LMI BASED DESIGN OF OPTIMAL PRECODER FOR MIMO CHANNELS WITH ERASURES

Dimuthu Jayasingha, Jingxin Zhang

Department of Electrical and Computer Systems Engineering, Monash University
Clayton, VIC 3800, Australia. Email:dimuthu.jayasingha@monash.edu, jingxin.zhang@monash.edu

ABSTRACT

This paper is concerned with the design of optimal linear precoder for MIMO communication systems with channel erasures. The erasures are modeled as Bernoulli process with known probability distribution and are incorporated explicitly in design formulation. The linear correlating transform precoding and redundant multiple channels are used to enhance the robustness to erasures, and the linear minimum mean square error (MMSE) estimation is used to reconstruct the source signal at the receiver end. The design of optimal precoder is formulated as minimizing the expected value of estimation distortion, and the solution is obtained by converting the nonconvex optimization problem into a convex optimization problem subject to linear matrix inequality (LMI) constraints. Numerical examples demonstrate the effectiveness and advantage of the obtained solution in enhancing the robustness to channel erasures.

***Index Terms*—** Erasure channel, Precoder, MIMO, MMSE, LMI

1. INTRODUCTION

In transmitting signal over wireless medium, often the transmission is subjected to interference and fading which causes channel erasures. Though current communication systems are capable of delivering reliable transmission to many applications, it is still a question whether the performance is adequate for the systems (e.g. networked control systems) that require time critical responses. For example, TCP based communication protocols overcome the channel erasure by retransmission of lost packets. But the delay due to retransmission can has an adverse effect on the system performance.

One of the solutions to the above problem is to introduce redundancy in transmission by precoding the source signal. The redundant information transmitted through erasure channels enhances the robustness to channel loss. In recent years, there has been a significant research into transmitting redundant information to enhance the reconstruction of source signal in the event of channel loss [1, 2].

Multiple description coding (MDC) has proven to be an effective method to recover the source signal after multiple erasures [3]. In MDC a single data stream is encoded to multiple parallel streams such that the information of the source is spread across the coded descriptions. The encoded data is transmitted over multiple independent paths, so the receiver can recover the source even when only one description is received.

One approach to MDC is through correlating transform where the statistical dependencies of the transmitted coefficients are increased by a linear transformation. The optimal correlating transform coding for a 2×2 coder for a quantized source is presented in [3, 4]. Though these do not explicitly consider channel erasure process it is suggested that the increase in cross correlation due to linear precoding enhances robustness to channel loss. The general case of $N \times N$ (N sources and N coder outputs) optimal correlation transform to overcome the channel erasures based on Wiener filter has been studied in [5]. It has been extended to represent the redundant precoding, i.e. $M \times N$ coder (N sources and M coder outputs with $M > N$) in [6] with MMSE estimator at the receiver for source signal reconstruction. In [6] the redundant precoder is designed by repeating the eigenvectors in the source covariance. Both the works [5] and [6] propose using the gradient based computation to design the optimal precoder. Since the objective functions used in these designs are nonconvex, this computation method can be problematic in convergence and efficiency.

Another approach to redundant information transmission is to expand the source signals by an overcomplete frame and then transmit the frame coefficients across multiple channels. This approach uses the dual frame (left pseudo inverse) of the expansion (analysis) frame at receiver side to reconstruct source signal, and can achieve perfect reconstruction if the total channel losses are less than the redundancy of the frame [7, 8].

In this paper we investigate the design of optimal correlating transform precoder for MIMO channels with probabilistic erasures. Similarly to [6], we model the channel erasures as Bernoulli random process, use the linear MMSE estimation for source signal reconstruction under erasures, and use the expected value of estimation distortion as the design objective function. In contrast to [6], we explicitly incorporate the MIMO channel model in problem formulation, convert the nonconvex objective function into a convex objective function subject to LMI constraints, and hence derive a convex optimization based design that can be efficiently and reliably computed.

The notations of the paper are standard. For a matrix A, its transpose and trace are denoted by A^T and $Tr(A)$. I_N denotes the $N \times N$ identity matrix.

2. PROBLEM STATEMENT

Figure 1 gives an MIMO communication system with M transmitters and K receivers, where $K \geq M$. The source signal $x_k \in \mathbb{R}^N$ is a non-white Gaussian vector with zero mean and covariance $E(xx^T) = R_x$. $T \in \mathbb{R}^{M \times N}$ is a linear precoder that encodes the N dimensional source vector x_k to M correlated descriptions $y_k \in \mathbb{R}^M$, with $M > N$. Since $M > N$, precoding adds redundancy to the transmission. The coded descriptions are transmitted through M transmitter antennas and received by K receiver antennas. The communication channel $H \in \mathbb{R}^{K \times M}$ is a constant matrix with known parameters and full column rank M.

The MIMO channel is subject to erasures. In this work we assume that the loss of a channel is equivalent to a receiver failure. That is, the loss of subchannel i results in null output at the receiver i. As a result of channel erasure the decoder receives L descriptions, with $L \leq K$. The matrix $P_e \in R^{L \times K}$ denotes the channel erasure state. $P_e = I_K$ when there are no erasure events. In an event of erasure the respective rows from P_e are removed to denote the subchannel losses. The number of received descriptions, L, depends on the channel state at the time. For K receivers, $L \in [1, 2, \cdots, K]$ and hence there are total 2^K channel states. At the receiver the MMSE estimator $V_e \in R^{N \times L}$ reconstructs source signal based on the signal received after the erasure.

We further assume that the source quantization is sufficiently fine such that the effect of the quantization noise is insignificant to the signal reconstruction at the receiver. Also the transmitted coefficients are assumed to be either completely lost or received without error. The receiver has the full information of the channel matrix H and the channel state information P_e at each transmission. But the channel and erasure information is not available at the transmitter during the communication.

In the sequel, we will investigate the design of the precoder T and the MMSE estimator V_e for the minimization of signal reconstruction error in the MIMO communication system described above.

3. PRECODER DESIGN FORMULATION

From the system in Figure 1 the signal received by the estimator is

$$z_k = P_e H T x_k \tag{1}$$

and the estimator output is

$$\hat{x}_k = V_e z_k. \tag{2}$$

There are two possible cases of equation (1) depending on the descriptions received by the receivers. Case 1: The number of received descriptions, L, satisfies $N \leq L \leq K$. In this case, the matrix $P_e HT$ is a tall matrix with the column rank equal to N. Hence, (1) is solvable for x_k by taking the left pseudo inverse of $P_e HT$, and the estimator V_e is simply $V_e = (T^T H^T P_e^T P_e H T)^{-1} T^T H^T P_e^T$. It yields $\hat{x}_k = x_k$ with no reconstruction error. Case 2: The number of received descriptions $L < N$. In this case, the matrix $P_e HT$ is a fat matrix with the rank equal to L. The equation (1) is no longer solvable for x_k by using the left pseudo inverse of $P_e HT$, since it does not exist. In this situation we will estimate x_k using the MMSE estimation described in Appendix.

Since x_k has zero mean and covariance R_x, z_k also has zero mean and covariance $R_z = P_e H T R_x T^T H^T P_e^T$. The cross covariances of x_k and z_k are $\sum_{xz} = R_x T^T H^T P_e^T$ and $\sum_{zx} = P_e H T R_x$. Then by the MMSE estimation method given in Appendix, the estimate of x_k knowing z_k is given by

$$\hat{x}_k = R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} z_k \tag{3}$$

and the MMSE estimator is given by

$$V_e = R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1}. \tag{4}$$

The covariance and distortion (energy) of the estimation error are respectively

$$R_e = R_x - R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x \tag{5}$$

and

$$D_e = Tr[R_x - R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x]. \tag{6}$$

Above derivation is for a specific state of channel erasure matrix P_e. Since P_e is a stochastic variable that depends on the erasure process, it can be modeled as a Bernoulli process. Denote λ the probability of a single channel erasure at the receiver, and let w_e be the probability of the channel in a particular (erasure) state. Then $w_e = \lambda^{K-L} \times (1-\lambda)^L$, where L is the number of received descriptions and $L = 1, 2, \cdots, K$.

Since the channel can be in different states depending on the erasure, the precoder should be designed to address all possible erasure states. Therefore the expected value of the estimation distortion (energy of estimation error) is used as the criterion for precoder design. Excluding Case 1 discussed above, the expected value of distortion is given by

$$\begin{aligned} D = \sum_{e=1}^{E} w_e \times D_e = \sum_{e=1}^{E} & w_e Tr[R_x - R_x T^T H^T P_e^T \\ & (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x]. \\ = Tr \sum_{e=1}^{E} w_e & \Big[R_x - R_x T^T H^T P_e^T \\ & (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x \Big]. \end{aligned} \tag{7}$$

where $E := 2^K - \sum_{j=N}^{K} \begin{pmatrix} K \\ j \end{pmatrix}$.

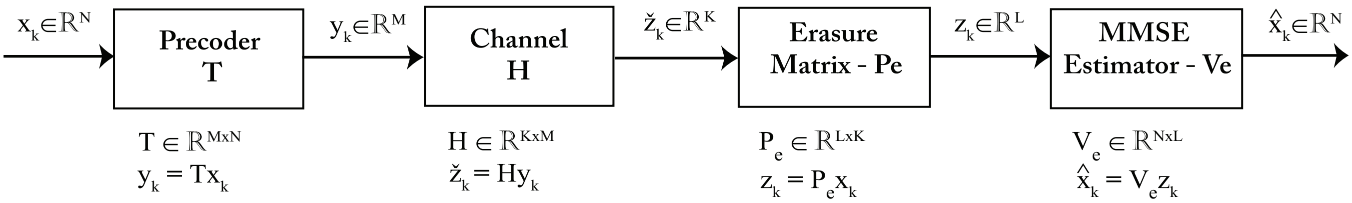


Fig. 1. System Configuration

Thus the optimal precoder for MMSE estimation can be obtained from the following optimization

$$\min_T D = \min_T Tr\left\{\left[\sum_{e=1}^{E} w_e\, R_x - R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x\right]\right\}. \quad (8)$$

The optimal precoder T that solves (8) is a full column rank matrix.

A problem that affects the solvability of the minimization in (8) is the scenario of 'total channel loss', which is when information on all the subchannels is lost. Since the erasure matrix P_e loses the respective rows corresponding to the lost subchannels in modeling, this situation will create a null erasure matrix P_e and hence the distortion will tend to infinity. Therefore, (8) will be dominated by the 'total channel loss' scenario. The probability of 'total channel loss', λ^K, is very low relative to other channel states; also, no precoder design can effectively deal with such scenario. Based on these facts, this scenario is excluded from the precoder design to give the following design formulation

$$\min_T \quad D = \min_T \quad Tr\left\{\tilde{\Lambda}_c R_x - \sum_{e=1}^{\tilde{E}} w_e \left[R_x T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T R_x\right]\right\} \quad (9)$$

subject to *rank*$[T] = N$, where $\tilde{\Lambda}_c := \sum_{e=1}^{\tilde{E}} w_e$ and $\tilde{E} = E - 1$.

4. CONVEX EQUIVALENT OF DESIGN OPTIMIZATION

The objective function in (9) is a nonlinear and nonconvex function of the optimization variable T, and hence its minimization is a rather difficult problem. We will present in this section a convex relaxation approach to solving this nonlinear and non-convex problem.

Let $U^T U = R_x$ be the Cholesky decomposition of R_x and define

$$S := T R_x T^T \quad (10)$$

$$W := \sum_{e=1}^{\tilde{E}} w_e \left[U T^T H^T P_e^T (P_e H T R_x T^T H^T P_e^T)^{-1} P_e H T U^T\right] = \sum_{e=1}^{\tilde{E}} w_e \left[U T^T H^T P_e^T (P_e H S H^T P_e^T)^{-1} P_e H T U^T\right] \quad (11)$$

Substituting (10) and (11) into (9), the design optimization can be equivalently written as

$$\min_T \quad D = \min_T \quad Tr\{\tilde{\Lambda}_c R_x - U^T W U\}$$

subject to

$$W = \sum_{e=1}^{\tilde{E}} w_e \left[U T^T H^T P_e^T (P_e H S H^T P_e^T)^{-1} P_e H T U^T\right],$$

$$S = T R_x T^T, \quad S \geq 0, \quad S = S^T, \quad W \geq 0, \quad W = W^T.$$

Using the convex relaxation [9] and Schur complement [10]

$$S \geq T R_x T^T \iff \begin{bmatrix} S & TU^T \\ UT^T & I_N \end{bmatrix} \geq 0,$$

$$W = \sum_{e=1}^{\tilde{E}} w_e \left[U T^T H^T P_e^T (P_e H S H^T P_e^T)^{-1} P_e H T U^T\right] \iff W \geq \sum_{e=1}^{\tilde{E}} w_e \left[U T^T H^T P_e^T (P_e H S H^T P_e^T)^{-1} P_e H T U^T\right],$$

the design optimization can be further converted to

$$\min_{T,W,S} \quad Tr(\tilde{\Lambda}_c R_x - U^T W U) + \delta_S Tr(S) + \delta_W Tr(W) \quad (12)$$

subject to

$$\tilde{\Lambda}_c R_x - U^T W U \geq 0, \quad W \geq 0, \quad W = W^T, \quad (13)$$

$$\begin{bmatrix} S & TU^T \\ UT^T & I_N \end{bmatrix} \geq 0, \quad S \geq 0, \quad S = S^T, \quad (14)$$

$$\begin{bmatrix} W & \sqrt{\omega_1} U T^T H^T P_1^T & \sqrt{\omega_2} U T^T H^T P_2^T & \cdots & \sqrt{\omega_{\tilde{E}}} U T^T H^T P_{\tilde{E}}^T \\ \sqrt{\omega_1} P_1 H T U^T & P_1 H S H^T P_1^T & 0 & \cdots & 0 \\ \sqrt{\omega_2} P_2 H T U^T & 0 & P_2 H S H^T P_2^T & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{\omega_{\tilde{E}}} P_{\tilde{E}} H T U^T & 0 & 0 & \cdots & P_{\tilde{E}} H S H^T P_{\tilde{E}}^T \end{bmatrix} \geq 0 \quad (15)$$

where $\delta_S, \delta_T > 0$ are small positive numbers for reducing the variable slack of W and S. The design optimization (12)-(15) is a standard convex optimization problem subject to linear matrix inequality constraints, which can be solved for globally optimal solution by using the interior point method in MATLAB LMI and CVX toolboxes [11].

Scaling for maximal power transmission: It can be seen from (9) that the coefficients of T may have arbitrarily small values without affecting the average distortion D. Therefore the transmission power of the resultant precoder may become very small. To overcome this problem, one can scale the optimal precoder T by $\alpha = \sqrt{P_{TX}/P_{out}}$ and use $T_\alpha = \alpha T$ as the optimal precoder. Here P_{TX} is the maximal total transmission power physically allowed, and $P_{out} = Tr[y\ y^T] = Tr[TR_xT^T]$ is the transmission power of the precoder T obtained from solving (12)-(15). The scaled optimal precoder T_α yields the maximal transmission power physically allowed. From the second line of (9), it can be readily seen that this scaling does not affect the distortion D.

5. NUMERICAL EXAMPLE

Example 1 $[N = 2, M = 3, K = 3]$*:* Consider the source correlation matrix and the channel matrix

$$R_x = \begin{bmatrix} 0.78 & 0.12 \\ 0.12 & 0.78 \end{bmatrix}, \quad H = \begin{bmatrix} 0.4 & 0.7 & 0.9 \\ 0.6 & 0.35 & 0.2 \\ 0.5 & 0.15 & 0.45 \end{bmatrix}$$

with the erasure probability $\lambda = 0.2$. The channel loss states are $P_1 = [1\ 0\ 0]$, $P_2 = [0\ 1\ 0]$ and $P_2 = [0\ 0\ 1]$ for $L = 1$. For this channel, an optimal precoder T is obtained from the optimization (12)-(15) with small δ_S and δ_W. The optimization is computed using MATLAB CVX toolbox. The optimal precoder T is scaled up to achieve the maximal total transmission power of 5watt. The scaled optimal precoder is shown below.

$$T_\alpha^T = \begin{bmatrix} 0.4544 & -0.6408 & -1.4695 \\ 0.4532 & -0.6424 & -1.4700 \end{bmatrix}.$$

The optimal precoder T_α attains an average distortion of $D_T = 0.0634$ and equal side distortions of 0.66.

The optimal precoder T_α is compared with the Mercedes-Benz frame expansion precoder and the uniform length, non-Parseval frame expansion precoder given in [8], and the results are shown in Figure 2. The optimal precoder has better performance in general than the frame expansion precoders. Also, the performance of the optimal precoder is compared with 50,000 random transform precoders of the same size. As shown in Figure 3, the optimal precoder T_α always performances better than those of the random precoders. In contrast to the unpredictable performances of random precoders, the performance of optimal precoder is always guaranteed by design.

Example 2 $[N = 3, M = 4, K = 4]$*:* Consider the source correlation matrix and the channel matrix

$$R_x = \begin{bmatrix} 0.78 & 0.12 & 0.54 \\ 0.12 & 0.78 & 0.12 \\ 0.54 & 0.12 & 0.78 \end{bmatrix}, \quad H = \begin{bmatrix} 0.4 & 0.7 & 0.9 & 0.55 \\ 0.6 & 0.35 & 0.2 & 0.95 \\ 0.5 & 0.15 & 0.45 & 0.8 \\ 0.35 & 0.8 & 0.55 & 0.6 \end{bmatrix}$$

with erasure probability $\lambda = 0.2$. The channel loss states are $P_1 = [1\ 0\ 0\ 0]$, $P_2 = [0\ 1\ 0\ 0]$, $P_3 = [0\ 0\ 1\ 0]$, $P_4 = [0\ 0\ 0\ 1]$ for $L = 1$ and $P_5 = \left[\begin{smallmatrix} 1&0&0&0 \\ 0&1&0&0 \end{smallmatrix}\right]$, $P_6 = \left[\begin{smallmatrix} 1&0&0&0 \\ 0&0&1&0 \end{smallmatrix}\right]$, $P_7 = \left[\begin{smallmatrix} 1&0&0&0 \\ 0&0&0&1 \end{smallmatrix}\right]$, $P_8 = \left[\begin{smallmatrix} 0&1&0&0 \\ 0&0&1&0 \end{smallmatrix}\right]$, $P_9 = \left[\begin{smallmatrix} 0&1&0&0 \\ 0&0&0&1 \end{smallmatrix}\right]$, $P_{10} = \left[\begin{smallmatrix} 0&0&1&0 \\ 0&0&0&1 \end{smallmatrix}\right]$, for $L = 2$. By solving (12)-(15) the optimal precoder with maximal total transmission power 10watt is obtained

$$T_\alpha^T = \begin{bmatrix} 0.7881 & 1.0404 & 0.7958 & 1.0038 \\ 0.1615 & 0.1478 & -0.0059 & 0.8473 \\ 0.8581 & 1.0380 & 0.7933 & 0.9655 \end{bmatrix}.$$

The performance of this preocder against 50,000 random precoders are shown in Figure 4.

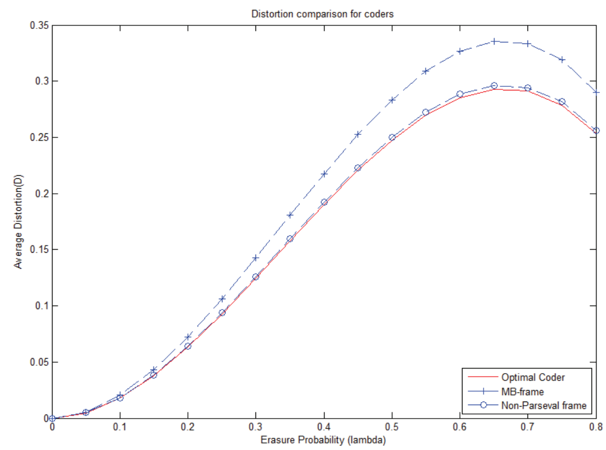


Fig. 2. Comparison with the frame expansion precoder

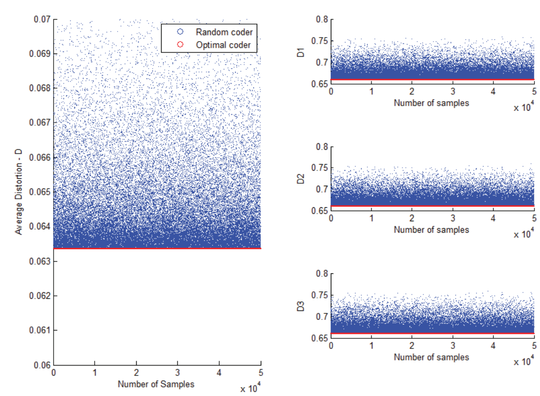
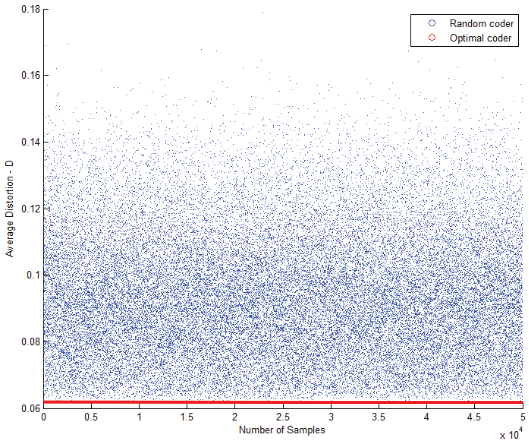


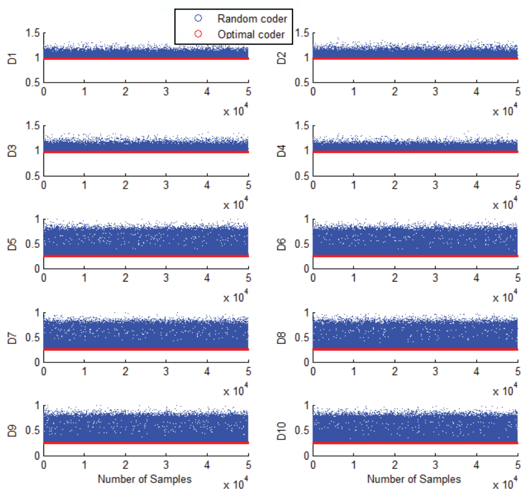
Fig. 3. Comparison of 3×2 optimal precoder with random precoders

6. CONCLUSION

A novel method has been presented to design the optimal precoder for MIMO communication systems subject to channel erasures. The method is in the form of convex optimization subject to LMI constraints, and hence it can be easily computed by the well developed computation tools, such as the interior point method, in MATLAB and other convex optimization software. Numerical examples have



(a) Average distortion



(b) Side distortion

Fig. 4. Comparison of 4×3 optimal precoder with random coders

demonstrated the effectiveness and advantage of the design method in enhancing the robustness to channel erasures. Because the design is dependant on the selection of the weighting factors δ_S and δ_W, the optimal precoder is not unique.

7. APPENDIX (MMSE ESTIMATION)

Conditional PDF of Multivariate Gaussian [12]: Let X and Y be jointly Gaussian random variables with means $\bar{x}$ and $\bar{y}$ and variances Σ_x and Σ_y respectively . Let $Z = [X'Y']'$, then Z is Gaussian variable with mean and covariance of $\left(\begin{smallmatrix} \bar{x} \\ \bar{y} \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{smallmatrix}\right)$ respectively. Then X conditioned on $Y = y$ is also a Gaussian distribution with mean $\bar{x} + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \bar{y})$ and covariance $\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$.

Posterior PDF for Baysian Linear Model [12]: For the observed data model of $y = Hx + w$ where $y \in R^N$ is the data vector, $H \in R^{N \times p}$ is a known matrix, $x \in R^p$ is a random vector with prior PDF $\mathcal{N}(\mu_x, C_x)$ and $w \in R^N$ is a noise vector independent of x with PDF $\mathcal{N}(0, C_w)$. Then the posterior PDF is Gaussian with mean $\mu_x + C_x H^T (HC_x H^T + C_w)^{-1}(y - H\mu_x)$ and covariance $C_x - C_x H^T (HC_x H^T + C_w)^{-1} HC_x$. The linear MMSE estimator is the mean of the posterior PDF and the error variance of the estimator is the covariance of the posterior PDF.

8. REFERENCES

[1] A.R. Mesquita, J.P. Hespanha, and G. Nair, "Redundant data transmission in control/estimation over wireless networks," in *American Control Conference, 2009. ACC '09.*, 2009.

[2] R. Blind, S. Uhlich, B. Yang, and F. Allgower, "Robustification and optimization of a kalman filter with measurement loss using linear precoding," in *American Control Conference, 2009. ACC '09.*, 2009.

[3] V.K Goyal, "Multiple description coding: compression meets the network," *Signal Processing Magazine, IEEE*, vol. 18, pp. 74 –93, September 2001.

[4] V.K Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *Information Theory, IEEE Transactions on*, vol. 47, pp. 2199 –2224, September 2001.

[5] G. Romano, P. Salvo Rossi, and F. Palmieri, "Optimal correlating transform for erasure channels," *Signal Processing Letters, IEEE*, vol. 12, pp. 677 – 680, October 2005.

[6] S. Uhlich and Bin Yang., "A generalized optimal correlating transform for multiple description coding and its theoretical analysis," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008, pp. 2969 –2972.

[7] J. Kovacevic, P.L. Dragotti, and V.K. Goyal, "Filter bank frame expansions with erasures," *Information Theory, IEEE Transactions on*, vol. 48, pp. 1439 –1450, jun 2002.

[8] J. Leng, D. Han, and T. Huang, "Optimal dual frames for communication coding with probabilistic erasures," *Signal Processing, IEEE Transactions on*, vol. 59, pp. 5380 –5389, nov. 2011.

[9] Y. Kehu, S. Cai, and Z.Q. Luo, "Convex relaxation approaches to maximum likelihood doa estimation in ula's and uca's with unknown mutual coupling," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 2556 –2559.

[10] S. Boyd, L.E. Ghaoui, E. Feron, and V. Balakrishnan, "Linear matrix inequalities in system and control theory," *SIAM*, 1994.

[11] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," `http://cvxr.com/cvx/`, Apr. 2011.

[12] S. M. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory," *Englewood Cliffs, N.J. : PTR Prentice-Hall*, 1993.