On Achievable Rates of MIMO Systems with Nonlinear Receivers

G. Psaltopoulos F. Troesch A. Wittneben

Communication Technology Laboratory
Wireless Communications Group
ETH Zurich

July 2006 — COST 289, Aveiro

Agenda

1. Introduction
2. System Model
3. Rate Computation
4. Simulation Results
5. Conclusions

Motivation

- MIMO systems achieve high rate gains.
- Multiple antennas = multiple receiver chains.
- **Expensive** and **power consuming** receiver circuitry (mixers, frequency synthesis, PLL's, linear amplifiers etc.).
- Applications like sensor networks require very energy-efficient and cheap solutions in order to maximize lifetime.
- **Amplitude** and **phase** detection are known as low-cost/low-power alternatives to receiver design.
- We explore **achievable rates** of **coherent** MIMO systems that use amplitude or phase detection receivers.

Motivation

- MIMO systems achieve high rate gains.
- Multiple antennas = multiple receiver chains.
- **Expensive** and **power consuming** receiver circuitry (mixers, frequency synthesis, PLL's, linear amplifiers etc.).
- Applications like sensor networks require very energy-efficient and cheap solutions in order to maximize lifetime.
- **Amplitude** and **phase** detection are known as low-cost/low-power alternatives to receiver design.
- We explore **achievable rates** of **coherent** MIMO systems that use amplitude or phase detection receivers.

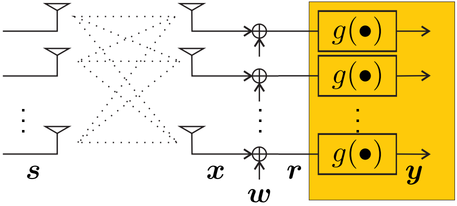
Motivation

- MIMO systems achieve high rate gains.
- Multiple antennas = multiple receiver chains.
- **Expensive** and **power consuming** receiver circuitry (mixers, frequency synthesis, PLL's, linear amplifiers etc.).
- Applications like sensor networks require very energy-efficient and cheap solutions in order to maximize lifetime.
- **Amplitude** and **phase** detection are known as low-cost/low-power alternatives to receiver design.
- We explore **achievable rates** of **coherent** MIMO systems that use amplitude or phase detection receivers.

Channel State Information

- We assume that channel knowledge is available at the receiver.
- In a Time Division Duplex system, CSI can be signaled from the more complex access point to the node.
- The node could employ a linear receiver for channel estimation and switch to a nonlinear receiver for receiving data.
- For phase detection, it is possible to estimate the channel amplitude from the variation of the phase samples [Althaus, Wittneben].

System Model



(See flip board!)

- Linear model: $\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}$.
- The channel matrix $\boldsymbol{H}$ has i.i.d. $\mathcal{CN}(0,1)$ entries.
- Transmit signal — $\boldsymbol{s} \sim \mathcal{CN}(\mathbf{0}, \sigma_s^2 \boldsymbol{I}_{N_\mathsf{T}})$.
- AWG noise — $\boldsymbol{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \boldsymbol{I}_{N_\mathsf{R}})$.
- The nonlinear operator g extracts either the phase or the amplitude of $\boldsymbol{r}$.

Reference Model - Linear MIMO

Capacity of linear MIMO System [Telatar, Foschini-Gans]

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

- Average SNR per receive antenna

$$\mathsf{SNR} = \frac{\mathrm{tr}(\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}})}{\sigma_w^2} = \frac{N_{\mathsf{T}}\sigma_s^2}{\sigma_w^2}.$$

- Capacity scales with $N_{\mathsf{min}} = \mathsf{min}(N_{\mathsf{T}}, N_{\mathsf{R}})$ at high SNR

$$C_{\mathsf{lin}} \simeq N_{\mathsf{min}} \log\frac{\mathsf{SNR}}{N_{\mathsf{T}}} + \sum_{i=1}^{N_{\mathsf{min}}} \mathrm{E}\left[\log\lambda_i^2\right].$$

Reference Model - Linear MIMO

Capacity of linear MIMO System [Telatar, Foschini-Gans]

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

- Average SNR per receive antenna

$$\mathsf{SNR} = \frac{\operatorname{tr}(\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}})}{\sigma_w^2} = \frac{N_{\mathsf{T}}\sigma_s^2}{\sigma_w^2}.$$

- Capacity scales with $N_{\mathsf{min}} = \mathsf{min}(N_{\mathsf{T}}, N_{\mathsf{R}})$ at high SNR

$$C_{\mathsf{lin}} \simeq N_{\mathsf{min}} \log \frac{\mathsf{SNR}}{N_{\mathsf{T}}} + \sum_{i=1}^{N_{\mathsf{min}}} \mathrm{E}\left[\log \lambda_i^2\right].$$

Evaluating Nonlinear MIMO Mutual Information (1)

- Mutual Information given the channel realization H

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

- Chain rule for Mutual Information

$$\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) = \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H),$$
$$\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) = \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H).$$

- Since $\boldsymbol{s} \to \boldsymbol{x} \to \boldsymbol{y}$, we have that $\mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H) = 0$.

Evaluating Nonlinear MIMO Mutual Information (1)

- Mutual Information given the channel realization H

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

- Chain rule for Mutual Information

$$\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) = \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H),$$
$$\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) = \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H).$$

- Since $\boldsymbol{s} \rightarrow \boldsymbol{x} \rightarrow \boldsymbol{y}$, we have that $\mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H) = 0$.

Evaluating Nonlinear MIMO Mutual Information (1)

- Mutual Information given the channel realization H

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

- Chain rule for Mutual Information

$$\begin{aligned}
\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) &= \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H),\\
\mathrm{I}(\boldsymbol{s},\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H) &= \mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{H}=H) + \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H).
\end{aligned}$$

- Since $\boldsymbol{s} \to \boldsymbol{x} \to \boldsymbol{y}$, we have that $\mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}=H) = 0$.

Evaluating Nonlinear MIMO Mutual Information (2)

- Meanwhile, since $\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s}$

$$\begin{aligned}
\mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H) &= \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H) - \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) \\
&= \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) - \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) \\
&= 0.
\end{aligned}$$

- Hence

$$\mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{H}=H) = \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H).$$

and

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

Evaluating Nonlinear MIMO Mutual Information (2)

- Meanwhile, since $\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s}$

$$\begin{aligned} \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H) &= \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{H}=H) - \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) \\ &= \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) - \mathrm{h}(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{x},\boldsymbol{H}=H) \\ &= 0. \end{aligned}$$

- Hence

$$\mathrm{I}(\boldsymbol{s};\boldsymbol{y}|\boldsymbol{H} = H) = \mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H} = H).$$

and

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

Evaluating $\mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H)$

- Definition

$$\begin{aligned}\mathrm{I}(\boldsymbol{x};\boldsymbol{y}) &= \mathrm{h}(\boldsymbol{y}) - \mathrm{h}(\boldsymbol{y}|\boldsymbol{x}) \\ &= -\int p(\boldsymbol{y})\log(p(\boldsymbol{y}))\mathrm{d}\boldsymbol{y} \\ &\quad + \iint p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x})\log(p(\boldsymbol{y}|\boldsymbol{x})\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{x}.\end{aligned}$$

- Since $g(\bullet)$ acts element wise

$$p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{N_\mathsf{R}} p(y_i|x_i).$$

Evaluating $\mathrm{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}=H)$

- Definition

$$\begin{aligned}\mathrm{I}(\boldsymbol{x};\boldsymbol{y}) &= \mathrm{h}(\boldsymbol{y})-\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})\\ &= -\int p(\boldsymbol{y})\log(p(\boldsymbol{y}))\mathrm{d}\boldsymbol{y}\\ &\quad+\iint p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x})\log(p(\boldsymbol{y}|\boldsymbol{x})\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{x}.\end{aligned}$$

- Since $g(\bullet)$ acts element wise

$$p(\boldsymbol{y}|\boldsymbol{x})=\prod_{i=1}^{N_{\mathsf{R}}}p(y_i|x_i).$$

$\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ — conditional pdf $p(y_i|x_i)$ (1)

- The nonlinear operator for amplitude/phase detection

$$
\begin{aligned}
y_{i,\text{ampl.}} &= g_{\text{ampl.}}(r_i) = \sqrt{\Re(r_i)^2 + \Im(r_i)^2}, \\
y_{i,\text{phase}} &= g_{\text{phase}}(r_i) = \arg(r_i) = \tan^{-1}\left(\frac{\Im(r_i)}{\Re(r_i)}\right).
\end{aligned}
$$

- Given x_i, $r_i \sim \mathcal{CN}(x_i, \sigma_w^2)$, and the norm of r is Rice distributed

$$
p_{\text{ampl.}}(y_i|x_i) = \frac{2y_i}{\sigma_w^2} e^{-\frac{y_i^2 + |x_i|^2}{\sigma_w^2}} I_0\left(\frac{2y_i|x_i|}{\sigma_w^2}\right).
$$

(Use approximation of $I_0(z)$ for high z.)

$\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ — conditional pdf $p(y_i|x_i)$ (1)

- The nonlinear operator for amplitude/phase detection

$$\begin{aligned} y_{i,\text{ampl.}} &= g_{\text{ampl.}}(r_i) = \sqrt{\Re(r_i)^2 + \Im(r_i)^2}, \\ y_{i,\text{phase}} &= g_{\text{phase}}(r_i) = \arg(r_i) = \tan^{-1}\left(\frac{\Im(r_i)}{\Re(r_i)}\right). \end{aligned}$$

- Given x_i, $r_i \sim \mathcal{CN}(x_i, \sigma_w^2)$, and the norm of r is Rice distributed

$$p_{\text{ampl.}}(y_i|x_i) = \frac{2y_i}{\sigma_w^2} e^{-\frac{y_i^2 + |x_i|^2}{\sigma_w^2}} I_0\left(\frac{2y_i|x_i|}{\sigma_w^2}\right).$$

(Use approximation of $I_0(z)$ for high z.)

$\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ — conditional pdf $p(y_i|x_i)$ (2)

- The phase has a more involved distribution

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

where $\rho_i = \frac{|x_i|^2}{\sigma_w^2}$ and $\Delta\phi_i = y_{i,\text{phase}} - \arg(x_i) \in [0, 2\pi)$.

- We estimate the entropy $\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ with **Monte Carlo (MC) integration**

$$\mathrm{h}(\boldsymbol{y}|\boldsymbol{x}) = -\iint p(\boldsymbol{x}, \boldsymbol{y}) \log(p(\boldsymbol{y}|\boldsymbol{x})) \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{x}$$

$$= -\mathrm{E}_{p(\boldsymbol{x},\boldsymbol{y})}[\log(p(\boldsymbol{y}|\boldsymbol{x}))] \simeq -\frac{1}{N} \sum_{i=1}^{N} \log(p(\boldsymbol{y}_i|\boldsymbol{x}_i)),$$

where $(\boldsymbol{x}_i, \boldsymbol{y}_i) \sim p(\boldsymbol{x}, \boldsymbol{y})$ and we average over N samples.

$\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ — conditional pdf $p(y_i|x_i)$ (2)

- The phase has a more involved distribution

$$p_{\mathsf{phase}}(\Delta\phi_i|x_i) = \frac{e^{-\rho_i}}{\sigma_w^2} + \sqrt{\frac{\rho_i}{4\pi}} e^{-\rho_i \sin^2 \Delta\phi_i} \cdot \cos\Delta\phi_i \,\mathrm{erfc}(-\sqrt{\rho_i}\cos\Delta\phi_i),$$

where $\rho_i = \frac{|x_i|^2}{\sigma_w^2}$ and $\Delta\phi_i = y_{i,\mathsf{phase}} - \arg(x_i) \in [0, 2\pi)$.

- We estimate the entropy $\mathrm{h}(\boldsymbol{y}|\boldsymbol{x})$ with **Monte Carlo (MC) integration**

$$\begin{aligned}\mathrm{h}(\boldsymbol{y}|\boldsymbol{x}) &= -\iint p(\boldsymbol{x},\boldsymbol{y}) \log(p(\boldsymbol{y}|\boldsymbol{x}))\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{x} \\ &= -\mathrm{E}_{p(\boldsymbol{x},\boldsymbol{y})}[\log(p(\boldsymbol{y}|\boldsymbol{x}))] \simeq -\frac{1}{N}\sum_{i=1}^{N} \log(p(\boldsymbol{y}_i|\boldsymbol{x}_i)),\end{aligned}$$

where $(\boldsymbol{x}_i, \boldsymbol{y}_i) \sim p(\boldsymbol{x}, \boldsymbol{y})$ and we average over N samples.

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_\mathsf{R} = 1$

- In the MISO case $r = \sum_{i=1}^{N_\mathsf{T}} h_{1i} s_i + w$, and for Gaussian input,

$$r \sim \mathcal{CN}(0, \sigma_r^2), \qquad \sigma_r^2 = \sum_{i=1}^{N_\mathsf{T}} |h_{1i}|^2 \sigma_s^2 + \sigma_w^2.$$

- The norm of r is Rayleigh distributed and the entropy is

$$\mathrm{h}(r)[\text{bits}] = \left(1 + \ln \sqrt{\frac{\sigma_r^2}{4}} + \frac{\gamma}{2}\right) \log_2 e$$

- The phase of r is uniform distributed in $[0, 2\pi)$

$$\mathrm{h}(r)[\text{bits}] = \log_2(2\pi)$$

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_{\mathsf{R}} = 1$

- In the MISO case $r = \sum_{i=1}^{N_{\mathsf{T}}} h_{1i} s_i + w$, and for Gaussian input,

$$r \sim \mathcal{CN}(0, \sigma_r^2), \qquad \sigma_r^2 = \sum_{i=1}^{N_{\mathsf{T}}} |h_{1i}|^2 \sigma_s^2 + \sigma_w^2.$$

- The norm of r is Rayleigh distributed and the entropy is

$$\mathrm{h}(r)[\text{bits}] = \left(1 + \ln \sqrt{\frac{\sigma_r^2}{4}} + \frac{\gamma}{2}\right) \log_2 e$$

- The phase of r is uniform distributed in $[0, 2\pi)$

$$\mathrm{h}(r)[\text{bits}] = \log_2(2\pi)$$

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_{\mathsf{R}} = 1$

- In the MISO case $r = \sum_{i=1}^{N_{\mathsf{T}}} h_{1i} s_i + w$, and for Gaussian input,

$$r \sim \mathcal{CN}(0, \sigma_r^2), \qquad \sigma_r^2 = \sum_{i=1}^{N_{\mathsf{T}}} |h_{1i}|^2 \sigma_s^2 + \sigma_w^2.$$

- The norm of r is Rayleigh distributed and the entropy is

$$\mathrm{h}(r)[\text{bits}] = \left(1 + \ln \sqrt{\frac{\sigma_r^2}{4}} + \frac{\gamma}{2}\right) \log_2 e$$

- The phase of r is uniform distributed in $[0, 2\pi)$

$$\mathrm{h}(r)[\text{bits}] = \log_2(2\pi)$$

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_\mathsf{R} > 1$ (1)

- When $N_\mathsf{R} > 1$, the amplitude is a *correlated multivariate Rayleigh distribution* [Mallik, Miller]. For $N_\mathsf{R} = 2$, $N_\mathsf{R} = 3$

$$p(y_1, y_2) = 4 y_1 y_2 \det(\boldsymbol{S}) e^{-(\boldsymbol{S}_{11} y_1^2 + \boldsymbol{S}_{22} y_2^2)} I_0(2|\boldsymbol{S}_{12}| y_1 y_2),$$

$$\begin{aligned} p(y_1, y_2, y_3) &= 8 y_1 y_2 y_3 \det(\boldsymbol{S}) e^{-(\boldsymbol{S}_{11} y_1^2 + \boldsymbol{S}_{22} y_2^2 + \boldsymbol{S}_{33} y_3^2)} \\ &\quad \sum_{m=0}^{\infty} \left[\varepsilon_m (-1)^m I_m(2|\boldsymbol{S}_{12}| y_1 y_2) I_m(2|\boldsymbol{S}_{23}| y_2 y_3)\right. \\ &\quad \left.\cdot I_m(2|\boldsymbol{S}_{31}| y_3 y_1) \cos(m(\chi_{12} + \chi_{23} + \chi_{31}))\right], \end{aligned}$$

where $\boldsymbol{S} = (\mathrm{E}[\boldsymbol{r}\boldsymbol{r}^\mathrm{H}])^{-1} = (\sigma_s^2 \boldsymbol{H}\boldsymbol{H}^\mathrm{H} + \sigma_w^2 \boldsymbol{I})^{-1}$ and $\chi_{ij} = \arg(\boldsymbol{S}_{ij})$.

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_\mathsf{R} > 1$ (1)

- When $N_\mathsf{R} > 1$, the amplitude is a *correlated multivariate Rayleigh distribution* [Mallik, Miller]. For $N_\mathsf{R} = 2$, $N_\mathsf{R} = 3$

$$p(y_1, y_2) = 4 y_1 y_2 \det(\boldsymbol{S}) e^{-(\boldsymbol{S}_{11} y_1^2 + \boldsymbol{S}_{22} y_2^2)} I_0(2|\boldsymbol{S}_{12}| y_1 y_2),$$

$$\begin{aligned} p(y_1, y_2, y_3) = \; & 8 y_1 y_2 y_3 \det(\boldsymbol{S}) e^{-(\boldsymbol{S}_{11} y_1^2 + \boldsymbol{S}_{22} y_2^2 + \boldsymbol{S}_{33} y_3^2)} \\ & \sum_{m=0}^{\infty} \left[\varepsilon_m (-1)^m I_m(2|\boldsymbol{S}_{12}| y_1 y_2) I_m(2|\boldsymbol{S}_{23}| y_2 y_3) \right. \\ & \left. \cdot I_m(2|\boldsymbol{S}_{31}| y_3 y_1) \cos(m(\chi_{12} + \chi_{23} + \chi_{31})) \right], \end{aligned}$$

where $\boldsymbol{S} = (\mathrm{E}[\boldsymbol{r}\boldsymbol{r}^\mathrm{H}])^{-1} = (\sigma_s^2 \boldsymbol{H}\boldsymbol{H}^\mathrm{H} + \sigma_w^2 \boldsymbol{I})^{-1}$ and $\chi_{ij} = \arg(\boldsymbol{S}_{ij})$.

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_\mathsf{R} > 1$ (2)

- When $N_\mathsf{R} > 1$, the phase is a *correlated multivariate uniform distribution* (uniform marginals). For $N_\mathsf{R} = 2$

$$p(y_1, y_2) = \frac{\det(\boldsymbol{S})}{8\pi^2 S_{11} S_{22}} \left[\frac{1}{1-\lambda^2} - \frac{\lambda \cos^{-1} \lambda}{(1-\lambda^2)^{3/2}} \right]$$

and

$$\lambda = \frac{S_{12}}{\sqrt{S_{11} S_{22}}} \cos(y_1 - y_2 - \chi_{12}).$$

- In both cases we calculate the entropy with MC integration.

$\mathrm{h}(\boldsymbol{y})$ — distribution of $\boldsymbol{y}$, $N_\mathsf{R} > 1$ (2)

- When $N_\mathsf{R} > 1$, the phase is a *correlated multivariate uniform distribution* (uniform marginals). For $N_\mathsf{R} = 2$

$$p(y_1, y_2) = \frac{\det(\boldsymbol{S})}{8\pi^2 S_{11} S_{22}} \left[\frac{1}{1-\lambda^2} - \frac{\lambda \cos^{-1} \lambda}{(1-\lambda^2)^{3/2}} \right]$$

and

$$\lambda = \frac{S_{12}}{\sqrt{S_{11} S_{22}}} \cos(y_1 - y_2 - \chi_{12}).$$

- In both cases we calculate the entropy with MC integration.

Entropy of $\boldsymbol{y}$ in the general case

- In all other cases we estimate the distribution of $\boldsymbol{y}$ numerically

$$p(\boldsymbol{y}_i) = \int p(\boldsymbol{x})p(\boldsymbol{y}_i|\boldsymbol{x})\mathrm{d}\boldsymbol{x} \simeq \frac{1}{M}\sum_{j=1}^{M} p(\boldsymbol{y}_i|\boldsymbol{x}_j)$$

  where $\boldsymbol{x}_j \sim p(\boldsymbol{x})$. The entropy of $\boldsymbol{y}$ is given by

$$\mathrm{h}(\boldsymbol{y}) = -\int p(\boldsymbol{y})\log(p(\boldsymbol{y}))\mathrm{d}\boldsymbol{y} \simeq -\frac{1}{N}\sum_{i=1}^{N}\log(p(\boldsymbol{y}_i))$$

  where $\boldsymbol{y}_i \sim p(\boldsymbol{y})$.

- This method is rather cumbersome with case specific peculiarities (use importance sampling in specific cases).

Entropy of $\boldsymbol{y}$ in the general case

- In all other cases we estimate the distribution of $\boldsymbol{y}$ numerically

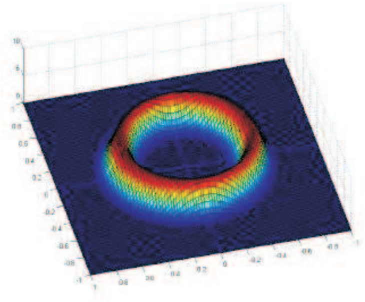
$$p(\boldsymbol{y}_i) = \int p(\boldsymbol{x})p(\boldsymbol{y}_i|\boldsymbol{x})\mathrm{d}\boldsymbol{x} \simeq \frac{1}{M}\sum_{j=1}^{M} p(\boldsymbol{y}_i|\boldsymbol{x}_j)$$

  where $\boldsymbol{x}_j \sim p(\boldsymbol{x})$. The entropy of $\boldsymbol{y}$ is given by

$$\mathrm{h}(\boldsymbol{y}) = -\int p(\boldsymbol{y})\log(p(\boldsymbol{y}))\mathrm{d}\boldsymbol{y} \simeq -\frac{1}{N}\sum_{i=1}^{N}\log(p(\boldsymbol{y}_i))$$

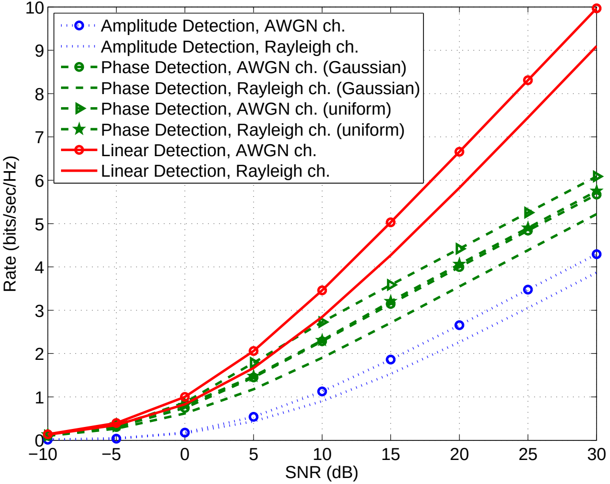
  where $\boldsymbol{y}_i \sim p(\boldsymbol{y})$.
- This method is rather cumbersome with case specific peculiarities (use importance sampling in specific cases).

Example: Amplitude detection, 20 dB SNR, $y = 0.5$



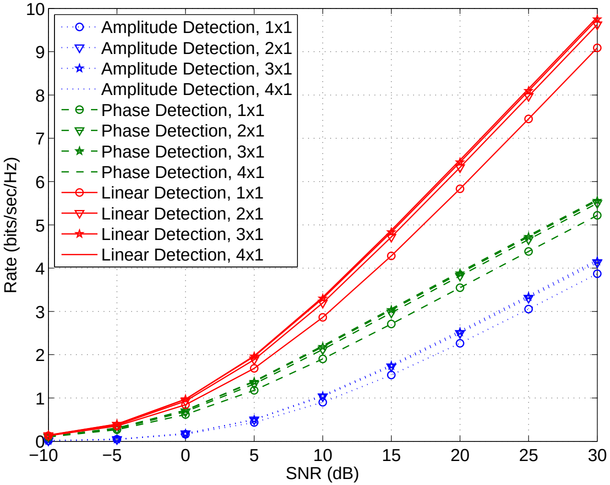
- pdf: $p(y = 0.5|x)$ for different x (x is complex valued).
- If we create samples for x Gaussian distributed, we miss the "important" area.
- Instead, we use an auxiliary function that captures the ring around $y = 0.5$.

SISO System

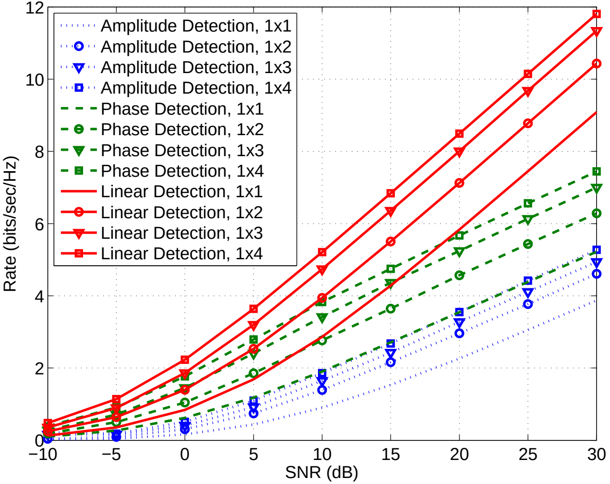


Amplitude Detection, AWGN ch.
Amplitude Detection, Rayleigh ch.
Phase Detection, AWGN ch. (Gaussian)
Phase Detection, Rayleigh ch. (Gaussian)
Phase Detection, AWGN ch. (uniform)
Phase Detection, Rayleigh ch. (uniform)
Linear Detection, AWGN ch.
Linear Detection, Rayleigh ch.
Rate (bits/sec/Hz)
SNR (dB)
0
1
2
3
4
5
6
7
8
9
10
−10
−5
0
5
10
15
20
25
30

MISO System

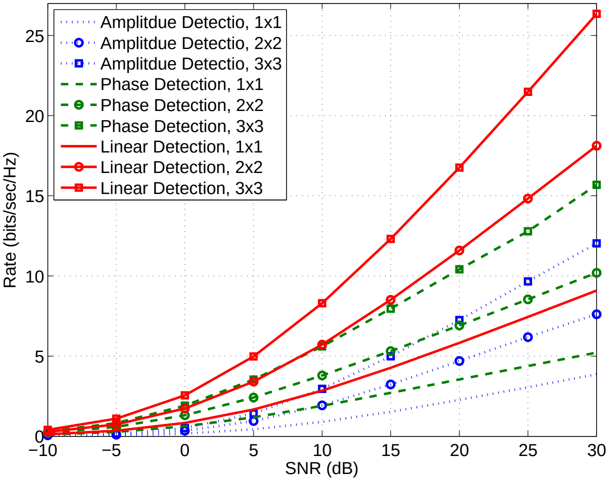

Amplitude Detection, 1x1
Amplitude Detection, 2x1
Amplitude Detection, 3x1
Amplitude Detection, 4x1
Phase Detection, 1x1
Phase Detection, 2x1
Phase Detection, 3x1
Phase Detection, 4x1
Linear Detection, 1x1
Linear Detection, 2x1
Linear Detection, 3x1
Linear Detection, 4x1
Rate (bits/sec/Hz)
SNR (dB)
0
1
2
3
4
5
6
7
8
9
10
−10
−5
0
5
10
15
20
25
30

SIMO System

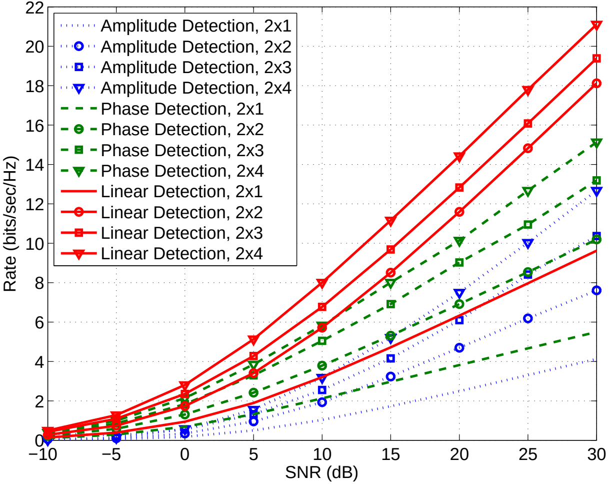


Amplitude Detection, 1x1
Amplitude Detection, 1x2
Amplitude Detection, 1x3
Amplitude Detection, 1x4
Phase Detection, 1x1
Phase Detection, 1x2
Phase Detection, 1x3
Phase Detection, 1x4
Linear Detection, 1x1
Linear Detection, 1x2
Linear Detection, 1x3
Linear Detection, 1x4
Rate (bits/sec/Hz)
SNR (dB)
0
2
4
6
8
10
12
−10
−5
0
5
10
15
20
25
30

MIMO $N \times N$ system



MIMO $2 \times N$ system



Conclusions

- The performance of nonlinear receivers is clearly inferior to linear reception.
- Achievable rates of nonlinear receiver behave in a similar way as linear receivers (SIMO, MISO, etc.).
- Additional receive antennas improve the performance of nonlinear receivers (resolve more dimensions)!
- It may be cheaper to employ more nonlinear receivers than linear ones.

Conclusions

- The performance of nonlinear receivers is clearly inferior to linear reception.
- Achievable rates of nonlinear receiver behave in a similar way as linear receivers (SIMO, MISO, etc.).
- Additional receive antennas improve the performance of nonlinear receivers (resolve more dimensions)!
- It may be cheaper to employ more nonlinear receivers than linear ones.

Thank you for your attention!

Appendix

(M,N) MIMO	Degrees of freedom N_{min}	Real degrees of freedom $2N_{\text{min}}$	Slope of linear detection	Slope of nonlinear detection	Real degrees of freedom $2N_{\text{min}}$ - 1
(1,1)	1	2	2	1	1
(1,2),(2,1),. . .	1	2	2	1	1
(2,2)	2	4	4	2	3
(2,3)	2	4	4	~ 2.4	3
(2,4)	2	4	4	~ 2.8	3