Channel Estimation using 1-bit Quantization and Oversampling for Large-scale Multiple-antenna Systems

Zhichao Shao
Lukas T. N. Landau
Rodrigo C. de Lamare

Center for Telecommunications Studies
PUC-Rio

May 2019

Outline

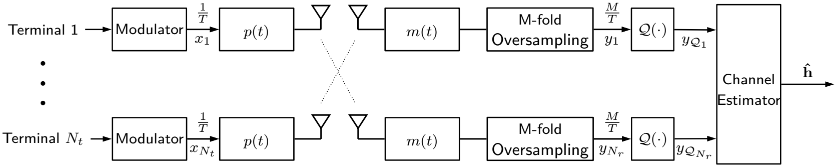
Introduction

- One key technology for 5G: Large-scale MIMO or massive MIMO [1].
  - High spectral efficiency
  - High reliability
- Problem: with the increasing number of BS antennas the energy consumption will grow, especially when using high resolution ADCs for each antenna.
- Solution: low-cost and low-resolution ADCs. As one extreme case, 1-bit ADCs at the front-end can dramatically decrease the receiver energy consumption.
- Linear receive processing can be adjusted to the case of low-resolution ADCs.
- The loss of information due to the quantization can be partially compensated by oversampling [2].

[1] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, "Massive MIMO for next generation wireless systems," in IEEE Communications Magazine, vol. 52, no. 2, pp. 186-195, February 2014.

[2] L. Landau, M. Dörpinghaus and G. P. Fettweis, "1-Bit Quantization and Oversampling at the Receiver: Communication Over Bandlimited Channels With Noise," in IEEE Communications Letters, vol. 21, no. 5, pp. 1007-1010, May 2017.

System Model



$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ is the $NN_t \times 1$ column vector transmitted by N_t terminals for a block of symbols with length N given by

$$\mathbf{x} = \left[x_{1,1} \cdots x_{N,1} \middle| x_{1,2} \cdots x_{N,2} \middle| x_{1,3} \cdots x_{N,N_t}\right]^T, \tag{2}$$

where $x_{i,j}$ corresponds to the transmitted symbol of terminal j at time instant i.

System Model

The vector $\mathbf{n}$ is the filtered oversampled noise vector of size $MN_rN \times 1$ expressed by

$$\mathbf{n} = \left(\mathbf{I}_{N_r} \otimes \mathbf{G}\right)\mathbf{w}. \tag{3}$$

where the noise vector $\mathbf{w} \sim \mathcal{CN}\left(\mathbf{0}_{3MN_rN}, \sigma_n^2 \mathbf{I}_{3MN_rN}\right)$ contains independent and identically distributed (IID) complex Gaussian random variables with zero mean and variance σ_n^2.

$\mathbf{G}$ is a Toeplitz matrix, which contains the coefficients of the matched filter $m(t)$ at different time instants

$$\mathbf{G} = \begin{bmatrix} m(-NT) & m(-NT+\frac{1}{M}T) & \dots & m(NT) & 0 & \dots & 0 \\ 0 & m(-NT) & \dots & m(NT-\frac{1}{M}T) & m(NT) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m(-NT) & m(-NT+\frac{1}{M}T) & \dots & m(NT) \end{bmatrix}_{MN \times 3MN} \tag{4}$$

System Model

The equivalent channel matrix $\mathbf{H}$ is described as

$$\mathbf{H} = (\mathbf{I}_{N_r} \otimes \mathbf{Z})\, \mathbf{U}\, (\mathbf{H}' \otimes \mathbf{I}_N)\,, \tag{5}$$

where $\mathbf{H}'$ is an $N_r \times N_t$ matrix whose element in the ith row and jth column corresponds to the channel coefficient between terminal j and receive antenna i. $\mathbf{U}$ is an oversampling matrix, which can be calculated as

$$\mathbf{U} = \mathbf{I}_{N_r N} \otimes [0 \quad \cdots \quad 0 \quad 1]^T_{1\times M}\,. \tag{6}$$

$\mathbf{Z}$ is a Toeplitz matrix that contains the coefficients of $z(t)$ at different time instants, where $z(t)$ is the convolution product of $p(t)$ and $m(t)$, and is given by

$$\mathbf{Z} = \begin{bmatrix} z(0) & z(\frac{T}{M}) & \dots & z(NT - \frac{1}{M}T) \\ z(-\frac{T}{M}) & z(0) & \dots & z(NT - \frac{2}{M}T) \\ \vdots & \vdots & \ddots & \vdots \\ z(-NT + \frac{1}{M}T) & z(-NT + \frac{2}{M}T) & \dots & z(0) \end{bmatrix}_{MN\times MN}\,. \tag{7}$$

Bayesian Bound

The system model in (1) can be vectorized as

$$\mathbf{y} = [\mathbf{x}^T \otimes \mathbf{I}_{N_r} \otimes \mathbf{Z}(\mathbf{I}_N \otimes \mathbf{u})]\text{vec}(\mathbf{H}' \otimes \mathbf{I}_N) + \mathbf{n} \tag{8}$$

with the property of vectorization and Kronecker products

$$\text{vec}(\mathbf{H}' \otimes \mathbf{I}_N) = [\mathbf{I}_{N_t} \otimes (\mathbf{e}_1 \otimes \mathbf{I}_{N_r} \otimes \mathbf{e}_1 + \cdots + \mathbf{e}_N \otimes \mathbf{I}_{N_r} \otimes \mathbf{e}_N)]\,\text{vec}(\mathbf{H}'), \tag{9}$$

where $\mathbf{e}_n$ is an all-zero column vector except that the nth element is one. Eq.(8) can be written in the following simplified form

$$\mathbf{y} = \mathbf{\Phi}\text{vec}(\mathbf{H}') + \mathbf{n} = \mathbf{\Phi}\mathbf{h}' + \mathbf{n}, \tag{10}$$

where $\mathbf{\Phi} \in \mathbb{C}^{MNN_r \times N_r N_t}$ is called the equivalent transmit matrix.
Eq.(10) can be rewritten in the real-valued form as

$$\begin{bmatrix} \mathbf{y}^R \\ \mathbf{y}^I \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}^R & -\mathbf{\Phi}^I \\ \mathbf{\Phi}^I & \mathbf{\Phi}^R \end{bmatrix} \begin{bmatrix} \mathbf{h}'^R \\ \mathbf{h}'^I \end{bmatrix} + \begin{bmatrix} \mathbf{n}^R \\ \mathbf{n}^I \end{bmatrix}. \tag{11}$$

Bayesian Bound

Considering the unknown parameter vector $\tilde{\mathbf{h}'} = [\mathbf{h'}^R; \mathbf{h'}^I]$ and with the independence of $\mathbf{y}^R$ and $\mathbf{y}^I$, the Bayesian information matrix (BIM) is defined as

$$\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'}) = \mathbf{J}_{\mathbf{y}_\mathcal{Q}^R}(\tilde{\mathbf{h}'}) + \mathbf{J}_{\mathbf{y}_\mathcal{Q}^I}(\tilde{\mathbf{h}'}), \tag{12}$$

where

$$[\mathbf{J}_{\mathbf{y}_\mathcal{Q}^{R/I}}(\tilde{\mathbf{h}'})]_{ij} = E_{\mathbf{y}_\mathcal{Q}^{R/I},\tilde{\mathbf{h}'}} \left\{ \frac{\partial \ln p(\mathbf{y}_\mathcal{Q}^{R/I}, \tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_i} \frac{\partial \ln p(\mathbf{y}_\mathcal{Q}^{R/I}, \tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_j} \right\}, \tag{13}$$

with $[\tilde{\mathbf{h}'}]_i$ and $[\tilde{\mathbf{h}'}]_j$ being the elements of $\tilde{\mathbf{h}'}$ and $\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'})$ is arranged as follows:

$$\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'}) = \begin{bmatrix} [\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'})]_{RR} & [\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'})]_{RI} \\ [\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'})]_{IR} & [\mathbf{J}_{\mathbf{y}_\mathcal{Q}}(\tilde{\mathbf{h}'})]_{II} \end{bmatrix}. \tag{14}$$

H. L. Van Trees, K. L. Bell, and Z. Tian, Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory, Detection Estimation and Modulation Theory. Wiley, 2013.

Bayesian Bound

Eq.(13) can be divided into two parts

$$[\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})]_{ij} = [\mathbf{J}^{D}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})]_{ij} + [\mathbf{J}^{P}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})]_{ij}, \tag{15}$$

where

$$[\mathbf{J}^{D}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})]_{ij} \triangleq E_{\mathbf{y}_{\mathcal{Q}}^{R/I}|\tilde{\mathbf{h}'}} \left\{ \frac{\partial \ln p(\mathbf{y}_{\mathcal{Q}}^{R/I} \mid \tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_i} \frac{\partial \ln p(\mathbf{y}_{\mathcal{Q}}^{R/I} \mid \tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_j} \right\} \tag{16}$$

and

$$[\mathbf{J}^{P}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})]_{ij} \triangleq E_{\tilde{\mathbf{h}'}} \left\{ \frac{\partial \ln p(\tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_i} \frac{\partial \ln p(\tilde{\mathbf{h}'})}{\partial [\tilde{\mathbf{h}'}]_j} \right\}. \tag{17}$$

To transform the real value $\tilde{\mathbf{h}'}$ back to the complex domain $\mathbf{h}'$, we apply the chain rule to get:

$$\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\mathbf{h}') = \frac{1}{4}\left([\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'})]_{RR} + [\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'})]_{II}\right) + \frac{j}{4}\left([\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'})]_{RI} - [\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'})]_{IR}\right). \tag{18}$$

The variance of the LMMSE estimator $\hat{\mathbf{h}'}(\mathbf{y}_{\mathcal{Q}})$ is lower bounded by

$$\text{var}[\hat{h'}_i(\mathbf{y}_{\mathcal{Q}})] \geq [\mathbf{J}^{-1}_{\mathbf{y}_{\mathcal{Q}}}(\mathbf{h}')]_{ii}. \tag{19}$$

BIM for Non-oversampled Systems

For non-oversampled system, the noise vector $\mathbf{n}$ has the covariance matrix $\mathbf{C}_{\mathbf{n}} = \sigma_n^2 \mathbf{I}_{NN_r}$. The conditional log-likelihood function can be expressed as

$$\ln p(\mathbf{y}_{\mathcal{Q}} \mid \tilde{\mathbf{h}'}) = \sum_{k=1}^{NN_r} \left[\ln p([\mathbf{y}_{\mathcal{Q}}^R]_k \mid [\tilde{\mathbf{h}'}]_k) + \ln p([\mathbf{y}_{\mathcal{Q}}^I]_k \mid [\tilde{\mathbf{h}'}]_k) \right], \qquad (20)$$

Inserting (20) into (16) we obtain

$$[\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}^D(\tilde{\mathbf{h}'})]_{ij} = -E\left\{ \frac{\partial^2 \ln p(\mathbf{y}_{\mathcal{Q}} \mid \tilde{\mathbf{h}'})}{\partial[\tilde{\mathbf{h}'}]_i \partial[\tilde{\mathbf{h}'}]_j} \right\} = [\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}^R}^D(\tilde{\mathbf{h}'})]_{ij} + [\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}^I}^D(\tilde{\mathbf{h}'})]_{ij}. \qquad (21)$$

With the assumption $\tilde{\mathbf{h}'} \in \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\mathbf{h}'}} = \frac{1}{2}\mathbf{I}_2 \otimes \mathbf{C}_{\mathbf{h}'})$, $\ln p(\tilde{\mathbf{h}'})$ yields

$$\ln p(\tilde{\mathbf{h}'}) = -\frac{1}{2} N_r N_t \ln\left[(2\pi)^{2N_r N_t} \det(\mathbf{C}_{\tilde{\mathbf{h}'}}) \right] - \frac{1}{2} \tilde{\mathbf{h}'}^T \mathbf{C}_{\tilde{\mathbf{h}'}}^{-1} \tilde{\mathbf{h}'} \qquad (22)$$

and inserted into (17) we obtain

$$\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}^P(\tilde{\mathbf{h}'}) = 2\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}^P(\tilde{\mathbf{h}'}) = 2\mathbf{C}_{\tilde{\mathbf{h}'}}^{-1}. \qquad (23)$$

The BIM is the summation of (21) and (23) as described by

$$\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'}) = \mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}^D(\tilde{\mathbf{h}'}) + \mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}^P(\tilde{\mathbf{h}'}). \qquad (24)$$

BIM for Oversampled Systems

When $M \geq 2$ the noise vector $\mathbf{n}$ contains correlated noise samples. Computing the exact form of $p(\mathbf{y}_{\mathcal{Q}}^{R/I} \mid \tilde{\mathbf{h}'})$ is not available. Instead, the authors in [3] have given a lower bound of $\mathbf{J}^{D}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'})$, which is

$$\mathbf{J}^{D}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'}) \geq \left(\frac{\partial \mu_{\mathbf{y}_{\mathcal{Q}}^{R/I}}}{\partial \tilde{\mathbf{h}'}}\right)^{T} \mathbf{C}^{-1}_{\mathbf{y}_{\mathcal{Q}}^{R/I}} \left(\frac{\partial \mu_{\mathbf{y}_{\mathcal{Q}}^{R/I}}}{\partial \tilde{\mathbf{h}'}}\right) = \tilde{\mathbf{J}}^{D}_{\mathbf{y}_{\mathcal{Q}}^{R/I}}(\tilde{\mathbf{h}'}). \tag{25}$$

where the equality holds for $M = 1$. Based on [4], the mean value of the real part of the kth received symbol is given by

$$\begin{aligned}
[\mu_{\mathbf{y}_{\mathcal{Q}}^{R}}]_k &= \frac{1}{\sqrt{2}} p\left([\mathbf{y}_{\mathcal{Q}}^{R}]_k = +1 \mid \tilde{\mathbf{h}'}\right) - \frac{1}{\sqrt{2}} p\left([\mathbf{y}_{\mathcal{Q}}^{R}]_k = -1 \mid \tilde{\mathbf{h}'}\right) \\
&= \frac{1}{\sqrt{2}}\left[1 - 2Q\left(\frac{[\mathbf{\Phi}^{R}\mathbf{h'}^{R} - \mathbf{\Phi}^{I}\mathbf{h'}^{I}]_k}{\sqrt{[\mathbf{C}_{\mathbf{n}}]_{kk}/2}}\right)\right],
\end{aligned} \tag{26}$$

[3] M. Stein, A. Mezghani, and J. A. Nossek, "A Lower Bound for the Fisher Information Measure," IEEE Signal Process. Lett., vol. 21, no. 7, pp. 796–799, Jul. 2014.

[4] M. Schlüter and M. Dörpinghaus and G. P. Fettweis, "Bounds on Channel Parameter Estimation with 1-Bit Quantization and Oversampling," in 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Jun. 2018, pp. 1–5.

BIM for Oversampled Systems

The derivative of (26) is

$$\frac{\partial[\mu_{\mathbf{y}_{\mathcal{Q}}^R}]_k}{\partial[\tilde{\mathbf{h}'}]_i} = \frac{2\text{exp}\left(-\frac{[\boldsymbol{\Phi}^R\mathbf{h'}^R-\boldsymbol{\Phi}^I\mathbf{h'}^I]_k^2}{[\mathbf{C}_\mathbf{n}]_{kk}}\right)\frac{\partial[\boldsymbol{\Phi}^R\mathbf{h'}^R-\boldsymbol{\Phi}^I\mathbf{h'}^I]_k}{\partial[\tilde{\mathbf{h}'}]_i}}{\sqrt{2\pi[\mathbf{C}_\mathbf{n}]_{kk}}}. \quad (27)$$

The diagonal elements of the covariance matrix are given by

$$[\mathbf{C}_{\mathbf{y}_{\mathcal{Q}}^R}]_{kk} = \frac{1}{2} - [\mu_{\mathbf{y}_{\mathcal{Q}}^R}]_k^2, \quad (28)$$

while the off-diagonal elements are calculated as

$$[\mathbf{C}_{\mathbf{y}_{\mathcal{Q}}^R}]_{kn} = p(z_k > 0, z_n > 0) + p(z_k \leq 0, z_n \leq 0) - \frac{1}{2} - [\mu_{\mathbf{y}_{\mathcal{Q}}^R}]_k[\mu_{\mathbf{y}_{\mathcal{Q}}^R}]_n, \quad (29)$$

where $[z_k, z_n]^T$ is a bi-variate Gaussian random vector

$$\begin{bmatrix} z_k \\ z_n \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} [\boldsymbol{\Phi}^R\mathbf{h'}^R - \boldsymbol{\Phi}^I\mathbf{h'}^I]_k \\ [\boldsymbol{\Phi}^R\mathbf{h'}^R - \boldsymbol{\Phi}^I\mathbf{h'}^I]_n \end{bmatrix}, \frac{1}{2}\begin{bmatrix} [\mathbf{C}_\mathbf{n}]_{kk} & [\mathbf{C}_\mathbf{n}]_{kn} \\ [\mathbf{C}_\mathbf{n}]_{nk} & [\mathbf{C}_\mathbf{n}]_{nn} \end{bmatrix}\right).$$

The lower bound for the imaginary part is derived in the same way. We get the lower bound of the BIM as

$$\mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}(\tilde{\mathbf{h}'}) \geq \tilde{\mathbf{J}}_{\mathbf{y}_{\mathcal{Q}}}^D(\tilde{\mathbf{h}'}) + \mathbf{J}_{\mathbf{y}_{\mathcal{Q}}}^P(\tilde{\mathbf{h}'}). \quad (30)$$

Oversampling based LRA-LMMSE Channel Estimation

During the training phase, all terminals simultaneously transmit τ pilot sequences to the BS, which yields

$$\mathbf{y}_{\mathcal{Q}_p} = \mathcal{Q}(\mathbf{\Phi}_p \mathbf{h}' + \mathbf{n}_p) = \tilde{\mathbf{\Phi}} \mathbf{h}' + \tilde{\mathbf{n}}_p, \quad (31)$$

where $\tilde{\mathbf{\Phi}}_p = \mathbf{A}_p \mathbf{\Phi}_p$ and $\tilde{\mathbf{n}}_p = \mathbf{A}_p \mathbf{n}_p + \mathbf{n}_q$. The vector $\mathbf{n}_q$ is the statistically equivalent quantizer noise. The matrix $\mathbf{A}_p$ is the linear operator chosen independently from $\mathbf{y}_p$, which yields

$$\mathbf{A}_p = \mathbf{C}^H_{\mathbf{y}_p \mathbf{y}_{\mathcal{Q}_p}} \mathbf{C}^{-1}_{\mathbf{y}_p} = \sqrt{\frac{2}{\pi}} \mathrm{diag}\left(\mathbf{C}_{\mathbf{y}_p}\right)^{-\frac{1}{2}}, \quad (32)$$

where $\mathbf{C}_{\mathbf{y}_p \mathbf{y}_{\mathcal{Q}_p}}$ denotes the cross-correlation matrix between the received signal $\mathbf{y}_p$ and the quantized signal $\mathbf{y}_{\mathcal{Q}_p}$ and $\mathbf{C}_{\mathbf{y}_p}$ is given by

$$\mathbf{C}_{\mathbf{y}_p} = \mathbf{\Phi}_p \mathbf{C}_{\mathbf{h}'} \mathbf{\Phi}_p^H + \sigma_n^2 \mathbf{I}_{N_r} \otimes \mathbf{G}\mathbf{G}^H. \quad (33)$$

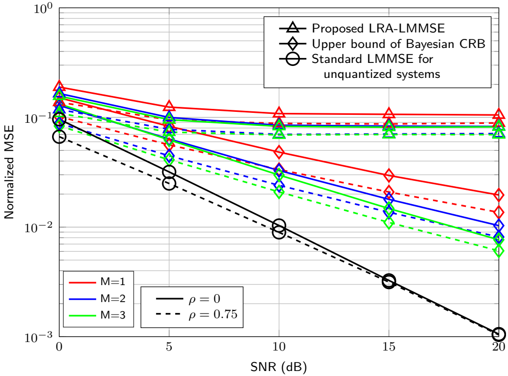
Based on the equivalent linear model (31), the proposed oversampling based LRA-LMMSE channel estimator is given by

$$\hat{\mathbf{h}'}_{\text{LMMSE}} = \mathbf{C}_{\mathbf{h}'} \tilde{\mathbf{\Phi}}^H \mathbf{C}^{-1}_{\mathbf{y}_{\mathcal{Q}_p}} \mathbf{y}_{\mathcal{Q}_p}. \quad (34)$$

J. J. Bussgang, "Crosscorrelation functions of amplitude distorted Gaussian signals," Res. Lab. Elec., Mas. Inst. Technol., vol. Tech. Rep. 216, Mar. 1952.

Numerical Results

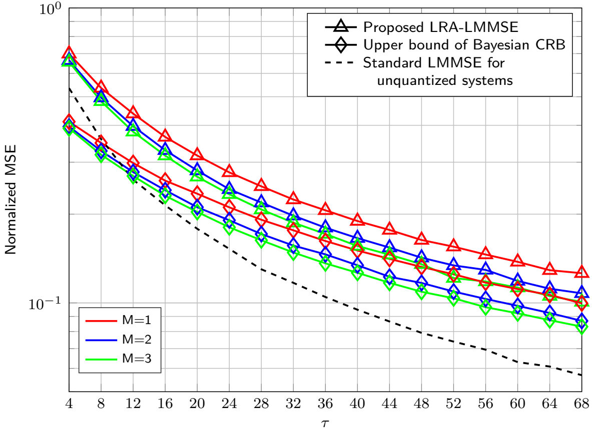
- $N_t = 4$ and $N_r = 16$. Modulation scheme: QPSK.
- Block fading channel with the Kronecker model $\mathbf{H}' = \mathbf{R}_r^{\frac{1}{2}} \mathbf{H}'_w \mathbf{R}_t^{\frac{1}{2}}$ and $\tau = 40$.
- Normalized RRC pulse filter with a roll-off factor of 0.8.



Normalized MSE comparison as a function of SNR.

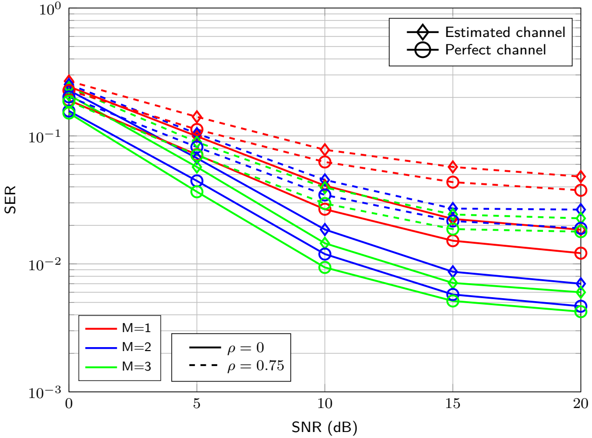
Da-Shan Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," IEEE Trans. Commun., vol. 48, no. 3, pp. 502–513, Mar. 2000.

Numerical Results



Normalized MSE comparison as a function of τ when $\text{SNR} = 0\text{dB}$ and $\rho = 0$.

Numerical Results



SER performance comparison between different oversampling factors.

Conclusion

- This work has proposed the LRA-LMMSE channel estimator for uplink large-scale MIMO systems with 1-bit quantization and oversampling at the receiver.
- We have further characterized the system performance analytically in terms of the Bayesian information.
- The simulation results have shown that the proposed oversampling based channel estimator outperforms the existing non-oversampled BLMMSE channel estimator in terms of the MSE and the SER performances.