Defining the Wavelet Bispectrum

Julian Newman

Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom

Department of Mathematics, University of Exeter, Exeter EX4 4QF, United Kingdom

Aleksandra Pidde

Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

Aneta Stefanovska*

Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom

Abstract

Bispectral analysis is an effective signal processing tool for analysing interactions between oscillations, and has been adapted to the continuous wavelet transform for time-evolving analysis of open systems. However, one unaddressed question for the wavelet bispectrum is quantification of the bispectral content of an area of scale-scale space. This makes the capacity for quantitative rather than merely qualitative interpretation of wavelet bispectrum computations very limited. In this paper, we overcome this limitation by providing suitable normalisations of the wavelet bispectrum formula that enable it to be treated as a density to be integrated. These are roughly analogous to the normalisation for second-order wavelet spectral densities. We prove that our definition of the wavelet bispectrum matches the traditional bispectrum of sums of sinusoids, in the limit as the frequency resolution tends to infinity. We illustrate the improved quantitative power of our definition with numerical and experimental data.

Keywords: continuous wavelet transform, wavelet bispectrum, bispectral analysis, time-frequency analysis, lognormal wavelets

1. Introduction

We begin by outlining the background to wavelet bispectral analysis and its applications, the limitations of current definitions, and how we overcome these limitations in terms of suitably defined notions of *wavelet bispectral density*.

Background to wavelet bispectral analysis

Power-spectral analysis of time-series data may be used to investigate oscillatory influences in a signal, and likewise cross-spectral analysis to investigate common oscillatory influences between

*Corresponding author

Email addresses: j.newman1@lancaster.ac.uk (Julian Newman), j.m.i.newman@exeter.ac.uk (Julian Newman), a.pidde@lancaster.ac.uk (Aleksandra Pidde), aneta@lancaster.ac.uk (Aneta Stefanovska)

two simultaneous signals. Power spectra and cross-spectra are referred to as *second-order spectra*. In a similar manner, *bispectral* (i.e. third-order spectral) analysis of time-series data may be used to investigate influences – either in one signal or common among two or three simultaneous signals – coming from a nonlinearly combined effect of two oscillators at different[1] frequencies. In particular, this can include influence from a pair of *interacting* oscillators (illustrated in Fig. 1), although it can also include a nonlinear response to a superposition of two non-interacting oscillatory influences. Just as at the heart of cross-spectral analysis is the coherence of phases associated to a frequency f, so by exact analogy, at the heart of bispectral analysis is the coherence between the sum of phases associated to two frequencies f_1, f_2 (each of which can be positive or negative) and the phase associated to the sum of the frequencies $f_1 + f_2$.

Bispectra (and more general higher-order spectra) of stationary stochastic processes were introduced in 1953 by John Tukey [70], and first applied in the study of nonlinear superposition of ocean waves [69, 24], subsequent to which many diverse applications have been found [65]. Bispectra have also been defined for deterministic signals [46, 2]. An in-depth exposition of bispectra and more general higher-order spectra in both the stationary stochastic setting and the deterministic setting can be found in [47] (where other uses of bispectra than investigating interaction of physical processes are also discussed).

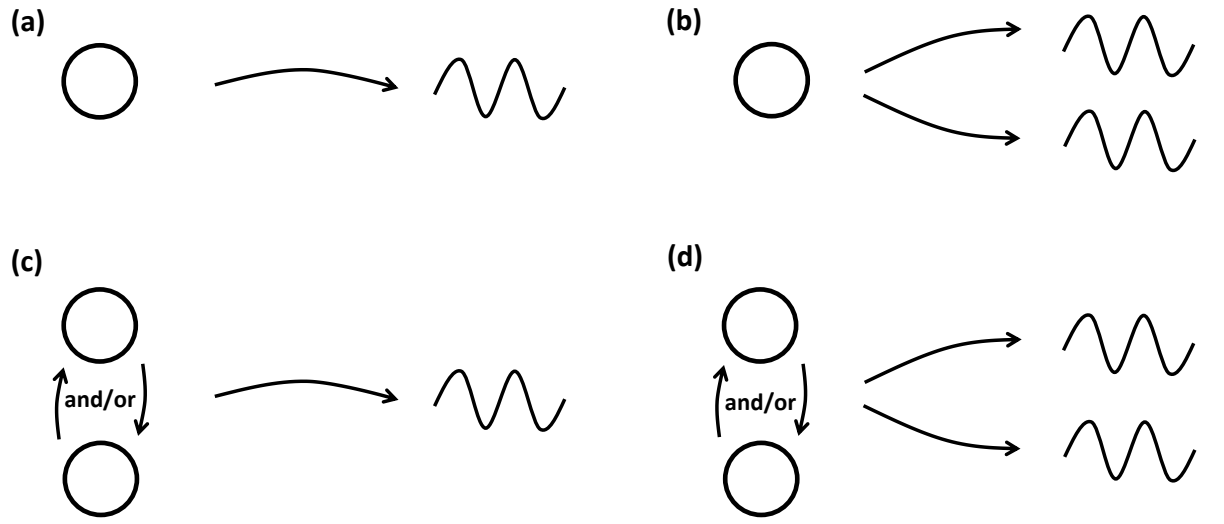


Figure 1: Illustration of what can be investigated by (a,b) second-order, and (c,d) third-order spectral analysis. (a) An oscillatory process contributing to a signal, as detected by power-spectral analysis. (b) An oscillatory process contributing simultaneously to two signals, as detected by cross-spectral analysis. (c) An interacting pair of oscillatory processes contributing to a signal, as detected by autobispectral analysis of the signal. (d) An interacting pair of oscillatory processes contributing simultaneously to two signals, as detected by cross-bispectral analysis of the two signals; cross-bispectral analysis can also be applied to three simultaneous signals.

[1]Bispectral results on the diagonal of frequency-frequency space indicate the potential influence of individual nonlinear oscillations rather than nonlinear combinations of oscillations [46, Sec. IIIB]. Likewise, bispectral results at rational ratios between the two coordinates of frequency-frequency space may be due to individual nonlinear oscillations. See also Sec. 5.2.

Traditional bispectra, as with traditional second-order spectra, are defined as functions of frequency variables but no time variable. However, thermodynamically open systems – such as biological systems – have time-variable characteristics, such as temporal intermittency of interaction between oscillatory components [72, 73], and temporal variation in the basic frequencies of the oscillatory components themselves [29]. Extensions of traditional bispectra of stationary stochastic processes to the setting of periodic [16, 61] and almost-periodic [10] stochastic processes are also not capable of dealing with "free-form" temporal variations, as is necessary for investigating typical open systems such as biological systems.

Now in the context of second-order spectral analysis, for a time-resolved investigation of common oscillatory influences between two simultaneous signals, one can:

- define a time-evolving cross-spectrum derived from a *time-frequency representation* of the signals (that is, a time-evolving representation of the frequency content of the signals);
- then use this time-evolving cross-spectrum to analyse coherence of phases within a sliding time-window, where the temporal spread of the window is
  - on the one hand, sufficiently short that changes in behaviour can be resolved and located in time reasonably well,
  - but on the other hand, sufficiently long that coherence of phases over the window can be meaningfully considered.

This approach to time-evolving second-order cross-spectral analysis was extended to bispectral analysis in the pioneering papers [72, 73]. The time-frequency representation used there for defining time-evolving bispectra was the continuous wavelet transform, and the time-evolving bispectra defined in this way are called *wavelet bispectra*. The advantage of the wavelet transform over other time-frequency representations is that it rescales its time-localisation in accordance with each frequency under investigation, thereby enabling the simultaneous resolving of oscillatory contributions on a broad range of timescales. Wavelet bispectra, as with wavelet cross-spectra, are defined for deterministic signals rather than stochastic processes, but such objects can still be applied to stochastic processes – both stationary and non-stationary – by sample-pathwise application [9]; in this case, a statistical descriptor [40] may then be obtainable by taking expectations (and factoring out the signal or window duration as appropriate).

There do exist various time-frequency representations other than the continuous wavelet transform [4], an important example being the windowed Fourier transform. Various time-evolving bispectra other than the wavelet bispectrum of [72, 73] have also been introduced for both deterministic and stochastic processes (in both discrete and continuous time) [46, 18, 64, 15, 51, 63, 1, 28, 56, 58, 54, 5, 53, 68].

Overview of application of wavelet bispectral analysis

Since its introduction, wavelet bispectral analysis has found effective application in diverse areas, including: the identification of relations between different oscillations in turbulent flows [17], especially of plasmas [72, 73, 12, 34, 60], and in laminar flows at the boundary of transition to turbulence [23]; investigation of water wave interactions beyond linear superposition [11, 75, 6]; analysis of electrostatic fluctuations in the ionosphere [41, 13]; interactions between brain waves measured in EEG signals during sleep [42], anaesthesia [7], and burst suppression while sedated [55] (where a slightly modified version of the wavelet transform is used), as well as characterisation of pain from

EEG signals [21]; prediction of epileptic seizures from real-time analysis of electrocorticography signals [71]; characterisation of wheezes from their audio recording [66], and characterisation of abnormal cardiac sounds [35]; investigation of cardiorespiratory interaction from respiration signals and ECG and/or blood flow signals [27, 33]; analysis of oscillations in photovoltaic current of surface state electrons on liquid helium [8]; coordination of renal autoregulation mechanisms via analysis of kidney blood flow signals [57]; and analysis of vibrations to detect mechanical faults [74, 43, 44]. Surrogate testing for wavelet bispectral analysis results is discussed in [38, 57, 20, 33].

Current definition and its limitation

In the papers [72, 73] where wavelet bispectral analysis was introduced, given two signals $x(t)$ and $y(t)$ (which could be the same or different), the wavelet bispectrum over a time-interval $I = [T, T + \Delta T]$ was defined by replacing Fourier transform terms in the traditional definition of bispectral density of deterministic finite-energy signals with time-dependent continuous wavelet transform terms, and integrating over the time-interval. That is,

$$B^I_{xxy}(s_1, s_2) \;:=\; \int_T^{T+\Delta T} W_x(s_1, t) W_x(s_2, t) \overline{W_y((s_1^{-1} + s_2^{-1})^{-1}, t)}\, dt \tag{1}$$

where s_1 and s_2 are the input timescales, and $W_x(s, t)$ and $W_y(s, t)$ are the continuous wavelet transforms of the signals $x(\cdot)$ and $y(\cdot)$ respectively (which depend on the choice of mother wavelet for the wavelet transform). One can easily extend this definition to cover three signals $x(t)$, $y(t)$ and $z(t)$ by

$$B^I_{xyz}(s_1, s_2) \;:=\; \int_T^{T+\Delta T} W_x(s_1, t) W_y(s_2, t) \overline{W_z((s_1^{-1} + s_2^{-1})^{-1}, t)}\, dt. \tag{2}$$

The wavelet bispectrum defined in this manner makes it possible for the concept of a third-order coherence index, as had been defined for stationary stochastic processes [47, Sec. 2.3.9], to be translated to the setting of deterministic signals with no assumption of stationary statistics: specifically, as a third-order analogue of wavelet coherence, *wavelet bicoherence* was defined in [72, 73] by

$$\mathfrak{c}^I_{xxy}(s_1, s_2) \;:=\; \frac{|B^I_{xxy}(s_1, s_2)|^2}{\int_T^{T+\Delta T} |W_x(s_1, t) W_x(s_2, t)|^2\, dt \int_T^{T+\Delta T} |W_y((s_1^{-1} + s_2^{-1})^{-1}, t)|^2\, dt}. \tag{3}$$

Eq. (1) has continued to persist as the standard definition of the wavelet bispectrum [22]. However, one basic question that cannot be addressed by wavelet bispectral analysis based on this definition of the bispectrum is how to evaluate quantitatively, rather than just describe qualitatively, the bispectral content within a given region in scale-scale space. Even for just one isolated bispectral peak in the scale-scale space, the *Heisenberg-Gabor uncertainty principle* [31, Sec. 2.2] implies that this peak will not be a Dirac mass but rather a blur; and this peak will only be yet more blurry if the frequencies of the oscillations involved are time-varying. Naturally, the inability to quantify the bispectral content of a region in scale-scale space also implies the inability to describe how bispectral content is distributed across different regions in scale-scale space.

The reason why this basic question has received almost no attention is likely because wavelet bispectral analysis has tended to focus specifically on bicoherence analysis using Eq. (3), for which Eq. (1) is merely a preliminary definition necessary in the construction of the wavelet bicoherence $\mathfrak{c}^I_{xxy}$. However, only analysing the wavelet bicoherence $\mathfrak{c}^I_{xxy}$ directly makes it difficult to tell whether

high bicoherence-values correspond to interacting oscillatory influences or just the absence of oscillatory components and associated harmonics, and in the former case, also makes it difficult to locate the approximate frequencies of the oscillatory influences. The actual bispectral values prior to the factoring out of norms in Eq. (3) may contain important information.

Illustration of the limitation

To illustrate the problem, in Fig. 2 we plot over scale-scale space the integrand of Eq. (2) at a fixed time τ, i.e. we plot the "instantaneous wavelet bispectrum"

$$b^{\tau}_{xyz}(s_1, s_2) \;:=\; W_x(s_1, \tau) W_y(s_2, \tau) \overline{W_z((s_1^{-1} + s_2^{-1})^{-1}, \tau)}, \tag{4}$$

using the conventional $s^{-\frac{1}{2}}$ normalisation in the wavelet terms as used in [72, 73], for signals of the form

$$\left\{ \begin{array}{rcl} x(t) & = & \cos(2\pi\nu_1 t) \\ y(t) & = & \cos(2\pi\nu_2 t) + \cos(2\pi\nu_3 t) \\ z(t) & = & \cos(2\pi(\nu_1 + \nu_2)t) + \cos(2\pi(\nu_1 + \nu_3)t - \theta). \end{array} \right. \tag{5}$$

For our choice of signals and parameters, the dependence on the fixed time τ is negligible, so the results are essentially the same as if integrating over a time-interval I as in Eq. (2).

We use a "lognormal" mother wavelet as described in Sec. 5. For the frequency resolutions used here, this is very similar to the more conventional Morlet wavelet; but all our later numerical illustrations will use lognormal mother wavelets, and so we also use it here. (The reason for our choice of lognormal wavelets in this paper are that they are analytic – which is required for the theory developed in this paper – and that the lognormal wavelet transform has good time-frequency resolution properties [25], including rapidly decaying wavelet power in the high-frequency tails.)

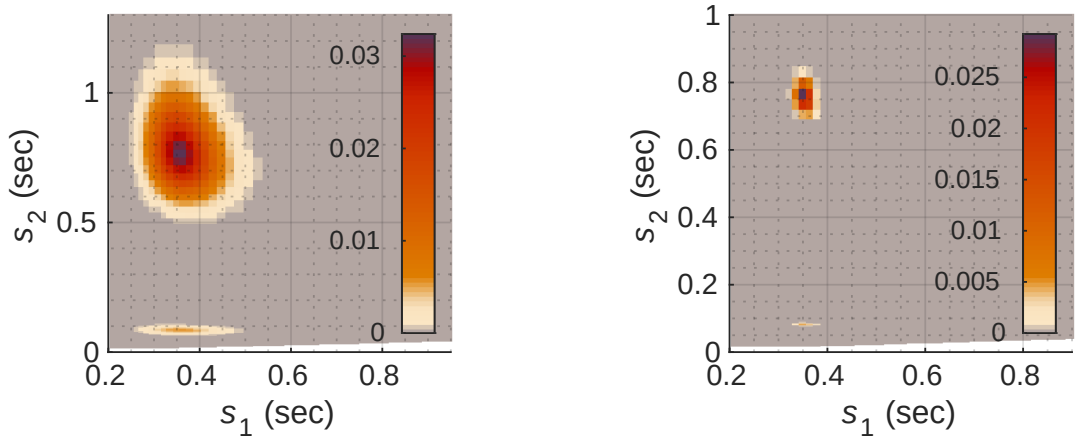
In Fig. 2 (both (a) and (b)), we see clearly a bispectral contribution associated to the frequency pair (ν_1, ν_2) and a bispectral contribution associated to the frequency pair (ν_1, ν_3), both appearing as "blurry" peaks as described above; but it is not at all immediately clear how to compare the weight of the two bispectral contributions, let alone how to quantify them absolutely. Certainly one would not immediately expect from the figure that the two bispectral contributions should be considered equal in magnitude. But from a Fourier analysis perspective, they are indeed equal: in both cases the three contributing sinusoidal components all have amplitude equal to 1.

How we solve the problem

Traditional time-independent spectral analysis defines *spectral densities* that can be integrated over regions of the space of input frequencies. Likewise, time-dependent second-order spectral analysis, via the windowed Fourier transform or the continuous wavelet transform, defines spectral densities that can be integrated over a region of time-frequency space or time-scale space.

Accordingly, to solve the problem described above, we seek to define a notion of *wavelet bispectral density*[2] by suitable normalisation of the formula (1)/(2), which can then be integrated over a region of time-frequency-frequency space or time-scale-scale space to give the bispectral content of that region. A definition of wavelet bispectral density is suitable insofar as it simultaneously fulfils the following two requirements:

[2]The term "wavelet bispectral density" is used in [39], but there the "density" just refers to "density with respect to time", i.e. where the duration ΔT of the time-interval is factored out.



(a) $\sigma = 1$ (b) $\sigma = 4$

Figure 2: Wavelet bispectrum according to the current standard definition. Plots show magnitude $|b^{\tau}_{xyz}(s_1, s_2)|$ of the instantaneous wavelet bispectrum (4) in units of $\sec^{\frac{3}{2}}$ for the signals x, y, z in (5) with $\nu_1 = 2.8$ Hz, $\nu_2 = 1.3$ Hz, $\nu_3 = 12$ Hz and $\theta = \frac{\pi}{3}$, using conventional $p = \frac{1}{2}$ definition of the wavelet transform, with a lognormal mother wavelet (Sec. 5) with frequency resolution parameter (a) $\sigma = 1$ and (b) $\sigma = 4$. With these lognormal mother wavelets, we consider scale s and frequency f to be related by $f = \frac{1}{s}$. The signals are simulated over a time-interval $[0\,\text{s}, 60\,\text{s}]$ and the time-instant τ is taken at $\tau = 30$ s.

- For strictly quasiperiodic signals,
  - the wavelet bispectral content of each contribution to the bispectrum should match well that given by Fourier bispectral analysis;
  - and the "blurriness" of the peaks in the wavelet bispectral density should be in accordance with the frequency resolution of the underlying wavelet transform.
- For more general signals, the bispectral content of a given region of frequency-frequency space over a given time-interval I should be very little affected by the behaviour of the signals outside of the time-interval I.

In seeking to obtain such a definition of wavelet bispectral density, we specifically consider wavelets that are analytic with non-negative-valued Fourier transform; this is a reasonable assumption for time-frequency analysis, and non-analytic wavelets can be made analytic by convolution with the inverse Fourier transform of a suitable cut-off function. We assume a simple inverse relationship $f = \frac{\kappa}{s}$ between scale and frequency. This is very natural for time-frequency analysis using wavelets with unimodal Fourier transform, and is also necessary for the application of wavelet bispectral analysis to investigate nonlinear oscillatory interactions [73]. We obtain:

- a "global" definition of wavelet bispectral density with respect to logarithmic frequency axes, which enables one to define the time-evolving distribution of bispectral content across frequency-frequency space (and also to compare this across signals);
- a "local" definition of wavelet bispectral density with respect to logarithmic frequency axes, which is (very slightly) better suited to the investigation of bispectral content around a prespecified location in frequency-frequency space than the "global" definition is.

In the "global" definition, the normalisation factor in front of the product of wavelet terms in (2) is itself dependent on the ratio of the two input frequencies or scales. (If $p = 1$ is used for the s^{-p} normalisation in the individual wavelet terms themselves, then it is *only* the ratio, and not the two individual input frequencies, that is needed in the extra normalisation factor.) An important question is whether this non-constancy of the normalisation factor in our "global" definition causes the distribution of bispectral content over frequency-frequency space to be misrepresented. Accordingly, Theorem 8 justifies the soundness of our definition in a limit as the frequency resolution of the mother wavelet tends to ∞, while the numerics in Sec. 5.1 indicate that our definition gives suitable results even when the frequency resolution is not high at all.

The wavelet bispectral densities described above are with respect to logarithmic frequency, but can easily be converted into densities with respect to linear frequency. Nonetheless, the derivation of these densities is based on consideration of logarithmic frequency. A formula derived analogously from consideration of linear frequency can also be obtained, but this is only able to give a "local" definition of wavelet bispectral density, not a "global" one. This is presented in Appendix C.

Relation to existing work

The significance of the question of how to be able to compare wavelet bispectral content across different regions of scale-scale space has been highlighted in [29], when comparing wavelet bispectral analysis with traditional Fourier-based bispectral analysis. No clear mathematical answer to this question is obtained, but working with a modified Morlet wavelet transform, the paper does suggest a wavelet normalisation that essentially corresponds to taking $p = 1$ in the s^{-p} normalisation for wavelet transforms, and appears to hint implicitly at the fact that doing so normalises the modal magnitude of a fixed-frequency bispectral contribution in proportion with its Fourier bispectral content; we describe this fact in Sec. 3.3.

Normalisation of wavelet bispectra with reference to traditional bispectra has also been addressed in [39], from quite a different point of view from our question of how to compute the bispectral content of a region of time-frequency-frequency space. Namely, for application to stochastic processes of constant expectation, a normalised wavelet bispectrum formula (given explicitly for discretely sampled finite-duration time-series recordings) is presented, that is, roughly speaking, derived so that if applied to a stationary stochastic process of flat bispectral density, the infinite-time-average expected wavelet bispectrum at each pair of frequencies matches the traditional bispectral density.

Significance of our result

By defining wavelet bispectral density, we enable a quantitative evaluation of the time-localised distribution of bispectral content over frequency-frequency space, in a manner that can still take advantage of the multiscale resolution capacity of the continuous wavelet transform. This has the potential to lead to more refined indicators of the various physical and biological phenomena in open systems that wavelet bispectral analysis is used to investigate.

Structure of the paper

Sec. 2 covers preliminaries on the (continuous) wavelet transform, second-order Fourier and wavelet spectra, and third-order Fourier spectra. After starting with notational conventions, we introduce the wavelet transform, first according to the $p = 1$ convention which we consider most simple and natural, and then with general p. We then discuss the basic scaling property of the

wavelet transform, particularly as compared with the windowed Fourier transform. We then present definitions and basic results concerning second-order Fourier spectra of finite-energy deterministic signals, followed by analogous definitions and results for wavelet second-order spectra of deterministic signals. Finally, we present definitions and basic results concerning Fourier bispectra of finite-energy deterministic signals.

Sec. 3 covers our definitions and results pertaining to wavelet bispectral analysis. We start by describing how bispectral density would be defined for the windowed Fourier transform, and the technical difficulty in trying to do the same for the wavelet transform. In view of the impossibility of obtaining a "perfect" definition of wavelet bispectra, we explain our criteria for a suitable definition, after which we introduce our main definition of wavelet bispectral density and the wavelet bispectrum. Additional theoretical justification (Theorem 8) for the suitability of this definition is provided. We then discuss frequency-localised adaptation of our definition of the wavelet bispectrum.

In Sec. 4, we prove Theorem 8.

In Sec. 5, we provide numerical illustration of our wavelet bispectrum formula. First, we apply it to signals with pure sinusoidal components, to illustrate the match with Fourier analysis. Then, we apply it to a coupled phase-oscillator pair, to illustrate the usage of our new formula for investigating interactions within time-dependent oscillatory dynamics.

In Sec. 6, we illustrate an application of our wavelet bispectrum formula to recorded time-series of cell membrane potentials.

In Appendix A, we describe the relationship between the lognormal wavelet parameter σ and corresponding time and frequency resolution properties.

In Appendix B, we discuss definitions of coherence, bicoherence and "bicorrelation" (introduced in [33] under the name "real wavelet biphase"). In particular, we introduce a new definition of "wavelet-phase bicoherence". We also briefly describe how wavelet bispectral analysis can be used to suggest unidirectionality of coupling [30].

In Appendix C, we present an alternative frequency-localised wavelet bispectrum formula to that given in Sec. 3.3.

Availability of codes

Codes for the wavelet bispectral analysis methods developed and used in this paper will be made available on the Lancaster Publications and Research system Pure.

2. Preliminaries

In this section, we discuss the wavelet transform and its basic properties, introduce second-order spectra for both Fourier and wavelet transforms, and introduce third-order Fourier spectra. This then lays the foundation for introducing third-order wavelet spectra in Sec. 3.

Some conventions and notations

Fourier transform. In this paper, the Fourier transform is formally defined according to the convention

$$(\mathcal{F}x)(f) \;=\; \hat{x}(f) \;=\; \int_{-\infty}^{\infty} x(\tau)e^{-2\pi i f\tau}\,d\tau.$$

It will also be very useful to define

$$\mathring{x}(f) \;:=\; \hat{x}\big(\tfrac{1}{f}\big).$$

One-point compactification of the real line. In accordance with the usual notation, we write $\hat{\mathbb{R}}$ for the one-point compactification of $\mathbb{R}$, namely $\mathbb{R} \cup \{\infty\}$ where ∞ is regarded as the same as $-\infty$. (The $\hat{}$ here is unrelated to Fourier transforms.)

Dimensional analysis. Generally speaking, variables denoted t or τ will correspond to physical quantities of the time dimension, variables denoted f, ν or ξ will correspond to physical quantities of the frequency dimension, variables denoted ζ will correspond to a logarithmic representation of quantities of the frequency dimension, and variables denoted r will correspond to dimensionless values. The variable λ (parametrising "projective frequency-frequency space") and wavelet-related parameters κ and σ are also dimensionless.

Spaces of functions. Given a space of complex-valued functions $f\colon U \to \mathbb{C}$ whose notation follows the style "$F(U)$" (e.g. Lebesgue spaces $L^p(U)$, Sobolev spaces $W^{k,p}(U)$, etc.), for any $X \subset \mathbb{C}$ we use the style of notation "$F(U, X)$" to mean the set of functions $f\colon U \to X$ that belong to $F(U)$ when the codomain is expanded from X to $\mathbb{C}$. (For example, $L^2(\mathbb{R})$ is the set of square-integrable functions $f\colon \mathbb{R} \to \mathbb{C}$ while $L^2(\mathbb{R}, \mathbb{R})$ is the set of square-integrable functions $f\colon \mathbb{R} \to \mathbb{R}$.)

Delta-ring of finite integration. Given a Borel set $X \subset \mathbb{R}^d$ and a measurable function $f\colon X \to \mathbb{C}$, we define

$$\mathcal{B}_f \;:=\; \left\{ A \in \mathcal{B}(X) \;:\; \int_A |f(x)|\, dx < \infty \right\}.$$

One can easily show that $\mathcal{B}_f$ is a δ-ring (i.e. is closed under pairwise union, countable intersection and relative complement) on which $A \mapsto \int_A f(x)\, dx$ is σ-additive.

Summary of main other notations: ψ denotes a (mother) wavelet function; $W_{\psi,\kappa,x}$ is the corresponding wavelet transform (with $p = 1$) of the signal $x(t)$, where κ is the constant of inverse proportion between scale s and frequency f; $W^{[p]}_{\psi,\kappa,x}$ is the wavelet transform for general p in the s^{-p} normalisation; second-order Fourier spectral densities of finite-energy signals x and y are denoted $\mathcal{P}_{xy}$ (with $x = y$ being "energy spectral density" and $x \neq y$ being "cross-energy spectral density"); $\tilde{p}_{\psi,\kappa,xy}$ likewise denotes second-order wavelet spectral density (as a function of both frequency and time) with respect to a linear frequency axis, and $p_{\psi,\kappa,xy}$ with respect to a logarithmic frequency axis; $\mathfrak{p}_{\psi,\kappa,xy}$ is the second-order wavelet spectrum obtained by integrating the second-order wavelet spectral density; $\mathcal{B}_{xyz}$ denotes the Fourier bispectral density for a triplet of finite-energy signals (x, y, z); $b_{\psi,\kappa,xyz}$ denotes globally defined wavelet bispectral density with respect to logarithmic frequency axes, and $\tilde{b}_{\psi,\kappa,xyz}$ is the transformed version for linear frequency axes; $\mathfrak{b}_{\psi,\kappa,xyz}$ is the globally defined wavelet bispectrum obtained by integrating the wavelet bispectral density; "local" wavelet bispectral density and the "local" wavelet bispectrum have an extra subscript $_\lambda$ for the frequency ratio $\lambda = \frac{f_1}{f_1+f_2}$ about which the localisation takes place.

2.1. Definition of the wavelet transform

In our definition of the wavelet transform, we will specifically work with absolutely integrable Hermitian analytic wavelet functions with non-negative-valued Fourier transform; these assumptions are reasonable for time-frequency analysis. Any non-analytic wavelet (meaning that the Fourier transform contains negative-frequency content) can be made analytic by convolution with the inverse Fourier transform of a suitable cut-off function: provided the frequency resolution of the original non-analytic wavelet is not too poor, this analyticisation can be achieved with very little modification to the wavelet itself. Whereas the wavelet transform traditionally has scale s as its first argument, throughout this paper the wavelet transform's first argument is a frequency

variable f that is reciprocally proportional to scale s (with a constant of proportion denoted κ), and all densities considered are with respect to f or $\log f$ rather than with respect to s.

A (*mother*) *wavelet function* will mean a Hermitian function $\psi \in L^1(\mathbb{R}) \setminus \{0\}$ with $\hat{\psi} \geq 0$, $\hat{\psi}\big|_{(-\infty,0]} = 0$, and $\int_0^\infty \frac{\hat{\psi}(r)}{r}\, dr < \infty$. (This implies in particular that $\int_{\mathbb{R}} \psi(r)\, dr = 0$.) Given a wavelet function ψ and a value $\kappa > 0$, we define the associated (*continuous*) *wavelet transform* of any $x \in L^\infty(\mathbb{R}, \mathbb{R})$ by

$$W_{\psi,\kappa,x}(f,t) \;=\; \frac{f}{\kappa} \int_{\mathbb{R}} x(\tau) \overline{\psi\Big(\tfrac{(\tau-t)f}{\kappa}\Big)}\, d\tau \tag{6}$$

for all $f > 0$ and $t \in \mathbb{R}$. In practice, one often chooses the wavelet function ψ such that $\hat{\psi}$ is unimodal, and then takes κ to be where $\hat{\psi}$ is maximised [25, 73].

We can also take the wavelet transform of an unbounded function $x \in L^q(\mathbb{R}, \mathbb{R})$ for any $q \in [1, \infty)$, in which case (6) is well-defined for almost all $(f,t) \in (0,\infty) \times \mathbb{R}$. When $x \in L^1 \cup L^2$, the wavelet transform can be computed from the Fourier transform of its temporal sections,

$$(\mathcal{F} W_{\psi,\kappa,x}(f,\cdot))(\xi) \;=\; \hat{x}(\xi) \hat{\psi}\Big(\tfrac{\kappa\xi}{f}\Big) \tag{7}$$

for almost all $\xi \in \mathbb{R}$, for each $f > 0$.

Although not typically necessary, one can choose to extend the definition of $W_{\psi,\kappa,x}(f,t)$ to include negative f, by

$$W_{\psi,\kappa,x}(f,t) \;=\; \frac{|f|}{\kappa} \int_{\mathbb{R}} x(\tau) \overline{\psi\Big(\tfrac{(\tau-t)f}{\kappa}\Big)}\, d\tau \;=\; \overline{W_{\psi,\kappa,x}(-f,t)}$$

for all $f \neq 0$ and $t \in \mathbb{R}$. In this case, (7) holds for negative as well as positive f.

The wavelet transform provides a means of *time-localised frequency analysis*: in the case that $x(t) = A\cos(2\pi\nu t + \phi)$ with $\nu > 0$, we have

$$W_{\psi,\kappa,x}(f,t) \;=\; \tfrac{1}{2} \hat{\psi}\Big(\tfrac{\kappa\nu}{f}\Big)\, A e^{i(2\pi\nu t + \phi)} \tag{8}$$

for all $f > 0$ and $t \in \mathbb{R}$. If we include negative f, this becomes

$$W_{\psi,\kappa,x}(f,t) \;=\; \tfrac{1}{2} \hat{\psi}\Big(\tfrac{\kappa\nu}{|f|}\Big)\, A e^{\operatorname{sgn}(f) i(2\pi\nu t + \phi)}. \tag{9}$$

From (8) we see that, unlike the Fourier transform, the representation of a sine wave given by the wavelet transform is, at each time t, not a sharp peak but rather a *blurry peak* in accordance with the shape of $\hat{\psi}$. This is an inherent property of time-localised frequency analysis: very fine time-localisation necessitates poor frequency resolution, and very high frequency resolution necessitates poor time-localisation. This is due to the Heisenberg-Gabor uncertainty principle mentioned in the Introduction.

We refer to the modulus and the argument of $W_{\psi,\kappa,x}(f,t)$ respectively as the *wavelet amplitude* and the *wavelet phase* associated to the frequency f at time t.

Remark 1. It is always possible to rescale ψ such that κ becomes 1: namely, defining $\psi_1(r) = \frac{1}{\kappa}\psi\big(\frac{r}{k}\big)$, we have $W_{\psi,\kappa,x} = W_{\psi_1,1,x}$. Thus, from a mathematical point of view, we could simply take $\kappa = 1$ without any loss of generality. However, in practice, the conventional formulations of some wavelets (such as Morlet wavelets) do not have their Fourier transform $\hat{\psi}$ maximised at or near 1. Because of this fact, we retain the presence of κ in our definitions.

Definition with general p

An alternative, more commonly found, definition of the wavelet transform is

$$W^{[\frac{1}{2}]}_{\psi,\kappa,x}(f,t) \;=\; \sqrt{\tfrac{|f|}{\kappa}} \int_{\mathbb{R}} x(\tau)\overline{\psi\Big(\tfrac{(\tau-t)f}{\kappa}\Big)}\,d\tau \;=\; \sqrt{\tfrac{\kappa}{|f|}} \cdot W_{\psi,\kappa,x}(f,t) \tag{10}$$

for all $f \neq 0$ and $t \in \mathbb{R}$; and more generally, for any $p \geq 0$, one can define

$$W^{[p]}_{\psi,\kappa,x}(f,t) \;=\; \Big(\tfrac{|f|}{\kappa}\Big)^p \int_{\mathbb{R}} x(\tau)\overline{\psi\Big(\tfrac{(\tau-t)f}{\kappa}\Big)}\,d\tau \tag{11}$$

for all $f \neq 0$ and $t \in \mathbb{R}$. So $W_{\psi,\kappa,x}$ corresponds to $p = 1$.

One justification for defining the wavelet transform according to (10) rather than (6) is that the L^2 norm of the map $\tau \mapsto \sqrt{\frac{f}{\kappa}}\,\overline{\psi\Big(\frac{\tau f}{\kappa}\Big)}$ is constant over varying f, being simply equal to the L^2 norm of ψ itself. However, for the context of higher-order spectra the meaningfulness of such a justification is doubtful. We will see that when working with logarithmic frequency axes in accordance with the logarithmic frequency resolution of the wavelet transform (Sec. 2.2), the formulae (25), (26), (45) and (49) for both second- and third-order densities are simplest when we take $p = 1$. (By contrast, in the "linear-frequency derived" wavelet spectral density formulae, while the second-order densities (23) and (24) are simplest with $p = \frac{1}{2}$ as in (10), the third-order densities (C.8) are instead simplest with $p = \frac{1}{3}$.) Two important further practical advantages of working with $p = 1$ are:

- Assuming that κ is the mode of $\hat{\psi}$, the maximal amplitude for the wavelet transform of a sinusoidal input will occur precisely at the frequency of the input, if and only if $p = 1$.
- Computation of the wavelet transform via (7) is simplest with $p = 1$: an additional pre-factor $\Big(\frac{|f|}{\kappa}\Big)^{p-1}$ is needed for any other value of p.

Finite-time signals

The theoretical definition of the wavelet transform works with infinite-time signals. For a finite-time signal $x\colon (a,b) \to \mathbb{R}$, a typical procedure is as follows: First define an extension $x_{\text{ext}}\colon \mathbb{R} \to \mathbb{R}$ of x; typical examples include the *zero-padding extension* ($x_{\text{ext}}(t) = 0$ for $t \notin (a,b)$) and the *periodic padding extension* (x_{ext} is $(b-a)$-periodic), as well as more sophisticated forms of "predictive" padding [25]. Then define the wavelet transform of x to be the restriction of $W_{\psi,\kappa,x_{\text{ext}}}$ to some subset $C \subset (0,\infty) \times (a,b)$ of (f,t)-space, where C is chosen such that the behaviour of x_{ext} outside the time-interval (a,b) has little influence on the values of $W_{\psi,\kappa,x_{\text{ext}}}(f,t)$ for $(f,t) \in C$. This set C is called the *cone of influence*. See [25] for further details (as well as discussion of other practical aspects of wavelet analysis of digital signals).

Throughout the rest of this paper, we present theory for infinite-time signals; application to finite-time signals can be achieved by the above procedure.

2.2. Comparison with the windowed Fourier transform

Given an even function $w \in L^1(\mathbb{R},\mathbb{R})$ with $\int_{\mathbb{R}} w(t)\,dt > 0$ (called a *window function*), we define the associated *windowed Fourier transform* of a function $x \in L^\infty(\mathbb{R},\mathbb{R})$ by

$$F_{w,x}(f,t) \;=\; \int_{\mathbb{R}} x(\tau)w(\tau-t)e^{2\pi i f(t-\tau)}\,d\tau$$

for all $f, t \in \mathbb{R}$. In other words,

$$F_{w,x}(\cdot, t) \;=\; \mathcal{F}(\tau \mapsto x(\tau + t)w(\tau)).$$

Note that $\hat{w}$ is an even real-valued function. Typically, one uses a window function w such that $\hat{w}$ is maximised at 0 and such that at any other local maximum of $|\hat{w}|$ the value of $|\hat{w}|$ is very small compared to $\hat{w}(0)$. Thus the windowed Fourier transform provides a means of time-localised frequency analysis, over frequency ranges that do not come too close to 0: for $x(t) = A\cos(2\pi\nu t + \phi)$ with $\nu > 0$, we have

$$F_{w,x}(f, t) \;=\; \tfrac{1}{2}\hat{w}(f - \nu)Ae^{i(2\pi\nu t + \phi)} \;+\; \underbrace{\tfrac{1}{2}\hat{w}(f + \nu)Ae^{-i(2\pi\nu t + \phi)}}_{\to 0 \text{ uniformly over } f > 0 \text{ as } \nu \to \infty} \tag{12}$$

for all $f, t \in \mathbb{R}$. We see from (12) that the smallest frequencies for which the windowed Fourier transform is able to give meaningful results depends on how quickly the tails of $\hat{w}$ decay.

At the heart of the difference between the windowed Fourier transform and the wavelet transform is the following: In the windowed Fourier transform, the "envelope" $w(\tau - t)$ of the function against which the input signal is integrated is independent of the frequency f under investigation. By contrast, in the wavelet transform the function $\tau \mapsto \frac{f}{\kappa}\,\overline{\psi\Big(\frac{(\tau - t)f}{\kappa}\Big)}$ against which the input signal is integrated is a linearly rescaled version of $\overline{\psi}$ where the temporal rescaling is inversely proportional to the frequency f under investigation. Note that, from the point of view of dimensional analysis, the inputs of a window function w are of the *time* dimension, while the inputs of a wavelet function ψ are dimensionless.

This rescaling of the wavelet in accordance with the frequency under investigation is the key property that makes the wavelet transform able to investigate different oscillatory components across multiple simultaneous timescales. A basic consequence of this rescaling is that if we define $x_c(t) = x(ct)$ for any constant $c > 0$, then

$$W_{\psi,\kappa,x_c}(f, t) \;=\; W_{\psi,\kappa,x}\big(\tfrac{f}{c}, ct\big) \tag{13}$$

for all $f \neq 0$ and $t \in \mathbb{R}$. (For more general p, the right-hand side has an extra pre-factor c^{p-1}.)

Eq. (13) can be interpreted as saying that the wavelet transform has *logarithmic frequency resolution*: For signals of the form

$$x(t) \;=\; A_1\cos(2\pi\nu_1 t) \;+\; A_2\cos(2\pi\nu_2 t),$$

the resolvability of the two sinusoidal components by the wavelet transform depends only on the ratio between the two frequencies ν_1 and ν_2. Indeed, if one plots $|W_{\psi,\kappa,x}|$ over time-frequency space with a logarithmic frequency axis, multiplying the frequencies ν_1 and ν_2 by the same number c would simply shift the entire diagram up the frequency axis by an amount proportional to $\log c$. This is in contrast with the *linear frequency resolution* of the windowed Fourier transform, where the resolvability of the two sinusoidal components would depend essentially on the *difference* in frequency $|\nu_2 - \nu_1|$ rather than the ratio $\frac{\nu_2}{\nu_1}$.

It is precisely the wavelet transform's "logarithmic" rather than "linear" treatment of frequencies that leads to the main non-triviality in defining the wavelet bispectrum, because the definition of bispectra involves a linear sum of the two input frequencies. (See the start of Sec. 3 for details.)

2.3. Second-order Fourier spectra of deterministic signals

Throughout this paper, the abbreviation "ESD" stands for "energy spectral density", and "cross-ESD" is to be interpreted "cross-energy spectral density".

The second-order Fourier spectral densities of finite-energy signals are defined as the integrands in the "frequency-domain integral formulae" for the L^2-norm and L^2-inner product. Namely, for any $x, y \in L^2(\mathbb{R}, \mathbb{R})$, we have

$$\int_{\mathbb{R}} x(t)y(t)\,dt \;=\; \int_{\mathbb{R}} \mathcal{P}_{xy}(f)\,df \;=\; 2\,\mathrm{Re}\biggl(\int_0^\infty \mathcal{P}_{xy}(f)\,df\biggr) \tag{14}$$

and in particular

$$\int_{-\infty}^{\infty} x(t)^2\,dt \;=\; \int_{-\infty}^{\infty} \mathcal{P}_{xx}(f)\,df \;=\; 2\int_0^\infty \mathcal{P}_{xx}(f)\,df \tag{15}$$

where $\mathcal{P}_{xx}\colon f \mapsto |\hat{x}(f)|^2$ is the *Fourier ESD* of x, and $\mathcal{P}_{xy}\colon f \mapsto \hat{x}(f)\overline{\hat{y}(f)}$ is the *Fourier cross-ESD* of x with y. We are careful to refer to these functions as "spectral *densities*", with a "spectrum" itself referring to the measure (for the energy spectrum) or complex-valued measure (for the cross-energy spectrum) obtained by integrating the associated spectral density. Note the equalities

$$\mathcal{P}_{xy}(-f) \;=\; \overline{\mathcal{P}_{xy}(f)} \;=\; \mathcal{P}_{yx}(f) \tag{16}$$

for all $f \in \mathbb{R}$, from which we have the following two consequences (which will also hold for second-order wavelet spectra):

- for second-order spectral analysis it is sufficient to consider only positive frequencies;
- for a pair of functions x and y there are essentially three distinct second-order spectra, namely the energy spectra of x and y and the cross-energy spectrum of x with y.

The above spectra are defined for real-valued signals. We also recall that for general complex-valued functions $x, y \in L^2(\mathbb{R})$,

$$\int_{\mathbb{R}} x(t)\overline{y(t)}\,dt \;=\; \int_{\mathbb{R}} \hat{x}(f)\overline{\hat{y}(f)}\,df \tag{17}$$

$$\int_{\mathbb{R}} x(t)y(t)\,dt \;=\; \int_{\mathbb{R}} \hat{x}(f)\hat{y}(-f)\,df, \tag{18}$$

and due to (17), for all $x \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$,

$$\int_{-\infty}^{\infty} |x(t)|^2\,dt \;=\; \int_{-\infty}^{\infty} |\hat{x}(f)|^2\,df. \tag{19}$$

2.4. Second-order wavelet spectra

Now let us consider a wavelet analogue of the second-order Fourier spectra. For any wavelet function ψ, define

$$C_\psi \;=\; \int_0^\infty \frac{\hat{\psi}(r)^2}{r}\,dr \;=\; \int_0^\infty \frac{\mathring{\psi}(r)^2}{r}\,dr \;=\; \int_{-\infty}^\infty \hat{\psi}(e^r)^2\,dr \;=\; \int_{-\infty}^\infty \mathring{\psi}(e^r)^2\,dr. \tag{20}$$

The value C_ψ is often known as the *admissibility constant*. It can be regarded as a "power-normalisation constant" for the wavelet transform: given a signal $x(t) = A\cos(2\pi\nu t + \phi)$ with $\nu \neq 0$, Eq. (8) implies

$$\tfrac{1}{4}A^2 \;=\; \frac{1}{\kappa C_\psi}\int_0^\infty |W^{[\frac{1}{2}]}_{\psi,\kappa,x}(f,t)|^2\,df \;=\; \frac{1}{C_\psi}\int_{-\infty}^\infty |W_{\psi,\kappa,x}(e^\zeta,t)|^2\,d\zeta \tag{21}$$

for any $t \in \mathbb{R}$. More generally, if we have signals

$$\begin{aligned} x(t) &= A_1\cos(2\pi\nu t + \phi) \\ y(t) &= A_2\cos(2\pi\nu t + \phi - \theta) \end{aligned}$$

with $\nu > 0$, Eq. (8) implies

$$\tfrac{1}{4}A_1A_2e^{i\theta} \;=\; \frac{1}{\kappa C_\psi}\int_0^\infty W^{[\frac{1}{2}]}_{\psi,\kappa,x}(f,t)\overline{W^{[\frac{1}{2}]}_{\psi,\kappa,y}(f,t)}\,df \tag{22a}$$

$$\tfrac{1}{4}A_1A_2e^{i\theta} \;=\; \frac{1}{C_\psi}\int_{-\infty}^\infty W_{\psi,\kappa,x}(e^\zeta,t)\overline{W_{\psi,\kappa,y}(e^\zeta,t)}\,d\zeta \tag{22b}$$

for any $t \in \mathbb{R}$. (We mention now that although (22a) and (22b) describe the same object C_ψ, the objects described by their "third-order analogues" – namely $D_\psi(\lambda)$ in Eq. (44), and $\tilde{D}_\psi(\lambda)$ in (A)–(C) in Appendix C – are not the same as each other.)

Given functions $x, y\colon \mathbb{R} \to \mathbb{R}$ with wavelet transforms $W_{\psi,\kappa,x}$ and $W_{\psi,\kappa,y}$, we define the *linear-frequency wavelet ESD* $\tilde{p}_{\psi,\kappa,xx}\colon (\mathbb{R}\setminus\{0\}) \times \mathbb{R} \to [0,\infty)$ of x by

$$\tilde{p}_{\psi,\kappa,xx}(f,t) \;:=\; (\kappa C_\psi)^{-1}|W^{[\frac{1}{2}]}_{\psi,\kappa,x}(f,t)|^2 \tag{23}$$

and the *linear-frequency wavelet cross-ESD* $\tilde{p}_{\psi,\kappa,xy}\colon (\mathbb{R}\setminus\{0\}) \times \mathbb{R} \to \mathbb{C}$ of x with y by

$$\tilde{p}_{\psi,\kappa,xy}(f,t) \;:=\; (\kappa C_\psi)^{-1}W^{[\frac{1}{2}]}_{\psi,\kappa,x}(f,t)\overline{W^{[\frac{1}{2}]}_{\psi,\kappa,y}(f,t)}, \tag{24}$$

and we define the *logarithmic-frequency wavelet ESD* $p_{\psi,\kappa,xx}\colon (0,\infty) \times \mathbb{R} \to [0,\infty)$ of x by

$$p_{\psi,\kappa,xx}(f,t) \;:=\; C_\psi^{-1}|W_{\psi,\kappa,x}(f,t)|^2 \tag{25}$$

and the *logarithmic-frequency wavelet cross-ESD* $p_{\psi,\kappa,xy}\colon (0,\infty) \times \mathbb{R} \to \mathbb{C}$ of x with y by

$$p_{\psi,\kappa,xy}(f,t) \;:=\; C_\psi^{-1}W_{\psi,\kappa,x}(f,t)\overline{W_{\psi,\kappa,y}(f,t)}. \tag{26}$$

At each time t, these spectral densities as a function of f satisfy the same relations (16) satisfied by the Fourier second-order spectral densities. Regarding the cross-energy spectral densities, note that

$$\frac{\tilde{p}_{\psi,\kappa,xy}(f,t)}{|\tilde{p}_{\psi,\kappa,xy}(f,t)|} \;=\; \frac{p_{\psi,\kappa,xy}(f,t)}{|p_{\psi,\kappa,xy}(f,t)|} \;=\; e^{i(\phi_x(f,t)-\phi_y(f,t))} \tag{27}$$

where $\phi_x(f,t) = \arg(W_{\psi,\kappa,x}(f,t))$ and $\phi_y(f,t) = \arg(W_{\psi,\kappa,y}(f,t))$; we refer to $\phi_x(f,t) - \phi_y(f,t)$ as the *phase difference* associated to f at time t.

Whereas the Fourier spectral densities in Sec. 2.3 are densities over the frequency axis, these wavelet spectral densities are densities over *time-frequency* space, with a logarithmic frequency axis

in the case of $p_{\psi,\kappa,xy}$ and a linear frequency axis in case of $\tilde{p}_{\psi,\kappa,xy}$. The *wavelet energy spectrum* of x is the measure $\mathfrak{p}_{\psi,\kappa,xx}$ on $(0,\infty)\times\mathbb{R}$ given by

$$\begin{aligned}\mathfrak{p}_{\psi,\kappa,xx}(A) &= \int_{\mathbb{R}^2} \mathbb{1}_A(e^\zeta,t)p_{\psi,\kappa,xx}(e^\zeta,t)\,d(\zeta,t) &(28)\\ &= \int_A \tilde{p}_{\psi,\kappa,xx}(f,t)\,d(f,t). &(29)\end{aligned}$$

This measure can also be extended to include the negative frequency axis (i.e. to be a measure on the whole of $\mathbb{R}^2$) using (29). Similarly, for any $A\subset(0,\infty)\times\mathbb{R}$ with $A\in\mathcal{B}_{\tilde{p}_{\psi,\kappa,xy}}$ we define

$$\begin{aligned}\mathfrak{p}_{\psi,\kappa,xy}(A) &= \int_{\mathbb{R}^2} \mathbb{1}_A(e^\zeta,t)p_{\psi,\kappa,xy}(e^\zeta,t)\,d(\zeta,t) &(30)\\ &= \int_A \tilde{p}_{\psi,\kappa,xy}(f,t)\,d(f,t). &(31)\end{aligned}$$

Again we can extend $\mathfrak{p}_{\psi,\kappa,xy}$ to all sets $A\in\mathcal{B}(\mathbb{R}^2)$ with $A\setminus(\{0\}\times\mathbb{R})\in\mathcal{B}_{\tilde{p}_{\psi,\kappa,xy}}$ using (31). We refer to $\mathfrak{p}_{\psi,\kappa,xy}$ as the *wavelet cross-energy spectrum* of x with y. If $x,y\in L^\infty(\mathbb{R},\mathbb{R})$ then for any compact $K\subset(0,\infty)\times\mathbb{R}$ the restriction of $\mathfrak{p}_{\psi,\kappa,xy}$ to $\mathcal{B}(K)$ defines a complex-valued measure on K.

In analogy to the energy-preservation and cross-energy-preservation properties of the Fourier transform (14)-(15), we have the following well-known fact [31, Sec. 3.2]. (We include a proof, particularly for the sake of the discussion at the start of Sec. 3.)

Proposition 2 (Wavelet energy preservation)**.** *For any $x,y\in L^2(\mathbb{R},\mathbb{R})$ we have*

$$\int_{\mathbb{R}} x(t)y(t)\,dt \;=\; \mathfrak{p}_{\psi,\kappa,xy}(\mathbb{R}^2) \;=\; 2\,\mathrm{Re}(\mathfrak{p}_{\psi,\kappa,xy}((0,\infty)\times\mathbb{R})), \tag{32}$$

and in particular,

$$\int_{-\infty}^{\infty} x(t)^2\,dt \;=\; \mathfrak{p}_{\psi,\kappa,xx}(\mathbb{R}^2) \;=\; 2\,\mathfrak{p}_{\psi,\kappa,xx}((0,\infty)\times\mathbb{R}). \tag{33}$$

Proof. We first prove (33):

$$\begin{aligned}C_\psi\mathfrak{p}_{\psi,\kappa,xx}(\mathbb{R}^2) &= \int_{\mathbb{R}^2}\frac{|W_{\psi,\kappa,x}(f,t)|^2}{|f|}\,d(f,t)\\ &= \int_{\mathbb{R}^2}\frac{|\hat{x}(\xi)|^2\,\mathring{\psi}\left(\frac{f}{\xi\kappa}\right)^2}{|f|}\,d(f,\xi) \qquad \text{by (7) and (19)}\\ &= \int_{\mathbb{R}}|\hat{x}(\xi)|^2\int_{\mathbb{R}}\frac{\mathring{\psi}(r)^2}{|r|}\,dr\,d\xi \qquad \text{by substitution } f=\xi\kappa r\\ &= C_\psi\int_{\mathbb{R}}|\hat{x}(\xi)|^2\,d\xi\\ &= C_\psi\int_{\mathbb{R}}|x(t)|^2\,dt \qquad\qquad \text{by (15)/(19).}\end{aligned}$$

One then obtains (32) by repeating the calculation with (17) in place of (19), with the validity of the steps justified by the fact that $(f,t)\mapsto\frac{W_{\psi,\kappa,x}(f,t)}{\sqrt{|f|}}$ and $(f,t)\mapsto\frac{W_{\psi,\kappa,y}(f,t)}{\sqrt{|f|}}$ are square-integrable as proved above. □

Of the differences between the Fourier cross-energy spectrum and the wavelet cross-energy spectrum, one that we emphasise is that due to Eqs. (27) and (30)/(31), the wavelet cross-energy spectrum $\mathfrak{p}_{\psi,\kappa,xy}$ is able to take into account the level of *coherence* of phases over time, by which we mean how constant the phase difference $\phi_x(f,t) - \phi_y(f,t)$ associated to each frequency f remains over time. This is particularly useful when analysing signals recorded from open systems, where oscillatory components have time-dependent characteristics and there will be various time-dependent background effects picked up in the signals being measured. Detailed discussion of coherence of wavelet phases and its quantification is given in Appendix B.1.

2.5. Fourier bispectra of deterministic signals

Throughout this paper, the abbreviation "BD" stands for "bispectral density", and "autoBD" and "cross-BD" are to be interpreted "autobispectral density" and "cross-bispectral density".

For any functions $x, y, z\colon \mathbb{R} \to \mathbb{R}$ with Fourier transforms $\hat{x}, \hat{y}, \hat{z}\colon \mathbb{R} \to \mathbb{C}$ (defined at least up to almost-everywhere equality), define

$$\mathcal{B}_{xyz}(f_1, f_2) \;=\; \hat{x}(f_1)\hat{y}(f_2)\overline{\hat{z}(f_1+f_2)} \;=\; \hat{x}(f_1)\hat{y}(f_2)\hat{z}(-(f_1+f_2)). \tag{34}$$

Note the symmetries $\mathcal{B}_{yxz}(f_1,f_2) = \mathcal{B}_{xyz}(f_2,f_1)$ and $\mathcal{B}_{xzy}(f_1,f_2) = \mathcal{B}_{xyz}(f_1,-(f_1+f_2))$ for all $f_1, f_2 \in \mathbb{R}$. The definition of $\mathcal{B}_{xyz}(f_1,f_2)$ makes reference to three frequencies, namely the two input frequencies f_1 and f_2 and their sum f_1+f_2. As shown in Fig. 3, we define the following 6 regions of $\mathbb{R}^2$ according to the signs of these three frequencies:

$$\begin{array}{ll} \Gamma_1 = \{f_1, f_2 > 0\} & -\Gamma_1 = \{f_1, f_2 < 0\} \\ \Gamma_2 = \{0 < f_1 < -f_2\} & -\Gamma_2 = \{0 > f_1 > -f_2\} \\ \Gamma_3 = \{0 < f_2 < -f_1\} & -\Gamma_3 = \{0 > f_2 > -f_1\}. \end{array}$$

It will also be useful to define the subset $\Gamma_{1-} := \{0 < f_2 < f_1\}$ of Γ_1. We refer to the two-dimensional space of inputs (f_1, f_2) of the function $\mathcal{B}_{xyz}$ as "frequency-frequency space". We also define *projective frequency-frequency space* as the set of all straight lines through the origin $(0,0)$. For our purposes, projective frequency-frequency space is naturally parametrised by a value $\lambda \in \hat{\mathbb{R}}$ corresponding to $\frac{f_1}{f_1+f_2}$. With this parametrisation, we have that

- $\Gamma_1 \cup -\Gamma_1$ is covered by $\lambda \in (0,1)$;
- $\Gamma_2 \cup -\Gamma_2$ is covered by $\lambda \in (-\infty, 0)$;
- $\Gamma_3 \cup -\Gamma_3$ is covered by $\lambda \in (1, \infty)$.

Now by iterated application of the convolution theorem, one obtains that for all "sufficiently controlled" functions $x, y, z\colon \mathbb{R} \to \mathbb{R}$ (e.g. if these belong to the Sobolev space $W^{2,1}(\mathbb{R})$), the following "third-order versions" of (14) and (15) hold:

$$\int_{\mathbb{R}} x(t)y(t)z(t)\,dt \;=\; \int_{\mathbb{R}^2} \mathcal{B}_{xyz}(f_1,f_2)\,d(f_1,f_2) \tag{35}$$

$$\;=\; 2\,\mathrm{Re}\Big(\int_{(0,\infty)^2} \mathcal{B}_{xyz}(f_1,f_2) + \mathcal{B}_{yzx}(f_1,f_2) + \mathcal{B}_{zxy}(f_1,f_2)\,d(f_1,f_2) \Big), \tag{36}$$

and in particular,

$$\int_{\mathbb{R}} x(t)^3\,dt \;=\; \int_{\mathbb{R}^2} \mathcal{B}_{xxx}(f_1,f_2)\,d(f_1,f_2) \;=\; 12\,\mathrm{Re}\left(\int_{\Gamma_{1-}} \mathcal{B}_{xxx}(f_1,f_2)\,d(f_1,f_2)\right).$$

The functions $\mathcal{B}_{uvw}$ with $u,v,w \in \{x,y,z\}$ represent *third-order spectra* or *bispectra*. However, the functions themselves are densities: we refer to each function $\mathcal{B}_{uvw}$ with $u,v,w \in \{x,y,z\}$ as a *Fourier BD*, and we refer to $\mathcal{B}_{xxx}$ as the *Fourier autoBD* of x; a function $\mathcal{B}_{uvw}$ for which it is *not* the case that $u = v = w$ is referred to as a *Fourier cross-BD*. Although these functions $\mathcal{B}_{uvw}$ are often referred to simply as bispectra, we regard "bispectra" as referring to the complex-valued measures on frequency-frequency space whose densities are given by the functions $\mathcal{B}_{uvw}$.

Note that it is possible to restrict all bispectral analysis to positive frequencies only:[3]

$$\left\{\begin{array}{lcl}
(f_1,f_2) \in -\Gamma_1 & \Rightarrow & \mathcal{B}_{xyz}(f_1,f_2) = \overline{\mathcal{B}_{xyz}(|f_1|,|f_2|)} \\
(f_1,f_2) \in \Gamma_2 & \Rightarrow & \mathcal{B}_{xyz}(f_1,f_2) = \mathcal{B}_{xzy}(f_1,|f_2|-f_1) \\
(f_1,f_2) \in -\Gamma_2 & \Rightarrow & \mathcal{B}_{xyz}(f_1,f_2) = \overline{\mathcal{B}_{xzy}(|f_1|,f_2-|f_1|)} \\
(f_1,f_2) \in \Gamma_3 & \Rightarrow & \mathcal{B}_{xyz}(f_1,f_2) = \mathcal{B}_{zyx}(|f_1|-f_2,f_2) \\
(f_1,f_2) \in -\Gamma_3 & \Rightarrow & \mathcal{B}_{xyz}(f_1,f_2) = \overline{\mathcal{B}_{zyx}(f_1-|f_2|,|f_2|)}.
\end{array}\right. \tag{37}$$

With the symmetries of the bispectrum, the complete list of possible cross-bispectra given a pair of signals or a triplet of signals is as follows:

- For a pair of signals x and y, there are essentially 4 different bispectra ($\mathcal{B}_{xxx}$, $\mathcal{B}_{xxy}$, $\mathcal{B}_{xyy}$, $\mathcal{B}_{yyy}$), or 6 if we restrict to positive frequencies ($\mathcal{B}_{xxx}$, $\mathcal{B}_{xxy}$, $\mathcal{B}_{xyx}$, $\mathcal{B}_{xyy}$, $\mathcal{B}_{yyx}$, $\mathcal{B}_{yyy}$).
- For a triplet of signals x, y and z, there are essentially 10 different bispectra ($\mathcal{B}_{xxx}$, $\mathcal{B}_{xxy}$, $\mathcal{B}_{xxz}$, $\mathcal{B}_{xyy}$, $\mathcal{B}_{xyz}$, $\mathcal{B}_{xzz}$, $\mathcal{B}_{yyy}$, $\mathcal{B}_{yyz}$, $\mathcal{B}_{yzz}$, $\mathcal{B}_{zzz}$), or 18 if we restrict to positive frequencies ($\mathcal{B}_{xxx}$, $\mathcal{B}_{xxy}$, $\mathcal{B}_{xxz}$, $\mathcal{B}_{xyx}$, $\mathcal{B}_{xyy}$, $\mathcal{B}_{xyz}$, $\mathcal{B}_{xzx}$, $\mathcal{B}_{xzy}$, $\mathcal{B}_{xzz}$, $\mathcal{B}_{yyx}$, $\mathcal{B}_{yyy}$, $\mathcal{B}_{yyz}$, $\mathcal{B}_{yzx}$, $\mathcal{B}_{yzy}$, $\mathcal{B}_{yzz}$, $\mathcal{B}_{zzx}$, $\mathcal{B}_{zzy}$, $\mathcal{B}_{zzz}$).

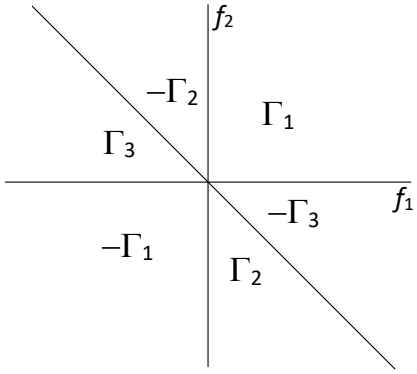


Figure 3: The six regions in frequency-frequency space defined by the signs of the three frequencies under consideration by the bispectrum.

[3]This fact does not hold for higher-than-third-order spectral analysis of continuous-time signals.

Finally, note that Eq. (35) can be extended to complex-valued signals as

$$\int_{\mathbb{R}} x(t)y(t)z(t)\,dt \;=\; \int_{\mathbb{R}^2} \hat{x}(f_1)\hat{y}(f_2)\hat{z}(-(f_1+f_2))\,d(f_1,f_2), \tag{38}$$

which is the third-order analogue of (18).

3. Wavelet bispectra

As in Sec. 2.5, the abbreviation "BD" stands for "bispectral density", and "autoBD" and "cross-BD" are to be interpreted "autobispectral density" and "cross-bispectral density" respectively.

In this section, we will see how a "perfect" definition of the wavelet bispectrum is not achievable, but how we can nonetheless provide a "suitable" definition that fulfils the needs of time-evolving bispectral analysis.

The difficulty in defining wavelet bispectral density

We seek to obtain a "third-order analogue" of the second-order wavelet spectrum $\mathfrak{p}_{\psi,\kappa,xy}$. As described in Sec. 2.2, the difficulty in achieving this lies in the contrast between linear sum of frequencies inherent to bispectral analysis and the logarithmic treatment of frequencies inherent to wavelet analysis. We will now explain this in detail.

We first consider how to define the *windowed-Fourier-transform bispectrum* [46]. Let w be a window function as defined in Sec. 2.2, which for simplicity we will assume to belong to the Schwartz space $\mathcal{S}(\mathbb{R},\mathbb{R})$, and assume additionally that

$$K_w := \int_{\mathbb{R}} w(t)^3\,dt > 0.$$

(Window functions w used in practice typically fulfil the stronger condition that either $w \geq 0$ everywhere or $\hat{w} \geq 0$ everywhere.) We use the same notation for the windowed Fourier transform as in Sec. 2.2. We present two ways to derive the definition of *windowed-Fourier-transform bispectral density* (also called a *bispectrogram*):

- In the first instance, we consider finite-energy signals (which, for simplicity's sake, we will assume to be Schwartz functions).
- In the second instance, we consider infinite-time sinusoidal signals.

Consider signals $x, y, z \in \mathcal{S}(\mathbb{R},\mathbb{R})$. In analogy to (7), we have

$$(\mathcal{F}F_{w,x}(f,\cdot))(\xi) \;=\; \hat{x}(\xi)\hat{w}(f-\xi) \tag{39}$$

for all $\xi, f \in \mathbb{R}$. Consequently, following analogous steps to the proof of Proposition 2 (and noting

that $\hat{z}$ and $F_{w,z}(\cdot,t)$ are Hermitian), we have

$$\begin{aligned}
&\int_{\mathbb{R}^3} F_{w,x}(f_1,t)F_{w,y}(f_2,t)\overline{F_{w,z}(f_1+f_2,t)}\,d(f_1,f_2,t)\\
&= \int_{\mathbb{R}^4} \hat{x}(\xi_1)\hat{w}(f_1-\xi_1)\hat{y}(\xi_2)\hat{w}(f_2-\xi_2)\overline{\hat{z}(f_1+f_2)}\hat{w}(f_1+f_2-(\xi_1+\xi_2))\,d(f_1,f_2,\xi_1,\xi_2)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{by (39) and (38)}\\
&= \int_{\mathbb{R}^2} \mathcal{B}_{xyz}(\xi_1,\xi_2)\int_{\mathbb{R}^2}\hat{w}(\nu_1)\hat{w}(\nu_2)\hat{w}(\nu_1+\nu_2)\,d(\nu_1,\nu_2)\,d(\xi_1,\xi_2)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{by substitution } f_i=\xi_i+\nu_i\\
&= K_w\int_{\mathbb{R}^2}\mathcal{B}_{xyz}(\xi_1,\xi_2)\,d(\xi_1,\xi_2)\qquad \text{by (35)/(38)}\\
&= K_w\int_{\mathbb{R}} x(t)y(t)z(t)\,dt\qquad\qquad \text{by (35)/(38).}
\end{aligned}$$

Therefore we define the windowed-Fourier-transform bispectral density as

$$\text{WFT-BD}_{w,x,y,z}(f_1,f_2,t) \;=\; K_w^{-1}F_{w,x}(f_1,t)F_{w,y}(f_2,t)\overline{F_{w,z}(f_1+f_2,t)} \tag{40}$$

for $f_1,f_2,t\in\mathbb{R}$. Now instead consider sinusoidal input signals

$$\begin{aligned}
x(t) &= A_1\cos(2\pi\nu_1 t+\phi_1)\\
y(t) &= A_2\cos(2\pi\nu_2 t+\phi_2)\\
z(t) &= A_3\cos(2\pi(\nu_1+\nu_2)t+\phi_1+\phi_2-\theta),
\end{aligned}$$

with $\nu_1,\nu_2>0$. Using (12) one can show that at each $t\in\mathbb{R}$,

$$\int_{\mathbb{R}^2} F_{w,x}(f_1,t)F_{w,y}(f_2,t)\overline{F_{w,z}(f_1+f_2,t)}\,d(f_1,f_2) \;=\; \tfrac{1}{4}K_w A_1A_2A_3\cos(\theta)+c(t) \tag{41}$$

where $c(t)$ is a finite sum of zero-mean circular motions in $\mathbb{C}$. And by expressing x,y,z in terms of complex exponentials, one can easily check that the average value of the function $t\mapsto x(t)y(t)z(t)$ is indeed $\frac{1}{4}A_1A_2A_3\cos(\theta)$. Thus, consideration of sinusoidal input signals also justifies the definition (40) for windowed-Fourier-transform bispectral density.

The windowed-Fourier-transform bispectrum for input functions $x,y,z\in L^\infty(\mathbb{R},\mathbb{R})$ is then defined by integrating the density, i.e. it is the map

$$A\mapsto K_w^{-1}\int_A F_{w,x}(f_1,t)F_{w,y}(f_2,t)\overline{F_{w,z}(f_1+f_2,t)}\,d(f_1,f_2,t)$$

on all sets $A\in\mathcal{B}(\mathbb{R}^3)$ over which the integrand is absolutely integrable.

Having defined the windowed-Fourier-transform bispectral density, through two approaches that produce the same result, let us now consider how the analogous calculations would look for attempting to define *wavelet bispectral density*. More specifically, at this point we will only consider the calculation for finite-energy signals; the case of sinusoidal signals is treated in Sec. 3.1.

For simplicity, take $\kappa = 1$ and assume that $\psi \in \mathcal{S}(\mathbb{R})$. For signals $x, y, z \in \mathcal{S}(\mathbb{R}, \mathbb{R})$, we have

$$\begin{aligned}
&\int_{\mathbb{R}^3} \frac{W_{\psi,1,x}(f_1,t) W_{\psi,1,y}(f_2,t) \overline{W_{\psi,1,z}(f_1+f_2,t)}}{f_1 f_2}\, d(f_1,f_2,t) \\
&= \int_{\mathbb{R}^4} \frac{\hat{x}(\xi_1)\mathring{\psi}(\frac{f_1}{\xi_1})\hat{y}(\xi_2)\mathring{\psi}(\frac{f_2}{\xi_2})\overline{\hat{z}(f_1+f_2)}\mathring{\psi}(\frac{f_1+f_2}{\xi_1+\xi_2})}{f_1 f_2}\, d(f_1,f_2,\xi_1,\xi_2) \\
&= \int_{\mathbb{R}^2} \mathcal{B}_{xyz}(\xi_1,\xi_2) \int_{\mathbb{R}^2} \frac{\mathring{\psi}(\frac{f_1}{\xi_1})\mathring{\psi}(\frac{f_2}{\xi_2})\mathring{\psi}(\frac{f_1+f_2}{\xi_1+\xi_2})}{f_1 f_2}\, d(f_1,f_2)\, d(\xi_1,\xi_2).
\end{aligned}$$

If the term $\frac{f_1+f_2}{\xi_1+\xi_2}$ were instead $\frac{f_1}{\xi_1} + \frac{f_2}{\xi_2}$, then as in the proof of Proposition 2 the change of variables $r_i = \frac{f_i}{\xi_i}$ $(i = 1, 2)$ would cause the above expression to simplify to

$$C_\psi^{(3)} \int_{\mathbb{R}} x(t) y(t) z(t)\, dt$$

where

$$C_\psi^{(3)} \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathring{\psi}(e^{r_1})\mathring{\psi}(e^{r_2})\mathring{\psi}(e^{r_1}+e^{r_2})\, dr_1 dr_2 \;=\; \int_0^{\infty}\int_0^{\infty} \frac{\mathring{\psi}(r_1)\mathring{\psi}(r_2)\mathring{\psi}(r_1+r_2)}{r_1 r_2}\, dr_1 dr_2.$$

Thus it is the discrepancy between $\frac{f_1+f_2}{\xi_1+\xi_2}$ and $\frac{f_1}{\xi_1} + \frac{f_2}{\xi_2}$ that prevents the calculation from going through.

How to define a "reasonable" notion of wavelet bispectra

Although we cannot obtain a "perfect" third-order analogue of the second-order wavelet spectra, nonetheless we can still obtain somewhat "reasonable" notions of a wavelet bispectrum.

A valid candidate for a definition of the "wavelet bispectrum" is one expressible through integration of a formula for "wavelet bispectral density", whose value at (f_1, f_2, t) depends on the input signals only via values of their wavelet transforms at time t. This implies in particular that the wavelet bispectrum can serve as a time-localised measure of the bispectral content of signals with time-dependent oscillatory characteristics.

Of course, the "reasonableness" of such a candidate definition of the wavelet bispectrum is a rather vague question, but a natural way of assessing it is by the following (also somewhat vague) question:

For signals x, y and z that include prominent sinusoidal contributions $A_1 \cos(2\pi\nu_1 t + \phi_1)$, $A_2 \cos(2\pi\nu_2 t + \phi_2)$ and $A_3 \cos(2\pi(\nu_1+\nu_2)t + \phi_1 + \phi_2 - \theta)$ respectively,

- *is the bispectral contribution within positive frequency-frequency space made by this triplet of sinusoidal components over a time-interval of duration ΔT approximately equal to $\frac{1}{8} A_1 A_2 A_3 \Delta T e^{i\theta}$ [or some other pre-determined constant multiple of $A_1 A_2 A_3 \Delta T e^{i\theta}$], and*
- *is this bispectral contribution concentrated, in due proportion with the frequency resolution of the underlying wavelet ψ, around (ν_1, ν_2)?*

A valid candidate for the definition of the wavelet bispectrum is considered to be a "reasonable" definition in so far as it provides an affirmative answer to the above question.

We go on to give a definition of the wavelet bispectrum whose "reasonableness" is formalised in Theorem 8, in a limit as frequency resolution tends to ∞. Our numerics in Sec. 5.1 will indicate that the frequency resolution needed for "reasonable" results is not high at all.

3.1. Definition of the wavelet bispectrum (global version)

Given a wavelet function ψ as defined in Sec. 2.1, for each $\lambda \in [0,1]$, define

$$D_\psi(\lambda) := \int_0^\infty\int_0^\infty \frac{\mathring{\psi}(r_1)\mathring{\psi}(r_2)\mathring{\psi}(\lambda r_1 + (1-\lambda)r_2)}{r_1 r_2}\, dr_1 dr_2 \tag{42}$$

$$= \int_{-\infty}^\infty\int_{-\infty}^\infty \mathring{\psi}(e^{r_1})\mathring{\psi}(e^{r_2})\mathring{\psi}(\lambda e^{r_1} + (1-\lambda)e^{r_2})\, dr_1 dr_2. \tag{43}$$

Note that $D_\psi(\lambda)$ is strictly positive (by fixing $r_1 = r_2 \in \mathring{\psi}^{-1}((0,\infty))$ in the integrand) and bounded above by $\big(\int_0^\infty \frac{\hat{\psi}(r)}{r}\, dr\big)^2 \max \hat{\psi}(\cdot)$. It serves as a kind of third-order analogue of C_ψ by fulfilling the following third-order analogue of (22b): Given a wavelet function ψ, a value $\kappa > 0$ and signals

$$\begin{aligned} x(t) &= A_1 \cos(2\pi\lambda\nu t + \phi_1) \\ y(t) &= A_2 \cos(2\pi(1-\lambda)\nu t + \phi_2) \\ z(t) &= A_3 \cos(2\pi\nu t + \phi_1 + \phi_2 - \theta) \end{aligned}$$

with $\nu > 0$, we have

$$\tfrac{1}{8}A_1A_2A_3e^{i\theta} = \frac{1}{D_\psi(\lambda)}\int_{\mathbb{R}^2} W_{\psi,\kappa,x}(e^{\zeta_1},t)W_{\psi,\kappa,y}(e^{\zeta_2},t)\overline{W_{\psi,\kappa,z}(e^{\zeta_1}+e^{\zeta_2},t)}\, d(\zeta_1,\zeta_2) \tag{44}$$

for any $t \in \mathbb{R}$. This fact is essentially immediate from (8).

Remark 3. If we chose any p-value other than 1 for the wavelet transform in (44), then we would need to change $D_\psi(\lambda)$ to an expression taking the form $\nu^{3(p-1)}E_\psi(\lambda)$ where $E_\psi(\lambda)$ is independent of ν. (This can be derived explicitly but also follows immediately from dimensional analysis.)

At the start of Sec. 3 we showed the difficulty in defining the wavelet bispectrum, using finite-energy input signals. If instead we consider sinusoidal signals as above, then this impossibility of a perfect third-order analogue of the second-order wavelet spectra is now represented by the fact that the factor D_ψ *is not constant but depends on the ratio of the frequencies* as represented by λ. This is in contrast with the windowed-Fourier-transform bispectrum, where the factor K_w is constant. Nonetheless, using $D_\psi(\cdot)$ as defined above, we will define wavelet bispectral density by "pointwise" application of the normalisation D_ψ^{-1}:

Definition 4. Given a wavelet function ψ, a value $\kappa > 0$ and functions $x, y, z \in L^\infty(\mathbb{R},\mathbb{R})$, define the *logarithmic-frequency wavelet BD* $b_{\psi,\kappa,xyz}\colon (0,\infty)^2 \times \mathbb{R} \to \mathbb{C}$ by

$$b_{\psi,\kappa,xyz}(f_1,f_2,t) = D_\psi\Big(\tfrac{f_1}{f_1+f_2}\Big)^{-1} W_{\psi,\kappa,x}(f_1,t)W_{\psi,\kappa,y}(f_2,t)\overline{W_{\psi,\kappa,z}(f_1+f_2,t)}. \tag{45}$$

The normalisation factor $D_\psi\Big(\frac{f_1}{f_1+f_2}\Big)$ can be expressed as

$$D_\psi\Big(\tfrac{f_1}{f_1+f_2}\Big) = \int_0^\infty\int_0^\infty \frac{\hat{\psi}\big(\frac{f_1}{\xi_1}\big)\hat{\psi}\big(\frac{f_2}{\xi_2}\big)\hat{\psi}\big(\frac{f_1+f_2}{\xi_1+\xi_2}\big)}{\xi_1\xi_2}\, d\xi_1 d\xi_2. \tag{46}$$

We refer to $b_{\psi,\kappa,xxx}$ as the *logarithmic-frequency wavelet autoBD* of x. As in Sec. 2.5, a BD $b_{\psi,\kappa,uvw}$ for which it is *not* the case that $u = v = w$ is referred to as a cross-BD.

One can convert these densities to linear-frequency densities: namely, we define $\tilde{b}_{\psi,\kappa,xyz}\colon (0,\infty)^2\times\mathbb{R}\to\mathbb{C}$ by

$$\tilde{b}_{\psi,\kappa,xyz}(f_1,f_2,t) \;=\; \frac{b_{\psi,\kappa,xyz}(f_1,f_2,t)}{f_1 f_2}.$$

For any $A\in\mathcal{B}_{\tilde{b}_{\psi,\kappa,xyz}}\subset\mathcal{B}((0,\infty)^2\times\mathbb{R})$, we define

$$\mathfrak{b}_{\psi,\kappa,xyz}(A) \;=\; \int_{\mathbb{R}^3}\mathbb{1}_A(e^{\zeta_1},e^{\zeta_2},t)b_{\psi,\kappa,xyz}(e^{\zeta_1},e^{\zeta_2},t)\,d(\zeta_1,\zeta_2,t) \;=\; \int_A \tilde{b}_{\psi,\kappa,xyz}(f_1,f_2,t)\,d(f_1,f_2,t). \tag{47}$$

We refer to $\mathfrak{b}_{\psi,\kappa,xyz}$ as the *wavelet bispectrum*. (The terminology regarding wavelet autobispectra and wavelet cross-bispectra is analogous to the above terminologies for autoBD and cross-BD.) For any compact $K\subset(0,\infty)^2\times\mathbb{R}$ the restriction of $\mathfrak{b}_{\psi,\kappa,xyz}$ to $\mathcal{B}(K)$ defines a complex-valued measure on K.

Remark 5. Since we constructed our definition of wavelet bispectra on the basis of logarithmic-frequency consideration as reflected in (44), we have not incorporated negative frequencies in our definition. This is not problematic or restrictive, since in general, all bispectral analysis can be restricted to positive frequencies (see Sec. 2.5).

Remark 6. The integral in (46), as well as the integral in (42) and the integral in (43), can be computed by integrating only below the diagonal and then doubling the result. Note also that $D_\psi(\lambda)=D_\psi(1-\lambda)$, or equivalently that the expression $D_\psi\Big(\frac{f_1}{f_1+f_2}\Big)$ is symmetric in f_1 and f_2. It follows in particular that the wavelet autobispectral density $b_{\psi,\kappa,xxx}$ or $\tilde{b}_{\psi,\kappa,xxx}$ is symmetric in its frequency inputs.

Remark 7. The dependence of $D_\psi(\lambda)$ on λ is continuous, and hence in particular $\{D_\psi(\lambda)\}_{\lambda\in[0,1]}$ is bounded away from 0. This follows from the dominated convergence theorem since the integrand in (42) is dominated by the integrable function

$$(r_1,r_2)\;\mapsto\;\Big(\max_{r>0}\hat{\psi}(r)\Big)\cdot\frac{\mathring{\psi}(r_1)}{r_1}\cdot\frac{\mathring{\psi}(r_2)}{r_2}.$$

Fig. 4 illustrates the dependence of $D_{\psi_\sigma}(\lambda)$ on λ for the "lognormal" wavelet function ψ_σ defined by $\hat{\psi}_\sigma(r)=e^{-2(\pi\sigma\log r)^2}$ on $r>0$, where the parameter σ is taken over the range $[0.5,3]$. (Lognormal wavelets are discussed further at the start of Sec. 5.)

Now using the same wavelet phase notation as was used in Eq. (27), we have that

$$\frac{b_{\psi,\kappa,xyz}(f,t)}{|b_{\psi,\kappa,xy}(f,t)|} \;=\; e^{i(\phi_x(f_1,t)+\phi_y(f_2,t)-\phi_z(f_1+f_2,t))}. \tag{48}$$

We refer to the argument $\phi_x(f_1,t)+\phi_y(f_2,t)-\phi_z(f_1+f_2,t)$ of $b_{\psi,\kappa,xyz}(f,t)$ as the *biphase* associated to the frequency pair (f_1,f_2) at time t. Of the differences between Fourier bispectra and wavelet bispectra, one that we emphasise is that due to Eqs. (48) and (47), the wavelet bispectrum $\mathfrak{b}_{\psi,\kappa,xyz}$ is able to take into account how constant the biphase associated to each frequency pair remains over time. Just as described in Sec. 2.4 for second-order wavelet cross-spectra, this is particularly useful when analysing signals recorded from open systems. Constancy of biphase is referred to as *bicoherence*. Detailed discussion of bicoherence and its quantification is given in Appendix B.2.

Usage of our new definition of the wavelet bispectrum is illustrated in Sec. 5.2 and Sec. 6.

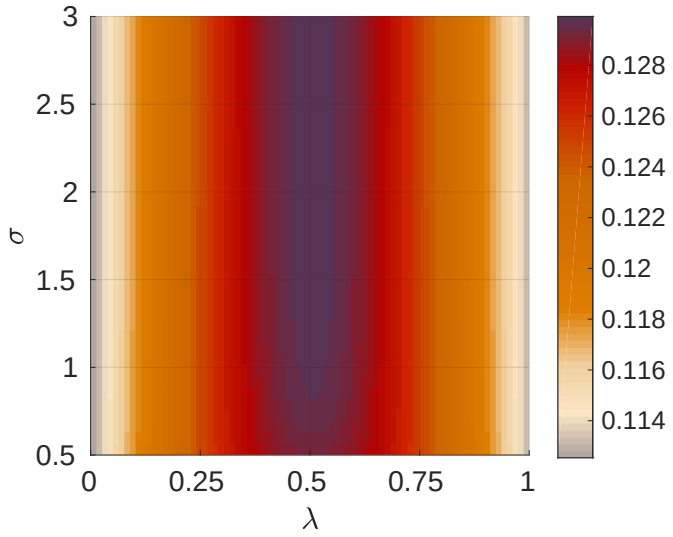


Figure 4: Depedence of $D_{\psi_\sigma}(\lambda)$ on λ for the lognormal wavelet function ψ_σ, across varying σ. The plot specifically shows $\sigma^2 D_{\psi_\sigma}(\lambda)$ as a function of σ and λ. For each σ, we see that $D_{\psi_\sigma}(\lambda)$ ranges from about $0.11\sigma^{-2}$ to about $0.13\sigma^{-2}$.

3.2. Theoretical justification for the wavelet bispectral density formula (45)

In this section, we consider the wavelet bispectrum of a triplet of signals (any two of which or all three of which could be the same), each consisting of a finite sum of sinusoids. We work with a one-parameter family of wavelets whose parameter σ represents the "frequency resolution" (or inversely represents "frequency uncertainty"), and consider the limit as $\sigma \to \infty$. So, in principle, the setup we consider here is

$$\begin{aligned} x(t) &= \sum_{i=1}^{N_1} A_{i1} \cos(2\pi\nu_{i1} t + \phi_{i1}) \\ y(t) &= \sum_{i=1}^{N_2} A_{i2} \cos(2\pi\nu_{i2} t + \phi_{i2}) \\ z(t) &= \sum_{i=1}^{N_3} A_{i3} \cos(2\pi\nu_{i3} t + \phi_{i3}) \\ \hat{\psi}(r) &= \begin{cases} g(r^\sigma) & r > 0 \\ 0 & r \leq 0 \end{cases} \end{aligned}$$

for some $g\colon (0,\infty) \to [0,\infty)$, and (taking $\kappa = 1$ without loss of generality) we consider the limiting behaviour of $\mathfrak{b}_{\psi,1,xyz}$ as $\sigma \to \infty$.

However, one slight issue is that for a given function g, even if the inverse Fourier transform of the function

$$g_{\text{ext}} = \begin{cases} g(r) & r > 0 \\ 0 & r \leq 0 \end{cases}$$

exists and fulfils all the conditions of being a wavelet function (as defined in Sec. 2.1), this might not guarantee that the inverse Fourier transform of $\hat{\psi}$ is an L^1 function for all σ, and therefore

might not guarantee that the wavelet transforms $W_{\psi,1,x}$, $W_{\psi,1,y}$ and $W_{\psi,1,z}$ as given by (6) are well-defined for all large σ. Therefore, in the result below we consider the "wavelet transform" defined by (8) rather than (6).

Theorem 8. *Fix a continuous function $g\colon (0,\infty) \to [0,\infty)$ (other than the constant zero function) such that the "finite uncertainty" condition $\int_0^\infty \frac{|\log r| g(r)}{r}\, dr < \infty$ holds and $\log(r)^2 g(r) \to 0$ as r tends to 0 and to ∞.*[4] *For each $\sigma \geq 1$, define $g_\sigma\colon (0,\infty) \to [0,\infty)$ by $g_\sigma(r) = g(r^\sigma)$.*

Now for each $j \in \{1,2,3\}$, fix a positive integer N_j, values $A_{1j}, \ldots, A_{N_j j} \neq 0$, values $\phi_{1j}, \ldots, \phi_{N_j j} \in \mathbb{R}$, and distinct values $\nu_{1j}, \ldots, \nu_{N_j j} > 0$.[5] *For each $\sigma \geq 1$ and $j \in \{1,2,3\}$, define*

$$W_{\sigma,j}(f,t) \;=\; \tfrac{1}{2} \sum_{i=1}^{N_j} g_\sigma\Big(\tfrac{\nu_{ij}}{f}\Big) A_{ij} e^{i(2\pi\nu_{ij} t + \phi_{ij})}$$

for all $f > 0$ and $t \in \mathbb{R}$. For each $\sigma \geq 1$, define $D_\sigma\colon (0,1) \to (0,\infty)$ by

$$D_\sigma\Big(\tfrac{f_1}{f_1+f_2}\Big) \;=\; \int_0^\infty \int_0^\infty \frac{g_\sigma\big(\frac{f_1}{\xi_1}\big) g_\sigma\big(\frac{f_2}{\xi_2}\big) g_\sigma\big(\frac{f_1+f_2}{\xi_1+\xi_2}\big)}{\xi_1 \xi_2}\, d\xi_1 d\xi_2$$

with $f_1, f_2 > 0$, and define

$$b_\sigma(f_1,f_2,t) \;=\; D_\sigma\Big(\tfrac{f_1}{f_1+f_2}\Big)^{-1} W_{\sigma,1}(f_1,t) W_{\sigma,2}(f_2,t) \overline{W_{\sigma,3}(f_1+f_2,t)}$$

for all $f_1, f_2 > 0$ and $t \in \mathbb{R}$. Let

$$P \;=\; \{(n_1,n_2,n_3) \in \{1,\ldots,N_1\} \times \{1,\ldots,N_2\} \times \{1,\ldots,N_3\} \,:\, \nu_{n_1 1} + \nu_{n_2 2} = \nu_{n_3 3}\}.$$

(I) For any $S \subset (0,\infty)^2$ such that $(\nu_{n_1 1}, \nu_{n_2 2}) \notin \bar{S}$ for all $(n_1,n_2,n_3) \in P$, we have $b_\sigma(f_1,f_2,t) \to 0$ uniformly across $(f_1,f_2,t) \in S \times \mathbb{R}$ as $\sigma \to \infty$. (II) For each σ,

$$\int_{-\infty}^\infty \int_{-\infty}^\infty \sup_{t \in \mathbb{R}} |b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)|\, d\zeta_1 d\zeta_2 \;<\; \infty\,;$$

and for any $S \in \mathcal{B}((0,\infty)^2)$ such that $(\nu_{n_1 1}, \nu_{n_2 2}) \notin \partial S$ for all $(n_1,n_2,n_3) \in P$, we have

$$\int_{\mathbb{R}^2} \mathbb{1}_S(e^{\zeta_1}, e^{\zeta_2}) b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)\, d(\zeta_1,\zeta_2) \;\to \sum_{\substack{(n_1,n_2,n_3) \in P \\ \text{with } (\nu_{n_1 1}, \nu_{n_2 2}) \in S}} \tfrac{1}{8} A_{n_1 1} A_{n_2 2} A_{n_3 3} e^{i(\phi_{n_1 1} + \phi_{n_2 2} - \phi_{n_3 3})}$$

uniformly across $t \in \mathbb{R}$ as $\sigma \to \infty$.

The proof is given in Sec. 4.

Remark 9. In Theorem 8, the limit as frequency uncertainty tends to zero is only taken *within* a one-parameter family $(g_\sigma)_{\sigma \geq 1}$. The natural next question is whether some condition on a general sequence of wavelet functions $(\psi_n)_{n \in \mathbb{N}}$ can be found that guarantees the same conclusions as in Theorem 8. We conjecture that under some suitable quantitative definitions of frequency uncertainty and time-frequency uncertainty for wavelet functions, the conclusions of Theorem 8 will hold for any sequence $(\psi_n)_{n \in \mathbb{N}}$ of wavelet functions for which the frequency uncertainty tends to zero while the time-frequency uncertainty remains bounded.

[4] If g is unimodal then the finite uncertainty condition automatically implies that $\log(r)^2 g(r) \to 0$ as r tends to 0 and to ∞.

[5] We emphasise that the sets $\{\nu_{i1}\}_{i=1}^{N_1}$, $\{\nu_{i2}\}_{i=1}^{N_2}$ and $\{\nu_{i3}\}_{i=1}^{N_3}$ need not be disjoint.

3.3. Frequency-localised bispectra

Suppose we have a region of frequency-frequency space $S \subset (0,\infty)\times(0,\infty)$ contained within a strip of the form

$$\{(f_1,f_2) : \tfrac{f_1}{f_1+f_2} \in [\lambda-\varepsilon,\lambda+\varepsilon]\}$$

for some fixed $\lambda \in (0,1)$. One may wish to consider the *local logarithmic-frequency wavelet BD* $b_{\psi,\kappa,xyz;\lambda}\colon S\times\mathbb{R}\to\mathbb{C}$ defined by

$$b_{\psi,\kappa,xyz;\lambda}(f_1,f_2,t) \;=\; D_\psi(\lambda)^{-1} W_{\psi,\kappa,x}(f_1,t) W_{\psi,\kappa,y}(f_2,t)\overline{W_{\psi,\kappa,z}(f_1+f_2,t)}. \tag{49}$$

The point here is that, in contrast to the "global" definition (45), we do not vary the normalisation factor $D_\psi(\lambda)^{-1}$ in accordance with the input frequency pair (f_1,f_2). We then define

$$\mathfrak{b}_{\psi,\kappa,xyz;\lambda}(A) \;=\; \int_{\mathbb{R}^3} \mathbb{1}_A(e^{\zeta_1},e^{\zeta_2},t) b_{\psi,\kappa,xyz;\lambda}(e^{\zeta_1},e^{\zeta_2},t)\, d(\zeta_1,\zeta_2,t) \tag{50}$$

for any $A \in \mathcal{B}_{\tilde{b}_{\psi,\kappa,xyz}}$ contained in $S\times\mathbb{R}$. Note that the localised bispectrum $\mathfrak{b}_{\psi,\kappa,xyz;\lambda}$ is meaningless if the frequency resolution of the wavelet is so low that S does not capture virtually all the wavelet bispectral content arising from bispectral contributions of interest within S.

Recall that our notion of the wavelet bispectrum was constructed from logarithmic-frequency consideration as in (44). Frequency-localised bispectra constructed from analogous linear-frequency consideration will be presented in Appendix C.

If one is not concerned with "absolute" values for bispectral results but only comparisons between bispectral results, then the normalising factor $D_\psi(\lambda)^{-1}$ can be removed from (49). Nonetheless, due to Remark 3, the wavelet transform terms must still be with $p=1$. The unnormalised function $(f_1,f_2,t) \mapsto W_{\psi,\kappa,x}(f_1,t)W_{\psi,\kappa,y}(f_2,t)\overline{W_{\psi,\kappa,z}(f_1+f_2,t)}$ has an additional possible advantage for locating the frequencies of oscillatory contributions: If $\hat{\psi}$ is unimodal with its mode at κ, then for input signals

$$\begin{aligned} x(t) &= A_1\cos(2\pi\nu_1 t+\phi_1)\\ y(t) &= A_2\cos(2\pi\nu_2 t+\phi_2)\\ z(t) &= A_3\cos(2\pi(\nu_1+\nu_2)t+\phi_1+\phi_2-\theta)\end{aligned}$$

with $A_1,A_2,A_3,\nu_1,\nu_2>0$, at any time t the mode of

$$(f_1,f_2)\mapsto |W_{\psi,\kappa,x}(f_1,t)W_{\psi,\kappa,y}(f_2,t)W_{\psi,\kappa,z}(f_1+f_2,t)|$$

will occur *precisely* at (ν_1,ν_2). (This of course does not hold for any $p\neq 1$.) Moreover, if we define the normalised version

$$b'_{\psi,\kappa,xyz}(f_1,f_2,t) \;=\; 8\hat{\psi}(\kappa)^{-3} W_{\psi,\kappa,x}(f_1,t)W_{\psi,\kappa,y}(f_2,t)\overline{W_{\psi,\kappa,z}(f_1+f_2,t)}, \tag{51}$$

then at any time t the value $b'_{\psi,\kappa,xyz}(\nu_1,\nu_2,t)$ – whose magnitude is the peak value of $|b'_{\psi,\kappa,xyz}|$ – is precisely equal to $A_1A_2A_3e^{i\theta}$.

4. Proof of Theorem 8

Define $h(r) = g(\frac{1}{r})$ and $h_\sigma(r) = g_\sigma(\frac{1}{r})$. So

$$D_\sigma(\lambda) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_\sigma(e^{r_1})h_\sigma(e^{r_2})h_\sigma(\lambda e^{r_1}+(1-\lambda)e^{r_2})\,dr_1dr_2$$

for all $\lambda \in (0,1)$, and we can extend this also to $\lambda \in \{0,1\}$. For convenience we also write $N_\sigma(f_1,f_2) := D_\sigma\Big(\frac{f_1}{f_1+f_2}\Big)$. As in Remark 7, $\{D_\sigma(\lambda)\}_{\lambda\in[0,1]}$ is bounded away from 0. Also define

$$D'(\lambda) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(e^{r_1})h(e^{r_2})h(e^{\lambda r_1+(1-\lambda)r_2})\,dr_1dr_2$$

for all $\lambda \in [0,1]$. Taking $r_1 = r_2 \in \log(h^{-1}(0,\infty))$ in the integrand yields that $D'(\lambda) \neq 0$. Since h is bounded and $\int_{-\infty}^{\infty} h(e^r)\,dr < \infty$, the dominated convergence theorem gives that $\lambda \mapsto D'(\lambda)$ is continuous. Again for convenience write $N'(f_1,f_2) := D'\Big(\frac{f_1}{f_1+f_2}\Big)$.

Lemma 10. *We have*

$$\sigma^2 D_\sigma(\lambda) \;\to\; D'(\lambda)$$

uniformly across $\lambda \in [0,1]$ *as* $\sigma \to \infty$.

Proof. For all λ, we have

$$\sigma^2 D_\sigma(\lambda) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_\sigma(e^{\frac{r_1}{\sigma}})h_\sigma(e^{\frac{r_2}{\sigma}})h_\sigma(\lambda e^{\frac{r_1}{\sigma}}+(1-\lambda)e^{\frac{r_2}{\sigma}})\,dr_1dr_2 \tag{52}$$

$$\;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(e^{r_1})h(e^{r_2})h([\lambda e^{\frac{r_1}{\sigma}}+(1-\lambda)e^{\frac{r_2}{\sigma}}]^\sigma)\,dr_1dr_2. \tag{53}$$

It is easy to show that for any $r_1, r_2 \in \mathbb{R}$,

$$\sigma\log[\lambda e^{\frac{r_1}{\sigma}}+(1-\lambda)e^{\frac{r_2}{\sigma}}] \;\to\; \lambda r_1+(1-\lambda)r_2$$

uniformly across $\lambda \in [0,1]$ as $\sigma \to \infty$, and therefore

$$h([\lambda e^{\frac{r_1}{\sigma}}+(1-\lambda)e^{\frac{r_2}{\sigma}}]^\sigma) \;\to\; h(e^{\lambda r_1+(1-\lambda)r_2})$$

uniformly across $\lambda \in [0,1]$ as $\sigma \to \infty$. Since $r \mapsto h(e^r)$ is integrable and h is bounded, the dominated convergence theorem applied to (53) gives that $\sigma^2 D_\sigma(\lambda) \;\to\; D'(\lambda)$ uniformly across $\lambda \in [0,1]$ as $\sigma \to \infty$. □

Let A be the maximum of the amplitudes A_{ij}, and let $\gamma = \max_{r>0} g(r)$.
(I) Fix $S \subset (0,\infty)^2$ with $(\nu_{n_1 1}, \nu_{n_2 2}) \notin \bar{S}$ for all $(n_1,n_2,n_3) \in P$. Let

$$d \;:=\; \min\{\max(|\log(\nu_{n_1 1})-\log(f_1)|, |\log(\nu_{n_2 2})-\log(f_2)|, |\log(\nu_{n_3 3})-\log(f_1+f_2)|) \;:\; (f_1,f_2)\in S, 1\le n_1\le N_1, 1\le n_2\le N_2, 1\le n_3\le N_3)\}.$$

Then for all $(f_1,f_2,t) \in S\times\mathbb{R}$,

$$|b_\sigma(f_1,f_2,t)| \;\le\; \frac{\frac{1}{8}N_1N_2N_3A^3\gamma^2\max(g_\sigma(e^d),g_\sigma(e^{-d}))}{N_\sigma(f_1,f_2)}.$$

So using Lemma 10, one can find $c > 0$ independent of σ such that for sufficiently large σ, for all $(f_1, f_2, t) \in S \times \mathbb{R}$,

$$|b_\sigma(f_1, f_2, t)| \;\le\; c\sigma^2 \max(g_\sigma(e^d), g_\sigma(e^{-d})).$$

Since $r^2 g(e^r) \to 0$ as $r \to \pm\infty$, the right-hand side tends to 0 as $\sigma \to \infty$.

(II) For each σ, for all $(\zeta_1, \zeta_2, t) \in \mathbb{R}^3$, we have

$$|b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)| \;\le\; \frac{\frac{1}{8}N_3 A\gamma}{N_\sigma(e^{\zeta_1}, e^{\zeta_2})} \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2} g_\sigma(\nu_{i_1 1}e^{-\zeta_1}) g_\sigma(\nu_{i_2 2}e^{-\zeta_2}) \tag{54}$$

$$\le\; \frac{\frac{1}{8}N_3 A\gamma}{\min_{\lambda\in[0,1]} D_\sigma(\lambda)} \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2} g_\sigma(\nu_{i_1 1}e^{-\zeta_1}) g_\sigma(\nu_{i_2 2}e^{-\zeta_2}). \tag{55}$$

The right-hand side of (55) is independent of t and, as a function of (ζ_1, ζ_2), is integrable over $\mathbb{R}^2$. Now to prove the remainder of part (II), in view of part (I) it will be sufficient to show that

(A) there exists a compact set $K \subset \mathbb{R}^2$ such that

$$\int_{\mathbb{R}^2\setminus K} \sup_{t\in\mathbb{R}} |b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)|\, d(\zeta_1, \zeta_2) \;\to\; 0 \quad \text{as} \quad \sigma \to \infty\,;$$

(B) for each $(n_1, n_2, n_3) \in P$ and $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\eta \in (0, \delta)$ there exists $M > 1$ such that if $\sigma > M$ then for all $t \in \mathbb{R}$,

$$\left| \tfrac{1}{8}A_{n_1 1}A_{n_2 2}A_{n_3 3}e^{i(\phi_{n_1 1}+\phi_{n_2 2}-\phi_{n_3 3})} - \int_{\log(\nu_{n_2 2})-\eta}^{\log(\nu_{n_2 2})+\eta}\int_{\log(\nu_{n_1 1})-\eta}^{\log(\nu_{n_1 1})+\eta} b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)\, d\zeta_1 d\zeta_2 \right| \;<\; \varepsilon.$$

(A) Applying Lemma 10 to (54), one can find $c > 0$ independent of σ such that for sufficiently large σ, for all $(\zeta_1, \zeta_2, t) \in \mathbb{R}^3$,

$$|b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)| \;\le\; c\sigma^2 \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2} g_\sigma(\nu_{i_1 1}e^{-\zeta_1}) g_\sigma(\nu_{i_2 2}e^{-\zeta_2}). \tag{56}$$

Pick any value $d > 0$ and define

$$\begin{aligned} a_1 &= \log(\min\{\nu_{11}, \ldots, \nu_{N_1 1}\}) - d \qquad & b_1 &= \log(\max\{\nu_{11}, \ldots, \nu_{N_1 1}\}) + d \\ a_2 &= \log(\min\{\nu_{12}, \ldots, \nu_{N_2 2}\}) - d \qquad & b_2 &= \log(\max\{\nu_{12}, \ldots, \nu_{N_2 2}\}) + d, \end{aligned}$$

and take $K = [a_1, b_1] \times [a_2, b_2]$. Writing

$$G_1 := \int_{-\infty}^{\infty} g(e^r)\, dr \;=\; \sigma \int_{-\infty}^{\infty} g_\sigma(e^r)\, dr \quad \text{and} \quad G_2 := \int_0^{\infty} |r| g(e^r)\, dr \;=\; \sigma^2 \int_0^{\infty} |r| g_\sigma(e^r)\, dr,$$

we obtain from (56) (using Markov's inequality) that

$$\begin{aligned} &\int_{(-\infty, a_1)\times\mathbb{R}} \sup\nolimits_{t\in\mathbb{R}} |b_\sigma(e^{\zeta_1}, e^{\zeta_2}, t)|\, d(\zeta_1, \zeta_2) \\ &\qquad\le\; cN_2\sigma^{-1}G_1G_2 \sum_{i=1}^{N_1} \left((\log(\nu_{i1}) - a_1)^{-1}\right) \\ &\qquad\le\; cN_1N_2\sigma^{-1}G_1G_2 d^{-1} \;\to\; 0 \text{ as } \sigma \to \infty. \end{aligned}$$

One obtains a similar conclusion for the integrals over $(b_1,\infty)\times\mathbb{R}$, $\mathbb{R}\times(-\infty,a_2)$ and $\mathbb{R}\times(b_2,\infty)$.

(B) Fix $(n_1,n_2,n_3)\in P$, and for convenience, write $B := \frac{1}{8}A_{n_1 1}A_{n_2 2}A_{n_3 3}e^{i(\phi_{n_1 1}+\phi_{n_2 2}-\phi_{n_3 3})}$. Define

$$\begin{aligned}\beta_\sigma(f_1,f_2) &= g_\sigma\Big(\tfrac{\nu_{n_1 1}}{f_1}\Big)\,g_\sigma\Big(\tfrac{\nu_{n_2 2}}{f_2}\Big)\,g_\sigma\Big(\tfrac{\nu_{n_3 3}}{f_1+f_2}\Big)\\ b^1_\sigma(f_1,f_2,t) &= b_\sigma(f_1,f_2,t)-N_\sigma(f_1,f_2)^{-1}B\beta_\sigma(f_1,f_2).\end{aligned}$$

For each $\eta>0$, write

$$S_\eta := [\log(\nu_{n_1 1})-\eta,\log(\nu_{n_1 1})+\eta]\times[\log(\nu_{n_2 2})-\eta,\log(\nu_{n_2 2})+\eta].$$

Without loss of generality, fix $0<\varepsilon\ll|B|$. Let $\delta>0$ be sufficiently small that $(\nu_{\tilde{n}_1 1},\nu_{\tilde{n}_2 2})\notin S_\delta$ for all $(\tilde{n}_1,\tilde{n}_2,\tilde{n}_3)\in P\setminus\{(n_1,n_2,n_3)\}$ and

$$\left\{\frac{N'(\nu_{n_1 1},\nu_{n_2 2})}{N'(e^{\zeta_1},e^{\zeta_2})}\,:\,(\zeta_1,\zeta_2)\in S_\delta\right\}\;\subset\;(1-\tfrac{\varepsilon}{7|B|},1+\tfrac{\varepsilon}{7|B|}). \tag{57}$$

On the basis of Lemma 10, let M_1 be such that for $\sigma>M_1$,

$$\left\{\frac{\sigma^2 N_\sigma(e^{\zeta_1},e^{\zeta_2})}{N'(e^{\zeta_1},e^{\zeta_2})}\,:\,(\zeta_1,\zeta_2)\in S_\delta\right\}\;\subset\;(1-\tfrac{\varepsilon}{7|B|},1+\tfrac{\varepsilon}{7|B|})$$

and therefore by (57),

$$\left\{\frac{N_\sigma(\nu_{n_1 1},\nu_{n_2 2})}{N_\sigma(e^{\zeta_1},e^{\zeta_2})}\,:\,(\zeta_1,\zeta_2)\in S_\delta\right\}\;\subset\;(1-\tfrac{4\varepsilon}{7|B|},1+\tfrac{4\varepsilon}{7|B|}). \tag{58}$$

By the same reasoning as in the proof of part (I), we have that $b^1_\sigma(e^{\zeta_1},e^{\zeta_2},t)\to 0$ uniformly across $(\zeta_1,\zeta_2,t)\in S_\delta\times\mathbb{R}$ as $\sigma\to\infty$; so let M_2 be such that for $\sigma>M_2$,

$$\int_{S_\delta}\sup_{t\in\mathbb{R}}|b^1_\sigma(e^{\zeta_1},e^{\zeta_2},t)|\,d(\zeta_1,\zeta_2)\;<\;\tfrac{\varepsilon}{7}. \tag{59}$$

Now fix $\eta\in(0,\delta)$. By the same reasoning as in the proof of part (A), we have that

$$N_\sigma(\nu_{n_1 1},\nu_{n_2 2})^{-1}\int_{\mathbb{R}^2\setminus S_\eta}\beta_\sigma(e^{\zeta_1},e^{\zeta_2})\,d(\zeta_1,\zeta_2)\;\to\;0$$

as $\sigma\to\infty$. So choose $M>\max(M_1,M_2)$ such that for $\sigma>M$,

$$N_\sigma(\nu_{n_1 1},\nu_{n_2 2})^{-1}\int_{\mathbb{R}^2\setminus S_\eta}\beta_\sigma(e^{\zeta_1},e^{\zeta_2})\,d(\zeta_1,\zeta_2)\;<\;\tfrac{\varepsilon}{7|B|}$$

and therefore

$$N_\sigma(\nu_{n_1 1},\nu_{n_2 2})^{-1}\int_{S_\eta}\beta_\sigma(e^{\zeta_1},e^{\zeta_2})\,d(\zeta_1,\zeta_2)\;\in\;\big(1-\tfrac{\varepsilon}{7|B|},1\big].$$

By (58), this implies that

$$\int_{S_\eta}N_\sigma(e^{\zeta_1},e^{\zeta_2})^{-1}\beta_\sigma(e^{\zeta_1},e^{\zeta_2})\,d(\zeta_1,\zeta_2)\;\in\;\big(1-\tfrac{6\varepsilon}{7|B|},1+\tfrac{4\varepsilon}{7|B|}\big)$$

and so

$$\left| B - \int_{S_\eta} N_\sigma(e^{\zeta_1}, e^{\zeta_2})^{-1} B\beta_\sigma(e^{\zeta_1}, e^{\zeta_2})\, d(\zeta_1, \zeta_2) \right| \;<\; \tfrac{6\varepsilon}{7}.$$

Combining this with (59) gives that for all $t \in \mathbb{R}$,

$$\left| B - \int_{S_\eta} b_\sigma(e^{\zeta_1}, e^{\zeta_2})\, d(\zeta_1, \zeta_2) \right| \;<\; \varepsilon$$

as required.

5. Numerical examples

The purposes of this section are:

(a) to illustrate that our definition of the wavelet bispectrum in Sec. 3.1 gives appropriate results for fixed-frequency oscillatory components (in accordance with the "reasonableness" criteria in the discussion immediately preceding Sec. 3.1);

(b) to illustrate the usage of our wavelet bispectrum formula for investigating interactions within time-dependent oscillatory dynamics, with the example of a coupled phase-oscillator model.

The first purpose is primarily addressed by Sec. 5.1, and the second is addressed by Sec. 5.2.

For our numerics, we will use a family of wavelets ψ_σ, indexed by a parameter $\sigma > 0$, given by

$$\hat{\psi}_\sigma(r) \;=\; \mathring{\psi}_\sigma(r) \;=\; e^{-2(\pi\sigma\log r)^2} \;=\; \hat{f}_\sigma(\log r)$$

for $r > 0$, where f_σ is the probability density function of a normal distribution of mean zero and variance σ^2; and we take $\kappa = 1$, which is where $\hat{\psi}$ is maximised for each σ. These wavelets are known as *lognormal* or *log Gabor* wavelets [32, 14, 45, 25]. How the parameter σ relates quantitatively to the frequency resolution and time localisation of the lognormal wavelet is presented in Appendix A. But the key point is that larger σ corresponds to higher frequency resolution and poorer time localisation. If the frequencies of oscillatory components in the signals being analysed are not very far from each other, then a high value of σ needs to be used, otherwise the representations of the oscillatory components in the wavelet transform will overlap with and interfere with each other. If the frequency of an oscillatory component of interest is quickly varying over time, then one must not use too high a value for σ, otherwise temporal variations will be smeared and will interfere with each other in the wavelet representation.

In any case, in application to real data, the value chosen for σ should generally not be less than about 0.7, otherwise the wavelet transform is too highly time-localised to be able to "see oscillations". (This is analogous to the fact that the human ear cannot hear pressure vibrations of frequency less than about 20 Hz due to the ear's time-frequency analyser being too time-localised to pick up such low-frequency oscillations.) For σ larger than about 0.7, the lognormal wavelet may be approximated reasonably well by the Gabor function

$$\psi_\sigma(r) \;\approx\; f_\sigma(r)e^{2\pi i r}. \tag{60}$$

This is illustrated in Fig. 5.

The admissibility constant (20) of the lognormal wavelet function is given by

$$C_{\psi_\sigma} \;=\; \int_{-\infty}^{\infty} e^{-(2\pi\sigma r)^2}\, dr \;=\; \frac{1}{2\sqrt{\pi}\sigma}.$$

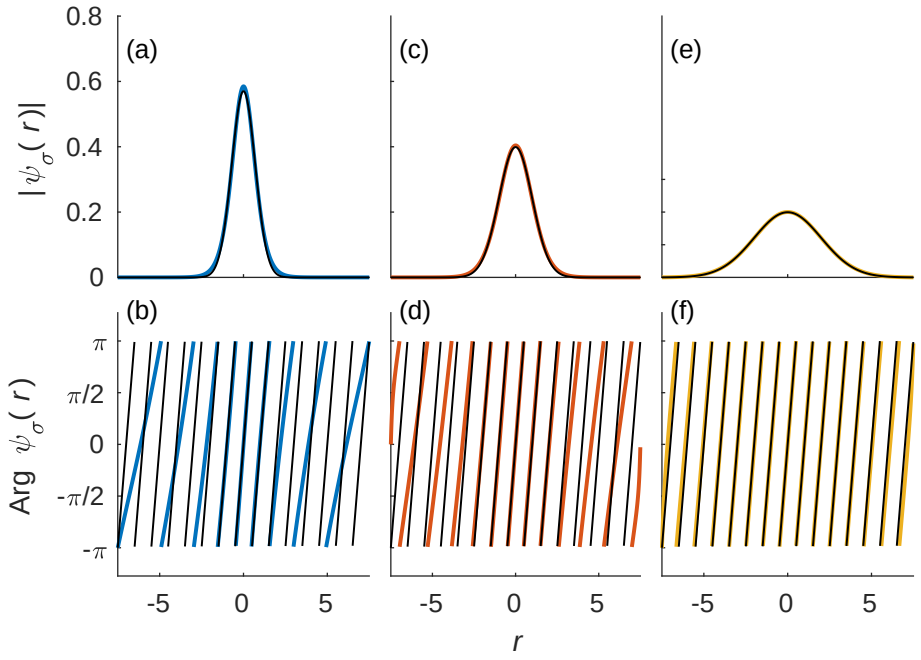


Figure 5: Lognormal wavelet function ψ_σ for (a,b) $\sigma = 0.7$, (c,d) $\sigma = 1$, (e,f) $\sigma = 2$. In (a,c,e) the modulus $|\psi_\sigma(r)|$ is shown in colour, against which $f_\sigma(r)$ is shown in black. In (b,d,f) the argument of $\psi_\sigma(r)$ is shown in colour, against which the wrapped angle $2\pi r$ is shown in black. We see that $|\psi_\sigma(r)|$ is approximately equal to $f_\sigma(r)$, and we see that where $f_\sigma(r)$ is not very small (compared to its maximum value $\frac{1}{\sqrt{2\pi}\sigma}$), the argument of $\psi_\sigma(r)$ is approximately equal to $2\pi r$ modulo 2π.

5.1. Sinusoidal oscillations

In this section, we define the "instantaneous wavelet bispectral content" of a set $A \in \mathcal{B}((0,\infty)^2)$ at a time τ by

$$\mathfrak{b}^{\text{inst}}_{\psi,\kappa,xyz}(A,\tau) \;:=\; \int_{\mathbb{R}^2} \mathbb{1}_A(e^{\zeta_1}, e^{\zeta_2}) b_{\psi,\kappa,xyz}(e^{\zeta_1}, e^{\zeta_2}, \tau)\, d(\zeta_1,\zeta_2)$$

provided the integrand is absolutely integrable. All the signals in this section are simulated over a time interval $[0\text{ s}, 60\text{ s}]$, and instantaneous wavelet bispectral results are considered at $\tau = 30$ s.

We have seen that whereas the second-order normalisation C_ψ derived in (21) or (22b) is a constant, the analogous third-order calculation represented by (44) yields a quantity $D_\psi(\lambda)$ that is dependent on the distance (in terms of logarithmic axes) from the diagonal of frequency-frequency space. Since the quantity $D_\psi(\lambda)$ is derived precisely in terms of integration of bispectral results for sinusoids, all "imperfect" results in the application of our wavelet bispectrum definition to sinusoids will be specifically due to the variability of $D_\psi(\lambda)$ with respect to λ. Theorem 8 justifies that the imperfection tends to 0 in the limit as frequency resolution tends to ∞. The goal of this section is essentially to illustrate that for the lognormal wavelet, any choice of $\sigma \geq 0.5$ will only lead to small imperfection. (As stated at the start of Sec. 5, using σ much less than 0.7 is physically meaningless.)

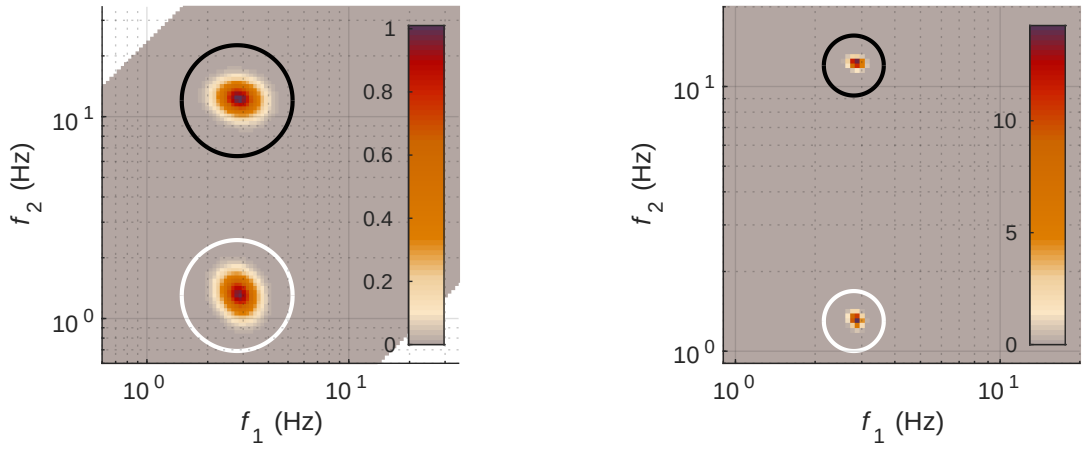
This fact is already shown to some degree by Fig. 4: it is seen that for $\sigma \geq 0.5$, the variability

of $D_{\psi_\sigma}(\lambda)$ with respect to λ is fairly small. In the limit as $\sigma \to \infty$, Lemma 10 gives that

$$\sigma^2 D_{\psi_\sigma}(\lambda) \;\to\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{f}_1(r_1)\hat{f}_1(r_2)\hat{f}_1(\lambda r_1 + (1-\lambda)r_2)\, dr_1 dr_2 \;=\; \frac{1}{\pi\sqrt{8(1-\lambda+\lambda^2)}},$$

for which the range is between $(\sqrt{8}\pi)^{-1} \approx 0.1125$ and $(\sqrt{6}\pi)^{-1} \approx 0.1299$.

We start by considering exactly the same signals (5) as considered in the Introduction with the same parameters, namely $\nu_1 = 2.8$ Hz, $\nu_2 = 1.3$ Hz, $\nu_3 = 12$ Hz and $\theta = \frac{\pi}{3}$ (as in the caption of Fig. 2). The logarithmic-frequency wavelet bispectral density $b_{\psi_\sigma,1,xyz}(f_1, f_2, \tau)$ is shown in Fig. 6, using $\sigma = 1$ and $\sigma = 4$ just as in Fig. 2.



(a) $\sigma = 1$ (b) $\sigma = 4$

Figure 6: Wavelet bispectral density according to our new Definition 4, for the signals considered in the Introduction. Plots show magnitude of the logarithmic-frequency wavelet bispectral density $b_{\psi_\sigma,1,xyz}(f_1, f_2, \tau)$ for the signals x, y, z in (5) with $\nu_1 = 2.8$ Hz, $\nu_2 = 1.3$ Hz, $\nu_3 = 12$ Hz and $\theta = \frac{\pi}{3}$, with (a) $\sigma = 1$ and (b) $\sigma = 4$. In (a), the white circle marks the boundary of S_1 and the black circle the boundary of S_2. In (b), the white circle marks the boundary of S_3 and the black circle the boundary of S_4.

We see a blurry peak around (ν_1, ν_2) and around (ν_1, ν_3) for both $\sigma = 1$ and $\sigma = 4$. For $\sigma = 1$, we mark a region S_1 containing the visible blur around (ν_1, ν_2) and a region S_2 containing the visible blur around (ν_1, ν_3), both indicated in Fig. 6. A numerical computation yields that

$$\mathfrak{b}^{\text{inst}}_{\psi_1,1,xyz}(S_1, \tau) \approx 0.1251 \quad \text{and} \quad \mathfrak{b}^{\text{inst}}_{\psi_1,1,xyz}(S_2, \tau) \approx 0.1249 e^{0.3333\pi i}.$$

Likewise, for $\sigma = 4$, we mark a region S_3 containing the visible blur around (ν_1, ν_2) and a region S_4 containing the visible blur around (ν_1, ν_3), again both indicated in Fig. 6. A numerical computation yields that

$$\mathfrak{b}^{\text{inst}}_{\psi_4,1,xyz}(S_3, \tau) \approx 0.1250 \quad \text{and} \quad \mathfrak{b}^{\text{inst}}_{\psi_4,1,xyz}(S_4, \tau) \approx 0.1250 e^{0.3333\pi i}.$$

So we see that all four of these regions have instantaneous bispectral content equal in magnitude to approximately $\frac{1}{8}$. In particular, for both frequency resolutions we are able to show that the two

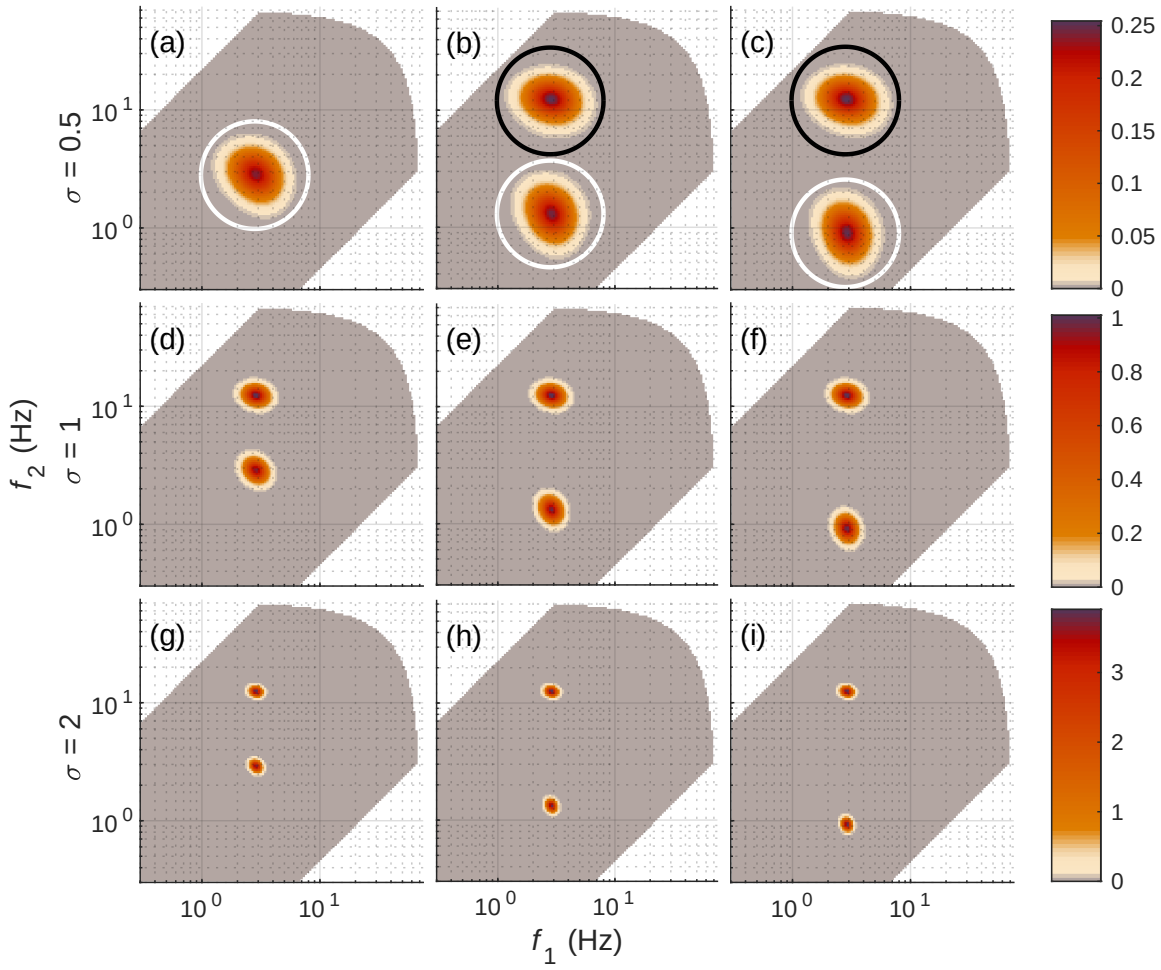


Figure 7: More wavelet bispectra for fixed-frequency oscillatory components. In (a), $|b_{\psi_\sigma,1,xxy}(f_1,f_2,\tau)|$ is shown for x,y as in (61) with $\nu = 2.8$ Hz, with $\sigma = 0.5$. In (d,g) is shown, respectively with $\sigma = 1$ and $\sigma = 2$, $|b_{\psi_\sigma,1,xyz}(f_1,f_2,\tau)|$ for x,y,z as in (5) with $\nu_1 = \nu_2 = 2.8$ Hz, $\nu_3 = 12$ Hz and $\theta = 0$. In (b,e,h) is shown, respectively with σ equal to 0.5, 1 and 2, $|b_{\psi_\sigma,1,xyz}(f_1,f_2,\tau)|$ for x,y,z as in (5) with $\nu_1 = 2.8$ Hz, $\nu_2 = 1.3$ Hz, $\nu_3 = 12$ Hz and $\theta = 0$. In (c,f,i) is shown, respectively with σ equal to 0.5, 1 and 2, $|b_{\psi_\sigma,1,xyz}(f_1,f_2,\tau)|$ for x,y,z as in (5) with $\nu_1 = 2.8$ Hz, $\nu_2 = 0.9$ Hz, $\nu_3 = 12$ Hz and $\theta = 0$. All nine plots use the same frequency axes. The instantaneous wavelet bispectral content of the circled regions in (a–c) are given in the text.

bispectral contributions are of essentially the same magnitude as each other, which could not be seen from the current standard definition of wavelet bispectra as described in the Introduction.

Even if we consider lower frequency resolution, our wavelet bispectrum definition gives good results. Fig. 7(a) shows $b_{\psi_\sigma,1,xxy}(f_1,f_2,\tau)$ with $\sigma = 0.5$ for the pair of signals

$$\begin{cases} x(t) &= \cos(2\pi\nu t) \\ y(t) &= \cos(4\pi\nu t) \end{cases} \tag{61}$$

with $\nu = 2.8$ Hz. The instantaneous wavelet bispectral content of the white-circled region in Fig. 7(a) is approximately 0.1255, again remarkably close to the ideal value of $\frac{1}{8}$. Likewise Fig. 7(b–c) shows $b_{\psi_\sigma,1,xxy}(f_1, f_2, \tau)$ with $\sigma = 0.5$ for signals x, y, z of the form (5), with $\theta = 0$. The instantaneous wavelet bispectral content of the white-circled region and the black-circled region in Fig. 7(b) are approximately $0.1260 + 0.0002i$ and $0.1254 - 0.0002i$ respectively. The instantaneous wavelet bispectral content of the white-circled region and the black-circled region in Fig. 7(c) are approximately $0.1253 + 0.0001i$ and $0.1250 - 0.0001i$ respectively. All of these values are very close to $\frac{1}{8}$, as desired. We also see in the remaining plots of Fig. 7 how the size of the blurs in frequency-frequency space decrease with increasing σ.

Let us now illustrate autobispectra for signals $x(t) = \cos(2\pi\nu_1 t) + \cos(2\pi\nu_2 t) + \cos(2\pi(\nu_1+\nu_2)t)$; specifically, we consider the cases

$$x(t) = 2\cos(2\pi\nu t) + \cos(4\pi\nu t) \tag{62a}$$

$$x(t) = \cos(2\pi\nu t) + \cos(3\pi\nu t) + \cos(5\pi\nu t) \tag{62b}$$

$$x(t) = \cos(2\pi\nu t) + \cos(4\pi\nu t) + \cos(6\pi\nu t) \tag{62c}$$

where $\nu = 2.4$ Hz. Recall that wavelet autobispectra are symmetric in the diagonal. Fig. 8 shows $b_{\psi_\sigma,1,xxx}(f_1, f_2, \tau)$ with $\sigma = 3$ for each of the three signals $x(t)$ in (62). For (62a), the instantaneous wavelet bispectral content of the circled region in Fig. 8(a) is approximately 0.5001. For (62b), the instantaneous wavelet bispectral content of the circled region in Fig. 8(b) is approximately 0.1250. For (62c), the instantaneous wavelet bispectral content of each of the circled regions in Fig. 8(c) is approximately 0.1250.

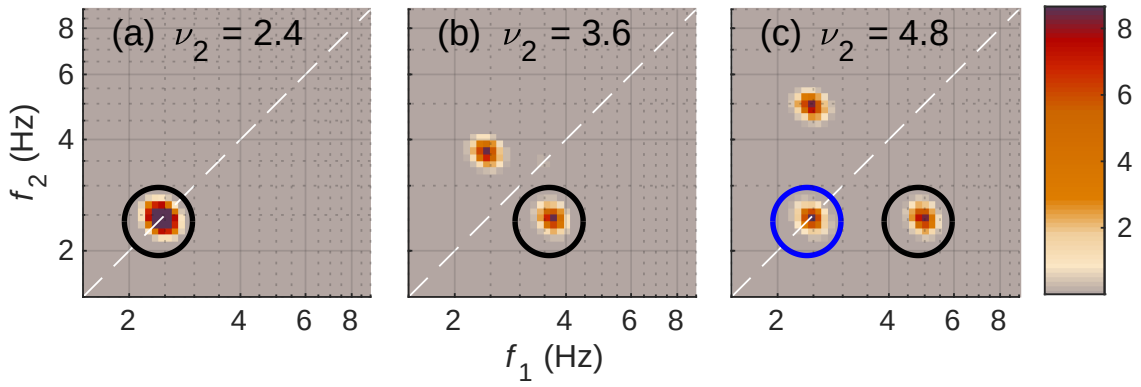


Figure 8: Wavelet autobispectra for signals with fixed-frequency oscillatory components. In (a), (b) and (c) are shown $|b_{\psi_\sigma,1,xxx}(f_1, f_2, \tau)|$ for $x(t)$ as in (62a), (62b) and (62c) respectively, with $\nu = 2.4$ Hz, with $\sigma = 3$. Shown in dashed white is the diagonal. The instantaneous wavelet bispectral content of the circled regions are given in the text.

5.2. Coupled phase oscillators

In this section, given a time interval $I \subset \mathbb{R}$ we define the "time-marginalised bispectral density"

$$b^I_{\psi,\kappa,xyz}(f_1, f_2) := \int_I b_{\psi,\kappa,xyz}(f_1, f_2, t)\, dt. \tag{63}$$

Except where stated otherwise, the signals in this section are simulated over a time interval $[0\text{ s}, 500\text{ s}]$, and bispectral results are considered over the central time-subinterval

$$I_0 := [233\tfrac{1}{3}\text{ s}, 266\tfrac{2}{3}\text{ s}]$$

of duration $33\frac{1}{3}$ s. (This corresponds to one time period of the frequency modulation that will be introduced later.)

One form of interaction between oscillators that bispectral analysis is likely to be able to detect is *dynamical coupling* in the form of added coupling terms in the differential equations of motion [73, 28]. Here, we consider Kuramoto-type symmetric coupling between two phase oscillators. The simplest scenario would be to introduce the coupling between two linear phase oscillators, as in the usual Kuramoto model; however, as we shall explain later, bispectral analysis is probably unable to detect such coupling between two linear phase oscillators. Therefore, instead, we introduce the Kuramoto coupling between two highly nonlinear phase oscillators.

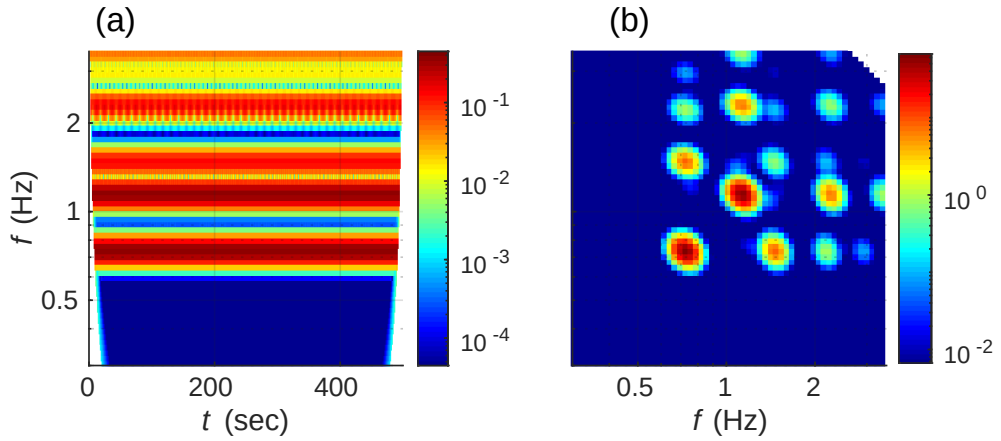


Figure 9: Wavelet and wavelet-bispectral representation of a linear superposition of two non-interacting nonlinear oscillators. In (a) and (b) are shown respectively $|W_{\psi_\sigma,1,x}(f,t)|$ and $|b^{I_0}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ (the latter in units of seconds), for the fixed-frequency uncoupled oscillators as in (64), with $\sigma = 3$.

We consider two phase oscillators $\theta_1(t)$ and $\theta_2(t)$ and define the signal

$$x(t) \;=\; \cos(\theta_1(t)) + \cos(\theta_2(t))$$

on which autobispectral analysis will be carried out. In the first instance, we suppose the phase oscillators are uncoupled and have fixed basic frequency; specifically, we take them to follow the differential equation

$$\begin{cases} \dot\theta_1(t) &= 2\pi\nu_1(1+0.6\cos(\theta_1)) \\ \dot\theta_2(t) &= 2\pi\nu_2(1+0.6\cos(\theta_2)) \end{cases} \tag{64}$$

with $\nu_1 = 0.9$ Hz and $\nu_2 = 1.4$ Hz. The oscillator θ_1 is strictly periodic with frequency $\tilde\nu_1 := 0.8\nu_1 = 0.72$ Hz, and the oscillator θ_2 is strictly periodic with frequency $\tilde\nu_2 := 0.8\nu_2 = 1.12$ Hz. For this case, Fig. 9(a) shows the wavelet transform of x and Fig. 9(b) shows $b^{I_0}_{\psi_\sigma,1,xxx}(f_1,f_2)$, both with $\sigma = 3$. The six most prominent bispectral contributions that we see in Fig. 9(b) are:

the diagonal contribution around $(\tilde{\nu}_1, \tilde{\nu}_1)$, and the off-diagonal contribution around $(2\tilde{\nu}_1, \tilde{\nu}_1)$ and its mirror-reflection around $(\tilde{\nu}_1, 2\tilde{\nu}_1)$; and likewise, the diagonal contribution around $(\tilde{\nu}_2, \tilde{\nu}_2)$, and the off-diagonal contribution around $(2\tilde{\nu}_2, \tilde{\nu}_2)$ and its mirror-reflection around $(\tilde{\nu}_2, 2\tilde{\nu}_2)$. These bispectral contributions are simply due to the nonlinearity of the two individual oscillators; they do not signify any kind of interaction between the oscillators.

However, when we introduce coupling, new peaks appear that indicate an interaction between the oscillators. Specifically, consider now the coupled system

$$\begin{cases} \dot{\theta}_1(t) &= 2\pi\nu_1(1 + 0.6\cos(\theta_1)) + K\sin(\theta_2 - \theta_1) \\ \dot{\theta}_2(t) &= 2\pi\nu_2(1 + 0.6\cos(\theta_2)) + K\sin(\theta_1 - \theta_2) \end{cases} \tag{65}$$

with ν_1 and ν_2 as before, and $K = 0.2$ rad/s. This coupling is not very strong, but it has a significant effect. Once again, Fig. 10(a) shows the wavelet transform of x and Fig. 10(b) shows $b^{I_0}_{\psi_\sigma,1,xxx}(f_1, f_2)$, both with $\sigma = 3$. We see in Fig. 10(b) prominent bispectral contributions around roughly the same points as in Fig. 9(b), plus some new peaks, the most prominent being a bispectral contribution roughly around $(\tilde{\nu}_2, \tilde{\nu}_1)$ as well as its mirror-reflection roughly around $(\tilde{\nu}_1, \tilde{\nu}_2)$.

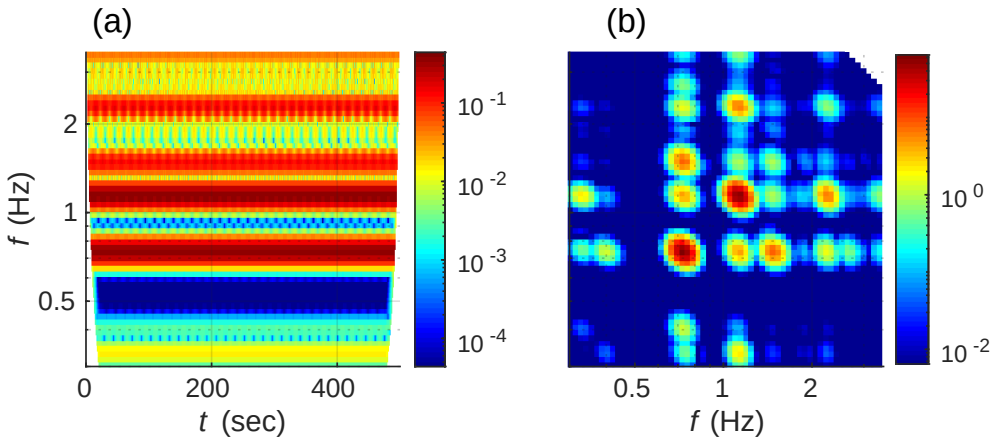


Figure 10: Wavelet and wavelet-bispectral representation of a linear superposition of two interacting nonlinear oscillators. In (a) and (b) are shown respectively $|W_{\psi_\sigma,1,x}(f,t)|$ and $|b^{I_0}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ (the latter in units of seconds), for the coupled fixed-frequency oscillators as in (65), with $\sigma = 3$.

So we have illustrated how the bispectrum can detect the introduction of an interaction between the oscillators. But all of this so far could have been illustrated by traditional non-time-evolving Fourier bispectral analysis. So we now introduce slow frequency modulation. For simplicity, the modulation is sinusoidal in shape, but any shape of (not too fast) frequency modulation can be treated by wavelet bispectral analysis. Consider the system

$$\begin{cases} \dot{\theta}_1(t) &= 2\pi\nu_1(t)(1 + 0.6\cos(\theta_1)) + K\sin(\theta_2 - \theta_1) \\ \dot{\theta}_2(t) &= 2\pi\nu_2(1 + 0.6\cos(\theta_2)) + K\sin(\theta_1 - \theta_2) \end{cases} \tag{66}$$

with ν_2 and K as before, and

$$\nu_1(t) \;=\; 0.9 \text{ Hz} \times (1 + 0.1\sin(2\pi f_m t))$$

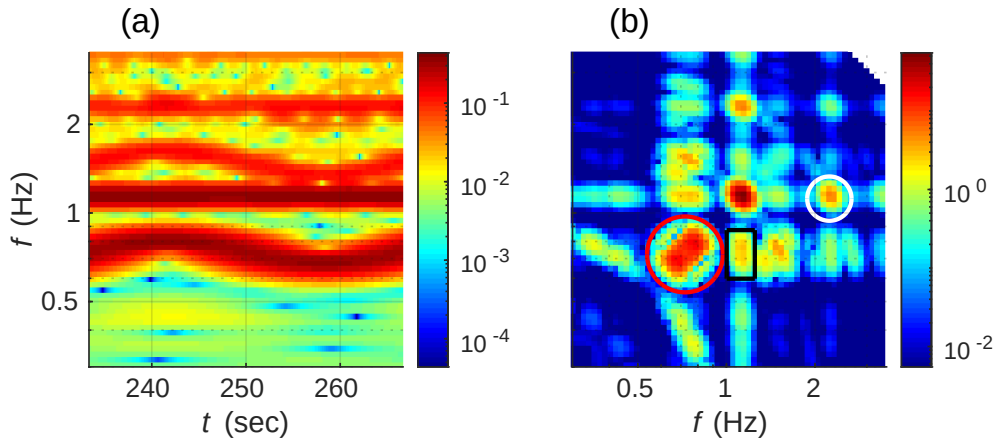


Figure 11: Wavelet and wavelet-bispectral representation of a linear superposition of two interacting nonlinear oscillators, with one having slowly time-varying inherent frequency. In (a) and (b) are shown respectively $|W_{\psi_\sigma,1,x}(f,t)|$ and $|b^{I_0}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ (the latter in units of seconds), for the coupled oscillators as in (66), with $\sigma = 3$. The bispectral content associated to the three marked frequency-frequency regions over the time interval I_0 are given in the text.

where $f_m = 0.03$ Hz. So now θ_1 has slowly varying inherent frequency, approximately of the sinusoidal shape $0.8\nu_1(t)$. The time interval I_0 corresponds to exactly one time period of $\nu_1(t)$. Fig. 11(a) shows the wavelet transform of x over I_0 and Fig. 11(b) shows $b^{I_0}_{\psi_\sigma,1,xxx}(f_1,f_2)$, both with $\sigma = 3$. Some bispectral contribution such as that around $(\tilde{\nu}_1,\tilde{\nu}_1)$ are now bimodal rather than unimodal. This is because the sinusoidal temporal variation of $\nu_1(t)$ spends more time near its extreme values 0.81 Hz and 0.99 Hz than it does around each value in between.

With the above example, we now illustrate how our new definition of the wavelet bispectrum is able to quantify and compare bispectral contributions over a given time-interval, across different parts of frequency-frequency space; as described in the Introduction, this was not possible under the previous state-of-the-art of wavelet bispectral analysis.

In Fig. 11(b) are marked regions R_{purple}, R_{grey} and R_{black} respectively around the blurry peaks containing $(\tilde{\nu}_1,\tilde{\nu}_1)$, $(2\tilde{\nu}_2,\tilde{\nu}_2)$ and $(\tilde{\nu}_2,\tilde{\nu}_1)$. Once again, the first two regions contain contributions due to the nonlinearity of the individual oscillators, while the third represents the interaction that has been introduced between the oscillators. The bispectral content associated to each of the three indicated bispectral contributions are:

$$\begin{aligned}
\mathfrak{b}_{\psi_\sigma,1,xxx}(R_{\text{purple}} \times I_0) &\approx 0.976e^{0.08\pi i} \text{ s} \\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R_{\text{grey}} \times I_0) &\approx 0.097e^{0.03\pi i} \text{ s} \\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R_{\text{black}} \times I_0) &\approx 0.086e^{-0.45\pi i} \text{ s}.
\end{aligned}$$

Moreover, one can actually track over time the motion of the peaks and associated bispectral content. Define the time intervals $I_1, I_2, I_3 \subset I_0$ to be the 5-second intervals centred on $7.25f_m^{-1}$, $7.5f_m^{-1}$ and $7.75f_m^{-1}$ respectively; these are marked in Fig. 12. In Fig. 13 are shown $b^{I_1}_{\psi_\sigma,1,xxx}(f_1,f_2)$, $b^{I_2}_{\psi_\sigma,1,xxx}(f_1,f_2)$ and $b^{I_3}_{\psi_\sigma,1,xxx}(f_1,f_2)$; and on these plots are marked respectively the regions $R^{(1)}_{\text{black}}$, $R^{(2)}_{\text{black}}$ and $R^{(3)}_{\text{black}}$ which help to trace over time the bispectral contribution to the green-marked

region in Fig. 11(b). Again, our new definition of the wavelet bispectrum enables quantification of the time-evolving bispectral contributions:

$$\begin{aligned}
\mathfrak{b}_{\psi_\sigma,1,xxx}(R^{(1)}_{\text{black}} \times I_1) &\approx 0.008e^{-0.52\pi i} \text{ s} \\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R^{(2)}_{\text{black}} \times I_2) &\approx 0.017e^{-0.47\pi i} \text{ s} \\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R^{(3)}_{\text{black}} \times I_3) &\approx 0.011e^{-0.45\pi i} \text{ s}
\end{aligned}$$

Here, we have sampled three time-subintervals. A more continuous-time tracking of bispectral content can be achieved by following the frequency variation of oscillatory components of interest directly from the wavelet transforms themselves via "ridge-extraction" methods, as described in Remark 12 (in Appendix B). Bicoherence analysis (Appendix B.2) can then also be carried out.

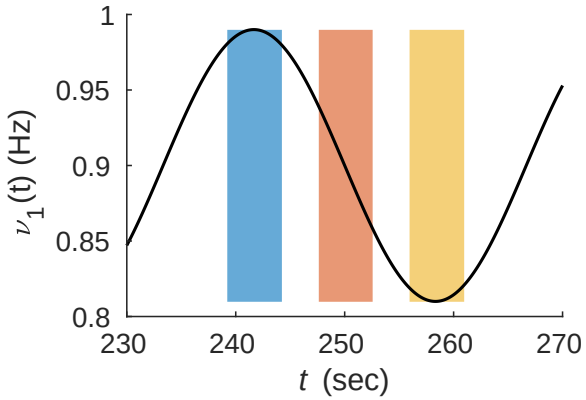


Figure 12: Graph of $\nu_1(t)$ over I_0, on which are marked the time intervals I_1, I_2 and I_3.

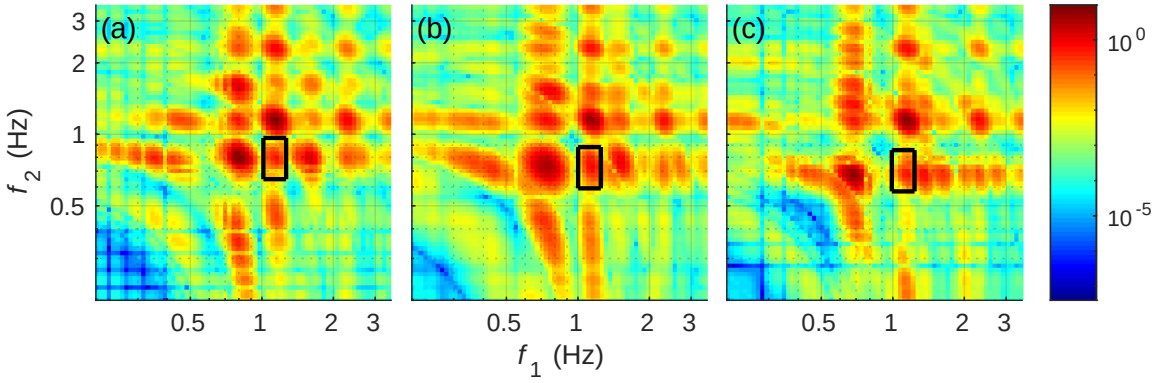


Figure 13: Tracing a bispectral contribution over time. In (a), (b) and (c) are shown respectively $|b^{I_1}_{\psi_\sigma,1,xxx}(f_1,f_2)|$, $|b^{I_2}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ and $|b^{I_3}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ (all in units of seconds), for the coupled oscillators as in (66), with $\sigma = 3$. The bispectral content associated to the marked regions of frequency-frequency over the respective time-intervals are given in the text.

Why the bispectrum does not detect the coupling for linear phase oscillators

Still working with $x(t) = \cos(\theta_1(t)) + \cos(\theta_2(t))$, we consider the Kuramoto model

$$\begin{cases} \dot{\theta}_1(t) &= 2\pi\nu_1 + K\sin(\theta_2 - \theta_1) \\ \dot{\theta}_2(t) &= 2\pi\nu_2 + K\sin(\theta_1 - \theta_2) \end{cases} \tag{67}$$

with $\nu_1 = 0.9$ Hz, $\nu_2 = 1.4$ Hz and $K = 0.2$ rad/s. The signal $x(t)$ was simulated over the time interval $[0\text{ s}, 10^4\text{ s}]$. Fig. 14 shows the magnitude of $X(\cdot)$ defined as the result of taking the Fourier transform of x restricted to the interval $[200\text{ s}, 10^4\text{ s}]$ and dividing by the time-duration 9800 s. (The first 200 seconds were removed to avoid any possible transient dynamical behaviour of the system (67).)

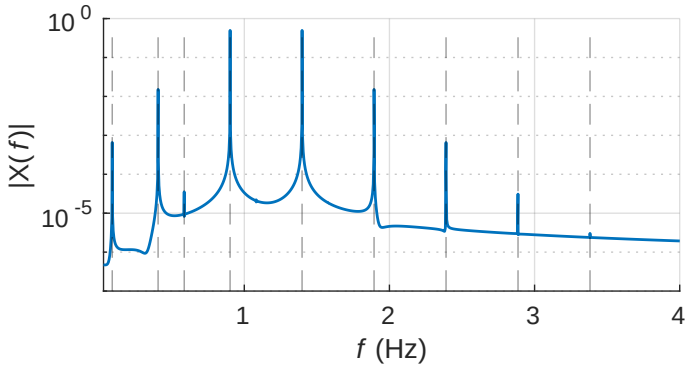


Figure 14: The Fourier components of $x(t)$. Here, $|X(f)|$ as plotted on the vertical axis is the Fourier amplitude associated to f, normalised by the duration of the signal. The frequency-values $|0.406+0.496n|$ Hz are marked for integers n ranging from -2 to 6.

We see in Fig. 14 that the Fourier components of $x(t)$ have frequencies that all lie within a set of the form $\{|\xi_1 + n\xi_2| : n \in \mathbb{Z}\}$. If ξ_1 and ξ_2 are rationally independent, then no two frequencies in this set can have a sum that also lies in this set. Thus bispectral analysis would not be able to detect the coupling. Instead, at least some kind of trispectral (i.e. fourth-order spectral) analysis would be needed: perhaps the simplest and most direct approach, as proposed in [28], would be to investigate the presence of frequency triplets with equal spacing between the two consecutive pairs of frequencies, using a time-evolving version of the "modified" Fourier bispectrum (cf. also [62])

$$\mathcal{T}_{xxx}(f_1, f_2) \;=\; \frac{\hat{x}(f_1)^2\,\overline{\hat{x}(f_2)\hat{x}(2f_1 - f_2)}}{|\hat{x}(f_1)|}$$

on $\Gamma_{1-} := \{0 < f_2 < f_1\}$ or equivalently on $\{0 < f_1 < f_2 < 2f_1\}$.

6. Application to cell membrane potential data

Cell membrane potential oscillations and coupling of cellular oscillators are of considerable importance in understanding cellular biology [19, 3, 48, 49]. In this section, we illustrate application of our new definition of the wavelet bispectrum to a cell membrane potential time-series, recorded with the free-running patch clamp technique; full details of the experimental setup are given in [50]. For a much more extensive analysis of cell membrane potential time-series, see [52].

The membrane potential time-series $x_{\mathrm{raw}}(t)$ that we consider is a digital signal recorded over a time-interval $I = [0\text{ s}, 600\text{ s}]$, shown in Fig. 15(b). We apply the MATLAB® `detrend` function, which subtracts a best-fit linear trend, to obtain a linearly detrended signal $x(t)$ defined over the same time-interval I. Fig. 15(a) shows the wavelet transform of $x(t)$, using a lognormal wavelet ψ_σ (as used throughout Sec. 5) with frequency resolution $\sigma = 3$.

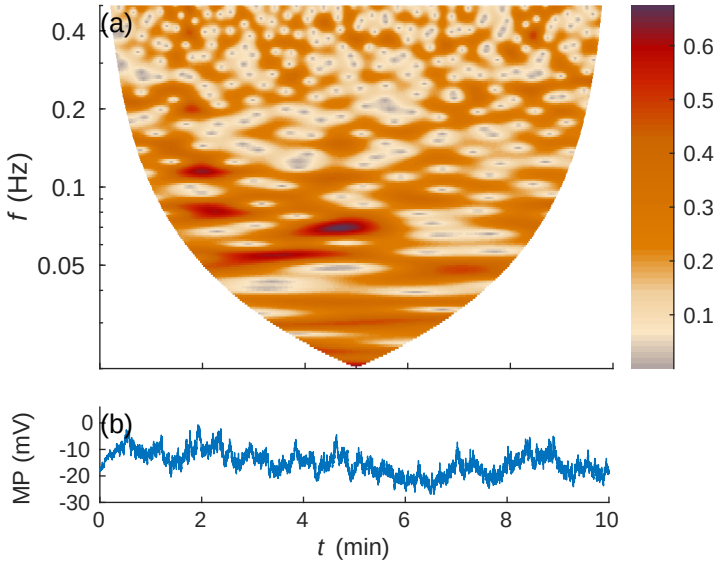


Figure 15: Cell membrane potential time-series and wavelet transform. In (a) is shown the wavelet amplitude $|W_{\psi_\sigma,1,x}(f,t)|$ (in units of mV) of the detrended signal $x(t)$, with $\sigma = 3$. In (b) is shown the membrane potential $x_{\mathrm{raw}}(t)$ itself.

Along similar lines to Sec. 5.2, before seeking to track time-evolving bispectral values, we first give a plot of the bispectral content of x taken "over the whole time-interval", in order to ascertain where peaks in the bispectrum lie. Since the signal is of finite duration and finite sampling frequency, there is only a subset of time-frequency space over which the wavelet transform can be calculated (as seen in Fig. 15(a)). Therefore, for each pair of frequencies (f_1, f_2), we define $I(f_1, f_2) \subset I$ to be the interval of times t at which the product of wavelet terms $W_{\psi_3,1,x}(f_1,t)W_{\psi_3,1,x}(f_2,t)\overline{W_{\psi_3,1,x}(f_1+f_2,t)}$ can be computed. We write $T(f_1, f_2)$ for the duration of the time-interval $I(f_1, f_2)$. In Fig. 16(a), we plot the magnitude of the "time-averaged bispectral density"

$$b_{\mathrm{ave},xxx}(f_1,f_2) \;:=\; \frac{b^{I(f_1,f_2)}_{\psi_3,1,xxx}(f_1,f_2)}{T(f_1,f_2)} \;=\; \frac{1}{T(f_1,f_2)}\int_{I(f_1,f_2)} b_{\psi_3,1,xxx}(f_1,f_2,t)\,dt, \tag{68}$$

where the notation $b^I_{\psi,\kappa,xyz}$ denotes "time-marginalised bispectral density" as in Eq. (63).

To help give an indication of whereabouts in frequency-frequency space the bispectral values shown in Fig. 16(a) are significantly high, we carry out a 95%-significance surrogate test using WIAAFT surrogates [38]: 19 digital signals $x_n(t)$, $n = 1, \ldots, 19$, of the same length as the digital signal $x(t)$, are sampled independently of each other from the WIAAFT surrogate distribution generated by $x(t)$, and the quantity $b_{\mathrm{ave},x_nx_nx_n}(f_1, f_2)$ is computed for these 19 signals exactly as it

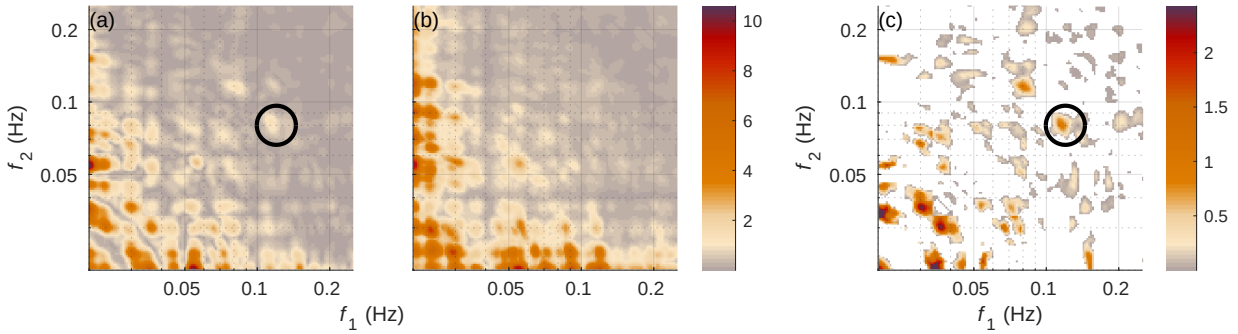


Figure 16: Time-averaged wavelet bispectral density of $x(t)$. In (a) is shown the magnitude $|b_{\mathrm{ave},xxx}(f_1, f_2)|$ (in units of mV^3) of the signal's time-averaged bispectral density $b_{\mathrm{ave},xxx}(f_1, f_2)$ defined in Eq. (68). In (b) is shown the significance threshold $b_{\mathrm{thresh}}(f_1, f_2)$ (in units of mV^3) for 95% significance in a surrogate test with 19 WIAAFT surrogates [38], as given by Eq. (69). In (c) is shown $|b_{\mathrm{ave},xxx}(f_1, f_2)| - b_{\mathrm{thresh}}(f_1, f_2)$ (in units of mV^3), wherever $|b_{\mathrm{ave},xxx}(f_1, f_2)|$ is larger than $b_{\mathrm{thresh}}(f_1, f_2)$. The black circle in plots (a) and (c) marks one particular region R where we see significantly high bispectral content.

was computed for $x(t)$. At each pair of frequencies (f_1, f_2), the 95% significance threshold is given by

$$b_{\mathrm{thresh}}(f_1, f_2) \;=\; \max_{n\in\{1,\ldots,19\}} |b_{\mathrm{ave},x_n x_n x_n}(f_1, f_2)|. \tag{69}$$

Fig. 16(b) shows this critical threshold. Fig. 16(c) shows the region of (f_1, f_2)-values for which $|b_{\mathrm{ave},xxx}(f_1, f_2)| > b_{\mathrm{thresh}}(f_1, f_2)$, with the colour-coding representing the difference

$$|b_{\mathrm{ave},xxx}(f_1, f_2)| - b_{\mathrm{thresh}}(f_1, f_2).$$

We note that all three plots in Fig. 16 are symmetric in the diagonal, because these plots are for autobispectra.

One of the significant bispectral peaks that we see in Fig. 16(c) occurs roughly around the frequency pair $f_1 = 0.12$ Hz and $f_2 = 0.08$ Hz. Having identified this peak, let us now look more at how the bispectral content in that region of frequency-frequency space evolves in time. Define the following five consecutive time-intervals of 30 seconds each: $I_1 = [90\ \mathrm{s}, 120\ \mathrm{s}]$, $I_2 = [120\ \mathrm{s}, 150\ \mathrm{s}]$, $\ldots$, $I_5 = [210\ \mathrm{s}, 240\ \mathrm{s}]$. The five plots in Fig. 17 show the time-marginalised bispectral densities over these five time-intervals. We specify a region R around the abovementioned frequency pair; this region is marked by a black circle in Fig. 16(a,c) and all the plots in Fig. 17. The time-evolution of the bispectral content of R can be traced as follows:

$$\begin{aligned}
\mathfrak{b}_{\psi_\sigma,1,xxx}(R \times I_1) &\approx 9.6e^{0.20\pi i}\ \mathrm{mV}^3\mathrm{s}\\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R \times I_2) &\approx 5.8e^{0.25\pi i}\ \mathrm{mV}^3\mathrm{s}\\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R \times I_3) &\approx 1.4e^{0.20\pi i}\ \mathrm{mV}^3\mathrm{s}\\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R \times I_4) &\approx 0.6e^{-0.65\pi i}\ \mathrm{mV}^3\mathrm{s}\\
\mathfrak{b}_{\psi_\sigma,1,xxx}(R \times I_5) &\approx 0.5e^{-0.37\pi i}\ \mathrm{mV}^3\mathrm{s}.
\end{aligned}$$

The sequence of above values reflects the general trend seen in Fig. 17 that the contributions to the wavelet autobispectrum of x are decreasing in intensity over time. Indeed, by the third minute of the experiment, the main bispectral contribution in R has essentially disappeared altogether. This

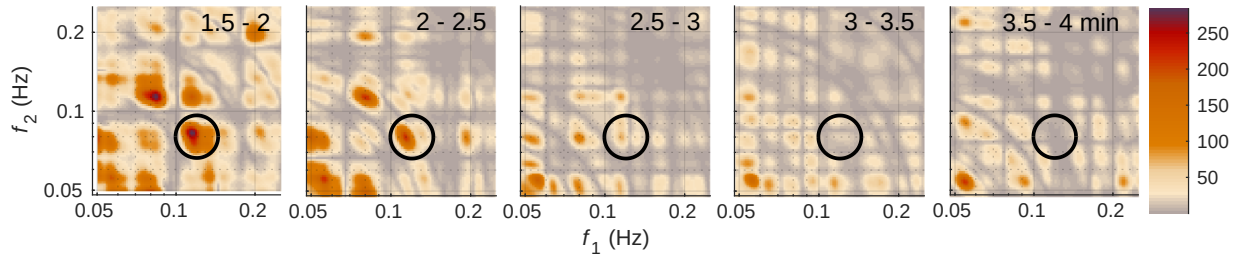


Figure 17: Time-evolving bispectral content of $x(t)$. The plots show the magnitude $|b^{I_i}_{\psi_\sigma,1,xxx}(f_1,f_2)|$ (in units of $\mathrm{mV}^3\mathrm{s}$) of the time-marginalised bispectral density as defined in Eq. (63) over each of the time-intervals $I_1,\ldots,I_5$ (consecutive 30-second intervals starting at 90 s), with $\sigma = 3$. The black circle marks the region R that was marked in Fig. 16.

general trend seen in Fig. 17 may be due to the weakening activity of cellular oscillatory processes as the cell gradually dies during the patch clamp experiment [52].

In this section, we have illustrated how our new definition of the wavelet bispectrum $\mathfrak{b}_{\psi,\kappa,xyz}$ can be used to study the time-evolving bispectral content of experimental time-series data.

Appendix A. Resolution properties of lognormal wavelets

We describe how, for lognormal wavelets, the parameter σ relates quantitatively to the frequency resolution and time localisation of the wavelet transform; for reference, we compare this with the frequency resolution and time localisation properties of the Gaussian-windowed Fourier transform. A numerical study of the performance of lognormal wavelets is also included in [25], which uses an alternative approach to quantifying resolution properties from what we present below.

For reference we start by considering windowed Fourier transforms with Gaussian windows. First let us define our approach to quantification of time localisation and frequency resolution for windowed Fourier transforms. Given an even function $w \in L^1(\mathbb{R},[0,\infty)) \cap L^2(\mathbb{R},[0,\infty)) \setminus \{0\}$, probably the most standard way to quantify the time localisation and frequency resolution of the associated windowed Fourier transform is, inversely, by the *Heisenberg uncertainty* properties of w. Writing $\|w\|_2 := \int_{-\infty}^{\infty} w(t)^2\,dt = \int_{-\infty}^{\infty} \hat{w}(f)^2\,df$, we define

- the *Heisenberg time uncertainty* of w by

$$\epsilon_{\mathrm{time}}(w) \;=\; \left(\frac{\int_{-\infty}^{\infty} t^2 w(t)^2\,dt}{\|w\|_2}\right)^{\frac{1}{2}};$$

- the *Heisenberg frequency uncertainty* of w by

$$\epsilon_{\mathrm{freq}}(w) \;=\; \left(\frac{\int_{-\infty}^{\infty} f^2 \hat{w}(f)^2\,df}{\|w\|_2}\right)^{\frac{1}{2}}.$$

Now for any $\tau > 0$ let f_τ be the probability density function of the normal distribution of mean zero and variance τ^2. The squared Heisenberg uncertainties of the Gaussian function f_τ are given by

$$\epsilon_{\mathrm{time}}(f_\tau)^2 = \tfrac{1}{2}\tau^2 \quad \text{and} \quad \epsilon_{\mathrm{freq}}(f_\tau)^2 = \frac{1}{8\pi^2\tau^2}.$$

The product $\epsilon_{\text{time}}(f_\tau)\epsilon_{\text{freq}}(f_\tau)$, called the *Heisenberg area* or the *Heisenberg time-frequency uncertainty* of f_τ, is equal to $\frac{1}{4\pi}$ (independently of τ). This is the lowest possible value for the Heisenberg time-frequency uncertainty of a window function, and is uniquely obtained by Gaussian windows [31, Sec. 2.2].

We now describe how the concept of Heisenberg uncertainties can be translated to the setting of wavelets and the wavelet transform. Given a wavelet function ψ with $\|\psi\|_2 := \int_{-\infty}^{\infty} |\psi(r)|^2\,dr < \infty$ and with admissibility constant $C_\psi := \int_{-\infty}^{\infty} \hat{\psi}(e^r)^2\,dr$ as in (20), together with a value $\kappa > 0$ (typically taken to be where $\hat{\psi}$ is maximised), we define

- the *Heisenberg time uncertainty* of (ψ, κ) by

$$\epsilon_{\text{time}}(\psi,\kappa) \;=\; \kappa \left(\frac{\int_{-\infty}^{\infty} r^2 |\psi(r)|^2\,dr}{\|\psi\|_2} \right)^{\frac{1}{2}};$$

- the *Heisenberg linear frequency uncertainty* of (ψ, κ) by

$$\epsilon_{\text{linfreq}}(\psi,\kappa) \;=\; \kappa^{-1} \left(\frac{\int_{0}^{\infty} (r-\kappa)^2 \hat{\psi}(r)^2\,dr}{\|\psi\|_2} \right)^{\frac{1}{2}};$$

- the *Heisenberg logarithmic frequency uncertainty* of (ψ, κ) by

$$\epsilon_{\text{freq}}(\psi,\kappa) \;=\; \left(\frac{\int_{-\infty}^{\infty} (r-\log\kappa)^2 \hat{\psi}(e^r)^2\,dr}{C_\psi} \right)^{\frac{1}{2}}.$$

The logarithmic-frequency approach in the definition of $\epsilon_{\text{freq}}(\psi, \kappa)$ is very natural for wavelets, due to the inherently logarithmic nature of frequency resolution for the wavelet transform as described in Sec. 2.2.

Let ψ_σ be the lognormal wavelet as defined in Sec. 5. We will take $\kappa = 1$, which is where $\hat{\psi}_\sigma$ is maximised. It is not hard to show that ψ_σ belongs to the Schwartz space $\mathcal{S}(\mathbb{R})$ for all σ. One can compute (via Plancherel's theorem) that

$$\epsilon_{\text{time}}(\psi_\sigma, 1)^2 \;=\; \tfrac{1}{2}\sigma^2 + \tfrac{1}{16\pi^2}.$$

This is very slightly more than the classical Heisenberg time uncertainty of a Gaussian window with variance σ^2, namely $\frac{1}{2}\sigma^2$. The Heisenberg logarithmic frequency uncertainty of $(\psi_\sigma, 1)$ is (essentially by definition) the same as the classical Heisenberg frequency uncertainty of a Gaussian window with variance σ^2; that is,

$$\epsilon_{\text{freq}}(\psi_\sigma, 1)^2 \;=\; \frac{1}{8\pi^2\sigma^2}.$$

Finally, one can compute that

$$\epsilon_{\text{linfreq}}(\psi_\sigma, 1)^2 \;=\; e^{\frac{1}{2\pi^2\sigma^2}} - 2e^{\frac{3}{16\pi^2\sigma^2}} + 1.$$

For σ not too small, this is very close to $\frac{1}{8\pi^2\sigma^2}$.

Appendix B. Coherence and bicoherence

In this appendix, we describe some notions of coherence (for second-order cross-spectral analysis) and bicoherence (for autobispectral and cross-bispectral analysis) derived from the cross-energy spectral and bispectral densities defined in Sec. 2.4 and Sec. 3 respectively. We also briefly address unidirectional dynamical coupling, and address "quadratic phase coupling" as detected by "bicorrelation" analysis.

Appendix B.1. Coherence

The wavelet cross-energy spectrum can be used to search for common oscillatory influences between two signals and track temporal variations thereof. For instance, one can integrate a sliding time-frequency window function with respect to the wavelet cross-energy spectrum and look for where the absolute value is high; that is, one can plot a function taking the form

$$(f,t) \mapsto \left| \int_{\mathbb{R}^2} S_f(\zeta, \tau - t) p_{\psi,\kappa,xy}(e^{\zeta}, \tau)\, d(\zeta,\tau) \right|$$

(or, as in [59], one can just smooth the wavelet cross-ESD with respect to time rather than time-frequency) and look for high values. However, high values in a region of time-frequency space can easily be due to high wavelet amplitudes happening to occur by fluke at similar frequencies in both signals. If the frequencies are time-varying, one way to help overcome this issue is to use a quantity that is invariant under rescaling of wavelet amplitudes and gives an indication of how constant the wavelet phase difference remains around a given point in time-frequency space.

We now describe two possible such quantities; although varying terminologies exist and both of these quantities have been referred to as "phase coherence", here (as in [37]) we will refer to the first of them as *wavelet coherence*, and the second as *wavelet-phase coherence* (since it works directly with the wavelet phases).

Remark 11. We give definitions of wavelet coherence $\mathfrak{c}^{(1)}_{\psi,\kappa,xy}(f,t)$ and wavelet-phase coherence $\mathfrak{c}^{(2)}_{\psi,\kappa,xy}(f,t)$ defined over (f,t) space. It may be useful to follow these specifically over a "time-frequency curve" $(f_t)_{t\in[a,b]}$. In particular, by looking at the wavelet transform of x or of y one may be able to trace the temporal variations in frequency f_t of some oscillatory component, e.g. by "ridge-extraction" methods as detailed in [25, 26]. Then, having traced f_t over a time-interval $[a,b]$, one can investigate whether this component represents a process influencing both x and y, by plotting $t \mapsto \mathfrak{c}^{(i)}_{\psi,\kappa,xy}(f_t,t)$ over $[a,b]$, where $\mathfrak{c}^{(i)}_{\psi,\kappa,xy}$ $(i=1,2)$ is the time-frequency-smoothed version of coherence as in (B.1)/(B.3). (Alternatively one can smooth only in time and plot the coherence over a neighbourhood of the curve $\{(f_t,t) : t \in [a,b]\}$ in (f,t) space.)

Wavelet coherence

Wavelet coherence (under the approach introduced in [67]) is defined in terms of the "angle" in the Hilbert space $L^2(\mathbb{R}^2)$ between time-frequency-smoothed versions of the wavelet transforms of the two signals. Since the smoothing should not blur together far apart scales, it should not really matter whether one integrates with respect to linear frequency or logarithmic frequency; here, we write the definition in terms of logarithmic frequency. But the smoothing function itself should generally be rescaled in accordance with the frequency under investigation, both in its temporal and frequency smoothing.

Given a family $(S_f)_{f>0}$ of bounded integrable functions $S_f\colon \mathbb{R}^2 \to [0,\infty)$ satisfying $S_f(\zeta,\tau) = S_f(\zeta,-\tau)$ for all $\zeta,\tau \in \mathbb{R}$, define the wavelet coherence between signals x and y by

$$\mathfrak{c}^{(1)}_{\psi,\kappa,xy}(f,t) \;=\; \frac{\left|\int_{\mathbb{R}^2} S_f(\zeta,\tau-t) W^{[p]}_{\psi,\kappa,x}(e^\zeta,\tau)\overline{W^{[p]}_{\psi,\kappa,y}(e^\zeta,\tau)}\,d(\zeta,\tau)\right|^2}{\int_{\mathbb{R}^2} S_f(\zeta,\tau-t)|W^{[p]}_{\psi,\kappa,x}(e^\zeta,\tau)|^2\,d(\zeta,\tau)\int_{\mathbb{R}^2} S_f(\zeta,\tau-t)|W^{[p]}_{\psi,\kappa,y}(e^\zeta,\tau)|^2\,d(\zeta,\tau)} \tag{B.1}$$

for all $f>0$ and $t\in\mathbb{R}$ where this is well-defined (which includes in particular that the two terms in the denominator are non-zero). The family $(S_f)_{f>0}$ serves as a smoothing kernel. The *default* dependence of S_f on f should be

$$S_f(\zeta,\tau) \;=\; S(\zeta-\log f, f\tau) \tag{B.2}$$

for some $S\colon \mathbb{R}^2 \to [0,\infty)$; a simple approach would be to take S to be a bivariate Gaussian probability density function with mean $(0,0)$ and to use $p=1$ in the wavelet transform. However, if one already has information regarding how the quickly time-varying the frequencies of the different oscillatory components in the signals are, then it may make sense to alter the dependence of S_f on f accordingly. If one wishes simply to follow the coherence along an individual time-frequency curve as described in Remak 11, then it may make sense not to have any dependence of S_f on f.

If well-defined, the wavelet coherence $\mathfrak{c}^{(1)}_{\psi,\kappa,xy}(f,t)$ takes values between 0 and 1. Larger values may indicate the presence of a common oscillatory influence upon the signals x and y around time t, with the frequency at time t being approximately equal to either f or (in the case of a nonlinear oscillatory influence) $\frac{f}{n}$ for some integer n. One can then look at the wavelet energy spectra of x and y to gauge the strength and the linearity or nonlinearity of this oscillatory influence.

Note that the "autocoherence" $\mathfrak{c}^{(1)}_{\psi,\kappa,xx}(f,t)$, wherever it is well-defined, is equal to 1.

One can also choose not to smooth in time-frequency space but just to smooth in time (in which case the value of p makes no difference), and track over time the presence of frequency-intervals of high values. Additionally, one can fix a time-interval $[T_0,T_1]$ over which to look at the overall coherence at each frequency, by integrating against the rectangular window function $\frac{1}{T_1-T_0}\mathbb{1}_{[T_0,T_1]}$ instead of smoothing.

Wavelet-phase coherence

An alternative approach to coherence is to reverse the order of the "smoothing" and the "angle-extracting" operations, so that wavelet amplitudes $|W^{[p]}_{\psi,\kappa,x}(f,\tau)|$ play no role whatsoever: one simply smooths the wavelet phase difference between the signals. The key difference between this approach and the previous approach is that in the previous approach, the measure of coherence between phases is weighted over time in proportion with the instantaneous wavelet amplitudes. (So while the previous approach *is* invariant under time-independent rescaling of wavelet amplitudes, it would not be invariant under time-dependent rescaling of wavelet amplitudes.)

The "phases-only" approach was introduced in [36], using temporal smoothing with rectangular windows. More generally one can smooth in time-frequency space: Given a family $(S_f)_{f>0}$ of integrable functions $S_f\colon \mathbb{R}^2 \to [0,\infty)$ satisfying $S_f(\zeta,\tau) = S_f(\zeta,-\tau)$ for all $\zeta,\tau\in\mathbb{R}$, define the wavelet-phase coherence between x and y by

$$\mathfrak{c}^{(2)}_{\psi,\kappa,xy}(f,t) \;=\; \frac{\left|\int_{\mathbb{R}^2} S_f(\zeta,\tau-t)\langle\, W_{\psi,\kappa,x}(e^\zeta,\tau)\overline{W_{\psi,\kappa,y}(e^\zeta,\tau)}\,\rangle\, d(\zeta,\tau)\right|}{\int_{\mathbb{R}^2} S_f(\zeta,\tau)\,d(\zeta,\tau)} \tag{B.3}$$

where $\langle v \rangle := v/|v|$, for all $f > 0$ and $t \in \mathbb{R}$ where this is well-defined. Note that changing the value of p makes no difference to the results. The default dependence of S_f on f should be as in (B.2), in which case

$$\int_{\mathbb{R}^2} S_f(\zeta, \tau)\, d(\zeta, \tau) \;=\; \frac{1}{f} \int_{\mathbb{R}^2} S(\zeta, \tau)\, d(\zeta, \tau).$$

But again, if one wishes simply to follow the coherence along an individual time-frequency curve as described in Remak 11, then it may make sense not to have any dependence of S_f on f.

As with wavelet coherence, the wavelet-phase coherence $\mathfrak{c}^{(2)}_{\psi,\kappa,xy}(f,t)$ takes a value between 0 and 1, where larger values may indicate to the presence of a common oscillatory influence just as described above for wavelet coherence. Wherever the "autocoherence" $\mathfrak{c}^{(2)}_{\psi,\kappa,xx}(f,t)$ is well-defined, it is equal to 1. Once again, one can choose to smooth only in time and track the presence of frequency-intervals of high values, and one can also fix a time-interval $[T_0, T_1]$ over which to look at the overall coherence at each frequency.

Appendix B.2. Bicoherence

Given either a single signal or two or three simultaneous signals, the wavelet bispectra can be used to search in a time-localised manner for influences coming from a pair of interacting oscillatory processes. For instance, one can integrate a sliding time-frequency-frequency window function with respect to the wavelet bispectrum and look for where the absolute value is high; that is, one can plot a function taking the form

$$(f_1, f_2, t) \;\mapsto\; \left| \int_{\mathbb{R}^2} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau - t) b_{\psi,\kappa,xyz}(e^{\zeta_1}, e^{\zeta_2}, \tau)\, d(\zeta_1, \zeta_2, \tau) \right|$$

(or one can just smooth the wavelet BD with respect to time rather than time-frequency-frequency), and look for high values. However, high values can easily be due to high wavelet amplitudes happening to occur by fluke at a triplet of frequencies satisfying the frequency-sum rule. If the frequencies are time-varying, one way to help overcome this issue is to use a quantity that is invariant under rescaling of wavelet amplitudes and gives an indication of how constant the biphase remains around a given point in time-frequency-frequency space. Constancy of biphase is generally referred to as *bicoherence*. We will describe two possible such quantities measuring bicoherence, namely the third-order analogues of the measures of coherence described in Appendix B.1.

Remark 12. One can extend Remark 11 to the setting of bicoherence: We will give definitions of wavelet coherence $\mathfrak{c}^{(1)}_{\psi,\kappa,xyz}(f_1, f_2, t)$ and wavelet-phase coherence $\mathfrak{c}^{(2)}_{\psi,\kappa,xyz}(f_1, f_2, t)$ defined over (f_1, f_2, t) space, but it may be useful to follow these specifically over a "time-frequency-frequency curve" $((f_{1,t}, f_{2,t}))_{t\in[a,b]}$. In particular, by looking at the wavelet transform of x, y or z, one may be able to trace the temporal variations in frequency of some oscillatory component, e.g. by "ridge-extraction" methods as detailed in [25, 26]. Accordingly, one can:

(i) extract time-frequency curves $(f_{1,t})_{t\in[a,b]}$ and $(f_{2,t})_{t\in[a,b]}$ from the wavelet transforms of x and y respectively; or

(ii) extract time-frequency curves $(f_{1,t})_{t\in[a,b]}$ and $(f_{3,t})_{t\in[a,b]}$ from the wavelet transforms of x and z respectively, with $f_{3,t} > f_{1,t}$, and then define $f_{2,t} = f_{3,t} - f_{1,t}$; or

(iii) extract time-frequency curves $(f_{2,t})_{t\in[a,b]}$ and $(f_{3,t})_{t\in[a,b]}$ from the wavelet transforms of y and z respectively, with $f_{3,t} > f_{2,t}$, and then define $f_{1,t} = f_{3,t} - f_{2,t}$.

Then one can plot $t \mapsto \mathfrak{c}^{(i)}_{\psi,\kappa,xyz}(f_{1,t}, f_{2,t}, t)$ over $[a, b]$, where $\mathfrak{c}^{(i)}_{\psi,\kappa,xyz}$ $(i = 1, 2)$ is the time-frequency-frequency-smoothed version of bicoherence as in (B.4) or (B.8). The same also applies to the "bicorrelations" $\rho^{(i)}_{\psi,\kappa,xyz}$ $(i = 1, 2)$ introduced further below.

Remark 13. A slightly simpler approach to looking for interacting oscillatory contributions via time-evolving bispectral analysis, as described in [28], is as follows: Plot over frequency-frequency space the modulus of the time-averaged wavelet BD over a time-interval $[T_0, T_1]$ of interest, identify points (ν_1, ν_2) around which there is high bispectral content, and then plot the time-evolution of the biphase associated to (ν_1, ν_2) over the time-interval $[T_0, T_1]$. Subintervals during which this biphase remains roughly constant may correspond to interaction between oscillatory processes. One can also plot the "biamplitude" $|b_{\psi,\kappa,xyz}(\nu_1, \nu_2, t)|$ over $t \in [T_0, T_1]$ to look at the time-localised intensity of the bispectral contribution during the time-subintervals on which the biphase remains roughly constant. The limitation of this approach in its "raw form" is that it cannot follow variations of frequency in oscillatory components. Therefore a possible modification as described in [29] is to extract frequencies of oscillatory components as a function of time (e.g. as described in Remark 12) so as to give a time-dependent point $(\nu_{1,t}, \nu_{2,t})$ in frequency-frequency space, and then – provided these frequencies are sufficiently slowly varying – plot the biphase at $(\nu_{1,t}, \nu_{2,t}, t)$ against time t and once again look for time-subintervals during which this remains roughly constant.

In the below, we fix three signals $x(t)$, $y(t)$ and $z(t)$, any two or which or all three of which could be the same.

Wavelet bicoherence

The definition of wavelet bicoherence was introduced in [72, 73], where temporal evolution of bicoherence was tracked by partitioning the signal duration into small time-intervals on each of which the bicoherence at each point in frequency-frequency space was computed according to Eq. (3). The definition can be adapted to the more general framework of time-frequency-frequency smoothing in analogy to the time-frequency smoothing described in Appendix B.1, as follows:

Given a family $(S_{f_1,f_2})_{f_1,f_2>0}$ of bounded integrable functions $S_{f_1,f_2}\colon \mathbb{R}^3 \to [0,\infty)$ satisfying $S_{f_1,f_2}(\zeta_1, \zeta_2, \tau) = S_{f_1,f_2}(\zeta_1, \zeta_2, -\tau)$ for all $\zeta_1, \zeta_2, \tau \in \mathbb{R}$, define the *wavelet bicoherence* of x and y with z by

$$\mathfrak{c}^{(1)}_{\psi,\kappa,xyz}(f_1, f_2, t) \;=\; \frac{\left|\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau - t) W^{[p]}_{\psi,\kappa,x}(e^{\zeta_1}, \tau) W^{[p]}_{\psi,\kappa,y}(e^{\zeta_2}, \tau) \overline{W^{[p]}_{\psi,\kappa,z}(e^{\zeta_1} + e^{\zeta_2}, \tau)}\, d(\zeta_1, \zeta_2, \tau)\right|^2}{\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau - t) |W^{[p]}_{\psi,\kappa,x}(e^{\zeta_1}, \tau) W^{[p]}_{\psi,\kappa,y}(e^{\zeta_2}, \tau)|^2\, d(\zeta_1, \zeta_2, \tau) \int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau - t) |W^{[p]}_{\psi,\kappa,z}(e^{\zeta_1} + e^{\zeta_2}, \tau)|^2\, d(\zeta_1, \zeta_2, \tau)} \tag{B.4}$$

for all $f_1, f_2 > 0$ and $t \in \mathbb{R}$, provided this is well-defined. A natural default dependence of S_{f_1,f_2} on (f_1, f_2) would be

$$S_{f_1,f_2}(\zeta_1, \zeta_2, \tau) \;=\; S(\zeta_1 - \log f_1, \zeta_2 - \log f_2, \min(f_1, f_2)\tau) \tag{B.5}$$

for some $S\colon \mathbb{R}^3 \to [0,\infty)$; a simple approach would be to take S to be a trivariate Gaussian probability density function with mean $(0, 0, 0)$ and to use $p = 1$ in the wavelet transform. However, if one already has information regarding how the quickly time-varying the frequencies of the different oscillatory components in the signals are, then it makes sense to alter the dependence of S_{f_1,f_2}

on (f_1, f_2) accordingly. If one wishes simply to follow the bicoherence along an individual time-frequency-frequency curve as described in Remak 12, then it may make sense not to have any dependence of S_{f_1,f_2} on (f_1, f_2).

If well-defined, the wavelet bicoherence $\mathfrak{c}^{(1)}_{\psi,\kappa,xyz}(f,t)$ takes values between 0 and 1. Larger values may indicate the presence of a coupled-oscillator influence upon the signals x, y and z around time t, with the frequencies at time t being approximately equal to f_1 and f_2, or f_1 and $f_1 + f_2$, or f_2 and $f_1 + f_2$, or possibly more generally any pair of frequencies whose integer span includes f_1 and f_2. One can then look at the wavelet bispectrum $\mathfrak{b}_{\psi,\kappa,xyz}$ to gauge the strength and nonlinearity of this coupled-oscillator influence.

Note that unlike for second-order spectra, the *wavelet autobicoherence* $\mathfrak{c}^{(1)}_{\psi,\kappa,xxx}(f_1, f_2, t)$ is just as free to take any value in $[0, 1]$ as wavelet cross-bicoherences are.

One can also choose not to smooth in time-frequency-frequency space but just to smooth in time (in which case the value of p makes no difference), and track over time the presence of frequency-frequency regions of high values. Following the same temporal rescaling as in (B.5), this would be as follows: given a bounded integrable even function $S' \colon \mathbb{R} \to [0, \infty)$, one defines the associated wavelet bicoherence by

$$(f_1, f_2, t) \mapsto \frac{\left| \int_{\mathbb{R}} S'(\min(f_1, f_2)\tau - t) W_{\psi,\kappa,x}(f_1, \tau) W_{\psi,\kappa,y}(f_2, \tau) \overline{W_{\psi,\kappa,z}(f_1 + f_2, \tau)}\, d\tau \right|^2}{\int_{\mathbb{R}} S'(\min(f_1, f_2)\tau - t) |W_{\psi,\kappa,x}(f_1, \tau) W_{\psi,\kappa,y}(f_2, \tau)|^2\, d\tau \int_{\mathbb{R}} S'(\min(f_1, f_2)\tau - t) |W_{\psi,\kappa,z}(f_1 + f_2, \tau)|^2\, d\tau} \tag{B.6}$$

wherever this is well-defined. Likewise, one can look at the overall wavelet bicoherence across a time-interval $[T_0, T_1]$, namely

$$(f_1, f_2) \mapsto \frac{\left| \int_{T_0}^{T_1} W_{\psi,\kappa,x}(f_1, t) W_{\psi,\kappa,y}(f_2, t) \overline{W_{\psi,\kappa,z}(f_1 + f_2, t)}\, dt \right|^2}{\int_{T_0}^{T_1} |W_{\psi,\kappa,x}(f_1, t) W_{\psi,\kappa,y}(f_2, t)|^2\, dt \int_{T_0}^{T_1} |W_{\psi,\kappa,z}(f_1 + f_2, t)|^2\, dt} \tag{B.7}$$

wherever this is well-defined.

Wavelet-phase bicoherence

We introduce here an alternative approach to bicoherence which works directly with the wavelet phases, which we accordingly call "wavelet-phase bicoherence". (It may also be regarded as the phase-lag-invariant analogue of the "real wavelet biphase" introduced in [33].) The key difference between this approach and the previous approach is that in the previous approach, the measure of bicoherence is weighted over time in proportion with the instantaneous wavelet amplitudes. (So while the previous approach is invariant under time-independent rescaling of wavelet amplitudes, it would not be invariant under time-dependent rescaling of wavelet amplitudes.)

Given a family $(S_{f_1,f_2})_{f_1,f_2>0}$ of integrable functions $S_{f_1,f_2} \colon \mathbb{R}^3 \to [0, \infty)$ satisfying $S_{f_1,f_2}(\zeta_1, \zeta_2, \tau) = S_{f_1,f_2}(\zeta_1, \zeta_2, -\tau)$ for all $\zeta_1, \zeta_2, \tau \in \mathbb{R}$, define the *wavelet-phase bicoherence* of x and y with z by

$$\mathfrak{c}^{(2)}_{\psi,\kappa,xyz}(f_1, f_2, t) \;=\; \frac{\left| \int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau - t) \langle\, W_{\psi,\kappa,x}(e^{\zeta_1}, \tau) W_{\psi,\kappa,y}(e^{\zeta_2}, \tau) \overline{W_{\psi,\kappa,z}(e^{\zeta_1} + e^{\zeta_2}, \tau)} \,\rangle\, d(\zeta_1, \zeta_2, \tau) \right|}{\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1, \zeta_2, \tau)\, d(\zeta_1, \zeta_2, \tau)} \tag{B.8}$$

where $\langle v \rangle := v/|v|$, provided this is well-defined. Note that changing the value of p would make no difference whatsoever to the results. Again, a natural default dependence of S_{f_1,f_2} on (f_1, f_2)

would be as in (B.5), in which case

$$\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau)\,d(\zeta_1,\zeta_2,\tau) \;=\; \frac{1}{\min(f_1,f_2)}\int_{\mathbb{R}^3} S(\zeta_1,\zeta_2,\tau)\,d(\zeta_1,\zeta_2,\tau).$$

But again, if one wishes simply to follow the bicoherence along an individual time-frequency-frequency curve as described in Remak 12, then it may make sense not to have any dependence of S_{f_1,f_2} on (f_1,f_2).

Again, this takes a value between 0 and 1, where larger values may indicate to the presence of a coupled-oscillator influence just as described above for wavelet bicoherence. One can look at the wavelet bispectrum $\mathfrak{b}_{\psi,\kappa,xyz}$ to gauge the strength and nonlinearity of this influence. Once again, the *wavelet-phase autobicoherence* $\mathfrak{c}^{(2)}_{\psi,\kappa,xxx}(f,t)$ is just as free to take any value in $[0,1]$ as wavelet-phase cross-bicoherences are.

One can also choose just to smooth temporally, and look for regions of high values. Following the same temporal rescaling as in (B.5), this would be as follows: given an integrable even function $S'\colon \mathbb{R}\to[0,\infty)$, one defines the associated wavelet-phase bicoherence by

$$(f_1,f_2,t) \;\mapsto\; \frac{\min(f_1,f_2)\left|\int_{\mathbb{R}} S'(\min(f_1,f_2)\tau - t)\langle\, W_{\psi,\kappa,x}(f_1,\tau)W_{\psi,\kappa,y}(f_2,\tau)\overline{W_{\psi,\kappa,z}(f_1+f_2,\tau)}\,\rangle\,d\tau\right|}{\int_{\mathbb{R}} S'(\tau)\,d\tau} \tag{B.9}$$

wherever this is well-defined. Likewise, one can look at the overall wavelet-phase bicoherence across a time-interval $[T_0,T_1]$, namely

$$(f_1,f_2) \;\mapsto\; \frac{\left|\int_{T_0}^{T_1}\langle\, W_{\psi,\kappa,x}(f_1,t)W_{\psi,\kappa,y}(f_2,t)\overline{W_{\psi,\kappa,z}(f_1+f_2,t)}\,\rangle\,dt\right|}{T_1-T_0} \tag{B.10}$$

wherever this is well-defined.

A remark about unidirectional coupling

Sometimes unidirectionality of coupling of oscillators can be suggested by analysis of bicoherences and bispectral densities. Slightly generalising the description in [30]: Suppose we have signals x, y and z, where y and z could be the same as each other or different, recorded from a system that contains an oscillatory process consisting of two phase-coupled oscillators. Suppose that

- analysis of properties of the bispectrum $\mathfrak{b}_{\psi,\kappa,xyz}$ around (f_1,f_2,t) detects the presence of this oscillatory process;
- analysis of properties of the bispectrum $\mathfrak{b}_{\psi,\kappa,xyx}$ around (f_1,f_2,t) does not detect the presence of this oscillatory process;
- analysis of properties of the cross-energy spectrum $\mathfrak{p}_{\psi,\kappa,xy}$ around (f_2,t) does not detect the presence of this oscillatory process.

This may indicate that only one of the two oscillators is detected within the signal x, and that this oscillator is unidirectionally driving the other oscillator.

Quadratic phase coupling and bicorrelation

In some contexts where bicoherence is observed, i.e. where the biphase remains roughly constant over a time-interval of interest, it may also be useful to ascertain whether or not the constant value at which the biphase roughly lies is zero, since this may help indicate in the context the type of interaction between the oscillations [57]. Bispectral content around a point in frequency-frequency space where the biphase remains constant at 0 is often known as *quadratic phase coupling* (although the terminology is not entirely uniform[6]).

One can define a version of wavelet bicoherence and wavelet-phase bicoherence that specifically measures *low-phase-shift bicoherence*, by replacing the modulus with the real part in the definitions of bicoherence. We will refer to such quantities as "bicorrelations". Using the notation of (B.4), we define the *wavelet bicorrelation* by

$$\rho^{(1)}_{\psi,\kappa,xyz}(f_1,f_2,t) \;=\; \frac{\mathrm{Re}\Big(\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau-t) W^{[p]}_{\psi,\kappa,x}(e^{\zeta_1},\tau) W^{[p]}_{\psi,\kappa,y}(e^{\zeta_2},\tau) \overline{W^{[p]}_{\psi,\kappa,z}(e^{\zeta_1}+e^{\zeta_2},\tau)}\, d(\zeta_1,\zeta_2,\tau)\Big)}{\sqrt{\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau-t)|W^{[p]}_{\psi,\kappa,x}(e^{\zeta_1},\tau)W^{[p]}_{\psi,\kappa,y}(e^{\zeta_2},\tau)|^2\, d(\zeta_1,\zeta_2,\tau) \int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau-t)|W^{[p]}_{\psi,\kappa,z}(e^{\zeta_1}+e^{\zeta_2},\tau)|^2\, d(\zeta_1,\zeta_2,\tau)}} \tag{B.11}$$

and using the notation of (B.8), we define the *wavelet-phase bicorrelation* by

$$\rho^{(2)}_{\psi,\kappa,xyz}(f_1,f_2,t) \;=\; \frac{\mathrm{Re}\Big(\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau-t) \langle\, W_{\psi,\kappa,x}(e^{\zeta_1},\tau) W_{\psi,\kappa,y}(e^{\zeta_2},\tau) \overline{W_{\psi,\kappa,z}(e^{\zeta_1}+e^{\zeta_2},\tau)}\,\rangle\, d(\zeta_1,\zeta_2,\tau)\Big)}{\int_{\mathbb{R}^3} S_{f_1,f_2}(\zeta_1,\zeta_2,\tau)\, d(\zeta_1,\zeta_2,\tau).} \tag{B.12}$$

Once again, we can also choose just to smooth in time, as well as to look at the overall bicorrelation over a fixed time-interval $[T_0,T_1]$. These temporally smoothed or averaged versions of the wavelet-phase bicorrelation were introduced in [33] under the name "real wavelet biphase".

All of these bicorrelation values are between -1 and 1: values close to 1 are indicative of quadratic phase coupling.

Appendix C. A linear-frequency-derived definition of local wavelet bispectra

Recall the six regions $\Gamma_1, -\Gamma_1, \Gamma_2, -\Gamma_2, \Gamma_3, -\Gamma_3$ defined in Sec. 2.5, as shown in Fig. 3.

The definition of the wavelet bispectrum in Sec. 3.1, and hence also of the local wavelet bispectrum in Sec. 3.3, was derived by logarithmic-frequency consideration as in (44). We will now derive an alternative possible definition of local wavelet bispectra, in terms of the analogous linear-frequency consideration. In analogy to Remark 3, consideration of dimensional analysis suggests that this will require using $p = \frac{1}{3}$ in the definition of the wavelet transform.

Throughout this appendix, we fix a wavelet function ψ fulfilling

$$\int_0^1 \frac{\hat{\psi}(r)}{r^{\frac{4}{3}}}\, dr \;<\; \infty \quad \text{and} \quad \int_1^\infty \frac{\hat{\psi}(r)}{r^{\frac{2}{3}}}\, dr \;<\; \infty. \tag{C.1}$$

[6] For example, [33] uses the term *quadratic phase coupling with delay* for bispectral content where the biphase is constant at a non-zero value, and *quadratic phase uncoupling* for power components at frequencies f_1, f_2 and f_1+f_2 without the associated constancy of biphase.

For each $\lambda \in \mathbb{R} \setminus \{0,1\}$, define

$$\tilde{D}_\psi(\lambda) \;=\; \begin{cases} \int_0^\infty\int_0^\infty \frac{\mathring{\psi}\left(\frac{r_1}{\lambda}\right)\mathring{\psi}\left(\frac{r_2}{1-\lambda}\right)\mathring{\psi}(r_1+r_2)}{[r_1r_2(r_1+r_2)]^{\frac{2}{3}}}\,dr_1dr_2 & \lambda \in (0,1) \\ \int_0^\infty\int_0^\infty \frac{\mathring{\psi}\left(\frac{r_1}{-\lambda}\right)\mathring{\psi}(r_2)\mathring{\psi}\left(\frac{r_1+r_2}{1-\lambda}\right)}{[r_1r_2(r_1+r_2)]^{\frac{2}{3}}}\,dr_1dr_2 & \lambda \in (-\infty,0) \\ \int_0^\infty\int_0^\infty \frac{\mathring{\psi}(r_1)\mathring{\psi}\left(\frac{r_2}{\lambda-1}\right)\mathring{\psi}\left(\frac{r_1+r_2}{\lambda}\right)}{[r_1r_2(r_1+r_2)]^{\frac{2}{3}}}\,dr_1dr_2 & \lambda \in (1,\infty). \end{cases} \tag{C.2}$$

This can equivalently be expressed as

$$\tilde{D}_\psi\Big(\tfrac{f_1}{f_1+f_2}\Big) \;=\; \begin{cases} \int_0^\infty\int_0^\infty \frac{\hat{\psi}\left(\frac{|f_1|}{\xi_1}\right)\hat{\psi}\left(\frac{|f_2|}{\xi_2}\right)\hat{\psi}\left(\frac{|f_1+f_2|}{\xi_1+\xi_2}\right)}{[\xi_1\xi_2(\xi_1+\xi_2)]^{\frac{2}{3}}}\,d\xi_1d\xi_2 & (f_1,f_2) \in \Gamma_1 \cup -\Gamma_1 \\ \int_0^\infty\int_0^\infty \frac{\hat{\psi}\left(\frac{|f_1|}{\xi_1}\right)\hat{\psi}\left(\frac{|f_2|}{\xi_1+\xi_2}\right)\hat{\psi}\left(\frac{|f_1+f_2|}{\xi_2}\right)}{[\xi_1\xi_2(\xi_1+\xi_2)]^{\frac{2}{3}}}\,d\xi_1d\xi_2 & (f_1,f_2) \in \Gamma_2 \cup -\Gamma_2 \\ \int_0^\infty\int_0^\infty \frac{\hat{\psi}\left(\frac{|f_1|}{\xi_1+\xi_2}\right)\hat{\psi}\left(\frac{|f_2|}{\xi_2}\right)\hat{\psi}\left(\frac{|f_1+f_2|}{\xi_1}\right)}{[\xi_1\xi_2(\xi_1+\xi_2)]^{\frac{2}{3}}}\,d\xi_1d\xi_2 & (f_1,f_2) \in \Gamma_3 \cup -\Gamma_3. \end{cases} \tag{C.3}$$

It is easy to see from (C.3) that $\tilde{D}_\psi\Big(\frac{f_1}{f_1+f_2}\Big)$ is symmetric in f_1 and f_2, and thus $\tilde{D}_\psi(\lambda) = \tilde{D}_\psi(1-\lambda)$ for all $\lambda \in \mathbb{R} \setminus \{0,1\}$. We see that $\tilde{D}_\psi(\lambda) > 0$ for all $\lambda \in \mathbb{R} \setminus \{0,1\}$ by considering

$$\begin{array}{ll} (r_1,r_2) = (\lambda c, (1-\lambda)c) & \lambda \in (0,1) \\ (r_1,r_2) = (-\lambda c, c) & \lambda \in (-\infty,0) \\ (r_1,r_2) = (c, (\lambda-1)c) & \lambda \in (-\infty,0) \end{array}$$

with $c \in \mathring{\psi}^{-1}((0,\infty))$ in the integrand in (C.2).

Proposition 14. *The map $\lambda \mapsto \tilde{D}_\psi(\lambda)$ is continuous on $\mathbb{R} \setminus \{0,1\}$, and $\tilde{D}_\psi(\lambda) \to 0$ as $\lambda \to 0$ and as $\lambda \to 1$. Moreover, if*

$$\sup_{r>0} \hat{\psi}(r) r^{\frac{2}{3}} \;<\; \infty \tag{C.4}$$

then $\tilde{D}_\psi(\lambda) \to 0$ as λ tends to ∞ in $\hat{\mathbb{R}}$.

Proof. Since $\tilde{D}_\psi(\lambda) = \tilde{D}_\psi(1-\lambda)$, we can restrict to $\lambda \in (-\infty,0) \cup (0,1)$ without loss of generality. For $\lambda \in (0,1)$, applying the transformation $\frac{\lambda}{r_1} =: r_1$ and $\frac{1-\lambda}{r_2} =: r_2$ gives

$$\lambda^{-\frac{1}{3}}(1-\lambda)^{\frac{1}{3}}\tilde{D}_\psi(\lambda) \;=\; \int_0^\infty\int_0^\infty \frac{\hat{\psi}(r_1)\hat{\psi}(r_2)\mathring{\psi}\left(\frac{\lambda}{r_1}+\frac{1-\lambda}{r_2}\right)}{(r_1r_2)^{\frac{4}{3}}\left(\frac{\lambda}{(1-\lambda)r_1}+\frac{1}{r_2}\right)^{\frac{2}{3}}}\,dr_1dr_2, \tag{C.5}$$

and for $\lambda \in (-\infty,0)$, applying the transformation $\frac{-\lambda}{r_1} =: r_1$ and $\frac{1}{r_2} =: r_2$ gives

$$(-\lambda)^{-\frac{1}{3}}\tilde{D}_\psi(\lambda) \;=\; \int_0^\infty\int_0^\infty \frac{\hat{\psi}(r_1)\hat{\psi}(r_2)\mathring{\psi}\left(\frac{1}{1-\lambda}\left(\frac{-\lambda}{r_1}+\frac{1}{r_2}\right)\right)}{(r_1r_2)^{\frac{4}{3}}\left(\frac{-\lambda}{r_1}+\frac{1}{r_2}\right)^{\frac{2}{3}}}\,dr_1dr_2. \tag{C.6}$$

The integrands in (C.5) and (C.6) are dominated by the λ-independent integrable function

$$(r_1, r_2) \;\mapsto\; \left(\max_{r>0} \hat{\psi}(r)\right) \cdot \frac{\hat{\psi}(r_1)}{r_1^{\frac{4}{3}}} \cdot \frac{\hat{\psi}(r_2)}{r_2^{\frac{2}{3}}}.$$

Therefore $\lambda \mapsto \tilde{D}_\psi(\lambda)$ is continuous on $(-\infty, 0) \cup (0, 1)$ and $\tilde{D}_\psi(\lambda) \to 0$ as $\lambda \to 0$. Now suppose (C.4) is satisfied. We can re-express (C.6) as

$$(-\lambda)^{-\frac{1}{3}}(1-\lambda)^{\frac{2}{3}}\tilde{D}_\psi(\lambda) \;=\; \int_0^\infty\!\!\int_0^\infty \frac{\hat{\psi}(r_1)\hat{\psi}(r_2)\mathring{\psi}\big(\frac{1}{1-\lambda}\big(\frac{-\lambda}{r_1}+\frac{1}{r_2}\big)\big)}{(r_1 r_2)^{\frac{4}{3}}\big(\frac{1}{1-\lambda}\big(\frac{-\lambda}{r_1}+\frac{1}{r_2}\big)\big)^{\frac{2}{3}}}\, dr_1 dr_2. \tag{C.7}$$

Here, the integrand is dominated by the λ-independent integrable function

$$(r_1, r_2) \;\mapsto\; \left(\sup_{r>0} \hat{\psi}(r) r^{\frac{2}{3}}\right) \cdot \frac{\hat{\psi}(r_1)}{r_1^{\frac{4}{3}}} \cdot \frac{\hat{\psi}(r_2)}{r_2^{\frac{4}{3}}},$$

and so $\tilde{D}_\psi(\lambda) \to 0$ as λ tends via the negative axis to ∞. □

Remark 15. For $\lambda \in (0, 1)$, the formula (42) for $D_\psi(\lambda)$ can be expressed in a similar form to (C.2), namely

$$D_\psi(\lambda) \;=\; \int_0^\infty\!\!\int_0^\infty \frac{\mathring{\psi}(\frac{r_1}{\lambda})\mathring{\psi}(\frac{r_2}{1-\lambda})\mathring{\psi}(r_1+r_2)}{r_1 r_2}\, dr_1 dr_2,$$

in which case the integrand converges pointwise to 0 as $\lambda \to 0$ or 1, just as it does for (C.2); however, the conditions of the dominated convergence theorem are not fulfilled as they were for the expression (42) (as described in Remark 7).

For a third-order analogue of (22a), since the integration is against linear frequency, we can consider negative frequencies as well as positive frequencies. (Regarding (22a) itself, second-order spectral densities on the negative-frequency axis are simply the complex conjugate of their reflection, and so the same value C_ψ applies to both positive and negative frequencies.) Following (37) to be able to include negative frequencies, we can define the following third-order analogue of (22a), whose veracity is straightforwardly verifiable using (9):

(A) Given signals

$$\begin{aligned} x(t) &= A_1 \cos(2\pi\nu_1 t + \phi_1) \\ y(t) &= A_2 \cos(2\pi\nu_2 t + \phi_2) \\ z(t) &= A_3 \cos(2\pi(\nu_1+\nu_2)t + \phi_1 + \phi_2 - \theta) \end{aligned}$$

with $\nu_1, \nu_2 > 0$, writing $(f_1, f_2) := (\nu_1, \nu_2) \in \Gamma_1$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{i\theta} \;=\; \frac{1}{\kappa^2\tilde{D}_\psi\big(\frac{f_1}{f_1+f_2}\big)} \int_{\Gamma_1} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1, t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2, t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2, t)}\, d(\tilde{f}_1, \tilde{f}_2)$$

and writing $(f_1, f_2) := (-\nu_1, -\nu_2) \in -\Gamma_1$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{-i\theta} \;=\; \frac{1}{\kappa^2\tilde{D}_\psi\big(\frac{f_1}{f_1+f_2}\big)} \int_{-\Gamma_1} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1, t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2, t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2, t)}\, d(\tilde{f}_1, \tilde{f}_2)$$

for any $t \in \mathbb{R}$.

(B) Given signals

$$\begin{aligned} x(t) &= A_1 \cos(2\pi\nu_1 t + \phi_1) \\ y(t) &= A_3 \cos(2\pi(\nu_1+\nu_2)t + \phi_1 + \phi_2 - \theta) \\ z(t) &= A_2 \cos(2\pi\nu_2 t + \phi_2) \end{aligned}$$

with $\nu_1, \nu_2 > 0$, writing $(f_1, f_2) := (\nu_1, -(\nu_1+\nu_2)) \in \Gamma_2$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{i\theta} = \frac{1}{\kappa^2 \tilde{D}_\psi\left(\frac{f_1}{f_1+f_2}\right)} \int_{\Gamma_2} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1,t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2,t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2,t)}\, d(\tilde{f}_1,\tilde{f}_2)$$

and writing $(f_1, f_2) := (-\nu_1, \nu_1+\nu_2) \in -\Gamma_2$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{-i\theta} = \frac{1}{\kappa^2 \tilde{D}_\psi\left(\frac{f_1}{f_1+f_2}\right)} \int_{-\Gamma_2} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1,t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2,t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2,t)}\, d(\tilde{f}_1,\tilde{f}_2)$$

for any $t \in \mathbb{R}$.

(C) Given signals

$$\begin{aligned} x(t) &= A_3 \cos(2\pi(\nu_1+\nu_2)t + \phi_1 + \phi_2 - \theta) \\ y(t) &= A_2 \cos(2\pi\nu_2 t + \phi_2) \\ z(t) &= A_1 \cos(2\pi\nu_1 t + \phi_1) \end{aligned}$$

with $\nu_1, \nu_2 > 0$, writing $(f_1, f_2) := (-(\nu_1+\nu_2), \nu_2) \in \Gamma_3$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{i\theta} = \frac{1}{\kappa^2 \tilde{D}_\psi\left(\frac{f_1}{f_1+f_2}\right)} \int_{\Gamma_3} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1,t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2,t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2,t)}\, d(\tilde{f}_1,\tilde{f}_2)$$

and writing $(f_1, f_2) := (\nu_1+\nu_2, -\nu_2) \in -\Gamma_3$ we have

$$\tfrac{1}{8}A_1A_2A_3e^{-i\theta} = \frac{1}{\kappa^2 \tilde{D}_\psi\left(\frac{f_1}{f_1+f_2}\right)} \int_{-\Gamma_3} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(\tilde{f}_1,t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(\tilde{f}_2,t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(\tilde{f}_1+\tilde{f}_2,t)}\, d(\tilde{f}_1,\tilde{f}_2)$$

for any $t \in \mathbb{R}$.

From here, the definition of linear-frequency wavelet bispectral density that would be analogous to (45) would be

$$(f_1, f_2, t) \mapsto \left(\kappa^2 \tilde{D}_\psi\left(\tfrac{f_1}{f_1+f_2}\right)\right)^{-1} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(f_1,t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(f_2,t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(f_1+f_2,t)}$$

defined (modulo null sets) on the whole of $\mathbb{R}^3$. However, since the normalisation $\tilde{D}_\psi\left(\frac{f_1}{f_1+f_2}\right)^{-1}$ blows up towards ∞ at the boundaries of the six regions of frequency-frequency space (Proposition 14), it is highly inadvisable to take this approach to defining a wavelet bispectrum. Nonetheless, one can still define *local* bispectra analogous to those defined in Sec. 3.3.

Suppose we have a region of frequency-frequency space $S \subset \mathbb{R}^2$ contained within a cone of the form

$$\{(f_1, f_2) : \tfrac{f_1}{f_1+f_2} \in [\lambda - \varepsilon, \lambda + \varepsilon]\}$$

where the closed interval $[\lambda - \varepsilon, \lambda + \varepsilon]$ is contained in $\mathbb{R} \setminus \{0, 1\}$. One can define a notion of local wavelet bispectral (linear-frequency) density $\tilde{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda} \colon S \times \mathbb{R} \to \mathbb{C}$ by

$$\tilde{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda}(f_1, f_2, t) \;=\; (\kappa^2 \tilde{D}_\psi(\lambda))^{-1} W^{[\frac{1}{3}]}_{\psi,\kappa,x}(f_1, t) W^{[\frac{1}{3}]}_{\psi,\kappa,y}(f_2, t) \overline{W^{[\frac{1}{3}]}_{\psi,\kappa,z}(f_1 + f_2, t)}. \tag{C.8}$$

We then define

$$\mathfrak{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda}(A) \;=\; \int_A \tilde{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda}(f_1, f_2, t)\, d(f_1, f_2, t) \tag{C.9}$$

for any $A \in \mathcal{B}_{\tilde{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda}}$. (This can also be computed in terms of logarithmic-frequency integration by multiplying the integrand by the exponential of the sum of the two integrator log-frequencies.) Note as in Sec. 3.3 that the localised bispectrum $\mathfrak{b}^{\mathrm{lin}}_{\psi,\kappa,xyz;\lambda}$ is meaningless if the frequency resolution of the wavelet is so low that S does not capture virtually all the wavelet bispectral content arising from bispectral contributions of interest within S. If one is not concerned with "absolute" values for bispectral results but only comparisons between bispectral results, then the normalising factor $(\kappa^2 \tilde{D}_\psi(\lambda))^{-1}$ can be removed from (C.8). Nonetheless, analogously to Sec. 3.3, the wavelet transform terms must still be with $p = \frac{1}{3}$.

Acknowledgements

The authors are grateful for the comments of two anonymous referees, which led to considerable improvement in the readability and flow of the paper as well as the inclusion of Sec. 6. The research reported in this paper has been funded the EPSRC grant EP/M006298/1 *A device to detect and measure the progression of dementia by quantifying the interactions between neuronal and cardiovascular oscillations*, and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 642563.

References

[1] A. Akan and R. B. U. Artan. Time-varying bispectral analysis of nonstationary signals. In *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings.*, volume 1, pages 569–572, 2003.

[2] H. Bartelt, A. W. Lohmann, and B. Wirnitzer. Phase and amplitude recovery from bispectra. *Appl. Opt.*, 23(18):3121–3129, 1984. doi: 10.1364/AO.23.003121.

[3] M. J. Berridge and P. E. Rapp. A comparative survey of the function, mechanism and control of cellular oscillators. *J. Exp. Biol.*, 81(1):217–279, 1979.

[4] B. Boashash, editor. *Time-Frequency Signal Analysis and Processing* (Second Edition). Academic Press, Oxford, 2016.

[5] V. Chandran. Time-varying bispectral analysis of visually evoked multi-channel EEG. *EURASIP J. Adv. Signal Process.*, 2012(1):140, 2012. doi: 10.1186/1687-6180-2012-140.

[6] H. Chen, X. Tang, R. Zhang, and J. Gao. Effect of bottom slope on the nonlinear triad interactions in shallow water. *Ocean Dyn.*, 68:469–483, 2018. doi: 10.1007/s10236-018-1143-y.

[7] P. Clemson, G. Lancaster, and A. Stefanovska. Reconstructing time-dependent dynamics. *Proc. IEEE*, 104(2):223–241, 2016. doi: 10.1109/JPROC.2015.2491262.

[8] P. T. Clemson and A. Stefanovska. Discerning non-autonomous dynamics. *Phys. Rep.*, 542 (4):297–368, 2014. doi: 10.1016/j.physrep.2014.04.001.

[9] P. T. Clemson, S. Petkoski, T. Stankovski, and A. Stefanovska. Coupled nonautonomous oscillators. In P. E. Kloeden and C. Pötzsche, editors, *Nonautonomous Dynamical Systems in the Life Sciences*, pages 163–197. Springer International Publishing, Cham, 2013. ISBN 978-3-319-03080-7. doi: 10.1007/978-3-319-03080-7_5.

[10] A. V. Dandawate and G. B. Giannakis. Nonparametric polyspectral estimators for kth-order (almost) cyclostationary processes. *IEEE Trans. Inf. Theory*, 40(1):67–84, 1994. doi: 10.1109/18.272456.

[11] G. Dong, Y. Ma, M. Perlin, X. Ma, B. Yu, and J. Xu. Experimental study of wave-wave nonlinear interactions using the wavelet-based bicoherence. *Coast. Eng.*, 55(9):741–752, 2008. doi: 10.1016/j.coastaleng.2008.02.015.

[12] T. Dudok de Wit and V. V. Krasnosel'skikh. Wavelet bicoherence analysis of strong plasma turbulence at the Earth's quasiparallel bow shock. *Phys. Plasmas*, 2(11):4307–4311, 1995. doi: 10.1063/1.870985.

[13] L. Dyrud, B. Krane, M. Oppenheim, H. L. Pécseli, K. Schlegel, J. Trulsen, and A. W. Wernik. Low-frequency electrostatic waves in the ionospheric E-region: a comparison of rocket observations and numerical simulations. *Ann. Geophys.*, 24(11):2959–2979, 2006. doi: 10.5194/angeo-24-2959-2006.

[14] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394, 1987. doi: 10.1364/JOSAA.4.002379.

[15] J. R. Fonollosa and C. T. Nikias. Wigner higher order moment spectra: Definition, properties, computation and application to transient signal analysis. *IEEE Trans. Signal Process.*, 41(1): 245–266, 1993. doi: 10.1109/TSP.1993.193143.

[16] W. A. Gardner and C. M. Spooner. The cumulant theory of cyclostationary time-series. I. Foundation. *IEEE Trans. Signal Process.*, 42(12):3387–3408, 1994. doi: 10.1109/78.340775.

[17] Z. Ge. Significance testing for wavelet bicoherence and its application in analyzing nonlinearity in turbulent shear flows. *Phys. Rev. E*, 81:056311, 2010. doi: 10.1103/PhysRevE.81.056311.

[18] N. L. Gerr. Introducing a third-order Wigner distribution. *Proc. IEEE*, 76(3):290–292, 1988. doi: 10.1109/5.4410.

[19] T. H. Goldsmith and M. H. M. Goldsmith. The interpretation of intracellular measurements of membrane potential, resistance, and coupling in cells of higher plants. *Planta*, 143(3):267–274, 1978. doi: 10.1007/BF00391997.

[20] K. Gurley, T. Kijewski, and A. Kareem. First- and higher-order correlation detection using wavelet transforms. *J. Eng. Mech.*, 129(2):188–201, 2003. doi: 10.1061/(ASCE)0733-9399(2003)129:2(188).

[21] L. J. Hadjileontiadis. EEG-based tonic cold pain characterization using wavelet higher order spectral features. *IEEE Trans. Biomed. Eng.*, 62(8):1981–1991, 2015. doi: 10.1109/TBME.2015.2409133.

[22] L. J. Hadjileontiadis. Continuous wavelet transform and higher-order spectrum: combinatory potentialities in breath sound analysis and electroencephalogram-based pain characterization. *Philos. Trans. R. Soc. A*, 376(2126), 2018. doi: 10.1098/rsta.2017.0249.

[23] J. Han and N. Jiang. Dynamics evolution investigation of Mack mode instability in a hypersonic boundary layer by bicoherence spectrum analysis. *Chin. Phys. Lett.*, 29(7):074703, 2012. doi: 10.1088/0256-307x/29/7/074703.

[24] K. Hasselman, W. Munk, and G. MacDonald. Bispectra of ocean waves. In M. Rosenblatt, editor, *Proceedings of the Symposium on Time Series Analysis (Brown University, June 11-14, 1962)*, pages 125–139. John Wiley & Sons, New York, 1963.

[25] D. Iatsenko, P. V. McClintock, and A. Stefanovska. Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms. *Digit. Signal Process.*, 42:1–26, 2015. doi: 10.1016/j.dsp.2015.03.004.

[26] D. Iatsenko, P. McClintock, and A. Stefanovska. Extraction of instantaneous frequencies from ridges in time–frequency representations of signals. *Signal Process.*, 125:290–303, 2016. doi: 10.1016/j.sigpro.2016.01.024.

[27] J. Jamšek and A. Stefanovska. The cardio-respiratory couplings observed in the LDF signal using wavelet bispectrum. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4072–4075, 2007. doi: 10.1109/IEMBS.2007.4353228.

[28] J. Jamšek, A. Stefanovska, P. V. E. McClintock, and I. A. Khovanov. Time-phase bispectral analysis. *Phys. Rev. E*, 68:016201, 2003. doi: 10.1103/PhysRevE.68.016201.

[29] J. Jamšek, A. Stefanovska, and P. V. E. McClintock. Wavelet bispectral analysis for the study of interactions among oscillators whose basic frequencies are significantly time variable. *Phys. Rev. E*, 76:046221, 2007. doi: 10.1103/PhysRevE.76.046221.

[30] J. Jamšek, M. Paluš, and A. Stefanovska. Detecting couplings between interacting oscillators with time-varying basic frequencies: Instantaneous wavelet bispectrum and information theoretic approach. *Phys. Rev. E*, 81:036207, 2010. doi: 10.1103/PhysRevE.81.036207.

[31] G. Kaiser. *A Friendly Guide to Wavelets*. Birkhäuser Boston, Inc., Boston, MA, 1994. ISBN 0-8176-3711-7.

[32] H. Knutsson, C.-F. Westin, and G. Granlund. Local multiscale frequency and bandwidth estimation. In *Proceedings of 1st International Conference on Image Processing*, volume 1, pages 36–40, 1994. doi: 10.1109/ICIP.1994.413270.

[33] S. Kontaxis, J. Lázaro, E. Gil, P. Laguna, and R. Bailón. Assessment of quadratic nonlinear cardiorespiratory couplings during tilt-table test by means of real wavelet biphase. *IEEE Trans. Biomed. Eng.*, 66(1):187–198, 2019. doi: 10.1109/TBME.2018.2821182.

[34] A. A. Koronovskiĭ and A. E. Khramov. Wavelet bicoherence analysis as a method for investigating coherent structures in an electron beam with an overcritical current. *Plasma Phys. Rep.*, 28(8):666–681, 2002. doi: 10.1134/1.1501324.

[35] D. Kumar, R. Jadeja, and S. Pande. Wavelet bispectrum-based nonlinear features for cardiac murmur identification. *Cogent Eng.*, 5(1):1–12, 2018. doi: 10.1080/23311916.2018.1502906.

[36] J.-P. Lachaux, E. Rodriguez, M. Le Van Quyen, A. Lutz, J. Martinerie, and F. J. Varela. Studying single-trials of phase-synchronous activity in the brain. *Int. J. Bifurc. Chaos Appl. Sci. Eng.*, 10(10):2429–2439, 2000. doi: 10.1142/S0218127400001560.

[37] J.-P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le Van Quyen, J. Martinerie, and F. Varela. Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Neurophysiol. Clin.*, 32(3):157–174, 2002. doi: 10.1016/S0987-7053(02)00301-5.

[38] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska. Surrogate data for hypothesis testing of physical systems. *Phys. Rep.*, 748:1–60, 2018. doi: 10.1016/j.physrep.2018.06.001.

[39] Y. Larsen. Wavelet-Polyspectra. Master's thesis, University of Tromsø, Department of Physics, Norway, 1999.

[40] Y. Larsen and A. Hanssen. Wavelet-polyspectra: analysis of non-stationary and non-Gaussian/non-linear signals. In *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No.00TH8496)*, pages 539–543, 2000.

[41] Y. Larsen, A. Hanssen, B. Krane, H. L. Pécseli, and J. Trulsen. Time-resolved statistical analysis of nonlinear electrostatic fluctuations in the ionospheric E region. *J. Geophys. Res. Space Phys.*, 107(A1), 2002. doi: 10.1029/2001JA900125.

[42] X. Li, D. Li, L. J. Voss, and J. W. Sleigh. The comodulation measure of neuronal oscillations with general harmonic wavelet bicoherence and application to sleep analysis. *NeuroImage*, 48 (3):501–514, 2009. doi: 10.1016/j.neuroimage.2009.07.008.

[43] Y. Li, J. Lin, X. Wang, and Y. Lei. Biphase randomization wavelet bicoherence for mechanical fault diagnosis. *Measurement*, 49:407–420, 2014. ISSN 0263-2241. doi: 10.1016/j.measurement.2013.12.012.

[44] Y. Li, X. Wang, J. Lin, and S. Shi. A wavelet bicoherence-based quadratic nonlinearity feature for translational axis condition monitoring. *Sensors*, 14:2071–2088, 2014. doi: 10.3390/s140202071.

[45] J. M. Lilly and S. C. Olhede. Generalized Morse wavelets as a superfamily of analytic wavelets. *IEEE Trans. Signal Process.*, 60(11):6036–6041, 2012. doi: 10.1109/TSP.2012.2210890.

[46] A. W. Lohmann and B. Wirnitzer. Triple correlations. *Proc. IEEE*, 72(7):889–901, 1984. doi: 10.1109/PROC.1984.12946.

[47] C. Nikias and A. Petropulu. *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*. Prentice Hall signal processing series. PTR Prentice Hall, 1993. ISBN 9780136782100.

[48] Y. Okada, Y. Doida, G. Roy, W. Tsuchiya, K. Inouye, and A. Inouye. Oscillations of membrane potential in L cells. *J. Membrane Biol.*, 35(1):319–335, 1977. doi: 10.1007/BF01869957.

[49] Y. Okada, G. Roy, W. Tsuchiya, Y. Doida, and A. Inouye. Oscillations of membrane potential in L cells. *J. Membrane Biol.*, 35(1):337–350, 1977. doi: 10.1007/BF01869958.

[50] S. Patel. *The role of membrane potential dynamics in cell behaviours: Investigating the membrane potential dynamics in the Jurkat and HMEC-1 cell lines using the continuous wavelet transform*. PhD thesis, Division of Biomedical and Life Sciences, Lancaster University, 2015.

[51] R. J. Perry and M. G. Amin. On computing and implementing the running bispectra. *IEEE Trans. Signal Process.*, 43(4):1017–1021, 1995. doi: 10.1109/78.376859.

[52] A. Pidde. *Dynamics of the membrane potential: Studies of the membrane potential of Jurkat cells*. PhD thesis, Department of Physics, Lancaster University, 2020.

[53] P. D. Purnamasari, A. A. P. Ratna, and B. Kusumoputro. Relative wavelet bispectrum feature for alcoholic EEG signal classification using artificial neural network. In *2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering*, pages 154–158, 2017. doi: 10.1109/QIR.2017.8168473.

[54] R. Raghavan, X. Chen, K.-P. Yip, D. J. Marsh, and K. H. Chon. Interactions between TGF-dependent and myogenic oscillations in tubular pressure and whole kidney blood flow in both SDR and SHR. *Am. J. Physiol. Renal Physiol.*, 290(3):F720–F732, 2006. doi: 10.1152/ajprenal.00205.2005. PMID: 16219915.

[55] B. Schack, H. Witte, M. Helbig, C. Schelenz, and M. Specht. Time-variant non-linear phase-coupling analysis of EEG burst patterns in sedated patients during electroencephalic burst suppression period. *Clin. Neurophysiol.*, 112(8):1388–1399, 2001. doi: 10.1016/S1388-2457(01)00577-6.

[56] K. Schwab, M. Eiselt, C. Schelenz, and H. Witte. Time-variant parametric estimation of transient quadratic phase couplings during electroencephalographic burst activity. *Methods Inf. Med.*, 44(3):374–383, 2005. doi: 10.1267/METH05030374.

[57] C. G. Scully, N. Mitrou, B. Braam, W. A. Cupples, and K. H. Chon. Detecting interactions between the renal autoregulation mechanisms in time and space. *IEEE Trans. Biomed. Eng.*, 64(3):690–698, 2017. doi: 10.1109/TBME.2016.2569453.

[58] M. Shen, Y. Liu, F. H. Y. Chan, and P. J. Beadle. Novel approach for time-varying bispectral analysis of non-stationary EEG signals. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 829–832, 2005. doi: 10.1109/IEMBS.2005.1616543.

[59] L. W. Sheppard, A. Stefanovska, and P. V. E. McClintock. Testing for time-localized coherence in bivariate data. *Phys. Rev. E*, 85:046205, 2012. doi: 10.1103/PhysRevE.85.046205.

[60] A. Spicher, W. J. Miloch, L. B. N. Clausen, and J. I. Moen. Plasma turbulence and coherent structures in the polar cap observed by the ICI-2 sounding rocket. *J. Geophys. Res. Space Phys.*, 120:10,959–10,978, 2015. doi: 10.1002/2015JA021634.

[61] C. M. Spooner and W. A. Gardner. The cumulant theory of cyclostationary time-series. II. Development and applications. *IEEE Trans. Signal Process.*, 42(12):3409–3429, 1994. doi: 10.1109/78.340776.

[62] J. R. Stack, R. G. Harley, and T. G. Habetler. An amplitude modulation detector for fault diagnosis in rolling element bearings. In *IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02*, volume 4, pages 3377–3382, 2002.

[63] T. Subba Rao and K. C. Indukumar. Spectral and wavelet methods for the analysis of nonlinear and nonstationary time series. *J. Franklin Inst.*, 333(3):425–452, 1996. doi: 10.1016/0016-0032(96)00011-7.

[64] A. Swami. Third-order Wigner distributions: definitions and properties. In *[Proceedings] ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 3081–3084, 1991. doi: 10.1109/ICASSP.1991.150106.

[65] A. Swami, G. B. Giannakis, and G. Zhou. Bibliography on higher-order statistics. *Signal Process.*, 60(1):65–126, 1997. doi: 10.1016/S0165-1684(97)00065-0.

[66] S. A. Taplidou and L. J. Hadjileontiadis. Nonlinear analysis of wheezes using wavelet bicoherence. *Comput. Biol. Med.*, 37(4):563–570, 2007. doi: 10.1016/j.compbiomed.2006.08.007.

[67] C. Torrence and P. J. Webster. Interdecadal changes in the ENSO-monsoon system. *J. Clim.*, 12(8):2679–2690, 1999. doi: 10.1175/1520-0442(1999)012⟨2679:ICITEM⟩2.0.CO;2.

[68] R. K. Tripathy and S. Dandapat. Automated detection of heart ailments from 12-lead ECG using complex wavelet sub-band bi-spectrum features. *Healthc. Technol. Lett.*, 4(2):57–63, 2017. doi: 10.1049/htl.2016.0089.

[69] J. W. Tukey. What can data analysis and statistics offer today? In N. A. of Sciences, editor, *Ocean Wave Spectra: Proceedings of a Conference (Easton, Maryland, May 1–4, 1961)*, pages 347–350. Prentice Hall, Englewood Cliffs, New Jersey, 1963.

[70] J. W. Tukey. The spectral representation and transformation properties of the higher moments of stationary time series. In D. R. Brillinger, editor, *The Collected Works of John W. Tukey, Volume 1*, pages 165–184. Chapman & Hall, London, 1984.

[71] Z. Vahabi, R. Amirfattahi, F. Shayegh, and F. Ghassemi. Online epileptic seizure prediction using wavelet-based bi-phase correlation of electrical signals tomography. *Int. J. Neural Syst.*, 25(06):1550028, 2015. doi: 10.1142/S0129065715500288.

[72] B. P. van Milligen, C. Hidalgo, and E. Sánchez. Nonlinear phenomena and intermittency in plasma turbulence. *Phys. Rev. Lett.*, 74:395–398, 1995. doi: 10.1103/PhysRevLett.74.395.

[73] B. P. van Milligen, E. Sánchez, T. Estrada, C. Hidalgo, B. Brañas, B. Carreras, and L. García. Wavelet bicoherence: A new turbulence analysis tool. *Phys. Plasmas*, 2(8):3017–3032, 1995. doi: 10.1063/1.871199.

[74] D.-M. Yang. Wavelet-based bispectra for motor rotor fault detection. In *2008 Eighth International Conference on Intelligent Systems Design and Applications*, volume 1, pages 603–607, 2008. doi: 10.1109/ISDA.2008.23.

[75] B. Yu, Y. Ma, X. Ma, and G. Dong. Laboratory study of the nonlinear transformation of irregular waves over a mild slope. *China Ocean Eng.*, 28:489–500, 2014. doi: 10.1007/s13344-014-0040-5.