RECONFIGURABLE INTELLIGENT SURFACES AND HOLOGRAPHIC MASSIVE MIMO

VISION, FUNDAMENTALS, AND KEY OPEN PROBLEMS

Emil Björnson

KTH Royal Institute of Technology

and **Linköping University**, Sweden

Outline

Introduction

- Beamforming and wave propagation
- Evolution of multi-antenna technology

Reconfigurable intelligent surfaces

- Fundamentals and vision
- System model and optimization

Holographic Massive MIMO

- Properties and benefits

Spatial channel modeling

Open problem 1: What are suitable use cases?

Open problem 2: Channel estimation

Summary

INTRODUCTION

Physics of Wireless Signal Propagation

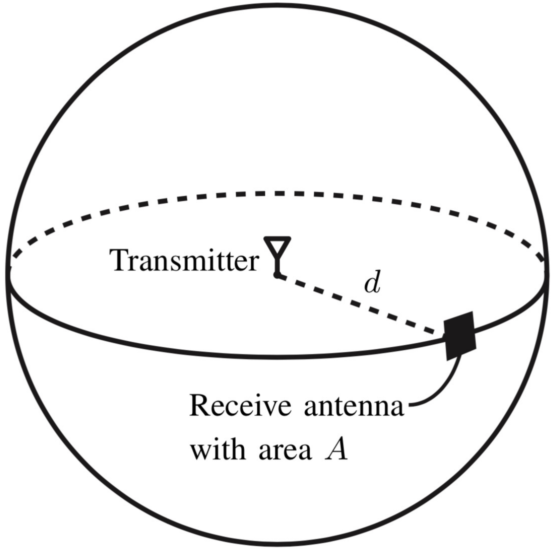
- Electromagnetic waves travel at speed of light
  - Spreads out in all directions

- Friis' propagation formula:

$$\text{Receive power} = \text{Transmit power} \cdot \frac{A}{4\pi d^2}$$

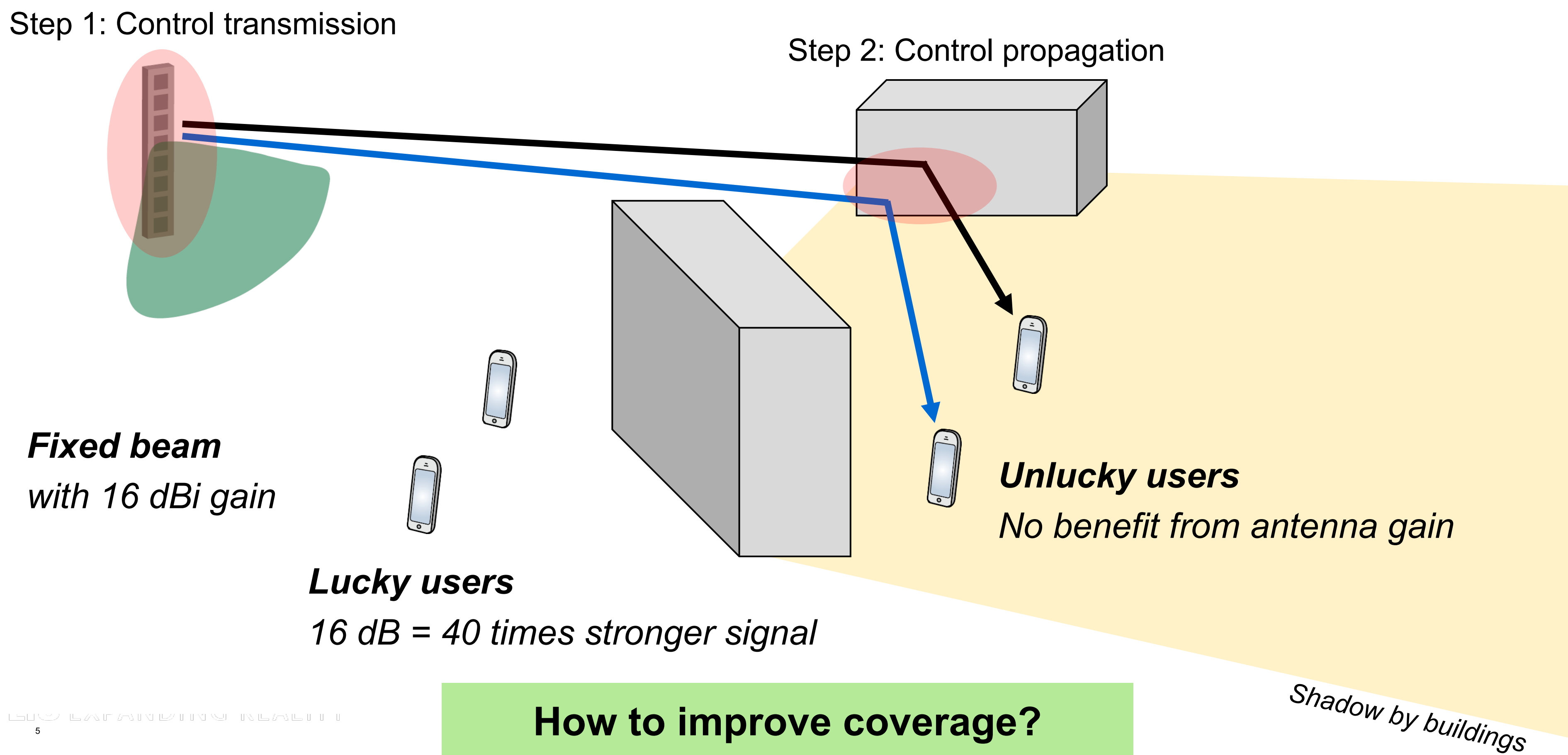
Example: $A = \left(\frac{\lambda}{4}\right)^2$, $\lambda = 0.1$ m (3 GHz)

0.005% received at 1 m $(-43$ dB)
0.00005% received at 10 m $(-63$ dB)



Only a tiny fraction of transmit power is received!

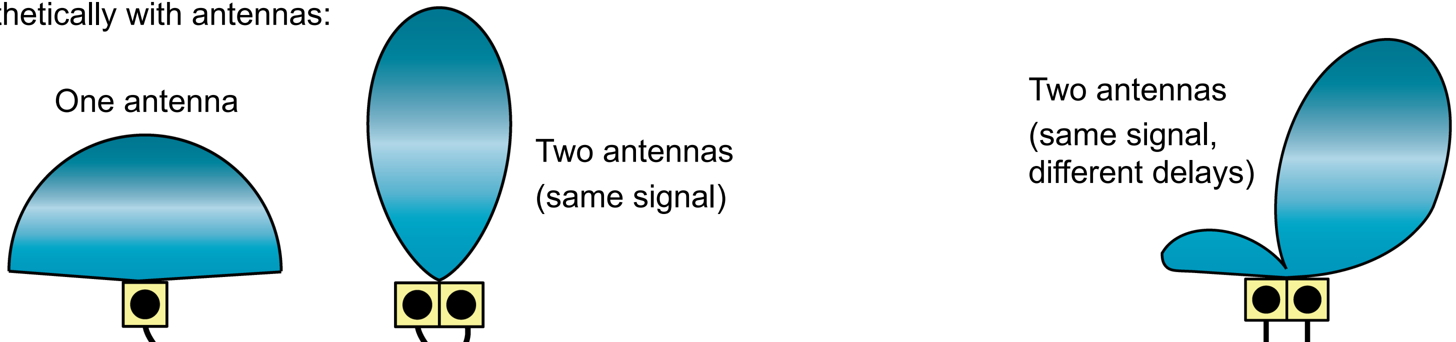
Conventional Base Station Antennas



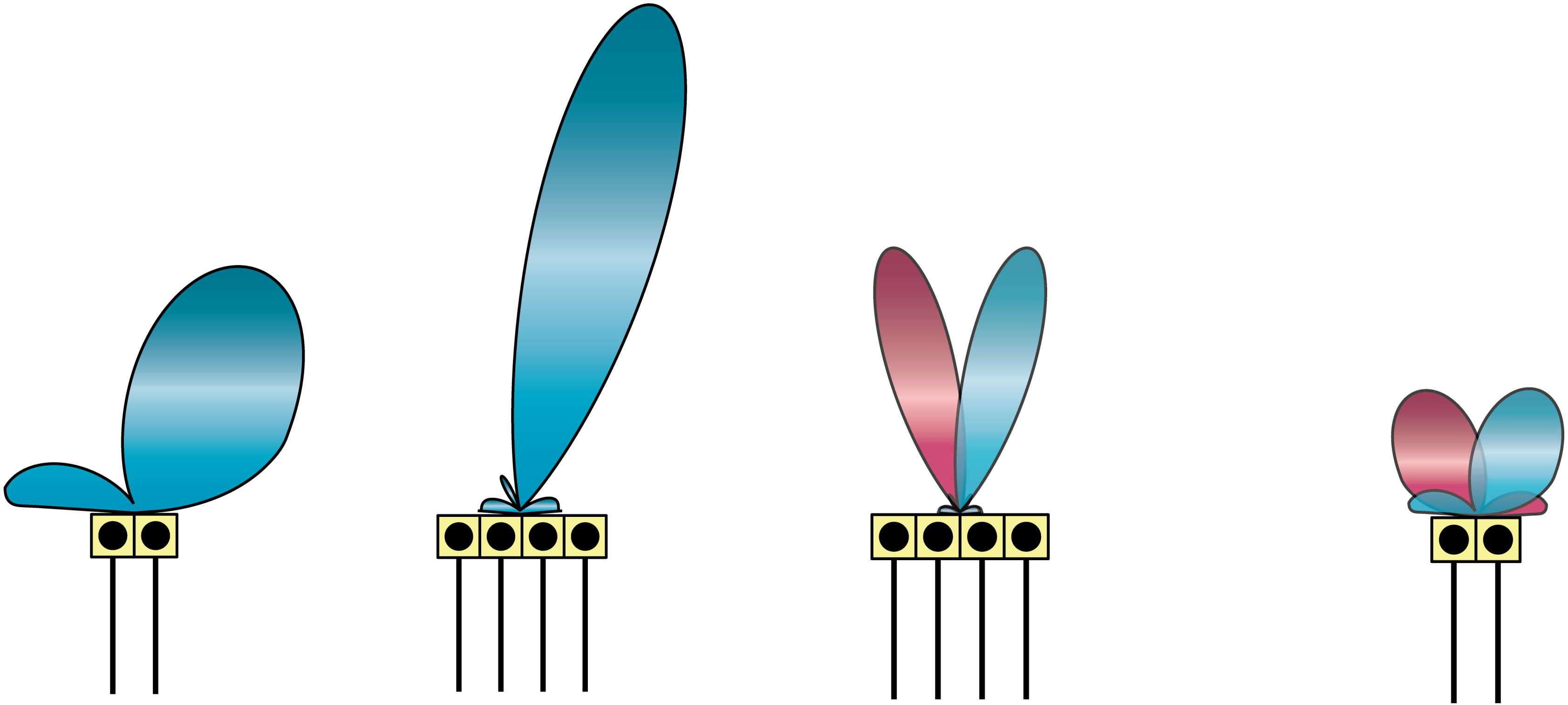
How to improve coverage?

Beamforming by Constructive Interference

Generate synthetically with antennas:



Array Gain and Spatial Multiplexing



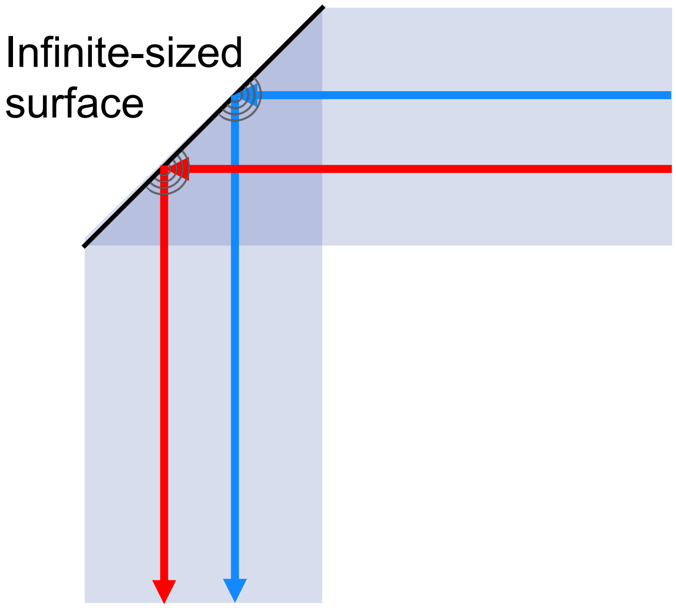
Two antennas:
2x signal signal

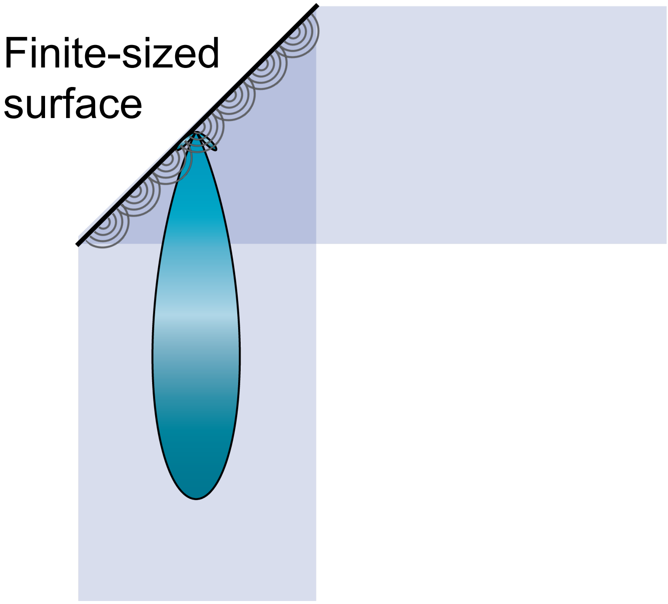
Four antennas:
4x signal gain

Four antennas, two signals:
Divide power between two beams
4x signal gain

Two antennas, two signals:
Divide power between two beams
2x signal gain

Constructive Interference with Reflection





Huygens–Fresnel principle

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

Beamforming

Direction determined by geometry
Beamwidth ∝ 1/(surface width)

EVOLUTION OF MULTI-ANTENNA TECHNOLOGY

Pioneering Work on Adaptive Beamforming

Grimeton Radio Station

Outside Varberg, Sweden

UNESCO World Heritage Site

UNITED STATES PATENT OFFICE.

ERNST F. W. ALEXANDERSON, OF SCHENECTADY, NEW YORK, ASSIGNOR TO GENERAL ELECTRIC COMPANY, A CORPORATION OF NEW YORK.

ANTENNA.

1,360,167. Specification of Letters Patent. Patented Nov. 23, 1920.

Application filed September 13, 1917. Serial No. 191,110.

To all whom it may concern:

Be it known that I, ERNST F. W. ALEXANDERSON, a citizen of the United States, residing at Schenectady, in the county of Schenectady, State of New York, have invented certain new and useful Improvements in Antennæ, of which the following is a specification.

My present invention relates to antennæ for radio signaling systems and more particularly to the manner in which the radiation of the transmitting system is effected. My present application is a continuation in part of my prior application Serial No. 123,276 filed Oct. 2, 1916.

The antenna of a radio signaling system as previously constructed has consisted of an elevated electrical conductor or a network of conductors which is charged by a source of high frequency energy in such a way that it becomes a source of energy radiation of the type known as the Hertzian oscillator. The theory for the radiation of an antenna has therefore been universally treated by the mathematical theory of the Hertzian oscillator.

The usual radiating antenna system may be considered as a single Hertzian oscillator. It has been proposed heretofore to employ a plurality of such oscillators in order to increase the amount of radiation or to secure directive effects but no practical use has

current and voltage supplied to the radiator. The energy consumed by the radiator is proportional to the product of current and voltage and is consequently proportional to the square of the field intensity of the radiated wave. The general theory of wave motion, however, teaches that if several systems of waves are superimposed, they combine in such a way that the field intensity in any one place is the algebraic sum of the momentary intensities of all the separate waves. If a system consisting of a plurality of separate radiators is controlled in such a way that the relative phase of the oscillations from the individual radiators is made to combine in a predetermined desired way, it will be possible to operate the system in such a manner that the field intensity in the receiving station is the arithmetic sum of the field intensities produced by all of the individual radiators. A radiation of unity intensity from a station with a single radiator may be said to produce a field intensity of unity in the receiving station. If a system comprising a plurality of radiators is operated in such a way that each of the radiators emits a wave of unity intensity, the effect on the receiving station will be the same as that of a wave with a field strength of as many times unity as there are individual radiators.

The energy consumption of the single ...

MIMO Communication (Multiple-Input Multiple-Output)

Space-division multiple access

Multi-user MIMO

Concept from late 80s, early 90s

Information theory in 00s

Patents submitted in early 90s

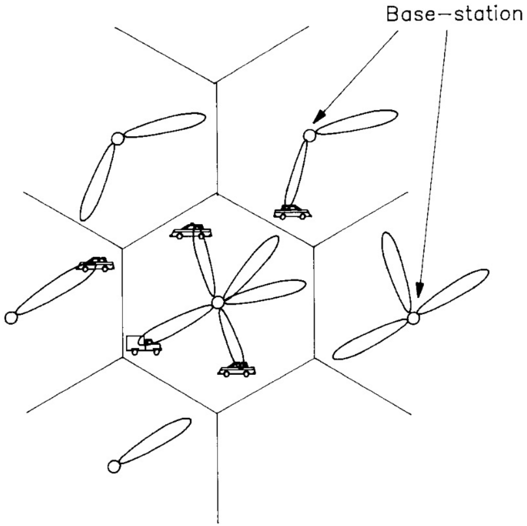
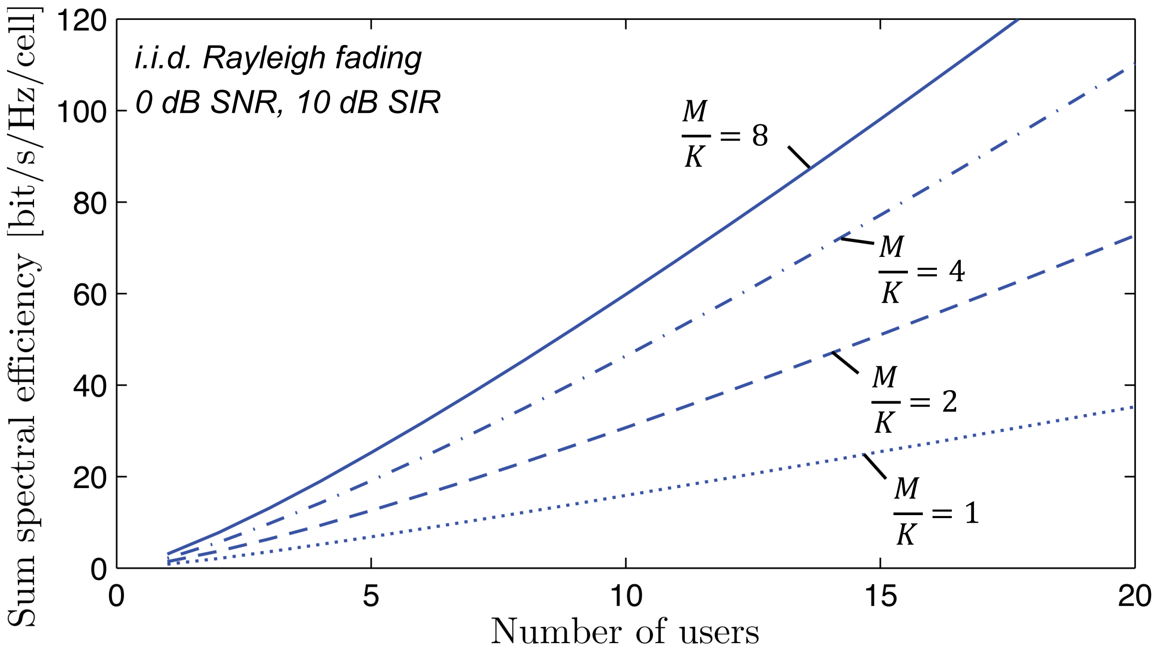


Fig. 2. Tracking of mobiles with multiple beams.

Motivation: Data rate grows with number of antennas M and users K



MIMO Communication (Multiple-Input Multiple-Output)

Space-division multiple access
Multi-user MIMO

Concept from late 80s, early 90s
Information theory in 00s
Patents submitted in early 90s

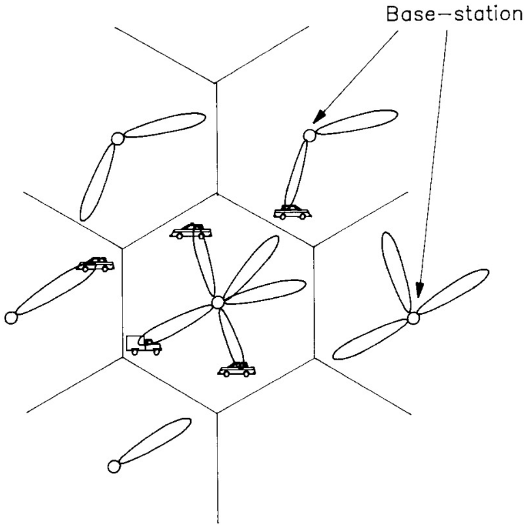
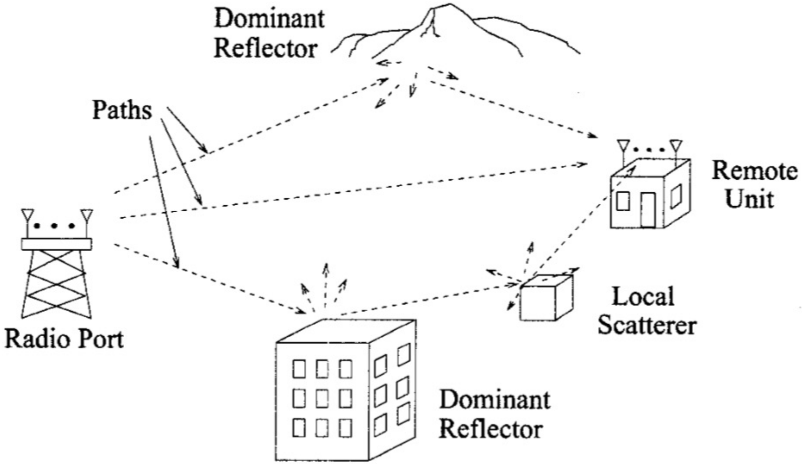


Fig. 2. Tracking of mobiles with multiple beams.

Layered space-time architecture
Single-user MIMO

Concept and patents from late 90s
Information theory at the same time



Motivation: Same benefit for one user

From Science Fiction to Mainstream

Skeptical voices…

Too large and expensive

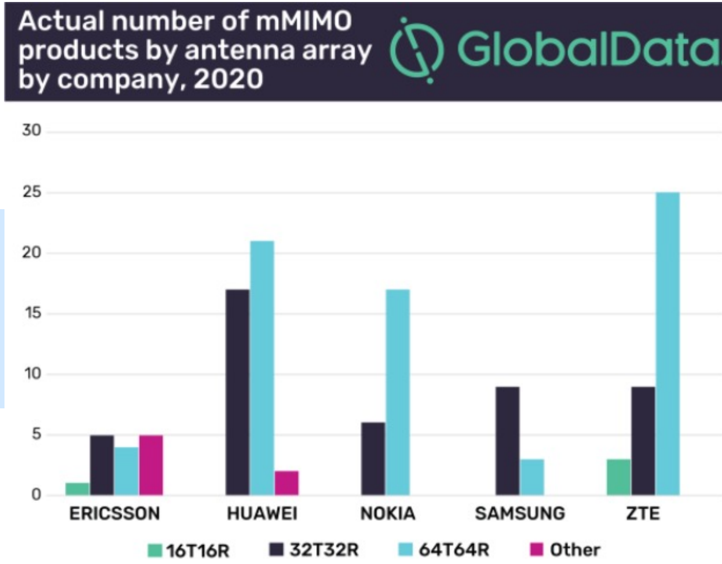
Too high complexity

No practical gains

Not competitive

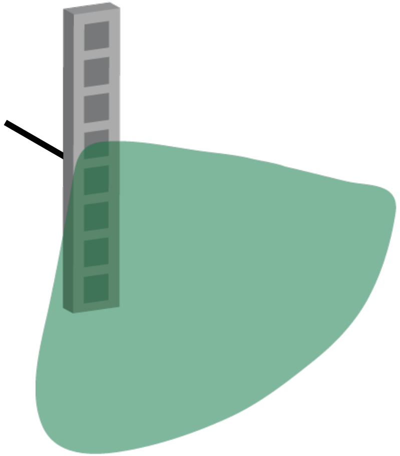
- Easier to deploy more base stations
- Too small: $M \approx K \approx 8$
- No information theory behind it
- Voice calls, rather than data packets

Today:
A mainstream
5G technology



Base Station Development in 5G

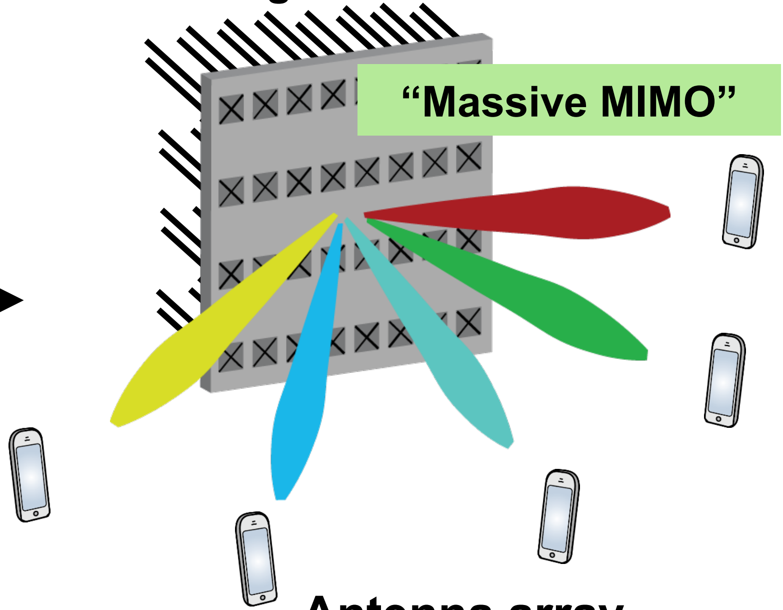
1 high-gain antenna



Passive antenna

Constant directivity

64 low-gain antennas



Antenna array

Antenna-integrated radios

Strong, adaptive directivity

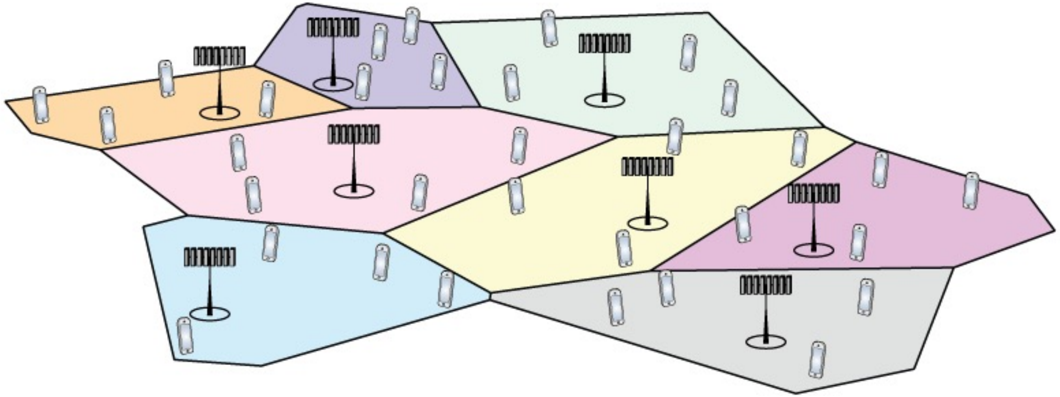
Separate users in space

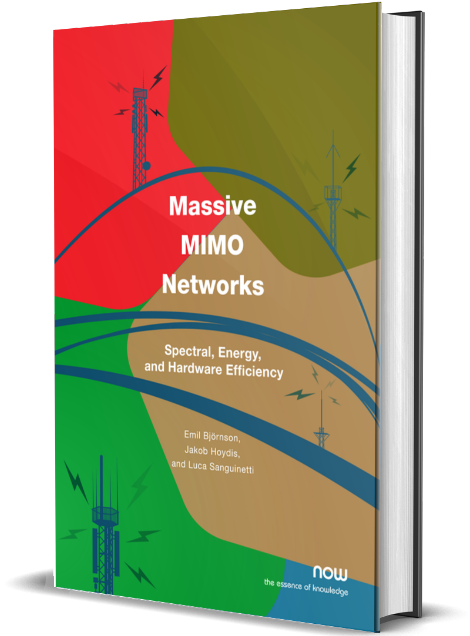
Definition of Massive MIMO

- Cellular network
  - Many antennas M per base station
  - Spatial multiplexing of many users K
  - Antenna-user ratio: $M/K > 1$

$M \geq 64$

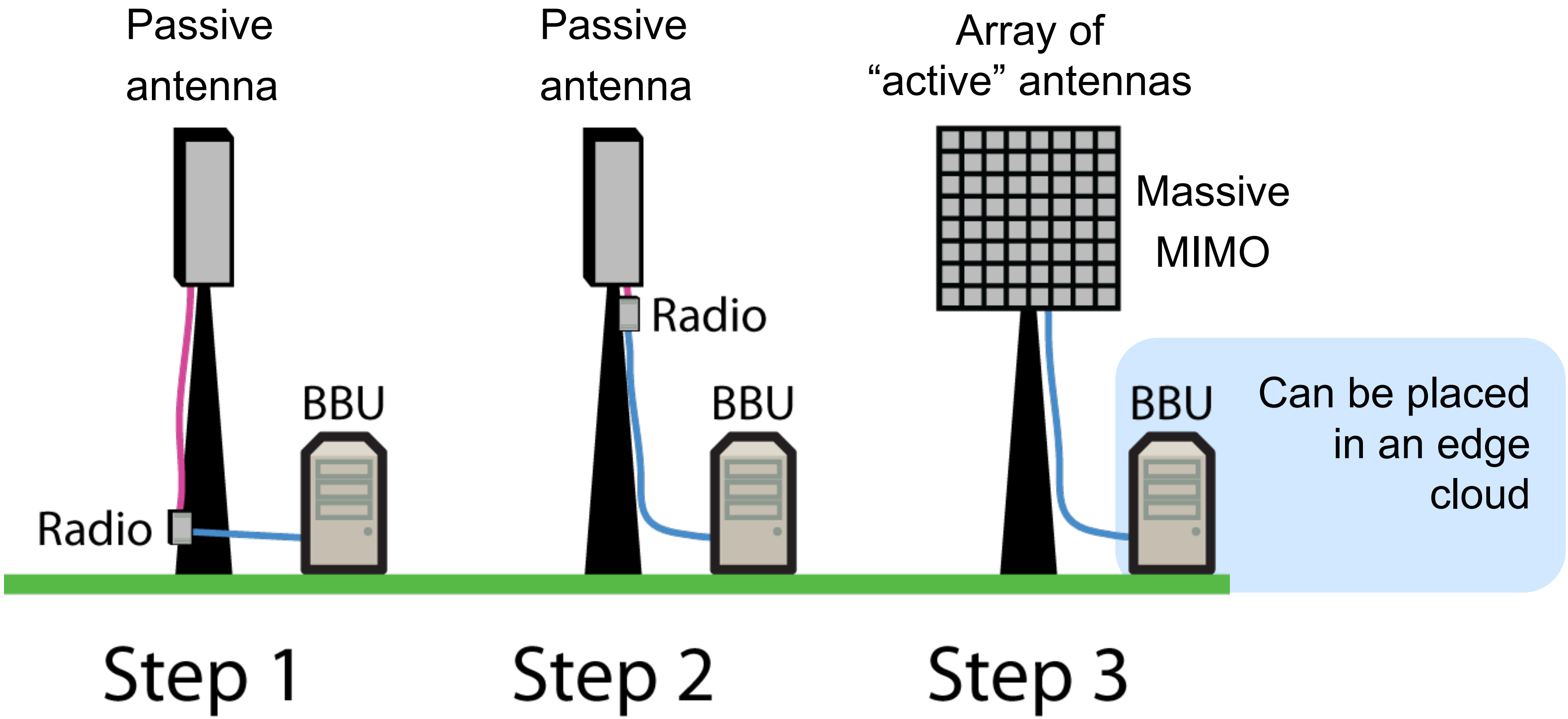
$K \geq 8$



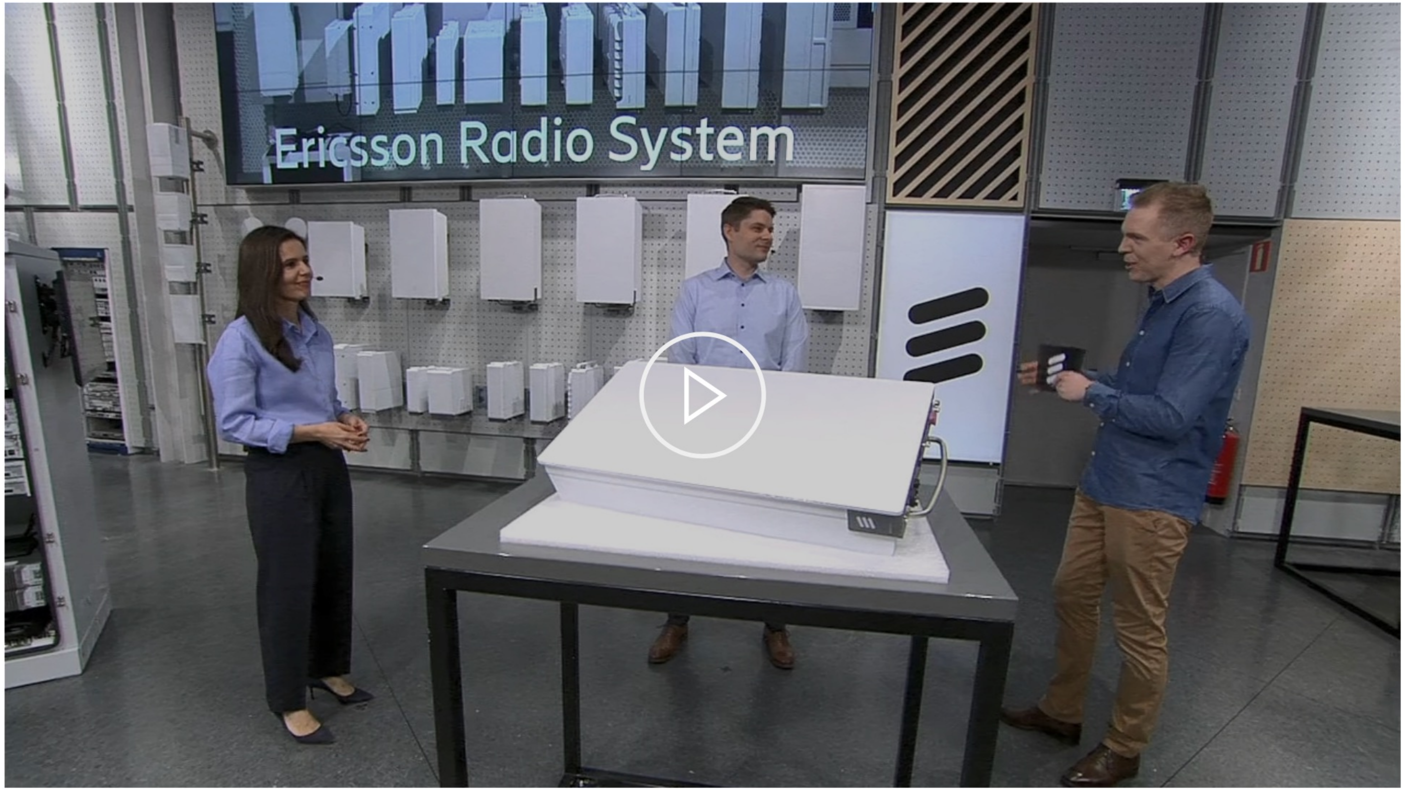


Seminal work: Thomas Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” IEEE Trans. Wireless Communications, 2010

Hardware Evolution Enabling Massive MIMO



Massive MIMO is a Reality: How ”Massive” is it?



AIR 6419

64 antennas
(192 elements)

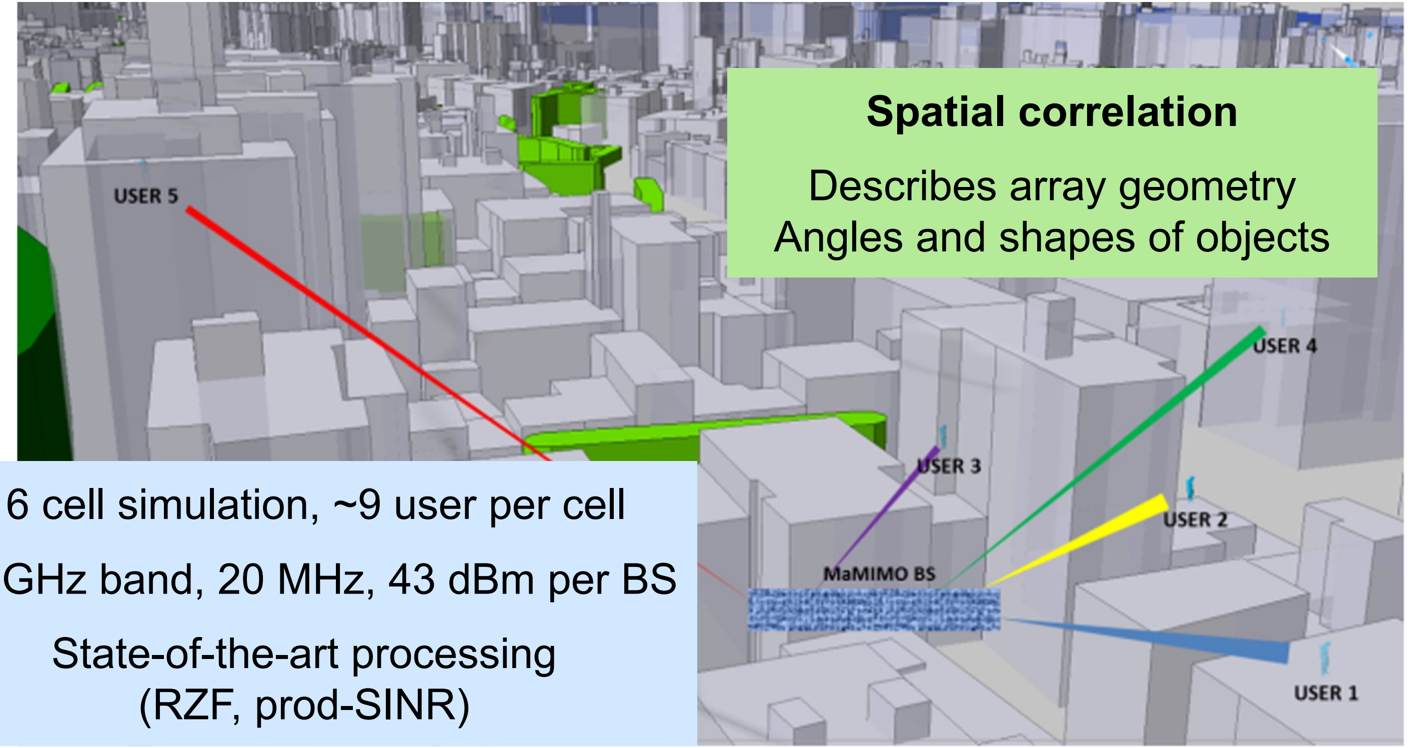
20 kg

320 W over
200 MHz

Integrated
circuitry

Ericsson Live Broadcast, March 1, 2021

Does Massive MIMO Perform as Expected?



Spatial correlation

Describes array geometry
Angles and shapes of objects

16 cell simulation, ~9 user per cell

3.5 GHz band, 20 MHz, 43 dBm per BS

State-of-the-art processing
(RZF, prod-SINR)

M. Aslam, Y. Corre, E. Björnson, E. G. Larsson, “Performance of a Dense Urban Massive MIMO Network From a Simulated Ray-Based Channel,” EURASIP Journal on Wireless Communications and Networking, 2019.

Answer: Depends on Antenna Deployment

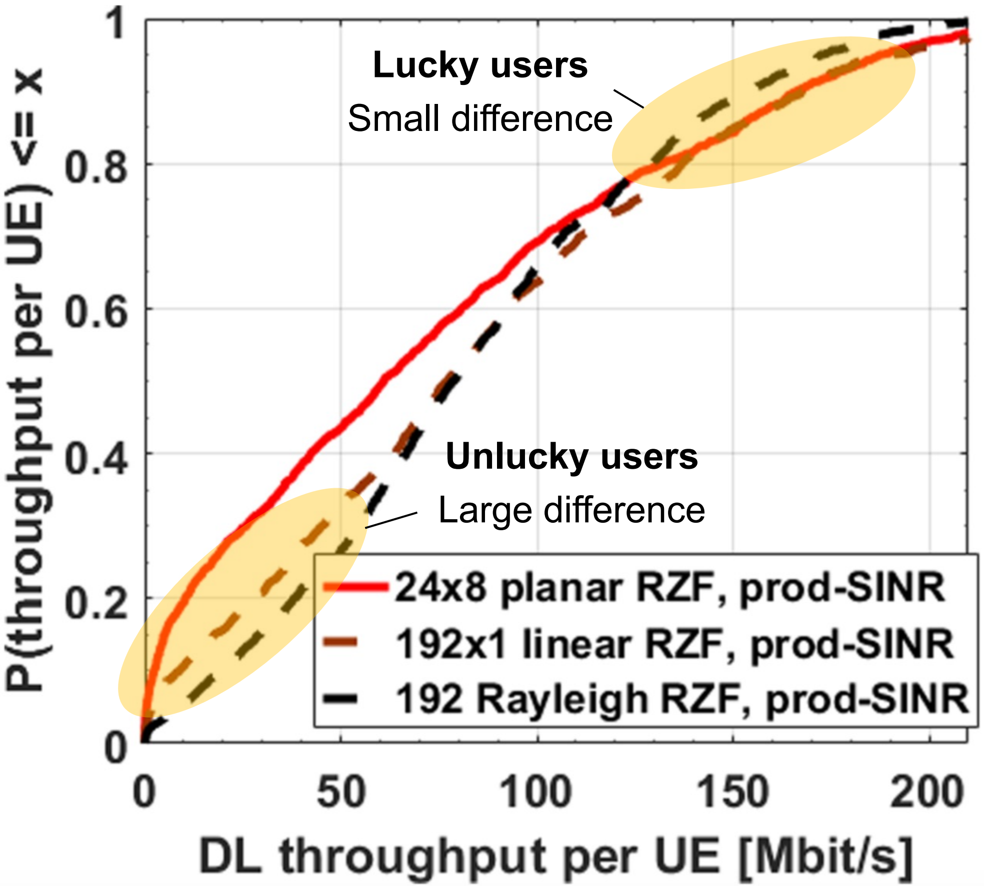
Marzetta's baseline:

i.i.d. Rayleigh fading

Two deployments:

Planar array (24 x 8): 1 m x 0.34 m

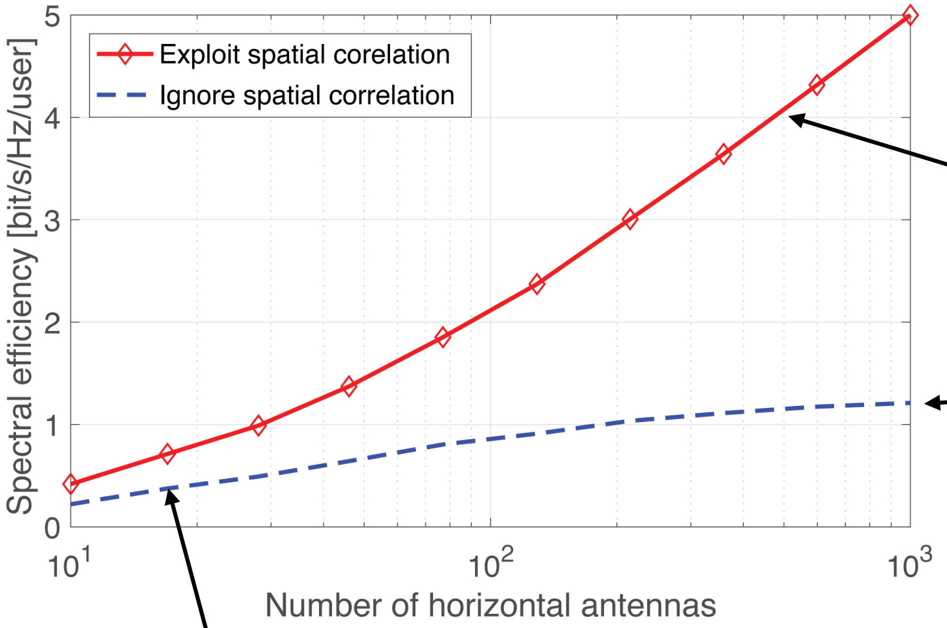
Linear array (192 x 1): 8 m x 0.04 m



We need a very large horizontal array!

Users are mainly separable horizontally

We are Far From *Truly* Massive MIMO



No upper limit,
MMSE, correlated fading
(SBH, TWC18)

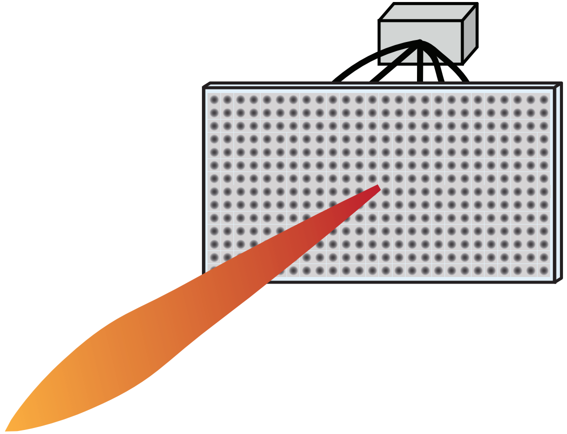
Finite upper bound
MR, i.i.d. fading
(Marzetta, TWC10)

5G base stations: 8-16 antennas per row

T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," TWC10.
E. Björnson, J. Hoydis, L. Sanguinetti, "Massive MIMO Has Unlimited Capacity," TWC18.

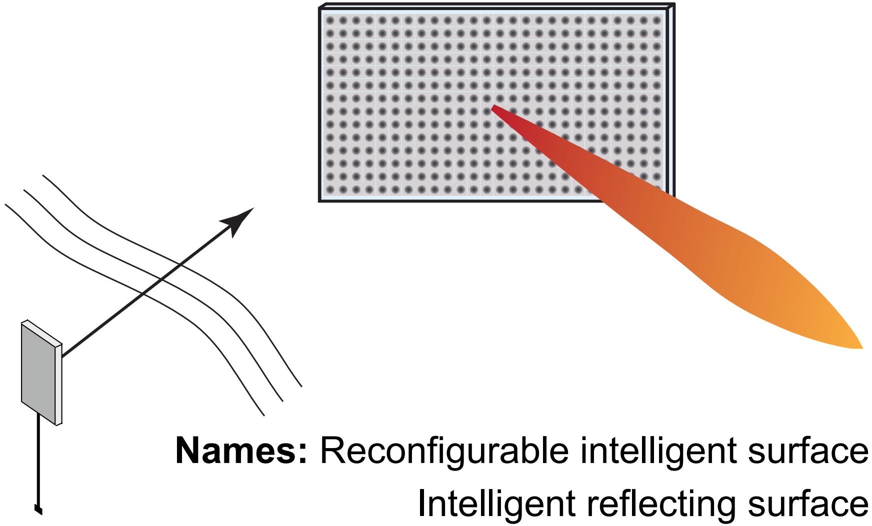
Two Ways to Build Much Larger Arrays

Active arrays



Names: Holographic Massive MIMO
Large intelligent surface
Reconfigurable reflectarrays

Semi-passive arrays

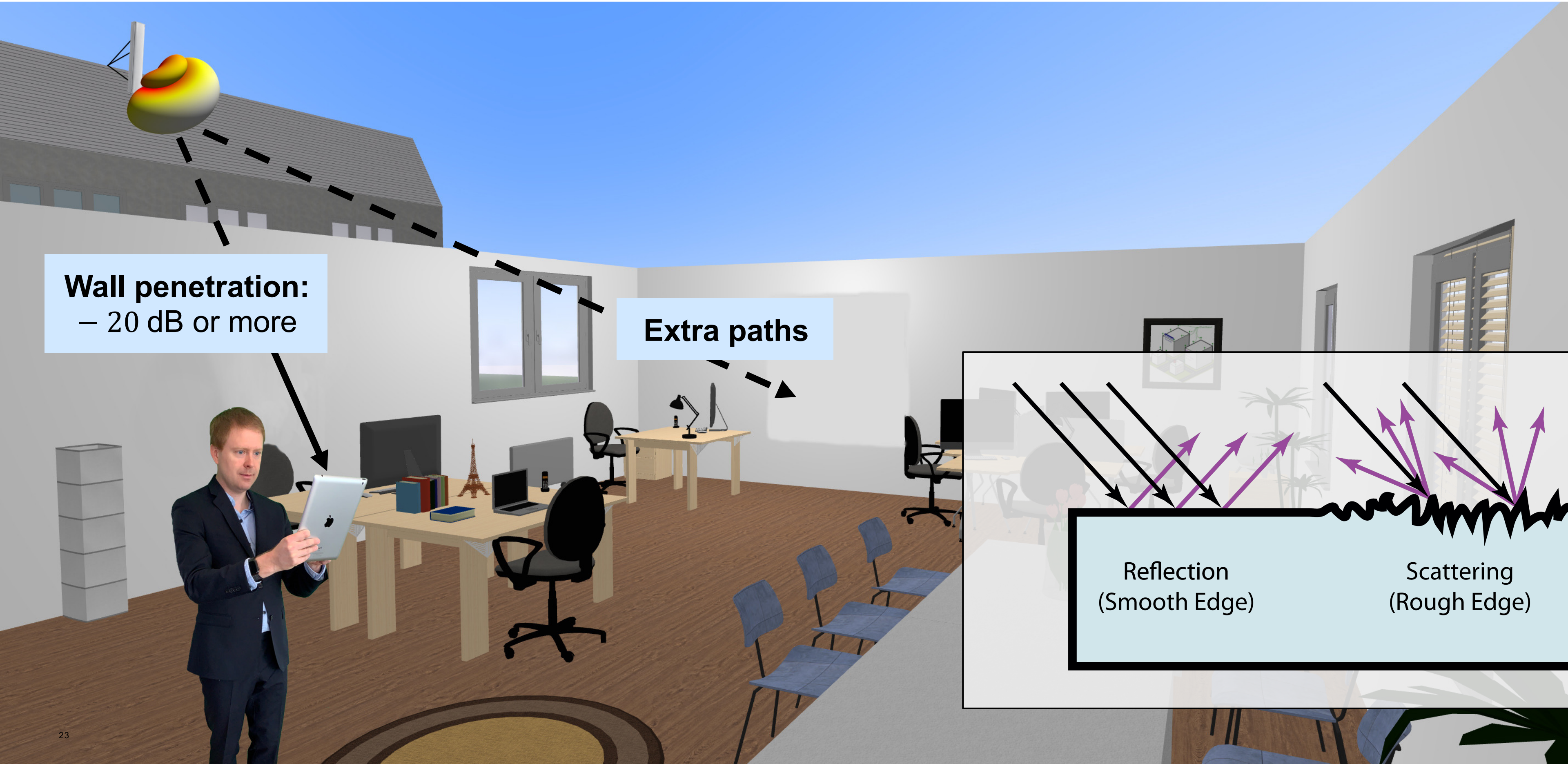


Names: Reconfigurable intelligent surface
Intelligent reflecting surface
Software-controlled metasurface

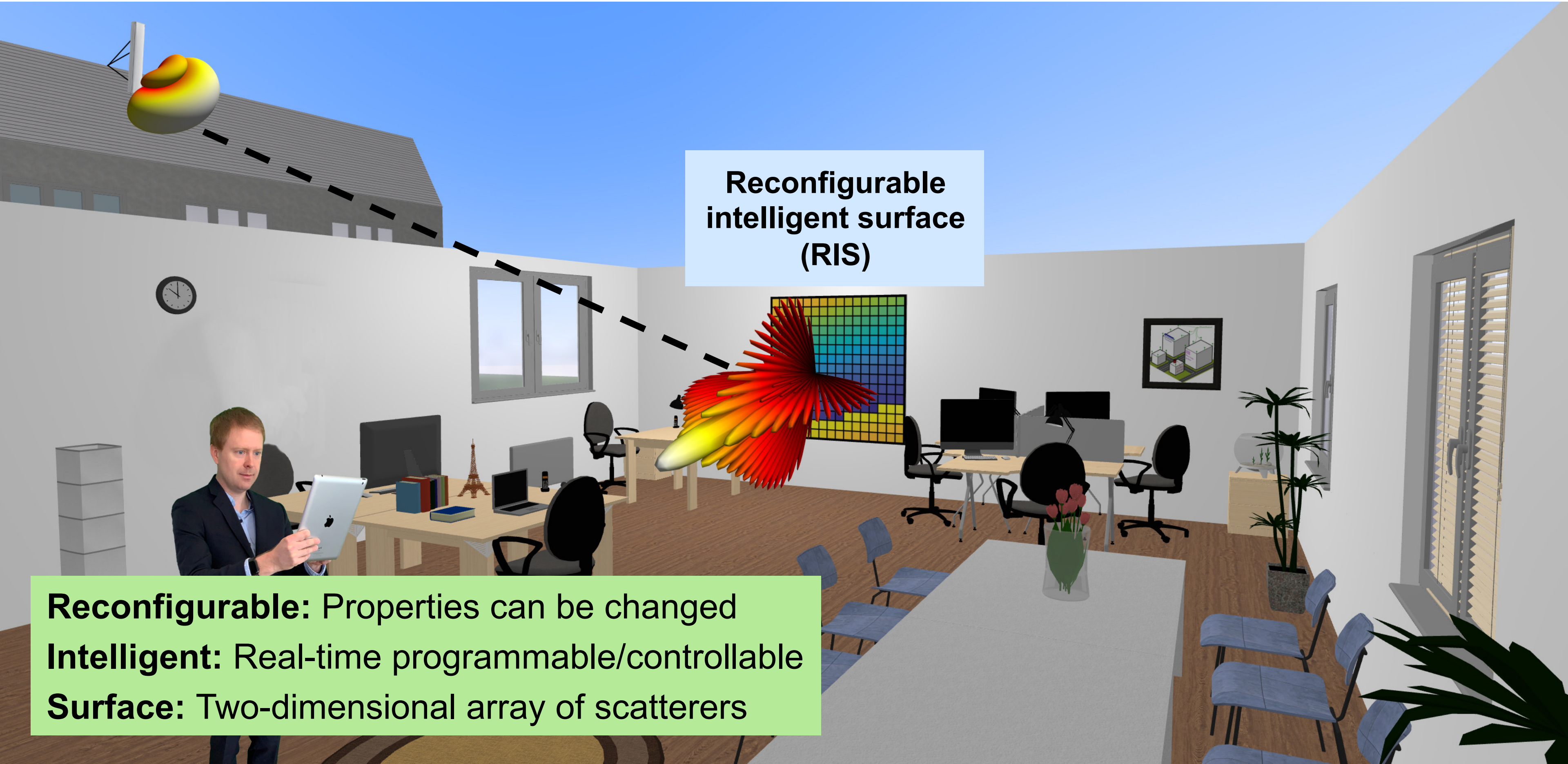
Common feature: Physically larger than current technology

FUNDAMENTALS AND VISION OF RECONFIGURABLE INTELLIGENT SURFACES (RIS)

No Direct Path: Even Larger Propagation Losses



Shaping Scattering Towards the Receiver

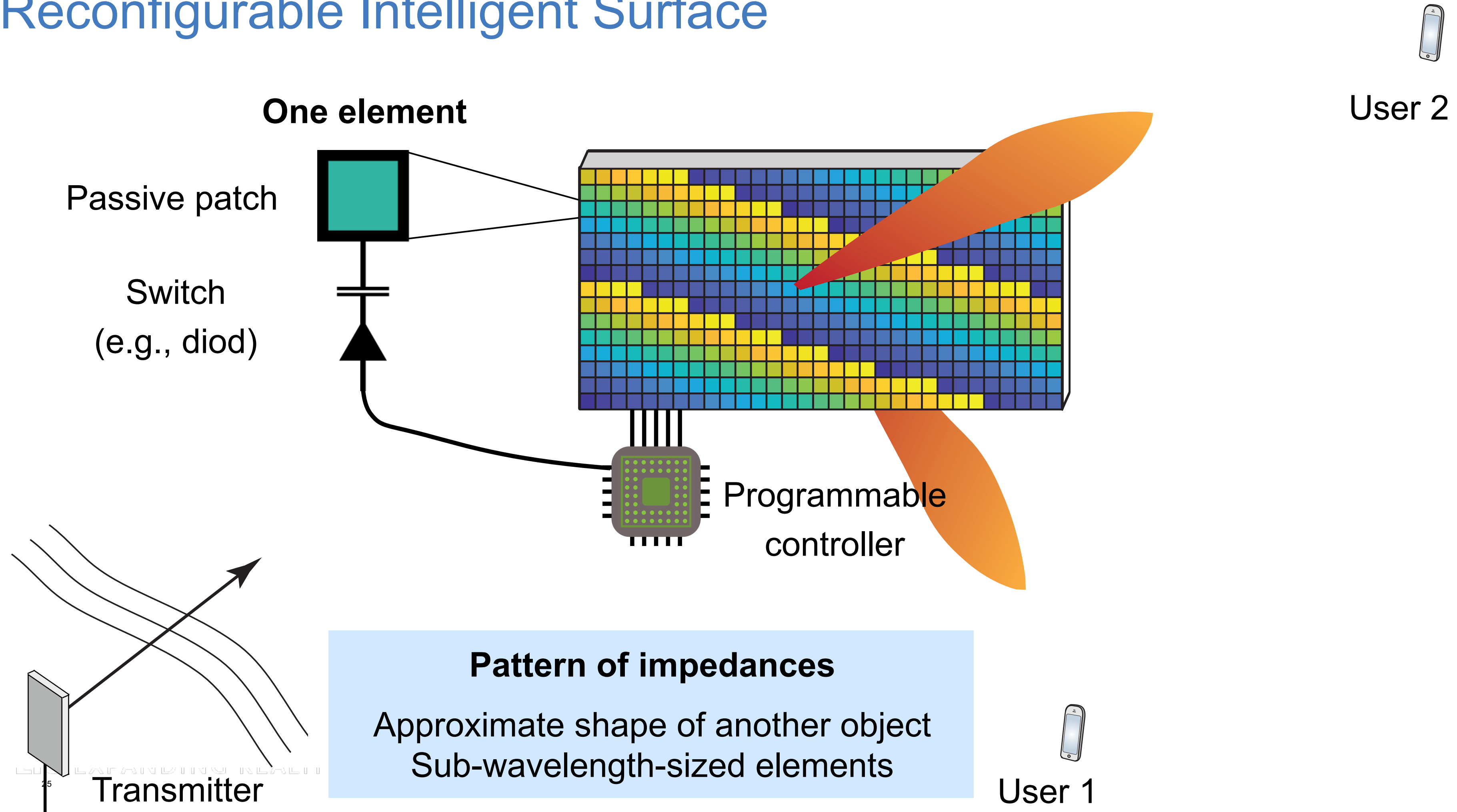


Reconfigurable: Properties can be changed

Intelligent: Real-time programmable/controllable

Surface: Two-dimensional array of scatterers

Reconfigurable Intelligent Surface

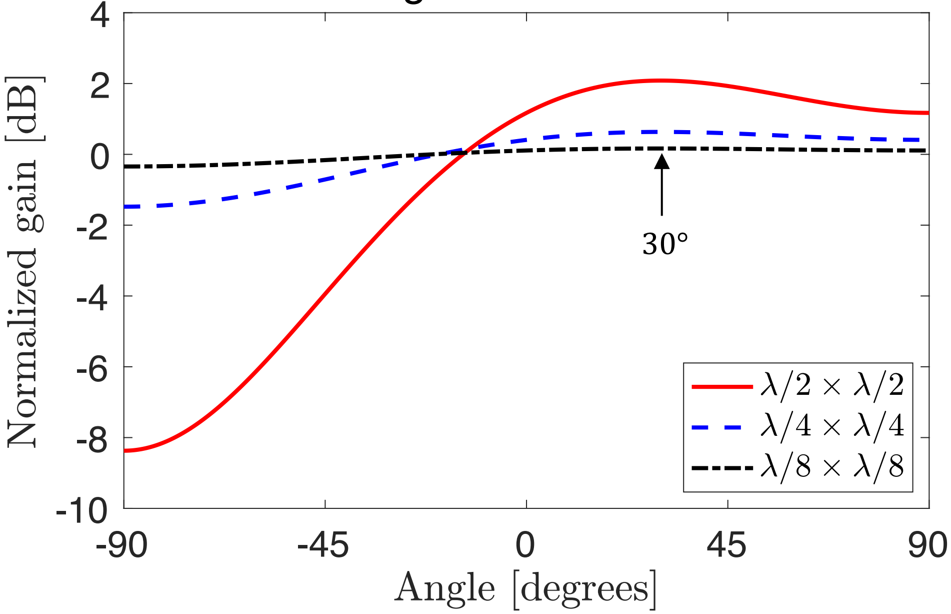


Pattern of impedances

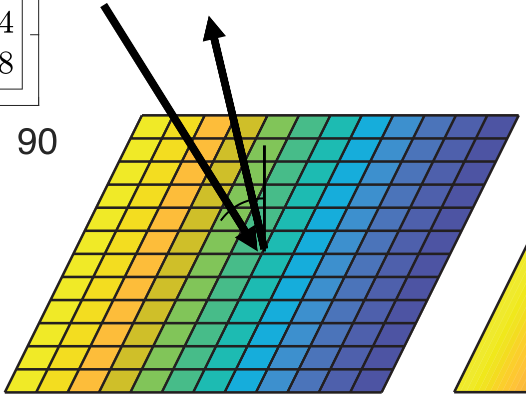
Approximate shape of another object
Sub-wavelength-sized elements

How Large are the Elements?

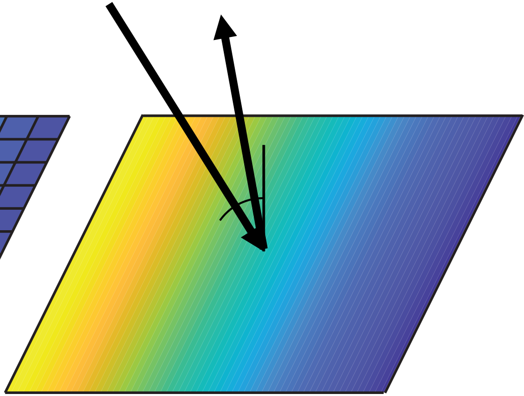
When a signal arrives from $-30°$:



Discretized reconfiguration

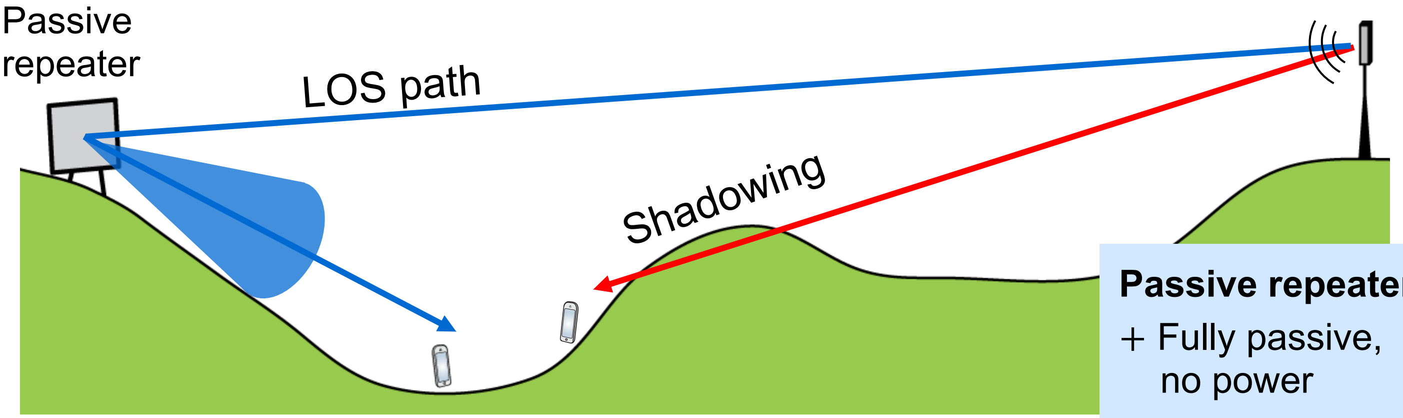


Ideal continuous reconfiguration



Each element should scatter signals almost isotropically

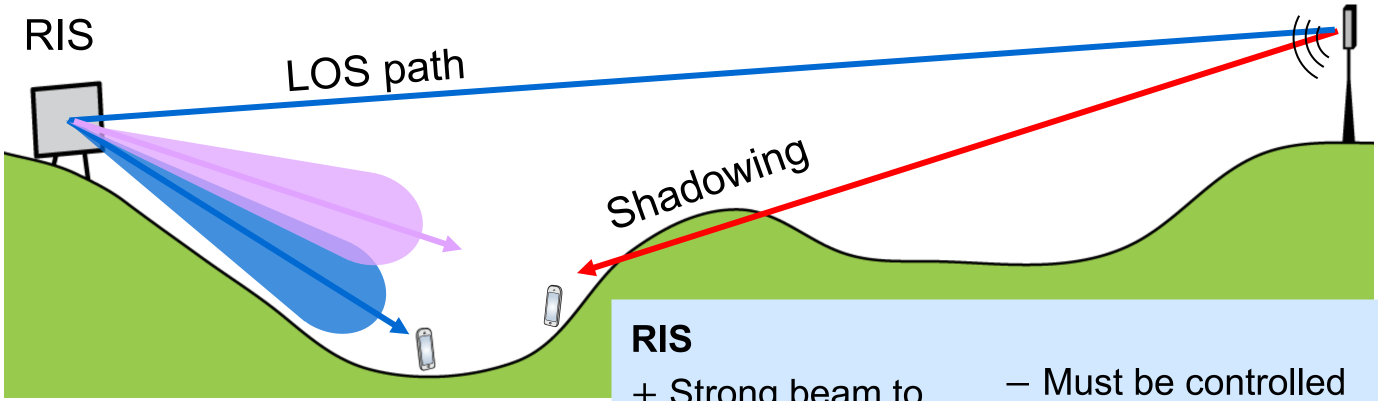
Passive Repeaters versus RIS



Passive repeater

+ Fully passive, no power

– Fixed beam pattern: wide or strong?



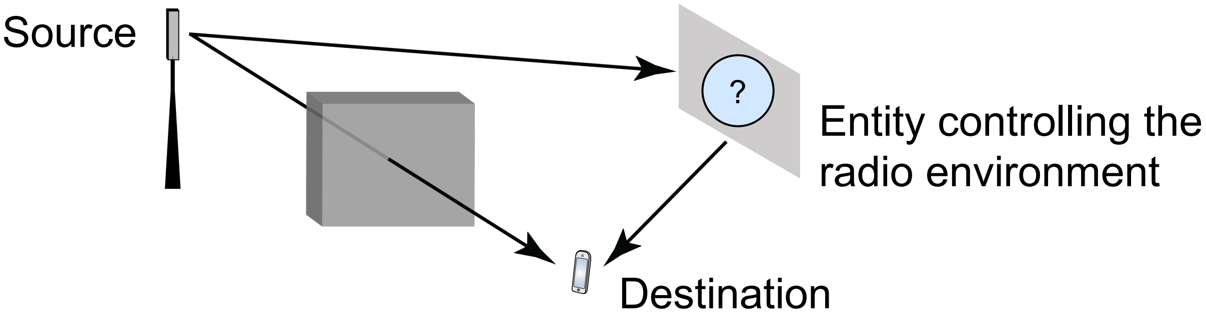
RIS

+ Strong beam to desired location

– Must be controlled

– Requires power

Taxonomy of “Cooperative Communications”



Used in cellular networks

Old-school analog repeater

	Regenerative	Transparent
Half duplex	Decode-and-forward (DF) relay	Amplify-and-forward (AF) relay
Full duplex	**AF and DF relay**	Passive repeater, **RIS**

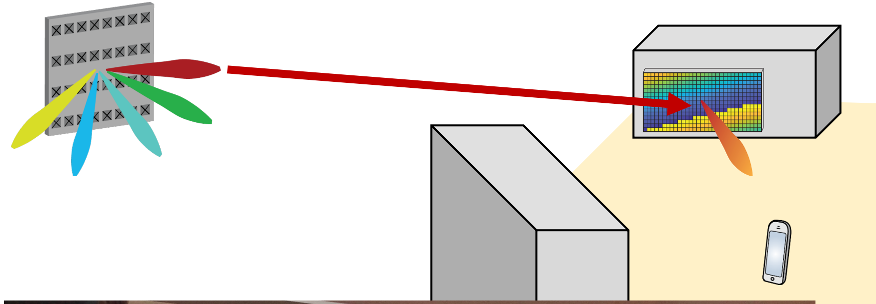
Emerging technologies

Not one or the other!
Different choice in different setups

Vision: Controllable Propagation

RIS ***as a whole*** can control

- Directivity of scattered signal
- Signal absorption
- Change polarization



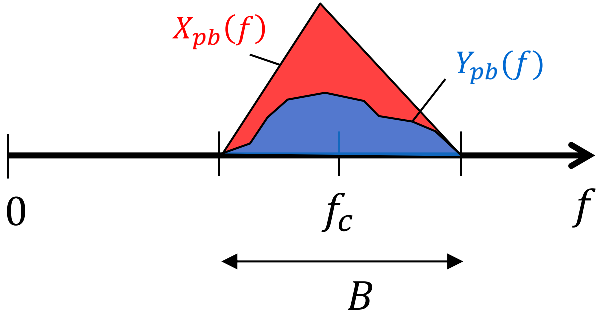
Improved indoor coverage

Mitigate shadow fading

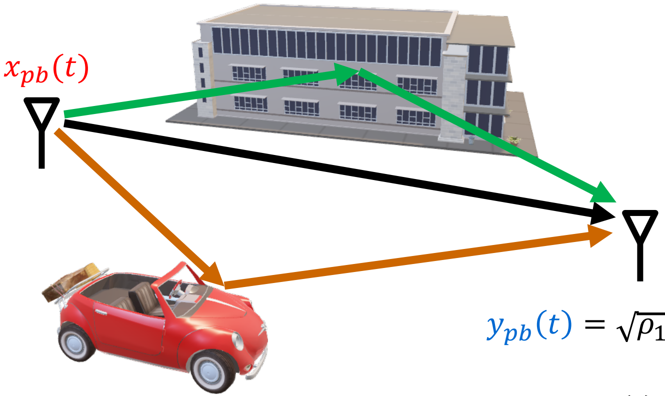
SYSTEM MODEL FOR
RIS-AIDED COMMUNICATIONS

Basics of Signals and Systems

$x_{pb}(t) \longrightarrow$ **Linear-time invariant:** *Impulse response* $h_{pb}(t)$ $\longrightarrow$ $y_{pb}(t) = (h_{pb} * x_{pb})(t) = \int_{\infty}^{-\infty} h_{pb}(u) x_{pb}(t-u)\,du$



- Communication channels are systems/filters:



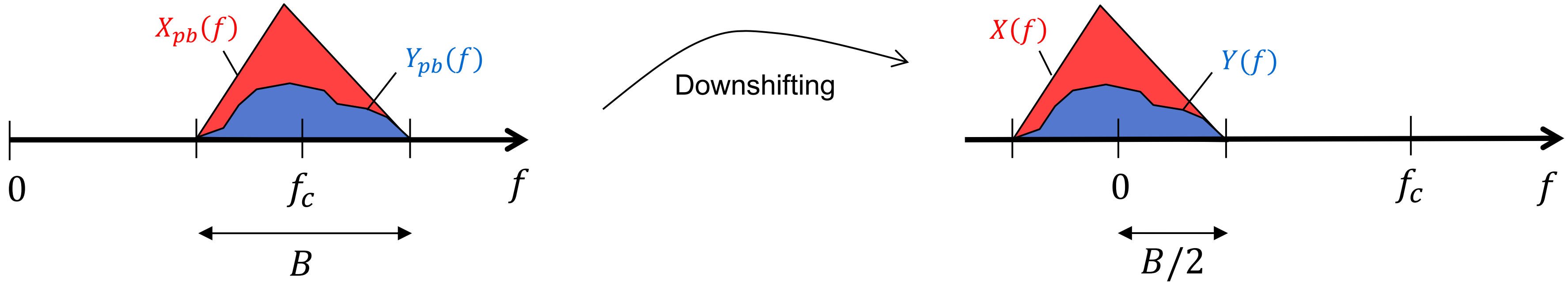
$$y_{pb}(t) = \sqrt{\rho_1} x_{pb}(t-\tau_1) + \sqrt{\rho_2} x_{pb}(t-\tau_2) + \sqrt{\rho_3} x_{pb}(t-\tau_3)$$

Impulse response: $h_{pb}(t) = \sqrt{\rho_1}\delta(t-\tau_1) + \sqrt{\rho_2}\delta(t-\tau_2) + \sqrt{\rho_3}\delta(t-\tau_3)$

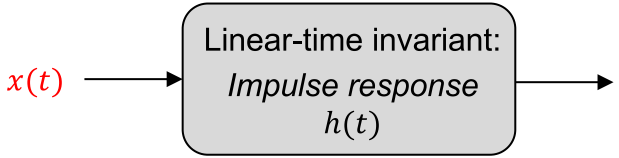
Complex Baseband Representation

- Communication theory is developed for the baseband

Real passband | **Complex baseband**



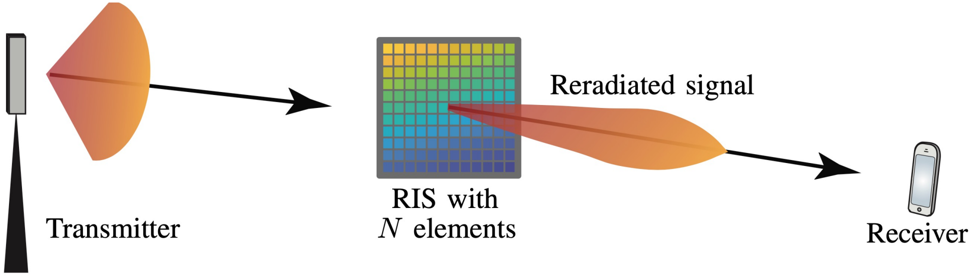
- Connection: $X_{pb}(f) = \frac{X(f-f_c)+X^*(-f-f_c)}{\sqrt{2}}$, $Y_{pb}(f) = \frac{Y(f-f_c)+Y^*(-f-f_c)}{\sqrt{2}}$



$y(t) = \int_{\infty}^{-\infty} h(u)x(t-u)du$

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

Analyzing Reconfigurable Intelligent Surface



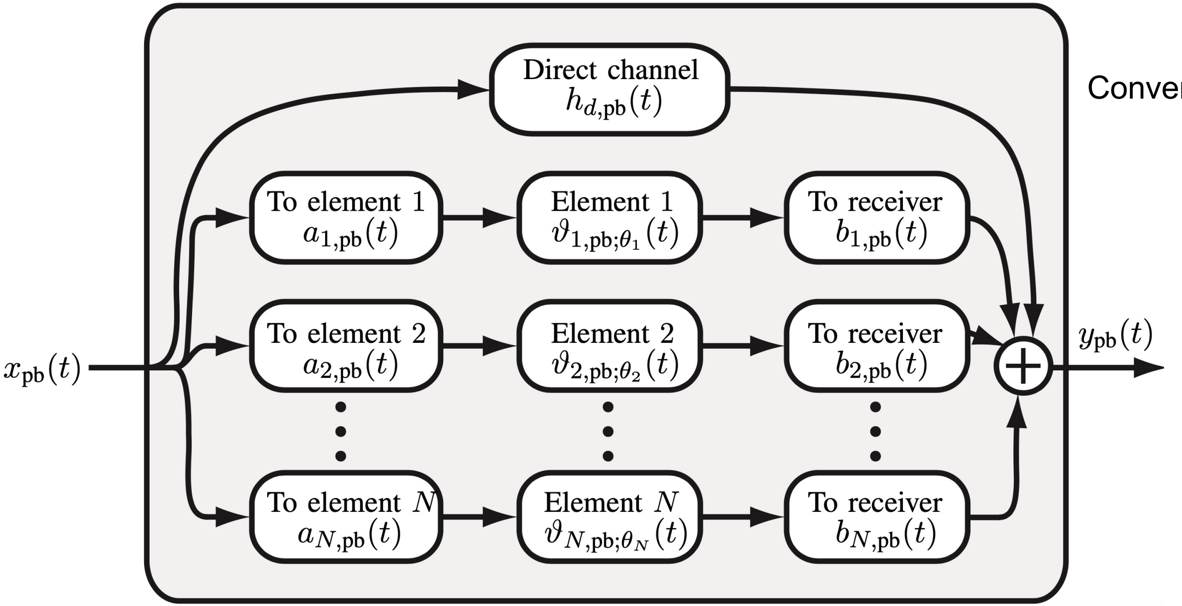
End-to-end impulse response:

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

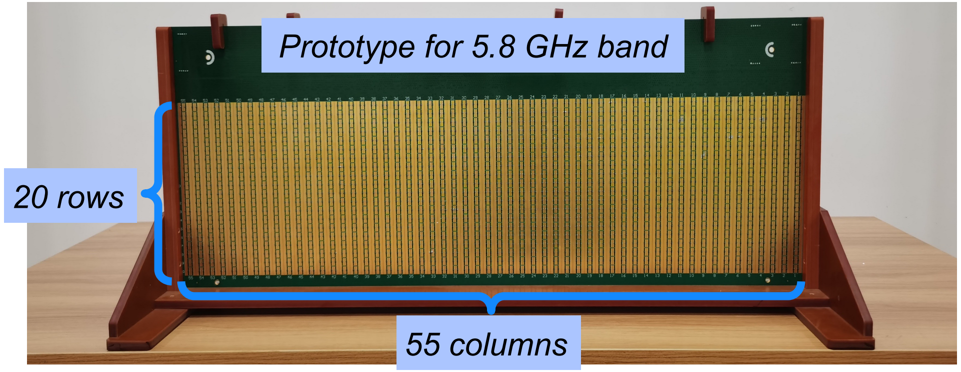
Conventional channel models

Controlled by RIS using $\theta_1, \ldots, \theta_N$

End-to-end system with impulse response $h_{\mathrm{pb}}(t)$



How Will an RIS Element Filter the Signal?

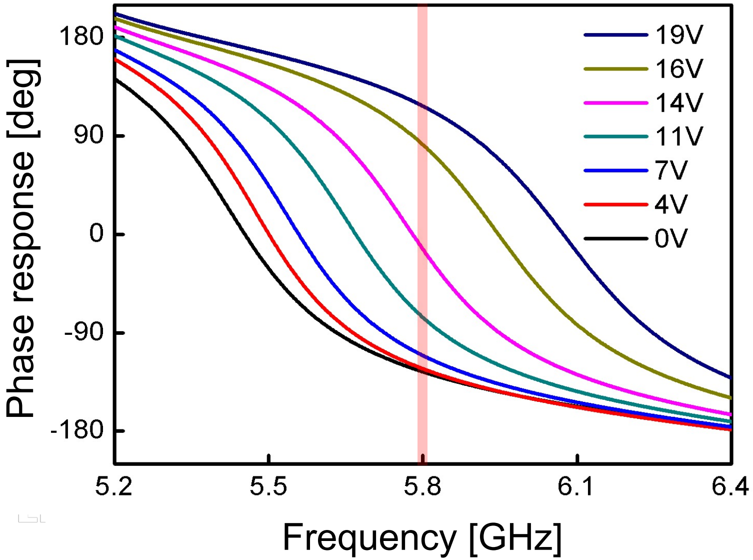


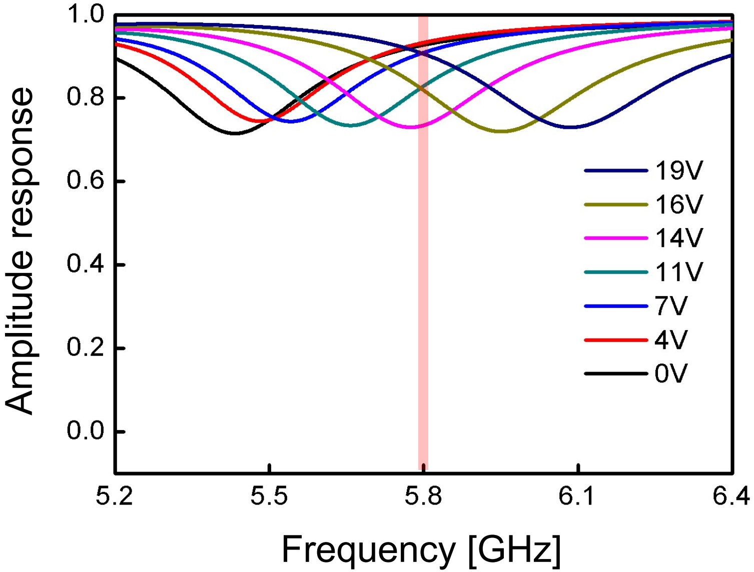
Example: Patch with bias voltage V

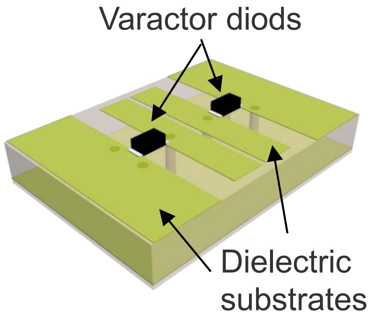
Reflection coefficient:

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

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," arXiv:2103.00534







How To Transmit Data?

- Pulse amplitude modulation:

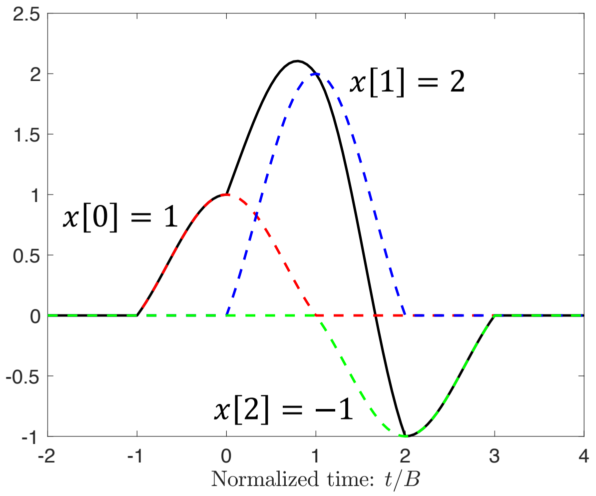
$$x(t) = \sum_m x[m] p\left(t - \frac{m}{B}\right)$$

  - Transmit discrete sequence: $x[m], \; m = \text{integer}$

Use a pulse-form $p(t)$ satisfying the Nyquist criterion:

$p\left(\frac{m}{B}\right) = 0$ for integer $m \neq 0$ and non-zero for $m = 0$

  - **Example:** $p(t) = \sqrt{B}\,\mathrm{sinc}(Bt)$



Sampling of received signal $y(t) = x(t)$:

$$y\left(\frac{k}{B}\right) = x\left(\frac{k}{B}\right) = \sum_m x[m] p\left(\frac{k-m}{B}\right) = x[k]$$

Reception with Channel and Noise

- Received signal (with Gaussian noise):

$$y(t) = (h * x)(t) + w(t)$$

1. Filter using $p(t) = \sqrt{B}\mathrm{sinc}(Bt)$:

$$z(t) = (p * y)(t) = \sum_m x[m]\,(p * h * p)\left(t - \frac{m}{B}\right) + (p * w)(t)$$

2. Sample received signal:

$$\underbrace{z\left(\frac{k}{B}\right)}_{\text{Call it } z[k]} = \sum_m x[m]\,\underbrace{(p * h * p)}_{\text{Effective pulse function}}\left(\frac{k-m}{B}\right) + \underbrace{(p * w)\left(\frac{k}{B}\right)}_{\text{Complex Gaussian noise } CN(0, N_0)}$$

Narrowband channel: $h \approx \text{constant} \cdot \delta(t - \tau)$ in the band, Nyquist criterion satisfied

$$z[k] = \text{constant} \cdot\; x[k] + \text{Gaussian noise}$$

Narrowband: Putting the Pieces Together

- Direct channel: $h_{d,pb}(t) = \sqrt{\rho}\,\delta(t - \tau_d) \quad \rightarrow h_d(t) = \sqrt{\rho}\,e^{-j2\pi f_c t}\,\delta(t - \tau_d)$

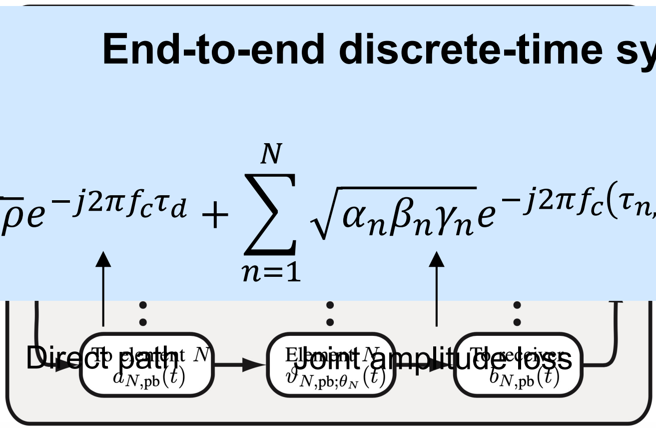
- Related to element n:

$$a_{n,pb}(t) = \sqrt{\alpha_n}\,\delta(t - \tau_{n,a}) \quad \rightarrow \quad a_n(t) = \sqrt{\alpha_n}\,e^{-j2\pi f_c t}\,\delta(t - \tau_{n,a})$$

$$\vartheta_{n,pb;\theta_n}(t) = \sqrt{\gamma_n}\,\delta(t - \tau_{\theta_n}) \quad \rightarrow \quad \vartheta_{n;\theta_n}(t) = \sqrt{\gamma_n}\,e^{-j2\pi f_c t}\,\delta(t - \tau_{\theta_n})$$

$$b_{n,pb}(t) = \sqrt{\beta_n}\,\delta(t - \tau_{n,b}) \quad \rightarrow \quad b_n(t) = \sqrt{\beta_n}\,e^{-j2\pi f_c t}\,\delta(t - \tau_{n,b})$$

End-to-end system with impulse response $h_{\mathrm{pb}}(t)$



End-to-end discrete-time system model:

$$z[k] = \left(\sqrt{\rho}\,e^{-j2\pi f_c \tau_d} + \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n}\,e^{-j2\pi f_c(\tau_{n,a}+\tau_{\theta_n}+\tau_{n,b})}\right) x[k] + \text{Noise}$$

Direct path

Joint amplitude loss

Joint delay

Tunable!

OPTIMIZING NARROWBAND COMMUNICATION PERFORMANCE

Maximizing Performance Without a Direct Path

Received signal without direct path:

$$y = \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} e^{-j2\pi f_c (\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b})} \cdot \text{signal} + \text{noise}$$

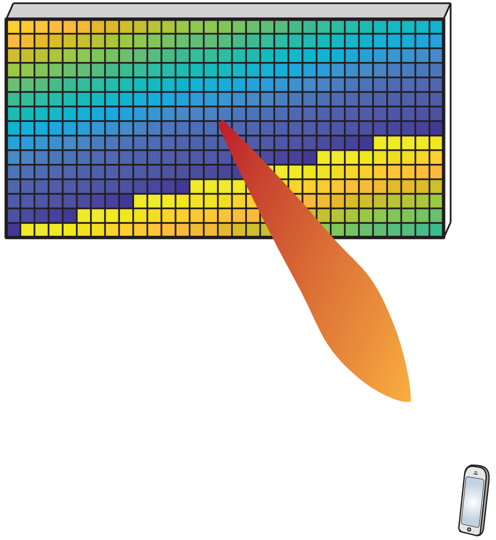
Signal processing problem:

Maximize the signal-to-noise ratio

$$\left| \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} e^{-j2\pi f_c (\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b})} \right|^2 \leq \left| \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} \right|^2 \approx N^2 \alpha\beta\gamma$$

If $\alpha_n \beta_n \gamma_n = \alpha\beta\gamma$ for all n

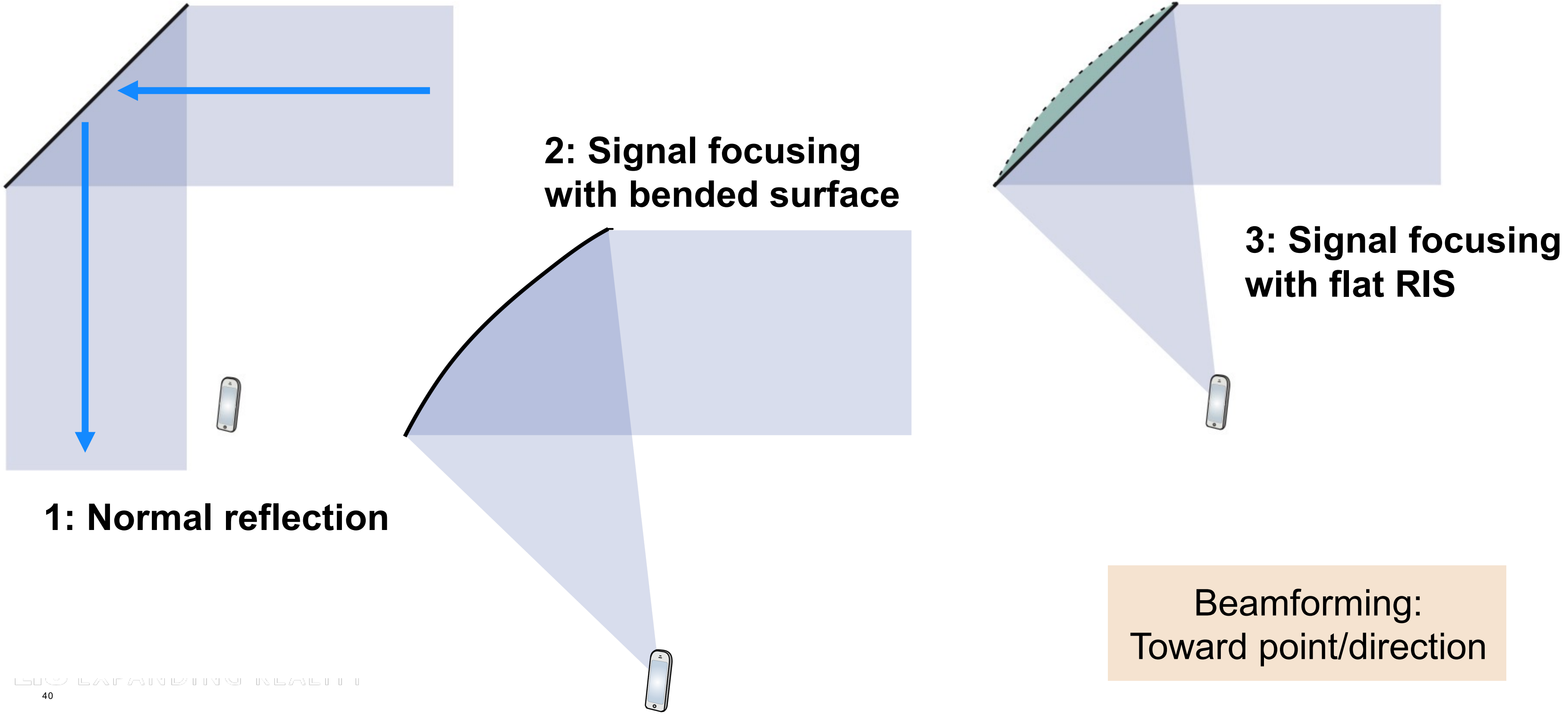
Cauchy–Schwarz inequality
$(\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b} = \text{constant})$



Minimum positive delay solution:

$$\tau_{\theta_n} = \max_m \left(\tau_{m,a} + \tau_{m,b} \right) - \left(\tau_{n,a} + \tau_{n,b} \right)$$

Example: Synthesizing Surface Shapes



Beamforming:
Toward point/direction

Maximizing Performance With a Direct Path

Received signal with direct path:

$$y = \left(\sqrt{\rho} e^{-j2\pi f_c \tau_d} + \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} e^{-j2\pi f_c \left(\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b}\right)} \right) \cdot \text{signal} + \text{noise}$$

Maximize channel gain:

$$\left| \sqrt{\rho} e^{-j2\pi f_c \tau_d} + \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} e^{-j2\pi f_c \left(\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b}\right)} \right|^2 \leq \left| \sqrt{\rho} + \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} \right|^2$$

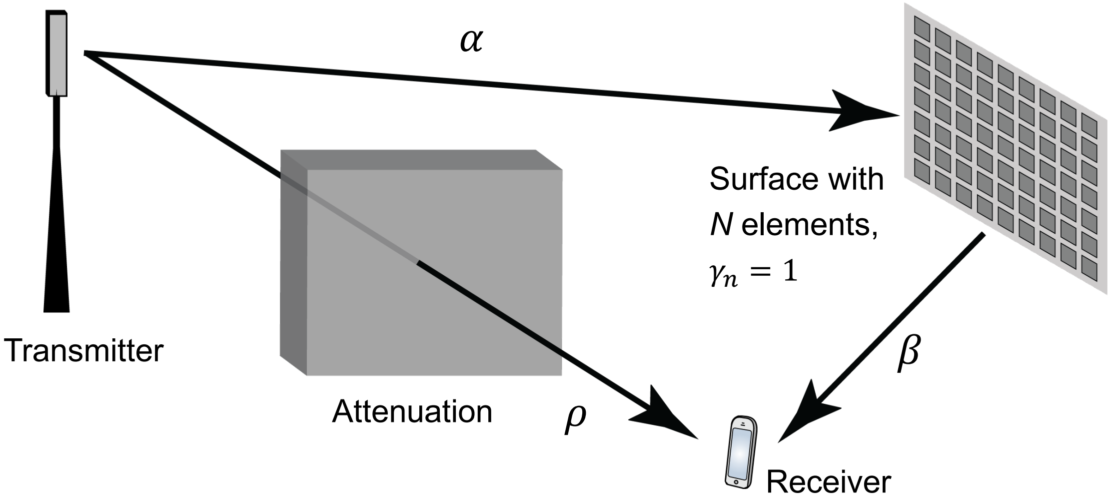
Achieved when:

$$\tau_{n,a} + \tau_{\theta_n} + \tau_{n,b} = \tau_d$$

Minimum positive delay solution:

$$\tau_{\theta_n} = \tau_d - \left(\tau_{n,a} + \tau_{n,b}\right) + \frac{\text{integer}}{f_c}$$

Basic Use Case: Coverage Extension



Product of pathlosses

Low SNR with one element

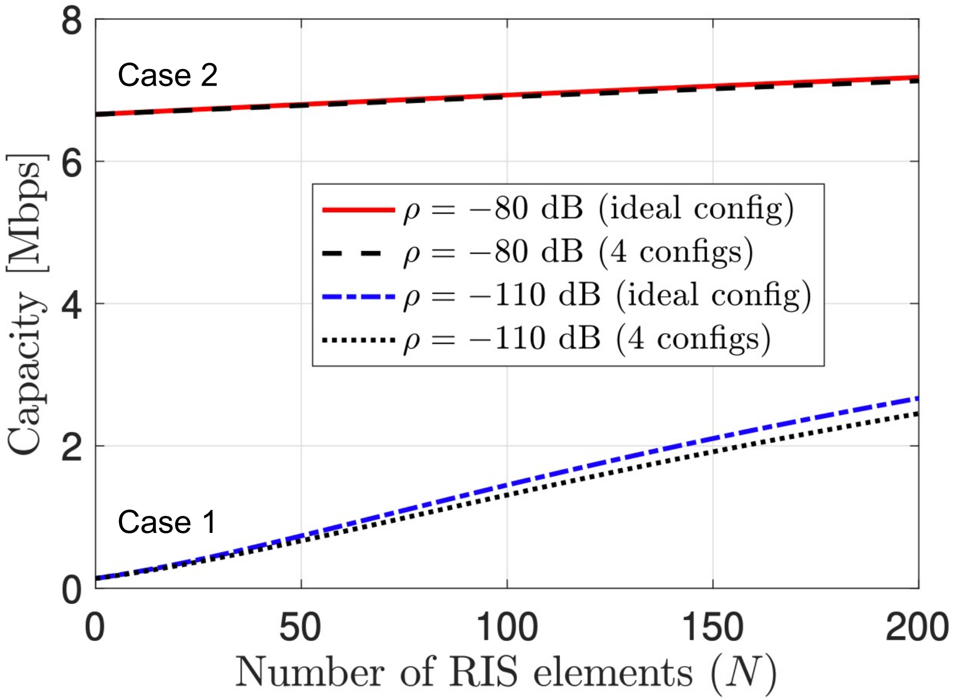
Squaring effect

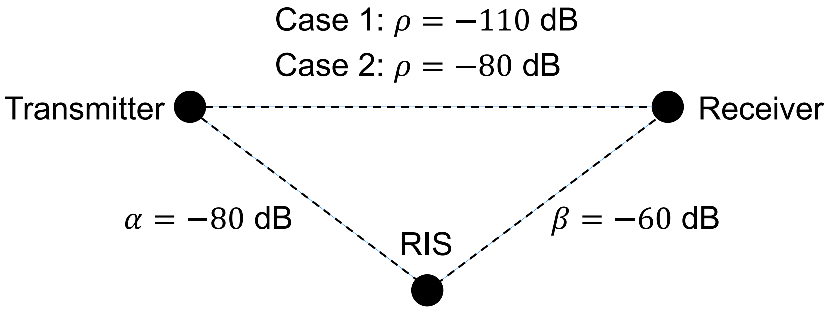
SNR grows with square of N

Spectral efficiency: $\log_2(1 + \text{SNR})$

$$\text{SNR} = \left| \sqrt{\rho} + \sum_{n=1}^{N} \sqrt{\alpha_n \beta_n \gamma_n} \right|^2 = \left| \sqrt{\rho} + N\sqrt{\alpha\beta} \right|^2 \approx \begin{cases} \rho, & \text{small } N \\ N^2 \alpha\beta, & \text{large } N \end{cases}$$

Basic Performance Benefit





Bandwidth: $B = 1$ MHz

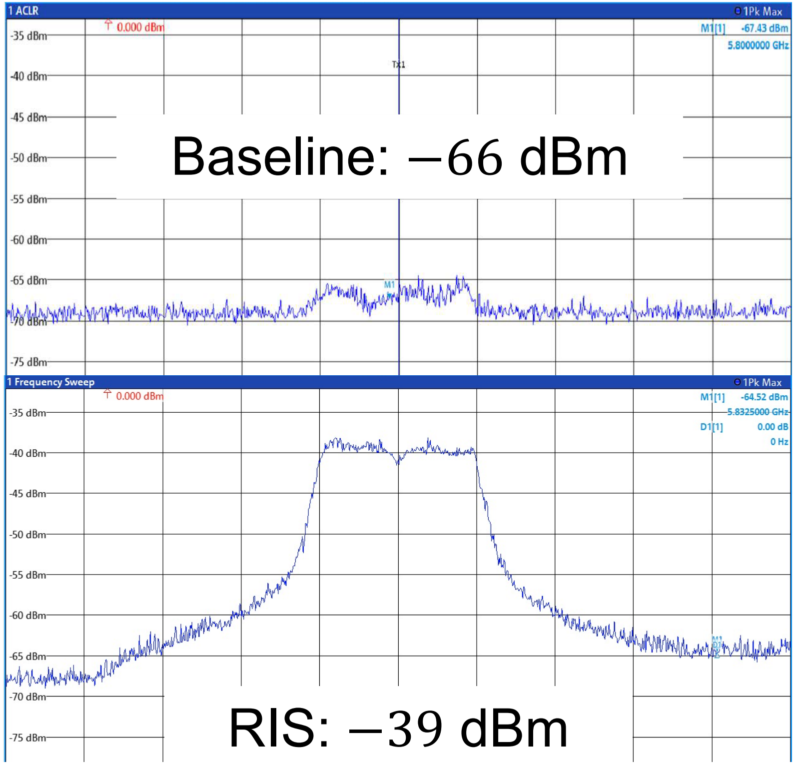
Transmit SNR: 100 dB

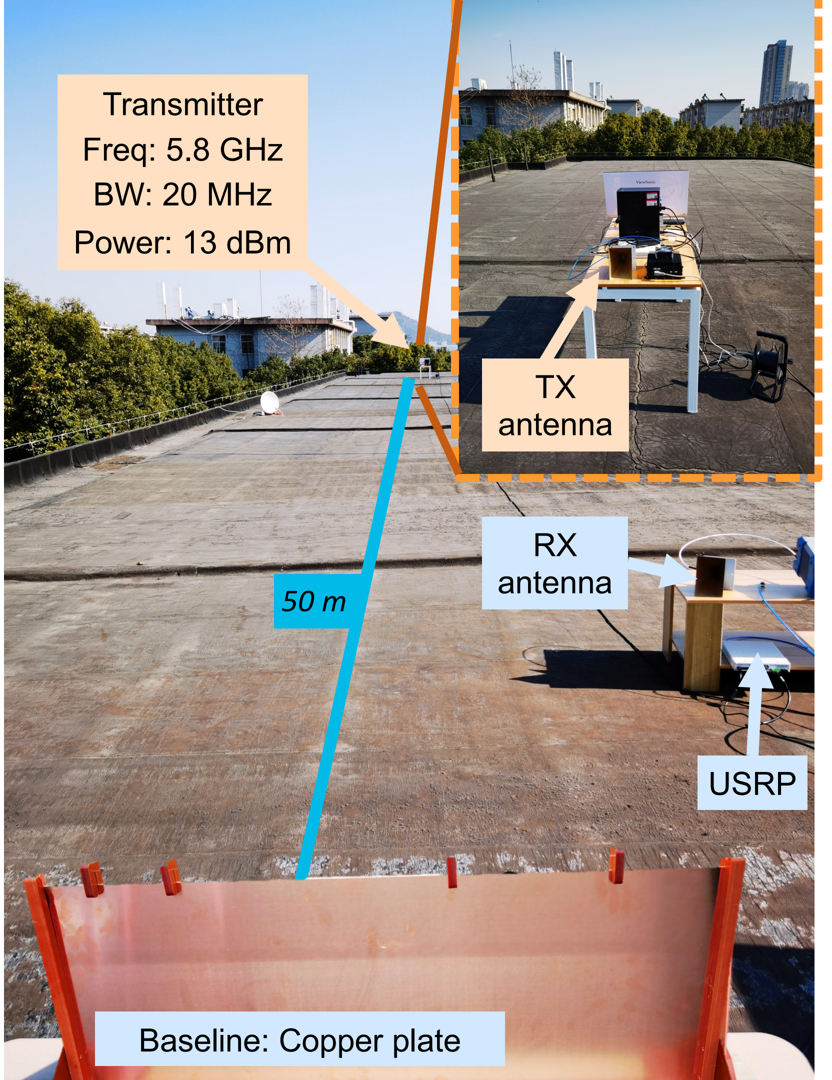
RIS is particularly helpful

when direct path is relatively weak

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.*

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," arXiv:2103.00534

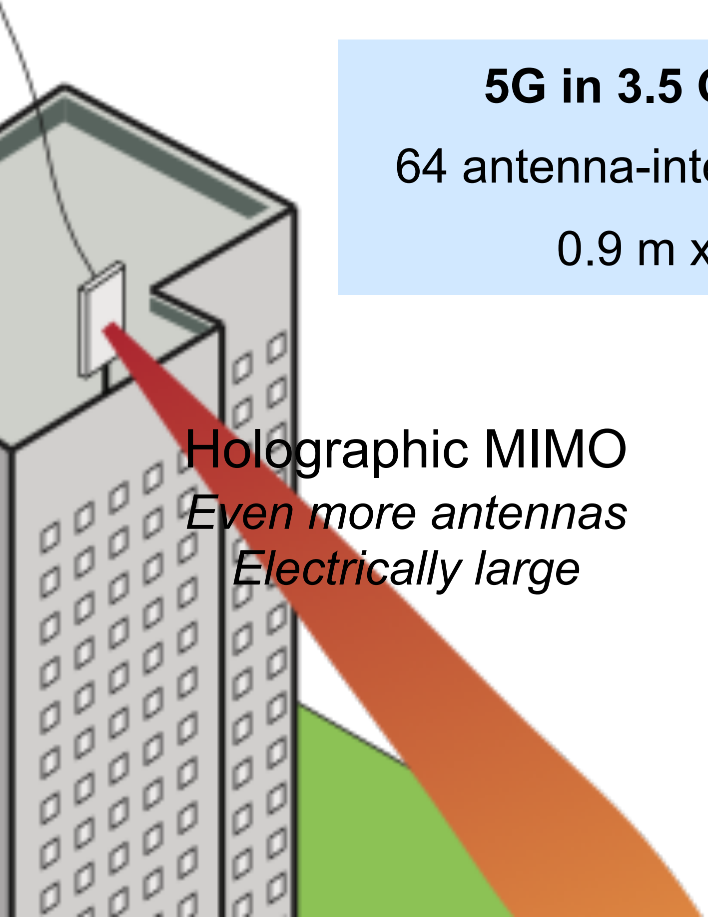
HOLOGRAPHIC MASSIVE MIMO

Massive MIMO vs. Holographic Massive MIMO

5G in 3.5 GHz band

64 antenna-integrated radios

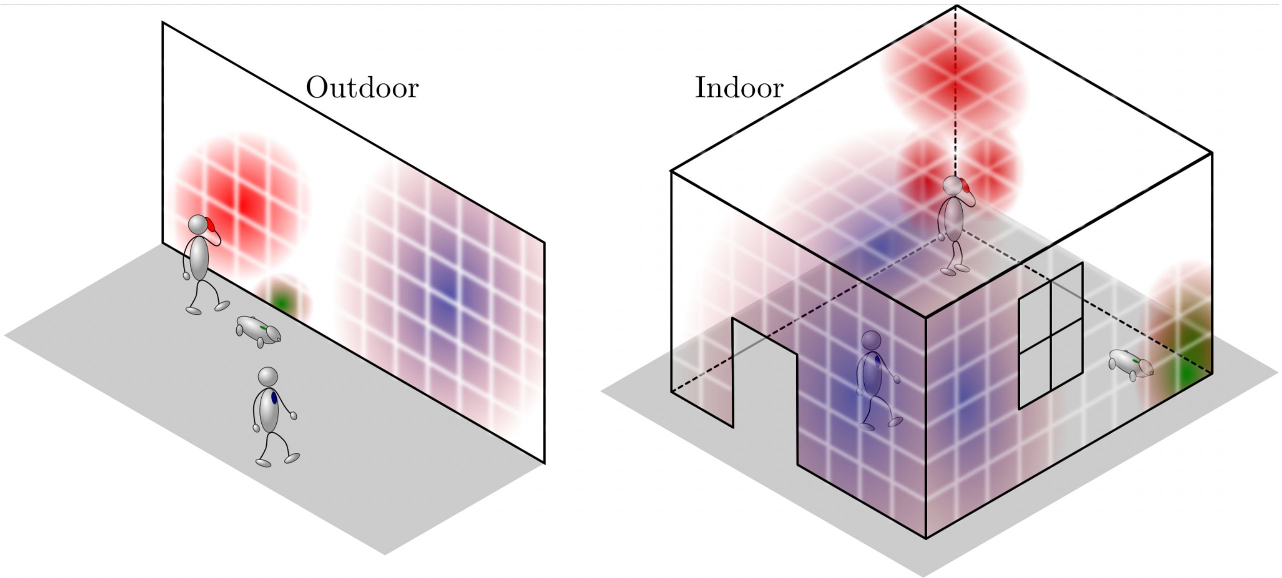
0.9 m x 0.5 m



Array gain

Spatial multiplexing gain

Making Surrounding Surfaces Intelligent



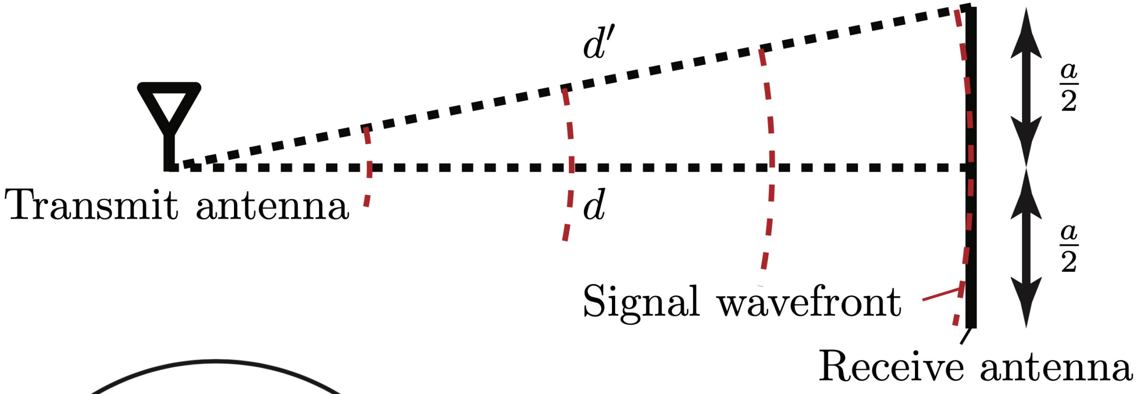
Use cases
Data transmission
Positioning
Sensing

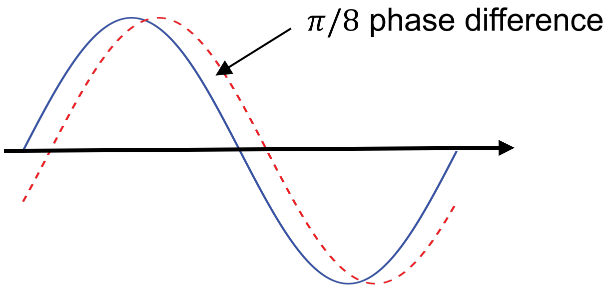
High spatial resolution: New propagation models needed

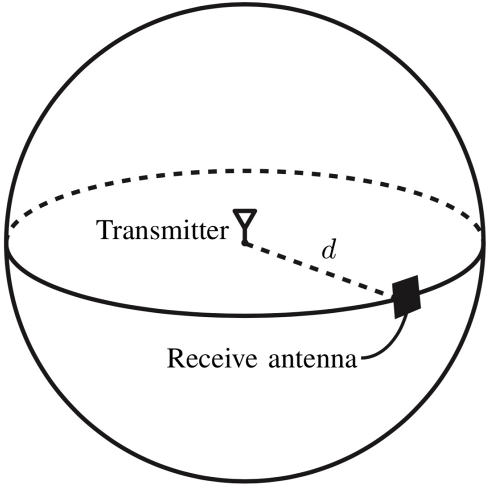
Near-field effects, realistic scattering conditions

S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," IEEE TSP, 2018.

Near-Field and Far-Field of an Antenna







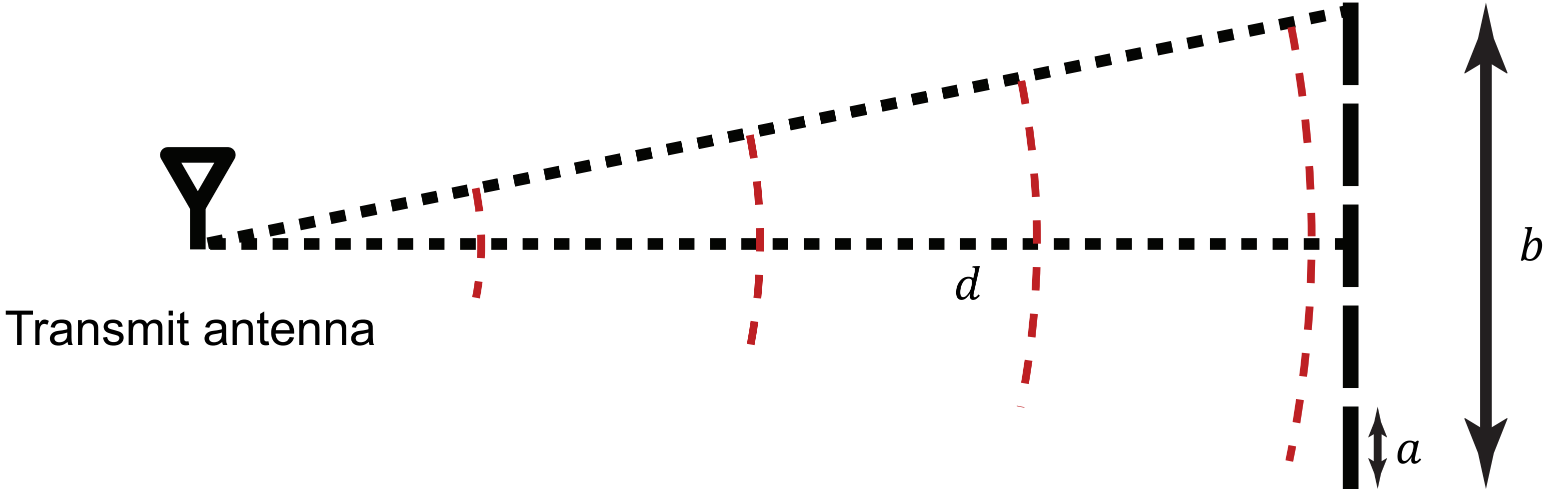
- Length difference: $d' - d = \sqrt{d^2 + \left(\frac{a}{2}\right)^2} - d \approx \frac{a^2}{8d}$
- At most $\pi/8$ phase diff: $\frac{\pi}{8} \geq \frac{2\pi}{\lambda}\frac{a^2}{8d} \rightarrow d \geq \frac{2a^2}{\lambda}$

Fraunhofer distance

Other electromagnetic near-field effects

Evanescent waves, Magnetic induction, …

Near-Field When Having an Antenna Array



Example: 3 GHz, $\lambda = 0.1$ m

$b = 10\lambda = 1$ m: $\frac{2b^2}{\lambda} = 20$ m

$b = 100\lambda = 10$ m: $\frac{2b^2}{\lambda} = 2$ km

Practical operating regime

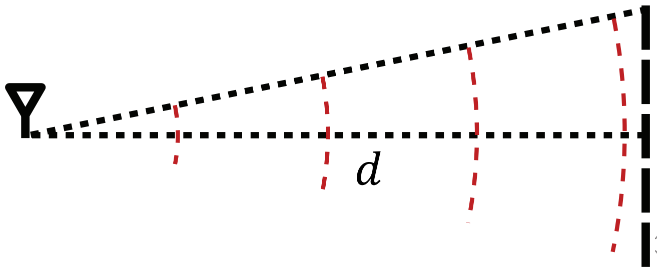
Far-field of antenna: $d \gg 2a^2/\lambda$

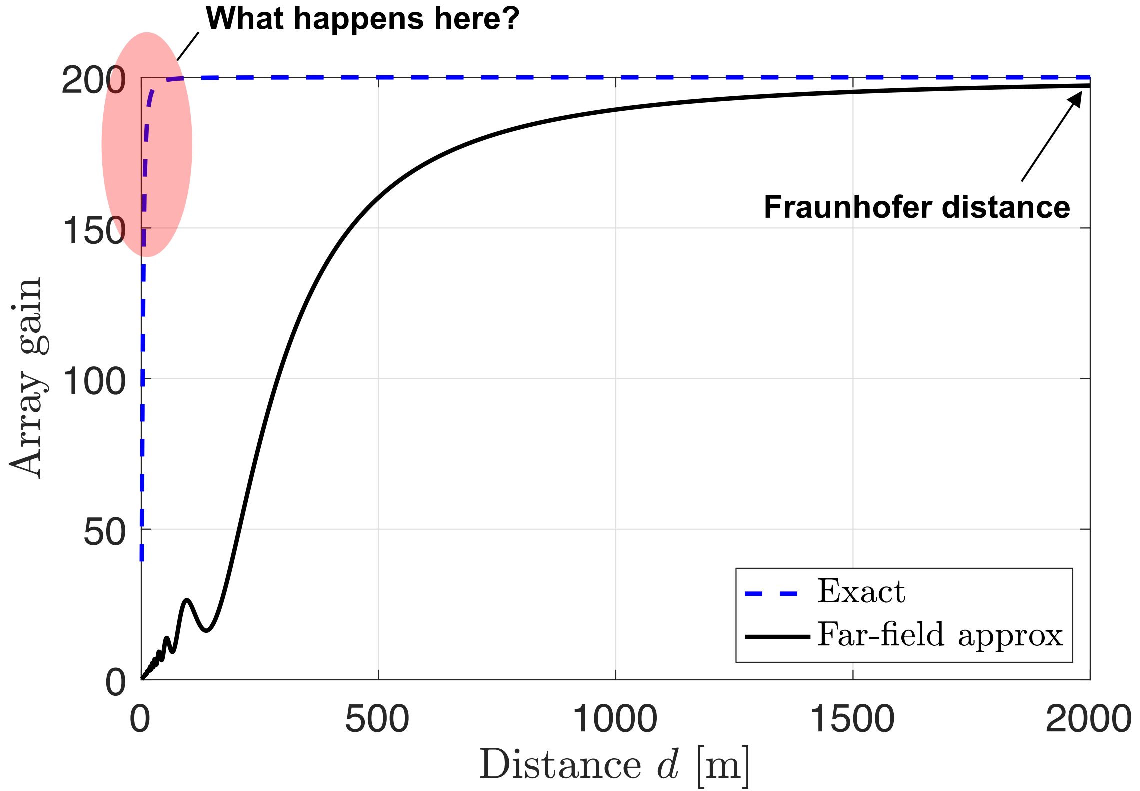
Near-field of array if $d \leq 2b^2/\lambda$

Spherical Waves in the Near-Field

Example: 3 GHz, $\lambda = 0.1$ m

200 antennas, $\lambda/2$ spacing





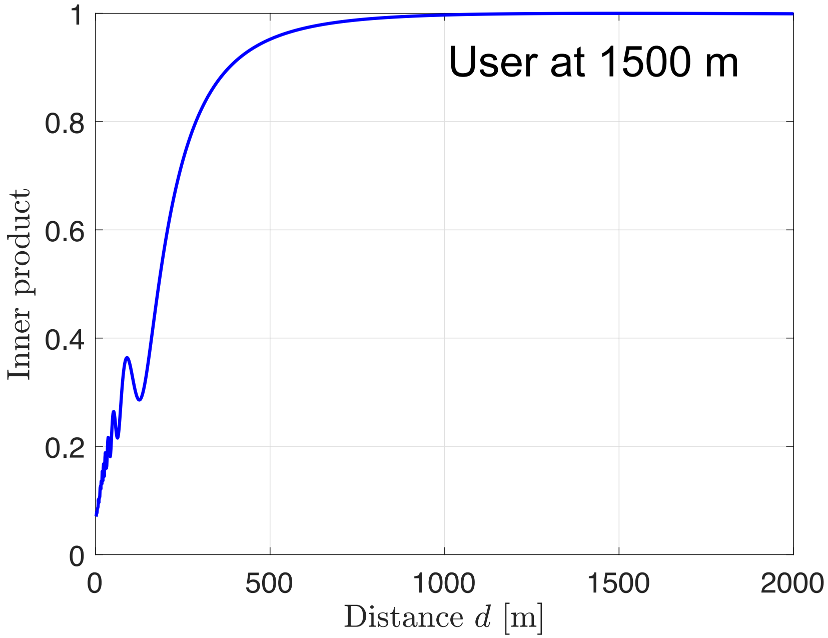
Array response

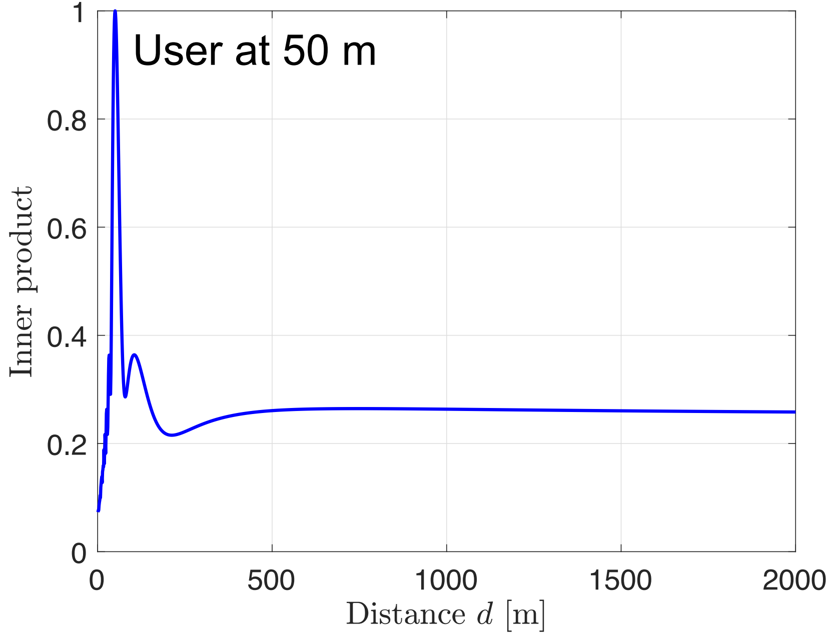
Depends on angle *and* distance

Depth Perception

Two users in same direction

One at varying distance d



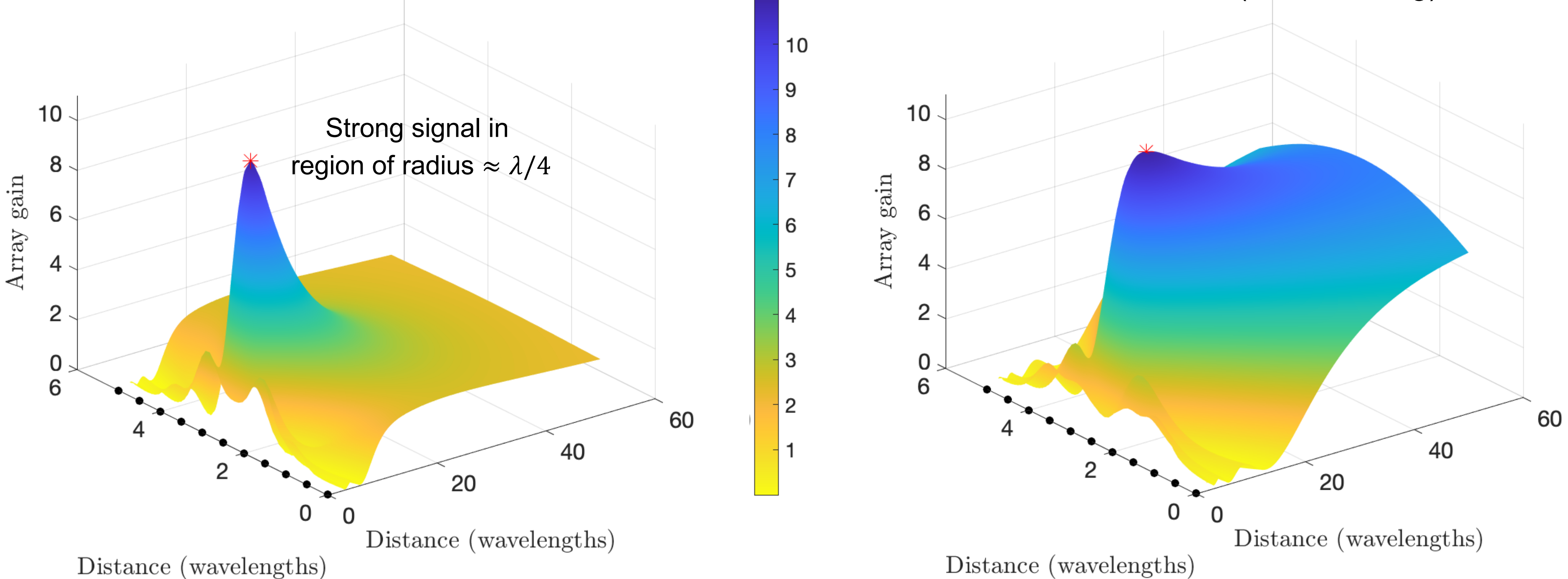


Near-Field Focusing: Array Gain

Fraunhofer distance with $a = 5\lambda$: $\frac{2a^2}{\lambda} = 50\lambda$

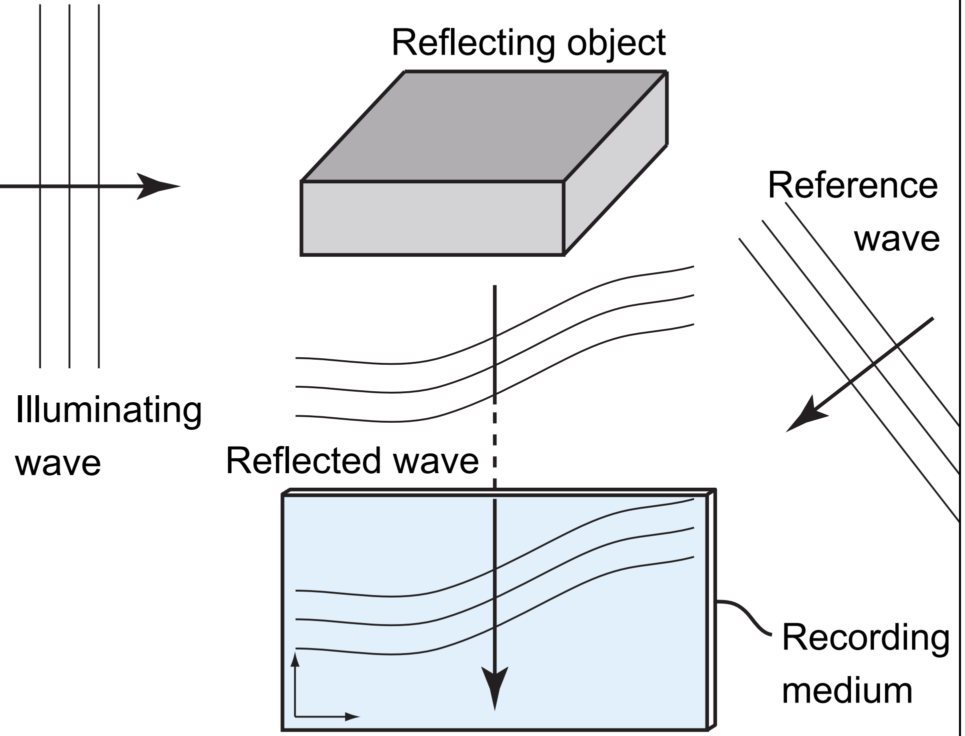
Distance: 5λ (focusing)

Distance: 10λ (beamforming)

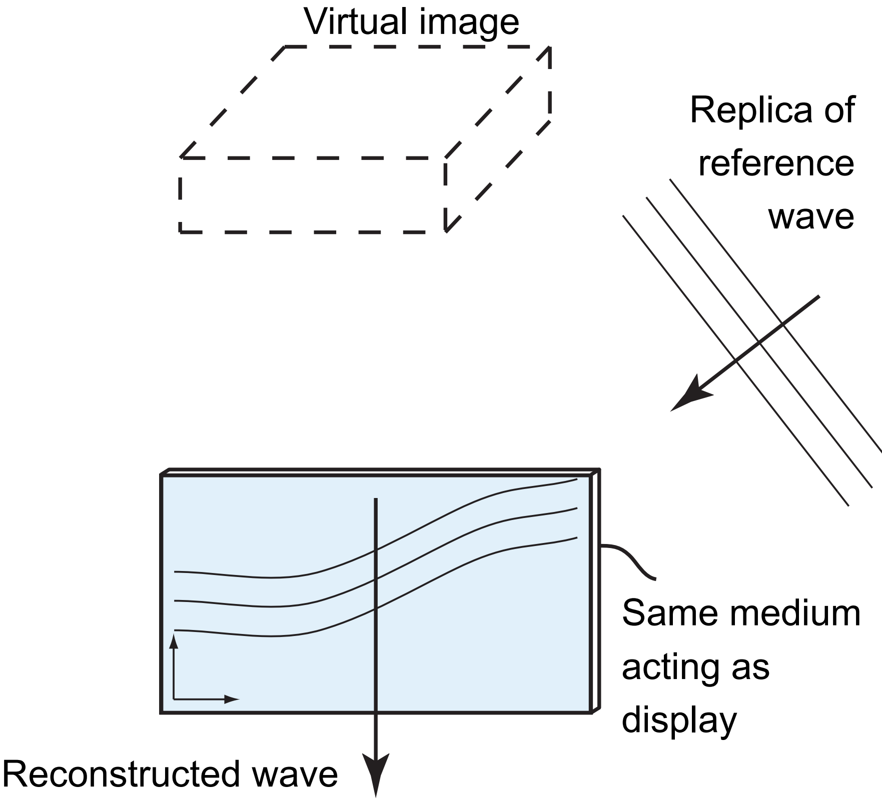


Connection to Holography

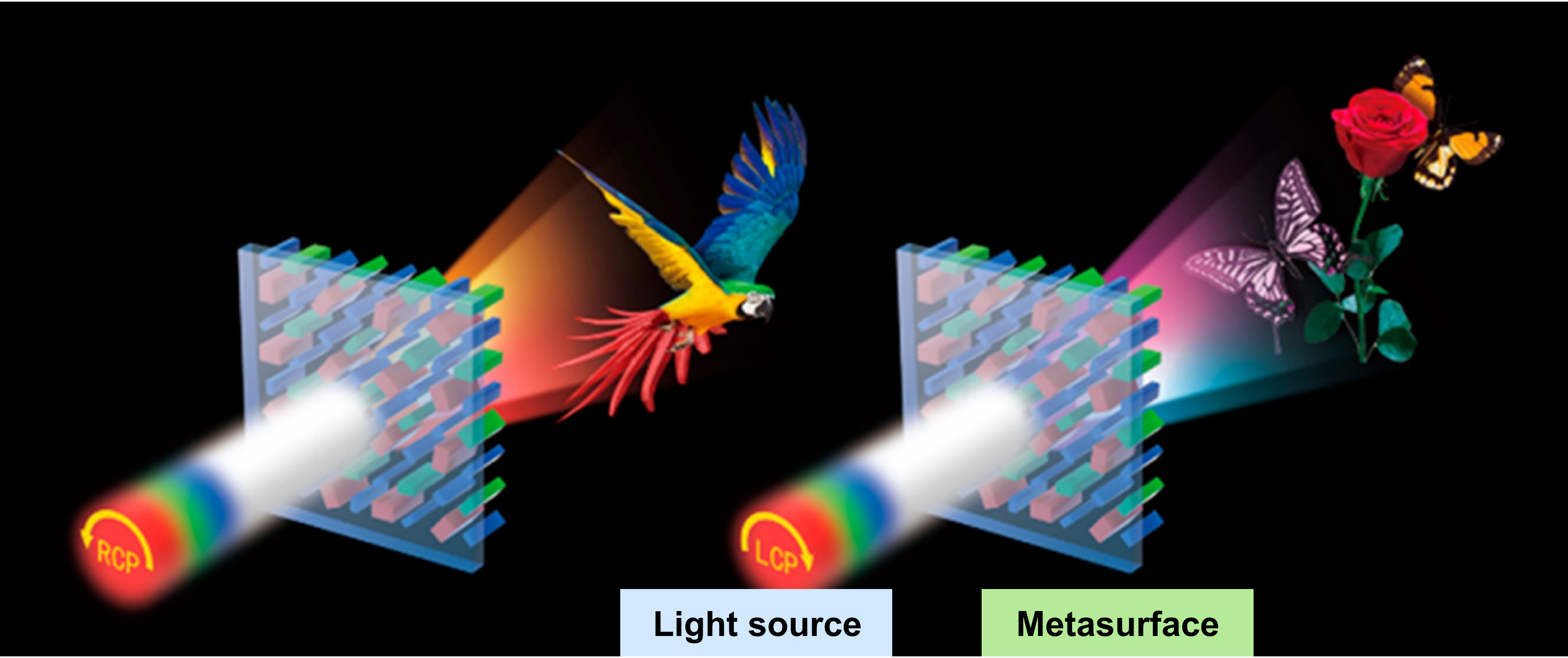
Holographic recording:



Reconstruction:



Holography Using Metasurfaces



Reference: H. Feng, Q. Li, W. Wan, J. Song, Q. Gong, M. L. Brongersma, Y. Li, “Spin-Switched Three-Dimensional Full-Color Scenes Based on a Dielectric Meta-hologram,” ACS Photonics, 2019.

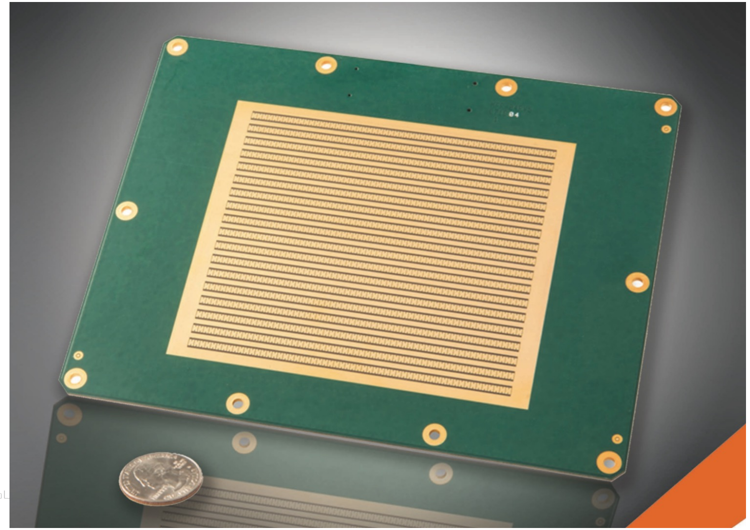
Holographic Beamforming

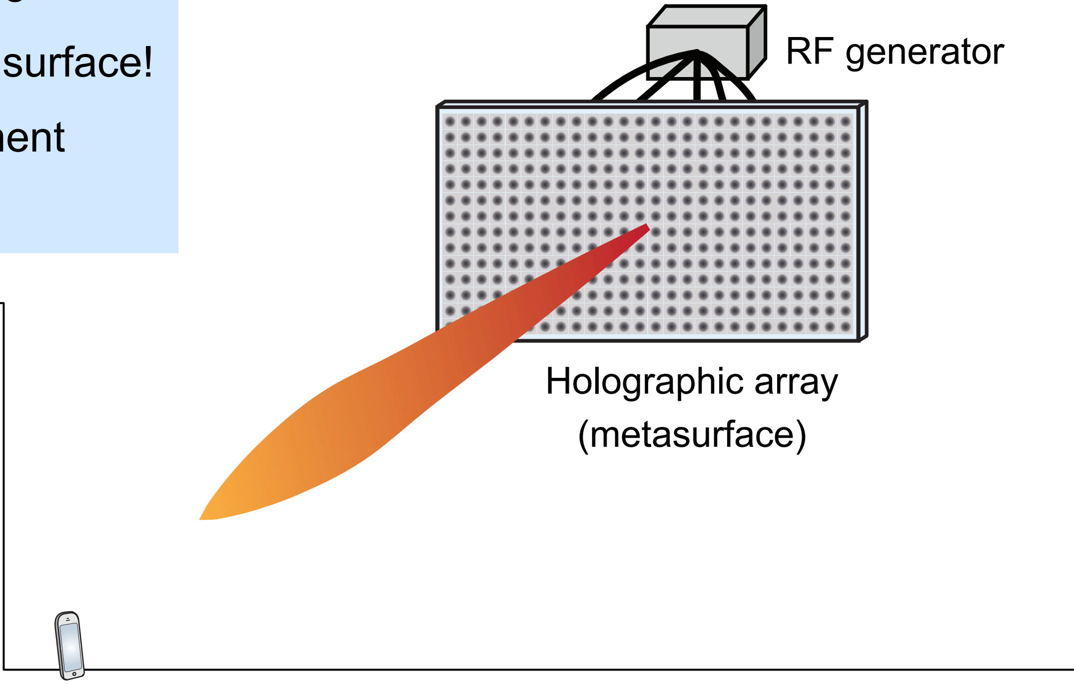
Metasurface with built-in source

No pathloss between RF generator and surface!

Thin form factor – invisible deployment

Ideally a contiguous surface

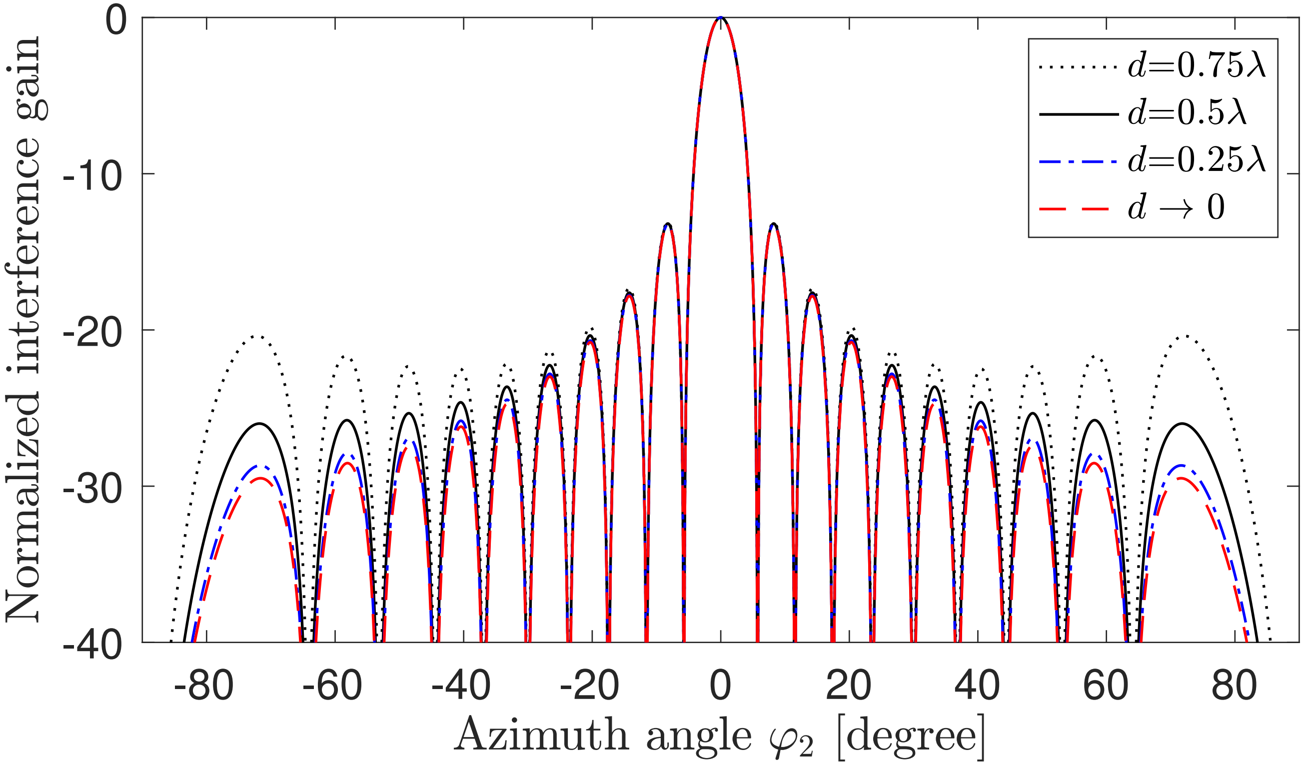




PIVOTAL COMMWARE

28 GHz, Small array, one RF chain

Towards a Continuous Aperture



Fixed surface dimensions
$L = Md, H = Nd$

Beamwidth

Unaffected as $d \to 0$

Determined by L and H

Sidelobes reduce

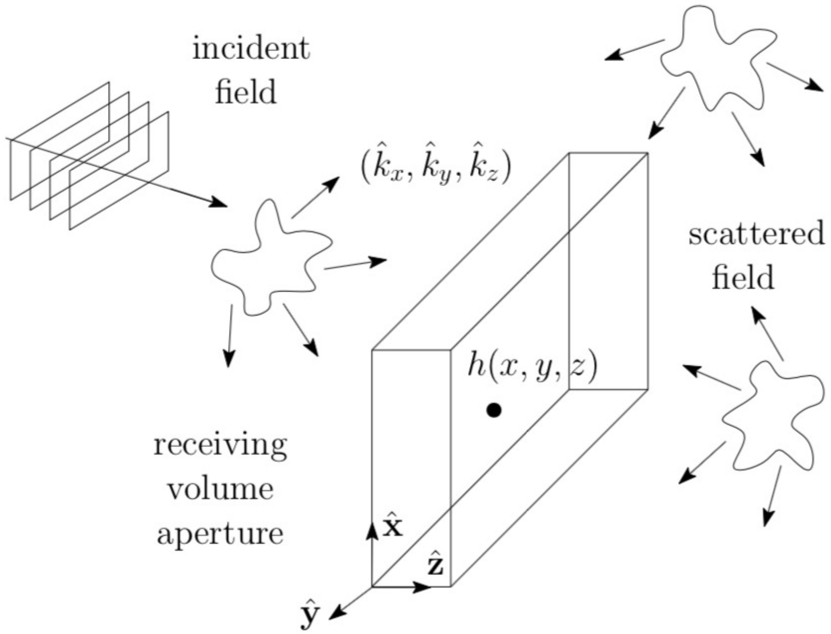
Small gain below
$d = 0.25\lambda$

Reason to consider $d \to 0$: Analysis with integrals instead of summations

How Many Signals Can Be Multiplexed?

- Need physically-meaningful propagation model
  - Satisfying the Helmholtz equation
  - Modeling practical scattering environments

Random scattered
non-line-of-sight scenario:



Independent Rayleigh fading
Not physically meaningful

Spatial Degrees-of-Freedom

Definition:
If the sum capacity behaves as $\eta \log_2(\mathrm{SNR}) + \mathrm{constant}$ at high SNR, then η is called the spatial degrees-of-freedom (DoF)

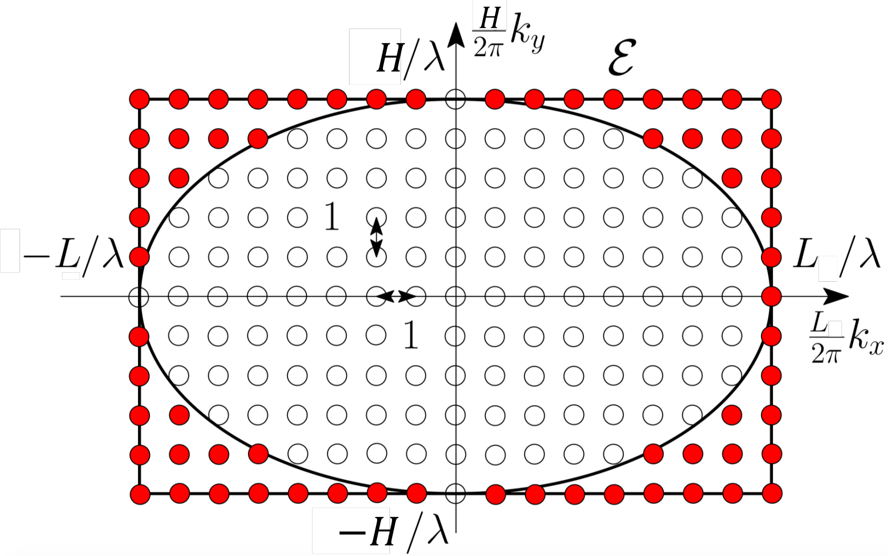
- **One dimension:**
  Uniform linear array (length L):

$$\eta = \frac{2}{\lambda} L$$

Like sampling of a band-limited signal, but we have a space-limited signal

- **Two dimensions:**
  Holographic MIMO (length L, height H):

$$\eta = \frac{\pi}{\lambda^2} LH$$



A. Pizzo, T. L. Marzetta, L. Sanguinetti, "Spatially-Stationary Model for Holographic MIMO Small-Scale Fading," 2020.

Huge DoF Differences

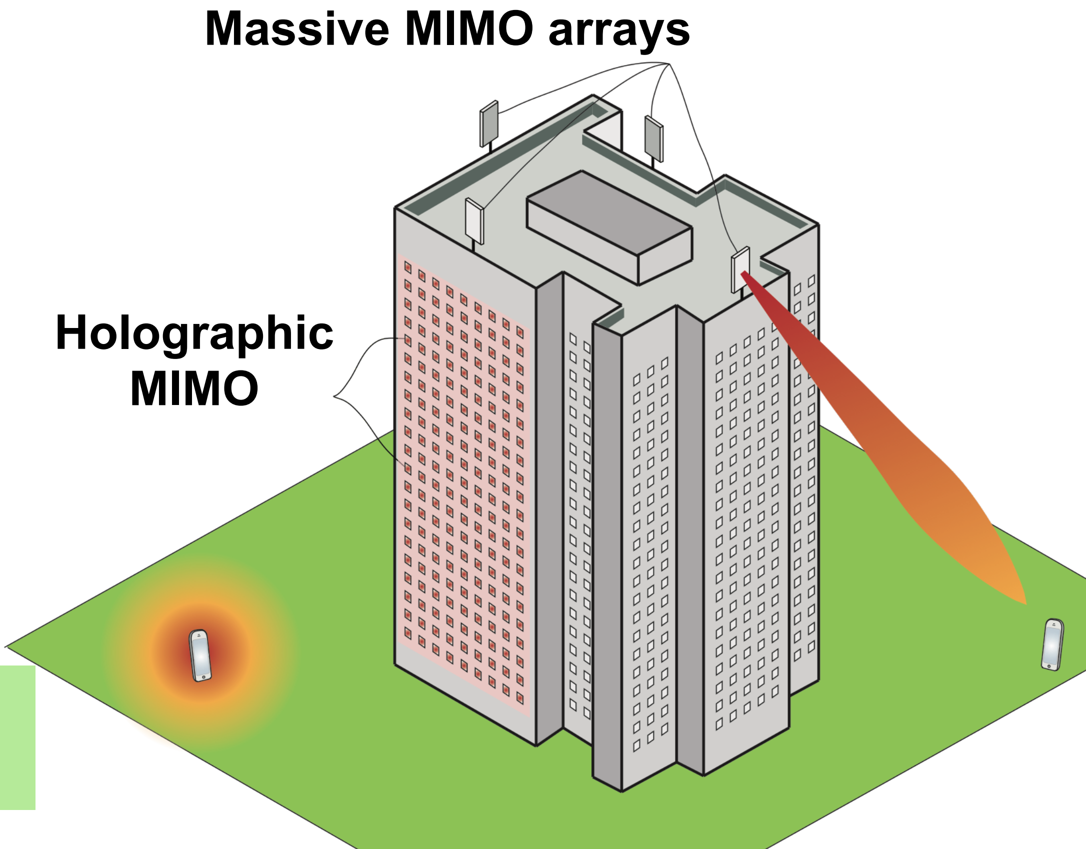
- Example: $\lambda = 0.1$ m (3 GHz)

- 5G Massive MIMO panel (64 antennas):

$$\eta = 64$$

- Large intelligent surface ($L = 10$ m, $H = 30$ m):

$$\eta \approx \frac{\pi}{\lambda^2} LH \approx 100000$$

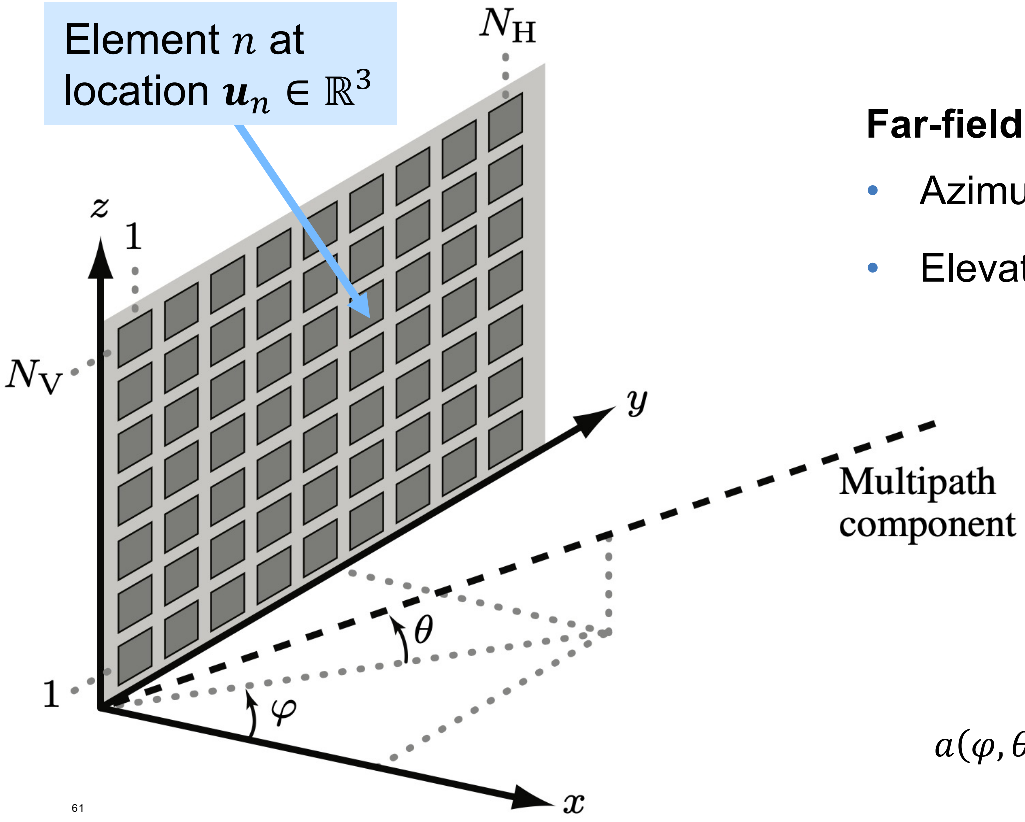


We are far from the limits!

The richness of the channels is key

SPATIAL CHANNEL MODELING

Array Response Vector



Far-field: Incoming/outgoing plane wave determined by

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

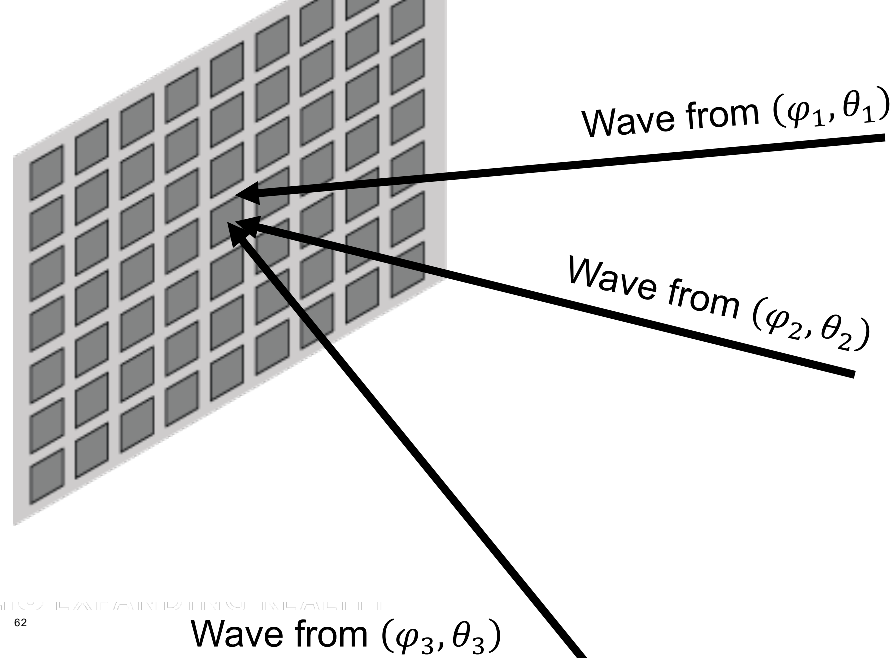
Channel vector for one plane wave

$\text{constant} \cdot a(\varphi, \theta)$

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

Multipath Channel Model

RIS with N elements,
Element area A



Channel vector for L plane wave

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

Rayleigh fading: $\boldsymbol{h} \sim CN(\boldsymbol{0}, \boldsymbol{R})$

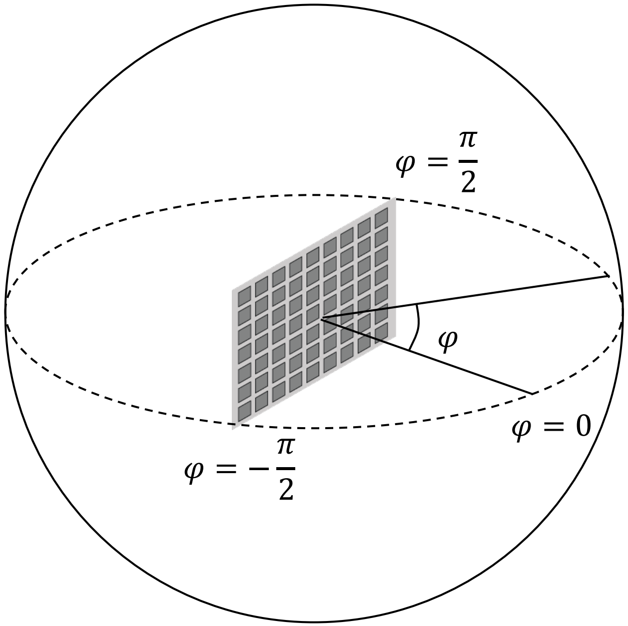
Achieved as $L \to \infty$ if

c_l i.i.d. with zero mean, variance $A\mu$

(φ_l, θ_l) i.i.d. with PDF $f(\varphi, \theta)$

$$[\boldsymbol{R}]_{nm} = A\mu \, E\left\{ e^{j\boldsymbol{k}(\varphi,\theta)^T(\boldsymbol{u}_n - \boldsymbol{u}_m)} \right\}$$

Isotropic Scattering



Uniform distribution in one half-space

(φ_l, θ_l) i.i.d. with PDF

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

Spatial correlation matrix:

$$[\boldsymbol{R}]_{nm} = A\mu \, E\left\{e^{j\boldsymbol{k}(\varphi,\theta)^T(\boldsymbol{u}_n - \boldsymbol{u}_m)}\right\}$$
$$= A\mu \cdot \mathrm{sinc}\left(\frac{2}{\lambda}\|\boldsymbol{u}_n - \boldsymbol{u}_m\|\right)$$

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

Natural Spatial Correlation

Square array and isotropic scattering

$$\frac{\text{rank}(\boldsymbol{R})}{\pi N A/\lambda^2} \to 1$$

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

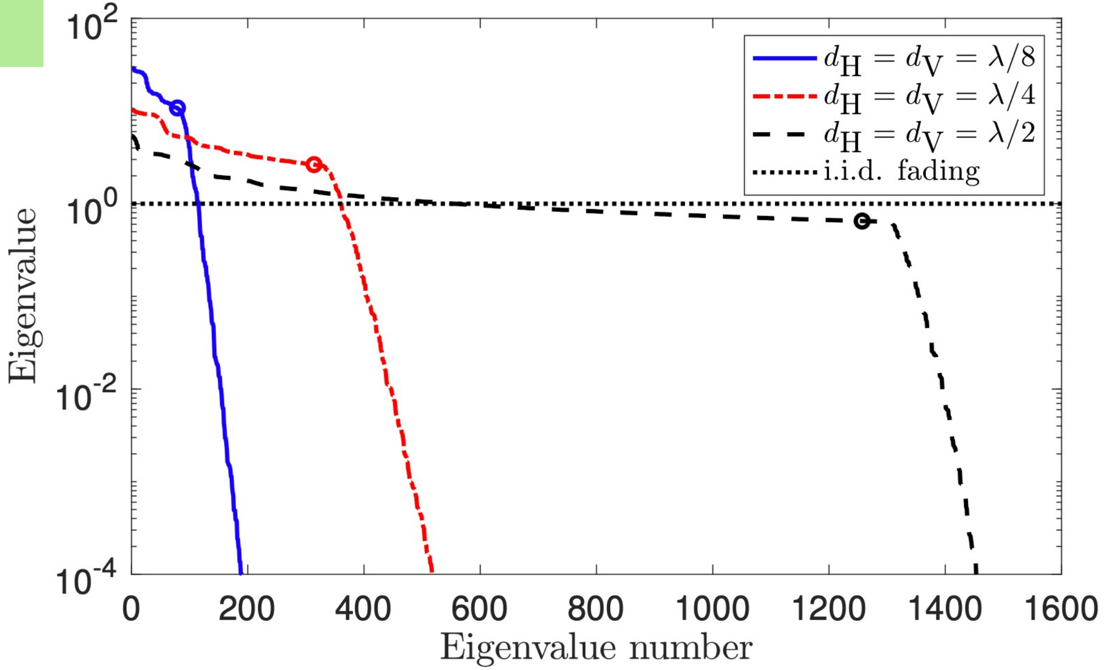
Channel sparsity

Even in isotropic fading!

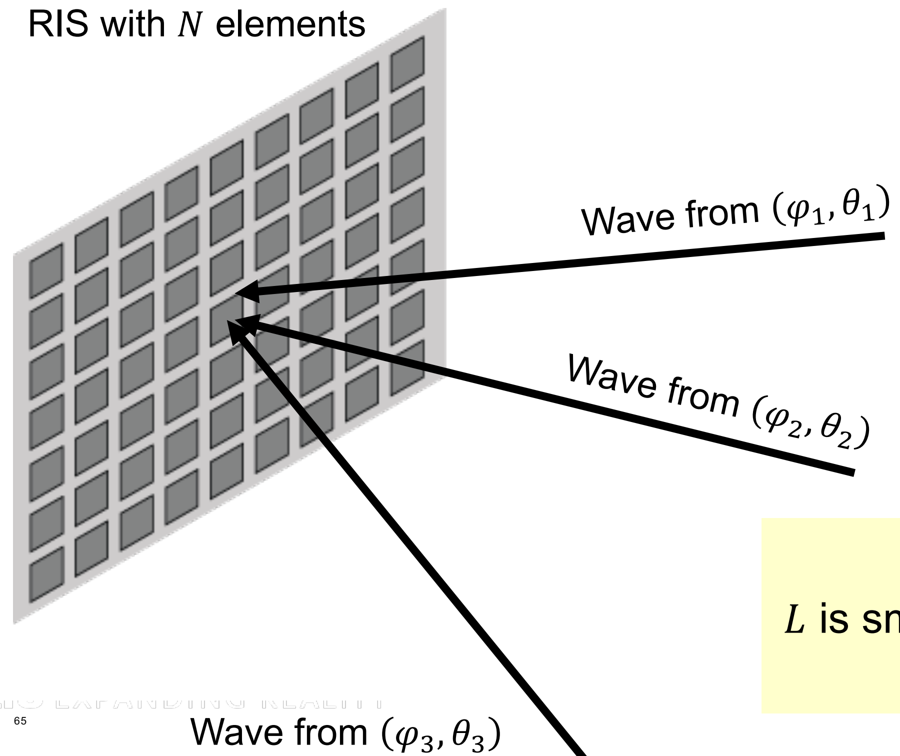
Two contributing factors

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

$N = 1600$ elements (40×40 array), $A = d_H d_V$



Even sparser channels



Channel vector for L plane wave

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

N complex parameters in $\boldsymbol{h}$

or $3L$ real parameters

Sparse channels

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

Example: Line-of-sight channels

Sparse Channels with RIS

Recall the channel gain:

$$\sqrt{\rho}e^{-j2\pi f_c\tau_d}+\sum_{n=1}^{N}\sqrt{\alpha_n\beta_n\gamma_n}e^{-j2\pi f_c(\tau_{n,a}+\tau_{\theta_n}+\tau_{n,b})}$$

$$=\sqrt{\rho}e^{-j2\pi f_c\tau_d}+\left(\underbrace{\begin{bmatrix}\sqrt{\beta_1}e^{-j2\pi f_c\tau_{1,b}}\\ \vdots\\ \sqrt{\beta_N}e^{-j2\pi f_c\tau_{N,b}}\end{bmatrix}}_{\text{Sparse channel from RIS}}\underset{\text{Element-wise multiplication}}{\odot}\underbrace{\begin{bmatrix}\sqrt{\alpha_1}e^{-j2\pi f_c\tau_{1,a}}\\ \vdots\\ \sqrt{\alpha_N}e^{-j2\pi f_c\tau_{N,a}}\end{bmatrix}}_{\text{Sparse channel to RIS}}\right)^T\underbrace{\begin{bmatrix}\sqrt{\gamma_1}e^{-j2\pi f_c\tau_{\theta_1}}\\ \vdots\\ \sqrt{\gamma_N}e^{-j2\pi f_c\tau_{\theta_N}}\end{bmatrix}}_{\text{Controllable vector}}$$

Unknowns *without* sparsity: $N+1$

Unknowns *with* sparsity: Much fewer?

Near-Field Channel Modeling

Near-field phenomena

1. Different distances to antennas
2. Different effective areas
3. Different polarization losses

Channel gain (right in front, distance d, N elements):

$$\alpha_d = \frac{N\beta_d}{3(N\beta_d\pi + 1)\sqrt{2N\beta_d\pi + 1}} + \frac{2}{3\pi}\tan^{-1}\left(\frac{N\beta_d\pi}{\sqrt{2N\beta_d\pi + 1}}\right)$$

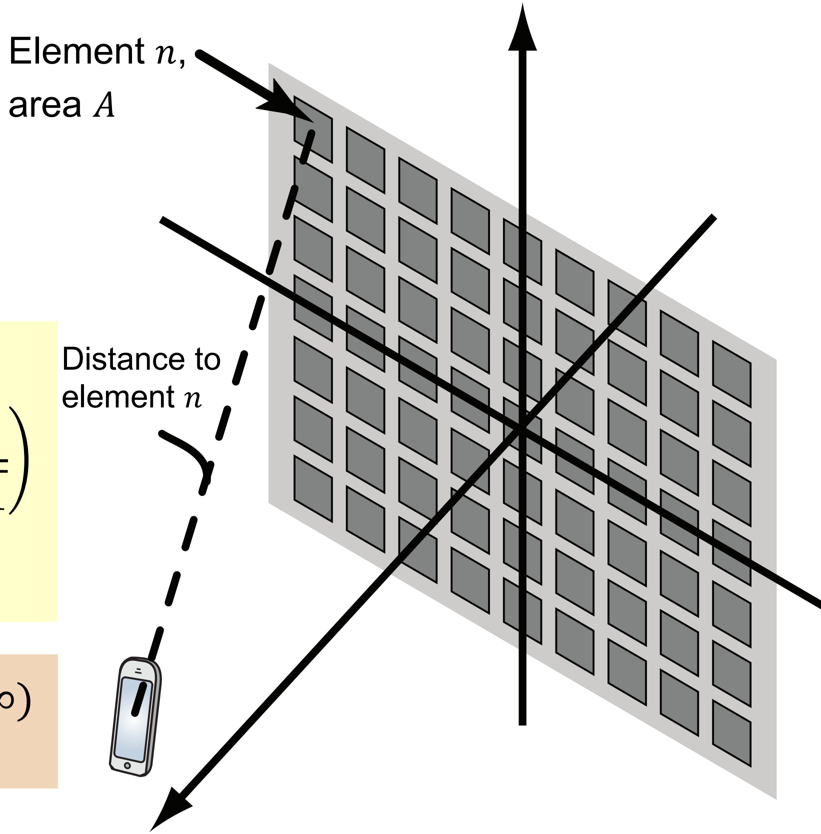
where $\beta_d = \frac{A}{4\pi d^2}$

Far-field ($d \gg \sqrt{NA}$)

$\alpha_d \approx N\beta_d$

Asymptotic limit ($N \to \infty$)

$\alpha_d \to 1/3$



How is the Expression Computed?

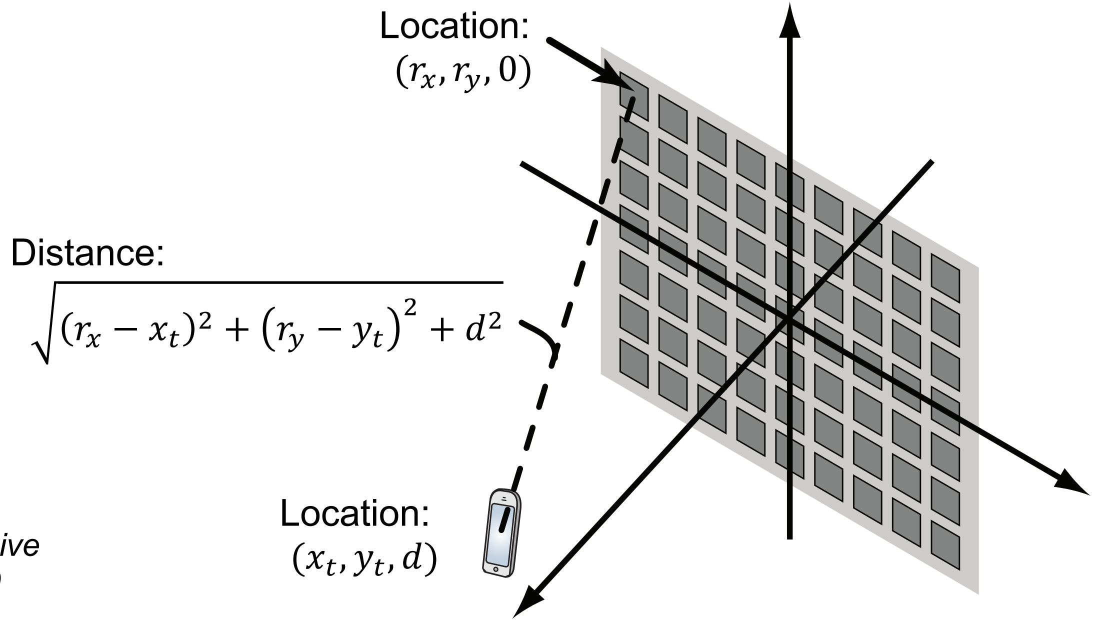
$$\iint_{\text{Antenna area}} \underbrace{\frac{d}{\sqrt{(r_x - x_t)^2 + (r_y - y_t)^2 + d^2}}}_{\text{Reduction in effective area from directivity}} \underbrace{\frac{(r_x - x_t)^2 + d^2}{(r_x - x_t)^2 + (r_y - y_t)^2 + d^2}}_{\text{Polarization loss factor}} \underbrace{\frac{\partial r_x \partial r_y}{4\pi((r_x - x_t)^2 + (r_y - y_t)^2 + d^2)}}_{\text{Free-space pathloss}}$$

If we ignore polarization loss

Converges to 1/2 instead of 1/3

If we ignore reduction in effective area:

Divergence



Reference: *E. Björnson, L. Sanguinetti, "Power Scaling Laws and Near-Field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces," 2020*

OPEN PROBLEM 1:
WHAT ARE SUITABLE USE CASES?

Hard to speculate about use cases!

- **We still don't even know what the 5G use cases will be!**

All papers with 6G applications are just speculations!

Some applications appear earlier, other not at all

3G was introduced in 1998

Believed 3G Applications

Video calling

Facetime in 2010

Mobile e-commerce

Location-based services

Games and sports events

Broadband Internet access

Broadband video services

iPhone 3G in 2008

OECD (2004-09-14), "Development of Third-Generation Mobile Services in the OECD", OECD Digital Economy Papers.

Start from Performance Metrics – Identify new technology solutions

Cannot do all of this simultaneously

Must look for new *radical* solutions

10x improvements, not 10%!

Must compare against state-of-the-art benchmarks

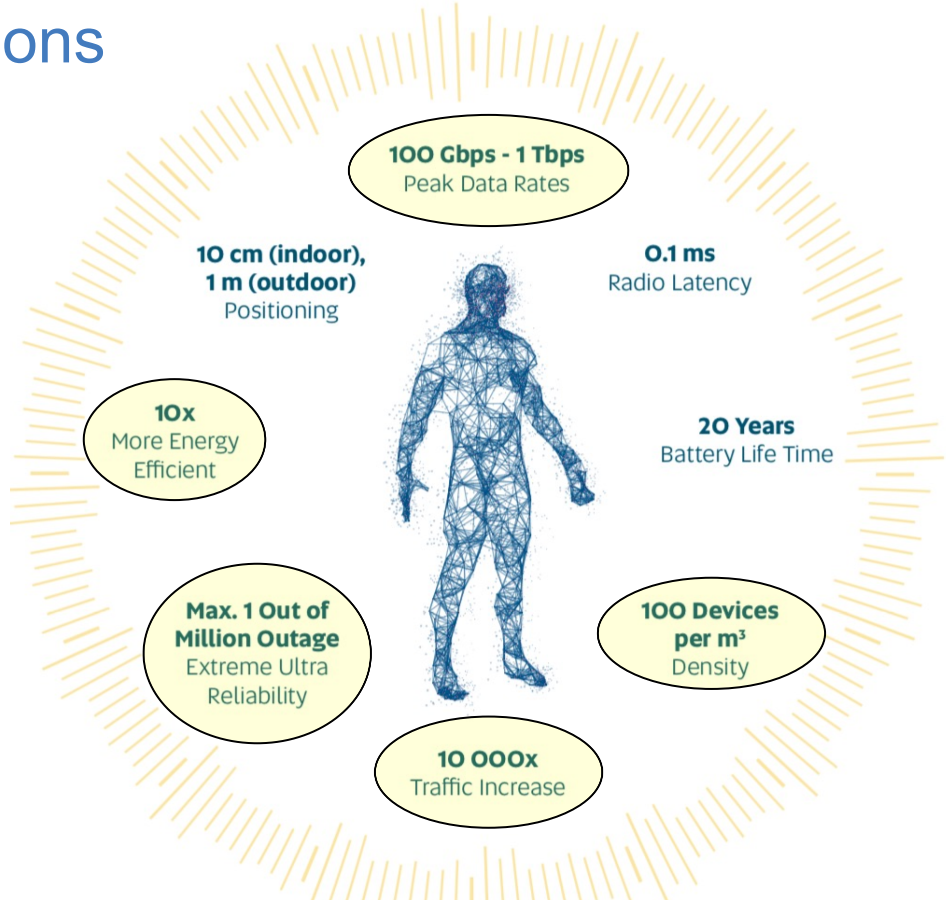


Image source: 6G Flagship, “Key Drivers and Research Challenges for 6G Ubiquitous Wireless Intelligence,” 2019

Ways to Increase Capacity

- Capacity of a communication system [bit/s]:

$$\text{Multiplexed layers} \cdot \text{Bandwidth} \cdot \log_2\left(1 + \frac{\text{Signal power} \cdot \text{Beamforming gain} \cdot \text{Pathloss}}{N_0 \cdot \text{Bandwidth}}\right)$$

- Three ways to increase it:

Multiplexed more layers
To one or multiple users

Use more bandwidth
Must be associated with higher power or beamforming gain

Reduce range
Improve pathloss

Rate Expressions with Holographic MIMO

Basically a Massive MIMO system

- Reuse existing theory
- Utilize different channel models

Don’t use i.i.d. Rayleigh fading!

Asymptotic issues finally appear

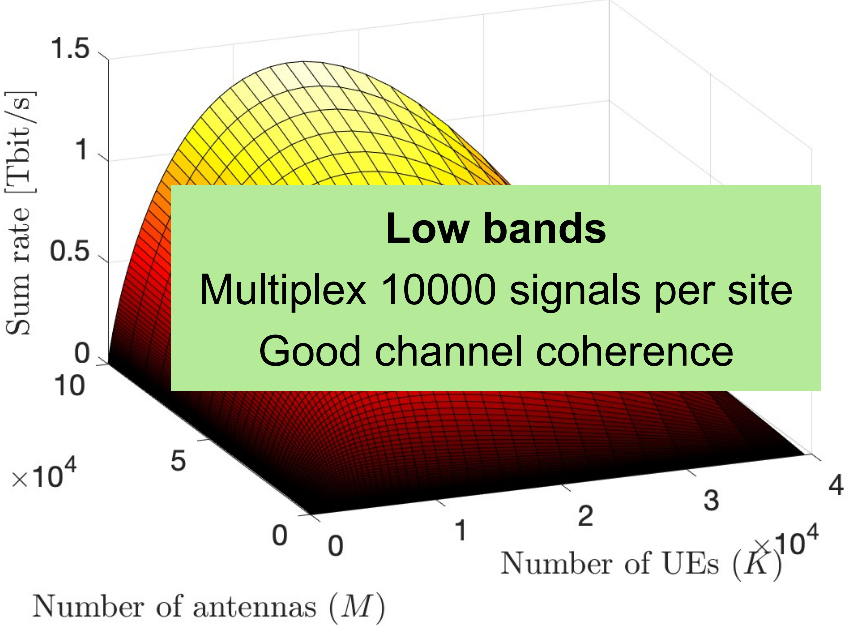
- Pilot contamination
- Computational complexity
- Data shuffling
- Hybrid beamforming

Near-field possibility

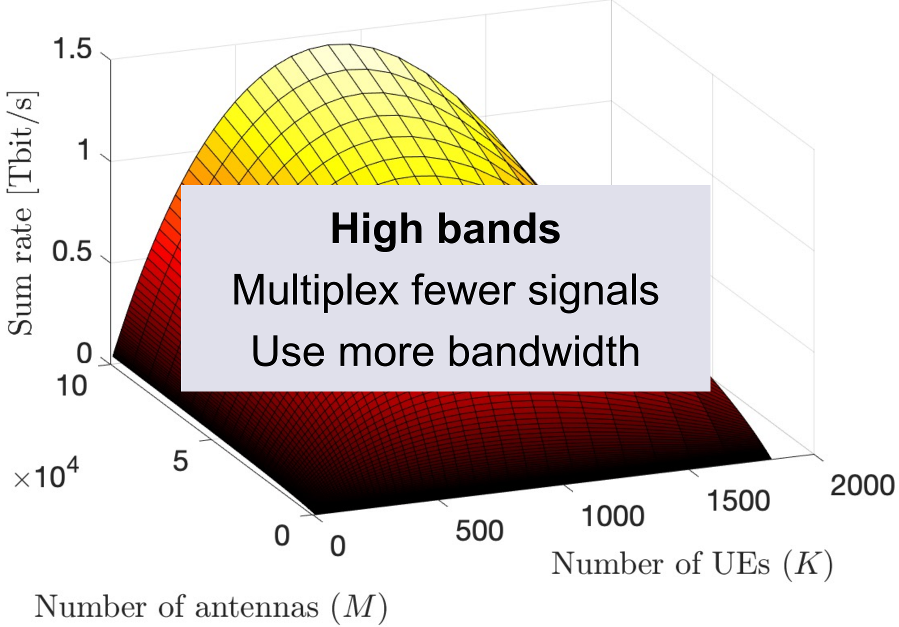
Many layers per user!

Reaching 1 Tb/s by Holographic MIMO

50 MHz bandwidth, 3 GHz band

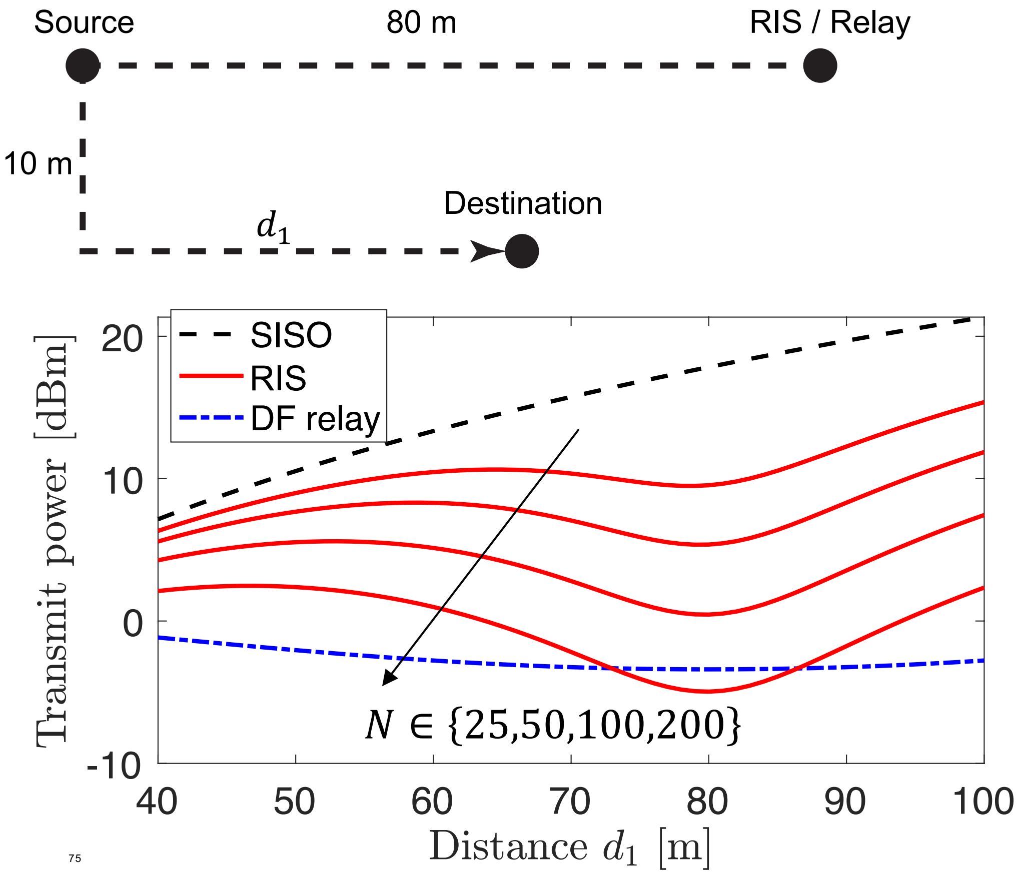


1 GHz bandwidth, 60 GHz band



E. Björnson, L. Van der Perre, S. Buzzi, E. G. Larsson, "Massive MIMO in Sub-6 GHz and mmWave: Physical, Practical, and Use-Case Differences," IEEE Wireless Communications, 2019

Maximize SNR: RIS vs. DF Relay



Goal: Achieve 4 bit/s/Hz

Channel gain: 3GPP Urban Micro (NLOS direct path, LOS otherwise)

Bandwidth: 10 MHz

Observations

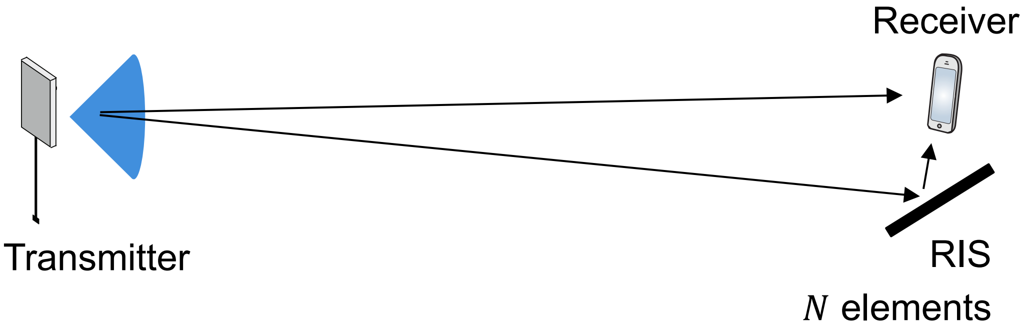
Outperforms SISO case

200 elements needed to beat relay

But fits into $1 \times 0.5 \text{ m}^2$ at 3 GHz carrier

Reference: E. Björnson, Ö. Özdogan, E. G. Larsson, "Intelligent Reflecting Surface vs. Decode-and-Forward: How Large Surfaces Are Needed to Beat Relaying?," IEEE Wireless Commun. Letters, vol. 9, no. 2, pp. 244-248, February 2020.

Improving Channel Properties

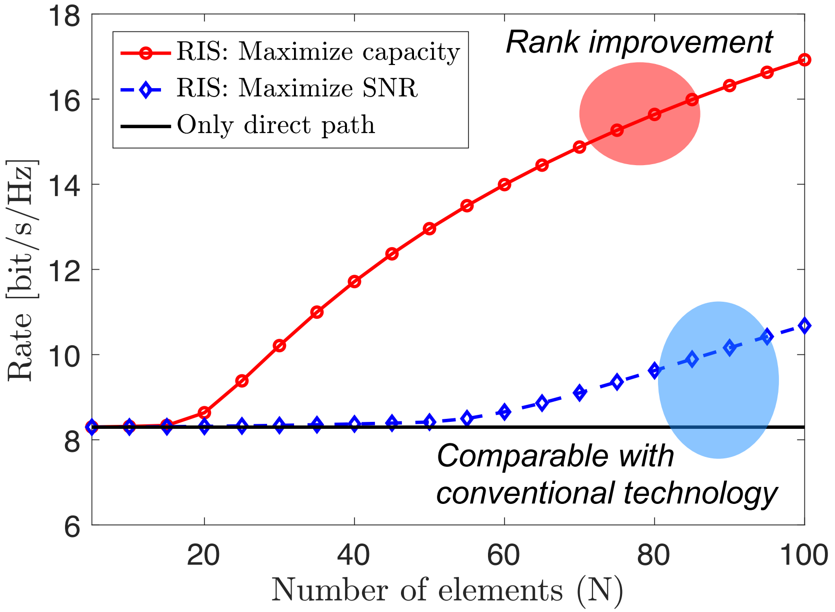


Two antennas at each device

Line-of-sight channels: Rank 1

Improve propagation conditions

More than just SNR gain!



Reference:

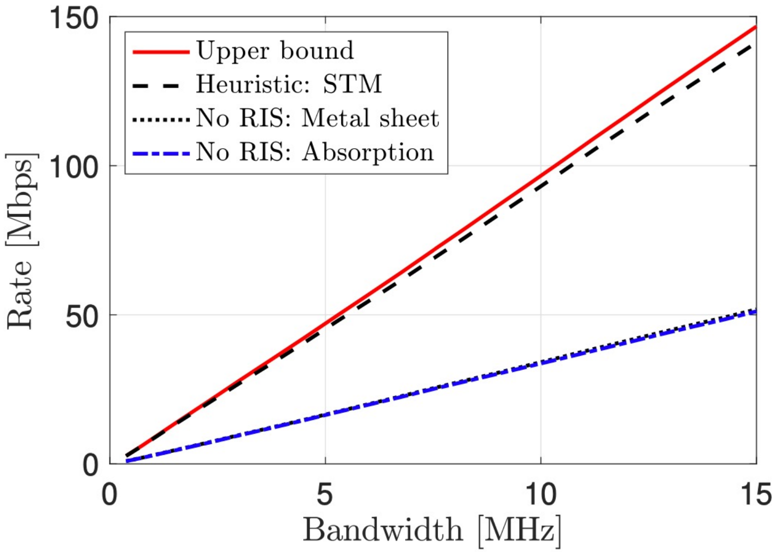
[R2] Ö. Özdogan, E. Björnson, E. G. Larsson, "Using Intelligent Reflecting Surfaces For Rank Improvement in MIMO Communications"

RIS with Wideband Channels

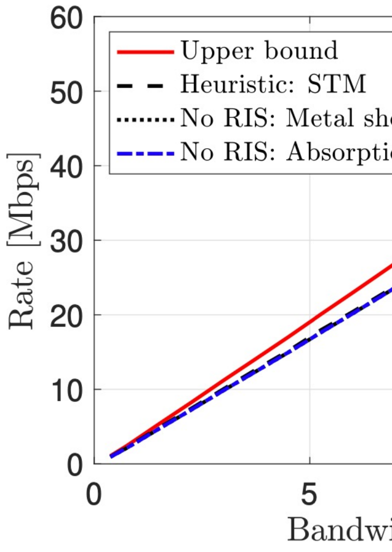
Different channels on each subcarrier

Only one RIS configuration

Heuristic: Strongest tap maximization (STM)



(a) LOS channel from transmitter to RIS.



We need LOS channels to/from RIS!

(b) NLOS channel from transmitter to RIS.

B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," IEEE WCL, 2020.

S. Lin, et al., "Adaptive transmission for reconfigurable intelligent surface-assisted OFDM wireless communications," IEEE JSAC, 2020

Exploit the Large Surface Area

Prospective use cases:

1. Mitigate fading through channel hardening

Example: Emil Björnson, Luca Sanguinetti, “Rayleigh Fading Modeling and Channel Hardening for Reconfigurable Intelligent Surfaces,” IEEE Wireless Communications Letters, 2021

2. Counteract Doppler effects

Example: Bho Matthiesen, Emil Björnson, Elisabeth De Carvalho, Petar Popovski, “Intelligent Reflecting Surface Operation under Predictable Receiver Mobility: A Continuous Time Propagation Model,” IEEE Wireless Communications Letters, 2021.

3. Macro diversity

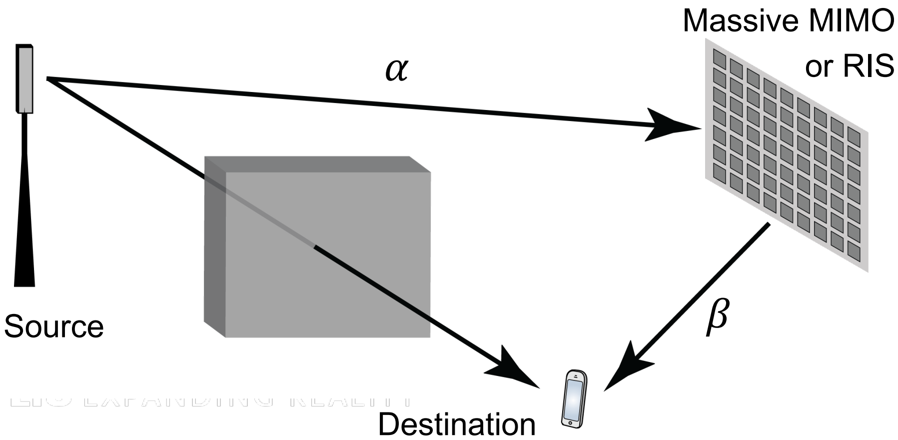
Increase channel rank, mitigate blocking, particularly at (sub-)terahertz bands

RIS versus Holographic MIMO

Spectral efficiency: $\log_2(1 + \rho N^2 \alpha \beta)$

An RIS has worse SNR for the same N

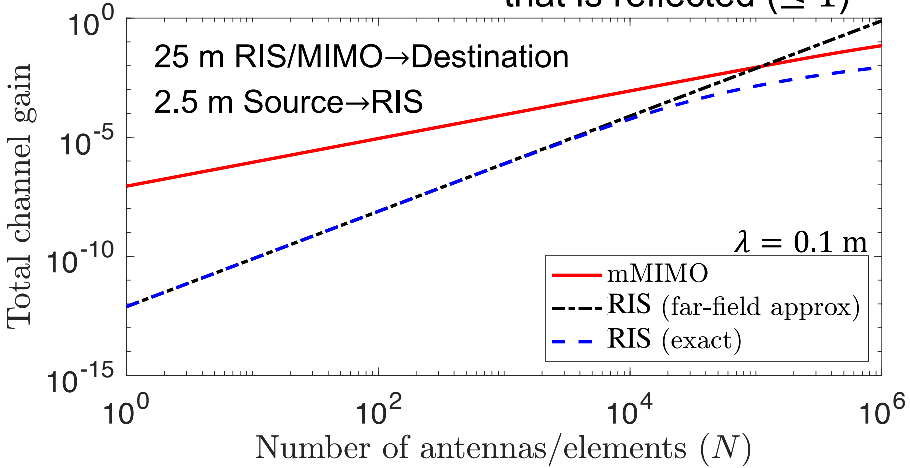
Trade larger area to get passive operation



$$\rho N^2 \alpha \beta = \rho N \beta \cdot N \alpha$$

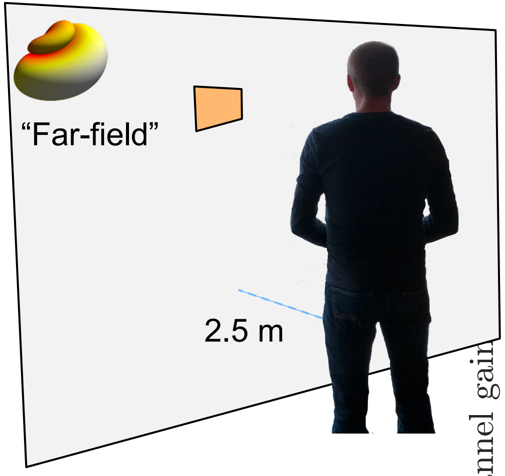
SNR with Massive MIMO

Fraction of transmit power that is reflected (≤ 1)



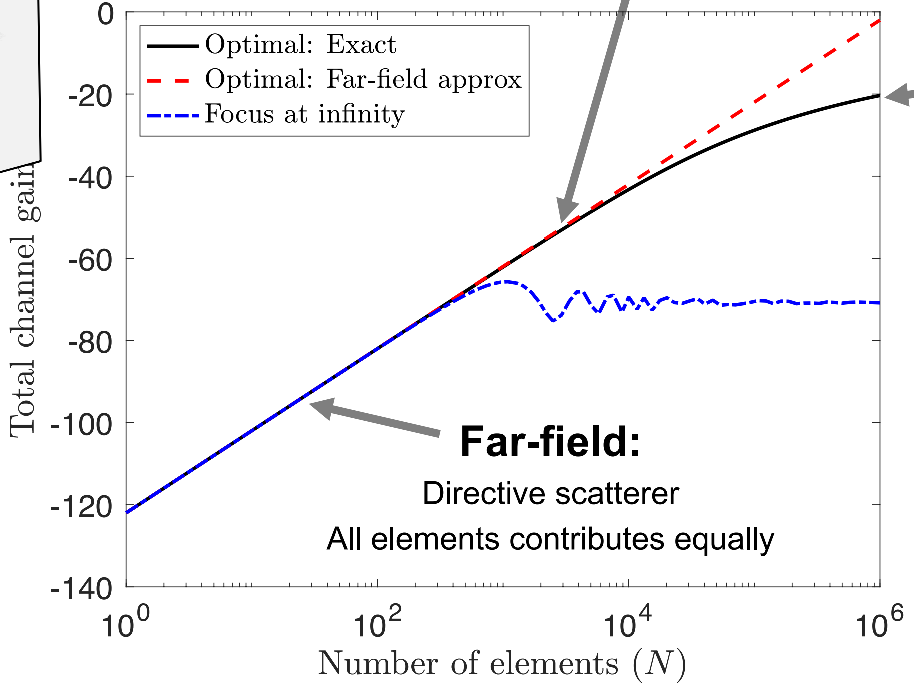
Reference: E. Björnson, L. Sanguinetti, “Power Scaling Laws and Near-Field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces,” 2020.

RIS with Asymptotically Large Surface



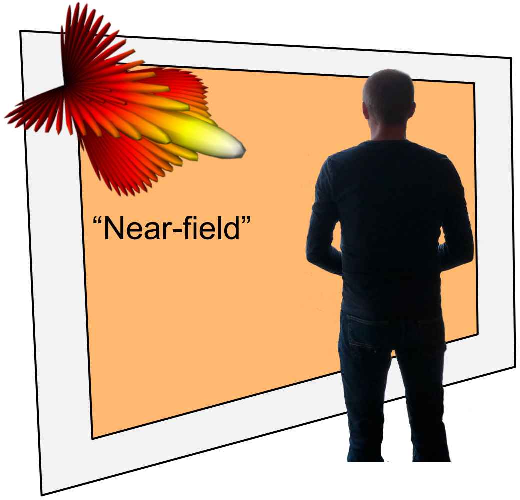
Near-field:

Focus at a point
(not in a direction)



Asymptotic limit

Receive almost all
transmit power



Far-field:

Directive scatterer

All elements contributes equally

Assumptions:

25 m from transmitter

2.5 m to receiver

Isotropic antennas (3 GHz)

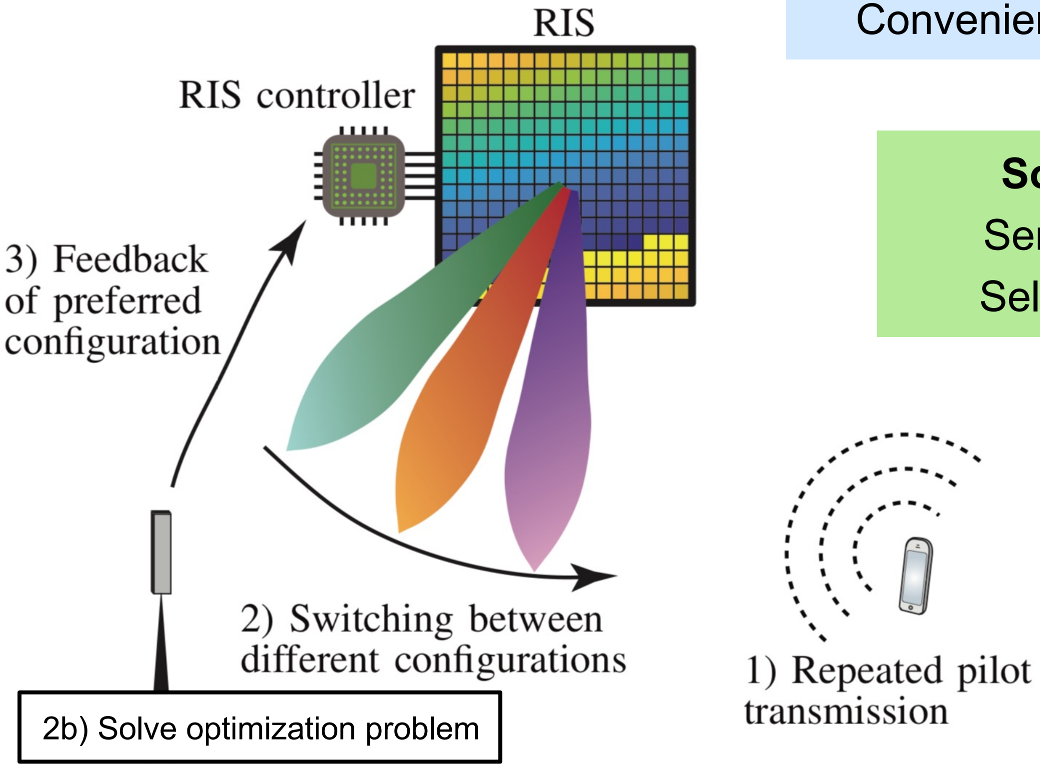
OPEN PROBLEM 2: CHANNEL ESTIMATION

RIS Reconfigurability is Complicated

The RIS is blind!

Solution 1: Have a fraction of active RIS elements
Convenient, but removes the “low-cost” argument

Solution 2: Codebook approach
Send pilots and switch configuration
Select the best configuration in a set



System Model for Channel Estimation

Recall the channel gain:

$$\sqrt{\rho}e^{-j2\pi f_c\tau_d} + \sum_{n=1}^{N}\sqrt{\alpha_n\beta_n\gamma_n}e^{-j2\pi f_c(\tau_{n,a}+\tau_{\theta_n}+\tau_{n,b})}$$

$$= \underbrace{\sqrt{\rho}e^{-j2\pi f_c\tau_d}}_{h\text{: Direct channel}} + \underbrace{\left(\begin{bmatrix}\sqrt{\beta_1}e^{-j2\pi f_c\tau_{1,b}}\\ \vdots \\ \sqrt{\beta_N}e^{-j2\pi f_c\tau_{N,b}}\end{bmatrix} \odot \begin{bmatrix}\sqrt{\alpha_1}e^{-j2\pi f_c\tau_{1,a}}\\ \vdots \\ \sqrt{\alpha_N}e^{-j2\pi f_c\tau_{N,a}}\end{bmatrix}\right)^T}_{\boldsymbol{g}\text{: Cascaded channel}} \underbrace{\begin{bmatrix}\sqrt{\gamma_1}e^{-j2\pi f_c\tau_{\theta_1}}\\ \vdots \\ \sqrt{\gamma_N}e^{-j2\pi f_c\tau_{\theta_N}}\end{bmatrix}}_{\boldsymbol{\omega}_\theta\text{: Controllable vector}}$$

Received signal:

$$z[n] = (h + \boldsymbol{g}^T\boldsymbol{\omega}_\theta)x[n] + w[n]$$

Basic Codebook-Based Channel Estimation

Repeat a known pilot signal L times

- $x[n] = s, \; n = 1, \dots, L$
- Use different configurations: $\boldsymbol{\omega}_{\theta_1}, \dots, \boldsymbol{\omega}_{\theta_L}$

- Received signal: $z[n] = (h + \boldsymbol{g}^T \boldsymbol{\omega}_{\theta}) x[n] + w[n]$

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

$$\boldsymbol{Z} = \begin{bmatrix} h \\ \boldsymbol{g} \end{bmatrix}^T \begin{bmatrix} 1 & \dots & 1 \\ & \boldsymbol{\Omega} & \end{bmatrix} s + \boldsymbol{W}$$

Received signal

If $\widetilde{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & \dots & 1 \\ & \boldsymbol{\Omega} & \end{bmatrix}$ is invertible

Requires $L = N + 1$

Example: $\widetilde{\boldsymbol{\Omega}}$ =DFT matrix

$$\boldsymbol{Z} \widetilde{\boldsymbol{\Omega}}^{-1} s^{-1} = \begin{bmatrix} h \\ \boldsymbol{g} \end{bmatrix}^T + \boldsymbol{W} \widetilde{\boldsymbol{\Omega}}^{-1} s^{-1}$$

Channel estimate

Simplify by Exploiting Channel Sparsity

$$\boldsymbol{Z} = \begin{bmatrix} h \\ \boldsymbol{g} \end{bmatrix}^T \begin{bmatrix} 1 & \cdots & 1 \\ & \boldsymbol{\Omega} & \end{bmatrix} s + \boldsymbol{W}$$

Received signal

Sparsity will exist in $\boldsymbol{g}$!

An open research challenge to utilize it most effectively

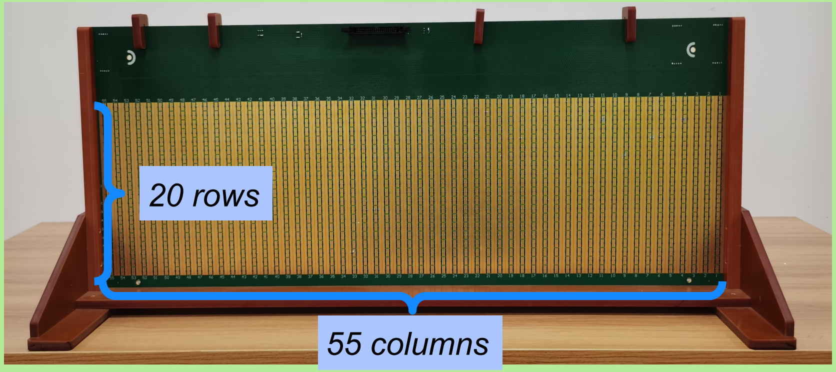
References: "*Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions*"

Use sparse channel models

Group elements together

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," arXiv:2103.00534

Basic configuration algorithm

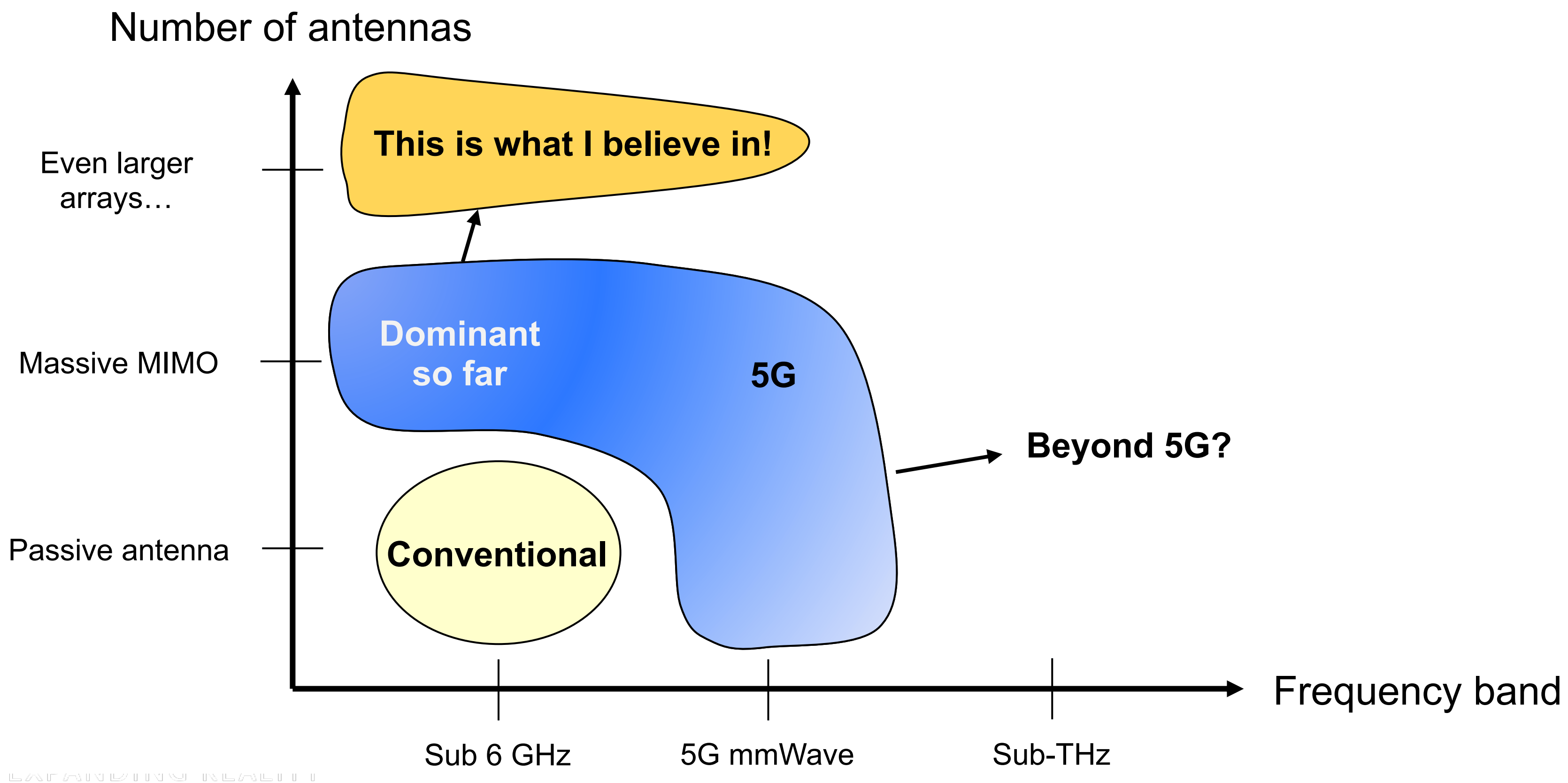


Two states per element: ± 90 degrees

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

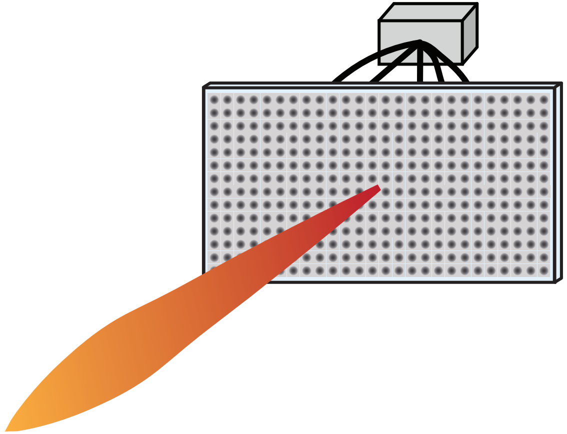
SUMMARY

Massive MIMO is Only the Beginning…



Future Requires Larger Arrays!

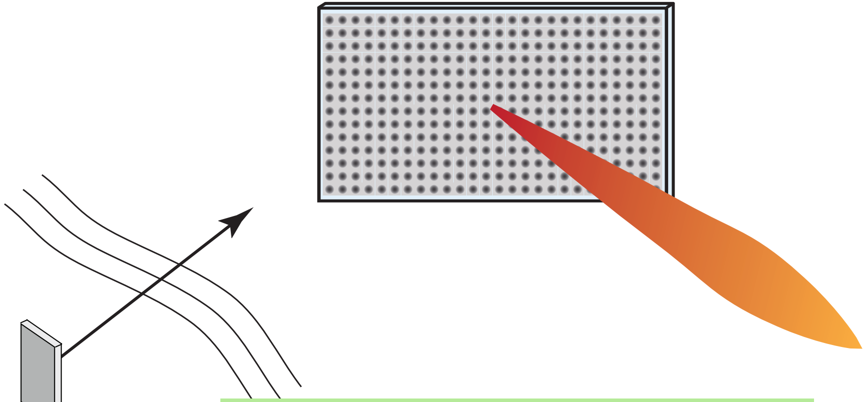
Holographic MIMO



Massive spatial DoF

Many active components

Reconfigurable Intelligent Surface



Simpler surface hardware

Challenge to control it

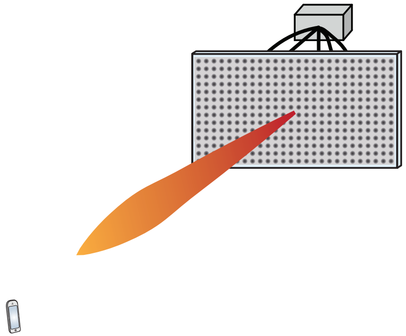
Many interesting research directions!

A lot of fundamental problems remain open!

When to use which technology?

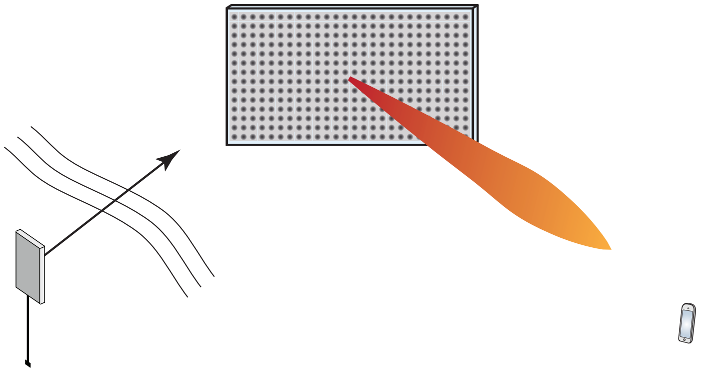
Holographic Massive MIMO

- Spatial multiplexing
- Utilize near-field effects
- Any propagation environment



Reconfigurable intelligent surfaces

- Improve coverage and reduce fading
- Improve channel rank
- Deploy to get LOS channels



Key References

Overview papers

1. E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, “Massive MIMO is a reality—What is next? Five promising research directions for antenna arrays,” Digital Signal Processing, vol. 94, pp. 3–20, Nov. 2019.
2. 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.
3. E. Björnson, Ö. Özdogan, E. G. Larsson, “Reconfigurable Intelligent Surfaces: Three Myths and Two Critical Questions,” IEEE Communications Magazine, 2020.
4. M. D. Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead,” https://arxiv.org/abs/2004.09352

Channel modeling

5. Ö. Özdogan, E. Björnson, E. G. Larsson, “Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling,” IEEE Wireless Commun. Letters, 2020.
6. E. Björnson, L. Sanguinetti, “Power Scaling Laws and Near-Field Behaviors of Massive MIMO and Intelligent Reflecting Surfaces,” IEEE O. J. Commun. Soc. 2020.

Key References

Prototyping of RIS

5. 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," Available on arXiv:2102.00742

Holographic MIMO

6. S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2746–2758, 2018.
7. A. Pizzo, T. L. Marzetta, L. Sanguinetti, "Spatially-Stationary Model for Holographic MIMO Small-Scale Fading," IEEE Journal on Selected Areas in Communications, to appear, 2020.
8. E. Björnson, L. Sanguinetti, "Utility-based Precoding Optimization Framework for Large Intelligent Surfaces," Asilomar Conference on SSC, 2019.
9. A. de Jesus Torres, Luca Sanguinetti, Emil Björnson, "Near- and Far-Field Communications with Large Intelligent Surfaces," Asilomar Conference on SSC 2020.

My papers are published with code:

https://github.com/emilbjornson

Thanks to all my collaborators!

Particularly:

- Erik G. Larsson, Özgecan Özdogan (Linköping University)
- Luca Sanguinetti, Andrea de Jesus Torres (University of Pisa)
- Petar Popovski, Elisabeth de Carvalho (Aalborg University)
- Henk Wymeersch (Chalmers University)
- Bho Matthiesen (University of Bremen)

Swedish Foundation for Strategic Research

Future Research Leader Grant

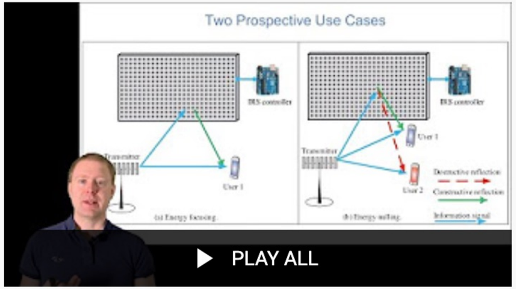
Knut and Alice Wallenberg Foundation

Wallenberg Academy Fellow Grant

Swedish Research Council

Project Grant

YouTube Videos

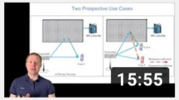


Intelligent reflecting surfaces for 6G

5 videos • 1,361 views • Last updated on Jan 4, 2021

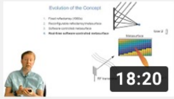
Wireless Future / Communication Systems

SUBSCRIBE

1 

Fundamentals of Intelligent Reflecting Surfaces

Wireless Future / Communication Systems

2 

Reconfigurable intelligent surfaces: Myths and realities

Wireless Future / Communication Systems

3 

A Programmable Wireless World With Reconfigurable Intelligen

Wireless Future / Communication Systems

4 

Communication Using Reconfigurable Intelligent Surfaces: Fun Recent Insights

Wireless Future / Communication Systems

5 

Power Scaling Laws and Near-Field Behaviors of Massive MIM Intelligent Reflecting Surfaces

IEEEComSoc

Podcast:

WIRELESS FUTURE