2004 IEEE 5th Workshop on Signal Processing Advances in Wireless Communications

# Canonical Coordinate Geometry of Precoder and Equalizer Designs for Multichannel Communication [1]

Zhifei Fan
Electrical and Computer Engr
Colorado State University
Fort Collins, CO. 80523
Email: fanz@engr.colostate.edu

Louis Scharf
Electrical and Computer Engr and Statistics
Colorado State University
Fort Collins, CO. 80523
Email: scharf@engr.colostate.edu

Timothy N. Davidson
Electrical and Computer Engr
McMaster University
Hamilton, Ontario, L8S 4K1
Canada
Email: davidson@mcmaster.ca

*Abstract* — **For vector communication over a matrix channel, precoders and equalizers are now known for designs that minimize MSE, maximize SNR, minimize BER, or maximize information rate, under a power constraint [1] - [5]. Some of these designs may be made to be zero-ISI. For reduced-rank signal processing, designs are known for estimation at minimum MSE and maximum information rate [6] - [10]. These signal processing designs show that reduced-rank estimation must be done in a system of canonical coordinates. What connection is there, then, between reduced-rank signal processing and precoder/equalizer design under a power constraint? In this paper we show that all known designs for precoder/equalizer design are, in fact, decompositions of a virtual two-channel problem into a system of canonical coordinates, wherein whitened variables in the canonical message channel are correlated only pairwise with whitened variables in the canonical measurement channel. This finding clarifies the geometry of known precoder/equalizer designs and illustrates that these designs decompose the two-channel communication problem into the Shannon channel [9], where its geometry is revealed.**

## I. Introduction

We know from [3], [4], [13], [14] that when a SISO channel is an FIR filter or a MIMO channel is a matrix of FIR filters, source block coding may be used to convert the FIR filers to matrix maps. Such block or matrix transmission schemes include multi-carrier MIMO systems and discrete multitone modulation. In these transmission systems, the channel state information is assumed known to both transmitter and receiver. Then for specific signal constellations and coding schemes, joint linear precoder and equalizer designs are considered in [1] - [5] to further optimize the system and exploit joint transmit-receive diversity. The criteria for which designs are known are maximum SNR, maximum information rate, minimum MSE and minimum BER, under a power constraint.

Canonical correlations measure cosines of principal angles between random vectors. The principal angles between the message and measurement determine error covariance, information rate, and capacity. For reduced-rank and/or quantized estimation of one vector from another, designs are known for reduced rank estimation at minimum MSE and maximum information rate [6] - [10]. These results show that reduced-rank estimation must be done in a system of canonical coordinates.

[1]This work was supported by the National Science Foundation under Grant CCR-0112573.

In this paper we show that all known results for precoder/equalizer design are, in fact, decompositions of a virtual two-channel problem into a system of canonical coordinates, wherein whitened variables in the canonical message channel are correlated only pairwise with whitened variables in the canonical measurement channel. This finding clarifies the geometry of known precoder/equalizer designs and illustrates that these designs decompose the two-channel communication problem into the Shannon channel [9], where its geometry is revealed.

## II. Precoder and equalizer designs based on SNR criteria and Canonical Coordinate Geometry

The block or matrix transmission system is illustrated in Figure 1, where $u \in \mathbb{C}^m$ is the source vector with covariance $R_{uu}$, $x \in \mathbb{C}^t$ is the transmit vector, $n \in \mathbb{C}^r$ is proper additive Gaussian noise with covariance $R_{nn}$, $y \in \mathbb{C}^r$ is the received vector, $v \in \mathbb{C}^n$ is the vector after equalization, $G$ is a $t \times m$ precoder matrix, $H$ is an $r \times t$ channel matrix, and $F$ is an $n \times r$ equalizer matrix. Generally we assume that $t \geq m$, and $r \geq n$.

We define the SNR matrix after precoding and equalization as

$$S(G,F) := R_{uu}^{H/2}(FHG)^H(FR_{nn}F^H)^{-1}FHGR_{uu}^{1/2},$$

where the inverse term is the inverse noise covariance and the terms on either side are symmetric factors of the signal covariance. The coherence matrix after precoding and equalization is [9]

$$\begin{aligned} C(G,F) &:= E(R_{uu}^{-1/2}u)(R_{vv}^{-1/2}v)^H \\ & R_{uu}^{H/2}(FHG)^H(FR_{nn}F^H)^{-H/2}\Big((FR_{nn}F^H)^{-1/2}FHGR_{uu} \\ & \qquad (FHG)^H(FR_{nn}F^H)^{-H/2}+I\Big)^{-H/2}. \end{aligned}$$

The coherence matrix is the cross correlation between whitened versions of the message $u$ and measurement after equalization $v$. The fundamental identity between $S(G,F)$ and $C(G,F)C^H(G,F)$ is [9]

$$I - C(G,F)C^H(G,F) = (I + S(G,F))^{-1}.$$

The SNR matrix can be factored as [3],

$$\begin{aligned} S(G,F) &:= R_{uu}^{H/2}G^HH^HF^H(FR_{nn}F^H)^{-1}FHGR_{uu}^{1/2} \\ &= R_{uu}^{H/2}G^HH^HR_{nn}^{-1/2}\Pi_{R_{nn}^{1/2}F^H}R_{nn}^{-1/2}HGR_{uu}^{1/2} \\ &\leq R_{uu}^{H/2}G^HH^HR_{nn}^{-1}HGR_{uu}^{1/2} := S \end{aligned}$$

where $\Pi_{R_{nn}^{1/2}F^H} = R_{nn}^{1/2}F^H(FR_{nn}F^H)^{-1}FR_{nn}^{1/2} \leq I$. Thus $S$ is an upper bound of $S(G,F)$. The equality holds when $F =$

$\Gamma R_{uu}^{1/2} G^H H^H R_{nn}^{-1}$, where $\Gamma$ is an arbitrary $m \times m$ invertible matrix [3]. Then, with this optimal $F$ the dimension $n$ of $v$ is the same as the dimension $m$ of $u$. From the bound on $S(G, F)$ it follows that

$$C(G,F)C^H(G,F) \leq CC^H,$$

where

$$R_{uu}^{H/2} G^H H^H R_{nn}^{-1/2} (R_{nn}^{-1/2} H G R_{uu} G^H H^H R_{nn}^{-1/2} + I)^{-1/2} := C.$$

The identity between $S$ and $CC^H$ is

$$I - CC^H = (I + S)^{-1}. \quad (1)$$

Next, we introduce the eigenvalue decomposition

$$H^H R_{nn}^{-1} H = V \Lambda V^H.$$

Here $V$ is a $t \times t$ unitary matrix, and $\Lambda$ is a $t \times t$ diagonal matrix with diagonal elements in decreasing order. W.l.o.g., assume $G = V \Phi R_{uu}^{-1/2}$ where $\Phi$ is an arbitrary $t \times m$ matrix. Then the bounding SNR matrix and coherence matrix may be written as

$$S = \Phi^H \Lambda \Phi, \quad (2)$$

$$C = \Phi^H \Lambda^{1/2} (\Lambda^{1/2} \Phi \Phi^H \Lambda^{1/2} + I)^{-1/2}. \quad (3)$$

We call any criterion that only depends on $S$ an SNR-based criterion. It's easy to see that any SNR-based criterion is also a $CC^H$-based, or Coherence-based criterion.

The power constraints will take one of the following forms:

$$\mathrm{tr}(G R_{uu} G^H) = \mathrm{tr}(\Phi \Phi^H) \leq \mathcal{P}, \quad (4)$$

$$\mathrm{ev}_{\max}(G R_{uu} G^H) = \mathrm{ev}_{\max}(\Phi \Phi^H) \leq \mathscr{P}. \quad (5)$$

Here $ev_{max}(A)$ is the maximum eigenvalue of the Hermitian matrix $A$. For all SNR-based criteria, the optimal precoder $G = V \Phi R_{uu}^{-1/2}$ will use a matrix $\Phi$ that is a $t \times m$ diagonal matrix with nonzero elements only on its main diagonal. Furthermore, the SNR matrix $S$, and the squared coherence matrix $CC^H$ will be diagonal. It is the connection between $S$ and $CC^H$ in (1) that brings the geometry we are after.

First, we'll prove this lemma from [4], which is useful for the optimizations that follow.

**Lemma 1** *If the optimal $S = \Phi^H \Lambda \Phi$ is diagonal with diagonal elements in decreasing order, then the optimal $\Phi$ will have nonzero elements only on its main diagonal.*

Proof: If $S$ is diagonal, then $\Phi$ has the representation $\Phi = \Lambda^{-1/2} Q \Sigma$, where $Q$ is an arbitrary $t \times t$ unitary matrix and $\Sigma$ is a $t \times m$ diagonal matrix with nonzero elements only on its main diagonal, and $\Sigma^H \Sigma = S$. The power constraint in (4) is then $\mathrm{tr}(\Lambda^{-1} Q \Sigma \Sigma^H Q^H) \leq \mathcal{P}$. By Theorem 9.H.1.h. in [12], $\mathrm{tr}(\Lambda^{-1} Q \Sigma \Sigma^H Q^H) \geq \sum_{i=1}^{m} \lambda_i^{-1} \mathrm{ev}_i(Q \Sigma \Sigma^H Q^H)$, where, $\mathrm{ev}_i(A)$ is the $i$th eigenvalue of the Hermitian matrix $A$ and eigenvalues are sorted in decreasing order. The equality holds when $Q \Sigma \Sigma^H Q^H$ is diagonal with diagonal elements in decreasing order. Since $\Sigma^H \Sigma = S$, $\Sigma \Sigma^H$ is diagonal with diagonal elements in decreasing order. So, we can choose $Q = I$ to get the lower transmit power. Similarly, the power constraint in (5) is $\mathrm{ev}_{max}(\Lambda^{-1} Q \Sigma \Sigma^H Q^H) \leq \mathscr{P}$. By Theorem 9.H.1.i. in [12], $\mathrm{ev}_{\max}(\Lambda^{-1} Q \Sigma \Sigma^H Q^H) \geq \lambda_{\max}^{-1} \mathrm{ev}_{\max}(Q \Sigma \Sigma^H Q^H)$. So, in this case, we can also choose $Q = I$ to lower the transmit power. Then, the optimal $\Phi = \Lambda^{-1/2} \Sigma$ has nonzero elements only on its main diagonal. [1]

**Weighted SNR Design** [4]. The objective function for maximizing the weighed SNR is

$$\max \mathrm{tr}(WS)$$

Here $W$ is a diagonal weighting matrix, and w.l.o.g., we assume the diagonal elements of $W$ in decreasing order. By Theorem 9.H.1.g. in [12],

$$\mathrm{tr}(WS) \leq \sum_{i=1}^{m} w_i \mathrm{ev}_i(S),$$

where $\mathrm{ev}_i(S)$ is the $i$th eigenvalue of $S$ and the eigenvalues are sorted in decreasing order. The equality holds when $S = \Phi^H \Lambda \Phi$ is diagonal with diagonal elements in decreasing order. Note that the optimal $CC^H$ is also diagonal with diagonal elements in decreasing order. So, according to Lemma 1, the optimal $\Phi$ is diagonal. Then the optimization problem is simplified to a scalar optimization problem and the solution is easy to derive using standard optimization procedures [4]. Under constraint (4) the design is

$$|\phi_1|^2 = \mathcal{P}, |\phi_2|^2 = \cdots = |\phi_m|^2 = 0.$$

Here $\phi_i$ is the $i$th diagonal element of $\Phi$. The optimal weighted SNR is

$$\mathrm{tr}(WS) = w_1 \lambda_1 \mathcal{P}.$$

Note that in this case, the channel only uses one subchannel. Only one symbol is transmitted at one time. Under constraint(5),

$$|\phi_1|^2 = \cdots = |\phi_m|^2 = \mathscr{P}.$$

The optimal weighted SNR is then

$$\mathrm{tr}(WS) = \mathscr{P} \sum_{i=1}^{m} w_i \lambda_i.$$

**Max Info Rate Design** [2] - [4]. The information rate between $u$ and $v$ is

$$\begin{aligned} I(u;v) &= \log\det(I + S) \\ &= \log\det(I + \Phi^H \Lambda \Phi) \\ &= \log\det(I + \Lambda^{1/2} \Phi \Phi^H \Lambda^{1/2}). \end{aligned}$$

Define $\hat{S} := \Lambda^{1/2} \Phi \Phi^H \Lambda^{1/2}$, so, $\Phi \Phi^H = \Lambda^{-1/2} \hat{S} \Lambda^{-1/2}$. Then the power constraint is $\mathrm{tr}(\Lambda^{-1} \hat{S}) \leq \mathcal{P}$, or $\mathrm{ev}_{\max}(\Lambda^{-1} \hat{S}) \leq \mathscr{P}$. According to Lemma 1, we know the optimal $\hat{S}$ should be diagonal with diagonal elements in decreasing order, which means the optimal $\Phi \Phi^H$ should be diagonal. W.l.o.g., the optimal $\Phi$ can be diagonal. Then $S$ and $CC^H$ are diagonal. Standard optimization procedures may be used to solve the simplified scalar optimization problem [4], [3]. Under constraint (4) the design is [4]

$$\begin{aligned} |\phi_i|^2 &= (\frac{1}{\mu} - \frac{1}{\lambda_i})^+, \\ \mu &= \frac{1}{\frac{1}{m}\mathcal{P} + \frac{1}{m}\sum_{i=1}^{m} \frac{1}{\lambda_i}}, \end{aligned}$$

[1]In this paper, if matrix $A$ is not full rank, $A^{-1}$ is the pseudo inverse of $A$.

where $(x)^+ = \max(x, 0)$, $\bar{m}$ is the number of the subchannels where $|\phi_i|^2 > 0$ and $\lambda_i$ are diagonal elements of $\boldsymbol{\Lambda}$. The maximum mutual information is

$$I(\boldsymbol{u}; \boldsymbol{v}) = \sum_{i=1}^{m} \max(\log(\frac{\lambda_i}{\mu}), 0).$$

Under constraint (5) the design is [3],

$$|\phi_1|^2 = \cdots = |\phi_m|^2 = \mathscr{P},$$

and the optimal mutual information is

$$I(\boldsymbol{u}; \boldsymbol{v}) = \sum_{i=1}^{m} \log(1 + \mathscr{P}\lambda_i).$$

So, in summary, under these SNR-based criteria, the optimal precoder is $\boldsymbol{G} = \boldsymbol{V}\boldsymbol{\Phi}\boldsymbol{R}_{uu}^{-1/2}$, where $\boldsymbol{\Phi}$ has nonzero elements only on its main diagonal elements. $\boldsymbol{S}$ is diagonal with diagonal elements $|\phi_i|^2\lambda_i$ [2] - [4], and $\boldsymbol{C}\boldsymbol{C}^H$ is diagonal with diagonal elements $|\phi_i|^2/(|\phi_i|^2 + \frac{1}{\lambda_i})$.

**Canonical Coordinates and Geometry**. W.l.o.g., $\boldsymbol{C}$ in (3) is diagonal. This means the combined effects of $\boldsymbol{G}$ and $\boldsymbol{F}$ have been to transform the whitened source $\hat{\boldsymbol{u}} = \boldsymbol{R}_{uu}^{-1/2}\boldsymbol{u}$ and equalized and whitened measurement $\hat{\boldsymbol{v}} = \boldsymbol{R}_{vv}^{-1/2}\boldsymbol{v}$ into a system of canonical coordinate where variables are pairwise correlated with correlations

$$\begin{aligned} |E\hat{u}_i\bar{\hat{v}}_i|^2 &= (\boldsymbol{C}\boldsymbol{C}^H)_i \\ &= \frac{|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}. \end{aligned}$$

Because $\hat{u}_i$ ($i$th element of $\hat{\boldsymbol{u}}$) and $\hat{v}_i$ ($i$th element of $\hat{\boldsymbol{v}}$) have unit variance, we may call these cosine-squared of the angle between $\hat{u}_i$ and $\hat{v}_i$. It is as if the whitened source $\hat{\boldsymbol{u}}$ is communicating over $\bar{m}$ uncorrelated channels, where $\bar{m}$ is the number of channels with $|\phi_i| > 0$. In each such channel, the SNR and squared canonical correlations are

$$\mathrm{SNR}_i = \frac{|\phi_i|^2}{\frac{1}{\lambda_i}},$$

$$(\boldsymbol{C}\boldsymbol{C}^H)_i = \frac{|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}.$$

The geometry of each channel is illustrated in Figure 2. In this Pythagorean decomposition of a subchannel, $|\phi_i|^2$ is the message power, $\frac{1}{\lambda_i}$ the noise power, and $\frac{|\phi_i|^2}{|\phi_i|^2+\frac{1}{\lambda_i}}$ is the cosine-squared of the angle between equalized and whitened measurement and whitened message.

## III. Precoder and equalizer designs based on a minimum mean square error (MSE) criterion and Canonical Coordinate Geometry

We define the error covariance matrix after precoding and equalization as

$$\begin{aligned} \boldsymbol{M}(\boldsymbol{G}, \boldsymbol{F}) &:= E(\boldsymbol{F}\boldsymbol{H}\boldsymbol{G}\boldsymbol{u} + \boldsymbol{F}\boldsymbol{n} - \boldsymbol{u})(\boldsymbol{F}\boldsymbol{H}\boldsymbol{G}\boldsymbol{u} + \boldsymbol{F}\boldsymbol{n} - \boldsymbol{u})^H \\ &= (\boldsymbol{F}\boldsymbol{H}\boldsymbol{G} - \boldsymbol{I})\boldsymbol{R}_{uu}(\boldsymbol{F}\boldsymbol{H}\boldsymbol{G} - \boldsymbol{I})^H + \boldsymbol{F}\boldsymbol{R}_{nn}\boldsymbol{F}^H, \end{aligned}$$

and the half coherence matrix after precoding and equalization as the cross covariance between $\boldsymbol{u}$ and whitened $\boldsymbol{v}$ [6]:

$$\begin{aligned} &\boldsymbol{D}(\boldsymbol{G}, \boldsymbol{F}) := E\boldsymbol{u}(\boldsymbol{R}_{vv}^{-1/2}\boldsymbol{v})^H = \\ &\boldsymbol{R}_{uu}^H(\boldsymbol{F}\boldsymbol{H}\boldsymbol{G})^H(\boldsymbol{F}\boldsymbol{R}_{nn}\boldsymbol{F}^H)^{-1/2}\Big((\boldsymbol{F}\boldsymbol{R}_{nn}\boldsymbol{F}^H)^{-1/2}\boldsymbol{F}\boldsymbol{H}\boldsymbol{G}\boldsymbol{R}_{uu} \\ &\qquad (\boldsymbol{F}\boldsymbol{H}\boldsymbol{G})^H(\boldsymbol{F}\boldsymbol{R}_{nn}\boldsymbol{F}^H)^{-1/2} + \boldsymbol{I}\Big)^{-1/2}. \end{aligned}$$

Then the fundamental identity between $\boldsymbol{M}(\boldsymbol{G}, \boldsymbol{F})$ and $\boldsymbol{D}(\boldsymbol{G}, \boldsymbol{F})$ is [6]

$$\boldsymbol{R}_{uu} - \boldsymbol{D}(\boldsymbol{G}, \boldsymbol{F})\boldsymbol{D}^H(\boldsymbol{G}, \boldsymbol{F}) = \boldsymbol{M}(\boldsymbol{G}, \boldsymbol{F}).$$

The objective function for minimum MSE is

$$\min \mathrm{tr}\boldsymbol{M}(\boldsymbol{G}, \boldsymbol{F}).$$

From [1] [4], we know that the optimum equalizer is the Wiener filter

$$\boldsymbol{F} = \boldsymbol{R}_{uu}\boldsymbol{G}^H\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{G}\boldsymbol{R}_{uu}\boldsymbol{G}^H\boldsymbol{H}^H + \boldsymbol{R}_{nn})^{-1}.$$

With this choice, $\boldsymbol{M}(\boldsymbol{G}, \boldsymbol{F})$ is

$$\begin{aligned} &\boldsymbol{R}_{uu}^{1/2}(\boldsymbol{I} + \boldsymbol{R}_{uu}^{1/2}\boldsymbol{G}^H\boldsymbol{H}^H\boldsymbol{R}_{nn}^{-1}\boldsymbol{H}\boldsymbol{G}\boldsymbol{R}_{uu}^{1/2})^{-1}\boldsymbol{R}_{uu}^{1/2} \\ = \;&\boldsymbol{R}_{uu}^{1/2}(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{R}_{uu}^{1/2} := \boldsymbol{M}, \end{aligned} \tag{6}$$

and the half coherence matrix is

$$\begin{aligned} &\boldsymbol{R}_{uu}^H\boldsymbol{G}^H\boldsymbol{H}^H\boldsymbol{R}_{nn}^{-1/2}(\boldsymbol{R}_{nn}^{-1/2}\boldsymbol{H}\boldsymbol{G}\boldsymbol{R}_{uu}\boldsymbol{G}^H\boldsymbol{H}^H\boldsymbol{R}_{nn}^{-1/2} + \boldsymbol{I})^{-1/2} \\ = \;&\boldsymbol{R}_{uu}^{1/2}\boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2}(\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Phi}\boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2} + \boldsymbol{I})^{-1/2} := \boldsymbol{D}. \end{aligned} \tag{7}$$

The identity between $\boldsymbol{M}$ and $\boldsymbol{D}$ is still

$$\boldsymbol{R}_{uu} - \boldsymbol{D}\boldsymbol{D}^H = \boldsymbol{M}.$$

For minimum mean squared error under power constraints (4) or (5), the optimal precoder will be $\boldsymbol{G} = \boldsymbol{V}\boldsymbol{\Phi}\boldsymbol{U}^H\boldsymbol{R}_{uu}^{-1/2}$, where $\boldsymbol{V}$ is a $t \times t$ unitary matrix, $\boldsymbol{U}$ is an $m \times m$ unitary matrix, and $\boldsymbol{\Phi}$ is a $t \times m$ diagonal matrix with nonzero elements only on its main diagonal. (The notation here conflicts with the earlier notation $\boldsymbol{G} = \boldsymbol{V}\boldsymbol{\Phi}\boldsymbol{R}_{uu}^{-1/2}$ just above equation (2). But rather than introduce a new symbol, we take $\boldsymbol{\Phi}\boldsymbol{U}^H$ to be a replacement for $\boldsymbol{\Phi}$.) Furthermore, the MSE matrix $\boldsymbol{M}$ and the squared half coherence matrix $\boldsymbol{D}\boldsymbol{D}^H$ will have the form $\boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^H$ where $\boldsymbol{T}$ is diagonal. It is the connection between $\boldsymbol{M}$ and $\boldsymbol{D}\boldsymbol{D}^H$ that will bring the geometry we are after.

First, we'll optimize $\mathrm{tr}\boldsymbol{M}$ under power constraints. Let us introduce the singular value decomposition

$$\boldsymbol{R}_{uu} = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{U}^H,$$

where $\boldsymbol{U}$ is an $m \times m$ unitary matrix, and $\boldsymbol{\Delta}$ is an $m \times m$ diagonal matrix with diagonal elements $\delta_i$ in decreasing order. So the objective function $\mathrm{tr}\boldsymbol{M}$ can be written as

$$\mathrm{tr}(\boldsymbol{\Delta}\boldsymbol{U}^H(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{U}).$$

By Theorem 9.H.1.h. in [12], $\mathrm{tr}(\boldsymbol{\Delta}\boldsymbol{U}^H(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{U}) \geq \sum_i \delta_i \mathrm{ev}_{m-i+1}(\boldsymbol{U}^H(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{U})$, where $\mathrm{ev}_{m-i+1}(\boldsymbol{A})$ is the $(m-i+1)$th eigenvalue of Hermitian matrix $\boldsymbol{A}$, with eigenvalues sorted in decreasing order. The equality holds when $\boldsymbol{U}^H(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{U}$ is diagonal with diagonal elements in increasing order. Suppose the optimal $\boldsymbol{U}^H(\boldsymbol{I} + \boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{U} = \boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is diagonal with diagonal elements in increasing order. Then $\boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi} = \boldsymbol{U}(\boldsymbol{\Omega}^{-1} - \boldsymbol{I})\boldsymbol{U}^H$, and the optimal $\boldsymbol{\Phi}$ is $\boldsymbol{\Phi} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{U}^H$. Here $\boldsymbol{P}$ is an arbitrary $t \times t$ unitary matrix, and $\boldsymbol{\Sigma}$ is a $t \times m$ matrix with nonzero elements only on its main diagonal in decreasing order and $\boldsymbol{\Sigma}^H\boldsymbol{\Sigma} = \boldsymbol{\Omega}^{-1} - \boldsymbol{I}$. Put the optimal $\boldsymbol{\Phi}$ into (4) to get $\mathrm{tr}(\boldsymbol{\Lambda}^{-1}\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\boldsymbol{P}^H) \leq \mathcal{P}$. Similar to the proof of Lemma 1, we can choose $\boldsymbol{P} = \boldsymbol{I}$ to lower the transmit power. And it's easy to show that we can also choose $\boldsymbol{P} = \boldsymbol{I}$ to lower the transmit power in (5). So, the optimal $\boldsymbol{\Phi}$ is $\boldsymbol{\Phi} = \hat{\boldsymbol{\Phi}}\boldsymbol{U}^H$, where $\hat{\boldsymbol{\Phi}}$ is a $t \times m$ matrix with nonzero elements only on its main diagonal. Then the optimal $\boldsymbol{G}$ is

$G = V\hat{\Phi}U^H R_{uu}^{-1/2}$. To be unified, we use $\Phi$ to replace $\hat{\Phi}$, and write $G = V\Phi U^H R_{uu}^{-1/2}$, where $\Phi$ has nonzero elements only on its main diagonal. Then the MSE and squared half-coherence matrix are

$$M = U\Delta^{1/2}(I + \Phi\Lambda\Phi^H)\Delta^{1/2}U^H,$$

$$DD^H = U\Delta^{1/2}\Phi^H\Lambda^{1/2}(\Lambda^{1/2}\Phi\Phi^H\Lambda^{1/2}+I)^{-1}\Lambda^{1/2}\Phi\Delta^{1/2}U^H.$$

These have the form $UTU^H$ with $T$ diagonal. Under constraint (4) the solution for $\min \mathrm{tr}M$ is [1] [3] [4]

$$|\phi_i|^2 = \Big(\frac{\mathcal{P} + \sum_{k=1}^{\tilde{m}} \lambda_k^{-1}}{\sum_{k=1}^{\tilde{m}} \sqrt{\delta_k/\lambda_k}} \sqrt{\delta_i/\lambda_i} - \lambda_i^{-1}\Big)^+.$$

The minimum MSE is

$$\mathrm{tr}M = \frac{(\sum_{i=1}^{\tilde{m}} \sqrt{\delta_i/\lambda_i})^2}{\mathcal{P} + \sum_{i=1}^{\tilde{m}} \frac{1}{\lambda_i}}.$$

Under constraint (5) the solution is [3],

$$|\phi_1|^2 = \cdots = |\phi_m|^2 = \mathscr{P},$$

and the optimal MSE is

$$\sum_{i=1}^{m} \frac{\delta_i}{1 + \mathscr{P}\lambda_i}.$$

In summary, under a minimum mean squared error criterion, the optimal precoder is $G = V\Phi U^H R_{uu}^{-1/2}$ and $M$, $DD^H$ have the form $UTU^H$ with $T$ diagonal.

**Half Canonical Coordinates and Geometry.** $M$ and $DD^H$ have the form $UTU^H$ with $T$ diagonal. The squared half canonical correlation matrix is the diagonal matrix

$$\begin{aligned} LL^H &= U^H DD^H U \\ &= \Delta^{1/2}\Phi^H\Lambda^{1/2}(\Lambda^{1/2}\Phi\Phi^H\Lambda^{1/2} + I)^{-1}\Lambda^{1/2}\Phi\Delta^{1/2}. \end{aligned}$$

This means the combined effects of precoder and equalizer have been to transform the decorrelated source $\tilde{u} := U^H u$, with covariance $\Delta$, and equalized and whitened measurement $\hat{v} := R_{vv}^{-1/2} v$ into a system of half canonical coordinates where variable are only pairwise correlated with correlation

$$\begin{aligned} |E\tilde{u}_i\bar{\hat{v}}_i|^2 &= (LL^H)_i \\ &= \frac{\delta_i|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}. \end{aligned}$$

If we normalize the variable $(\tilde{u})_i$ by its standard deviation $\sqrt{\delta_i}$, then the cosine-square of the angle between $(\tilde{u})_i/\sqrt{\delta_i}$ and $(\hat{v})_i$ is

$$|E\frac{\tilde{u}_i}{\sqrt{\delta_i}}\bar{\hat{v}}_i|^2 = \frac{|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}.$$

Again we see that the channel is decomposed into parallel uncorrelated channels where the SNR in each channel is $\frac{|\phi_i|^2}{\frac{1}{\lambda_i}}$ and the cosine-squared of the angle between $(\hat{u})_i$ and $(\hat{v})_i$ remains $\frac{|\phi_i|^2}{|\phi_i|^2+\frac{1}{\lambda_i}}$.

## IV. Precoder and equalizer design based on minimum bit error rate (MBER) criterion and Canonical Coordinate Geometry

Here, we use the model and result of [5]. The transmit symbols are assumed to be equiprobable antipodal symbols with $R_{uu} = I$. The noise is additive white Gaussian noise with $R_{nn} = \sigma^2 I$. The equalizer $F = (HG)^\dagger$ is a zero forcing equalizer. Under the minimum bit error rate criterion and under all of the above assumptions, the optimum $G$ is $G = V\Phi D_F^H$, where $V$ is a $t \times t$ unitary matrix, $\Phi$ is a $t \times m$ diagonal matrix with nonzero elements only on its main diagonal, and $D_F^H$ is an $m \times m$ unitary inverse DFT matrix [5]. So, the squared coherence matrix has the form $D_F T D_F^H$ with $T$ diagonal. This brings the geometry we are after.

The probability of error is [5]

$$\begin{aligned} P_e &= \frac{1}{2m}\sum_{i=1}^{m} \mathrm{erfc}\Big(\frac{1}{\sqrt{2\sigma^2(FF^H)_{ii}}}\Big) \\ &= \frac{1}{2m}\sum_{i=1}^{m} \mathrm{erfc}\Big(\frac{1}{\sqrt{2\sigma^2(\Phi^H\Lambda\Phi)_{ii}^{-1}}}\Big). \end{aligned} \tag{8}$$

This function is convex with respect to $(\Phi^H\Lambda\Phi)_{ii}^{-1}$ if and only if $(\Phi^H\Lambda\Phi)_{ii}^{-1} \le \frac{1}{3\sigma^2}$ [5]. When it is convex, using Jensen's Inequality,

$$\begin{aligned} P_e &= \frac{1}{2m}\sum_{i=1}^{m} \mathrm{erfc}\Big(\frac{1}{\sqrt{2\sigma^2(\Phi^H\Lambda\Phi)_{ii}^{-1}}}\Big) \\ &\ge \frac{1}{2}\mathrm{erfc}\Big(\sqrt{\frac{m}{2\sigma^2\mathrm{tr}(\Phi^H\Lambda\Phi)^{-1}}}\Big) = P_{e,LB}. \end{aligned} \tag{9}$$

The equality holds when $(\Phi^H\Lambda\Phi)^{-1}$ has equal diagonal elements. So the optimization problem becomes

$$\min \mathrm{tr}(\Phi^H\Lambda\Phi)^{-1}. \tag{10}$$

The constraints are both (4) and $(\Phi^H\Lambda\Phi)_{ii}^{-1} \le \frac{1}{3\sigma^2}$. It has been proved in [5] that if and only if $\mathrm{tr}(\Phi^H\Lambda\Phi)^{-1} \le \frac{m}{3\sigma^2}$, their exists a unitary matrix $D$ such that $(D(\Phi^H\Lambda\Phi)^{-1}D^H)_{ii} \le \frac{1}{3\sigma^2}$. So, the problem is simplified to minimizing (10) subject to (4). And the solution is feasible if and only if $\mathrm{tr}(\Phi^H\Lambda\Phi)^{-1} \le \frac{m}{3\sigma^2}$. We introduce the singular value decomposition $\Phi = PTQ^H$, where $P$ is a $t \times t$ unitary matrix, $T$ is a $t \times m$ diagonal matrix with nonzero elements only on its main diagonal in increasing order, and $Q$ is an $m \times m$ unitary matrix. Since we have assumed that $t \ge m$, $T = \begin{bmatrix} I \\ 0 \end{bmatrix}\hat{T}$. Here $\hat{T}$ is a $m \times m$ diagonal matrix with diagonal elements in increasing order. So (10) becomes

$$\begin{aligned} \mathrm{tr}(\Phi^H\Lambda\Phi)^{-1} &= \mathrm{tr}(\hat{T}^{-H}(\begin{bmatrix} I & 0 \end{bmatrix} P^H\Lambda P \begin{bmatrix} I \\ 0 \end{bmatrix})^{-1}\hat{T}^{-1}) \\ &= \mathrm{tr}((\hat{T}^H\hat{T})^{-1}(\begin{bmatrix} I & 0 \end{bmatrix} P^H\Lambda P \begin{bmatrix} I \\ 0 \end{bmatrix})^{-1}). \end{aligned}$$

The power constraint (4) becomes $\mathrm{tr}(\hat{T}\hat{T}^H) \le \mathcal{P}$. Define $\Gamma = (\hat{T}^H\hat{T})^{-1}$. So, $\Gamma$ is diagonal with diagonal elements in decreasing order. Define $Z = (\begin{bmatrix} I & 0 \end{bmatrix} P^H\Lambda P \begin{bmatrix} I \\ 0 \end{bmatrix})^{-1}$. According to Theorem 9.H.1.h. in [12], $\mathrm{tr}(\Gamma Z) \ge \sum_{i=1}^{m} \gamma_i \mathrm{ev}_i(Z)$, where $\gamma_i$ is the $i$th diagonal element of $\Gamma$, and $\mathrm{ev}_i(Z)$ is the $i$th eigenvalue of $Z$, with eigenvalues in increasing order. The equality holds when $Z$ is diagonal with diagonal elements in increasing order. Since $\Lambda$ is diagonal with diagonal elements in decreasing order, the optimal $P$ is $P = I$. The problem is simplified to a scalar optimization problem, and we can easily get the solution $|t_i|^2 = \frac{\mathcal{P}\sqrt{\lambda_i^{-1}}}{\sum_{k=1}^{m}\sqrt{\lambda_k^{-1}}}$. If and only if $\frac{(\sum_{i=1}^{m}\sqrt{\lambda_i^{-1}})^2}{\mathcal{P}} \leq \frac{m}{3\sigma^2}$, the solution is feasible. In case the solution is feasible, we should find a unitary matrix $\boldsymbol{D}$ such that $(\boldsymbol{D}(\boldsymbol{\Phi}^H\boldsymbol{\Lambda}\boldsymbol{\Phi})^{-1}\boldsymbol{D}^H)_{ii} = (\boldsymbol{D}\boldsymbol{Q}(\boldsymbol{T}^H\boldsymbol{\Lambda}\boldsymbol{T})^{-1}\boldsymbol{Q}^H\boldsymbol{D}^H)_{ii} \leq \frac{1}{3\sigma^2}$. From [5], $\boldsymbol{D} = \boldsymbol{D}_F\boldsymbol{Q}^H$ is an optimal choice since it makes the diagonal elements equal. Here $\boldsymbol{D}_F$ is a unitary DFT matrix. So the optimal $\boldsymbol{\Phi}$ is $\boldsymbol{\Phi} = \boldsymbol{T}\boldsymbol{D}_F^H$. The optimal $\boldsymbol{G}$ is $\boldsymbol{G} = \boldsymbol{V}\boldsymbol{T}\boldsymbol{D}_F^H$. To be unified, we use $\boldsymbol{\Phi}$ in place of $\boldsymbol{T}$, so $\boldsymbol{G} = \boldsymbol{V}\boldsymbol{\Phi}\boldsymbol{D}_F^H$. Here $\boldsymbol{\Phi}$ has nonzero elements only on its main diagonal. With this optimum $\boldsymbol{G}$, the optimum squared coherence matrix is

$$\boldsymbol{C}\boldsymbol{C}^H = \boldsymbol{D}_F\boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2}(\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Phi}\boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2} + \boldsymbol{I})^{-1}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Phi}\boldsymbol{D}_F^H,$$

which has the form $\boldsymbol{D}_F\boldsymbol{T}\boldsymbol{D}_F^H$, with $\boldsymbol{T}$ diagonal. The squared canonical correlation matrix is

$$\boldsymbol{K}\boldsymbol{K}^H = \boldsymbol{D}_F^H\boldsymbol{C}\boldsymbol{C}^H\boldsymbol{D}_F = \boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2}(\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Phi}\boldsymbol{\Phi}^H\boldsymbol{\Lambda}^{1/2}+\boldsymbol{I})^{-1}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Phi}^H.$$

This means the combined effects of $\boldsymbol{G}$ and $\boldsymbol{F}$ have been to transform the white source $\hat{\boldsymbol{u}} = \boldsymbol{D}^H\boldsymbol{u}$ and equalized and equalized and whitened measurement $\hat{\boldsymbol{v}} = \boldsymbol{R}_{vv}^{-1/2}\boldsymbol{v}$ into a system of canonical coordinates where variable are only pairwise correlated:

$$\begin{aligned} |E\hat{u}_i\bar{\hat{v}}_i|^2 &= (\boldsymbol{K}\boldsymbol{K}^H)_i \\ &= \frac{|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}. \end{aligned}$$

It is as if the white source after inverse DFT, $\boldsymbol{D}^H\boldsymbol{u}$ is communicating over $m$ uncorrelated channels. In each, the cosine-squared are

$$(\boldsymbol{K}\boldsymbol{K}^H)_i = \frac{|\phi_i|^2}{|\phi_i|^2 + \frac{1}{\lambda_i}}.$$

Again, the geometry of each channel is that of previous designs.

## V. Conclusion

We have analyzed optimum precoder and equalizer designs under various criteria for block transmission over a matrix channel. To derive the known results of [1] - [5] we have used the theory of majorization, as in [4]. More importantly, we have shown that, under all criteria, the optimum precoder and equalizer decompose the channel into parallel Shannon channels where the whitened message and equalized and whitened measurement are canonical coordinates. The canonical correlations (or cosines) between these canonical coordinates determine SNR, information rate, or MSE in each channel. So, the geometry of optimum precoding and equalizing is the geometry of canonical coordinates [9]. For each of the known precoder and equalizer designs we have reviewed, the message is precoded and the measurement is equalized in such a way that their whitened versions are canonical. In each design the cosine-squared of the angle between a pair of canonical message and measurement coordinates is $\frac{|\phi_i|^2}{|\phi_i|^2+\frac{1}{\lambda_i}}$. The details of power allocation $|\phi_i|^2$ depend on the criterion of optimality.

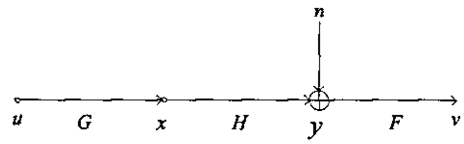


Figure 1: Block or matrix transmission system

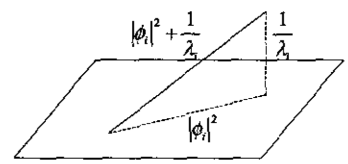


Figure 2: Pythagorean decomposition of a subchannel

## References


[1] A. Scaglione, G.B. Giannakis, S.Barbarossa, "Redundant filterbank precoders and equalizers part 1: unification and optimal designs", *IEEE Transactions on Signal Processing* , vol. 47, no. 7, pp. 1988-2006, July 1999.

[2] A. Scaglione, S. Barbarossa, G.B. Giannakis, "Filterbank Transceivers Optimizing Information Rate in Block Transmissions over Dispersive Channels," *IEEE Trans. on Information Theory*, vol. 45, no.3, pp.1019-1032, April 1999.

[3] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, H. Sampath, "Optimal Designs for Space-Time Linear Precoders and Decoders," *IEEE Transactions on Signal Processing* , vol. 50, no. 5, pp. 1051-1064, May 2002.

[4] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, "Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: a Unified Framework for Convex Optimization," *IEEE Trans. on Signal Processing*, vol. 51, no.9, pp 2381-2401, September 2003.

[5] Y. Ding, T.N. Davidson, Z. Luo, K.M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. on Signal Processing*, vol. 51, no.9, pp 2410-2423, September 2003.

[6] L.L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1990.

[7] L.L. Scharf, J.K. Thomas, "Wiener Filters in Canonical Coordinates for Transform Coding, Filtering, and Quantizing," *IEEE Trans. on Signal Processing*, vol. 46, no.3, pp.647-654, March 1998.

[8] Y. Hua, M. Nikpour, P. Stoica, "Optimal Reduced-Rank Estimation and Filtering," *IEEE Trans. on Signal Processing*, vol. 49, No.3, pp.457-469, March 2001.

[9] L.L. Scharf, C.T. Mullis, "Canonical Coordinates and the Geometry of Inference, Rate, and Capacity," *IEEE Trans. on Signal Processing*, vol. 48, no.3, pp.824-831, March 2000.

[10] P.J. Schreier, L.L. Scharf, "Canonical coordinates for reduced-rank estimation of improper complex random vectors," *Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on*, vol. 2 , 2002 Page(s): 1153 -1156

[11] J. Yang, S. Soy, "On Joint Tranmitter and Receiver Optimization for Multiple-Input-Multiple-Output (MIMO) Transmission Systems," *IEEE Trans. on Communications*, vol. 42, No.12, pp.3221-3231, December 1994.

[12] A.W. Marshall, I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, NY, 1979.

[13] Z. Wang, G.B. Giannakis, "Wireless Multicarrier Communications," *IEEE Signal Processing Magazine*, pp.29-48, May 2000.

[14] G.G. Raleigh, V.K. Jones, "Multivariate Modulation and Coding for Wireless Communication," *IEEE Journal on Selected Areas in Communications*, vol. 17, No. 5, pp.851-866, May 1999.