An Analytic Wavelet Transform with a Flexible Time-Frequency Covering

İlker Bayram
ibayram@itu.edu.tr

Abstract—We develop a rational-dilation wavelet transform for which the dilation factor, the Q-factor and the redundancy can be easily specified. The introduced transform contains Hilbert transform pairs of atoms, therefore it is also suitable for oscillatory signal processing. The transform may be modified to obtain a tight chirplet frame for discrete-time signals. A fast implementation, that makes use of an equivalent filter bank, makes the transform suitable for long signals. Examples on natural signals are provided to demonstrate the utility of the transform.

I. INTRODUCTION

Wavelet transforms with a high Q-factor allow a multiscale analysis with a high frequency resolution. Typical domains of application involve oscillatory signals (like audio, various biomedical signals, etc.). Since the analytic signal is instrumental for processing or extracting information from such signals [17], it is of interest that the wavelet transforms contain Hilbert transform pairs of atoms. In this paper, we introduce a wavelet transform that hosts Hilbert transform pairs of atoms and also allows easy control over parameters like the dilation factor, redundancy and the Q-factor. We also discuss a modification that leads to a tight chirplet frame for discrete-time signals.

A typical covering of the time-frequency plane by a wavelet frame is shown in Fig. 1. Three parameters stand out in this figure : (i) the dilation factor, 'd'; (ii) the Q-factor, '$f/\Delta\omega$'; (iii) the shift parameter, 'Δt'. These parameters are not independent of each other. If we ask that the frame be tight, the dilation factor sets an upper bound on the Q-factor. If we ask that wavelets have small time-frequency supports (subject to the uncertainty principle [17]), the Q-factor sets an upper bound on the shift parameter. We also note that the redundancy of the transform is a function of these three parameters. The introduced wavelet transform allows to easily set these parameters, subject to the outlined constraints.

There are close relations between the rational dilation wavelet transform (RADWT) [7], the tunable-Q wavelet transform (TQWT) [33] and the introduced transform. Both the RADWT and TQWT are real transforms obtained by iterating a filter bank (FB) with two channels. Due to
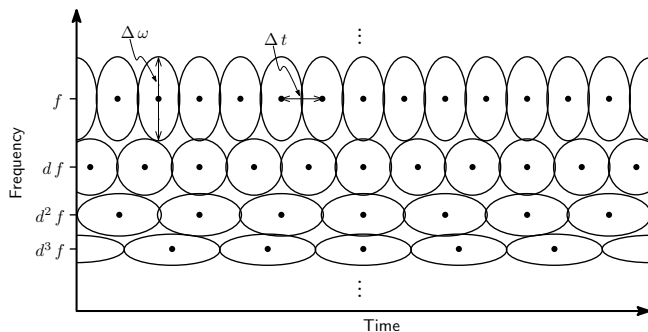

Fig. 1. Time-frequency covering for a wavelet frame. Here, we emphasize three parameters : the dilation factor d, the Q-factor $\Delta f/f$, the shift parameter Δt. The proposed transform allows to easily set these three parameters.

difficulties presented by employing a highpass filter in a rational rate-changer, the highpass channel of the RADWT consists of a filter, followed by a downsampler; i.e. the sampling factors for the highpass channel are restricted to rational numbers of the form $1/s$ where s is an integer. This, in turn, leads to a transform that is rather rigid – it is not easy to control the number of oscillations in the mother wavelet (which is related to the Q-factor) and the redundancy of the frame. With respect to Fig. 1, the RADWT is able to employ arbitrary rational dilation factors, but the set of allowed $\Delta\omega$ and Δt values is rather sparse. The TQWT, introduced by Selesnick in [33], overcomes the mentioned difficulty (of employing a highpass filter in a rational rate changer) by introducing a novel operation called 'highpass scaling'. TQWT, obtained this way, allows to easily set the Q-factor of the wavelet and the redundancy of the frame. In addition, the gained flexibility can be used to obtain a fast implementation. However, given the Q-factor and the redundancy, TQWT does not allow to select the dilation factor. This is not a very desirable feature for music signal processing where the dilation factor can be tuned to decompose octaves into equal number of bins [32], [19]. Here, we split the positive and negative frequency parts of the filter used in the highpass channel – this allows us to employ arbitrary sampling rates in the highpass channels. This in turn leads to a flexible transform where the Q-factor, redundancy and the dilation factor can be easily set. This
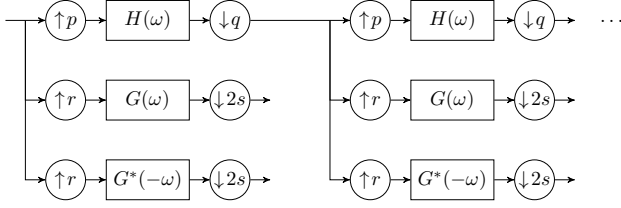
Fig. 2. The proposed transform consists of an iterated filter bank. This filter bank with (almost) analytic filters shown in Fig. 4 gives a complex transform. To obtain a real transform we take the sum and differences of the positive-frequency and negative-frequency subbands (with proper weighting in order to preserve the tightness of the frame).

Fig. 3. The analysis and synthesis filter banks used in the proposed transform.

flexibility can also be used to obtain a fast implementation with exact perfect reconstruction for finite length signals. In addition to these, separation of the positive and negative frequency allows to easily obtain a real transform hosting Hilbert transform pairs of atoms – a feature that is not available for either RADWT or TQWT.

The proposed transform is realized by the iterated FB shown in Fig. 2. The underlying FB (shown in Fig. 3) consists of one lowpass and two highpass channels. One of these highpass channels analyze 'positive frequencies'; the other analyzes 'negative frequencies'. Typical frequency response characteristics, showing the passbands, transition bands of the filters are sketched in Fig. 4.

We note that the construction outlined above cannot be obtained as a special case of the dual-tree type transforms [34]. Dual-tree type transforms employ two tight frames and therefore they are redundant by a factor of at least two. In contrast, the transform in this paper can achieve an arbitrary redundancy. The construction also differs from the realization in [29], which applies a wavelet transform on the analytic signal derived from the input.

As one of the reviewers noted, the construction in this paper can be easily extended so as to employ M band-pass/highpass filters covering positive frequencies and M bandpass/highpass filters covering negative frequencies. In that case, we would have M different Hilbert-transform pairs of mother wavelets. In this paper, we restrict our attention to the case $M = 1$.

Related Work

Rational-dilation wavelet transforms can be obtained by iterating filter banks with rational sampling factors [10]. Wavelet transforms with different constraints have been proposed following this general schema. Critically sampled FBs with finite impulse response (FIR) filters are discussed in [28], [25], [12], [6]. Reference [5] discusses critically
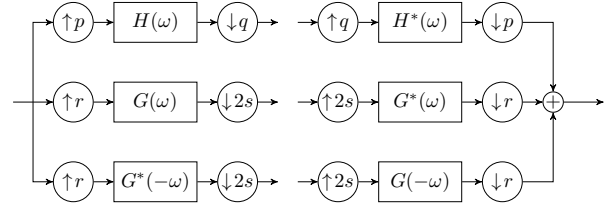
sampled filter banks with ideal filters (implemented using the fast Fourier transform (FFT) – leading to Shannon-like wavelets). Overcomplete filter banks with FIR filters are discussed in [8], [6]. References [7], [33] discuss overcomplete filter banks designed in the frequency domain, implemented using FFTs. Also, [9] develops an analytic rational dilation wavelet transform, based on the RADWT, using the dual-tree framework [34].

Another line of work that adresses constant-Q analysis can be found in [36], [30], [13], [14], [32]. The idea is to modify the window so that for different center-frequencies, different windows are used, achieving a constant-Q analysis in the end. For an earlier paper, discussing a more general form, see [21]. In a similar vein, a general family of frames, called 'nonstationary Gabor frames' were recently introduced in [2] (also see [19]). This family generalizes the notion of a Gabor frame by employing multiple (possibly unrelated) windows and window-dependent sampling frequencies.
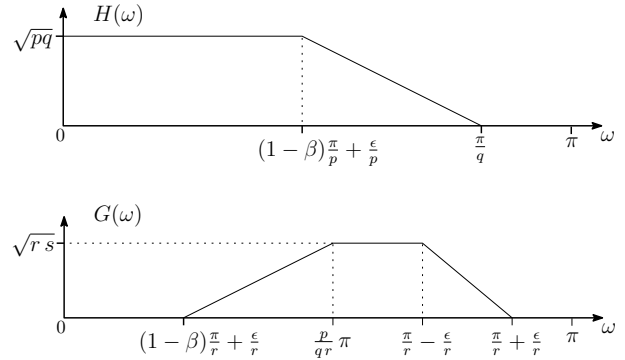

Fig. 4. The transition bands of the filters used in the FB in Figure 3.

Notation

Discrete-time sequences are denoted by small case letters, as in $h(n)$. DTFTs of discrete-time sequences are denoted by capital letters as in $H(\omega)$, where,

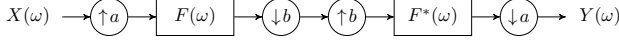$$H(\omega) = \sum_{n \in \mathbb{Z}} h(n)\, e^{-j\omega n}. \tag{1}$$

Fig. 5. The system in Fig. 3 consists of linear combinations of systems like the one shown above.
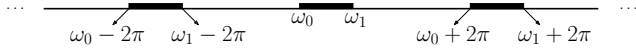

Fig. 6. If the frequency support of the the filter $F(\omega)$ in Fig. 5 is restricted as indicated by the thick segments, then the system in Fig. 5 becomes linear time-invariant.

Fig. 7. We study the system in Fig. 5 in two steps, using the two systems above.

We note that DTFTs are periodic by 2π. Therefore it is sufficient to specify a DTFT on an interval of length 2π. We use the intervals $[-\pi, \pi)$ or $[0, 2\pi)$ in different places in the manuscript.

Outline

In Section II, we derive the perfect reconstruction conditions for the FB in Fig. 3 and propose filters that satisfy these conditions. We study the iterated filter bank and describe the atoms (discrete-time wavelets) in Section III. In Section IV we discuss how chirplets can be obtained by introducing a phase term to given PR filters and provide an example use of the obtained chirplet frame. Section V provides the details for the realization of the proposed transform in a computationally efficient way. In Section VI, we apply the proposed transform on natural signals to demonstrate its utility. Section VII is the conclusion.

II. PERFECT RECONSTRUCTION

To derive the perfect reconstruction conditions for the system in Fig. 3, we start by studying the system in Fig. 5. Here we take $a, b \in \mathbb{Z}$, and assume that $b > a$. Provided $F(\omega)$ is appropriately bandlimited, we will show that this system is LTI and we will derive the equivalent filter.

Specifically, consider a filter $F(\omega)$ whose frequency response is supported on a single interval $[\omega_0, \omega_1]$ for $\omega \in [\omega_0, \omega_0 + 2\pi)$ (see Fig. 6). Also, assume that,

$$\omega_1 - \omega_0 \leq \frac{2\pi}{b} \leq \frac{2\pi}{a}. \tag{2}$$

This implies,

$$F^*(\omega) F\left(\omega + k\frac{2\pi}{b}\right)$$
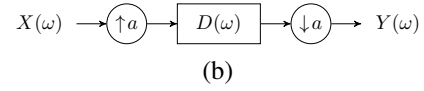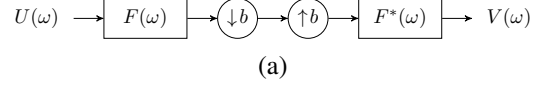$$= \begin{cases} 0, & \text{for} \quad 1 \leq k \leq b-1, \\ |F(\omega)|^2, & \text{for} \quad k = 0. \end{cases} \tag{3}$$

We will carry out our study of the system in Fig. 5 in two steps. Consider first the system in Fig. 7a. For this system, we have,

$$V(\omega) = F^*(\omega) \frac{1}{b} \sum_{k=0}^{b-1} U\left(\omega + k\frac{2\pi}{b}\right) F\left(\omega + k\frac{2\pi}{b}\right). \tag{4}$$

By (3), we obtain,

$$V(\omega) = \frac{1}{b} |F(\omega)|^2 U(\omega). \tag{5}$$

In words, provided that (3) holds, the system in Fig. 7a is LTI with frequency response $|F(\omega)|^2/b$.

Consider now the system in Fig. 7b. This time, the output is related to the input as,

$$Y(\omega) = X(\omega) \frac{1}{a} \sum_{k=0}^{a-1} D\left(\frac{\omega}{a} + k\frac{2\pi}{a}\right). \tag{6}$$

Therefore, the system is LTI with frequency response

$$S(\omega) = \frac{1}{a} \sum_{k=0}^{a-1} D\left(\frac{\omega}{a} + k\frac{2\pi}{a}\right). \tag{7}$$

Combining these two observations on the systems in Fig. 7, we reach the following lemma.

Lemma 1. If $F(\omega)$ is bandlimited (with a bandwidth less than $2\pi/b$), as indicated in Fig. 6, then the system in Fig. 5 is linear time-invariant with frequency response,

$$T(\omega) = \frac{1}{ab} \sum_{k=0}^{a-1} \left| F\left(\frac{\omega}{a} + k\frac{2\pi}{a}\right) \right|^2. \tag{8}$$

$\square$

From this lemma, we obtain our main result about the perfect reconstruction condition of the system in Fig. 3.

Proposition 1. Let $H(\omega)$, $G(\omega)$ be the frequency responses of two filters. If

(i) there exist ω_0, ω_1 with $|\omega_1 - \omega_0| < 2\pi/q$, such that $H(\omega)$ is supported on a single interval $[\omega_0, \omega_1]$ for $\omega \in [\omega_0, \omega_0 + 2\pi)$,

(ii) there exist ω_2, ω_3 with $|\omega_3 - \omega_2| < \pi/s$, such that $G(\omega)$ is supported on a single interval $[\omega_2, \omega_3]$ for $\omega \in [\omega_2, \omega_2 + 2\pi)$,

then the system in Fig. 3 is linear time-invariant with frequency response,

$$T(\omega) = \frac{1}{p\,q} \sum_{k=0}^{p-1} \left| H\left(\frac{\omega}{p} + k\frac{2\pi}{p}\right) \right|^2$$
$$+ \frac{1}{2\,r\,s} \sum_{k=0}^{r-1} \left| G\left(\frac{\omega}{r} + k\frac{2\pi}{r}\right) \right|^2 + \left| G\left(-\frac{\omega}{r} + k\frac{2\pi}{r}\right) \right|^2. \tag{9}$$

If, in addition, $T(\omega)$ as defined above is unity for $\omega \in [0, 2\pi)$, then the filter bank is perfect reconstruction.

Proof: The proposition follows by applying Lemma 1 to each branch of the system in Fig. 3. ∎

As a corollary of this result, we can construct a perfect reconstruction filter bank as described in the following proposition.

Proposition 2. Let $\theta(\omega)$ be a function defined for $\omega \in [0, \pi]$ that satisfies,

$$\left| \theta(\omega) \right|^2 + \left| \theta(\pi - \omega) \right|^2 = 1, \tag{10}$$

and β, ϵ positive constants constrained as

$$1 - p/q \leq \beta \leq r/s, \quad \epsilon \leq \left(\frac{p - q + \beta\,q}{p + q} \right) \pi. \tag{11}$$

Also, let $H(\omega)$, $G(\omega)$ be the frequency responses of two filters (therefore, periodic by 2π) defined as (see Fig. 4),

$$H(\omega) = \begin{cases} \sqrt{p\,q}, & |\omega| < \omega_p, \\ \sqrt{p\,q}\,\theta\left((\omega_s - \omega_p)^{-1}(\omega - \omega_p)\right), \\ & \omega_p \leq \omega \leq \omega_s, \\ \sqrt{p\,q}\,\theta\left((\omega_s - \omega_p)^{-1}(\pi - \omega + \omega_p)\right), \\ & -\omega_s \leq \omega \leq -\omega_p, \\ 0, & |\omega| \geq \omega_s, \end{cases} \tag{12}$$

where $\omega_p = (1 - \beta)\,\pi/p + \epsilon/p$, $\omega_s = \pi/q$,

$$G(\omega) = \begin{cases} \sqrt{r\,s}\,\theta\left((\omega_1 - \omega_0)^{-1}(\pi - \omega - \omega_0)\right), \\ & \omega_0 \leq \omega \leq \omega_1, \\ \sqrt{r\,s}, & \omega_1 \leq \omega < \omega_2, \\ \sqrt{r\,s}\,\theta\left((\omega_3 - \omega_2)^{-1}(\omega - \omega_2)\right), \\ & \omega_2 \leq \omega \leq \omega_3, \\ 0, & \omega \in [0, \omega_0) \cup (\omega_2, 2\,\pi), \end{cases} \tag{13}$$

where $\omega_0 = (1 - \beta)(\pi + \epsilon)/r$, $\omega_1 = p\,\pi/(q\,r)$, $\omega_2 = (\pi - \epsilon)/r$, $\omega_3 = (\pi + \epsilon)/r$. Then, the system in Fig. 3 is perfect reconstruction.

Proof: The proposition follows by applying Prop. 1 for the system with the filters $H(\omega)$, $G(\omega)$ specified as above. ∎

An example for a $\theta(\omega)$ function satisfying (10) is given by [7], [33],

$$\theta(\omega) = \frac{1}{2}(1 + \cos\omega)\sqrt{2 - \cos\omega} \quad \text{for } \omega \in [0, \pi]. \tag{14}$$

In addition to this choice of the transition function $\theta(\omega)$, we also set

$$\epsilon = \frac{1}{32}\left(\frac{p - q + \beta\,q}{p + q}\right)\pi \tag{15}$$

to produce the figures in this paper.

Since the phases of the filters do not appear in the perfect reconstruction condition '$T(\omega) = 1$' (see (9)), we can modify the phase of $G(\omega)$ without affecting the perfect reconstruction property. We will make use of this observation to construct chirplets in Section IV, so we state it here as a corollary.

Corollary 1. Let $H(\omega)$, $G(\omega)$ be defined as in Prop. 2 so that the FB is PR. If we replace $G(\omega)$ with $G(\omega)\,e^{-j\phi(\omega)}$ (and similarly, replace $G^*(-\omega)$ with $G^*(-\omega)\,e^{j\phi(-\omega)}$) where $\phi(\omega)$ is a 2π periodic function, then the resulting FB is also PR. □

III. ATOMS OF THE TRANSFORM

The iterated filter bank in Fig. 2 computes inner products of the input with the atoms of the transform (or the so-called discrete-time wavelets). In this section we describe these atoms and study their properties.

We denote the kth atom in the ith stage as $d_k^i(n)$. Here, the parameter 'i' is associated with frequency and 'k' is associated with time. To find an expression for $d_k^i(n)$, we study the system shown in Fig. 8b – notice that the diamond block represents a correlator (see Fig. 8a). This system computes the inner product of the input with $d_k^i(n)$. In order to derive the expression for $D_k^i(\omega)$ (the DTFT of $d_k^i(n)$), we will use the following auxiliary result. The proof is an application of Parseval's theorem.

Lemma 2. The system in Fig. 9 computes $\frac{1}{2\pi}\int_{-\pi}^{\pi} U(\omega)\,E^*(\omega)\,d\omega$ for some $E^*(\omega)$. In particular,

(a) If $F(\omega) = 0$ for $2\pi/b \leq \omega \leq 2\pi$, then,

$$E(\omega) = \frac{1}{a}F^*\left(\frac{\omega}{a}\right)D\left(\frac{b}{a}\omega\right) \quad \text{for } 0 \leq \omega < 2\pi. \tag{16}$$
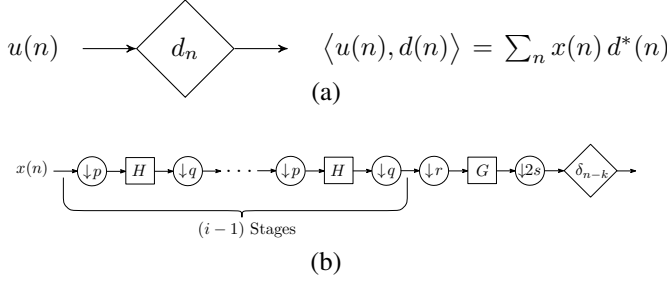
Fig. 8. (a) The diamond block indicates a correlator, which computes the inner product of the input with the sequence $d(n)$. (b) This system computes the inner product of $x(n)$ with $d_k^i(n)$ (the kth atom in the ith stage).
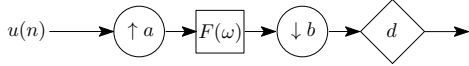

Fig. 9. This system computes the inner product of $u(n)$ with some $e(n)$. Lemma 2 gives a description of $E(\omega)$ for two cases of interest.

(b) If $F(\omega) = 0$ for $\pi/b \le |\omega| \le \pi$, then,

$$E(\omega) = \frac{1}{a} F^* \left(\frac{\omega}{a} \right) D \left(\frac{b}{a} \omega \right) \quad \text{for } -\pi \le \omega < \pi. \tag{17}$$

$\square$

Particularly, if we set $d = \delta(n-k)$, $a = r$, $b = 2s$, $F(\omega) = G(\omega)$, then part (a) can be used to derive the atoms of the first stage as,

$$D_k^1(\omega) = \frac{1}{r} G^* \left(\frac{\omega}{r} \right) e^{-j\omega \, k \, 2s/r} \quad \text{for } 0 \le \omega < 2\pi. \tag{18}$$

Similarly, if we set $d = d_k^i$ (the kth atom in the ith stage), $a = p$, $b = q$, $F(\omega) = H(\omega)$, we can obtain d_k^{i+1} (the kth atom in the $(i+1)^{\text{st}}$ stage), by invoking part (b) as,

$$D_k^{i+1}(\omega) = \frac{1}{p} H^* \left(\frac{\omega}{p} \right) D_k^i \left(\frac{q}{p} \omega \right) \quad \text{for } |\omega| \le \pi. \tag{19}$$

Noting that the frequency support of $D_k^1(\omega)$ is $\left[(1-\beta)\pi + \epsilon, \pi + \epsilon \right]$ for $0 \le \omega < 2\pi$ and that the frequency support of $H(\omega)$ is $[-\pi/q, \pi/q]$ for $-\pi \le \omega < \pi$, we obtain the frequency support of $D_k^2(\omega)$ as $\left[\frac{p}{q}\left((1-\beta)\pi + \epsilon \right), \frac{p}{q}\pi \right]$. By an induction argument, we can obtain the frequency support of $D_k^i(\omega)$ as,

Frequency Support of $D_k^i(\omega)$

$$= \left[\left(\frac{p}{q} \right)^{i-1} \left((1-\beta)\pi + \epsilon \right), \left(\frac{p}{q} \right)^{i-1} \pi \right]. \tag{20}$$

This note on the frequency support of $D_k^i(\omega)$ allows us to express it as[1],

$$D_k^i(\omega) = \begin{cases} D_k^1 \left(\left(\frac{q}{p} \right)^{i-1} \omega \right) \prod_{m=0}^{i-2} \frac{1}{p} H^* \left(\left(\frac{q}{p} \right)^m \frac{\omega}{p} \right), \\ \qquad\qquad\qquad\qquad 0 \le \omega < \left(\frac{p}{q} \right)^{i-1} \pi, \\ 0, \qquad\qquad \left(\frac{p}{q} \right)^{i-1} \pi \le \omega < 2\pi. \end{cases} \tag{21}$$

The Quality Factor of the Atoms

In [33], Selesnick obtains the Q-factor of the discrete-time wavelets through some mild approximations. Following [33], and noting that $\epsilon \approx 0$, we take the center frequency of $D_k^i(\omega)$ to be the midpoint of its frequency support (recall (20)) as,

$$\omega_c^{(i)} = \left(\frac{p}{q} \right)^{i-1} \frac{2-\beta}{2} \pi. \tag{22}$$

Taking the bandwidth as half of the measure of the frequency support, we have,

$$\text{BW}_i = \frac{1}{2} \left(\frac{p}{q} \right)^{i-1} \beta \pi. \tag{23}$$

Therefore, the Q-factor, defined as the ratio of the center frequency to the bandwidth, is given by,

$$\text{Q} = \frac{\omega_c^{(i)}}{\text{BW}_i} = \frac{2-\beta}{\beta}. \tag{24}$$

We see that β determines the quality factor when plugged into the function $(2-x)/x$. Since $(2-x)/x$ maps $(0,1]$ onto $[1,\infty)$, we can theoretically achieve any desired Q-factor (greater than unity) by adjusting β.

Redundancy of the Transform

Noting that the ratio of the number of output samples of the ith stage to the number of input samples is $\left(\frac{p}{q} \right)^{i-1} (r/s)$, we find the redundancy factor of the transform as,

$$\text{R} = (r/s) \sum_{i=1}^{\infty} \left(\frac{p}{q} \right)^{i-1} = (r/s) \frac{1}{1 - p/q}. \tag{25}$$

---

[1]Here, we take $\prod_{m=0}^{-1} f(m) = 1$.

5

Redundancy, Dilation, and Quality Factors as Parameters

From the preceding discussion, we see that the parameters p, q, r, s, β allow us to easily construct a wavelet transform with a given dilation factor, Q-factor and redundancy. Indeed, suppose we are given a desired dilation factor d, Q-factor 'Q', and redundancy 'R'. First, we choose p, q, so that $p/q \approx d$. Next we choose β so that,

$$Q \approx \frac{2 - \beta}{\beta} \quad \Longleftrightarrow \quad \beta \approx \frac{2}{Q + 1}, \tag{26}$$

Recall that the only constraint on β is $p/q + \beta \geq 1$, hence the dilation factor sets a lower bound on β, thus an upper bound on the Q-factor. Now for the given p, q (with $p/q \approx d$), we choose r, s so that

$$R \approx (r/s) \frac{1}{1 - p/q} \quad \Longleftrightarrow \quad r/s \approx R\,(1 - d). \tag{27}$$

Recall that given β, (r, s) are constrained as $\beta \leq r/s$. Therefore the redundancy has to be greater than $\beta/(1 - d)$ – the Q-factor and the dilation factor set a lower bound for redundancy.

We provide two examples in Fig. 10 where given the Q-factor and redundancy, p, q, β are determined using (26) and (27).

This discussion partially parallels the motivation for the introduction of the 'Tunable-Q Wavelet Transform' (TQWT) in [33]. The TQWT provides a construction flexible enough to allow direct control over the Q-factor and the redundancy. However, given the Q-factor, TQWT adjusts the redundancy by selecting the dilation factor. In applications involving music, the dilation factor can also be a desired parameter (see [19] for instance where 48 bins in an octave are used, requiring $d \approx 2^{1/48}$). In such scenarios, it helps if we have a more flexible transform that allows to set (subject to constraints outlined above) the dilation factor, Q-factor and the redundancy separately. The introduced transform possesses this flexibility. We make use of this in a 'music transposition' application in Section VI (see Experiment 2).

Selection of the Parameters: The parameters of the transform should be chosen by taking into account the signal of interest. Specifically, for a music signal, in order to separate semitones, the dilation factor should not exceed $2^{1/12}$. However, for other applications, like speech coding, dilation factors like $5/6$ have been previously proposed [11].

Following the selection of the dilation factor, the Q-factor determines the frequency resolution of the transform. High-Q values lead to a fine analysis in the frequency domain. However, it should be kept in mind that a very fine frequency analysis leads to poor time-resolution, due to the uncertainty principle.
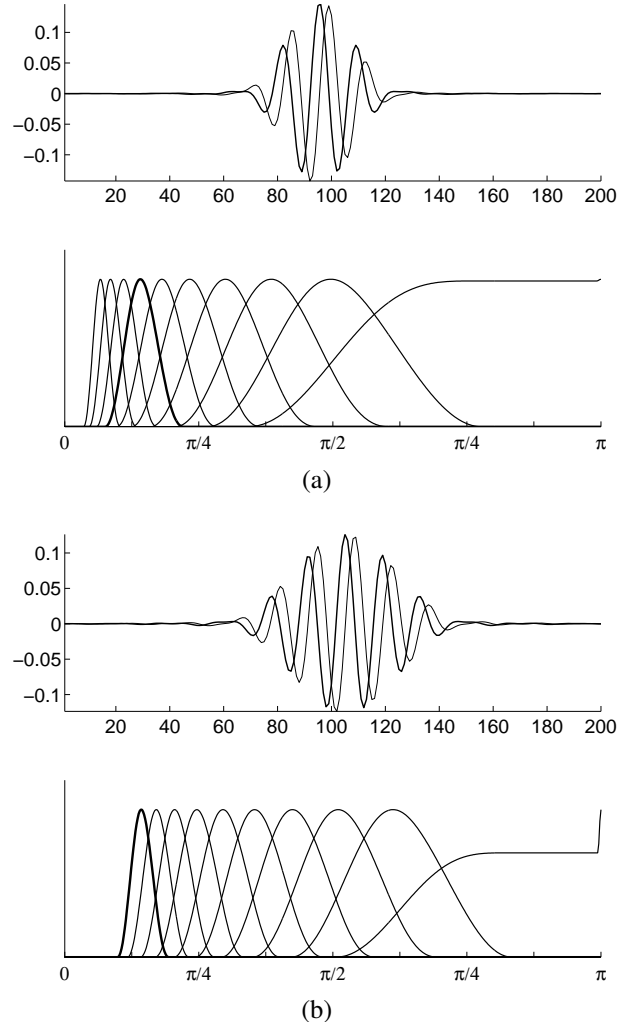

(a)

(b)

Fig. 10. Wavelets from the iterated filter bank. Top panels show the time domain signals (the thick and thin lines are Hilbert transform pairs). Bottom panels show the frequency decomposition obtained by the iterated filter bank. The thick frequency response in each figure corresponds to the time domain wavelet on the top panel. (a) Q = 2, Red = 3; Sampling parameters : $(p, q, r, s) = (7, 9, 2, 3)$, $\beta = r/s$. (b) Q = 3, Red = 3; Sampling parameters : $(p, q, r, s) = (5, 6, 1, 2)$, $\beta = r/s$.

Finally, for fixed dilation and Q-factors, the redundancy allows to vary the time-shift between the wavelets at the same scale/subband. As the redundancy increases, the wavelets are placed closer to each other. For some specific choices in practice, we refer to the experiments in Section VI.

IV. A Tight Discrete-Time Chirplet Frame

Most popular frames cover the time-frequency plane with atoms which have time-frequency supports that are aligned with the time and the frequency axes. In that respect, 'chirplets' [3], [15], [27], [4] are an exception, as thy employ
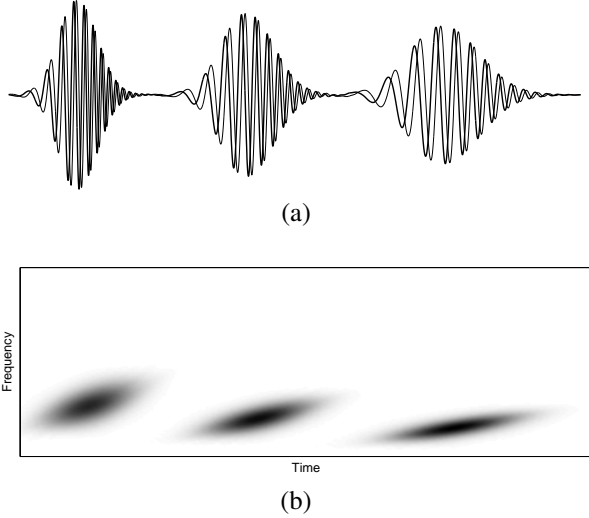
(a)

(b)

Fig. 11. (a) A few atoms from the introduced chirplet frame in the time-domain (the thick and thin functions are Hilbert transform pairs). The parameters are $(p, q, r, s) = (3, 4, 5, 6)$, $\beta = r/s$, $\gamma = 7$. The frame consists of chirplets at different scales. The frame hosts Hilbert-pairs so as to enhance shift-invariance. (b) The (zoomed-in) spectrogram of the atoms.

atoms with oriented time-frequency supports. Chirplets can be useful in a number of different applications ranging from sparsely representing audio [22], analyzing, characterizing visual evoked potentials [18], modelling auditory processing [16], to instantaneous frequency estimation [1]. In this section, we describe how to modify the FB to obtain a tight chirplet frame for discrete-time signals.

As noted in Corollary 1, given a PR FB with filters $H(\omega)$ and $G(\omega)$, we can modify the phase response of the filters without altering the PR property. In particular, by adding a linear group delay to $G(\omega)$, a tight chirplet frame can be obtained. More precisely, for $G(\omega)$ given as in Prop. 2, we can obtain a chirplet transform by employing, instead of $G(\omega)$, a modified filter $\tilde{G}(\omega)$, given by,

$$\tilde{G}(\omega) = G(\omega) \exp\big(-j\,\gamma\,(\omega - \omega_c/r)^2\big) \quad \text{for } 0 \le \omega \le 2\pi,\tag{28}$$

where $\omega_c = \pi\big(1 - \beta/2\big)$ (see (22)) and γ is a 'chirp parameter'.

Multiplication with the term '$\exp\big(-j\,\gamma\,(\omega - \omega_c/r)^2\big)$' shifts (in time) the different frequency components of $G(\omega)$ by different amounts – this is referred to as a 'time-shear' in [27], [4].
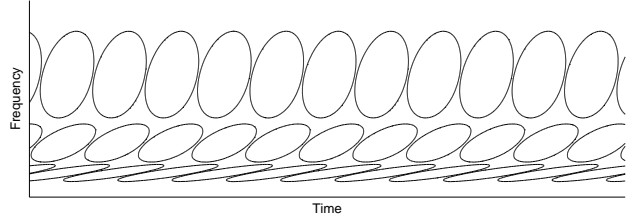

Fig. 12. Tiling of the time-frequency plane with the proposed chirplet atoms. Notice that due to scaling, the time-frequency orientations of the atoms are different in each subband.
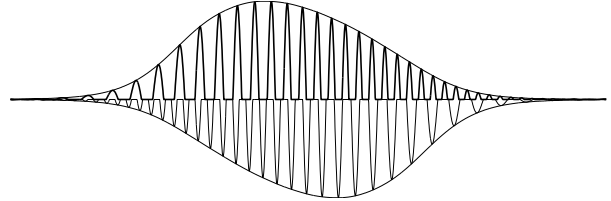

Fig. 13. Above the zero line, the positive part of an atom $d_i^k(n)$ with chirp parameter γ is shown. Below the zero line, the negative part of an atom described by the same stage and position parameters (i, k) but with chirp parameter $-\gamma$, is shown. We see that the time-support and center of the atom is not affected by the sign of the chirp parameter – this is a desired property especially if one wants to form groups of atoms (say, for a signal prior) with different chirp parameters.

Here, the new atoms can be expressed as,

$$\tilde{D}_k^i(\omega) = D_k^i(\omega)\,\exp\!\left(-j\gamma\left(\frac{q}{p}\right)^{2(i-1)}\big(\omega - \omega_c^{(i)}\big)^2\right),$$
$$\text{for } 0 \le \omega \le 2\pi \quad (29)$$

where $D_k^i(\omega)$ is described in (21). A few atoms from the transform with the parameters $(p, q, r, s) = (3, 4, 7, 8)$, $\beta = r/s$, $\gamma = 5$ are shown in Fig. 11. Notice that the time-frequency supports of the atoms are directional. We remark that due to scaling, the actual directions of the time-frequency supports differ from scale to scale. This leads to a covering of the time-frequency plane as depicted in Fig. 12.

The phase modification term for $\tilde{D}_k^i(\omega)$, that is, '$\exp\big(-j\gamma(q/p)^{2(i-1)}\big(\omega - \omega_c^{(i)}\big)\big)$' is centered around the center frequency of the atom. Therefore, the center frequency of the atom is not shifted. This is demonstrated in Fig. 13 which shows two atoms obtained by setting $\gamma = \gamma_0$ and $\gamma = -\gamma_0$ in (29). Because the atoms have similar time supports, they can be used to form groups (this property is useful when certain signal priors, like mixed norms [26] are used). Notice also that the envelopes are skewed. This is primarily due to the fact that the frequency response magnitudes are skewed to start with; shifting (in time) components with different weights leads to a skewed time-domain atom.

To see the effects of a directional time-frequency analysis, consider the speech signal whose wavelet coefficients are shown in Fig. 14a. The signal is a male speaker asking 'Hmm?'. The analysis coefficients for $(p,q,r,s) = (31, 32, 1, 4)$, $\beta = r/s$ and $\gamma = 30$ are shown in the left panel of Fig. 14b. The right panel shows the wavelet coefficients for the same sampling parameters but $\gamma = -30$. Notice that the analysis coefficients for $\gamma = 30$ are better concentrated than those for $\gamma = -30$. This is expected because the atoms for positive chirp parameters have increasing pitch – they are *aligned* (in the time-frequency plane) with utterances where the components have increasing pitch, as is the case for the signal considered here; atoms for negative chirp parameters have decreasing pitch – they lie *across* components with increasing pitch, and therefore we end up with a greater number of significantly large coefficients for this example. Next, we try to find a sparse representation using these frames (we employ a variant of the algorithm in [31]). The $\gamma = 30$ frame leads to a sparser representation compared to the $\gamma = -30$ frame (compare the left and the right panels of 14c).

In brief, chirplets can be useful in applications where the atoms' time-frequency orientation is important, or carries information [24], [16]. Experiment 3 in Section VI makes use of the observations above to decompose a bird call signal into its components.

V. IMPLEMENTATION

In the following, we will use ideas from [33], [19], [2] to describe a fast implementation. We denote the input by x. We assume that the length of the signal, N, is even.

The transform is implemented using DFTs. We denote the orthonormal DFT of x as X. For a length-N sequence x, the relation between x and X is,

$$X(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \, e^{-j \frac{2\pi}{N} n k}, \ 0 \le k \le N-1,$$

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) \, e^{j \frac{2\pi}{N} n k}, \ 0 \le n \le N-1.$$

(30)

Although the transform is an iterated FB, the equivalent system in Fig. 15b leads to a faster implementation. Here, the parameters (p_i, q_i, r_i, s_i) are defined as follows. We set $q_1 = N$, $s_1 = N/2$ and define, for $i \ge 1$,

$$p_i = 2 \left\lfloor \frac{p^i}{2\,q^i} N \right\rfloor, \qquad q_{i+1} = p_i,$$

$$r_i = \left\lfloor \frac{r}{2s} \frac{p^{i-1}}{q^{i-1}} N \right\rfloor, \quad s_{i+1} = p_i/2.$$
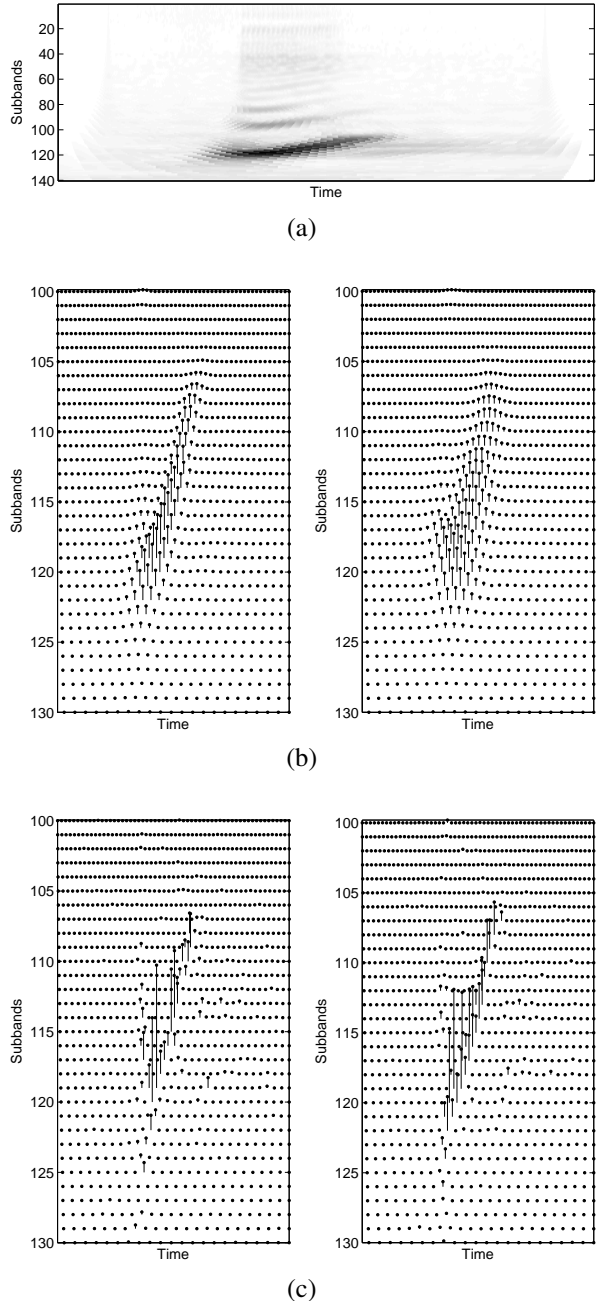
(31)

(a)

(b)

(c)

Fig. 14. A chirplet analysis of a speech signal : a male speaker asking 'Hmm?'. The effects of varying the chirp parameter γ is investigated in this example. The parameters common for all of the subfigures are $(p,q,r,s) = (31, 32, 1, 4)$, $\beta = r/s$. Recall that, a positive γ yields atoms with increasing pitch; a negative γ yields atoms with decreasing pitch. (a) Magnitudes of the analytic wavelet coefficients ($\gamma = 0$). Notice the increase in pitch towards the end of the utterance. In (b) and (c) we zoom into the dominant ridge (subbands 100–130). (b) Magnitudes of the analysis chirplet coefficients for $\gamma = 30$ (left panel) and $\gamma = -30$ (right panel). The coefficients for $\gamma = 30$ are more concentrated than those for $\gamma = -30$. (c) Magnitudes of the 'sparsified' synthesis coefficients for $\gamma = 30$ (left panel) and $\gamma = -30$ (right panel). Compared to $\gamma = -30$, the frame with $\gamma = 30$ requires fewer significant coefficients to represent the signal.
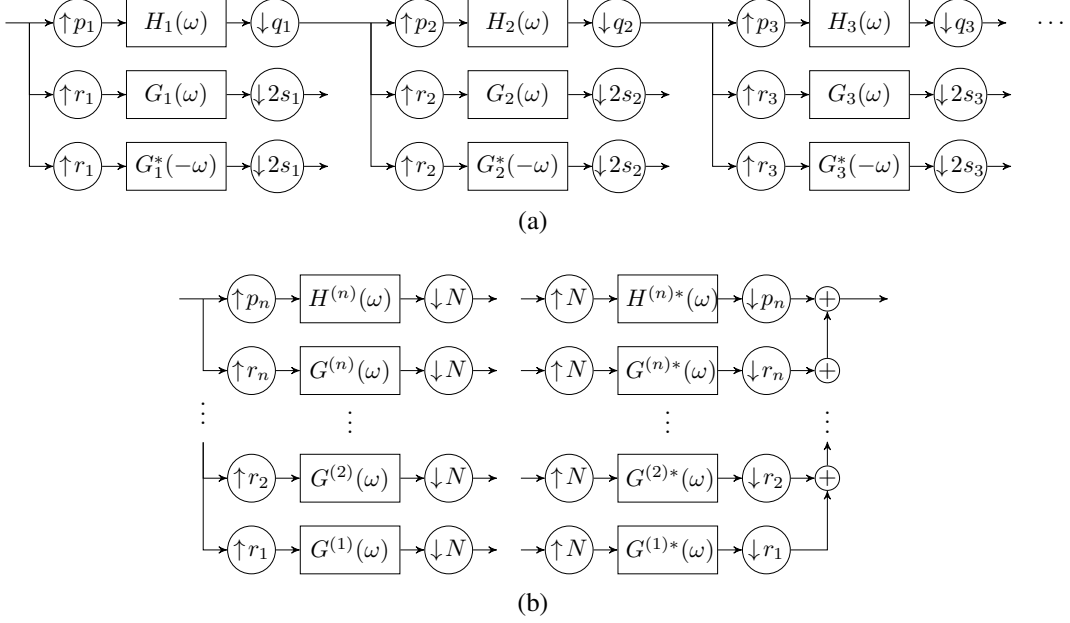
8

(a)

(b)

Fig. 15. For implementation, we approximate the original transform by the iterated FB shown in (a). This system is equivalent to the system in (b). See the text for the description of the sampling parameters and the filters. Regarding the system in (a) as a bank of rational rate changers as in (b), leads to a faster implementation. We note that in (b), only the 'positive -frequency' filters are shown to save space.

Also H_i, G_i, denote the filters for the FB with parameters $(p_i, q_i, r_i, s_i, \beta)$ (see Section II for the description) – we assume that β satisfies the constraints described in Prop. 2 for all i. Given these, we set $H^{(1)}(\omega) = H_1(\omega)$, $G^{(1)}(\omega) = G_1(\omega)$ and

$$
H^{(n)}(\omega) = \begin{cases} \frac{1}{p_{n-1}} H^{(n-1)}\left(\frac{p_n}{p_{n-1}}\omega\right) H_n\left(\frac{q_1}{p_{n-1}}\omega\right) \\ \qquad\qquad \text{for } |\omega| \leq \pi/N, \\ 0 \quad \text{for } \pi/N < |\omega| \leq \pi. \end{cases}
$$
(32)

$$
G^{(n)}(\omega)
$$
$$
= \begin{cases} 0, \quad 0 \leq \omega \leq \frac{p_{n-1}}{q_1}\left[(1-\beta)\pi/(p_{n-1}q_1) + \epsilon/r_n\right], \\ \frac{1}{p_{n-1}} H^{(n-1)}\left(\frac{r_n}{p_{n-1}}\omega\right) G\left(\frac{q_1}{p_{n-1}}\omega\right), \\ \qquad \frac{p_{n-1}}{q_1}\left[\frac{1-\beta}{p_n r_n}\pi + \frac{\epsilon}{r_n}\right] \leq \omega \leq \frac{p_{n-1}}{q_1 r_n}\pi, \\ 0, \quad \frac{p_{n-1}}{q_1 r_n}\pi \leq \omega \leq \pi/q_1. \end{cases}
$$
(33)

for $n > 1$.

Remark 1. Given the parameters (p_i, q_i, r_i, s_i) and the filters H_i, G_i, equivalence of the two systems in Fig. 15 can be shown by noting the frequency support of the filters and using the noble identities [35].

Note that the system in Fig. 15b consists of a bank of
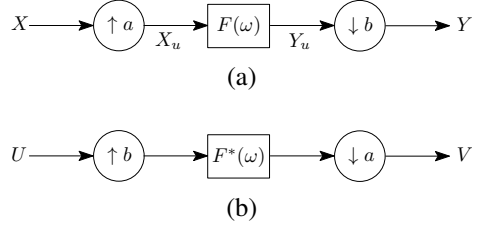

(a)

(b)

Fig. 16. (a) A rational rate changer, (b) The transpose of the system in (a).

rational rate changers. Below we describe an efficient implementation for rational rate changers.

Rational Rate Changers For Finite Length Signals

Consider the system in Fig. 16a. For simplicity, let us start with a specific example.

A Specific Example: For $a = 2$, $b = 3$, $N = 6$, consider the input DFT sequence shown in Fig. 17a. X_u, the upsampled signal's DFT, is obtained by repeating X twice (since $a = 2$). Then, $Y_u = X_u F$. To get Y from Y_u, we add the blocks Y_1, Y_2, Y_3 :

$$
Y(k) = \frac{1}{3}\sum_{n=0}^{3} Y_u(k + 4\,n) \quad \text{for } k = 0, 1, 2, 3. \tag{34}
$$

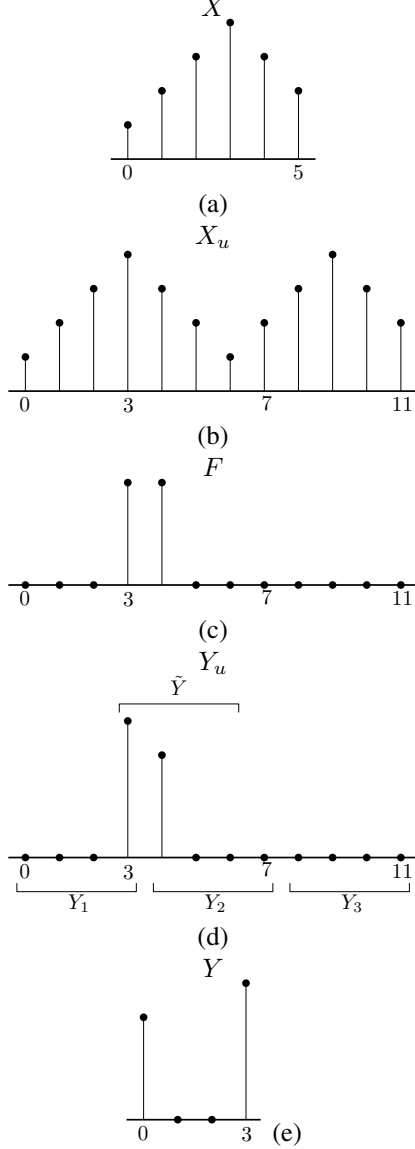Notice that the length of Y is given by $K := N\,a/b = 4$.

9

Fig. 17. The signals in Fig. 16a for a specific input X and filter F. The sampling parameters are taken as $a = 2$, $b = 3$ in Fig. 16a.

Although the description above is correct, it is not computationally efficient. Below is an alternative implementation that exploits the small support of F.

The support of F (which is $I = [3, 4]$), is contained in an interval of length K. This allows us to obtain Y by permuting the non-zero samples of Y_u. Specifically, let us define $\tilde{Y}$ as (see Fig. 17d)

$$\tilde{Y}(k) = \frac{1}{3} Y(3 + k) \quad \text{for } k = 0, 1, 2, 3. \tag{35}$$

Also set $s = \text{mod}(3, K)$. Then, Y is obtained by circularly shifting $\tilde{Y}$ right by s samples. Other than the factor

$1/3$, this implementation requires $|I| = 2$ multiplications, which is more efficient than the description above (which requires $N\,a = 12$ multiplications and $N\,(a/b)\,(b-1) = 8$ additions).

General Case: Let us now take $a = K$, $b = N$ in Fig. 16a. In this case, the length of Y is K. The filter $F(k)$ is a DFT-sequence of length KN, obtained by sampling $F(\omega)$ at $k\,2\pi/(KN)$ for $k = 0, \ldots, KN - 1$. Assume that $F(k)$ is supported on $f_0 < k \leq f_1$ where $f_1 - f_0 < K$. This condition ensures that there is no 'aliasing'. If we define a sequence of length K as,

$$\tilde{Y}(k) = \begin{cases} \frac{1}{N} X\big(\text{mod}(k + f_0 + 1, N)\big) F\big(f_0 + 1 + k\big), \\ \qquad\qquad \text{for } 0 \leq k < f_1 - f_0, \\ 0, \quad \text{for } f_1 - f_0 \leq k \leq K - 1, \end{cases} \tag{36}$$

then, Y is equal to the sequence obtained by circularly shifting $\tilde{Y}$ by $\text{mod}(f_0, K)$ samples to the right.

Remark 2. If $f_1 < K$, then

$$Y(k) = \begin{cases} 0, & \text{for } 0 \leq k < f_0, \\ \frac{1}{N} X(k)\,F(k), & \text{for } f_0 \leq k < f_1, \\ 0, & \text{for } f_1 \leq k \leq K - 1. \end{cases} \tag{37}$$

Notice that in this case, we are effectively applying the sequence $F(k)$ for $0 \leq k \leq K - 1$ on the input. $\qquad \square$

The transpose of this system is depicted in Fig. 16b. In this case, the output (which is of length N) can be obtained as follows. First define the length-N sequence

$$\tilde{U}(k) = \begin{cases} U\,(\text{mod}(k + f_0 + 1, K))\,F^*(k + f_0 + 1), \\ \qquad\qquad \text{for } 0 \leq k < f_1 - f_0 \\ 0, \quad \text{for } f_1 - f_0 \leq k \leq N - 1. \end{cases} \tag{38}$$

V is obtained by circularly shifting $\tilde{U}$ right by $\text{mod}(f_0, N)$ samples to the right.

Remark 3. If $f_1 < N$, then

$$V(k) = \begin{cases} 0, & \text{for } 0 \leq k \leq f_0, \\ \frac{1}{K} U\,(\text{mod}(k, K))\,F^*(k + f_0 + 1), \\ \qquad\qquad \text{for } f_0 \leq k < f_1, \\ 0, & \text{for } f_1 \leq k \leq N - 1 \end{cases} \tag{39}$$

VI. EXPERIMENTS AND APPLICATIONS

Matlab code that implements the transform and the experiments below is available at `web.itu.edu.tr/ibayram/AnDWT`.

Experiment 1. It is well known that in practice, one can reconstruct (with little perceptual difference) an audio signal from the magnitudes of its STFT coefficients. Such a reconstruction can be achieved by iteratively modifying the coefficients of some initial signal (like noise) [23]. One algorithm to achieve this is :

> **Algorithm 1.** Let $|Y|$ be the given STFT magnitudes. Let x be an initial signal, and let S be the STFT operator. Repeat until convergence,
> (i) Modify $X = S(x)$ ◁ Compute the STFT of x
> (ii) Modify $Z = |Y| \exp(i \angle X)$ ◁ Force the STFT magnitude to be $|Y|$
> (iii) Modify $x = S^{-1}(Z)$ ◁ Modify the signal using the new STFT, Z

It turns out that a similar phenomenon occurs for the wavelet transform introduced in this paper. That is, we can reconstruct a signal from the magnitudes of its analytic wavelet coefficients. For this, we replace S and S^{-1} by the wavelet analysis and synthesis operators respectively. In particular, for a low dilation factor (for $p = 69$, $q = 70$; $69/70 \approx (1/2)^{1/48}$), and high enough Q-factor ($r = 1$, $s = 10$; Q = 19) we achieved a high-quality reconstruction.

Being able to reconstruct from wavelet coefficient magnitudes leads to simple algorithms for audio processing. The actual subband coefficients are oscillatory wheras the subband magnitudes are slowly-changing (or smooth). In turn, they are easier to interpret and they can be modified more easily to achieve desired effects. For instance, if we time-scale (say, by α) the magnitude of each subband using bicubic interpolation, the reconstructed signal is a time-scaled (by α) version of the original signal. In a similar fashion, for a dilation factor of d, if we time-scale each subband magnitude of an instrumental piece by d^K and shift the modified subband magnitudes by K subbands, the reconstruction is a 'transposed signal'[2]. This is demonstrated in Fig. 18. Here, the dilation factor is $d = 69/70 \approx (1/2)^{1/48}$ and the input piece is raised by two semitones (which corresponds to shifting up by 8 subbands for the given dilation factor). We also refer to [19] for an example where the interpolation of the coefficients can be avoided, at the expense of increased redundancy.

[2]Transposing music corresponds to moving a piece up or down in pitch by some amount while maintaining the relative tone structure. For a more detailed discussion from a DSP perspective along with algorithms, see [37].
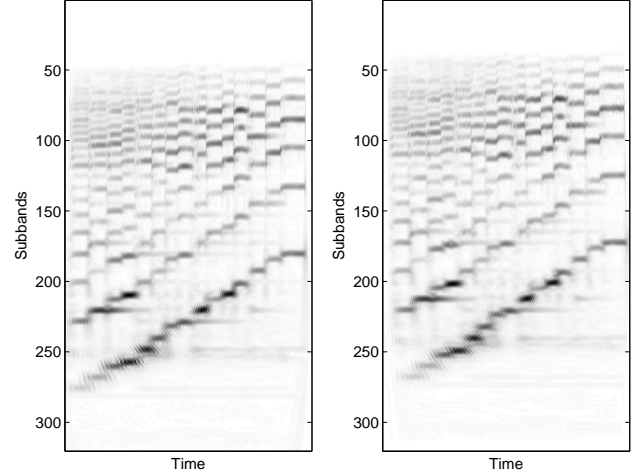

Fig. 18. Music transposition experiment. For a rational dilation factor of (approximately) $(1/2)^{1/48}$ the magnitudes of the analytic wavelet coefficients are shifted by 8 subbands (without changing the time-duration – this is achieved by bicubic interpolation on the magnitudes). The reconstruction is a transposed piece, raised by 2 semitones. Left panel : Subbands of the original signal. Right Panel : Subbands of the transposed signal.

This experiment/application makes use of the availability of quadrature pairs of wavelets in the introduced transform. We note therefore that it is not possible to adapt this scheme to previously introduced transforms like RADWT or TQWT, since these transforms do not contain quadrature pairs of wavelets. □

Experiment 2. The flexibility of the transform can also be exploited in order to facilitate processing. In this experiment, we 'transpose' the input piece in the previous experiment using a different approach.

Consider Fig. 19a. Here, the black dots mark the time-frequency centers of the atoms from a wavelet frame. Suppose we shift the wavelet coefficients of a music piece up by one subband, following the arrows. It can be seen from the figure that in addition to a modification along the frequency axis, this also leads to a scaling in time (we end up with a signal lengthened by a factor of d^{-1} where $d = p/q$ is the dilation factor). This is due to the fact that the sampling period is different in each subband. If the sampling period were the same in each subband (like in STFT), such a problem would not occur. However, in that case, a constant-Q transform would have a very high redundancy. The flexibility of the introduced transform allows us to get around this problem. We compute the coefficients of the signal for a wavelet transform with parameters p, q, r, s, β and then move the coefficients up by one subband. For reconstruction, we employ a wavelet transform with the same dilation (i.e. p, q) and Q-factor (i.e. β) but we
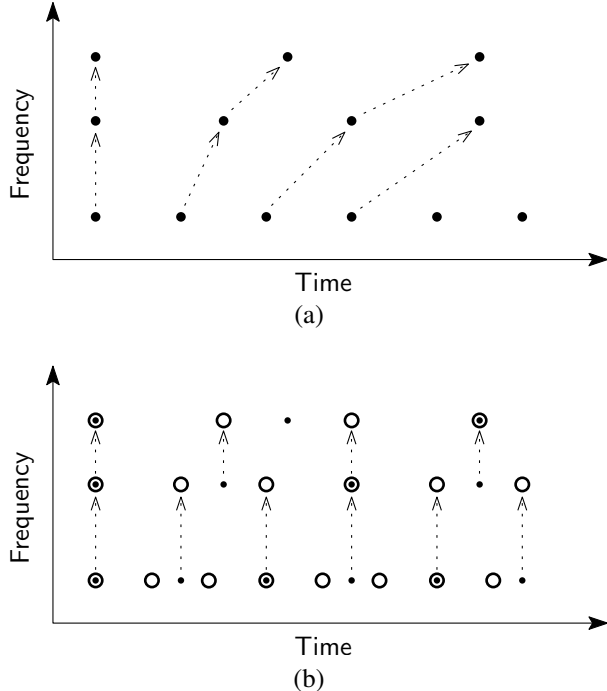
11

Fig. 19. (a) Black dots indicate the time-frequency centers of the atoms from a wavelet frame. Assume that the dilation factor is p/q, Q-factor is Q and highpass sampling factor is r/s. When we shift the coefficients up by one subband, we get a transposed *and* time-scaled signal. (b) In order to eliminate time-scaling, we employ a different transform for reconstruction. Here, the coefficients from the first frame (shifted up by one subband) are synthesized using a wavelet transform with the same dilation and Q-factor but with a modified highpass sampling factor of $(p\,r)/(q\,s)$. The time-frequency centers of the atoms from this frame are shown using circles.

modify the highpass sampling parameters as $\tilde{r} = r\,p$, $\tilde{s} = s\,q$. The time-frequency centers of the atoms for this new transform are shown in Fig. 19b (hollow circles). The operation described above corresponds to translating the time-frequency samples marked with black dots in Fig. 19b up by one subband, placing them into the time-frequency points marked with hollow circles (pointed to by the arrows). In turn, there is no change in the duration of the signal but the frequency content is modified as desired. This method can easily be generalized to transpose the signal by more than one subband as well.

The method desribed above makes use of the flexibility of the introduced transform. We note that it is not possible to employ RADWT or TQWT in such a scheme because those transforms do not allow simultaneous control over the dilation factor, the Q-factor and the redundancy. □

Experiment 3. In this experiment, we apply the chirplet transform on a bird call signal. The signal is shown in Fig. 20a. Notice that there are three components with increasing pitch and a last component with decreasing

pitch (towards the end of the last component, the pitch is approximately constant). We aim to separate the increasing and decreasing pitch components. For this, we employ two chirplet frames, with chirp parameters $\gamma_1 = 200$ and $\gamma_2 = -200$. Other than the chirp parameters, the two frames share the same sampling factors $p = 5$, $q = 6$, $r = 1$, $s = 4$; that leads to Q = 7, redundancy = 3/2. Suppose that the synthesis operators for these two frames are denoted as F_1 and F_2. Given the input bird call signal x, we consider the minimization problem :

$$\min_{\theta_1, \theta_2} \quad \left\| x - (F_1\,\theta + F_2\,\theta_2) \right\|_2^2 + \lambda_1 \left\| \theta_1 \right\|_1 + \lambda_2 \left\| \theta_2 \right\|_1 \quad (40)$$

If we denote the minimizers as θ_1^* and θ_2^*, we take the separated components as $x_1^* = F_1\,\theta_1^*$, $x_2^* = F_2\,\theta_2^*$ – we refer the reader to [20] for a more detailed discussion on the variants of this approach.

The signals x_1^* and x_2^* obtained this way are shown in Fig. 20b and Fig. 20c respectively. We see that the components with different time-frequency orientations are successfully separated. This experiment supports the discussion at the end of Section IV (Fig. 14) : the chirp parameter can be tuned to match the signal characteristics. □

VII. CONCLUSION

We developed a wavelet transform that possesses a number of desirable properties like good frequency resolution and easy control over parameters like the dilation factor, redundancy and Q-factor. In addition, the frame contains quadrature pairs of atoms, which makes it suitable for oscillatory signal processing. As a consequence of the particular construction, it is also easy to form a tight chirplet frame for discrete-time signals, which could be useful in applications where time-frequency orientation is of interest. Although the transform is obtained by iterated filter banks, an equivalent filter bank leads to a fast realization using FFTs. We tested the transform on real signals and demonstrated how it can be used for applications like time-scaling, music transposition and component separation.

REFERENCES

[1] L. Angrisani and M. D'Arco. A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation. *IEEE Trans. Instrumentation and Measurement*, 51(4):704–711, August 2002.
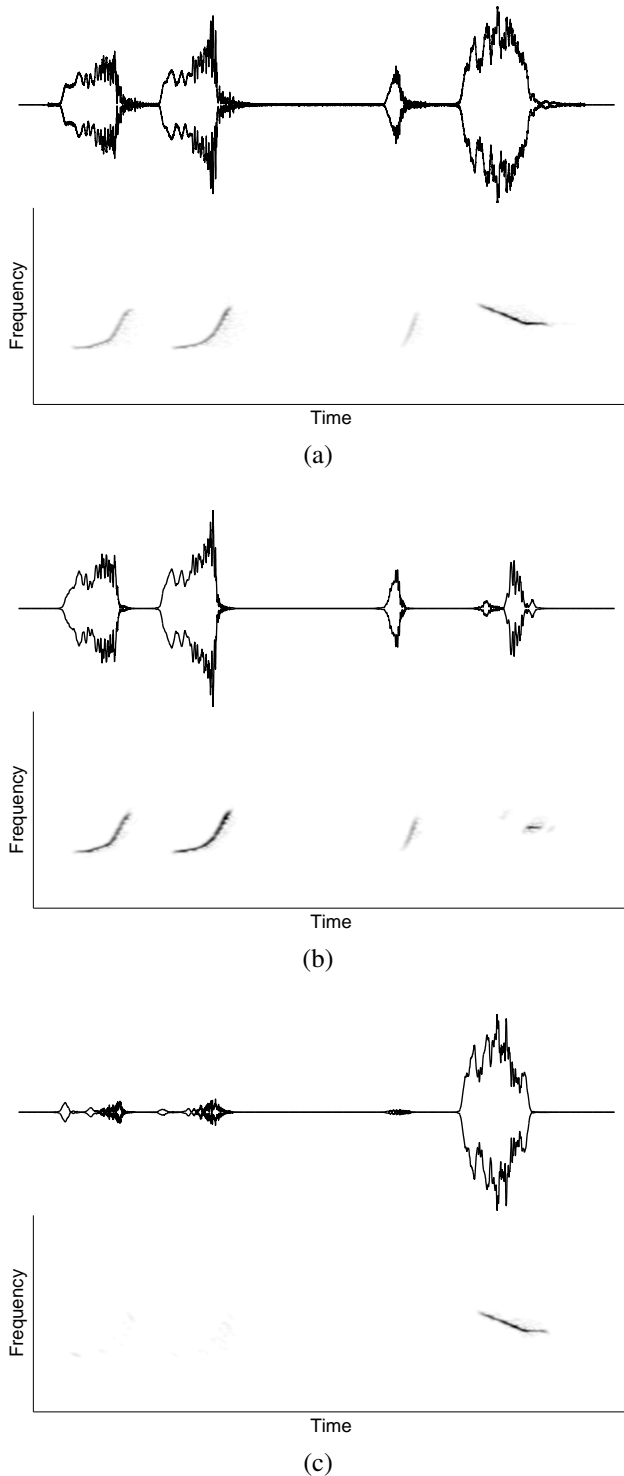
(a)

(b)

(c)

Fig. 20. (a) The original bird call signal (the envelope is shown). Notice that the signal consists of three components with increasing pitch and a last component with initially decreasing, then constant pitch. We separate these components making use of the chirplet transforms in Experiment 3. (b) The components with increasing pitch (corresponding to a positive chirp parameter). (c) The components with decreasing pitch (corresponding to a positive chirp parameter).

[2] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. Velasco. Theory, implementation and applications of nonstationary Gabor frames. *J. of Appl. and Comp. Harm. Analysis*, 236(6):1481–1496, October 2011.

[3] R. G. Baraniuk and D. L. Jones. Shear madness : New orthonormal bases and frames using chirp functions. *IEEE Trans. Signal Processing*, 41(12):3543–3549, December 1993.

[4] R. G. Baraniuk and D. L. Jones. Wigner-based formulation of the chirplet transform. *IEEE Trans. Signal Processing*, 44(12):3129–3135, December 1996.

[5] A. Baussard, F. Nicolier, and F. Truchetet. Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising. *Signal Processing*, 84(10):1735–1747, October 2004.

[6] İ. Bayram and I. Selesnick. Design of orthonormal and overcomplete wavelet transforms based on rational sampling factors. In *Proc. Fifth SPIE Conference on Wavelet Applications in Industrial Processing*, 2007.

[7] İ. Bayram and I. W. Selesnick. Frequency-domain design of overcomplete rational dilation wavelet transforms. *IEEE Trans. Signal Processing*, 57(8):2957–2972, August 2009.

[8] İ. Bayram and I. W. Selesnick. Overcomplete discrete wavelet transforms with rational dilation factors. *IEEE Trans. Signal Processing*, 57(1):131–145, January 2009.

[9] İ. Bayram and I. W. Selesnick. A dual-tree rational dilation complex wavelet transform. *IEEE Trans. Signal Processing*, 59(12):6251–6256, December 2011.

[10] T. Blu. Iterated filter banks with rational rate changes–connection with discrete wavelet transforms. *IEEE Trans. Signal Processing*, 41(12):3232–3244, December 1993.

[11] T. Blu. An iterated rational filter bank for audio coding. In *Proc. IEEE Int. Symposium on Time-frequency and Time-scale Analysis*, 1996.

[12] T. Blu. A new design algorithm for two-band orthonormal rational filter banks and orthonormal rational wavelets. *IEEE Trans. Signal Processing*, 46(6):1494–1504, June 1998.

[13] J. C. Brown. Calculation of a constant Q spectral transform. *J. Acoust. Soc. Amer.*, 89(1):425–434, January 1991.

[14] J. C. Brown and M. S. Puckette. An efficient algorithm for the calculation of a constant Q transform. *J. Acoust. Soc. Amer.*, 92(5):2698–2701, November 1992.

[15] A. Bultan. A four-parameter atomic decomposition of chirplets. *IEEE Trans. Signal Processing*, 47(3):731–745, March 1999.

[16] T. Chi, P. Ru, and S. A. Shamma. Multiresolution spectrotemporal analysis of complex sounds. *J. Acoust. Soc. Amer.*, 118(2):887–906, August 2005.

[17] L. Cohen. *Time-Frequency Analysis*. Prentice Hall, 1995.

[18] J. Cui and W. Wongl. The adaptive chirplet transform and visual evoked potentials. *IEEE Trans. Biomedical Engineering*, 53(7):1378–1384, July 2006.

[19] M. Dörfler, N. HoligHaus, T. Grill, and G. Velasco. Constructing an invertible constant-Q transform with nonstationary Gabor frames. In *Proc. Int. Conf. on Digital Audio Effects (DAFx)*, 2011.

[20] M. J. Fadili, J.-L. Starck, J. Bobin, and Y. Moudden. Image decomposition and separation using sparse representations : An overview. *Proc. IEEE*, 98(6):983 – 994, 2010.

[21] G. Gambardella. A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters. *IEEE Trans. Circuit Theory*, 18(4):455–460, July 1971.

[22] R. Gribonval. Fast matching pursuit with a multiscale dictionary of gaussian chirps. *IEEE Trans. Signal Processing*, 49(5):994–1001, May 2001.

[23] D. W. Griffin and J. S. Lim. Signal estimation from modified short-time fourier transform. *IEEE Trans. Acoust., Speech, and Signal Proc.*, 32(2):236–243, April 1984.

[24] O. Kalinli and S. Narayanan. Prominence detection using auditory attention cues and task-dependent high level information. *IEEE Trans. Audio, Speech and Language Processing*, 17(5):1009–1024, July 2009.

[25] J. Kovačević and M. Vetterli. Perfect reconstruction filter banks with rational sampling factors. *IEEE Trans. Signal Processing*, 41(6):2047–2066, June 1993.

[26] M. Kowalski. Sparse regression using mixed norms. *J. of Appl. and Comp. Harm. Analysis*, 27(3):303–324, November 2009.

[27] S. Mann and S. Haykin. Chirplet transform : Physical considerations. *IEEE Trans. Signal Processing*, 43(11):2745–2761, November 1995.

[28] K. Nayebi, T. P. Barnwell III, and M. J. T. Smith. The design of perfect reconstruction nonuniform band filter banks. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 1991.

[29] H. Olkkonen, J. T. Olkkonen, and P. Pesola. FFT-Based computation of shift invariant analytic wavelet transform. *IEEE Signal Processing Letters*, 14(3):177–180, 2007.

[30] T. L. Petersen and S. F. Boll. Critical band analysis-synthesis. *IEEE Trans. Acoust., Speech, and Signal Proc.*, 31(3):656–663, June 1983.

[31] T.H. Reeves and N. G. Kingsbury. Overcomplete image coding using iterative projection-based noise shaping. In *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, 2002.

[32] C. Schörkhuber and A. Klapuri. Constant-Q toolbox for music processing. In *Proc. SMC Conference*, 2010.

[33] I. W. Selesnick. Wavelet transform with tunable Q-factor. *IEEE Trans. Signal Processing*, 59(8):3560–3575, August 2009.

[34] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury. The dual-tree complex wavelet transform - A coherent framework for multiscale signal and image processing. *IEEE Signal Processing Magazine*, 22(6):123–151, November 2005.

[35] P.P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, 1992.

[36] J. E. Youngberg and S. F. Boll. Constant-Q signal analysis and synthesis. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 1978.

[37] U. Zölzer, editor. *DAFX: Digital Audio Effects*. Wiley, 2011.