

Intelligent Reflecting Surfaces for Wireless Communications: Fundamentals, Designs, and Open Issues

Emil Björnson

Rui Zhang

Professor

KTH Royal Institute of Technology
Sweden



Professor

National University of Singapore
Singapore



Outline of Part 1

Introduction to intelligent reflecting surface (IRS)

- Active and passive beamforming

System modeling and optimization

- Derive system model, optimization of IRS

Narrowband channel modeling

- Pathloss, array responses, correlation

Channel estimation

Wideband modeling and optimization

Competing technologies and weaknesses

My research is sponsored by



Swedish
Foundation for
Strategic
Research

Future Research Leader Grant



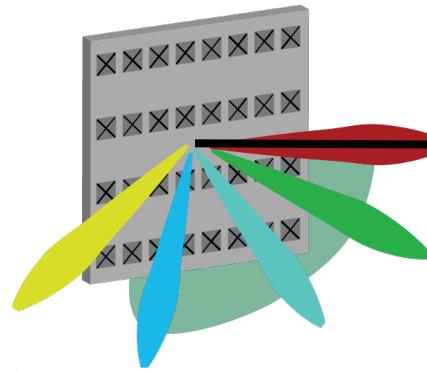
Swedish
Research
Council

Project Grant

INTRODUCTION

Evolution of Wireless Infrastructure

5G: Adaptive multi-user beamforming

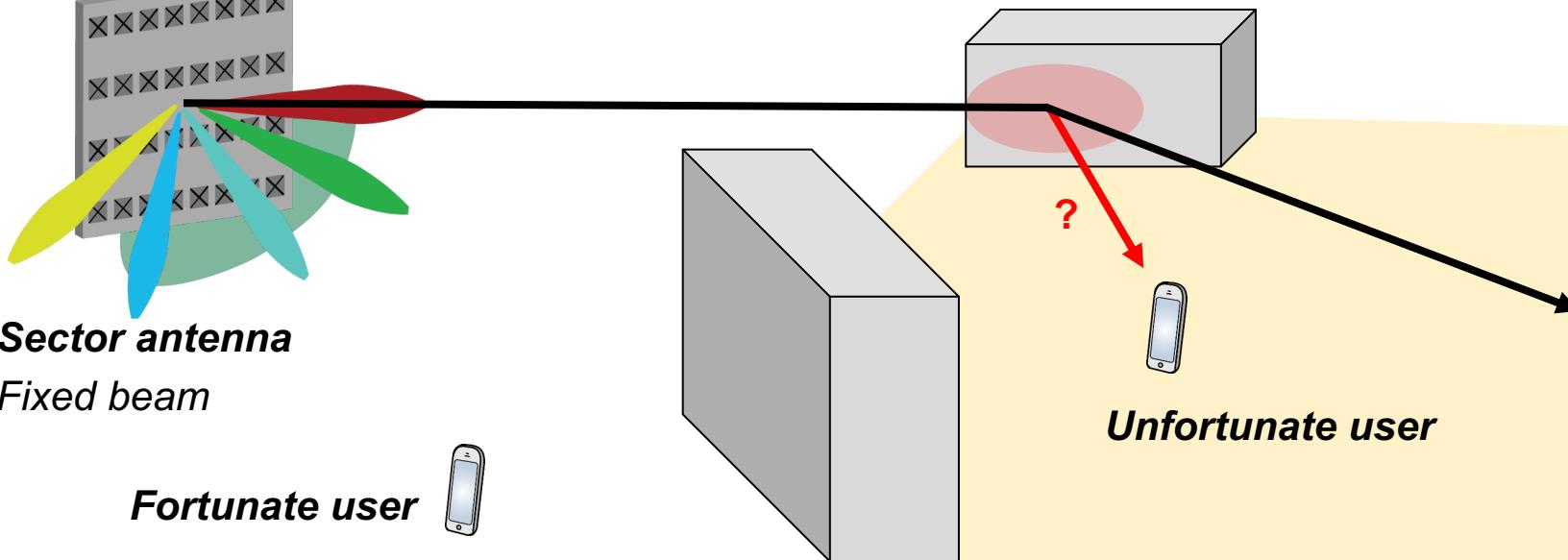


Sector antenna

Fixed beam

Fortunate user

6G: Control objects in the environment?



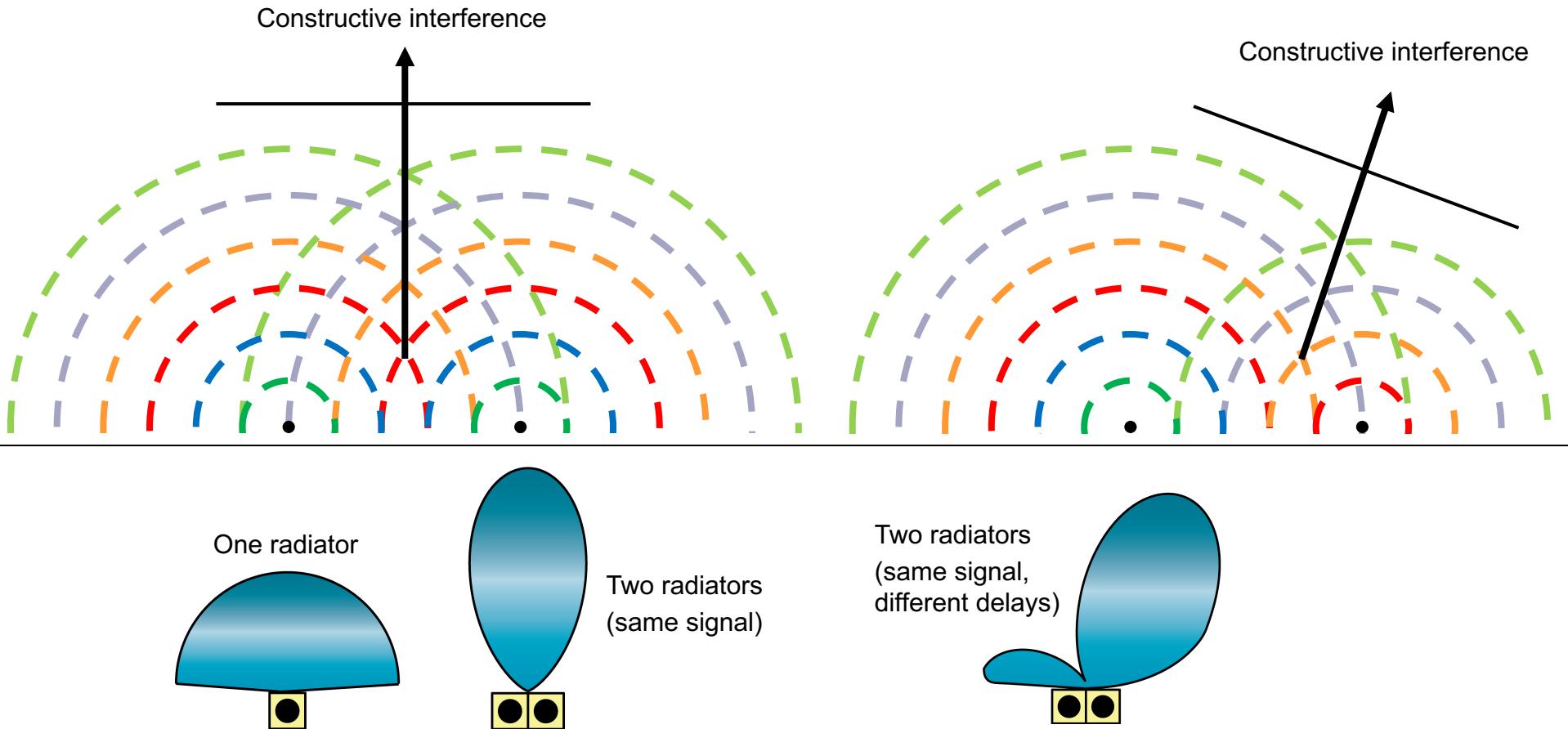
Unfortunate user

Reconfigurable metasurfaces

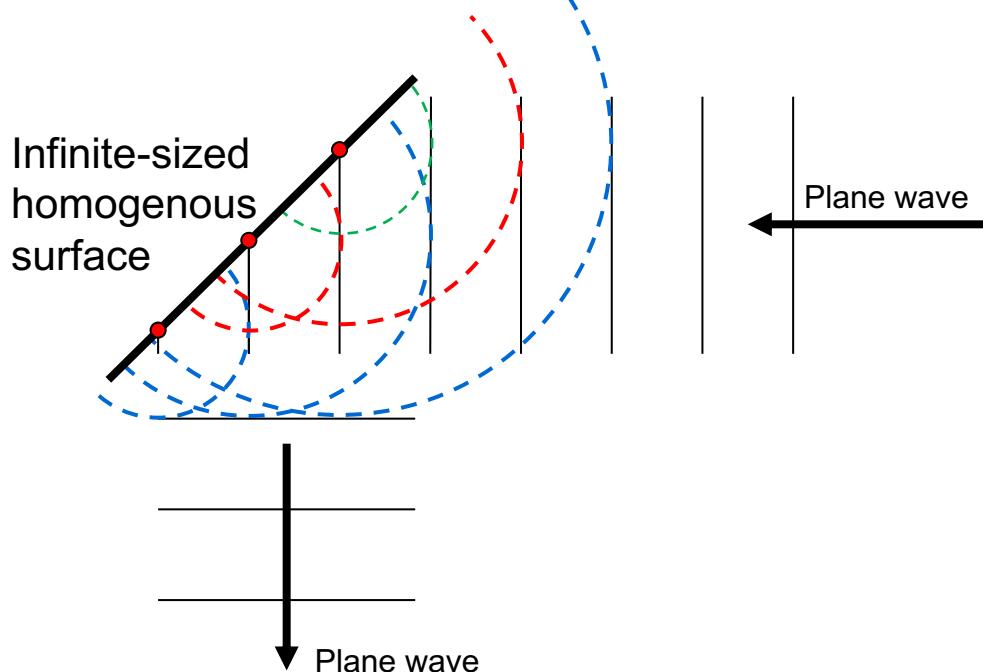
Intelligent reflecting surface (IRS)

Reconfigurable intelligent surface (RIS)

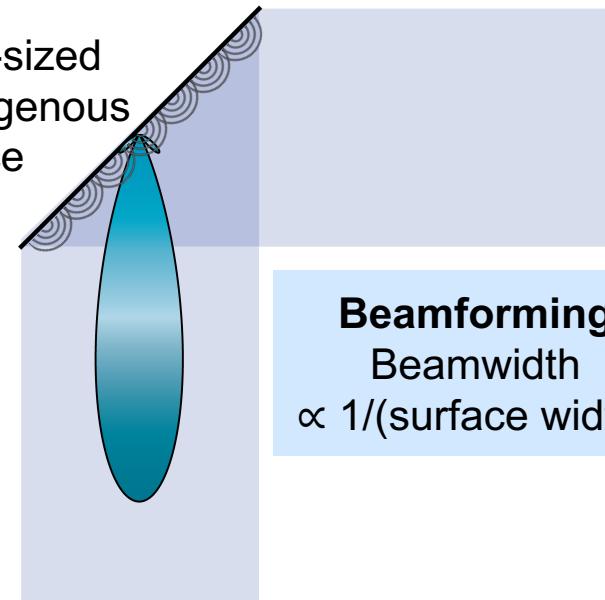
Beamforming: Directivity by Constructive Interference



Interpreting Reflection via the Huygens-Fresnel Principle



Finite-sized
homogenous
surface

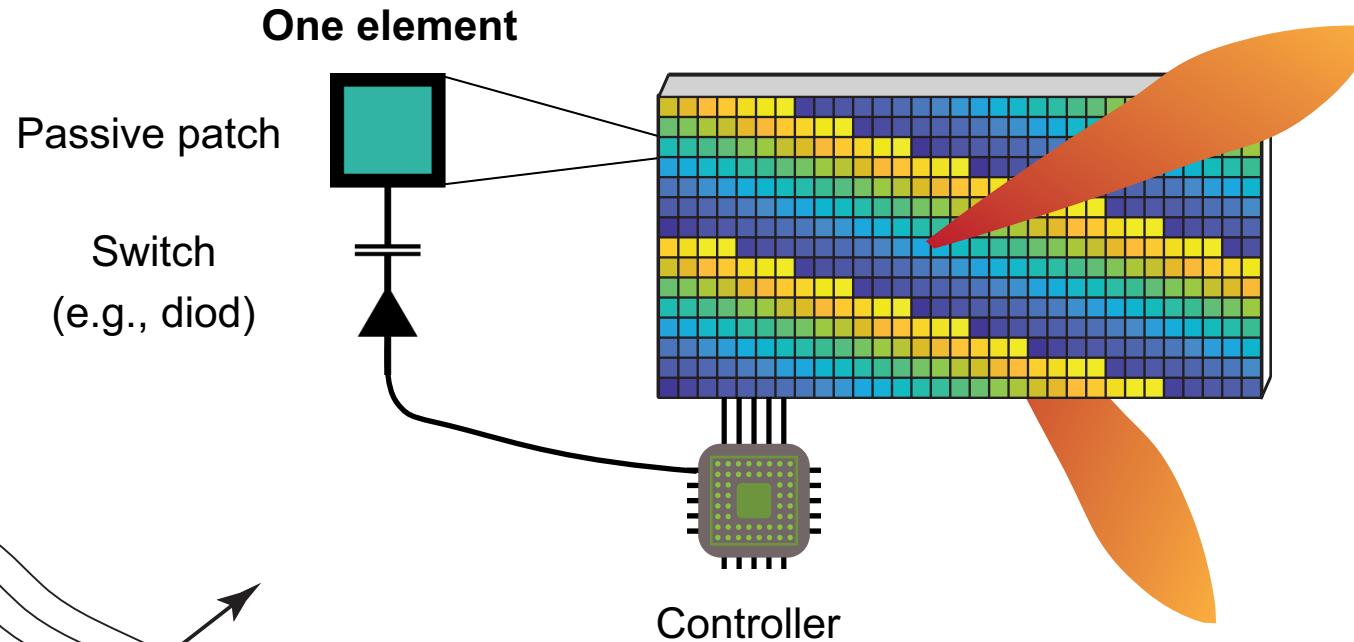


Every point “scatters” a spherical wave:
Constructive interference determines direction

“Passive” Beamforming With IRS

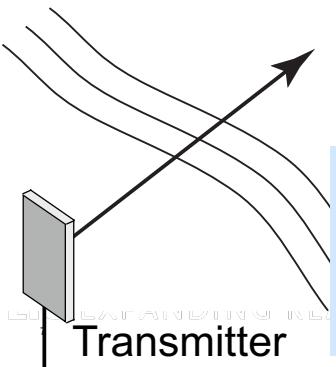


User 2



Phase-shifting pattern

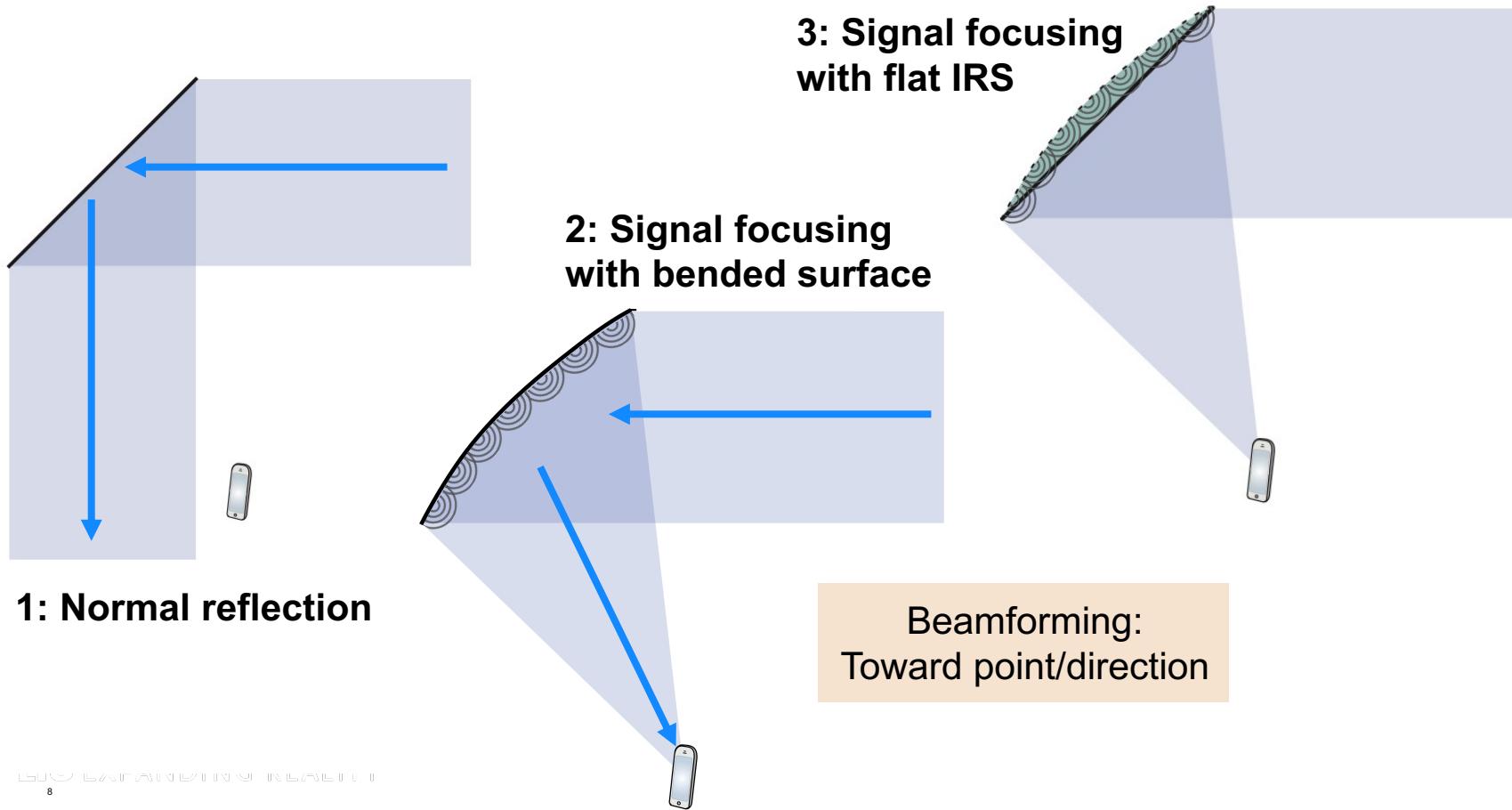
Sub- λ -sized elements, varying impedance
Scattering with varying amplitude and phase



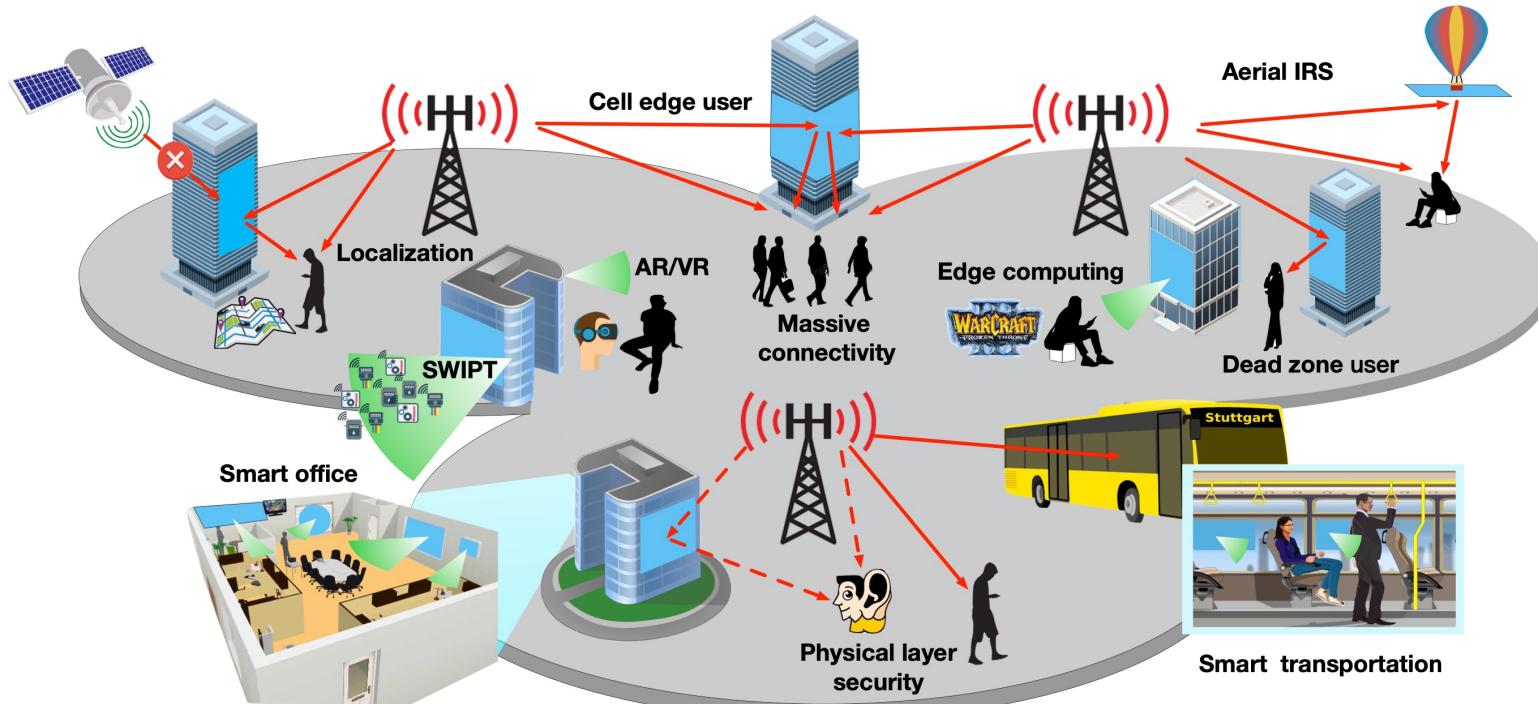
User 1



Geometrical Interpretation at the Global Level

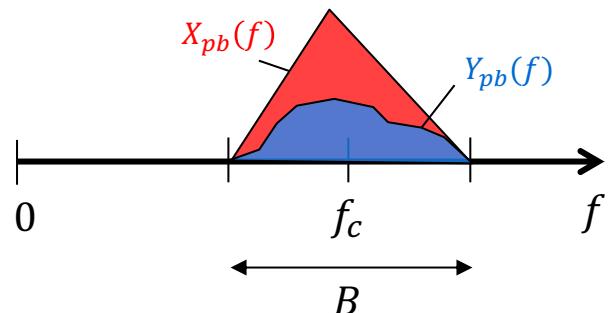
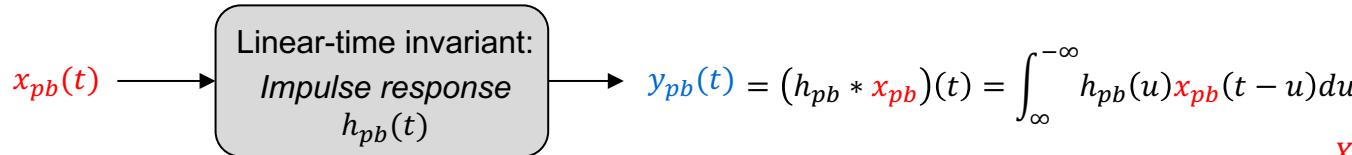


A Wealth of Prospective Use Cases

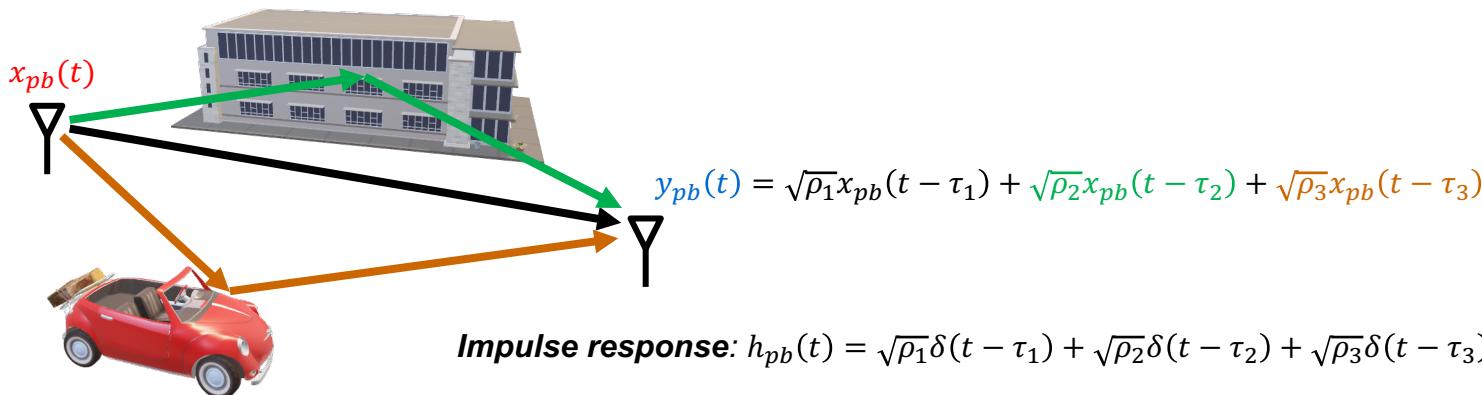


SYSTEM MODELING AND IRS OPTIMIZATION

Introduction to Signals and Systems



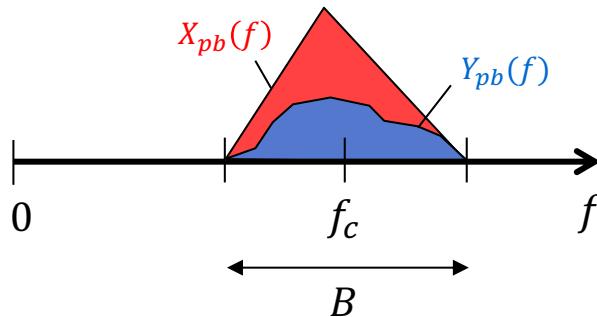
- Communication channels are systems/filters:



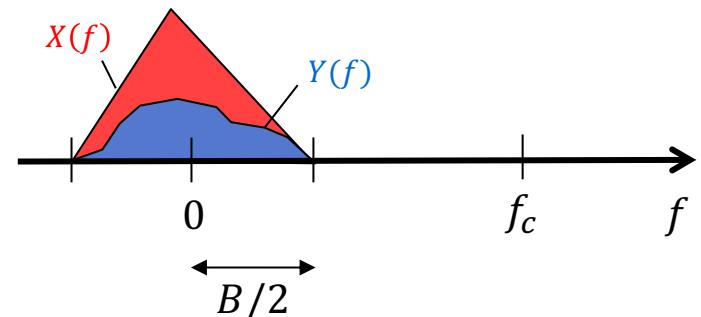
Complex Baseband Representation

- Communication theory is developed for the baseband

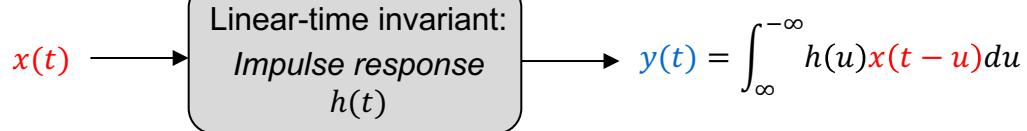
Real passband



Complex baseband



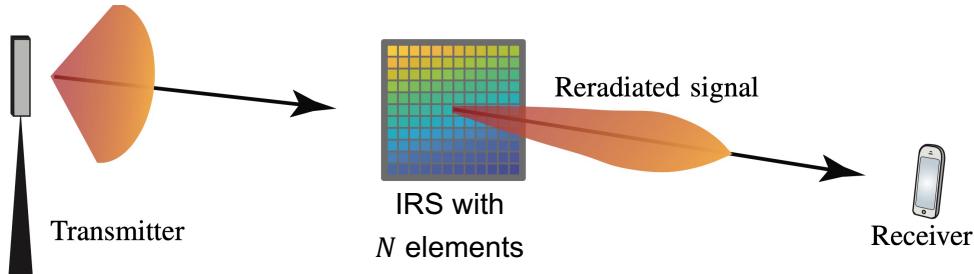
- Connection: $X_{pb}(f) = \frac{X(f-f_c)+X^*(-f-f_c)}{\sqrt{2}}$, $Y_{pb}(f) = \frac{Y(f-f_c)+Y^*(-f-f_c)}{\sqrt{2}}$



Downshifted channel:

$$h(t) = h_{pb}(t)e^{-j2\pi f_c t}$$

Analyzing System With Intelligent Reflecting Surface



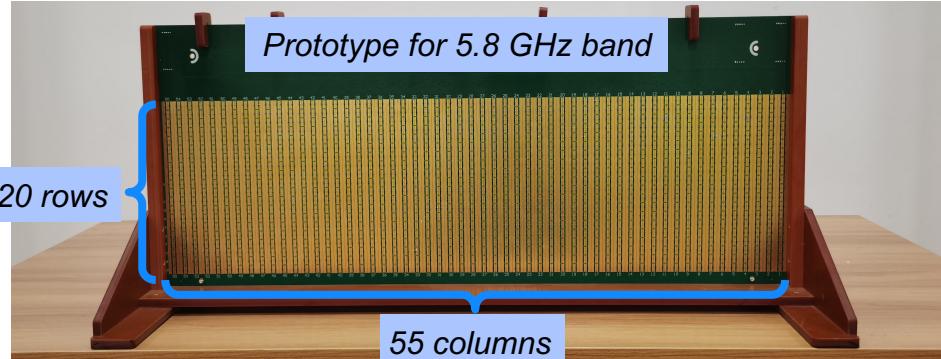
IRS control variables $\theta_1, \dots, \theta_N$ in
 $\vartheta_{n,pb;\theta_n}$

End-to-end impulse response:

$$h_{pb}(t) = h_{d,pb}(t) + \sum_{n=1}^N (b_{n,pb} * \vartheta_{n,pb;\theta_n} * a_{n,pb})(t)$$

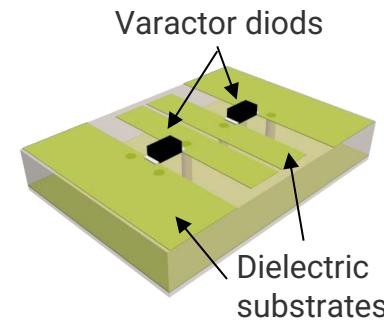
Conventional channel models for
 $a_{n,pb}, b_{n,pb}, h_{d,pb}$

How Will an IRS Element Filter the Signal?

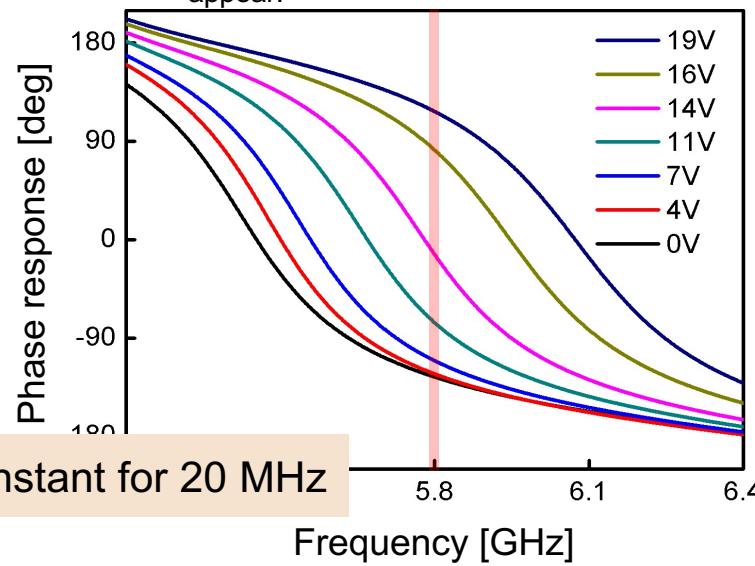
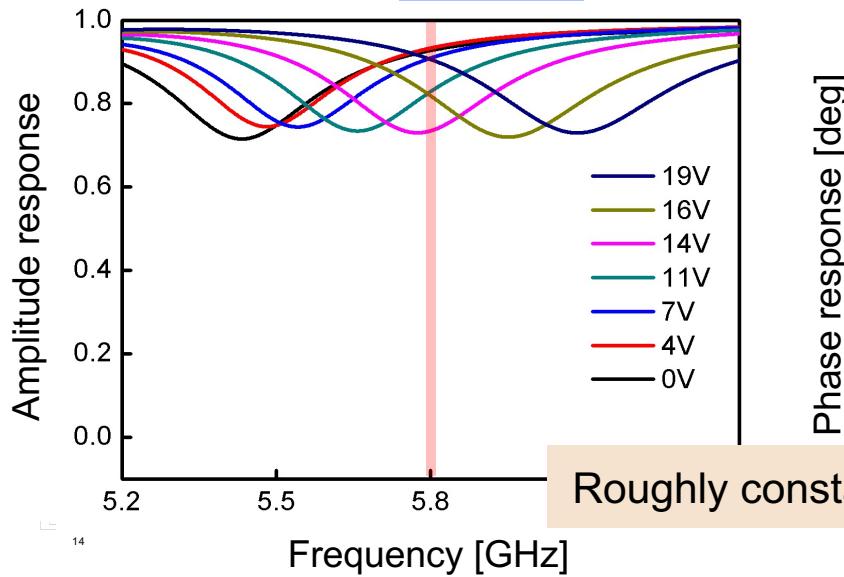


Example: Patch with bias voltage V
Reflection coefficient:

$$\frac{Z_n(V) - Z_0}{Z_n(V) + Z_0}$$

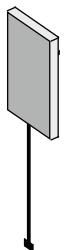


Reference: X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, E. Björnson, "RIS-Aided Wireless Communications: Prototyping, Adaptive Beamforming, and Indoor/Outdoor Field Trials," IEEE Trans. Commun., to appear.

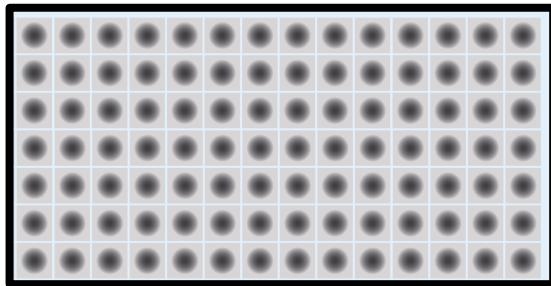


Narrowband System Modelling: N IRS elements

Transmitter



$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}$$



Receiver



$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix}$$

Phase-shifts: $\boldsymbol{\omega}_\theta = \begin{bmatrix} e^{-j\theta_1} \\ \vdots \\ e^{-j\theta_N} \end{bmatrix}$

Configure IRS to maximize $|\gamma|^2$:

Opposite phases in $\boldsymbol{\omega}_\theta$ and $\mathbf{a} \odot \mathbf{b}$:

$$\theta_n = \arg(a_n b_n)$$

$$|\gamma|^2 = \left| \sum_{n=1}^N a_n e^{-j\theta_n} b_n \right|^2 \leq \left| \sum_{n=1}^N |a_n| \cdot |b_n| \right|^2 = \|\mathbf{a} \odot \mathbf{b}\|_1^2$$

End-to-end-channel

$$\gamma = \sum_{n=1}^N a_n e^{-j\theta_n} b_n = (\mathbf{a} \odot \mathbf{b})^T \boldsymbol{\omega}_\theta$$

Element-wise product

Adding a “Direct” Path

$$|\mathbf{h}_d + \gamma|^2 = \left| \mathbf{h}_d + \sum_{n=1}^N a_n e^{-j\theta_n} b_n \right|^2 \leq \left| |\mathbf{h}_d| + \sum_{n=1}^N |a_n| |b_n| \right|^2$$

Match phases to direct path: $\theta_n = \arg(a_n b_n) - \arg(h_d)$

Power scaling behavior

- Assume $|a_n| = \sqrt{\alpha}$, $|b_n| = \sqrt{\beta}$, $|h_d| = \sqrt{\rho}$
- SNR proportional to

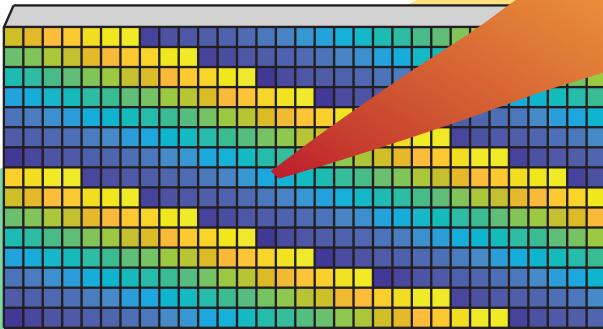
$$|\sqrt{\rho} + N\sqrt{\alpha\beta}|^2 = N^2\alpha\beta + \rho + 2N\sqrt{\alpha\beta\rho}$$

$$\approx \begin{cases} N^2\alpha\beta, & \text{if } N\sqrt{\alpha\beta} \gg \sqrt{\rho} \\ \rho, & \text{if } N\sqrt{\alpha\beta} \ll \sqrt{\rho} \end{cases} \quad \text{“Squaring law”}$$

Why is SNR Proportional to N^2 ?

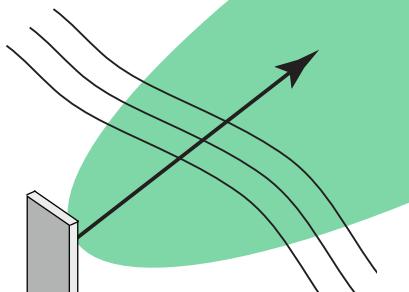


Receiver



Beamforming gain
proportional to IRS size (N)

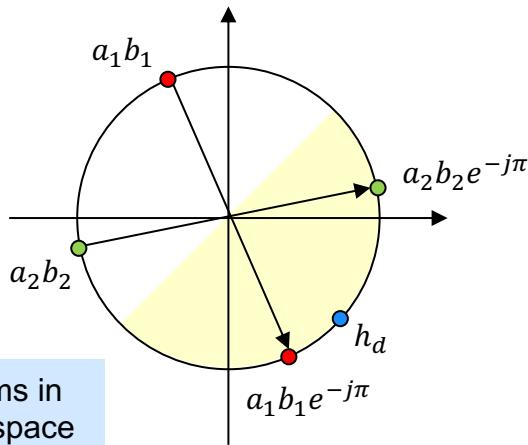
Aperture gain: Collect energy
proportional to IRS size (N)



Beamforming gain: Only for desired signal
Aperture gain: Applies to any signal and noise

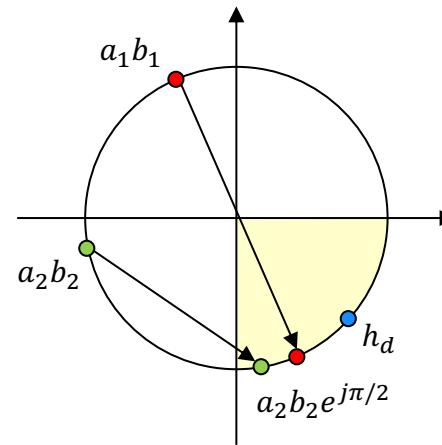
Finite-Resolution Phase Shift Configurations

Two configurations: $\theta_n \in \{0, \pi\}$



Gather all terms
in the same half space

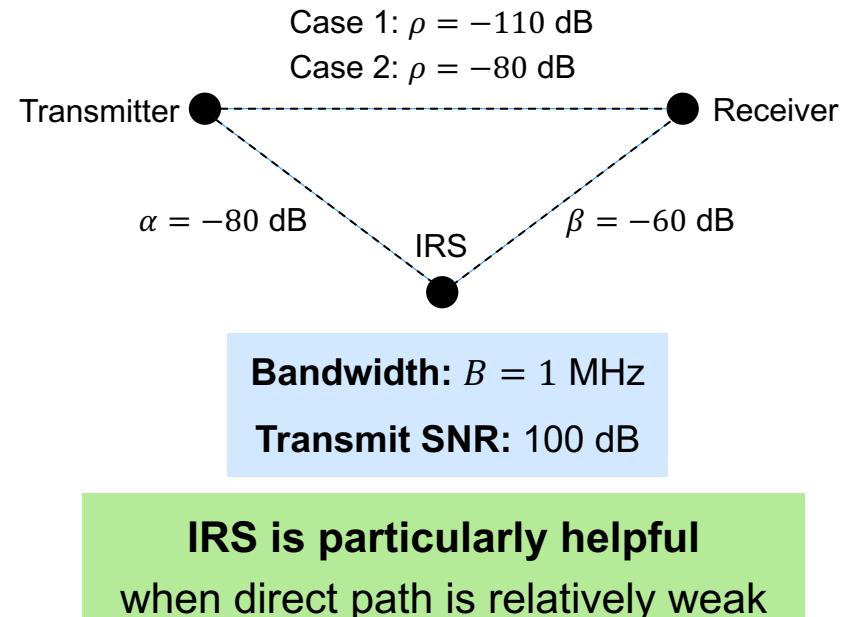
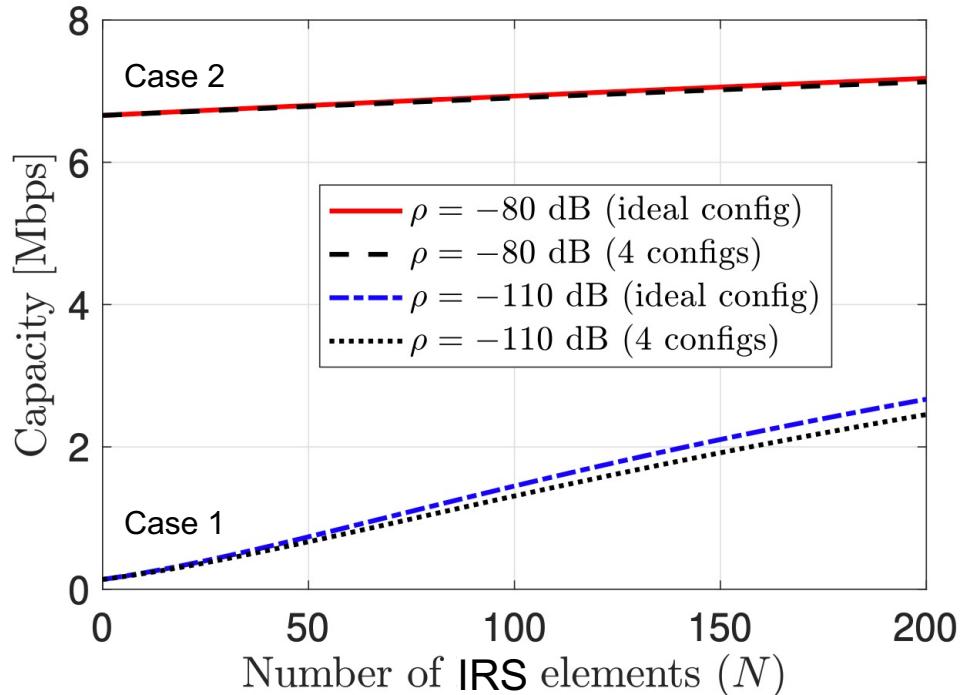
Four configurations: $\theta_n \in \left\{-\frac{\pi}{2}, 0, \frac{\pi}{2}, \pi\right\}$



Gather all terms
in the same
quadrant

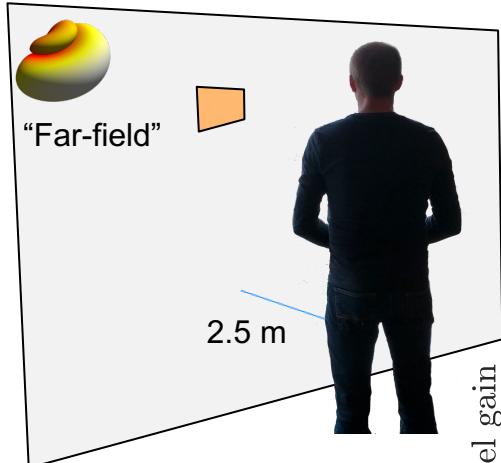
$$|h_d + a_1 b_1 e^{-j\theta_1} + a_2 b_2 e^{-j\theta_2}|^2$$

Basic Performance Benefit



Reference: E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, E. de Carvalho, “A Signal Processing Perspective on Reconfigurable Intelligent Surfaces With Wireless Applications”, Available on arXiv:2102.00742.

What Happens Asymptotically?

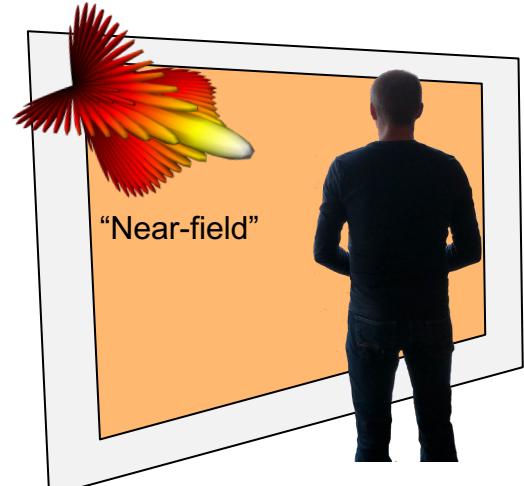
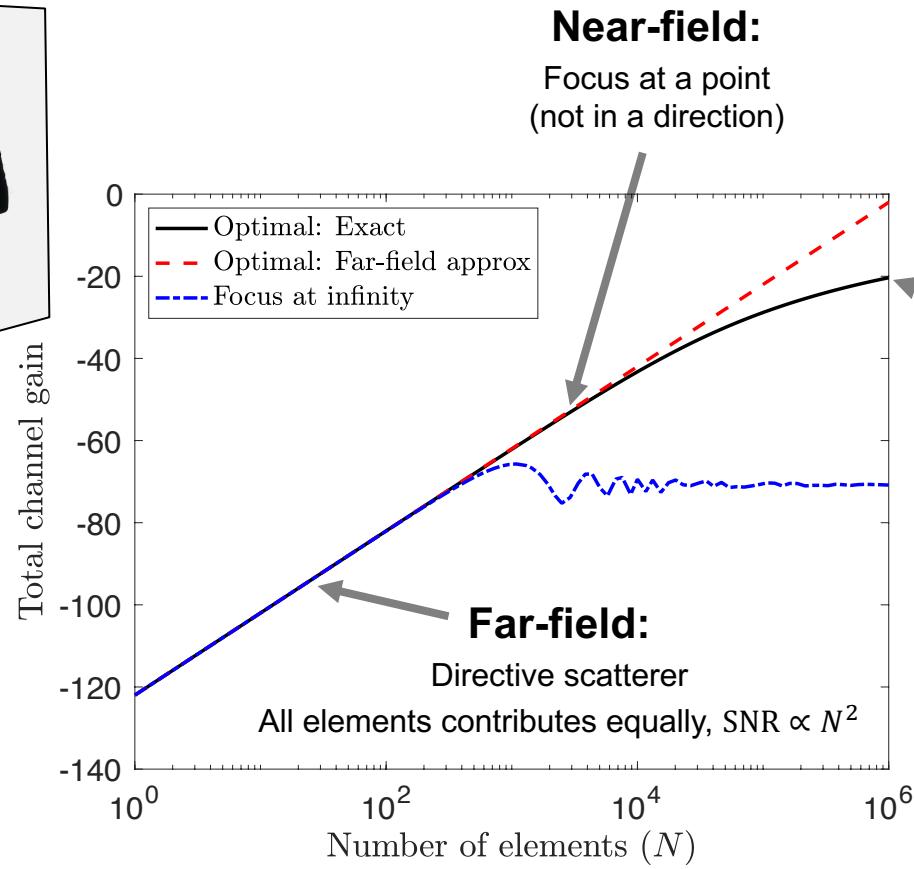


Assumptions:

25 m from transmitter

2.5 m to receiver

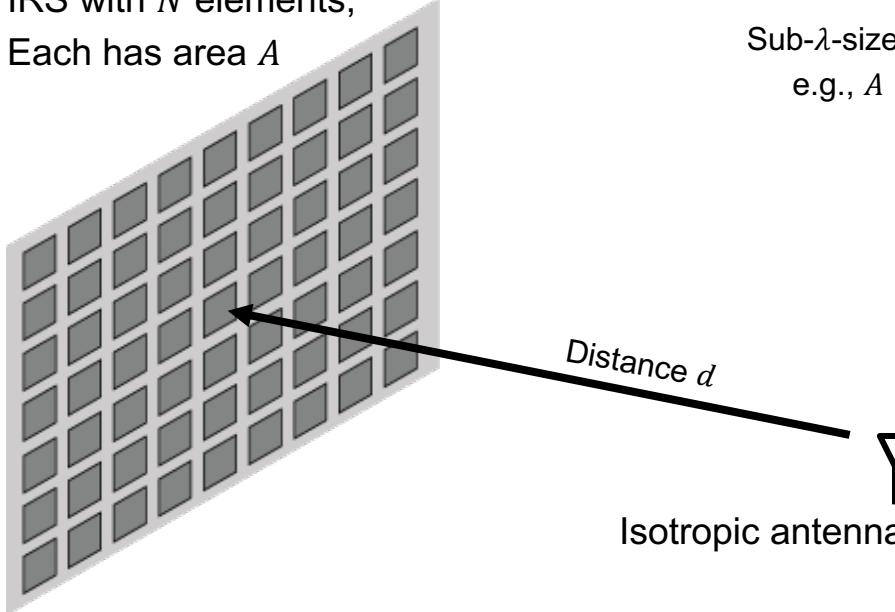
Isotropic antennas (3 GHz)



NARROWBAND CHANNEL MODELING

Basic Pathloss Modeling

IRS with N elements,
Each has area A



$$\text{Pathloss: } \alpha = \frac{\text{Effective area}}{4\pi d^2} \approx \frac{A}{4\pi d^2}$$

Sub- λ -sized elements,
e.g., $A = (\lambda/5)^2$

Compare to isotropic receive antenna:

$$\alpha = \alpha_{\text{iso}} \frac{A}{\lambda^2/4\pi}$$

Smaller than 1

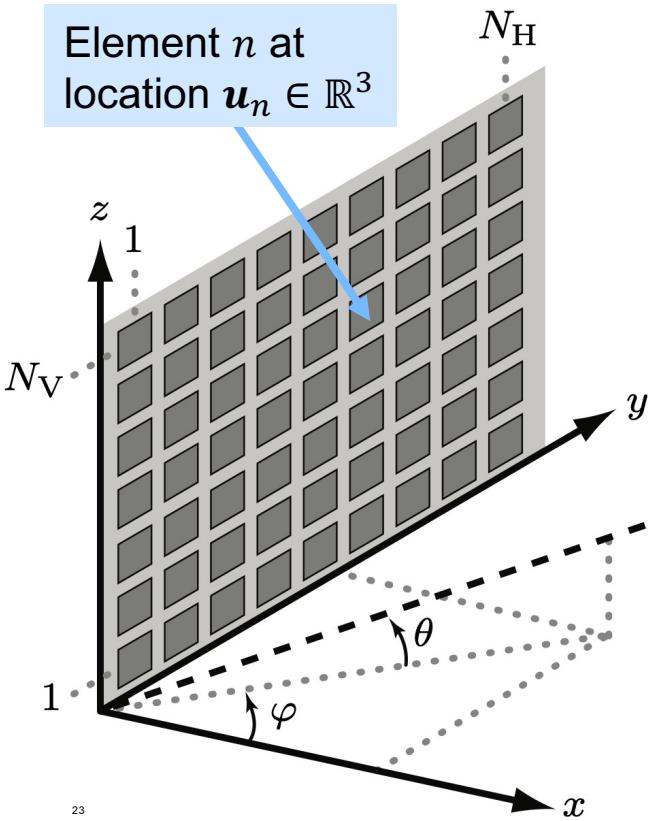
Example: Use 3GPP Urban Micro Model

$$\alpha_{\text{iso}} [\text{dB}] = -37.5 - 22 \log_{10}(d/1 \text{ m})$$

$$\alpha [\text{dB}] = \alpha_{\text{iso}} [\text{dB}] + 10 \log_{10} \left(\frac{A}{\lambda^2/4\pi} \right)$$

Extra reduction

Channel Modeling Using Array Response Vector



Far-field: Incoming/outgoing plane wave determined by

- Azimuth angle $\varphi \in [-\pi, \pi]$
- Elevation angle $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

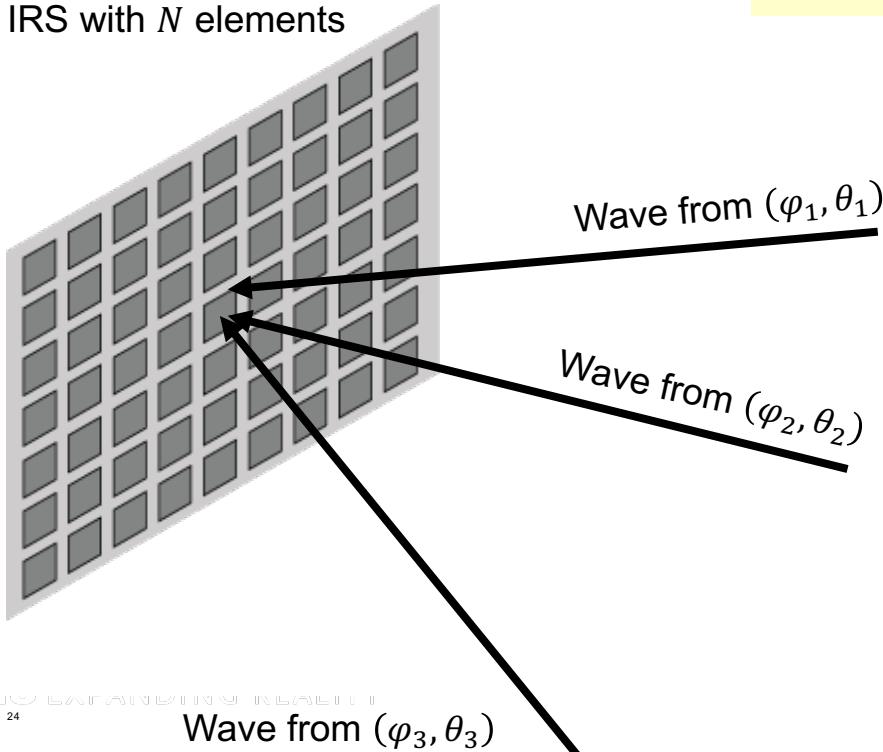
Channels a or b for wave in direction (φ, θ)
constant $\cdot \mathbf{r}(\varphi, \theta)$

Array response vector:

$$\mathbf{r}(\varphi, \theta) = \begin{bmatrix} e^{j\mathbf{k}(\varphi, \theta)^T \mathbf{u}_1} \\ \vdots \\ e^{j\mathbf{k}(\varphi, \theta)^T \mathbf{u}_N} \end{bmatrix}$$
$$\mathbf{k}(\varphi, \theta) = \frac{2\pi}{\lambda} \begin{bmatrix} \cos(\theta) \cos(\varphi) \\ \cos(\theta) \sin(\varphi) \\ \sin(\theta) \end{bmatrix}$$

Basic Multipath Channels

IRS with N elements



Sparse channels

L is small or one path is much stronger than all other

Example: Line-of-sight channels

Channel vector for L plane wave

$$\mathbf{a} = \sum_{l=1}^L \frac{c_l}{\sqrt{L}} \cdot \mathbf{r}(\varphi_l, \theta_l)$$

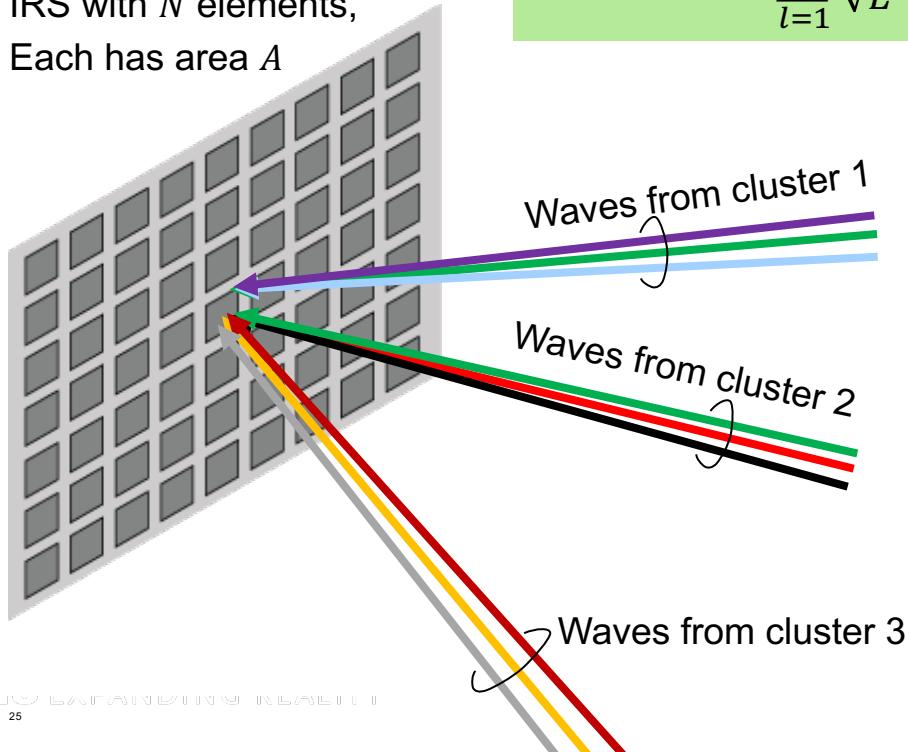
N complex parameters in \mathbf{a}
or $4L$ real parameters

Rich Multipath Channels

Channel vector for L plane wave

$$\mathbf{a} = \sum_{l=1}^L \frac{c_l}{\sqrt{L}} \cdot r(\varphi_l, \theta_l)$$

IRS with N elements,
Each has area A

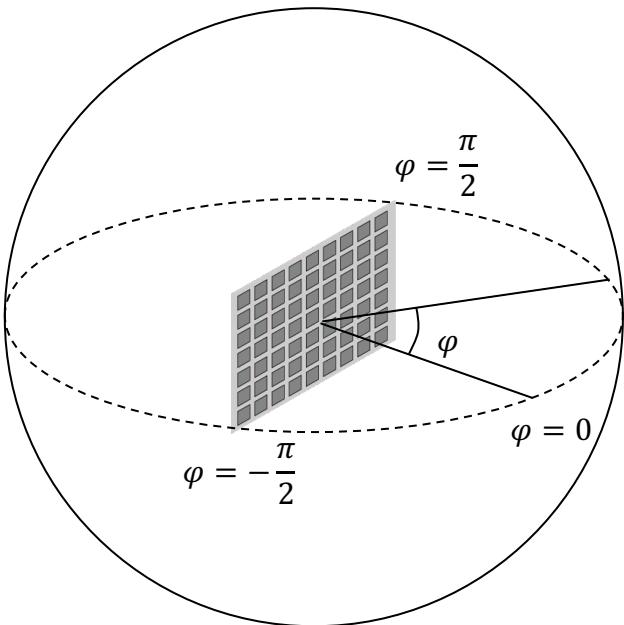


Rayleigh fading: $\mathbf{a} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$

Achieved as $L \rightarrow \infty$ if
 c_l i.i.d. with zero mean, variance $A\mu$
 (φ_l, θ_l) i.i.d. with PDF $f(\varphi, \theta)$

$$[\mathbf{C}]_{nm} = A\mu E \left\{ e^{jk(\varphi, \theta)^T (\mathbf{u}_n - \mathbf{u}_m)} \right\}$$

Isotropic Scattering



Uniform distribution in one half-space

(φ_l, θ_l) i.i.d. with PDF

$$f(\varphi, \theta) = \begin{cases} \cos(\theta)/2\pi, & \varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ 0, & \text{otherwise} \end{cases}$$

Spatial correlation matrix:

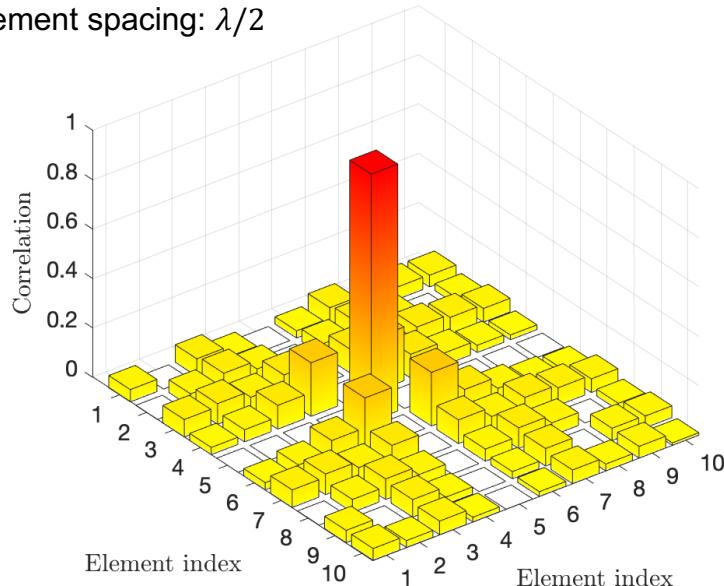
$$\begin{aligned} [\mathbf{C}]_{nm} &= A\mu E \left\{ e^{jk(\varphi, \theta)^T (\mathbf{u}_n - \mathbf{u}_m)} \right\} \\ &= A\mu \cdot \text{sinc}\left(\frac{2}{\lambda} \|\mathbf{u}_n - \mathbf{u}_m\|\right) \end{aligned}$$

Misconception: “We can assume i.i.d. Rayleigh fading”

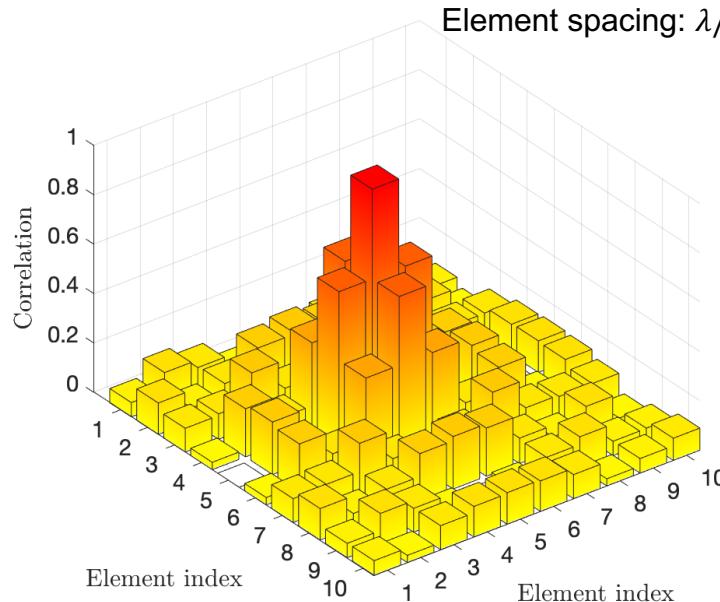
Spatial correlation matrix in isotropic scattering:

$$[\mathbf{C}]_{nm} = A\mu \cdot \text{sinc}\left(\frac{2}{\lambda} \|\mathbf{u}_n - \mathbf{u}_m\|\right)$$

Element spacing: $\lambda/2$



Element spacing: $\lambda/4$



Reference: E. Björnson, L. Sanguinetti, “Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces,” IEEE Wireless Communications Letters, 2021.

Spatial Correlation Appears Naturally

Square array, isotropic scattering

$$\frac{\text{rank}(C)}{\pi NA/\lambda^2} \rightarrow 1$$

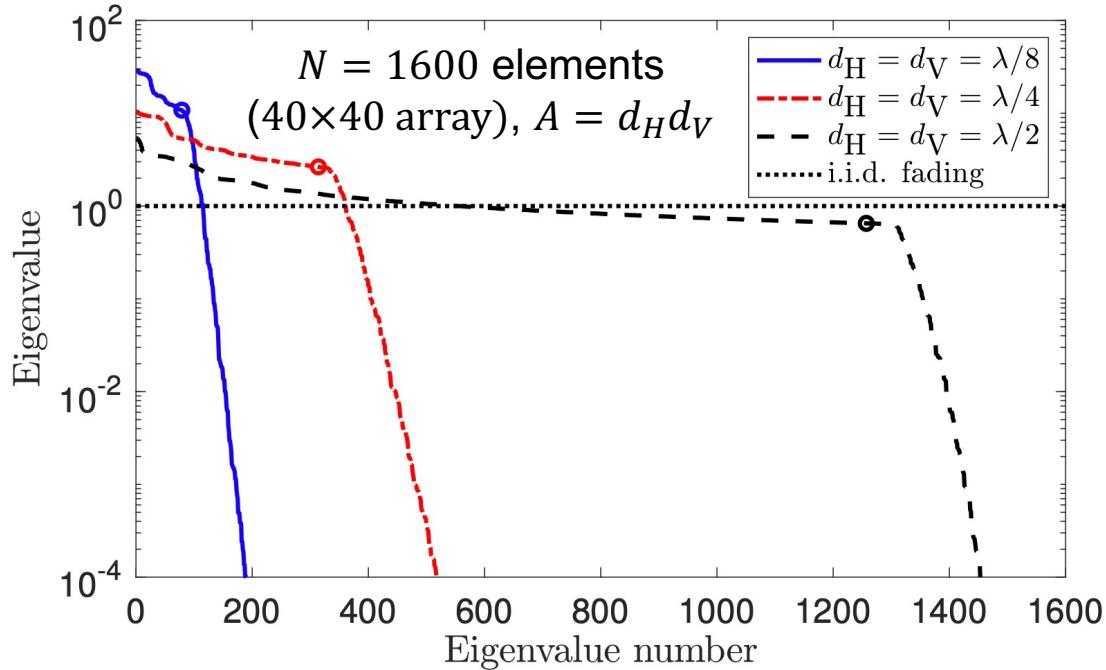
when $N \rightarrow \infty, A \rightarrow 0$ so that $NA \rightarrow \infty$

Channel sparsity

Even in isotropic fading!

Two contributing factors

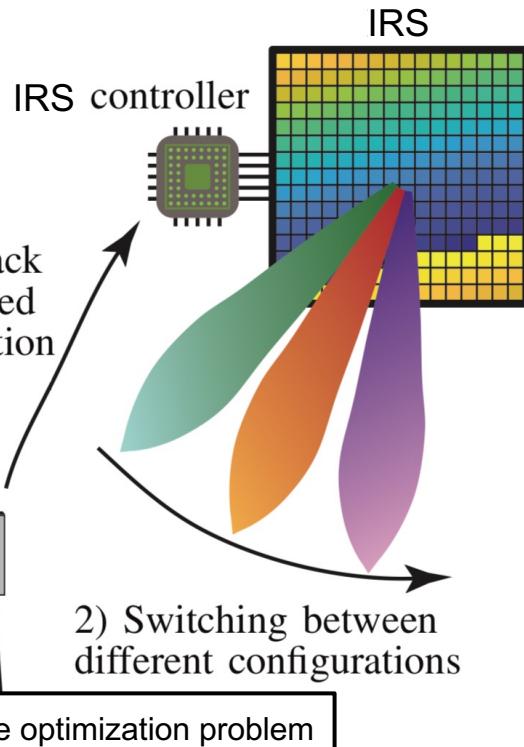
1. Sub-wavelength antenna spacing
2. Two-dimensional surface



CHANNEL ESTIMATION

IRS Reconfigurability is Complicated

The IRS is blind!



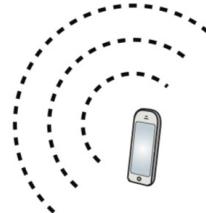
Solution 1: Have a fraction of active IRS elements

Convenient, but removes the “low-cost” argument

Solution 2: Codebook approach

Send pilots and switch configuration

Select the best configuration in a set



1) Repeated pilot transmission

Basic Codebook-Based Channel Estimation

Received signal: $z[n] = (h_d + (\mathbf{a} \odot \mathbf{b})^T \boldsymbol{\omega}_\theta)x[n] + w[n]$

Repeat a known pilot signal L times: $x[n] = s, n = 1, \dots, L$

- Call cascaded channel $\mathbf{c} = \mathbf{a} \odot \mathbf{b}$
- Use different configurations: $\boldsymbol{\omega}_{\theta_1}, \dots, \boldsymbol{\omega}_{\theta_L}$

- Received signal: $z[n] = (h_d + \mathbf{c}^T \boldsymbol{\omega}_{\theta_n})x[n] + w[n]$

$$\underbrace{[z[1], \dots, z[L]]}_{= \mathbf{Z}} = h_d[1 \dots 1]s + \underbrace{\mathbf{c}^T [\boldsymbol{\omega}_{\theta_1}, \dots, \boldsymbol{\omega}_{\theta_L}]s}_{= \boldsymbol{\Omega}} + \underbrace{[w[1], \dots, w[L]]}_{= \mathbf{W}}$$

$$\mathbf{Z} = [h_d]^T \begin{bmatrix} 1 & \cdots & 1 \\ \mathbf{c} & \Omega \end{bmatrix} s + \mathbf{W}$$

If $\tilde{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & \cdots & 1 \\ \boldsymbol{\Omega} \end{bmatrix}$ is invertible

$$\mathbf{Z}\tilde{\boldsymbol{\Omega}}^{-1}s^{-1} = [h_d]^T + \mathbf{W}\tilde{\boldsymbol{\Omega}}^{-1}s^{-1}$$

Received signal

Requires $L = N + 1$
Example: $\tilde{\boldsymbol{\Omega}} = \text{DFT or Hadamard matrix}$

Channel estimate

Exploiting Channel Sparsity

$$\mathbf{z} = \begin{bmatrix} h_d \\ \mathbf{c} \end{bmatrix}^T \begin{bmatrix} 1 & \dots & 1 \\ \boldsymbol{\Omega} \end{bmatrix} s + \mathbf{w}$$

Received signal

Sparsity will exist in c !

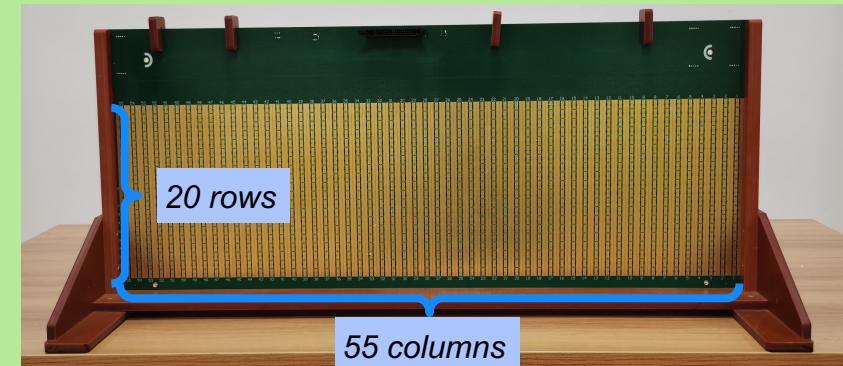
An open research challenge to utilize it most effectively

References: "Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions", 2020

Use sparse channel models

Group elements together

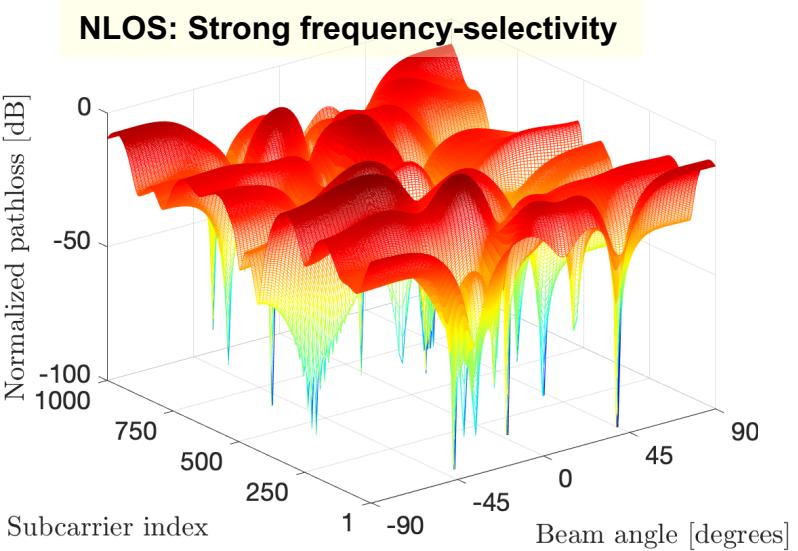
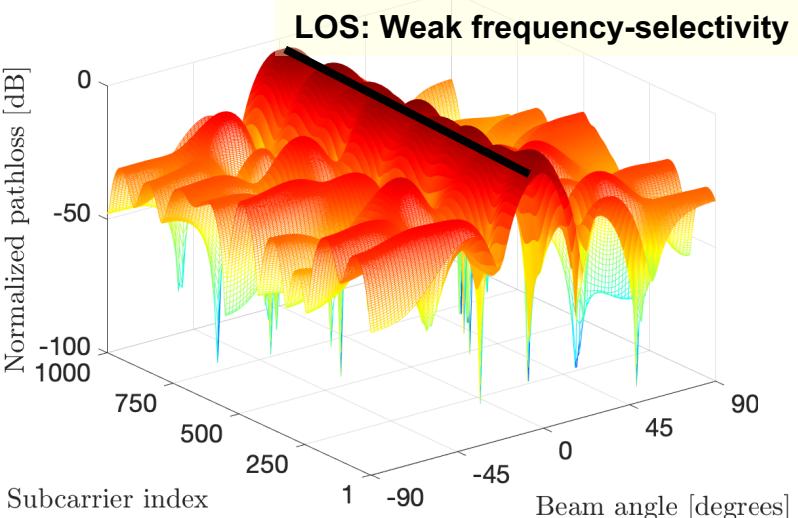
Basic configuration algorithm



Two states per element: ± 90 degrees

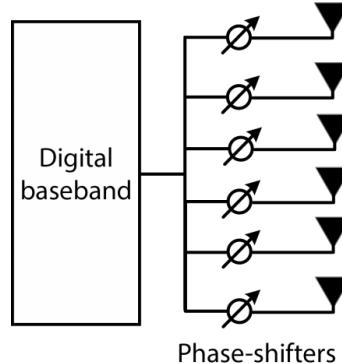
- Flip one row or column, pick best option
- Sweep horizontally, then vertically

WIDEBAND MODELING

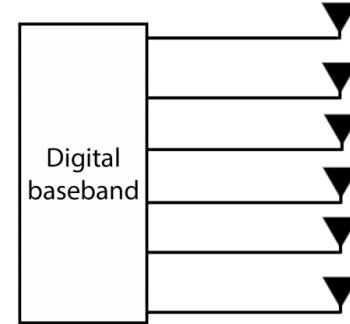


Frequency selective fading

Analog beamforming



Digital beamforming



One angular beam

Same on all subcarriers

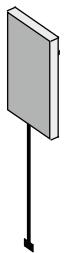
Same principle as for IRS

Sum of many beams

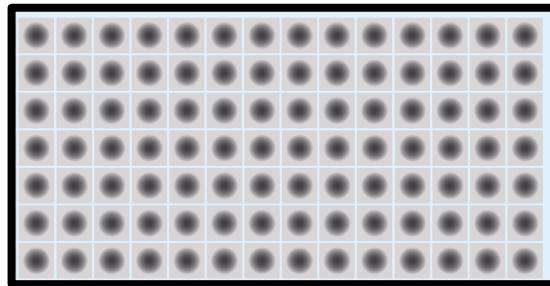
Different over subcarriers

OFDM System Modeling

Transmitter



Distinct paths:
 $\mathbf{a}_1, \dots, \mathbf{a}_{L_h}$



Phase-shifts: $\boldsymbol{\omega}_\theta = \begin{bmatrix} e^{-j\theta_1} \\ \vdots \\ e^{-j\theta_N} \end{bmatrix}$

Receiver



Distinct paths:
 $\mathbf{b}_1, \dots, \mathbf{b}_{L_g}$

K subcarrier channels:

$$\mathbf{y} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_K \end{bmatrix} = \mathbf{F} \mathbf{V}^T \boldsymbol{\omega}_\theta$$

DFT matrix ($K \times M$)

Time-domain end-to-end channel matrix (M taps)

$$\mathbf{V} = \sum_{l=1}^{L_h} \sum_{i=1}^{L_g} (\mathbf{a}_l \odot \mathbf{b}_i) \left[\delta[\tau_{h,l} + \tau_{g,i}] \dots \delta[\tau_{h,l} + \tau_{g,i} - (M-1)] \right]$$

↑ ↑
Integer delays

IRS Optimization for OFDM system

Sum rate expression:

$$R = \frac{B}{K + M - 1} \sum_{k=1}^K \log_2 \left(1 + \frac{P}{\sigma^2} |\gamma_k|^2 \right) = \frac{B}{K + M - 1} \sum_{k=1}^K \log_2 \left(1 + \frac{P}{\sigma^2} |\mathbf{f}_k^H \mathbf{V}^T \boldsymbol{\omega}_\theta|^2 \right)$$


Row k of \mathbf{F}

Upper bound: Optimize $\boldsymbol{\omega}_\theta$ for each subcarrier:

$$|\mathbf{f}_k^H \mathbf{V}^T \boldsymbol{\omega}_\theta|^2 \leq \|\mathbf{f}_k^H \mathbf{V}^T\|_1^2$$

Heuristic solution: Strongest-path maximization (STM),

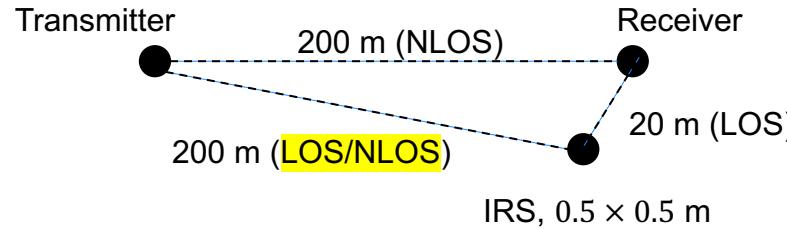
$$m = \arg \max_i \|\mathbf{v}_i\|_1^2 \text{ where } \mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_M]$$

$$\boldsymbol{\omega}_\theta = \arg \max_{\boldsymbol{\omega}} |\mathbf{v}_m^T \boldsymbol{\omega}|^2$$

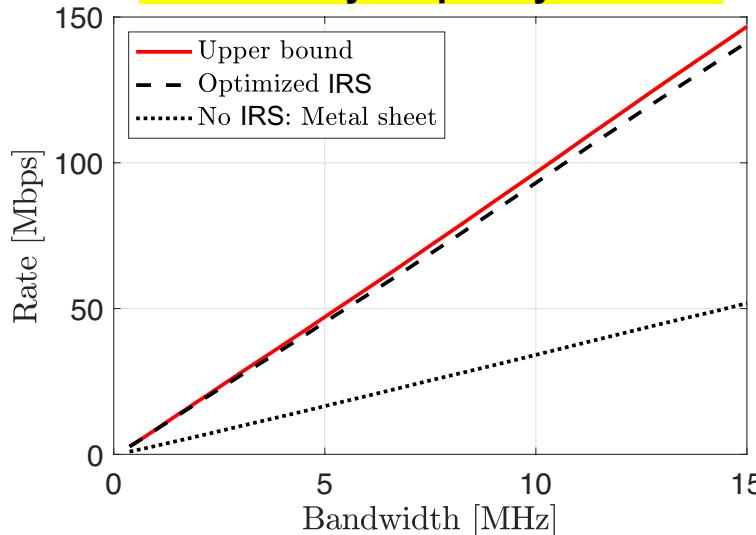
Reference: B. Zheng, R. Zhang, "Intelligent Reflecting Surface-Enhanced OFDM: Channel Estimation and Reflection Optimization," IEEE Wireless Communications Letters, 2020

IRS in Frequency Selective Channels

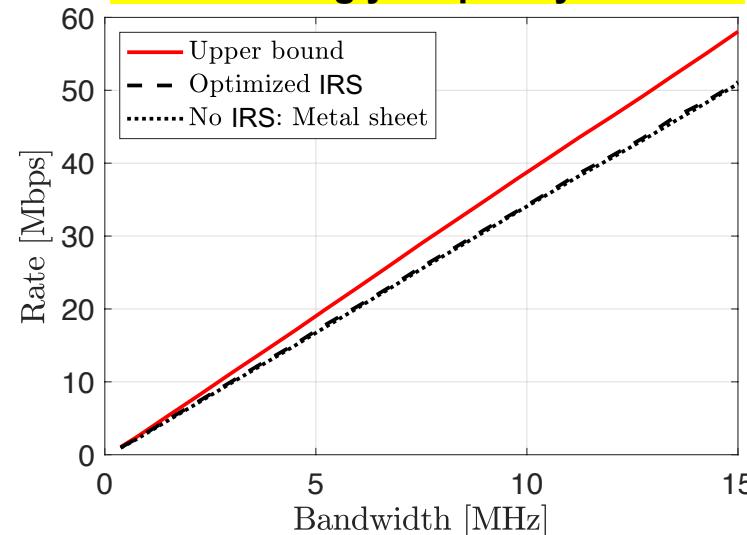
IRS mainly effective with LOS
to transmitter and receiver!



LOS: Weakly frequency-selective

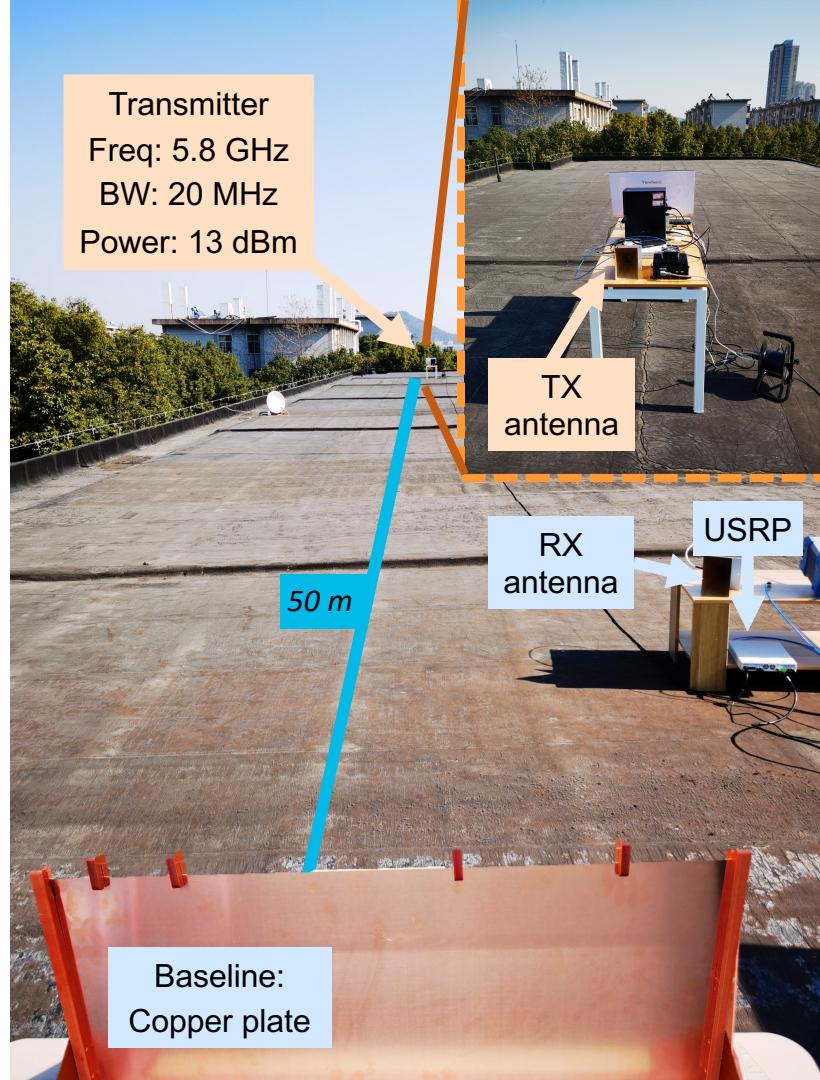
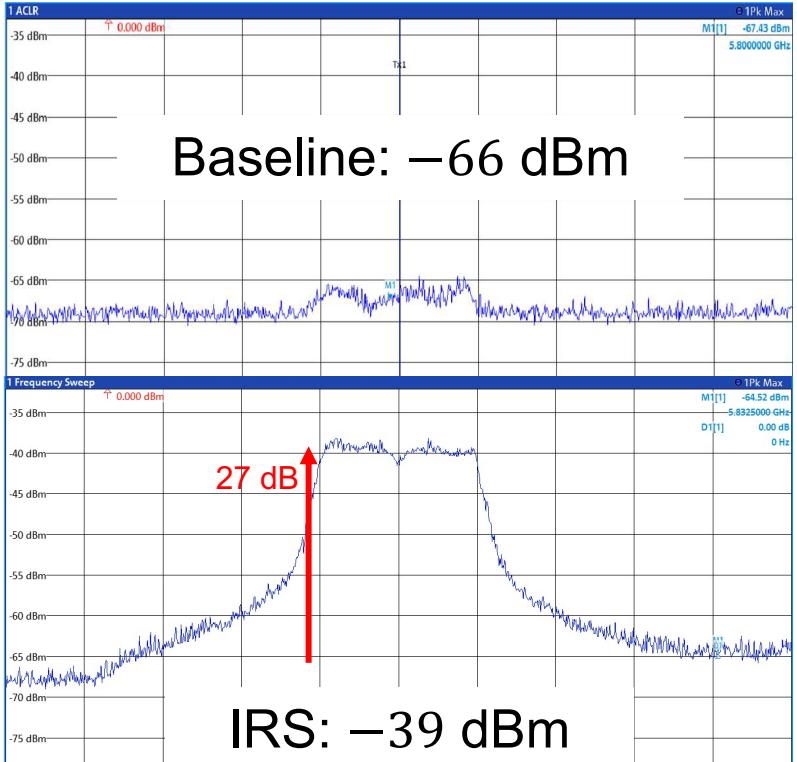


NLOS: Strongly frequency-selective



Reference: E. Björnson, et al., "Reconfigurable Intelligent Surfaces: A Signal Processing Perspective With Wireless Applications", Available on arXiv:2102.00742.

Experimental Validation



Reference: X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, E. Björnson, "RIS-Aided Wireless Communications: Prototyping, Adaptive Beamforming, and Indoor/Outdoor Field Trials," to appear

COMPETING TECHNOLOGIES AND POTENTIAL WEAKNESSES

Passive Repeaters versus IRS

Passive
repeater

LOS path

Shadowing

Passive repeater
+ Fully passive,
no power

- Fixed beam pattern:
wide or strong?

IRS

LOS path

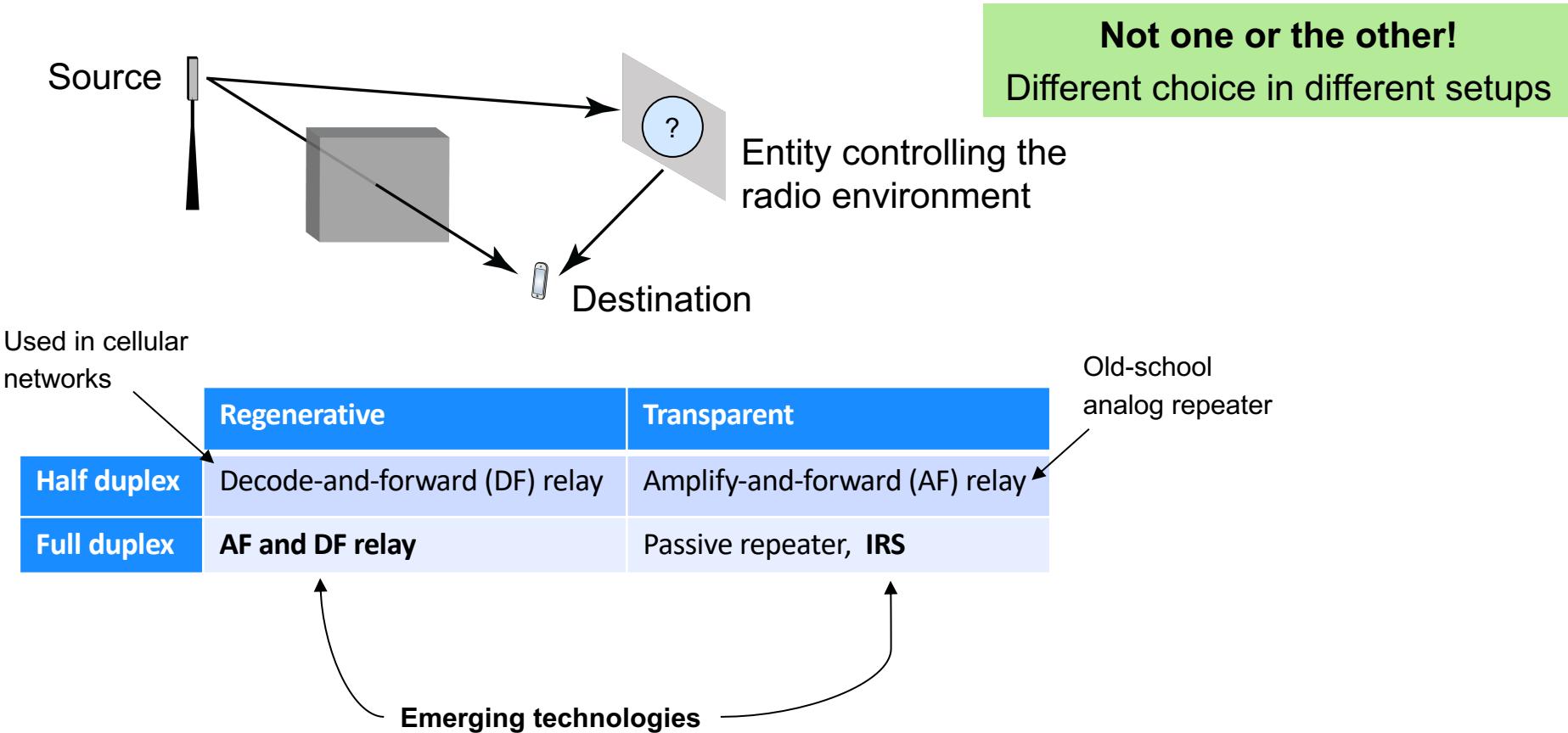
Shadowing

IRS

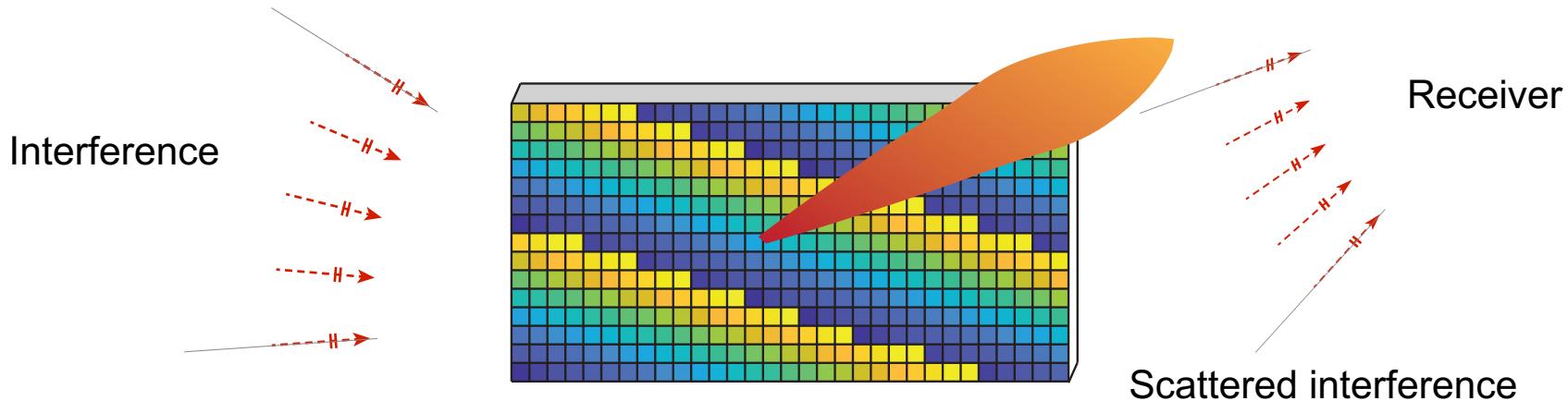
+ Strong beam to
desired location

- Must be controlled
- Requires power

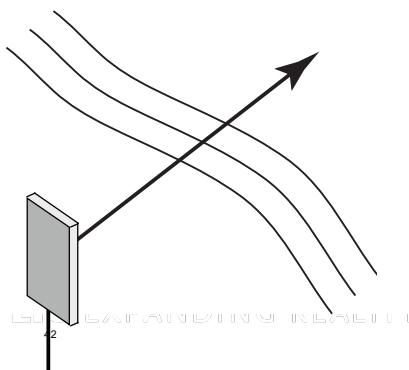
Taxonomy of “Cooperative Communications”



Electromagnetic Interference



Like painting a room with bright colors
Interference reflected by IRS proportional to N



IRS Might Reduce “SNR”

Simulation setup

SNR at IRS elements: 20 dB

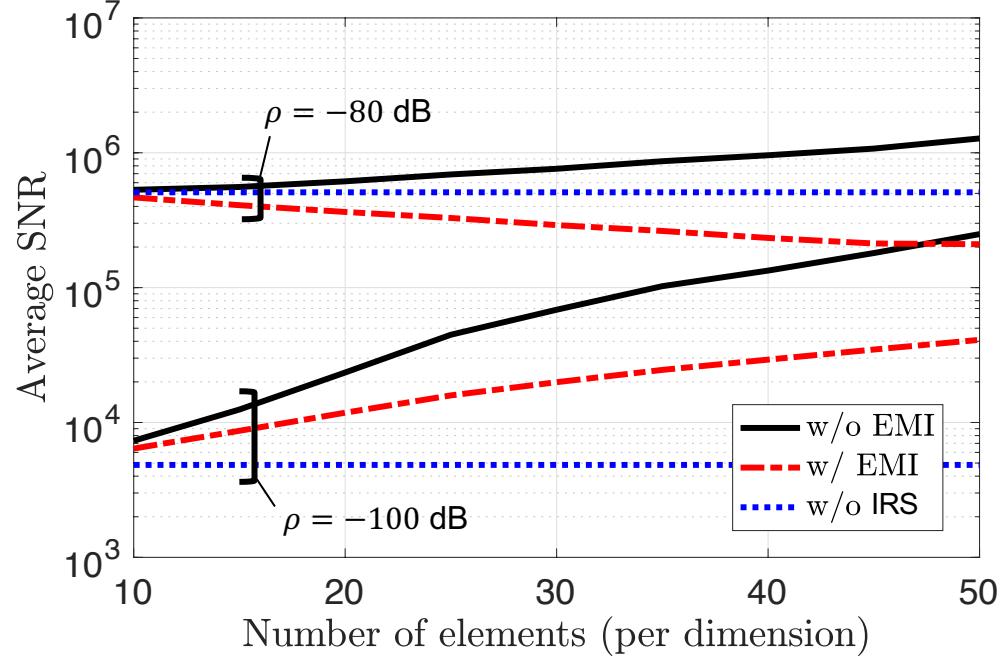
Isotropic electromagnetic interference (EMI)

IRS can sometimes reduce performance,
Asymptotic SNR proportional to N , not N^2

Careful deployment is necessary!

$$\text{SNR} \approx \frac{P(\sqrt{\rho} + N\sqrt{\alpha\beta})^2}{N \cdot \text{constant} + \sigma^2}$$


EMI



The Most Promising Example

	NLOS to/from IRS	LOS to/from IRS
Strong direct path	IRS not effective in OFDM Difficult to estimate	IRS adds little to direct path
Weak direct path		Large gains in OFDM Few parameters to estimate

END OF PART 1