Achievable Rates of Nonlinear MIMO Systems with Noisy Channel State Information

Georgios K. Psaltopoulos and Armin Wittneben
Wireless Communications Group, ETH Zurich, CH-8092 Zurich, Switzerland
Email: {psaltopoulos, wittneben}@nari.ee.ethz.ch

***Abstract*— *Multiple input multiple output* (MIMO) communication systems that employ *nonlinear detectors*, i.e., amplitude-only or phase-only detectors, where first introduced in [1]. Such systems are low-complexity and low-power alternatives to legacy MIMO systems in applications, e.g. sensor networks, where low power consumption and low cost are very stringent constraints. This work extends the achievable rates study presented in [1], to cover the case where the channel estimation at the receiver contains an additive Gaussian noise term.**

I. Introduction

Multiple Input Multiple Output (MIMO) communication systems are well known for providing high data rates through exploiting the spatial dimension of the medium [2], [3]. Compared to Single Input Single Output (SISO) systems, MIMO systems achieve higher data rates for the same signal-to-noise ratio, or equivalently, spend less energy for achieving the same data rate. However, MIMO systems are not used in practice as a means to reduce transmit energy and hence power consumption, because the required MIMO circuitry renders them power-hungry. As a matter of fact, such systems require multiple transmitter/receiver chains, which means multiple low-noise amplifiers, filters, mixers, analog-to-digital converters etc. [4]. This leads to very high power consumption and hinders employment of MIMO systems for low-power applications, whereas techniques that try to minimize power consumption lead to high implementation complexity [5], [6].

Nonlinear MIMO systems offer a way to reduce both complexity and power consumption of MIMO systems, by employing nonlinear detection methods, i.e. amplitude and phase detection. These are known as low-cost approaches to receiver design, since they can be implemented with much simpler and more energy-efficient circuitry. In fact, phase detectors are not prone to system nonlinearities, since amplitude carries no information, while amplitude detection can be simply implemented using an envelope detector, thus avoiding the power consuming down-conversion. We refer to conventional MIMO systems as *linear* MIMO systems, meaning that an I/Q demodulator yields a complex baseband received signal. Nonlinear MIMO systems are especially lucrative for systems that require high numbers of very low-cost and low-power devices, e.g., sensor networks. In general, nonlinear MIMO systems are relevant in a future environment where wireless access is enabled by an ubiquitous network of sensors and nodes.

In [1] we presented achievable rates of nonlinear MIMO systems. It was shown that $N \times N$ nonlinear MIMO systems exhibit half the spatial multiplexing of linear MIMO systems, while additional receive antennas further increase the spatial multiplexing gain. Furthermore, nonlinear MIMO systems exploit the full MIMO diversity. However, our study of achievable rates was based on the assumption that the receiver has perfect channel state information (CSI). Such an assumption is very optimistic in real systems, and especially for nonlinear MIMO systems, where legacy channel estimation techniques are not applicable. In this paper we extend the work presented in [1] with the assumption that channel estimation at the receiver is not perfect. We use the popular MMSE channel estimation model employed in [7] and [8], and quantify the rate degradation caused by channel estimation errors.

In Section II, we present the nonlinear MIMO system model and revise the case of perfect channel state information at the receiver. Section III, introduces the noisy channel state information system model and presents upper and lower bounds for the capacity of the linear MIMO reference system. Section IV describes the computation of the achievable rates of nonlinear MIMO systems with noisy channel state information and the simulation results in Section V show that nonlinear techniques behave similar to channel estimation errors as linear detection.

Throughout the paper, bold lower and upper case letters stand for random vectors and matrices, respectively. Realizations of random variables are denoted with upright letters. $\mathcal{E}[\bullet]$, $\mathrm{I}(\bullet)$, $\mathrm{h}(\bullet)$ and $\mathrm{f}(\bullet)$ denote expectation, mutual information, differential entropy and probability density function, respectively. $\mathbf{A}_{ij}$ is the (i, j)th element of matrix $\mathbf{A}$, a_i is the ith element of the vector $\mathbf{a}$ and $(\bullet)^{\mathrm{H}}$ denotes complex-conjugate transposition. $\boldsymbol{e}_i$ is the ith column of the $N \times N$ identity matrix $\boldsymbol{I}_N$.

II. Perfect Channel State Information

The nonlinear MIMO system model with perfect CSI was presented in [1]. The transmitter employs N_{T} transmit and the receiver N_{R} receive antennas. The system model is depicted in Fig. 1. The transmitter sends the signal vector $\boldsymbol{s} \in \mathbb{C}^{N_{\mathrm{T}}}$ over the memoryless flat fading channel $\boldsymbol{H} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$, with tap-gain $\boldsymbol{H}_{ij}$ from the jth transmit to the ith receive antenna. The received vector $\boldsymbol{x} \in \mathbb{C}^{N_{\mathrm{R}}}$ is perturbed by a zero-mean circularly symmetric Gaussian noise vector $\boldsymbol{w} \in \mathbb{C}^{N_{\mathrm{R}}}$, with autocorrelation function $\mathcal{E}[\boldsymbol{w}_k \boldsymbol{w}_l^{\mathrm{H}}] = \sigma_w^2 \boldsymbol{I}_{N_{\mathrm{R}}} \delta(k-l)$, where k and l are time instants. The nonlinear detector operates

978-1-4244-2571-6/08/$25.00 ©2008 IEEE

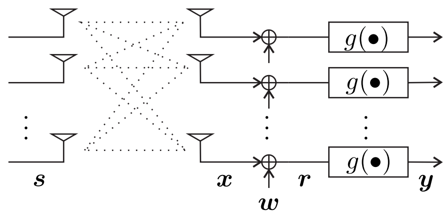

Fig. 1. MIMO System Reference Model

element wise on the perturbed vector $\boldsymbol{r}$ and extracts either the amplitude or the phase of the complex signal,

$$y_{i,\text{ampl.}} = g_{\text{ampl.}}(r_i) = \sqrt{\Re\{r_i\}^2 + \Im\{r_i\}^2}, \quad (1)$$

$$y_{i,\text{phase}} = g_{\text{phase}}(r_i) = \arg(r_i) = \tan^{-1}\left(\frac{\Im\{r_i\}}{\Re\{r_i\}}\right), \quad (2)$$

$i = 1, \ldots, N_{\mathrm{R}}$. The real observed vector $\boldsymbol{y} = g(\boldsymbol{r}) \in \mathbb{R}^{N_{\mathrm{R}}}$, $g : \mathbb{C}^{N_{\mathrm{R}}} \mapsto \mathbb{R}^{N_{\mathrm{R}}}$, misses one dimension of the two-dimensional (complex) transmitted signal, and the input-output relation is given by

$$\boldsymbol{y} = g(\boldsymbol{r}) = g(\boldsymbol{x} + \boldsymbol{w}) = g(\boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}). \quad (3)$$

The receiver has perfect knowledge of the channel matrix realization $\boldsymbol{H}$ at any time.

The achievable rates of this system where computed in [1], for a Gaussian input distribution. The rate[1]

$$\mathrm{I}(\boldsymbol{s}; \boldsymbol{y}, \boldsymbol{H}) = \mathcal{E}_{\boldsymbol{H}}[\mathrm{I}(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H} = \mathbf{H})]. \quad (4)$$

was numerically evaluated, using Monte Carlo integration (cf. Sec. IV) and the fact that the conditional probabilities $\mathrm{f}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N_{\mathrm{R}}} \mathrm{f}(y_i|x_i)$ are known, both for amplitude and phase detection:

$$\mathrm{f}_{\text{ampl.}}(y_i|x_i) = \frac{2y_i}{\sigma_w^2} e^{-\frac{y_i^2 + |x_i|^2}{\sigma_w^2}} I_0\left(\frac{2y_i|x_i|}{\sigma_w^2}\right), \quad (5)$$

$$\mathrm{f}_{\text{phase}}(\Delta\phi_i|x_i) = \frac{e^{-\rho_i}}{2\pi} + \sqrt{\frac{\rho_i}{4\pi}} e^{-\rho_i \sin^2 \Delta\phi_i} \cdot \cos\Delta\phi_i \operatorname{erfc}(-\sqrt{\rho_i}\cos\Delta\phi_i), \quad (6)$$

where $\rho_i = |x_i|^2/\sigma_w^2$ and $\Delta\phi_i = y_{i,\text{phase}} - \arg(x_i) \in [0, 2\pi)$ [9]. $I_0(\bullet)$ is the zeroth order modified Bessel function of the first kind and $\operatorname{erfc}(\bullet)$ is the complementary error function.

When $\boldsymbol{y} = \boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}$, we obtain the linear reference system. The capacity with a total transmit power constraint P and perfect receive CSI has been characterized in [2], [3]. For a stationary memoryless channel, the ergodic capacity is given by

$$C_{\text{lin}} = \mathcal{E}_{\boldsymbol{H}}\left[\log\det\left(\boldsymbol{I}_{N_{\mathrm{R}}} + \frac{\mathrm{SNR}}{N_{\mathrm{T}}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)\right], \quad (7)$$

and the capacity achieving input distribution for i.i.d. Rayleigh fading is circularly symmetric complex Gaussian, $\boldsymbol{s} \sim \mathcal{CN}(\mathbf{0}, \sigma_s^2 \boldsymbol{I}_{N_{\mathrm{T}}})$. The average SNR per receive antenna is defined as

$$\mathrm{SNR} = \frac{P}{\sigma_w^2} = \frac{N_{\mathrm{T}}\sigma_s^2}{\sigma_w^2}. \quad (8)$$

[1]We equivalently compute the mutual information between $\boldsymbol{x}$ and $\boldsymbol{y}$, since $\boldsymbol{s}, \boldsymbol{x}$ and $\boldsymbol{y}$ form a Markov chain (cf. [1]).

III. Noisy Channel State Information Reference System

We now assume the receiver performs minimum mean square error channel estimation of $\boldsymbol{H}$, which leads to a non-perfect estimate $\widehat{\boldsymbol{H}}$. We assume the entries of the fading channel $\boldsymbol{H}$ are i.i.d. zero-mean circularly symmetric complex Gaussian with variance σ_h^2. The estimation error $\boldsymbol{E}$

$$\boldsymbol{H} = \widehat{\boldsymbol{H}} + \boldsymbol{E} \quad (9)$$

is uncorrelated with the estimate $\widehat{\boldsymbol{H}}$, following a property of MMSE estimation. The estimation error variance $\sigma_e^2 = \mathcal{E}[|\boldsymbol{H}_{i,j}|^2] - \mathcal{E}[|\widehat{\boldsymbol{H}}_{i,j}|^2]$ is a measure of the quality of the estimation. The entries of $\widehat{\boldsymbol{H}}$ are also i.i.d. zero-mean circularly symmetric complex Gaussian with variance $\sigma_h^2 - \sigma_e^2$.

The capacity of the linear MIMO system with noisy channel estimation is not known. However, there exist lower and upper bounds, which will serve as a reference for comparison with the nonlinear MIMO achievable rates. We assume the input distribution is white Gaussian, and only consider bounds on the mutual information. A lower bound on the mutual information is computed choosing a Gaussian input distribution and the entropy of a Gaussian random variable, whose variance is given by the mean square error of the MMSE estimate of $\boldsymbol{s}$ given $\boldsymbol{y}$ and $\widehat{\boldsymbol{H}}$ [8]

$$\mathrm{I}_{\text{lower}}(\boldsymbol{s}; \boldsymbol{y}|\widehat{\boldsymbol{H}}) = \mathcal{E}_{\widehat{\boldsymbol{H}}}\left[\log_2\left|\boldsymbol{I}_{N_{\mathrm{R}}} + \frac{\mathrm{SNR}}{1 + \sigma_e^2\mathrm{SNR}}\frac{1}{N_{\mathrm{T}}}\widehat{\boldsymbol{H}}\widehat{\boldsymbol{H}}^{\mathrm{H}}\right|\right], \quad (10)$$

where we used that $\mathcal{E}[\boldsymbol{s}^{\mathrm{H}}\boldsymbol{s}] = \frac{P}{N_{\mathrm{T}}}\boldsymbol{I}$ and the SNR definition from (8). An upper bound on the mutual information is computed using the fact that the Gaussian distribution maximizes entropy. It can be written as a function of the lower bound

$$\mathrm{I}_{\text{upper}}(\boldsymbol{s}; \boldsymbol{y}|\widehat{\boldsymbol{H}}) = \mathrm{I}_{\text{lower}}(\boldsymbol{s}; \boldsymbol{y}|\widehat{\boldsymbol{H}}) + N_{\mathrm{R}} \cdot \mathcal{E}_{\boldsymbol{s}}\left[\log_2 \frac{1 + \sigma_e^2\mathrm{SNR}}{1 + \sigma_e^2\|\boldsymbol{s}\|^2}\right]. \quad (11)$$

IV. Nonlinear MIMO Systems with Noisy CSI

Our goal is to compute the mutual information between $\boldsymbol{s}$ and $\boldsymbol{y}$ given a channel estimate $\widehat{\boldsymbol{H}}$ when the observation is a nonlinear function of $\boldsymbol{r}$:

$$\mathrm{I}(\boldsymbol{s}; \boldsymbol{y}, \widehat{\boldsymbol{H}}) = \mathcal{E}_{\widehat{\boldsymbol{H}}}[\mathrm{I}(\boldsymbol{s}; \boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})]. \quad (12)$$

Unlike in [1], we can not use the mutual information between $\boldsymbol{x}$ and $\boldsymbol{y}$, since we do not have perfect knowledge of $\boldsymbol{H}$. The steps we follow are however similar, the only difference being that they include the effect of the noisy channel estimate on $\mathrm{h}(\boldsymbol{y}|\widehat{\boldsymbol{H}})$ and $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}})$.

From Eqs. (3) and (9), the input output relation is written as

$$\boldsymbol{y} = g(\widehat{\boldsymbol{H}}\boldsymbol{s} + \boldsymbol{E}\boldsymbol{s} + \boldsymbol{w}) \quad (13)$$

For the computation of $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}})$, we write (13) as

$$\boldsymbol{y} = g(\widehat{\boldsymbol{H}}\boldsymbol{s} + \boldsymbol{w}') = g(\boldsymbol{z}), \quad (14)$$

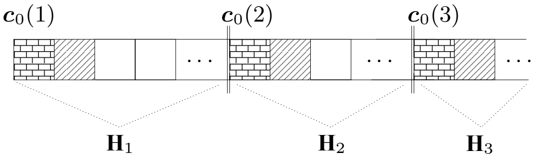

Fig. 2. Coding Scheme

where $\boldsymbol{w}' = \boldsymbol{w} + \boldsymbol{E}\boldsymbol{s}$. Here we assume that $\boldsymbol{w}$ and $\boldsymbol{E}$ are statistically independent. To ensure this, we use the coding scheme depicted in Fig. 2, which depicts a block fading channel. The channel remains unchanged for a certain period, allowing for appropriate estimation. The estimation yields an estimate $\widehat{\mathbf{H}}_i$ and an error $\mathbf{E}_i$ for every channel realization $\mathbf{H}_i$. $\mathbf{E}_i$ remains constant during a block. However, every codeword, eg. $\boldsymbol{c}_0$ in Fig. 2, is spread over many channel realizations/estimations, such that $\boldsymbol{w}$ and $\boldsymbol{E}$ in Eq. (13) remain independent. Hence, for a given realization of $\boldsymbol{s} = \mathbf{s}$, $\boldsymbol{w}'$ is Gaussian distributed, as

$$\boldsymbol{w}' \sim \mathcal{CN}\left(0, \left(\sigma_w^2 + \sigma_e^2 \sum_{j=1}^{N_\text{T}} |s_j|^2\right)\boldsymbol{I}\right). \tag{15}$$

Furthermore, for a given realization of $\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}$, $\boldsymbol{z}$ is also Gaussian distributed, as

$$\boldsymbol{z} \sim \mathcal{CN}\left(\widehat{\mathbf{H}}\mathbf{s}, \left(\sigma_w^2 + \sigma_e^2 \sum_{j=1}^{N_\text{T}} |s_j|^2\right)\boldsymbol{I}\right), \tag{16}$$

and the distribution $\mathrm{f}(y_i|\boldsymbol{s} = \mathbf{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ is given by Eqs. (5) and (6), evaluated by setting $x_i \leftarrow \boldsymbol{e}_i^\text{T}\widehat{\mathbf{H}}\mathbf{s}$ and $\sigma_w^2 \leftarrow \sigma_w^2 + \sigma_e^2 \sum_{j=1}^{N_\text{T}} |s_j|^2$ in those equations. Since the composite noise component $\boldsymbol{w}'$ remains white, i.e., its correlation matrix is diagonal, the distribution of $\boldsymbol{y}$ can be split as

$$\mathrm{f}(\boldsymbol{y}|\boldsymbol{s} = \mathbf{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) = \prod_{i=1}^{N_\text{R}} \mathrm{f}(y_i|\boldsymbol{e}_i^\text{T}\widehat{\mathbf{H}}\mathbf{s}), \tag{17}$$

Thus, we can compute $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}})$ using the same distribution as for the perfect channel estimation case. The estimation error is captured in the variance of the composite noise $\boldsymbol{w}'$ through the estimation error variance σ_e^2, and is a function of the instantaneous signal realization $\mathbf{s}$. Finally, the Monte Carlo estimate of the entropy $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ for a given channel estimate is computed as

$$\begin{aligned}
&\mathrm{h}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) = \\
&- \int \mathrm{f}(\boldsymbol{s}, \boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) \log(\mathrm{f}(\boldsymbol{y}|\boldsymbol{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})) \mathrm{d}\boldsymbol{s}\mathrm{d}\boldsymbol{y} \\
&\simeq -\frac{1}{N} \sum_{i=1}^{N} \log(\mathrm{f}(\boldsymbol{y} = \mathbf{y}_i|\boldsymbol{s} = \mathbf{s}_i, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})),
\end{aligned}$$

where $(\boldsymbol{s}_i, \boldsymbol{y}_i)$ are N sample tuples generated according to the joint distribution $\mathrm{f}(\boldsymbol{s}, \boldsymbol{y})$ for a given $\widehat{\mathbf{H}}$. We generate these samples by randomly creating N realizations of $\boldsymbol{s}$, and compute $\boldsymbol{y}$ by varying the composite noise $\boldsymbol{w}'$. Note that the only source of uncertainty is the noise term $\boldsymbol{w}'$, which includes the estimation error uncertainty.

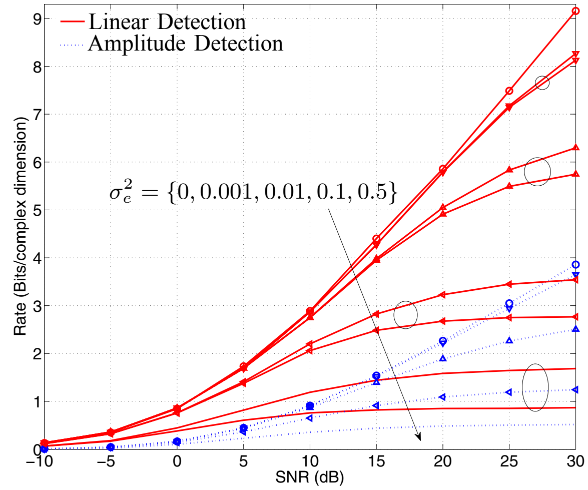

Fig. 3. SISO system — Amplitude Detection.

The distribution of $\boldsymbol{y}$ given the perfect channel estimate was given in analytical (or semi-analytical) form for various cases, like SISO, $N_\text{R} = 2, 3$. This is not the case with noisy channel estimation. The distribution of $\boldsymbol{y}$ is now a nonlinear function of the sum of a Gaussian random variable ($\boldsymbol{w}$) and a term that consists of the sum of products of Gaussian random variables ($\boldsymbol{E}\boldsymbol{s}$), and is not known in analytical form. Hence, we resort to Monte Carlo estimation of the distribution $\mathrm{f}(\boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ for specific values of $\boldsymbol{y} = \mathbf{y}_i$, which are generated from the distribution itself. We can create such samples simply by passing Gaussian input symbols through the system of Fig. 1 for a given channel estimate $\widehat{\mathbf{H}}$, and letting the estimation error $\boldsymbol{E}$ and the noise $\boldsymbol{w}$ vary. The entropy $\mathrm{h}(\boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ is computed with Monte Carlo integration

$$\mathrm{h}(\boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) \simeq -\frac{1}{N} \sum_{i=1}^{N} \log(\mathrm{f}(\mathbf{y}_i|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})). \tag{18}$$

Finally, the values $\mathrm{f}(\mathbf{y}_i|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ themselves are computed again with Monte Carlo integration as follows

$$\begin{aligned}
\mathrm{f}(\mathbf{y}_i|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) &= \int \mathrm{f}(\boldsymbol{s})\, \mathrm{f}(\mathbf{y}_i|\boldsymbol{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}) \mathrm{d}\boldsymbol{s} \\
&\simeq \frac{1}{M} \sum_{j=1}^{M} \mathrm{f}(\mathbf{y}_i|\boldsymbol{s} = \mathbf{s}_j, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}).
\end{aligned} \tag{19}$$

Here, $\mathbf{s}_j$ are M samples of $\boldsymbol{s}$, normally distributed and generated independent of $\mathbf{y}_i$.

V. Simulation Results

In the following plots, solid, dotted and dashed lines stand for linear, amplitude and phase detection, respectively. Curves marked with a circle refer to the case of perfect channel estimation ($\sigma_e^2 = 0$). For linear detection, we plot the lower and upper bounds given in Eqs. (10) and (11), respectively. The upper and lower bound pairs corresponding to the same estimation error variance are encircled. The input distribution is circularly symmetric complex Gaussian. The

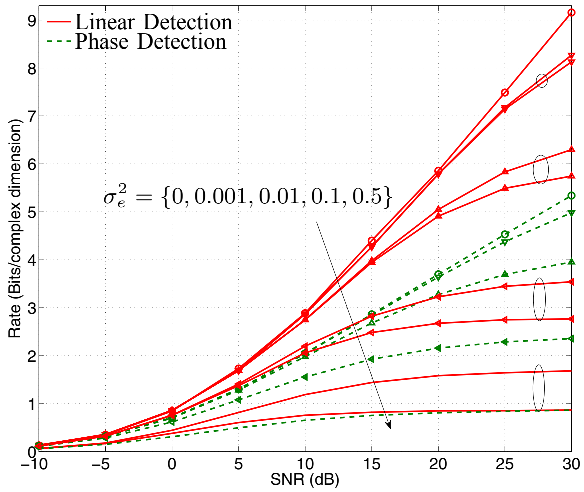

Fig. 4. SISO system — Phase Detection.

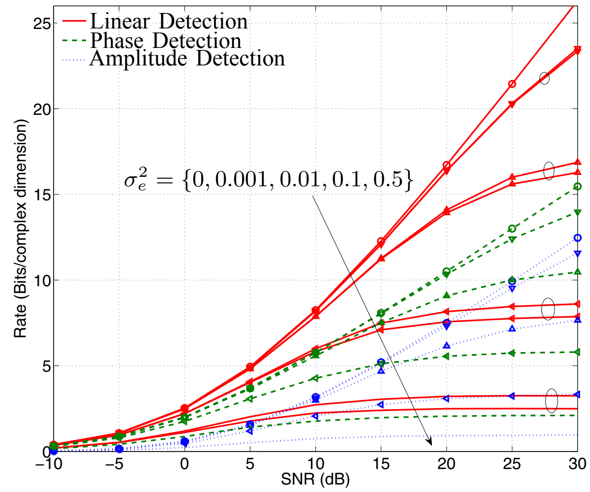

Fig. 6. MIMO system, $N_{\mathrm{T}} = N_{\mathrm{R}} = 3$.

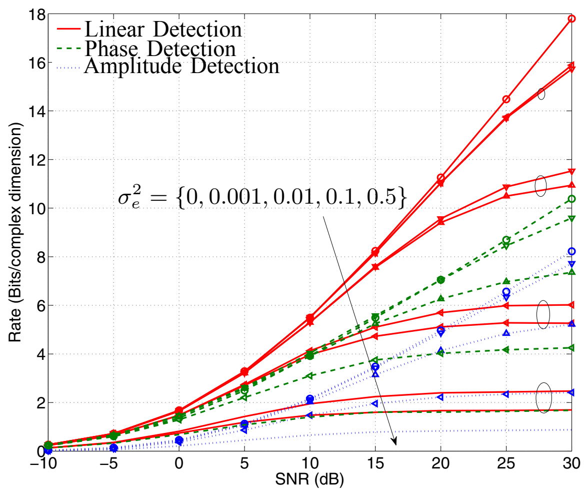

Fig. 5. MIMO system, $N_{\mathrm{T}} = N_{\mathrm{R}} = 2$.

SNR is defined in (8) and the channel matrix has i.i.d. zero-mean unit-variance ($\sigma_h^2 = 1$) circularly symmetric complex Gaussian entries. The values of the estimation error variance are $\sigma_e^2 = \{0.5, 0.1, 0.01, 0.001\}$. Curves with the same marker correspond to the same σ_e^2 value. In our simulations, the estimation error variance is kept constant for the whole SNR range. This assumption is motivated by scenarios where the channel estimate deterioration is due to outdated use of the channel estimate, i.e., σ_e^2 depends only on the speed the channel changes between channel estimations.

Figs. 3 and 4 depict the rates for a SISO system for amplitude and phase detection, respectively. Similar to the linear reference system, the achievable rates reach a threshold for high SNR. This threshold is reached sooner for large estimation error variances. The rate loss at $\sigma_e^2 = 0.001$, which implies a very good channel estimate, is very low. Phase detection remains always better than amplitude detection for the same σ_e^2, similar to the perfect channel knowledge case. This result holds throughout the following antenna configurations, pointing out that both techniques are equally robust to channel estimation errors. Figs. 5 and 6 depict the achievable rates for $N_{\mathrm{T}} = N_{\mathrm{R}} = 2$ and $N_{\mathrm{T}} = N_{\mathrm{R}} = 3$ MIMO systems, respectively. The behavior is similar to the SISO case. Phase detection performed well with respect to linear detection at low SNR for perfect CSI. Now we can see that phase detection can even outperform linear detection, when the channel estimation quality of the former is somewhat better than that of the latter, e.g., $\sigma_e^2 = 0.01$ for phase detection and $\sigma_e^2 = 0.1$ for linear detection. Throughout the various plots, amplitude detection exhibits the lowest rate loss of all techniques at $\sigma_e^2 = 0.001$ with respect to the perfect CSI case. However, this advantage diminishes at lower channel estimation variances. Concluding, we can say that both amplitude and phase detection exhibit the same bahavior as linear detection when the channel estimation is not perfect.

References

[1] G. K. Psaltopoulos, F. Trösch, and A. Wittneben, "On Achievable Rates of MIMO Systems with Nonlinear Receivers," in *IEEE International Symposium on Information Theory*, June 2007, pp. 1071–1075.

[2] I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," in *European Transactions on Telecommunications*, November 1999, pp. 585–595.

[3] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," in *Wireless Personal Communications*, vol. 6, no. 3, March 1998, pp. 311–355.

[4] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and Cooperative MIMO Techniques in Sensor Networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1089–1098, August 2004.

[5] S. K. Jayaweera, "Energy analysis of MIMO techniques in wireless sensor networks," in *38th Annual Conf. on Information Science and Systems*, Princeton, NJ, March 2004.

[6] H. Dai, L. Xiao, and Q. Zhou, "Energy Efficiency of MIMO Transmission Strategies in Wireless Networks," in *International Conference on Computing, Communications and Control Technologies*, Austin, TX, August 2004.

[7] M. Medard, "The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel," *IEEE Transactions on Information Theory*, vol. 46, pp. 933–946, May 2000.

[8] T. Yo and A. Goldsmith, "Capacity and Power Allocation for Fading MIMO Channels With Channel Estimation Error," in *IEEE Transactions on Information Theory*, vol. 52, no. 5, May 2006, pp. 2203–2214.

[9] W. R. Bennett, "Methods of solving noise problems," in *Proceedings of the IRE*, vol. 44, May 1956, pp. 609–638.