ROBUST FIR PRECODER DESIGN WITH IMPERFECT CHANNEL KNOWLEDGE FOR BROADBAND MIMO WIRELESS SYSTEMS

Yongfang Guo and *Bernard C. Levy*

Department of Electrical and Computer Engineering
University of California, Davis, CA 95616
E-mails: yguo, levy@ece.ucdavis.edu

ABSTRACT

This paper considers a robust FIR precoder design problem for frequency-selective multiple-input multiple-output (MIMO) wireless systems under transmit power constraints with imperfect channel knowledge at the transmitter. The robust design problem requires finding the optimal power constrained FIR precoder for the worst-case channel model within a neighborhood of the estimated channel. The neighborhood is formed by placing a bound on Kullback-Leibler (KL) divergence between the actual and estimated channels. We show that the optimal power constrained FIR precoder designed for the estimated channel under the assumption of perfect channel knowledge is robust when channel uncertainties are measured by the KL divergence “metric”.

1. INTRODUCTION

For the case of a left coprime FIR channel, which arises generically when the number p of transmit antennas is larger than the number q of receive antennas, the matrix Bezout identity [1] can be employed to design a FIR MIMO precoder that equalizes exactly the channel at the transmitter. Unfortunately, when deep fades are present in the MIMO channel frequency response, Bezout precoders usually increase the transmit power. To overcome this problem, a method was proposed in [2] for the optimal design of FIR precoders under transmit power constraints. It reduces the design of FIR precoders to the minimization of a quadratic convex objective function under a quadratic convex constraint, so it can be solved by using Lagrangian duality [3]. The design technique assumes, however, that the the channel is perfectly known at the transmitter. Channel state information (CSI) can be obtained at the receiver through the use of either traning sequences or blind/semi-blind estimation. At the transmit end, CSI can be acquired either by using a feedback channel between the receiver and transmitter, or from previous receive measurements due to the channel reciprocity property. However, in real systems, channel estimates at the transmitter tend to be inaccurate, owing to the time varying nature of channels and feedback delay. The impact of channel estimation errors is important in practice and needs to be taken into account when designing a precoder. In this paper we examine the design of a robust FIR precoder with transmit power constraints and imperfect channel knowledge. The goal is to synthesize a power constrained precoder which minimizes the power of the residual intersymbol interference (ISI) and interchannel interference (ICI) for the worst channel located in a neighborhood of the estimated channel. Following [4], the neighborhood is formed by placing a bound on the Kullback-Leibler (KL) divergence between the actual channel and the estimated channel. The use of the KL divergence is rather natural as a metric for model mismatch since it is commonly used by statisticians [5] for fitting statistical models, and by using a differential geometric viewpoint it is argued in [6] that KL divergence is the natural geometric ”distance” between systems. Therefore, the design of a robust power constrained FIR precoder reduces to the solution of a min-max problem where we seek to find the optimal precoder for the least favorable channel within a neighborhood specified by the KL divergence between the actual and estimated channels. We show that the optimal precoder for the estimated channel under the assumption of perfect channel knowledge is robust, but its performance is impacted by the presence of uncertainties. This result is consistent with those obtained in [4, 7] for robust estimation and detection with a KL criterion. A precise characterization of the performance loss in function of the allowable KL tolerance is provided and simulations are presented.

This research was supported by NSF Grant ECS-0121469.

2. KL DIVERGENCE FOR THE ROBUST PRECODER DESIGN

Given a $q \times p$ channel $\mathbf{H}(z) = \sum_{l=0}^{d} \mathbf{H}(l) z^{-l}$ of order d, the error between $\mathbf{H}(z)$ and the estimated channel $\hat{\mathbf{H}}(z) = \sum_{l=0}^{d} \hat{\mathbf{H}}(l) z^{-l}$ can be expressed as

$$\mathbf{H}_{\triangle}(z) = \sum_{l=0}^{d} \mathbf{H}_{\triangle}(l) z^{-l} = \mathbf{H}(z) - \hat{\mathbf{H}}(z) \,. \tag{2.1}$$

When the number p of transmit antennas is larger than the number q of receive antennas, a $p \times q$ FIR precoder $\mathbf{F}(z) = \sum_{n=0}^{r} \mathbf{F}(n) z^{-n}$ of degree r can be used to equalize the given channel at the transmitter. The sampled received vector sequence $\{\mathbf{y}(n) \in \mathbb{C}^q \,,\ n \in \mathbb{Z}\}$ can be written as

$$\mathbf{y}(n) = \mathbf{H}(n) * \mathbf{F}(n) * \mathbf{x}(n) + \mathbf{v}(n) \,, \tag{2.2}$$

where $*$ represents the convolution operation, $\mathbf{v}(n)$ is a vector complex circular WGN sequence independent of $\mathbf{x}(n)$, with zero-mean and variance $\sigma_v^2 I_q$, and the input vector sequence $\mathbf{x}(n)$ is assumed to have zero mean and covariance matrix $\sigma_x^2 \mathbf{I}_q$. For the actual channel $\mathbf{H}(z)$, the probability density $f_{\mathbf{H}}(\mathbf{y}(n)|\mathbf{x}_n)$ of $\mathbf{y}(n)$ conditioned on $\mathbf{x}_n$ is a Gaussian distribution with mean $\mathbf{H}\Gamma(\mathbf{F})\mathbf{x}_n$ and variance $\sigma_v^2 \mathbf{I}_q$, where the $(d+r+1)q \times 1$ column vector $\mathbf{x}_n$ and the $q \times (d+1)p$ row block matrix $\mathbf{H}$ are given by

$$\begin{aligned} \mathbf{x}_n &= \begin{bmatrix} \mathbf{x}^H(n) & \mathbf{x}^H(n-1) & \dots & \mathbf{x}^H(n-r-d) \end{bmatrix}^H \\ \mathbf{H} &= \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(d) \end{bmatrix} \,, \end{aligned} \tag{2.3}$$

and the $(d+1)p \times (d+r+1)q$ block Toeplitz matrix $\Gamma(\mathbf{F})$ takes the form

$$\Gamma(\mathbf{F}) = \begin{bmatrix} \mathbf{F}(0) & \dots & \mathbf{F}(r) & 0 & \dots 0 \\ 0 & \mathbf{F}(0) & \dots & \mathbf{F}(r) & \dots 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots 0 & \mathbf{F}(0) & \dots & \mathbf{F}(r) \end{bmatrix} . \tag{2.4}$$

Similarly, for the estimated channel $\hat{\mathbf{H}}(z)$, the conditional probability density $f_{\hat{\mathbf{H}}}(\mathbf{y}(n)|\mathbf{x}_n)$ is a Gaussian distribution with mean $\hat{\mathbf{H}}\Gamma(\mathbf{F})\mathbf{x}_n$ and variance $\sigma_v^2 \mathbf{I}_q$, where the matrix $\hat{\mathbf{H}}$ has the same form as $\mathbf{H}$ except that $\mathbf{H}(n)$ is replaced by $\hat{\mathbf{H}}(n)$. Since these two Gaussian densities have the same variance matrix, but different means, the Kullback-Leibler (KL) divergence or relative entropy of $f_{\mathbf{H}}(\mathbf{y}(n)|\mathbf{x}_n)$ with respect to $f_{\hat{\mathbf{H}}}(\mathbf{y}(n)|\mathbf{x}_n)$ can be expressed as

$$\begin{aligned} & D(\mathbf{H}, \hat{\mathbf{H}}|\mathbf{x}_n) \\ =\ & \int \ln\Big[\frac{f_{\mathbf{H}}\big(\mathbf{y}(n)|\mathbf{x}_n\big)}{f_{\hat{\mathbf{H}}}\big(\mathbf{y}(n)|\mathbf{x}_n\big)}\Big] f_{\mathbf{H}}\big(\mathbf{y}(n)|\mathbf{x}_n\big) \mathrm{d}\mathbf{y}(n) \\ =\ & ||\mathbf{H}_{\triangle}\Gamma(\mathbf{F})\mathbf{x}_n||_2^2/(2\sigma_v^2) \,, \end{aligned} \tag{2.5}$$

where $||.||_2^2$ denotes the squared Euclidean norm and $\mathbf{H}_{\triangle} = \mathbf{H} - \hat{\mathbf{H}}$. Then the KL divergence of $f_{\mathbf{H}}(\mathbf{y}(n))$ with respect to $f_{\hat{\mathbf{H}}}(\mathbf{y}(n))$ is given by

$$D(\mathbf{H}, \hat{\mathbf{H}}) = E_{\mathbf{x}_n}[D(\mathbf{H}, \hat{\mathbf{H}}|\mathbf{x}_n)] = \sigma_x^2 ||\mathbf{H}_{\triangle}\Gamma(\mathbf{F})||_F^2/(2\sigma_v^2), \tag{2.6}$$

where $||\mathbf{M}||_F^2 = \mathrm{tr}(\mathbf{M}\mathbf{M}^H)$ denotes the squared Frobenius norm.

The neighborhood $\mathcal{H}$ formed by all channels whose KL divergence with respect to the estimated channel $\hat{\mathbf{H}}(n)$ is less c can be characterized as

$$\sigma_x^2 ||\mathbf{H}_{\triangle}\Gamma(\mathbf{F})||_F^2/(2\sigma_v^2) \le c \Leftrightarrow \mathrm{tr}\big[\Gamma(\mathbf{F})^H \mathbf{H}_{\triangle}^H \mathbf{H}_{\triangle}\Gamma(\mathbf{F})\big] \le c', \tag{2.7}$$

where $c' = 2c\sigma_v^2/\sigma_x^2$. Our goal in this paper is to find the optimal power constrained precoder for the least favorable channel within the neighborhood of the estimated channel specified by (2.7).

3. ROBUST PRECODER DESIGN WITH A KL DIVERGENCE BOUND

The system we consider has p transmit and q receive antennas and we assume $p > q$. For a left coprime channel, which arises generically when $p > q$, there exists a $p \times q$ FIR matrix $\mathbf{F}(z)$ obeying the matrix Bezout identity

$$\mathbf{H}(z)\mathbf{F}(z) = \mathbf{E}(z) \stackrel{\triangle}{=} \mathrm{diag}\{z^{-k_i} \,,\ 1 \le i \le q\} \,, \tag{3.1}$$

where the integer delays k_i can be selected arbitrarily. Precoders $\mathbf{F}(z)$ obeying (3.1) form the class of Bezout precoders discussed in [1]. However Bezout precoders may increase the transmit power significantly to compensate for deep fades in the singular values of $\mathbf{H}(e^{j\theta})$. To overcome this defect, [2] proposed a precoder design technique that achieves the optimal trade-off between channel equalization and transmit power constraints when CSI is perfectly known at the transmitter. When the estimated channel $\hat{\mathbf{H}}(z)$ is different from the actual channel $\mathbf{H}(z)$, the objective function evaluating the residual ISI and ICI power can be written as

$$\begin{aligned} & J(\mathbf{F}, \mathbf{H}_{\triangle}) \\ =\ & \frac{1}{2\pi} \int_{-\pi}^{\pi} ||[\hat{\mathbf{H}}(e^{j\theta}) + \mathbf{H}_{\triangle}(e^{j\theta})]\mathbf{F}(e^{j\theta}) - \mathbf{E}(e^{j\theta})||_F^2 d\theta \\ =\ & ||\hat{\mathbf{H}}\Gamma(\mathbf{F}) + \mathbf{H}_{\triangle}\Gamma(\mathbf{F}) - \mathbf{E}||_F^2 \,, \end{aligned} \tag{3.2}$$

where the target zero-forcing matrix $\mathbf{E}$ of dimension $q \times (d+r+1)q$ is given by

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}(0) & \mathbf{E}(1) & \dots & \mathbf{E}(d+r) \end{bmatrix} . \tag{3.3}$$

Given the objective function, the robust precoder design with a KL bound can be formulated as a min-max problem of the form

$$\begin{aligned} \min_{\mathbf{F}} \max_{\mathbf{H}_\triangle} \quad & J(\mathbf{F}, \mathbf{H}_\triangle) \\ s.t. \quad & \text{tr}[\Gamma(\mathbf{F})^H \mathbf{H}_\triangle^H \mathbf{H}_\triangle \Gamma(\mathbf{F}))] \le c' , \\ & \text{tr}[\mathbf{F}^H \mathbf{F}] \le P_T . \end{aligned} \tag{3.4}$$

where $\mathbf{F} = \begin{bmatrix} \mathbf{F}^H(0) & \mathbf{F}^H(1) & \dots & \mathbf{F}^H(r) \end{bmatrix}^H$ has dimension $(r+1)p \times q$. To solve this min-max problem, we can first maximize $J(\mathbf{F}, \mathbf{H}_\triangle)$ with respect to $\mathbf{H}_\triangle$ under the constraint (2.7) and then once the least favorable channel has been identified, we can synthesize the optimal precoder $\mathbf{F}$ for this channel. The Lagrangian associated with maximization of $J(\mathbf{F}, \mathbf{H}_\triangle)$ under the constraint (2.7) takes the form

$$L(\mathbf{F}, \mathbf{H}_\triangle, \lambda) = J(\mathbf{F}, \mathbf{H}_\triangle) + \lambda(c' - ||\mathbf{H}_\triangle \Gamma(\mathbf{F})||_F^2), \tag{3.5}$$

where λ is chosen such that the Hessian $\bigtriangledown^2_{\mathbf{H}_\triangle} L(\mathbf{F}, \mathbf{H}_\triangle, \lambda) = (1-\lambda)\Gamma(\mathbf{F})\Gamma(\mathbf{F})^H$ is negative semi-definite, i.e. $\lambda \ge 1$. In this case, the function $L(\mathbf{F}, \mathbf{H}_\triangle, \lambda)$ is concave in $\mathbf{H}_\triangle$ and the maximum is unique and obeys the gradient condition

$$\begin{aligned} & \bigtriangledown_{\mathbf{H}_\triangle} L(\mathbf{F}, \mathbf{H}_\triangle, \lambda) \\ = \; & \big[\hat{\mathbf{H}}\Gamma(\mathbf{F}) + \mathbf{H}_\triangle \Gamma(\mathbf{F}) - \mathbf{E} - \lambda \mathbf{H}_\triangle \Gamma(\mathbf{F})\big] \boldsymbol{\Gamma}(\mathbf{F})^{\mathbf{H}} \\ = \; & 0 . \end{aligned} \tag{3.6}$$

Since $\Gamma(\mathbf{F})^{\mathbf{H}}$ has full row rank, we have

$$\hat{\mathbf{H}}\Gamma(\mathbf{F}) + \mathbf{H}_\triangle^{\max}\Gamma(\mathbf{F}) - \mathbf{E} - \lambda^* \mathbf{H}_\triangle^{\max}\Gamma(\mathbf{F}) = \mathbf{0} ,$$

which yields

$$\mathbf{H}_\triangle^{\max}\Gamma(\mathbf{F}) = \frac{\mathbf{E} - \hat{\mathbf{H}}\Gamma(\mathbf{F})}{1 - \lambda^*} . \tag{3.7}$$

Here λ^* satisfies the Karush-Kuhn-Tucker condition $c' = ||\mathbf{H}_\triangle^{\max}\Gamma(\mathbf{F})||_F^2$, which results in

$$\lambda^* = 1 + \sqrt{\frac{||\mathbf{E} - \hat{\mathbf{H}}\Gamma(\mathbf{F})||_F^2}{c'}} . \tag{3.8}$$

Substituting the expressions (3.7) and (3.8) for $\mathbf{H}_\triangle \Gamma(\mathbf{F})$ and λ inside (3.5), the worst-case objective function is given by

$$M(\mathbf{F}) = L(\mathbf{F}, \mathbf{H}_\triangle^{\max}, \lambda^*) = [\sqrt{||\mathbf{E} - \hat{\mathbf{H}}\Gamma(\mathbf{F})||_F^2} + \sqrt{c'}]^2 . \tag{3.9}$$

This function represents the residual ISI and ICI power for the least favorable channel in the neighborhood $\mathcal{H}$ centered around the estimated channel $\hat{\mathbf{H}}$. The next step is to find the precoder minimizing $M(\mathbf{F})$ under a transmit power constraint. This is equivalent to minimizing $||\mathbf{E} - \hat{\mathbf{H}}\Gamma(\mathbf{F})||_F^2$, which is also equivalent to minimizing $||\mathbf{E}' - \Gamma(\hat{\mathbf{H}})\mathbf{F}||_F^2$ where the $(d+r+1)q \times (r+1)p$ block Toeplitz matrix $\Gamma(\hat{\mathbf{H}})$ takes the form

$$\Gamma(\hat{\mathbf{H}}) = \begin{bmatrix} \hat{\mathbf{H}}(0) & 0 & \ddots & 0 \\ \vdots & \hat{\mathbf{H}}(0) & \ddots & \vdots \\ \hat{\mathbf{H}}(d) & \vdots & \ddots & 0 \\ 0 & \hat{\mathbf{H}}(d) & \ddots & \hat{\mathbf{H}}(0) \\ \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \hat{\mathbf{H}}(d) \end{bmatrix} \tag{3.10}$$

and the block column matrix

$$\mathbf{E}' = \begin{bmatrix} \mathbf{E}^H(0) & \mathbf{E}^H(1) & \dots & \mathbf{E}^H(r+d) \end{bmatrix}^H \tag{3.11}$$

has dimension $(d+r+1)q \times q$. Then the optimal precoder design for the worst-case channel model reduces to

$$\begin{aligned} \min \quad & W(\mathbf{F}) = ||\mathbf{E}' - \Gamma(\hat{\mathbf{H}})\mathbf{F}||_F^2 \\ s.t \quad & \text{tr}(\mathbf{F}^H\mathbf{F}) \le P_T . \end{aligned} \tag{3.12}$$

Thus the robust precoder design with a KL divergence bound and a transmit power constraint coincides with the standard power constrained optimal precoder design problem for the estimated channel under perfect channel information, which is consistent with the results of [4, 7]. Since the Hessian $\bigtriangledown^2 W(\mathbf{F}) = \Gamma(\hat{\mathbf{H}})^H\Gamma(\hat{\mathbf{H}})$ is positive semi-definite, the objective function W is convex. The constraint is also convex, so the resulting convex optimization problem can be solved either in primal space or dual space [3]. Because of the large number of coefficients of $\mathbf{F}$, the primal problem has a large dimension, whereas the dual problem is scalar problem, since there is only one constraint. Therefore, we solve the dual form of the optimization problem.

Following the approach of [2], The Lagrangian associated with the minimization problem (3.12) takes the form

$$\Lambda(\mathbf{F}, \mu) = W(\mathbf{F}) + \mu\big(\text{tr}(\mathbf{F}^H\mathbf{F}) - P_T\big) . \tag{3.13}$$

In order to find the saddle point $(\mathbf{F}^*, \mu^*)$, we first fix μ and minimize the Lagrangian over $\mathbf{F}$, which yields

$$\mathbf{F}^*(\mu) = (\mathbf{M} + \mu\mathbf{I})^{-1}\Gamma(\hat{\mathbf{H}})^H\mathbf{E}' , \tag{3.14}$$

where $\mathbf{M} \triangleq \Gamma(\hat{\mathbf{H}})^H\Gamma(\hat{\mathbf{H}})$. The dual function is then given by

$$\begin{aligned} & G(\mu) = \Lambda(\mathbf{F}^*(\mu), \mu) = \\ & \text{tr}\Big[-\mathbf{E}'^H\boldsymbol{\Gamma}(\hat{\mathbf{H}})(\mathbf{M}+\mu\mathbf{I})^{-1}\boldsymbol{\Gamma}(\hat{\mathbf{H}})^H\mathbf{E}'\Big] + q - \mu P_T \end{aligned} \tag{3.15}$$

over the domain $\mathcal{D} = \{\mu : \mu \ge 0, \mathbf{M} + \mu\mathbf{I} > 0\}$. The unique solution μ^* of the dual problem is obtained by maximizing the concave function $G(\mu)$ over $\mathcal{D}$. This can be accomplished by using Newton techniques. Then, substituting μ^* in (3.14) gives the solution

$$\mathbf{F}^* = (\mathbf{M} + \mu^*\mathbf{I})^{-1}\Gamma(\hat{\mathbf{H}})^H\mathbf{E}' \tag{3.16}$$

of the robust precoder design problem

4. SIMULATION RESULT

In the simulations below, a 4-input-2-output channel with length 5 is considered and uncoded QPSK symbols are transmitted. The SNR is defined as $\mathrm{SNR} = \mathrm{tr}(\mathbf{F}\mathbf{F}^H)\sigma_x^2/\sigma_v^2$. Fig. 1 shows the BER performance of robust precoders vs. the bound c' on the KL divergence when $\mathrm{SNR} = 10\mathrm{dB}$ where c' is normalized, i.e., $c' = \frac{||\mathbf{H}_\triangle\Gamma(\mathbf{F})||_F^2}{||\hat{\mathbf{H}}\Gamma(\mathbf{F})||_F^2}$. The dashed and star curves represent the BER performance of the optimal precoder if the signal is transmitted over the estimated channel and the worst-case channel, respectively. For the worst-case channel, as the bound inceases, the BER performance of robust precoders degrades, i.e., if the desired BER is less than 10^{-2}, the KL divergence bound c' must be less than 0.1. Fig. 2 shows the BER performance of the robust precoder vs. SNR when c' is fixed as 0.1.

5. CONCLUSION

In this paper, we have considered the robust design of a power constrained FIR precoder for an imperfectly known frequency-selective MIMO channel. The channel is assumed to be located within a neighborhood centered on the estimated channel and obtained by bounding the KL divergence between the actual and estimated channels. This design technique is applicable to unbalanced channels with $p > q$. It formulates the design of robust precoders as a min-max problem where the goal is to find the power constrained precoder minimizing the residual ISI and ICI for the least favorable channel in the given neighborhood. By solving this min-max problem, the robust precoder design with imperfect CSI is reduced to a standard power constrained precoder design problem for the estimated channel under perfect CSI.

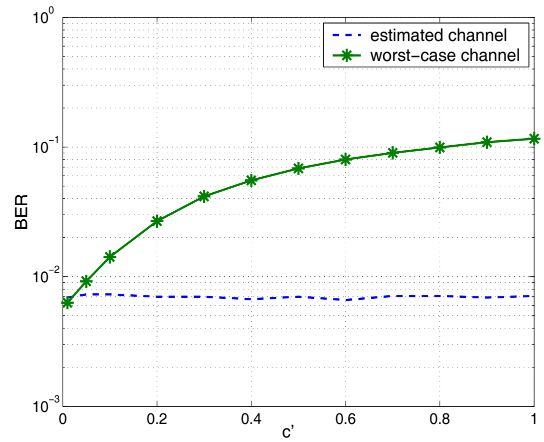

Fig. 1. BER vs c'(SNR=10dB)

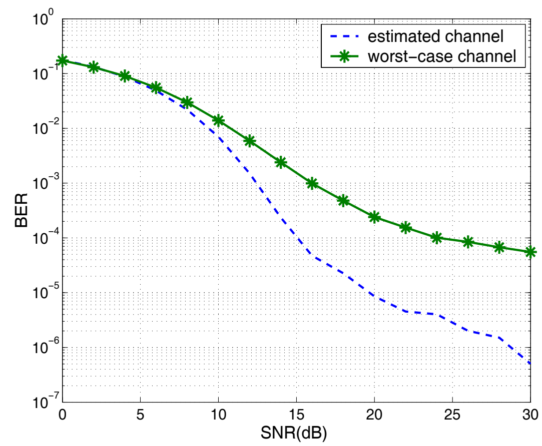

Fig. 2. BER vs. SNR (c'=0.1)

6. REFERENCES

[1] S. Y. Kung, Y. Wu, and X. Zhang, "Bezout space-time precoders and equalizers for MIMO channels," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2499–2514, Oct. 2002.

[2] Y. Guo and B. C. Levy, "Design of FIR precoders and equalizers for broadband MIMO wireless channels with power constraints," in *Proc. of 37th Asilomar Conf. on Signal, System, and Computers*, (Pacific Grove, CA), Nov. 2003.

[3] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[4] B. C. Levy and R. Nikoukhah, "Robust least-squares estimation with a relative entropy constraint." to appear in IEEE Trans. Informat. Theory, 2003.

[5] R. H. Shumway and D. Stoffer, *Time Series Analysis and its Applications*. New York, NY: Springer Verlag, 2000.

[6] S.-I. Amari and H. Nagaoka, *Methods of information Geometry*. Providence, RI: American Mathematical Society, 2000.

[7] A. G. Dabak and D. H. Johnson, "Geometrically based robust detection," in *Proc. Conf. Information Sciences and Systems*, (The Johns Hopkins Univ.), 1993.