

Characterization of Analytic Wavelet Transforms and a New Phaseless Reconstruction Algorithm

Nicki Holighaus, Günther Koliander, Zdeněk Průša, and Luis Daniel Abreu

Abstract—We obtain a characterization of all wavelets leading to analytic wavelet transforms (WT). The characterization is obtained as a by-product of the theoretical foundations of a new method for wavelet phase reconstruction from magnitude-only coefficients. The cornerstone of our analysis is an expression of the partial derivatives of the continuous WT, which results in phase-magnitude relationships similar to the short-time Fourier transform (STFT) setting and valid for the generalized family of Cauchy wavelets. We show that the existence of such relations is equivalent to analyticity of the WT up to a multiplicative weight and a scaling of the mother wavelet. The implementation of the new phaseless reconstruction method is considered in detail and compared to previous methods. It is shown that the proposed method provides significant performance gains and a great flexibility regarding accuracy versus complexity. Additionally, we discuss the relation between scalogram reassignment operators and the wavelet transform phase gradient and present an observation on the phase around zeros of the WT.

Index Terms—gradient theorem, numerical integration, phase reconstruction, short-time Fourier transform, wavelet transform, phase derivative, Cauchy-Riemann equations.

I. INTRODUCTION

Time-frequency and time-scale representations are fundamental tools in many areas of signal analysis and signal processing, ranging from medical data [1], [2], damage or fault detection in materials [3] and machines [4], to image [5], [6] and audio processing [7]–[9]. Such representations are usually complex-valued and admit a natural decomposition into magnitude and phase components, with the phase containing crucial information about the analyzed signal. However, phase information is often discarded in favor of the magnitude information, from which, supposedly, the desired information is more readily obtained. On the other hand, whenever synthesis from the representation coefficients is desired, the phase is crucial for the quality of the synthesized signal.

In recent years, the problems of phase retrieval and phaseless reconstruction for time-frequency and time-scale dictionaries have attracted considerable attention, leading to theoretical results for the feasibility of phase retrieval in different

contexts [10]–[15] and various algorithms [16]–[21] (see also the survey [22]). Many of these algorithms, e.g., [20], attempt to construct, iteratively or directly, an appropriate phase to match the magnitude-only representation coefficients, before finally performing a regular synthesis step.

In many applications, in particular in audio signal processing, phaseless reconstruction in time-frequency and time-scale dictionaries as proposed in, e.g., [20], [23], is successfully applied. These applications include, but are not restricted to audio and speech synthesis [24]–[26], source separation [27]–[29], as well as pitch and time-scale modifications [30]–[32]. In these scenarios, either reconstruction from a phaseless representation is desired or the given phase has been invalidated in the course of processing and must be replaced.

Despite a relevant research activity regarding the phase behavior of time-frequency and time-scale representations, with particular incidence in the short-time Fourier transform (STFT) [33]–[35], applications that consider and/or modify the phase are relatively scarce. Prime examples of analysis tools that successfully use phase information are the so-called reassignment and synchrosqueezing methods [36]–[40] that deform the signal representation using a vector field obtained from the phase gradient of the representation.

For the STFT with Gaussian generator, the notion that phase and magnitude carry equally important information has been made precise by Portnoff [34] and later by Auger and Flandrin [33]. They show that in this specific case, the phase gradient is completely characterized by the gradient of the (logarithmically scaled) magnitude.

Contributions: Our two main contributions are a characterization of all analytic wavelet transforms (WT) and a new method for wavelet phase reconstruction from magnitude-only coefficients. Here, the notion of analytic WTs is used in the sense that the WT is an analytic function of the upper half-plane (up to a signal-independent factor). It is common to use the same terminology for wavelet transforms generated by analytic wavelets [41], i.e., wavelets whose Fourier transform vanishes at negative frequencies. We discuss the connection between these two notions in detail.

The result characterizing all analytic wavelet transforms shows that the most general wavelet leading to an analytic transform has a Fourier transform given by

$$\xi^{\frac{\alpha-1}{2}} e^{-2\pi\gamma\xi} e^{i\beta\log\xi} \quad (1)$$

for positive frequencies and we assume that the Fourier transform vanishes for negative frequencies. The appearance of the β -dependent hyperbolic chirp in (1) may be surprising since, to our knowledge, only the case $\beta = 0$ has been associated with analytic functions in the literature (see [42]). It is commonly

N. Holighaus*, G. Koliander, Z. Průša, and L. D. Abreu are with the Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12–14, 1040 Vienna, Austria, email: nicki.holighaus@oeaw.ac.at (corresponding address), guenther.koliander@oeaw.ac.at, zdenek.prusa@oeaw.ac.at, labreu@kfs.oeaw.ac.at

Accompanying web page (sound examples, Matlab code, color figures) <http://lthfat.github.io/notes/053>

This work was supported by the Austrian Science Fund (FWF): Y 551-N13, I 3067-N30, and P 31225-N32 and the Vienna Science and Technology Fund (WVTF): MA16-053.

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

referred to as *Cauchy wavelet* and historically associated with affine coherent states. However, the more general wavelets given by (1) have been shown to minimize a time-scale counterpart of Heisenberg uncertainty and are also known as “Klauder wavelets” [43]. The problem of characterizing all analytic WTs has been considered and partially solved in [44], where it was also shown that the Gaussian is essentially the only window leading to an analytic STFT.

We first discuss several aspects of the WT phase, loosely following the structure of [33]. Specifically, we express the phase and log-magnitude derivatives in terms of pointwise quotients of 3 different WTs. The mother wavelets used in these transforms can be derived directly from the mother wavelet of the original WT. These expressions can be used, e.g., to estimate the local group delay and instantaneous scale (or frequency). Furthermore, they impose conditions on the wavelets leading to analytic wavelet transforms, and this is then used to show that the class of wavelets satisfying this condition are the generalized Cauchy wavelets (1). The corresponding Cauchy Riemann (CR) equations provide a relation between the phase gradient and the (log-)magnitude gradient. Following [35], we also discuss the singular behavior of the phase close to zeros of the WT and the implications of our results for wavelet reassignment and ridge analysis. Finally, we discuss the relation between analytic wavelets, i.e., wavelets that can be extended to analytic functions on the upper half plane, and wavelets resulting in an analytic WT.

In the second part of the contribution, we implement a method for reconstruction from magnitude-only wavelet coefficients. More specifically, we use a discrete approximation of the derived phase-magnitude relations to modify the phase gradient heap integration algorithm [45]–[47]. The method is evaluated using Cauchy wavelets of multiple orders, obtaining favorable results. In the evaluation, we further examine (uniform) decimation and the number of scales at which the WT is sampled, controlling the transform redundancy.

Structure of the Paper: In Section II, we present our expressions for the derivatives of the log-magnitude and the phase of the WT. These expressions are used in Section III to obtain our result characterizing the wavelets that lead to analytic WTs. Furthermore, we prove direct relations between the log-magnitude and phase derivatives. In Section IV, we provide a time-frequency interpretation of the corresponding WT. Further applications to pole behavior and scalogram reassignment of the expression provided in Section II are presented in Sections V-A and V-B, respectively. In Section V-C, we discuss the relation between analytic wavelets and our analytic WT. Finally, we apply the phase-magnitude relations to the problem of phaseless reconstruction. In Section VI, we formally introduce the discretization of our results and describe the phaseless reconstruction algorithm; in Section VII, we conduct experiments on real data and compare our method to previous approaches toward phaseless reconstruction.

II. LOG-MAGNITUDE AND PHASE DERIVATIVES OF THE WAVELET TRANSFORM

Fix a function $\psi \in \mathbf{L}^2(\mathbb{R})$ such that its Fourier transform $\widehat{\psi}$ vanishes almost everywhere on \mathbb{R}^- . The continuous WT

(CWT) of a function (or signal) $s \in \mathbf{L}^2(\mathbb{R})$ with respect to the *mother wavelet* ψ is defined as

$$W_\psi s(x, y) = \langle s, \mathbf{T}_x \mathbf{D}_y \psi \rangle = \frac{1}{\sqrt{y}} \int_{\mathbb{R}} s(t) \overline{\psi\left(\frac{t-x}{y}\right)} dt, \quad (2)$$

for all¹ $x \in \mathbb{R}$, $y \in \mathbb{R}^+$. Here, \mathbf{T}_x and \mathbf{D}_y denote the translation and dilation operators, respectively, given by $(\mathbf{T}_x s)(t) = s(t-x)$, and $(\mathbf{D}_y s)(t) = y^{-1/2} s(t/y)$ for all $t \in \mathbb{R}$.

The CWT can be represented in terms of its magnitude $M_\psi^s := |W_\psi s| \geq 0$ and phase $\phi_\psi^s := \arg(W_\psi s) \in \mathbb{R}$ as usual. With this convention, $\log(W_\psi s) = \log(M_\psi^s) + i\phi_\psi^s$. This straightforward relation is the basis of the following expressions for the partial derivatives of the log-magnitude and phase components, derived in Appendix A.

Theorem 1. *Let $\psi \in \mathbf{L}^2(\mathbb{R})$ with $\widehat{\psi}(\xi) = 0$ for $\xi < 0$ and assume that ψ is continuously differentiable with $\psi', \mathbf{T}\psi' \in \mathbf{L}^2(\mathbb{R})$, where \mathbf{T} , without subscript, denotes the time-weighting operator $(\mathbf{T}s)(t) = ts(t)$. Then, for all $x \in \mathbb{R}$ and $y \in \mathbb{R}^+$ satisfying $M_\psi^s(x, y) > 0$,*

$$\nabla \log(M_\psi^s)(x, y) = \begin{pmatrix} 0 \\ \frac{1}{2y} \end{pmatrix} - \operatorname{Re} \left(\frac{\frac{W_{\psi'} s(x, y)}{y W_\psi s(x, y)}}{\frac{W_{(\mathbf{T}\psi')} s(x, y)}{y W_\psi s(x, y)}} \right) \quad (3)$$

and

$$\nabla \phi_\psi^s(x, y) = -\operatorname{Im} \left(\frac{\frac{W_{\psi'} s(x, y)}{y W_\psi s(x, y)}}{\frac{W_{(\mathbf{T}\psi')} s(x, y)}{y W_\psi s(x, y)}} \right). \quad (4)$$

The partial phase derivatives of the WT are often related to the local instantaneous scale, as well as the local group delay, of the analyzed signal, although at least the latter notion is not entirely clear for WTs, see Section V-B. The second order derivatives, which can be obtained similar to the first order derivatives in Appendix A, describe the variation of these quantities across phase space and thus proved useful as well, see [48]–[50].

Formulas (3)–(4) can be used for the computation of the partial derivatives by using efficient implementations of the WT. Furthermore, the possible accuracy of direct numerical differentiation of the wavelet transform is limited if the WT can be computed only at certain positions and not everywhere, e.g., in the presence of decimation in either coordinate. In contrast, even if a closed form expression for the derivative of the (time-weighted) wavelet is not available, numerical differentiation of the wavelet is not limited by this constraint.

III. THE PHASE-MAGNITUDE RELATIONSHIP

In general, the observations in Section II do not yield a direct connection between the partial derivatives of the log-magnitude and phase components. However, if the WT is analytic, we can characterize the phase gradient by the log-magnitude gradient. Based on the Cauchy-Riemann (CR) equations, we can construct conditions on the mother wavelet ψ such that the WT of any signal is an analytic function.

¹Although we restrict here to positive scales, there is no technical obstruction to allowing $y \in \mathbb{R} \setminus \{0\}$.

Similar to the analysis of analytic STFTs studied in [44] and the partial study of analytic wavelet transforms in the same contribution, we allow for an (x, y) -dependent factor $f(x, y)$ that is independent of the signal s . This factor can easily be accounted for when applying results from complex analysis to the transformed signal and leads to a significantly less restrictive class of analyticity-inducing wavelets. Furthermore, we want the class of analyticity-inducing wavelets to be invariant under the natural transforms associated with the WT, namely dilation and translation. Thus, we also allow for a constant dilation by $b \in \mathbb{R}^+$ and a translation specified by $a \in \mathbb{R}$ in our analysis.

Theorem 2. *Let $\psi \in \mathbf{L}^2(\mathbb{R})$ with $\widehat{\psi}(\xi) = 0$ for $\xi < 0$. There exist constants $a \in \mathbb{R}$, $b \in \mathbb{R}^+$ and a C^∞ function $f: \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{C}$ with $f(x, y) \neq 0$ such that*

$$h: \{z \in \mathbb{C} : \text{Im}(z) > 0\} \rightarrow \mathbb{C} \quad (5)$$

$$x + iy \mapsto f(x, y) W_\psi s(x - aby, by)$$

is analytic for all $s \in \mathbf{L}^2(\mathbb{R})$, if and only if

$$\widehat{\psi}(\xi) = c \xi^{\frac{\alpha-1}{2}} e^{-2\pi\gamma\xi} e^{i\beta \log \xi} \quad (6)$$

for all $\xi > 0$ and some constants $c \in \mathbb{C}$, $\alpha > -1$, $\beta \in \mathbb{R}$, and $\gamma \in \mathbb{C}$ with $\text{Re}(\gamma) > 0$.

A proof of the theorem is provided in Appendix B. The wavelets ψ specified by (6) are known as Klauder wavelets [43] and are a minor generalization of Cauchy wavelets $\psi^{(\alpha)}$ [42], which are recovered for the choice $\beta = 0$ and $\gamma = 1$. Among other effects, modifying β results in a proportional shift of the temporal concentration of the mother wavelet away from time zero. Furthermore, a change in γ results only in a scale change, dependent on $\text{Re}(\gamma)$, and a time shift, dependent on $\text{Im}(\gamma)$. Disregarding the constant factor $c \in \mathbb{C}$ in (6), we denote the generalized Cauchy wavelets in Theorem 2 by $\psi^{(\alpha, \beta, \gamma)}$ or $\psi^{(\alpha, \beta)}$, if $\gamma = 1$.

If ψ is given by (6), a corresponding choice for h being analytic is $f(x, y) = y^{-\frac{\alpha}{2}} e^{i\beta \log y}$, $a = \text{Im}(\gamma)$, and $b = 1/\text{Re}(\gamma)$, i.e.,

$$x + iy \mapsto y^{-\frac{\alpha}{2}} e^{i\beta \log y} W_\psi s\left(x - \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} y, \frac{y}{\text{Re}(\gamma)}\right). \quad (7)$$

We note that the wavelets are admissible only for $\alpha > 1$.

Theorem 2 can easily be modified to allow for wavelets ψ where $\widehat{\psi}$ does not vanish for negative frequencies. In this case, $\widehat{\psi}$ must satisfy (6) for $\xi > 0$ and

$$\widehat{\psi}(\xi) = c_n(-\xi)^{\frac{\alpha_n-1}{2}} e^{2\pi\gamma_n\xi} e^{i\beta_n \log(-\xi)} \quad (8)$$

for $\xi < 0$ and some constants c_n , α_n , γ_n , and β_n satisfying the same constraints as c , α , γ , and β , respectively.

Based on the CR equations, we obtain a phase-magnitude relation for the WTs using $\psi^{(\alpha, \beta, \gamma)}$.

Theorem 3. *Let ψ be given by (6) with $c = 1$. Then*

$$\frac{\partial}{\partial x} \phi_\psi^s = \frac{\alpha}{2y \text{Re}(\gamma)} - \frac{\frac{\partial}{\partial y} \log(M_\psi^s)}{\text{Re}(\gamma)} + \frac{\text{Im}(\gamma) \frac{\partial}{\partial x} \log(M_\psi^s)}{\text{Re}(\gamma)} \quad (9)$$

and

$$\begin{aligned} \frac{\partial}{\partial y} \phi_\psi^s &= \frac{\alpha \text{Im}(\gamma) - 2\beta}{2y \text{Re}(\gamma)} + \frac{|\gamma|^2 \frac{\partial}{\partial x} \log(M_\psi^s)}{\text{Re}(\gamma)} \\ &\quad - \frac{\text{Im}(\gamma) \frac{\partial}{\partial y} \log(M_\psi^s)}{\text{Re}(\gamma)}. \end{aligned} \quad (10)$$

For $\gamma = 1$, these relations simplify to

$$\frac{\partial}{\partial x} \phi_\psi^s(x, y) = -\frac{\partial}{\partial y} \log(M_\psi^s)(x, y) + \frac{\alpha}{2y} \quad (11)$$

and

$$\frac{\partial}{\partial y} \phi_\psi^s(x, y) = \frac{\partial}{\partial x} \log(M_\psi^s)(x, y) - \frac{\beta}{y}. \quad (12)$$

A proof of the theorem is provided in Appendix C. As a simple consequence of Theorem 3, we obtain for the second order derivatives in the case $\gamma = 1$

$$\frac{\partial^2}{\partial x^2} \log(M_\psi^s)(x, y) + \frac{\partial^2}{\partial y^2} \log(M_\psi^s)(x, y) = -\frac{\alpha}{2y^2} \quad (13)$$

and

$$\frac{\partial^2}{\partial x^2} \phi_\psi^s(x, y) + \frac{\partial^2}{\partial y^2} \phi_\psi^s(x, y) = \frac{\beta}{y^2}. \quad (14)$$

For Cauchy wavelets, it is known that the magnitude uniquely determines the phase up to a constant phase factor. Moreover, this statements even holds after decimation in the scale component and in certain discretized settings [10].

IV. THE CAUCHY WAVELET TRANSFORM AS TIME-FREQUENCY REPRESENTATION

If the mother wavelet ψ is frequency-localized around frequency ξ_b , then we can interpret $W_\psi s(x, y)$ as a time-frequency measurement at frequency $\xi = \xi_b/y$. In the case of the wavelets $\psi^{(\alpha, \beta)}$, this leads to a particularly convenient form of the phase-magnitude relationship. Here, we consider the unique peak of $|\psi^{(\alpha, \beta)}|$ as center frequency, i.e., $\xi_b = \frac{\alpha-1}{4\pi}$. Additionally, instead of the unitary dilation \mathbf{D}_y , we consider the dilation $\tilde{\mathbf{D}}_y s(t) = y^{-1} s(t/y)$ to define

$$\tilde{W}_\psi s(x, \xi) = \langle s, \mathbf{T}_x \tilde{\mathbf{D}}_{\xi_b/\xi} \psi \rangle = \sqrt{\frac{\xi}{\xi_b}} W_\psi s(x, \xi_b/\xi).$$

Using the relations in Section III, it is easy to derive

$$\begin{aligned} \frac{\partial}{\partial x} \tilde{\phi}_\psi^s(x, \xi) &= \frac{\xi^2}{\xi_b} \frac{\partial}{\partial \xi} \log(\tilde{M}_\psi^s)(x, \xi) + \frac{\alpha-1}{2\xi_b} \xi \\ &= \frac{4\pi\xi^2}{\alpha-1} \frac{\partial}{\partial \xi} \log(\tilde{M}_\psi^s)(x, \xi) + 2\pi\xi \end{aligned} \quad (15)$$

and

$$\begin{aligned} \frac{\partial}{\partial \xi} \tilde{\phi}_\psi^s(x, \xi) &= -\frac{\xi_b}{\xi^2} \left(\frac{\partial}{\partial x} \log(\tilde{M}_\psi^s)(x, \xi) - \frac{\beta\xi}{\xi_b} \right) \\ &= -\frac{\alpha-1}{4\pi\xi^2} \frac{\partial}{\partial x} \log(\tilde{M}_\psi^s)(x, \xi) + \frac{\beta}{\xi}, \end{aligned} \quad (16)$$

where \tilde{M}_ψ and $\tilde{\phi}_\psi$ denote the magnitude and phase of \tilde{W}_ψ , respectively. Interestingly, for standard Cauchy wavelets, i.e., $\beta = 0$, these formulas coincide with the phase-magnitude relations for the STFT with a dilated Gaussian (see [45,

Section III) up to the simple change that the constant time-frequency ratio λ is replaced by the frequency-dependent term $\frac{\alpha-1}{4\pi\xi^2}$. The above form (15)–(16) will enable us to adapt the phase reconstruction method presented in [47] to the WT more easily in Section VI below.

The additive term β/ξ in (16) compensates for the fact that $\psi^{(\alpha,\beta)}$ is time-localized around $\frac{2\beta}{\alpha-1}$. In other words, the frequency (or scale) bands in $\tilde{W}_{\psi^{(\alpha,\beta)}}$ (or $W_{\psi^{(\alpha,\beta)}}$) are not time-aligned. Time-alignment can be restored by choosing $\text{Im}(\gamma) = \frac{2\beta}{\alpha-1}$, thus removing the additive term at the cost of introducing directional derivatives of $\log(\tilde{M}_\psi^s)$ in the expression of the phase gradient:

$$\frac{\partial}{\partial x} \tilde{\phi}_\psi^s(x, \xi) = \nabla_{d_1(\xi)} \log(\tilde{M}_\psi^s)(x, \xi) + 2\pi\xi, \quad (17)$$

with $d_1(\xi) = \left(\frac{2\beta}{\alpha-1}, \frac{4\pi\xi^2}{\alpha-1}\right)$ and

$$\frac{\partial}{\partial \xi} \tilde{\phi}_\psi^s(x, \xi) = -\nabla_{d_2(\xi)} \log(\tilde{M}_\psi^s)(x, \xi), \quad (18)$$

with $d_2(\xi) = \left(\frac{\alpha-1}{4\pi\xi^2} + \frac{\beta^2}{(\alpha-1)\pi\xi^2}, \frac{2\beta}{\alpha-1}\right)$.

V. DERIVATIVES AND ANALYTICITY OF THE CONTINUOUS WAVELET TRANSFORM—FURTHER OBSERVATIONS

A. The Phase Around Zeros of the Wavelet Transform

For the STFT with Gaussian window, it was remarked by Auger et al. [33], that the phase has characteristic poles where the STFT is zero and that this fact can be derived from the analyticity of the Bargman transform. In [35], the characteristic pole behavior was proven under weaker conditions, as long as the STFT is smooth enough, more specifically \mathcal{C}^2 or \mathcal{C}^3 .

In fact, the techniques used therein apply to any complex-valued function of two real variables, as long as its higher-order partial derivatives are continuous. In the case of the WT, this can be ensured by selecting a sufficiently smooth and decaying mother wavelet ψ . In particular, if the k -th derivative of ψ weighted by t^l is square integrable, i.e., $\mathbf{T}^l(\frac{\partial^k}{\partial t^k} \psi) \in \mathbf{L}^2(\mathbb{R})$, for all $l, k \in \{0, \dots, K\}$, then $W_\psi s \in \mathcal{C}^K(\mathbb{R} \times \mathbb{R}^+, \mathbb{C})$, for all $s \in \mathbf{L}^2(\mathbb{R})$, cf. Appendix D. This implies the following result.

Theorem 4. *Let $\psi \in \mathbf{L}^2(\mathbb{R})$ and assume that $\mathbf{T}^l(\frac{\partial^k}{\partial t^k} \psi) \in \mathbf{L}^2(\mathbb{R})$, for all $l, k \in \{0, \dots, 2\}$. If $W_\psi s(x_0, y_0) = 0$ and the Jacobian determinant $J_{x_0, y_0} := \det(D(W_\psi s))(x_0, y_0)$ of $W_\psi s$ at (x_0, y_0) is nonzero, then for $\varepsilon > 0$ converging to 0*

$$\text{sgn}(J_{x_0, y_0}) \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial x} \phi_\psi^s(x_0, y_0 + \varepsilon) = \infty, \quad (19)$$

$$\text{sgn}(J_{x_0, y_0}) \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial x} \phi_\psi^s(x_0, y_0 - \varepsilon) = -\infty \quad (20)$$

and

$$\text{sgn}(J_{x_0, y_0}) \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial y} \phi_\psi^s(x_0 + \varepsilon, y_0) = \infty, \quad (21)$$

$$\text{sgn}(J_{x_0, y_0}) \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial y} \phi_\psi^s(x_0 - \varepsilon, y_0) = -\infty. \quad (22)$$

If even $\mathbf{T}^l(\frac{\partial^k}{\partial t^k} \psi) \in \mathbf{L}^2(\mathbb{R})$, for all $l, k \in \{0, \dots, 3\}$, then the limits

$$\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial x} \phi_\psi^s(x_0 + \varepsilon, y_0) = \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial x} \phi_\psi^s(x_0 - \varepsilon, y_0) \quad (23)$$

and

$$\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial y} \phi_\psi^s(x_0, y_0 + \varepsilon) = \lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial y} \phi_\psi^s(x_0, y_0 - \varepsilon) \quad (24)$$

exist and are finite.

After noting that the assumptions imply that $W_\psi s \in \mathcal{C}^2(\mathbb{R} \times \mathbb{R}^+, \mathbb{C})$ (or $\mathcal{C}^3(\mathbb{R} \times \mathbb{R}^+, \mathbb{C})$), the proof of the above result is identical to the proofs of [35, Theorem 4.7–4.9], which only rely on continuous differentiability locally.

B. Scalogram Reassignment and Ridge Points

Reassignment is a technique for sharpening time-frequency and time-scale representations [36]. The reassignment map is a vector field that is used to deform the representation of choice and depends on the representation and the input signal. It has been derived using different methods, e.g., as a constant phase deformation derived from group theoretical properties [37], [38] or by a center of gravity argument relying on the Wigner distribution [36]. In the case of the spectrogram, both notions can be shown to lead to the same reassignment map. However, this is no longer true for the scalogram, i.e., the squared modulus of the WT.

The reassignment map given in [38] relies only on the phase gradient. It is defined as

$$(x, y) \mapsto \left(x + \frac{y^2 \frac{\partial}{\partial y} \phi_\psi^s(x, y)}{\xi_b}, \frac{\xi_b}{\frac{\partial}{\partial x} \phi_\psi^s(x, y)} \right). \quad (25)$$

Using (4), the map can be rewritten leading to

$$(x, y) \mapsto \left(x - \frac{y \text{Im} \left(\frac{W_{(\mathbf{T}\psi)'s}(x, y)}{W_\psi s(x, y)} \right)}{\xi_b}, -\frac{y\xi_b}{\text{Im} \left(\frac{W_{(\mathbf{T}\psi)'s}(x, y)}{W_\psi s(x, y)} \right)} \right). \quad (26)$$

Similar to the expressions provided in [36], (26) can prove useful for the efficient calculation of reassigned scalograms. In particular, a direct computation of the phase gradient from samples of the phase may be quite inaccurate. If the phase-magnitude relations (9) and (10) are satisfied, then inserting these relations into the reassignment map (25) allows scalogram reassignment from the scalogram itself.

In [36], a different reassignment map is given by²

$$(x, y) \mapsto \left(x + y \text{Re} \left(\frac{W_{\mathbf{T}\psi s}(x, y)}{W_\psi s(x, y)} \right), -\frac{y\xi_b}{\text{Im} \left(\frac{W_{\mathbf{T}\psi s}(x, y)}{W_\psi s(x, y)} \right)} \right). \quad (27)$$

The two reassignment maps coincide in the second coordinate, which can be considered a notion of local *instantaneous scale*. However, in the first coordinate, which is often considered an estimate of the local *group delay*, they are quite different. In particular, there seems to be no connection between the first coordinate of (27) and partial derivatives of the WT $W_\psi s$ in general. However, assuming $\mathbf{T}\psi = d_0\psi + d_1\psi'$ for some constants $d_0, d_1 \in \mathbb{C}$, enables an expression of the first coordinate in (27) as a linear combination of the partial derivatives of $\log(M_\psi^s)$ and ϕ_ψ^s . A mother wavelet

²Note that the equality $\psi = h(\bullet) \exp(i\xi_b(\bullet))$ converts between the different WT conventions used.

satisfying this differential equation is, e.g., the *Gabor wavelet* $\psi_G(t) = e^{-t^2/2 + i\xi_b t}$, with $\mathbf{T}\psi_G = i\xi_b \psi_G - (\psi_G)'$. In this case, we can rewrite

$$\begin{aligned} \operatorname{Re} \left(\frac{W_{\mathbf{T}\psi_G} s(x, y)}{W_{\psi_G} s(x, y)} \right) &= -\operatorname{Re} \left(\frac{W_{(\psi_G)'} s(x, y)}{W_{\psi_G} s(x, y)} \right) \\ &= y \frac{\partial}{\partial x} \log(M_{\psi_G}^s)(x, y) \end{aligned} \quad (28)$$

and therefore (27) becomes

$$(x, y) \mapsto \left(x + y^2 \frac{\partial}{\partial x} \log(M_{\psi_G}^s)(x, y), \frac{\xi_b}{\frac{\partial}{\partial x} \phi_{\psi_G}^s(x, y)} \right). \quad (29)$$

Note that the Gabor wavelet does not vanish at negative frequencies. However, the proof of Theorem 1 does not rely on this property, such that the derivation of (29) from (27) remains valid.

The expression of the first coordinate in (29) is not too surprising, since $\mathbf{D}_y \psi_G$ is simply a dilated, modulated Gaussian, and thus this expression could also be obtained using the reassignment operators and phase-magnitude relationship for the Gaussian STFT, as observed in [33].

Similar to reassignment, wavelet ridge analysis [41], [51] attempts to identify a *skeleton* of essential time-scale positions in the wavelet transform. The notion of magnitude ridge points is defined as the points (x, y) , such that $\frac{\partial}{\partial y} \log(y^{-\frac{1}{2}} M_{\psi}^s)(x, y) = 0$ and $\frac{\partial^2}{\partial y^2} \log(y^{-\frac{1}{2}} M_{\psi}^s)(x, y) < 0$. Similarly, phase ridge points are defined as the points (x, y) , such that $\frac{\partial}{\partial x} \phi_{\psi}^s(x, y) - \frac{\xi_b}{y} = 0$ and $\frac{\partial}{\partial y} \left(\frac{\partial}{\partial x} \phi_{\psi}^s(x, y) - \frac{\xi_b}{y} \right) > 0$. It is straightforward to verify that Theorems 2 and 3 imply that the phase and magnitude ridge points coincide if and only if $\psi = c\psi^{(\alpha, \beta, \gamma)}$, with $c \neq 0$, $\alpha > -1$, $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}^+$.

C. Analytic Wavelets and the Analytic Wavelet Transform

To prevent confusion related to other works on WTs, we want to point out the connection between our analytic WT and the WT using analytic wavelets [41], [52]. A wavelet is called analytic if it vanishes almost everywhere on \mathbb{R}^- , i.e., $\hat{\psi}(\xi) = 0$ for $\xi < 0$. The reason for this terminology is that these wavelets can be extended to an analytic function on the upper half-plane by the Paley-Wiener theorem. Furthermore, the WT using an analytic wavelet ψ and at a fixed scale y_0 also has the property that it can be extended to an analytic function on the upper half-plane in a complex variable w , i.e., the function $x \mapsto W_{\psi} s(x, y_0)$ can be extended to an analytic function

$$w \mapsto W_{\psi}^{(a)} s(w, y_0). \quad (30)$$

Now, for $\psi = \psi^{(\alpha, \beta, \gamma)}$, the function in (7) is also analytic. The two functions in (7) and (30) coincide (up to a constant) for $y = \operatorname{Re}(\gamma)y_0$ and $x = w + \operatorname{Im}(\gamma)y_0$ for all $w \in \mathbb{R}$. More specifically,

$$\begin{aligned} (\operatorname{Re}(\gamma)y_0)^{-\frac{\alpha}{2}} e^{i\beta \log(\operatorname{Re}(\gamma)y_0)} W_{\psi}^{(a)} s(w, y_0) \\ = h(w + \operatorname{Im}(\gamma)y_0 + i(\operatorname{Re}(\gamma)y_0)) \end{aligned} \quad (31)$$

for all $w \in \mathbb{R}$ and where both sides are analytic functions in w on the upper half-plane. Thus, they have to coincide everywhere and we see that h describes the analytic continuation

$W_{\psi}^{(a)} s(w, y_0)$ of an arbitrary scale up to some constants and shifts. In particular, for the case of Cauchy wavelets $\psi^{(\alpha)}$, we have

$$y_0^{-\frac{\alpha}{2}} W_{\psi}^{(a)} s(w, y_0) = h(w + iy_0) \quad (32)$$

and the analytic continuation at a given scale only differs by a constant multiple from the analytic continuation at any other scale. This *equivalence* of all analytic continuations (30) is unique to the wavelets $\psi^{(\alpha, \beta, \gamma)}$.

VI. APPLICATION—PHASELESS RECONSTRUCTION FOR THE DISCRETE CONTINUOUS WAVELET TRANSFORM

In the following, we propose and evaluate a method for signal reconstruction from magnitude-only wavelet coefficients. More specifically, the proposed algorithm computes a phase estimate from the given magnitude-only coefficients. After combining the magnitude-only coefficients with the estimated phase, the wavelet transform must be inverted to obtain a time-domain signal. Here, any method that implements reconstruction from wavelet coefficients can be used.

Since arbitrary dilations cannot be naturally transferred to the discrete domain, discrete implementations of the WT can be quite different from each other, see [53]–[57] and references therein. For illustrative purposes and to clarify notation, we shortly sketch an implementation of the discrete WT that follows [58]–[60] closely, with some of the modifications introduced in [9]. In particular, we mimic the dilation operator by sampling the continuous frequency response of the mother wavelet $\psi \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^1(\mathbb{R})$ at the appropriate density. The following description reflects the implementation used in our experiments, see Section VII. We use the terminology *discrete continuous WT (DCWT)* to distinguish this type of discrete WT from methods based on a multiresolution analysis and wavelet bases [53], [57], that are commonly known as discrete wavelet transform (DWT).

A. Discrete Continuous Wavelet Transform

We will denote discretizations of continuous signals by brackets, e.g., the discretized signal $s_d[l] \in \mathbb{C}$ for $l \in \{1, \dots, L\}$ and some $L \in \mathbb{N}$. In this discrete domain, the translation operator acts circularly, i.e., $s_d[l - m]$ is interpreted as $s_d[\operatorname{mod}(l - m, L)]$. Assuming the sampling rate ξ_s , the wavelet at scale $y = \xi_b/\xi$ is derived from the frequency response of the mother wavelet $\psi \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^1(\mathbb{R})$ as

$$\widehat{\psi}_y[k] = \widehat{\psi} \left(\frac{y \xi_s k}{L} \right) = \widehat{\psi} \left(\frac{\xi_b \xi_s k}{L \xi} \right),$$

for $k \in \{-\lfloor L/2 \rfloor, \dots, \lfloor L/2 \rfloor - 1\}$. Naturally, only a finite range of scales can be considered before the wavelet deteriorates either due to the sampling density being too coarse (y large) or its bandwidth approaching ξ_s (y small). Hence, in order to cover the entire frequency range, we introduce an additional low-pass function in the style of [9, Section 3.1.2].

The entire wavelet system is characterized by the minimum scale $y_m \in \mathbb{R}^+$, the scale step³ $2^{1/B}$, with $B \in \mathbb{R}^+$, the

³We choose the commonly used geometric spacing of center frequencies, but the proposed phase reconstruction method remains valid for any center frequency spacing.

number of scales $K \in \mathbb{N}$, and the decimation factor $a_d \in \mathbb{N}$, with $a_d | L$. The corresponding scaled and shifted wavelets are given as

$$\psi_{n,k} = \mathbf{T}_{na_d} \psi_{2^{k/B} y_m} \quad (33)$$

for $k \in \{0, \dots, K-1\}$ and $n \in \{0, \dots, L/a_d - 1\}$. A plateau function $P_{lp} \in \mathbb{C}^L$, centered at 0, specifies the low-pass function as

$$\widehat{\psi}_{lp} = a_d^{-1} P_{lp} \Psi_{lp}, \quad (34)$$

where

$$\Psi_{lp} = \sqrt{\max(\Psi) - \Psi}, \quad \Psi = \sum_{k=0}^{K-1} |\widehat{\psi}_{0,k}|^2. \quad (35)$$

An analysis with the constructed system yields LK/a_d complex-valued coefficients for the wavelet scales and additional L/a_d real-valued coefficients for the low-pass function, for a total redundancy of $(2K+1)/a_d$ when analyzing real-valued signals. With a slight abuse of terminology, we will from now on refer to the proportional quantity K/a_d as the redundancy.

If ψ is smooth and $a_d, 1/B$ are small enough, then the results in [58] imply that the DCWT is invertible. Inversion can be achieved by interpreting the wavelet transform as a filter bank analysis and invoking the frame theory of uniform filter banks [61]–[64] to compute a dual filter bank synthesis. This can be done either directly using dual filters $\tilde{\psi}_k$, or iteratively [9], [65] using conjugate gradient iterations. The consideration of general uniform filter banks is necessary: It is not always possible to find a dual filter bank, which is required to achieve perfect reconstruction, with wavelet structure. Nonetheless, the dual filter bank shares the number of channels $K+1$ and the decimation factor a of the wavelet analysis.

B. Application to Phaseless Reconstruction

For the phase-magnitude relations presented in Theorem 3 to hold, we have to assume a wavelet $\psi^{(\alpha, \beta, \gamma)}$ as in (6). In particular, we will restrict to the case $\gamma = 1$ for simplicity and drop the superscript (α, β) for notational convenience. For the WT and the phase-magnitude relations, we will use the convention introduced in Section IV. The generalization to the full class of wavelets described by (6) is straightforward.

Note that Theorem 3 provides only the phase-derivative and indeed reconstruction can at best be expected to be accurate up to a global phase factor. Furthermore, the reconstruction quality is expected to be worse for low-magnitude areas and thus the proposed algorithm only reconstructs the phase down to a certain magnitude-threshold. Coefficients below that threshold are expected to have little effect on the synthesis and can thus be assigned a random phase. As a consequence of the local, adaptive integration scheme, the reconstructed phase is in fact only expected to be consistent locally with changes by a constant phase factor between these local components. On audio signals, such as the chosen corpus of test data, this change is not expected to have notable perceptual effects. Nonetheless, it is visible in the phase difference (between original and reconstructed phase) in Fig. 3. At low redundancy, which is not well-suited for phase reconstruction in general,

the phase distortion may become more severe (see the lower right corner in Fig. 3), sometimes leading to perceivable distortion.

Assume that the continuous-time signal s is approximately band- and time-limited on $[0, \xi_s)$ and $[0, L/\xi_s)$, respectively. Then, with $s_d[l] = s(l/\xi_s)$, for $l \in \{0, \dots, L-1\}$, $a_d = a\xi_s \in \mathbb{N}$, and $\xi_k = 2^{-k/B} \xi_b/y_m$, we obtain the approximation

$$M_s[n, k] := |\langle s_d, \psi_{n,k} \rangle| \approx \xi_s \tilde{M}_\psi^s(na, \xi_k). \quad (36)$$

Note that, after taking the logarithmic derivative of (36), the normalization by ξ_s becomes irrelevant.

Hence, we have

$$\begin{aligned} \frac{\partial}{\partial x} \tilde{\phi}_\psi^s(na, \xi_k) &= \frac{4\pi\xi_k^2}{\alpha-1} \frac{\partial}{\partial \xi} \log(\tilde{M}_\psi^s)(na, \xi_k) + 2\pi\xi_k \\ &\approx \Delta_{\tilde{\phi}_\psi^{s,x}}[n, k] := \frac{4\pi\xi_k^2}{\alpha-1} \Delta_k(\log(M_s))[n, k] + 2\pi\xi_k, \end{aligned} \quad (37)$$

and

$$\begin{aligned} \frac{\partial}{\partial \xi} \tilde{\phi}_\psi^s(na, \xi_k) &= -\frac{\alpha-1}{4\pi\xi_k^2} \frac{\partial}{\partial x} \log(\tilde{M}_\psi^s)(na, \xi_k) + \frac{\beta}{\xi_k} \\ &\approx \Delta_{\tilde{\phi}_\psi^{s,\xi}}[n, k] := -\frac{\alpha-1}{4\pi\xi_k^2} \Delta_n(\log(M_s))[n, k] + \frac{\beta}{\xi_k}. \end{aligned} \quad (38)$$

Here, Δ_n and Δ_k are appropriate discrete differentiation schemes. For Δ_n , we can use centered differences, i.e.,

$$\Delta_n(M)[n, k] := \frac{\xi_s(M[n+1, k] - M[n-1, k])}{2a_d}. \quad (39)$$

The sampling step in the scale coordinate changes depends on k and weighted centered differences can be used:

$$\begin{aligned} \Delta_k(M)[n, k] &:= \frac{M[n, k+1] - M[n, k]}{2(\xi_{k+1} - \xi_k)} \\ &\quad + \frac{M[n, k] - M[n, k-1]}{2(\xi_k - \xi_{k-1})}. \end{aligned} \quad (40)$$

Now, from $\Delta_{\tilde{\phi}_\psi^{s,x}}$ and $\Delta_{\tilde{\phi}_\psi^{s,\xi}}$, an estimate of the phase of $\tilde{W}_\psi s$ at the sampling points $\{(na, \xi_k)\}_{n,k}$ can be obtained using a quadrature rule considering the variable sampling intervals. That even simple 1-dimensional trapezoidal quadrature provides satisfactory results is illustrated by our experiments, see Section VII.

The integration itself can be performed by a slightly modified Phase Gradient Heap Integration (PGHI) algorithm [45], [47], [66], see Algorithm 1, using, e.g., the following integration rule on the set of neighbors $(n_n, k_n) \in \mathcal{N}_{n,k} := \{(n \pm 1, k), (n, k \pm 1)\}$ of (n, k)

$$\begin{aligned} is(\tilde{\phi}_\psi^s)_{\text{est}}[n_n, k_n] &= (\tilde{\phi}_\psi^s)_{\text{est}}[n, k] + \frac{\xi_{k_n} - \xi_k}{2} \left(\Delta_{\tilde{\phi}_\psi^{s,\xi}}[n, k] + \Delta_{\tilde{\phi}_\psi^{s,\xi}}[n_n, k_n] \right) \\ &\quad + \frac{a_d(n_n - n)}{2\xi_s} \left(\Delta_{\tilde{\phi}_\psi^{s,x}}[n, k] + \Delta_{\tilde{\phi}_\psi^{s,x}}[n_n, k_n] \right). \end{aligned} \quad (41)$$

When inserting (39) and (40) into (41), the absolute scale of the center frequencies ξ_k and sampling rate ξ_s becomes unimportant and only their ratio enters the quadrature (41). Hence, by considering relative frequencies ξ_k/ξ_s , the algorithm is valid independent of the assumed sampling rate.

Algorithm 1: Wavelet Phase Gradient Heap Integration

Input: Magnitude M_s of wavelet coefficients, estimates $\Delta_{\psi}^{\tilde{\phi},x,s}$ and $\Delta_{\psi}^{\tilde{\phi},\xi,s}$ of the partial phase derivatives, relative tolerance tol .

Output: Phase estimate $(\tilde{\phi}_{\psi}^s)_{\text{est}}$.

```

1  $abstol \leftarrow tol \cdot \max(M_s[n, k]);$ 
2 Create set  $\mathcal{I} = \{(n, k) : M_s[n, k] > abstol\};$ 
3 Assign random values to  $(\tilde{\phi}_{\psi}^s)_{\text{est}}(n, k)$  for  $k \notin \mathcal{I}$ ;
4 Construct a self-sorting max heap [67] for  $(n, k)$  pairs;
5 while  $\mathcal{I}$  is not  $\emptyset$  do
6   if heap is empty then
7     Move  $(n_m, k_m) = \arg \max_{(n,k) \in \mathcal{I}} (M_s[n, k])$  from  $\mathcal{I}$ 
       into the heap;
8      $(\tilde{\phi}_{\psi}^s)_{\text{est}}(n_m, k_m) \leftarrow 0$ ;
9   end
10  while heap is not empty do
11     $(n, k) \leftarrow$  remove the top of the heap;
12    foreach  $(n_n, k_n)$  in  $\mathcal{N}_{n,k} \cap \mathcal{I}$  do
13      Compute  $(\tilde{\phi}_{\psi}^s)_{\text{est}}(n_n, k_n)$  by means of (41);
14      Move  $(n_n, k_n)$  from  $\mathcal{I}$  into the heap;
15    end
16  end
17 end

```

Once the phase estimate $(\tilde{\phi}_{\psi}^s)_{\text{est}}$ has been computed, it is combined with the magnitude by $W_{s,\text{est}} := M_s e^{i(\tilde{\phi}_{\psi}^s)_{\text{est}}}$. Subsequently, a time-domain signal can be synthesized as usual, e.g., using a dual filter bank.

VII. EXPERIMENTS

To test and evaluate the proposed method, we performed two experiments, described and discussed below. Both experiments were run on the first 5 seconds of all 70 test signals from the Sound Quality Assessment Material recordings for subjective tests provided by the European Broadcasting Union (SQAM database) [68]. For wavelet analysis and synthesis, we used the filter bank methods in the open source Large Time-Frequency Analysis Toolbox (LTFAT [69], <http://lftat.github.io/>), where our implementation of *Wavelet Phase Gradient Heap Integration* (WPGHI) is available by using the ‘wavelet’ flag in `filterbankconstphase`. A function to generate the wavelet filters and scripts for generating the individual experiments and figures are provided on the manuscript website <http://lftat.github.io/notes/053/>, where the resulting audio files for all experiment conditions can be found as well. Experimental conditions were restricted to classical Cauchy wavelets, i.e., $\beta = 0$ and $\gamma = 1$.

Thus, the WT parameters used in the experiments are (α, a_d, K) , where α is the order of the Cauchy wavelet, a_d is the decimation step and K is the number of frequency channels (or scales) used before adding the lowpass filter. As quantitative error measure, we employ (wavelet) spectral convergence [70], i.e., the relative mean squared error (in dB)

between the wavelet coefficient magnitude of the target signal s_t and the proposed solution s_p :

$$SC(s_p, s_t) = 20 \log_{10} \frac{\|M_{s_p} - M_{s_t}\|}{\|M_{s_t}\|}.$$

It should be noted that the wavelet coefficient magnitude in the above formula was computed using the same parameter set (α, a_d, K) for which phaseless reconstruction was attempted.⁴

A. Experiment I—Comparison to Previous Methods

To study the performance of the proposed algorithm in comparison with previous methods for phaseless recovery from wavelet coefficients, we selected three settings of the WT parameters (α, a_d, K) . For all settings, the channel center frequencies were geometrically spaced in $\frac{\xi_s}{20} \cdot [2^{-6}, 2^{3.3}]$. We considered the following tuples of parameters: $(30, 5, 100)$, $(300, 12, 240)$, and $(3000, 20, 400)$. Here, the ratio $K/a_d = 20$ was fixed in all cases, but a_d and K were adjusted to accommodate for bandwidth variations with changing α .

The dimensionality of the considered audio data renders a systematic comparison to existing implementations of some established methods, e.g., [16], [19] unfeasible, such that we resort to *fast Griffin-Lim* [20], [23] as baseline method. We compare four different methods: wavelet PGHI (WPGHI, *proposed*), filter bank PGHI (FBPGHI, [47]), fast Griffin-Lim with random initialization (R-FGLIM, [23]) and fast Griffin-Lim initialized with the result of WPGHI (W-FGLIM, *proposed*). Fast Griffin-Lim was restricted to at most 150 iterations.⁵ Spectral convergence of the four methods on all test signals is shown in Figure 1 for the different parameter sets (α, a_d, K) . The means and standard deviation across all signals, for every method and parameters set are shown in Table I.

It can be seen that on average, plain WPGHI (*proposed*) outperforms both R-FGLIM and FBPGHI on all parameter sets, although FBPGHI approaches the other methods for larger values of α . Moreover, W-FGLIM (*proposed*) shows significant improvements over either WPGHI or R-FGLIM. Looking at the individual signals more closely, we see in Figure 1 that there are only very few cases in which R-FGLIM yields a better result than W-FGLIM. For $\alpha = 3000$, all methods show comparable performance, with the exception of W-FGLIM, which still provides a clear advantage.

The figures and the computed standard deviations both suggest that methods that perform well on average are prone to larger performance fluctuation between individual signals. However, there is a small set of signals on which all methods perform badly, indicating that the fault is with the wavelet

⁴Although spectral convergence is in some cases sensitive to parameter changes, preliminary tests showed that, for fixed α , comparable spectral convergence is achieved with respect to representations with varying a_d and K . On the other hand, when α is changed as well, then the value of $SC(s_p, s_t)$ may change dramatically, such that comparing the results across different choices of α may be misleading.

⁵Although we are mainly interested in the reconstruction quality and not in computational performance, it is worth mentioning that the solutions of plain WPGHI and FBPGHI are computed in a small fraction of the time required for executing either R-FGLIM or W-FGLIM, even if, at the cost of reconstruction quality, the maximum number of iterations was significantly reduced.

representation rather than the method applied for phaseless reconstruction. Notably, signal 65 (corresponding to the right-most signal in Figures 1–2) from the SQAM database, which, in the considered range, only contains a sustained, extremely low-pitched note is badly resolved by the employed WT and yields the worst spectral convergence of all signals, for any of the employed methods.

Informal listening showed that, for $\alpha = 3000$ and an untrained listener, the obtained reconstructions are mostly indistinguishable from the original signal. For $\alpha = 300$ and, more prominently, $\alpha = 30$, FBPGHI often introduces a characteristic pitch-shift, likely due to a wrongly estimated time-direction phase derivative, while R-FGLIM suffers from undesired frequency modulation artifacts; both types of distortion are most easily audible in simple signals, such as signal 1 (sine wave) and 4 (electronic gong) of the SQAM database and not present in the reconstructions provided by WPGHI. For $\alpha = 30$ and some select cases, e.g., signals 16 (clarinet) and 32 (triangle), audible distortions were present in the solutions by W-FGLIM, but not those by plain WPGHI, indicating that the observed improvement in terms of spectral convergence does not necessarily provide a perceptual improvement. To confirm our observations and to form their own opinion, the reader is invited to visit the manuscript webpage <http://lftfat.github.io/notes/053/>.

As a side note, the audio examples⁶ provided with [45] have clearly audible artifacts for signal 54 (male German speech), for STFT-based PGHI and some competing algorithms. These artifacts are not present in any of the reconstructions we obtained using WPGHI, R-FGLIM, or W-FGLIM, for any considered parameter set, indicating that in some cases, usage of the WT may provide a genuine advantage over the STFT.

α	method	WPGHI	FBPGHI	R-FGLIM	W-FGLIM
30	mean	-33.9512	-18.2916	-28.9164	-40.2932
	std	7.7632	2.0358	3.2641	7.0526
300	mean	-36.6393	-26.8642	-29.3628	-42.7550
	std	7.9517	3.4095	4.0747	9.0264
3000	mean	-38.5527	-34.8240	-32.0108	-44.9260
	std	6.7961	5.5137	5.4821	10.6349

Table I: Means and standard deviation of spectral convergence for the considered methods and parameter sets.

B. Experiment II—Changing the Redundancy

In a second set of experiments, we investigate the influence of the redundancy K/a_d on the performance of the proposed methods WPGHI and W-FGLIM. For this purpose, we fixed an intermediate value for the order parameter, setting $\alpha = 1000$ and consider redundancies $K/a_d \in \{3, 5, 10, 30\}$. Here, however, the reconstruction quality does not only depend on the accuracy of WPGHI on the given magnitude coefficients, but also on the robustness of the synthesis by the dual system. This robustness can be quantified by the so-called frame bound ratio of the respective wavelet system (for details see [53], [71], [72]).

In the ranges considered and for fixed K/a_d , the number K of frequency channels has a larger influence on

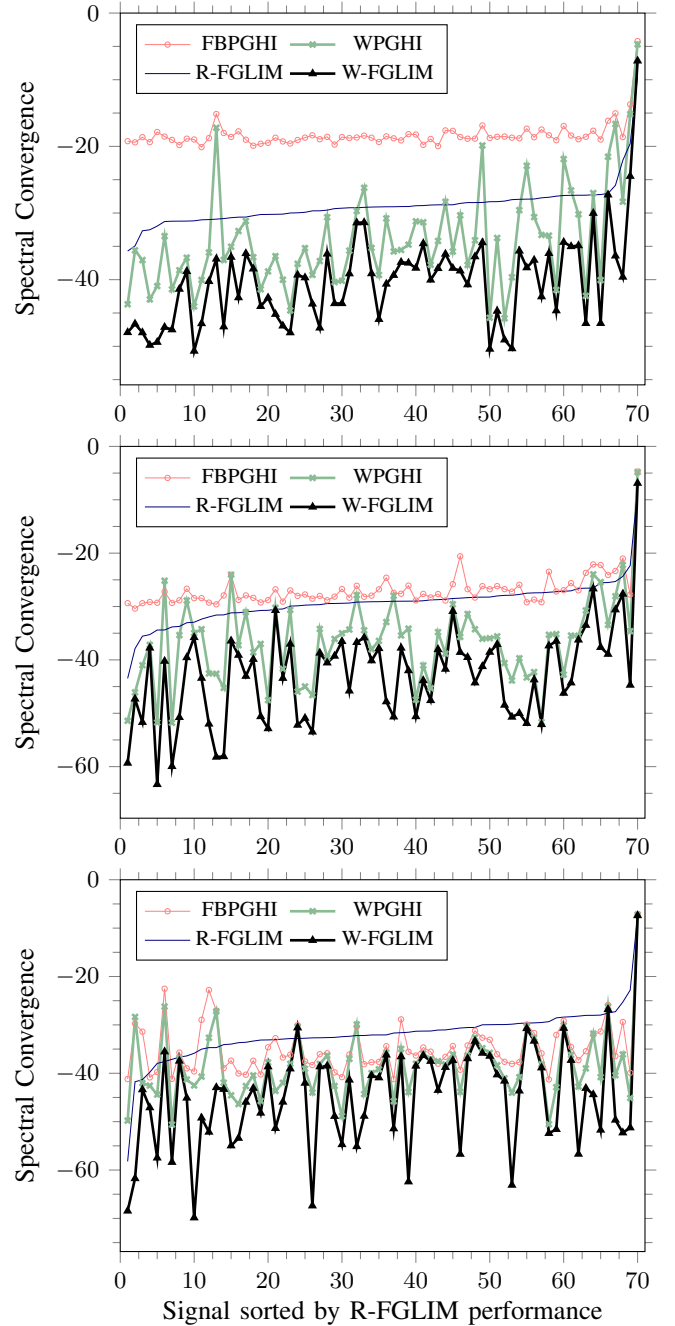


Figure 1: 70 signals for $\alpha = 30$ (top), $\alpha = 300$ (middle), and $\alpha = 3000$ (bottom) sorted by R-FGLIM performance; the sort sequences are available on the manuscript webpage.

WPGHI performance than the decimation step a_d , which was generally small. On the other hand, the wavelet frame bound ratio deteriorates very quickly⁷ for too large decimation steps a_d . Hence, the choice of wavelet parameters was a trade-off between the two factors with no clear optimal solution. After some preliminary testing, we fixed the following parameter sets (α, a_d, K) : Low redundancy (1000, 30, 90) (low), Medium redundancy (1000, 25, 125)

⁷Large frame bound ratios also decrease numerical stability, such that audio file generation from the obtained reconstructions is prone to clipping artifacts.

⁶Available at <http://lftfat.github.io/notes/040/>

(medium), Medium high redundancy (1000, 18, 180) (med-high), High redundancy (1000, 10, 300) (high).

Similar to Experiment I, mean value and standard deviation over all signals are presented in Table II, for all parameter sets, with detailed results for all test signals shown in Figure 2. Additionally, Figure 3 shows an example of the difference between the target phase and the WPGHI-proposed phase estimate at different redundancies.

As expected, performance of both proposed methods increases with increasing redundancy. These improvements are apparent in the average performance over the whole signal set, but in most cases also on the level of individual signals, as can be seen in Figure 2. Similar to Experiment I, the average improvement in terms of spectral convergence of W-FGLIM over WPGHI is significant, ranging between 10 (low) and 6 dB (high).

K/a_d	method	WPGHI	W-FGLIM
30	mean	-38.2012	-44.1790
	std	6.6897	9.8989
10	mean	-32.5835	-40.7522
	std	6.5999	6.9075
5	mean	-26.8400	-36.4934
	std	5.9723	5.6917
3	mean	-20.7873	-29.7372
	std	5.1843	6.6363

Table II: Means and standard deviation of spectral convergence for WPGHI and WPGHI-FGLIM for redundancies $K/a = 30, 10, 5, 3$, with $\alpha = 1000$.

The performance difference between the different redundancies is also apparent in Figure 3. The phase reconstruction quality can be visually estimated from the characteristics of the difference between the target phase and the proposed estimate. Large areas of flat color indicate good quality; as the quality decreases, the phase difference becomes more *patchy*, with stronger fluctuation within patches. The figure shows that this *patchiness* is closely linked to redundancy of the underlying wavelet representation, or more generally, the employed sampling scheme. This behavior is characteristic for WPGHI and was previously observed for STFT-based PGHI [45] as well.

For all redundancies, informal listening has shown good to excellent reconstruction quality. Usually, the reconstructions using medium, medium high, or high redundancy and either of the proposed methods were indistinguishable from one another. In the case where distortions are audible at all, they were never found to be irritating. At low redundancy, the results are generally good as well, but a larger number of examples has clearly audible artifacts, e.g., signals 14 (oboe), 15 (cor anglais) and 16 (clarinet).

VIII. CONCLUSION

We have highlighted a number of interesting properties of the partial derivatives of the continuous WT. In particular, we showed how the log-magnitude and phase partial derivatives are related to wavelet transforms with modified mother wavelets. We characterize the class of wavelets that generate analytic WTs and obtain as a result a generalization

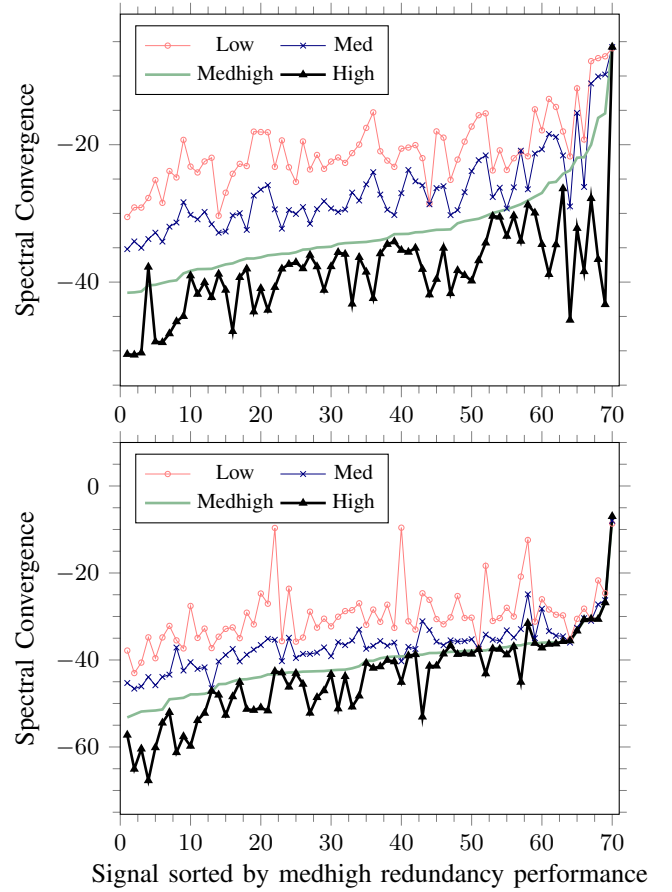


Figure 2: Comparison for different redundancies in WPGHI (top) and W-FGLIM (bottom) for 70 signals sorted by medhigh redundancy performance; the sort sequences are available on the manuscript webpage.

of the Cauchy wavelets. Based on the analyticity of these WTs, we obtain explicit phase-magnitude relations. Similar to the Gaussian STFT, the phase-magnitude relationship can be used as a basis for implementing magnitude-only reassignment and phaseless reconstruction. We explored the second application, providing very good results when applied to complex audio data, in most cases free of any perceptible distortion. We demonstrated that reconstruction quality is often notably improved over the established Griffin-Lim algorithm and our own previous implementation, relying on approximate phase-magnitude relations for Gaussian filter banks.

Future work will be concerned with real-time implementation of the proposed algorithm in the style of RTPGHI [66] and its generalization to nonuniform decimation schemes. The new phase-magnitude relations for WTs are crucial for the derivation of more appropriate phase approximation schemes for more general time-frequency filters, improving and extending the previously proposed filter bank PGHI algorithm [47]. We also plan to investigate phase-magnitude relations for the polyanalytic generalizations of the Cauchy WT [73]. Finally, the proposed scheme can serve as starting point for a wavelet-based phase vocoder for time-stretching and pitch-shifting of audio in the spirit of [74].

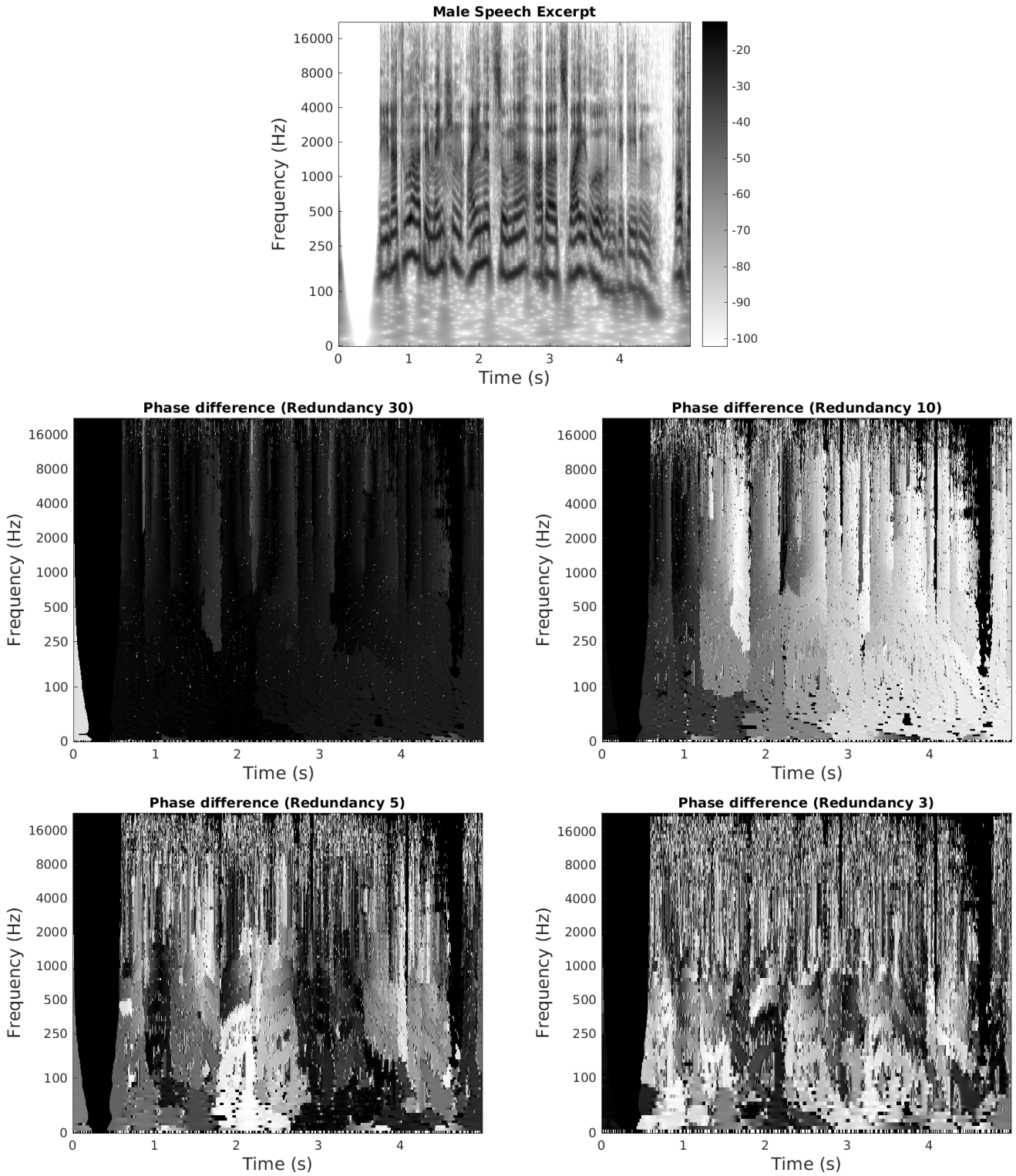


Figure 3: Top: Wavelet scalogram of a male German speech recording (derived from signal 54 of the SQAM dataset). Other panels: Difference between the true WT phase the test signal and the phase estimate proposed by WPGHI. Gray level indicates the difference in the range 0 (black) to π (white). Whenever the WT magnitude is below a tolerance level, the phase difference is also set to zero. Phase difference fluctuation clearly increases with decreasing redundancy.

ACKNOWLEDGMENTS

We wish to thank Andrés Marafioti for performing a small listening test confirming the reported observations on the perceptual quality of the reconstructed audio samples. Furthermore, we would like to thank Patrick Flandrin for pointing us to the literature on Klauder wavelets which coincide with the analyticity inducing wavelets given by (6). Finally, we thank the reviewers, whose comments helped us improve our results and their presentation.

APPENDIX A
PROOF OF THEOREM 1

Proof. Under the assumption that $\psi \in \mathbf{L}^2(\mathbb{R})$ is continuously differentiable with $\psi', \mathbf{T}\psi' \in \mathbf{L}^2(\mathbb{R})$, we can exchange differentiation and integration in (2), see Appendix D. Thus, the partial derivatives of $W_\psi s$ can be expressed as

$$\begin{aligned} \frac{\partial}{\partial x} W_\psi s(x, y) &= -\frac{1}{y\sqrt{y}} \int_{\mathbb{R}} s(t) \psi' \left(\frac{t-x}{y} \right) dt \\ &= -\frac{1}{y} W_{\psi'} s(x, y) \end{aligned} \quad (42)$$

and

$$\begin{aligned} \frac{\partial}{\partial y} W_\psi s(x, y) &= -\frac{1}{y\sqrt{y}} \int_{\mathbb{R}} s(t) \left(\frac{\psi(\bullet)}{2} + (\bullet) \psi'(\bullet) \right) \left(\frac{t-x}{y} \right) dt \\ &= -\frac{1}{y} \left(\frac{W_\psi s(x, y)}{2} + W_{\mathbf{T}(\psi')} s(x, y) \right) \\ &= -\frac{1}{y} \left(-\frac{W_\psi s(x, y)}{2} + W_{(\mathbf{T}\psi)'} s(x, y) \right). \end{aligned} \quad (43)$$

Here, we used that $(\mathbf{T}\psi)' = \psi + \mathbf{T}\psi'$ and that the WT is conjugate linear with respect to the chosen wavelet. Using that $\frac{\partial}{\partial x} \log(W_\psi s) = \frac{\frac{\partial}{\partial x} W_\psi s}{W_\psi s}$ and similarly for the partial derivative with respect to y , we obtain for all (x, y) with $W_\psi s(x, y) \neq 0$,

$$\frac{\partial}{\partial x} \log(W_\psi s)(x, y) = -\frac{1}{y} \frac{W_{\psi'} s(x, y)}{W_\psi s(x, y)} \quad (44a)$$

$$\frac{\partial}{\partial y} \log(W_\psi s)(x, y) = \frac{1}{2y} - \frac{1}{y} \frac{W_{(\mathbf{T}\psi)'} s(x, y)}{W_\psi s(x, y)}. \quad (44b)$$

Taking real and imaginary parts in (44) results in (3) and (4), respectively. ■

APPENDIX B
PROOF OF THEOREM 2

We first argue that analyticity of h and differentiability of f already imply that ψ satisfies the assumptions in Theorem 1, i.e., ψ is continuously differentiable with $\psi, \psi', \mathbf{T}\psi' \in \mathbf{L}^2(\mathbb{R})$. To this end, we note that the assumptions imply that $W_\psi s(x - aby, by)$ must be \mathcal{C}^∞ in x and y for an arbitrary $s \in \mathbf{L}^2(\mathbb{R})$. The same holds true for $W_\psi s(x, y)$.

At $x = 0$ and $y = 1$, the derivative $\frac{\partial}{\partial x} W_\psi s$ can be written as

$$\frac{\partial}{\partial x} W_\psi s(0, 1) = \lim_{x_0 \rightarrow 0} \left\langle s, \frac{\psi - \mathbf{T}_{-x_0} \psi}{x_0} \right\rangle, \quad (45)$$

which converges for every fixed $s \in \mathbf{L}^2(\mathbb{R})$. Thus, by a variant of the Banach-Steinhaus theorem [75, Ch. II.1, Corollary 2], the limit $\lim_{x_0 \rightarrow 0} \frac{\psi - \mathbf{T}_{-x_0} \psi}{x_0}$ represents a continuous linear functional and hence, by Riesz representation theorem, an element ψ_x in $\mathbf{L}^2(\mathbb{R})$. Rewriting the derivative for compactly supported, smooth s alternatively as

$$\frac{\partial}{\partial x} W_\psi s(0, 1) = \lim_{x_0 \rightarrow 0} \left\langle \frac{s - \mathbf{T}_{x_0} s}{x_0}, \psi \right\rangle = -\langle s', \psi \rangle, \quad (46)$$

we see that ψ_x is the weak derivative of ψ , i.e., the weak derivative of ψ exists and belongs to $\mathbf{L}^2(\mathbb{R})$. Repeating the argument for higher derivatives guarantees that weak derivatives of arbitrary order exist and by standard Sobolev embeddings so do continuous derivatives.

Similarly, the derivative $\frac{\partial}{\partial y} W_\psi s$ at $x = 0$ and $y = 1$ can be written as

$$\frac{\partial}{\partial y} W_\psi s(0, 1) = \lim_{y_0 \rightarrow 0} \left\langle s, \frac{\psi - \mathbf{D}_{-y_0} \psi}{y_0} \right\rangle. \quad (47)$$

Again we have the convergence $\lim_{y_0 \rightarrow 0} \frac{\psi - \mathbf{D}_{-y_0} \psi}{y_0} = \psi_y$. Rewriting the derivative for $s \in \mathcal{C}_0^\infty$ alternatively as

$$\begin{aligned} \frac{\partial}{\partial y} W_\psi s(0, 1) &= \lim_{y_0 \rightarrow 0} \left\langle \frac{s - \mathbf{D}_{1/-y_0} s}{y_0}, \psi \right\rangle \\ &= \langle s + \mathbf{T}s', \psi \rangle \\ &= \langle (\mathbf{T}s)', \psi \rangle. \end{aligned} \quad (48)$$

Now the operator $s \mapsto \mathbf{T}s'$ is well defined for compactly supported, smooth s with adjoint $s \mapsto -(\mathbf{T}s)'$. Furthermore, ψ is in the domain of this operator because $\langle -(\mathbf{T}s)', \psi \rangle = \langle s, \psi_y \rangle$ for all s in a dense subset. Thus, $\mathbf{T}\psi' = \psi_y \in \mathbf{L}^2(\mathbb{R})$. Hence, we established all assumptions of Theorem 1.

Analyticity of the function h is equivalent to it satisfying the CR equations that can be compactly expressed as $\frac{\partial}{\partial x} h = -i \frac{\partial}{\partial y} h$. To rewrite the CR equations for the function h in (5), we use (42), to obtain

$$\begin{aligned} \frac{\partial}{\partial x} h &= \left(\frac{\partial}{\partial x} f(x, y) \right) W_\psi s(x - aby, by) \\ &\quad - \frac{f(x, y)}{by} W_{\psi'} s(x - aby, by). \end{aligned} \quad (49)$$

Similarly, by (43), we have

$$\begin{aligned} \frac{\partial}{\partial y} h &= \left(\frac{\partial}{\partial y} f(x, y) \right) W_\psi s(x - aby, by) \\ &\quad - \frac{f(x, y)}{y} \left(-\frac{W_\psi s(x - aby, by)}{2} \right. \\ &\quad \left. + W_{(\mathbf{T}\psi)'} s(x - aby, by) - a W_{\psi'} s(x - aby, by) \right) \\ &= \left(\frac{\partial}{\partial y} f(x, y) + \frac{f(x, y)}{2y} \right) W_\psi s(x - aby, by) \\ &\quad - \frac{f(x, y)}{y} W_{(\mathbf{T}\psi)'} s(x - aby, by) \\ &\quad + \frac{f(x, y)}{y} a W_{\psi'} s(x - aby, by). \end{aligned} \quad (50)$$

Inserting these expressions into the CR equations results in

$$\left(\frac{y \frac{\partial}{\partial x} f(x, y) + iy \frac{\partial}{\partial y} f(x, y)}{f(x, y)} + \frac{i}{2} \right) W_\psi s(x - aby, by)$$

$$= \frac{1 - iab}{b} W_{\psi'} s(x - aby, by) + i W_{(\mathbf{T}\psi)'} s(x - aby, by). \quad (51)$$

We note that this condition depends on f only via the function $g(x, y) = \frac{y \frac{\partial}{\partial x} f(x, y) + iy \frac{\partial}{\partial y} f(x, y)}{f(x, y)} + \frac{i}{2}$. Moreover, using the definition of the WT in (51) implies that

$$\int_{\mathbb{R}} s(t) \left(g(x, y) \overline{\psi\left(\frac{t-x}{y}\right)} - \frac{1 - iab}{b} \overline{\psi'\left(\frac{t-x}{y}\right)} - i(\mathbf{T}\psi)' \left(\frac{t-x}{y}\right) \right) dt = 0 \quad (52)$$

for all $s \in \mathbf{L}^2(\mathbb{R})$ and thus

$$g(x, y) \overline{\psi\left(\frac{t-x}{y}\right)} - \frac{1 - iab}{b} \overline{\psi'\left(\frac{t-x}{y}\right)} - i(\mathbf{T}\psi)' \left(\frac{t-x}{y}\right) = 0 \quad (53)$$

as a function of t in $\mathbf{L}^2(\mathbb{R})$. In particular, this implies that $g(x, y)$ must be a constant $w \in \mathbb{C}$ and further

$$\overline{w}\psi - \frac{1 + iab}{b} \psi' + i(\mathbf{T}\psi)' = 0. \quad (54)$$

To solve this differential equation, it is more convenient to consider the Fourier transformed equivalent of (54). Using the standard properties of the Fourier transform $\widehat{\psi}' = 2\pi i \mathbf{T}\widehat{\psi}$ and $\widehat{\mathbf{T}\psi} = -(2\pi i)^{-1}(\widehat{\psi})'$, this is easily seen to be given by

$$\overline{w}\widehat{\psi} - 2\pi i \frac{1 + iab}{b} \mathbf{T}\widehat{\psi} - i\mathbf{T}(\widehat{\psi})' = 0. \quad (55)$$

We first note that our assumption $\widehat{\psi}(\xi) = 0$ for $\xi < 0$ satisfies this differential equation on this domain. For $\xi > 0$ and $\widehat{\psi} \neq 0$, we can easily solve the differential equation by rewriting

$$\frac{(\widehat{\psi})'(\xi)}{\widehat{\psi}(\xi)} = \frac{\overline{w}}{i\xi} - 2\pi \frac{1 + iab}{b}, \quad (56)$$

which gives

$$\begin{aligned} \widehat{\psi}(\xi) &= c e^{-i\overline{w} \log \xi} e^{-2\pi \frac{1 + iab}{b} \xi} \\ &= c \xi^{-\text{Im}(w)} e^{-2\pi(\frac{1}{b} + ia)\xi} e^{-i \text{Re}(w) \log \xi} \end{aligned} \quad (57)$$

for an arbitrary constant $c \in \mathbb{C}$. To obtain the parameters used in the theorem, we substitute $-\text{Im}(w) = \frac{\alpha - 1}{2}$, $\frac{1}{b} = \text{Re}(\gamma)$, $a = \text{Im}(\gamma)$, and $\text{Re}(w) = -\beta$. Based on our assumptions, we obtain the constraints $\text{Re}(\gamma) > 0$ and $\alpha > -1$ to guarantee $\widehat{\psi} \in \mathbf{L}^2(\mathbb{R})$.

APPENDIX C PROOF OF THEOREM 3

We can use the CR equations to obtain relationships between the derivatives of real and imaginary parts of the WT, which we will show to result in (11) and (12). For an arbitrary analytic function $h = u + iv$ the CR equations hold and are given by $\frac{\partial}{\partial x} u = \frac{\partial}{\partial y} v$ and $\frac{\partial}{\partial y} u = -\frac{\partial}{\partial x} v$. Writing $h = M e^{i\phi}$, the CR equations imply that

$$\frac{\partial}{\partial x} \phi = -\frac{\partial}{\partial y} \log M \quad (58)$$

and

$$\frac{\partial}{\partial y} \phi = \frac{\partial}{\partial x} \log M. \quad (59)$$

Using (58) and (59) for the function given by (7), yields

$$\begin{aligned} &\frac{\partial}{\partial x} \left(\phi_{\psi}^s \left(x - \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} y, \frac{y}{\text{Re}(\gamma)} \right) + \beta \log y \right) \\ &= -\frac{\partial}{\partial y} \log \left(y^{-\frac{\alpha}{2}} M_{\psi}^s \left(x - \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} y, \frac{y}{\text{Re}(\gamma)} \right) \right) \end{aligned} \quad (60)$$

and

$$\begin{aligned} &\frac{\partial}{\partial y} \left(\phi_{\psi}^s \left(x - \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} y, \frac{y}{\text{Re}(\gamma)} \right) + \beta \log y \right) \\ &= \frac{\partial}{\partial x} \log \left(y^{-\frac{\alpha}{2}} M_{\psi}^s \left(x - \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} y, \frac{y}{\text{Re}(\gamma)} \right) \right). \end{aligned} \quad (61)$$

These are equivalent to

$$\begin{aligned} \frac{\partial}{\partial x} \phi_{\psi}^s &= \frac{\alpha}{2y \text{Re}(\gamma)} - \frac{1}{\text{Re}(\gamma)} \frac{\partial}{\partial y} \log(M_{\psi}^s) \\ &\quad + \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} \frac{\partial}{\partial x} \log(M_{\psi}^s) \end{aligned} \quad (62)$$

and

$$-\frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} \frac{\partial}{\partial x} \phi_{\psi}^s + \frac{1}{\text{Re}(\gamma)} \frac{\partial}{\partial y} \phi_{\psi}^s + \frac{\beta}{y \text{Re}(\gamma)} = \frac{\partial}{\partial x} \log(M_{\psi}^s). \quad (63)$$

Inserting (62) into (63) finally results in

$$\begin{aligned} &-\frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} \left(\frac{\alpha}{2y \text{Re}(\gamma)} - \frac{1}{\text{Re}(\gamma)} \frac{\partial}{\partial y} \log(M_{\psi}^s) \right. \\ &\quad \left. + \frac{\text{Im}(\gamma)}{\text{Re}(\gamma)} \frac{\partial}{\partial x} \log(M_{\psi}^s) \right) + \frac{1}{\text{Re}(\gamma)} \frac{\partial}{\partial y} \phi_{\psi}^s + \frac{\beta}{y \text{Re}(\gamma)} \\ &= \frac{\partial}{\partial x} \log(M_{\psi}^s), \end{aligned} \quad (64)$$

which is equivalent to

$$\begin{aligned} \text{Re}(\gamma) \frac{\partial}{\partial y} \phi_{\psi}^s &= \frac{\alpha \text{Im}(\gamma)}{2y} - \frac{\beta}{y} + |\gamma|^2 \frac{\partial}{\partial x} \log(M_{\psi}^s) \\ &\quad - \text{Im}(\gamma) \frac{\partial}{\partial y} \log(M_{\psi}^s). \end{aligned} \quad (65)$$

This concludes the proof.

APPENDIX D ON DIFFERENTIABILITY OF THE WAVELET TRANSFORM

Denote by Φ_z and Ψ_z the difference operators

$$\Phi_z \psi = \frac{\mathbf{T}_z \psi - \psi}{z} \quad \text{and} \quad \Psi_z \psi = \frac{\mathbf{D}_{1+(z-1)} \psi - \psi}{z-1},$$

and note that

$$\begin{aligned} \frac{\partial}{\partial x} W_{\psi} s(x, y) &= \lim_{x_0 \rightarrow 0} \frac{W_{\psi} s(x + x_0, y) - W_{\psi} s(x, y)}{x_0} \\ &= \lim_{x_0 \rightarrow 0} \frac{1}{y} \langle s, \mathbf{T}_x \mathbf{D}_y (\Phi_{x_0/y} \psi) \rangle, \end{aligned}$$

as well as

$$\begin{aligned} \frac{\partial}{\partial y} W_{\psi} s(x, y) &= \lim_{y_0 \rightarrow 0} \frac{W_{\psi} s(x, y + y_0) - W_{\psi} s(x, y)}{y_0} \\ &= \lim_{y_0 \rightarrow 0} \frac{1}{y} \langle s, \mathbf{T}_x \mathbf{D}_y (\Psi_{1+y_0/y} \psi) \rangle, \end{aligned}$$

provided that the right-hand sides converge.

We first show that, for $z \rightarrow 0$

$$\Phi_z \psi \rightarrow -\psi' \text{ and } \Psi_{1+z} \psi \rightarrow -(\psi/2 + \mathbf{T}\psi') \quad (66)$$

as functions in $\mathbf{L}^2(\mathbb{R})$, provided that ψ , ψ' , and $\mathbf{T}\psi'$ are elements of $\mathbf{L}^2(\mathbb{R})$.

Since ψ' is assumed to be in $\mathbf{L}^2(\mathbb{R})$, there is for every $\tilde{\epsilon} > 0$ an $r_{\tilde{\epsilon}} > 0$, such that $\|\psi' \chi_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}}(0)}}\| < \tilde{\epsilon}$. Moreover, by the fundamental theorem of calculus,

$$\begin{aligned} |\Phi_z \psi(t) + \psi'(t)| &= \left| \frac{\int_t^{t-z} \psi'(s) ds}{z} + \psi'(t) \right| \\ &= \left| -\int_0^1 \psi'(t-sz) + \psi'(t) ds \right| \end{aligned} \quad (67)$$

Using (67), we obtain

$$\begin{aligned} &\left\| (\Phi_z \psi + \psi') \chi_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}+|z|}(0)}} \right\|^2 \\ &\leq \int_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}+|z|}(0)}} \left(\int_0^1 |\psi'(t-sz) - \psi'(t)| ds \right)^2 dt \\ &\leq \int_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}+|z|}(0)}} \int_0^1 3|\psi'(t-sz)|^2 + 3|\psi'(t)|^2 ds dt \\ &< 6\tilde{\epsilon}^2 \end{aligned} \quad (68)$$

where we used Jensen's inequality and Fubini's theorem. Furthermore, we have $|\psi'(t-s) - \psi'(t)| < \epsilon_r(|s|)$ for all $|t| < r$ and $|s| < 1$ by uniform continuity of ψ' on the compact set $B_{r+1}(0)$, where $\epsilon_r(\delta) \searrow 0$ for $\delta \rightarrow 0$. Thus, similar to (68), we obtain

$$\left\| (\Phi_z \psi + \psi') \chi_{\overline{B_{r_{\tilde{\epsilon}}+|z|}(0)}} \right\|^2 < 2(r_{\tilde{\epsilon}} + |z|) \epsilon_{r_{\tilde{\epsilon}}+1}^2(|z|) \quad (69)$$

for $|z| < 1$. Finally, for every $\tilde{\epsilon}$, there is a $z_{\tilde{\epsilon}} \in (0, 1)$, such that $\epsilon_{r_{\tilde{\epsilon}}+1}(z_{\tilde{\epsilon}}) < \tilde{\epsilon}/\sqrt{2(r_{\tilde{\epsilon}}+1)}$, which implies $\|\Phi_z \psi + \psi'\| < \sqrt{7}\tilde{\epsilon}$ for all $|z| \leq z_{\tilde{\epsilon}}$.

A similar, but slightly more complicated argument shows that $\Psi_{1+z} \psi \rightarrow -(\psi/2 + \mathbf{T}\psi')$, provided $\psi, \mathbf{T}\psi' \in \mathbf{L}^2(\mathbb{R})$. Here, we start with an $r_{\tilde{\epsilon}} > 0$, such that $\|\mathbf{T}\psi' \chi_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}}(0)}}\| < \tilde{\epsilon}$ and $\|\psi \chi_{\mathbb{R} \setminus \overline{B_{r_{\tilde{\epsilon}}}(0)}}\| < \tilde{\epsilon}$. As above, we use the fundamental theorem of calculus to obtain

$$\begin{aligned} &|\Psi_{1+z} \psi(t) + \psi(t)/2 + \mathbf{T}\psi'(t)| \\ &= \left| \frac{\frac{1}{(1+z)^{1/2}} \psi\left(\frac{t}{1+z}\right) - \psi(t)}{z} + \frac{\psi(t)}{2} + t\psi'(t) \right| \\ &= \left| \frac{\int_0^z \left[\frac{1}{(1+\bullet)^{1/2}} \psi\left(\frac{t}{1+\bullet}\right) \right]'(s) ds}{z} + \frac{\psi(t)}{2} + t\psi'(t) \right| \\ &= \left| -\int_0^z \frac{\frac{1}{2} \psi\left(\frac{t}{1+s}\right) + \frac{t}{1+s} \psi'\left(\frac{t}{1+s}\right)}{(1+s)^{3/2}} ds + \frac{\psi(t)}{2} + t\psi'(t) \right| \\ &= \left| -\int_0^1 \frac{\frac{1}{2} \psi\left(\frac{t}{1+sz}\right) + \frac{t}{1+sz} \psi'\left(\frac{t}{1+sz}\right)}{(1+sz)^{3/2}} + \frac{\psi(t)}{2} + t\psi'(t) ds \right| \\ &\leq \int_0^1 \left| \frac{\frac{1}{2} \psi\left(\frac{t}{1+sz}\right)}{(1+sz)^{3/2}} - \frac{\psi(t)}{2} \right| + \left| \frac{t}{1+sz} \psi'\left(\frac{t}{1+sz}\right) - t\psi'(t) \right| ds \end{aligned} \quad (70)$$

Similar to (68), this results in

$$\left\| (\Psi_{1+z} \psi + \psi/2 + \mathbf{T}\psi') \chi_{\mathbb{R} \setminus \overline{B_{(1+|z|)r_{\tilde{\epsilon}}}(0)}} \right\|^2 < 16\tilde{\epsilon}^2.$$

Furthermore, we can bound for $t \in \overline{B_{2r}(0)}$

$$\begin{aligned} &\left| \frac{\frac{1}{2} \psi\left(\frac{t}{1+s}\right)}{(1+s)^{3/2}} - \frac{\psi(t)}{2} \right| \\ &\leq \frac{1}{2|(1+s)^{3/2}|} \left| \psi\left(\frac{t}{1+s}\right) - \psi(t) \right| + \frac{|\psi(t)|}{2} \left| \frac{1}{(1+s)^{3/2}} - 1 \right| \\ &\leq \epsilon_r(t|s|) + \frac{|\psi(t)|}{2} \epsilon_r(|s|) \end{aligned} \quad (71)$$

for any $|s| \leq 1/2$, where $\epsilon_r(\delta) \rightarrow 0$ monotonically for $\delta \rightarrow 0$. Analogously, we obtain

$$\left| \frac{\frac{t}{1+s} \psi'\left(\frac{t}{1+s}\right)}{(1+s)^{3/2}} - t\psi'(t) \right| \leq 2\epsilon_r(t|s|) + |\psi(t)| \epsilon_r(|s|). \quad (72)$$

Equations (70)–(72) imply

$$\left| \Psi_{1+z} \psi(t) + \frac{\psi(t)}{2} + \mathbf{T}\psi'(t) \right| \leq 3\epsilon_r(t|z|) + 2|\psi(t)| \epsilon_r(|z|) \quad (73)$$

and thus

$$\begin{aligned} &\left\| \left(\Psi_{1+z} \psi + \frac{\psi}{2} + \mathbf{T}\psi' \right) \chi_{\overline{B_{(1+|z|)r_{\tilde{\epsilon}}}(0)}} \right\| \\ &< \sqrt{4r_{\tilde{\epsilon}}} \left(3 + 2 \sup_{t \in \overline{B_{2r_{\tilde{\epsilon}}}(0)}} |\psi(t)| \right) \epsilon_{r_{\tilde{\epsilon}}}(2r_{\tilde{\epsilon}}|z|). \end{aligned} \quad (74)$$

for $|z| < 1/(4r_{\tilde{\epsilon}})$ and where we assumed for simplicity $2r_{\tilde{\epsilon}} > 1$. Again, choosing $|z|$ sufficiently small implies the proposed convergence.

Thus, we finished the proof of (66), which implies

$$\left\langle s, \lim_{x_0 \rightarrow 0} \mathbf{T}_x \mathbf{D}_y (\Phi_{x_0/y} \psi) \right\rangle = -\langle s, \mathbf{T}_x \mathbf{D}_y \psi' \rangle$$

and

$$\left\langle s, \lim_{y_0 \rightarrow 0} \mathbf{T}_x \mathbf{D}_y (\Psi_{1+y_0/y} \psi) \right\rangle = -\langle s, \mathbf{T}_x \mathbf{D}_y (\psi/2 + \mathbf{T}\psi') \rangle.$$

Because convergence is in $\mathbf{L}^2(\mathbb{R})$, we can exchange limit and integral by continuity of the inner product.

Clearly, this argument can be repeated to obtain higher order derivatives, provided that ψ has sufficient regularity and decay.

REFERENCES

- [1] R. D. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1408–1419, Oct. 1999.
- [2] S. Chaplot, L. M. Patnaik, and N. R. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," *Biomed. Signal Process. Control*, vol. 1, no. 1, pp. 86–92, Jan. 2006.
- [3] A. Marec, J.-H. Thomas, and R. El Guerjouma, "Damage characterization of polymer-based composite materials: Multivariable analysis and wavelet transform for clustering acoustic emission data," *Mech. Syst. Sig. Process.*, vol. 22, no. 6, pp. 1441–1464, Aug. 2008.
- [4] J. Cusido, J. A. Rosero, L. Romeral, J. A. Ortega, and A. Garcia, "Fault detection in induction machines using power spectral density in wavelet decomposition," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 633–643, Feb. 2008.

- [5] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
- [6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, vol. 1, no. 2, pp. 205–220, Apr. 1992.
- [7] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830–1847, Jul. 2004.
- [8] S. Chu, S. Narayanan, and C.-C. J. Kuo, "Environmental sound recognition with time-frequency audio features," *IEEE Audio, Speech, Language Process.*, vol. 17, no. 6, pp. 1142–1158, Aug. 2009.
- [9] T. Necciarri, N. Holighaus, P. Balazs, Z. Průša, P. Majdak, and O. Derrien, "Audlet filter banks: A versatile analysis/synthesis framework using auditory frequency scales," *Appl. Sci.*, vol. 8, no. 1(96), Jan. 2018.
- [10] S. Mallat and I. Waldspurger, "Phase retrieval for the Cauchy wavelet transform," *J. Fourier Anal. Appl.*, vol. 21, no. 6, pp. 1251–1309, Dec. 2015.
- [11] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Appl. Comput. Harmon. Anal.*, vol. 20, no. 3, pp. 345–356, May 2006.
- [12] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, "Saving phase: Injectivity and stability for phase retrieval," *Appl. Comput. Harmon. Anal.*, vol. 37, no. 1, pp. 106–125, Jul. 2014.
- [13] R. Alaifari, I. Daubechies, P. Grohs, and R. Yin, "Stable phase retrieval in infinite dimensions," *Found. Comput. Math.*, 2018.
- [14] R. Alaifari, I. Daubechies, P. Grohs, and G. Thakur, "Reconstructing real-valued functions from unsigned coefficients with respect to wavelet and other frames," *J. Fourier Anal. Appl.*, vol. 23, no. 6, pp. 1480–1494, Dec. 2017.
- [15] R. Alaifari and P. Grohs, "Phase retrieval in the general setting of continuous frames for Banach spaces," *SIAM J. Math. Anal.*, vol. 49, no. 3, pp. 1895–1911, 2017.
- [16] I. Waldspurger, "Phase retrieval for wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2993–3009, May 2017.
- [17] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug. 1982.
- [18] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
- [19] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, Aug. 2013.
- [20] D. Griffin and J. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.
- [21] Y. Shechtman, A. Beck, and Y. C. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 928–938, Feb. 2014.
- [22] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 87–109, May 2015.
- [23] N. Perraudin, P. Balazs, and P. L. Søndergaard, "A fast Griffin-Lim algorithm," in *Proc. IEEE Appl. Sig. Process. Audio Acoustics*, New Paltz, NY, USA, Oct. 2013.
- [24] E. Moulines and F. Charpentier, "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones," *Speech communication*, vol. 9, no. 5-6, pp. 453–467, Dec. 1990.
- [25] Y. Wang, R. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio et al., "Tacotron: Towards end-to-end speech synthesis," *arXiv preprint arXiv:1703.10135*, 2017.
- [26] A. Marafioti, N. Holighaus, N. Perraudin, and P. Majdak, "Adversarial generation of time-frequency features with application in audio synthesis," *arXiv preprint arXiv:1902.04072*, 2019.
- [27] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Audio, Speech, Language Process.*, vol. 15, no. 3, pp. 1066–1074, Mar. 2007.
- [28] F. R. Bach and M. I. Jordan, "Learning spectral clustering, with application to speech separation," *J. Mach. Learn. Res.*, vol. 7, pp. 1963–2001, Oct. 2006.
- [29] J. Bruna, P. Sprechmann, and Y. LeCun, "Source separation with scattering non-negative matrix factorization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.* South Brisbane, Australia: IEEE, 2015, pp. 1876–1880.
- [30] E. Moulines and J. Laroche, "Non-parametric techniques for pitch-scale and time-scale modification of speech," *Speech communication*, vol. 16, no. 2, pp. 175–205, Feb. 1995.
- [31] J. Laroche and M. Dolson, "Improved phase vocoder time-scale modification of audio," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 3, pp. 323–332, May 1999.
- [32] Z. Průša and N. Holighaus, "Phase vocoder done right," in *Proc. Eur. Signal Process. Conf. EUSIPCO*, Kos island, Greece, Aug. 2017, pp. 976–980.
- [33] F. Auger, E. Chassande-Mottin, and P. Flandrin, "On phase-magnitude relationships in the short-time Fourier transform," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 267–270, May 2012.
- [34] M. R. Portnoff, "Magnitude-phase relationships for short-time Fourier transforms based on Gaussian analysis windows," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, Washington, D. C., USA, Apr 1979, pp. 186–189.
- [35] P. Balazs, D. Bayer, F. Jalliet, and P. Søndergaard, "The pole behavior of the phase derivative of the short-time Fourier transform," *Appl. Comput. Harmon. Anal.*, vol. 40, no. 3, pp. 610–621, May 2016.
- [36] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May 1995.
- [37] L. Daudet, M. Morvidone, and B. Torresani, "Time-frequency and time-scale vector fields for deforming time-frequency and time-scale representations," in *Proc. SPIE*, vol. 3813, San Diego, CA, USA, Jul. 1999.
- [38] M. Morvidone and B. Torresani, "Variations on Hough-wavelet transforms for time-frequency chirp detection," in *Proc. SPIE*, vol. 5207, Nov. 2003.
- [39] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013.
- [40] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 243–261, Mar. 2011.
- [41] J. M. Lilly and S. C. Olhede, "On the analytic wavelet transform," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 4135–4156, Aug. 2010.
- [42] I. Daubechies and T. Paul, "Time-frequency localisation operators—a geometric phase space approach: II The use of dilations," *Inverse Prob.*, vol. 4, no. 3, pp. 661–680, Aug. 1988.
- [43] P. Flandrin, "Separability, positivity, and minimum uncertainty in time-frequency energy distributions," *J. Math. Phys.*, vol. 39, no. 8, pp. 4016–4040, Aug. 1998.
- [44] G. Ascensi and J. Bruna, "Model space results for the Gabor and wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2250–2259, May 2009.
- [45] Z. Průša, P. Balazs, and P. L. Søndergaard, "A noniterative method for reconstruction of phase from STFT magnitude," *IEEE Audio, Speech, Language Process.*, vol. 25, no. 5, May 2017.
- [46] Z. Průša and P. L. Søndergaard, "Real-time spectrogram inversion using phase gradient heap integration," in *Proc. Int. Conf. Digital Audio Effects*, Brno, Czech Republic, Sep. 2016.
- [47] Z. Průša and N. Holighaus, "Non-iterative filter bank phase (re)construction," in *Proc. Eur. Signal Process. Conf. EUSIPCO*, Kos island, Greece, Aug. 2017, pp. 952–956.
- [48] D. J. Nelson, "Cross-spectral methods for processing speech," *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2575–2592, Nov. 2001.
- [49] —, "Instantaneous higher order phase derivatives," *Digital Signal Process.*, vol. 12, no. 2–3, pp. 416–428, 2002.
- [50] F. Auger, E. Chassande-Mottin, and P. Flandrin, "Making reassignment adjustable: The Levenberg-Marquardt approach," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Kyoto, Japan, Mar. 2012, pp. 3889–3892.
- [51] N. Delprat, B. Escudé, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torresani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 644–664, Mar. 1992.
- [52] J. M. Lilly and S. C. Olhede, "Generalized Morse wavelets as a superfamily of analytic wavelets," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6036–6041, Nov. 2012.
- [53] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. Burlington, MA: Academic press, 2008.
- [54] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 569–586, Mar. 1992.

- [55] M. Unser, A. Aldroubi, and S. J. Schiff, "Fast implementation of the continuous wavelet transform with integer scales," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3519–3523, Dec. 1994.
- [56] M. J. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, Oct. 1992.
- [57] Z. Průša, P. L. Søndergaard, and P. Rajmic, "Discrete wavelet transforms in the large time-frequency analysis toolbox for Matlab/GNU Octave," *ACM Trans. Math. Softw.*, vol. 42, no. 4, pp. 32:1–32:23, Jun. 2016.
- [58] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. Velasco, "Theory, implementation and applications of nonstationary Gabor frames," *J. Comput. Appl. Math.*, vol. 236, no. 6, pp. 1481–1496, Oct. 2011.
- [59] N. Holighaus, M. Dörfler, G. A. Velasco, and T. Grill, "A framework for invertible, real-time constant-Q transforms," *IEEE Audio, Speech, Language Process.*, vol. 21, no. 4, pp. 775–785, Apr. 2013.
- [60] C. Schörkhuber, A. Klapuri, N. Holighaus, and M. Dörfler, "A Matlab toolbox for efficient perfect reconstruction time-frequency transforms with log-frequency resolution," in *Proc. AES Conf. Semantic Audio*, London, UK, Jan. 2014.
- [61] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3256–3268, Dec. 1998.
- [62] Z. Cvetkovic and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1245–1255, May 1998.
- [63] M. Fickus, M. L. Massar, and D. G. Mixon, "Finite frames and filter banks," in *Finite Frames*, P. G. Casazza and G. Kutyniok, Eds. Basel, Switzerland: Birkhäuser, 2013, pp. 337–379.
- [64] T. Strohmer, "Numerical algorithms for discrete Gabor expansions," in *Gabor Analysis and Algorithms*, H. G. Feichtinger and T. Strohmer, Eds. Boston, MA, USA: Birkhäuser, 1998, pp. 267–294.
- [65] K. Gröchenig, "Acceleration of the frame algorithm," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3331–3340, Dec. 1993.
- [66] Z. Průša and P. Rajmic, "Toward high-quality real-time signal reconstruction from STFT magnitude," *IEEE Signal Process. Lett.*, vol. 24, no. 6, pp. 892–896, Jun. 2017.
- [67] J. W. J. Williams, "Algorithm 232: Heapsort," *Communications of the ACM*, vol. 7, no. 6, pp. 347–348, Jun. 1964.
- [68] "Tech 3253: Sound Quality Assessment Material recordings for subjective tests," Eur. Broadc. Union, Geneva, Tech. Rep., Sept. 2008.
- [69] Z. Průša, P. L. Søndergaard, N. Holighaus, C. Wiesmeyr, and P. Balazs, "The large time-frequency analysis toolbox 2.0," in *Sound, Music, and Motion*, M. Aramaki, O. Derrien, R. Kronland-Martinet, and S. Ystad, Eds. Cham, Switzerland: Springer, 2014, pp. 419–442.
- [70] N. Sturm and L. Daudet, "Signal reconstruction from STFT magnitude: A state of the art," in *Proc. Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 375–386.
- [71] O. Christensen, *An Introduction to Frames and Riesz Bases*. Basel, Switzerland: Birkhäuser, 2016.
- [72] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, 1992.
- [73] L. D. Abreu, "Superframes and polyanalytic wavelets," *J. Fourier Anal. Appl.*, vol. 23, no. 1, pp. 1–20, Feb. 2017.
- [74] Z. Průša and N. Holighaus, "Phase vocoder done right," in *Proc. Eur. Signal Process. Conf. EUSIPCO*, Kos island, Greece, Aug. 2017, pp. 1006–1010.
- [75] K. Yosida, *Functional Analysis*, 6th ed. Berlin: Springer, 1980.