Approximation Error Analysis for Transform-based Lossless Audio Coding

Yoshikazu Yokotani*, Soontorn Oraintara*, Ralf Geiger†, Gerald Schuller† and K.R. Rao*
*University of Texas at Arlington, Electrical Engineering Dept., Arlington,TX, U.S.A.
Emails: [yoshi,oraintar,rao]@uta.edu
†Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany
Email: [ggr,shl]@idmt.fraunhofer.de

***Abstract*— The Integer Modified Discrete Cosine Transform (IntMDCT), an integer approximation of the MDCT, is a reversible transform realized by the lifting scheme and thus is a useful transform for lossless audio coding. Because of the integer approximation, however, the approximation error appears as "noise floor" in the transform domain and limits the lossless coding efficiency.**

In this paper, a theoretical analysis of the approximation error of the IntMDCT is discussed. The result is then used to design a simple test filter applied to each rounding operation of the IntMDCT in such a way that the error spectrum is shaped towards the low frequencies. As a result, especially when the spectral energy of an input signal is concentrated in the low frequency domain, the lossless coding efficiency is improved.

I. INTRODUCTION

The lifting scheme [1]-based integer transforms are quite useful for lossless coding applications such as audio [2], [3] and image [4] compression. These transforms are composed of approximated plane rotations, each of which is realized by three lifting steps associated with multiplications and rounding operations. Every rounding operation introduces rounding noise, and it is accumulated in the transform domain. The accumulated noise is interpreted as the approximation error of the original floating-point transform and it appears as "noise floor" in the transform domain. Although the error is cancelled by the inverse transform, it is desirable that the noise floor level is kept small since it has a significant impact on the coding efficiency. This is critically important especially for lossless audio coding since it requires a large size of the transform, which has many stages of approximated plane rotations in its fast implementation.

To improve the efficiency, the multi-dimensional lifting (MDL) scheme was recently proposed [5]. This technique can reduce the number of lifting steps for computing the integer modified discrete cosine transform (IntMDCT) significantly. As a result, the approximation error is lowered significantly as presented in [6].

It is possible to improve the efficiency further by shaping the approximation error spectrum towards the low frequencies especially when the spectral energy of an input signal is concentrated at the low frequencies. Since the approximation error occupies the entire band, as will be proven in Section III, shaping the error can reduce the cost to encode the noise by shifting the error under the audio signal.

Conventional quantization noise shaping has been used to shape the quantization noise towards the high frequency bands to make the noise as minimally audible as possible [7]. Therefore, the noise shaping filter to be applied to the rounding operations of the IntMDCT can be treated similarly, but in the opposite way. However, as will be seen in Section III, the rounding noise shaping is applied to rounding operations for signals in the time and frequency domains, whereas the conventional scheme is used for a quantization operation for the signal only in the time domain. Therefore, it is not clear if simply applying the conventional scheme to the rounding operations in the opposite way can shape the approximation error as expected, and thus a theoretical analysis of the approximation error is necessary before such a noise shaping scheme is considered.

In this paper, a mathematical analysis of the approximation error for the MDL scheme-based stereo IntMDCT [5] is presented. It is shown that the conventional noise shaping can be applied to rounding operations with a lowpass filter designed in the odd discrete Fourier transform (ODFT) [8] domain. Furthermore, an experimental test shows an improvement in the lossless coding efficiency when the spectral energy of an input signal is concentrated at the low frequency bands.

This paper is organized as follows: in Section II, the structure of the MDL scheme-based stereo IntMDCT is described. In Section III, the approximation error is calculated before and after rounding noise shaping is applied. In Section IV, a simulation is carried out in order to illustrate an improvement in the lossless coding efficiency by applying a simple test filter into the IntMDCT-based lossless audio codec. Section V concludes the paper.

II. MDL SCHEME-BASED STEREO INTMDCT

The MDL scheme-based stereo IntMDCT [5] transforms $2N$ stereo audio samples $x_{bL}[n]$ and $x_{bR}[n]$ for $n = 0, \ldots, N-1$ into N spectral lines $X_L[k]$ and $X_R[k]$ for $k = 0, \ldots, N-1$ in the left and right channel, respectively, where $N = 1024$ for lossless audio coding applications. The subscript L and R indicate the left and right channel, respectively. b denotes the frame block number. The IntMDCT for the stereo signal, as illustrated in Fig. 1, is composed of (a) an identical sine window and time domain aliasing operation realized by the conventional three lifting steps [1] and (b)

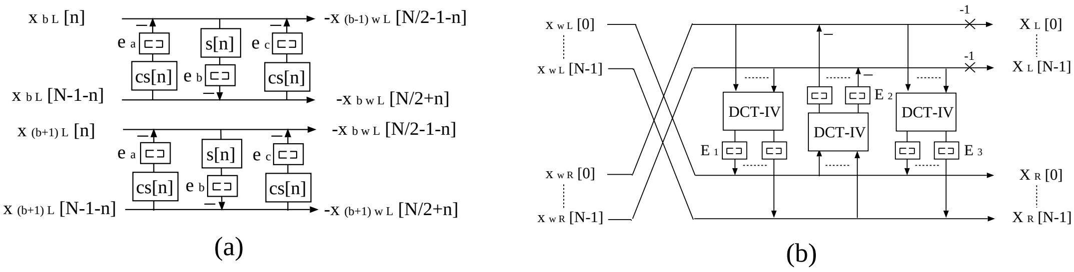

Fig. 1. Structure of the MDL scheme-based stereo IntMDCT. [] symbolizes a rounding operation. (a) the three lifting step structure for the sine window and time domain aliasing operation and (b) the MDL structure for the IntDCT-IV operation.

the integer discrete cosine transform of type IV (IntDCT-IV) operation realized by the MDL steps [5]. In Fig. 1 (a), only the left channel case is drawn. $cs[n]$ and $s[n]$ are the lifting coefficients and given by $cs[n] = (c[n]-1)/s[n]$, $c[n] = \cos\theta[n]$, and $s[n] = \sin\theta[n]$, where $\theta[n] = \pi/(2N)(n+0.5)$ for $n = 0, ..., N/2-1$. e_a, e_b, and e_c is the rounding noise introduced in the rounding operation associated with the first, second, and third lifting step, respectively. Since the lifting coefficients are floating-point values and not quantized, the rounding noises are assumed to be white and in the range of -0.5 and 0.5. The outputs of the window and time domain aliasing operation are flipped and the sign is changed, and become $x_{wL}[n]$ and $x_{wR}[n]$, which are the inputs of the IntDCT-IV in Fig. 1 (b). Note that the subscript b is omitted from x_{bwL} and x_{bwR} and the rest of the signals in Fig. 1 (b) since all the signals processed by the IntDCT-IV are in the b^{th} frame. As these signals go through the MDL steps each of which has the floating-point DCT-IV and rounding operations, the rounding noises E_1, E_2, and E_3 are injected into the signals. Because of using floating-point DCT-IV, the rounding noises are assumed to have the same statistics as e_a, e_b, and e_c.

III. A Theoretical Analysis of the Approximation Error

In this section, we mathematically analyze the approximation error of the MDL scheme-based IntMDCT before and after rounding noise shaping is applied. In each case, we compare the approximation errors obtained by results of the analysis and actual implementations to verify the results of the analysis. The input audio signals for the implementations are chosen to be the 15 audio items used in MPEG-4 task group for lossless audio coding (the sampling frequency 48kHz and quantized at 16bit PCM) [9]. In each audio item, frames representing silence are excluded since audio samples in these frames are so small that the rounding noise added in lifting steps cannot be simply assumed to be white and have a uniform distribution.

A. Case 1: No Rounding Noise Shaping

First of all, the rounding noise introduced in the three lifting steps in Fig. 1 (a) is treated. Since the noise introduced in each lifting step is multiplied by the lifting coefficient $cs[n]$ or $s[n]$ in the following steps until it reaches the output, the accumulated noise at the output can be given by:

$$e_{wu}[n] = -c[n]e_a + cs[n]e_b - e_c, \quad (1)$$
$$e_{w\ell}[n] = s[n]e_a - e_b, \quad (2)$$

where e_{wu} and $e_{w\ell}$ are the accumulated rounding noises at the upper and lower port of the output, respectively. Let e_{wL} and e_{wR} be the approximation errors associated with x_{wL} and x_{wR}, respectively. Then, these estimated variances are given by the following equations:

$$E[e^2_{wL}[n]] = E[e^2_{wR}[n]] = \begin{cases} \frac{c^2\left[\frac{N}{2}-1-n\right]+cs^2\left[\frac{N}{2}-1-n\right]+1}{12} \\ \quad \text{for} \quad n = 0, \ldots, \frac{N}{2}-1. \\ \frac{s^2\left[n-\frac{N}{2}\right]+1}{12} \\ \quad \text{for} \quad n = \frac{N}{2}, \ldots, N-1. \end{cases} \quad (3)$$

Since both channels have the identical structure, only the left channel case in (3) is proved briefly. Since e_{wL} is obtained by flipping the value and the sign changing of e_{wu} and $e_{w\ell}$, it can be given as follows:

$$e_{wL}[n] = \begin{cases} -e_{wu}\left[\frac{N}{2}-1-n\right] & \text{for } n = 0, \ldots, \frac{N}{2}-1. \\ -e_{w\ell}\left[n-\frac{N}{2}\right] & \text{for } n = \frac{N}{2}, \ldots, N-1. \end{cases} \quad (4)$$

Since e_a, e_b, and e_c are uncorrelated with one another and each variance is equal to $1/12$, $E[e^2_{wu}[n]]$ and $E[e^2_{w\ell}[n]]$ can be computed by $1/12(c^2[n]+cs^2[n]+1)$ and $1/12(s^2[n]+1)$, respectively. From these equations with (1) and (2), the left channel can be proved.

Let us start analyzing the approximation error spectra at the output of the IntDCT-IV when the inputs are x_{wL} and x_{wR} with the approximation errors e_{wL} and e_{wR}. The spectra at the spectral line index k, $E_L[k]$ and $E_R[k]$, are given by

$$E_L[k] = \mathbf{C}^k_{IV}(\mathbf{e}_{wL} + \mathbf{E}_1) + E_2[k], \quad (5)$$
$$E_R[k] = \mathbf{C}^k_{IV}(\mathbf{e}_{wR} - \mathbf{E}_2) + E_3[k], \quad (6)$$

where for $k = 0, \ldots, N-1$. $\mathbf{e}_{wL}$, $\mathbf{e}_{wR}$, $\mathbf{E}_1$, and $\mathbf{E}_2$ are $N \times 1$ column vectors of $e_{wL}[n]$, $e_{wR}[n]$, $E_1[k]$, $E_2[k]$, respectively. $\mathbf{C}^k_{IV}$ is the k^{th} row vector of the $N \times N$ DCT-IV matrix whose n^{th} element is given by

$$\mathbf{C}_{IV,(k,n)} = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right). \quad (7)$$

(5) and (6) can be derived from the structure of the MDL steps in Fig. 1 (b) and $\mathbf{C}_{IV}^{-1} = \mathbf{C}_{IV}$. From these two equations, the variance of E_L and E_R in case of no rounding noise shaping can be approximately estimated by the following theorem:

Theorem 1: In case of no rounding noise shaping, the variances of the approximation error in the IntMDCT domain in both left and right channels are approximately given as follows:

$$E[E_L^2[k]] \approx \overline{\sigma_{ew}^2} + \overline{\sigma_{ec}^2} + E[E_2^2[k]], \quad (8)$$
$$E[E_R^2[k]] \approx \overline{\sigma_{ew}^2} + \overline{\sigma_{ec}^2} + E[E_3^2[k]], \quad (9)$$

where

$$\overline{\sigma_{ew}^2} = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wR}^2[n]], \quad (10)$$
$$\overline{\sigma_{ec}^2} = \frac{1}{N}\sum_{k=0}^{N-1} E[E_1^2[k]] = \frac{1}{N}\sum_{k=0}^{N-1} E[E_2^2[k]]. \quad (11)$$

Proof: From (3), e_{wL} and e_{wR} have the same variance. In addition, E_1, E_2, and E_3 have the same variance as well. Thus, both (8) and (9) are expected to have the same value, and hence we will only prove the left channel case.

Since e_{wL} is a colored noise from (3) and E_1 and E_2 are white, they are uncorrelated with one another. Thus, the variance of $E_L[k]$ can be computed by (5):

$$\begin{aligned} E[E_L^2[k]] &= E\left[\left|\sum_{n=0}^{N-1} e_{wL}[n]\mathbf{C}_{IV,(k,n)}\right|^2\right] \\ &\quad + E\left[\left|\sum_{k'=0}^{N-1} E_1[k']\mathbf{C}_{IV,(k,k')}\right|^2\right] + E[E_2^2[k]] \\ &= \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]]\left(1 + \mathbf{C}_{A,(k,n)}\right) \\ &\quad + \frac{1}{N}\sum_{k'=0}^{N-1} E[E_1^2[k']]\left(1 + \mathbf{C}_{A,(k,k')}\right) + E[E_2^2[k]], \quad (12) \end{aligned}$$

where $\mathbf{C}_{A,(k,n)} = \cos\left(\frac{2\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right)$. Since $E[E_1^2[k']] = 1/12$ constant for $k' = 0, \ldots, N-1$ and $\sum_{k'=0}^{N-1} \mathbf{C}_{A,(k,k')} = 0$, (12) is simplified as follow:

$$E[E_L^2[k]] = \overline{\sigma_{ew}^2} + \overline{\sigma_{ec}^2} + E[E_2^2[k]] + \epsilon[k], \quad (13)$$

where $\epsilon[k]$ is given by the following:

$$\epsilon[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e_{wL}^2[n]] \cos\left(\frac{2\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right).$$

In fact, it is quite complex to simplify $\epsilon[k]$ further by using (3). Fortunately, $\epsilon[k]$ is a deterministic function, and the absolute value is bounded by around 0.015. On the other hand, a summation of the other terms in (13), $\overline{\sigma_{ew}^2} + \overline{\sigma_{ec}^2} + E[E_2^2[k]]$, is around 0.2937 constant for $k = 0, \ldots, N-1$. Thus, $\epsilon[k]$ can be assumed to be approximately zero. Consequently, the proof for the left channel is completed. For the right channel, the same procedure can be taken as prove. ♣

(8) and (9) show that each of $E[E_L^2[k]]$ and $E[E_R^2[k]]$ is approximately given by a summation of an average variance of the rounding noise introduced in the window and time domain aliasing operation and two white noises introduced in the MDL steps. Consequently, the estimated error variance for each channel is approximately flat as shown in Fig. 2.

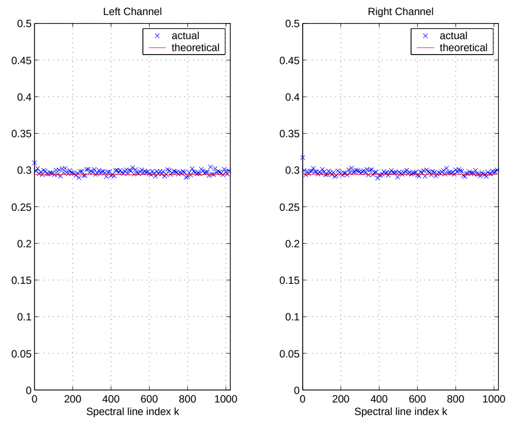

Fig. 2. The variance of the approximation error at the output of the IntDCT-IV computation. $E[E_L^2[k]]$ for the left channel and $E[E_R^2[k]]$ for the right channel.

B. Case 2: Rounding Noise Shaping

In this subsection, the approximation error is re-calculated after rounding noise shaping is employed. Fig. 3 shows a block diagram of the conventional noise shaping scheme applied to a rounding operation in a lifting step.

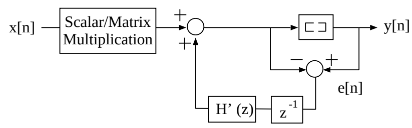

Fig. 3. A block diagram of rounding noise shaping for a lifting step. [] symbolizes a rounding operation.

In this figure, $x[n]$ is an input of the lifting step which is multiplied by a scalar constant. This type of operations appears at the three lifting step of the window operation. The scalar multiplication is replaced by a DCT-IV multiplication for the case of MDL steps. After each multiplication, the signal is added by a filtered version of the rounding noise $e[n]$. The result is then rounded and becomes the output $y[n]$.

$H'(z)$ is a causal filter of order M. The noise shaping filter formulated in Fig. 3 can be represented by $H(z) = 1 + H'(z)$ where $H(z) = \sum_{n=0}^{M} h[n]z^{-n}$ and $h[0] = 1$. The filtered noise $e[n]$, $e_H[n]$, and the DCT-IV coefficient at the spectral line index k, $E_H[k]$, can be computed by

$$e_H[n] = \sum_{m=0}^{M} h[m]e[n-m], \quad (14)$$

and

$$E_H[k] = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} e_H[n] \cos\left(\frac{\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right)$$
$$= \sqrt{\frac{2}{N}} Re\left\{W^k \sum_{m=0}^{M} h[m] e^{-j\frac{\pi}{N}m\left(k+\frac{1}{2}\right)} \times \sum_{n=-m}^{M-1-m} e[n] e^{-j\frac{\pi}{N}n\left(k+\frac{1}{2}\right)}\right\},$$

where $N-1 > M$ and both $h[m]$ and $e[n-m]$ are real. $W^k = e^{-j\frac{\pi}{2N}\left(k+\frac{1}{2}\right)}$. Since $e[n]$ is assumed to be stationary white noise, the second summation of above equation is approximately the same in case of any n as long as $N-1-M > 0$. Thus, $E_H[k]$ can be computed by the following approximation:

$$E_H[k] \approx \sqrt{\frac{2}{N}} Re\left\{W^k H_O[k] E_O[k]\right\}, \quad (15)$$

where $E_O[k]$ and $H_O[k]$ are the ODFT [8] coefficients of $e[n]$ and $h[m]$, respectively.

We now consider applying this noise shaping filter h into all the rounding operations in realization of the IntMDCT except the rounding operations in the third MDL step. In other words, the filter is used to shape $e_{wL}[n]$, $e_{wR}[n]$, $E_1[k]$, and $E_2[k]$. Let $e_{wLH}[n]$, $e_{wRH}[n]$, $E_{1H}[k]$, and $E_{2H}[k]$ be the corresponding shaped noise, convolved by the noise shaping filter h, respectively. From (5) and (6), the approximation error at the spectral line index k in the IntMDCT domain, $E_{LH}[k]$ and $E_{RH}[k]$, can be given as follows:

$$E_{LH}[k] = \mathbf{C}^k_{IV}(\mathbf{e}_{wLH} + \mathbf{E}_{1H}) + E_{2H}[k], \quad (16)$$
$$E_{RH}[k] = \mathbf{C}^k_{IV}(\mathbf{e}_{wRH} - \mathbf{E}_{2H}) + E_3[k], \quad (17)$$

where $\mathbf{e}_{wLH}$, $\mathbf{e}_{wRH}$, $\mathbf{E}_{1H}$, and $\mathbf{E}_{2H}$ are $N \times 1$ column vectors of $e_{wLH}[n]$, $e_{wRH}[n]$, $E_{1H}[k]$ and $E_{2H}[k]$, respectively. $e_{wLH}[n]$, $e_{wRH}[n]$, $E_{1H}[k]$, and $E_{2H}[k]$ are uncorrelated with one another, since each one of them is, as it is given by (14), a linear combination of an uncorrelated noise, $e_{wL}[n-m]$, $e_{wR}[n-m]$, $E_1[k-m]$, and $E_2[k-m]$, respectively for $m = 0, \ldots, M$. Hence, the variance of $E_{LH}[k]$ and $E_{RH}[k]$ can be given by

$$E[E^2_{LH}[k]] = E[E^2_{wLH}[k]] + E[E'^2_{1H}[k]] + E[E^2_{2H}[k]], \quad (18)$$
$$E[E^2_{RH}[k]] = E[E^2_{wRH}[k]] + E[E'^2_{2H}[k]] + E[E^2_3[k]]. \quad (19)$$

where $E_{wLH}[k]$, $E_{wRH}[k]$, $E'_{1H}[k]$, and $E'_{2H}[k]$ are the DCT-IV coefficients of $e_{wLH}[n]$, $e_{wRH}[n]$, $E_{1H}[k]$, and $E_{2H}[k]$, respectively. They can be obtained similar to (15) and substituted to simplify (18) and (19). However, due to a limitation of the paper space, only the left channel case is shown. In the left channel, $E_{wLH}[k]$ and $E'_{1H}[k]$ are given by

$$E_{wLH}[k] \approx \sqrt{\frac{2}{N}} Re\left\{W^k H_O[k] E_{wLO}[k]\right\}, \quad (20)$$
$$E'_{1H}[k] \approx \sqrt{\frac{2}{N}} Re\left\{W^k H_O[k] E'_{1O}[k]\right\}, \quad (21)$$

where E_{wLO} and E'_{1O} are the ODFT coefficients of $e_{wL}[n]$ and $E_1[k]$, respectively. Then, $E[E^2_{wLH}[k]]$ can be calculated as follow:

$$E[E^2_{wLH}[k]] \approx \frac{2}{N}\left[Re\{W^k H_O[k]\}\right]^2 \sum_{n=0}^{2N-1} E[e^2_{wL}[n]] \times \mathbf{C}^2_{B,(k,n)} + \frac{2}{N}\left[Im\{W^k H_O[k]\}\right]^2 \sum_{m=0}^{2N-1} E[e^2_{wL}[m]] \times \mathbf{S}^2_{B,(k,m)} - \frac{4}{N} Re\{W^k H_O[k]\} Im\{W^k H_O[k]\} \times \sum_{n=0}^{2N-1}\sum_{m=0}^{2N-1} E[e_{wL}[n] e_{wL}[m]] \mathbf{C}_{B,(k,n)} \mathbf{S}_{B,(k,m)},$$

where $Re\{x\}$ and $Im\{x\}$ takes a real and an imaginary part of a complex number x, respectively. $\mathbf{C}_{B,(k,n)} = \cos\left(\frac{\pi}{N}n\left(k+\frac{1}{2}\right)\right)$ and $\mathbf{S}_{B,(k,m)} = \sin\left(\frac{\pi}{N}m\left(k+\frac{1}{2}\right)\right)$. Since $E[e_{wL}[n] e_{wL}[m]] = 0$ for $n \neq m$, the approximation above can be simplified as follow:

$$E[E^2_{wLH}[k]] \approx |H_O[k]|^2 \overline{\sigma^2_{ew}} + \left(\left[Re\{W^k H_O[k]\}\right]^2 - \left[Im\{W^k H_O[k]\}\right]^2\right)\phi_w[k] + 2\left(Re\{W^k H_O[k]\} Im\{W^k H_O[k]\}\right)\psi_w[k], \quad (22)$$

where $\overline{\sigma^2_{ew}}$ is given by (10) and

$$\phi_w[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e^2_{wL}[n]] \cos\left(\frac{\pi}{N}n(2k+1)\right),$$
$$\psi_w[k] = \frac{1}{N}\sum_{n=0}^{N-1} E[e^2_{wL}[n]] \sin\left(\frac{\pi}{N}n(2k+1)\right).$$

Likewise, $E[E'^2_{1H}[k]]$ can be simplified and given by the following approximation which is similar to (22):

$$E[E'^2_{1H}[k]] \approx |H_O[k]|^2 \overline{\sigma^2_{ec}} + \left(\left[Re\{W^k H_O[k]\}\right]^2 - \left[Im\{W^k H_O[k]\}\right]^2\right)\phi_c[k] + 2\left(Re\{W^k H_O[k]\} Im\{W^k H_O[k]\}\right)\psi_c[k], \quad (23)$$

where $\overline{\sigma^2_{ec}}$ is given by (11) and

$$\phi_c[k] = \frac{1}{N}\sum_{k'=0}^{N-1} E[E^2_1[k']] \cos\left(\frac{\pi}{N}k'(2k+1)\right),$$
$$\psi_c[k] = \frac{1}{N}\sum_{k'=0}^{N-1} E[E^2_1[k']] \sin\left(\frac{\pi}{N}k'(2k+1)\right).$$

Finally, (18) and (19) can be re-written as the following approximations:

Theorem 2: In case of rounding noise shaping, the variances of the approximation error in the IntMDCT domain in both left and right channels are approximately given as follows:

$$E[E_{LH}^2[k]] \approx |H_O[k]|^2(\overline{\sigma_{ew}^2}+\overline{\sigma_{ec}^2}) + E[|h[k]*E_2[k]|^2] + \alpha[k](\phi_w[k]+\phi_c[k]) + \beta[k](\psi_w[k]+\psi_c[k]), \quad (24)$$

$$E[E_{RH}^2[k]] \approx |H_O[k]|^2(\overline{\sigma_{ew}^2}+\overline{\sigma_{ec}^2}) + E[E_3^2[k]] + \alpha[k](\phi_w[k]+\phi_c[k]) + \beta[k](\psi_w[k]+\psi_c[k]), \quad (25)$$

where $*$ indicates the convolution operator and

$$\alpha[k] = \left[Re\{W^k H_O[k]\}\right]^2 - \left[Im\{W^k H_O[k]\}\right]^2,$$
$$\beta[k] = 2Re\{W^k H_O[k]\}Im\{W^k H_O[k]\}.$$

Proof: (24) can be directly derived from (18), (22), and (23). (25) can be obtained similarly by the fact that $E[e_{wL}^2[m]] = E[e_{wR}^2[m]]$ and $E[E_1^2[k]] = E[E_2^2[k]]$. ♣

(24) and (25) show that e_{wL} and E_1 in the left channel and e_{wR} and E_2 in the right channel can be shaped by $H_O[k]$. However, the third and forth terms are sub-products of the noise shaping. Thus, it is necessary to evaluate the impact on the noise shaping numerically. Since $\alpha[k] \le |H_O[k]|$ and $\beta[k] \le |H_O[k]|$ for $k = 0, \ldots, N-1$, how large $\phi_w[k]+\phi_c[k]$ and $\psi_w[k]+\psi_c[k]$ compared to $\overline{\sigma_{ew}^2}+\overline{\sigma_{ec}^2}$ are evaluated by using (3) and $E[E_1^2[k]] = E[E_2^2[k]] = 1/12$. $\phi_w[k]+\phi_c[k]$, and $\psi_w[k]+\psi_c[k]$ are bounded by around 0.13 at only a few low and high frequencies and the values are close to zero elsewhere, whereas $\overline{\sigma_{ew}^2}+\overline{\sigma_{ec}^2}$ is approximately 0.2937 constant. This indicates that the impact is minor and it is possible to shape the noise injected into the window and time domain aliasing operation and the first and second MDL step by using the noise shaping filter $H_O[k]$.

IV. Lossless Coding Implementation

In this section, a simple test filter $H(z) = 1 + z^{-1}$ is incorporated into the stereo IntMDCT followed by a context based arithmetic encoder [10] to evaluate an improvement of the lossless coding efficiency due to the noise shaping. The input audio signals are the 15 MPEG audio items used in previous section and same 15 items with sampling frequency 96kHz.

In order to confirm the noise shaping effect by $H(z)$, a comparison between theoretical curves calculated by (24), (25), and numerical data of H_O and the actual data obtained by using the input audio items, and it is shown in Fig. 4. One observes that the filter lowers the error spectrum at the high frequencies (approximately $k \ge 700$ for both channels).

Table I shows average bit rates of the losslessly compressed audio items in case of no noise shaping and when the noise shaping filter is present. An improvement is observed especially when the sampling frequency of the input signals is 96kHz. In this case, the spectral energy is more concentrated at the low frequencies.

V. Conclusion

In this paper, a theoretical analysis of the approximation error for the MDL scheme-based stereo IntMDCT was discussed. It was shown that the conventional noise shaping scheme can be applied for rounding noise shaping in the opposite way by using a filter designed in the ODFT domain. An experimental test was carried out in the IntMDCT based lossless audio codec. The result showed an improvement in the coding efficiency especially when the spectral energy of the input signal is mainly concentrated at the low frequencies.

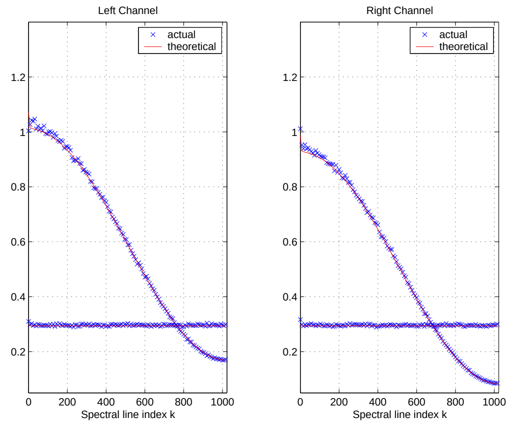

Fig. 4. The variances of the approximation error before (flat lines) and after (curves) the noise shaping filter $H(z) = 1 + z^{-1}$ is applied.

TABLE I

Average bit rates of losslessly compressed test audio items(bits/sample)

	no filter	$H(z)$
48kHz 16bit	7.755	7.749
96kHz 16bit	5.389	5.364

References

[1] I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps," *J. of Fourier Anal. Appl.*, vol. 4, no. 3, pp. 247–269, 1998.

[2] R. Geiger, T. Sporer, J. Koller, and K. Brandenburg, "Audio Coding based on Integer Transforms," *111th AES Convention*, December 2001, Preprint 5471.

[3] T. Krishnan and S. Oraintara, "A Fast and Lossless Forward and Inverse Structure for the MDCT in MPEG Audio Coding," *Proc. of IEEE ISCAS*, vol. 2, pp. 181–184, May 2002.

[4] J. Liang and T. D. Tran, "Fast Multiplierless Approximations of the DCT with the Lifting Scheme," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3032–3044, December 2001.

[5] R. Geiger, Y. Yokotani, and G. Schuller, "Improved Integer Transforms for Lossless Audio Coding," *Proc. of the Asilomar Conference*, vol. 2, pp. 2119–2123, November 2003.

[6] R. Geiger, Y. Yokotani, G. Schuller, and J. Herre, "Improved Integer Transforms using Multi-Dimensional Lifting," *Proc. of IEEE ICASSP*, vol. 2, pp. 1005–1008, May 2004.

[7] M. Gerzon and P. G. Craven, "Optimal Noise Shaping and Dither of Digital Signals," *87th AES Convention*, October 1989, Preprint 2822.

[8] A. J. S. Ferreira, "Accurate Estimation in the ODFT Domain of the Frequency, Phase and Magnitude of Stationary Sinusoids," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 47–50, October 2001.

[9] ISO/IEC JTC1/SC29/WG11 Moving Pictures Expert Group, "Final Call for Proposals on MPEG-4 Lossless Audio Coding," October 2002, Shanghai, China, N5208.

[10] T. Qiu, "Lossless Audio Coding Based on High Order Context Modeling," *IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 575–580, 2001.