Precoded Integer-Forcing Equalization: A Matrix Double-Sided Diophantine Approximation Problem

Uri Erez
Joint work with Elad Domanovitz, Anatoly Khina and Or Ordentlich

July 5, 2016, University of York

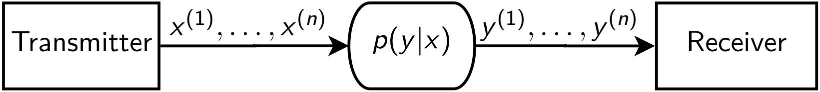
Part I

A Bit of Information Theory and Coding

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory		
SISO		
MIMO		

Unicast: Point-to-Point Communication



Memoryless channel

$$p\left(y^{(1)}, \ldots, y^{(n)} \middle| x^{(1)}, \ldots, x^{(n)}\right) = \prod_{t=1}^{n} p\left(y^{(t)} \middle| x^{(t)}\right)$$

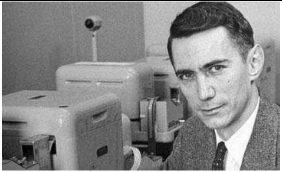
Channel capacity [Shannon '48]

Best achievable rate over memoryless channel $p(y|x)$:

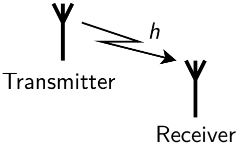
$$C = \max_{p(x)} I(x; y)$$

- Maximization over all admissible input distributions $p(x)$

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory	 ✓	
SISO		
MIMO		

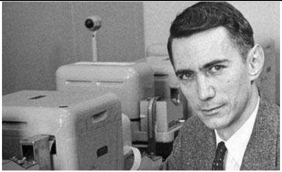
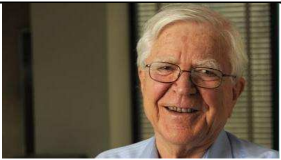
Gaussian Single-Input Single-Output (SISO) Unicast



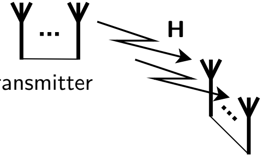
$$y = hx + z$$

- x – Input, power constraint $Px = \mathbb{E}(x^2) \leq \text{SNR}$
- y – Output
- h – Channel gain
- z – White Gaussian noise $\sim \mathcal{CN}(0, 1)$
- Optimal communication rate (capacity):
  $C = \log(1 + \text{SNR}|h|^2)$
- **Good *practical* codes that approach capacity are known!**

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory	 ✓	
SISO	 ✓	
MIMO		

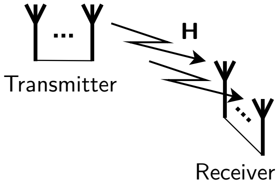
Multiple-Input Multiple-Output (MIMO) Unicast



$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{z}$$

- $\mathbf{x} \in \mathbb{C}^{N_t}$ subject to power constraint: $\mathbb{E}\|\boldsymbol{x}\|^2 \leq N_t \cdot \mathrm{SNR}$
- $\mathbf{y} \in \mathbb{C}^{N_r}$ is the channel output vector
- $\mathbf{H}$ is an $N_r \times N_t$ complex channel matrix
  - $\rightarrow$ Fixed over entire block length
  - $\rightarrow$ $\mathbf{H}_{k\ell}$ – Gain from transmit-antenna ℓ to receive-antenna k
- $\boldsymbol{z} \sim \mathcal{CSCN}(0, \mathbf{I})$
- Capacity: $C = \max_{\mathbf{C}_{\boldsymbol{x}}} \log\det\left(\mathbf{I} + \mathbf{H}\mathbf{C}_{\boldsymbol{x}}\mathbf{H}^{\dagger}\right)$

MIMO Unicast



$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$$

Capacity

$$C = \max_{\mathbf{C_x}} \log\det\left(\mathbf{I} + \mathbf{H}\mathbf{C_x}\mathbf{H}^\dagger\right)$$

- But how is this rate achieved in practice?

What Do We Mean by Practical?

(1) Capacity is closely approached

(2) Black box approach: Reduce MIMO to SISO

- Using simple signal processing (linear algebra):
  - linear operations
  - (later also alow modulo operation)
- Over resulting SISO channels: Use any ‘off-the-shelf” standard encoder and decoder (e.g., turbo/LDPC/convolutional code)
- Gap-to-capacity dictated by gap-to-capacity of SISO codes

Singular-Value Decomposition (SVD) Scheme [Telatar '99]

- $\mathbf{H} = \boldsymbol{U}\Sigma\boldsymbol{V}^{\dagger}$, $\boldsymbol{U}$ and $\boldsymbol{V}$ – unitary
- Tx applies $\boldsymbol{V}$ and Rx applies $\boldsymbol{U}^{\dagger}$
- $$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma_{N-1} & 0 \\ 0 & \cdots & 0 & 0 & \sigma_N \end{pmatrix} \Rightarrow \begin{matrix} y_1 = \sigma_1 x_1 + z_1 \\ y_2 = \sigma_2 x_2 + z_2 \\ \vdots \\ y_N = \sigma_N x_N + z_N \end{matrix}$$
- Results in parallel scalar sub-channels (each sub-channel has a different SNR)
- Allocate power P_i to channel $i \to C_i = \log(1 + P_i \sigma_i^2)$.
- Use standard SISO code with appropriate rate over each sub-channel

Degrees of Freedom

DoF

$$\text{DoF} = \lim_{\mathsf{SNR}\to\infty} \frac{C(\mathsf{SNR})}{\log(\mathsf{SNR})}$$

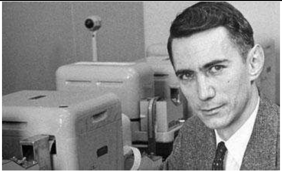
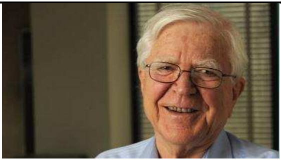
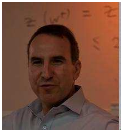
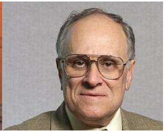
At high SNR

Setting $P_i = \mathsf{SNR}$ is (DoF sense) close to optimal

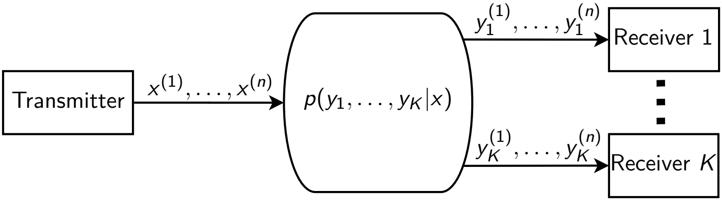
$$\Rightarrow C(\mathsf{SNR}) = \sum_i \log\left(1 + \mathsf{SNR}\sigma_i^2\right)$$
$$\sim (\# \quad \text{non-zero} \quad \text{singular} \quad \text{values}) \cdot \log(\mathsf{SNR})$$

$\Rightarrow$ DoF$=$ $\#$ of non-zero singular values$\leq \min(N_r, N_t) \log(\mathsf{SNR})$

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory	 ✓	
SISO	 ✓	
MIMO	  ✓	

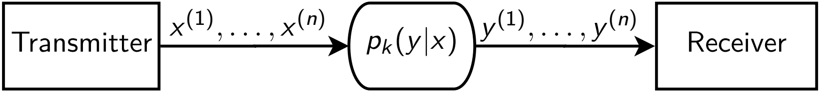
Multicast: Communication over a Compound Channel



Physical-layer multicast

- Transmit same message to K receivers: $y_1, \ldots, y_K$
- All receivers recover message with negligible error probability

Multicast: Communication over a Compound Channel



Compound channel

- K possible channel realizations: $\{p_k(y|x) | k = 1, \ldots K\}$
- Transmitter does not know k
- Error probability is negligible for all $p_k(y|x)$ simultaneously

Multicast: Communication over a Compound Channel

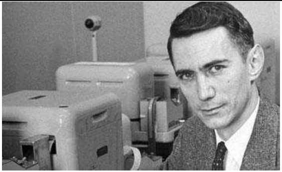
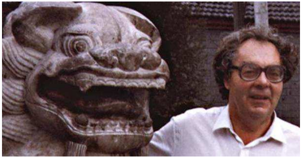
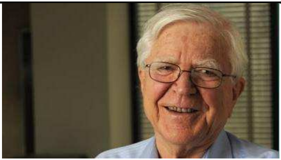
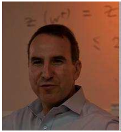
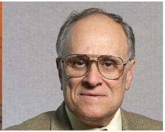
Compound channel / multicast capacity
[Dobrushin '59][Blackwell-Breiman-Thmoasian '59][Wolfowitz '60]

Best achievable rate over K-user memoryless channel $\{p(y_i|x)\}$:

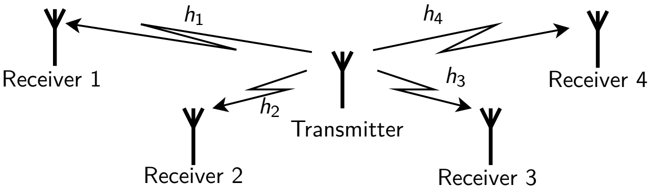
$$C = \max_{p(x)} \min_{i=1,\ldots,K} I(x; y_i)$$

- Maximization over all admissible input distribution $p(x)$
- Minimization over all users

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory	 ✓	 ✓
SISO	 ✓	
MIMO	  ✓	

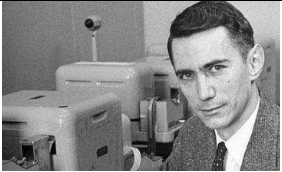
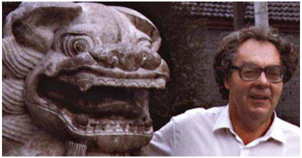
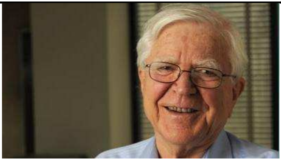
SISO Multicast



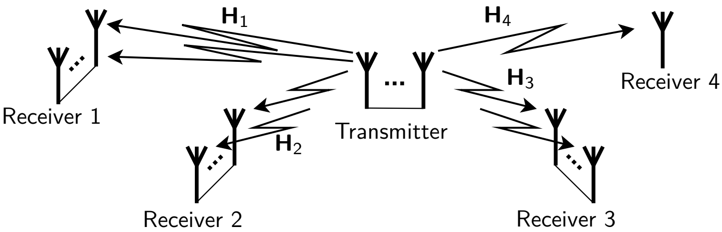
$$y_i = h_i x + z_i \quad i = 1, \ldots, K$$

- x – Input subject to power constraint: $P_x = \mathbb{E}(x^2) \leq \mathrm{SNR}$
- y_i – Output of user i
- h_i – Channel gain to user i
- z_i – White Gaussian noise $\sim \mathcal{CN}(0, 1)$
- Capacity: $C = \min_i \log(1 + \mathrm{SNR}|h_i|^2)$

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

	Unicast	Multicast
Theory	 ✓	 ✓
SISO	 ✓	 ✓
MIMO	 ✓	

Gaussian MIMO Multicast



$$\boldsymbol{y}_i = \mathbf{H}_i \boldsymbol{x} + \boldsymbol{z}_i \quad i = 1, \ldots, K$$

- $\boldsymbol{x}$ – $N_t \times 1$ subject to power constraint: $\mathbb{E}\|\boldsymbol{x}\|^2 \leq N_t \cdot \mathrm{SNR}$
- $\boldsymbol{y}_i$ – Output vector of user i
- $\mathbf{H}_i$ – Channel matrix to user i
- $\boldsymbol{z}_i$ – White Gaussian noise vector $\sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$
- "Closed loop" (Full channel knowledge everywhere)

Optimal Achievable Rate (Capacity)

Multicast capacity

$$C = \max_{\mathbf{C}_{\boldsymbol{x}}} \min_{i=1,\ldots,K} \log\det\left(\mathbf{I} + \mathbf{H}_i \mathbf{C}_{\boldsymbol{x}} \mathbf{H}_i^{\dagger}\right)$$

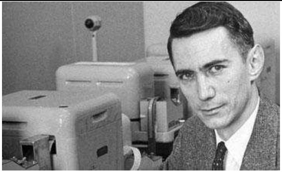
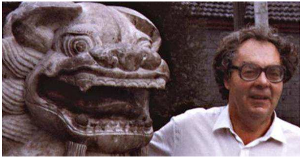
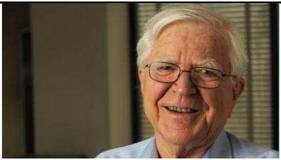
- Optimization over covariance matrices $\mathbf{C}_{\boldsymbol{x}}$ satisfying the power constraint

Achievable rate with isotropic input

- Setting $\mathbf{C}_{\boldsymbol{x}} = \mathrm{SNR} \cdot \mathbf{I}$ does incur a large penalty. Results in:

$$C_{\mathrm{WI}} = \min_{i=1,\ldots,K} \log\det\left(\mathbf{I} + \mathrm{SNR}\,\mathbf{H}_i \mathbf{H}_i^{\dagger}\right)$$

SISO/MIMO Unicast/Multicast Communication: Orientation and State of the Art

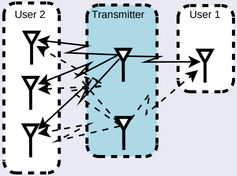
	Unicast	Multicast
Theory	 ✓	 ✓
SISO	 ✓	 ✓
MIMO	 ✓	 ?

Part II

Goal: Achieving MIMO Multicast Capacity

Objective

- Can we find a scheme that is:
  - Practical
    - Linear complexity in the block length
    - Uses off-the-shelf SISO codes
  - Has provable good performance guarantees
  - Universal: Is good for all channels with same white-input mutual information (WI-MI)
- Does such a scheme exist?
- Universal $\implies$ needs to deal with DoF mismatch



Compound MIMO Channel Model

- The mutual information of the compound channel is maximized by a Gaussian input with covariance matrix $\mathbf{C}_{\boldsymbol{X}}$

$$C = \max_{\mathbf{C}_{\boldsymbol{X}} : \operatorname{Tr} \mathbf{Q} \leq N_t \mathrm{SNR}} \log\det\left(\mathbf{I}_{N_r \times N_r} + \mathbf{H}_c \mathbf{C}_{\boldsymbol{X}} \mathbf{H}_c^{\dagger}\right)$$

- For ease of notation we set $\mathrm{SNR} = 1 \implies$ we "absorb" the SNR into the channel matrix: $\mathbf{H}_c = \mathbf{H}_c \sqrt{\mathrm{SNR}}$

- Taking $\mathbf{C}_{\boldsymbol{X}} = I_{N_t \times N_t}$ the isotropic "white-input" (WI) mutual information is given by

$$C_{\mathrm{WI}} = \log\det\left(\mathbf{I}_{N_r \times N_r} + \mathbf{H}_c \mathbf{H}_c^{\dagger}\right)$$

- Define the set

$$\mathbb{H}(C_{\mathrm{WI}}) = \left\{\mathbf{H}_c \in \mathbb{C}^{N_r \times N_t} : \log\det\left(I + \mathbf{H}_c \mathbf{H}_c^{\dagger}\right) = C_{\mathrm{WI}}\right\},$$

Multicast via Generalization of SVD-based Scheme?

$$\mathbf{H}_1^{\text{eff}} = \boldsymbol{U}_1 \Sigma_1 \boldsymbol{V}_1^{\dagger}$$
$$\mathbf{H}_2^{\text{eff}} = \boldsymbol{U}_2 \Sigma_2 \boldsymbol{V}_2^{\dagger}$$

- Precoding matrix $\boldsymbol{V}_i$ depends on the channel matrix $\mathbf{H}_i^{\text{eff}}$
- But $\boldsymbol{V}$ is shared by all users!
- Cannot be used for multi-user case ☹

Diagonal Matrices

Even if all matrices are diagonal $\Rightarrow$ **Rate bottleneck problem!**

Candidate Scheme: Integer Forcing (Zhan et al. 2010)

- Consider the (SU) channel

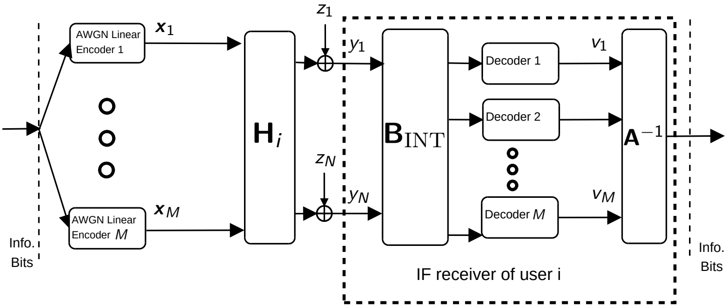
$$\mathbf{H} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

- At high SNR linear receiver front-end *inverts* the channel (ZF) thus resulting in *noise amplification*

$$\mathbf{H}^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

- Can we avoid noise amplification?
- IF idea: If all streams are coded with **same** linear code $\Longrightarrow$ signal at each antenna is a **valid codeword**
- However, normal channels do not consist only of integers
- Integer Forcing (IF) equalizes to "closest channel" with integer entries

Integer-Forcing Equalization: Review



- Information bits are fed into M encoders, each of which uses the same scalar (SISO) *AWGN linear code*
- Linear equalization matrix $\mathbf{B}_{\text{INT}}$ is applied, such that the resulting equivalent channel $\mathbf{A} = \mathbf{B}_{\text{INT}}\mathbf{H}$ is approximately an integer matrix
- Integer matrix $\Longrightarrow$ the output of the channel (without noise) after applying a modulo operation is a valid codeword

Notational Intermezzo: Complex to Real Conversion...

- Recall that any equation of the form

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{x}_c + \mathbf{z}_c$$

over the complex field can be expressed via its real-valued representation,

$$\begin{bmatrix} \mathrm{Re}(\mathbf{y}_c) \\ \mathrm{Im}(\mathbf{y}_c) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(\mathbf{H}_c) & -\mathrm{Im}(\mathbf{H}_c) \\ \mathrm{Im}(\mathbf{H}_c) & \mathrm{Re}(\mathbf{H}_c) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{x}_c) \\ \mathrm{Im}(\mathbf{x}_c) \end{bmatrix} + \begin{bmatrix} \mathrm{Re}(\mathbf{z}_c) \\ \mathrm{Im}(\mathbf{z}_c) \end{bmatrix}$$

- We denote

$$\mathbf{H} = \begin{bmatrix} \mathrm{Re}(\mathbf{H}_c) & -\mathrm{Im}(\mathbf{H}_c) \\ \mathrm{Im}(\mathbf{H}_c) & \mathrm{Re}(\mathbf{H}_c) \end{bmatrix}$$

IF Performance

- Using MMSE the equalization matrix is

$$\mathbf{B}_{\mathrm{INT}} = \mathbf{A}\mathbf{H}^T\left(\mathbf{I}+\mathbf{H}\mathbf{H}^T\right)^{-1}$$

- The effective SNR at the m'th subchannel is

$$\mathsf{SNR}_{\mathrm{eff,m}} = \left(\boldsymbol{a}_m^T(\mathbf{I}+\mathbf{H}^T\mathbf{H})^{-1}\boldsymbol{a}_m\right)^{-1}$$

- Achievable rate (Zhan '10):

$$R_{\mathrm{IF}}(\mathbf{H}) = \max_{\substack{\mathbf{A}\in\mathbb{Z}^{2N_t\times 2N_t}\\ \det\mathbf{A}\neq 0}} \min_{m=1,\ldots,2N_t} 2N_t\frac{1}{2}\log\left(\mathsf{SNR}_{\mathrm{eff,m}}\right)$$

$$= -N_t\log\left(\min_{\substack{\mathbf{A}\in\mathbb{Z}^{2N_t\times 2N_t}\\ \det\mathbf{A}\neq 0}} \max_{m=1,\ldots,2N_t} \boldsymbol{a}_m^T(\mathbf{I}+\mathbf{H}^T\mathbf{H})^{-1}\boldsymbol{a}_m\right)$$

IF Performance

- Consider the singular value decomposition

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^T \Longrightarrow (\mathbf{I} + \mathbf{H}^T\mathbf{H})^{-1} = \mathbf{V}\mathbf{D}^{-1/2}\mathbf{V}^T$$

- We have

$$R_{\mathrm{IF}}^{\mathbf{a}_m}(\mathbf{H}) \triangleq R_{\mathrm{IF}}^{\mathbf{a}_m}(\mathbf{D}, \mathbf{V}) = -\frac{1}{2}\log\left(\|\mathbf{D}^{-1/2}\mathbf{V}^T\mathbf{a}_m\|^2\right)$$

- Let Λ be the lattice spanned by $\mathbf{G} = \mathbf{D}^{-1/2}\mathbf{V}^T$

Definition

Let $\Lambda(\mathbf{G})$ be a lattice spanned by the full-rank matrix $\mathbf{G} \in \mathbb{R}^{K\times K}$. for $k = 1, ..., K$, we define the kth **successive minima** as

$$\lambda_k(\mathbf{G}) \triangleq \inf\left\{r : dim\left(span\left(\Lambda(\mathbf{G}) \cap \mathbb{B}(\mathbf{0}, r)\right)\right) \geq k\right\}$$

where $\mathbb{B}(\mathbf{0}, r) = \left\{x \in \mathbb{R}^K : \|\mathbf{x}\| \leq r\right\}$

- Thus, we have

$$R_{\mathrm{IF}}(\mathbf{D}, \mathbf{V}) = -2N_t\frac{1}{2}\log\left(\lambda_{2N_t}^2(\Lambda)\right) = N_t\log\left(\frac{1}{\lambda_{2N_t}^2(\Lambda)}\right)$$

Bad Channels for IF/Linear Equalization

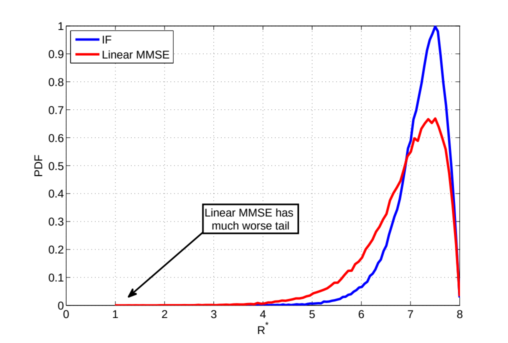
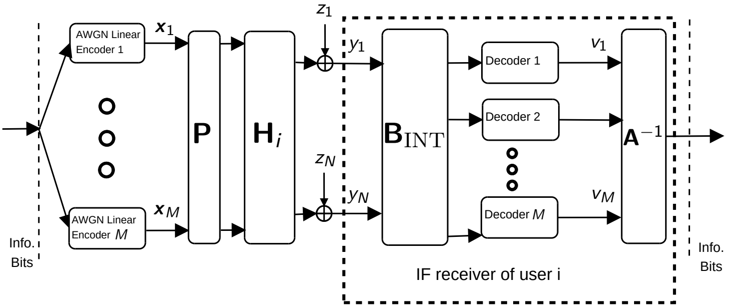


Figure: PDF of 2×2 Rayleigh channels normalized to WI=8 bits

- Worst channel $\mathbf{H}_{\text{worst}} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Combating Bad Channels via Random Precoding



Precoded IF: Apply linear precoding matrix (known to the receiver as well)

Combating Bad Channels via Random Precoding

- What can we do against nature?
- Apply random precoding

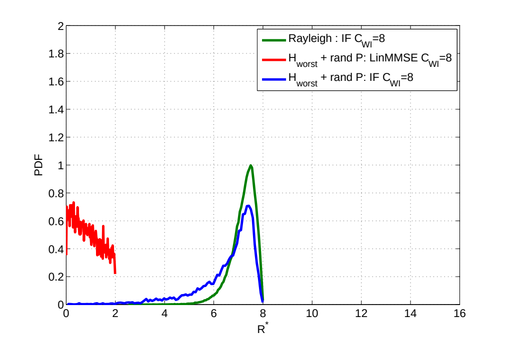


Figure: PDF of Random Unitary Precoding to $\mathbf{H}_{\text{worst}}$

Combating Bad Channels via Random Precoding

- What can we do against nature?
- Apply random precoding

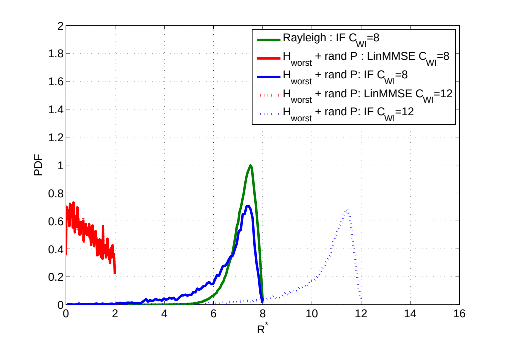


Figure: PDF of Random Unitary Precoding to $\mathbf{H}_{\text{worst}}$

Combating Bad Channels via Random Precoding

- What can we do against nature?
- Apply random precoding

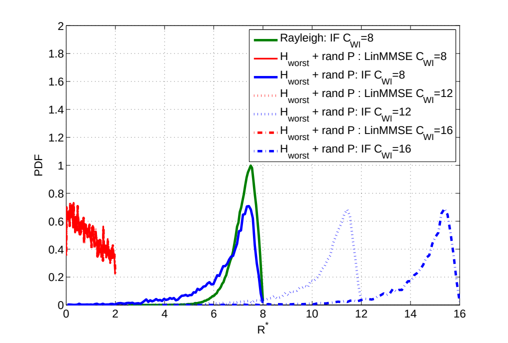


Figure: PDF of Random Unitary Precoding to $\mathbf{H}_{\text{worst}}$

Combating Bad Channels via Random Precoding

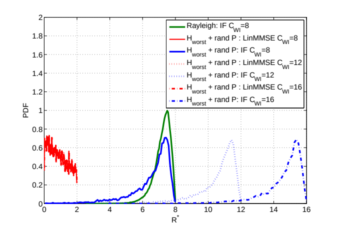


Figure: PDF of Random Unitary Precoding to $\mathbf{H}_{\text{worst}}$

- No precoding can salvage linear Eq. when channel is singular
- IF copes well with "most" singular channels
- But can we prove this...?

Integer-Forcing: Performance Guarantees

Scenario Recap

- Transmission scheme:
  - Rx side - integer forcing equalization
  - Tx side - random space-only unitary linear precoding
- Transmitter knows only the WI-MI of the channel

Performance Guarantees

- For Rayleigh distribution IF achieves optimal receive DMT (Zhan et al. 14')
- With space-time precoding, achieves capacity to within a constant gap (OE 14')

- Q: What is the measure of bad precoding martrices P for IF?
- A: Zero, IF achieves optimal degrees-of-freedom for almost all $N_r \times N_t$ channels (OE 14')

Proof Idea: Diophantine Approximation Problem

- Using Banaszczyk 93', SNR of IF may be expressed via *first* of the successive minima of the dual lattice:

$$\mathrm{SNR}_{\mathrm{eff}} > \frac{1}{4N_t^2} \min_{\boldsymbol{a} \in \mathbb{Z}^{2N_t} \setminus \mathbf{0}} \left(\boldsymbol{a}^T (\mathbf{I} + \mathbf{H}^T \mathbf{H}) \boldsymbol{a} \right)$$

- May be further lower bounded as:

$$\mathrm{SNR}_{\mathrm{eff}} > \frac{1}{4N_t^2} \min_{L=1,2,\ldots} \left(L^2 + \mathrm{SNR} d_{\min}^2(\mathbf{H}, L) \right)$$

where

$$d_{\min}^2(\mathbf{H}, L) = \min_{\boldsymbol{a} \in \mathrm{PAM}(L) \setminus \mathbf{0}} \|\mathbf{H}\boldsymbol{a}\| \quad ,$$

$$\mathrm{PAM}(L) = \{-L, -L+1, \ldots, L-1, L\}$$

Proof Idea: Diophantine Approximation Problem

Key Step

By Hussain and Levesley 13' the set of bad matrices

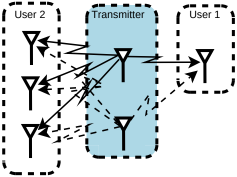
$$\mathcal{H}_0(N_t, N_r) = \Big\{ \mathbf{H} \in [-1,1]^{N_r \times N_t} : d^2_{\min}(\mathbf{H}, L) < L^{-2\left(\frac{N_t+\epsilon}{N_r}-1\right)} \text{ for i.m. } L \in \mathbb{N} \Big\}$$

has zero measure

Implication

Integer Forcing is guaranteed to be good at high SNR for multicasting to any (finite) set of users as long as their channels have same number of DoF.

Beyond DoF Guarantees



- DoF characterization is useful
- But still: two systems can have a different of DoF but same capacity, or same number of DoF but very different capacities
- Would like to show that IF in a non-asymptotic sense
- In other words, we'd like to characterize the PDF/CDF of the achievable rate of IF

Statistical Characterization of Performance of IF

- Characterize universal (over MIMO compound channel) outage probability of IF (practical scheme) for random precoding
- Outage probability is taken w.r.t. *random precoding* ensemble, not w.r.t. to *channel statistics*
- We consider the worst-case performance over the whole compound class of channels

$$\mathbb{H}(C_{\mathrm{WI}}) = \left\{ \mathbf{H}_c \in \mathbb{C}^{N_r \times N_t} : \log\det\left(I + \mathbf{H}_c \mathbf{H}_c^{\dagger} \right) = C_{\mathrm{WI}} \right\},$$

- Hence, in case of randomized integer-forcing our target is to bound

$$P_{\mathrm{out}}^{\mathrm{WC}}\left(C_{\mathrm{WI}}, \Delta C\right) = \sup_{\mathbf{H}_c \in \mathbb{H}(C_{\mathrm{WI}})} P\left(R_{\mathrm{IF}}\left(\mathbf{H}_c \cdot \mathbf{P}\right) < C - \Delta C\right)$$

Universal Bound on Outage

Theorem (1)

For any $N_r \times N_t$ complex channel with WI-MI C, and for $\mathbf{V}_c$ drawn from the CUE (which induces a real-valued precoding matrix $\mathbf{V}$), we have

$$P_{\text{out}}^{\text{WC}}(C, \Delta C) \leq c(N_t) 2^{-\Delta C},$$

where

$$c(N_t) = \left[\left(2 + \frac{\sqrt{2N_t}}{2}\right)^{2N_t} - \left(1 - \frac{\sqrt{2N_t}}{2}\right)^{2N_t}\right] N_t \alpha(N_t)^{N_t} \frac{\pi^{N_t}}{\Gamma(N_t+1)}$$

and

$$\alpha(N_t) = \frac{2N_t + 3}{4} \left(\frac{2}{\pi} \Gamma\left(2 + N_t\right)^{1/N_t}\right)^2.$$

Thus , $c(N_t)$ is a constant that depends only on N_t.

P-IF Performance

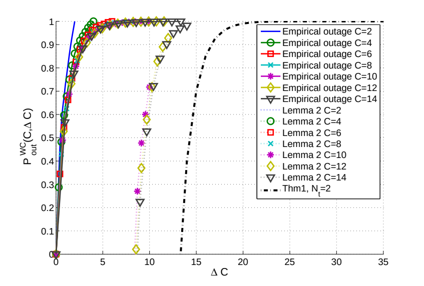


Figure: $N_r \times 2$ complex channel achievable probability

Main Step in Proof: Statistics of Successive Minima

- Recall $\mathbf{o} \sim \mathrm{Unif}(\mathcal{S})$ where $\mathcal{S}$ is the unit sphere. Denote $\mathbf{o}_{\|\mathbf{a}\|} \sim \mathrm{Unif}(\mathcal{S} \cdot \|\mathbf{a}\|)$
- Using Banaszczyk, we saw that

$$P\left(\|\mathbf{D}^{1/2}\mathbf{V}^T\mathbf{a}\| < \sqrt{\beta}\right) = P(\|\mathbf{D}^{1/2}\mathbf{o}_{\|\mathbf{a}\|}\| < \sqrt{\beta})$$

$$P\left(\|\mathbf{D}^{1/2}\mathbf{o}_{\|\mathbf{a}\|}\| < \sqrt{\beta}\right) = \frac{|\mathbf{D}^{1/2}\mathcal{S} \cdot \|\mathbf{a}\| \cap \sqrt{\beta}\mathcal{S}|}{|\mathbf{D}^{1/2}\mathcal{S}\|\mathbf{a}\||} = \frac{\mathrm{CAP}_{\mathrm{ell}}}{L}$$

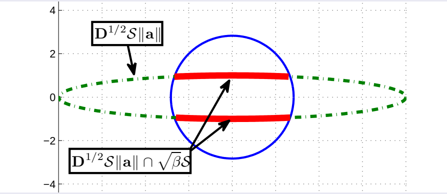


Figure: 2×2 real channel

Outlook

Beyond Random Precoding: "Two-Sided Diophantine Approximation"

- So far, receiver was smart but precoding was dumb (random)
- Can't do much better if all that is known is $C_W I$...
- But what if we have a a finite and known compound class of MIMO channels $\mathbf{H}_k$, $k = 1, \ldots, K$
- We can then do much better by judiciously choosing the precoding matrix $\mathbf{P}$ to help in pre-equalizing to an integer matrix
- Let's consider the case of $K = 1$

Outlook

Beyond Random Precoding: "Two-Sided Diophantine Approximation"

- So far, receiver was smart but precoding was dumb (random)
- Can't do much better if all that is known is the capacity C_{WI}...
- But what if we have a a finite and known compound class of MIMO channels $\mathbf{H}_k$, $k = 1, \ldots, K$
- Can then do much better by judiciously choosing the precoding matrix $\mathbf{P}$ to help in pre-equalizing to an integer matrix
- Let's consider the case of $K = 1$ first

IF Performance as a Function of Precoding Matrix

- Recall the singular value decomposition

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^T \implies (\mathbf{I} + \mathbf{H}^T\mathbf{H})^{-1} = \mathbf{V}\mathbf{D}^{-1/2}\mathbf{V}^T$$

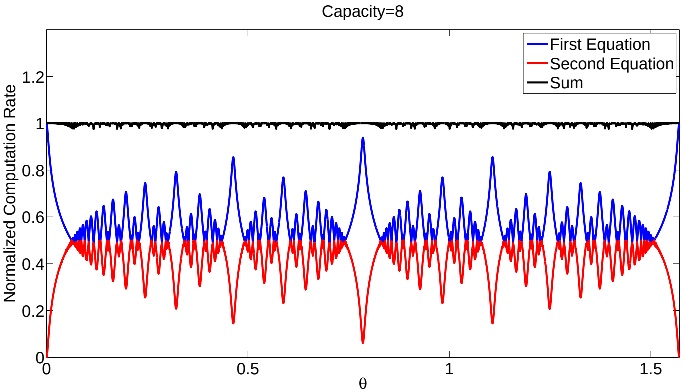
- We saw that

$$R_{\mathrm{IF}}^{\mathbf{a}_m}(\mathbf{H}) \triangleq R_{\mathrm{IF}}^{\mathbf{a}_m}(\mathbf{D}, \mathbf{V}) = -\frac{1}{2}\log\left(\|\mathbf{D}^{-1/2}\mathbf{V}^T\mathbf{a}_m\|^2\right)$$

- Let us plot the two best rates (successive minima) as function of rotation angle θ for the case of a 2×2 singular channel
- Precoding matrix is given by

$$\mathbf{V} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

Performance With Optimized Precoding



And Let Us Conclude with a Movie