Channel Estimation Using Feedback

Gilad Bukai
Gzeht Investments
Tel Aviv, Israel
Email: bgilad2002@gmail.com

Neri Merhav
Technion - Israel Institute of Technology
Haifa, Israel
Email: merhav@ee.technion.ac.il

***Abstract*—It is well known from the Information Theory literature that in the presence of a memoryless channel, as well as some channels with memory, the existence of feedback does not improve channel capacity, but can contribute dramatically to the simplification of the encoder and the decoder. In this paper, we focus on estimating the parameters of an unknown channel when a feedback link from the decoder to the encoder is given. We examine different types of channels and study the conditions under which feedback improves the estimation in the sense of the Cramér-Rao bound (CRB). We additionally find what is the best strategy for designing channel input signals and the channel parameter estimator of the receiver. By assuming a certain class of parametric models for the channel, we show that feedback is essential when some nuisance parameters of the model are unknown, but it is superfluous when there are no unknown nuisance parameters.**

I. Introduction

The need to estimate the parameters of a communication channel, in order to improve decoder performance and communication reliability, is present in many communication applications. Some of present-day techniques use a training phase, in which the transmitter sends a known signal to the receiver, allowing the latter to estimate the channel based on the received signal. Several works deal with the design of the training sequence. In [1] least squared filtering was used to determine the optimal time domain training sequences, which minimize the mean squared error (MSE) of the estimation. Optimal sequences for frequency domain estimation, which minimize the variance of the estimation error, can be found in [2]. In [3], it was shown that a time domain channel estimator achieves superior MSE compared to a frequency domain estimator. In this paper, we focus on the estimation of the parameters of the communication channel with feedback and under the constraint of limited transmission power. One approach in parameter estimation is to divide the parameter set into two groups: the group of parameters we wish to estimate and an unknown nuisance parameters group. The subject of finding optimal channel inputs for different kinds of channels was deeply discussed in the past, both in the presence of nuisance parameters and without them. Specifically, for the case of an intersymbol interference (ISI) channel with an additive white noise, it was proved that the optimal input channel has flat spectrum and for the case where the noise is colored [4], the optimal signal transmitted should use the frequency at which the noise power spectral density (psd) function is minimum. Several works used feedback in order to design optimal channel inputs. In [5], it was proved that the optimal input channel for the case of all-pole model structure is given by a minimum variance feedback regulator with a white noise disturbance. In [6], it was proved that for the case where feedback is given, the optimal channel input, which minimizes a weighted CRB, has no correlation with the channel noise. The optimal training sequence set, for the case of MIMO channels with jamming signals, was obtained in [7]. It was shown that the optimal training depends on the eigenvalues of the correlation matrix of the jamming signals. When this matrix is unavailable to the transmitter, a feedback link was used to convey the estimated correlation matrix.

This work is part of the M.Sc. thesis of G. Bukai, carried out in the Technion - I.I.T., Haifa, Israel.

The main novelty in this paper is that it examines the necessity of feedback, both in the presence of nuisance parameters and without them, for the purpose of optimal channel input design, where the channel is modeled as an ISI channel with colored noise. We show that feedback is essential to achieve optimal error performance for the case where nuisance parameters are present and we show under which conditions, the error performance for the case where all the nuisance parameters are known, is asymptotically achievable by using feedback.

II. Notation, Problem Formulation and Preliminaries

Throughout this paper, scalar random variables (RV's) are represented by capital letters and their sample values by the respective lower case letters. A similar convention applies to random vectors and their sample values, represented by upper and lower case symbols in bold face font. Thus, for example, a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ may assume a specific vector value $\boldsymbol{x} = (x_1, \ldots, x_n)$. All vectors are column vectors (row vectors are indicated by the transposition sign T in superscript; e.g, $\boldsymbol{X}^T$ is the transpose of $\boldsymbol{X}$). $\boldsymbol{X}^d$ denotes a vector of dimension d whose components are the first d RV's of a random process $\{X_i\}$, i.e, $\boldsymbol{X}^d = [X_1, X_2, \cdots, X_d]^T$. The identity matrix is denote by I. The expectation operator is denote by $E\{\cdot\}$. The cross correlation matrix of two random vectors, $\boldsymbol{X}^N$ and $\boldsymbol{Y}^N$, will be represented by: $(\boldsymbol{R}_{xy})_{ij} = \frac{1}{N}\sum_{n=1}^{N} E\left[X(n-i)Y(n-j)\right]$.

The most general model we consider, in this paper, is illustrated in Fig. 1 and is represented by:

$$y_n = \sum_{i=0}^{M-1} \theta_i x_{n-i} + z_n, \tag{1}$$

where $\{x_n\}$ is the transmitted signal, $\{y_n\}$ is the channel output, and the colored noise $\{z_n\}$ is generated by:

$$z_n = \sum_{k=1}^{p} c_k z_{n-k} + v_n, \tag{2}$$

where $\{v_n\}$ is an i.i.d. process, with a marginal density function: $g_\sigma(v) = \frac{1}{\sigma} f\left(\frac{v}{\sigma}\right)$, where σ is a positive scaling parameter, $f(\cdot)$ is a known density with $\int_{-\infty}^{\infty} v f(v) dv = 0$ and $\int_{-\infty}^{\infty} v^2 f(v) dv = 1$. It is assumed that the second derivative of $f(\cdot)$ exists and it has finite expectation.

We assume that, at time instant n, the channel input, x_n, may depend on the past channel outputs and the past channel inputs, $\boldsymbol{y}^{n-1}$ and $\boldsymbol{x}^{n-1}$, respectively, through a conditional density $f_n(x_n|\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1})$ and in addition, $(\boldsymbol{x}^n, \boldsymbol{y}^n)$ is available to the receiver. In order to justify the last assumption, $\{x_n\}$ is assumed to be a deterministic sequence, known to both the receiver and the transmitter, in the case where feedback is not available, while for the case of feedback, the receiver randomly generates x_n, based on $\boldsymbol{y}^{n-1}$ and $\boldsymbol{x}^{n-1}$, and sends it back, through the noiseless feedback link, to the transmitter.

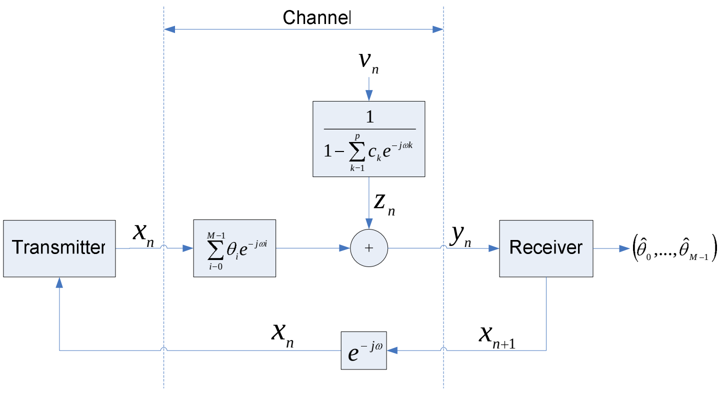

Fig. 1. System configuration

Under the assumed parametric model, our focus is on designing the channel input signal, i.e., the sequence of conditional densities $\left\{f_n\left(x_n|\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1}\right)\right\}_{n\geq 1}$, in order to achieve optimum estimation of the parameters vector $\boldsymbol{\theta} = [\theta_0, \cdots, \theta_{M-1}]^T$, subject to an average power constraint, P, which is given by:

$$\frac{1}{N} \sum_{n=1}^{N} E\left[X_n^2\right] \leq P, \tag{3}$$

where the expectation is w.r.t. the channel noise and the possible randomization in the choice of x_n based on $\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1}$. Specifically, the parameter vector $\boldsymbol{\theta}$ is the one we wish to estimate, whereas the parameters $\boldsymbol{c} = [c_1, \cdots, c_p]^T$ and σ are treated as nuisance parameters. This approach is useful when the goal is to identify a linear system which describes the channel and it can be extended for the cases where we are interested also in the estimation of the noise characteristics. The noise-shaping filter, $H(e^{j\omega})$, is represented by:

$$H(e^{j\omega}) = \frac{1}{1 - \sum_{k=1}^{p} c_k e^{-j\omega k}} \tag{4}$$

and it is assumed to be stable, i.e., all its poles are strictly inside the unit circle. The MSE of the estimation at time instant N, is defined as: $\varepsilon_N = E\left\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^N\right\|^2$, where $\|.\|$ represents the Euclidean norm and $\hat{\boldsymbol{\theta}}^N$ is the estimate of $\boldsymbol{\theta}$ at time N.

For the purpose of our subsequent analysis, we now derive the Fisher Information (FI) matrix of $\boldsymbol{\theta}$. The probability density function (pdf) of $(\boldsymbol{x}^N, \boldsymbol{y}^N) = (x_1, \cdots, x_N, y_1, \cdots, y_N)$ as a function of $\boldsymbol{\theta}$ obeys:

$$f_\theta(\boldsymbol{x}^N, \boldsymbol{y}^N) = \prod_{n=1}^{N} f_\theta\left(x_n, y_n|\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1}\right)$$

$$= \prod_{n=1}^{N} f\left(x_n|\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1}\right) f_\theta\left(y_n|\boldsymbol{x}^n, \boldsymbol{y}^{n-1}\right), \tag{5}$$

where we use the Bayes theorem for the two equalities. Now, for each time instant n, the term $f(x_n|\boldsymbol{x}^{n-1}, \boldsymbol{y}^{n-1})$ stands for a feedback-dependent transmitter (which is independent of $\boldsymbol{\theta}$) and the term $f_\theta(y_n|\boldsymbol{x}^n, \boldsymbol{y}^{n-1})$ stands for the action of the channel, whose exact form depends on the specific channel model, which in our case, is given by:

$$f_\theta(y_n|\boldsymbol{x}^n, \boldsymbol{y}^{n-1}) = g_\sigma\left(w_y(n) - \sum_{i=0}^{M-1} \theta_i w_x(n-i)\right), \tag{6}$$

where $w_x(n) \triangleq x_n - \sum_{k=1}^{p} c_k x_{n-k}$ and $w_y(n) \triangleq y_n - \sum_{k=1}^{p} c_k y_{n-k}$. Now, let us denote:

$$\lambda_\sigma(u) \triangleq \left[\frac{\partial \ln g_\sigma(v)}{\partial v}\right]_{v=u}. \tag{7}$$

Under the assumption of the channel defined by eqs. (1), (2), the FI matrix of the parameter vector $[\boldsymbol{\theta}, \boldsymbol{c}, \sigma]$ is given by:

$$\boldsymbol{J}(\boldsymbol{\theta}, \boldsymbol{c}, \sigma) = \begin{pmatrix} \boldsymbol{J}_{11} & \boldsymbol{J}_{12} & \boldsymbol{J}_{13} \\ \boldsymbol{J}_{21} & \boldsymbol{J}_{22} & \boldsymbol{J}_{23} \\ \boldsymbol{J}_{31} & \boldsymbol{J}_{32} & J_{33} \end{pmatrix} \tag{8}$$

where:

$$\begin{aligned} \boldsymbol{J}_{11} &= NE\left[\lambda_\sigma^2(V)\right] \boldsymbol{R}_{w_x w_x} \\ \boldsymbol{J}_{12} &= \boldsymbol{J}_{21}^T = NE\left[\lambda_\sigma^2(V)\right] \boldsymbol{R}_{w_x z} \\ \boldsymbol{J}_{22} &= NE\left[\lambda_\sigma^2(V)\right] \boldsymbol{R}_{zz} \\ (\boldsymbol{J}_{13})_j &= \left(\boldsymbol{J}_{31}^T\right)_j = \frac{K_1(\sigma)}{\sigma^2} \sum_{n=1}^{N} E\left[W_x(n-j+1)\right] \\ \boldsymbol{J}_{23} &= \boldsymbol{J}_{32}^T = 0 \\ J_{33} &= -\frac{N}{\sigma^2}\left(1 + K_2(\sigma)\right) \end{aligned} \tag{9}$$

where $K_1(\sigma) = E\left[\left(h \frac{\partial^2 \ln f(h)}{\partial h^2}\right)_{h=\frac{V}{\sigma}}\right]$, and $K_2(\sigma) = E\left[\left(h^2 \frac{\partial^2 \ln f(h)}{\partial h^2} + 2h \frac{\partial \ln f(h)}{\partial h}\right)_{h=\frac{V}{\sigma}}\right]$.

As mentioned before, our goal is to estimate the ISI vector of coefficients $\boldsymbol{\theta}$. The optimal channel input is derived by the minimization of the trace of the first block in the inverse FI matrix (8), under the power constraint (3). Specifically, in the case where $[\boldsymbol{c}, \sigma]$ are known parameters, the first block is given by:

$$\boldsymbol{J}^{-1}(\theta) = \boldsymbol{J}_{11}^{-1} \tag{10}$$

III. Memoryless Channels

In this section, we find the minimum CRB for estimating θ for the special case where $M = 1$ and $p = 0$. Specifically, the equation that describes the system is given by:

$$y_n \quad = \theta x_n + v_n, \tag{11}$$

A. Optimal Transmitter

Substituting $M = 1$, $p = 0$ in eq. (10) and by using the power constraint in eq. (3), we get:

$$J(\theta) = E\left[\lambda_\sigma^2(V)\right] \sum_{n=1}^{N} E\left[X_n^2\right] \leq NPE\left[\lambda_\sigma^2(V)\right] \tag{12}$$

One can observe that equality is achieved iff the transmitted signal $\{X_n\}_{n=1}^N$ meets the power constraint with equality. Hence, the maximum of FI is achieved without using feedback at all and regardless of the input signal shape. Moreover, it can be proved [8] that the additive white Gaussian noise (AWGN) channel is worst-case in the FI sense.

B. The Maximum Likelihood Estimator

We now move on to the implementation of the receiver. The ML estimator of θ at time N, $\hat{\theta}_N$, is given by:

$$\hat{\theta}_N = \arg\max_\theta \sum_{n=1}^{N} \ln g_\sigma(y_n - \theta x_n). \tag{13}$$

By differentiating the last equation with respect to θ and by using eq. (7), we get an equation that $\hat{\theta}_N$ must satisfy:

$$\sum_{n=1}^{N} x_n \lambda_\sigma \left(y_n - \hat{\theta}_N x_n\right) = 0. \tag{14}$$

In some cases, the solution to (14) depends on σ. When σ is unknown, the natural approach is to estimate jointly σ and θ at the decoder. The ML estimator of σ for a given θ at time N, $\hat{\sigma}_N$, is given by:

$$\hat{\sigma}_N = \arg\max_\sigma \sum_{n=1}^{N} \ln\left[\frac{1}{\sigma} f\left(\frac{y_n - \theta x_n}{\sigma}\right)\right] \tag{15}$$

Differentiation of eq. (15) with respect to σ gives likelihood equation:

$$-\frac{1}{N} \sum_{n=1}^{N} (y_n - \theta x_n)\, \lambda_{\hat{\sigma}_N}(y_n - \theta x_n) = 1. \tag{16}$$

The equations used for estimating θ and σ (eq. (16) and eq. (14)) are, in general, dependent, i.e., in order to estimate θ, one must know σ and vice versa. In order to solve these equations, the receiver is required to iterate between the two equations, until some convergence criterion is met. In general, the convergence is not guaranteed and depends on the specific form of $\lambda_\sigma(v)$. Further, we can conclude that although the nuisance parameters are present, feedback is not required to achieve optimum performance.

IV. Channels with Memory

We next move on to the case of a colored noise channel with ISI. The ML estimator of $\boldsymbol{\theta}$, based on all observations and channel inputs until time n, is achieved by solving the following equations:

$$\forall m = 0, \ldots, (M-1): \quad \frac{\partial}{\partial \theta_m} \ln f_\theta(\boldsymbol{x}^N, \boldsymbol{y}^N) = 0, \tag{17}$$

Using the fact that $f_\theta(\boldsymbol{x}^N, \boldsymbol{y}^N)$ depends on $\boldsymbol{\theta}$ only through $f_\theta(y_n|\boldsymbol{x}^n, \boldsymbol{y}^{n-1})$ (5) and using (6), the ML estimator of $\boldsymbol{\theta}$ must satisfy::

$$\sum_{n=1}^{N} \lambda_\sigma(\hat{v}_n) \left(x_{n-m} - \sum_{k=1}^{p} c_k x_{n-m-k}\right) = 0 \tag{18}$$

where:

$$\hat{v}_n = y_n - \sum_{i=0}^{M-1} \hat{\theta}_i x_{n-i} - \sum_{k=1}^{p} c_k \left(y_{n-k} - \sum_{i=0}^{M-1} \hat{\theta}_i x_{n-i-k}\right). \tag{19}$$

When $\{c_k\}_{k=1}^p$ are unknown nuisance parameters, in order to solve (18), an estimation of these nuisance parameters is required. The ML estimator of $\{c_k\}_{k=1}^p$ is obtained by solving the following equations:

$$\frac{\partial}{\partial c_l} \sum_{n=1}^{N} \ln f_\theta(y_n|\boldsymbol{x}^n, \boldsymbol{y}^{n-1}) = 0, \qquad l = 1, \ldots, p. \tag{20}$$

Differentiating with respect to c_l, the ML estimator of $\boldsymbol{c}$ must satisfy:

$$\sum_{n=1}^{N} \lambda_\sigma(\hat{v}_n) \left(y_{n-l} - \sum_{i=0}^{M-1} \hat{\theta}_i x_{n-l-i}\right) = 0, \tag{21}$$

where $\hat{v}_n$ is defined as in (19), with $\hat{c}_k$ replacing c_k. The ML estimator of $\boldsymbol{\theta}$ is achieved by jointly solving eq. (18) and eq. (21).

Second, we wish to find the minimum CRB under the power constraint (3) and derive the optimal input signal which achieves it. To simplify the problem, we examine several different cases of the channel.

A. AR(p) Channel - No ISI

Let us begin with the case of a general order p, but we assume $M = 1$. The optimal channel input for this case was first found in [4]. Considering eq. (10), the FI now becomes:

$$J(\theta) = NE\left[\lambda_\sigma^2(V)\right] R_{w_x w_x}(0). \tag{22}$$

An upper bound on the averaged power of $\{w_x(n)\}$ is:

$$R_{w_x w_x}(0) \leq P \left|\frac{1}{H(e^{j\omega_{min}})}\right|^2, \tag{23}$$

where ω_{min} is the frequency in which $|H(e^{j\omega})|$ attains its minimum. The equality is achieved iff:

$$x_n = \sqrt{2P}\cos(\omega_{min}n + \phi), \tag{24}$$

where ϕ is a random phase in the range $[-\pi, \pi)$.

The optimal strategy for the transmitter is to send a single tone at the frequency where the noise, $\{z_n\}$, has its minimum power spectral density. For the case of white noise (i.e., $c = 0$), the power spectral density of the noise is flat, and so, any signal of power P equally good to achieve the maximum FI. Again, we can conclude that the maximum of FI is achieved for this case without the use of feedback.

When $\{c_k\}_{k=1}^{p}$ are unknown nuisance parameters, the receiver does not know at which frequency $|H(e^{j\omega})|$ is minimum and the problem becomes harder. We suggest an algorithm, which iteratively estimates both $\{c_k\}_{k=1}^{p}$ and θ by using feedback. The algorithm is as follows.

1. For $n = 1$, the transmitter sends:

$$x_1 = \cos(w_{min}^1 + \phi), \tag{25}$$

where w_{min}^1 and ϕ are known to both the receiver and the transmitter.

For each index time $n \geq 1$:

2. The receiver estimates iteratively the parameter vector $[\theta, \boldsymbol{c}]$ using (18) and (21).
3. The receiver finds the frequency which minimizes the current estimate of the noise-shaping frequency response and sends it to the transmitter:

$$\hat{w}_{min}^n = \arg\min_{\omega} \left| 1 - \sum_{k=1}^{p} \hat{c}_k^n e^{-j\omega k} \right|.$$

4. The transmitter sends:

$$x_n = \sqrt{2P}\cos\left(\sum_{k=0}^{n-1} \hat{w}_{min}^k + \phi\right), \tag{26}$$

where the summation over the estimates of the frequency approximates a sine wave with a time varying frequency, which when stabilizes around the correct frequency, is equivalent to a sine wave with frequency w_{min}. For the special case, where $g_\sigma(V)$ is the Gaussian pdf, it can be shown that the FI matrix for the suggested algorithm is a block diagonal matrix for $N \to \infty$ and in addition, the algorithm achieves asymptotically the CRB for the case of known parameters, which makes the algorithm optimal.

Since we show that we can achieve asymptotically the CRB of the known parameters case by using feedback and in addition the optimal channel input (24) is unique and depends on the nuisance parameter $\boldsymbol{c}$, or in other words, the knowledge of $\boldsymbol{c}$ is essential - we can conclude that feedback is essential to achieve optimal error performance.

We examine the error performance of the algorithm for four cases of unknown channels, whose frequency responses are depicted in Fig. 2. We assume that the colored noise is an AR(p) process, where $\{v_n\}$ is a white Gaussian noise with zero mean and variance σ^2. In Fig. 3, the normalized MSE,

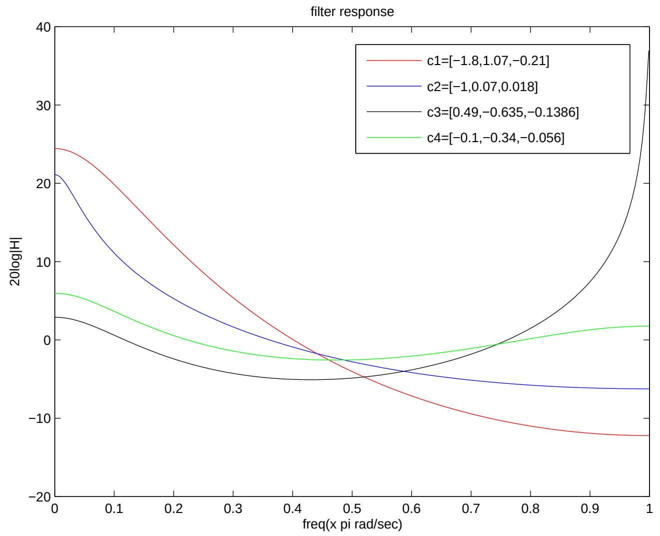

Fig. 2. Channel response of different channels - unknown AR(p) channel

$N \cdot \epsilon_N$, is compared to the CRB of each channel for the case $\theta = 0.5$, $P = 1$ and $\sigma = 0.25$. It can be seen that the CRB is achieved asymptotically and that the convergence rate of the error depends directly on the frequency characteristics of the channel: The smaller is the minimum response of the noise-shaping filter, the CRB is smaller too, and the convergence rate is faster.

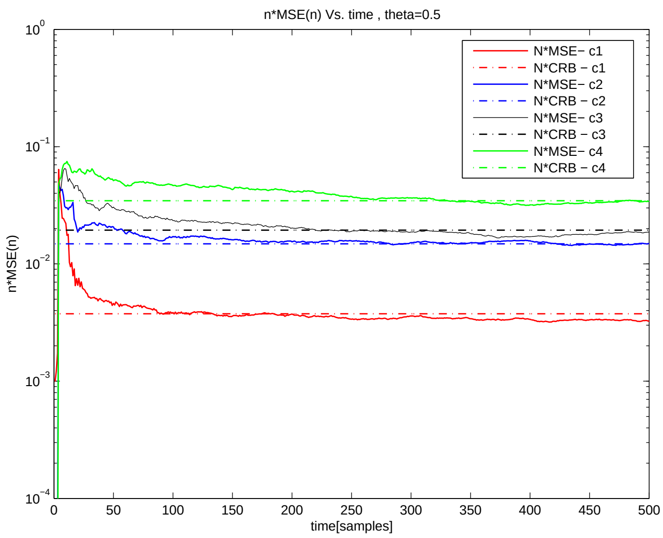

Fig. 3. $N \cdot \varepsilon_N$ vs. time for different values of $\boldsymbol{c}$ - unknown AR(p) channel

B. *ISI Channel With White Noise*

We next consider the case of ISI channel with white additive noise ($M > 1, p = 0$). Using eq. (10), the FI matrix for this special case is given by: $J(\boldsymbol{\theta}) = NE\left[\lambda_\sigma^2(V)\right]\boldsymbol{R}_{xx}$.

A lower bound on the CRB of each θ_i, $J_{ii}^{-1}(\boldsymbol{\theta})$, is given by:

$$J_{ii}^{-1}(\boldsymbol{\theta}) \geq \frac{1}{NPE\left[\lambda_\sigma^2(V)\right]}, \tag{27}$$

where the right term stands for the CRB of the estimation of only one ISI coefficient (12). The equality in (27) is achieved for the case of:

$$R_{xx}(k) = P\delta_k, \tag{28}$$

where $\delta_k = 1$ if $k = 0$ and for $k \neq 0$, $\delta_k = 0$. One observes that the optimal channel input does not depend on the past channel outputs, thus feedback is not helpful in this case.

C. AR(p) Channel - With ISI

The most general case we consider is the case of AR(p) channel with ISI ($M > 0$). The FI matrix for this case is given in (10): $J(\boldsymbol{\theta}) = NE\left[\lambda_\sigma^2(V)\right]\boldsymbol{R}_{w_x}$.

While the problem is stated for a general M, we solve only the case of very long ISI, i.e. $M \to \infty$, because of its simplicity and since, practically, it is the common case in most of the problems in reality. Using the approximation below (see [9]):

$$\left[\boldsymbol{R}_{w_x}^{-1}\right]_{k,j} \cong \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|H(e^{j\omega})\right|^2}{S_x(e^{j\omega})}e^{(k-j)\omega}d\omega \tag{29}$$

Our optimization problem can be written as:

$$\begin{aligned} \min \quad & \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\left|H(e^{j\omega})\right|^2}{S_x(e^{j\omega})}d\omega \\ \text{s.t.} \quad & \frac{1}{2\pi}\int_{-\pi}^{\pi}S_x(e^{j\omega})d\omega \leq P \end{aligned} \tag{30}$$

The solution to this problem is stated in the next theorem:

Theorem 1: Under the assumption of an average power constraint P at the transmitter and signals which are included in Wiener class, the minimum CRB is given by:

$$CRB^* = \frac{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|H(e^{j\omega})|d\omega\right)^2}{NPE\left[\lambda_\sigma^2(V)\right]} \tag{31}$$

where $H(e^{j\omega})$ is given in (4). The input signal, which achieves the minimum CRB, has a power spectral density function given by:

$$S_x^*(e^{j\omega}) = \frac{P|H(e^{j\omega})|}{\frac{1}{2\pi}\int_{-\pi}^{\pi}|H(e^{j\omega})|d\omega}. \tag{32}$$

Once again, we can conclude that if the transmitter knows the magnitude of the noise-shaping filter, $|H\left(e^{j\omega}\right)|$, there is no need for feedback to achieve the minimum CRB.

If $\{c_k\}_{k=1}^{p}$ are unknown nuisance parameters, the transmitter cannot generate the optimal input signal in (32). Once again, the approach that can be taken is to let the receiver estimate $\{c_k\}_{k=1}^{p}$ and send them to the transmitter through the feedback link. The estimation of $[\boldsymbol{\theta}, \boldsymbol{c}]$ is done using iterations between eq. (18) and eq. (21).

We examined the error performance for the case of estimating $\boldsymbol{\theta}$ only. We assumed an infinite ISI filter, in which the ISI coefficients are represented by $\theta_i = \alpha^i$, where $|\alpha| < 1$. In this case, the frequency response of the ISI sequence can be described as an infinite impulse response (IIR) filter with one pole: $\sum_{i=0}^{\infty}\alpha^i e^{-j\omega i} = \frac{1}{1-\alpha e^{-j\omega}}$.

When the receiver estimates $M(N)$ coefficients, the MSE at time instant N, can be represented by:

$$\varepsilon_N = E\left[\sum_{i=0}^{M(N)-1}\left(\theta_i - \hat{\theta}_i^N\right)^2 + \sum_{i=M(N)}^{\infty}\alpha^{2i}\right] \tag{33}$$

where we assume that, for all N: $\hat{\theta}_i^N = 0, \quad \forall i \geq M(N)$.

The results of $N \cdot MSE$ are drawn in Fig. 4 for the case of colored noise, which is modeled by an AR(p) process with unknown nuisance parameters $\boldsymbol{c} = [c_1, \cdots, c_p]$, where $\{v_n\}$ is a white Gaussian noise with zero mean, and variance σ^2. The results are shown for different selections of $M(N)$: $M_1(N) = \ln N$, $M_2(N) = \sqrt{N}$, $M_3(N) = N^{\frac{1}{3}}$ and for the parameters: $\theta = 0.5$, $\sigma = 0.25$, $P = 1$ and $\boldsymbol{c} = [-0.5, 0.4, 0.5]$. The error result is the average of 100 experiments. One can observe that

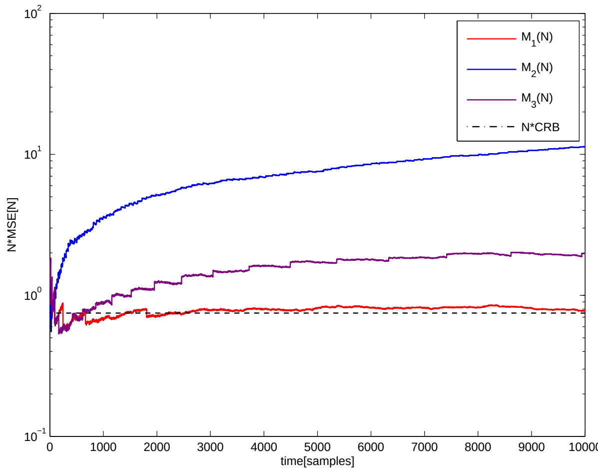

Fig. 4. $N \cdot \varepsilon_N$ vs. time - unknown AR(P) + ISI channel

the CRB is achieved for $M_1(N) = \ln N$.

References

[1] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud, "Least sum of squared errors (LSSE) channel estimation," *IEE Proceedings-F*, vol. 138, pp. 371–378, Aug. 1991.

[2] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions," *IEEE Signal Processing Mag.*, vol. 21, pp. 12–25, Nov. 2004.

[3] W. Chen and U. Mitra, "Training sequence optimization: comparison and an alternative criterion, " *IEEE Trans. Commun.*, vol. 48, pp. 1987–1991, Dec. 2000.

[4] R. K. Mehra, "Optimal input signals for parameter estimation in dynamic systems - survey and new results," *IEEE Trans. on Automatic Control*, AC-19(6), pp. 753–768, 1974.

[5] G. Goodwin, and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York, 1977.

[6] L. Ljung, *System Identification: Theory for the User*, 2nd Edition. Prentice-Hall. Englewood Cliffs, NJ ,1999.

[7] T. F. Wong and B. Park, "Training sequence optimization in MIMO systems with colored interference," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1939–1947, Nov. 2004.

[8] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[9] R. M. Gray, "Toeplitz and Circulant Matrices: A review," *Foundations and Trends in Communication and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.