t)d\tilde{\omega},$$

$$T_s(\omega,t) = C_\psi^{-1}\int_0^\infty \delta(\omega-\nu_W(\tilde{\omega},t))W_s(\tilde{\omega},t)\frac{d\tilde{\omega}}{\tilde{\omega}}, \tag{3.11}$$

where C_g, C_ψ are defined in (3.2), (3.6). Similarly to the underlying WFT/WT themselves, their synchrosqueezed versions also represent invertible transforms: integrating (3.11) over $d\omega$ and using (3.2), (3.6), one can show that signal can be reconstructed from its SWFT/SWT as

$$s^a(t) = \int_0^\infty V_s(\omega,t)d\omega = \int_0^\infty T_s(\omega,t)d\omega. \tag{3.12}$$

However, in contrast to the WFT/WT, there is no possibility of reconstructing directly the signal's FT from its SWFT/SWT.

Since the SWFT and SWT are generally not analytic (e.g. in theory for a single tone signal they are δ-functions), in practice one computes not the true SWFT/SWT as defined in (3.11), but their integrals $\widetilde{V}_s(\omega_k,t)$ and $\widetilde{T}_s(\omega_k,t)$ over each frequency bin:

$$\widetilde{V}_s(\omega_k,t) \equiv \int_{(\omega_{k-1}+\omega_k)/2}^{(\omega_k+\omega_{k+1})/2} V_s(\omega,t)d\omega,$$

$$\widetilde{T}_s(\omega_k,t) \equiv \int_{\sqrt{\omega_{k-1}\omega_k}}^{\sqrt{\omega_k\omega_{k+1}}} T_s(\omega,t)d\omega, \tag{3.13}$$

where the difference between integral limits is on account of centers ω_k of the frequency bins usually being linearly and logarithmically spaced for the SWFT and SWT, respectively. The computational cost of synchrosqueezing is of the same order as for the usual WFT/WT, and fast algorithm for performing it is discussed in Supplementary Section 2. Note that in terms of a numerical SWFT/SWT (3.13) the reconstruction formula (3.12) becomes simply $s^a(t) = \sum_k \widetilde{V}_s(\omega_k,t) = \sum_k \widetilde{T}_s(\omega_k,t)$.

3.4. Other time–frequency representations

The (S)WFT and (S)WT represent only a few out of the many existing TFR types. The latter include different quadratic representations (e.g. the Wigner–Ville [48,49], Rihaczek [50] and Choi–Williams [51] distributions: see [20,21,26] for a more comprehensive list), as well as various WFT/WT modifications (the windowed fractional Fourier transform [52–54], the local polynomial FT [19, 55], the chirplet transform [56,57] etc.). These are designed in

such a way as to provide better representation for particular signal types, e.g. chirps.

There are also techniques that allow one to improve certain properties of the TFR by applying some sort of processing to it. For example, the "sharpness" of different TFRs can be increased by use of the so-called reassignment methods [30,58], which include synchrosqueezing as a particular case. Furthermore, many methods have been proposed for adaptive signal representation, e.g. the method of frames [59], (orthogonal) matching pursuit [60,61], best orthogonal basis [62] and basis pursuit [63]: for overviews see [63] and [7].

Even this list is far from being exhaustive, and there exist many other TFR types and related techniques (see e.g. [64–71], and an overview in [9]), including those devoted exclusively to some specific goal, such as instantaneous frequency estimation. Each representation, however, has its own advantages and drawbacks, thus deserving a separate review. In the present work we concentrate solely on the (S)WFT and (S)WT.

4. Time-, frequency- and time–frequency resolution

We have made extensive use of the notions of time, frequency and time–frequency resolutions, although having defined them only briefly. In this section different resolution characteristics are considered in detail, and the definitions are made more precise. The TFR resolution properties and their classical treatment are discussed in detail in [7,20,21,25], so below we will focus only on some novel views and related issues. Furthermore, since synchrosqueezing does not seem to change the resolution properties of the transform (see Sections 8 and 11 below), in what follows we will concentrate mainly on the WFT and WT, assuming that the same qualitative and quantitative considerations are applicable respectively to the SWFT and SWT.

4.1. General formulation

Consider a signal $s_{\nu\nu}(t)$ ($s_{\tau\tau}(t)$) consisting of two frequency events – tones (time events – delta-peaks), so that its WFT (3.1) and WT (3.4) are

$$s_{\nu\nu}(t)=\cos(\nu t)+\cos((\nu+\Delta\nu)t+\Delta\varphi) \quad\Rightarrow\quad \begin{cases} G_s(\omega,t)=\frac{1}{2}\big[\hat g(\omega-\nu) \\ \qquad\qquad +\hat g(\omega-\nu-\Delta\nu)e^{i\Delta\nu t}e^{i\Delta\varphi}\big]e^{i\nu t}, \\ W_s(\omega,t)=\frac{1}{2}\Big[\hat\psi^*\Big(\frac{\omega_\psi\nu}{\omega}\Big) \\ \qquad\qquad +\hat\psi^*\Big(\frac{\omega_\psi(\nu+\Delta\nu)}{\omega}\Big)e^{i\Delta\nu t}e^{i\Delta\varphi}\Big]e^{i\nu t}, \end{cases} \tag{4.1}$$

$$s_{\tau\tau}(t)=\delta(t-\tau)+\delta(t-\tau-\Delta\tau)e^{i\Delta\varphi} \quad\Rightarrow\quad \begin{cases} G_s(\omega,t)\cong\big[g(\tau-t) \\ \qquad\qquad +g(\tau+\Delta\tau-t)e^{-i\omega\Delta\tau}e^{i\Delta\varphi}\big]e^{i\omega(t-\tau)}, \\ W_s(\omega,t)\cong\frac{\omega}{\omega_\psi}\Big[\psi^*\Big(\frac{\omega(\tau-t)}{\omega_\psi}\Big) \\ \qquad\qquad +\psi^*\Big(\frac{\omega(\tau+\Delta\tau-t)}{\omega_\psi}\Big)e^{i\Delta\varphi}\Big]. \end{cases} \tag{4.2}$$

Everywhere in this section, e.g. in (4.2), the symbol "$\cong$" denotes equality up to an error associated with the difference between the WFT/WT (3.1), (3.4) and their forms as calculated using the full signal $s(t)$ instead of its positive-frequency part $s^+(t)$; note that, if redefining WFT/WT to use the full signal, the approximate equality will migrate from $s_{\tau\tau}(t)$ (4.2) to $s_{\nu\nu}(t)$ (4.1). See Supplementary Section 4 for a detailed discussion of these issues and the quality of the approximation in (4.2).

It seems reasonable to define the time (frequency) resolution γ_t (γ_ω) of the transform as the reciprocal of the minimum time (frequency) difference $\Delta\tau$ ($\Delta\nu$) in $s_{\tau\tau}(t)$ ($s_{\nu\nu}(t)$) for which two delta-peaks (tones) can still be reliably resolved in the TFR:

$$\gamma_t(\nu,\tau)=1/\Delta\tau_{\min}(\nu,\tau),\quad \gamma_\omega(\nu,\tau)=1/\Delta\nu_{\min}(\nu,\tau). \tag{4.3}$$

However, the meaning of "reliably resolved" still remains imprecise, and will be dealt with later.

The joint time–frequency resolution $\gamma_{\omega t}$ is most often defined as the reciprocal of the area of the minimal resolvable square $[\nu,\nu+\Delta\nu_{\min}(\nu,\tau)]\times[\tau,\tau+\Delta\tau_{\min}(\nu,t)]$, being equal to the product of the time and frequency resolutions. However, within such a square $\gamma_t(\nu,\tau)$ and $\gamma_\omega(\nu,\tau)$ can vary considerably, so that these variations should be taken into account to make $\gamma_{\omega t}$ more meaningful. The latter can therefore be defined as

$$\gamma_{\omega t}(\nu,\tau)=\frac{\Delta\tau_{\min}(\nu,\tau)\Delta\nu_{\min}(\nu,\tau)}{\int_\nu^{\nu+\Delta\nu_{\min}(\nu,\tau)}\Delta\tau_{\min}(\omega,\tau)d\omega\int_\tau^{\tau+\Delta\tau_{\min}(\nu,\tau)}\Delta\nu_{\min}(\nu,t)dt}. \tag{4.4}$$

Note that, if neither $\Delta\nu_{\min}$ nor $\Delta\tau_{\min}$ depends on time or frequency, then one has the traditional $\gamma_{\omega t}=\gamma_\omega\gamma_t$.

4.2. Classical definitions and their flaws

Although $\Delta\tau_{\min}$ ($\Delta\nu_{\min}$) in (4.3) was defined as the minimum time (frequency) difference which can reliably be resolved in the TFR, the meaning of "reliably resolved" remains mathematically unclear. Based on how it is defined, one can characterize the resolution properties of the TFRs in different ways.

The traditional approach [7,25] implicitly assumes two delta-peaks (tones) to be well-resolved if the time (frequency) distance between them exceeds some number of standard deviations of the squared window/wavelet function in the time (frequency) domain. Within this framework, for the WFT one has

$$\begin{aligned} &\Delta\nu_{\min}^{(cl)}=k_1\Delta_\omega,\quad \Delta_\omega^2=E_g^{-1}\frac{1}{2\pi}\int(\omega-\omega_c)^2|\hat g(\omega)|^2d\omega, \\ &\Delta\tau_{\min}^{(cl)}=k_2\Delta_t,\quad \Delta_t^2=E_g^{-1}\int(t-t_c)^2|g(t)|^2dt, \\ &\omega_c\equiv E_g^{-1}\frac{1}{2\pi}\int\omega|\hat g(\omega)|^2d\omega,\quad t_c\equiv E_g^{-1}\int t|g(t)|^2dt, \end{aligned} \tag{4.5}$$

where $E_g\equiv\int|g(t)|^2dt$ ($=\frac{1}{2\pi}\int|\hat g(\xi)|^2d\xi$ by Parseval's identity), and $k_{1,2}$ are implicitly assumed to be the same for all window functions. Obviously, such a definition is far from universal, since for different windows different number of standard deviations are needed to resolve the two tones/delta-peaks.

For the WT, the classic variance-based framework is

$$\begin{aligned} &\Delta\nu_{\min}^{(cl)}(\nu)=k_1\frac{\nu}{\omega_\psi}\Delta_\omega, \\ &\Delta_\omega^2=E_\psi^{-1}\frac{1}{2\pi}\int(\omega-\omega_c)^2|\hat\psi(\omega)|^2d\omega, \\ &\Delta\tau_{\min}^{(cl)}(\nu)=k_2\frac{\omega_\psi}{\nu}\Delta_t,\quad \Delta_t^2=E_\psi^{-1}\int(t-t_c)^2|\psi(t)|^2dt, \\ &\omega_c\equiv E_\psi^{-1}\frac{1}{2\pi}\int\omega|\hat\psi(\omega)|^2d\omega,\quad t_c\equiv E_\psi^{-1}\int t|\psi(t)|^2dt, \end{aligned} \tag{4.6}$$

where $E_\psi\equiv\int|\psi(t)|^2dt$, and $k_{1,2}$ are again implicitly assumed to be the same for all wavelet functions. This approach has the same drawbacks as (4.5) for the WFT. However, in the case of the WT it is actually not appropriate at all, at least in terms of the frequency resolution. Thus, the tones are represented in the WT as terms $\sim\hat\psi(\omega_\psi\nu/\omega)$, so that the decay of their contribution as $\omega\to\infty$, determined by the behavior of $\hat\psi(\xi)$ as $\xi\to0$, will obviously have a big effect on the frequency resolution. At the same time, the

usual variance Δ_ω^2 (4.6) takes no account of this fact, e.g. being invariant under $\hat{\psi}(\xi) \to \hat{\psi}(\xi + \omega_\psi)$, that makes the wavelet inadmissible (in which case even infinitely distant in frequency tones interfere, so that the frequency resolution becomes effectively zero). Hence, for wavelets it is more appropriate to study at least the variance of $|\hat{\psi}(\omega_\psi/\xi)|^2$, but by no means that of $|\hat{\psi}(\xi)|^2$.

For both WFT and WT, the classic time–frequency resolution measure is taken as $\gamma_{\omega t}^{(cl)} = [\Delta_\omega \Delta_t]^{-1}$, with Δ_ω, Δ_t being given by (4.5) for the WFT and by (4.6) for the WT. It can be shown [7,8,25], that this measure attains its maximum for the Gaussian window (3.3) and (up to the effect of the admissibility term $\sim e^{-(2\pi f_0)^2/2}$) for the Morlet wavelet (3.8). However, as follows from the discussion above, only in the case of the WFT does the classic $\gamma_{\omega t}^{(cl)}$ make some sense, though even in this case it remains highly non-universal.

4.3. Notion of the window/wavelet ϵ-support

Before proceeding to a reconsideration of the classic definitions, it is useful to introduce the notions of the window/wavelet ϵ-supports in frequency $[\xi_1(\epsilon), \xi_2(\epsilon)]$ and time $[\tau_1(\epsilon), \tau_2(\epsilon)]$, which will be used frequently in what follows. These ϵ-supports are defined as the widest intervals containing the $(1-\epsilon)$ part of the total integrals of the window/wavelet function which appear in $C_{g,\psi}$ and $\widetilde{C}_{g,\psi}$ (3.2), (3.6). As will be seen below, they are directly related to the accuracy with which the components can be recovered from the WFT/WT.

Considering first the WFT, for an arbitrary window function, including functions that are not always positive and can be oscillating or complex, the corresponding definitions are

$$
\begin{aligned}
&R_g(\omega) \equiv \frac{\int_{-\infty}^{\omega} \hat{g}(\xi)d\xi}{\int \hat{g}(\xi)d\xi} = C_g^{-1}\frac{1}{2}\int_{-\infty}^{\omega} \hat{g}(\xi)d\xi,\\
&\xi_{1,2}(\epsilon):\ |R_g(\xi \le \xi_1)| \le \epsilon/2,\ |1 - R_g(\xi \ge \xi_2)| \le \epsilon/2,\\
&P_g(\tau) \equiv \frac{\int_{-\infty}^{\tau} g(t)dt}{\int g(t)dt} = \widetilde{C}_g^{-1}\int_{-\infty}^{\tau} g(t)dt,\\
&\tau_{1,2}(\epsilon):\ |P_g(\tau \le \tau_1)| \le \epsilon/2,\ |1 - P_g(\tau \ge \tau_2)| \le \epsilon/2.
\end{aligned}
\tag{4.7}
$$

Evidently, $|R_g(\omega)|$ and $|1 - R_g(\omega)|$ quantify the relative parts of $\hat{g}(\xi)$ that are contained in the ranges $\xi < \omega$ and $\xi > \omega$, respectively, while the values $\xi_{1,2}(\epsilon)$ specify the limits within which the $(1-\epsilon)$ part of the window FT resides. In the same manner, $|P_g(\tau)|$ and $|1 - P_g(\tau)|$ reflect the relative parts of $g(t)$ contained in the ranges $t < \tau$ and $t > \tau$, respectively, while $[\tau_1(\epsilon), \tau_2(\epsilon)]$ represents the region encompassing its $(1-\epsilon)$ part. The inequalities in the definitions of $\xi_{1,2}(\epsilon)$ ($\tau_{1,2}(\epsilon)$) are needed only for complex or real but not strictly positive $\hat{g}(\xi)$ ($g(t)$) to ensure that the its integral over any frequency (time) region containing $[\xi_1(\epsilon), \xi_2(\epsilon)]$ ($[\tau_1(\epsilon), \tau_2(\epsilon)]$) will always approximate the corresponding full integral with relative error not higher than ϵ.

Similarly to the case of the WFT, the ϵ-supports for the WT are defined based on (3.6) as

$$
\begin{aligned}
&R_\psi(\omega) \equiv \frac{\int_0^{\omega} \hat{\psi}^*(\xi)\frac{d\xi}{\xi}}{\int_0^{\infty} \hat{\psi}^*(\xi)\frac{d\xi}{\xi}} = C_\psi^{-1}\frac{1}{2}\int_0^{\omega} \hat{\psi}^*(\xi)\frac{d\xi}{\xi},\\
&\xi_{1,2}(\epsilon):\ |R_\psi(\xi \le \xi_1)| \le \epsilon/2,\ |1 - R_\psi(\xi \ge \xi_2)| \le \epsilon/2,\\
&P_\psi(\tau) \equiv \frac{\int_{-\infty}^{\tau} \psi^*(t)e^{i\omega_\psi t}dt}{\int \psi^*(t)e^{i\omega_\psi t}dt} = \widetilde{C}_\psi^{-1}\int_{-\infty}^{\tau} \psi^*(t)e^{i\omega_\psi t}dt,\\
&\tau_{1,2}(\epsilon):\ |P_\psi(\tau \le \tau_1)| \le \epsilon/2,\ |1 - P_\psi(\tau \ge \tau_2)| \le \epsilon/2.
\end{aligned}
\tag{4.8}
$$

Like $|P_g(\tau)|$ in (4.8), $|P_\psi(\tau)|$ ($|1 - P_\psi(\tau)|$) quantifies the relative part of $\psi(t)e^{-i\omega_\psi t}$ contained at $t < \tau$ ($t > \tau$), with $[\tau_1(\epsilon), \tau_2(\epsilon)]$ specifying the interval encompassing its $(1-\epsilon)$ part. In the same manner, $|R_\psi(\omega)|$ and $\xi_{1,2}(\epsilon)$ are related to the relative part of $\hat{\psi}(\xi)$, taken on a logarithmic scale.

The quantities (4.7), (4.8) are very convenient and will be used extensively below, not only in the present section. For simplicity, $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ denote the respective ϵ-supports both for the window function in the WFT and for the wavelet function in the WT. The meaning will always be clear from the context. Note that the full supports of the window/wavelet in time ($g(t), \psi(t)$) and frequency ($\hat{g}(\xi), \hat{\psi}(\xi)$), whether finite or not, are $[\tau_1(0), \tau_2(0)]$ and $[\xi_1(0), \xi_2(0)]$, respectively.

4.4. Reconsidered definitions

A more universal and appropriate approach (than the traditional variance-based one) is to regard two components as being reliably resolved if they can each be accurately identified and reconstructed from the signal's TFR, i.e. can be recovered with a relative error not exceeding some threshold. Consider the WFT of the two-tone signal (4.1), from which one wants to find the individual analytic signals $x^a_{\nu\nu;1}(t) = e^{i\nu t}$, $x^a_{\nu\nu;2}(t) = e^{i(\nu+\Delta\nu)t}e^{i\Delta\varphi}$ for each of the two tones. At any time t, this can be done by first dividing the frequency range at some $\omega = \omega_x(t)$ into two parts, each responsible for a separate tone, and then integrating the WFT as in (3.2), but over the corresponding frequency ranges. This will give the reconstructed analytic signals $\tilde{x}^a_{\nu\nu;1}(t) = C_g^{-1}\int_{-\infty}^{\omega_x(t)} G_s(\omega,t)d\omega$ and $\tilde{x}^a_{\nu\nu;2}(t) = C_g^{-1}\int_{\omega_x(t)}^{\infty} G_s(\omega,t)d\omega$ which, using (4.1) and (4.7), can be represented as

$$
\begin{aligned}
\tilde{x}^a_{\nu\nu;1}(t) &= e^{i\nu t}\big[\big(1 - R_g(\nu - \omega_x(t))\big) + R_g(\omega_x(t)\\
&\quad - \nu - \Delta\nu)e^{i(\Delta\nu t + \Delta\varphi)}\big],\\
\tilde{x}^a_{\nu\nu;2}(t) &= e^{i\nu t}\big[R_g(\nu - \omega_x(t)) + \big(1 - R_g(\omega_x(t)\\
&\quad - \nu - \Delta\nu)\big)e^{i(\Delta\nu t + \Delta\varphi)}\big].
\end{aligned}
\tag{4.9}
$$

Obviously, the reconstruction errors $x^a_{\nu\nu;1,2}(t) - \tilde{x}^a_{\nu\nu;1,2}(t)$ generally depend on the phase-shift $\Delta\varphi$. Therefore, in the corresponding expressions one should take the average over $\Delta\varphi$, which will be denoted as $\langle\ldots\rangle_{\Delta\varphi}$. The relative errors of each tone's reconstruction $\varepsilon_{\nu\nu;1,2}(\nu, t, \Delta\nu)$ then become

$$
\begin{aligned}
\varepsilon^2_{\nu\nu;1,2}(\nu, t, \Delta\nu) &\equiv \frac{\langle|x^a_{\nu\nu;1,2}(t) - \tilde{x}^a_{\nu\nu;1,2}(t)|^2\rangle_{\Delta\varphi}}{\langle|x^a_{\nu\nu;1,2}(t)|^2\rangle_{\Delta\varphi}}\\
&= |R_g(\nu - \omega_x(t))|^2 + |R_g(\omega_x(t) - \nu - \Delta\nu)|^2.
\end{aligned}
\tag{4.10}
$$

Note that, in the present case, averaging over $\Delta\varphi$ and time-averaging will give the same results, but in general the TFR resolution properties can depend on time, and taking the mean over phase-shifts allows one to localize these errors at each t.

The minimum resolvable frequency difference $\Delta\nu_{\min}(\nu, t)$ is then the minimum $\Delta\nu$ in (4.10) for which the total error $\varepsilon_{\nu\nu}(\nu, t, \Delta\nu)$ remains smaller than some accuracy threshold ϵ_r:

$$
\begin{aligned}
&\varepsilon_{\nu\nu}(\nu, t, \Delta\nu \ge \Delta\nu_{\min}(\nu, t))\\
&\quad \equiv \Big[\varepsilon^2_{\nu\nu;1}(\nu, t, \Delta\nu) + \varepsilon^2_{\nu\nu;2}(\nu, t, \Delta\nu)\Big]^{1/2} \le \epsilon_r.
\end{aligned}
\tag{4.11}
$$

It can be expressed through the ϵ_r-support of the window in frequency (4.7). Thus, consider the WFT with real, positive and symmetric $\hat{g}(\omega)$, e.g. a Gaussian (3.3). Then it follows from (4.1) that the minimum WFT amplitude between the peaks corresponding to two tones will always appear at $\omega = \nu + \Delta\nu/2$ (unless these two

peaks are merged into a single one at some times, which might happen if $\Delta\nu$ is too small). Therefore, in practice the respective frequency regions of the tones will be separated exactly at their average frequency (see Section 5 below), so that one should use $\omega_x(t) = \nu + \Delta\nu/2$ when estimating the errors (4.10). The overall reconstruction error is then

$$\begin{aligned}\varepsilon_{\nu\nu}(\Delta\nu) &= \left[2|R_g(-\Delta\nu/2)|^2 + 2|R_g(-\Delta\nu/2)|^2\right]^{1/2} \\ &= 2|R_g(-\Delta\nu/2)|,\end{aligned} \tag{4.12}$$

and, taking into account that $\xi_1(\epsilon) = -\xi_2(\epsilon)$ due to the assumed window symmetry, it follows from (4.11), (4.7) that the frequency difference for which two tones are recovered with inaccuracy ϵ is exactly equal to the ϵ-support of $\hat{g}(\xi)$. For other window forms (e.g. asymmetric $\hat{g}(\xi)$) all becomes more complicated, but one can still expect to get an overall error of around ϵ when $\Delta\nu = \xi_2(\epsilon) - \xi_1(\epsilon)$. Note that the above considerations hold for reasonably small ϵ, so that $\Delta\nu$ is high enough and there are always two distinct peaks in the WFT amplitude.

The case of two delta-peaks (4.2) is closely similar to that of two tones, so the same considerations apply, with just $\xi_{1,2}(\epsilon) \to \tau_{1,2}(\epsilon)$. Hence, setting ϵ_r as the maximum allowable reconstruction error for which two tones/delta-peaks can still be regarded as resolved, the resolution characteristics for the WFT become

$$\begin{aligned}&\Delta\nu_{\min} = \xi_2(\epsilon_r) - \xi_1(\epsilon_r), \quad \Delta\tau_{\min} \cong \tau_2(\epsilon_r) - \tau_1(\epsilon_r), \\ &\gamma_\omega \equiv \Delta\nu_{\min}^{-1}, \quad \gamma_t = \Delta\tau_{\min}^{-1}, \\ &\gamma_{\omega t} \equiv \gamma_\omega\gamma_t = [\Delta\nu_{\min}\Delta\tau_{\min}]^{-1},\end{aligned} \tag{4.13}$$

where $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ are defined in (4.7). Accurate reconstruction can reasonably be assumed as being at 95% precision, so one can set $\epsilon_r = 0.05$ in (4.13). The resolution characteristics of different windows are listed in Supplementary Section 7.

The same approach straightforwardly extends to the WT case, where one applies similar considerations in terms of (4.8). Thus, it can be shown that two tones with frequency ratio $\frac{\nu+\Delta\nu}{\nu} = 1 + \Delta\nu/\nu = \frac{\xi_2(\epsilon)}{\xi_1(\epsilon)}$ are reconstructed from the WT with an overall relative error of around ϵ. This estimate is exact if $\hat{\psi}(\xi)$ is real, positive and symmetric on a logarithmic scale (such as the lognormal wavelet (3.9)), in which case the tones will always be separated at $\omega_x(t) = \exp[\log\nu + \log(\nu + \Delta\nu)] = \sqrt{\nu(\nu + \Delta\nu)}$. For the resolution of two delta-peaks, it follows from (4.2) that one should consider the ϵ-supports corresponding to $\psi^*(\omega t/\omega_\psi)$, so that the related error will be different at each frequency ω, characterized by ϵ calculated from $\omega\Delta\tau/\omega_\psi = \tau_2(\epsilon) - \tau_1(\epsilon)$. Hence, the resolution parameters (4.3), (4.4) for the WT are

$$\begin{aligned}&\Delta\nu_{\min}(\nu) = \nu\Big(\frac{\xi_2(\epsilon_r)}{\xi_1(\epsilon_r)} - 1\Big), \\ &\Delta\tau_{\min}(\nu) \cong \frac{\omega_\psi}{\nu}\big(\tau_2(\epsilon_r) - \tau_1(\epsilon_r)\big), \\ &\gamma_\omega(\nu) \equiv [\Delta\nu_{\min}(\nu)]^{-1}, \quad \gamma_t(\nu) = [\Delta\tau_{\min}(\nu)]^{-1}, \\ &\gamma_{\omega t} \cong \left[\omega_\psi\big(\tau_2(\epsilon_r) - \tau_1(\epsilon_r)\big)\log\frac{\xi_2(\epsilon_r)}{\xi_1(\epsilon_r)}\right]^{-1},\end{aligned} \tag{4.14}$$

where $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ are defined in (4.8), and ϵ_r is the maximum allowable reconstruction error, which can be set to $\epsilon_r = 0.05$, similarly to that in (4.13). The resolution characteristics of different wavelets are listed in Supplementary Section 7.

Summarizing, in contrast to the classic resolution measures (4.5), (4.6), the quantities in (4.13), (4.14) are very universal and have clear physical meaning, being directly related to the accuracy with which two time or frequency events can be recovered from the resultant TFR.

5. Extraction of components from the TFR

One of the main purposes of time–frequency analysis is the identification and quantification of the AM/FM components present in a signal. Being reliably represented in the TFR, they can be identified and reconstructed. In this respect the TFR can be used to decompose a signal into its constituent components, or just to recover some particular components of interest. We now discuss the two steps of component extraction from the TFR, namely its identification and reconstruction.

5.1. Identification of the components

AM/FM components visually appear as "curves" in the TFR plots (see e.g. Fig. 1). Thus, if the construction of the (S)WFT or (S)WT is well-matched to the signal's structure, then each component will be represented by a unique sequence of TFR amplitude peaks – the ridge curve $\omega_p(t)$, being mapped to a particular time–frequency region around it – the time–frequency support (TFS) $[\omega_-(t), \omega_+(t)]$. Hence, to identify the component one first needs to find its associated ridge curve and TFS.

This is not a trivial task, since in real cases there exist many peaks in the TFR amplitude at each time and, moreover, their number can vary in time (e.g. due to noise). The problem of finding the appropriate sequence of peaks is widely discussed in the context of ridge analysis [1–3,72], but it is a separate topic, which we consider in detail in [73]. For the remainder of this section, we will therefore assume that the ridge curve $\omega_p(t)$ for the component of interest has already been found. The TFS can then be defined as the widest region of unimodal and non-zero (WFT/WT) or just non-zero (SWFT/SWT) TFR amplitude around $\omega_p(t)$, as discussed below.

Definition 1. The time–frequency support $[\omega_-(t), \omega_+(t)]$ of the component in the current TFR $H_s(\omega, t)$ (WFT, WT, SWFT or SWT) is defined as the widest frequency region around the corresponding ridge curve $\omega_p(t)$ where:

1. The TFR amplitude is higher than zero:

$$|H_s(\omega, t)| > 0, \quad \forall t, \omega \in [\omega_-(t), \omega_+(t)].$$

2. [WFT/WT only] The TFR amplitude is unimodal:

$$\operatorname{sign}(\omega - \omega_p(t))\partial_\omega|H_s(\omega, t)| \le 0, \quad \forall t, \omega \in [\omega_-(t), \omega_+(t)].$$

Conversely, the ridge curve is then the sequence of highest TFR amplitude peaks in the corresponding TFS at each time:

$$\omega_p(t) = \operatorname*{argmax}_{\omega\in[\omega_-(t),\omega_+(t)]} |H_s(\omega, t)|.$$

Examples of ridge curves and TFSs are shown in Fig. 2.

In relation to the WFT/WT, the first criterion in the TFS definition establishes that the support is cut when the WFT/WT amplitude drops to zero, usually implying that the region to which some AM/FM component is mapped has ended. The second criterion establishes that if two supports are overlapping in the WFT/WT (i.e. the corresponding components interfere with each other), then they will be nearly-optimally separated at the frequency where the amplitude minimum between the two peaks occurs. Note that such a TFS definition is fully appropriate only for unimodal $|\hat{g}(\xi)|$ and $|\hat{\psi}(\xi > 0)|$, while for windows and wavelets that are multimodal in frequency, the WFT/WT-based separation of the components becomes much more problematic (and therefore such windows/wavelets are rarely used).

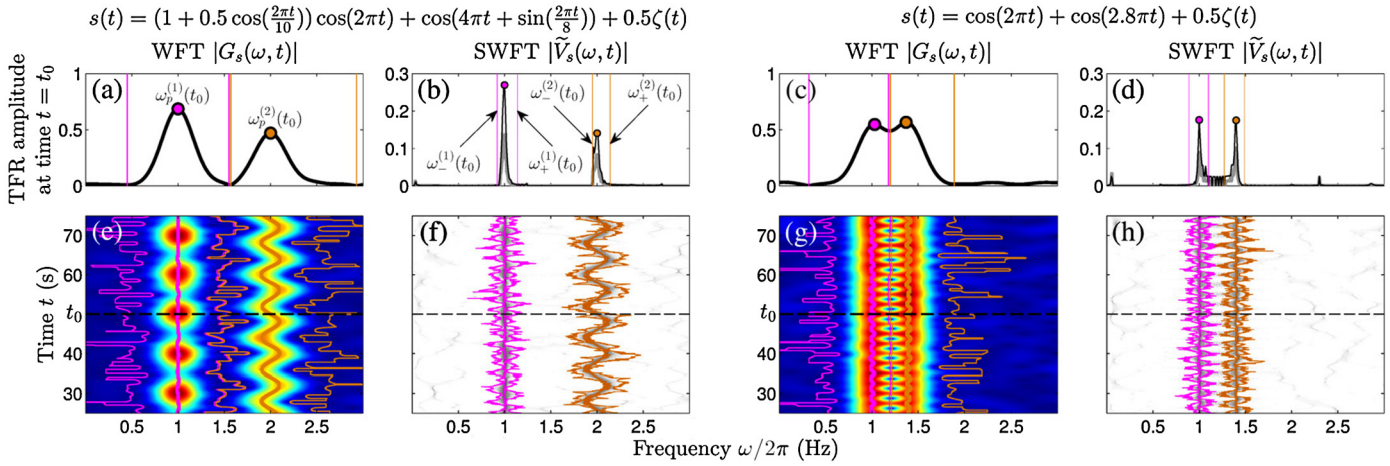

Fig. 2. Examples of the extracted components' ridge curves $\omega_p(t)$ and time–frequency supports $[\omega_-(t), \omega_+(t)]$ in the WFT and SWFT (the case of the WT and SWT is similar) for two different signals, defined by the equations at the top (with $\zeta(t)$ denoting unit-deviation Gaussian white noise). Each signal consists of two components, and in each panel the colors magenta and orange refer to the first and second components, respectively. (a–d): Snapshots of the WFT and SWFT amplitudes (thick black lines) at time $t_0 = 50$ s, with the corresponding ridge points $\omega_p(t_0)$ and the TFSs borders $\omega_\pm^{(1,2)}(t_0)$ being shown by filled colored circles and thin colored lines, respectively. (e–h): Full time-evolutions of the WFT and SWFT amplitudes (t_0, corresponding to the snapshots in (a–d), are indicated by the dashed black lines), with the components' TFSs $[\omega_-^{(1,2)}(t), \omega_+^{(1,2)}(t)]$ being indicated by thin colored lines, and their ridge curves $\omega_p^{(1,2)}(t)$ by thick solid lines of the same colors. The gray background lines in (e, f) show the amplitudes of the SWFTs calculated with a frequency discretization step $\Delta\omega/2\pi = 0.005$, which is half that of the original. The signals were sampled at 100 Hz for 100 s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For the SWFT/SWT, on the other hand, the first criterion is enough, so that it is appropriate to extract the TFS simply as the widest region of non-zero amplitude around the given peaks. Thus, one cannot require strict unimodality as synchrosqueezed transforms are non-smooth and often demonstrate saw-tooth patterns (see Fig. 2(b, d)). Neglecting the second criterion might seem to cause interfering components to be assigned to the same support (as would be the case for the WFT/WT), but numerical tests indicate that this does not happen if the corresponding peaks are distinguishable (i.e. not merged together completely) in the underlying WFT/WT.

In general, the behavior of the synchrosqueezed TFRs is much more complex than that of the WFT/WT and, even in the noise-free case, the SWFT/SWT will contain a lot of spurious TFSs represented by small-amplitude spikes (see Fig. 2(b, d)). Thus, noise, interference or amplitude/frequency modulation might cause some of the SWFT/SWT power of the component to be dissipated over other frequencies in the form of spiky TFSs. Such "leakage" is usually very small and does not lead to large errors in signal reconstruction (see Section 8), but the existence of so many TFSs greatly complicates the interpretation of the resultant SWFT/SWT, representing a significant drawback in comparison with the usual WFT/WT. Moreover, in contrast to the latter, the amplitude of the synchrosqueezed TFRs depends on the width of the frequency bins in a quite sophisticated way (conditioned on the signal's structure and characteristics of the components), as illustrated in Fig. 2(b, d). This reflects the important fact that, for the SWFT/SWT, only the integral over the whole support makes sense, so that e.g. the component's amplitude cannot be estimated based on the heights of the corresponding peaks, as can be done for the WFT/WT (see below).

Remark 5.1. For the SWFT/SWT, it has also been proposed [5,30] to reconstruct the components from the time–frequency area of constant width Δ around the corresponding ridge curve, i.e. to define the TFS as $\omega_\pm(t) = \omega_p(t) \pm \Delta/2$. Such an approach, however, is highly non-universal, as well as non-adaptive. Thus, the spread of the SWFT/SWT around each peak is often asymmetric and varies in time (see e.g. Fig. 2(f,h)), being determined by the signal's structure and TFR resolution properties.

5.2. Estimation of the component's parameters

Having found the ridge curve $\omega_p(t)$ and TFS $[\omega_-(t), \omega_+(t)]$ corresponding to the chosen AM/FM component $A(t)\cos\phi(t)$, one then usually needs to estimate its amplitude $A(t)$, phase $\phi(t)$ and frequency $\nu(t)$. This can be done by two methods: direct and ridge. Both these approaches are considered below, while their relative performance and the corresponding errors will be studied in detail in Sections 8 and 10.

5.2.1. Direct reconstruction

Given its TFS, the amplitude and phase of the component can be reconstructed using the inversion formulas (3.2), (3.6), (3.12), but with the integration over ω being restricted to only the corresponding time–frequency region $[\omega_-(t), \omega_+(t)]$. The expressions for estimating instantaneous frequency, and generally any order time-derivatives of amplitude and phase, can be derived in a similar way (see Supplementary Section 5). Such a method will be referred to as *direct reconstruction*, and the corresponding formulas are:

direct[**WFT**]:

$$A(t)e^{i\phi(t)} = C_g^{-1}\int_{\omega_-(t)}^{\omega_+(t)} G_s(\omega,t)d\omega,$$

$$\nu(t) = \mathrm{Re}\left[\frac{\int_{\omega_-(t)}^{\omega_+(t)}\omega G_s(\omega,t)d\omega}{\int_{\omega_-(t)}^{\omega_+(t)} G_s(\omega,t)d\omega} - \overline{\omega}_g\right],\quad \overline{\omega}_g \equiv \frac{C_g^{-1}}{2}\int \xi\,\hat{g}(\xi)d\xi,$$

$(\overline{\omega}_g = 0$ for symmetric $\hat{g}(\xi)$, such as Gaussian window),

direct[**WT**]:

$$A(t)e^{i\phi(t)} = C_\psi^{-1}\int_{\omega_-(t)}^{\omega_+(t)} W_s(\omega,t)\frac{d\omega}{\omega},$$

$$\nu(t) = \mathrm{Re}\left[\frac{D_\psi^{-1}\int_{\omega_-(t)}^{\omega_+(t)}\omega W_s(\omega,t)\frac{d\omega}{\omega}}{C_\psi^{-1}\int_{\omega_-(t)}^{\omega_+(t)} W_s(\omega,t)\frac{d\omega}{\omega}}\right],\quad D_\psi \equiv \frac{\omega_\psi}{2}\int_0^\infty \frac{1}{\xi}\hat{\psi}^*(\xi)\frac{d\xi}{\xi},$$

direct[**SWFT**]:

$$A(t)e^{i\phi(t)} = \int_{\omega_-(t)}^{\omega_+(t)} V_s(\omega,t)d\omega, \quad \nu(t) = \mathrm{Re}\left[\frac{\int_{\omega_-(t)}^{\omega_+(t)} \omega V_s(\omega,t)d\omega}{\int_{\omega_-(t)}^{\omega_+(t)} V_s(\omega,t)d\omega}\right],$$

direct[**SWT**]:

$$A(t)e^{i\phi(t)} = \int_{\omega_-(t)}^{\omega_+(t)} T_s(\omega,t)d\omega, \quad \nu(t) = \mathrm{Re}\left[\frac{\int_{\omega_-(t)}^{\omega_+(t)} \omega T_s(\omega,t)d\omega}{\int_{\omega_-(t)}^{\omega_+(t)} T_s(\omega,t)d\omega}\right], \tag{5.1}$$

where the formula for reconstructing $\nu(t)$ from the SWFT/SWT was introduced simply by analogy to the WFT/WT case, thus being derived rather empirically. In practice, the frequency axis is partitioned onto bins, so the formulas (5.1) should also be discretized, and the correct way of doing this is discussed in Supplementary Section 1.2.

Except for the SWFT/SWT-based $\nu(t)$, all direct estimates (5.1) by definition give exact values (up to the error of the analytic approximation (2.4)) in the "ideal" case when TFS contains all the energy of the component and no other contributions. The SWFT/SWT-based instantaneous frequency estimates, though not rigorously derived, are also almost exact in this case, as will be seen in Section 8.

Remark 5.2. The direct estimation of instantaneous frequency (5.1) is not possible for windows and wavelets with infinite $\overline{\omega}_g$ and D_ψ (5.1), respectively. While for the WFT this situation arises only for a very exotic windows with $|\int \xi \hat{g}(\xi)d\xi| = \infty$, for the WT it is more common, occurring e.g. for the Morlet wavelet. In such circumstances one can use a kind of empirical approximation which we refer to as hybrid reconstruction:

$$\textbf{hybrid}[\textbf{WFT}]\text{: } \nu(t) = \mathrm{Re}\left[\frac{\int_{\omega_-(t)}^{\omega_+(t)} \nu_G(\omega,t)G_s(\omega,t)d\omega}{\int_{\omega_-(t)}^{\omega_+(t)} G_s(\omega,t)d\omega}\right],$$

$$\textbf{hybrid}[\textbf{WT}]\text{: } \nu(t) = \mathrm{Re}\left[\frac{\int_{\omega_-(t)}^{\omega_+(t)} \nu_W(\omega,t)W_s(\omega,t)\frac{d\omega}{\omega}}{\int_{\omega_-(t)}^{\omega_+(t)} W_s(\omega,t)\frac{d\omega}{\omega}}\right]. \tag{5.2}$$

When both direct and hybrid reconstructions are possible, direct is to be preferred (as it was rigorously derived, see Supplementary Section 5), though the difference between the two estimates is usually negligible.

5.2.2. Ridge reconstruction

The other widely used possibility is to reconstruct the component's parameters using TFR values at the ridge points $\omega_p(t)$ [1,7,72], which will be referred to as the *ridge reconstruction*. Thus, by considering the WFT/WT of a single tone signal $s(t) = A\cos(\nu t + \varphi)$ it can be seen that the tone amplitude and phase can be perfectly reconstructed from the TFR value at any frequency as $Ae^{i(\nu t+\varphi)} = 2G_s(\omega,t)/\hat{g}(\omega-\nu)$ (WFT) or $Ae^{i(\nu t+\varphi)} = 2W_s(\omega,t)/\hat{\psi}^*(\omega_\psi \nu/\omega)$ (WT). Generalizing such an approach to the case of any AM/FM component, one obtains the ridge reconstruction formulas:

ridge[**WFT**]:

$$\nu(t) = \omega_p(t) + \delta\nu_d(t), \quad A(t)e^{i\phi(t)} = \frac{2G_s(\omega_p(t),t)}{\hat{g}(\omega_p(t)-\nu(t))},$$

ridge[**WT**]:

$$\nu(t) = \omega_p(t)e^{\delta\log\nu_d(t)}, \quad A(t)e^{i\phi(t)} = \frac{2W_s(\omega_p(t),t)}{\hat{\psi}^*(\omega_\psi \nu(t)/\omega_p(t))},$$

ridge[**SWFT**]:

$$\nu(t) = \omega_p(t), \quad \phi(t) = \arg\left[V_s(\omega_p(t),t)\right],$$

ridge[**SWT**]:

$$\nu(t) = \omega_p(t), \quad \phi(t) = \arg\left[T_s(\omega_p(t),t)\right], \tag{5.3}$$

where $\delta\nu_d(t)$ and $\delta\log\nu_d(t)$ are the corrections for discretization of the frequency scale (see below). There are no expressions for the SWFT/SWT-based amplitude estimates because there is no possibility of reconstructing amplitude from the ridges of synchrosqueezed TFRs. This is because single points in the latter do not reflect the component's amplitude and will generally depend on the widths of frequency bins, as illustrated in Fig. 2 (see also [73,74]).

The corrections $\delta\nu_d(t)$ in (5.3) arise because, in practice, the TFRs are calculated at the discrete frequencies (see Section 6.2), so the peak positions $\omega_p(t)$ are determined only up to the half-width of a frequency bin, leading to inaccuracies in all estimate. For the WFT/WT, these discretization errors can greatly be reduced by using quadratic interpolation to better locate the peak, so that in (5.3) one sets

$$\textbf{WFT}\text{: } \delta\nu_d(t) = \frac{\Delta\omega}{2}\frac{a_3-a_1}{2a_2-a_1-a_3},$$

$$\textbf{WT}\text{: } \delta\log\nu_d(t) = \frac{n_\nu^{-1}\log 2}{2}\frac{a_3-a_1}{2a_2-a_1-a_3},$$

$$a_{\{1,2,3\}} \equiv |G_s(\omega_{\{k_p(t)-1,k_p(t),k_p(t)+1\}},t)| \text{ or } |W_s(\omega_{\{k_p(t)-1,k_p(t),k_p(t)+1\}},t)|, \tag{5.4}$$

where $k_p(t)$ denotes the discrete index of the peak at each time: $\omega_p(t) \equiv \omega_{k_p(t)}$. The corrections (5.4) then propagate to the amplitude and phase estimates (5.3), making them more accurate as well. The interpolation cannot, however, be performed for the SWFT/SWT due to their non-smoothness, and therefore in this case one uses the "uncorrected" estimates in (5.3). The discretization errors of ridge reconstruction for each TFR are discussed in detail in Supplementary Section 1.2.

In contrast to direct reconstruction (5.1), ridge estimates (5.3) are not exact, having errors proportional to the strengths of the amplitude and frequency modulations of the component [1,72]. On the other hand, they are more robust to noise and interference than direct estimates (5.1). These and other related issues will be discussed in Section 10.

6. Practical issues

In theory, one has infinite time and frequency scales, and both of these variables are continuous. In practice, however, everything is finite and discrete, which has specific consequences in terms of the resultant TFRs. In this section, the issues that arise while dealing with real signals are briefly reviewed, while the technical details and a more detailed discussion are provided in Supplementary Section 1.

6.1. Signal preprocessing

To obtain a reliable TFR, an initial preprocessing of the signal should be performed. First, one should remove trends, which can otherwise spoil the resultant representation due to having non-negligible power at a range of frequencies; we detrend the signal by subtracting a third order polynomial fit from it, though other possibilities can also be considered. Secondly, the signal should be filtered in the frequency band of interest to eliminate the interference with the other components lying outside, and we use simple bandpass filtering to do so. The motivation behind these two steps

and their effects are thoroughly discussed in Supplementary Section 1.1.

6.2. Frequency discretization

In theory one has a continuous frequency variable ω, but in practice the TFRs are calculated at the chosen discrete values $\omega_k = (k-k_0)\Delta\omega$ (WFT and SWFT: linear scale) or $\omega_k/2\pi = 2^{(k-k_0)/n_v}$ (WT and SWT: logarithmic scale). Therefore, the widths of the frequency bins, determined by the discretization parameters $\Delta\omega$ or n_v, set an upper bound on the TFR frequency resolution (as it becomes impossible to distinguish components with frequencies in the same bin), as well as introducing related errors into the direct and ridge estimates.

It can be argued, that the optimal discretization parameters, for which bin widths are small enough so as not to degrade the theoretical resolution properties of the TFR, but not smaller, are

$$\textbf{(S)WFT}\text{: } \Delta\omega = \frac{\xi_2(0.5)-\xi_1(0.5)}{N_b},$$

$$\textbf{(S)WT}\text{: } n_v = \frac{N_b \log 2}{\log\xi_2(0.5) - \log\xi_1(0.5)}, \tag{6.1}$$

where N_b is a number larger than one (we use $N_b = 10$ by default). Furthermore, with the choice (6.1) the discretization errors in both direct and ridge estimates also become negligible. See Supplementary Section 1.2 for more details.

6.3. Finite signal length effects

Theoretically, one integrates over an infinite time while computing the WFT (3.1) and WT (3.4), but in practice the signal has a finite time duration. Consequently, the resultant TFR becomes ill-defined near the signal's time borders (when t is close either to 0 or to the overall time-length T), and one needs to find an appropriate way to estimate it there.

This problem is usually tackled by first calculating the TFR for a longer signal obtained by padding the original one with specifically chosen values at both ends (thus continuing it to $t<0$ and $t>T$ according to some rule), and then "trimming" the resultant TFR to the original time limits. The commonest padding rules include zero padding (just set $s(t<0) = s(t>T) = 0$), symmetric/periodic padding (reflect/periodically-continue the signal at both ends) and predictive padding (infer past and forecast future signal values, with Supplementary Section 6 presenting one possible way of doing so).

However, irrespectively of the padding scheme being used, there often appear some distortions of the WFT/WT near both signal ends – *boundary effects* – which obviously propagate to the SWFT/SWT as well. For the case of zero padding, the relative boundary error $\epsilon_b(\omega,t)$ in the TFR coefficient at frequency ω and time t can be estimated as (see Supplementary Section 1.3):

$$\begin{aligned}&\textbf{(S)WFT}\text{: } \epsilon_b(\omega,t) = \big|P_g(-t) + P_g(t-T)\big|,\\ &\textbf{(S)WT}\text{: } \epsilon_b(\omega,t) = \big|P_\psi(-\omega t/\omega_\psi) + P_\psi(\omega(t-T)/\omega_\psi)\big|.\end{aligned} \tag{6.2}$$

For the other schemes boundary errors are harder to estimate, but the above expressions can serve as a good proxy for their upper bounds. Predictive padding typically leads to the smallest distortions, and we therefore use it by default.

Nevertheless, for better accuracy it is advisable to estimate all quantities of interest from the time–frequency region where boundary distortions are negligible (i.e. smaller than some chosen accuracy threshold ϵ) – the so-called *cone-of-influence*:

$$\textbf{Cone-of-influence } \{\omega,t\}_{coi}^{(\epsilon)}\text{: } (\omega,t)\in\{\omega,t\}_{coi}^{(\epsilon)} \Leftrightarrow \epsilon_b(\omega,t)\le\epsilon. \tag{6.3}$$

A more detailed discussion of the padding schemes and boundary effects, as well as the expressions for cones-of-influence in the case of the (S)WFT/(S)WT, can be found in Supplementary Section 1.3 (see also [7,8]).

6.4. TFR frequency range

For completeness, the restrictions on the frequency range $[\omega_{\min},\omega_{\max}]$ over which to calculate the TFR, i.e. how high/low in frequency one can in principle go, should also be discussed. For a signal $s(t)$ sampled at f_s Hz for T seconds, one has

$$\begin{aligned}&\omega_{\min}/2\pi \ge 1/T, \quad \omega_{\max}/2\pi \le f_s/2, \quad \omega_{\min}^{(\text{stat})}/2\pi = 5/T,\\ &\textbf{WFT}\text{: } \omega_{\min}^{(\epsilon)} = -\infty; \quad \textbf{WT}\text{: } \omega_{\min}^{(\epsilon)} \le \omega_\psi\big[\tau_2(\epsilon)-\tau_1(\epsilon)\big]/T;\\ &\textbf{SWFT}\text{: } \omega_{\min}^{(\epsilon)} = -\infty; \quad \textbf{SWT}\text{: } \omega_{\min}^{(\epsilon)} \le \xi_2(\epsilon)\big[\tau_2(\epsilon)-\tau_1(\epsilon)\big]/T,\end{aligned} \tag{6.4}$$

where $\omega_{\min}^{(\text{stat})}$ is the minimal frequency for which one can do any statistically meaningful conclusions based on the behavior of the corresponding TFR coefficients, while $\omega_{\min}^{(\epsilon)}$ is the minimal frequency for which some TFR coefficients are determined with an accuracy better than ϵ in terms of the time boundary effects.

Finally, it should also be noted that to calculate SWFT/SWT in the region $[\omega_{\min},\omega_{\max}]$ with accuracy ϵ in terms of the frequency boundary effects, the underlying WFT/WT should be calculated over a slightly wider range $[\widetilde{\omega}_{\min},\widetilde{\omega}_{\max}]$ given by

$$\begin{aligned}&\textbf{WFT}\text{: } [\widetilde{\omega}_{\min},\widetilde{\omega}_{\max}] = [\omega_{\min}+\xi_1(\epsilon),\omega_{\max}+\xi_2(\epsilon)],\\ &\textbf{WT}\text{: } [\widetilde{\omega}_{\min},\widetilde{\omega}_{\max}] = \Big[\omega_{\min}\frac{\xi_1(\epsilon)}{\xi_2(\epsilon)},\omega_{\max}\frac{\xi_2(\epsilon)}{\xi_1(\epsilon)}\Big].\end{aligned} \tag{6.5}$$

The derivations of (6.4) and (6.5) are provided in Supplementary Section 1.4 together with a more detailed discussion.

Part 2. Advanced topics

Having reviewed the main aspects of the (S)WFT and (S)WT, we now consider three more sophisticated but important issues, namely: the effects of window/wavelet parameters on the resultant TFR, the relative performance of different reconstruction methods, and the advantages/drawbacks of synchrosqueezing. We start by discussing the questions to be addressed. Their numerical study is then performed in four different cases, when the signal is represented as: two interfering tones; an amplitude-modulated component; a frequency-modulated component; and a single tone corrupted by noise. In this way, accounting for all possible complications, one can build up quite a complete picture of how the issues in question manifest themselves for an arbitrary signal. Using the acquired knowledge, we summarize the results, provide additional discussion and draw conclusions for each issue being considered.

For this second part of the work, we introduce few additional assumptions and conventions:

- From now on, $\hat{g}(\xi)$ and $\hat{\psi}(\xi>0)$ are assumed to be real, positive and unimodal, thereby automatically implying the realness and positivity of $R_{g,\psi}(\omega)$ (4.7), (4.8).
- The boundary effects for all TFRs (see Section 6.3) will be minimized by padding the signal with exact values (i.e. simulating it for a longer period and considering only the central part, with the rest used as padding).

- Where needed, the ridge curve $\omega_p(t)$ will be selected so as to provide the most realistic component extraction. This will usually correspond to selecting the highest TFR amplitude peaks at each time: $\omega_p(t) = \mathrm{argmax}_\omega |H_s(\omega,t)|$. Such an approach is very fast, but works only for simulated examples, while for real signals a more universal schemes should be used [73].
- In all simulations of Section 8 we will use very fine frequency binning: $\Delta\omega/2\pi = 0.002$ for (S)WFT and $n_v = 256$ for (S)WT. These values are chosen to guarantee that the discretization-related errors of SWFT/SWT-based frequency estimates (see Section 6.2) are $\Delta\nu_d/2\pi \le 0.001$ (SWFT) and $\Delta\nu_d/\nu \le 0.0015$ (SWT). We will therefore not discuss reconstruction errors that are < 0.001.

Furthermore, since the only significant difference between the WFT and WT lies in their linear and logarithmic frequency resolutions, one can unify many related expressions. Denoting the WFT/WT as $H_s(\omega,t)$, we therefore write

$$H_s(\omega,t) = \int s^+(t) h_{u-t}(\omega)dt = \frac{1}{2\pi}\int_0^\infty e^{i\xi t}\hat{s}(\xi)\hat{h}_\xi(\omega)d\xi,$$

$$\textbf{WFT}:\ h_t(\omega) = g(t)e^{-i\omega t},\quad \hat{h}_\xi(\omega) = \hat{g}(\omega-\xi),$$

$$\textbf{WT}:\ h_t(\omega) = \big(\omega/\omega_\psi\big)\psi^*(\omega t/\omega_\psi),\quad \hat{h}_\xi(\omega) = \hat{\psi}^*(\omega_\psi\xi/\omega). \tag{6.6}$$

The WFT/WT of the multitone signal is then

$$s(t) = \sum_n A_n\cos(\nu_n t+\varphi_n)$$

$$\Rightarrow\ H_s(\omega,t) = \sum_n \frac{A_n e^{i\varphi_n}}{2}\hat{h}_{\nu_n}(\omega)e^{i\nu_n t}, \tag{6.7}$$

and, denoting $\phi_n \equiv \nu_n t+\varphi_n$ and $\Delta\phi_{nm}\equiv\phi_n(t)-\phi_m(t)$, one gets

$$|H_s(\omega,t)|^2 = \sum_n \frac{A_n^2\hat{h}^2_{\nu_n}(\omega)}{4} + \sum_{n,m>n}\frac{A_nA_m\hat{h}_{\nu_n}(\omega)\hat{h}_{\nu_m}(\omega)}{2}\cos\Delta\phi_{nm}(t),$$

$$\arg[H_s(\omega,t)] = \arctan\frac{\sum_n A_n\hat{h}_{\nu_n}(\omega)\sin\phi_n(t)}{\sum_n A_n\hat{h}_{\nu_n}(\omega)\cos\phi_n(t)},$$

$$\nu_H(\omega,t)\equiv\partial_t\arg[H_s(\omega,t)] = \sum_n\frac{A_n^2\hat{h}^2_{\nu_n}(\omega)}{4|H_s(\omega,t)|^2}\nu_n + \sum_{n,m>n}\frac{A_nA_m\hat{h}_{\nu_n}(\omega)\hat{h}_{\nu_m}(\omega)}{4|H_s(\omega,t)|^2}[\nu_n+\nu_m]\cos\Delta\phi_{nm}(t). \tag{6.8}$$

Note that the expressions for the synchrosqueezed TFRs are more complicated and cannot be obtained in a simple form.

Finally, the formulas (3.2) and (3.6) for reconstructing the signal in the time-domain are generalized as

$$s^a(t) = C_h^{-1}\int H_s(\omega,t)d\mu(\omega),\quad C_h = \frac{1}{2}\int\hat{h}_\xi(\omega)d\mu(\omega)$$

$$\textbf{WFT}:\ C_h = C_g,\ \ \mu(\omega)=\omega\in(-\infty,\infty),$$

$$\textbf{WT}:\ C_h = C_\psi,\ \ \mu(\omega)=[\log\omega]\in(-\infty,\infty)\quad\big(\omega\in(0,\infty)\big). \tag{6.9}$$

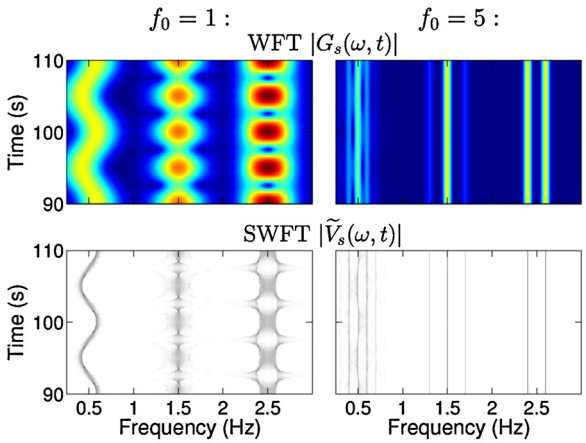

Fig. 3. The WFT and SWFT calculated at different f_0 for a signal consisting of four components $s(t) = s_1(t)+s_2(t)+s_3(t)+s_4(t)$: 1) FM component $s_1(t) = \cos(2\pi\cdot 0.5(t+0.25\cos(2\pi\cdot 0.1t)))$; 2) AM component $s_2(t) = (1+0.5\cos(2\pi\cdot 0.2t))\cos(2\pi\cdot 1.5t)$; 3,4) two tones with close frequencies $s_{3,4}(t) = \cos(2\pi\cdot(2.5\mp 0.1)t)$. The situation for the WT and SWT is qualitatively the same. The signal was sampled at 50 Hz for 200 s.

7. The questions to be addressed

7.1. Choice of the window/wavelet parameters

The parameters of the window/wavelet, e.g. the resolution parameter f_0, determine the time and frequency resolution of the TFR (see Section 4.3), thus having great impact on the resultant representation. In particular, f_0 "instructs" the TFR on how to treat AM/FM components present in the signal: either as single entities, or as sums of tones [75]. This is illustrated in Fig. 3, where different TFRs are shown for signals consisting of one FM component, one AM component, and two tones. As can be seen, for $f_0 = 1$, the TFRs treat both AM/FM components as individual entities, while at $f_0 = 5$ they are represented as sums of tones; at the same time, two tones present in the signal that appear indistinguishable at $f_0 = 1$ can be resolved at $f_0 = 5$. For the case considered in the figure, neither $f_0 = 1$ nor $f_0 = 5$ is suitable, and one needs to choose the resolution parameter based on a compromise between reliable representation of the AM/FM components and reliable representation of the tones.

Generally, the effects of changing f_0 are not always fully understood and appreciated. We will investigate them in detail by studying the TFR behavior for different signals, with the aim of answering the questions: How high should f_0 be to distinguish between two tones? How low should it be to reliably represent components with particular amplitude and/or frequency modulation? Within what range can it be chosen?

Since the AM/FM component can be represented as a sum of tones (see Section 2), consideration can be restricted to the TFR for a multitone signal $s(t) = \sum_{m=1}^M a_m\cos(\nu_m t+\varphi_m)$. For WFT/WT, one way to classify its behavior is to consider the time-averaged number $\langle N_p\rangle$ of TFR amplitude peaks

$$\langle N_p\rangle = \langle\#\omega_p(t):\ |\partial_\omega H_s(\omega_p(t),t)| = 0,\ |\partial^2_\omega H_s(\omega_p(t),t)| < 0\rangle, \tag{7.1}$$

where only peaks larger than 10^{-6} of the summed amplitude of all peaks are taken into account (to discard spurious peaks, e.g. due to round-off errors). In addition, the degree of interference between the tones will be quantified by η (with $\eta\to 0$ and $\eta\to 1$ corresponding to no and complete interference, respectively); it will be defined for each particular case separately.

The qualitative behavior of the WFT/WT for a multitone signal can then be partitioned into four different regimes, as illustrated

Table 1
Regimes of possible WFT/WT behavior for an M-tone signal $s(t)=\sum_{m=1}^{M} a_m \cos(\nu_m t+\varphi_m)$; the ϵ_c ($=0.001$ by default) denotes a precision threshold used for regarding the tones as completely resolved or completely merged together.

Regime	Description	Condition
I	All tones are fully resolved, i.e. at each time there are M well-separated peaks in the WFT/WT amplitude, and the time-variations of the latter are zero or negligible.	$\langle N_p\rangle = M,\ \eta \le \epsilon_c$
II	Tones partly interfere with each other, so the WFT/WT amplitude varies in time, but there are always M distinct peaks.	$\langle N_p\rangle = M,\ \eta > \epsilon_c$
III	Tones severely interfere, so that the nearest ones sometimes merge, i.e. there exist moments when there are fewer than M (non-negligible) peaks in the WFT/WT amplitude.	$\langle N_p\rangle < M,\ \eta < 1-\epsilon_c$
IV	Tones are completely merged so that, although the WFT/WT amplitude varies strongly in time, it usually has only one dominant peak.	$\langle N_p\rangle \approx 1,\ \eta \ge 1-\epsilon_c$

in Table 1. Evidently, for a few unrelated tones, one should aim at Regime I, while for AM/FM components – at Regime IV. The SWFT/SWT, on the other hand, cannot be classified using the same measures (e.g. $\langle N_p\rangle$) due to being non-smooth but, as will be seen below, synchrosqueezing does not seem to change the qualitative aspects of the TFR behavior.

7.2. Choice of the reconstruction method

As discussed in Section 5, there are two possible methods – direct and ridge – by which a component can be reconstructed from its support in a TFR. However, to the best of our knowledge, there are no works comparing their performance, so it is not clear in what cases which method should be used. This issue will be thoroughly investigated in the following sections. We will quantify the relative error of reconstruction $\varepsilon_{a,\phi,f}$ of amplitude, phase and frequency of the AM/FM component as

$$\varepsilon_a \equiv \frac{\sqrt{\langle [A_{rec}(t)-A_{true}(t)]^2\rangle}}{\sqrt{\langle [A_{true}(t)]^2\rangle}},$$
$$\varepsilon_\phi \equiv \sqrt{1-|\langle e^{i(\phi_{rec}(t)-\phi_{true}(t))}\rangle|^2},$$
$$\varepsilon_f \equiv \frac{\sqrt{\langle [\nu_{rec}(t)-\nu_{true}(t)]^2\rangle}}{\nu_0} \tag{7.2}$$

where $A_{rec,true}(t)$, $\phi_{rec,true}(t)$, $\nu_{rec,true}(t)$ denote the reconstructed and true parameters, while $\nu_0 \equiv 2\pi$ for the (S)WFT and $\nu_0 \equiv \langle \nu_{true}(t)\rangle$ for the (S)WT (to account for the linear and logarithmic frequency resolutions of these transforms). The form of ε_ϕ was chosen so as to magnify the corresponding error and make it comparable to the others.

7.3. Advantages/drawbacks of synchrosqueezing

As can be seen from Fig. 3, if the components are not reliably represented in the WFT/WT, they will be not well reflected in the SWFT/SWT either. Therefore, it is not immediately obvious what are the advantages of SWFT/SWT over the usual WFT/WT, apart from being more visually appealing. Does synchrosqueezing improve the time or frequency resolution of the transform? Or does it allow more accurate reconstruction of the components? And generally, does the concentration of the TFR represent the primary characteristic of its performance? To answer these questions, we compare not only the performance of different reconstruction methods, but also the accuracy of estimates obtained by these methods from the usual and synchrosqueezed TFRs. For example, to understand whether synchrosqueezing improves frequency resolution it is sufficient to study the case of two interfering components: the resolution can be regarded as being increased only if one is able to extract the parameters of these components more accurately from the SWFT/SWT than from the underlying WFT/WT.

8. Simulation study

8.1. Resolution of two tones

Consider a two-tone signal, and define

$$s(t)=\cos(\nu_1 t+\varphi_1)+r\cos(\nu_2 t+\varphi_2), \quad \nu_1<\nu_2;$$
$$\Delta\nu \equiv \nu_2-\nu_1, \quad \Delta\phi(t)\equiv \Delta\nu t+(\varphi_2-\varphi_1), \quad \bar\nu \equiv \frac{\nu_1+\nu_2}{2}. \tag{8.1}$$

Then, according to (6.8) and (6.7), one has

$$2H_s(\omega,t)=e^{i(\nu_1 t+\varphi_1)}\left[\hat h_{\nu_1}(\omega)+r\hat h_{\nu_2}(\omega)e^{i\Delta\phi(t)}\right],$$
$$4|H_s(\omega,t)|^2=\hat h_{\nu_1}^2(\omega)+r^2\hat h_{\nu_2}^2(\omega)+2r\hat h_{\nu_1}(\omega)\hat h_{\nu_2}(\omega)\cos\Delta\phi(t),$$
$$\nu_H(\omega,t)=\frac{\nu_1\hat h_{\nu_1}^2(\omega)+\nu_2 r^2\hat h_{\nu_2}^2(\omega)+2r\bar\nu\hat h_{\nu_1}(\omega)\hat h_{\nu_2}(\omega)\cos\Delta\phi(t)}{\hat h_{\nu_1}^2(\omega)+r^2\hat h_{\nu_2}^2(\omega)+2r\hat h_{\nu_1}(\omega)\hat h_{\nu_2}(\omega)\cos\Delta\phi(t)}. \tag{8.2}$$

As can be seen, the supports of different tones in the WFT/WT can overlap with each other, so that $\hat h_{\nu_1}(\omega)\hat h_{\nu_2}(\omega)\neq 0$. In this situation one says that the two tones interfere in the TFR, which is reflected in the appearance of terms $\sim\hat h_{\nu_1}(\omega)\hat h_{\nu_2}(\omega)\cos\Delta\phi(t)$, causing time-variations in the squared TFR amplitude $|H_s(\omega,t)|^2$ and an instantaneous frequency $\nu_H(\omega,t)$. These terms will be called *interference terms*, and they are mainly responsible for the different types of TFR behavior, considered below.

Effects of the window/wavelet parameters. Different types of (S)WFT behavior (see Table 1) for a two-tone signal (8.1) are illustrated in Fig. 4 (for the (S)WT all is similar). As can be seen, for sufficiently high f_0, the WFT amplitude has two well-separated peaks at all times (Regime I). For lower f_0, although there are still two distinct peaks in the WFT at all times, "bridges" begin to appear between them at certain moments, reflecting interference between components and causing localized corruption of the WFT (Regime II). Regime III behavior appears when we further decrease f_0, so the window/wavelet frequency resolution becomes insufficient to resolve the two tones, and they become mixed, i.e. sometimes there are two peaks in the WFT amplitude and sometimes only one. Finally, for a very low value of the resolution parameter, the frequency resolution becomes so poor (although the time resolution is extremely sharp) that the two tones are completely merged, appearing as a single AM/FM component in TFR (Regime IV).

Comparing (a–d) and (m–p) in Fig. 4, it can be seen that, in agreement with what was said above, synchrosqueezing does not change the qualitative behavior of the TFR. Thus, when interference is present, it affects both the TFR amplitude and instantaneous frequency (8.2), which are used in the SWFT/SWT construction. Note that, for an SWT based on the wavelet with compact frequency support, the resolution of two tones was also considered in [76].

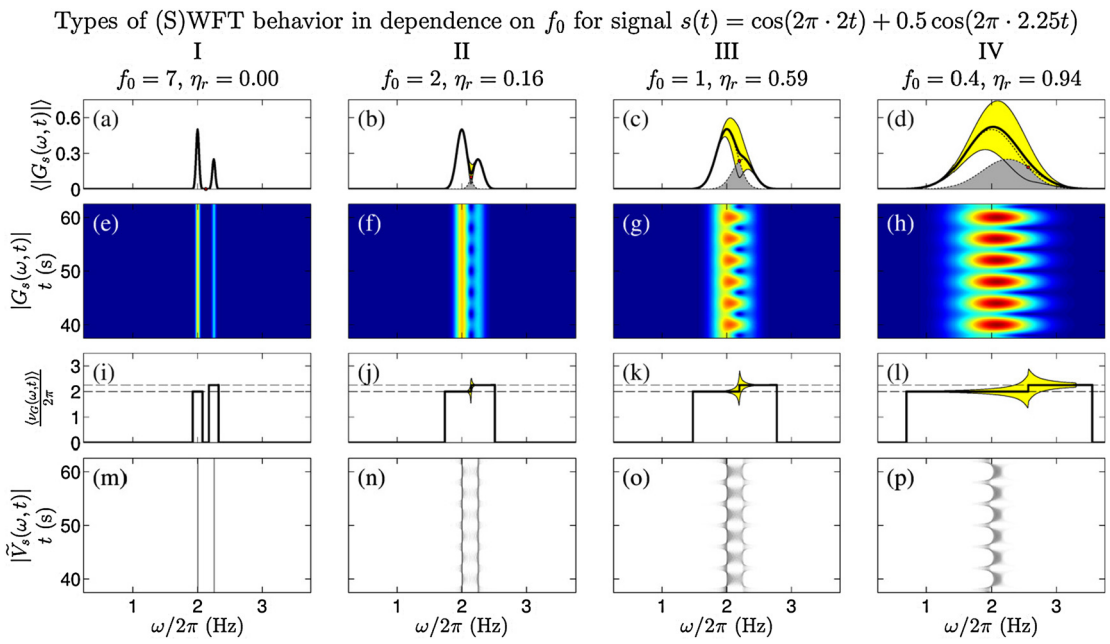

Fig. 4. Possible behaviors of the (S)WFT for a two-tone signal, sampled at 20 Hz for 500 s; for illustrational purposes, the (S)WFT was calculated using a compact support Gaussian window: $\hat{g}(\xi \notin [\xi_1(0.001), \xi_2(0.001)]) = 0$. (a–d): Time-averaged WFT amplitudes; dotted lines indicate WFTs of each tone separately, with the area shared by both of them being shown in gray. (e–h): WFT amplitudes in time–frequency domain. (i–l): Time-averaged WFT frequency $\nu_G(\omega, t)$, with dashed lines showing the frequencies of each tone ν_1, ν_2. (m–p): SWFT amplitudes in time–frequency domain. In (a–d) and (i–l), the yellow regions indicate $\pm\sqrt{2}$ standard deviations around the corresponding average values. The interference measure η_r is defined in (8.3). Note that (d, h, l, p) correspond to Regime IV only if we take $\epsilon_c > 0.06$ in its condition (see Table 1), while for lower precision smaller f_0 are needed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For the first and second tone separately (i.e. in the absence of the other one), the "interference-free" WFT/WT amplitude will equal $h_{\nu_1}(\omega)/2$ and $rh_{\nu_2}(\omega)/2$ (dotted lines in Fig. 4(a–d)). The area shared by these two interference-free amplitudes (shown in gray in Fig. 4(a–d)) in relation to the total area under $h_{\nu_1}(\omega)/2$ (first tone) or $rh_{\nu_2}(\omega)/2$ (second tone) can therefore serve as a measure of the interference-related corruption for each tone. The overall degree of interference η_r can then be taken as the maximum between the two:

$$\eta_r \equiv \max[\eta_1, \eta_2],$$
$$\eta_1 = r\eta_2 = \frac{\int \min[\hat{h}_{\nu_1}(\omega), r\hat{h}_{\nu_2}(\omega)]d\mu(\omega)}{\int \hat{h}_{\nu_1}(\omega)d\mu(\omega)}. \tag{8.3}$$

Proceeding as in Section 4.4, one can show that the maxima (over time) of the direct reconstruction errors of the tones are proportional to $\eta_{1,2}$. The condition $\eta_r \le \epsilon_c$ for the first type of TFR behavior (see Table 1) may therefore be approximated as

$$\textbf{(S)WFT}: \Delta\nu \ge \xi_2(\epsilon_c) - \xi_1(\epsilon_c); \quad \textbf{(S)WT}: \frac{\nu_2}{\nu_1} \ge \frac{\xi_2(\epsilon_c)}{\xi_1(\epsilon_c)}. \tag{8.4}$$

As will be seen below, this approximation is very accurate even for small and large r, especially in the case of the WFT. In fact, (8.4) implies $\eta_1 \le (1+r)\epsilon/2$ and $\eta_2 \le (1+r^{-1})\epsilon/2$.

Fig. 5 shows the parameter regions (found by numerical simulations) corresponding to each type of TFR behavior (see Table 1) for the two-tone signal (8.1). Clearly, the behavior of the (S)WFT (Fig. 5(a)) depends only on r and the product $f_0\Delta\nu$, so that the corresponding pictures remain the same for any f_0; they also do not change if one exchanges the tones $r \to 1/r$, which for symmetric $\hat{g}(\xi)$ is equivalent to $\eta_{1,2} \to \eta_{2,1}$. A similar situation will appear for the lognormal wavelet (S)WT (except $f_0\Delta\nu \to f_0 \log\frac{\nu_2}{\nu_1} = f_0\frac{\Delta\nu}{\nu_1} + O(\frac{\Delta\nu^2}{\nu_1^2})$). The Morlet wavelet, however, changes its form when one reduces f_0 (see Section 3.2), so the behavior of the corresponding (S)WT (Fig. 5(b)) also changes and becomes progressively asymmetric under $r \to 1/r$.

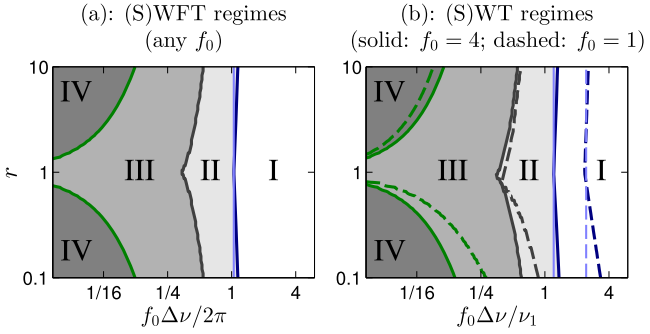

Fig. 5. Dependence of the TFR behavior on the parameters $r, \nu_1, \nu_2 \equiv \nu_1 + \Delta\nu$ of the two-tone signal (8.1) and Gaussian window (a) or Morlet wavelet (b) resolution parameter f_0. In both plots, light-blue lines show approximate Regime I boundaries according to (8.4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Direct vs ridge reconstruction. We now investigate numerically the performance of the different reconstruction methods (direct and ridge-based) for the two-tone signal (8.1). For each tone, the ridge curve (see Section 5.1) is extracted by finding at each time the two most dominant peaks in the TFR amplitude and picking the one nearest to the actual tone frequency. Note that this is done for each tone separately so, when they merge into a single peak, both will have the same extracted support. We then apply the direct (5.1) and ridge (5.3) reconstruction methods to obtain parameters of each tone and calculate the respective errors $\varepsilon_{a,\phi,f}$ (7.2).

Fig. 6 shows examples of amplitudes, phases and frequencies as reconstructed from the (S)WFT. As can be seen, all methods fail in Regime III (Fig. 6(d, e, f)), so any comparison does not make sense in that case. On the other hand, when the interference is not so strong (Regime II – Fig. 6(a, b, c)) the ridge estimates are clearly superior to the direct ones, both for the WFT and for the SWFT. The robustness of the ridge reconstruction to interference is furthermore confirmed by Supplementary Figs. 10 and 11, where

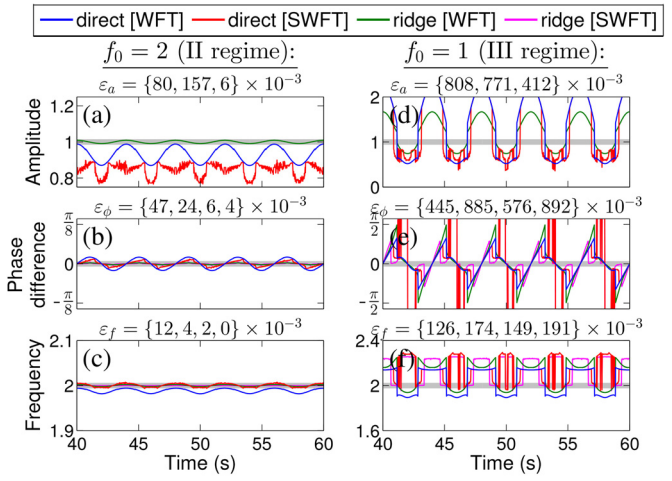

Fig. 6. Instantaneous parameters of the first tone in a signal (8.1), as reconstructed from the signal's WFT and SWFT (colored lines); the true values are indicated by thick gray lines. Reconstruction errors $\varepsilon_{a,\phi,f}$ (7.2) are shown in the same order as lines in legend, corresponding to direct[WFT] (blue), direct[SWFT] (red), ridge[WFT] (green), ridge[SWFT] (brown). In (a, d), ridge reconstruction from the SWFT is not shown as it is not appropriate for amplitude (see Section 5). In (b, e), the difference between the reconstructed and true phases is shown. The signal (8.1) was sampled at 50 Hz for 100 s, and it was simulated with $\nu_1 = \nu_2 - \pi/2 = 4\pi$, $r = 1.25$ and $\varphi_1 = \varphi_2 = 0$. Note that ordinate scale in (d, e, f) is four times zoomed out as compared to (a, b, c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the (S)WFT-based and (S)WT-based estimation errors are shown in dependence on f_0 and parameters of the signal (8.1). There one can also see that, for both methods, reconstruction errors are largely determined by the type of TFR behavior, being negligible in Regime I, increasing in II and becoming very large in III–IV.

WFT/WT vs SWFT/SWT. From Fig. 4 it is clear that the interference between components propagates from the WFT/WT to the SWFT/SWT, but it remains to be established whether it is reduced/magnified in the process of synchrosqueezing, and to which extent. In practice, however, it appears that synchrosqueezing does not change the degree of interference between the components. Thus, when the interference is not overwhelmingly strong, the WFT and SWFT provide almost identical ridge estimates for all parameters (Fig. 6(a, b, c)). On the other hand, the SWFT-based direct estimates of amplitude (phase and frequency) are slightly less (more) accurate than the WFT-based ones (Fig. 6(a, b, c)), but this is likely to be on account of the difference between the TFS definitions for these transforms (see Section 5.1). Supplementary Figs. 10 and 11 provide detailed information regarding relative performance of the WFT/WT and SWFT/SWT (in terms of component reconstruction) in dependence on signal parameters. It can be concluded that synchrosqueezing does not allow better resolution of two interfering tones, and therefore does not increase the frequency resolution of the TFR.

8.2. Amplitude modulation

We now consider the AM component with sinusoidal amplitude modulation, and define

$$\begin{aligned} s(t) &= [1 + r_a \cos(\nu_a t + \varphi_a)] \cos(\nu t + \varphi) \\ &= \cos\phi(t) + \frac{r_a}{2}\Big[\cos\big(\phi(t) + \phi_a(t)\big) + \cos\big(\phi(t) - \phi_a(t)\big)\Big]; \\ \phi(t) &\equiv \nu t + \varphi, \quad \phi_a(t) \equiv \nu_a t + \varphi_a, \end{aligned} \tag{8.5}$$

where the amplitude definition implies $r_a \le 1$, $0 \le \nu_a < \nu$.

It is evident that the AM component (8.5) can be represented as a sum of three tones: the main one at frequency ν and two equal amplitude side tones at $\nu \pm \nu_a$; the latter appear as a result of the amplitude modulation and therefore will be called "AM-induced". Therefore, all the formulas and classification for multitone signals also apply for the AM component (8.5), for which (6.8) become

$$\begin{aligned} 2H_s(\omega,t) &= e^{i\phi(t)}\Big[\hat h_\nu(\omega) + \frac{r_a}{2}\hat h_{\nu+\nu_a}(\omega)e^{i\phi_a(t)} \\ &\quad + \frac{r_a}{2}\hat h_{\nu-\nu_a}(\omega)e^{-i\phi_a(t)}\Big], \\ 4|H_s(\omega,t)|^2 &= \hat h_{\nu_1}^2(\omega) + \frac{r_a^2}{4}\big[\hat h_{\nu+\nu_a}^2(\omega) + \hat h_{\nu-\nu_a}^2(\omega)\big] \\ &\quad + r_a\hat h_\nu(\omega)\big[\hat h_{\nu+\nu_a}(\omega) + \hat h_{\nu-\nu_a}(\omega)\big]\cos\phi_a(t) \\ &\quad + \frac{r_a^2}{2}\hat h_{\nu+\nu_a}(\omega)\hat h_{\nu-\nu_a}(\omega)\cos 2\phi_a(t), \\ \nu_H(\omega,t) &= \nu + \frac{r_a\nu_a}{2}\frac{\hat h_{\nu+\nu_a}(\omega) - \hat h_{\nu-\nu_a}(\omega)}{4|H_s(\omega,t)|^2} \\ &\quad \times \Big(\frac{r_a}{2}[\hat h_{\nu+\nu_a}(\omega) + \hat h_{\nu-\nu_a}(\omega)] + \hat h_\nu(\omega)\cos\phi_a(t)\Big). \end{aligned} \tag{8.6}$$

Comparing with the two-tone case (8.2), we now have two interference terms, $\sim\cos\phi_a(t)$ and $\sim\cos 2\phi_a(t)$, responsible for interference between the main tone and the AM-induced ones, and between the two AM-induced ones, respectively.

Effects of the window/wavelet parameters. From (8.5) it is clear that one can look at the AM component from two different perspectives: either as an oscillation with amplitude modulation, or as the superposition of three tones with particular amplitude, phase and frequency relationships. The way in which the component (8.5) is represented in the TFR is determined by the window/wavelet parameters: if the time resolution is high (f_0 small), it will be treated as a single component; if the frequency resolution is high (f_0 large), it will be treated as three independent tones. This is illustrated in Fig. 7, where different types of (S)WFT behavior (see Table 1) are shown for an AM signal (8.5). Note that the TFR behavior is not changed qualitatively after synchrosqueezing, as usual.

In the present case, we define an interference measure η_a as the ratio of the area shared by the interference-free TFR amplitude for the main tone, $\hat h_\nu(\omega)/2$, and the sum of such amplitudes for the AM-induced tones, $\frac{r_a}{2}[\hat h_{\nu+\nu_a}(\omega) + \hat h(\nu-\nu_a)]/2$, to the total area under the latter:

$$\eta_a = \frac{\int \min\Big(\hat h_\nu(\omega),\, \frac{r_a}{2}\big[\hat h_{\nu+\nu_a}(\omega) + \hat h_{\nu-\nu_a}(\omega)\big]\Big)d\mu(\omega)}{\frac{r_a}{2}\int\big[\hat h_{\nu+\nu_a}(\omega) + \hat h_{\nu-\nu_a}(\omega)\big]d\mu(\omega)}. \tag{8.7}$$

This definition is motivated by (8.3) and the analogy between (8.6) and (8.2). Note that η_a in fact quantifies the degree to which the component behaves as a single entity in the TFR.

Numerically found parameter regions corresponding to each type of TFR behavior (see Table 1) for the AM signal (8.5) are shown in Fig. 8. Similarly to the case of a two-tone signal, the (S)WFT behavior in the present case (Fig. 8(a)) depends only on r_a and $f_0\nu_a$, and not on f_0 or ν_a separately (this, however, ceases to be true for asymmetric $\hat g(\xi)$). In contrast, the behavior of the (S)WT depends, apart from r_a, on both f_0 and ν_a/ν. This will generally be true for any wavelet, since AM-induced tones are located around the main one symmetrically on a linear, rather than logarithmic, frequency scale. Note also that $\nu_a < \nu$ is needed for the analytic approximation (2.4) to hold, which is required for a reliable time–frequency analysis (see Section 2); this implies

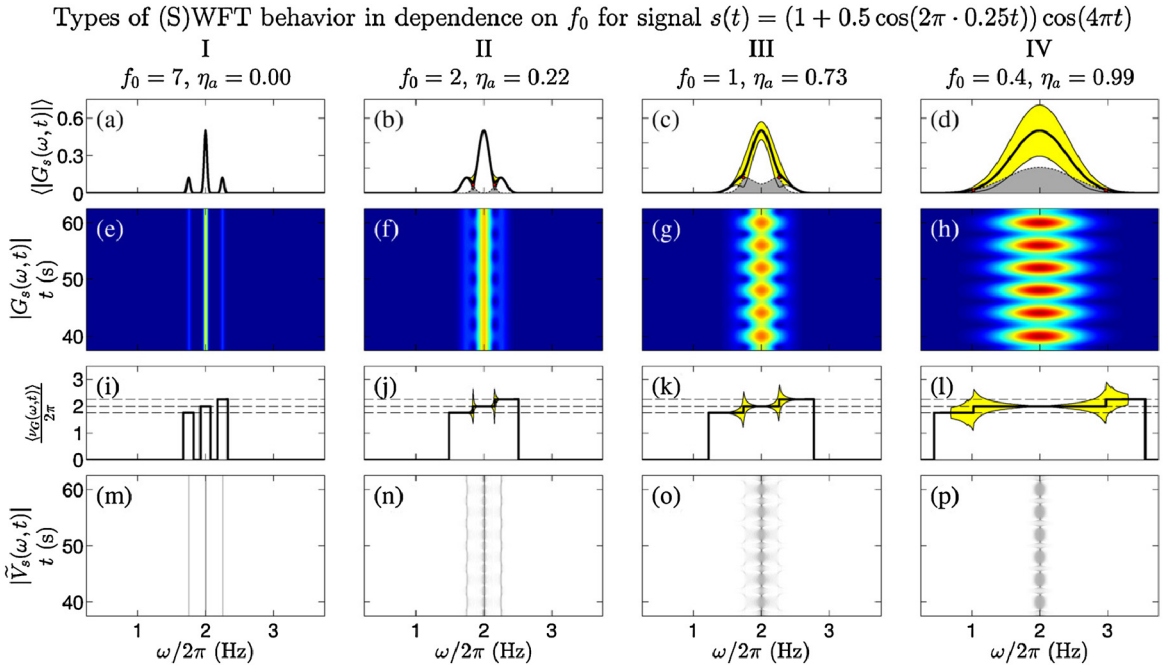

Fig. 7. Possible behaviors of the (S)WFT for an AM signal, sampled at 20 Hz for 500 s. The interference measure η_a is defined in (8.7), and the dotted lines in (a–d) show $\hat{h}_\nu(\omega)/2$ and $\frac{r_a}{2}[\hat{h}_{\nu+\nu_a}(\omega)+\hat{h}_{\nu-\nu_a}(\omega)]/2$ (see text), with the area shared by both of them being shown in gray. All other details are the same as in Fig. 4.

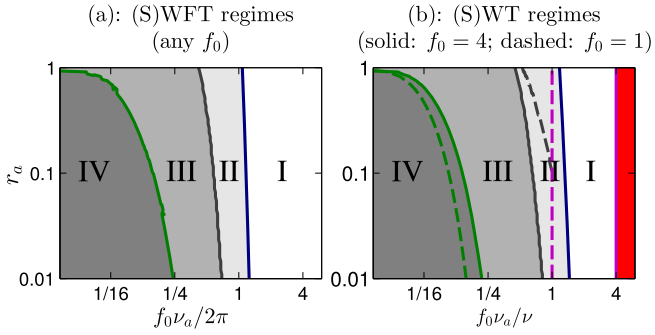

Fig. 8. Dependence of the TFR behavior on the parameters r_a, ν, ν_a of the AM signal (8.5) and Gaussian window (a) or Morlet wavelet (b) resolution parameter f_0. In (b), the red area separated by magenta line indicates region where $\nu_a > \nu$, which is prohibited (see text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

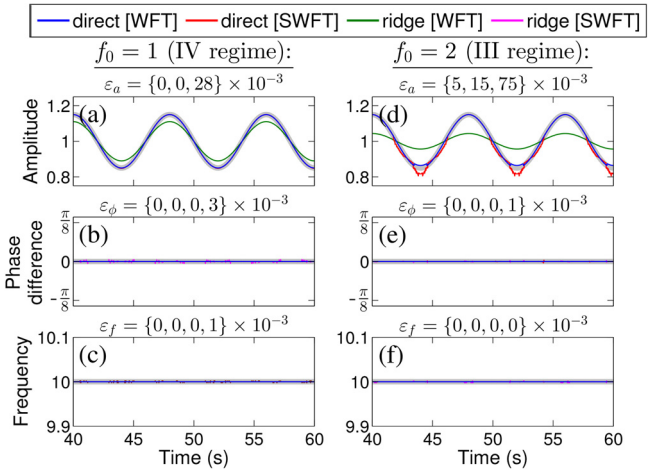

Fig. 9. Instantaneous parameters of the AM component (8.5) with $\nu/2\pi = 10$, $r_a = 0.15$, $\nu_a = \pi/8$ and $\varphi = \varphi_a = 0$ (sampled at 50 Hz for 100 s), as reconstructed from its WFT and SWFT (colored lines); the true values are indicated by thick gray lines. All other details are the same as in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$f_0\nu_a/\nu \le f_0$, and the corresponding "prohibited" region is shown in red in Fig. 8(b).

Direct vs ridge reconstruction. To investigate the performance of different reconstruction methods in the present case, we first extract the ridge curve by selecting the highest TFR amplitude peaks at each time, and then estimate the component's parameters by the direct (5.1) and ridge (5.3) methods. Examples of amplitudes, phases and frequencies reconstructed from the (S)WFT are shown in Fig. 9. It is clear that, for both the WFT and SWFT, direct amplitude estimates are much more accurate than ridge estimates (Fig. 9(a, d)). Next, for constant frequency components, amplitude modulation does not affect phase and frequency reconstruction, which can therefore be recovered exactly by both methods (Fig. 9(b, c, e, f)). This, however, is a unique property of windows with symmetric $\hat{g}(\xi)$ only, for which the WFT of an AM component at each time has peak at $\omega = \nu$ and is symmetric around it. For asymmetric windows and for the WT, phase and frequency estimates will contain errors, with direct estimates being the more accurate.

The full dependences of the reconstruction errors on the parameters of an AM signal (8.5) are shown in Supplementary Figs. 12 and 13 for the (S)WFT and (S)WT, respectively. Although the errors depend very much on the type of TFR behavior, the direct estimates are almost always superior in the present case (for Regime IV, they are exact).

WFT/WT vs SWFT/SWT. The issues related to amplitude representation seem to affect the WFT/WT and the SWFT/SWT to a similar extent (see Fig. 7). This is confirmed by Fig. 9 and Supplementary Figs. 12 and 13, which show that the quality of the SWFT/SWT-based estimates (both direct and ridge) for all parameters is either the same as, or worse than, that of the WFT/WT-based ones. Hence, synchrosqueezing does not improve the representation of amplitude modulation.

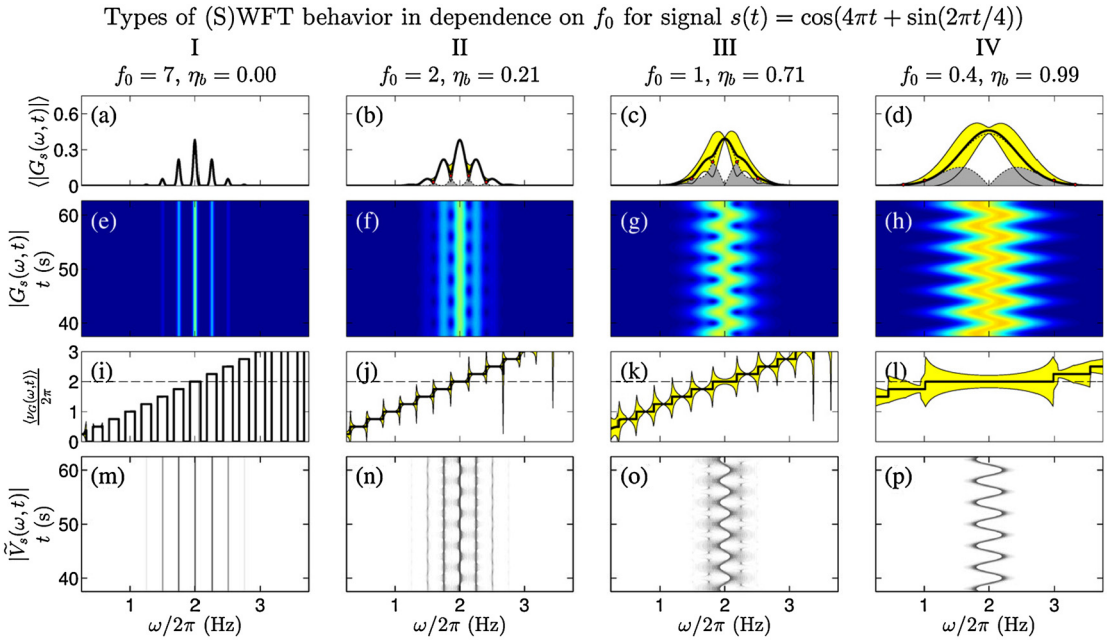

Fig. 10. Possible behaviors of the (S)WFT for an FM signal, sampled at 20 Hz for 500 s. The interference measure η_b is defined in (8.10), and the dotted lines in (a–d) show $\hat{h}^{(+)}(\omega)/2$ and $|\hat{h}^{(-)}(\omega)|/2$ (see text), with the area shared by both of them being shown in gray. All other details are the same as in Fig. 4.

8.3. Frequency modulation

We now consider the FM component with sinusoidal frequency modulation, and define

$$\begin{aligned} s(t) &= \cos[\phi_0(t) + r_b \sin\phi_b(t)] \\ &= \sum_{n=-\infty}^{\infty} J_n(r_b)\cos[\phi_0(t) + n\phi_b(t)]; \\ \phi_0(t) &\equiv \nu t + \varphi, \quad \phi_b(t) \equiv \nu_b t + \varphi_b, \end{aligned} \tag{8.8}$$

where we have used the expansion $e^{ia\sin\theta} = \sum_{n=-\infty}^{\infty} J_n(a)e^{in\theta}$.

According to the definition of phase (2.1), one should have $\partial_t[\phi_0(t) + r_b\sin\phi_b(t)] > 0 \Rightarrow \nu > r_b\nu_b$. Furthermore, the analytic approximation (2.4) to hold, which is a prerequisite for any time–frequency analysis (see Section 2), all non-negligible terms in the expansion (8.8) should correspond to positive frequencies $\nu + n\nu_b > 0$. This can be formulated as $\nu > n_J(r_b)\nu_b$, where $n_J(r_b)$ is the maximal order of $J_n(r_b)$ for which it is non-negligible, i.e. higher than some specified threshold ϵ_J:

$$n_J(r_b) = \max[n : |J_n(r_b)| > \epsilon_J] \tag{8.9}$$

Thus, one can effectively reduce the sum in (8.8) to just the terms with $|n| \le n_J(r_b)$. Under precision $\epsilon_J = 0.01$, one has $n_J(r_b \in [0.02, 0.3]) \approx 1$ and $n_J(r_b \in [0.3, 0.8]) \approx 2$.

Clearly, the FM component (8.8) can be represented in the form of a multitone signal: the main tone at frequency ν, and many side tones of pairwise-equal amplitudes at $\nu \pm n\nu_b$; the latter appear due to the frequency modulation and will be called "FM-induced". Therefore, the classification and formulas for the multitone signals apply here as well, and the expressions for $|H_s(\omega,t)|^2$ and $\nu_H(\omega,t)$ follow from (6.8); they are quite complicated and do not give much insight, so we omit them.

Effects of the window/wavelet parameters. Similarly to the previously considered case of AM component, the FM component (8.8) can be perceived either as an oscillation with frequency modulation, or as a multitone signal with particular relationships between the characteristics of the tones. The way it is represented in the TFR is determined by the window/wavelet parameters, as illustrated in Fig. 10 for the (S)WFT. Note that the behaviors of the WFT and SWFT are qualitatively the same.

Rather intuitively, we define the interference measure η_b as

$$\begin{aligned} \eta_b &= \frac{\int \min(\hat{h}^{(+)}(\omega), |\hat{h}^{(-)}(\omega)|)d\mu(\omega)}{\int |\hat{h}^{(-)}(\omega)|d\mu(\omega)}; \\ \hat{h}^{(+)}(\omega) &= J_0(r_b)\hat{h}_\nu(\omega) \\ &\quad + \sum_{n=1}^{\infty} J_{2n}(r_b)\Big[\hat{h}_{\nu+2n\nu_a}(\omega) + \hat{h}_{\nu-2n\nu_a}(\omega)\Big], \\ \hat{h}^{(-)}(\omega) &= \sum_{n=1}^{\infty} J_{2n-1}(\omega)\Big[\hat{h}_{\nu+(2n-1)\nu_a}(\omega) - \hat{h}_{\nu-(2n-1)\nu_a}(\omega)\Big]. \end{aligned} \tag{8.10}$$

One motivation behind considering such η_b, $\hat{h}^{(+)}(\omega)$ and $\hat{h}^{(-)}(\omega)$ is that in the limit of vanishing frequency modulation $\nu_b \to 0$, where FM component reduces to a single tone, one has $\eta_b \to 1$, as desired. Furthermore, in this limit $\hat{h}^{(+)}(\omega) \to \hat{h}_\nu(\omega)$ and $\hat{h}^{(-)}(\omega) \to 0$ for any r_b, i.e. all energy concentrates in $\hat{h}^{(+)}(\omega)$.

In what follows, we consider $r_b \in [0,1]$, in which case $n_J \lesssim 2$ and consideration can be restricted to only five terms $\sim J_{0,\pm1,\pm2}(r_b)$ in (8.8), corresponding to one dominant tone at the main frequency ν and two pairs of FM-induced ones with smaller amplitudes. For higher r_b, e.g. $r_b \gtrsim 1.45$, $J_1(r_b)$ becomes higher than $J_0(r_b)$ so, instead of one main tone, there will be two dominant side tones, as well as additional FM-induced tones to be considered (as $J_{n>2}(r_b)$ will no longer be negligible). With further increase of r_b, the FM-induced tones with frequencies further distant from ν become dominant, and additional terms become non-negligible. Note, however, that η_b (8.10) remains appropriate for any r_b.

Fig. 11 shows the numerically found parameter regions corresponding to each type of TFR behavior (see Table 1) for the FM signal (8.8). As usual, the behavior of the (S)WFT depends on r_b and the product $f_0\nu_b$ (at least for symmetric $\hat{g}(\xi)$), while the behavior of the (S)WT depends on all r_b, f_0 and ν_b/ν. Note also the red area in Fig. 8(b), which indicates the region where one of the

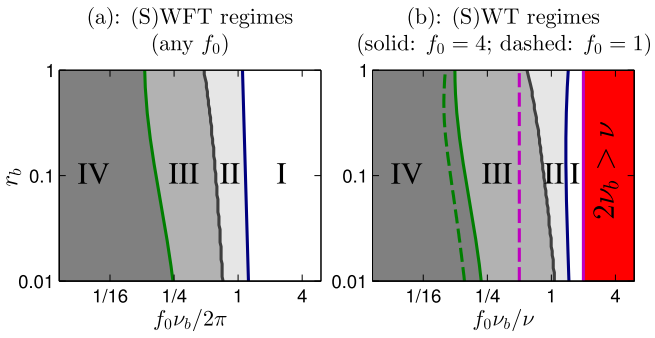

Fig. 11. Dependence of the TFR behavior on the parameters r_b, ν, ν_b of the FM signal (8.8) and Gaussian window (a) or Morlet wavelet (b) resolution parameter f_0. In (b), the red area separated by magenta line indicates region where $2\nu_a > \nu$, which is prohibited (see text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

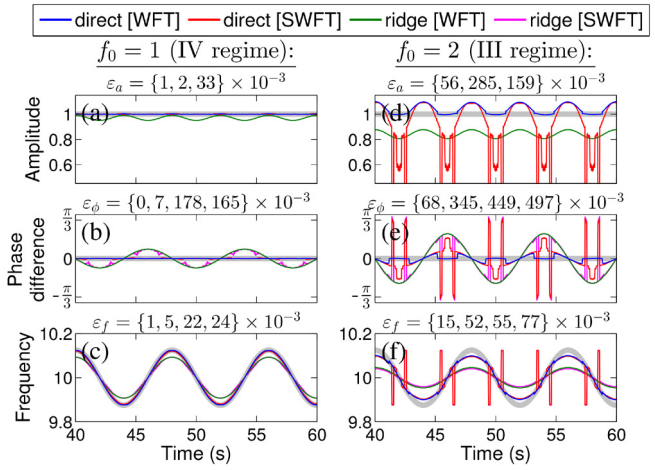

Fig. 12. Instantaneous parameters of the FM component (8.8) with $\nu/2\pi = 10$, $r_b = 1$, $\nu_b = \pi/8$ and $\varphi = \varphi_b = 0$ (sampled at 50 Hz for 100 s), as reconstructed from its WFT and SWFT (colored lines); the true values are indicated by thick gray lines. All other details are the same as in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

prominent FM-induced tones has negative frequency, so that everything becomes ill-defined (see Section 2).

Direct vs ridge reconstruction. We investigate the performance of different reconstruction methods for the FM signal (8.8) by first forming the ridge curve $\omega_p(t)$ from the highest TFR amplitude peaks at each time and then, based on it, estimating the component's parameters by the direct (5.1) and ridge (5.3) methods. Examples of amplitudes, phases and frequencies reconstructed from the (S)WFT are shown in Fig. 12, while Supplementary Figs. 14 and 15 show (S)WFT-based and (S)WT-based reconstruction errors in dependence on signal parameters. Similarly to the case of the AM component, the estimates obtained using direct methods are exact in Regime IV and are almost always preferred to ridge estimates. All reconstruction errors, as always, are significantly correlated with the TFR behavior type. Note that, while amplitude modulation does not introduce errors into the (S)WFT-based frequency estimates if $\hat{g}(\xi)$ is symmetric (as was seen in Fig. 9), frequency modulation can seriously affect amplitude estimates, especially if reconstructed from ridges (see Fig. 12(a, d)).

WFT/WT vs SWFT/SWT. From Fig. 10 it is clear that synchrosqueezing does not change qualitative aspects of the representation of frequency modulation in TFR. Furthermore, Fig. 12 and Supplementary Figs. 14 and 15 indicate that, for both reconstruction methods, the SWFT/SWT-based estimates of the FM component's parameters are either the same as the WFT/WT-based ones, or can even be much worse. Thus, in contrast to the WFT/WT, in the case of strong frequency modulation, side peaks in the SWFT/SWT (see Fig. 10(o)) can become dominant at certain times, and selecting them as part of the ridge curve introduces rapid jumps into the estimated parameters (Fig. 12(d, e, f)). In conclusion, synchrosqueezing does not allow better reconstruction of frequency or amplitude (see Section 8.2) variations, and therefore it does not improve the TFR's time resolution.

8.4. Noise and its effects

We study the effects of noise by considering the signal

$$s(t) = \cos \nu t + \frac{\sigma}{\sqrt{2}}\zeta(t) \tag{8.11}$$

where $\zeta(t)$ denotes white Gaussian noise of unit variance, and $\sigma/\sqrt{2}$ is its standard deviation. The $1/\sqrt{2}$ multiplier is introduced to make σ equivalent to the noise-to-signal ratio (standard deviation of the noise divided by that of the signal).

Effects of the window/wavelet parameters. In general, noise can be regarded as a superposition of infinitely many tones with random independent phase shifts and frequency-dependent amplitudes (for white noise these amplitudes are the same: $|\hat{\zeta}(\omega)| = \text{const}$). Thus, in some sense the case of noise is similar to the case of interfering components; but, for noise, we have (in theory) infinitely many tones with infinitely close frequencies. It is therefore hard to devise any useful classification of the TFR behavior, so we restrict ourselves to qualitative and illustrative considerations.

The (S)WFT of the signal (8.11) is presented in Fig. 13 for different values of f_0. As can be seen, both WFT and SWFT are equally affected by noise. Furthermore, we observe that the higher f_0 is, the better one can distinguish a genuine tone within the noise. This is because the frequency range where noise tones are picked up while calculating the WFT or WT is determined by the spread of $\hat{g}(\xi)$ or $\hat{\psi}(\xi)$, which is inversely proportional to f_0. Note that, for this reason, white noise intensity in the WT amplitude $|W_s(\omega, t)|$ is proportional to $\sqrt{\omega}$ (not shown), since the spread of $\hat{\psi}(\omega_\psi \xi/\omega)$ (which is used in the WT computation (3.4)) increases with ω.

We conclude that, for the single tone corrupted by noise (8.11), the higher f_0 is the better (in fact, the best one can use here is the usual Fourier transform, which provides the maximum possible frequency resolution). However, if the signal represents a noise-corrupted component with amplitude or frequency modulation, then one needs to choose f_0 as a compromise between reducing the effect of noise and ensuring that the AM/FM component is still represented as a single entity in the TFR; for the previously considered cases (8.5) and (8.8) this will correspond to selecting f_0 near the border of Regime IV.

Direct vs ridge reconstruction. We now investigate the noise-robustness of different reconstruction methods by estimating the characteristics of the tone in (8.11) from the signal's TFR. At each time, we find the corresponding ridge curve $\omega_p(t)$ as the frequency of the maximum among the TFR amplitude peaks (in the case of the WT, they are multiplied by $1/\sqrt{\omega}$ to account for the white noise power scaling mentioned above). Examples of the resultant estimates are presented in Fig. 14, while Supplementary Figs. 16 (WFT/SWFT) and 17 (WT/SWT) show the full dependences of the reconstruction errors on f_0 and the noise intensity σ. Similarly to the case of two tones (Section 8.1), ridge estimates of all parameters (amplitude, phase and frequency) appear to be more accurate than the corresponding direct estimates. This is true for any f_0, σ and TFR type, which indicates the exceptional noise-robustness of the ridge reconstruction method.

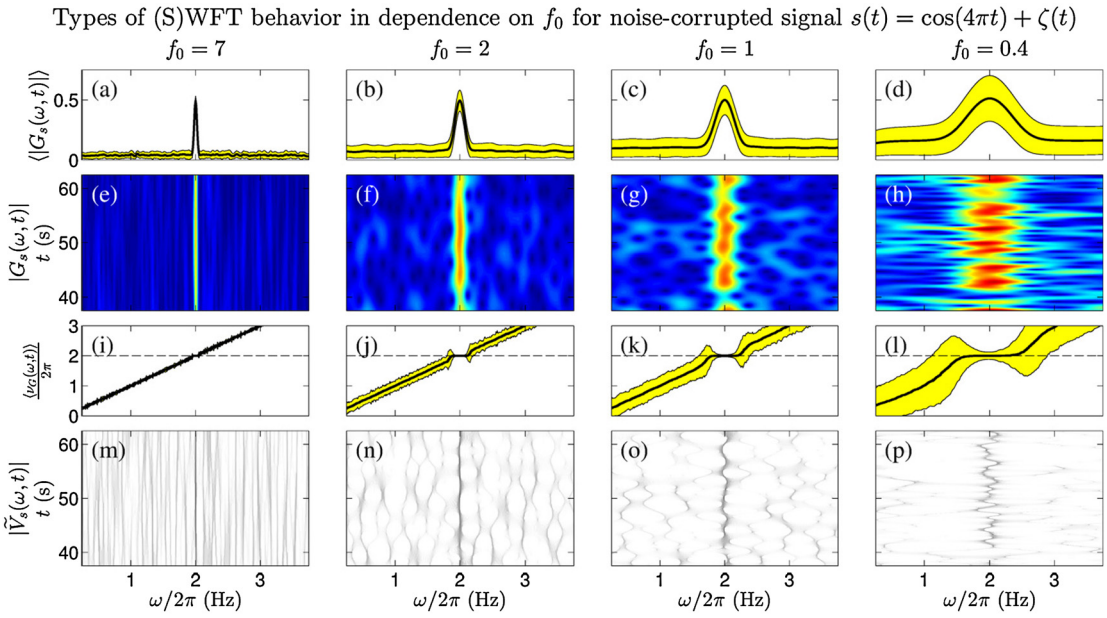

Fig. 13. Behavior of the (S)WFT for a single tone plus noise signal (sampled at 20 Hz for 500 s) in dependence on the Gaussian window resolution parameter f_0. All other details are the same as in Fig. 4.

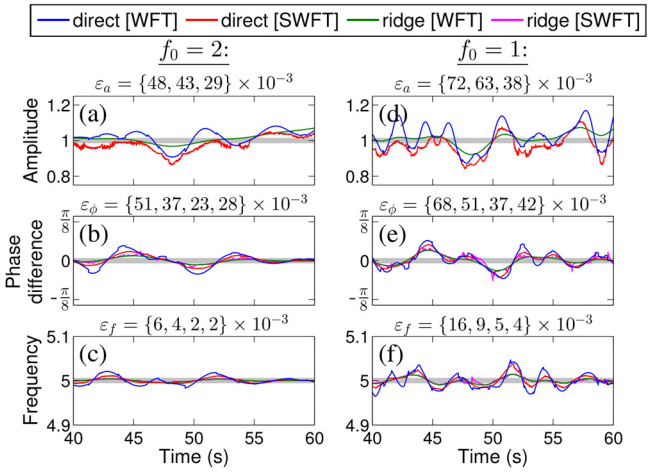

Fig. 14. Instantaneous parameters of the tone embedded in noise (8.11) with $\nu/2\pi = 5$ and $\sigma = 0.5$ (sampled at 50 Hz for 100 s), as reconstructed from its WFT and SWFT (colored lines); the true values are indicated by thick gray lines. All other details are the same as in Fig. 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

WFT/WT vs SWFT/SWT. Fig. 13 suggests that synchrosqueezing does not change the noise corruption of the TFR, and this is confirmed by Fig. 14 and Supplementary Figs. 16 and 17. Thus, the quality of the SWFT/SWT-based estimates is the same as that of the WFT/WT-based ones (there are slight differences between the corresponding direct estimates, but they can be attributed to inequivalence of the TFS definitions, see Section 5.1). Hence, the SWFT/SWT and WFT/WT are susceptible to noise to the same extent.

9. Selecting window/wavelet parameters

As seen in previous sections, the TFR behavior depends drastically on the choice of window/wavelet parameters, which in turn strongly affects the extraction and reconstruction of the components. Moreover, in the analysis of real data, inappropriate resolution can even lead to wrong conclusions, e.g. that there exist a number of oscillations while there is actually only a single one but with amplitude and/or frequency modulation. Thus, the choice of appropriate window/wavelet parameters is crucial for both the interpretation and quantitative estimation of the signal structure.

Unfortunately, there is no universal choice, and the optimal resolution parameter depends on the signal composition. It should be chosen in such a way as to resolve the independent components (Regime I), e.g. tones, but at the same time to represent AM/FM components as single curves in the TFR (Regime IV). In other words, each component, with or without amplitude/frequency modulation, at each time should ideally be mapped to a single individual peak and time–frequency support.

Usually, however, this cannot be achieved for all components simultaneously, so the optimal choice requires compromise. Furthermore, in real situations, one often does not know for sure whether it is appropriate to represent the underlying process as a single AM/FM component, or as separate tones. This arbitrariness is unavoidable, but the TFR-based methods give the possibility of specifying a criterion by choosing appropriate window/wavelet parameters. Methods that can be used for this task are considered below. An extensive overview of the methods themselves and of the related literature is presented in [9]; here we extend those works by concentrating on some significant drawbacks which can greatly restrict the application of the methods in practice. Note that, because synchrosqueezing does not change the resolution properties of the transform (see Sections 8 and 11), we consider below only the case of the WFT/WT (the optimal parameters for the SWFT/SWT will be the same).

Remark 9.1. Although in what follows we consider optimization of the (single) resolution parameter f_0, the same methods can be used quite generally for choosing any set of window/wavelet parameters. For example, it might be advantageous to use a chirped window $g(t) \sim e^{-t^2/2f_0^2} e^{i\alpha t^2}$ in the WFT [19,52–55,77–83], and selection of the optimal pair $\{f_0, \alpha\}$ in this particular case does not differ qualitatively from the selection of f_0 alone. However, as will be seen, even a single f_0 is hard to choose appropriately.

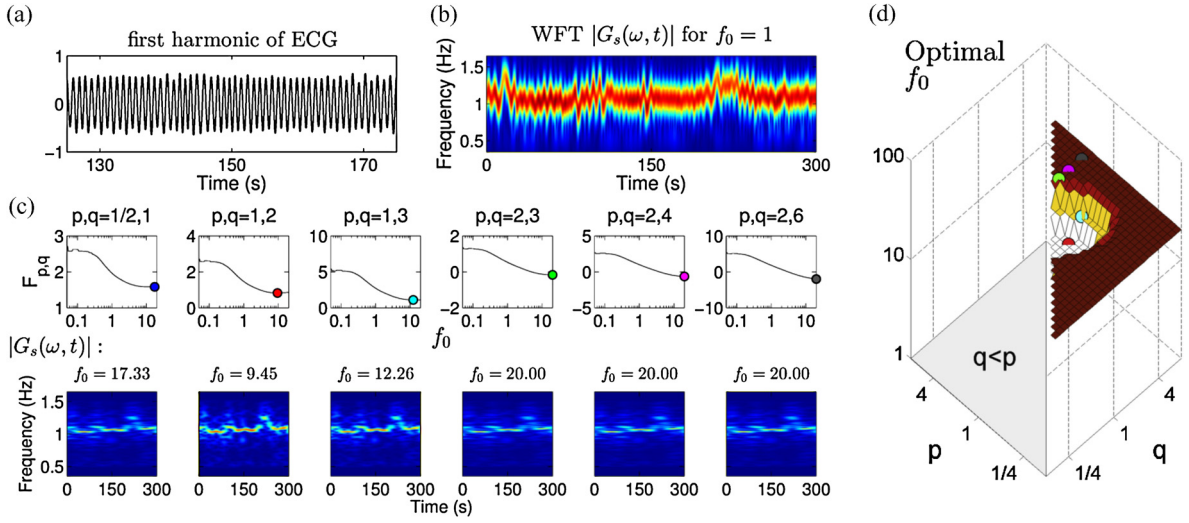

Fig. 15. Adapting the WFT according to (9.2) for the first harmonic of the human ECG signal (3-lead, with electrodes placed on shoulders and the lowest left rib [86]), as obtained by bandpass filtering it in the region [0.5, 1.5] Hz. (a): Signal in the time domain. (b): The signal's WFT calculated for $f_0 = 1$. (c): Examples of the dependencies of $F_{p,q}$ (9.2) on f_0 for different pairs of p, q (upper row); optimal f_0 minimizing $F_{p,q}$ are shown by filled circles, with the WFTs calculated for these values being presented at the bottom. (d): Optimal f_0 in dependence on p and q in (9.2). The signal was sampled at 40 Hz for 30 min, but only central 5 min part is taken for analysis, while other is used for padding to eliminate boundary effects. The WFT was calculated in the frequency range [0, 2] Hz, and the optimal f_0 was searched in a range [0.05, 20].

9.1. "Monocomponent" approaches

Optimization of the TFR parameters for the representation of a single AM/FM component (possibly embedded in noise) was considered in [78,84,85]. The main aim of these approaches is usually to estimate the component's instantaneous frequency $\nu(t) = \phi'(t)$ (see [21,42,43] for an overview of related concepts and algorithms), which is reconstructed from the optimized TFR. Real signals, however, are rarely monocomponent, and such an assumption is often too restrictive. In fact, it is multicomponent signals for which time–frequency analysis is most useful (see Section 2) and, as will be shown below, the main difficulty of adapting f_0 lies in estimating the number of components. We will therefore not consider the monocomponent case but will proceed with a more general approach.

9.2. Functional approaches

A particular class of methods for selecting an appropriate f_0 based on the signal's structure was considered in [77,82,87–89]. According to these works, the optimal window/wavelet parameters f_0^* can be selected as being those that minimize a suitably chosen functional $F[\cdot]$ of the signal's TFR:

$$f_0^* = \operatorname{argmin}_{f_0}\Big(F[H_s\{f_0\}(\omega,t)]\Big) \tag{9.1}$$

where $H_s\{f_0\}(\omega,t)$ denotes the WFT/WT calculated using the chosen f_0. Almost all functional approaches proposed so far [77,82,87–89] can be reduced to the minimization of

$$F_{p,q}[H_s(\omega,t)] = \log\frac{\left(\int |H_s(\omega,t)|^p d\mu(\omega)dt\right)^{q/p}}{\int |H_s(\omega,t)|^q d\mu(\omega)dt}, \quad q > p > 0. \tag{9.2}$$

For example, the Renyi entropy of order α, which was thoroughly investigated in [87,90–92], corresponds to $p = 2$, $q = 2\alpha$; the measures proposed in [88] correspond to $p = 1$, $q = 2$; while maximization of the ratio of the fourth power of L_4-norm to the squared L_2-norm proposed in [77,82,89] corresponds to the minimization of $F_{2,4}$. Note that the numerator power q/p in (9.2) guarantees that the functional does not depend on window/wavelet normalization (that can depend on f_0), while the inequality $q > p$ establishes that $f_0^* \to \infty$ ($f_0^* \to 0$) for the single tone (delta-peak) signal.

For signals consisting of Gaussian pulses $\sim e^{-a(t-t_0)^2}\cos(bt+\varphi)$, delta-peaks, tones and/or components whose frequency modulation contains a trend-like term $\sim t^a$, $a > 0$ (with all components being allowed to exist only during certain time intervals), it was shown [77,82,87–89] that using the functional (9.2) one can indeed select appropriate window/wavelet parameters. An interesting feature is that for $p \gtrsim 2$ one gets an f_0^* that maximizes the concentration of only one of the signal's components, while for $p \lesssim 2$ the optimal resolution parameter is chosen based on a compromise between representing well all components. This was previously noted in [88], and gives a good reason to use the functional $F_{1,2}$ (instead e.g. of the previously proposed $F_{2,4}$), which we would also recommend.

The component's amplitude/frequency, however, might simply oscillate around some mean value, not subject to any trends. In this case, which often appears in practice, $F_{p,q}$ (9.2) will usually prefer a more concentrated representation of such a component as a sum of tones (Regime I) rather than as a single entity (Regime IV), being therefore unable to resolve the corresponding ambiguity. For example, it can easily be verified that for any $p > 0$ and $q > p$ in (9.2) one will get $f_0^* \to \infty$ for simple AM and FM components (8.5) and (8.8). This issue is not restricted to sinusoidal amplitude/frequency modulation, but occurs quite generally for persistent AM/FM components occupying well-defined frequency bands, as illustrated in Fig. 15 on the example of ECG signal. As can be seen, although the heart rate modulation is generally quite complex, minimizing $F_{p,q}$ still tries to separate it into tones, giving too large f_0^* for which the TFR becomes unsuitable for the analysis (e.g. one cannot extract the instantaneous heart frequency from it).

In addition to (9.2), another measure of TFR performance was suggested in [93–95]. It is based on the relationships between various characteristics of the different TFR peaks, such as their frequencies, widths and heights. The optimal TFR can then be selected by minimizing this measure, which can therefore be taken as a functional in (9.1). This approach was shown to perform very well on signals containing chirps, but it is not clear how well it will perform when the component can be represented both as a single entity and as a sum of tones. Most likely, it will also prefer a

more concentrated representation as a sum of tones, thus suffering from the same drawback as the previously considered class of measures (9.2).

9.3. Other approaches

Apart from the methods already considered, there exist many other ways of selecting optimal window/wavelet parameters. For example, it was proposed [67,79,80] to construct an adaptive TFR based on certain characteristics (e.g. ridge frequencies) estimated from some initial TFR. It is easy to see, however, that these approaches are susceptible to the choice of initial f_0 whereas, if the latter is chosen adaptively based on some variation of the functional (9.2), as in [81], then the drawbacks of this functional approach will apply.

Another idea is to optimize the WFT/WT based on its local moments in the time–frequency plane [83]. This method, however, is very expensive computationally, being $O(N^3)$, though the cost might possibly be reduced if estimating the global resolution parameter and not its time–frequency varying version $f_0(\omega,t)$, as originally. Some "monocomponent" methods can also be generalized to the case of multicomponent signals by introducing frequency dependence into the resolution parameter [96], but this will at the same time increase the computational complexity and give rise to additional issues.

Finally, there are various methods [87–89,96–103] for optimizing types of TFR other than the (S)WFT/(S)WT considered here. Most of them represent a modification of one of the approaches already mentioned, but tailored for a particular TFR, e.g. the Wigner–Ville distribution.

Nevertheless, to the best of the authors' knowledge, there are at present no universal methods for selecting the optimal window/wavelet parameters. Thus, the existing approaches can work very well for a particular class of signals (usually the one discussed in Section 9.2), but will fail for other classes. Generally, given the duality of the representation of an AM/FM component as a single entity and as a sum of tones, it is questionable whether any universal approach can in principle be developed. However, what suggests that this may in fact be feasible is that an adaptive signal decomposition method known as basis pursuit [63] seems to have all the desired properties, albeit at a computational cost of $O(N^3)$. Therefore, it might be possible that, for example, the functional approach can be made more universal by choosing instead of (9.2) a better performance measure, which will distinguish between independent components and AM/FM-induced ones by implicitly taking into account the specific relationships between their amplitudes, phases and frequencies. This remains an open problem.

10. Selecting reconstruction method

In previous sections we have investigated the relative performance of the direct (5.1) and ridge (5.3) methods for reconstructing components from the signal's TFR. This section summarizes related results and considers the errors of each method in more detail, as well as devising the scheme for adaptively choosing between the direct and ridge estimates.

10.1. Errors of different estimates

As we saw in previous sections, synchrosqueezing does not usually change the accuracy of parameter estimates, so in what follows we concentrate on the WFT/WT-based reconstruction errors only. Suppose we are given a signal that consists of $M+1$ AM/FM components corrupted by some noise $\zeta(t)$,

$$s(t) = A(t)\cos\phi(t) + \sum_{n=1}^{M} \tilde{A}_n(t)\cos\tilde{\phi}_n(t) + \sigma\zeta(t), \tag{10.1}$$

and we want to estimate $A(t)$, $\phi(t)$ and $\nu(t)\equiv\phi_0'(t)$ from the signal's TFR (assume for simplicity that $A(t)\cos\phi(t)$ is the dominant component in the signal). As described in Section 5, we do this by first finding the component's ridge curve $\omega_p(t)$ and the associated TFS $[\omega_-(t),\omega_+(t)]$, and then applying the direct or ridge reconstruction method to obtain parameter estimates. The latter will be denoted as

$$P^{(d,r)}(t) \equiv \{A^{(d,r)}(t), \phi^{(d,r)}(t), \nu^{(d,r)}(t)\}, \tag{10.2}$$

where (d) and (r) stand for direct and ridge reconstruction.

We are now interested in the errors of the estimates (10.2). They can be represented as

$$\Delta P^{(d,r)}(t) \equiv P^{(d,r)}(t) - P^{(true)}(t) \equiv \Delta P_T^{(d,r)}(t) + \Delta P_I^{(d,r)}(t), \tag{10.3}$$

where $\Delta P_T^{(d,r)}(t)$ is the theoretical error (the one arising in the case when there is only one component and no noise), while $\Delta P_I^{(d,r)}(t)$ denotes interference-related error (the one attributable to noise and the interferences with other components).

If the TFR can represent the component of interest $A(t)\cos\phi(t)$ reliably (Regime IV), then direct reconstruction is by definition exact (see Section 5.2), while the theoretical ridge errors are given in [72], so one has

Regime IV: $\Delta P_T^{(d)}(t) = 0$

$$\Delta P^{(r)}(t) \approx \frac{Q^2(\nu(t))}{2}\left\{A''(t), \phi''(t), \phi'''(t) + 2\frac{A'(t)}{A(t)}\phi''(t)\right\},$$
$$Q^2(\omega) \equiv \left[\frac{\partial_\nu^2 \hat{h}_\nu(\omega)}{\hat{h}_\nu(\omega)}\right]_{\nu=\omega} = \begin{cases} -\frac{\hat{g}''(0)}{\hat{g}(0)} \text{ for the WFT,} \\ -\frac{\omega_\psi^2}{\omega^2}\frac{\hat{\psi}''(\omega_\psi)}{\hat{\psi}(\omega_\psi)} \text{ for the WT.} \end{cases} \tag{10.4}$$

In the above expression, the quality of approximation for $\Delta P^{(r)}(t)$ is $\{O(\delta_{N_T}^2), O(\delta_{N_T}^2), O(\delta_{N_T}^3)\}$, where the value of δ_{N_T} is determined by the strength of the amplitude/frequency modulation in relation to the window/wavelet parameters (see [72]). For example, in the higher order approximation the ridge estimates of amplitude also contain errors proportional to $\phi''(t)$, which can become dominant if δ_{N_T} is large and/or $A''(t)$ is small (this becomes especially clear for chirps).

Remark 10.1. Only the WT was considered in [72], but the expressions for the WFT-based ridge errors (10.4) can be inferred simply by analogy based on comparison of the linear/logarithmic frequency resolution of the WFT/WT; the correctness of (10.4) was also confirmed numerically.

All becomes more complicated if the TFR cannot represent the component reliably, i.e. its behavior in the case when there is only one component and no noise is of the I–III type. However, from the analysis performed in Sections 8.2 and 8.3 one can see that the errors of the direct estimates are still smaller than those of the ridge ones, so that in all cases

$$|\Delta P_T^{(d)}(t)| \lesssim |\Delta P_T^{(r)}(t)| \tag{10.5}$$

The interference-related errors $\Delta P_I^{(d,r)}$ are also hard to estimate, though some analytic expressions can be derived under certain assumptions. These expressions are not very informative, however, so we do not present them here. Nevertheless, the results of

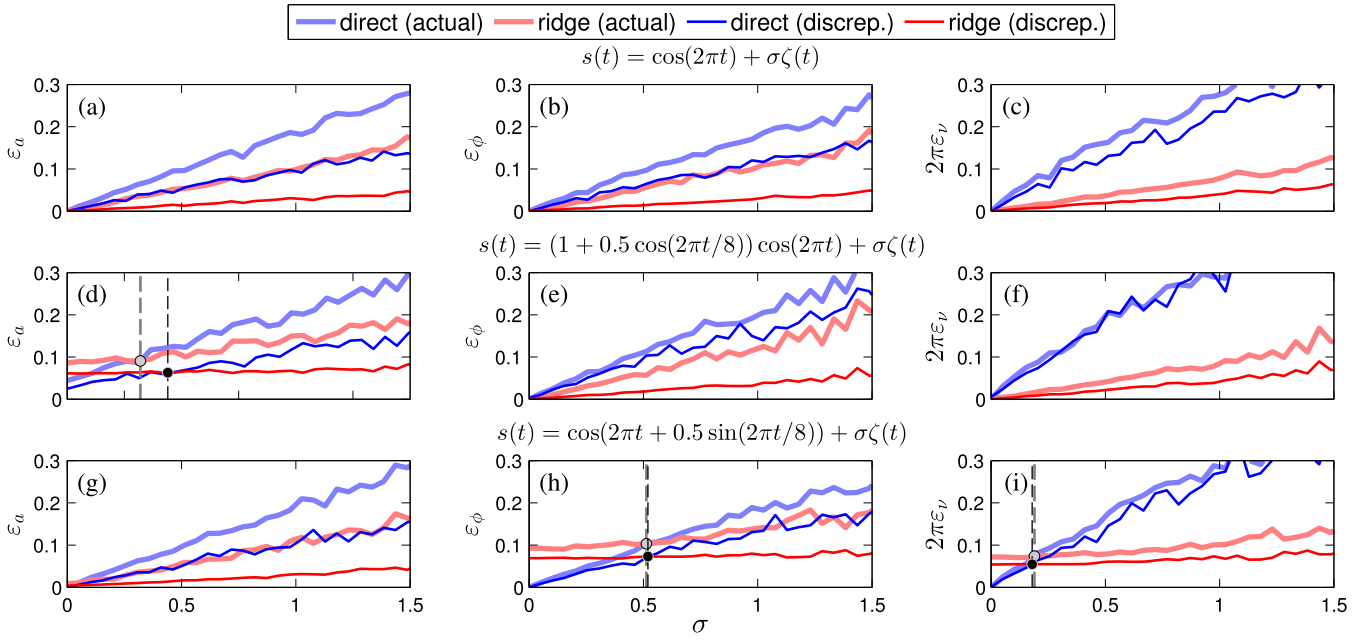

Fig. 16. The actual reconstruction errors (7.2) of the direct and ridge methods (light-blue and light-red lines, respectively) and the corresponding discrepancies (10.9) (direct – blue, ridge – red) in their dependence on the noise level σ. (a, d, g): Amplitude reconstruction errors. (b, e, h): Phase reconstruction errors. (c, f, i): Frequency reconstruction errors. The signals associated with each row are given by the equations above the central panels (b, e, h), with $\zeta(t)$ denoting unit-deviation Gaussian white noise; each signal was sampled at 50 Hz for 200 s. Where present, the gray (or black) points with the corresponding dashed lines indicate the intersections between the true errors (7.2) (or the discrepancies (10.9)) of the direct and ridge methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Sections 8.1 and 8.4 indicate that ridge estimates are more robust to noise and interference than direct ones, so that

$$|\Delta P_I^{(d)}(t)| \gtrsim |\Delta P_I^{(r)}(t)| \tag{10.6}$$

Summarizing, direct reconstruction allows more accurate estimation of amplitude/frequency variations but is less robust to noise and interferences between the components in comparison to ridge reconstruction. Therefore, the choice of the method depends on the signal under consideration: when the noise is small and different components are well-separated in the TFR, then the direct method should be used; if the noise and/or interference is strong, or the component's amplitude/frequency modulation is weak, the ridge method is the better choice. For example, the ridge method is always preferable for estimating parameters of the tone, as the corresponding theoretical errors (10.4) become zero in this case. Note also that the (S)WFT with symmetric $\hat{g}(\xi)$ provides additional advantages in terms of parameter estimation (see Fig. 9 and related discussion). However, since the reconstruction errors very much depend on the type of TFR behavior, as shown in previous sections, the window/wavelet resolution properties are what matters the most.

Remark 10.2. Using ridge error estimates (10.4), one can in fact devise a criterion for good representation of a given component in the TFR by requiring these errors to be small

$$\sqrt{\frac{\langle[\Delta A^{(r)}(t)]^2\rangle}{\langle A^2(t)\rangle}} \le \epsilon_p, \quad \sqrt{\langle[\Delta\phi^{(r)}(t)/\pi]^2\rangle} \le \epsilon_p, \tag{10.7}$$

where ϵ_p is some chosen precision (for determination of Regime IV, $\epsilon_p \in [0.05, 0.1]$ works quite well). For a Gaussian window WFT and lognormal wavelet WT this gives

$$f_0^2 \le \frac{2\epsilon_p}{\max\left[\sqrt{\langle[A''(t)]^2\rangle/\langle A^2(t)\rangle}, \sqrt{\langle[\phi''(t)/\pi]^2\rangle}\right]} \quad \text{(WFT)},$$

$$f_0^2 \le \frac{2\epsilon_p/4\pi^2}{\max\left[\sqrt{\langle[A''(t)/\nu^2(t)]^2\rangle/\langle A^2(t)\rangle}, \sqrt{\langle[\phi''(t)/(\pi\nu^2(t))]^2\rangle}\right]} \quad \text{(WT)}. \tag{10.8}$$

10.2. Adaptive choice of the method

To choose the best method automatically, one can devise an empirical criterion as follows. Suppose we have found direct and ridge parameter estimates $A^{(d,r)}(t)$, $\phi^{(d,r)}(t)$ and $\nu^{(d,r)}(t)$. To understand which of them are more accurate, we calculate the TFR (using the same window/wavelet as originally) of the signal $s^{(d)}(t) = A^{(d)}(t)\cos\phi^{(d)}(t)$, extract the ridge curve and TFS from it (taking simple maxima $\omega_p(t) = \operatorname{argmax}_\omega |H_s(\omega, t)|$ is sufficient here), and reconstruct by the direct method the "refined" parameters $\tilde{A}^{(d)}(t), \tilde{\phi}^{(d)}(t), \tilde{\nu}^{(d)}(t)$. The same procedure is performed for the "ridge" signal $s^{(r)}(t) = A^{(r)}(t)\cos\phi^{(r)}(t)$, now using the ridge method to reconstruct the refined estimates.

Obviously, if e.g. the direct estimates are accurate, one should have $\{\tilde{A}^{(d)}(t), \tilde{\phi}^{(d)}(t), \tilde{\nu}^{(d)}(t)\} \approx \{A^{(d)}(t), \phi^{(d)}(t), \nu^{(d)}(t)\}$. Therefore, one can assess which method is better on the basis of the *discrepancies* between the original and refined estimates, which can be quantified as

$$\tilde{\varepsilon}_a^{(d,r)} \equiv \kappa_a^{(d,r)} \frac{\sqrt{\langle[\tilde{A}^{(d,r)}(t) - A^{(d,r)}(t)]^2\rangle}}{\langle[A^{(d,r)}(t)]^2\rangle},$$

$$\tilde{\varepsilon}_\phi^{(d,r)} \equiv \kappa_\phi^{(d,r)} \sqrt{1 - |\langle e^{i[\tilde{\phi}^{(d,r)}(t) - \tilde{\phi}^{(d,r)}(t)]}\rangle|^2},$$

$$\tilde{\varepsilon}_\nu^{(d,r)} \equiv \kappa_\nu^{(d,r)} \frac{\sqrt{\langle[\tilde{\nu}^{(d,r)}(t) - \nu^{(d,r)}(t)]^2\rangle}}{2\pi}, \tag{10.9}$$

where $\kappa_{a,\phi,\nu}^{(d,r)}$ are the coefficients that can be used to tune the performance of the approach (they were found empirically to be $\kappa_{a,\phi,\nu}^{(d)} = \{3, 4, 2\}$, $\kappa_{a,\phi,\nu}^{(r)} = 1$). For each parameter, the choice between its direct and ridge estimate is then made based on the

corresponding discrepancy (10.9): the smaller it is, the more accurate the reconstructed parameter is expected to be.

Despite being empirical, the approach outlined above works very well in practice, selecting the best estimates in the majority of cases. This is illustrated in Fig. 16, where the discrepancies (10.9) are shown together with the actual reconstruction errors (7.2) for each method. As can be seen, the values of $\tilde{\varepsilon}^{(d,r)}_{a,\phi,\nu}$ are proportional to the true errors and allow one to judge reliably about the relative performance of the two reconstruction methods. Thus, as discussed previously, for a single tone signal embedded in noise the ridge estimates are always preferred, and the criterion based on (10.9) correctly reflects this fact (see Fig. 16(a–c)). Next, when amplitude/frequency modulation is present, at low noise levels the direct estimates are preferred, but with increasing noise strength their inaccuracy grows faster than in the case of ridge reconstruction. Therefore, beyond some threshold noise level (indicated by gray vertical dashed lines in Fig. 16) ridge estimates become the more accurate; this threshold and the optimal method in each case can be well recovered from the behavior of the discrepancies (10.9), as is clear from Fig. 16(g, h, i).

11. Do we really need synchrosqueezing?

As we have seen in Sections 8.1–8.4, synchrosqueezing does not change the qualitative behavior of the TFR and does not allow better resolution of components or more accurate estimation of their amplitude/frequency variations (for both direct and ridge methods). Therefore, it does not increase the time or frequency resolution of the transform, as might have seemed the case at first glance: all interferences and other complications are just summed into a more compact supports. In other words, synchrosqueezing improves only the "readability" of the TFR [58], providing a more visually appealing picture.

While not providing significant advantages, the SWFT/SWT has a few drawbacks in comparison to the WFT/WT, namely:

- The behavior of the SWFT/SWT is more complicated than that of the WFT/WT, being harder to study analytically and practically (e.g. see Fig. 2 and related discussion). In some cases this may lead to drastically reduced quality of the resultant parameter estimates (Fig. 12).
- The estimation of amplitude by the ridge method is not possible for the SWFT/SWT (in contrast to the WFT/WT), because the SWFT/SWT amplitude depends on the discretization of the frequency scale (see Section 5).
- The SWFT/SWT-based direct and ridge estimates of instantaneous frequency suffer greatly from the discretization effects (in contrast to the WFT/WT-based estimates, see Section 6.2), thus requiring very small frequency bins to make discretization errors negligible, which increases the computational cost of such estimation.

Hence, the usefulness of synchrosqueezing is questionable, because in terms of the components' reconstruction it only introduces additional complications.

Remark 11.1. For windows/wavelets with multimodal $|\hat{g}(\xi)|$ and $|\hat{\psi}(\xi>0)|$ (which case is outside of the assumptions made at the beginning of Part 2), synchrosqueezing has the advantageous property of joining together the component's power contained in all sidelobes into one TFS in the SWFT/SWT, hence making the latter more interpretable than the underlying WFT/WT. Nevertheless, in practice it appears that such a property is greatly affected even by small interference between components (or by considerable amplitude/frequency modulation). However, the advantages of synchrosqueezing for multimodal $|\hat{g}(\xi)|$ and $|\hat{\psi}(\xi>0)|$ is a separate topic, which is not of much relevance since such windows/wavelets are rarely used.

The fact that synchrosqueezing increases the concentration of the TFR, but at the same time does not improve its resolution properties, leads to reconsideration of a more general question: does concentration of the TFR alone represent the main measure of its performance, as is often believed? Our results argue against such a view, indicating that the most important characteristics of the TFR are its resolution properties.

To understand the difference between concentration and resolution, consider the "ideal" TFR $I(\omega,t)\sim\sum_k A_k(t)\delta(\omega-\phi'_k(t))$ of the signal $s(t)=\sum_k A_k(t)e^{i\phi_k(t)}\cos\phi_k(t)$. Then the inverse of the (somehow defined) "distance" between this perfect representation $I(\omega,t)$ and the calculated TFR can be considered as a measure of its performance. What one aims to achieve, therefore, is not just to increase the TFR concentration, but to increase it *around the instantaneous frequencies* $\phi'_k(t)$ and/or to improve the representation of the amplitude variations. For example, a TFR having peaks at $\omega=\phi'_k(t)$, but not being too concentrated, is obviously to be preferred to an extremely concentrated TFR with peaks distant from the true instantaneous frequencies. In other words, the main goal is to represent appropriately all the components present in the signal, so that their parameters can accurately be recovered.

Considering synchrosqueezing, from the previous sections it is clear that the positions of the SWFT/SWT ridges are no closer to the actual frequencies than the WFT/WT ridges. Thus, "curves" in synchrosqueezed TFRs, although being more concentrated, are not located around the actual instantaneous frequencies of the components (though the latter can be fully recovered from the full TFS). This is illustrated in Fig. 17, where the SWFT is compared with the ridge/direct WFT skeletons. For each time, the latter are constructed by partitioning the WFT into regions of unimodal amplitude (time–frequency supports $[\omega^{(m)}_-(t),\omega^{(m)}_+(t)]$), reconstructing from them the amplitudes $A^{(m)}(t)$, phases $\phi^{(m)}(t)$ and frequencies $\nu^{(m)}(t)$ using the chosen method, and then assigning $A^{(m)}(t)e^{i\phi^{(m)}(t)}$ to the frequency bin where the estimated frequency $\nu^{(m)}(t)$ lies; the WT skeletons can be constructed in the same way. For example, up to frequency discretization effects, the ridge-based WFT/WT skeleton is simply the WFT/WT with only peaks left (and multiplied by $2/\hat{g}(0)$), while other coefficients are set to zero. The MatLab codes for calculating TFR skeletons can be downloaded from [32] together with the other codes used in this work.

Pairwise comparison of Fig. 17(b, f) and (c, g) shows that the SWFT is similar to a simple ridge-based WFT skeleton (which is additionally more concentrated and easier to interpret): in both former and latter cases, the "curves" are not located around the true component frequency, but have similar deviations from it. On the other hand, the direct WFT skeleton in the noiseless case provides an almost perfect representation (Fig. 17(d)), being clearly superior to the SWFT or ridge skeleton, though the picture becomes more complicated when noise is present (Fig. 17(e–h)). However, both skeletons are constructed from the original WFT and obviously have the same resolution properties (as the accuracy of the parameters' estimates remains does not change), providing advantages mainly in terms of visual appearance, similarly to the case of the SWFT.

12. Summary and conclusions

In Part 1 of this work we provided a thorough review of the linear and synchrosqueezed TFRs and their properties. Additionally, we introduced some new concepts and procedures to advance the theory and applicability of these methods, as well as revisiting certain related issues. For example, in Section 4 we argued that the

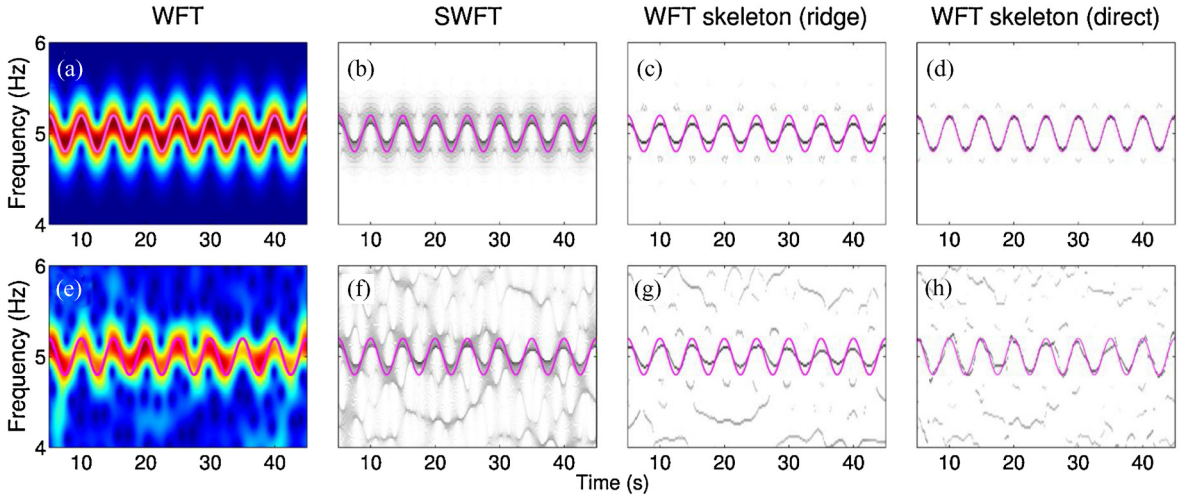

Fig. 17. Comparison of the WFT, SWFT and WFT skeletons based on ridge and direct reconstruction for: (a–d) the FM component $s(t) = \cos(10\pi t + \sin(2\pi t/5))$; (e–h) the same component additionally corrupted by noise $\sqrt{2}\zeta(t)$. Magenta lines show the true frequency of the component. The small intermittent components appearing at both sides of the main frequency in (c, d) are due to III type of TFR behavior. The signal was sampled at 100 Hz for 50 s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

classical resolution measures (4.5), (4.6) are not really appropriate for characterizing the resolution properties of a TFR (especially for the (S)WT), and new, more physically meaningful measures (4.13), (4.14) have been designed.

Practical aspects of using the (S)WFT and (S)WT were considered in Section 6 (technical details being provided in Supplementary Section 1). This included devising criteria (6.1) for selecting the optimal frequency bin width (Section 6.2); derivation of expressions for boundary errors (6.2) and related quantities; and introduction of the predictive padding scheme (Supplementary Sections 1.3 and 6). Finally, we discussed ways of reconstructing the components' parameters from different TFRs and derived the corresponding formulas (Section 5).

Part 2 presented a detailed numerical and theoretical study of the specific issues discussed in Section 7. The results can be summarized as:

- Parameters of the window/wavelet, e.g. the resolution parameter f_0, determine the resolution properties of the TFR, so their appropriate choice is of crucial importance. Thus, different choices lead to different TFR behaviors, and therefore different quantitative and qualitative results: if the frequency resolution is too high, the AM/FM component might be represented as a number of independent tones, whereas if it is too low, the two interfering tones can be merged together and appear as a single component in the TFR. We considered and illustrated these issues on numerous examples in Sections 8.1–8.4.
- The optimal f_0 depends on the signal. For a particular class of signals it can be selected using some of the approaches reviewed in [9] and Section 9 here (e.g. based on $F_{1,2}$ (9.2)), but there does not seem to be any universal approach suited to the general case. The question of how best to select the appropriate window/wavelet parameters for a given signal therefore remains open.
- In the absence of an adaptation scheme, one can choose f_0 based on the desired resolution properties of the TFR using conditions (8.4) and (10.7) for resolving two tones and for representing the AM/FM component reliably. Note that, in respect of different kinds of signals (with different characteristic frequency bands), the choice of f_0 for the WT seems to be slightly more universal than for the WFT, with the commonest value being $f_0 = 1$.
- The relative performance of the direct (5.1) and ridge (5.3) reconstruction methods depends on the signal. The direct method better estimates amplitude/frequency variations but is less robust to noise and interference than the ridge method. An automatic procedure for selecting optimal reconstruction method was suggested in Section 10.2. It remains possible, however, that some universal estimate combining the advantages of both direct and ridge methods could be designed, which would be highly desirable.
- Synchrosqueezing does not provide significant advantages in terms of components' reconstruction, but it introduces additional complications. Thus, the concentration and resolution are not always the same: although being more concentrated, the SWFT/SWT actually has the same time and frequency resolutions as the WFT/WT from which it is constructed, so it does not enable more accurate estimation of time-variations in the components' parameters, and nor does it enable the resolution of components that lie closer in frequency. On the other hand, if one is mainly interested in concentration and/or a visually appealing picture, then the most natural choice would be the TFR skeletons (see Section 11).

Appendix A. Nomenclature

Abbreviations

ECG	Electrocardiogram
FT	Fourier Transform
FFT	Fast Fourier Transform (algorithm)
IFFT	Inverse Fast Fourier Transform (algorithm)
WFT	Windowed Fourier Transform (see Section 3.1)
WT	Wavelet Transform (see Section 3.2)
SWFT	Synchrosqueezed WFT (see Section 3.3)
SWT	Synchrosqueezed WT (see Section 3.3)
TFR	Time–Frequency Representation (there are many different types, but in this work we refer to the (S)WFT and (S)WT only)
TFS	Time–Frequency Support (see Section 5)

Terminology

AM/FM component, or simply component	Sinusoidal oscillation $A(t)\cos\phi(t)$ with amplitude and/or frequency modulation (see Section 2 for a detailed discussion).

Tone	AM/FM component with constant amplitude and frequency, i.e. a simple sine $x(t) = A\cos(\nu t + \varphi)$.
Gaussian window	Window function of the form (3.3).
Morlet wavelet	Wavelet function of the form (3.8).
Lognormal wavelet	Wavelet function of the form (3.9).
Time, frequency and joint time–frequency resolution of the TFR	The measures of how fast changes in time and how close differences in frequencies can be resolved in the TFR, and the tradeoff between the two, respectively; their precise definitions are discussed in Section 4.

Main notation

$\hat{f}(\xi)$, $f^{\pm}(t)$	Fourier Transform and positive/negative frequency part of $f(t)$, see (2.2).
$\langle f(t)\rangle$, std$[f(t)]$	Mean and standard deviation of $f(t)$, see (2.2).
$f^*(t)$	Complex conjugate of $f(t)$.
c.c.	Complex conjugate of preceding expression.
Re$[x]$, Im$[x]$	Real and imaginary parts of x.
$s^a(t) = 2s^+(t)$	Analytic signal (2.3), see Section 2.
sign(x)	Sign function: sign$(x>0)=1$, sign$(x<0)=-1$, sign$(0)=0$.
$J_n(x)$	nth order Bessel function of the first kind.
$I_n(x)$	nth order modified Bessel function of the first kind.
$\Gamma(a)$	Gamma function $\Gamma(a) \equiv \int_0^\infty x^{a-1}e^{-x}dx$.
erf(x)	Gauss error function $\mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}du$.
$n_G(\epsilon)$	Number of standard deviations of the Gaussian distribution within which its $(1-\epsilon)$ part is contained: $\frac{1}{\sqrt{2\pi}}\int_{-n_G}^{n_G} e^{-u^2/2}du = \mathrm{erf}(n_G/\sqrt{2}) = 1-\epsilon$. For example, $n_G(0.05)\approx 2$, $n_G(0.01)\approx 2.5$.
$G_s(\omega,t)$, $W_s(\omega,t)$	WFT (3.1) and WT (3.4) of a given signal.
$g(t)$, $\hat{g}(\xi)$	Time domain and frequency domain forms of the window function used for computation of the WFT (3.1), respectively.
$\psi(t)$, $\hat{\psi}(\xi)$	Time domain and frequency domain forms of the wavelet function used for computation of the WT (3.4), respectively.
ω_ψ	Wavelet peak frequency (3.5).
C_g, C_ψ	Integration constants defined in (3.2), (3.6).
$\overline{\omega}_g$, D_ψ	Integration constants defined in (5.1).
f_0	Window/Wavelet resolution parameter which determines its resolutions properties (see Supplementary Section 7 for a list of different window and wavelet functions).
$R_g(\omega)$, $P_g(\tau)$	Quantitative measures of the area below $\hat{g}(\xi<\omega)$ and $g(t<\tau)$, as defined in (4.7).
$R_\psi(\omega)$, $P_\psi(\tau)$	Quantitative measures of the area below $\hat{\psi}(0<\xi<\omega)$ and $\psi(t<\tau)$, as defined in (4.8).
$\xi_{1,2}(\epsilon)$, $\tau_{1,2}(\epsilon)$	ϵ-supports of the window/wavelet function in time and frequency, defined in (4.7) for windows and in (4.8) for wavelets.
$V_s(\omega,t)$, $T_s(\omega,t)$	SWFT and SWT (3.11) of a given signal.
$\widetilde{V}_s(\omega,t)$, $\widetilde{T}_s(\omega,t)$	Numerical SWFT and SWT (3.13).
$[\omega_{\min},\omega_{\max}]$	Frequency range for which the currently considered TFR is calculated.
$\omega_p(t)$	Ridge curve of some component in the currently considered TFR (see Section 5).
$[\omega_-(t),\omega_+(t)]$	Time–frequency support of some component in the currently considered TFR (see Section 5).
$H_s(\omega,t)$	Either WFT $G_s(\omega,t)$ or WT $W_s(\omega,t)$, see (6.6).
$h_t(\omega)$, $\hat{h}_\xi(\omega)$	The functions with which the signal is convoluted respectively in time and in frequency while constructing its WFT/WT, see (6.6).
C_h	Either C_g (WFT) or C_ψ (WT), see (6.9).
$\mu(\omega)$	The integration measure for the WFT ($\mu(\omega)=\omega$) or WT ($\mu(\omega)=\log\omega$), see (6.9).

Conventions

- Where undefined, integrals are taken over $(-\infty,\infty)$, or over the full range of the variable: for a signal $s(t)$, one has $\int s(t)dt \equiv \int_{-\infty}^{\infty} s(t)dt$ theoretically and $\int s(t)dt \equiv \int_0^T s(t)dt$ practically, with T denoting the overall time duration of the signal.
- All TFRs are computed for a real signal $s(t)$ assuming, without loss of generality that $\omega_g \equiv \mathrm{argmax}_\xi|\hat{g}(\xi)| = 0$.
- Unless otherwise specified, the (S)WFT and (S)WT are calculated using the Gaussian window (3.3) and Morlet wavelet (3.8), with resolution $f_0 = 1$. The frequency axis is discretized as $\omega_k = (k-k_0)\Delta\omega$ for the (S)WFT, and as $\omega_k/2\pi = 2^{(k-k_0)/n_v}$ for the (S)WT, with the discretization parameters $\Delta\omega$ and n_v selected according to (6.1). To reduce boundary effects (see Section 6.3), predictive padding is used by default; in most cases, however, only the "distortion-free" TFR parts (i.e. lying within the cones-of-influence (6.3) with $\epsilon = 0.001$) are shown. In plots of time-averaged TFR amplitudes or phase velocities, the yellow regions indicate $\pm\sqrt{2}$ of the corresponding standard deviations.
- For simplicity, circular frequencies (rad/s) are mainly used, denoted by ω, ν, ξ; to convert them to Hz, divide by 2π. Note, however, that the sampling frequency f_s is always in Hz.
- The additional assumptions and conventions used in Part 2 are summarized just before Section 7.

Appendix B. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.dsp.2015.03.004.

References

[1] N. Delprat, B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, B. Torrésani, Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies, IEEE Trans. Inf. Theory 38 (2) (1992) 644–664.

[2] R.A. Carmona, W.L. Hwang, B. Torresani, Characterization of signals by the ridges of their wavelet transforms, IEEE Trans. Signal Process. 45 (10) (1997) 2586–2590.

[3] R.A. Carmona, W.L. Hwang, B. Torresani, Multiridge detection and time–frequency reconstruction, IEEE Trans. Signal Process. 47 (2) (1999) 480–492.

[4] D. Iatsenko, A. Bernjak, T. Stankovski, Y. Shiogai, P.J. Owen-Lynch, P.B.M. Clarkson, P.V.E. McClintock, A. Stefanovska, Evolution of cardio-respiratory interactions with age, Philos. Trans. R. Soc. Lond. A 371 (1997) (2013) 20110622.

[5] G. Thakur, E. Brevdo, N.S. Fuckar, H.-T. Wu, The Synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications, Signal Process. 93 (5) (2013) 1079–1094.

[6] I. Daubechies, J. Lu, H. Wu, Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool, Appl. Comput. Harmon. Anal. 30 (2) (2011) 243–261.

[7] S. Mallat, A Wavelet Tour of Signal Processing, 3rd edn., Academic Press, Burlington, 2008.

[8] P.S. Addison, The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, IOP Publishing, Bristol, 2002.

[9] E. Sejdić, I. Djurović, J. Jiang, Time–frequency feature representation using energy concentration: an overview of recent advances, Digit. Signal Process. 19 (1) (2009) 153–183.

[10] Y. Shiogai, A. Stefanovska, P.V.E. McClintock, Nonlinear dynamics of cardiovascular ageing, Phys. Rep. 488 (2) (2010) 51–110.

[11] M. Wacker, H. Witte, Time–frequency techniques in biomedical signal analysis. A tutorial review of similarities and differences, Methods Inf. Med. 52 (4) (2013) 279–296.
[12] A. Stefanovska, M. Bračič, H.D. Kvernmo, Wavelet analysis of oscillations in the peripheral blood circulation measured by laser Doppler technique, IEEE Trans. Biomed. Eng. 46 (10) (1999) 1230–1239.
[13] C. Torrence, G.P. Compo, A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc. 79 (1) (1998) 61–78.
[14] P. Kumar, E. Foufoula-Georgiou, Wavelet analysis for geophysical applications, Rev. Geophys. 35 (4) (1997) 385–412.
[15] A. Stefanovska, Coupled oscillators: complex but not complicated cardiovascular and brain interactions, IEEE Eng. Med. Biol. Mag. 26 (6) (2007) 25–29.
[16] D. Labat, Recent advances in wavelet analyses: part 1. A review of concepts, J. Hydrol. 314 (1) (2005) 275–288.
[17] D. Labat, J. Ronchail, J.L. Guyot, Recent advances in wavelet analyses: part 2. Amazon, Parana, Orinoco and Congo discharges time scale variability, J. Hydrol. 314 (1) (2005) 289–311.
[18] L. Aguiar-Conraria, P.C. Magalhães, M.J. Soares, Cycles in politics: wavelet analysis of political time series, Am. J. Polit. Sci. 56 (2) (2012) 500–518.
[19] L. Stankovic, M. Dakovic, T. Thayaparan, Time–Frequency Signal Analysis with Applications, Artech House, Norwood, MA, 2013.
[20] F. Hlawatsch, F. Auger, Time–Frequency Analysis, vol. 36, John Wiley & Sons, 2010.
[21] B. Boashash, Time Frequency Signal Analysis and Processing, Elsevier, 2003.
[22] L. Cohen, Time–Frequency Analysis, vol. 778, Prentice Hall PTR, New Jersey, 1995.
[23] I. Daubechies, Ten Lectures on Wavelets, vol. 61, SIAM, Philadelphia, 1992.
[24] K. Gröchenig, Foundations of Time–Frequency Analysis, Springer, 2001.
[25] G. Kaiser, A Friendly Guide to Wavelets, Birkhäuser, Boston, 1994.
[26] F. Hlawatsch, G.F. Boudreaux-Bartels, Linear and quadratic time–frequency signal representations, IEEE Signal Process. Mag. 9 (2) (1992) 21–67.
[27] L. Cohen, Time–frequency distributions – a review, Proc. IEEE 77 (7) (1989) 941–981.
[28] S. Qian, D. Chen, Joint time–frequency analysis, IEEE Signal Process. Mag. 16 (2) (1999) 52–67.
[29] G. Thakur, H.-T. Wu, Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples, SIAM J. Math. Anal. 43 (5) (2011) 2078–2095.
[30] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, Time–frequency reassignment and synchrosqueezing: an overview, IEEE Signal Process. Mag. 30 (6) (2013) 32–41.
[31] I. Daubechies, S. Maes, A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models, in: Wavelets in Medicine and Biology, 1996, pp. 527–546.
[32] The MatLab codes for computing WFT, WT, SWFT and SWT, as well as for direct and ridge reconstruction of components' parameters from them, are freely available at http://www.physics.lancs.ac.uk/research/nbmphysics/diats/tfr/.
[33] A. Grinsted, J.C. Moore, S. Jevrejeva, Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonlinear Process. Geophys. 11 (5/6) (2004) 561–566.
[34] J.P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le van Quyen, J. Martinerie, F. Varela, Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence, Clin. Neurophysiol. 32 (3) (2002) 157–174.
[35] A. Bandrivskyy, A. Bernjak, P.V.E. McClintock, A. Stefanovska, Wavelet phase coherence analysis: application to skin temperature and blood flow, Cardiovasc. Eng. 4 (1) (2004) 89–93.
[36] D.S. Bloomfield, R.T.J. McAteer, B.W. Lites, P.G. Judge, M. Mathioudakis, F.P. Keenan, Wavelet phase coherence analysis: application to a quiet-sun magnetic element, Astrophys. J. 617 (1) (2004) 623.
[37] L.W. Sheppard, A. Stefanovska, P.V.E. McClintock, Testing for time-localised coherence in bivariate data, Phys. Rev. E 85 (2012) 046205.
[38] L.W. Sheppard, A. Stefanovska, P.V.E. McClintock, Detecting the harmonics of oscillations with time-variable frequencies, Phys. Rev. E 83 (1) (2011) 016206.
[39] J. Jamšek, A. Stefanovska, P.V.E. McClintock, Wavelet bispectral analysis for the study of interactions among oscillators whose basic frequencies are significantly time variable, Phys. Rev. E 76 (4) (2007) 046221.
[40] J. Jamšek, M. Paluš, A. Stefanovska, Detecting couplings between interacting oscillators with time-varying basic frequencies: instantaneous wavelet bispectrum and information theoretic approach, Phys. Rev. E 81 (3) (2010) 036207.
[41] B. Picinbono, On instantaneous amplitude and phase of signals, IEEE Trans. Signal Process. 45 (3) (1997) 552–560.
[42] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals, Proc. IEEE 80 (4) (1992) 520–538.
[43] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications, Proc. IEEE 80 (4) (1992) 540–568.
[44] E. Bedrosian, A product theorem for Hilbert transforms, Proc. IEEE 51 (5) (1963) 868–869.
[45] D. Vakman, On the analytic signal, the Teager–Kaiser energy algorithm, and other methods for defining amplitude and frequency, IEEE Trans. Signal Process. 44 (4) (1996) 791–797.
[46] D. Gabor, Theory of communication, J. IEEE 93 (1946) 429–457.
[47] J. Morlet, Sampling theory and wave propagation, in: C.H. Chen (Ed.), Issues on Acoustic Signal/Image Processing and Recognition, vol. I, in: NATO ASI Series, vol. 1, Springer, Berlin, 1983, pp. 233–261.
[48] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (5) (1932) 749.
[49] J.d. Ville, et al., Théorie et applications de la notion de signal analytique, Câables et Transmission 2 (1) (1948) 61–74.
[50] A.W. Rihaczek, Signal energy distribution in time and frequency, IEEE Trans. Inf. Theory 14 (3) (1968) 369–374.
[51] H.-I. Choi, W.J. Williams, Improved time–frequency representation of multicomponent signals using exponential kernels, IEEE Trans. Acoust. Speech Signal Process. 37 (6) (1989) 862–871.
[52] R. Tao, Y.-L. Li, Y. Wang, Short-time fractional Fourier transform and its applications, IEEE Trans. Signal Process. 58 (5) (2010) 2568–2580.
[53] L.B. Almeida, The fractional Fourier transform and time–frequency representations, IEEE Trans. Signal Process. 42 (11) (1994) 3084–3091.
[54] C. Capus, K. Brown, Short-time fractional Fourier methods for the time–frequency representation of chirp signals, J. Acoust. Soc. Am. 113 (6) (2003) 3253–3263.
[55] V. Katkovnik, A new form of the Fourier transform for time-varying frequency estimation, Signal Process. 47 (2) (1995) 187–200.
[56] S. Mann, S. Haykin, The chirplet transform: physical considerations, IEEE Trans. Signal Process. 43 (11) (1995) 2745–2761.
[57] S. Mann, S. Haykin, The chirplet transform: a generalization of Gabor's logon transform, in: Vision Interface, vol. 91, 1991, pp. 205–212.
[58] F. Auger, P. Flandrin, Improving the readability of time–frequency and time-scale representations by the reassignment method, IEEE Trans. Signal Process. 43 (5) (1995) 1068–1089.
[59] I. Daubechies, Time–frequency localization operators: a geometric phase space approach, IEEE Trans. Inf. Theory 34 (4) (1988) 605–612.
[60] S.G. Mallat, Z. Zhang, Matching pursuits with time–frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (1993) 3397–3415.
[61] Y.C. Pati, R. Rezaiifar, P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, in: Signals, Systems and Computers, IEEE, 1993, pp. 40–44.
[62] R.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Trans. Inf. Theory 38 (2) (1992) 713–718.
[63] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159.
[64] L. Stanković, M. Daković, T. Thayaparan, A real-time time–frequency based instantaneous frequency estimator, Signal Process. 93 (5) (2013) 1392–1397.
[65] L. Stanković, T. Alieva, M.J. Bastiaans, Time–frequency signal analysis based on the windowed fractional Fourier transform, Signal Process. 83 (11) (2003) 2459–2468.
[66] H.K. Kwok, D.L. Jones, Improved instantaneous frequency estimation using an adaptive short-time Fourier transform, IEEE Trans. Signal Process. 48 (10) (2000) 2964–2972.
[67] H. Kawahara, I. Masuda-Katsuse, A. de Cheveigné, Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based f0 extraction: possible role of a repetitive structure in sounds, Speech Commun. 27 (3) (1999) 187–207.
[68] L. Stanković, A method for time–frequency analysis, IEEE Trans. Signal Process. 42 (1) (1994) 225–229.
[69] B. Barkat, B. Boashash, A high-resolution quadratic time–frequency distribution for multicomponent signals analysis, IEEE Trans. Signal Process. 49 (10) (2001) 2232–2239.
[70] V. DeBrunner, J.P. Havlicek, T. Przebinda, M. Ozaydin, Entropy-based uncertainty measures for $L^2(R^n)$, $l^2(Z)$, and $l^2(Z/NZ)$ with a Hirschman optimal transform for $l^2(Z/NZ)$, IEEE Trans. Signal Process. 53 (8) (2005) 2690–2699.
[71] G. Liu, V. DeBrunner, High-resolution non-parametric spectral estimation using the Hirschman optimal transform, in: ICASSP 2012, IEEE, 2012, pp. 3721–3724.
[72] J.M. Lilly, S.C. Olhede, On the analytic wavelet transform, IEEE Trans. Inf. Theory 56 (8) (2010) 4135–4156.
[73] D. Iatsenko, P.V.E. McClintock, A. Stefanovska, On the extraction of instantaneous frequencies from ridges in time–frequency representations of signals, submitted for publication, Preprint arXiv:1310.7276.
[74] L.A. Montejo, A.L. Vidot-Vega, Synchrosqueezed wavelet transform for frequency and damping identification from noisy signals, Smart Struct. Syst. 9 (5) (2012) 441–459.
[75] G.R. Putland, B. Boashash, Can a signal be both monocomponent and multicomponent?, in: Third Australasian Workshop on Signal Processing Applications, WoSPA 2000, 2000, pp. 14–15.
[76] H.-T. Wu, P. Flandrin, I. Daubechies, One or two frequencies? The synchrosqueezing answers, Adv. Adapt. Data Anal. 3 (01n02) (2011) 29–39.
[77] D.L. Jones, R.G. Baraniuk, A simple scheme for adapting time–frequency representations, IEEE Trans. Signal Process. 42 (12) (1994) 3530–3535.

[78] M.K. Emresoy, A. El-Jaroudi, Iterative instantaneous frequency estimation and adaptive matched spectrogram, Signal Process. 64 (2) (1998) 157–165.
[79] J. Zhong, Y. Huang, Time–frequency representation based on an adaptive short-time Fourier transform, IEEE Trans. Signal Process. 58 (10) (2010) 5118–5128.
[80] T.K. Hon, A. Georgakis, Enhancing the resolution of the spectrogram based on a simple adaptation procedure, IEEE Trans. Signal Process. 60 (10) (2012) 5566–5571.
[81] S.-C. Pei, S.-G. Huang, STFT with adaptive window width based on the chirp rate, IEEE Trans. Signal Process. 60 (8) (2012) 4065–4080.
[82] D.L. Jones, T.W. Parks, A high resolution data-adaptive time–frequency representation, IEEE Trans. Acoust. Speech Signal Process. 38 (12) (1990) 2127–2135.
[83] G. Jones, B. Boashash, Generalized instantaneous parameters and window matching in the time–frequency plane, IEEE Trans. Signal Process. 45 (5) (1997) 1264–1275.
[84] V. Katkovnik, L. Stanković, Periodogram with varying and data-driven window length, Signal Process. 67 (3) (1998) 345–358.
[85] Q. Yin, L. Shen, M. Lu, X. Wang, Z. Liu, Selection of optimal window length using STFT for quantitative SNR analysis of LFM signal, J. Syst. Eng. Electron. 24 (1) (2013) 26–35.
[86] M.B. Lotric, A. Stefanovska, D. Stajer, V. Urbancic-Rovan, Spectral components of heart rate variability determined by wavelet analysis, Physiol. Meas. 21 (4) (2000) 441.
[87] R.G. Baraniuk, P. Flandrin, A.J.E.M. Janssen, O.J.J. Michel, Measuring time–frequency information content using the Rényi entropies, IEEE Trans. Inf. Theory 47 (4) (2001) 1391–1409.
[88] L. Stanković, A measure of some time–frequency distributions concentration, Signal Process. 81 (3) (2001) 621–631.
[89] D.L. Jones, R.G. Baraniuk, An adaptive optimal-kernel time–frequency representation, IEEE Trans. Signal Process. 43 (10) (1995) 2361–2371.
[90] V. Sucic, N. Saulig, B. Boashash, Estimating the number of components of a multicomponent nonstationary signal using the short-term time–frequency Rényi entropy, EURASIP J. Adv. Signal Process. 2011 (1) (2011) 1–11.
[91] N. Saulig, V. Sucic, Nonstationary signals information content estimation based on the local Rényi entropy in the time–frequency domain, in: Proceedings of the 11th International Conference on Telecommunications (ConTEL), IEEE, 2011, pp. 465–472.
[92] M.B. Malarvili, V. Sucic, M. Mesbah, B. Boashash, Renyi entropy of quadratic time–frequency distributions: effects of signal's parameters, in: 9th ISSPA, IEEE, 2007, pp. 1–4.
[93] B. Boashash, V. Sucic, Resolution measure criteria for the objective assessment of the performance of quadratic time–frequency distributions, IEEE Trans. Signal Process. 51 (5) (2003) 1253–1263.
[94] V. Sucic, B. Boashash, An approach for selecting a real-life signal best-performing time–frequency distribution, in: Proceedings of 7th International Symposium on Signal Processing and Its Applications, vol. 1, IEEE, 2003, pp. 125–128.
[95] V. Sucic, B. Boashash, K. Abed-Meraim, A normalised performance measure for quadratic time–frequency distributions, in: Proc. 2nd IEEE International Symposium on Signal Processing and Information Technology, ISSPIT'02, IEEE, 2002, pp. 463–466.
[96] L. Stanković, V. Katkovnik, The Wigner distribution of noisy signals with adaptive time–frequency varying window, IEEE Trans. Signal Process. 47 (4) (1999) 1099–1108.
[97] L. Stanković, V. Katkovnik, Instantaneous frequency estimation using higher order L-Wigner distributions with data-driven order and window length, IEEE Trans. Inf. Theory 46 (1) (2000) 302–311.
[98] R.G. Baraniuk, D.L. Jones, Signal-dependent time–frequency analysis using a radially Gaussian kernel, Signal Process. 32 (3) (1993) 263–284.
[99] R.G. Baraniuk, D.L. Jones, A signal-dependent time–frequency representation: optimal kernel design, IEEE Trans. Signal Process. 41 (4) (1993) 1589–1602.
[100] R.G. Baraniuk, D.L. Jones, A signal-dependent time–frequency representation: fast algorithm for optimal kernel design, IEEE Trans. Signal Process. 42 (1) (1994) 134–146.
[101] T.-H. Sang, W.J. Williams, Renyi information and signal-dependent optimal kernel design, in: Acoustics, Speech, and Signal Processing, vol. 2, ICASSP-95, IEEE, 1995, pp. 997–1000.
[102] L.J. Stankovic, V. Katkovnik, Algorithm for the instantaneous frequency estimation using time–frequency distributions with adaptive window width, IEEE Signal Process. Lett. 5 (9) (1998) 224–227.
[103] V. Katkovnik, L. Stanković, Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length, IEEE Trans. Signal Process. 46 (9) (1998) 2315–2325.

Dmytro Iatsenko is a Quantitative Analyst at Deutsche Bank. Following his BSc (2005–2009) and MSc (2009–2011) in Physics at Taras Shevchenko National University of Kyiv, he did his PhD (2011–2014) at Lancaster University. While doing his PhD research, he received the "Best International Student in the Physics Department" award (2012), the "Faculty of Science and Technology Dean's Award" (2012), and the "Best Student in the Physics Department" award (2014). His thesis, *Nonlinear Mode Decomposition: theory and applications*, has been selected for publication as a book by Springer. His research interests include quantitative finance, time–frequency analysis, and networks of coupled oscillators.

Peter V.E. McClintock is Emeritus Professor of Physics at Lancaster University. After reading for a BSc at Queen's University, Belfast (1958–1962) and research for a DPhil at Oxford University (1962–1966), followed by postdoctoral research at Duke University, North Carolina, he came to Lancaster in 1968. He was awarded a DSc by Queen's University in 1983, was a SERC/EPSRC Senior Fellow, 1990–1995, and served as Head of the Department of Physics 1997–2003. His research interests include low temperature physics, superfluidity, quantum turbulence, and nonlinear dynamics including its applications to biomedical problems. He is Editor-in-Chief of *Fluctuation and Noise Letters*.

Aneta Stefanovska is Professor of Biomedical Physics at Lancaster University. Following PhD (1992) research in the Universities of Ljubljana and Stuttgart, her work has centered on biological oscillations and their interactions in health and disease. New methods she has developed for studying their inherent time-variability include wavelet-based bispectral and phase coherence analysis, extraction of coupling direction and coupling functions using information theory and dynamical Bayesian inference, detection of high-harmonics, and nonlinear mode decomposition. She has recently proposed a theory of the chronotaxic systems that stabilize their phases under continuous perturbation. She is section editor-in-chief of *EPJ Nonlinear Biomedical Physics*.