ON MIMO CHANNEL ESTIMATION WITH SINGLE-BIT SIGNAL-QUANTIZATION

Michel T. Ivrlač and Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Munich University of Technology, Munich, Germany
{ivrlac, josef.a.nossek}@tum.de

ABSTRACT

The topic of this work is channel estimation for multi-input multi-output (MIMO) systems with very coarse signal quantization at the receiver. While coarse quantization of the received signal may only have a small to moderate impact on the channel capacity of MIMO systems, it is, however, necessary for the receiver to know the MIMO channel matrix. This motivates to study possible ways of estimating the channel, having available only the receive signals *after* quantization, especially fairly coarse ones. Starting from known results, we develop new insights into the behavior of MIMO channel estimators which work with a single-bit quantizer. Besides ultra-high-speed radio links, the application area includes on-chip and inter-chip bus-system, which employ a single-bit quantizer (logic comparator) at the receiving end of the bus.

1. INTRODUCTION

While the discrete-time nature of digital communications is well understood [1], the challenges which arise from quantization of the received signal have been largely neglected by the research community up to now. This situation is starting to change. Recently, it was demonstrated in [2, 3] that quantization of the received signal only has a small to moderate impact on the channel capacity of MIMO systems, provided that the receiver, and preferably also the transmitter, have knowledge about the MIMO channel matrix.

Because in general, the channel matrix cannot be assumed known apriori, a channel estimation has to be performed. In practice, it is highly desirable that the channel is estimated directly by the communication device. In this way, the channel estimator is restricted to use the received signal samples *after* quantization (analog to digital conversion).

Some interesting challenges arise when this quantization becomes rather coarse. This may happen in MIMO systems that have to operate at high-speed, such that quick-enough, high-resolution analog to digital converters (ADC) are either too power-hungry, too expensive, or even not available at all [4]. This may also happen in high-speed wireline MIMO systems, such as on-chip or chip-to-chip interconnects, where the receiver quantizes each wire's signal with a single-bit quantizer ("high"/"low"). This motivates investigation of channel estimation with coarse quantization. This problem was first addressed by [5], where a maximum likelihood (ML) channel estimation with quantized observation is presented. In general, the solution cannot be given in closed form, but requires an iterative numerical approach, which hampers the analysis of performance.

In this paper, we would like to help in providing some insight in what happens when channel estimation is performed for MIMO systems with coarse signal quantization. We focus on single-bit quantization as an extreme, yet practically interesting, case. We show that in contrast to unquantized estimation, different orthogonal pilot sequences (with same average total transmit power and same length) yield different performances. Especially, it turns out that establishing orthogonality in the time-domain, i.e. time-multiplexed pilots, can be preferable to orthogonality in space. By using a pilot that is multiplexed in time, only one transmit antenna is active at any time instant. This reduces the problem of a MIMO channel estimation to a SIMO (single-input multi-output) channel estimation. For independent receiver noise, one can reduce the problem still further to the SISO (single-input single-output) case, where only a channel coefficient between a pair of receive and transmit antennas is estimated. A closed-form solution can be found for the ML channel estimation problem, and the performance analyzed analytically.

2. RELATED WORK

The problem of channel estimation with quantized observation is also the focus of the work published in [5], which we briefly review for convenience. A quantized linear channel is estimated, which input-output relationship is given by:

$$\boldsymbol{y} = \mathbf{Q}(\boldsymbol{z}) \;, \quad \text{with} \tag{1}$$

$$\boldsymbol{z} = \mathbf{vec}\left[\boldsymbol{H}\boldsymbol{X}\right] + \boldsymbol{\nu} \;. \tag{2}$$

For simplicity, let $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ be the channel matrix of a frequency-flat MIMO system with N transmit and M receive ports, let $\boldsymbol{X} \in \mathbb{C}^{N\times n}$ contain n pilot-vectors of dimension N, while $\boldsymbol{\nu} \in \mathbb{C}^{(M\cdot n)\times 1}$, and $\boldsymbol{z} \in \mathbb{C}^{(M\cdot n)\times 1}$ shall denote the noise-samples and the unquantized received signal, both stacked into vectors. After quantization denoted by $\mathbf{Q}(.)$, the observation $\boldsymbol{y} \in \mathbb{C}^{(M\cdot n)\times 1}$ is obtained. The signal from (2) can also be written as

$$\boldsymbol{z} = \boldsymbol{\mathcal{X}}\boldsymbol{h} + \boldsymbol{\nu} \ , \tag{3}$$

where $\boldsymbol{h} = \mathbf{vec}[\boldsymbol{H}]$ and $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{(M\cdot n)\times(M\cdot N)}$ contains the pilot symbols placed in the proper places. For zero-mean white Gaussian noise, the maximum likelihood (ML) estimate for $\boldsymbol{h}$ based on the unquantized observation $\boldsymbol{z}$ is given by [6],

$$\widehat{\boldsymbol{h}'}_{\mathrm{ML}} = \left(\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{\mathcal{X}}\right)^{-1}\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{z} \ . \tag{4}$$

Since $\boldsymbol{z}$ itself cannot be observed, but only its quantized version $\boldsymbol{y}$, it can be shown [5] that the ML estimate of $\boldsymbol{h}$ is given as the solution of

$$\widehat{\boldsymbol{h}}_{\mathrm{ML}} = \left(\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{\mathcal{X}}\right)^{-1}\boldsymbol{\mathcal{X}}^{\mathrm{H}} \ \mathrm{E}\left[\boldsymbol{z} \mid \boldsymbol{y}, \widehat{\boldsymbol{h}}_{\mathrm{ML}}\right] \ . \tag{5}$$

This gives a system of non-linear equations which, in general, cannot be solved in closed form for $\widehat{\boldsymbol{h}}_{\mathrm{ML}}$. In [5] it is therefore suggested to perform a fix-point iteration:

① Set $\widehat{\boldsymbol{h}}_{\mathrm{ML}}$ to some initial value

② Compute $\boldsymbol{w} = \mathrm{E}\left[\boldsymbol{z} \mid \boldsymbol{y}, \widehat{\boldsymbol{h}}_{\mathrm{ML}}\right]$

③ $\widehat{\boldsymbol{h}}_{\mathrm{ML}} \leftarrow \left(\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{\mathcal{X}}\right)^{-1}\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{w}$

④ Continue at step ② until $\widehat{\boldsymbol{h}}_{\mathrm{ML}}$ has stabilized

3. TEMPORAL VS. SPATIAL PILOT MULTIPLEXING

For the case of unquantized ML channel estimation (based on $\boldsymbol{z}$) the mean square estimation error is given by

$$\epsilon' = \sigma_\nu^2 \cdot \mathrm{tr}\left(\left(\boldsymbol{\mathcal{X}}^{\mathrm{H}}\boldsymbol{\mathcal{X}}\right)^{-1}\right) \ , \tag{6}$$

where σ_ν^2 is the variance of the zero-mean, uncorrelated noise samples. The best performance is achieved when $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{X}}^{\mathrm{H}}$ is a scaled identity matrix, which means that the columns of $\boldsymbol{\mathcal{X}}$ are pairwise orthogonal and have the same norm (orthogonal pilot sequence). If for two pilot matrices $\boldsymbol{\mathcal{X}}_1$ and $\boldsymbol{\mathcal{X}}_2$ holds that $\boldsymbol{\mathcal{X}}_1^{\mathrm{H}}\boldsymbol{\mathcal{X}}_1 = \boldsymbol{\mathcal{X}}_2^{\mathrm{H}}\boldsymbol{\mathcal{X}}_2$, the performance will be the same, regardless if $\boldsymbol{\mathcal{X}}_1 \neq \boldsymbol{\mathcal{X}}_2$. This is, in general, *not* the case when the channel estimation is performed with the quantized received signal. To demonstrate this behavior, let us consider the following channel matrix

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1.1 \\ 0.9 & -1 \end{bmatrix} \ ,$$

and assume a single-bit quantizer, which merely returns the sign of the received signal, such that

$$\boldsymbol{y} = \mathbf{Q}(\boldsymbol{z}) = \begin{bmatrix} \mathrm{sign}\left(\mathbf{e}_1^{\mathrm{T}}\boldsymbol{z}\right) \\ \mathrm{sign}\left(\mathbf{e}_2^{\mathrm{T}}\boldsymbol{z}\right) \\ \vdots \\ \mathrm{sign}\left(\mathbf{e}_{2n}^{\mathrm{T}}\boldsymbol{z}\right) \end{bmatrix} \ , \tag{7}$$

where $\mathbf{e}_i$ is the i-th unit vector, and n is the pilot length, i.e. the number of two-dimensional pilot vectors which are transmitted during the pilot phase. The $\mathrm{sign}(.)$-function is defined to return -1 for negative arguments and 1 else. For simplicity, let the elements ν_i of the noise vector $\boldsymbol{\nu}$ be real valued with the probability density function (pdf):

$$\mathrm{pdf}_{\nu_i}(\nu_i) = \frac{\exp\left(-\frac{\nu_i^2}{2\sigma_\nu^2}\right)}{\sqrt{2\pi\sigma_\nu^2}}, \tag{8}$$

and the property

$$\mathrm{E}[\nu_i \cdot \nu_j] = \begin{cases} \sigma_\nu^2 & \text{for } i = j \\ 0 & \text{else} \end{cases} \ , \tag{9}$$

which makes the noise samples mutually independent. For a real-valued pilot matrix $\boldsymbol{\mathcal{X}}$, the conditional expectations from (5) can be written with (8) component wise as:

$$\mathrm{E}[z_i|y_i = 1, \widehat{h}_{\mathrm{ML}}] = \frac{\displaystyle\int_0^\infty z_i \, \mathrm{e}^{-\frac{\left(z_i - \mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}\right)^2}{2\sigma_\nu^2}} \mathrm{d}z_i}{\displaystyle\int_0^\infty \mathrm{e}^{-\frac{\left(z_i - \mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}\right)^2}{2\sigma_\nu^2}} \mathrm{d}z_i} \ , \tag{10}$$

and

$$\mathrm{E}[z_i|y_i = -1, \widehat{h}_{\mathrm{ML}}] = \frac{\displaystyle\int_{-\infty}^0 z_i \, \mathrm{e}^{-\frac{\left(z_i - \mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}\right)^2}{2\sigma_\nu^2}} \mathrm{d}z_i}{\displaystyle\int_{-\infty}^0 \mathrm{e}^{-\frac{\left(z_i - \mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}\right)^2}{2\sigma_\nu^2}} \mathrm{d}z_i} \ , \tag{11}$$

which can be computed in closed-form:

$$\mathrm{E}[z_i|y_i, \widehat{h}_{\mathrm{ML}}] = \mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}} + y_i\sqrt{\frac{2\sigma_\nu^2}{\pi}} \cdot \frac{\mathrm{e}^{-\frac{\left(\mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}\right)^2}{2\sigma_\nu^2}}}{\mathrm{erfc}\left(-\frac{y_i\mathbf{e}_i^{\mathrm{T}}\boldsymbol{\mathcal{X}}\widehat{\boldsymbol{h}}_{\mathrm{ML}}}{\sqrt{2\sigma_\nu^2}}\right)} \ , \tag{12}$$

where $\mathrm{erfc}(.)$ is the complementary error function, and z_i and y_i denote the i-th components of the vectors $\boldsymbol{z}$ and $\boldsymbol{y}$, respectively. The maximum likelihood channel estimate $\widehat{\boldsymbol{h}}_{\mathrm{ML}}$ can now be obtained by the iterative algorithm from Section 2. In order to develop more insight into how the pilot influences the performance of the channel estimation, let us distinguish the following two cases:

1. **Space-multiplexed pilots:** Both transmit antennas are used simultaneously and alternate between the two pilot symbols $\sqrt{P_\mathrm{T}/2}\cdot[1 \quad 1]^\mathrm{T}$ and $\sqrt{P_\mathrm{T}/2}\cdot[1 \quad -1]^\mathrm{T}$, where P_T is the total transmit power. The resulting pilot matrix $\boldsymbol{\mathcal{X}}_\mathrm{S}$ is given by:

$$\boldsymbol{\mathcal{X}}_\mathrm{S} = \sqrt{\frac{P_\mathrm{T}}{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \in \mathbb{R}^{(2n)\times 4} .$$

Note that $\boldsymbol{\mathcal{X}}_\mathrm{S}^\mathrm{H}\boldsymbol{\mathcal{X}}_\mathrm{S} = (nP_\mathrm{T}/2)\mathbf{I}_4$, where $\mathbf{I}_4$ is the 4×4 identity matrix. Hence, the pilot is orthogonal.

2. **Time-multiplexed pilots:** The transmit antennas are used only one at a time to transmit an all-ones pilot with transmit power P_T:

$$\boldsymbol{\mathcal{X}}_\mathrm{T} = \sqrt{P_\mathrm{T}} \begin{bmatrix} \mathbf{I}_4 \\ \mathbf{I}_4 \\ \vdots \end{bmatrix} \in \mathbb{R}^{(2n)\times 4} .$$

Since $\boldsymbol{\mathcal{X}}_\mathrm{T}^\mathrm{H}\boldsymbol{\mathcal{X}}_\mathrm{T} = (nP_\mathrm{T}/2)\mathbf{I}_4$, this is also an orthogonal pilot sequence.

Note that for unquantized channel estimation both pilot sequences would yield identical performance. However, this is *not* the case for quantized channel estimation. Figure 1 shows the relative mean square estimation error

$$\epsilon_\mathrm{rel} = \frac{\mathrm{E}\left[\left|\left|\widehat{\boldsymbol{h}}_\mathrm{ML} - \boldsymbol{h}\right|\right|_2^2\right]}{||\boldsymbol{h}||_2^2} , \tag{13}$$

which is computed numerically by executing the iterative algorithm described in Section 2, as a function of the ratio of total transmit power and receiver noise power for three different pilot lengths $n \in \{10^3, 3\times 10^3, 10^4\}$. For these pilot lengths, we can make the following key observations from Figure 1:

1. In general, the relative MSE is neither a monotonous, nor a convex function of $\log\left(P_\mathrm{T}/\sigma_\nu^2\right)$.

2. There is an optimum $P_\mathrm{T}/\sigma_\nu^2$ which yields the lowest relative MSE. This optimum appears to be almost independent of the pilot length, but depends whether time- or space-multiplexed pilots are used.

3. The time-multiplexed pilots yield a lower relative estimation error than the space-multiplexed pilots.

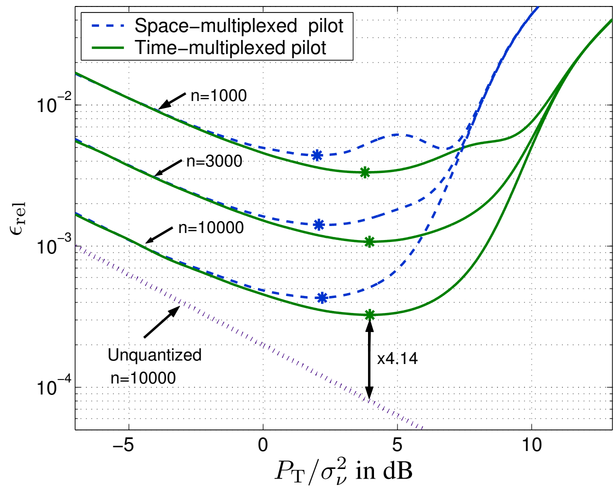

Fig. 1. Relative mean square error of the ML estimator with single-bit quantization, for two different orthogonal pilot sequences and three different pilot length n. For comparison also the performance of the unquantized ML estimator is shown (dotted curve).

Since for time-multiplexed pilots only one transmit antenna is active at any time, the problem is essentially reduced from the MIMO to the SIMO case. For independent receiver noise, one can simplify the problem still further to the SISO case, where only a channel coefficient between a pair of receive and transmit antennas is estimated. In this case, a closed-form solution can be found for the ML channel estimation problem, and the performance analyzed analytically. The good performance of the time-multiplexed pilot scheme motivates to have a closer look at the SISO channel estimation, which we will do in the following.

4. SYSTEM UNDER CONSIDERATION

Let us now restrict the problem to a scalar channel estimation in a single-input single-output system as shown in Fig. 2. A binary pilot sequence $(x_1, x_2, \ldots, x_n)$, where

$$x_i \in \{-1, 1\}, \quad \text{for} \quad i \in \{1, 2, \ldots, n\}, \tag{14}$$

is transmitted with power P_T over a communication channel, which is described by a constant but unknown scalar channel coefficient

$$h \in \mathbb{R}, \tag{15}$$

and perturbed by additive Gaussian noise $\nu_i \in \mathbb{R}$,with probability density function from (8). Assuming in addition the property (9), makes the noise samples mutually independent. The noisy signal is then provided to the input of a single-bit analog to digital converter, which only delivers the sign of the received signal. The input-output relationship can therefore

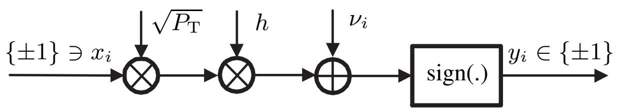

Fig. 2. A scalar single-bit quantized transmission system.

be written:

$$\{-1, 1\} \ni y_i = \mathrm{sign}\,(z_i) \ , \quad \text{with} \tag{16}$$

$$z_i = h \cdot \sqrt{P_\mathrm{T}} \cdot x_i + \nu_i \ . \tag{17}$$

By observing the sequence $(y_1, y_2, \ldots, y_n)$ of detected pilot bits, an estimation $\hat{h}$ of the channel coefficient h has to be obtained. In the following it is assumed that the ratio $P_\mathrm{T}/\sigma_\nu^2$ is constant and known error-free to the estimator a priori.

5. MAXIMUM LIKELIHOOD ESTIMATION

With the notation of (3), we have

$$\boldsymbol{h} = h \ , \quad \text{and} \tag{18}$$

$$\boldsymbol{X} = \sqrt{P_\mathrm{T}} \cdot \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\mathrm{T} \ , \tag{19}$$

and (12) becomes

$$\mathrm{E}[z_i | y_i, \widehat{h}_\mathrm{ML}] = \sqrt{P_\mathrm{T}} x_i \widehat{h}_\mathrm{ML} + y_i \sqrt{\frac{2\sigma_\nu^2}{\pi}} \cdot \frac{\mathrm{e}^{-\frac{\widehat{h}_\mathrm{ML}^2 P_\mathrm{T}}{2\sigma_\nu^2}}}{\mathrm{erfc}\left(-\frac{x_i y_i \widehat{h}_\mathrm{ML}}{\sqrt{2\sigma_\nu^2/P_\mathrm{T}}}\right)} \ . \tag{20}$$

Since $\left(\boldsymbol{X}^\mathrm{T}\boldsymbol{X}\right) = nP_\mathrm{T}$, we can rewrite (5) as:

$$\widehat{h}_\mathrm{ML} = \frac{1}{n\sqrt{P_\mathrm{T}}} \sum_{i=1}^{n} x_i \cdot \mathrm{E}[z_i | y_i, \widehat{h}_\mathrm{ML}] \ . \tag{21}$$

When we substitute (20) into (21), we find

$$\widehat{h}_\mathrm{ML} = \widehat{h}_\mathrm{ML} + (n - n_\mathrm{e}) \sqrt{\frac{2\sigma_\nu^2}{\pi}} \cdot \frac{\mathrm{e}^{-\frac{\widehat{h}_\mathrm{ML}^2 P_\mathrm{T}}{2\sigma_\nu^2}}}{\mathrm{erfc}\left(-\frac{\widehat{h}_\mathrm{ML}}{\sqrt{2\sigma_\nu^2/P_\mathrm{T}}}\right)} + - n_\mathrm{e} \sqrt{\frac{2\sigma_\nu^2}{\pi}} \cdot \frac{\mathrm{e}^{-\frac{\widehat{h}_\mathrm{ML}^2 P_\mathrm{T}}{2\sigma_\nu^2}}}{\mathrm{erfc}\left(\frac{\widehat{h}_\mathrm{ML}}{\sqrt{2\sigma_\nu^2/P_\mathrm{T}}}\right)} \ , \tag{22}$$

where n_e is the number of observed bit-errors in the detected pilot, i.e. the number of instances where $x_i y_i = -1$. From (22) then follows that

$$\frac{\mathrm{erfc}\left(-\widehat{h}_\mathrm{ML}\sqrt{\frac{P_\mathrm{T}}{2\sigma_\nu^2}}\right)}{\mathrm{erfc}\left(\widehat{h}_\mathrm{ML}\sqrt{\frac{P_\mathrm{T}}{2\sigma_\nu^2}}\right)} = \frac{n}{n_\mathrm{e}} - 1 \tag{23}$$

must hold. With the help of the equality

$$\mathrm{erfc}(-\tau) = 2 - \mathrm{erfc}(\tau) \tag{24}$$

we finally obtain the maximum likelihood channel estimation:

$$\widehat{h}_\mathrm{ML}(n_\mathrm{e}) = \sqrt{\frac{2\sigma_\nu^2}{P_\mathrm{T}}} \cdot \mathrm{erfc}^{-1}\left(2\frac{n_\mathrm{e}}{n}\right) \ , \tag{25}$$

where $\mathrm{erfc}^{-1}(.)$ is the inverse function of $\mathrm{erfc}(.)$, and $n_\mathrm{e} \in \{1, 2, \ldots, n-1\}$. Note that if there is *no* bit-error or *all* bits are wrong, we cannot obtain a meaningful channel estimate besides the sign of h. The maximum likely estimates would be plus or minus infinity. Hence, we exclude the cases where $n_\mathrm{e} \in \{0, n\}$. In practice this indicates that either the transmit power is too high or the pilot length is too low. The estimation has then to be restarted with lower transmit power and/or larger pilot length.

6. PERFORMANCE ANALYSIS

Let us now analyze how accurately the channel is estimated and which parameters influence the accuracy. A commonly used measure for accuracy is the mean square error (MSE):

$$\epsilon = \mathrm{E}_{n_\mathrm{e}}\left[\left(h - \hat{h}_\mathrm{ML}(n_\mathrm{e})\right)^2 \mid n_\mathrm{e} \notin \{0, n\}\right], \tag{26}$$

where $\mathrm{E}_{n_\mathrm{e}}[.|.]$ is the conditional expectation with respect to n_e, which is the only random variable of the channel estimator from (25). The condition $n_\mathrm{e} \notin \{0, n\}$ makes sure, that only those cases are taken into account where a valid estimate of the channel exists. However, note that this measure does not quantify very well the quality of the channel estimation, since a given value of ϵ may describe a low quality estimate if h is low in magnitude, and a high quality estimate if h has a large magnitude. Therefore, it is better to define a *relative mean square error*:

$$\epsilon_\mathrm{rel} = \frac{\epsilon}{h^2} \tag{27}$$

$$= \mathrm{E}_{n_\mathrm{e}}\left[\left(1 - \frac{\hat{h}_\mathrm{ML}(n_\mathrm{e})}{h}\right)^2 \mid n_\mathrm{e} \notin \{0, n\}\right]. \tag{28}$$

To proceed, note that the bit-error probability is given by:

$$p_\mathrm{b} = \frac{1}{2}\mathrm{erfc}\left(h\sqrt{\frac{P_\mathrm{T}}{2\sigma_\nu^2}}\right). \tag{29}$$

From (29) and (25) then follows

$$\frac{\hat{h}_{\mathrm{ML}}(n_{\mathrm{e}})}{h} = \frac{\mathrm{erfc}^{-1}\left(2\frac{n_{\mathrm{e}}}{n}\right)}{\mathrm{erfc}^{-1}\left(2p_{\mathrm{b}}\right)}. \tag{30}$$

Note that if the observed bit-error ratio n_{e}/n equals the true bit-error probability p_{b}, the ML estimate is perfect, otherwise an estimation error occurs. By substituting (30) into (28) we obtain for the relative mean square error:

$$\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n_{\mathrm{e}}) = \mathrm{E}_{n_{\mathrm{e}}}\left[\left(1 - \frac{\mathrm{erfc}^{-1}\left(2\frac{n_{\mathrm{e}}}{n}\right)}{\mathrm{erfc}^{-1}\left(2p_{\mathrm{b}}\right)}\right)^2 \middle| \, n_{\mathrm{e}} \notin \{0, n\}\right]. \tag{31}$$

The probability that we observe exactly $n_{\mathrm{e}} = i$ bit errors is given by

$$\Pr[\, n_{\mathrm{e}} = i \,] = \binom{n}{i} p_{\mathrm{b}}^{i} \cdot (1 - p_{\mathrm{b}})^{n-i}, \tag{32}$$

since the noise samples, and hence the bit-errors are independent. The probability of observing $n_{\mathrm{e}} = i$ bit-errors conditioned on $n_{\mathrm{e}} \notin \{0, n\}$ then becomes

$$\Pr[\, n_{\mathrm{e}} = i \mid n_{\mathrm{e}} \notin \{0, n\} \,] = \binom{n}{i} \frac{p_{\mathrm{b}}^{i} \cdot (1 - p_{\mathrm{b}})^{n-i}}{1 - p_{\mathrm{b}}^{n} - (1 - p_{\mathrm{b}})^{n}}. \tag{33}$$

Note that we have

$$\sum_{i=1}^{n-1} \Pr[\, n_{\mathrm{e}} = i \mid n_{\mathrm{e}} \notin \{0, n\} \,] = 1. \tag{34}$$

The expectation in (31) can now be written with the help of (33) in the following explicit form:

$$\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n) = \sum_{i=1}^{n-1}\left(1 - \frac{\mathrm{erfc}^{-1}\left(2\frac{i}{n}\right)}{\mathrm{erfc}^{-1}\left(2p_{\mathrm{b}}\right)}\right)^2 \cdot \binom{n}{i} \cdot \frac{p_{\mathrm{b}}^{i} \cdot (1-p_{\mathrm{b}})^{n-i}}{1-p_{\mathrm{b}}^{n}-(1-p_{\mathrm{b}})^{n}}. \tag{35}$$

We see that the relative mean square error is a function of the pilot length n and the bit-error probability p_{b}.

7. PERFORMANCE OPTIMIZATION AND RESULTS

Obviously, the performance will increase with increasing pilot length n. The situation is more complicated with p_{b}. Too low values are not good since there will be no or only a few bit errors such that n_{e}/n will not be a reliable estimate of the true bit-error probability p_{b}. On the other hand, too large values of p_{b} will also not be optimum, since n_{e} will not be very sensitive to the channel coefficient h, as the received signal would be dominated by noise, jamming the pilot signal. The

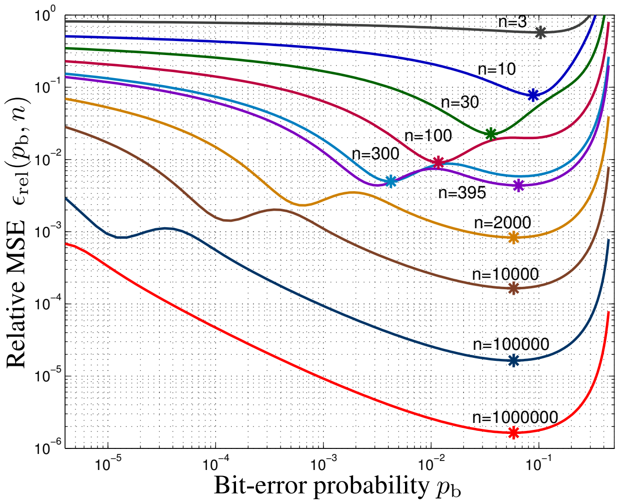

Fig. 3. Relative mean square error $\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n)$ as a function of the bit-error probability p_{b} for different pilot lengths n. The star-shaped markers indicate the position of the respective global minimum of the relative mean square error.

optimum bit-error probability for a given pilot length is given in general as:

$$p_{\mathrm{b,opt}}(n) = \arg\min_{p_{\mathrm{b}}} \; \epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n), \tag{36}$$

which leads to the *best achievable relative mean square error*:

$$\epsilon_{\mathrm{rel,opt}}(n) = \epsilon_{\mathrm{rel}}(p_{\mathrm{b,opt}}(n), n). \tag{37}$$

After substituting (35) into (36) it turns out that the optimization problem is unfortunately not convex and therefore several local minima exist. However, since the optimization has to be done with respect to one single variable p_{b}, a simple linear search through the range $(0 < p_{\mathrm{b}} < 1/2)$ can be performed efficiently. Figure 3 shows the result of this search of $\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n)$ for different values of n. Note that essentially the same behavior is seen as in the MIMO case example shown in Figure 1. It is interesting to observe that there is a *critical pilot length* n_{crit}, such that for

- ❶ $n \leq n_{\mathrm{crit}}$: the local minimum of $\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n)$ with *smallest* value of p_{b} is the global minimum, while for
- ❷ $n > n_{\mathrm{crit}}$: the local minimum of $\epsilon_{\mathrm{rel}}(p_{\mathrm{b}}, n)$ with the *largest* p_{b} is the global minimum.

From careful numerical analysis, we find:

$$n_{\mathrm{crit}} = 394, \tag{38}$$

our first "magic number". The position of the global minimum switches from the lower local minimum to the higher local minimum at the critical pilot length. Therefore, between $n = n_{\mathrm{crit}}$ and $n = n_{\mathrm{crit}} + 1$ there exists a sudden jump in $p_{\mathrm{b,opt}}$. This is best seen in Figure 4, which shows the opti-

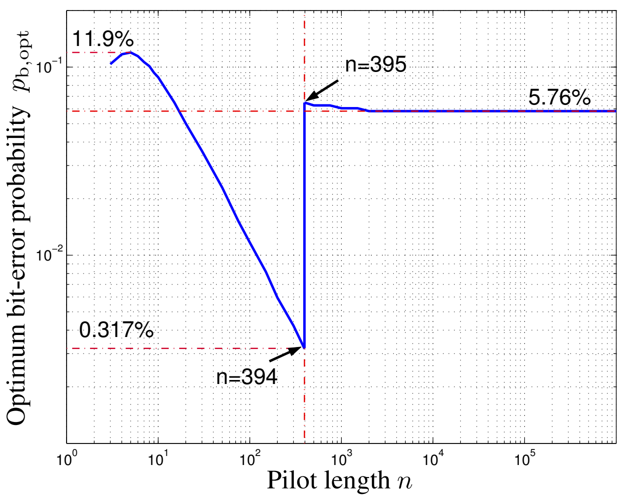

Fig. 4. Optimum bit-error probability $p_{\mathrm{b,opt}}$ as a function of the pilot length n.

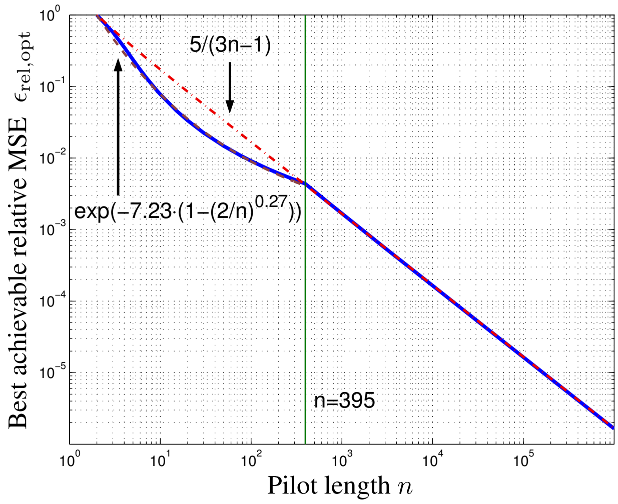

Fig. 5. Best achievable relative mean square error $\epsilon_{\mathrm{rel,opt}}(n)$ as a function of the pilot length n. The dash-dotted curves show an upper bound and an approximation.

mum bit-error probability $p_{\mathrm{b,opt}}$ that leads to the lowest relative mean square error for a given pilot length. The curve is obtained by taking the points marked with the star-shaped markers from Figure 3 for different values of n. We can see that for

❶ $n > n_{\mathrm{crit}}$: the value of $p_{\mathrm{b,opt}}(n)$ does not depend much on the pilot length n anymore. For

❷ $n \geq 2000$: the optimum bit-error probability is virtually independent of n and has the value:

$$p_{\mathrm{b,opt}}(n) = 5.76 \times 10^{-2}, \quad \text{for} \quad n > 2000, \tag{39}$$

our second "magic number". Notice that, this independence of n is also visible in the MIMO case example in Figure 1. Because the minimum of ϵ_{rel} is rather broad with respect to p_{b}, the same value of $5.76 \cdot 10^{-2}$ can be used as the bit-error probability for all pilot lengths with $n > n_{\mathrm{crit}} = 394$, in practice. For

❸ $n \leq n_{\mathrm{crit}}$: the value of $p_{\mathrm{b,opt}}(n)$ depends strongly on n and varies within the range

$$3.17 \times 10^{-3} < p_{\mathrm{b,opt}}(n) < 0.119, \quad \text{for} \quad n \leq n_{\mathrm{crit}} \tag{40}$$

in a non-monotonic way. The largest value applies for $n = 5$, while the smallest corresponds to $n = n_{\mathrm{crit}}$. The correct setting of p_{b} is also more critical than for $n > n_{\mathrm{crit}}$, since the minimum of ϵ_{rel} is more narrow.

When we make sure that the transmit power is chosen correctly, i.e. such that the corresponding p_{b} achieves its optimum value $p_{\mathrm{b,opt}}$ for a given pilot length, we obtain the best achievable relative mean square error $\epsilon_{\mathrm{ref,opt}}$ from (37). Figure 5 shows its dependency on the pilot length. It suggests that $\epsilon_{\mathrm{ref,opt}}$ can be bounded from above by:

$$\epsilon_{\mathrm{ref,opt}}(n) \leq \frac{5}{3n-1}, \tag{41}$$

for all values of n. For $n \leq n_{\mathrm{crit}}$ one can approximate by

$$\epsilon_{\mathrm{ref,opt}}(n) \approx \exp\left(-7.23\left(1-\left(\frac{2}{n}\right)^{0.27}\right)\right), \tag{42}$$

which is fairly accurate for $8 \leq n \leq n_{\mathrm{crit}}$.

8. UNQUANTIZED CHANNEL ESTIMATION

In order to obtain a better insight into the performance of the ML channel estimation after single-bit quantization, let us in the following have a look at what performance by using the unquantized received signal values

$$z_i = h\sqrt{P_{\mathrm{T}}}x_i + \nu_i \tag{43}$$

for the channel estimation. The ML estimate is well-known and given by

$$\hat{h}'_{\mathrm{ML}} = \frac{1}{n\sqrt{P_{\mathrm{T}}}}\sum_{i=1}^{n} z_i x_i. \tag{44}$$

When we compare (44) to (25), we see that the ML estimator without quantization does not depend on the variance σ_ν^2 of the noise, however the transmit power P_{T} has to be known. On the other hand, in the case of a single-bit quantization, the ratio $P_{\mathrm{T}}/\sigma_\nu^2$ has to be known, which usually means that both P_{T} and σ_ν^2 have to be known to the receiver.

Let us now have a look at the performance of the ML estimator from (44). By substituting (43) into (44), we obtain

$$\hat{h}'_{\mathrm{ML}} = h + \frac{1}{n\sqrt{P_{\mathrm{T}}}}\sum_{i=1}^{n} \nu_i x_i. \tag{45}$$

Since $\mathrm{E}[\hat{h}'_{\mathrm{ML}}] = h$, the ML estimator (44) is *unbiased*. Note that the ML estimator after single-bit quantization from (25) is however *biased*, in general.[1] This is another difference in the behavior of the ML estimators that is introduced by the non-linearity of the single-bit quantization. From (6) and (19) the estimation error becomes:

$$\epsilon' = \frac{\sigma_\nu^2}{nP_{\mathrm{T}}}, \tag{46}$$

while the relative mean square error becomes

$$\epsilon'_{\mathrm{rel}} = \frac{\epsilon'}{h^2} = \frac{\sigma_\nu^2}{nh^2 P_{\mathrm{T}}}. \tag{47}$$

From (29) we obtain

$$\frac{h^2 P_{\mathrm{T}}}{\sigma_\nu^2} = 2\left(\mathrm{erfc}^{-1}\left(2p_{\mathrm{b}}\right)\right)^2, \tag{48}$$

and by substituting into (47) we find that

$$\epsilon'_{\mathrm{rel}}(p_{\mathrm{b}}, n) = \frac{1}{2n\left(\mathrm{erfc}^{-1}\left(2p_{\mathrm{b}}\right)\right)^2}. \tag{49}$$

In contrast to the case of the single-bit quantization (35), the relative mean square error (49) when using the unquantized signal is strictly decreasing with decreasing p_{b} and approaching zero from above for p_{b} tending towards zero.

In order to obtain insight into the loss of performance that comes with the single-bit quantization, let us consider the following situation. A ML channel estimation is performed with single-bit quantized received signal and the bit-error probability p_{b} is set to the optimum value according to (36). Let us compare the performance of this estimation to the performance of the estimation with an unquantized signal. From (39) we know that for $n > 2000$ the optimum bit-error probability is given by $p_{\mathrm{b,opt}} = 5.76 \times 10^{-2}$. By substituting this value into (49) we obtain

$$\epsilon'_{\mathrm{rel}}(5.76 \times 10^{-2}, n) = \frac{0.403}{n}. \tag{50}$$

By comparison with (41) we see that

$$\frac{\epsilon_{\mathrm{rel,opt}}(n)}{\epsilon'_{\mathrm{rel}}(5.76 \times 10^{-2}, n)} \approx 4.14, \quad \text{for } n > 2000. \tag{51}$$

This shows that for a large enough pilot length ($n > 2000$), we could gain only about a factor 4 in terms of the relative mean square error, by taking the unquantized received signal instead of only its sign. Note that, the same effect can be observed for the MIMO case example in Figure 1. This corresponds to a factor of about 2 in terms of the relative root mean square error, which is a remarkably small loss for a single-bit quantization.

[1] One can set the bit-error probability such that the estimate becomes unbiased. However, for other values of the bit-error probability the estimate is biased.

9. CONCLUSION

Some insight into the effects of single-bit signal quantization on MIMO channel estimation is provided. It is demonstrated that, in contrast to unquantized channel estimation, different orthogonal pilot sequences (with same average total transmit power and same length) yield different performances. Especially, orthogonality in the time-domain (time-multiplexed pilots) can be preferable to orthogonality in space. With orthogonal pilots that are multiplexed in time, the problem can be reduced to the SIMO and finally to the SISO case, where a closed-form solution can be found for the maximum likelihood channel estimation problem, and the performance analyzed analytically. By focusing on binary pilot sequences, it turns out that the performance critically depends on the probability that a pilot symbol is detected wrongly. The optimum probability depends on the length of the pilot, however, for large pilots (2000 or more symbols), this optimum probability converges to about 5%. The lowest achievable mean square error turns out to be roughly only four times larger than the mean square error obtained from the unquantized signal at the same transmit power.

10. ACKNOWLEDGMENT

The authors would like to acknowledge the helpful comments of the anonymous reviewers, which helped in improving the quality of the paper.

11. REFERENCES

[1] A.V. Oppenheim and R.W. Schafer, *Discrete Time Signal Processing*, Prentice Hall, 1989.

[2] M.T. Ivrlač and J.A. Nossek, "Challenges in Coding for Quantized MIMO Systems," in *Proc. IEEE International Symposium on Information Theory, ISIT2006*, Seattle, U.S.A., July 2006, pp. 2114–2118.

[3] J.A. Nossek and Michel T. Ivrlač, "Capacity and Coding for Quantized MIMO Systems," in *Proc. International Wireless Communications & Mobile Computing Conference, IWCMC06*, Vancouver, Canada, July 2006, pp. 1387–1391.

[4] R. Schreier and G.C. Temes, *Understanding Delta-Sigma Data Converters*, IEEE Computer Society Press, 2004.

[5] T.M. Lok and V.K.W. Wei, "Channel Estimation with Quantized Observations," in *Proc. IEEE International Symposium on Information Theory, ISIT1998*, Cambridge, MA, U.S.A., August 1998, p. 333.

[6] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.