The Redundant Discrete Wavelet Transform and Additive Noise

James E. Fowler

Technical Report MSSU-COE-ERC-04-04

March 2004

GeoResources Institute—Mississippi State ERC
Mississippi State University

Abstract

The behavior under additive noise of the redundant discrete wavelet transform, a frame expansion that is essentially an undecimated discrete wavelet transform, is studied. Known prior results in the form of inequalities bound distortion energy in the original signal domain from additive noise in frame-expansion coefficients. In this paper, a more precise relationship between RDWT-domain and original-signal-domain distortion is developed. Specifically, it is determined that the contribution to distortion in the original signal domain from white noise in a single RDWT subband depends only on the decomposition scale at which the subband resides. Furthermore, the total noise distortion due to all subbands is found to be an equality rather than a bounding relationship involving frame bounds, which are shown, in fact, to widen as the number of decomposition scales increases.

1. Introduction

Linear transforms and expansions have a long history of serving as fundamental tools in the field of signal processing. Very often, it is necessary in signal-processing applications to be able to calculate distortion energy in the original signal domain from an equivalent quantity in the transform domain. That way, signal-processing operations can be performed in the transform domain with known effects in the original signal domain. For this reason, orthonormal sets are widely used, since, when the transform takes the form of an expansion using an orthonormal basis, Parseval's theorem guarantees

$$\|x\|^2 = \|X\|^2, \tag{1}$$

or that the energy of the original signal x can be determined from that of its transform X.

However, the constraints of orthonormal expansion sets can sometimes be too restrictive for some signal-processing applications. When one widens consideration to more general expansions, the increased functionality and flexibility unfortunately often comes at the cost of an exact energy relationship as above. Instead, one often has merely a bounding relationship in the form of

$$A\|x\|^2 \leq \|X\|^2 \leq B\|x\|^2 \tag{2}$$

that *frames* the energy in one domain with respect to that of the other domain for some constants $A > 0$ and $B < \infty$. Expansions with such energy bounds are hence known as *frame* expansions.

One of the more useful types of expansions which are more general than orthonormal sets are overcomplete, and thereby redundant, transforms. With redundancy, greater functionality often becomes possible. For instance, redundant transforms provide greater robustness to added noise and quantization [1–3] as well as increased numerical stability [1]. In addition, the redundancy can produce shift invariance, which can facilitate, among other tasks, feature detection [4, 5] and motion estimation [6–8].

In this paper, we focus on a specific redundant frame expansion known as the redundant discrete wavelet transform (RDWT), which is essentially an undecimated version of the discrete wavelet transform (DWT) ubiquitous to modern signal-processing applications. Since it is a frame expansion, the RDWT has energy bounds as in (2). As the initial contribution of this paper, we determine values for the frame-bound constants A and B assuming that an orthonormal filter pair underlies the RDWT. Then, as the primary contribution of this paper, we analyze the performance of the RDWT under additive noise. Our intuition concerning frames leads us to expect that an energy-bounding relationship would exist to describe the effect of additive noise. However, the somewhat unexpected result of our noise analysis is an *exact* relationship between added noise in the RDWT domain and its corresponding distortion energy in the original signal domain. That is to say, even though the RDWT is a highly redundant frame expansion, we can determine exactly the variance (i.e., expected distortion energy per sample) in the original signal domain of noise added in the RDWT domain. The most important aspect of our results is a per-subband noise relationship; that is, the contribution to the distortion in the original signal domain from noise in a single RDWT subband is found to depend only on the decomposition scale at which the subband resides and to be independent of that of other subbands.

The remainder of this paper is organized as follows. We present a number of preliminary mathematical concepts in Sec. 2 in which we briefly overview frame theory and known results concerning resilience of frames to added noise. Next, in Sec. 3, we describe the RDWT in detail. The main contributions of the paper are presented in Sec. 4—specifically, we derive frame bounds for the RDWT in Sec. 4.1 and analyze noise performance of the RDWT in Sec. 4.2. We present some considerations of a largely pragmatic nature in Sec. 5 in which we discuss the separable 2D RDWT and RDWTs based on biorthogonal filters, our main results having been derived for 1D RDWTs using orthonormal filters. Finally, we make some concluding remarks in Sec. 6.

2. Background

2.1 Mathematical Preliminaries and Notation

Consider a Hilbert space $\mathcal{H}$. Throughout this paper, we will focus our attention on discrete spaces, either $\mathcal{H} = \ell^2(\mathbb{Z})$, the space of square-summable sequences, or $\mathcal{H} = \mathbb{C}^N$, the N-dimensional complex space. We define the inner product in $\mathcal{H}$ to be $\langle x, y\rangle = \sum_k x[k]y^*[k]$, and the vector norm to be $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_k |x[k]|^2}$, where $x, y \in \mathcal{H}$, and α^* denotes the complex conjugate of $\alpha \in \mathbb{C}$.

For $x \in \ell^2(\mathbb{Z})$, the discrete-time Fourier transform (DTFT), $\widehat{x}(\omega)$ is defined as

$$\widehat{x}(\omega) = \sum_{k=-\infty}^{\infty} x[k]e^{-i\omega k}, \tag{3}$$

where $i = \sqrt{-1}$. Parseval's theorem for the DTFT is

$$\sum_{k=-\infty}^{\infty} \left|x[k]\right|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\widehat{x}(\omega)\right|^2 d\omega, \tag{4}$$

or, in terms of norms,

$$\|x\|^2 = \|\widehat{x}\|^2, \tag{5}$$

where the norm in the frequency domain, $\|\widehat{x}\|^2$, is defined as the right side of (4).

2.2 Frames

The following overview of frame theory summarizes discussion from [1, 2, 9]. A family of vectors $\Psi = \{\psi_n\}_{n\in\mathcal{I}} \subset \mathcal{H}$ is called a *frame* if there exist $A > 0$ and $B < \infty$ such that

$$A\|x\|^2 \leq \|X\|^2 \leq B\|x\|^2, \quad \forall x \in \mathcal{H}, \tag{6}$$

where $\mathcal{I}$ is a countable index set, $X \in \ell^2(\mathcal{I})$ such that $X[n] = \langle x, \psi_n \rangle$, and

$$\ell^2(\mathcal{I}) = \left\{ X \in \mathbb{C}^{|\mathcal{I}|} : \left\| X \right\|^2 < \infty \right\}.$$

Unless otherwise noted, we subsequently assume that frame Ψ is *uniform*, i.e., that $\|\psi_n\| = 1, \forall n \in \mathcal{I}$ [10]. A and B in (6) are called the *frame bounds*.

We define the *frame operator*, F, associated with frame Ψ as the mapping from $\mathcal{H}$ to $\ell^2(\mathcal{I})$,

$$Fx = X = \left\{ \langle x, \psi_n \rangle \right\}_{n \in \mathcal{I}}. \tag{7}$$

The adjoint, F^*, of F is the mapping from $\ell^2(\mathcal{I})$ to $\mathcal{H}$ defined as

$$F^* X = \sum_{n \in \mathcal{I}} X[n] \psi_n \tag{8}$$

for $X \in \ell^2(\mathcal{I})$. In the case that $\mathcal{H} = \mathbb{C}^N$, then $\ell^2(\mathcal{I}) = \mathbb{C}^M$ for some $M \geq N$. In this case, we define the *redundancy* of a frame as the ratio $r = M/N$.

The *dual frame*, $\widetilde{\Psi}$, of Ψ is defined as $\widetilde{\Psi} = \left\{ \widetilde{\psi}_n \right\}_{n \in \mathcal{I}}$ where

$$\widetilde{\psi}_n = \left(F^* F \right)^{-1} \psi_n, \quad \forall n \in \mathcal{I}. \tag{9}$$

It can be shown [1, 9] that the dual frame is itself a frame with bounds B^{-1} and A^{-1}; that is,

$$\frac{1}{B} \left\| x \right\|^2 \leq \sum_{n \in \mathcal{I}} |\langle x, \widetilde{\psi}_n \rangle|^2 \leq \frac{1}{A} \left\| x \right\|^2, \quad \forall x \in \mathcal{H}. \tag{10}$$

The frame operator and its adjoint associated with dual frame $\widetilde{\Psi}$ are $\widetilde{F}$ and $\widetilde{F}^*$, respectively. We note that, even if the frame Ψ is uniform, the dual frame $\widetilde{\Psi}$ may not be. In fact, it can be shown that

$$\frac{1}{B^2} \leq \left\| \widetilde{\psi}_n \right\|^2 \leq \frac{1}{A^2}, \quad \forall n \in \mathcal{I} \tag{11}$$

(e.g., see App. I-B of [2]).

If the two frame bounds are equal, $A = B$, the frame is called a *tight frame*. In a tight frame,

$$\left\| X \right\|^2 = A \left\| x \right\|^2. \tag{12}$$

Additionally, the vectors of the dual frame are

$$\widetilde{\psi}_n = \frac{1}{A} \psi_n. \tag{13}$$

For a tight frame in $\mathcal{H} = \mathbb{C}^N$, we have that the redundancy is given by the frame bound; i.e., $r = A$. A tight frame becomes orthonormal when $A = 1$, in which case, there is no redundancy ($r = 1$).

2.3 Frame Expansion, Reconstruction, and Additive Noise

Given a frame Ψ and its dual $\widetilde{\Psi}$, any $x \in \mathcal{H}$ can be expanded as

$$x = \widetilde{F}^* X = \sum_{n \in \mathcal{I}} X[n] \widetilde{\psi}_n, \tag{14}$$

where the expansion coefficients are $X[n] = (Fx)_n = \langle x, \psi_n \rangle$. In the case of a tight frame, we have

$$x = \frac{1}{A} \sum_{n \in \mathcal{I}} \langle x, \psi_n \rangle \psi_n. \tag{15}$$

We note that, given a set of coefficients $X \in \ell^2(\mathcal{I})$, the process of reconstructing $x \in \mathcal{H}$ is not unique; in fact, frame operator F has an infinite number of left inverses [9]. However, the reconstruction given by (14), the so-called *dual-frame* or *pseudo-inverse* [9] reconstruction, is special in that it provides the optimal least-squares reconstruction in the event of corruption of the expansion coefficients with noise [11]. That is, if $Y = Fy$ and $Y' = Y + X$, where X is a noise signal, then $\|Fy' - Y'\|^2$ is minimized if we reconstruct y as $y' = \widetilde{F}^* Y' = \widetilde{F}^* Y + \widetilde{F}^* X$ as specified in (14) [11].

One of the key benefits of the redundancy offered by a frame expansion lies precisely in the robustness of the dual-frame reconstruction to added noise. This property has been studied extensively in the past, e.g., [1–3]. The following theorem, due to Goyal *et al*. [2], embodies the resilience of frame expansions to added noise assuming a finite-dimensional space $\mathcal{H} = \mathbb{C}^N$.

Theorem 1 *Given a uniform frame $\Psi \subset \mathcal{H} = \mathbb{C}^N$, let $X \in \mathbb{C}^M$ be zero-mean white noise such that*

$$E\Big[X[n]X^*[m]\Big] = \begin{cases} 0, & n \neq m, \\ \sigma^2, & n = m, \end{cases} \tag{16}$$

for all $n, m \in \mathcal{I}$. Then

$$\frac{1}{B^2} E\left[\|X\|^2\right] \leq E\left[\|x\|^2\right] \leq \frac{1}{A^2} E\left[\|X\|^2\right], \tag{17}$$

where $x = \widetilde{F}^* X \in \mathbb{C}^N$ is the dual-frame reconstruction of (14).

Proof: See Proposition 1 of [2]. ■

Corollary 1 *In the case of a tight frame, Theorem 1 becomes*

$$E\left[\|x\|^2\right] = \frac{1}{A^2} E\left[\|X\|^2\right]. \tag{18}$$

Remark: If we observe that, since $X \in \mathbb{C}^M$,

$$E\left[\|X\|^2\right] = M\sigma^2, \tag{19}$$

then (17) becomes

$$\frac{M\sigma^2}{B^2} \leq E\left[\|x\|^2\right] \leq \frac{M\sigma^2}{A^2}. \tag{20}$$

In the case of a tight frame, we then have

$$E\left[\|x\|^2\right] = \frac{M\sigma^2}{A^2} = \frac{N\sigma^2}{r}. \tag{21}$$

From this last equation, we see how frames provide robustness to added noise—the greater the redundancy r of the frame, the less noise energy $E\left[\|x\|^2\right]$ will result in the original signal domain from added noise in the domain of the frame expansion. If the frame is not tight, then the amount of noise reduction is not known exactly but is bounded as indicated in (20). However, we are guaranteed that some noise reduction occurs due to redundancy even for non-tight frames, since $A > 1$ for redundant frames [9].

2.4 The Discrete Wavelet Transform and the Discrete-Time Wavelet Series

In this section, we present a few final background ideas—namely, mathematical results concerning the DWT and the related discrete-time wavelet series (DTWS)—before turning our attention to the RDWT in the next section. Suppose $h, g \in \ell^2(\mathbb{Z})$ are the scaling and wavelet filters, respectively, of an orthonormal DWT. The following are well-known properties of h and g (see, e.g., Chap. 5 of [12]):

$$\sum_k h[k] = \sqrt{2}, \tag{22}$$

$$\sum_k g[k] = 0, \tag{23}$$

$$\sum_k h[2k] = \sum_k h[2k+1], \tag{24}$$

$$g[k] = \pm(-1)^k h[K-k], \tag{25}$$

for some odd-valued integer K. Additionally, perfect reconstruction dictates that the filters are *power complementary*[1] such that

$$\left|\widehat{h}\left(2^\lambda \omega\right)\right|^2 + \left|\widehat{g}\left(2^\lambda \omega\right)\right|^2 = 2, \quad \forall \lambda \in \mathbb{Z}, \, \forall \omega \in \mathbb{R}. \tag{26}$$

It is well-known (e.g., [9, 14]) that the scaling and wavelet filters h and g from an orthonormal DWT—which is an expansion system for $L^2(\mathbb{R})$—can also be employed to implement a DTWS, an expansion system for $\ell^2(\mathbb{Z})$. In this case, an iterated, two-channel filter bank expands $x \in \ell^2(\mathbb{Z})$ using the orthonormal basis sequences,

$$\left\{\left\{\phi_J[k-2^J n]\right\}_{n\in\mathbb{Z}}, \left\{\psi_j[k-2^j n]\right\}_{1\le j\le J,\, n\in\mathbb{Z}}\right\}, \tag{27}$$

where J is depth of the filter-bank tree. The basis sequences are defined in the frequency domain as [9, 14]

$$\widehat{\phi}_j(\omega) = \prod_{\lambda=0}^{j-1} \widehat{h}(2^\lambda \omega), \tag{28}$$

$$\widehat{\psi}_j(\omega) = \widehat{g}(2^{j-1}\omega) \prod_{\lambda=0}^{j-2} \widehat{h}(2^\lambda \omega). \tag{29}$$

Since the DTWS basis is orthonormal if h and g come from an orthonormal DWT, we have

$$\left\|\phi_j\right\|^2 = \left\|\widehat{\phi}_j\right\|^2 = 1, \tag{30}$$

$$\left\|\psi_j\right\|^2 = \left\|\widehat{\psi}_j\right\|^2 = 1. \tag{31}$$

In the next section, we briefly overview the RDWT before arriving at the main contribution of this paper, a noise analysis of the RDWT, which follows in Sec. 4.

3. The Redundant Discrete Wavelet Transform

The RDWT was originally developed [15, 16] as a discrete approximation to the continuous wavelet transform. Subsequent formulations [17, 18] realized that removal of the downsampling operation from the traditional critically sampled DWT would produce an overcomplete representation with shift invariance, since the well known shift variance of the DWT arises from its use of downsampling, while the RDWT is

[1]The usual definition [13] of power complementary is for $\lambda = 0$, but this more general definition clearly is valid for any $\lambda \in \mathbb{Z}$.

shift invariant since the spatial sampling rate is fixed across scale. As a result, the size of each subband in an RDWT is the exactly the same as that of the input signal, facilitating a number of applications such as feature detection [4, 5], signal denoising [19, 20], motion estimation [6–8], and watermarking [21].

The RDWT[2] has a long history, having been independently discovered a number of times and given a number of different names, including the *algorithme à trous* [15, 16], the undecimated DWT (UDWT) [20], the overcomplete DWT (ODWT) [7], the shift-invariant DWT (SIDWT) [22], and discrete wavelet frames (DWF) [5]. There are several ways to implement the RDWT, and several ways to represent the resulting overcomplete set of coefficients. The original implementation was in form of the *algorithme à trous* [15, 16], which, in essence, removes the downsampling operator from the ubiquitous Mallat implementation [23] of the DWT. In this implementation, instead of signal downsampling, the filter responses themselves are upsampled, thereby inserting "holes" (*trous* in French) between nonzero filter taps. Mallat [24] independently proposed the *à trous* implementation, while the *à trous* implementation also arose in the form of iterated, nonsubsampled filter banks [14, 25]. Below, we present an overview of the mathematical details of the *à trous* implementations of the RDWT; for greater detail, refer to [9, 17].

An RDWT is defined in terms of an underlying DWT. Let $h \in \ell^2(\mathbb{Z})$ and $g \in \ell^2(\mathbb{Z})$ be the scaling and wavelet filters, respectively, of an orthonormal DWT. In essence, the *à trous* implementation of the RDWT removes the decimation operations from Mallat's algorithm [23] for the critically sampled DWT. To retain the proper multiresolution characteristic of the transform, the scaling and wavelet filters must be adjusted according to each scale. Specifically, define the upsampling operator as

$$x[k] \uparrow 2 = \begin{cases} x[k/2], & k \text{ even}, \\ 0, & k \text{ odd}, \end{cases} \tag{32}$$

and the scaling and wavelet filters at scale $j+1$ as

$$h_{j+1}[k] = h_j[k] \uparrow 2, \tag{33}$$

$$g_{j+1}[k] = g_j[k] \uparrow 2, \tag{34}$$

where

$$h_0[k] = h[k], \tag{35}$$

$$g_0[k] = g[k]. \tag{36}$$

In the frequency domain, the filters are then

$$\widehat{h}_j(\omega) = \widehat{h}_0\left(2^j\omega\right) \tag{37}$$

$$\widehat{g}_j(\omega) = \widehat{g}_0\left(2^j\omega\right). \tag{38}$$

The RDWT of $x \in \mathcal{H}$, where $\mathcal{H}$ is $\ell^2(\mathbb{Z})$ or $\mathbb{C}^N$, is implemented recursively with the filter-bank operations

$$c_{j+1}[k] = h_j[-k] * c_j[k], \tag{39}$$

$$d_{j+1}[k] = g_j[-k] * c_j[k], \tag{40}$$

where $c_0 = x$ and $j = 0, \ldots, J-1$. In the case of a finite-dimensional signal $x \in \mathcal{H} = \mathbb{C}^N$, the $*$ denotes circular convolution; otherwise, the $*$ is discrete-time convolution. The J-scale RDWT is the collection of subbands resulting from the recursive filtering operations, i.e.,

$$X^{(J)} = \mathrm{RDWT}_J\left[x\right] \tag{41}$$

$$= \begin{bmatrix} c_J & d_J & d_{J-1} & \cdots & d_1 \end{bmatrix}. \tag{42}$$

[2]The "RDWT" moniker comes from [12].

If $\mathcal{H} = \mathbb{C}^N$, then $X^{(J)} \in \mathbb{C}^M$, where $M = (J+1)N$. Otherwise, if $\mathcal{H} = \ell^2(\mathbb{Z})$, then $X^{(J)} \in \ell^2(\mathcal{I})$. We note that, in either case,

$$\left\| X^{(J)} \right\|^2 = \left\| c_J \right\|^2 + \sum_{j=1}^{J} \left\| d_j \right\|^2. \tag{43}$$

In order to reconstruct x in the original signal domain given $X^{(J)}$ in the RDWT domain, one recursively performs the following synthesis operation

$$c_j[k] = \frac{1}{2}\bigg(h_j[k] * c_{j+1}[k] + g_j[k] * d_{j+1}[k] \bigg). \tag{44}$$

The RDWT is a frame expansion, a fact we will verify below by calculating its frame bounds. The *à trous* synthesis procedure of (44) is the dual-frame reconstruction, (14), for this frame.

Assuming that $\mathcal{H} = \ell^2(\mathbb{Z})$, we can express the RDWT-filtering operations in the Fourier-frequency domain. That is, in the Fourier domain, (39), (40), and (44) become

$$\widehat{c}_{j+1}(\omega) = \widehat{h}_j^*(\omega)\widehat{c}_j(\omega), \tag{45}$$

$$\widehat{d}_{j+1}(\omega) = \widehat{g}_j^*(\omega)\widehat{c}_j(\omega), \tag{46}$$

$$\widehat{c}_j(\omega) = \frac{1}{2}\left(\widehat{h}_j(\omega)\widehat{c}_{j+1}(\omega) + \widehat{g}_j(\omega) d_{j+1}(\omega) \right), \tag{47}$$

respectively. If we expand the recursion of (45)-(46) and invoke (37) and (38), we result in

$$\begin{aligned} \widehat{c}_j(\omega) &= \left[\prod_{\lambda=0}^{j-1} \widehat{h}_j^*(\omega) \right] \widehat{x}(\omega) \\ &= \left[\prod_{\lambda=0}^{j-1} \widehat{h}_0^*(2^\lambda \omega) \right] \widehat{x}(\omega), \end{aligned} \tag{48}$$

$$\begin{aligned} \widehat{d}_j(\omega) &= \widehat{g}_{j-1}^*(\omega) \left[\prod_{\lambda=0}^{j-2} \widehat{h}_j^*(\omega) \right] \widehat{x}(\omega) \\ &= \widehat{g}_0^* \left(2^{j-1}\omega \right) \left[\prod_{\lambda=0}^{j-2} \widehat{h}_0^*(2^\lambda \omega) \right] \widehat{x}(\omega). \end{aligned} \tag{49}$$

We note that an alternative implementation of the RDWT was independently proposed by Shensa [17] and Beylkin [18]. In essence, this implementation employs filtering and downsampling as in the usual critically sampled DWT; however, all "phases" of downsampled coefficients are retained and arranged as "children" of the signal that was decomposed. The process is repeated on all the lowpass bands to achieve multiple decomposition scales that form a "tree" of decompositions. Although this alternative tree-based RDWT is a useful and common implementation in practice, we will focus on the *à trous* implementation here since it is much more amenable to mathematical analysis and derivation, a characteristic we exploit as we study the noise properties of the RDWT in the next section.

4. Noise Properties of the RDWT

In this section, we consider the effect of additive random noise in the RDWT domain. Specifically, we consider white-noise signals X in the RDWT domain and reconstruct x with the *à trous* synthesis algorithm of (44). We will find that, like all redundant frames, the RDWT exhibits a certain resilience to added noise. However, we will find that the RDWT permits a characterization of this resilience much stronger than that indicated by Theorem 1, which guarantees noise reduction, but leaves us unsure as to how much. We first derive frame bounds for the RDWT operating in $\ell^2(\mathbb{Z})$ in the next section before presenting our central results concerning noise and the RDWT in Sec. 4.2.

4.1 Frame Bounds of the RDWT

Lemma 1 *A single-scale RDWT operating in $\ell^2(\mathbb{Z})$ is a tight-frame expansion with frame bounds $A = B = 2$.*

Proof: Consider $x \in \ell^2(\mathbb{Z})$ with single-scale RDWT $X^{(1)} = \mathrm{RDWT}_1[x]$. Thus,

$$X^{(1)} = \left\|c_1\right\|^2 + \left\|d_1\right\|^2 \tag{50}$$

$$= \left\|\widehat{c}_1\right\|^2 + \left\|\widehat{d}_1\right\|^2 \tag{51}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\widehat{c}_1(\omega)\right|^2 d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\widehat{d}_1(\omega)\right|^2 d\omega. \tag{52}$$

From (45) and (46), we have

$$\left\|X^{(1)}\right\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\left|\widehat{h}_0(\omega)\right|^2 + \left|\widehat{g}_0(\omega)\right|^2\right]\left|\widehat{x}(\omega)\right|^2 d\omega \tag{53}$$

$$= 2\left\|\widehat{x}\right\|^2 \tag{54}$$

$$= 2\left\|x\right\|^2, \tag{55}$$

since h and g are power-complementary (26). Comparing to (12), we have that the frame bounds for the single-scale RDWT are $A = B = 2$. ■

Lemma 2 *If $X^{(J)}$ is the J-scale RDWT of $x \in \ell^2(\mathbb{Z})$, then*

$$\left\|X^{(J)}\right\|^2 = \left\|c_{J-1}\right\|^2 + \left\|X^{(J-1)}\right\|^2. \tag{56}$$

Proof: See App. A. ■

Theorem 2 *A J-scale RDWT operating in $\ell^2(\mathbb{Z})$ is a frame expansion with frame bounds $A = 2$ and $B = 2^J$.*

Proof: By the definition of a frame, it is sufficient to show that the frame bounds exist to show that the RDWT is a frame. To establish frame bounds, we use a proof by induction. The inductive basis is given by Lemma 1. The inductive step is as follows. Suppose that for $J \geq 2$, we have

$$2\left\|x\right\|^2 \leq \left\|X^{(J-1)}\right\|^2 \leq 2^{J-1}\left\|x\right\|^2 \tag{57}$$

for $X^{J-1} = \begin{bmatrix} c_{J-1} & d_{J-1} & d_{J-2} & \cdots & d_1 \end{bmatrix}$, where $X^{(J-1)}$ is the $(J-1)$-scale RDWT of x. Then, for the J-scale RDWT, we have from Lemma 2 and (57),

$$\left\|X^{(J)}\right\|^2 = \left\|c_{J-1}\right\|^2 + \left\|X^{(J-1)}\right\|^2 \tag{58}$$

$$\geq \left\|X^{(J-1)}\right\|^2 \tag{59}$$

$$\geq 2\left\|x\right\|^2, \tag{60}$$

which establishes inductively that the lower bound A satisfies

$$A \geq 2. \tag{61}$$

For the upper bound, we note that, from (43) and (57), we have

$$\begin{aligned} \|c_{J-1}\|^2 &= \|X^{(J-1)}\|^2 - \sum_{j=1}^{J-1} \|d_j\|^2 \\ &\leq 2^{J-1}\|x\|^2 - \sum_{j=1}^{J-1} \|d_j\|^2 \\ &\leq 2^{J-1}\|x\|^2. \end{aligned} \tag{62}$$

From Lemma 2, (57), and (62), we have

$$\left\|X^{(J)}\right\|^2 = \left\|c_{J-1}\right\|^2 + \left\|X^{(J-1)}\right\|^2 \tag{63}$$

$$\leq 2^{J-1}\|x\|^2 + 2^{J-1}\|x\|^2 \tag{64}$$

$$\leq 2^J\|x\|^2, \tag{65}$$

which establishes that the upper bound B satisfies

$$B \leq 2^J. \tag{66}$$

In App. B, we show that the bounds of $A = 2$ and $B = 2^J$ are the tightest possible frame bounds since we can find sequences of $x \in \ell^2(\mathbb{Z})$ that asymptotically meet these bounds. ■

4.2 Additive Noise in the RDWT Domain

In this section, we consider zero-mean, white-noise signals in the RDWT domain. Specifically, we define a zero-mean white-noise signal $X \in \mathbb{C}^\infty$ with variance σ^2 as

$$E\Big[X[n]\Big] = 0, \tag{67}$$

$$E\Big[X[n]X^*[m]\Big] = \begin{cases} 0, & n \neq m, \\ \sigma^2, & n = m. \end{cases} \tag{68}$$

Theorem 3 *Suppose we have $X^{(J)} \in \mathbb{C}^\infty$ such that*

$$X^{(J)} = \begin{bmatrix} c_J & d_J & d_{J-1} & \cdots & d_1 \end{bmatrix}, \tag{69}$$

where $c_J, d_j \in \mathbb{C}^\infty$. Suppose a single subband of $X^{(J)}$ consists of zero-mean white noise of variance σ^2 while all the other subbands are zero. Then, the reconstruction x due to the à trous synthesis algorithm of (44) is zero-mean noise with variance

$$E\left[\big|x[k]\big|^2\right] = \frac{\sigma^2}{4^j}, \tag{70}$$

where j is the scale of the subband in which the noise resides.

Proof: Establishing that x has zero mean is straightforward, so we will focus on the variance. The noise in X will be in either c_j or d_j while all the other subbands are zero. Let us consider first the case that noise is in c_j. Consequently, we have from (44),

$$x[k] = c_0[k] = \left(\frac{1}{2}\right)^j h_0[k] * h_1[k] * \cdots * h_{j-1}[k] * c_j[k]. \tag{71}$$

The power spectral density of the output of the synthesis operation will be

$$S_x(\omega) = \left| \left(\frac{1}{2}\right)^j \widehat{h}_0(\omega)\widehat{h}_1(\omega) \cdots \widehat{h}_{j-1}(\omega) \right|^2 \sigma^2 \tag{72}$$

$$= \frac{\sigma^2}{4^j} \left| \prod_{\lambda=0}^{j-1} \widehat{h}_0(2^\lambda \omega) \right|^2, \tag{73}$$

since the power spectral density of c_j is σ^2, and $\widehat{h}_j(\omega) = \widehat{h}_0(2^j\omega)$ from (37). Invoking (28), we have

$$S_x(\omega) = \frac{\sigma^2}{4^j} \left| \widehat{\phi}_j(\omega) \right|^2, \tag{74}$$

where ϕ_j is a basis vector of the DTWS underlying the RDWT. The variance of x is then

$$E\left[\left| x[k] \right|^2 \right] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(\omega)\, d\omega \tag{75}$$

$$= \frac{\sigma^2}{2\pi 4^j} \int_{-\pi}^{\pi} \left| \widehat{\phi}_j(\omega) \right|^2 d\omega \tag{76}$$

$$= \frac{\sigma^2}{4^j} \left\| \widehat{\phi}_j \right\|^2 \tag{77}$$

$$= \frac{\sigma^2}{4^j}, \tag{78}$$

where the last equality is due to (30). Thus, if the noise is in c_j, we have (70).

For d_j, the proof is similar. In this case, we have from (44),

$$x[k] = \left(\frac{1}{2}\right)^j h_0[k] * h_1[k] * \cdots * h_{j-2}[k] * g_{j-1}[k] * d_j[k], \tag{79}$$

while the power spectral density is

$$S_x(\omega) = \frac{\sigma^2}{4^j} \left| \widehat{g}_0(2^{j-1}\omega) \prod_{\lambda=0}^{j-2} \widehat{h}_0(2^\lambda \omega) \right|^2 \tag{80}$$

from (37) and (38). Invoking (29), we have

$$S_x(\omega) = \frac{\sigma^2}{4^j} \left| \widehat{\psi}_j(\omega) \right|^2. \tag{81}$$

The variance of x is then

$$E\left[\left|x[k]\right|^2\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_x(\omega)\,d\omega \tag{82}$$

$$= \frac{\sigma^2}{2\pi 4^j}\int_{-\pi}^{\pi}\left|\widehat{\psi}_j(\omega)\right|^2 d\omega \tag{83}$$

$$= \frac{\sigma^2}{4^j}\left\|\widehat{\psi}_j\right\|^2 \tag{84}$$

$$= \frac{\sigma^2}{4^j}, \tag{85}$$

where the last equality is due to (31). Thus, if the noise is in d_j, we have (70). ■

Theorem 4 *Suppose $X^{(J)} \in \mathbb{C}^\infty$ is a zero-mean, white-noise signal with variance σ^2. That is, suppose $X^{(J)}$ consists of mutually uncorrelated noise in all subbands c_J, d_J, $\ldots$, d_1. Then, the reconstruction x from (44) is zero-mean noise with variance*

$$E\left[\left|x[k]\right|^2\right] = \frac{\sigma^2}{3}\left[1 + 2\left(\frac{1}{4}\right)^J\right]. \tag{86}$$

Proof: Because the noise in a given subband is uncorrelated from that in the other subbands, the output of the synthesis operation (44) for that subband will be uncorrelated from the synthesis outputs for the other subbands. Thus, the total variance of the output is the sum of the output variances due to each individual subband as given by Theorem 3. Consequently, we have

$$E\left[\left|x[k]\right|^2\right] = \sigma^2\left(\frac{1}{4}\right)^J + \sigma^2\sum_{j=1}^{J}\left(\frac{1}{4}\right)^j. \tag{87}$$

Using the geometric-series theorem,

$$\sum_{n=m}^{\infty} r^n = \frac{r^m}{1-r}, \quad \forall r, |r| < 1, \tag{88}$$

we have

$$E\left[\left|x[k]\right|^2\right] = \sigma^2\left(\frac{1}{4}\right)^J + \sigma^2\sum_{j=1}^{\infty}\left(\frac{1}{4}\right)^j - \sigma^2\sum_{j=J+1}^{\infty}\left(\frac{1}{4}\right)^j$$

$$= \sigma^2\left(\frac{1}{4}\right)^J + \frac{\sigma^2}{3}\left[1 - \left(\frac{1}{4}\right)^J\right]$$

$$= \frac{\sigma^2}{3}\left[1 + 2\left(\frac{1}{4}\right)^J\right]. \tag{89}$$

■

Corollary 2 *For zero-mean, white-noise signal $X^{(J)}$ with J large, the variance of x is approximately $\sigma^2/3$.*

Proof: This result is a direct consequence of taking the limit as $J \to \infty$ of (86). ■

4.3 Discussion

Strictly speaking, Theorem 1 applies only to finite-dimensional spaces $\mathbb{C}^N$, whereas the frame bounds in Sec. 4.1 were derived assuming $\ell^2(\mathbb{Z})$, and the noise analysis of Sec. 4.2 concerned white-noise signals in $\mathbb{C}^\infty$. If we ignore for the moment these space differences, Theorem 1 would suggest that the noise variance (expected energy per signal sample) in the original signal domain for white-noise $X^{(J)}$ with variance σ^2 is bound as

$$\frac{M\sigma^2}{4^J N} \leq \frac{1}{N} E\left[\left\|x\right\|^2\right] \leq \frac{M\sigma^2}{4N}, \tag{90}$$

assuming $x \in \mathbb{C}^N$, $X^{(J)} \in \mathbb{C}^M$, and x is reconstructed from $X^{(J)}$ with the *à trous* synthesis procedure of (44). Since, for a J-scale RDWT, $M = (J+1)N$, we have

$$\frac{(J+1)\sigma^2}{4^J} \leq \frac{1}{N} E\left[\left\|x\right\|^2\right] \leq \frac{(J+1)\sigma^2}{4}. \tag{91}$$

We note that (91) suggests a limited ability to predict the effect in the original signal domain of noise added in the RDWT domain, particularly as J becomes large. This observation conforms to our intuition concerning frames—since the frame bounds given by Theorem 2 widen as J increases, we expect to be able to predict energy from one domain to the other with decreasing precision.

However, as it turns out, we can make a much stronger characterization of the noise variance in the original signal domain than we are led to believe from the above discussion. Corollary 2 suggests that, rather than being bound by ever widening bounds, the noise variance actually approaches $\sigma^2/3$ as J increases, with a more accurate result given by Theorem 4.

5. Other Considerations

In this section, we briefly consider several issues of practical relevance. Since wavelet transforms are widely used in image-processing applications, and the RDWT has been used to provide shift invariance for such applications [6, 8], we first generalize the above results to separable 2D RDWTs. Then, as biorthogonal filters are often used in practice in order to increase filter-design freedom (e.g., to obtain linear phase), we consider how the assumption of biorthogonality, rather than orthonormality, affects the above results. Finally, we reconsider the assumption that the RDWT frame is uniform, since, in certain practical settings, it is desired that the lowpass gain remain unity over all scales.

5.1 The Separable 2D RDWT

In order to provide a redundant transform for 2D image signals, the 1D RDWT as described in Sec. 3 can implement a 2D RDWT if the analysis and synthesis procedures are deployed in a separable fashion, that is, first along the rows of an image, and then along the columns, or vice versa. In this case, the *à trous* analysis procedure of (39) and (40) becomes

$$c_{j+1}[k,l] = h_j[-k] * h_j[-l] * c_j[k,l], \tag{92}$$
$$d^H_{j+1}[k,l] = g_j[-k] * h_j[-l] * c_j[k,l], \tag{93}$$
$$d^V_{j+1}[k,l] = h_j[-k] * g_j[-l] * c_j[k,l], \tag{94}$$
$$d^D_{j+1}[k,l] = g_j[-k] * g_j[-l] * c_j[k,l], \tag{95}$$

where d^H, d^V, and d^D are the horizontal, vertical, and diagonal subbands, respectively, of the 2D RDWT decomposition. The synthesis procedure of (44) becomes

$$c_j[k,l] = \frac{1}{4}\Big[h_j[k] * h_j[l] * c_{j+1}[k,l] + g_j[k] * h_j[l] * d^H_{j+1}[k,l] + h_j[k] * g_j[l] * d^V_{j+1}[k,l] + g_j[k] * g_j[l] * d^D_{j+1}[k,l]\Big]. \quad (96)$$

From these operations, it is straightforward to generalize the discussion of Sec. 4.2 to the separable 2D transform. Specifically, the single-subband result of Theorem 3 becomes

$$E\left[\left|x[k]\right|^2\right] = \sigma^2 \left(\frac{1}{16}\right)^j, \quad (97)$$

while the full-transform result of Theorem 4 becomes

$$E\left[\left|x[k]\right|^2\right] = \frac{\sigma^2}{5}\left[1 + 4\left(\frac{1}{16}\right)^J\right]. \quad (98)$$

5.2 The RDWT with Biorthogonal Filters

The *à trous* RDWT can be implemented with biorthogonal, rather than orthogonal, filters. In this case, we have primary scaling and wavelet filters h and g, respectively, as well as dual scaling and wavelet filters, $\widetilde{h}$ and $\widetilde{g}$, respectively. The *à trous* analysis procedure of (39) and (40) uses the primary filters h and g while the synthesis algorithm of (44) is modified to employ the dual filters $\widetilde{h}$ and $\widetilde{g}$ to become

$$c_j[k] = \frac{1}{2}\left(\widetilde{h}_j[k] * c_{j+1}[k] + \widetilde{g}_j[k] * d_{j+1}[k]\right). \quad (99)$$

In the case of biorthogonal filters, the derivation in the proof of Theorem 3 is modified slightly such that $\widehat{\widetilde{h}}_0$ and $\widehat{\widetilde{g}}_0$ replace $\widehat{h}_0$ and $\widehat{g}_0$, and we end up with

$$E\left[\left|x[k]\right|^2\right] = \sigma^2 \left(\frac{1}{4}\right)^j \left\|\widetilde{\phi}_j\right\|^2, \quad (100)$$

for noise in c_j, and

$$E\left[\left|x[k]\right|^2\right] = \sigma^2 \left(\frac{1}{4}\right)^j \left\|\widetilde{\psi}_j\right\|^2, \quad (101)$$

for noise in d_j. However, we fail to establish that either is $\sigma^2/4^j$ since, generally, $\widetilde{\phi}_j$ and $\widetilde{\psi}_j$ do not simultaneously have unit norm unless the corresponding DTWS basis is strictly orthonormal rather than biorthogonal.

All is not lost however. If the biorthogonal system is "near-orthonormal," or "snug," then

$$\left\|\widetilde{\phi}_j\right\|^2 \approx \left\|\widetilde{\psi}_j\right\|^2 \approx 1, \quad (102)$$

so that we very nearly achieve the result promised by Theorem 3. Near-orthonormal systems are frequently encountered in practice; for example, the 9-7 biorthogonal basis [26, 27] used ubiquitously in image-processing and other applications is very near to being orthonormal. For this system,

$$\left\|\widetilde{\phi}\right\|^2 = \sum_k \left|\widetilde{h}[k]\right|^2 \approx 0.98295286, \quad (103)$$

$$\left\|\widetilde{\psi}\right\|^2 = \sum_k \left|h[k]\right|^2 \approx 1.04043553, \quad (104)$$

Table 1: Empirical evaluation of the results of Theorems 3 and 4 for RDWTs with orthonormal and biorthogonal filters. "Theory" is value predicted by Theorem 3 for individual subbands, or by Theorem 4 for the entire signal. "Actual" is value from experiments run for a 5-scale RDWT on a white-noise signal of length 100,000 with zero mean and unit variance, with results averaged over 1000 trials. "D4" is the Daubechies length-4 orthonormal filter [28]; "CDF-9/7" is the Cohen-Daubechies-Feauveau length-9/7 biorthogonal filters [26, 27].

		D4		CDF-9/7	
Subband	Theory	Actual	% Error	Actual	% Error
c_5	0.0009765625	0.0009875311	1.1232%	0.0010488657	7.4038%
d_5	0.0009765625	0.0009625418	−1.4357%	0.0010503135	7.5521%
d_4	0.0039062500	0.0039076937	0.0370%	0.0041935694	7.3554%
d_3	0.0156250000	0.0156290960	0.0262%	0.0162325535	3.8883%
d_2	0.0625000000	0.0624991378	−0.0014%	0.0604332700	−3.3068%
d_1	0.2500000000	0.2499814210	−0.0074%	0.2600915191	4.0366%
Entire signal	0.3339843750	0.3340371660	0.0158%	0.3430821975	2.7240%

which are both very close to 1. Consequently, both Theorems 3 and 4 provide close estimates even though the 9-7 system is not orthonormal. In Table 1, we verify this observation empirically. In these results, we typically see a 3–7% error between the value predicted by Theorem 3 and that actually obtained for the 9-7 filter.

5.3 Alternate Normalizations

Throughout this paper, we have assumed that the h and g filters underlying the RDWT were normalized such that the corresponding frame was uniform. However, alternate filter normalizations are sometimes used in practice, in which case the preceding analysis is altered. For example, in some signal-processing applications, it is desired that the scaling filter not alter the dynamic range of the signal; consequently, one normalizes h such that

$$\sum_k h[k] = 1, \tag{105}$$

rather than using the orthonormal normalization of (22). In this case, the $\frac{1}{2}$ scaling factor is removed from the *à trous* synthesis procedure of (44). Since the RDWT frame is no longer uniform, Theorem 1 no longer applies—nor is it needed, because interestingly, under this normalization scheme, the RDWT becomes a tight frame with bounds $A = B = 1$, as was shown in [5]. As a consequence, the result of Theorem 4 simplifies to

$$E\left[\left|x[k]\right|^2\right] = \sigma^2, \tag{106}$$

while the result from Theorem 3 becomes

$$E\left[\left|x[k]\right|^2\right] = \frac{\sigma^2}{2^j}, \tag{107}$$

since it can be shown that the norms of the associated DTWS basis are

$$\left\|\phi_j\right\|^2 = \left\|\psi_j\right\|^2 = 2^{-j}. \tag{108}$$

Thus, this particular normalization offers a simple conservation-of-energy relationship ($\left\|x\right\|^2 = \left\|X\right\|^2$), which may be an advantage should total energy be the quantity of interest in an application. More frequently, however, applications need noise energy on a subband-by-subband basis, for which we have a result very similar to that of Theorem 3. Thus, both normalizations schemes offer equivalent performance potential from this perspective.

6. Conclusions

In this paper, we considered the effect of additive noise in RDWT expansions. We found that, contrary to expectations, we can very accurately characterize noise distortion in the original signal domain from distortion in the RDWT domain, despite the fact that the frame bounds of the RDWT widen as the number of decomposition scales increases. The most important aspect of our analytical results was a relationship for noise distortion on a subband-by-subband basis; that is, we determined that the noise resulting from a single RDWT subband depends only on the decomposition scale at which the subband resides and is independent of that from other subbands.

Appendices

A Proof of Lemma 2

Consider the sum

$$\left\|c_J\right\|^2 + \left\|d_J\right\|^2 - \left\|c_{J-1}\right\|^2 = \left\|\widehat{c}_J\right\|^2 + \left\|\widehat{d}_J\right\|^2 - \left\|\widehat{c}_{J-1}\right\|^2. \tag{109}$$

From (4), (48), and (49), we have

$$\left\|\widehat{c}_J\right\|^2 + \left\|\widehat{d}_J\right\|^2 - \left\|\widehat{c}_{J-1}\right\|^2 =$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\prod_{\lambda=0}^{J-1}\left|\widehat{h}_0\left(2^\lambda\omega\right)\right|^2 + \left|\widehat{g}_0\left(2^{J-1}\omega\right)\right|^2\prod_{\lambda=0}^{J-2}\left|\widehat{h}_0\left(2^\lambda\omega\right)\right|^2 - \prod_{\lambda=0}^{J-2}\left|\widehat{h}_0\left(2^\lambda\omega\right)\right|^2\right]\left|\widehat{x}(\omega)\right|^2 d\omega. \tag{110}$$

Since from (26),

$$\left|\widehat{g}_0\left(2^{J-1}\omega\right)\right|^2 = 2 - \left|\widehat{h}_0\left(2^{J-1}\omega\right)\right|^2, \tag{111}$$

we have

$$\left\|\widehat{c}_J\right\|^2 + \left\|\widehat{d}_J\right\|^2 - \left\|\widehat{c}_{J-1}\right\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\prod_{\lambda=0}^{J-2}\left|\widehat{h}_0\left(2^\lambda\omega\right)\right|^2\right]\left|\widehat{x}(\omega)\right|^2 d\omega \tag{112}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\widehat{c}_{J-1}(\omega)\right|^2 d\omega \tag{113}$$

$$= \left\|\widehat{c}_{J-1}\right\|^2 \tag{114}$$

$$= \left\|c_{J-1}\right\|^2, \tag{115}$$

where we again employ (48). Thus, we have

$$\left\|c_J\right\|^2 + \left\|d_J\right\|^2 - \left\|c_{J-1}\right\|^2 = \left\|c_{J-1}\right\|^2. \tag{116}$$

Now, expanding (43) in terms of $X^{(J-1)}$ yields

$$\left\|X^{(J)}\right\|^2 = \left\|c_J\right\|^2 + \sum_{j=1}^{J}\left\|d_j\right\|^2 \tag{117}$$

$$= \left\|c_J\right\|^2 + \left\|d_J\right\|^2 - \left\|c_{J-1}\right\|^2 + \left\|X^{(J-1)}\right\|^2, \tag{118}$$

and we arrive at (56) by substituting (116) into the above expression. ■

B Signals Satisfying RDWT Frame Bounds

We now show that the bounds of $A = 2$ and $B = 2^J$ are the tightest possible frame bounds since we can find sequences $x \in \ell^2(\mathbb{Z})$ that asymptotically meet these bounds. Specifically, consider a constant sequence $x[k] = 1$. Technically, this x is not in $\ell^2(\mathbb{Z})$; however, we define $x_N \in \ell^2(\mathbb{Z})$ as

$$x_N[k] = \begin{cases} \frac{1}{\sqrt{2N+1}}, & -N \leq k \leq N, \\ 0, & \text{else}, \end{cases} \tag{119}$$

for $N = 1, 2, \ldots$. We note that

$$\|x_N\|^2 = 1, \tag{120}$$

for all N. We have from (22) and (23),

$$h[-k] * x[k] = \sum_n h[n] = \sqrt{2}, \tag{121}$$

$$g[-k] * x[k] = \sum_n g[n] = 0. \tag{122}$$

Thus, we have in the limit,

$$\lim_{N\to\infty} \sqrt{2N+1}\Big(h[-k] * x_N[k]\Big) = h[-k] * \Big(\lim_{N\to\infty} x_N[k]\sqrt{2N+1}\Big) \tag{123}$$

$$= h[-k] * x[k] \tag{124}$$

$$= \sum_n h[n] = \sqrt{2}, \tag{125}$$

$$\lim_{N\to\infty} \sqrt{2N+1}\Big(g[-k] * x_N[k]\Big) = g[-k] * \Big(\lim_{N\to\infty} x_N[k]\sqrt{2N+1}\Big) \tag{126}$$

$$= g[-k] * x[k] \tag{127}$$

$$= \sum_n g[n] = 0, \tag{128}$$

which, in conjunction with (39) and (40), produce

$$\lim_{N\to\infty} c_j[k]\sqrt{2N+1} = 2^{j/2}, \tag{129}$$

$$\lim_{N\to\infty} d_j[k]\sqrt{2N+1} = 0, \tag{130}$$

for $j = 1, \ldots, J$. We see then that

$$\lim_{N\to\infty} c_j[k]\sqrt{2N+1} = 2^{j/2} \lim_{N\to\infty} x_N[k]\sqrt{2N+1}, \tag{131}$$

and so

$$\lim_{N\to\infty} \|c_j\|^2(2N+1) = 2^j \lim_{N\to\infty} \|x_N\|^2(2N+1), \tag{132}$$

or

$$\lim_{N\to\infty} \|c_j\|^2 = 2^j \lim_{N\to\infty} \|x_N\|^2 \tag{133}$$

$$= 2^j, \tag{134}$$

where the last equality derives from (120). Additionally, we have

$$\lim_{N\to\infty} \|d_j\|^2 = 0. \tag{135}$$

Now, consider the quantity β_N,

$$\beta_N = \frac{\left\|X_N^{(J)}\right\|^2}{\|x_N\|^2}, \tag{136}$$

where $X_N^{(J)} = \mathrm{RDWT}_J\left[x_N\right]$. From (120), (43), (134), and (135), we have

$$\lim_{N\to\infty} \beta_N = \lim_{N\to\infty} \frac{\left\|X_N^{(J)}\right\|^2}{\left\|x_N\right\|^2} \tag{137}$$

$$= \lim_{N\to\infty} \left\|X_N^{(J)}\right\|^2 \tag{138}$$

$$= \lim_{N\to\infty} \left[\left\|c_J\right\|^2 + \sum_{j=1}^{J} \left\|d_j\right\|^2\right] \tag{139}$$

$$= 2^J. \tag{140}$$

We note from (6), $\beta_N \le B,\ \forall N$, thus it is true in the limit. Consequently, we have $2^J \le B$ and, from (66), $B \le 2^J$. Thus, $B = 2^J$.

Alternatively, consider $x[k] = (-1)^k$. Again, note that $x \notin \ell^2(\mathbb{Z})$, but that we can define $x_N \in \ell^2(\mathbb{Z})$ as

$$x_N[k] = \begin{cases} \frac{(-1)^k}{\sqrt{2N+1}}, & -N \le k \le N, \\ 0, & \text{else}, \end{cases} \tag{141}$$

for $N = 1, 2, \ldots$. We note that

$$\left\|x_N\right\|^2 = 1, \tag{142}$$

for all N. Thus, we have in the limit,

$$\lim_{N\to\infty} \sqrt{2N+1}\Big(h[-k] * x_N[k]\Big) = h[-k] * \Big(\lim_{N\to\infty} x_N[k]\sqrt{2N+1}\Big) \tag{143}$$

$$= h[-k] * x[k] \tag{144}$$

$$= \sum_n h[-n] x[k-n] \tag{145}$$

$$= \sum_n h[-n](-1)^{k-n} \tag{146}$$

$$= (-1)^k \left(\sum_n h[2n] - \sum_n h[2n+1]\right) \tag{147}$$

$$= 0, \tag{148}$$

where the second-to-last equality comes from splitting the sum into even and odd indices, and the last equality comes from (24). From (25) and (22),

$$\lim_{N\to\infty} \sqrt{2N+1}\Big(g[-k] * x_N[k]\Big) = g[-k] * \Big(\lim_{N\to\infty} x_N[k]\sqrt{2N+1}\Big) \tag{149}$$

$$= g[-k] * x[k] \tag{150}$$

$$= \sum_n g[-n] x[k-n] \tag{151}$$

$$= \pm \sum_n (-1)^{-n} h[K+n](-1)^{k-n} \tag{152}$$

$$= \pm(-1)^k \sum_n h[n] \tag{153}$$

$$= \pm(-1)^k \sqrt{2}. \tag{154}$$

Thus, from (39) and (40),

$$\lim_{N\to\infty} c_j[k]\sqrt{2N+1} = 0, \quad j = 1, \ldots, J, \tag{155}$$

$$\lim_{N\to\infty} d_j[k]\sqrt{2N+1} = \begin{cases} \pm(-1)^k\sqrt{2}, & j = 1, \\ 0, & j = 2, \ldots, J. \end{cases} \tag{156}$$

We see that then

$$\lim_{N\to\infty} \left\|c_j\right\|^2 = 0, \tag{157}$$

and

$$\lim_{N\to\infty} \left\|d_1\right\|^2 (2N+1) = 2 \lim_{N\to\infty} \left\|x_N\right\|^2 (2N+1), \tag{158}$$

or

$$\lim_{N\to\infty} \left\|d_1\right\|^2 = 2 \lim_{N\to\infty} \left\|x_N\right\|^2 \tag{159}$$

$$= 2, \tag{160}$$

where the last equality derives from (142). Define α_N as

$$\alpha_N = \frac{\left\|X_N^{(J)}\right\|^2}{\left\|x_N\right\|^2}. \tag{161}$$

We note that from (6), $\alpha_N \geq A, \ \forall N$. From (142), (43), (157), and (160), we have

$$\lim_{N\to\infty} \alpha_N = \lim_{N\to\infty} \frac{\left\|X_N^{(J)}\right\|^2}{\left\|x_N\right\|^2} \tag{162}$$

$$= \lim_{N\to\infty} \left\|X_N^{(J)}\right\|^2 \tag{163}$$

$$= \lim_{N\to\infty} \left[\left\|c_J\right\|^2 + \sum_{j=1}^{J} \left\|d_j\right\|^2 \right] \tag{164}$$

$$= 2. \tag{165}$$

Consequently, from (61) and the derivation above, we have $2 \leq A \leq 2$, so $A = 2$. ■

Acknowledgment

The author thanks L. Hua for devising a preliminary proof of Theorem 2.

References

[1] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[2] V. K. Goyal, M. Vetterli, and N. T. Thao, "Quantized overcomplete expansions in $\mathbb{R}^n$: Analysis, synthesis, and algorithms," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 16–31, January 1998.

[3] Z. Cvetković, "Resilience properties of redundant expansions under additive noise and quantization," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 644–656, March 2003.

[4] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, July 1992.

[5] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, November 1995.

[6] H.-W. Park and H.-S. Kim, "Motion estimation using low-band-shift method for wavelet-based moving-picture coding," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 577–587, April 2000.

[7] G. Van der Auwera, A. Munteanu, P. Schelkens, and J. Cornelius, "Scalable wavelet video-coding with in-band prediction—The bottom-up overcomplete discrete wavelet transform," in *Proceedings of the International Conference on Image Processing*, Rochester, NY, September 2002, vol. 3, pp. 725–728.

[8] S. Cui, Y. Wang, and J. E. Fowler, "Mesh-based motion estimation and compensation in the wavelet domain using a redundant transform," in *Proceedings of the International Conference on Image Processing*, Rochester, NY, September 2002, vol. 1, pp. 693–696.

[9] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.

[10] J. Kovačević, P. L. Dragotti, and V. K. Goyal, "Filter bank frame expansions with erasures," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1439–1450, June 2002.

[11] M. N. Do and M. Vetterli, "Framing pyramids," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2329–2342, September 2003.

[12] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice-Hall, Upper Saddle River, NJ, 1998.

[13] P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Upper Saddle River, NJ, 1993.

[14] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[15] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets: Time-Frequency Methods and Phase Space*, J.-M. Combes, A. Grossman, and P. Tchamichian, Eds., pp. 286–297. Springer-Verlag, Berlin, Germany, 1989, Proceedings of the International Conference, Marseille, France, December 14–18, 1987.

[16] P. Dutilleux, "An implementation of the "algorithme à trous" to compute the wavelet transform," in *Wavelets: Time-Frequency Methods and Phase Space*, J.-M. Combes, A. Grossman, and P. Tchamichian, Eds., pp. 298–304. Springer-Verlag, Berlin, Germany, 1989, Proceedings of the International Conference, Marseille, France, December 14–18, 1987.

[17] M. J. Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithms," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2464–2482, October 1992.

[18] G. Beylkin, "On the representation of operators in bases of compactly supported wavelets," *SIAM Journal on Numerical Analysis*, vol. 29, no. 6, pp. 1716–1740, December 1992.

[19] Y. Xu, J. B. Weaver, D. Healy, Jr., and J. Lu, "Wavelet transform domain filters: A spatially selective noise filtration technique," *IEEE Transactions on Image Processing*, vol. 3, no. 6, pp. 747–758, November 1994.

[20] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, Jr., "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, January 1996.

[21] J.-G. Cao, J. E. Fowler, and N. H. Younan, "An image-adaptive watermark based on a redundant wavelet transform," in *Proceedings of the International Conference on Image Processing*, Thessaloniki, Greece, October 2001, vol. 2, pp. 277–280.

[22] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, Jr., "Nonlinear processing of a shift invariant DWT for noise reduction," in *Wavelet Applications II*, H. H. Szu, Ed., Orlando, FL, April 1995, Proc. SPIE 2491, pp. 640–651.

[23] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.

[24] S. Mallat, "Zero-crossings of a wavelet transform," *IEEE Transactions on Information Theory*, vol. 37, no. 4, pp. 1019–1033, July 1991.

[25] Z. Cvetković and M. Vetterli, "Oversampled filter banks," *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[26] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 45, no. 5, pp. 485–560, May 1992.

[27] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, April 1992.

[28] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, pp. 909–996, November 1988.