# SENSITIVITY ANALYSIS OF A SUBOPTIMAL PRECODING SCHEME FOR BLOCK CHANNELS WITH RESPECT TO CHANNEL INACCURACIES

*Athanasios P. Liavas*

Department of Electronic and Computer Engineering
Technical University of Crete
73100 Kounoupidiana, Chania, Greece
Email: liavas@telecom.tuc.gr

**ABSTRACT**

Water-filling is the technique that computes the input data covariance matrix that achieves capacity in Gaussian paraller or block channels. A suboptimal precoding scheme, that has been observed to perform quite close to water-filling in some cases of great interest is when all "active" eigenvectors of the input data covariance matrix receive the same power. Both techniques require perfect channel knowledge at the transmitter. We consider the suboptimal precoding scheme when at the transmitter we use a channel estimate as if it were the true channel and we derive a closed form expression relating the channel estimation error covariance matrix with the mean mutual information decrease. We observe that serious error magnification may happen if the channel matrix is badly conditioned and the SNR is high.

## 1. INTRODUCTION

Block-based transmission is common in communications. If the channel is known at the transmitter, due to, e.g., feedback, then it is possible to maximize the information rate by precoding the channel input. This problem has been considered for frequency selective channels in [1] and for flat fading MIMO channels in [2], which compute the input data covariance matrix that maximizes the mutual information between the channel input and output. It turns out that the eigenvectors of the optimal input covariance matrix are the right singular vectors of the channel matrix while its eigenvalues are computed through water-filling using the (squared) singular values of the channel matrix. An important observation of [1] is that a suboptimal input data covariance matrix having all its non-zero eigenvalues equal leads, in some important cases, e.g., DSL, to a mutual information that almost coincides with that obtained through water-filling.

A question that is directly related with the practical success of these precoding schemes concerns their sensitivity with respect to channel and noise statistics inaccuracies. In this work, we consider the sensitivity of the suboptimal approach with respect to channel estimation errors. This problem is of importance because (a) the suboptimal approach has been observed to be close to the optimal in many cases of great interest (DSL, high SNR) and (b) the study of the suboptimal approach is tractable and may be the first step toward the study of the water-filling, that is certainly more demanding. Our main result is a closed form expression relating the channel estimation error covariance matrix with the mean mutual information decrease. We observe that significant error magnification may arise when the channel matrix is badly conditioned and the SNR is high.

Supported in part by the EU under U-BROAD STREP #506790.

## 2. CAPACITY ANALYSIS OF BLOCK-BASED SYSTEMS

### 2.1. The channel model

We consider the baseband-equivalent discrete-time noisy communication channel modeled by the $\nu$-th order linear time-invariant system with input-output relation

$$y_k = \sum_{i=0}^{\nu} h_i x_{k-i} + n_k \tag{1}$$

where $h \triangleq [\ h_0 \ \cdots \ h_\nu\ ]^T$ (superscript $^T$ denotes transpose) is the channel impulse response vector and $x_k$, $n_k$ and $y_k$ denote, respectively, the samples of the channel input, noise and output. Considering the data vectors $y \triangleq [\ y_{k+N-1} \ \ldots \ y_k\ ]^T$, $x \triangleq [\ x_{k+N-1} \ \ldots \ x_{k-\nu}\ ]^T$ and $n \triangleq [\ n_{k+N-1} \ \ldots \ n_k\ ]^T$, we may rewrite (1) in matrix form as $y = Hx + n$, where $H$ is the $N \times (N+\nu)$ filtering matrix defined as

$$H \triangleq \begin{bmatrix} h_0 & \cdots & \cdots & h_\nu & & \\ & \ddots & & & \ddots & \\ & & h_0 & \cdots & \cdots & h_\nu \end{bmatrix}.$$

The noise samples are assumed to be samples of a complex-valued zero-mean white circularly symmetric Gaussian stationary stochastic process with covariance matrix $R_{nn} \triangleq \mathcal{E}[nn^*] = N_0 I_N$, where superscript $^*$ denotes Hermitian transpose and $I_i$ denotes the $i \times i$ identity matrix. The input symbols are assumed to be complex-valued zero-mean circularly symmetric Gaussian (in order to achieve capacity) with covariance matrix $R_{xx} \triangleq \mathcal{E}[xx^*]$.

### 2.2. Capacity analysis: the ideal case

A problem that has been considered in [1] and [2] is the computation of the input covariance matrix that maximizes the mutual information $I(X;Y)$ between the input $X$ and the output $Y$ of the above block channel. Toward this end, the following singular value decompositions are useful:

$$H = V_N \left[ \Sigma_N^{1/2} \ 0 \right] U_{N+\nu}^* \tag{2}$$

$$A \triangleq H^*H = U_{N+\nu} \begin{bmatrix} \Sigma_N & \\ & 0 \end{bmatrix} U_{N+\nu}^* = U_N \Sigma_N U_N^* \quad (3)$$

where $U_k$, $k = 1, \ldots, N+\nu$, is the matrix with columns the $k$ eigenvectors of $A$ associated with its $k$ largest eigenvalues, and $\Sigma_k \triangleq \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$, with $\sigma_i$ being the $i$-th largest eigenvalue of $A$. It turns out that the optimal solution is [1], [2]

$$R_{xx}^{\mathrm{opt}} = U_{N+\nu}\, \mathrm{diag}(\delta_1, \ldots, \delta_N, \underbrace{0, \ldots, 0}_{\nu})\, U_{N+\nu}^*$$

where terms $\delta_i$, $i = 1, \ldots, N$ are computed through water-filling using the eigenvalues of $A$. A suboptimal approach, which has been observed to be very close to the optimal in some cases of great interest [1], is to assume that all the non-zero eigenvalues of $R_{xx}^{\mathrm{opt}}$ are equal, i.e., $\delta = \delta_1 = \delta_2 = \cdots = \delta_{\bar{N}} = E/\bar{N}$, where $E$ is the total input power and $\bar{N} \le N$. In this case, the input data covariance matrix is $R_{xx} = \delta\, U_{\bar{N}} U_{\bar{N}}^*$. If we define the operator $\mathcal{D}[M]$ as the product of the non-zero eigenvalues of matrix $M$, then the mutual information per input sample between the channel input and output in the suboptimal case is given by [1]

$$\begin{aligned} I(X;Y) &= \frac{1}{N+\nu} \log_2 \mathcal{D}\left[\left(\frac{1}{N_0} H^*H + R_{xx}^{\sharp}\right) R_{xx}\right] \\ &= \frac{1}{N+\nu} \log_2 \mathcal{D}\left[\frac{\delta}{N_0}\left(\Sigma_{\bar{N}} + \frac{N_0}{\delta} I_{\bar{N}}\right)\right] \end{aligned} \quad (4)$$

where superscript $\sharp$ denotes the pseudo-inverse. If we define

$$\mathcal{A} \triangleq \Sigma_{\bar{N}} + \frac{N_0}{\delta} I_{\bar{N}} \quad (5)$$

then

$$I(X;Y) = \underbrace{\frac{\bar{N}}{N+\nu} \log_2\left(\frac{\delta}{N_0}\right)}_{\mathcal{C}} + \frac{1}{N+\nu} \sum_{i=1}^{\bar{N}} \log_2 \lambda_i(\mathcal{A}).$$

## 3. CAPACITY UNDER CHANNEL MISMATCH

### 3.1. The framework

In practice, we do not know the true channel $h$ but, instead, its estimate $\hat{h}$. Using at the transmitter $\hat{h}$ as if it were the true channel, we compute $\hat{A} \triangleq \hat{H}^*\hat{H} = \hat{U}_N \hat{\Sigma}_N \hat{U}_N^*$. In this case, we consider as "optimal" the input covariance matrix $\hat{R}_{xx} = \delta\, \hat{U}_{\bar{N}} \hat{U}_{\bar{N}}^*$.

We define the errors in $\hat{h}$, $\hat{H}$ and $\hat{U}_k$ and the first-order error in $\hat{A}$ as follows:

$$\Delta h \triangleq \hat{h} - h,\ \Delta H \triangleq \hat{H} - H,\ \Delta U_k \triangleq \hat{U}_k - U_k$$

$$\Delta A \triangleq \hat{A} - A = H^*\Delta H + \Delta H^* H + O(||\Delta h||^2) \quad (6)$$

where $||\cdot||$ denotes the norm of the vector or matrix argument (since all norms are equivalent in finite dimensional vector spaces, we do not fix the norm at this point). The corresponding mutual information between the channel input and output is given by

$$\begin{aligned} \hat{I}(X;Y) &= \frac{1}{N+\nu} \log_2 \mathcal{D}\left[\left(\frac{1}{N_0} H^*H + \hat{R}_{xx}^{\sharp}\right) \hat{R}_{xx}\right] \\ &= \frac{1}{N+\nu} \log_2 \mathcal{D}\left[\frac{\delta}{N_0}(\mathcal{A} + \Delta\mathcal{A})\right] \\ &= \mathcal{C} + \frac{1}{N+\nu} \sum_{i=1}^{\bar{N}} \log_2 \lambda_i(\mathcal{A} + \Delta\mathcal{A}) \end{aligned} \quad (7)$$

where

$$\Delta\mathcal{A} = \Delta U_{\bar{N}}^* U_{\bar{N}} \Sigma_{\bar{N}} + \Sigma_{\bar{N}} U_{\bar{N}}^* \Delta U_{\bar{N}} + \Delta U_{\bar{N}}^* U_N \Sigma_N U_N^* \Delta U_{\bar{N}}. \quad (8)$$

### 3.2. The tools from matrix perturbation theory

In order to study the influence of the channel estimation errors to the mutual information, we must relate the eigenvalues $\lambda_i(\mathcal{A})$ and $\lambda_i(\mathcal{A} + \Delta\mathcal{A})$ and the invariant subspaces $U_{\bar{N}}$ and $\hat{U}_{\bar{N}}$. To this end, we shall use two results of matrix perturbation theory. The first can be easily deduced from Theorem 2.7 of [3, p. 236].

*Theorem 1:* Let $[\,U_{\bar{N}}\ U_{\bar{N}}^c\,]$ be unitary and suppose that $\mathcal{R}(U_{\bar{N}})$, $\mathcal{R}(U_{\bar{N}}^c)$ are simple invariant subspaces of the $(N+\nu) \times (N+\nu)$ Hermitian matrix $A$. Then, under certain conditions, there exists a unique $(N+\nu-\bar{N}) \times \bar{N}$ matrix $P$, such that

$$\hat{U}_{\bar{N}} = (U_{\bar{N}} + U_{\bar{N}}^c P)(I_{\bar{N}} + P^*P)^{-1/2} \quad (9)$$

and

$$\hat{U}_{\bar{N}}^c = (U_{\bar{N}}^c - U_{\bar{N}} P^*)(I_{N+\nu-\bar{N}} + PP^*)^{-1/2} \quad (10)$$

form orthonormal bases for simple orthogonal invariant subspaces of $A + \Delta A$ (note: $\mathcal{R}(M)$ is the column space of matrix $M$).

*Remark:* The exact sufficient conditions guaranteeing the existence of such a $P$ are stated in [3, p. 236]. Loosely speaking, we can find such a $P$ if the perturbation $\Delta A$ is sufficiently smaller than the gap between the smallest eigenvalue associated with $U_{\bar{N}}$ and the largest eigenvalue associated with $U_{\bar{N}}^c$. In this case, $P$ can be computed by solving a *non-linear* matrix equation. More specifically, since $\mathcal{R}(\hat{U}_{\bar{N}})$ and $\mathcal{R}(\hat{U}_{\bar{N}}^c)$ are orthogonal invariant subspaces of $\hat{A}$, we have [3, p. 220]

$$\hat{U}_{\bar{N}}^{c*} \hat{A}\, \hat{U}_{\bar{N}} = 0$$

leading to

$$(U_{\bar{N}}^{c*} - P U_{\bar{N}}^*)(A + \Delta A)(U_{\bar{N}} + U_{\bar{N}}^c P) = 0$$

which is non-linear in $P$. Thus, it is very difficult to find a closed form expression for $P$. If a solution $P$ exists, then $P = O(||\Delta A||)$ [3, p. 236]. By ignoring higher-order terms, that is, terms involving products of $P$ and $\Delta A$, we construct the *linearized* version of the above equation for the first-order approximation $\tilde{P}$ of $P$, as:

$$\Sigma_{\bar{N}}^c \tilde{P} - \tilde{P} \Sigma_{\bar{N}} = -U_{\bar{N}}^{c*} \Delta A U_{\bar{N}} \quad (11)$$

where

$$\Sigma_{\bar{N}}^c = \mathrm{diag}(\sigma_{\bar{N}+1}, \ldots, \sigma_N, \underbrace{0, \ldots, 0}_{\nu}).$$

Note that $\tilde{P} = P + O(||\Delta A||^2)$. Using the property of the vectorization operation

$$\mathrm{vec}(ABC) = (C^T \otimes A)\mathrm{vec}(B)$$

where $\otimes$ denotes the Kronecker product, we obtain the closed form expression for $\tilde{P}$

$$(I_{\bar{N}} \otimes \Sigma_{\bar{N}}^c - \Sigma_{\bar{N}} \otimes I_{N+\nu-\bar{N}})\mathrm{vec}(\tilde{P}) = -\mathrm{vec}(U_{\bar{N}}^{c*} \Delta A U_{\bar{N}}). \quad (12)$$

The next result can be deduced from Theorem 2.3 of [3, p. 183].

*Theorem 2:* For Hermitian matrices $\mathcal{A}$ and $\Delta\mathcal{A}$, if $\lambda_i(\mathcal{A})$ is a simple eigenvalue of $\mathcal{A}$ with associated eigenvector $e_i$, then there exists $\lambda_i(\mathcal{A} + \Delta\mathcal{A})$ unique eigenvalue of $\mathcal{A} + \Delta\mathcal{A}$ such that

$$\lambda_i(\mathcal{A} + \Delta\mathcal{A}) = \lambda_i(\mathcal{A}) + e_i^* \Delta\mathcal{A} e_i + O(||\Delta\mathcal{A}||^2). \quad (13)$$

## 4. PERTURBATION ANALYSIS

### 4.1. Second-order approximation to $\Delta\mathcal{A}$

In this subsection, we derive a second-order approximation to $\Delta\mathcal{A}$. Initially, the approximation is given in terms of $P$, and since $P$ is difficult to compute in closed form, we derive the corresponding approximation in terms of $\tilde{P}$.

We start with an approximation to $(I_{\bar{N}} + P^*P)^{-1/2}$. Let the SVD of $P$ be $P = U'\Sigma'V'^*$ (note $||P||_2 = ||\Sigma'||_2 = O(||\Delta A||_2)$). Then,

$$I_{\bar{N}} + P^*P = I_{\bar{N}} + V'\Sigma'^*\Sigma'V'^* = V'(I_{\bar{N}} + \Sigma'^*\Sigma')V'^*.$$

Using the Taylor expansion $(1+x^2)^{-1/2} = 1 - \frac{1}{2}x^2 + O(x^4)$, we obtain

$$\begin{aligned}(I_{\bar{N}} + P^*P)^{-1/2} &= V'(I_{\bar{N}} + \Sigma'^*\Sigma')^{-1/2}V'^* \\ &= V'\left(I_{\bar{N}} - \frac{1}{2}\Sigma'^*\Sigma' + O(||\Delta A||^4)\right)V'^* \\ &= I_{\bar{N}} - \frac{1}{2}P^*P + O(||\Delta A||^4). \qquad (14)\end{aligned}$$

If we define

$$P_* \triangleq P(I_{\bar{N}} + P^*P)^{-1/2} \qquad (15)$$

then (9) and (14) give

$$\hat{U}_{\bar{N}} = U_{\bar{N}} - \frac{1}{2}U_{\bar{N}}P^*P + U^c_{\bar{N}}P_* + O(||\Delta A||^4)$$

leading to

$$\Delta U_{\bar{N}} = -\frac{1}{2}U_{\bar{N}}P^*P + U^c_{\bar{N}}P_* + O(||\Delta A||^4).$$

The error terms appearing in $\Delta\mathcal{A}$ become (see (8))

$$\Delta U^*_{\bar{N}}U_{\bar{N}} = -\frac{1}{2}P^*P + O(||\Delta A||^4)$$

and (after some calculations)

$$\Delta U^*_{\bar{N}}U_N\Sigma_N U^*_{\bar{N}}\Delta U_{\bar{N}} = P^*_*\Sigma^c_{\bar{N}}P_* + O(||\Delta A||^3)$$

leading to

$$\Delta\mathcal{A} = -\frac{1}{2}\Sigma_{\bar{N}}P^*P - \frac{1}{2}P^*P\Sigma_{\bar{N}} + P^*_*\Sigma^c_{\bar{N}}P_* + O(||\Delta A||^3).$$

Using the facts that $P = O(||\Delta A||)$ and $\tilde{P} = P + O(||\Delta A||^2)$, the expansion (14) and the definition of $P_*$ in (15), we can show

$$\tilde{P}^*\tilde{P} = P^*_*P_* + O(||\Delta A||^3) = P^*P + O(||\Delta A||^3)$$

$$\tilde{P}^*\Sigma^c_{\bar{N}}\tilde{P} = P^*_*\Sigma^c_{\bar{N}}P_* + O(||\Delta A||^3)$$

and thus we may approximate $\Delta\mathcal{A}$ as

$$\Delta\mathcal{A} = -\frac{1}{2}\Sigma_{\bar{N}}\tilde{P}^*\tilde{P} - \frac{1}{2}\tilde{P}^*\tilde{P}\Sigma_{\bar{N}} + \tilde{P}^*\Sigma^c_{\bar{N}}\tilde{P} + O(||\Delta A||^3). \qquad (16)$$

### 4.2. Second-order approximation to $\Delta\lambda_i(\mathcal{A})$

Since $\mathcal{A}$ is diagonal, its eigenvectors are the canonical vectors $e_i$, i.e., vectors with 1 at the $i$-th position and zeros elsewhere. Then, using Theorem 2 and (16), we obtain

$$\begin{aligned}\Delta\lambda_i(\mathcal{A}) &\triangleq \lambda_i(\mathcal{A}+\Delta\mathcal{A}) - \lambda_i(\mathcal{A}) = e^*_i\Delta\mathcal{A}e_i + O(||\Delta\mathcal{A}||^2) \\ &= -e^*_i\tilde{P}^*\underbrace{(\sigma_i I_{N+\nu-\bar{N}} - \Sigma^c_{\bar{N}})}_{\mathcal{S}_i}\tilde{P}e_i + O(||\Delta A||^3). \qquad (17)\end{aligned}$$

(Note that $|\Delta\lambda_i(\mathcal{A})| = O(||\tilde{P}||^2) = O(||\Delta A||^2)$).

### 4.3. Second-order approximation to $\Delta I(X;Y)$

Using (7) and (17), we obtain an approximation to $\hat{I}(X;Y)$ as:

$$\begin{aligned}\hat{I}(X;Y) &= \mathcal{C} + \frac{1}{N+\nu}\sum_{i=1}^{\bar{N}}\log_2\lambda_i(\mathcal{A}+\Delta\mathcal{A}) \\ &= \mathcal{C} + \frac{1}{N+\nu}\sum_{i=1}^{\bar{N}}\log_2(\lambda_i(\mathcal{A}) + \Delta\lambda_i(\mathcal{A})) \\ &= \mathcal{C} + \frac{1}{N+\nu}\sum_{i=1}^{\bar{N}}\log_2\left(\lambda_i(\mathcal{A})\left(1+\frac{\Delta\lambda_i(\mathcal{A})}{\lambda_i(\mathcal{A})}\right)\right) \\ &= I(X;Y) + \frac{1}{N+\nu}\sum_{i=1}^{\bar{N}}\log_2\left(1+\frac{\Delta\lambda_i(\mathcal{A})}{\lambda_i(\mathcal{A})}\right) \\ &\overset{(a)}{=} I(X;Y) + \frac{\log_2 e}{N+\nu}\sum_{i=1}^{\bar{N}}\frac{\Delta\lambda_i(\mathcal{A})}{\lambda_i(\mathcal{A})} + O(|\Delta\lambda_i|^2) \qquad (18)\end{aligned}$$

where in $(a)$ we used the expansion $\ln(1+\Delta x) = \Delta x + O(|\Delta x|^2)$. Thus, a second-order approximation of $\Delta I(X;Y)$ is

$$\Delta I(X;Y) = \frac{\log_2 e}{N+\nu}\sum_{i=1}^{\bar{N}}\frac{1}{\lambda_i(\mathcal{A})}e^*_i\tilde{P}^*\mathcal{S}_i\tilde{P}e_i + O(||\Delta A||^3).$$

In the sequel, in order to simplify notation, we shall omit the $O(\cdot)$ terms, which are obvious from the above analysis.

### 4.4. Computation of the mean mutual information decrease

We assume that the channel estimation error $\Delta h$ is zero-mean, circular, i.e., $\mathcal{E}\{\Delta h\,\Delta h^T\} = O_{(\nu+1)\times(\nu+1)}$, with covariance matrix $R_{\Delta h} \triangleq \mathcal{E}\{\Delta h\,\Delta h^*\}$ and we derive a second-order approximation of the mean mutual information degradation due to channel inaccuracies as

$$\mathcal{E}\{\Delta I(X;Y)\} = \frac{\log_2 e}{N+\nu}\sum_{i=1}^{\bar{N}}\frac{1}{\lambda_i(\mathcal{A})}\mathcal{E}\left\{\mathrm{Tr}(e^*_i\tilde{P}^*\mathcal{S}_i\tilde{P}e_i)\right\} \qquad (19)$$

where the expectation is with respect to the channel errors. Due to space limitation, we consider only the case $\bar{N} = N$ (the general case is treated in [5]).

#### 4.4.1. Case $\bar{N} = N$

In this case, $\Sigma^c_{\bar{N}} = O_{\nu\times\nu}$, which using (11), (6) and (17) gives

$$\tilde{P} = U^{c*}_{\bar{N}}\Delta H^*V_{\bar{N}}\Sigma^{-1/2}_{\bar{N}}, \quad \mathcal{S}_i = \sigma_i I_{N+\nu-\bar{N}}.$$

If we denote by $v_i$, $i = 1,\ldots,\bar{N}$, the $i$-th column of $V_N$, we obtain $\tilde{P}e_i = \sigma_i^{-1/2}U^{c*}_{\bar{N}}\Delta H^*v_i$, giving

$$\mathrm{vec}(\tilde{P}e_i) = \sigma_i^{-1/2}\underbrace{\left(v_i^T\otimes U^{c*}_{\bar{N}}\right)}_{\alpha_i}\mathrm{vec}(\Delta H^*).$$

It can easily be verified that

$$\mathrm{vec}(\Delta H^*) = \underbrace{\left[\begin{array}{c} 1_{N-1}\otimes\left[\begin{array}{c} I_{\nu+1} \\ O_{N\times(\nu+1)}\end{array}\right] \\ I_{\nu+1}\end{array}\right]}_{\mathcal{M}}\bar{\Delta h} \qquad (20)$$

where $1_i$ is the $i$-dimensional vector composed of 1's and $\bar{\Delta h}$ is the complex conjugate vector of $\Delta h$. Thus,

$$\begin{aligned}\mathrm{Tr}\left(e_i^* \tilde{P}^* \mathcal{S}_i \tilde{P} e_i\right) &= \sigma_i \mathrm{Tr}\left(\mathrm{vec}(\tilde{P} e_i)\mathrm{vec}(\tilde{P} e_i)^*\right) \\ &= \mathrm{Tr}\left(\alpha_i \mathcal{M} \bar{\Delta h} \Delta h^T \mathcal{M}^* \alpha_i^*\right).\end{aligned}$$

Using (19), we obtain

$$\begin{aligned}\mathcal{E}\{\Delta I(X;Y)\} &= \frac{\log_2 e}{N+\nu} \sum_{i=1}^{\bar{N}} \frac{1}{\lambda_i(\mathcal{A})} \mathrm{Tr}\left(R_{\Delta h}^T \mathcal{M}^* \alpha_i^* \alpha_i \mathcal{M}\right) \\ &= \frac{\log_2 e}{N+\nu} \mathrm{Tr}\left(R_{\Delta h}^T \mathcal{M}^* \left(\sum_{i=1}^{\bar{N}} \frac{1}{\lambda_i(\mathcal{A})} \alpha_i^* \alpha_i\right) \mathcal{M}\right).\end{aligned}$$

Using the relation $\alpha_i^* \alpha_i = \left(\bar{v}_i v_i^T \otimes U_{\bar{N}}^c U_{\bar{N}}^{c*}\right)$ and

$$\sum_{i=1}^{\bar{N}} \frac{1}{\lambda_i(\mathcal{A})} \alpha_i^* \alpha_i = \underbrace{\bar{V}_N \mathcal{A}^{-1} V_N^T \otimes U_{\bar{N}}^c U_{\bar{N}}^{c*}}_{\mathcal{X}}$$

we obtain

$$\boxed{\mathcal{E}\{\Delta I(X;Y)\} = \frac{\log_2 e}{N+\nu} \mathrm{Tr}\left(R_{\Delta h}^T \mathcal{M}^* \mathcal{X} \mathcal{M}\right).} \tag{21}$$

The term that may significantly magnify the channel estimation errors is $\mathcal{A}^{-1}$, which is large if $\sigma_{\bar{N}}$ and $\frac{N_0}{\delta}$ are small.

An intuitively satisfying fact that can be observed in the above expression is that, for fixed channel (that is, fixed $\sigma_i$, $V_N$ and $U_N^c$) and fixed channel estimation error covariance matrix, the mean mutual information degradation increases for increasing the SNR, that is, decreasing $\frac{N_0}{\delta}$.

## 5. SIMULATIONS

In this section, we use numerical simulations to illustrate our theoretical results. The power per input data sample is $P_x = 1$ and the total input power is $E = N + \nu$, with block length $N = 50$. The power of the additive white Gaussian noise is $N_0$ and the SNR is defined as $\mathrm{SNR} \triangleq \frac{P_x}{N_0}$. The true channel has order $\nu = 5$ and is

$$h = [\,0.3871\ -0.6087\ -0.1006\ -0.4675\ 0.1611\ -0.4713]^T.$$

In Fig. 1, we plot the mutual information $I(X;Y)$ that results from water-filling and the sub-optimal approach, for $\bar{N} = N$ (using the true channel $h$). We observe that for SNR higher than 4 dB, the two quantities practically coincide, supporting the observations of [1].

In order to check the accuracy of our approximation, we assume that the channel is estimated using $N_{\mathrm{tr}} = 12$ data samples of an ideal training sequence, giving that $R_{\Delta h} = \frac{N_0}{P_x(N_{\mathrm{tr}}-\nu)} I_{\nu+1}$ [4, p. 788]. In Fig. 2, we plot the experimentally computed (over $10^4$ independent noise realizations) mean mutual information degradation and the corresponding second-order approximation (21), for varying the SNR. We observe that for SNR higher than 15 dB the two quantities coincide, showing the usefulness of our approximation. A more extensise simulation study appears in [5].

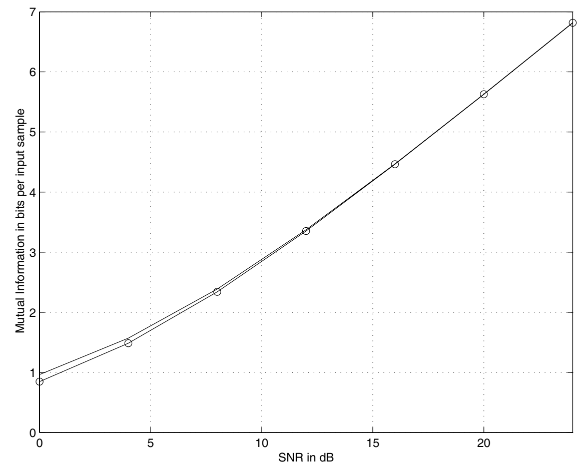


**Fig. 1**. Mutual information per sample: water-filling (solid line), suboptimal scheme ('o-') versus SNR.

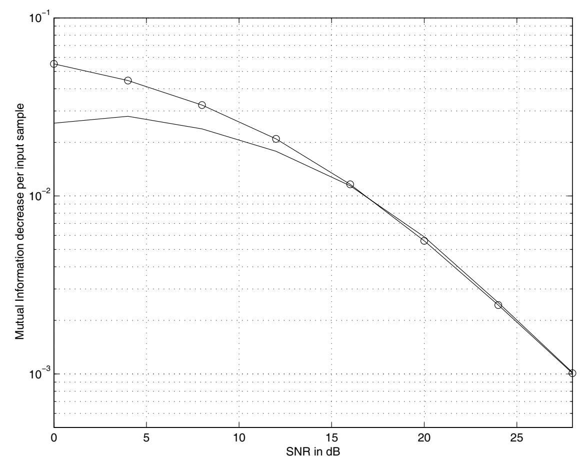


**Fig. 2**. Experimental mutual information decrease (solid line) and second-order approximation ('o-'), versus SNR.

## 6. REFERENCES

[1] N. Al-Dhahir and J. M. Cioffi, "Block transmission over dispersive channels: transmit filter optimization and realization, and MMSE-DFE receiver performance," *IEEE Trans. Information Theory*, vol. 42, no. 1, pp. 137–160, January 1996.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Communications,* vol. 10, no. 6, pp. 585–595, 1999.

[3] G. Stewart and Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

[4] H. Meyr, M. Moeneclaey and S. A. Fechtel. *Digital Communication Receivers.* John Wiley, 1998.

[5] A. P. Liavas, "On the sensitivity of a subptimal precoding scheme for block channels with respect to channel inaccuracies," submitted to the *IEEE Trans. Information Theory*, Sept. 2004.