16th IEEE International Conference on Control Applications
Part of IEEE Multi-conference on Systems and Control
Singapore, 1-3 October 2007

H_2 Optimal Signal Predictive Quantization in Oversampled FBs

Shenpeng Li*, Jingxin Zhang*†, and Cishen Zhang‡

*Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia

‡School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

†Author for correspondences

***Abstract*—This paper is concerned with the design of signal predictor in oversampled signal predictive subband coding. Spectral analysis of the prediction error system is used to derive an H_2 optimization formulation of the design problem, and an LMI based method is obtained to design the H_2 optimal signal predictor that minimizes the variance of quantizer input signals. Simulation examples are presented to demonstrate the obtained design method.**

I. INTRODUCTION

Signal prediction is a technique to reduce the quantization noise in A/D converter and digital coding, and has been studied extensively in different contexts, see eg [1, 9,11,12,14,15] and the references therein. Quantizer, often considered as the encoder, could convert the continuous signal to discrete signal or compress the high rate discrete-time signal to a low rate one. When filter bank (FB) is used for coding, the above procedure is known as subband coding. As a lossy compress technique, the quantization noise leads to signal distortion. To compensate the distortion caused by quantizer, signal prediction method is introduced to the subband coding. The function of signal prediction is to reduce the variance of the quantizer input, which is the variable to be quantized [8]. This technique is particularly useful at low bit rate, which is often required for data transmission over capacity limited communication channels.

In subband coding, it has been shown that for Gaussian sources, the signal prediction yields asymptotically rate-distortion-optimum encoder at high rate [15]. In [1], Bölcskei and Hlawatsch have presented a method to design a cross band predictor that attempts to minimize the variance of prediction error (quantizer input). This method assumes that the power spectral of source signal and quantization noises is known, and solves the design problem by truncated approximation of optimal solution using the linear predictor with finite impulse response. Because truncation is involved, the solution of this method is not optimal, which may result in degraded performance. Another problem with this method is that the order (length) of the FIR predictor, which controls the approximation and hence affects directly the performance, is not analytically predictable and has to be determined by trial and error.

This paper approaches the problem from systems point of view and uses the tools of optimal filtering to analyze the prediction error system and to obtain an H_2 optimization formulation for the design of signal predictor. From this formulation, an LMI based method is then derived to design the H_2 optimal linear predictor with infinite impulse response. The designed predictor minimizes the variance of quantizer input. The order of the predictor is determined directly by the order of the analysis FB and source signal, and therefore needs not to be determined by trial and error. Simulation examples are presented to demonstrate the obtained design method.

II. PRELIMINARY

A. Oversampled FBs and perfect reconstruction

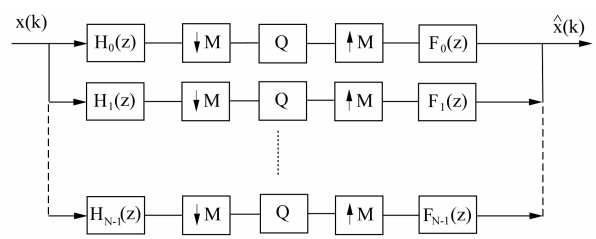

Fig.1 Oversampled FB with subband quantizers

Consider the N-Channel FB with decimator factor M shown in Fig 1. For $N \geq M$ it is called oversampled FB. In the figure, $H_n(z)$ and $F_n(z)$ are the analysis and synthesis filters, respectively, which are linear and bounded input and bounded output (BIBO) stable and in the form

$$H_n(z) = \sum_{m=-\infty}^{m=\infty} h_n(m) z^{-m} \text{ and } F_n(z) = \sum_{m=-\infty}^{m=\infty} f_n(m) z^{-m}$$

where $h_n(m)$ and $f_n(m)$ are the impulse responses of $H_n(z)$ and $F_n(z)$ respectively. Using the polyphase representation [4,13], $H_n(z)$ and $F_n(z)$ can be represented by the polyphase matrices $E(z) \in \mathbb{C}^{N\times M}$ and $R(z) \in \mathbb{C}^{N\times M}$ respectively, which are given by

$$E(z) = \begin{bmatrix} E_{00}(z) & \cdots & E_{0(M-1)}(z) \\ E_{10}(z) & \cdots & E_{1(M-1)}(z) \\ \vdots & \cdots & \vdots \\ E_{(N-1)0}(z) & \cdots & E_{(N-1)(M-1)}(z) \end{bmatrix} \quad (1)$$

$$R(z)=\begin{bmatrix} R_{00}(z) & \cdots & R_{(N-1)0}(z) \\ R_{01}(z) & \cdots & R_{(N-1)1}(z) \\ \vdots & \cdots & \vdots \\ R_{0(M-1)}(z) & \cdots & R_{(N-1)(M-1)}(z) \end{bmatrix} \tag{2}$$

The FB achieves prefect reconstruction (PR) when $\hat{x}(k)=x(k)$, which implies

$$R(z)E(z)=I_N \tag{3}$$

When $N>M$, the synthesis FB $R(z)$s that satisfy (3) for a given $E(z)$ is not unique. The one with minimum H_2 norm is given by [13]

$$R(z)=[\tilde{E}(z)E(z)]^{-1}\tilde{E}(z) \tag{4}$$

which is the left pseudo inverse of $E(z)$ and is in general noncausal. For the properties of the above $R(z)$ and the efficient direct computational method of $R(z)$, see [3].

B. Quantization noise

For a discrete-time signal, each sample value can be considered as an element in a set. The process of mapping a large set of elements into a smaller set is called quantization [7,8]. If each sample is quantized individually, it is called scalar quantization. Fig.2 depicts the structure of a scalar quantization.

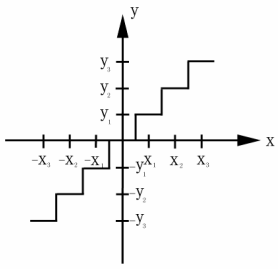

Fig. 2. Scalar Quantizer.

where the set $\{y_i\}$ is called codebook, $\{x_i\}$ is defined as decision levels. As shown in Fig.2, the input sample is scaled by decision levels $\{x_i\}$ and mapped into the codebook $\{y_i\}$. The step size of two adjacent elements in $\{y_i\}$ is defined as quantization interval, which indicates the quantization resolution. The overall number of elements in $\{y_i\}$ is regarded as quantizer length, which represents the bit rate in digital transmission or compression. If the quantization interval is all the same for any two adjacent elements in $\{y_i\}$, the quantizer is called uniform quantizer, which is widely used in practical applications.

Let $n(k)$ be the quantizer input and $\hat{n}(k)$ be the quantizer output. Then the error between input signal $n(k)$ and quantized signal $\hat{n}(k)$ is called quantization error or quantization noise

$$q(k)=n(k)-\hat{n}(k) \tag{5}$$

The variance of quantization noise is defined as

$$\sigma_q^2=E\left\{q(k)^2 \mid q(k)=n(k)-\hat{n}(k)\right\} \tag{6}$$

where $E\{\bullet\}$ is the mathematical expectation.

It is known that the spectrum of quantization error is very complex, but if the input signal $n(k)$ satisfies the uniform distribution, the quantization noise $q(k)$ can be assumed to be a white additive noise to $n(k)$ [7,8].

C. Signal and system norms

Let $\mathrm{x}(k)=\left[x_1(k) \quad x_2(k) \quad \cdots \quad x_n(k)\right]\in\mathbb{R}^m$ be a wide sense stationary (WSS) signal with zero mean. Then its auto-correlation and power spectral density (PSD) functions are given respectively by

$$R_{\mathrm{xx}}(n)=E\left[\mathrm{x}(k)\mathrm{x}^T(k+n)\right] \tag{7}$$

$$S_{\mathrm{xx}}\left(e^{j\omega}\right)=\sum_{n=-\infty}^{\infty}R_{\mathrm{xx}}(n)e^{-jn\omega} \tag{8}$$

and its variance σ_x^2, which equals its power P_x, is given by

$$\begin{aligned}\sigma_x^2&=E\left[\mathrm{x}^T(k)\mathrm{x}(k)\right]=Tr\left\{E\left[\mathrm{x}(k)\mathrm{x}^T(k)\right]\right\}\\&=\frac{1}{2\pi}\int_0^{2\pi}Tr\left[S_{\mathrm{xx}}\left(e^{j\omega}\right)\right]d\omega\\&=Tr\left[R_{\mathrm{xx}}(0)\right]=P_x\end{aligned} \tag{9}$$

If P_x is bounded, then $\mathrm{x}(k)$ is called bounded power signal and a semi-norm called power norm can be defined as

$$\begin{aligned}\|x\|_{\mathcal{P}}&:=\sqrt{E\left[\mathrm{x}(k)\mathrm{x}^T(k)\right]}\\&=\sqrt{\frac{1}{2\pi}\int_0^{2\pi}Tr\left[S_{\mathrm{xx}}\left(e^{j\omega}\right)\right]d\omega}\\&=\sqrt{Tr\left[R_{\mathrm{xx}}(0)\right]}=\sqrt{P_x}\end{aligned} \tag{10}$$

Let $T(z)$ be the transfer function of a stable discrete-time linear system with WSS input $\mathrm{x}(k)\in\mathbb{R}^m$ and output $\mathrm{y}(k)\in\mathbb{R}^n$. If $\mathrm{x}(k)$ is a bounded power signal, then $\mathrm{y}(k)$ is also a bounded power signal and the power norm of $\mathrm{y}(k)$ is given by

$$\|\mathrm{y}\|_{\mathcal{P}}=\|T\mathrm{x}\|_{\mathcal{P}}=\sqrt{\int_0^{2\pi}Tr\left[T\left(e^{j\omega}\right)S_{\mathrm{xx}}\left(e^{j\omega}\right)T^*\left(e^{j\omega}\right)\right]d\omega} \tag{11}$$

If $\mathrm{x}(k)$ is white with unit variance (power), then $S_{\mathrm{xx}}\left(e^{j\omega}\right)=I$ and $\|\mathrm{y}\|_{\mathcal{P}}$ becomes

$$\|\mathrm{y}\|_{\mathcal{P}}=\|T\mathrm{x}\|_{\mathcal{P}}=\sqrt{\int_0^{2\pi}Tr\left[T\left(e^{j\omega}\right)T^*\left(e^{j\omega}\right)\right]d\omega}:=\|T\|_2 \tag{12}$$

The above defined $\|T\|_2$ is called the H_2 norm of the system.

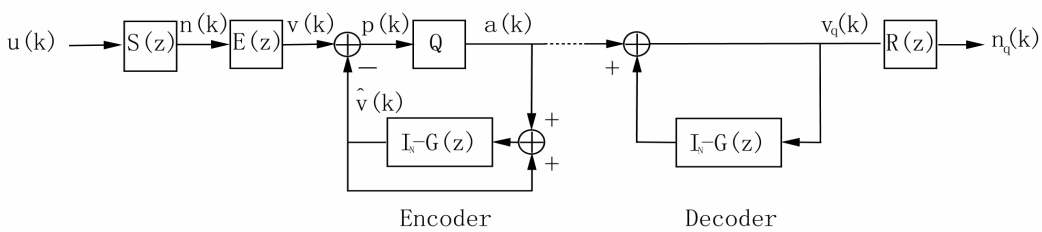

Fig.3 Signal-Predictive subband coder

III. Analysis of Signal Predictive Quantization

Fig.3 depicts a polyphase representation of an oversampled signal predictive subband coder, where $E(z)\in\mathbb{C}^{N\times M}$ and $R(z)\in\mathbb{C}^{M\times N}$ are the polyphase matrices of analysis and synthesis filter bank respectively; Q is an N-channel quantizer; $I_N-G(z)\in\mathbb{C}^{N\times N}$ is a linear strictly causal cross-band predictor; $\mathrm{n}(\mathrm{k})\in\mathbb{R}^M$ and $\mathrm{n_q}(\mathrm{k})\in\mathbb{R}^M$ are the input and reconstructed signals respectively; $\mathrm{v}(\mathrm{k})\in\mathbb{C}^N$ and $\mathrm{v_q}(\mathrm{k})\in\mathbb{C}^N$ are the signal to be encoded and the signal decoded; $\hat{\mathrm{v}}(\mathrm{k})\in\mathbb{C}^N$ is the prediction of $\mathrm{v}(\mathrm{k})$ and $\mathrm{p}(\mathrm{k})\in\mathbb{C}^N$ is the prediction error which is the quantizer input (the variable to be quantized); $\mathrm{a}(\mathrm{k})\in\mathbb{C}^N$ is the quantizer output. The input signal n(k) is assumed to be wide senses stationary with bounded power (variance) and power spectral density $S_{nn}(z)=S(z)S^*(z)$, where $S(z)$ is a stable factorization of $S_{nn}(z)$. Thus, it is represented in Fig 3 as the output of the filter $S(z)$ driven by $\mathrm{u}(\mathrm{k})\in\mathbb{R}^M$, which is assumed to be a Gaussian source. It is assumed that the FB has been designed to satisfy the perfect reconstruction condition $R(z)E(z)=I_N$.

The quantizer Q in the encoder is considered to be a device that injects an additive noise to the quantizer input $\mathrm{p(k)}$. Thus the encoder output can be written as

$$\mathrm{a}(\mathrm{k})=\mathrm{p}(\mathrm{k})+\mathrm{q}(\mathrm{k}) \tag{13}$$

where the quantization noise $\mathrm{q(k)}\in\mathbb{C}^{\mathrm{N}}$ is generally input p(k) dependent. If the quantizer length is fixed, the lower variance of a(k) leads to the lower quantization noise [7,8]. Using the quantizer model given in (13), the prediction error $\mathrm{p}(\mathrm{k})$ can be written as

$$\mathrm{p}(\mathrm{k})=\mathrm{v}(\mathrm{k})-\hat{\mathrm{v}}(\mathrm{k})=G(z)\mathrm{v}(\mathrm{k})-[I_N-G(z)]\mathrm{q}(\mathrm{k}) \tag{14}$$

Quantized by Q, the output signal of encoder is given by

$$\begin{aligned}\mathrm{a}(\mathrm{k})&=\mathrm{p}(\mathrm{k})+\mathrm{q}(\mathrm{k})=\mathrm{v}(\mathrm{k})-\hat{\mathrm{v}}(\mathrm{k})\\&=G(z)\mathrm{v}(\mathrm{k})+G(z)\mathrm{q}(\mathrm{k})\end{aligned} \tag{15}$$

Assume that the linear predictor $G(z)$ is invertible, then the decoder output is given by

$$\mathrm{v_q}(\mathrm{k})=\mathrm{v}(\mathrm{k})+\mathrm{q}(\mathrm{k}) \tag{16}$$

and the reconstructed signal is given by $\mathrm{n_q}(\mathrm{k})=R(z)\mathrm{v_q}(\mathrm{k})$. Define the reconstruction error $\mathrm{e}(\mathrm{k})=\mathrm{n_q}(\mathrm{k})-\mathrm{n}(\mathrm{k})$. Under the PR condition, the reconstruction error becomes

$$\begin{aligned}\mathrm{e}(\mathrm{k})&=\mathrm{n_q}(\mathrm{k})-\mathrm{n}(\mathrm{k})\\&=R(z)E(z)\mathrm{n}(\mathrm{k})+R(z)\mathrm{q}(\mathrm{k})-\mathrm{n}(\mathrm{k})\\&=R(z)\mathrm{q}(\mathrm{k})\end{aligned} \tag{17}$$

The above equation shows that for a perfect reconstruction FB, the reconstruction error $\mathrm{e}(\mathrm{k})$ is determined only by the quantization noise $\mathrm{q}(\mathrm{k})$. Thus, it is possible to reduce the reconstruction error $\mathrm{e}(\mathrm{k})$ by minimizing the quantization noise $\mathrm{q}(\mathrm{k})$.

It is known that [7,8] for a uniform scalar quantizer, the MSE of quantization noise is determined only by the quantization interval. One way to minimize the quantization noise is to decrease the quantization intervals, but this will increase the quantization length. If the quantizer input is a bounded power signal, another way to reduce the quantization noise is to reduce the variance of quantizer input. Because $\mathrm{n}(\mathrm{k})$ is bounded power, so is the quantizer input. The variance of $\mathrm{p}(\mathrm{k})$ can be reduced by designing a linear predictor $I_N-G(z)$ that minimize the power norm of $\mathrm{p}(\mathrm{k})$.

Define $E_S(z)=E(z)S(z)$ and substitute it into (14). Then the prediction error $\mathrm{p}(\mathrm{k})$ can be written as

$$\mathrm{p}(\mathrm{k})=G(z)E_S(z)\mathrm{u}(\mathrm{k})-[I_N-G(z)]\mathrm{q}(\mathrm{k}) \tag{18}$$

Define

$$T_{\mathrm{pw}}(z)=\left[G(z)E_S(z)\quad -\left(I-G(z)\right)\right] \tag{19}$$

and $\mathrm{w}(\mathrm{k})=\left[\mathrm{u}(\mathrm{k})^T\quad \mathrm{q}(\mathrm{k})^T\right]^T$, then (18) can be written as

$$\mathrm{p}(\mathrm{k})=T_{\mathrm{pw}}(z)\mathrm{w}(\mathrm{k}) \tag{20}$$

Let $S_{\mathrm{pp}}(e^{j\omega})$ and $S_{\mathrm{ww}}(e^{j\omega})$ be the PSDs of $\mathrm{p}(\mathrm{k})$ and $\mathrm{w}(\mathrm{k})$, respectively. Then the input and output PSDs are related by

$S_{\mathrm{pp}}\left(e^{j\omega}\right)=T_{\mathrm{pw}}\left(e^{j\omega}\right)S_{\mathrm{ww}}\left(e^{j\omega}\right)T_{\mathrm{pw}}\left(e^{j\omega}\right)^{*}$ and the variance (power) of $\mathrm{p}(\mathrm{k})$ is given by [16,17]

$$\|\mathrm{p}\|_{\mathcal{P}}^{2}=\frac{1}{2\pi}\int_{0}^{2\pi}Tr\left[S_{\mathrm{pp}}\left(e^{j\omega}\right)\right]d\omega=\left\|T_{\mathrm{pw}}\left(e^{jw}\right)\mathrm{w}\right\|_{\mathcal{P}}^{2} \quad (21)$$

If the quantization error q(k) is assumed to be a white noise, the input signal $\mathrm{w}(k)$ in (21) can be regarded as white with unit spectral density $S_{\mathrm{ww}}\left(e^{j\omega}\right)=I_N$. Then (21) becomes

$$\|\mathrm{p}\|_{\mathcal{P}}^{2}=\frac{1}{2\pi}\int_{0}^{2\pi}Tr\left[S_{\mathrm{pp}}\left(e^{j\omega}\right)\right]d\omega=\left\|T_{\mathrm{pw}}\left(e^{jw}\right)\right\|_{2}^{2} \quad (22)$$

From (22) it is clear that the minimization of $\|\mathrm{p}\|_{\mathcal{P}}$ amounts to the minimization of $\left\|T_{\mathrm{pw}}\left(e^{jw}\right)\right\|_{2}$, the H_2 norm of $T_{\mathrm{pw}}(z)$. Thus finding a $G(z)$ that minimizes $\|\mathrm{p}\|_{\mathcal{P}}$ is equivalent to finding a $G(z)$ that minimizes $\left\|T_{\mathrm{pw}}\left(e^{jw}\right)\right\|_{2}$, the H_2 norm of $T_{\mathrm{pw}}(z)$. Because the polyphase matrix $E_S(z)$ in $T_{\mathrm{pw}}(z)$ is generally an IIR transfer matrix, the optimal solution of $G(z)$ should also be IIR in general.

IV. H_2 Optimal Predictor Design

Define the state-space of $G(z)$ as $\left[\begin{array}{c|c}A_G & B_G\\ \hline C_G & D_G\end{array}\right]$. For practical realization, $I_N-G(z)$ must be strictly causal. To guarantee this, $G(z)$ is constrained to be $\left[\begin{array}{c|c}A_G & B_G\\ \hline C_G & I_N\end{array}\right]$. From(18), the signal model $E_S(z)$ is known and the its minimal state-space realization is given by $\left[\begin{array}{c|c}A_{ES} & B_{ES}\\ \hline C_{ES} & D_{ES}\end{array}\right]$. Then, the state-space realization of $T_{\mathrm{pw}}(z)$ can be written as

$$T_{\mathrm{pw}}(z)=\left[\begin{array}{ccc|cc}A_{ES} & 0 & 0 & B_{ES} & 0\\ B_GC_{ES} & A_G & 0 & B_GD_{ES} & 0\\ 0 & 0 & A_G & 0 & B_G\\ \hline C_{ES} & C_G & -C_G & D_{ES} & 0\end{array}\right] \quad (23)$$

Using the similarity transformation matrix

$$T=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix} \quad (24)$$

(23) can be transformed to

$$T_{\mathrm{pw}}(z)=\left[\begin{array}{ccc|cc}A_{ES} & 0 & 0 & B_{ES} & 0\\ B_GC_{ES} & A_G & 0 & B_GD_{ES} & -B_G\\ 0 & 0 & A_G & 0 & B_G\\ \hline C_{ES} & C_G & 0 & D_{ES} & 0\end{array}\right] \quad (25)$$

In (25), the unobservable part can be eliminated to give the minimum realization of $T_{\mathrm{pw}}(z)$ [6,16]

$$T_{\mathrm{pw}}(z)=\left[\begin{array}{c|c}A_T & B_T\\ \hline C_T & D_T\end{array}\right]=\left[\begin{array}{cc|cc}A_{ES} & 0 & B_{ES} & 0\\ B_GC_{ES} & A_G & B_GC_{ES} & -B_G\\ \hline C_{ES} & C_G & D_{ES} & 0\end{array}\right] \quad (26)$$

With (26), the design of H_2 optimal predictor can be converted to the following optimization problem.[2,6]

$$\min_{P,Q,A_T,B_T,C_T,D_T}\left\{Trace[Q]\right\} \quad (27)$$

subject to

$$\begin{bmatrix}P & PA_T & PB_T\\ A_T^TP & P & 0\\ B_T^TP & 0 & I\end{bmatrix}>0 \quad (28)$$

$$\begin{bmatrix}Q & C_T & D_T\\ C_T^T & P & 0\\ D_T^T & 0 & I\end{bmatrix}>0 \quad P=P^T>0 \quad (29)$$

It can be seen from (26) that $\{A_T \quad B_T \quad C_T\}$ contain undetermined variables $\{A_G \quad B_G \quad C_G\}$. Thus, the matrix inequalities (28) and (29) are nonlinear. Following [10], we can linearize (28) and (29) by linear variable change.

Let n be the dimension of A_{ES} and m the dimension of A_G. Define P and P^{-1} as follows

$$P=\begin{bmatrix}X & U\\ U^T & *\end{bmatrix} \quad P^{-1}=\begin{bmatrix}Y & V\\ V^T & *\end{bmatrix} \quad (30)$$

where X and Y are $n\times n$ symmetric matrices, U and V are $n\times m$ matrices, and $*$ denotes an arbitrary element. Define $\mathcal{Z}$

$$\mathcal{Z}=\begin{bmatrix}I & 0\\ Y & V\end{bmatrix} \quad (31)$$

By using congruence transformation on (28) and (29) with $diag(\mathcal{Z} \quad \mathcal{Z} \quad I)$ and $diag(I \quad \mathcal{Z} \quad I)$ respectively, we can get the following matrix inequalities:

$$\begin{bmatrix}\mathcal{Z}P\mathcal{Z}^T & \mathcal{Z}PA_T\mathcal{Z}^T & \mathcal{Z}PB_T\\ \mathcal{Z}A_T^TP\mathcal{Z}^T & \mathcal{Z}P\mathcal{Z}^T & 0\\ B_T^TP\mathcal{Z}^T & 0 & I\end{bmatrix}>0 \quad (32)$$

$$\begin{bmatrix}Q & C_T\mathcal{Z}^T & D_T\\ \mathcal{Z}C_T^T & \mathcal{Z}P\mathcal{Z}^T & 0\\ D_T^T & 0 & I\end{bmatrix}>0 \quad (33)$$

Substituting (26) into (32) and (33) gives

$$\begin{bmatrix}X & I & XA_{ES}+F_1C_{ES} & F_2 & XB_{ES}+F_1D_{ES} & -F_2\\ (\bullet) & Y & A_{ES} & A_{ES}Y & B_{ES} & 0\\ (\bullet) & (\bullet) & X & I & 0 & 0\\ (\bullet) & (\bullet) & (\bullet) & Y & 0 & 0\\ (\bullet) & (\bullet) & (\bullet) & (\bullet) & I & 0\\ (\bullet) & (\bullet) & (\bullet) & (\bullet) & (\bullet) & I\end{bmatrix}>0 \quad (34)$$

$$\begin{bmatrix}Q & C_{ES} & C_{ES}Y+F_3 & D_{ES} & 0\\ (\bullet) & X & I & 0 & 0\\ (\bullet) & (\bullet) & Y & 0 & 0\\ (\bullet) & (\bullet) & (\bullet) & I & 0\\ (\bullet) & (\bullet) & (\bullet) & (\bullet) & I\end{bmatrix}>0 \quad (35)$$

where $(\bullet)$ denotes the transpose of the corresponding elements in the upper triangle part. Three new variables have been introduced in these two matrix inequalities and their definitions are given by

$$F_1 = UB_G$$
$$F_2 = XA_{ES}Y + UB_G C_{ES} Y + UA_G V^T \quad (36)$$
$$F_3 = C_G V^T$$

With (34) and (35), the original optimization problem is converted to

$$\min_{Q,X,Y,F_1,F_2,F_3} \{Trace[Q]\} \quad (37)$$

subject to (34) and (35), which can be solved by the standard LMI algorithm. Using the following equation derived from $PP^{-1} = I$,

$$UV^T = I - XY \quad (38)$$

we can easily get a pair of solution U and V from X and Y. Then, $\{A_G \quad B_G \quad C_G\}$ can be calculated from (36) with the F_1, F_3 and F_2 obtained from the LMI optimization (37).

To guarantee the solvability and the unique solution of (36), the dimension of U and V must be $n \times n$. This means $m = n$, namely, the order of the predictor must be equal to that of $E_S(z)$.

V. Design Examples

Example 1: Considering the following two-channels critical sampling FB,

$$H_0(z) = 0.0022 - 0.0320z^{-1} + 0.0418z^{-2} + 0.4880z^{-3}$$
$$+0.4880z^{-4} + 0.0418z^{-5} - 0.0320z^{-6} + 0.0022z^{-7}$$
$$H_1(z) = 0.0022 + 0.0320z^{-1} + 0.0418z^{-2} - 0.4880z^{-3}$$
$$+0.4880z^{-4} - 0.0418z^{-5} - 0.0320z^{-6} - 0.0022z^{-7}$$

For the above analysis FB, a PR synthesis FB is calculated using the method given in [3]. The input signal to the FB is the AR-4 process given below

$$A(z) = \frac{1}{(1-0.8z^{-1})(1+0.25z^{-1})(1-0.6z^{-1})(1+0.15z^{-1})} \quad (39)$$

Using the H_2 optimization method, the following 6th order IIR optimal predictor $G_{H_2}(z)$ is designed

$$G_{H_2}(z) = \frac{1}{G_D(z)} \begin{bmatrix} G_{00}(z) & G_{01}(z) \\ G_{10}(z) & G_{11}(z) \end{bmatrix} \quad (40)$$

where

$$G_{00}(z) = 1 - 2.1739z^{-1} + 1.8161z^{-2} - 0.7321z^{-3}$$
$$+0.1474z^{-4} - 0.0142z^{-5} + 0.0007z^{-6}$$
$$G_{01}(z) = 0 + 0.1447z^{-1} - 0.2099z^{-2} + 0.1080z^{-3}$$
$$-0.0250z^{-4} + 0.0028z^{-5} - 0.0001z^{-6}$$
$$G_{10}(z) = 0 - 0.0227z^{-1} + 0.0303z^{-2} - 0.0141z^{-3}$$
$$+0.0031z^{-4} - 0.0004z^{-5} + 0.0000z^{-6}$$
$$G_{11}(z) = 1 - 1.4492z^{-1} + 0.7986z^{-2} - 0.2356z^{-3}$$
$$+0.0412z^{-4} - 0.0037z^{-5} + 0.0002z^{-6}$$
$$G_D(z) = 1 - 1.4531z^{-1} + 0.8043z^{-2} - 0.2387z^{-3}$$
$$+0.0420z^{-4} - 0.0038z^{-5} + 0.0002z^{-6}$$

Based on the structure shown in Fig.3, we established a simulation environment in SIMULINK with quantizers inserted between the analysis and synthesis FBs. The input signal was generated by (39) with white noise input. Fig. 4 compares the reconstruction SNR against the quantizer length for $G_{H_2}(z)$ and $G(z) = I_N$ (corresponding to no prediction). The reconstruction SNR is defined as $\|n\|_P^2 / \|e\|_P^2$ and the quantizer interval is fixed at 0.5.

As seen from Fig.4, H_2 optimal predictor significantly improves the SNR at low bit rate (quantizer length), and attains the peak SNR when quantizer length is 18. When the length is 16, which corresponds to a 4-bit quantizer, the SNR for H_2 design is only 1 dB below the peak SNR. In contrast, no prediction coder ($G(z) = I_N$) achieves the same SNR with 5-bit quantizer. Thus H_2 optimal predictor can achieve 1-bit saving in digital coding for this example.

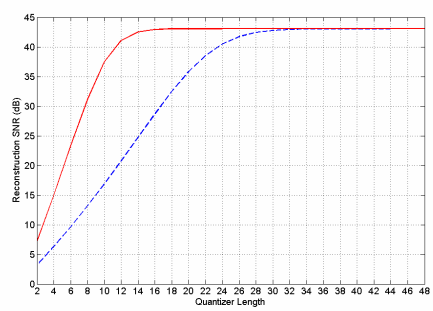

Fig.4 Reconstruction SNR vs. quantizer length with quantization interval=0.5. Solid: $G_{H_2}(z)$, Dashed: $G(z) = I_N$

Fig.5 shows the reconstruction SNR against quantization interval for a 4-bits quantizer. It can be seen that H_2 optimal prediction significantly improves SNR at all the reasonable quantization intervals.

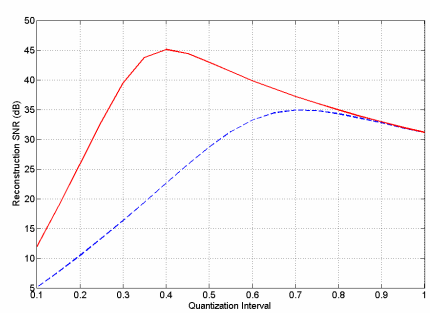

Fig.5 Reconstruction SNR vs. quantizer interval with quantizer length=4 bits. Solid: $G_{H_2}(z)$, Dashed: $G(z) = I_N$

Example 2: An oversampled FB is considered in this example. The signal model is the same as that in Example 1, the FB is a three channel one ($N = 3$) with decimator factor $M = 2$ and

$$H_0(z) = -0.0072 - 0.0126z^{-1} + 0.1183z^{-2} + 0.4015z^{-3}$$
$$+0.4015z^{-4} + 0.1183z^{-5} - 0.0126z^{-6} - 0.0072z^{-7}$$
$$H_1(z) = 0.0246 - 0.0866z^{-1} - 0.2350z^{-2} + 0.3609z^{-3}$$
$$+0.3609z^{-4} - 0.2350z^{-5} - 0.0866z^{-6} - 0.0246z^{-7}$$
$$H_2(z) = -0.0072 + 0.0126z^{-1} + 0.1183z^{-2} - 0.4015z^{-3}$$
$$+0.4015z^{-4} - 0.1183z^{-5} - 0.0126z^{-6} + 0.0072z^{-7}$$

Similar to Fig.4 and Fig.5, Fig.6 compares the reconstruction SNR against the quantizer length for the designed $G_{H_2}(z)$ and no prediction ($G(z) = I_N$). Fig.7 depicts the reconstruction SNR against quantizer interval in oversampled FB when quantizer length is fixed at 4-bits. The same improvement as in critically sampled case can be observed in these two figures.

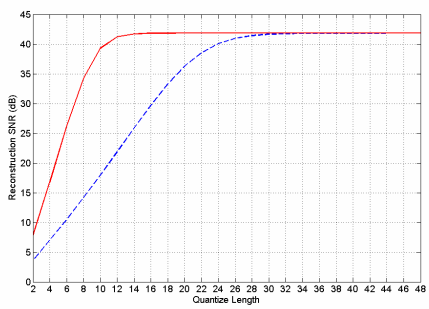

Fig.6 Reconstruction SNR vs. quantizer length in oversampled FB with quantization interval=0.5. Solid: $G_{H_2}(z)$, Dashed: $G(z) = I_N$

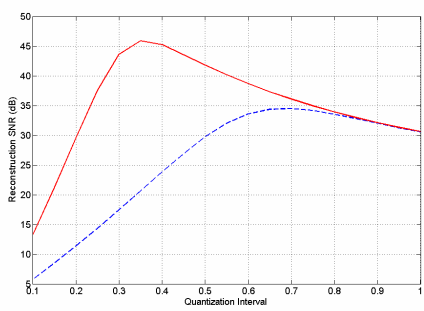

Fig.7 Reconstruction SNR vs. quantizer interval in oversampled FB with quantizer length=4 bits. Solid: $G_{H_2}(z)$, Dashed: $G(z) = I_N$

Example 3: In previous examples, only uniform quantizers are considered in channel coding. In digital coding applications, the optimal quantizers are often nonuniform ones. In this example, we investigate the performance of H_2 predictor in nonuniform quantization.

We replaced the uniform quantizer in Example 1 with a 4-bit nonuniform one, and then compared the SNRs achieved with and without H_2 optimal prediction. The nonunifrom quantizer used is given below, and the comparison results are given in Table 1, which shows the same improvement as in uniform quantizer case.

Nonuniform quantizer details

Decision levels:

$x = \{0 \quad \pm0.1 \quad \pm0.25 \quad \pm0.5 \quad \pm0.85 \quad \pm1.3 \quad \pm1.85 \quad \pm2.5 \quad \pm3.25\}$ C odebook:

$y = \{\pm0.05 \quad \pm0.15 \quad \pm0.35 \quad \pm0.65 \quad \pm1.05 \quad \pm1.55 \quad \pm2.15 \quad \pm2.85\}$

	Reconstruction SNR (dB)
$G_{H_2}(z)$	46.69
$G(z) = I_N$	23.0026

Table 1 Reconstruction SNR in nonuniform 4-bit quantizer

VI. Conclusion

Spectral analysis has been used to derive an H_2 optimization formulation of the design problem, and an LMI based method is obtained to design the H_2 optimal signal predictor with infinite impulse response. The designed predictor minimizes the variance of quantizer input. The order of the predictor is determined directly by the order of the analysis FB and source signal, and therefore needs not to be determined by trial and error. Simulation examples have been presented to demonstrate the obtained design method and to evaluate its performance.

References

[1] H. Bölcskei and F. Hlawatsch, "Noise Reduction in Oversampled Filter Banks Using Predictive Quantization", IEEE Transactions Information Theory, vol. 47, No.1, pp. 155-172, 2001.

[2] S. Boyd, L. El Gaoui, E. Feron and V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory. Society for Industrial and Applied Mathematics, 1994.

[3] L. Chai, J. Zhang, C. Zhang and E. Mosca, "Frame Theory Based Analysis and Design of Oversampled Filter Banks: Direct Computational Method", IEEE Transactions on Signal Processing, vol 55, No. 2, 2007.

[4] R. E. Crochiere and L. R. Rabiner, Multirate Digital Signal Processing. Prentice Hall, 1983

[5] B. De Schutter, "Minimal state-space realization in linear system theory: An overview", Journal of Computational and Applied Mathematics, Special Issue on Numerical Analysis in the 20th Century – Vol. I: Approximation Theory, vol. 121, No.1-2, pp. 331-354, 2000.

[6] P. Gahinet, A. Nemirovski, A.J. Laub and M. Chilali, LMI Control Toolbox for Use with MATLAB. The Mathworks, 1995.

[7] R. M. Gray, "Quantization Noise Spectra", IEEE Transactions on Information Theory, vol. 36, No.6, pp. 1220-1244, 1990.

[8] R. M. Gray and D. L. Neuhoff, "Quantization", IEEE Transactions on Information Theory, vol. 44, No. 6, pp. 2325-2383, 1998.

[9] D. E. Quevedo, G. C. Goodwin and H. Bölcskei, "Multi-step Optimal Quantization in Oversampled Filter Banks", 43rd IEEE Conference on Decision and Control, Atlantis, Paradise Island, Bahamas, December 14-17, 2004.

[10] C. Schere, "From single-channel LMI analysis to multi-channel mixed LMI synthesis: a general procedure." Selected topics in Identification and Control, Vol. &, pp.1-8, Delft University Press, Dec. 1994.

[11] C. L. Tan and T. R. Fischer, "Linear prediction of subband signals", IEEE Journal on Selected Areas in Communication, vol.12, pp. 1576-1589, 1994.

[12] S. K. Tewksbury and B. W. Hallock, "Oversampled, Linear Predictive and Noise-Shaping Coders of Order N>1", IEEE Transactions on Circuits and Systems, vol. CAS-25, No.7, pp. 436-447, 1978.

[13] P. P. Vaidyanathan, Multirate Systems and Filter Banks. Prentice Hall, 1993.

[14] L.Vandendorpe and B. Maison, "Multiple-input/multi-output prediction of subband signals", IEEE Transactions Signal Processing, vol.46, pp. 1230-1233, 1999.

[15] P. W. Wang, Rate distortion efficiency of subband coding with cross band prediction", IEEE Transactions Signal Processing, vol.43, pp. 352-356, 1997.

[16] J. Zhang, C. Zhang and E. Mosca, "Signal Reconstruction in practical Multirate Filter Banks", Proc the 4th Asian Control Conferrence, invited session, pp.1963-1969, Singapore, Sept, 2002.

[17] K. Zhou, J. C. Doyle and K. Glover, Robust and optimal control. Prentice Hall, 1996.