Optimal Zero-Forcing Precoding Design - Oversampled FB Frame Approach

Shengpeng Li*, Jingxin Zhang*, Li Chai†

*Department of Electrical and Computer Systems Engineering, Monash University
Clayton, Vic 3800, Australia. Email:shengpeng.li@eng.monash.edu.au jingxin.zhang@eng.monash.edu.au
†School of Information Science and Engineering, Wuhan University of Science and Technology
Wuhan 430081, P.R. China. Email:eechai@gmail.com

Abstract—This paper presents a frame analysis of Zero Forcing (ZF) precoding problem. It shows that from a frame point of view, the filter bank (FB) channel is a synthesis frame and its canonical dual analysis frame is the optimal ZF precoder under total power constraint. Based on this analysis, a new precoder design is presented. The new design removes the minimum phase assumption of polyphase FB channel in the existing methods and gives rise to optimal noncausal (stable) precoder. It provides a general framework for optimal zero-forcing precoding under total power constraint. The numerical results show that the new design outperforms existing causal stable design.

Index Terms—Zero-forcing precoding, oversampled filter banks, dual frame, state space approach

I. INTRODUCTION

Modern communication systems generally deploy multiple transmitters and receivers to maximize the data transmission rate and improve the reliability of transmission. Such systems are called multiple-input multiple-output (MIMO) system. In MIMO communication systems, inter-symbol interference (ISI) and inter-channel interference (ICI) have been considered as major defects to affect the transmission performance. Precoding is a common solution at the transmitter side to eliminate the ISI and ICI in channels. It can also reduce the complexity of receivers, which is a crucial feature for applications with mobile terminals.

Zero-forcing (ZF) precoding is a common linear precoding design with efficient hardware implementation, thus, it has attracted a great deal of attention, see [1], [2], [3], [4], [5] and the references therein. Based on the diffract model description, most current designs fall into two categories. (i) Constant matrix channel: In this category, the communication channel is characterized by an \( M \times N \) matrix with \( M \leq N \), where \( N \) is the number of channels and \( M \) is the dimension of the received signal. Each entry of matrix is equivalent to the gain between corresponding input and output. This description provides a convenient computational manipulation, but it cannot deal with the dynamics of channel. In [2], it is proved that the pseudo-inverse of communication channel \( H \) has the best performance under total power constraint. (ii) Polyphase FB channel: In this category, each entry of channel model is a filter. Some seminal FB designs were discussed in [3], [4]. In [5], a causal stable design is proposed for minimum phase channel using state-space based computation.

The frame theoretic analysis of oversampled FBs is an efficient tool to analyze the MIMO system [6], [7]. Recently, some efficient state-space based synthesis frame designs have been presented in [8], [9] for signal reconstruction, noise reduction and subband coding problems.

This paper presents a frame analysis of ZF precoding problem. It shows that from a frame point of view, the FB channel is a synthesis frame and its canonical dual analysis frame is the optimal ZF precoder under total power constraint. Based on this analysis, a new precoder design is presented. The new design removes the minimum phase assumption of polyphase FB channel in the existing methods and gives rise to optimal noncausal (stable) precoder. It provides a general framework for optimal zero-forcing precoding under total power constraint and includes some existing results as special cases. The numerical results show that the new design outperforms existing causal stable design.

The notations of the paper are standard. For a matrices \( T \), its transpose, conjugate transpose, trace and eigenvalues are denoted by \( T^T \), \( T^* \), \( T^*T \) and \( \lambda(T) \), respectively. Denote \( T^\dagger = (T^*T)^{-1}T^* \) if \( T \) is full column rank and \( T^\dagger = T^*(T^*T)^{-1} \) if \( T \) is full row rank.

II. PRELIMINARIES

A. PR oversampled FBs

Consider the \( N \)-channel oversampled FB with decimation factor \( M \leq N \) shown in Fig. 1. \( F_k(z) \) and \( H_k(z), \) \( k = 0, 1, \ldots, N-1 \), are analysis and synthesis filters in the form

\[
F_k(z) = \sum_{n=-\infty}^{\infty} f_k[n]z^{-n}, \quad H_k(z) = \sum_{n=-\infty}^{\infty} h_k[n]z^{-n}
\]

where \( f_k[n] \) and \( h_k[n] \) are the impulse response coefficients of \( F_k(z) \) and \( H_k(z) \), respectively. Denote \( F(z) \) and \( H(z) \) the polyphase matrices of analysis FB \( \{F_k(z)\} \) and synthesis FB \( \{H_k(z)\} \), respectively, with \( E_{0i}(z) = \sum_{n=-\infty}^{\infty} h_{i}[nM-j]z^{-n} \) and \( R_{0i}(z) = \sum_{n=-\infty}^{\infty} f_{i}[j-nM]z^{-n} \), for \( i = 0, \ldots, N-1 \) and \( j = 0, \ldots, M-1 \). The analysis and synthesis FBs that satisfy \( H(z)F(z) = I \) are called perfect reconstruction (PR) FB. In communication systems, the transmission channel is modeled by \( H(z) \). It can be seen from the FB

![Fig. 1. Oversampled FB representation of linear precoding with \( N \geq M \)](image-url)
trivial that the polyphase FB model is more general than constant matrix description. If a PR analysis FB $F(z)$ can be found, the ISI will be completely removed. Therefore, $F(z)$ is a zero-forcing precoder in communication system.

**B. State-Space Representation and Computation**

**Definition 1:** For any rational discrete-time transfer matrix $H(z) = \sum_{i=-\infty}^{\infty} H_i z^i \in \mathcal{C}^{M \times N}$. $H(z)$ is called causal if $H_i = 0$ for all $i < 0$. If there exists matrices $A \in \mathcal{C}^{N \times N}$, $B \in \mathcal{C}^{N \times M}$, $C \in \mathcal{C}^{N \times N}$, $D \in \mathcal{C}^{M \times N}$ satisfying $E(z) = D + C(zI - A)^{-1}B$, then $(A, B, C, D)$ is a state-space realization of $H(z)$ and denoted by $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. $H(z)$ is called anticausal if $H_i = 0$ for all $i > 0$. $H(z)$ is called strictly anticausal if $H_i = 0$ for all $i > 0$. If there exists matrices $A \in \mathcal{C}^{n \times n}$, $B \in \mathcal{C}^{n \times M}$, $C \in \mathcal{C}^{N \times n}$, $D \in \mathcal{C}^{M \times N}$ satisfying $E(z) = D + C(z^{-1}I - A)^{-1}B$, then $(A, B, C, D)$ is an anticausal state-space realization of $H(z)$ and denoted by $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{ac}$.

**Definition 2:** For any rational matrix $H(z)$, if it has a causal state-space realization $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ then $\hat{H}(z)$ has an anticausal state-space realization $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{ac}$ where $\hat{H}(z) = H^*(z^{-1})$ and $H^*(z)$ is the conjugate transpose of $H(z)$.

The following lemma from [10] is useful for computing the causal and anticausal cascaded discrete-time systems:

**Lemma 1:** Assume the compatibility of the operations. The product of cascaded causal and anticausal system is given by

\[
\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}_{ac} \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}_{ac} = \begin{bmatrix} A_1 & A_1YB_2 + B_1D_2 \\ C_1 & C_1D_2 + C_1YB_2 \end{bmatrix}_{c} + \begin{bmatrix} A_2 & B_2 \\ D_1C_2 + C_1YB_2 & 0 \end{bmatrix}_{ac},
\]

where $Y$ is given by Sylvester equation $A_1YA_2 - Y + B_1C_2 = 0$.

**C. The norms of signals and systems**

The following definition and lemma come from [11], [12]. Let $x(k)$ be a wide sense stationary (WSS) vector signal with finite power (variance) and power spectrum density (PSD) function $S_{xx}(e^{j\omega})$. The power (semi) norm of $x(k)$ is defined as

\[
\|x\|_p = \sqrt{\text{Tr} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{xx}(e^{j\omega})d\omega \right]}.
\]

Let $y(k)$ and $u(k)$ be the output and input of a system with transfer matrix $H(z)$. Assume that $u(k)$ is WSS with power norm $\|u\|_p$ and that $F(z)$ is stable (but not necessarily causal). Then $y(k)$ and $u(k)$ are related by $S_{yy}(e^{j\omega}) = H(e^{j\omega})S_{uu}(e^{j\omega})H^*(e^{j\omega})$. The $H_2$ (or $\ell_2$) norm of $H(z)$, denoted as $\|H\|_2$, is given by

\[
\|H\|_2 := \sqrt{\text{Tr} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega})H^*(e^{j\omega})d\omega \right]}.
\]

For $H(z)$ with minimal realization $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, $\|H\|_2$ can be computed by

\[
\|H\|_2 = \text{Tr} \left[ CPC^T + DD^T \right]
\]

subject to $APA^T - P + BB^T = 0, P = P^T > 0$.

**Definition 3:** For any rational noncausal $H(z)$, $H(z) = H_c(z) + H_{ac}(z)$, where $H_c(z)$ and $H_{ac}(z)$ are causal and strictly anticausal part of $H(z)$, respectively. If both $H_c(z)$ and $H_{ac}(z)$ are stable, then

\[
\|H\|_2^2 = \|H_c\|_2^2 + \|H_{ac}\|_2^2.
\]

**Lemma 2:** For any rational $H(z) \in \mathcal{C}^{M \times N}$, its pseudoinverse $H^T(z)$ has the minimum $H_2$ norm among all generalized inverses of $H(z)$.

**III. COMPUTATION OF DUAL ANALYSIS FRAME**

The synthesis frame theory and effective computational method have been discussed in [6], [7], [8], [9]. The definition of analysis FB frame is similar to the synthesis frame, but analysis frame is to characterize the right inverse of synthesis FB. The following lemma comes from [7], [8], [9], [11].

**Lemma 3:** Given an analysis FB $\{H_k(z)\}$ with polyphase matrix $H(z) \in \mathcal{C}^{n \times M}$, there exist PR analysis FBS if and only if $\text{rank}[F(e^{j\omega})] = N$ for all $\omega \in [0, 2\pi]$. If this condition holds, then all the PR analysis polyphase matrices $F(z) \in \mathcal{C}^{N \times M}$ are given by

\[
F(z) = F_0(z) + (I_N - F_0(Z)H(z))U(z)
\]

where $U(z)$ is a free parameter of any $N \times M$ stable transfer matrix. $F_0(z)$ is the right inverse of $H(z)$ and is defined as

\[
F_0(z) = \hat{H}(z)(H(z)\hat{H}(z))^{-1}
\]

The $F(z)$ and $F_0(z)$ given in (5) and (6) are called dual frame and canonical dual frame of analysis FBS, respectively.

**Lemma 4:** For $H(z) \in \mathcal{C}^M \times N$ with minimal causal realization $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, assume that

\[
A - e^{j\omega}I = \begin{bmatrix} 1 & B \\ C & D \end{bmatrix}
\]

has full row rank for all $\omega \in [0, 2\pi]$. Then $H(z)$ can be factorized in the form

\[
H(z) = M(z)^{-1}N(z)
\]

where $M(z)$ and $N(z)$ are rational transfer matrices, and $N(z)$ is causal stable satisfying $N(z)N(z) = I$. Moreover, $M(z)$ and $N(z)$ can be obtained from

\[
[M(z) \ N(z)] := \begin{bmatrix} A + LC & L \\ P + C & B + LD \end{bmatrix}
\]

where $P$ and $L$ are given by

\[
P = P^* = DD^* + CXC^*, \quad L = -(AXC^* + BD^*)P^{-1}
\]
and the semi-positive definite $X$ is the unique solution of algebraic Riccati equation

$$AXA^* - X + BB^* + L(CXA^* + DB^*) = 0.$$  \hspace{1cm} (10)

By Lemma 4, $N(z)$ and $M(z)$ are the left coprime factorization of $H(z)$. The following theorem presents an effective computational method to calculate canonical dual frame $F_0(z)$.

**Theorem 1:** Let $H(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ be the polyphase matrix of a synthesis FB satisfying condition A1 in Lemma 4. Let $P$, $L$ and $X$ be as given in (9) and (1). The canonical dual analysis frame $F_0(z)$ is given by

$$F_{0ac}(z) = \left[ \frac{(A + LC)^*}{(B + LD)^*} \right] \left[ \frac{C^*P^{-1}}{P^2} \right] \left[ \frac{A + LC}{B + LD} \right] Y^* \frac{L}{P^2} Y,$$

$$F_{0ac}(z) = \left[ \frac{(A + LC)^*}{(B + LD)^*} \right] \left[ \frac{C^*P^{-1}}{P^2} \right] \left[ \frac{A + LC}{B + LD} \right] Y^* \frac{L}{P^2} Y,$$

where $Y$ is the solution of the following Lyapunov equation

$$Y - (A + LC)^* Y (A + LC) = C^*P^{-1}C.$$

**Proof:** It can be seen from (8) that $N(z) = C^{\ast}P^{-1}C$ is given by

$$N(z) = \begin{bmatrix} A & LC \\ B + LD \\ P^2 \end{bmatrix}.$$

Because $N(z)$ is causal stable and $N(z) = N(z)$ is anticausal stable and given by

$$N(z) = \begin{bmatrix} A & LC \\ B + LD \\ P^2 \end{bmatrix}.$$

Substituting (7) into (6), the canonical dual frame $F_0(z)$ can be factorized as

$$F_0(z) = \hat{H}(z)[H(z)\hat{H}(z)]^{-1} = \hat{N}(z)\hat{M}(z)^{-1}[M^{-1}(z)N(z)\hat{N}(z)\hat{M}(z)M(z)]^{-1} = \hat{N}(z)\hat{M}(z),$$

where $\hat{N}(z)$ has an anticausal state-space realization and $\hat{M}(z)$ has a causal state-space realization. Thus, the state-space realization of cascaded $F_0(z)$ can be obtained by Lemma 1. \Box

**IV. FB FORMULATION OF PRECODING PROBLEM**

Fig. 2 gives a polyphase representation of oversampled $(N \times M)$ MIMO communication system with precoder $F(z)$. The input and received signals are $x(k) \in \mathbb{C}^M$ and $y(k) \in \mathbb{C}^M$, respectively; $w(k) \in \mathbb{C}^N$ is the additive noise in communication channel $H(z) \in \mathbb{C}^{M \times N}$; $F(z) \in \mathbb{C}^{N \times M}$ is the precoder to be designed. The system can be written as

$$y(k) = H(z)F(z)x(k) + w(k).$$

Denote $\hat{x}(k)$ the transmitted signal given by

$$\hat{x}(k) = H(z)F(z)x(k).$$

Assume that the additive noise $w(k)$ is uncorrelated to the transmitted signal $\hat{x}(k)$. Then the quality of the received signal $y(k)$ can be measured by the signal-to-noise ratio (SNR) given by

$$SNR = \frac{\sigma^2_y}{\sigma^2_w}.$$  \hspace{1cm} (17)

It can be seen from (17) that if the additive noise power $\sigma^2_w$ is determinate, it is possible to improve the SNR by increasing the signal power $\sigma^2_y$.

It has been well known in precoding theory that if the channel state information (CSI) of $H(z)$ is known at transmitter side, it is possible to design a precoder $F(z)$ to eliminate the inter-symbol interference (ISI) and inter-channel interference (ICI) with the zero-forcing (ZF) constraint $H(z)F(z) = I$. Under this constraint, the precoder $F(z)$ is a PR analysis FB and the input signal $x(k)$ and transmitted signal $\hat{x}(k)$ are exactly same. Moreover, the ZF constraint can be defined in broad sense

$$H(z)F(z) = \alpha I$$

where $\alpha$ is a positive constant. The $\alpha$ in (18) can be considered as a gain and would not affect all other properties of reconstructed signal $\hat{x}(k)$. It can be seen from (18) and (16) that the SNR of received signal is

$$SNR = \frac{\alpha \sigma^2_y}{\sigma^2_w}.$$  \hspace{1cm} (19)

Define $\alpha$ as the SNR index. It can be seen from (19) that the SNR can be improved by increasing the SNR index $\alpha$. However, due to the power constraint $P_t$ at the transmitter side, the $\alpha$ cannot be increased without limitation. Thus, the precoder optimization problem is

$$\max_{\frac{P_t}{P_z}} \{\alpha\} \hspace{1cm} (20)$$

subject to

$$\sigma^2_y \leq P_t.$$  \hspace{1cm} (21)

Suppose that the input signal $x(k)$ is WSS with known PSD $S_{xx}(z)$. Then the PSD of precoder output signal $b(k)$ is $S_{bb}(z) = F(z)S_{xx}(z)F^*(z)$ and its power norm is given by

$$\|b\|_p^2 = \frac{\alpha \sigma^2_y}{\sigma^2_w} \int_{-\pi}^{\pi} Tr\{F(e^{j\omega})S_{xx}(e^{j\omega})F^*(e^{j\omega})d\omega\}.$$  \hspace{1cm} (21)

From the definition of $H_2$ norm, the transmission power is given by

$$\sigma^2_y = \|b\|_2^2 = \|F\|_2^2 \sigma^2_w \leq P_t.$$  \hspace{1cm} (22)

**Theorem 2:** The system shown in Fig. 2, with total transmitter power constraint $P_t$ and input power $P_z$, achieves the maximum SNR index

$$\alpha = \sqrt{\frac{P_t}{P_z \|F_0(z)\|^2_2}},$$

where $F_0(z)$ is the canonical dual frame analysis FB given in (6) and the optimal precoder is

$$F(z) = \alpha F_0(z).$$  \hspace{1cm} (24)
Proof: It is known that the canonical dual frame $F_0(z)$ is a para-pseudo inverse of $H(z)$ [6], [7]. Thus, using Lemma 2, the system has the minimum transmission power $\bar{P}_m = \left\| F_0(z) \right\|^2 P_z$ subject to ZF condition $H(z) F(z) = I$. From $\left\| F_0(z) \right\|^2 P_z = P_m a^2 \leq P_t$, when equality is achieved, the system has the best SNR index $\alpha$ and the optimal precoder is $F_0(z)$. \hfill \Box

V. NUMERICAL EXAMPLE AND DISCUSSION

Example 1. Consider a redundant channel with channel $N = 3$ and output $M = 2$, its polyphase model $H(z)$ is given by

$$H(z) = \begin{bmatrix}
-0.0044 + 0.4235z^{-1} + 0.0808z^{-2} \\
-0.0373 + 0.4710z^{-1} - 0.1987z^{-2} \\
-0.0044 + 0.4235z^{-1} + 0.0808z^{-2} \\
0.0808 + 0.4235z^{-1} - 0.0044z^{-2} \\
-0.1987 + 0.4710z^{-1} - 0.0373z^{-2} \\
-0.0808 - 0.4235z^{-1} + 0.0044z^{-2}
\end{bmatrix}^T.
$$

By Theorem 2, the dual frame of $H(z)$ is a non-causal polyphase polyphase FB and is given by the causal part

$$F_0(z) = \begin{bmatrix}
-0.0781z^{-2} + 0.0251z^{-2} - 0.00207z^{-3} \\
0.0526z^{-3} + 0.00136z^{-2} - 0.00158z^{-6} \\
0.00121z^{-3} + 0.0173z^{-2} + 0.00397z^{-6} \\
0.00121z^{-3} + 0.0173z^{-2} + 0.00397z^{-6} \\
0.3449z^{-2} - 0.0447z^{-2} - 0.005z^{-6} \\
0.3449z^{-2} - 0.0447z^{-2} - 0.005z^{-6}
\end{bmatrix}.
$$

and anti-causal part

$$F_{0ac}(z) = \begin{bmatrix}
0.5736z^{-2} + 0.3459z^{-2} - 0.0569z^{-1} + 0.00228z^{-3} \\
0.5736z^{-2} + 0.3459z^{-2} - 0.0569z^{-1} + 0.00228z^{-3} \\
0.5099z^{-3} - 0.1375z^{-2} - 0.0443z^{-6} + 0.004508z^{-1} \\
0.5099z^{-3} - 0.1375z^{-2} - 0.0443z^{-6} + 0.004508z^{-1} \\
0.5736z^{-2} - 0.0771z^{-2} + 0.0157z^{-1} - 0.004429z^{-3} \\
0.5736z^{-2} - 0.0771z^{-2} + 0.0157z^{-1} - 0.004429z^{-3} \\
0.5099z^{-3} + 0.0511z^{-2} - 0.00836z^{-6} - 0.000328z^{-1} \\
0.5099z^{-3} + 0.0511z^{-2} - 0.00836z^{-6} - 0.000328z^{-1}
\end{bmatrix}.
$$

If the input power is 0.5watt and the total power of precoder is bounded by 1.5watt, then, by Theorem 2, the maximum SNR index $\alpha$ is given by

$$\alpha = \sqrt{\frac{1.5}{0.5(\left\| F_0(z) \right\|^2 + \left\| F_{0ac}(z) \right\|^2)}} = \sqrt{\frac{1.5}{0.5 \times 0.40570}} = 0.8599
$$

and the optimal precoder is given by $F(z) = 0.8599 F_0(z)$

Example 2. Consider the example used in [5] with

$$H(z) = \frac{1}{9} \begin{bmatrix}
1 + 2z^{-1} & 2.5z^{-1} & 2z^{-1} + z^{-2} \\
2 + z^{-1} & 1 + 2z^{-1} & 2.5z^{-1} \\
2.5 & 2z^{-1} & 1 + 2z^{-1}
\end{bmatrix}
$$

and

$$R(z) = \frac{1}{9} \begin{bmatrix}
16 & 56 & 22 \\
68 & 40 & 3 \\
96 & 50 & 2
\end{bmatrix}^T
$$

A precoder $F(z)$ satisfying $H(z) F(z) F^\dagger(z) = I$ is required. In [5], a causal stable design with $\|F\|_2 = 1.2292$ was obtained. By Theorem 1, we can find a pseudo inverse noncausal stable precoder $F(z)$ with minimum $\|F\|_2 = 1.0978$ satisfying ZF constraint. As shown in Theorem 1, the noncausal precoder can be decomposed into causal stable part and anticausal stable part, thus it is possible to implement it with current microprocessor technique. Therefore, the method presented in this paper provides a more general and better solution for zero-forcing precoding design.

VI. CONCLUSION

This paper has presented a general framework for zero-forcing precoding design of MIMO channel. It has made the following contributions: Firstly, The channel is modeled by polyphase FB, which includes more complicated channels than traditional matrix description. Secondly, with frame theoretic analysis, the minimum phase constraint on channel model is removed. Finally, it has shown that the canonical dual analysis frame has the minimum $H_2$ norm and outperforms some current FB precoder design. Most of the current designs can only achieve one or two aspects of these advantages.

REFERENCES