ROBUST MSE EQUALIZER DESIGN FOR MIMO FLAT-FADING CHANNELS IN THE PRESENCE OF MODEL UNCERTAINTIES

Yongfang Guo and *Bernard C. Levy*

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

ABSTRACT

This paper considers a robust mean-square error (MSE) equalizer design problem for flat-fading multiple-input multiple-output (MIMO) channels with imperfect channel and noise information at the receiver. When the channel state information (CSI) and the noise covariance are known exactly at the receiver, a minimum mean square error (MMSE) equalizer can be employed to estimate the transmitted signal. However, in actual systems, it is necessary to take into account channel and noise estimation errors. We consider here a worst-case equalizer design problem where the goal is to find the equalizer minimizing the equalization MSE for the least favorable channel model within a neighborhood of the estimated model. Lagrangian optimization is used to convert this min-max problem into a convex min-min problem over a convex domain which is solved by interchanging the minimization order.

1. INTRODUCTION

MIMO wireless communication systems have proved attractive due to their ability to exploit spatial diversity and multiplexing to achieve high data rate communications. Although various MIMO equalization methods have been proposed, one limitation of most existing techniques is that they assume that the CSI and noise distribution are known perfectly at the receiver. However, in practice, channel and noise estimates are subject to uncertainties. We consider in this paper the design of a worst-case MSE equalizer for an imperfectly known MIMO flat-fading channel. We seek to find the optimal equalizer minimizing the MSE for the least favorable model in a neighborhood of the estimated model. Following the approach of [1], the neighborhood is formed by placing a bound on the Kullback-Leibler (KL) divergence between the actual and estimated channel models. In this respect, note that the use of the KL divergence is rather natural as a metric for model mismatch since it is commonly used by statisticians for fitting statistical models, and used as the natural geometric "distance" between systems [2]. Equalizers designed in this manner can guarantee a fixed level of performance for any realization of actual channels in the neighborhood of the estimated channel. Then the design of a robust MSE equalizer reduces to the solution of a min-max problem. To solve this min-max problem, following an approach proposed in [3] for solving a regularized robust least-squares problem, we employ Lagrangian optimization to convert the min-max problem into a convex min-min problem over a convex domain. By partial minimization this problem reduces to a simple scalar minimization problem for the Lagrange multiplier which can be solved numerically by using the steepest descent method.

This research was supported by NSF Grant ECS-0121469.

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider here the case of flat-fading MIMO channels with n_T transmit and n_R receive antennas and assume that $n_R \geq n_T$, so that the complex $n_R \times n_T$ matrix $\mathbf{H}$ representing the channel in the baseband domain has full column rank with probability one. So we assume below that $\mathbf{H}$ has full column rank. Given a $n_R \times n_T$ matrix channel $\mathbf{H}$, the noisy observation is given by $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$, where $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_T})$ is the symbol vector transmitted during a signalling period and $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ is the noise vector. It is assumed that $\mathbf{R}$ is positive definite, so that the noise affects all observation components. When the channel matrix $\mathbf{H}$ and the noise covariance $\mathbf{R}$ are known perfectly at the receiver end, the MMSE equalizer is given by $\mathbf{F} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \mathbf{R})^{-1}$.

When the estimated channel $\hat{\mathbf{H}}$ and noise covariance matrix $\hat{\mathbf{R}}$ are different from the actual $\mathbf{H}$ and $\mathbf{R}$, the MSE objective function can be expressed as

$$J(\mathbf{F}, \mathbf{H}_\triangle, \mathbf{R}) = \mathrm{tr}\Big\{(\mathbf{I} - \mathbf{F}(\hat{\mathbf{H}} + \mathbf{H}_\triangle)) (\mathbf{I} - \mathbf{F}(\hat{\mathbf{H}} + \mathbf{H}_\triangle))^H + \mathbf{F}\mathbf{R}\mathbf{F}^H\Big\}, \quad (2.1)$$

where $\mathbf{H}_\triangle = \mathbf{H} - \hat{\mathbf{H}}$. We use the KL divergence to measure the "distance" between the actual model $(\mathbf{H}, \mathbf{R})$ and the estimated model $(\hat{\mathbf{H}}, \hat{\mathbf{R}})$, which takes the form

$$D(f(\mathbf{y}, \mathbf{x}), \hat{f}(\mathbf{y}, \mathbf{x})) = \int\int \ln\Big[\frac{f(\mathbf{y}, \mathbf{x})}{\hat{f}(\mathbf{y}, \mathbf{x})}\Big] f(\mathbf{y}, \mathbf{x}) d\mathbf{x} d\mathbf{y}.$$

By substituting the Gaussian probability densities of the actual and estimated models, it is easy to show (see e.g., [4]) that the KL divergence admits the expression

$$D(\mathbf{H}, \mathbf{R}; \hat{\mathbf{H}}, \hat{\mathbf{R}}) = \mathrm{tr}\{\mathbf{H}_\triangle^H \hat{\mathbf{R}}^{-1} \mathbf{H}_\triangle\} + \mathrm{tr}\{\hat{\mathbf{R}}^{-1}\mathbf{R} - \mathbf{I}_{n_R}\} - \ln\det(\hat{\mathbf{R}}^{-1}\mathbf{R}). \quad (2.2)$$

Let

$$\mathcal{B} = \{(\mathbf{H}_\triangle, \mathbf{R}) \,:\, D(\mathbf{H}, \mathbf{R}; \hat{\mathbf{H}}, \hat{\mathbf{R}}) \leq c\} \quad (2.3)$$

denote the 'ball' formed by the models whose KL divergence with respect to the nominal model $(\hat{\mathbf{H}}, \hat{\mathbf{R}})$ is less than or equal to c. The robust MSE equalizer design problem with a KL divergence bound can be formulated as a min-max problem of the form

$$\min_{\mathbf{F}} \max_{(\mathbf{H}_\triangle, \mathbf{R}) \in \mathcal{B}} J(\mathbf{F}, \mathbf{H}_\triangle, \mathbf{R}). \quad (2.4)$$

3. ROBUST MSE EQUALIZER DESIGN

To examine the min-max problem, we first consider the normalized channel matrix $\hat{\mathbf{H}}_s = \hat{\mathbf{R}}^{-1/2}\hat{\mathbf{H}}$, where $\hat{\mathbf{R}}^{1/2}$ satisfies $\hat{\mathbf{R}} = \hat{\mathbf{R}}^{1/2}(\hat{\mathbf{R}}^{1/2})^H$. Then we perform the singular value decomposition $\hat{\mathbf{H}}_s = \mathbf{U}[\mathbf{\Sigma}^T \;\; \mathbf{0}^T_{(n_R - n_T)\times n_T}]^T \mathbf{V}$, where $\mathbf{\Sigma} = \mathrm{diag}\{\sigma_i\,, 1 \leq i \leq n_T\}$ with $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_{n_T} > 0$. The main result of our paper is as follows.

Theorem 1 *The robust MSE equalizer $\mathbf{F}^{\mathrm{rob}}$ solving the min-max problem (2.4) has the structure*

$$\mathbf{F}^{\mathrm{rob}} = \mathbf{V}^H \left[\begin{array}{cc} \bar{\mathbf{F}}_1^{\mathrm{rob}} & \mathbf{0}_{n_T \times (n_R - n_T)} \end{array}\right] \mathbf{U}^H \hat{\mathbf{R}}^{-1/2}, \quad (3.1)$$

where depending on the value of the KL divergence bound c, $\bar{\mathbf{F}}_1^{\mathrm{rob}}$ and the least favorable model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ take the following form.

i) For $c \geq \sum_{i=1}^{n_T} \sigma_i^2$, $\bar{\mathbf{F}}_1^{\mathrm{rob}} = \mathbf{0}_{n_T \times n_T}$ and the least-favorable channel model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ is an arbitrary matrix pair in $\mathcal{B}$.

ii) For $c < \sum_{i=1}^{n_T} \sigma_i^2$, the robust equalizer $\bar{\mathbf{F}}_1^{\mathrm{rob}}$ is a diagonal matrix of the form

$$\bar{\mathbf{F}}_1^{\mathrm{rob}} = \mathrm{diag}\{f_1^{\mathrm{rob}}, f_2^{\mathrm{rob}}, \ldots, f_{n_T}^{\mathrm{rob}}\} \quad (3.2)$$

with $f_i^{\mathrm{rob}} = f_i(\lambda_0)$, where for $1 \leq i \leq n_T$, the function $f_i(\lambda)$ can be expressed as

$$f_i(\lambda) = \sqrt[3]{-\frac{q_i}{2} + \sqrt{(\frac{q_i}{2})^2 + (\frac{p_i}{3})^3}}\, w^2 + \sqrt[3]{-\frac{q_i}{2} - \sqrt{(\frac{q_i}{2})^2 + (\frac{p_i}{3})^3}}\, w - \frac{\sigma_i}{3}, \quad (3.3)$$

with

$$p_i = -\frac{\sigma_i^2}{3} - 1 - \lambda(\sigma_i^2 + 1)$$
$$q_i = \frac{\sigma_i}{3}[1 + \lambda(\sigma_i^2 + 1)] + \lambda\sigma_i + \frac{2\sigma_i^3}{27} \quad (3.4)$$

and $w = \exp(j2\pi/3)$, and where the Lagrange multiplier λ_0 is obtained numerically by minimizing the scalar convex function $C(\lambda)$ in (3.18) using the steepest descent method. Given λ_0 and $\mathbf{F}^{\mathrm{rob}}$, the least favorable model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ can be expressed as

$$\mathbf{H}_\triangle^{\max} = -(\lambda_0 \hat{\mathbf{R}}^{-1} - \mathbf{F}^{\mathrm{rob}H}\mathbf{F}^{\mathrm{rob}})^{-1} \mathbf{F}^{\mathrm{rob}H}(\mathbf{I}_{n_T} - \mathbf{F}^{\mathrm{rob}}\hat{\mathbf{H}})$$
$$\mathbf{R}^{\max} = (\hat{\mathbf{R}}^{-1} - \frac{1}{\lambda_0}\mathbf{F}^{\mathrm{rob}H}\mathbf{F}^{\mathrm{rob}})^{-1}. \quad (3.5)$$

To solve the min-max problem (2.4), we first convert the min-max problem into a min-min problem by using the theory of Lagrangian multipliers. Since $J(\mathbf{F}, \mathbf{H}_\triangle, \mathbf{R})$ and the neighborhood $\mathcal{B}$ are convex in $(\mathbf{H}_\triangle, \mathbf{R})$, the maximum of J is achieved at the boundary of $\mathcal{B}$, and a local maximum is necessarily a global maximum. Then the Lagrangian associated with the maximization of J with respect to $(\mathbf{H}_\triangle, \mathbf{R})$ under the constraint $D(\mathbf{H}, \mathbf{R}; \hat{\mathbf{H}}, \hat{\mathbf{R}}) = c$ takes the form

$$L(\mathbf{F}, \mathbf{H}_\triangle, \mathbf{R}, \lambda) = J(\mathbf{F}, \mathbf{H}_\triangle, \mathbf{R}) + \lambda(c - D(\mathbf{H}, \mathbf{R}; \hat{\mathbf{H}}, \hat{\mathbf{R}})).$$

Then, according to Proposition 3.2.1 of [5], $(\mathbf{H}_\triangle, \mathbf{R})$ will be a local maximum of J under the equality constraint if there exists a finite Lagrange multiplier λ such that the following sufficient conditions are satisfied:

$$\nabla_{\mathbf{H}_\triangle} L = -2\Big(\mathbf{F}^H\Big[\mathbf{I}_{n_T} - \mathbf{F}(\hat{\mathbf{H}} + \mathbf{H}_\triangle)\Big] + \lambda\hat{\mathbf{R}}^{-1}\mathbf{H}_\triangle\Big) = 0 \quad (3.6)$$

$$\nabla_{\mathbf{R}} L = \mathbf{F}^H\mathbf{F} - \lambda(\hat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}) = 0 \quad (3.7)$$

$$\nabla_\lambda L = c - D(\mathbf{H}, \mathbf{R}; \hat{\mathbf{H}}, \hat{\mathbf{R}}) = 0 \quad (3.8)$$

$$\begin{bmatrix} \nabla^2_{\mathbf{H}_\triangle} L & \nabla^2_{\mathbf{H}_\triangle,\mathbf{R}} L \\ \nabla^2_{\mathbf{R},\mathbf{H}_\triangle} L & \nabla^2_{\mathbf{R}} L \end{bmatrix} = \begin{bmatrix} 2(\mathbf{F}^H\mathbf{F} - \lambda\hat{\mathbf{R}}^{-1}) & \mathbf{0}_{n_T\times n_R} \\ \mathbf{0}_{n_R\times n_T} & -\lambda\mathbf{R}^{-2} \end{bmatrix} < \mathbf{0}\,. \quad (3.9)$$

Note that the condition (3.9) is is equivalent to

$$\lambda > ||\mathbf{F}\hat{\mathbf{R}}^{1/2}||_2^2\,, \quad (3.10)$$

where $||\mathbf{M}||_2^2$ denotes the largest singular value of $\mathbf{M}\mathbf{M}^H$. In this case, the conditions (3.6) and (3.7) yield

$$\begin{aligned} \mathbf{H}_\triangle^{\max}(\lambda) &= -(\lambda\hat{\mathbf{R}}^{-1} - \mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H(\mathbf{I}_{n_T} - \mathbf{F}\hat{\mathbf{H}}) \\ \mathbf{R}^{\max}(\lambda) &= (\hat{\mathbf{R}}^{-1} - \frac{1}{\lambda}\mathbf{F}^H\mathbf{F})^{-1}\,. \end{aligned} \quad (3.11)$$

Then we form the "dual" function

$$\begin{aligned} &G(\mathbf{F},\lambda) = L(\mathbf{F}, \mathbf{H}_\triangle^{\max}(\lambda), \mathbf{R}^{\max}(\lambda), \lambda) \\ &= \lambda\Big[c + \mathrm{tr}\big\{(\mathbf{I}_{n_T} - \mathbf{F}\hat{\mathbf{H}})^H(\lambda\mathbf{I}_{n_T} - \mathbf{F}\hat{\mathbf{R}}\mathbf{F}^H)^{-1} \\ &\quad (\mathbf{I}_{n_T} - \mathbf{F}\hat{\mathbf{H}})\big\} - \ln\det(\mathbf{I} - \frac{1}{\lambda}\mathbf{F}\hat{\mathbf{R}}\mathbf{F}^H)\Big]\,. \end{aligned} \quad (3.12)$$

If we can find $\lambda^* > ||\mathbf{F}\hat{\mathbf{R}}^{1/2}||_2^2$ minimizing the dual function, the maximization problem in (2.4) is equivalent to a minimization problem, i.e., $\max_{(\mathbf{H}_\triangle,\mathbf{R})\in\mathcal{B}} J(\mathbf{F},\mathbf{H}_\triangle,\mathbf{R}) = \min_{\lambda > ||\mathbf{F}\hat{\mathbf{R}}^{1/2}||_2^2} G(\mathbf{F},\lambda)$.

We have therefore transformed the original min-max problem into the equivalent min-min problem

$$\min_{(\mathbf{F},\lambda)\in\mathcal{D}} G(\mathbf{F},\lambda) \quad (3.13)$$

over the domain $\mathcal{D} \stackrel{\triangle}{=} \{(\mathbf{F},\lambda)\ :\ \lambda > ||\mathbf{F}\hat{\mathbf{R}}^{1/2}||_2^2\}$. The set $\mathcal{D}$ is convex [6]. Since $G(\mathbf{F},\lambda)$ is obtained by maximizing the Lagrangian function $L(\mathbf{F},\mathbf{H}_\triangle,\mathbf{R},\lambda)$ which is convex in $(\mathbf{F},\lambda)$, by invoking Proposition 1.2.4 c) of [7], we conclude that the function $G(\mathbf{F},\lambda)$ is convex over $\mathcal{D}$. So we have reduced the robust MSE equalizer design problem to a conventional convex minimization problem.

To perform this minimization, we form the partitioned matrix $\bar{\mathbf{F}} \stackrel{\triangle}{=} \mathbf{V}\mathbf{F}\hat{\mathbf{R}}^{1/2}\mathbf{U} = [\bar{\mathbf{F}}_1\ \bar{\mathbf{F}}_2]$, where $\bar{\mathbf{F}}_1$ and $\bar{\mathbf{F}}_2$ have size $n_T \times n_T$ and $n_T \times (n_R - n_T)$. Then, we have

$$G(\mathbf{F},\lambda) = \lambda\Big[c + \mathrm{tr}\big\{\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\big\} - \ln\det(\frac{1}{\lambda}\mathbf{B})\Big]\,,$$

where $\mathbf{A} = \mathbf{I}_{n_T} - \bar{\mathbf{F}}_1\mathbf{\Sigma}$ and $\mathbf{B} = \lambda\mathbf{I}_{n_T} - \bar{\mathbf{F}}\bar{\mathbf{F}}^H$. G is minimized by setting $\bar{\mathbf{F}}_2 = 0$. For this choice the function G only depends on $\bar{\mathbf{F}}_1$. Then, the first-order Gateaux derivative (see [7, p. 17] for a definition of the Gateaux derivative) of $G(\bar{\mathbf{F}}_1,\lambda)$ with respect to $\bar{\mathbf{F}}_1$ in the direction of $\mathbf{Y}_1$ is given by

$$\begin{aligned} \nabla_{\bar{\mathbf{F}}_1,\mathbf{Y}_1} G(\bar{\mathbf{F}}_1,\lambda) = 2\lambda\mathrm{tr}\Big\{\big[&\bar{\mathbf{F}}_1^H\mathbf{B}^{-1}\mathbf{A}\mathbf{A}^H\mathbf{B}^{-1} - \\ &\mathbf{\Sigma}\mathbf{A}^H\mathbf{B}^{-1} + \bar{\mathbf{F}}_1^H\mathbf{B}^{-1}\big]\mathbf{Y}_1\Big\}\,. \end{aligned} \quad (3.14)$$

Since in this case $\mathbf{B}$ is invertible, we have $\nabla_{\bar{\mathbf{F}}_1,\mathbf{Y}_1} G = 0$ for all $\mathbf{Y}_1$ if and only if

$$\bar{\mathbf{F}}_1^H\mathbf{B}^{-1}\mathbf{A}\mathbf{A}^H - \mathbf{\Sigma}\mathbf{A}^H + \bar{\mathbf{F}}_1^H = \mathbf{0}\,. \quad (3.15)$$

Multiplying (3.15) on the left by $\lambda\mathbf{I}_{n_R} - \bar{\mathbf{F}}_1^H\bar{\mathbf{F}}_1$ we find

$$(\bar{\mathbf{F}}_1^H - \lambda\mathbf{\Sigma})(\mathbf{I}_{n_R} - \bar{\mathbf{F}}_1\mathbf{\Sigma})^H + (\lambda\mathbf{I}_{n_R} - \bar{\mathbf{F}}_1^H\bar{\mathbf{F}}_1)\bar{\mathbf{F}}_1^H = \mathbf{0}\,. \quad (3.16)$$

For a convex function $G(\bar{\mathbf{F}}_1,\lambda)$ over the convex domain $\mathcal{D}$, if we can find $(\bar{\mathbf{F}}_1^{\mathrm{rob}},\lambda_0) \in \mathcal{D}$, which satisfies (3.16), then $(\bar{\mathbf{F}}_1^{\mathrm{rob}},\lambda_0)$ will be a global solution of the minimization problem. In order to solve (3.16), we assume that $\bar{\mathbf{F}}_1 = \mathrm{diag}\{f_i,\ 1 \le i \le n_T\}$. In this case, the pair $(\bar{\mathbf{F}}_1,\lambda)$ will be in the domain $\mathcal{D}$ provided $\lambda > f_i^2$ for all $1 \le i \le n_T$. Then (3.16) becomes decoupled cubic equations of the form

$$f_i(f_i^2 + f_i\sigma_i - 1) - \lambda\big(f_i(\sigma_i^2+1) - \sigma_i\big) = 0\,. \quad (3.17)$$

The root locus of this third order system is shown in Fig. 1. The roots $f_i(\lambda)$ corresponding to the left and right branches go from $1/2[-\sigma_i \pm (\sigma_i^2+4)^{1/2}]$ to ∞ as $\lambda \to \infty$, but the root corresponding to the middle branch goes from $f_i = 0$ at $\lambda = 0$ to $\sigma_i/(\sigma_i^2+1)$ as $\lambda \to \infty$. We can prove [6] that a root f_i is located in the domain $\lambda > f_i^2$ only if it belongs to the interval $(0,\ \sigma_i/(\sigma_i^2+1))$, so that only the middle branch of the root locus yields a solution $f_i(\lambda)$ of the cubic equation (3.17) such that the diagonal equalizer $\bar{\mathbf{F}}_1$ is in the domain $\mathcal{D}$.

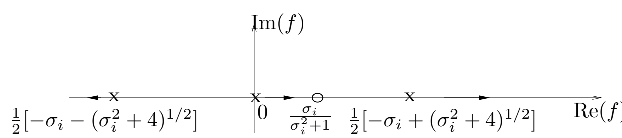

Fig. 1. Root locus of equation (3.17): poles and zeros are represented by x's and o's.

Then, let

$$\begin{aligned} C(\lambda) = \min_{\bar{\mathbf{F}}\ :\ \lambda > ||\bar{\mathbf{F}}||_2^2} G(\bar{\mathbf{F}},\lambda) = \lambda\Big[c + \\ \sum_{i=1}^{n_T}\frac{(1 - f_i(\lambda)\sigma_i)^2}{\lambda - f_i^2(\lambda)} - \sum_{i=1}^{n_T}\ln(1 - \frac{f_i^2(\lambda)}{\lambda})\Big] \end{aligned} \quad (3.18)$$

denote the function obtained by partial minimization of G with respect to $\bar{\mathbf{F}}$, where $f_i(\lambda)$ denotes the root of (3.17) corresponding to the middle branch of the root locus. Since $G(\bar{\mathbf{F}},\lambda)$ is convex over $\mathcal{D}$, according Proposition 2.3.6 of

[7], the function $C(\lambda)$ obtained by partial minimization of $G(\bar{\mathbf{F}}, \lambda)$ with respect to $\bar{\mathbf{F}}$ is a convex in λ.

In order to find the minimum of $C(\lambda)$, we need to explore the graph of $C(\lambda)$ depending on the value of the uncertainty bound c. We examine first the behavior of $C(\lambda)$ in the vicinity of $\lambda = 0$. Close to $\lambda = 0$, we can expand the root locus solution of (3.17) starting at $f_i = 0$ as

$$f_i(\lambda) = \mu_1 \lambda + \mu_2 \lambda^2 + \dots \,. \quad (3.19)$$

Substituting (3.19) into (3.17) and matching coefficients of λ and λ^2 yields $\mu_1 = -\mu_2 = \sigma_i$. So for small values of λ, we have $f_i(\lambda) = \sigma_i \lambda - \sigma_i \lambda^2$ and $\lim_{\lambda \to 0} \frac{dC}{d\lambda} = c - \sum_{i=1}^{n_T} \sigma_i^2$. Then depending on the value of the uncertainty bound c, we can identify two different cases.

1) When $c \geq \sum_{i=1}^{n_T} \sigma_i^2$, the convex function $C(\lambda)$ is minimized when $\lambda = 0$, so that $f_i^{\text{rob}} = 0$ for $0 \leq i \leq n_T$ and the least favorable channel model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ is an arbitrary matrix pair in $\mathcal{B}$.

2) When $c < \sum_{i=1}^{n_T} \sigma_i^2$, $C(\lambda)$ is minimized at a point $\lambda_0 > 0$. λ_0 can be found numerically by using the method of steepest descent. Then the least favorable model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ is obtained by substituting λ_0 and $\bar{\mathbf{F}}_1^{\text{rob}}$ inside (3.11).

This completes the proof of Theorem 1.

4. NUMERICAL SIMULATIONS

In the simulations shown below, we assume uncoded QPSK symbols are transmitted over a random flat fading MIMO channel with independent normalized complex Gaussian entries and the estimated channel noise covariance $\hat{\mathbf{R}} = \sigma_v^2 \mathbf{I}_{n_R}$. To find the least-favorable channel model, we apply the normalized KL divergence bound $\frac{D(\mathbf{H},\mathbf{R},\hat{\mathbf{H}},\hat{\mathbf{R}})}{\text{tr}\{\hat{\mathbf{H}}^H \hat{\mathbf{R}}^{-1} \hat{\mathbf{H}}\}} \leq b$. In the plot, the SNR (in dB) is defined as the total received signal power over the total noise power. Fig 2 compares the MSE performance of robust equalizers and MMSE equalizers over their corresponding least-favorable channel models for a 2-input-4-output system when b takes the values 0.01, 0.05 and 0.1. As we expected, the figure shows that the robust equalizer $\mathbf{F}^{\text{rob}}$ over its worst-case channel model $(\mathbf{H}_\triangle^{\max}, \mathbf{R}^{\max})$ achieves a smaller MSE than the MMSE equalizer $\mathbf{F}^{\text{MMSE}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\mathbf{R}})^{-1}$ over its worst-case channel model $(\mathbf{H}_\triangle^{\text{MMSE}}, \mathbf{R}^{\text{MMSE}})$.

5. CONCLUSION

We have considered the design of robust MSE equalizers for MIMO communications systems with imperfect channel and noise knowledge. Our results are applicable to channels with $n_T \leq n_R$ which have full column rank. The design problem is formulated as a min-max problem where the goal is to find the equalizer minimizing the equalization mean square error for the least favorable channel model within KL divergence bound. Lagrangian optimization is used to transform the min-max problem into a convex min-min problem over a convex domain, to which standard convex optimization methods apply.

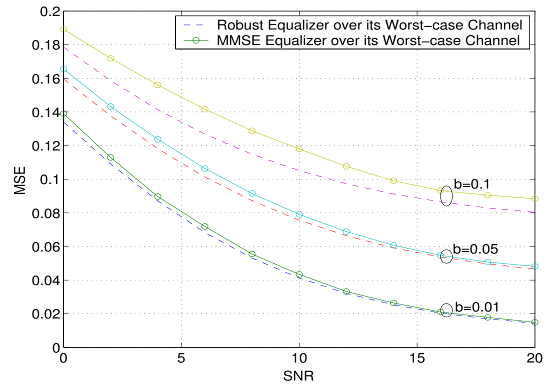

Fig. 2. MSE performance of robust equalizers $\mathbf{F}^{\text{rob}}$ and MMSE equalizers $\mathbf{F}^{\text{MMSE}}$ for their corresponding least favorable channel models.

6. REFERENCES

[1] B. C. Levy and R. Nikoukhah, "Robust least-squares estimation with a relative entropy constraint," *IEEE Trans. Informat. Theory*, vol. 50, pp. 89–104, Jan. 2004.

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

[3] A. H. Sayed, V. H. Nascimento, and F. A. M. Cipparone, "A regularized robust design criterion for uncertain data," *SIAM J. Matrix Anal. Appl.*, vol. 23, pp. 1120–1142, Apr. 2002.

[4] M. Basseville, "Information: Entropies, divergences et moyennes," Tech. Rep. 1020, Institut de Recherche en Informatique et Systèmes Aléatoires, Rennes, France, May 1996.

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

[6] Y. Guo, *Precoding and Equalization for MIMO Communication Systems*, Ph.D. thesis, Dept. of Elec. and Comp. Eng., UC Davis, Aug. 2004.

[7] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, Mass, 2003.