Why Linear MMSE Detection is not Necessarily Superior to Linear ZF Detection

Jan Mietzner and Peter A. Hoeher

Information and Coding Theory Lab, Faculty of Engineering, University of Kiel, Kaiserstrasse 2, D-24143 Kiel, Germany
URL: `www-ict.tf.uni-kiel.de` E-mail: `{jm,ph}@tf.uni-kiel.de`

***Abstract*— Based on a Bayesian view on linear detection, we demonstrate that in contrast to a common belief the symbol error rate (SER) performance of the linear minimum mean square error (MMSE) detector is not necessarily superior to that of the linear zero forcing (ZF) detector.**

I. Introduction

Consider the following matrix-vector communication model, which represents a discrete-time, non-dispersive multiple-input multiple-output (MIMO) system with M transmit and $N \geq M$ receive antennas[1]:

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

where $\mathbf{y} \in \mathbb{C}^{N\times 1}$ denotes the received vector, $\mathbf{H} \in \mathbb{C}^{N\times M}$ the channel matrix, $\mathbf{x} \in \mathbb{C}^{M\times 1}$ the transmitted vector, and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ an additive noise vector. The entries of $\mathbf{x}$ are assumed to be independent and identically distributed (i.i.d.) zero-mean random variables with variance σ_x^2 that are drawn from a finite symbol alphabet $\mathcal{A}$. (Correspondingly, $\mathbf{R}_\mathbf{x} := \mathsf{E}\{\mathbf{x}\mathbf{x}^\mathrm{H}\} = \sigma_x^2 \mathbf{I}_M$, where $\mathbf{I}_M$ denotes the $(M\times M)$-identity matrix.) The entries of $\mathbf{H}$ are assumed to be i.i.d. circularly symmetric complex Gaussian random variables with zero mean and unit variance. The entries of $\mathbf{n}$ are assumed to be i.i.d. circularly symmetric complex Gaussian random variables with zero mean and variance σ_n^2. Throughout this letter, we assume that the channel matrix $\mathbf{H}$ is known at the receiver. Moreover, we assume that $\mathbf{x}$, $\mathbf{H}$, and $\mathbf{n}$ are statistically independent.

Two popular solutions for linear detection of the transmitted vector $\mathbf{x}$ are the *linear ZF detector* and the *linear MMSE detector* [1]. The ZF detector performs an inversion of the channel matrix $\mathbf{H}$ and a subsequent element-wise hard decision (HD) with respect to the symbol alphabet $\mathcal{A}$:

$$\hat{\mathbf{x}}_{\mathrm{ZF}} := \mathrm{HD}_\mathcal{A}\left\{\mathbf{H}^+\,\mathbf{y}\right\}, \tag{2}$$

where $\mathbf{H}^+ := (\mathbf{H}^\mathrm{H}\mathbf{H})^{-1}\,\mathbf{H}^\mathrm{H}$ denotes the left-hand pseudo-inverse of $\mathbf{H}$. Correspondingly, the ZF detector (if it exists) exactly removes the spatial interference between the transmitted data symbols. However, if $\mathbf{H}^\mathrm{H}\mathbf{H}$ is ill-conditioned, the ZF detector is known to cause a significant enhancement and (spatial) coloring of the noise. The MMSE solution is generally considered to be superior to the ZF solution, because it provides an optimum trade-off (in the MMSE sense) between interference suppression and noise enhancement. The MMSE estimate of $\mathbf{x}$ is given by

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} := \mathrm{HD}_\mathcal{A}\left\{\left(\mathbf{H}^\mathrm{H}\mathbf{H} + \sigma_n^2/\sigma_x^2\cdot\mathbf{I}_M\right)^{-1}\mathbf{H}^\mathrm{H}\,\mathbf{y}\right\}. \tag{3}$$

For high signal-to-noise ratios (SNRs), i.e., $\sigma_x^2/\sigma_n^2 \rightarrow \infty$, the MMSE solution and the ZF solution are equivalent.

[1] In principle, the ideas presented here apply also for other matrix-vector communication systems, such as multi-user systems with a linear detector or dispersive systems with a linear equalizer.

In this letter, we demonstrate that the SER performance of the linear MMSE detector is not necessarily superior to that of the linear ZF detector. Motivated by a Bayesian view on linear detection, we give an example for which the ZF detector outperforms the MMSE detector.

II. Linear Bayesian Detector

If the a-priori probability distribution $\mathrm{P}(\mathbf{x})$ of the transmitted vector $\mathbf{x}$ is of form

$$\mathrm{P}(\mathbf{x}) \propto \exp\Big(-\frac{\mathbf{x}^\mathrm{H}\mathbf{x}}{\sigma_x^2}\Big), \tag{4}$$

the corresponding a-posteriori probability distribution $\mathrm{P}(\mathbf{x}|\mathbf{y})$ of the transmitted vector $\mathbf{x}$ given the received vector $\mathbf{y}$ is characterized by the following general expression [2]:

$$\mathrm{P}(\mathbf{x}|\mathbf{y}) \propto \exp\bigg(-\frac{1}{\sigma_n^2}\Big(\mathbf{x} - \mathbf{R}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)^\mathrm{H}\mathbf{R}^{-1}\Big(\mathbf{x} - \mathbf{R}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)\bigg), \tag{5}$$

where $\mathbf{R}$ is an arbitrary invertible $(M\times M)$-matrix. Given this general expression, the so-called *linear Bayesian detector* [2] is given by

$$\hat{\mathbf{x}}_\mathrm{B} := \mathrm{HD}_\mathcal{A}\Big\{\mathbf{R}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big\}. \tag{6}$$

Ignoring for the moment that the entries of the transmitted vector $\mathbf{x}$ are drawn from a finite symbol alphabet, we consider the following two special cases:

(i) The vector $\mathbf{x}$ is *uniformly distributed* on $\mathbb{C}^M$ (i.e., in (4) we let $\sigma_x^2 \rightarrow \infty$). In this case, $\mathrm{P}(\mathbf{x}|\mathbf{y})$ can be written as [2]

$$\begin{aligned}\mathrm{P}(\mathbf{x}|\mathbf{y}) \;\propto\; \exp\bigg(&-\frac{1}{\sigma_n^2}\Big(\mathbf{x} - \big(\mathbf{H}^\mathrm{H}\mathbf{H}\big)^{-1}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)^\mathrm{H} \times \\ &\times\; \mathbf{H}^\mathrm{H}\mathbf{H}\Big(\mathbf{x} - \big(\mathbf{H}^\mathrm{H}\mathbf{H}\big)^{-1}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)\bigg).\end{aligned} \tag{7}$$

Identifying $(\mathbf{H}^\mathrm{H}\mathbf{H})^{-1} =: \mathbf{R}$, we note that in this case the linear Bayesian detector (6) is equivalent to the linear ZF detector (2). This means that the ZF detector *implicitly* assumes a uniform distribution of $\mathbf{x}$ on $\mathbb{C}^M$ [2]. (Specifically, it does not exploit the finite-alphabet property of $\mathbf{x}$.)

(ii) The vector $\mathbf{x}$ is *Gaussian distributed* on $\mathbb{C}^M$ (with $\sigma_x^2 < \infty$). In this case, $\mathrm{P}(\mathbf{x}|\mathbf{y})$ can be written as [2]

$$\begin{aligned}&\mathrm{P}(\mathbf{x}|\mathbf{y}) \quad \propto \\ &\exp\bigg(-\frac{1}{\sigma_n^2}\Big(\mathbf{x} - \big(\mathbf{H}^\mathrm{H}\mathbf{H} + \sigma_n^2/\sigma_x^2\cdot\mathbf{I}_M\big)^{-1}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)^\mathrm{H} \times \\ &\quad\times\quad \big(\mathbf{H}^\mathrm{H}\mathbf{H} + \sigma_n^2/\sigma_x^2\cdot\mathbf{I}_M\big) \quad\times \\ &\quad\times\quad \Big(\mathbf{x} - \big(\mathbf{H}^\mathrm{H}\mathbf{H} + \sigma_n^2/\sigma_x^2\cdot\mathbf{I}_M\big)^{-1}\mathbf{H}^\mathrm{H}\,\mathbf{y}\Big)\bigg).\end{aligned} \tag{8}$$

Identifying $(\mathbf{H}^{\mathrm{H}}\mathbf{H}+\sigma_n^2/\sigma_x^2\cdot\mathbf{I}_M)^{-1}=:\mathbf{R}$, we note that in this case the linear Bayesian detector (6) is equivalent to the linear MMSE detector (3). This means that the MMSE detector *implicitly* assumes a Gaussian distribution of $\mathbf{x}$ on $\mathbb{C}^M$ (and does not exploit the finite-alphabet property of $\mathbf{x}$, similarly to the ZF detector) [2].

Given a *practical* symbol alphabet $\mathcal{A}$, such as a phase shift keying (PSK) or a quadrature amplitude modulation (QAM) constellation, the SER performance of the linear MMSE detector is generally superior to that of the linear ZF detector. Obviously, this means that for practical symbol alphabets the implicit Gaussian assumption of the MMSE detector is a better fit to the actual distribution of $\mathbf{x}$ than the uniform assumption made by the ZF detector. In the following, we present a counter example, for which the linear ZF detector (slightly) outperforms the linear MMSE detector.

III. Numerical Example

We consider the case of $M=2$ transmit and $N=2$ receive antennas. Moreover, we consider the cross-shaped (complex-valued) signal constellation $\mathcal{A}$ depicted in Fig. 1. The data symbols transmitted via the first and second antenna are independently drawn from the signal constellation $\mathcal{A}$, according to a uniform distribution. (The variance σ_x^2 of the data symbols results as $\sigma_x^2=33$.)

The average SERs resulting for the linear ZF detector and the linear MMSE detector are displayed in Fig. 2 as a function of σ_x^2/σ_n^2 in dB. These average SERs have been obtained by means of Monte Carlo simulations over 10^6 independent realizations of the channel matrix $\mathbf{H}$. As can be seen, the ZF detector indeed outperforms the MMSE detector (although the difference is small). For large values of σ_x^2/σ_n^2, both detectors exhibit the same performance, as expected.

IV. Conclusions

For the example of a discrete-time, non-dispersive MIMO system we have demonstrated that the performance of the linear MMSE detector is not necessarily superior to that of the linear ZF detector. Specifically, motivated by a Bayesian view on linear detection, we have given an example for which the ZF detector outperforms the MMSE detector. The main aim was to provide novel insights into these two fundamental types of linear detector. From a practical point of view, the presented example is of rather limited interest, because the observed performance difference is rather small and the signal constellation used is not very practicable.

Acknowledgement

The authors would like to thank Ronald Böhnke (Dept. of Communications Engineering, University of Bremen, Germany) for fruitful discussions during an all too long train journey.

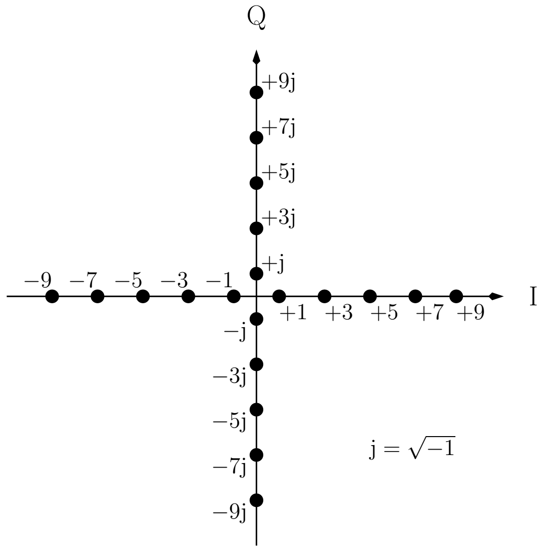

Fig. 1. Signal constellation $\mathcal{A}$ under consideration.

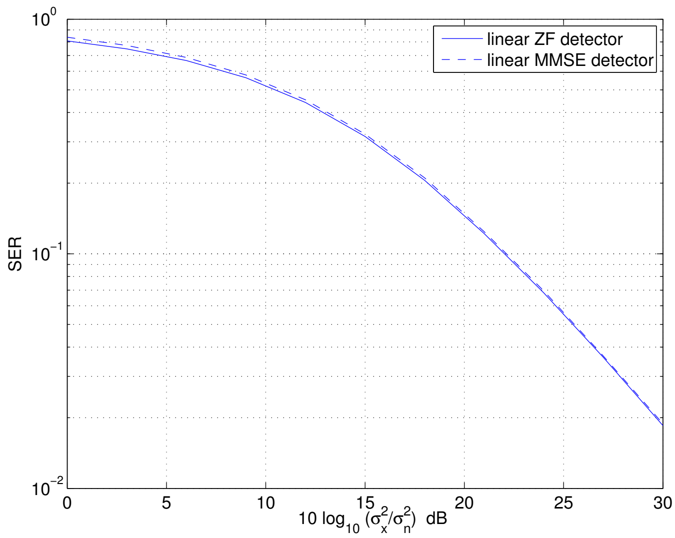

Fig. 2. Average SERs resulting for the linear ZF detector and the linear MMSE detector, as a function of σ_x^2/σ_n^2 in dB.

References

[1] E. Biglieri, G. Taricco, and A. Tulino, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1794–1803, July 2002.

[2] Y. Huang, J. Zhang, and P. M. Djurić, "Bayesian detection for BLAST," *IEEE Trans. Signal Processing*, vol. 53, no. 3, pp. 1086–1096, Mar. 2005.

◇ ◇ ◇