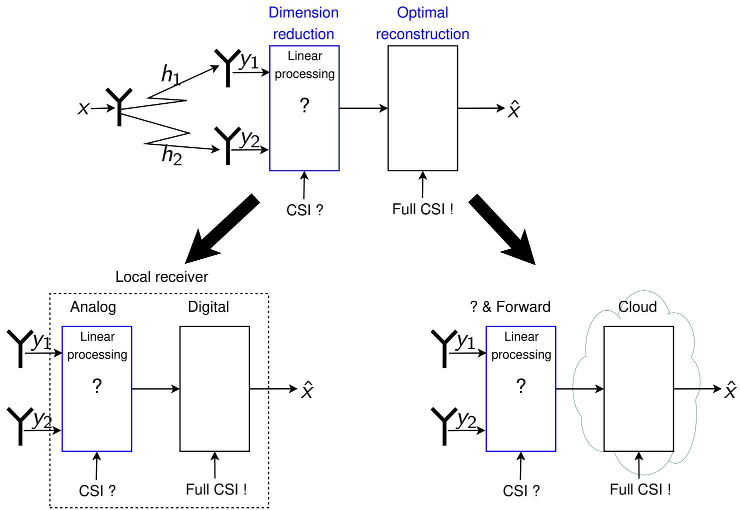
Diversity Combining via Universal Dimension-Reducing Space-Time Transformations

Elad Domanovitz and Uri Erez

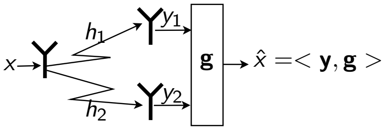
June 18th, 2018
2018 International Symposium on Information Theory

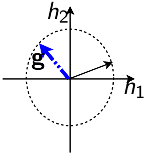
Scenario of interest



What can be guaranteed universally without **any** CSI at the dimension reduction transformation?

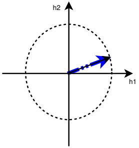
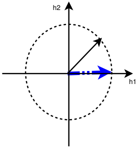
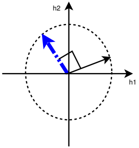
Scenario of interest



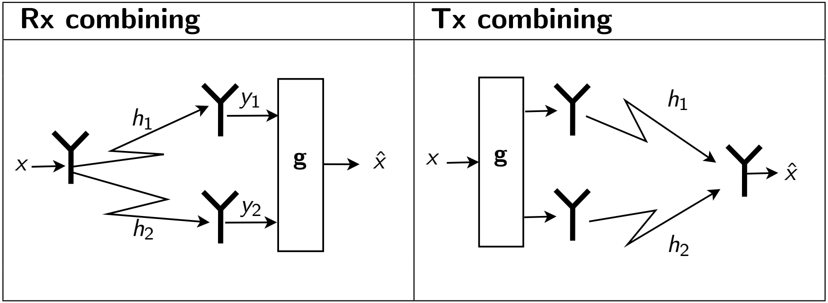


- Denote $\mathbf{y} = \mathbf{h}x + \mathbf{n}$
  - Signal $x \sim \mathcal{CN}(0, 1)$
  - Noise is $\mathbf{n} \sim \mathcal{CN}(0, \mathbf{I})$
- Assume
  - $\|\mathbf{h}\|^2 = const$, for simplicity $const = 1$
  - Receiver projects the received signal: $\hat{x} =< \mathbf{y}, \mathbf{g} >$
  - $\text{SNR} = | < \mathbf{h}, \mathbf{g} > |^2$
- Goal: maximize **worst-case** $\Longrightarrow \text{SNR}^* = \min_{\mathbf{h}} \left(\max_{\mathbf{g}=f(CSI)} | < \mathbf{h}, \mathbf{g} > |^2 \right)$

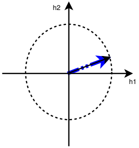
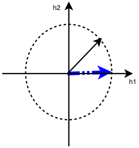
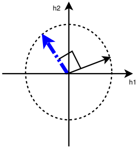
Scenario of interest

CSI	**Projection**	$\min\limits_{\mathbf{h}} \left(\max\limits_{\mathbf{g}=\mathbf{f}(\mathbf{CSI})} < \mathbf{y}, \mathbf{g} > \right)$	
Full ($\mathbf{h}$)	$\mathbf{g} = \mathbf{h}$ (MRC)	$\mathrm{SNR}(\mathbf{h}) = \Vert\mathbf{h}\Vert^2$ $\mathrm{SNR}^* = 1$	
1-bit $\left(\lvert h_1\rvert \overset{?}{\lessgtr} \lvert h_2\rvert\right)$	$\mathbf{g} = \begin{cases} [1\ 0]^T & \lvert h_1\rvert \geq \lvert h_2\rvert \\ [0\ 1]^T & \mathrm{O/W} \end{cases}$ (Selection)	$\mathrm{SNR}(\mathbf{h}) = \max(\lvert\mathbf{h_1}\rvert^2, \lvert\mathbf{h_2}\rvert^2)$ $\mathrm{SNR}^* = \frac{1}{2}$	
None	**?**	$\mathrm{SNR}(\mathbf{h}) = \lvert < \mathbf{h}, \mathbf{g} > \rvert^2$ $\mathrm{SNR}^* = 0$	

Is there something to learn from the dual problem?

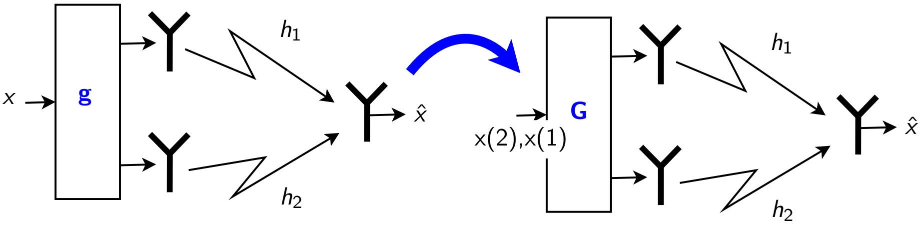


Performance of dual

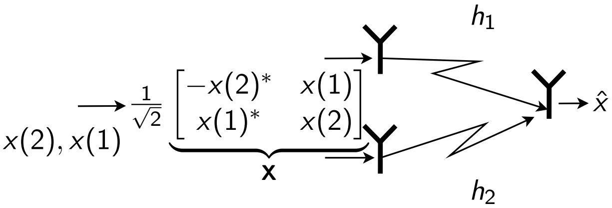
CSI	**Projection**	$\min\limits_{\mathbf{h}} \left(\max\limits_{\mathbf{g}=\mathbf{f}(\mathbf{CSI})} < \mathbf{y}, \mathbf{g} > \right)$	
Full ($\mathbf{h}$)	$\mathbf{g} = \mathbf{h}$ (Beamforming)	$\text{SNR}(\mathbf{h}) = \lVert\mathbf{h}\rVert^2$ $\text{SNR}^* = 1$	
1-bit $\left(\lvert h_1\rvert \overset{?}{\lessgtr} \lvert h_2\rvert\right)$	$\mathbf{g} = \begin{cases} [1\ 0]^T & \lvert h_1\rvert \geq \lvert h_2\rvert \\ [0\ 1]^T & \text{O/W} \end{cases}$ (Selection)	$\text{SNR}(\mathbf{h}) = \max(\lvert\mathbf{h_1}\rvert^2, \lvert\mathbf{h_2}\rvert^2)$ $\text{SNR}^* = \frac{1}{2}$	
None	**?**	$\text{SNR}(\mathbf{h}) = \lvert < \mathbf{g}, \mathbf{h} > \rvert^2$ $\text{SNR}^* = 0$	

Space-time codes to the rescue

- No matter what direction we choose, $\mathrm{SNR}^*(\mathbf{h}) = 0$
- **So we change the rules of the game**
- Assuming channel is fixed over multiple symbols $\Longrightarrow$ Unitary space-time codes
  - Still linear but over two or more time instances
- Recall Alamouti modulation



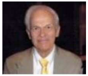
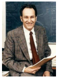
Alamouti modulation



- $$\begin{bmatrix} y(1) \\ y(2)^* \end{bmatrix} = \frac{1}{\sqrt{2}} \underbrace{\begin{bmatrix} h_1 & h_2 \\ -h_2^* & h_1^* \end{bmatrix}}_{\|\mathbf{h}\|\mathbf{H}_{\text{eff}}(h_1,h_2)} \begin{bmatrix} x(1) \\ x(2) \end{bmatrix} + \begin{bmatrix} n(1) \\ n(2) \end{bmatrix}$$
- $\mathbf{H}_{\text{eff}}(h_1, h_2)$ is an **orthonormal** matrix for **any** h_1, h_2: $\mathbf{H}_{\text{eff}}(h_1, h_2)\mathbf{H}_{\text{eff}}(h_1, h_2)^H = \mathbf{I}$
- Using an estimation of $\mathbf{H}_{\text{eff}}(h_1, h_2) \Longrightarrow \quad \hat{x} = \mathbf{H}_{\text{eff}}^H \mathbf{y} = \frac{\|\mathbf{h}\|}{\sqrt{2}} x + \mathbf{n}'$

$$\text{SNR}(\mathbf{h}) = \frac{\|\mathbf{h}\|^2}{2}, \quad \text{SNR}^* = \frac{1}{2}$$

Going back to Rx scenario

	Rx	**Tx**
Full CSI	MRC: $\mathrm{SNR}(\mathbf{h}) = \lVert\mathbf{h}\rVert^2$ $\mathrm{SNR}^* = 1$ 	Beamforming: $\mathrm{SNR}(\mathbf{h}) = \lVert\mathbf{h}\rVert^2$ $\mathrm{SNR}^* = 1$ 
1-bit CSI	Antenna selection: $\mathrm{SNR}(\mathbf{h}) = \max(\lvert h_1\rvert^2, \lvert h_2\rvert^2)$ $\mathrm{SNR}^* = \frac{1}{2}$  	Antenna selection: $\mathrm{SNR}(\mathbf{h}) = \max(\lvert h_1\rvert^2, \lvert h_2\rvert^2)$ $\mathrm{SNR}^* = \frac{1}{2}$ 
No CSI	?	Alamouti: $\mathrm{SNR}(\mathbf{h}) = \frac{\lVert\mathbf{h}\rVert^2}{2}$ $\mathrm{SNR}^* = \frac{1}{2}$ 

- We're missing a counterpart for Alamouti modulation
- Once the question is defined, the answer is quite evident...

So what is $\mathbf{G}$ in case of Alamouti?

- Alamouti modulation (complex): $\mathbf{X} = \frac{1}{\sqrt{2}} \begin{bmatrix} -x(2)^* & x(1) \\ x(1)^* & x(2) \end{bmatrix}$
- Can be written over the reals as:

$$\underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}}_{\mathbf{G}^T} \underbrace{\begin{bmatrix} x_R(1) \\ x_I(1) \\ x_R(2) \\ x_I(2) \end{bmatrix}}_{\mathbf{x}}$$

- Note - this operation amounts to **dimension expansion** $(4 \longrightarrow 8)$
- We want the other way around - **dimension reduction** $(8 \longrightarrow 4)$...

Linear universal combining at the receiver

- Signal received at antenna $i = 1, 2$, at time t : $s_i(t) = h_i x(t) + n_i(t)$
- Stack two receive symbols $\begin{bmatrix} s_1(1) & s_1(2) \\ s_2(1) & s_2(2) \end{bmatrix}$
- Apply $\mathbf{y} = \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\mathbf{G}} \begin{bmatrix} s_{1R}(1) \\ s_{1I}(1) \\ s_{2R}(1) \\ s_{2I}(1) \\ s_{1R}(2) \\ s_{1I}(2) \\ s_{2R}(2) \\ s_{2I}(2) \end{bmatrix}$
- Note that $\mathbf{G}^T$ is Alamouti modulation over the reals (**dimension expansion**$\rightarrow$ **dimension reduction**)

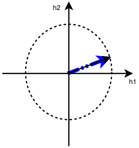
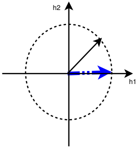
Linear universal combining at the receiver

- The following holds : $$\mathbf{y} = \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{U}(h_1, h_2)\mathbf{x} + \mathbf{Gn}$$ $$= \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{U}(h_1, h_2)\mathbf{x} + \mathbf{n}'$$

  where $\mathbf{U}(h_1, h_2) = \frac{1}{\|\mathbf{h}\|}\begin{bmatrix} h_{1R} & -h_{1I} & h_{2R} & -h_{2I} \\ h_{1I} & h_{1R} & -h_{2I} & -h_{2R} \\ h_{2R} & -h_{2I} & -h_{1R} & h_{1I} \\ h_{2I} & h_{2R} & h_{1I} & h_{R1} \end{bmatrix}$

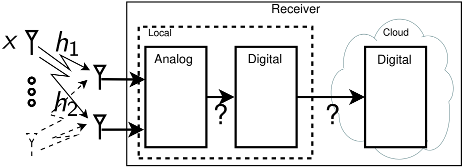
- $\mathbf{U}(h_1, h_2)$ is an **orthonormal** matrix for any h_1, h_2: $\mathbf{U}^T(h_1, h_2)\mathbf{U}(h_1, h_2) = \mathbf{I}$
- Using an estimation of $\mathbf{U} \Longrightarrow$ $$\hat{\mathbf{x}} = \mathbf{U}^T(h_1, h_2) \cdot \mathbf{y}$$ $$= \frac{\|\mathbf{h}\|}{\sqrt{2}}\mathbf{x} + \mathbf{n}''$$
- Remark: Channel needs to be estimated only at the end terminal

Rx combining

CSI	Projection	$\min_{\mathbf{h}} \left(\max_{\mathbf{g}=\mathbf{f}(\mathbf{CSI})} < \mathbf{y}, \mathbf{g} > \right)$	
Full ($\mathbf{h}$)	$\mathbf{g} = \mathbf{h}$ (MRC)	$\text{SNR}(\mathbf{h}) = \|\mathbf{h}\|^2$ $\text{SNR}^* = 1$	
1-bit $\left(\|h_1\| \overset{?}{\lessgtr} \|h_2\|\right)$	$\mathbf{g} = \begin{cases} [1\ 0]^T & \|h_1\| \geq \|h_2\| \\ [0\ 1]^T & \text{O/W} \end{cases}$ (Selection)	$\text{SNR}(\mathbf{h}) = \max(\|\mathbf{h_1}\|^2, \|\mathbf{h_2}\|^2)$ $\text{SNR}^* = \frac{1}{2}$	
None	$\mathbf{G}$ (Universal combining)	$\text{SNR}(\mathbf{h}) = \frac{\|\mathbf{h}\|^2}{2}$ $\text{SNR}^* = \frac{1}{2}$	

But can we think of any application?

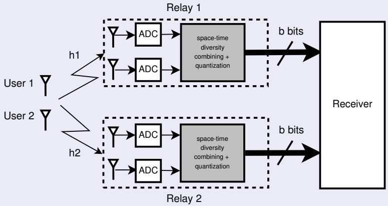
- We don't like loose ends...
- Why not make use of full CSI? After all, we're talking receiver side...
- Justification for 1-bit CSI (selection)
  - Reduce number of analog to digital converters (ADC)
  - Reduce number of bits in fronthaul
- Why is selection (1-bit CSI) not good enough?
  What is the benefit of universality?
  - **Minor**: in traditional scenarios, selection has some drawbacks (complexity, delay, errors)
  - **Major**: in case of multi-user detection, selection fails



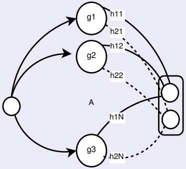
Potential applications - **multi user**

- Reduce the number of ADCs

- “Dumb” (low latency / enhanced diversity) relaying

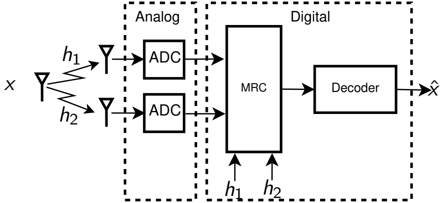


- Ultra-reliable, low-latency communication (ad-hoc netwroking)

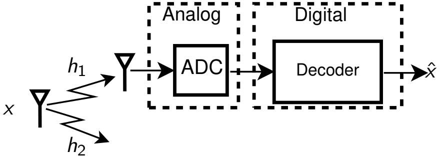


- Time-domain sub-Nyquist sampling

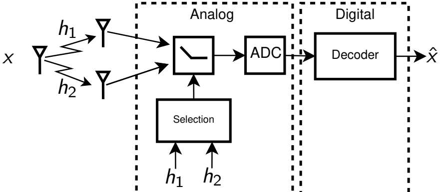
Application 1: ADC



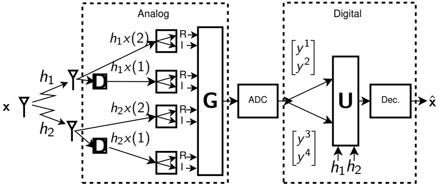
(a) MRC - $h_{\text{eff}} = \|\mathbf{h}\|$



(b) Arbitrary selection - $h_{\text{eff}} = h_1$



(c) Selection - $h_{\text{eff}} = \max(|h_1|, |h_2|)$



(d) Universal combining - $h_{\text{eff}} = \frac{\|\mathbf{h}\|}{\sqrt{2}}$

Application 1: reduce number of ADC, single user

Comparison of the mutual information $I_{\text{scheme}}(P) = \log\left(1 + h^2_{\text{eff,scheme}} P\right)$ attained by each of the schemes

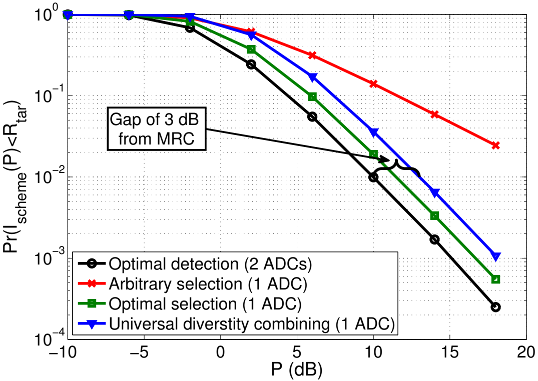


Figure: 2×1 i.i.d. Rayleigh fading channel, with a target rate of $R_{\text{tar}} = 2$ bits per complex symbol.

Application 1: reduce number of ADC, multi user

Comparison of the symmetric-capacity attained by each of the schemes

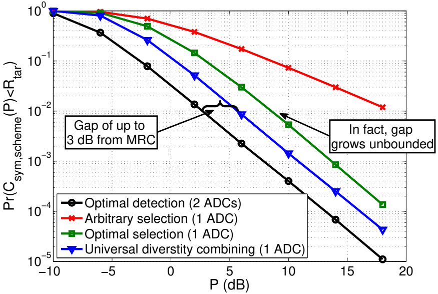


Figure: 8 transmitters, a common receiver equipped with two antennas. All users transmit at an equal rate R_{tar} such that $8R_{\text{tar}} = 2$ bits per complex symbol.

Theorem 1

For a Rayleigh fading $2 \times N$ MIMO-MAC, for any fixed (symmetric) target rate, at asymptotic high SNR, the universal combining scheme suffers a power penalty factor no greater than 2 with respect to an optimal receiver.

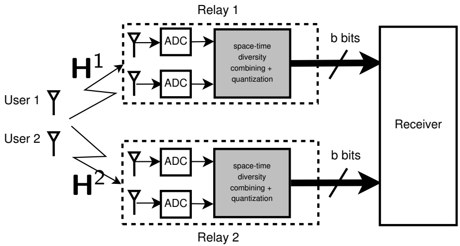
Outlook

- What about more than 2 Rx antennas?
  - Extensions to Alamouti: OSTBC
  - Straightforward implementation fails (rate-1 complex orthogonal designs do not exist beyond the case of two antennas)
  - The problem: Effective channel is non-square $\Longrightarrow$ not invertible
    - Extension 1: dither
    - Extension 2: quasi orthogonal codes
    - Other?
- Every Tx technique involving OSTBC can be considered ...
- What about more than a single antenna per user?
  - Is there a dual to Golden/Perfect codes??

Thank you for your attention

Application 2: "dumb" relaying

- "Dumb" relay = can only apply channel-independent linear processing followed by scalar quantization
- The output is fed into a rate-constrained bit pipe



- The signal received at relay $i = 1, 2$ and antenna $j = 1, 2$ is given by $s_j^i(t) = h_{j1}^i \cdot x_1(t) + h_{j2}^i \cdot x_2(t) + n_j^i(t)$.
- The corresponding channel matrix of relay i: $\mathbf{H}^i = \begin{bmatrix} h_{11}^i & h_{12}^i \\ h_{21}^i & h_{22}^i \end{bmatrix}$.

Application 2: “dumb” relaying

- The signal passed to the cloud from relay i:
  $\mathbf{y}^i = \mathbf{U}(h^i_{11}, h^i_{21})\mathbf{x}_1 + \mathbf{U}(h^i_{12}, h^i_{22})\mathbf{x}_2 + \mathbf{n}'^i$,
- Effective channel:

$$\begin{bmatrix} \mathbf{y}^1 \\ \mathbf{y}^2 \end{bmatrix} = \underbrace{\left[\begin{array}{c|c} \mathbf{U}(h^1_{11}, h^1_{21}) & \mathbf{U}(h^1_{12}, h^1_{22}) \\ \hline \mathbf{U}(h^2_{11}, h^2_{21}) & \mathbf{U}(h^2_{12}, h^2_{22}) \end{array}\right]}_{\mathcal{G}} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{n}'^1 \\ \mathbf{n}'^2 \end{bmatrix}.$$