NON-ITERATIVE PHASELESS RECONSTRUCTION FROM WAVELET TRANSFORM MAGNITUDE

Nicki Holighaus, Günther Koliander, Luis Daniel Abreu *

Acoustics Research Institute
Austrian Academy of Sciences
Vienna, Austria
{nholighaus,gkoliander,labreu}@kfs.oeaw.ac.at

Zdeněk Průša

LEWITT GmbH
Vienna, Austria
zprusa@kfs.oeaw.ac.at

ABSTRACT

In this work, we present an algorithm for phaseless reconstruction from magnitude-only wavelet coefficients. The method relies on an explicit relation between the log-magnitude and phase gradients of analytic wavelet transforms and an extension of the Phase-Gradient Heap Integration (PGHI) algorithm recently introduced for Gabor phaseless reconstruction. This relation is exact for a certain family of mother wavelets including Cauchy wavelets of arbitrary order, but only holds approximately otherwise. The presented experiments show that, in practice, the proposed wavelet PGHI method provides competitive quality for various mother wavelets. Furthermore, wavelet PGHI is a non-iterative scheme and thus computational performance is significantly better than established alternate projection methods.

1. INTRODUCTION

The analysis of data utilizing time-frequency or time-scale representations is prevalent in various scientific fields. Prominent examples are medicine [1] and image [2, 3] and audio processing [4, 5, 6]. Although these representations are often visualized by using magnitude-only measurements, they are usually complex-valued, i.e., provide an additional phase component. In general, reconstruction of the signal is only possible from the full complex-valued representation. Since manipulations of the signal are often performed in the magnitude-only representation domain and in some application we can even measure only the magnitudes, there is a need to construct a phase that matches a given magnitude-only representation.

This task is known as phase retrieval or phaseless reconstruction and has been considered from a theoretical [7, 8, 9, 10, 11] as well as an algorithmic [12, 13, 14, 15, 16, 17, 18] viewpoint. While theoretical results mainly deal with the feasibility of phase retrieval, most algorithms are based on iterative projection methods. An important result we will build on, is that although in general the full complex-valued representation is necessary for reconstruction, there are settings where the phase and magnitude components carry almost the same information. The first such case, the STFT with Gaussian generator, was considered by Portnoff [19] and later by Auger and Flandrin [20]. In this setting, the phase gradient and the gradient of the (logarithmically scaled) magnitude are in a one-to-one relationship. Recently, it was shown [21] that for certain mother wavelets $\psi \in \mathbf{L}^2(\mathbb{R})$ this is true for wavelet transforms (WT) as well. More specifically, the Fourier transform of the mother wavelet must satisfy

$$\hat{\psi}(\xi) = \begin{cases} c\xi^{\frac{\alpha-1}{2}} e^{-2\pi\gamma\xi} e^{i\beta\log\xi} & \xi \in \mathbb{R}^+, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

for some $c \in \mathbb{C}$, $\alpha > -1$, $\beta \in \mathbb{R}$, and $\gamma \in \mathbb{C}$ with $\mathrm{Re}(\gamma) > 0$. Similar to the STFT case with Gaussian generator, the resulting WTs give time-scale representations that are analytic functions for the time-scale parameter pair being interpreted as a single complex variable. This class includes the *Cauchy wavelet* ($\beta = 0$), see [22], for which analyticity of the WT was already shown in [23].

In this paper, we propose an algorithm relying on the phase-magnitude relationship of the WT to perform wavelet phaseless reconstruction. On large scale data, such as audio, this problem has previously been addressed with generic alternating projection methods, such as [16] and its variations. More recently a wavelet-adapted iterative scheme has been proposed in [12, 7]. By combining a discrete approximation of the phase-magnitude relationship with an adaptive integration scheme in the spirit of [24, 25, 26], we can forgo iteration and obtain a phase estimate directly from the magnitude-only coefficients.

We will first recall the results of [21], in particular, the characterization of the WT phase gradient by its log-magnitude gradient for wavelets of the form (1). To motivate the implementation in the discrete domain, we briefly sketch the transition from the continuous to the discrete realm, as well as the invertible WT implementation used in the experiments. Before proceeding to the experiments, which are the main focus of this work, we formally introduce the wavelet phase gradient heap integration algorithm (WPGHI) for phaseless reconstruction.

Although the phase-magnitude relationship only holds exactly for Cauchy wavelets, we demonstrate that for a broad class of mother wavelets the relations hold approximately and our phase reconstruction algorithm works surprisingly well. At the center of the manuscript is an extensive evaluation that demonstrates the performance of wavelet PGHI under variations of the mother wavelet, its time-frequency resolution trade-off and the oversampling rate. The dependence of reconstruction performance on the parameters is investigated and a comparison with the widely used (fast) Griffin-Lim algorithm [27] is performed. Furthermore, we also consider WPGHI as an initialization for fast Griffin-Lim. In a final experiment, we consider wavelet PGHI for a wavelet with compact support in the time domain. This can be considered the first step towards a bounded delay implementation of wavelet PGHI in

* 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 (WWTF): MA16-053.

the vein of RTPGHI [28] for the short-time Fourier transform. To complement this experiment, we further present a *causal* variant of wavelet PGHI that processes the wavelet coefficients one time position at a time, assuming that the phase at previous positions is already known. In conjunction with a bounded delay framework for wavelet analysis and synthesis, this algorithm can serve as the central building block for a real-time implementation in the future.

Notation: In this contribution, we consider signals as finite energy functions in a continuous or discrete variable, i.e. $s \in \mathbf{L}^2(\mathbb{R})$ or $s \in \mathbb{C}^L$ for some natural number $L \in \mathbb{N}$. For a differentiable function, we denote partial derivatives by $\frac{\partial}{\partial \bullet}$, where the variable with respect to which we differentiate is substituted for the placeholder $\bullet$. The Fourier transform on $\mathbf{L}^2(\mathbb{R})$ is the unitary operator derived in the usual way from the integral transform $\hat{s}(\xi) = \mathcal{F}(s)(\xi) = \int_{\mathbb{R}} s(t) e^{-2\pi i \xi t}\, dt$ that is defined for integrable signals $s \in \mathbf{L}^1(\mathbb{R})$. Finally, for a complex scalar $z \in \mathbb{C}$, we denote its real and imaginary parts by $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, respectively.

2. THE PHASE-MAGNITUDE RELATIONSHIP

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 all $x \in \mathbb{R}$, $y \in \mathbb{R}^+$. 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, $\log(W_\psi s) = \log(M_\psi^s) + i\phi_\psi^s$.

In [21], it was shown that, with ψ as in (1), the function

$$x + iy \mapsto y^{-\frac{\alpha}{2}} e^{i\beta \log y} W_\psi s\left(x - \frac{\mathrm{Im}(\gamma)}{\mathrm{Re}(\gamma)} y, \frac{y}{\mathrm{Re}(\gamma)}\right). \quad (3)$$

considered as a function in the complex variable $z = x + iy$ ($y > 0$) is analytic, i.e., complex differentiable on the upper half-plane. In this case, the following expressions linking the partial derivatives of the log-magnitude and phase components, hold:

Theorem 1 ([21, Th. 1])**.** *Let $\psi \in \mathbf{L}^2(\mathbb{R})$ be a function that satisfies*

$$\hat{\psi}(\xi) = \begin{cases} \xi^{\frac{\alpha-1}{2}} e^{-2\pi\gamma\xi} e^{i\beta \log \xi} & \xi \in \mathbb{R}^+, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

for some $\alpha > -1$, $\beta \in \mathbb{R}$, and $\gamma \in \mathbb{C}$ with $\mathrm{Re}(\gamma) > 0$. Then

$$\frac{\partial}{\partial x}\phi_\psi^s = \frac{\alpha}{2y\,\mathrm{Re}(\gamma)} - \frac{\frac{\partial}{\partial y}\log\left(M_\psi^s\right)}{\mathrm{Re}(\gamma)} + \frac{\mathrm{Im}(\gamma)\frac{\partial}{\partial x}\log\left(M_\psi^s\right)}{\mathrm{Re}(\gamma)} \quad (5)$$

and

$$\frac{\partial}{\partial y}\phi_\psi^s = \frac{\alpha\,\mathrm{Im}(\gamma) - 2\beta}{2y\,\mathrm{Re}(\gamma)} + \frac{|\gamma|^2 \frac{\partial}{\partial x}\log\left(M_\psi^s\right)}{\mathrm{Re}(\gamma)} - \frac{\mathrm{Im}(\gamma)\frac{\partial}{\partial y}\log\left(M_\psi^s\right)}{\mathrm{Re}(\gamma)}. \quad (6)$$

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

$$\frac{\partial}{\partial x}\phi_\psi^s(x,y) = -\frac{\partial}{\partial y}\log(M_\psi^s)(x,y) + \frac{\alpha}{2y} \quad (7)$$

and

$$\frac{\partial}{\partial y}\phi_\psi^s(x,y) = \frac{\partial}{\partial x}\log(M_\psi^s)(x,y) - \frac{\beta}{y}. \quad (8)$$

The wavelets ψ specified by (4) are also known as "Klauder wavelets" and minimize a time-scale counterpart of Heisenberg uncertainty [29, Prop. 16]. They are a minor generalization of Cauchy wavelets [22], which are recovered for the choice $\beta = 0$ and $\gamma = 1$. Because a change in γ results only in a scale change, dependent on $\mathrm{Re}(\gamma)$, and a time shift, dependent on $\mathrm{Im}(\gamma)$, we will only consider the case $\gamma = 1$. For most of this contribution, we will in fact consider only Cauchy wavelets.

In the following, ψ will refer to the wavelet specified by (4) with $\gamma = 1$. We can also interpret the wavelet coefficients as time-frequency measurements. More specifically, the Fourier transform of ψ has quick decay around its unique peak (or *center frequency*), located at $\xi_\mathrm{b} = \frac{\alpha-1}{4\pi}$. Considering the $\mathbf{L}^1$-normalized dilation $\tilde{\mathbf{D}}_y s(t) = y^{-1} s(t/y)$, we can define $\tilde{W}_\psi s$ by

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

For this form of the WT, straightforward calculations show that the phase-magnitude relations read as follows:

$$\frac{\partial}{\partial x}\tilde{\phi}_\psi^s(x,\xi) = \frac{4\pi\xi^2}{\alpha-1}\frac{\partial}{\partial \xi}\log(\tilde{M}_\psi^s)(x,\xi) + 2\pi\xi, \quad (9)$$

$$\frac{\partial}{\partial \xi}\tilde{\phi}_\psi^s(x,\xi) = -\frac{\alpha-1}{4\pi\xi^2}\frac{\partial}{\partial x}\log(\tilde{M}_\psi^s)(x,\xi) + \frac{\beta}{\xi}, \quad (10)$$

where $\tilde{M}_\psi$ and $\tilde{\phi}_\psi$ denote the magnitude and phase of $\tilde{W}_\psi$, respectively. In the following, we will discretize this form of the phase-magnitude relations to derive a discrete approximation for application within our proposed phaseless reconstruction algorithm.

It is notable that the formulas (9)–(10) almost exactly correspond to the phase-magnitude relations in the STFT case with a dilated Gaussian [24, Sec. III]. In particular, for $\beta = 0$ the only difference is that the constant time-frequency ratio λ is replaced by the frequency-dependent term $\frac{\alpha-1}{4\pi\xi^2}$.

3. PHASELESS RECONSTRUCTION

To perform phaseless reconstruction, we have to assume that the given magnitude coefficients originate from an invertible wavelet system, i.e., a wavelet frame. In this case, we can perform phase estimation followed either by direct synthesis via a dual frame [30, 31, 32, 33] or iterative synthesis via, e.g., conjugate gradient iteration [34, 6].

Although the performance of phaseless reconstruction should be largely independent of the particular implementation of the analysis and synthesis operations, we briefly sketch a potential implementation for illustrative purposes, largely following [35, 36, 37, 6]. We will denote discretizations of continuous signals by brackets, e.g., the discretized signal $s_\mathrm{d}[l] \in \mathbb{C}$ for $l \in \{1, \ldots, L-1\}$ and some $L \in \mathbb{N}$. In this discrete domain, the translation operator acts circularly, i.e., $s_\mathrm{d}[l-m]$ is interpreted as $s_\mathrm{d}[\mathrm{mod}(l-m, L)]$.

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: Assuming the sampling rate ξ_s, the frequency response of the wavelet at scale $y = \xi_\mathrm{b}/\xi$ is simply

$$\widehat{\psi_y}[k] = \widehat{\psi}\left(\frac{y\xi_\mathrm{s}k}{L}\right) = \widehat{\psi}\left(\frac{\xi_\mathrm{b}\xi_\mathrm{s}}{L}\frac{k}{\xi}\right),$$

for $k \in \{-\lfloor L/2 \rfloor, \dots, \lceil L/2 \rceil - 1\}$. In practice, we only cover a finite range of scales and to cover the entire frequency range, we introduce an additional low-pass function in the style of [6, Sec. 3.1.2].

The entire wavelet system is characterized by the minimum scale $y_\mathrm{m} \in \mathbb{R}^+$, the scale step $2^{1/B}$, with $B \in \mathbb{R}^+$, the number of scales $K \in \mathbb{N}$, and the decimation[1] factor $a_\mathrm{d} \in \mathbb{N}$, with $a_\mathrm{d} | L$. The corresponding scaled and shifted wavelets are given as

$$\psi_{n,k} = \mathbf{T}_{na_\mathrm{d}} \psi_{2^{k/B} y_\mathrm{m}} \tag{11}$$

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

$$\widehat{\psi_\mathrm{lp}} = a_\mathrm{d}^{-1} P_\mathrm{lp} \Psi_\mathrm{lp}, \tag{12}$$

where

$$\Psi_\mathrm{lp} = \sqrt{\max(\Psi) - \Psi}, \quad \Psi = \sum_{k=0}^{K-1} |\widehat{\psi_{0,k}}|^2. \tag{13}$$

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 $(2K+1)/a_\mathrm{d}$ when analyzing signals with no negative frequency content. With a slight abuse of terminology, we will from now on refer to the proportional quantity K/a_d as the redundancy.

In the following, we will assume that ψ satisfies (4) with $\beta = 0$ and $\gamma = 1$. The adaptation to general β and γ is straightforward, but lowers readability, so we leave it to the reader. For general wavelets, we have to determine the appropriate value of α by comparing the chosen mother wavelet to $\psi^{(\alpha)}$ for varying α and select the best match, see Section 4.

Assume that the continuous-time signal s is approximately band- and time-limited on $[0, \xi_\mathrm{s}[$ and $[0, L/\xi_\mathrm{s}[$, respectively. Then, with $s_\mathrm{d}[l] = s(l/\xi_\mathrm{s})$, for $l \in \{0, \dots, L-1\}$, $a_\mathrm{d} = a\xi_\mathrm{s} \in \mathbb{N}$, and $\xi_k = 2^{-k/B}\xi_\mathrm{b}/y_\mathrm{m}$, we obtain the approximation

$$M_s[n,k] := |\langle s_\mathrm{d}, \psi_{n,k} \rangle| \approx \xi_\mathrm{s} \tilde{M}^s_\psi(na, \xi_k). \tag{14}$$

Using (14), we can formulate a discrete approximation of the phase-magnitude relations (9) and (10). Note that normalization by ξ_s becomes irrelevant after taking the logarithmic derivative in (14). As a substitute for the continuous partial derivatives, we take Δ_n and Δ_k to be an appropriate discrete differentiation scheme. Our implementation relies on (weighted) centered differences:

$$\Delta_n(M_s)[n,k] := \frac{\xi_\mathrm{s}(M_s[n+1,k] - M_s[n-1,k])}{2a_\mathrm{d}}, \tag{15}$$

$$\begin{aligned}\Delta_k(M_s)[n,k] := &\frac{M_s[n,k+1] - M_s[n,k]}{2(\xi_{k+1} - \xi_k)} \\ &+ \frac{M_s[n,k] - M_s[n,k-1]}{2(\xi_k - \xi_{k-1})}.\end{aligned} \tag{16}$$

[1] For simplicity, we restrict here to uniform decimation.

Here, weighted centered differences are used in Δ_k, since the sampling step in the scale coordinate changes depends on k. For border points, i.e., $n \in \{0, N-1\}$ and $k \in \{0, K-1\}$, respectively, the appropriate forward or backward differences are used instead.

If we combine (14) with the phase-magnitude relations (9) and (10), we obtain

$$\begin{aligned}\frac{\partial}{\partial x}\tilde{\phi}^s_\psi(na, \xi_k) &= \frac{4\pi\xi_k^2}{\alpha - 1}\frac{\partial}{\partial \xi}\log(\tilde{M}^s_\psi)(na, \xi_k) + 2\pi\xi_k \\ &\approx \Delta^{\tilde{\phi},x,s}_\psi[n,k] := \frac{4\pi\xi_k^2}{\alpha-1}\Delta_k(\log(M_s))[n,k] + 2\pi\xi_k,\end{aligned} \tag{17}$$

and

$$\begin{aligned}\frac{\partial}{\partial \xi}\tilde{\phi}^s_\psi(na, \xi_k) &= -\frac{\alpha - 1}{4\pi\xi_k^2}\frac{\partial}{\partial x}\log(\tilde{M}^s_\psi)(na, \xi_k) \\ &\approx \Delta^{\tilde{\phi},\xi,s}_\psi[n,k] := -\frac{\alpha-1}{4\pi\xi_k^2}\Delta_n(\log(M_s))[n,k].\end{aligned} \tag{18}$$

Now, from $\Delta^{\tilde{\phi},x,s}_\psi$ and $\Delta^{\tilde{\phi},\xi,s}_\psi$, 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. The provided implementation relies on simple 1-dimensional trapezoidal quadrature. This results in the following integration rule on the set of neighbors of (n,k), i.e., $(n_\mathrm{n}, k_\mathrm{n}) \in \mathcal{N}_{n,k} := \{(n \pm 1, k), (n, k \pm 1)\} \cap \{0, \dots N-1\} \times \{0, \dots, K-1\}$.

$$\begin{aligned}&(\tilde{\phi}^s_\psi)_\mathrm{est}[n_\mathrm{n}, k_\mathrm{n}] \\ &= (\tilde{\phi}^s_\psi)_\mathrm{est}[n,k] + \frac{\xi_{k_\mathrm{n}} - \xi_k}{2}\left(\Delta^{\tilde{\phi},\xi,s}_\psi[n,k] + \Delta^{\tilde{\phi},\xi,s}_\psi[n_\mathrm{n}, k_\mathrm{n}]\right) \\ &\quad + \frac{a_\mathrm{d}(n_\mathrm{n} - n)}{2\xi_\mathrm{s}}\left(\Delta^{\tilde{\phi},x,s}_\psi[n,k] + \Delta^{\tilde{\phi},x,s}_\psi[n_\mathrm{n}, k_\mathrm{n}]\right).\end{aligned} \tag{19}$$

When inserting (15) and (16) into (19), the absolute scale of the center frequencies ξ_k and sampling rate ξ_s becomes unimportant and only their ratio enters the quadrature (19). Hence, by considering relative frequencies ξ_k/ξ_s, the algorithm is valid independent of the assumed sampling rate.

However, the integration step itself is not entirely straightforward. As discussed in [7] and [**?**], phase estimation from magnitude-only measurements is generally highly unstable when the coefficients are close to 0. To avoid these instabilities, the work [24] introduced the Phase Gradient Heap Integration (PGHI) algorithm, originally for Gabor phase reconstruction. The algorithm adaptively applies a given integration rule, starting at coefficients of large magnitude and avoiding areas of low magnitude. In the pseudo-code shown in Algorithm 1 it is assumed that all wavelet coefficients are available at all times, similar to [24] and [26].

Once the phase estimate $(\tilde{\phi}^s_\psi)_\mathrm{est}$ has been computed, it is combined with the magnitude by $W_s := M_s e^{i(\tilde{\phi}^s_\psi)_\mathrm{est}}$ and a time-domain signal is obtained by performing a regular synthesis step.

4. EXPERIMENTS

The performance of wavelet PGHI with Cauchy WTs was extensively evaluated on the EBU SQAM database [39] in [21], see also `http://ltfat.github.io/notes/053/`. Here, we evaluate the performance of wavelet PGHI with mother wavelets differing from the Cauchy wavelet. To this end, we selected a var-

Algorithm 1: Wavelet Phase Gradient Heap Integration

Input: Magnitude M_s of wavelet coefficients, estimates $\Delta^{\tilde{\phi},x,s}_{\psi}$ and $\Delta^{\tilde{\phi},\xi,s}_{\psi}$ of the partial phase derivatives, relative tolerance tol.
Output: Phase estimate $(\tilde{\phi}^s_{\psi})_{\text{est}}$.

$abstol \leftarrow tol \cdot \max(M_s[n,k])$;
Create set $\mathcal{I} = \{(n,k) : M_s[n,k] > abstol\}$;
Assign random values to $(\tilde{\phi}^s_{\psi})_{\text{est}}(n,k)$ for $(n,k) \notin \mathcal{I}$;
Construct a self-sorting max *heap* [38] for (n,k) pairs;
while $\mathcal{I}$ is not $\emptyset$ **do**
  if *heap* is empty **then**
    Move $(n_m,k_m) = \arg\max_{(n,k)\in\mathcal{I}} (M_s[n,k])$ from $\mathcal{I}$ into the *heap*;
    $(\tilde{\phi}^s_{\psi})_{\text{est}}(n_m,k_m) \leftarrow 0$;
  end
  while *heap* is not empty **do**
    $(n,k) \leftarrow$ remove the top of the *heap*;
    foreach $(n_{\text{n}},k_{\text{n}})$ in $\mathcal{N}_{n,k} \cap \mathcal{I}$ **do**
      Compute $(\tilde{\phi}^s_{\psi})_{\text{est}}(n_{\text{n}},k_{\text{n}})$ by means of (19);
      Move $(n_{\text{n}},k_{\text{n}})$ from $\mathcal{I}$ into the *heap*;
    end
  end
end

ied subset of 15 signals from the EBU SQAM database,[2] including signals that were reported as critical in previous contributions [24, 21]. The chosen subset contains synthetic signals, solo instruments, speech, and music.

We performed 3 experiments, derived from the experimental protocol in [21] so that results are comparable between studies. The experiments are described and discussed below. For wavelet analysis and synthesis, we used the filter bank methods in the open source Large Time-Frequency Analysis Toolbox (LTFAT [40], `http://ltfat.github.io/`), where our implementation of *wavelet PGHI* is available by using the `'wavelet'` flag in `filterbankconstphase`. A function to generate the wavelet filters and scripts for generating the individual experiments are provided on `http://ltfat.github.io/notes/055/`, together with audio examples for all experiment conditions. The functionality represented by the supplied code is to be integrated into the next release of the LTFAT Toolbox.

In Experiments I and II, we consider the following wavelet types in addition to the Cauchy wavelet: Morlet wavelets, generalized Morse wavelets [41] with symmetry parameter $\gamma \in \{2,3\}$ and bandlimited wavelets generated in the Fourier domain as cardinal B-spline of order $m \in \{3,5\}$. The latter have previously been called frequency B-spline wavelets [42]. Note that Cauchy wavelets are generalized Morse wavelets with $\gamma = 1$. For $\gamma = 3$, the generalized Morse wavelet is also known as Airy wavelet. In each case, the remaining parameters were adapted to match the bandwidth of Cauchy wavelets of the desired order $(\alpha-1)/2$. In Experiment III, we consider wavelets with compact support in the time domain, namely exponentially modulated B-splines. Here, we fixed the B-spline order $m = 4$. In Table 1, we list all the used wavelet types, their parameters and their Fourier transform $\hat{\psi}$.

The decimation step a_{d} and the number of frequency channels K (without the lowpass filter) were chosen equal to those used for the matched Cauchy wavelet. As quantitative error measure, we employ (wavelet) spectral convergence [43], 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}:

$$SC(s_{\text{p}},s_{\text{t}}) = 20\log_{10}\frac{\|M_{s_{\text{p}}} - M_{s_{\text{t}}}\|}{\|M_{s_{\text{t}}}\|}.$$

It should be noted that the wavelet coefficient magnitude in the above formula was computed using the same parameter set for which phaseless reconstruction was attempted. There is no unique method to match the Cauchy wavelet parameter α to another type of mother wavelet. We use a procedure that determines α such that the peak-normalized frequency responses of a given mother wavelet ψ and a Cauchy wavelet $\psi^{(\alpha)}$ of order $(\alpha-1)/2$ with the same central frequency have the same width at a given threshold height $h_{\text{thr}} > 0$. The value of α is computed by the MATLAB function `wpghi_findalpha.m`, supplied on the project webpage `http://ltfat.github.io/notes/055/`.

4.1. Experiment I—Comparison to Fast Griffin-Lim

To study the performance of the proposed algorithm for various mother wavelets and parameter settings, we compare wavelet PGHI to the iterative *fast Griffin-Lim* [16, 27] algorithm. The experimental protocol is similar to Experiment I in [21]. The Cauchy WT serves as a baseline comparison, it is specified by the parameter tuple $(\alpha, a_{\text{d}}, K)$. In this experiment, we considered the following settings: $(30,10,100)$, $(300,24,240)$, and $(3000,40,400)$, leading to a fixed redundancy[3] $K/a_{\text{d}} = 10$. For all settings, the channel center frequencies where geometrically spaced in $\frac{\xi_s}{20} \cdot [2^{-6}, 2^{3.3}]$. Matching the other wavelet types to the given values of α, we obtained the parameters given in Table 2.

We compare three different methods: wavelet PGHI (WPGHI, *proposed*), fast Griffin-Lim with zeros initialization (0-FGLIM, [27]) and fast Griffin-Lim initialized with the result of WPGHI (W-FGLIM). Fast Griffin-Lim was restricted to at most 40 iterations. Nonetheless, it should be noted that the execution time of WPGHI is a small fraction of the time required for either 0-FGLIM or W-FGLIM. Maximum, median, and minimum values for spectral convergence of the three methods are shown in Figure 1 for the different parameter sets $(\alpha, a_{\text{d}}, K)$.

The 0-FGLIM baseline shows the most stable performance across conditions, with little dependence on the mother wavelet or the parameter α. On the tested signals, it also shows the least dependence on the signal content, with 7–10 dB difference in spectral convergence between the best and worst result on any fixed condition. Notably, 0-FGLIM performs slightly better for the lowest value of α, i.e., the wavelets with worst frequency resolution.

The median performance of WPGHI is better than 0-FGLIM for $\alpha \in \{300, 3000\}$, but not for $\alpha = 30$, except when the Cauchy wavelet is used. At low values of α, the Cauchy wavelet is very asymmetric, and thus most different from the other considered mother wavelets. That the Morse wavelet with $\gamma = 3$ performs second best for $\alpha = 30$ gives further indication that the performance of WPGHI depends, as expected, on the closeness of the

[2] The chosen signals are as follows: 01, 02, 04, 14, 15, 16, 27, 39, 49, 50, 51, 52, 53, 54, 70. For each signal, tests were performed on the first 5 seconds of the signal.

[3] In [21], the redundancy $K/a_{\text{d}} = 20$ was used, but Experiment II in [21] suggests that $K/a_{\text{d}} = 10$ still provides excellent performance of all methods.

Wavelet name	Parameters	Fourier transforms $\hat{\psi}(\xi)$
Generalized Morse (M•)	$\alpha > 1, \gamma > 0$	$c_{\alpha,\gamma} \cdot \xi^{(\alpha-1)/2} e^{-2\pi\xi^{\gamma}}$
Morlet (M)	$\sigma > 0$	$c_{\sigma} \cdot (e^{-(\sigma-\xi)^2/2} - e^{-(\sigma^2+\xi^2)/2})$
Frequency B-spline (FB•)	$m \in \mathbb{N}, \xi_{fb} \geq 2$	$c_{m,\xi_{fb}} \cdot B_m(\xi - m\xi_{fb}/4)$
Modulated B-spline (MB•)	$m \in \mathbb{N}, \xi_{fm} \in \mathbb{N}$	$c_{m,\xi_{fm}} \cdot \sin(\pi(\xi - \xi_{fm}))^m / (\pi(\xi - \xi_{fm}))^m$

Table 1: Wavelets used in the experiments. The symbol • is a placeholder for γ or m in the Generalized Morse and B-Spline wavelets, i.e., M3 denotes a Morse wavelet with $\gamma = 3$. For $\gamma = 1$, the generalized Morse wavelet yields the Cauchy wavelet. The parameter σ of the Morlet wavelet is related to the center frequency and controls the time/frequency resolution trade-off. The parameter m of the B-Spline type wavelets denotes the order of the generating B-spline (a B-Spline of order m is a piecewise polynomial of order $m-1$). ξ_{fb} and ξ_{fm} denote the center frequency to bandwidth and center frequency to main lobe width, respectively. The restriction $\xi_{fm} \in \mathbb{N}$ guarantees that the modulated B-spline satisfies $\hat{\psi}(0) = 0$ and is admissible.

Wavelet / Cauchy α	≈ 30	≈ 300	≈ 3000	≈ 1000
M2, $\alpha =$	28.93	298.93	2998.93	999
M3, $\alpha =$	28.45	298.55	2998.45	999
M, $\sigma =$	3.79	12.25	38.78	22.38
FB3, $\xi_{fb} =$	—	4.32	13.70	7.90
FB5, $\xi_{fb} =$	—	3.25	10.30	5.94

Table 2: Wavelets parameters to match a given Cauchy parameter $\alpha \in \{30, 300, 3000\}$ used in Experiment I and $\alpha = 1000$ used in Experiment II. Note that $\alpha \approx 30$ cannot be achieved with the frequency B-Spline wavelet of order $m \in \{3, 5\}$.

mother wavelet to the Cauchy wavelet with the value of α, for which the phase-magnitude relations (17) and (18) are invoked.

As expected, WPGHI performs worse when the Cauchy wavelet is not used, but at large values of α, the difference in median and worst values is small. Over all values of α, the Frequency B-Spline wavelet of order $m = 3$ performs worst, and the Morse wavelet with $\gamma = 2$ performs closest to the Cauchy wavelet. The large range of values for WPGHI corroborates the observations from [21] that WPGHI performance can depend significantly on the signal content, for all chosen mother wavelets. Furthermore, the best performance seems to depend heavily on the chosen mother wavelet and its closeness to the Cauchy wavelet. Initializing fast Griffin-Lim with the result of WPGHI (W-FGLIM) shows significant improvements over either WPGHI or 0-FGLIM in all considered scenarios. In most cases, the final performance of W-FGLIM seems to be proportional to the quality of the WPGHI initialization.

Informal listening mostly confirmed the numerical results. At $\alpha \in \{300, 3000\}$, all methods produce little to no audible distortion, with the exception of 0-FGLIM for simple signals such as synthetic sine waves where the defect is still clearly audible. At $\alpha = 30$, WPGHI sometimes produces results that are perceptually worse when the Cauchy wavelet is not used, but nonetheless, distortions were often more severe in 0-FGLIM, despite contradicting numerical results. The results of W-FGLIM provide excellent quality, even at $\alpha = 30$, where the individual methods may fail to do so.

4.2. 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. Once more, the experiment follows the protocol established in Experiment II in [21], but with the main aim to compare performance across different mother wavelets. We fix $\alpha = 1000$ and once more match the parameters of all alternative mother wavelets, see Table 2. The considered redundancies are $K/a_{\mathrm{d}} \in \{3, 5, 10\}$. In contrast to [21], we do not consider $K/a_{\mathrm{d}} = 30$, as the results were rather close to the case $K/a_{\mathrm{d}} = 10$ and the same is expected here.

We fix the following parameter sets $(\alpha, a_{\mathrm{d}}, K)$ for the Cauchy wavelet baseline: low redundancy $(1000, 30, 90)$, medium redundancy $(1000, 25, 125)$, high redundancy $(1000, 18, 180)$. Similar to Experiment I, median value, maxima, and minima over the test set are presented in Figure 2 for all parameter sets and mother wavelets. As expected, performance of both proposed methods decreases at lower redundancy, but median performance at low redundancy is still decent. Both median and worst performance shows little dependence on the mother wavelet, due to the chosen large value of α, at which all considered mother wavelets are reasonably close to the Cauchy wavelet. Some influence of the mother wavelet is still apparent in the best values of spectral convergence. Especially for plain WPGHI, the Cauchy wavelet is still at an advantage. At any redundancy, W-FGLIM yields a significant improvement over plain WPGHI, but at low redundancy, the additional Griffin-Lim iteration only marginally improves the reconstruction on signals for which WPGHI performs badly. Perceptual quality is excellent over all redundancies. Only at low redundancy, minor distortions were observed.

4.3. Experiment III—Towards WPGHI with Bounded Delay

The implementation of a bounded delay framework for wavelet analysis and synthesis is more involved than for the short-time Fourier transform and not the objective of this contribution. Nonetheless, we want to indicate some steps that can be taken to enable the use of WPGHI within such a framework. First, we need to show that WPGHI produces good results in conjunction with mother wavelets that are compactly supported in the time domain and thus necessarily $\hat{\psi}(\xi) = 0$ for all $\xi \in \mathbb{R}^-$ cannot be satisfied.

To this end, we repeat Experiment I with a modulated B-spline of fixed order 4 as mother wavelet ψ. Due to the restriction of the parameter ξ_{fm} to positive integers, it is not possible to construct complex-modulated B-spline wavelets that match a Cauchy wavelet of arbitrary order $(\alpha-1)/2$. Instead we choose values for the center frequency to main lobe width ratio ξ_{fm} and compute the matching Cauchy parameter α. The resulting parameter values are shown in Table 3.

The Fourier transform $\hat{\psi}$ of the modulated B-Spline is symmetric around its peak. Hence, the lowest value of ξ_{fm} corre-

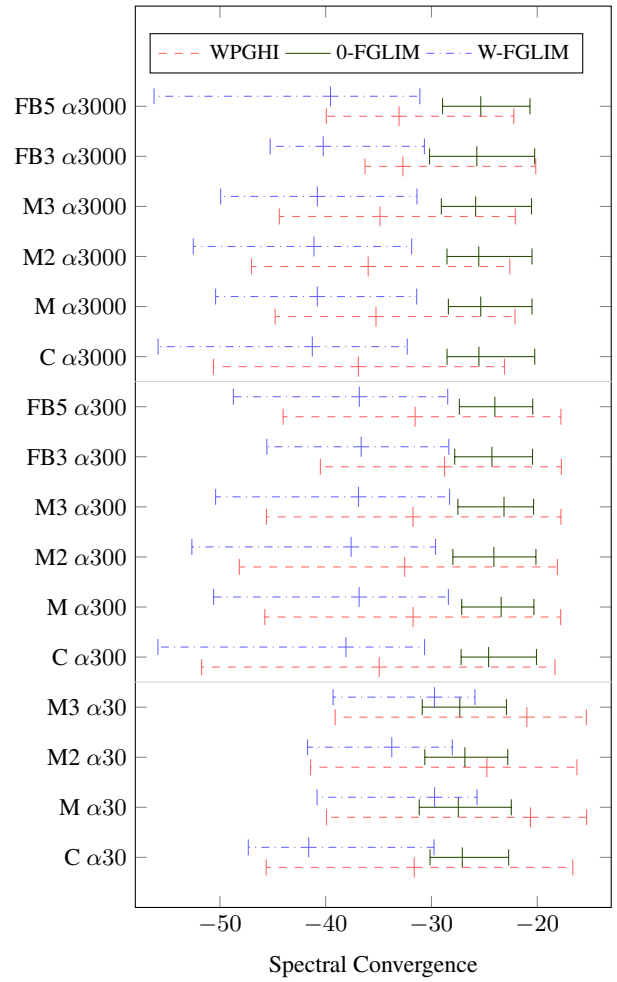

Figure 1: Results of Experiment I for wavelet PGHI and fast Griffin Lim for Cauchy $\alpha \in \{30, 300, 3000\}$ and all tested wavelets. The minimal, median, and maximal spectral convergences over the 15 test signals are depicted.

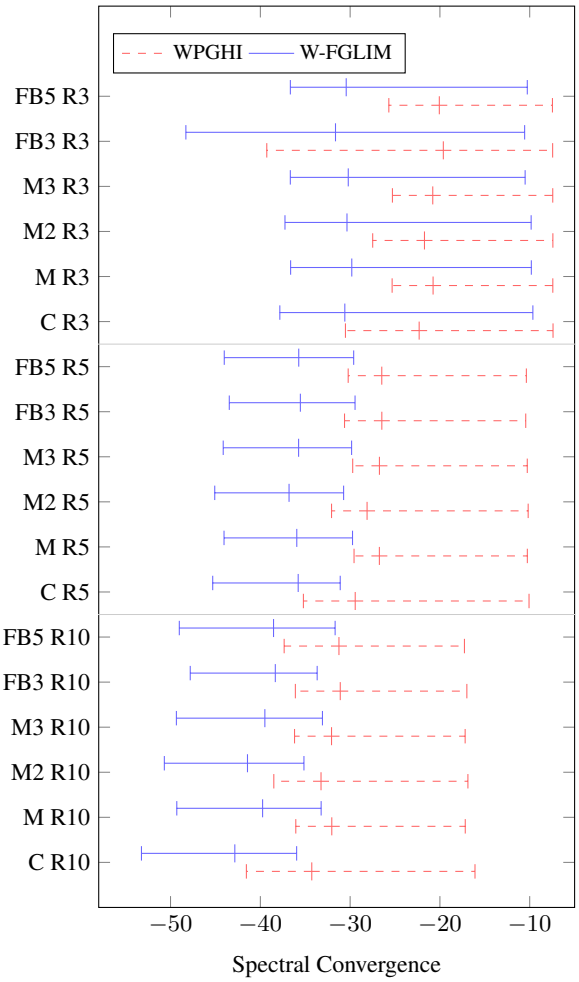

Figure 2: Results of Experiment II for wavelet PGHI and all tested wavelets at redundancies $K/a_d \in \{10, 5, 3\}$. The minimal, median, and maximal spectral convergences over the 15 test signals are depicted.

sponds to a small Cauchy wavelet parameter α, i.e., a rather asymmetric Cauchy wavelet. Thus, the average performance of WPGHI is expected to be inhibited for the modulated B-Spline. While this is certainly apparent in the results, see Figure 3, reconstruction quality is still decent. At higher values of ξ_{fm}, the difference between WPGHI performance for the two tested mother wavelets becomes increasingly negligible and reconstruction performance becomes competitive with or even better than the iterative fast Griffin-Lim algorithm. For the sake of completeness, we also show results for W-FGLIM, which continues to outperform both competing methods. Generally speaking, the results are, not unexpectedly, very similar to those obtained in Experiment I. The same is true for perceptual performance, where differences between the wavelets (in line with numerical results) have been observed almost exclusively for $\xi_{fm} = 1$.

Both 0-FGLIM and W-FGLIM involve iteration relying on many wavelet analysis and synthesis steps. On the other hand, the computations necessary for WPGHI are elementary and can be adapted to a real-time (bounded delay) setting easily. This has previously been shown for the short-time Fourier transform in [28].

We close this contribution with Algorithm 2, which demonstrates the changes necessary to adapt WPGHI to such a setting. This algorithm computes a phase estimate $(\tilde{\phi}^s_\psi)_{\text{est}}[n, \cdot]$ using the phase gradient at time positions $n - 1$ and n and relying on a previously computed phase estimate for $(\tilde{\phi}^s_\psi)_{\text{est}}[n - 1, \cdot]$, for $n = 1, \ldots, N - 1$. Assuming that the phase and phase derivative *at time* -1 are identically 0, it can also be used to initialize the phase estimate for $(\tilde{\phi}^s_\psi)_{\text{est}}[0, \cdot]$ from scratch.

5. CONCLUSION

We have presented a non-iterative method for reconstruction from magnitude-only wavelet coefficients, relying on the phase-magnitude relations for WTs with Cauchy-type mother wavelet, recently introduced in [21]. The resulting algorithm is computationally highly efficient and often performs on par or better than previous iterative schemes, for which it can also serve as an initialization. The latter initialization has been shown to boost the performance of either method used individually. In the presented experiments we showed that the theoretical restriction to Cauchy-type mother wavelets

Cauchy, $\alpha =$	29.7	257.0	2841.7
Mod. B-spline, $m = 4$, $\xi_{fm} =$	1	3	10

Table 3: Wavelet parameters used in Experiment III. Each column lists matching parameters between the Cauchy wavelet and the modulated B-spline of order $m = 4$.

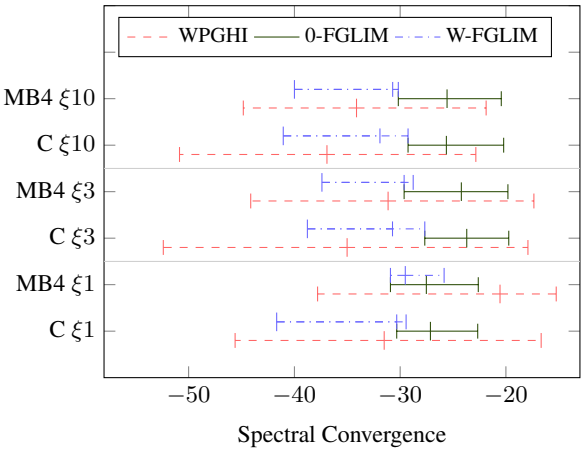

Figure 3: Results of Experiment III for wavelet PGHI and fast Griffin Lim for Cauchy and complex modulated B-spline wavelets. The minimal, median, and maximal spectral convergences over the 15 test signals are depicted.

becomes a soft restriction in practice. In other words, the method can be successfully applied for other types of mother wavelet, provided they are reasonably close to some Cauchy wavelet. Although this *closeness* to a Cauchy wavelet limits the performance of the algorithm when other wavelets are used, the obtained results for a small, but varied corpus of audio data are very promising. Finally, we indicated the steps that are necessary for introducing the proposed method into a bounded delay wavelet analysis/synthesis system, similar to what has been done for the short-time Fourier transform in [28].

Algorithm 2: *Causal* wavelet PGHI for the n-th time position

Input: Magnitude $M_s[m, \cdot]$ wavelet coefficients and time-direction phase derivatives $\Delta_{\tilde{\psi}}^{\tilde{\phi},x,s}[m, \cdot]$ for $m \in \{n-1, n\}$. Frequency-direction phase derivatives $\Delta_{\tilde{\psi}}^{\tilde{\phi},\xi,s}[n, \cdot]$, phase estimate $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}[n-1, \cdot]$ at time position $n-1$ and relative tolerance tol.

Output: Phase estimate $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}[n, \cdot]$ at time position n.

$abstol \leftarrow tol \cdot \max_{m \in \{n-1,n\}} (M_s[n, k])$;
Create set $\mathcal{I} = \{(n, k) : M_s[n, k] > abstol\}$;
Assign random values to $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}(n, k)$ for $k \notin \mathcal{I}$;
Construct a self-sorting max *heap* [38] for (n, k) pairs;
for $(n-1, k) \in \mathcal{I}$ **do**
    Move $(n-1, k)$ from $\mathcal{I}$ into the *heap*;
end
while $\mathcal{I}$ is not $\emptyset$ **do**
    if *heap* is empty **then**
        Move $(n_m, k_m) = \arg\max_{(n,k) \in \mathcal{I}} (M_s[n, k])$ from $\mathcal{I}$ into the *heap*;
        $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}(n_m, k_m) \leftarrow 0$;
    end
    while *heap* is not empty **do**
        $(n_h, k_h) \leftarrow$ remove the top of the *heap*;
        if $n_h = n - 1$ *and* $(n, k_h) \in \mathcal{I}$ **then**
            Compute $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}(n, k_h)$ by means of (19);
            Move (n, k_h) from $\mathcal{I}$ into the *heap*;
        end
        if $n_h = n$ **then**
            foreach (n_n, k_n) in $\{(n, k \pm 1)\} \cap \mathcal{I}$ **do**
                Compute $(\tilde{\phi}_{\tilde{\psi}}^{s})_{\text{est}}(n_\text{n}, k_\text{n})$ by means of (19);
                Move (n_n, k_n) from $\mathcal{I}$ into the *heap*;
            end
        end
    end
end

6. 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. 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.

[3] 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.

[4] 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.

[5] 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.

[6] T. Necciari, 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] I. Waldspurger, "Phase retrieval for wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2993–3009, May 2017.

[13] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug. 1982.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] N. Holighaus, G. Koliander, Z. Průša, and L. D. Abreu, "Characterization of analytic wavelet transforms and a new phaseless reconstruction algorithm," *submitted to IEEE Trans. Sig. Process.*, 2018. [Online]. Available: http://ltfat.github.io/notes/ltfatnote053.pdf

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] N. Perraudin, P. Balazs, and P. L. Sondergaard, "A fast Griffin-Lim algorithm," in *Proc. IEEE Appl. Sig. Process. Audio Acoustics*, New Paltz, NY, USA, Oct. 2013.

[28] 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.

[29] P. Flandrin, "Separability, positivity, and minimum uncertainty in time-frequency energy distributions," *J. Math. Phys.*, vol. 39, no. 8, pp. 4016–4040, Aug. 1998.

[30] 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.

[31] Z. Cvetkovic and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[32] 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.

[33] 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.

[34] K. Gröchenig, "Acceleration of the frame algorithm," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3331–3340, Dec. 1993.

[35] 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.

[36] 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.

[37] 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.

[38] J. W. J. Williams, "Algorithm 232: Heapsort," *Communications of the ACM*, vol. 7, no. 6, pp. 347–348, Jun. 1964.

[39] "Tech 3253: Sound Quality Assessment Material recordings for subjective tests," Eur. Broadc. Union, Geneva, Tech. Rep., Sept. 2008.

[40] 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.

[41] S. C. Olhede and A. T. Walden, "Generalized Morse wavelets," *IEEE Trans. Sig. Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002.

[42] R. X. Gao and R. Yan, *Wavelets: Theory and applications for manufacturing*. Springer Science & Business Media, 2010.

[43] N. Sturmel 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.