Nonlinear MIMO: Affordable MIMO Technology for Wireless Sensor Networks

Georgios K. Psaltopoulos and Armin Wittneben

***Abstract*—We consider a sensor network, where an access point (AP) communicates with many sensor nodes (SN), which are simple, cheap, low-complexity and low-power communication nodes. Such systems typically use nonlinear modulation and detection, due to their low power consumption. Increasing their performance by means of multiple antennas at the AP and the SNs has not been considered, since this would violate the stringent power and cost constraints at the SN. We consider SNs with MIMO receivers that perform a nonlinear operation on the complex-valued received signal (amplitude or phase detection). These receivers enjoy the low-cost, low-power, low-complexity characteristics that are crucial for a sensor network. Such *nonlinear MIMO* systems are first introduced and studied here. They bring the high-rate, high-performance world of MIMO systems and the low-cost, low-complexity world of sensor networks together. We only consider the single-user MIMO system between the AP and one SN, and study the fundamental limits of such systems. We compute achievable rates under perfect and noisy CSI at the SN, and observe that these systems also achieve spatial multiplexing gain, albeit different to legacy linear MIMO systems. We quantify and analyze these gains using numerical means, and give insight into the effect of the nonlinearity on the information theoretic limits of nonlinear MIMO systems.**

***Index Terms*—Wireless sensor networks, achievable rates, spatial multiplexing, robustness, nonlinear MIMO, nonlinear receivers, perfect CSIR, noisy CSIR.**

I. Introduction

WIRELESS sensor networks have been a hot research topic for over a decade. Such networks, composed of simple sensing nodes, densely deployed, can have a multitude of applications, spreading from health, home to military. Research on sensor networks has been focused on all communication levels, from routing protocols down to the physical layer. Design and optimization of sensor nodes is always connected to stringent power and complexity constraints which must be fulfilled. Since the number of nodes can be, and in many applications must be, very high, the extreme low cost of these devices is of paramount importance. Low cost implies limited complexity in terms of processing power and communications technology. On the other side, the nodes are mostly powered by batteries, in which case the power-efficiency defines the lifetime of the sensor. Hence, low power consumption is also a design target [1].

In order to assure low power consumption and modest complexity, nonlinear modulation techniques are very popular in realistic sensor node design. Modulations like On-Off Keying (OOK) [2] and Frequency-Shigt Keying (FSK) [3]–[5] both require an envelope detector which detects the presence of the signal by measuring energy. Both techniques lead to significant power savings in transmitter as well as receiver design by avoiding power-demanding components, like mixers, linear amplifiers, etc., which are required in an In-Phase/Quadrature-Phase (I/Q) architecture [4]. Another important aspect in power savings is the short start-up time and the associated power during this step (cf. [4]). This is very important, since wireless sensors usually operate in a low-duty cycle mode, enabling power savings while adapting to their sporadic communication needs. Hence, it is evident that the typical linear I/Q type of transceiver design is not suitable for sensor networks.

These characteristics carry over to Multiple Input Multiple Output (MIMO) systems. Such systems are well known for providing high data rates through exploiting the spatial dimension of the medium [6], [7]. 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, combining several I/Q branches, i.e. multiple low-noise amplifiers, filters, mixers, A/D converters, for every Tx/Rx antenna in a MIMO system, leads to very high power consumption [8], in fact orders of magnitude higher than that of efficient sensor nodes. Various techniques have been considered in literature to combat this shortcoming by enabling MIMO techniques virtually (virtual MIMO) through node cooperation [8]–[10]. The central aim of all these schemes is to achieve higher rates and thus reduce the transmit duration, using e.g. distributed space time codes. However these schemes require node cooperation in order to disseminate information locally, and synchronization in order to transmit/receive information jointly. Furthermore, continuous rate adaptation is used in order to minimize the transmit power. These assumptions lead to a complicated network structure with node cooperation, synchronization and complicated radios that allow rate adaptation. Hence, increased power consumption and high complexity hinder employment of MIMO or virtual MIMO systems in low-power short-range applications, like sensor networks.

Using multiple antennas in a sensor node is not straightforward applicable. However, combining the nonlinear techniques used in sensor nodes with multiple antennas can merge the benefits of both paradigms: low power consumption due to nonlinear modulation and detection, and higher rates due to spatial multiplexing using multiple antennas. We expect such a receiver to be low-cost and low-power, since employing

Manuscript received April 15, 2009; revised August 25, 2009; accepted November 21, 2009. The associate editor coordinating the review of this paper and approving it for publication was F. Takawira.

The authors are at the Communication Technology Laboratory, Swiss Federal Institute of Technology (ETH) Zurich, CH-8092 Zurich, Switzerland (e-mail: {psaltopoulos, wittneben}@nari.ee.ethz.ch).

Digital Object Identifier 10.1109/TWC.2010.02.090530

1536-1276/10$25.00 © 2010 IEEE

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

several nonlinear receiver chains adds minimal to cost and power consumption, remaining still below the budget of a single linear receiver chain. Furthermore, hosting multiple antennas on small sensor nodes is possible through advances in compact MIMO-antenna design, where diversity is harvested by various means, like polarization and antenna patterns, e.g. [11]–[13]. The resulting designs enable MIMO techniques on portable/handheld devices. Hence, a sensor node with multiple antennas that uses a nonlinear receiver can use the potential higher performance of MIMO systems by sacrificing power and complexity only modestly.

In this paper we consider this novel combination of MIMO systems with nonlinear detection. Such a combination is new: the properties, behavior and fundamental limits of *nonlinear MIMO systems* are not known. We study the theoretical limits of such a system and try to understand its similarities and differences to linear MIMO systems, which are well studied in literature. For our purpose, we consider a point-to-point link between an AP, with unlimited resources, and a simple sensor which employs a nonlinear MIMO receiver. We assume that such a receiver has access only to the amplitude (like an envelope detector) *or* to the phase of the complex-valued received signal on each antenna and allow for generic modulation techniques. Our contribution is summarized in the following points[1]:

- We provide achievable rates of nonlinear MIMO systems with either perfect or noisy channel knowledge at the receiver. We also numerically compute the rate of linear MIMO systems with noisy CSI and Gaussian input.
- We elaborate on how nonlinear MIMO systems exploit the spatial degrees of freedom and comment on the fundamental differences to linear MIMO systems.
- We explore the robustness of such systems with respect to channel estimation errors.
- We discuss why sensor networks can benefit from such a type of receiver design.

The paper is structured as follows. In Section II, we present a generic MIMO system model and discuss the differences between the linear reference model and our abstract nonlinear MIMO model. In Section III, we handle the perfect channel state information case, both for the linear reference model and for nonlinear MIMO. Section IV follows along the same lines with noisy channel state information. Simulation results are presented in Section V and the gains of nonlinear MIMO are discussed. Finally, Section VI discusses the added value of using nonlinear MIMO in a wireless sensor network.

Notation: Throughout the paper, bold-faced italic lower and upper case letters stand for random vectors and matrices, respectively, while their realizations are denoted with upright letters. $\mathcal{E}[\cdot]$, $\mathrm{I}(\cdot)$, $\mathrm{h}(\cdot)$ and $\mathrm{f}(\cdot)$ denote expectation, mutual information, differential entropy and probability density function, respectively. B_{ij} is the (i, j)th element of matrix $\mathbf{B}$, $|\mathbf{B}|$ denotes the determinant, b_i is the ith element of the vector $\mathbf{b}$, while $\mathbf{b}_i$ is a realization of the random vector $\boldsymbol{b}$, indexed with i. $(\cdot)^{\mathrm{H}}$ denotes complex-conjugate transposition. $\boldsymbol{e}_i$ is the ith column of the $N \times N$ identity matrix $\boldsymbol{I}_N$. The circularly symmetric complex Gaussian distributed random variable $\boldsymbol{x}$ with mean $\mathbf{m}$ and variance $\mathbf{R}$ is denoted as $\boldsymbol{x} \sim \mathcal{CN}_{\boldsymbol{x}}(\mathbf{m}, \mathbf{R})$. All logarithms are computed with base 2.

[1]Parts of the contribution have been presented in [14] and [15].

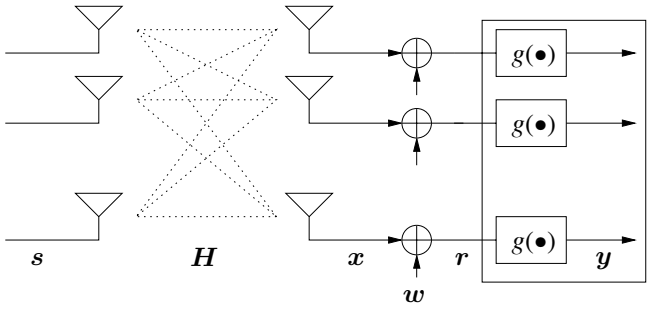

Fig. 1. MIMO system reference model.

II. System Model

We consider the generic MIMO system depicted in Fig. 1, with N_{T} transmit and N_{R} receive antennas. The transmitter emits the signal $\boldsymbol{s} \in \mathbb{C}^{N_{\mathrm{T}}}$ over the stationary memoryless flat fading channel $\boldsymbol{H} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$, with tap-gain H_{ij} from the j-th transmit to the i-th receive antenna. We use a block fading model for the channel $\boldsymbol{H}$. The channel remains constant for some (coherence) period, long enough to allow for accurate estimation, and changes to an independent realization in the next block. This can be achieved with a sufficiently long interleaver. 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 symbol instants. A function $g(\cdot)$ acts on every antenna, producing the output vector $\boldsymbol{y} = g(\boldsymbol{r}) = g(\boldsymbol{x} + \boldsymbol{w})$. The detector has access only to $\boldsymbol{y}$. Finally, we assume that we can code over many channel realizations, and hence we compute the ergodic capacity/achievable rates.

a) Linear MIMO: In legacy linear MIMO systems, the following simple linear input-output relationship is used:

$$\boldsymbol{y} = \boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}, \tag{1}$$

that is, the function $g(\cdot)$ does not modify the received signal. The linearity of (1) enables easy mathematical manipulation and analysis. However, this simple system model neglects several sources of imperfection which exist in real communication systems, such as I/Q imbalance, amplifier nonlinearities, phase noise and other RF imperfections [16].

b) Nonlinear MIMO: The function $g : \mathbb{C}^{N_{\mathrm{R}}} \mapsto \mathbb{R}^{N_{\mathrm{R}}}$ maps the complex-valued received signal $\boldsymbol{r}$ from the complex plane into one real dimension, element-wise on every antenna. This nonlinear processing allows for a very simplified receiver structure, which consumes significantly less power compared to a linear MIMO I/Q receiver. In particular, we consider two types of nonlinear functions known from literature and motivated by their efficient implementation: an amplitude and a phase function. These are described as follows:

$$y_{i,\mathrm{ampl.}} = g_{\mathrm{ampl.}}(r_i) = |r_i| = \sqrt{\Re\{r_i\}^2 + \Im\{r_i\}^2}, \tag{2}$$

$$y_{i,\mathrm{phase}} = g_{\mathrm{phase}}(r_i) = \angle(r_i) = \tan^{-1}\left(\frac{\Im\{r_i\}}{\Re\{r_i\}}\right), \tag{3}$$

for $i = 1, \ldots, N_{\mathrm{R}}$, for amplitude and phase detection, respectively. The inverse tangent in (3) considers the quadrant where

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

r_i lies in. In both cases, the main characteristic is that the signal $\boldsymbol{y}$ is missing half the dimensions after processing with the function g.

III. Achievable Rates with Perfect Channel State Information

First, we consider the case when the receiver has perfect knowledge of the channel realization. Although not realistic, this assumption is important as a first step for understanding nonlinear MIMO, and in order to compare with linear MIMO systems, where the case of perfect receive CSI is well understood.

A. Linear Reference System

The capacity of linear MIMO systems with a total transmit power constraint P and perfect receive CSI has been characterized in [6], [7]. For a stationary memoryless channel, the ergodic capacity is given by

$$C_{\text{lin}} = \mathcal{E}_{\boldsymbol{H}}\left[\log\det\left(\boldsymbol{I}_{N_{\text{R}}} + \frac{\text{SNR}}{N_{\text{T}}}\boldsymbol{H}\boldsymbol{H}^{\text{H}}\right)\right], \tag{4}$$

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

$$\text{SNR} = \frac{P}{\sigma_w^2} = \frac{N_{\text{T}}\sigma_s^2}{\sigma_w^2}, \tag{5}$$

with $\sigma_h^2 = 1$. In the high SNR regime, the capacity is approximately given by

$$C_{\text{lin}} \simeq N_{\min}\log\frac{\text{SNR}}{N_{\text{T}}} + \sum_{i=1}^{N_{\min}}\mathcal{E}\left[\log\lambda_i^2\right], \tag{6}$$

where $N_{\min} = \min(N_{\text{T}}, N_{\text{R}})$ and λ_i's are the singular values of $\boldsymbol{H}$. Linear MIMO systems exploit $N_{\min}$ degrees of freedom, or else, achieve $N_{\min}$ spatial multiplexing, meaning that capacity increases $N_{\min}$ bits for every 3 dB SNR increase in the high SNR regime [17].

B. Nonlinear MIMO

1) Rate Expressions: Since we assumed that the channel $\boldsymbol{H}$ is memoryless, the capacity is given by the maximum mutual information between the channel input $\boldsymbol{s}$ and the detector output $\boldsymbol{y}$. Furthermore, since $\boldsymbol{H} = \mathbf{H}$ is known to the receiver, the channel output is the pair $(\boldsymbol{y}, \boldsymbol{H})$ (cf. [6]). The mutual information can be written as

$$\text{I}(\boldsymbol{s}; \boldsymbol{y}, \boldsymbol{H}) \;=\; \mathcal{E}_{\boldsymbol{H}}\left[\text{I}(\boldsymbol{s}; \boldsymbol{y}|\boldsymbol{H} = \mathbf{H})\right], \tag{7}$$

since $\boldsymbol{s}$ and $\boldsymbol{H}$ are statistically independent. Given a channel realization $\boldsymbol{H} = \mathbf{H}$, we observe that $\boldsymbol{s}$, $\boldsymbol{x}$ and $\boldsymbol{y}$ form a Markov chain $\boldsymbol{s} \to \boldsymbol{x} \to \boldsymbol{y}$, since $\boldsymbol{s}$ and $\boldsymbol{y}$ are conditionally independent given $\boldsymbol{x} = \mathbf{H}\boldsymbol{s}$ [14]. This implies that

$$\text{I}(\boldsymbol{s}; \boldsymbol{y}|\boldsymbol{H}) = \mathcal{E}_{\boldsymbol{H}}[\text{I}(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H} = \mathbf{H})]. \tag{8}$$

Although (8) is straightforward for a bijective relationship between $\boldsymbol{s}$ and $\boldsymbol{x}$, it also holds when $\boldsymbol{H}$ is rank-deficient, and hence not invertible.

In the following, we evaluate the conditional mutual information from $\boldsymbol{x}$ to $\boldsymbol{y}$. For notational simplicity we will write $\text{I}(\boldsymbol{x}; \boldsymbol{y}|\mathbf{H})$ instead of $\text{I}(\boldsymbol{x}; \boldsymbol{y}|\boldsymbol{H} = \mathbf{H})$ in this subsection. The mutual information of interest $\text{I}(\boldsymbol{x}; \boldsymbol{y}|\mathbf{H})$ equals

$$\begin{aligned}\text{I}(\boldsymbol{x}; \boldsymbol{y}|\mathbf{H}) &= \text{h}(\boldsymbol{y}|\mathbf{H}) - \text{h}(\boldsymbol{y}|\boldsymbol{x}) \\ &= -\int \text{f}(\boldsymbol{y}|\mathbf{H})\log(\text{f}(\boldsymbol{y}|\mathbf{H}))\text{d}\boldsymbol{y} \\ &\quad + \iint \text{f}(\boldsymbol{x}, \boldsymbol{y})\log(\text{f}(\boldsymbol{y}|\boldsymbol{x}))\text{d}\boldsymbol{y}\text{d}\boldsymbol{x}. \end{aligned} \tag{9}$$

We used the fact that $\text{h}(\boldsymbol{y}|\boldsymbol{x}, \mathbf{H}) = \text{h}(\boldsymbol{y}|\boldsymbol{x})$. We need the distributions $\text{f}(\boldsymbol{y}|\mathbf{H})$ and $\text{f}(\boldsymbol{y}|\boldsymbol{x})$. The first distribution is known only in some cases. Otherwise, we estimate it numerically, as explained in the sequel. The second distribution is known analytically. Since the noise vector is i.i.d. and the nonlinear function $g(\cdot)$ acts on each antenna separately, the conditional distribution of $\boldsymbol{y}$ given $\boldsymbol{x}$ can be factorized as

$$\text{f}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N_{\text{R}}}\text{f}(y_i|x_i), \tag{10}$$

where y_i and x_i are the ith elements of the vectors $\boldsymbol{y}$ and $\boldsymbol{x}$, respectively. When $\boldsymbol{x}$ is known, $\boldsymbol{r}$ is complex Gaussian distributed like $\mathcal{CN}(\mathbf{x}, \sigma_w^2\boldsymbol{I}_{N_{\text{R}}})$. The amplitude $y_i = |r_i|$ is Ricean distributed [18]

$$\text{f}_{\text{ampl.}}(y_i|x_i) \;=\; \frac{2y_i}{\sigma_w^2}e^{-\frac{y_i^2+|x_i|^2}{\sigma_w^2}}\text{I}_0\left(\frac{2y_i|x_i|}{\sigma_w^2}\right), \tag{11}$$

while the distribution of the phase $y_i = \angle(r_i)$ is given by [19]

$$\begin{aligned}\text{f}_{\text{phase}}(\Delta\phi_i|x_i) \;=\; &\frac{e^{-\rho_i}}{\sigma_w^2} + \sqrt{\frac{\rho_i}{4\pi}}e^{-\rho_i\sin^2\Delta\phi_i}\cdot \\ &\cdot\cos\Delta\phi_i\,\text{erfc}\left(-\sqrt{\rho_i}\cos\Delta\phi_i\right), \end{aligned} \tag{12}$$

where $\rho_i = |x_i|^2/\sigma_w^2$ and $\Delta\phi_i = y_{i,\text{phase}} - \angle(x_i) \in [0, 2\pi)$. $\text{I}_0(\cdot)$ is the zeroth order modified Bessel function of the first kind and $\text{erfc}(\cdot)$ is the complementary error function.

In the following we discuss how to numerically compute achievable rates for fixed input distributions.

2) Numerical Computation: The integrals in (9) are computed using *Monte Carlo* (MC) integration [20]. MC methods are especially important for the computation of multi-dimensional integrals, i.e., more than 4 dimensions, where the traditional numerical methods become computationally prohibitive. Due to (10), (11) and (12), the distribution $\text{f}(\boldsymbol{y}|\boldsymbol{x})$ appearing in the second integral of (9) is known analytically for both amplitude and phase detection. To compute this integral we create N realizations of $\boldsymbol{x}$ and $\boldsymbol{y}$ pairs, $(\mathbf{x}_i, \mathbf{y}_i)$, $i = 1, \ldots, N$ which are distributed according to their joint distribution $\text{f}(\boldsymbol{x}, \boldsymbol{y})$. That is, $\mathbf{x}_i$ is drawn from a distribution $\text{f}(\boldsymbol{x})$ of our choice, e.g. by choosing a Gaussian input $\boldsymbol{s}$, and $\mathbf{y}_i$ is drawn from the distribution $\text{f}(\boldsymbol{y}|\mathbf{x}_i)$. Then, the MC estimate of the conditional entropy $\text{h}(\boldsymbol{y}|\boldsymbol{x})$ is given by

$$\text{h}(\boldsymbol{y}|\boldsymbol{x}) \simeq -\frac{1}{N}\sum_{i=1}^{N}\log\left(\text{f}(\mathbf{y}_i|\mathbf{x}_i)\right). \tag{13}$$

The distribution $\text{f}(\boldsymbol{y}|\mathbf{H})$ is necessary for the computation of $\text{h}(\boldsymbol{y}|\mathbf{H})$. However, $\text{f}(\boldsymbol{y}|\mathbf{H})$ is generally not known. For a Gaussian input alphabet $\boldsymbol{s}$ and a channel realization $\mathbf{H}$, $\boldsymbol{r} = \mathbf{H}\boldsymbol{s} + \boldsymbol{w}$ is Gaussian with $\boldsymbol{r}|\mathbf{H} \sim \mathcal{CN}_{\boldsymbol{r}}(\sigma_s^2\mathbf{H}\mathbf{H}^{\text{H}} + \sigma_w^2\boldsymbol{I})$. For amplitude

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

detection, $\mathrm{f}(\boldsymbol{y}|\mathbf{H}) = \mathrm{f}(g(\boldsymbol{r})|\mathbf{H})$ is a multivariate Rayleigh distribution, while for phase detection it is a multivariate distribution with uniformly distributed marginals. However, closed-form expressions exist only for specific values of N_R. For $N_\mathrm{R} = 1$, y is Rayleigh or uniformly distributed in $[0, 2\pi)$ distributed for amplitude detection and phase detection, respectively. The corresponding entropies are known. For $N_\mathrm{R} = 2$, the amplitude and phase distributions are given by [21], [22]

$$\mathrm{f}_{\mathrm{ampl.}}(\boldsymbol{y}|\mathbf{H}) = 4y_1y_2|\boldsymbol{\Phi}|e^{-(\Phi_{11}y_1^2+\Phi_{22}y_2^2)} \cdot \mathrm{I}_0(2|\Phi_{12}|y_1y_2), \quad (14)$$

$$\mathrm{f}_{\mathrm{phase}}(\boldsymbol{y}|\mathbf{H}) = \frac{|\boldsymbol{\Phi}|}{8\pi^2\Phi_{11}\Phi_{22}} \cdot \left[\frac{1}{1-\lambda^2} - \frac{\lambda\cos^{-1}\lambda}{(1-\lambda^2)^{3/2}}\right], \quad (15)$$

where $\boldsymbol{\Phi} = (\mathcal{E}[\boldsymbol{r}\boldsymbol{r}^\mathrm{H}])^{-1} = (\sigma_s^2\boldsymbol{H}\boldsymbol{H}^\mathrm{H} + \sigma_w^2\boldsymbol{I})^{-1}$,

$$\lambda = \frac{\Phi_{12}}{\sqrt{\Phi_{11}\Phi_{22}}} \cdot \cos(y_1 - y_2 - \chi_{12}), \quad (16)$$

and $\chi_{ij} = \angle(\Phi_{ij})$. An expression exists for the trivariate amplitude distribution (see [21], [23]). However, it consists of an infinite sum of alternating terms and proves to be numerically unstable. No closed-form expression exists for the trivariate phase distribution. Hence, for $N_\mathrm{R} \geq 3$, we will use MC integration to estimate the probability density of certain realizations $\mathbf{y}_i$. First, we create N realizations $\mathbf{y}_i$, distributed according to $\mathrm{f}(\boldsymbol{y}|\mathbf{H})$, for a given channel realization. We do this by passing samples of $\boldsymbol{s}$ through the system for a given $\mathbf{H}$. Depending on N_R, we evaluate the probability density from (14), (15), or numerically as follows:

$$\mathrm{f}(\mathbf{y}_i) = \int \mathrm{f}(\boldsymbol{x})\,\mathrm{f}(\mathbf{y}_i|\boldsymbol{x})\mathrm{d}\boldsymbol{x} \simeq \frac{1}{M}\sum_{j=1}^{M}\mathrm{f}(\mathbf{y}_i|\mathbf{x}_j), \quad (17)$$

where the samples $\mathbf{x}_j$ are generated according to the distribution $\mathrm{f}(\boldsymbol{x})$. Finally, the MC estimate $\mathrm{h}(\boldsymbol{y}|\mathbf{H})$ is given by

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

Finally, we average (13) and (18) over many channel realizations $\mathbf{H}$.

Eq. (11) is used in (13), (17) as well as in the next section. However, when the argument of the Bessel function in (11) is large, (11) becomes numerically unstable: the Bessel function grows to infinity and the exponential term converges to zero. This happens almost always when σ_w^2 is small, that is, at high SNR. We use the following approximation for (11) in these cases

$$\mathrm{f}_{\mathrm{ampl.}}(y_i|x_i) \simeq \sqrt{\frac{y_i}{\pi|x_i|\sigma_w^2}} \cdot e^{-\frac{(y_i-|x_i|)^2}{\sigma_w^2}}\left(1 - \frac{\sigma_w^2}{16y_i|x_i|^2} + \ldots\right) \quad (19)$$

which is computed using a high argument approximation of $\mathrm{I}_o(x)$ [24].

IV. Achievable Rates with Noisy Channel State Information

Next we consider the more realistic scenario, where channel knowledge is imperfect, as a result of practical channel estimation. A nonlinear MIMO sensor node can perform channel estimation, but the estimation model would presumably not be linear in that case. For the sake of simplicity we assume that the receiver has knowledge of a linear minimum mean square error (MMSE) estimate $\widehat{\boldsymbol{H}}$ of $\boldsymbol{H}$. A nonlinear MIMO sensor node could obtain such a channel estimate from the access point in a time division duplex system. Let us 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 $\boldsymbol{E}$

$$\boldsymbol{H} = \widehat{\boldsymbol{H}} + \boldsymbol{E} \quad (20)$$

is uncorrelated with the estimate $\widehat{\boldsymbol{H}}$, following a property of MMSE estimation. The estimation error variance $\sigma_e^2 = \mathcal{E}[|\boldsymbol{H}_{i,j}|^2] - \mathcal{E}[|\widehat{\boldsymbol{H}}_{i,j}|^2] \quad \forall i, j$, 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 $\sigma_h^2 - \sigma_e^2$.

We will consider two different scenarios in our simulation results. In one case, σ_e^2 is constant for all SNR. This corresponds to a scenario where the channel estimate deterioration is a result of an outdating, that is, the channel has changed from the last estimation. This type of degradation is independent of the operating SNR, and relates to the coherent time. The other scenario corresponds to a σ_e^2 that increases with SNR. This scenario relates the estimation quality to the operating SNR, as is usually the case in reality. In the following computation we treat σ_e^2 as a constant parameter, irrespective of the aforementioned scenarios.

A. Linear Reference System

The capacity of a linear MIMO system with noisy channel estimation is not known. Only upper and lower bounds on the mutual information are given in literature [25]. A lower bound on the mutual information is given by

$$\mathrm{I}_{\mathrm{lower}}\left(\boldsymbol{s}; \boldsymbol{y}\middle|\widehat{\boldsymbol{H}}\right) = \mathcal{E}_{\widehat{\boldsymbol{H}}}\left[\log\left|\boldsymbol{I}_{N_\mathrm{R}} + \frac{\mathrm{SNR}}{1+\sigma_e^2\mathrm{SNR}}\frac{1}{N_\mathrm{T}}\widehat{\boldsymbol{H}}\widehat{\boldsymbol{H}}^\mathrm{H}\right|\right], \quad (21)$$

where we used that $\mathcal{E}[\boldsymbol{s}^\mathrm{H}\boldsymbol{s}] = \frac{P}{N_\mathrm{T}}\boldsymbol{I}$ and the SNR definition from (5). An upper bound can be computed as a function of the lower bound as follows:

$$\mathrm{I}_{\mathrm{upper}}\left(\boldsymbol{s}; \boldsymbol{y}\middle|\widehat{\boldsymbol{H}}\right) = \mathrm{I}_{\mathrm{lower}}\left(\boldsymbol{s}; \boldsymbol{y}\middle|\widehat{\boldsymbol{H}}\right) + N_\mathrm{R} \cdot \mathcal{E}_{\boldsymbol{s}}\left[\log\frac{1+\sigma_e^2\mathrm{SNR}}{1+\sigma_e^2\|\boldsymbol{s}\|^2}\right]. \quad (22)$$

However, these bounds are tight only for small values of σ_e^2, and are not suited as reference curves for larger values of σ_e^2 (cf. Sec. V). For a better comparison with our nonlinear MIMO achievable rates, we will numerically compute the achievable rate of the linear reference system with a Gaussian input distribution. This computation is very similar to the nonlinear MIMO achievable rate computation that follows in the next session, and we will describe both in parallel.

B. Nonlinear MIMO

1) Rate Expressions: We compute the mutual information between $\boldsymbol{s}$ and $\boldsymbol{y}$ given a channel estimate $\widehat{\boldsymbol{H}}$:

$$\mathrm{I}\left(\boldsymbol{s}; \boldsymbol{y}\middle|\widehat{\boldsymbol{H}}\right) = \mathcal{E}_{\widehat{\boldsymbol{H}}}\left[\mathrm{I}\left(\boldsymbol{s}; \boldsymbol{y}\middle|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right)\right] = \mathcal{E}_{\widehat{\boldsymbol{H}}}\left[\mathrm{h}\left(\boldsymbol{y}\middle|\widehat{\mathbf{H}}\right) - \mathrm{h}\left(\boldsymbol{y}\middle|\boldsymbol{s}, \widehat{\mathbf{H}}\right)\right]. \quad (23)$$

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

The required distribution functions are $\mathrm{f}(\boldsymbol{y}|\widehat{\mathbf{H}})$ and $\mathrm{f}(\boldsymbol{y}|\boldsymbol{s},\widehat{\mathbf{H}})$. Unlike in (8), 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 $\mathrm{h}\left(\boldsymbol{y}\left|\widehat{\mathbf{H}}\right.\right)$ and $\mathrm{h}\left(\boldsymbol{y}\left|\boldsymbol{s},\widehat{\mathbf{H}}\right.\right)$. From Eq. (20), the general input-output relation is written as

$$\boldsymbol{y} = g\left(\widehat{\mathbf{H}}\boldsymbol{s} + \boldsymbol{E}\boldsymbol{s} + \boldsymbol{w}\right). \tag{24}$$

For the computation of the conditional entropy $\mathrm{h}\left(\boldsymbol{y}\left|\boldsymbol{s},\widehat{\mathbf{H}}\right.\right)$, we write (24) as

$$\boldsymbol{y} = g\left(\widehat{\mathbf{H}}\boldsymbol{s} + \widetilde{\boldsymbol{w}}\right) = g(\boldsymbol{r}), \tag{25}$$

where $\widetilde{\boldsymbol{w}} = \boldsymbol{w} + \boldsymbol{E}\boldsymbol{s}$. Here we assume that $\boldsymbol{w}$ and $\boldsymbol{E}$ are statistically independent. This is a valid assumption since coding is performed across different channel realizations, which is common when computing ergodic rates. Hence, for a given realization of $\boldsymbol{s} = \mathbf{s}$, $\widetilde{\boldsymbol{w}}$ is Gaussian distributed, as

$$\widetilde{\boldsymbol{w}}\,|\mathbf{s} \sim \mathcal{CN}_{\widetilde{\boldsymbol{w}}}\left(\mathbf{0}, \left(\sigma_w^2 + \sigma_e^2 \sum_{j=1}^{N_\mathrm{T}} |s_j|^2\right)\boldsymbol{I}\right). \tag{26}$$

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

$$\boldsymbol{r}\left|\mathbf{s},\widehat{\mathbf{H}} \sim \mathcal{CN}_{\boldsymbol{r}}\left(\widehat{\mathbf{H}}\mathbf{s}, \left(\sigma_w^2 + \sigma_e^2 \sum_{j=1}^{N_\mathrm{T}} |s_j|^2\right)\boldsymbol{I}\right)\right. . \tag{27}$$

We use the above distribution for the computation of all required entropies. In the case of the linear MIMO system, (27) corresponds to the distribution of $\boldsymbol{y}|\mathbf{s},\widehat{\mathbf{H}}$, since $\boldsymbol{y} = \boldsymbol{r}$. Hence, $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\widehat{\mathbf{H}})$ is readily given as

$$\mathrm{h}_{\mathrm{linear}}\left(\boldsymbol{y}\left|\boldsymbol{s},\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right) = \mathcal{E}_{\boldsymbol{s}}\left[\mathrm{h}\left(\boldsymbol{y}\left|\mathbf{s},\widehat{\mathbf{H}}\right.\right)\right] = \mathcal{E}_{\boldsymbol{s}}\left[N_\mathrm{R}\cdot\log\left(\pi e\left(\sigma_w^2 + \sigma_e^2\sum_{j=1}^{N_\mathrm{T}}|s_j|^2\right)\right)\right], \tag{28}$$

averaged over many realizations of $\boldsymbol{s}$. Note that (28) is independent of the channel estimate $\widehat{\mathbf{H}}$, which affects only the mean of the distribution in (27). This is true only for the linear MIMO case.

In the case of nonlinear MIMO, we can factorize the conditional distribution of $\boldsymbol{y}$ similar to (10) as follows

$$\mathrm{f}\left(\boldsymbol{y}\left|\boldsymbol{s} = \mathbf{s}, \widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right) = \prod_{i=1}^{N_\mathrm{R}} \mathrm{f}\left(y_i\left|\boldsymbol{e}_i^{\mathrm{T}}\widehat{\mathbf{H}}\mathbf{s}\right.\right), \tag{29}$$

using the fact that the composite noise $\widetilde{\boldsymbol{w}}$ remains white, i.e., the correlation matrix is diagonal. The distribution $\mathrm{f}(y_i|\boldsymbol{e}_i^{\mathrm{T}}\widehat{\mathbf{H}}\mathbf{s})$ per receive antenna is then given by (11) and (12), evaluated by setting $x_i \leftarrow \boldsymbol{e}_i^{\mathrm{T}}\widehat{\mathbf{H}}\mathbf{s}$ and $\sigma_w^2 \leftarrow \sigma_w^2 + \sigma_e^2\sum_{j=1}^{N_\mathrm{T}}|s_j|^2$ in those equations, using (27). Based on the above equation, we compute the remaining entropies numerically.

2) Numerical Computation: First, we compute $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\widehat{\mathbf{H}})$ for nonlinear MIMO, using MC integration. We create N realizations $(\mathbf{s}_i,\mathbf{y}_i)$ distributed according to $\mathrm{f}(\boldsymbol{s},\boldsymbol{y})$. We generate these samples by randomly creating N realizations of $\boldsymbol{s}$, and compute $\boldsymbol{y}$ by varying the composite noise $\widetilde{\boldsymbol{w}}$. Note that the only source of uncertainty is $\widetilde{\boldsymbol{w}}$, which includes the estimation error uncertainty. Then, we compute

$$\begin{aligned}\mathrm{h}\left(\boldsymbol{y}\left|\boldsymbol{s},\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right) &= -\iint \mathrm{f}\left(\boldsymbol{s},\boldsymbol{y}\left|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right)\log\left(\mathrm{f}\left(\boldsymbol{y}\left|\boldsymbol{s},\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right)\right)\mathrm{d}\boldsymbol{s}\mathrm{d}\boldsymbol{y}\\ &\simeq -\frac{1}{N}\sum_{i=1}^{N}\log\left(\mathrm{f}\left(\boldsymbol{y} = \mathbf{y}_i\left|\boldsymbol{s} = \mathbf{s}_i,\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right)\right),\end{aligned} \tag{30}$$

for a given channel estimate. The conditional distribution in (30) is described below (29). This concludes the computation of $\mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\widehat{\mathbf{H}})$ both for linear (see (28)) and nonlinear MIMO.

Finally, $\mathrm{h}(\boldsymbol{y}|\widehat{\mathbf{H}})$ is computed in the same way for linear and nonlinear MIMO. However, unlike in the perfect CSI case, the distribution $\mathrm{f}(\boldsymbol{y}|\widehat{\boldsymbol{H}})$ is not known analytically even for a SISO system. This is due to the term $(\boldsymbol{E}\boldsymbol{s})$ in (24), which consists of the sum of products of Gaussian random variables. Hence we resort to MC estimation of $\mathrm{f}(\boldsymbol{y}|\widehat{\boldsymbol{H}})$ as follows. First we create N samples $\boldsymbol{y} = \mathbf{y}_i$, generated according to $\mathrm{f}(\boldsymbol{y}|\widehat{\boldsymbol{H}})$. That is, for a fixed channel estimate $\widehat{\mathbf{H}}$, the source of uncertainty are $\boldsymbol{s}$, $\boldsymbol{E}$ and $\boldsymbol{w}$. The entropy $\mathrm{h}(\boldsymbol{y}|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ is computed with MC integration

$$\mathrm{h}\left(\boldsymbol{y}\left|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right) \simeq -\frac{1}{N}\sum_{i=1}^{N}\log\left(\mathrm{f}\left(\mathbf{y}_i\left|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right)\right). \tag{31}$$

The values $\mathrm{f}(\mathbf{y}_i|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}})$ themselves are computed again with MC integration. We use the model with the composite noise $\widetilde{\boldsymbol{w}}$ described above, and average only over $\boldsymbol{s}$:

$$\begin{aligned}\mathrm{f}\left(\mathbf{y}_i\left|\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right) &= \int \mathrm{f}(\boldsymbol{s})\,\mathrm{f}\left(\mathbf{y}_i\left|\boldsymbol{s},\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right)\mathrm{d}\boldsymbol{s}\\ &\simeq \frac{1}{M}\sum_{j=1}^{M}\mathrm{f}\left(\mathbf{y}_i\left|\boldsymbol{s} = \mathbf{s}_j,\widehat{\boldsymbol{H}} = \widehat{\mathbf{H}}\right.\right).\end{aligned} \tag{32}$$

Here, $\mathbf{s}_j$ are M samples of $\boldsymbol{s}$ distributed according to $\mathrm{f}(\boldsymbol{s})$, and generated independent of $\mathbf{y}_i$. The conditional distribution in (32) is given in (27) for linear MIMO, and in (10), (11) and (12) for nonlinear MIMO, as described below (29). The variation of $\boldsymbol{E}$ is captured in the variance of the composite noise $\widetilde{\boldsymbol{w}}$ through the estimation error variance σ_e^2, and its influence is a function of the instantaneous signal realization $\mathbf{s}$. Averaging only over $\boldsymbol{s}$ by using the composite noise $\widetilde{\boldsymbol{w}}$ is numerically preferable because it spares the two-fold averaging over both $\boldsymbol{s}$ and $\boldsymbol{E}$.

V. Simulation Results

In this section we present numerically computed achievable rates for various configurations. In all following plots, solid, dotted and dashed lines stand for linear, amplitude and phase detection, respectively. The input $\boldsymbol{s}$ is distributed according to $\boldsymbol{s} \sim \mathcal{CN}_{\boldsymbol{s}}(\mathbf{0},\sigma_s^2\boldsymbol{I})$, unless stated otherwise. The SNR is defined in (5). The channel matrix entries are modeled as i.i.d. zero-mean unit-variance circularly symmetric Gaussian random variables.

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

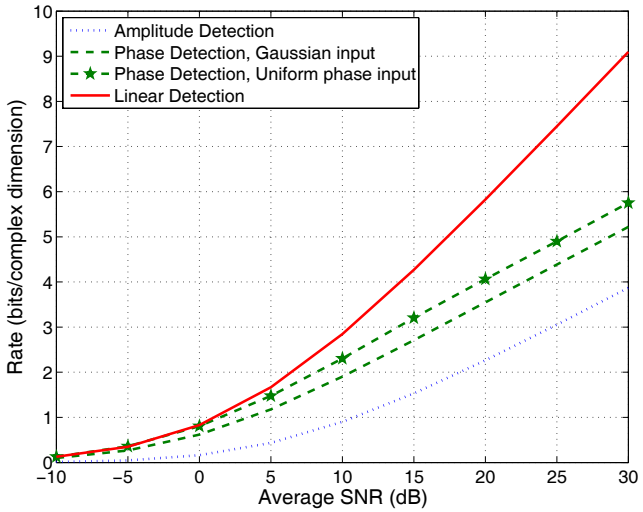

Fig. 2. SISO system, perfect CSI.

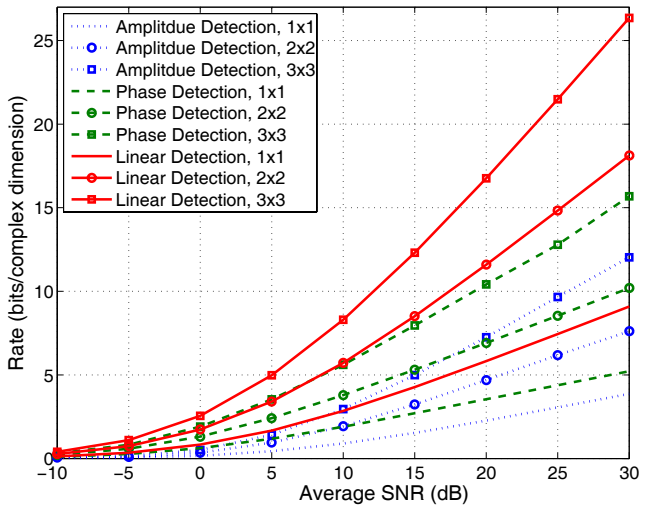

Fig. 3. Perfect CSI, $N_T = N_R = \{1, 2, 3\}$.

A. Perfect CSI

First we consider the case of perfect CSI at the receiver, as described in Section III. Fig. 2 depicts the rates for a SISO system, computed by averaging (8) over many channel realizations, using (13), (11), (12) and the analytic entropies for amplitude and phase detection (known in the SISO case only). Nonlinear detection is clearly inferior to linear due to the non-invertible, information-destroying processing by the function $g(\cdot)$. Furthermore, the slope of the rate curve for nonlinear detection at high SNR is approximately half the slope of linear detection, and thus, equal to the slope of linear detection using one-dimensional (real-valued) signals. We also plot the rate for phase detection when the input signal has unit amplitude and uniform phase. This distribution yields higher rate, although it is not capacity achieving, contrary to what is stated in [26]. For higher number of receive antennas, simulations show that the uniform distribution yields the same rates as the complex Gaussian distribution. This effect rises from the MIMO channel multiplexing. In the rest of the paper we use only the complex Gaussian input distribution. An observation which will also hold in the following plots is that phase detection outperforms amplitude detection at all SNR. In other words, phase conveys more information than amplitude. This is due to fact that the conditional PDF in (12) depends on the amplitude through ρ_i, and hence carries information regarding the state of the channel (weak of strong). On the other side, (11) is independent of the phase.

Fig. 3 depicts a MIMO system with $N_T = N_R$. As expected, the slope of the nonlinear rate curves increases when increasing the number of transmit and receive antennas. Furthermore, the slope of the nonlinear rates remains approximately *half* the slope of the respective linear capacity, like in the SISO case. For example, for $N_T = N_R = 2$, the slope of amplitude and phase detection at high SNR is the same as the slope of linear detection with $N_T = N_R = 1$. Hence, we conjecture that an $N \times N$ nonlinear MIMO system achieves $N/2$ spatial multiplexing gain.

Fig. 4(a) compares a system with $N_T = 2$ and variable number of receive antennas $N_R = 1, \ldots, 4$. For linear detection,

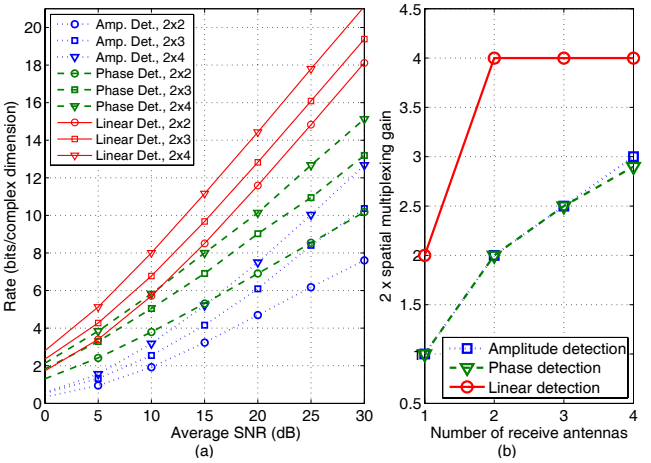

Fig. 4. $N_T = 2$ and varying N_R, perfect CSI. (a) Achievable rates and (b) spatial multiplexing gain in real-valued dimensions.

the slope of the capacity curves for $N_R \geq 2$ remains the same, since the degrees of freedom $N_{\min} = 2$ are unchanged and fully exploited. The parallel shift is the result of a power gain. However, the slope of the nonlinear rate curves continues to increase as we add more receive antennas, but always remains lower than the linear detection slope. We quantify this observation by approximating the rate curves with $R = \alpha \log(1 + \beta \mathrm{SNR})$ and plot twice the estimated spatial multiplexing gain α in Fig. 4(b) (this corresponds to real degrees of freedom). We see that the nonlinear detectors exploit the additional receive antennas to achieve higher spatial multiplexing gains. We conjecture that they always miss one real degree of freedom, i.e., they reach a spatial multiplexing gain up to $2N_{\min} - 1 = 3$ in real dimensions. We justify our observation by means of an example. In an $N_T = 2$, $N_R = 2$ system the received signal consists of two phases and two amplitudes. A phase detector can extract the two phases of the signal, but has no information about the two unknown amplitudes. However, if $N_R = 3$, we can deduce an equation relating the two amplitudes by combining the three observed phases. Thus, we only miss one more equation in

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

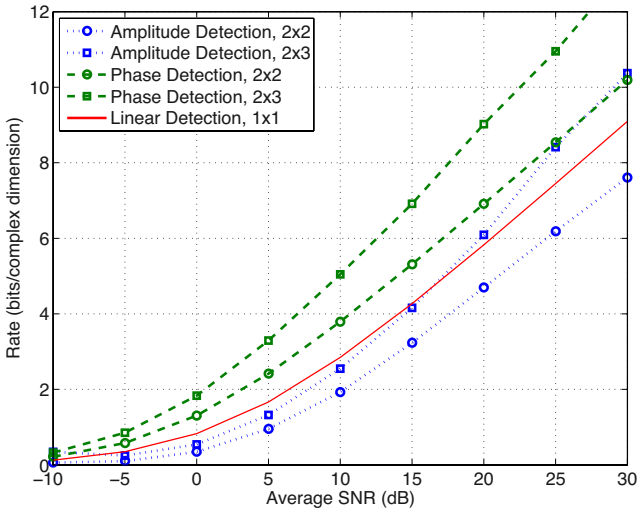

Fig. 5. Linear SISO versus nonlinear MIMO rates, perfect CSI.

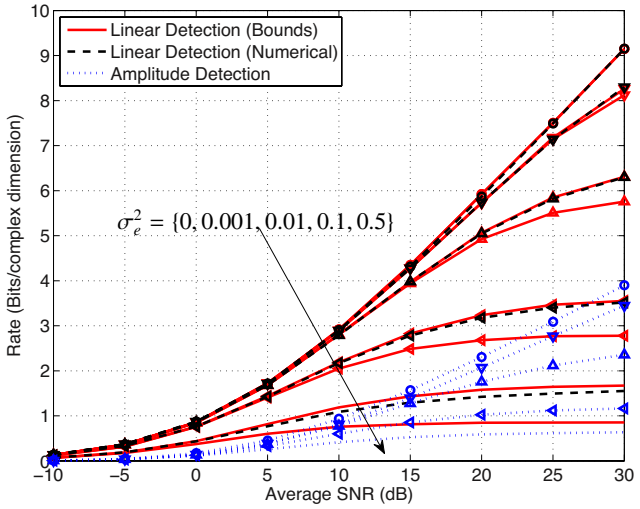

Fig. 6. SISO system, noisy CSI. Amplitude Detection.

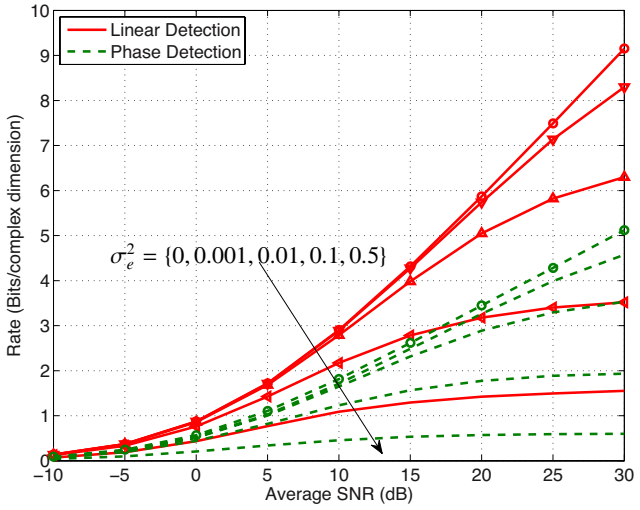

Fig. 7. SISO system, noisy CSI. Phase Detection.

order to also extract the two amplitudes. This manifests itself in the increased multiplexing gain. Nevertheless, there will always remain an ambiguity regarding the amplitudes, since multiplying the complex-valued received vector by a scalar would change the amplitudes but not the phases.

Fig. 5 compares the achievable rates of a linear SISO systems to those of nonlinear MIMO systems. As we see, phase detection outperforms linear detection, while amplitude detection is slightly worse than linear detection. This comparison makes sense if we consider that the power consumption of the linear SISO detection receiver is expected to be higher than that of the nonlinear MIMO receiver, although the latter employs multiple receiver branches, as explained in Section I. Hence, investing in more antennas and nonlinear receivers preserves the low power consumption requirements while not harming the achievable rates of these systems significantly.

B. Noisy CSI

1) Constant estimation error variance: We proceed to achievable rates with noisy channel state information. First, we consider a scenario where the channel estimate degradation is due to outdated information, i.e., σ_e^2 is constant for all SNR. The achievable rates of the linear reference system are computed as described in Sections IV-A and IV-B. We depict the lower and upper bounds of (21) and (22) only in Fig. 6 to illustrate how the numerically computed curves are more useful than the two bounds, especially at high σ_e^2. The estimation error variance is chosen as $\sigma_e^2 = \{0.5, 0.1, 0.01, 0.001\}$ and corresponds to an estimation SNR of $\{3, 10, 20, 30\}$ dB, respectively, while $\sigma_e^2 = 0$ corresponds to perfect CSI. Curves with the same marker correspond to the same value of σ_e^2.

Figs. 6 and 7 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 an excellent 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 for various antenna configurations, pointing out that both techniques are equally robust to channel estimation errors.

Fig. 8 depicts the achievable rates for a $N_T = N_R = 2$ MIMO system. The behavior is similar to the SISO case. Phase detection performs 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. Concluding, we can say that both amplitude and phase detection exhibit the same behavior as linear detection when the channel estimation is not perfect.

2) Variable estimation error variance: Finally, we simulate the achievable rates when the estimation $\mathrm{SNR}_{\mathrm{est}} = 10 \log_{10} \sigma_e^{-2}$ is a function of the operating SNR. In this case, operating at high SNR implies good quality channel estimation. We consider the case where $\mathrm{SNR}_{\mathrm{est}}$ is equal to the operating SNR. The results are depicted in Fig. 9, along with the case of perfect CSI (marked with a circle). As can be seen, the rate curves are parallel to the perfect CSI curves, and do

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

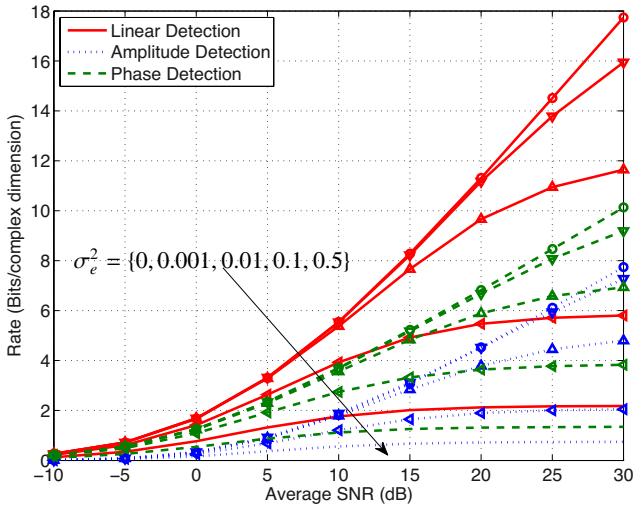

Fig. 8. MIMO system, $N_T = N_R = 2$, noisy CSI.

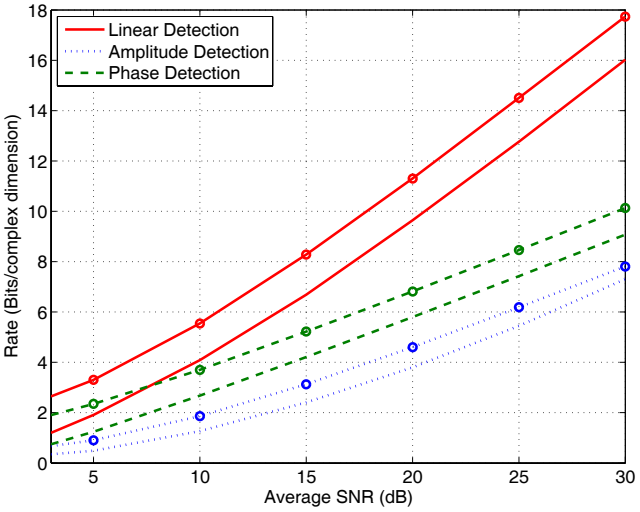

Fig. 9. MIMO system, $N_T = N_R = 2$, $SNR_{est} = SNR$.

not saturate for increasing SNR, since the channel estimation quality also increases.

VI. Impact of Nonlinear MIMO on Wireless Sensor Networks

In this section we discuss the impact of using a nonlinear MIMO receiver in a wireless sensor network. There are certain properties that are critical for the power consumption of the radio system on a sensor node. A sensor network is characterized by sporadic transmission (low duty-cycle) of short packets over small distances. This implies that the output power of the transmitter is relatively small, and does not dominate the Tx power consumption. Furthermore, the radio system must be waken-up every time a transmission is pending, which means that the power consumed during the start-up phase is very important. However, since the transmission time is short due to the small packet size, the start-up power consumption is equally important to the actual transmission. Hence, increasing the transmission rate alone is meaningful only if accompanied with a short start-up time, as elaborated in [3] and [27]. Concluding, the decisive factors for a low-power sensor node are *low Tx/Rx power consumption*, through suitable choice of (nonlinear) modulation, *short start-up time* and *high rate*. Our proposed nonlinear MIMO systems possess all these properties.

As we saw in Section V, nonlinear MIMO systems exploit spatial degrees of freedom and achieve *higher rates* than nonlinear SISO systems. Compared to a linear SISO system (cf. Fig. 5), the achievable rate of amplitude and phase detection are similar (or even better for phase detection) and the scaling with SNR (spatial multiplexing gain) is at least as good as that for linear SISO. However, nonlinear MIMO implementations have not been considered yet. As explained in Section I, facilitating multiple antennas on a node is possible with the use of compact MIMO-antennas. Regarding the RF circuitry, a nonlinear MIMO receiver would use multiple Tx/Rx chains of a nonlinear SISO design, e.g. [2] for amplitude detection. Associated properties like *very low start-up time* and reduced complexity will remain the same, despite the multiple antennas. The analog front-end power consumption will be increased, but still remain *low compared to a linear I/Q receiver*, since the low-power nonlinear SISO designs are orders of magnitude more efficient (cf. [2], [5]). Thus, nonlinear MIMO nodes would improve the power consumption in a sensor network without requiring any changes from the higher-layer network protocols.

VII. Conclusion

We proposed a new type of receiver that combines amplitude or phase detection and multiple antenna systems. We investigated the fundamental limits of such systems, and found that nonlinear MIMO systems also exploit the spatial degrees of freedom, although in a different way than linear MIMO systems. Furthermore, they exhibit a similar robustness to noisy CSI as their linear counterparts. Nonlinear MIMO systems are by design characterized by low power consumption, fast start-up times and high (relatively) rates. Properties, which render them a very promising alternative for short-range low-power systems, like sensor networks or RFID.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey of sensor networks," *IEEE Commun. Mag.*, vol. 40, pp. 102-114, Aug. 2002.

[2] D. C. Daly and A. P. Chandrakasan, "An energy-efficient OOK transceiver for wireless sensor networks," *IEEE J. Solid State Circuits*, vol. 42, pp. 1003-1011, May 2007.

[3] E. Shih, S. Cho, F. S. Lee, E. H. Calhoun, and A. P. Chandrakasan, "Design considerations for energy-efficient radios in wireless microsensor networks," *J. VLSI Signal Process.*, vol. 37, pp. 77-94, May 2004.

[4] A. Y. Wang, S. Cho, C. G. Sodini, and A. P. Chandrakasan, "Energy efficient modulation and MAC for asymmetric RF microsensor systems," in *Proc. International Symp. Low Power Electronics Design*, Aug. 2001, pp. 106-111.

[5] J. Polastre, R. Szewczyk, and D. Culler, "Telos: enabling ultra-low power wireless research," in *Proc. IEEE Symp. Inf. Process. Sensor Netw.*, Apr. 2005, pp. 364-369.

[6] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, pp. 585-595, Nov. 1999.

[7] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-355, Mar. 1998.

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.

[8] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1089-1098, Aug. 2004.

[9] S. K. Jayaweera, "Energy analysis of MIMO techniques in wireless sensor networks," in *38th Annual Conf. Inf. Science Syst.*, Princeton, NJ, Mar. 2004.

[10] H. Dai, L. Xiao, and Q. Zhou, "Energy efficiency of MIMO transmission strategies in wireless networks," in *International Conf. Computing, Commun. Control Technol.*, Austin, TX, Aug. 2004.

[11] C. Waldschmidt and W. Wiesbeck, "Compact wide-band multimode antennas for MIMO and diversity," *IEEE Trans. Antennas Propag.*, vol. 52, pp. 1963-1969, Nov. 2004.

[12] B. Getu and J. Andersen, "The MIMO cube—a compact MIMO antenna," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1136-1141, May 2005.

[13] D. Browne, M. Manteghi, M. Fitz, and Y. Rahmat-Samii, "Experiments with compact antenna arrays for MIMO radio communications," *IEEE Trans. Antennas Propagation*, vol. 54, pp. 3239-3250, Nov. 2006.

[14] G. K. Psaltopoulos, F. Trösch, and A. Wittneben, "On achievable rates of MIMO systems with nonlinear receivers," in *Proc. IEEE International Symp. Inf. Theory*, June 2007, pp. 1071-1075.

[15] G. K. Psaltopoulos and A. Wittneben, "Achievable rates of nonlinear MIMO systems with noisy channel state information," in *Proc. IEEE International Symp. Inf. Theory*, July 2008, pp. 2659-2662.

[16] G. Fettweis, M. Löhning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: a new paradigm," *International J. Wireless Inf. Netw.*, vol. 14, pp. 133-148, June 2007.

[17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[18] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw-Hill, 1991.

[19] W. R. Bennett, "Methods of solving noise problems," *Proc. IRE*, vol. 44, pp. 609-638, May 1956.

[20] S. Weinzierl, "Introduction to Monte Carlo methods," in Arxiv, Rep. Nr. NIKHEF-00-012, June 2000.

[21] K. S. Miller, *Multidimensional Gaussian Distributions*. Wiley, 1964.

[22] R. K. Mallik, "On multivariate Rayleigh and exponential distributions," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1499-1515, June 2003.

[23] Y. Chen and C. Tellambura, "Infinite series representaion of the trivariate and quadrivariate Rayleigh distribution and their applications," *IEEE Trans. Commun.*, vol. 53, pp. 2092-2101, Dec. 2005.

[24] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Table*, 1965.

[25] T. Yo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203-2214, May 2006.

[26] J. P. Aldis and A. G. Burr, "The channel capacity of discrete time phase modulation in AWGN," *IEEE Trans. Inf. Theory*, vol. 39, pp. 184-185, Jan. 1993.

[27] E. H. Calhoun, D. C. Daly, N. Verma, D. F. Finchelstein, D. D. Wentzloff, A. Wang, S.-H. Cho, and A. P. Chandrakasan, "Design considerations for ultra-low energy wireless microsensor nodes," *IEEE Trans. Computers*, vol. 54, pp. 727-740, June 2005.

Georgios K. Psaltopoulos (S'01) received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 2003, and the M.Sc. degree in communications engineering from the Munich University of Technology (TUM), Germany, in 2005. He has been with the Communication Technology Laboratory at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, since 2005, working towards the Ph.D. degree. His research interests include signal processing in wireless communications and communication theory, with emphasis on very low-power, affordable nonlinear MIMO systems.

Armin Wittneben Armin Wittneben received his M.S. and Ph.D. degrees in electrical engineering from the Technical University of Darmstadt, Germany, in 1983 and 1989 respectively. In 1989 he joined the central research unit ASCOM Tech of ASCOM Ltd., Switzerland, as postdoctoral researcher. In 1992 he became head of the Wireless Technology and Security Department. In 1996 he also served as co-director of ASCOM Tech. In 1997 he received the Venia Legendi (Habilitation) in Communication Technology from the TU-Darmstadt. In 1998 he accepted a position as full professor of Telecommunication at the Saarland University, Germany. Since 2002 he has been full professor of Wireless Technology at ETH Zurich, Switzerland and is serving as head of the Communication Technology Laboratory. His research interests include all aspects of wireless communication with special emphasis on communication theory, signal processing, cooperative wireless signaling and protocols, distributed MIMO systems and Ultra-Wideband/Ultra-Low Power communication, imaging and position location.

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on February 19,2010 at 08:39:33 EST from IEEE Xplore. Restrictions apply.