Reconfigurable Intelligent Surface for 6G: Communication, Sensing, and Localization

Boya Di[#], Hongliang Zhang[&], Lingyang Song[*], and Zhu Han[+],

[#]Department of Computing, Imperial College London, London, UK

[&]Department of Electrical Engineering, Princeton University, NJ, USA

[*] School of Electronics Engineering and Computer Science,

Peking University, Beijing, China

[+] Department of Electrical and Computer Engineering

University of Houston, Houston, TX, USA

Tutorial Presentation at ICCC'2020

http://wireless.egr.uh.edu/research.htm

ACK: Jingzhi Hu, Haobo Zhang, Shuhao Zeng, Xuelin Cao and Yali Chen

Objectives

To introduce RIS basics and potential RIS applications

- Communication/Internet of Things

To learn related mathematical tools to integrate RIS into future networks

- Optimization and machine learning

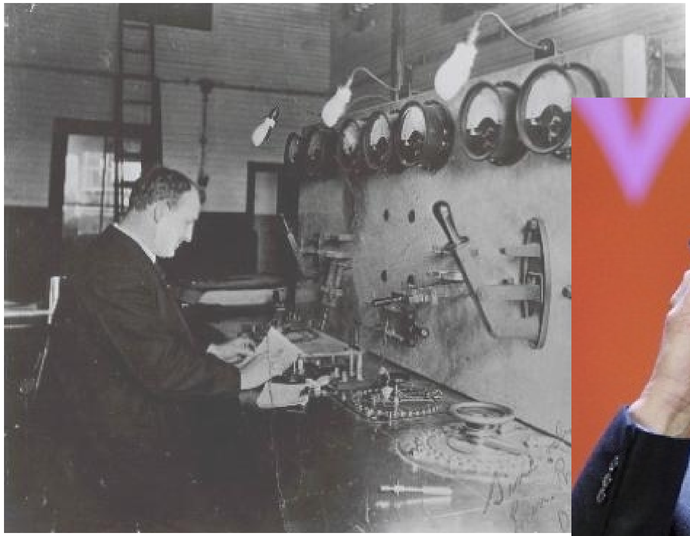
To understand how to optimize RIS aided networks

- Communication: beamforming and deployment
- Sensing: actively design multiple paths
- Localization: enlarge differences

Table of Contents

Mobile Communication Development

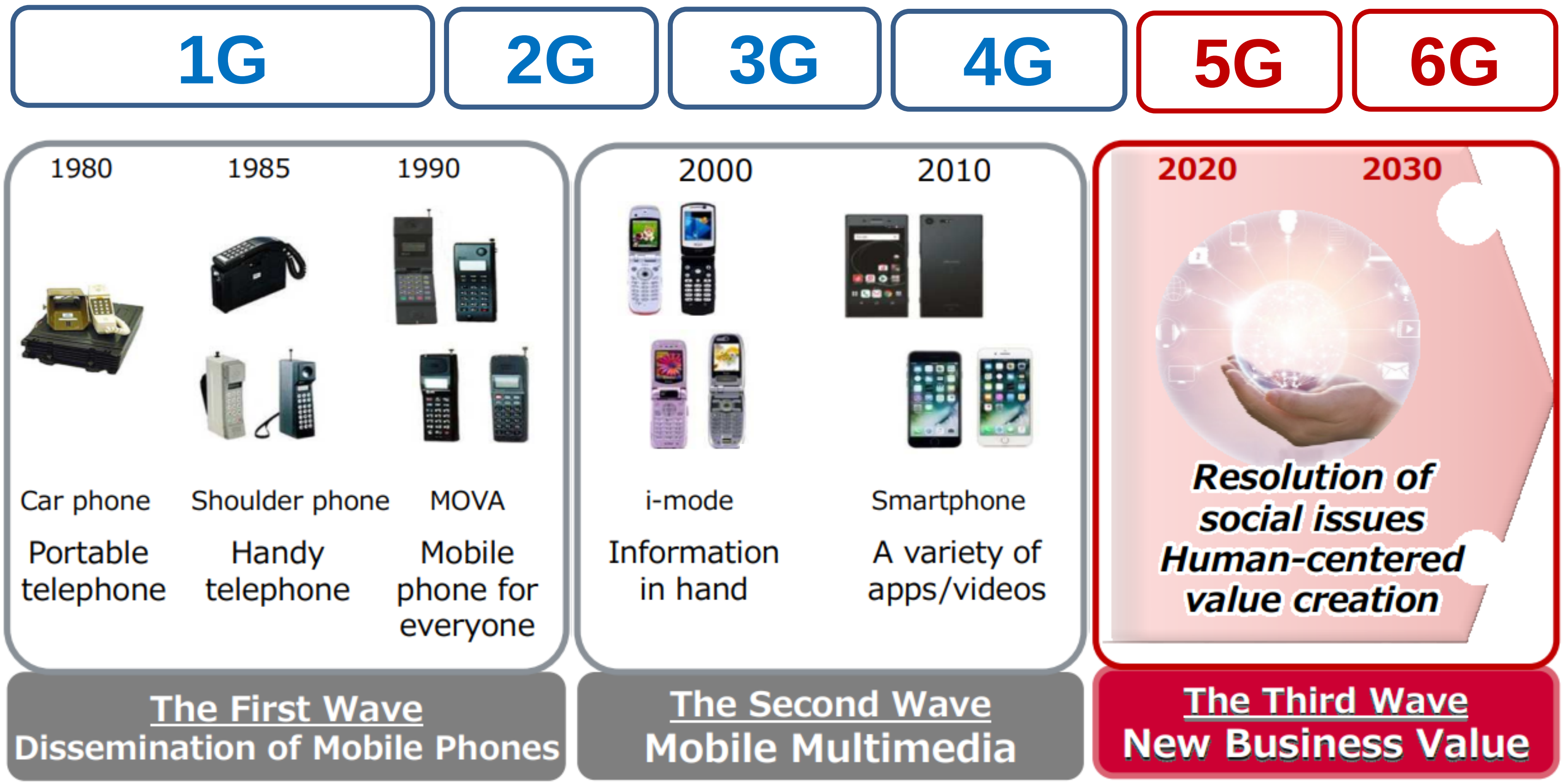
- A market needs both technology and application
- From texts … to voice … to data and everything





Standardization Facilitates Technology Evolution

- Each new evolution builds on the established market of the previous



Internet of Things

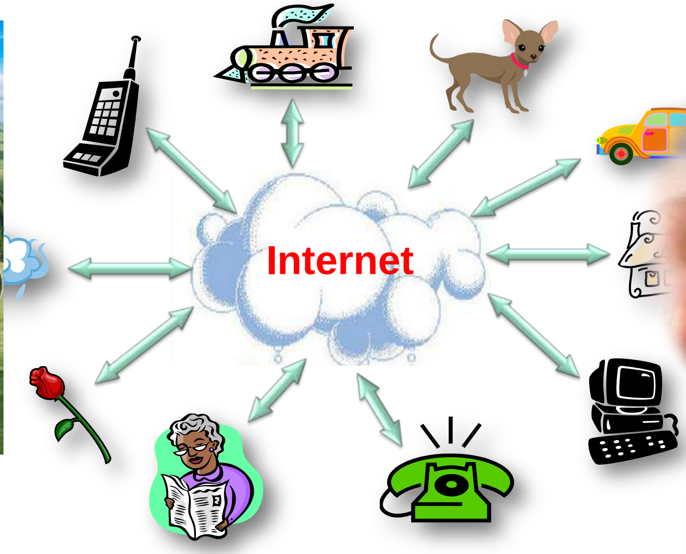
Mobile web services

- Faster
- More convenient

Internet of Things

- Connect physical world by sensors

Integrated Network

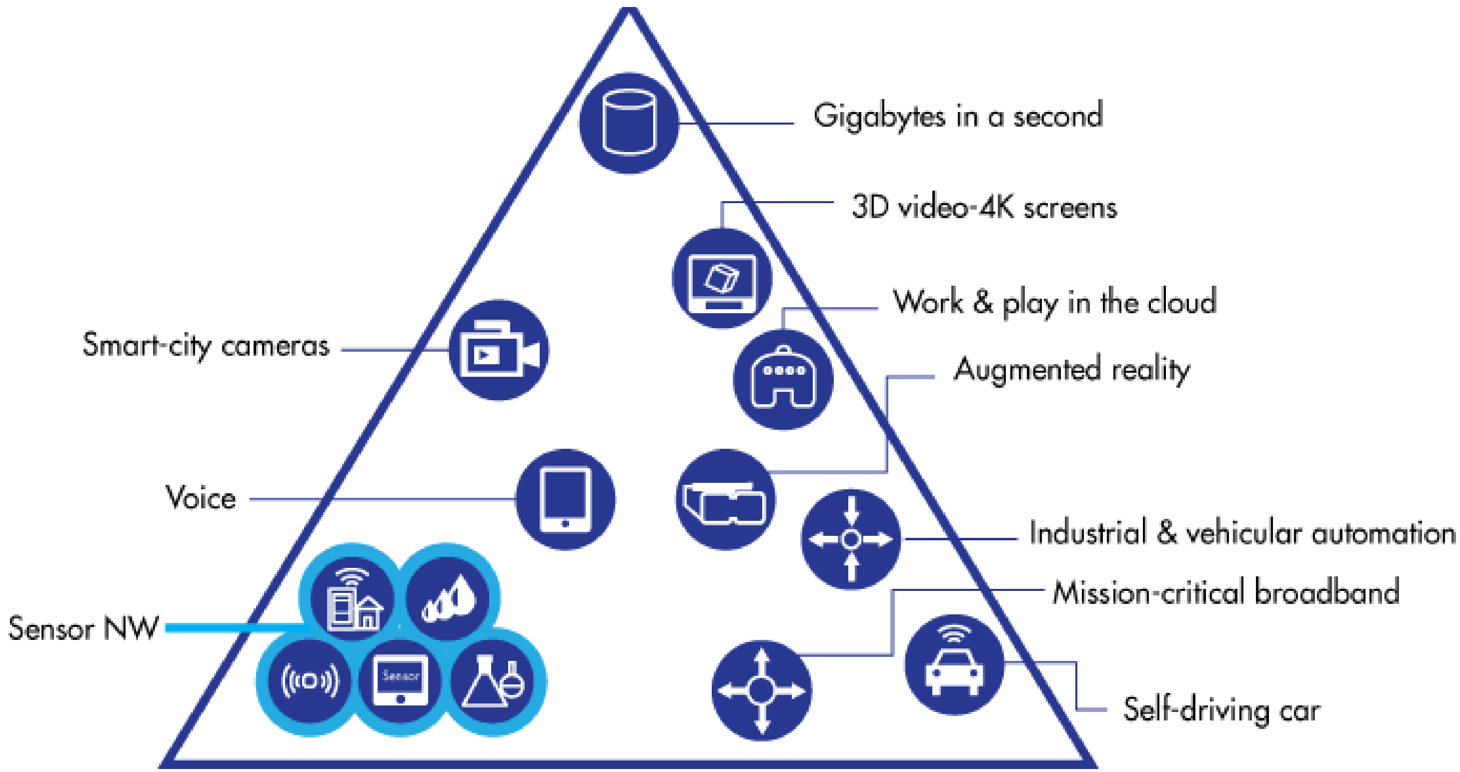


Smart Planet

Three Use Cases and KPI in 5G

Capacity Enhancement

eMBB (Gbps)



mMTC

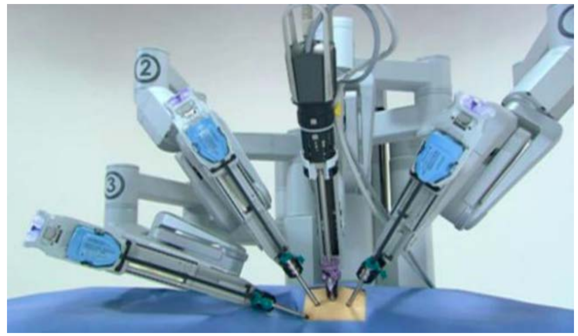
Massive Connectivity

(1M/km^2)

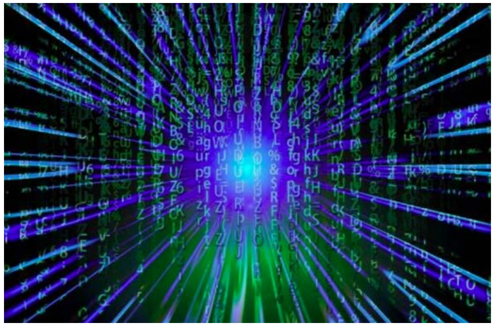
uRLLC

Ultra-high Reliability & Low Latency (1ms)

Typical Applications in 5G

VR/AR	Industry IoT	Autonomous Driving
		
AR for surgery	Auto-manufactoring	Assisted driving
		
VR for education	UAV transmissions	Smart coordination

Moving Towards 6G: Enabling Ubiquitous Intelligent Information Network



Higher data rates

Lower power consumption

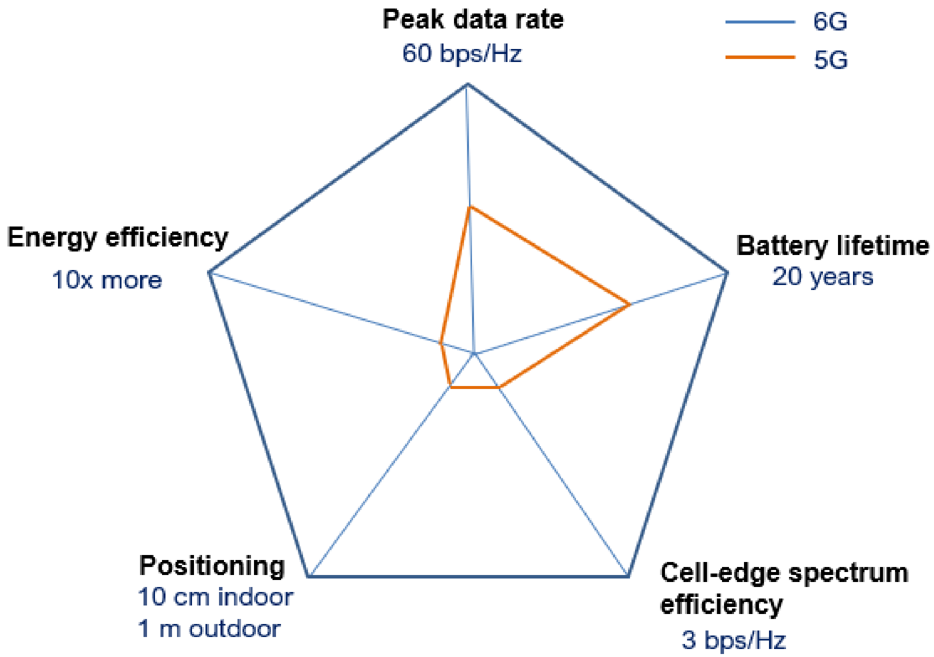
Larger Coverage

Smarter devices

Enhanced 6G KPI Targets Compared to 5G

Multi-dimensional Huge Jump

- Peak data rate (2x)
- Battery lifetime (1.5x)
- Cell-edge spectrum efficiency (10x)
- High resolution positioning (2-3x)
- Energy efficiency (10x)

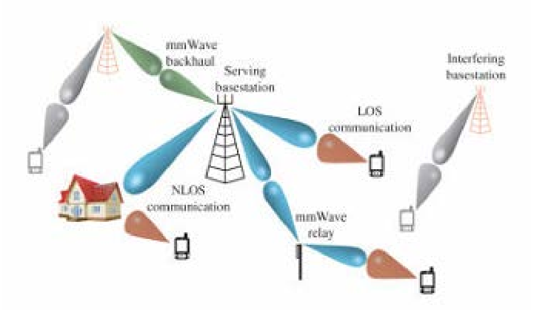


Data source: 6G White Paper, University of Oulu

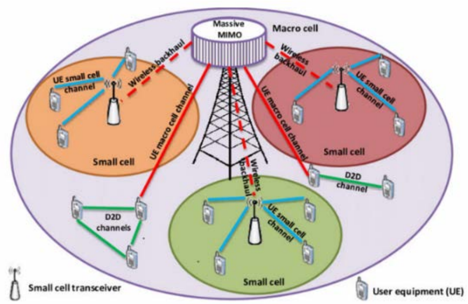
Dilemma in Current Technologies

1. Conflict between low hardware cost and high spatial resolution

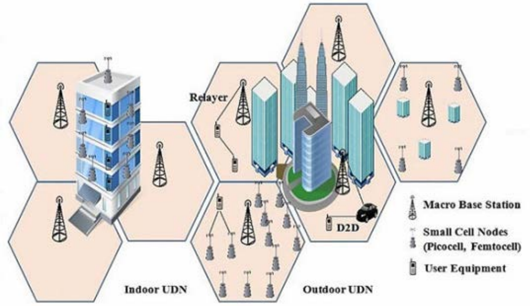
- High spatial resolution at a cost of expensive hardware
  - HF communication: dedicated RF chains lead to an increasing cost as the number of users grows
  - M-MIMO: a huge number of antennas each with a phase shifter
  - UD-Networking: a dense topology requires extremely high cost of deployment and coordination



High-Frequency Communications



Massive MIMO

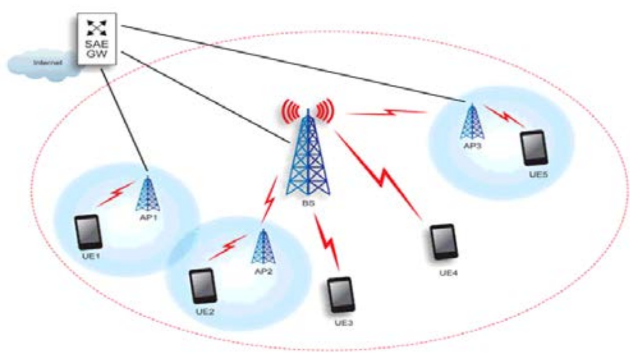


Ultra-Dense Networking

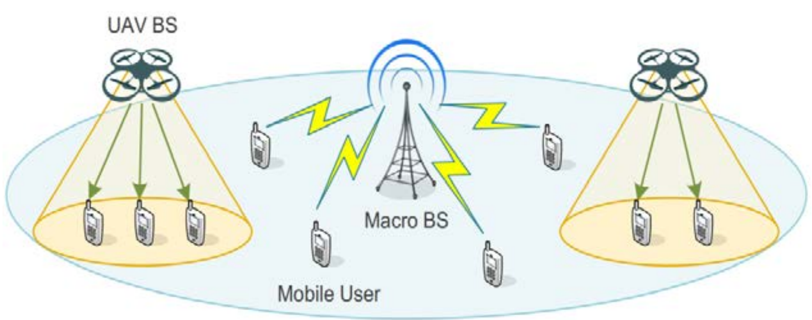
Dilemma in Current Technologies

2. Conflict between flexible network deployment and low energy consumption

- Fixed access points
  - Cannot guarantee to adapt to dynamically changing user traffic
- Moving access points
  - Involve high energy consumption (such as propulsion energy and transmission energy)



Fixed access points

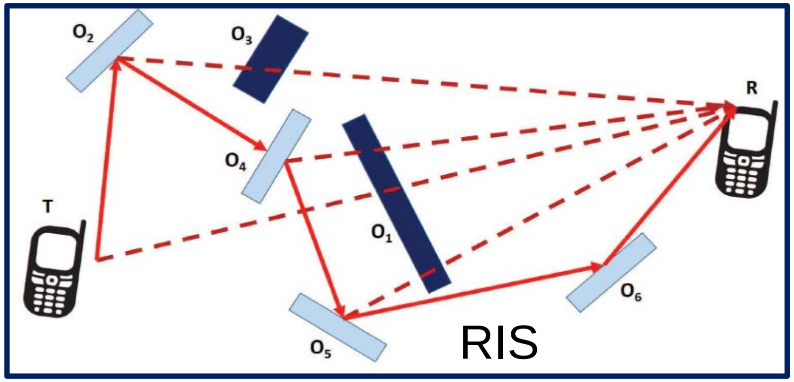


Moving access points

A Promising Solution: Reconfigurable Intelligent Surface (RIS)

Expectation on a new technology

- Low cost and high spatial resolution
- Easy to be deployed
- Compatible with 6G demands on communications and sensing



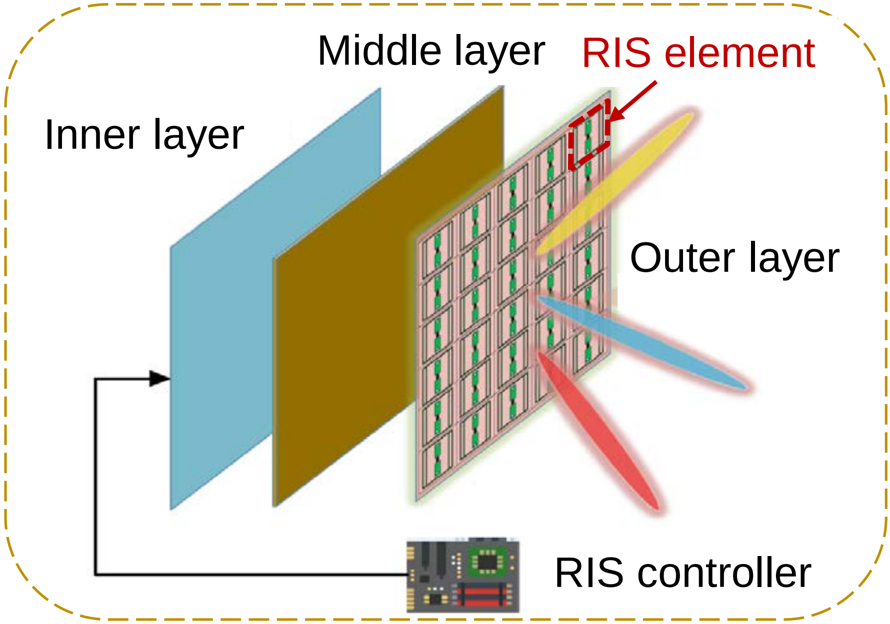
Reconfigurable intelligent surface

- Deploy easily between the transmitters and receivers
- Reflect waves into desired directions without extra hardware for communications
- Respond fast to changes in propagation environments for sensing

What is RIS?

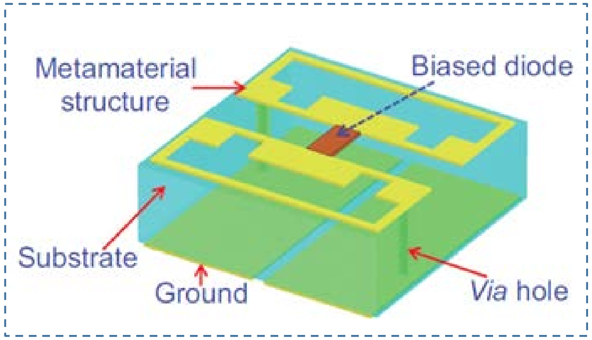
An ultra-thin meta-surface

- Outer layer：A dielectric substrate with RIS elements; directly interact with incident signals.
- Middle layer: A copper plate; avoid the signal energy leakage.
- Inner layer: A control circuit board.



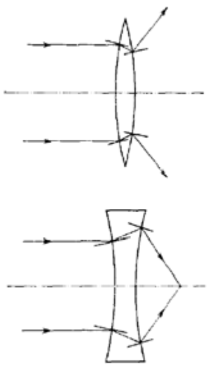
RIS element:

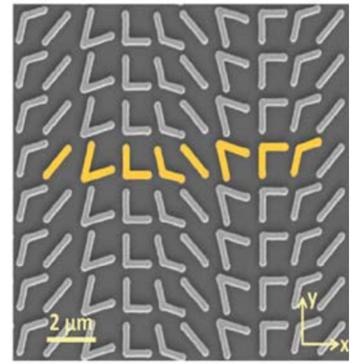
- Phase response: controlled by several PIN diodes' ON/OFF
- Working frequency: from sub-6 GHz to THz

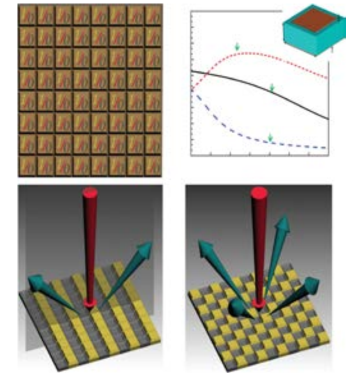


[1] L. Li, et al. "Machine-learning reprogrammable metasurface imager". Nature Communications, vol. 10, no.1, pp. 1-8, 2019.

RIS History







1967 **2000** **2006** **2011** **2014**

- 1967: Victor Veselago produced a material with negative refractive index
- 2000: Super lens could capture images below the diffraction limit
- 2006: First invisibility cloak was built
- 2011: Generalized laws of refraction and reflection was developed together with the metamaterial
- 2014: Programable metamaterial was proposed for real-time control

Why do We Need RIS?

Low-cost

- RIS is passive and does not process the signals
- There is no need to have ADC/DAC and amplifiers

High Spatial Resolution

- RIS integrates a large number of antenna elements into a compact space in the form of a spatially continuous electromagnetic aperture
- This makes it possible to generate a beam with any direction

Easy to Implement

- Thin surfaces make them flexible to deploy and extend
- Compatible with other existing technologies

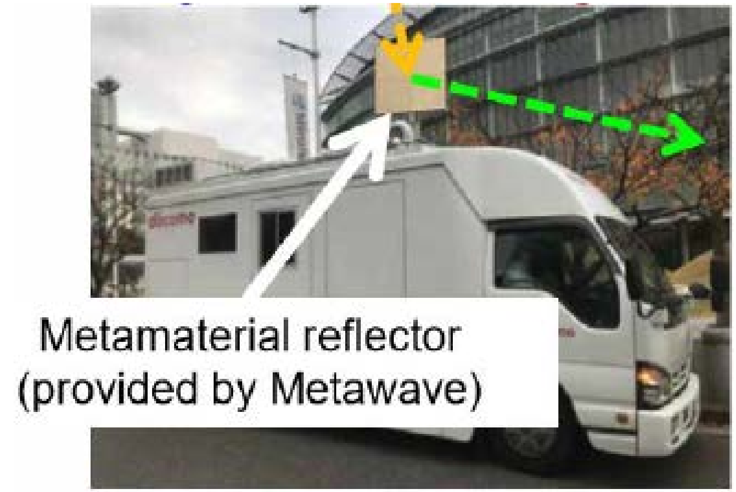
RIS vs. Existing Technologies

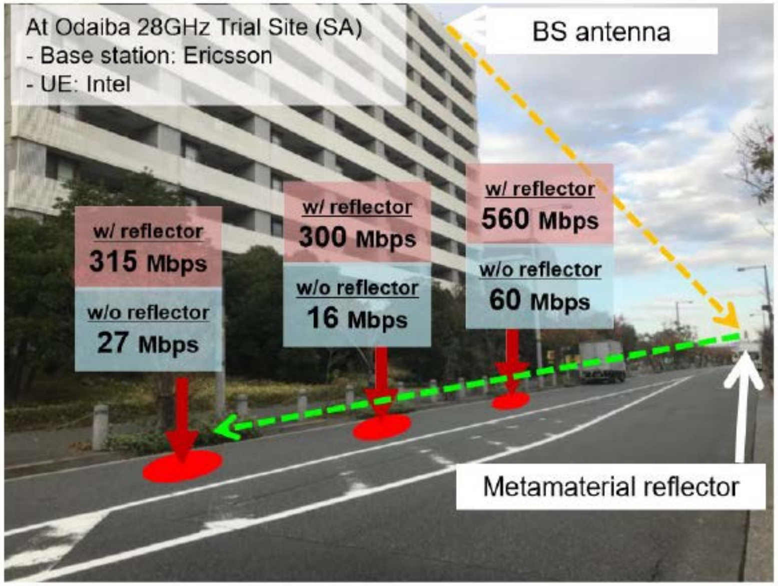
Technology	Operating mechanism	Duplex	No. of transmit RF chains needed	Hardware cost	Energy consumption
RIS	Passive/Active, reflection	Full	0	Low	Low
Massive MIMO	Active, transmission/ reception	Half/full	N	Very high	Very high
Relay	Active, reception and transmission	Half/full	N	High	High
Backscatter	Passive, reflection	Full	0	Very low	Very low

Prototypes

NTT Docomo: Metawave

- Extend 5G coverage in indoor and outdoor environments
- NTT Docomo demonstrated a 10x increase in communication speed in a true-to-life outdoor scenario in Tokyo.
- Applications: Extending 5G's range to service dead zones, and to bend and point signals around corners and connect to backhaul radios.

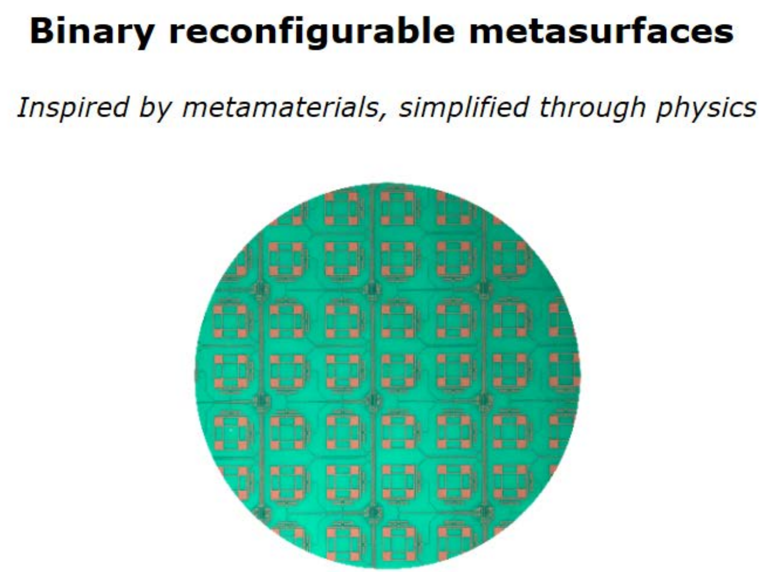




Prototypes

Duke University & Institut Langevin: Greenerwave

- Use binary RIS to detect and analyze motion
- Capture the temporal variations of motion and discern information
- Achieve Non line-of-sight motion detection
- Demonstrate that the use of RIS can substantially enhance the performance

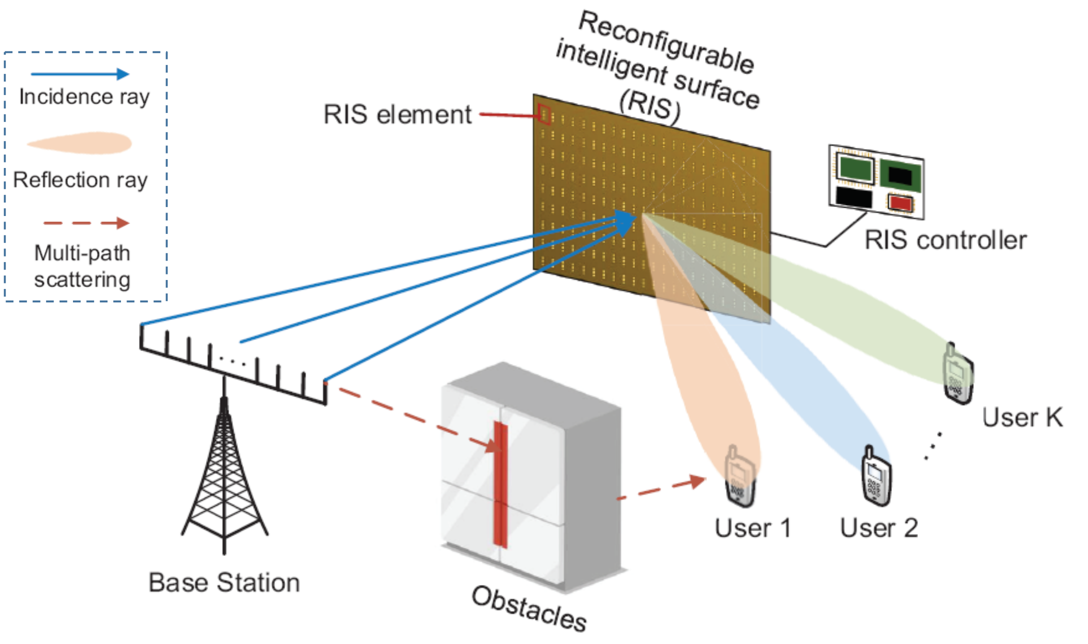


[2] P. del Hougne, et al, "Dynamic Metasurface Aperture as Smart Around-the-Corner Motion Detector", Science Report, vol. 8, no. 1, pp. 1-10, 2018.
[3] http://greenerwave.com/our-team/

Transmission Model

Transmission Process

- EM waves impinging on an RIS create induction current
- The RIS reflects the signals towards the users
- During the reflection, the RIS imposes additional response on the phase and the amplitude

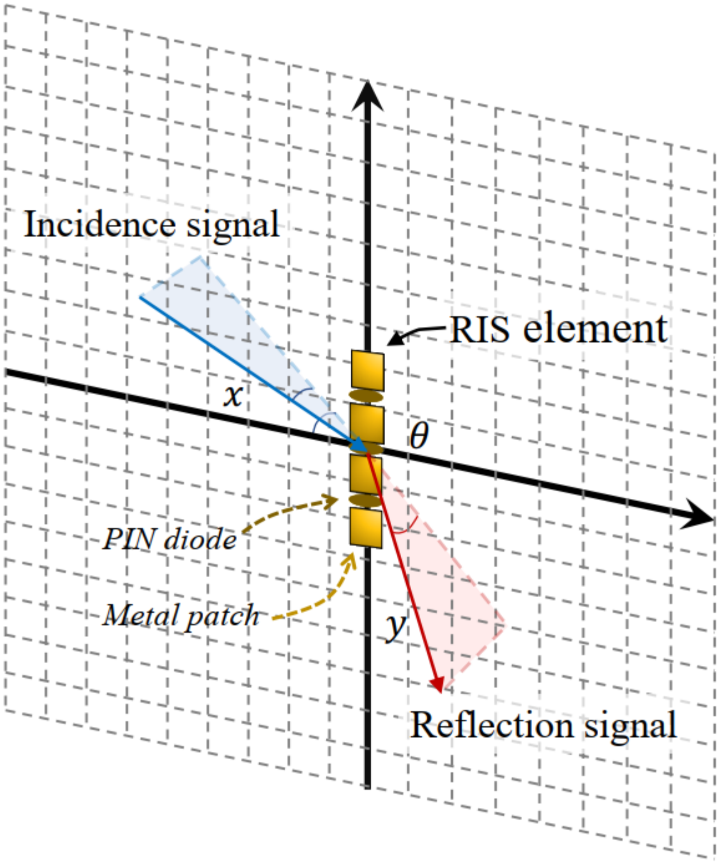


Reflection Model

Reflected signal

$$y = \Gamma e^{j\theta} x$$

- $\Gamma \in [0,1]$: reflection amplitude
  - $\Gamma = 0$: absorbed
  - $\Gamma = 0$: fully reflected
- $\theta \in [0,2\pi]$: phase shift for each element
  - In practical systems, phase shifts are controlled by PIN diodes, leading to the discrete phase shifts
  - K PIN diodes: 2^K phase shifts with uniform interval



Channel Model

Rician Model

- User-RIS-BS links act as the dominant LoS component
- All other paths contributes the NLoS

$$\tilde{h}_{m,n} = \sqrt{\frac{\kappa}{\kappa + 1}}\, h_{m,n} + \sqrt{\frac{1}{\kappa + 1}}\, \hat{h}_{m,n},$$

Ratio of LoS to NLoS LoS NLoS

- Product of distance path loss

$$|h_{m,n}|^2 \propto d_{m,n}^{-\alpha} D_{m,n}^{-\alpha}$$

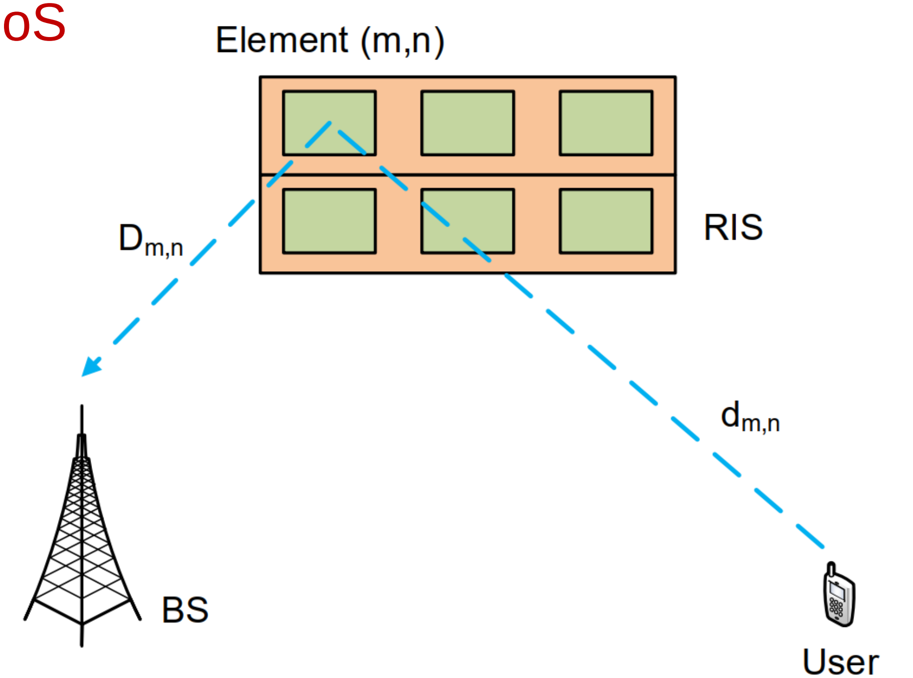
$$|\tilde{h}_{m,n}|^2 \propto d_{m,n}^{-\alpha} D_{m,n}^{-\alpha}$$

- Received signal

noise

$$y = \sum_{m,n} \underline{\Gamma e^{j\theta_{m,n}}}\, \underline{\tilde{h}_{m,n}}\, x + w$$

Reflection coefficient Channel gain



Channel Model

Rayleigh Model

- **Q**: Spatial channels of RIS-BS
- **F**: RIS response
- **G**: Spatial channels of user-RIS
- Composite channel

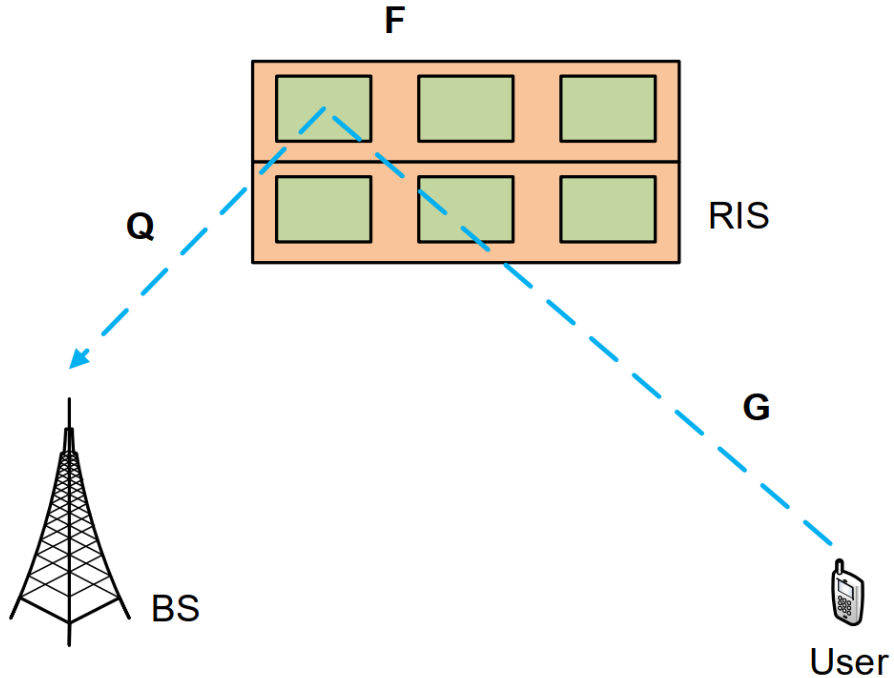
$$H = QFG$$

- Received signal

$$y = QFG\underline{s} + w$$

Transmitted signals noise

- Follows product of distance path loss



Applications: Communication

Spectrum Efficiency Enhancement

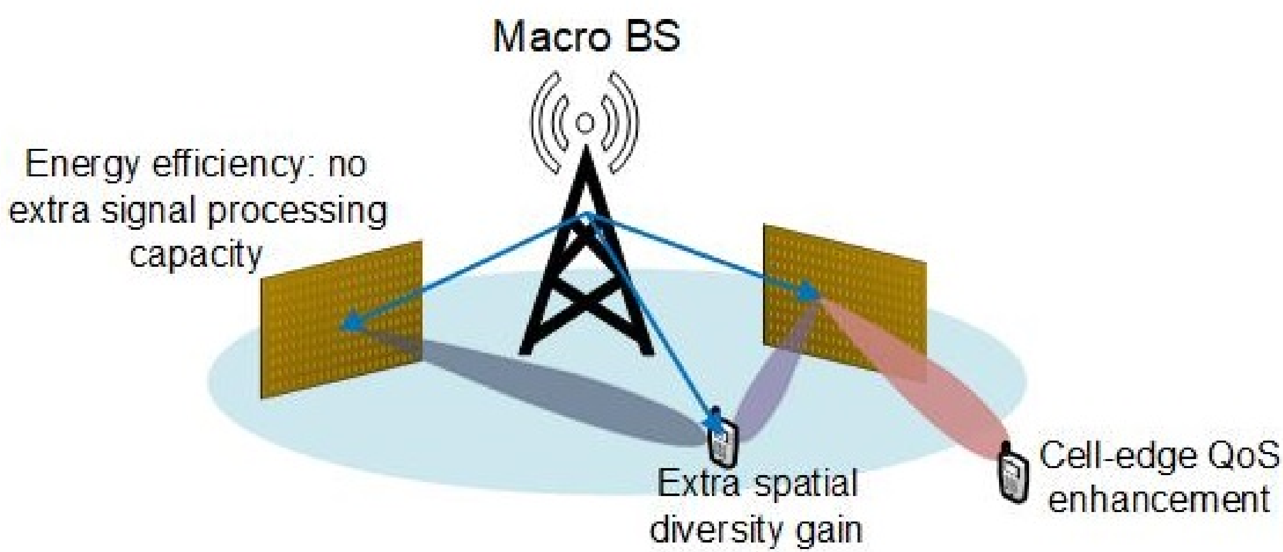
- RIS provides extra spatial diversity gain

Coverage Extension

- RIS as a passive relay can assist APs to serve cell-edge users

Energy Efficiency Improvement

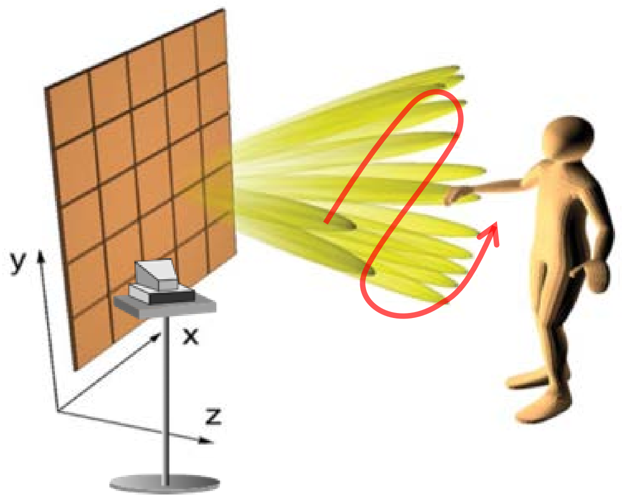
- RIS does not need extra energy-consuming hardware to be deployed

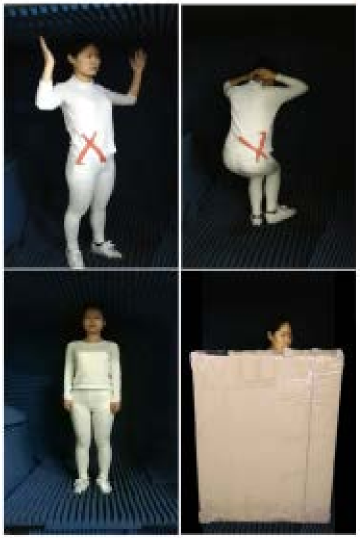


Applications: Sensing and Localization

Sense the changes in propagation environments

- Movement of targets leads to changes in propagation environments
- The receiver can infer such changes based on received information
- RIS actively creates various propagation environments to make target movements significantly distinguishable





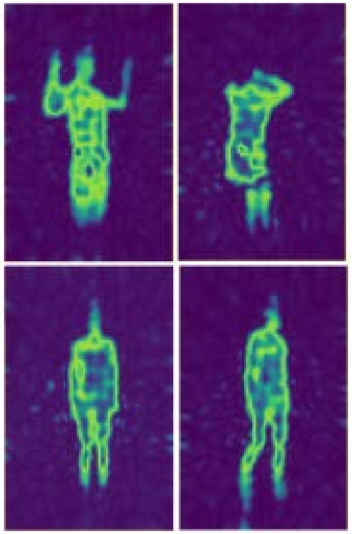


Table of Contents

- Background
  - 6G Communications and Requirements
  - RIS Basics and Potential Applications
- Mathematical Tools
  - Optimization Theory
  - Machine Learning
- RIS-aided Cellular Communications
  - Limited Phase Shifts Effect
  - RIS aided Coverage Extension
  - RIS aided MIMO Communications
  - RIS aided D2D Communications
  - RIS assisted MAC Access
- RIS-aided Internet of Things
  - RIS aided RF Sensing
  - RIS aided indoor Localization

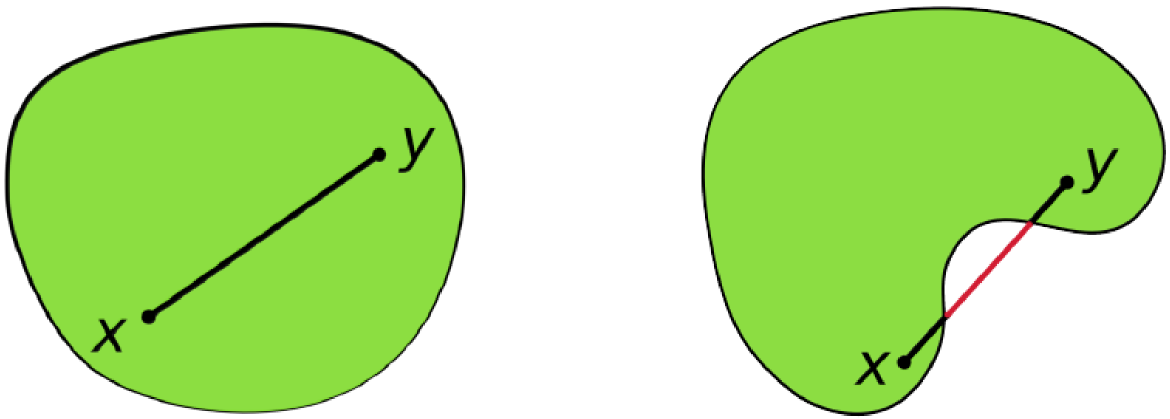
Optimization Theory

- Convex Set and Convex Functions
- Gradient Descent and Newton's Method
- Duality and KKT Condition

Convex Set

- A set $S \subseteq \mathbb{R}^n$ is convex if for any $x, y \in S$ and any $\lambda \in [0, 1]$, we have

$$\lambda x + (1 - \lambda) y \in S$$



- There are convex sets and non-convex sets
- **Note:** There is no such thing as a “concave set”

Convex Function

- Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies

$$f(\lambda \mathbf{a} + (1-\lambda)\mathbf{b}) \leq \lambda f(\mathbf{a}) + (1-\lambda) f(\mathbf{b}),\ \forall\ \lambda \in [0,1], \mathbf{a}, \mathbf{b} \in \mathbb{R}^n$$

  then f is called a convex function. [or $-f$ is called a concave function]
- Intuitively, a convex function "holds water"

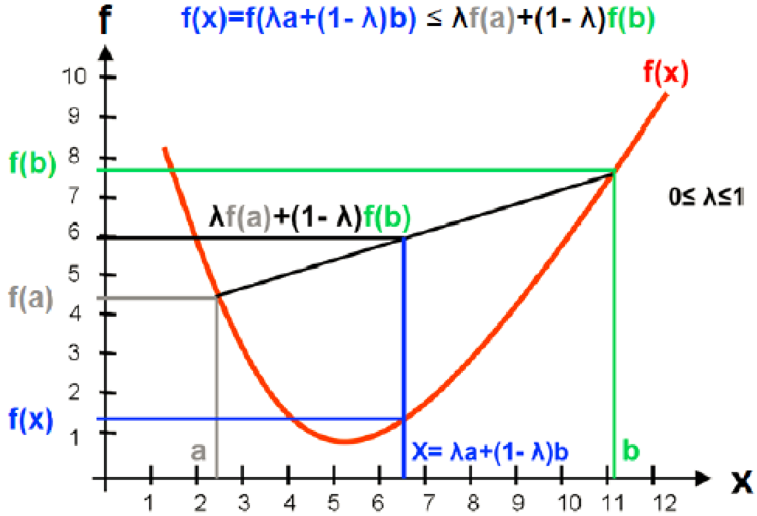


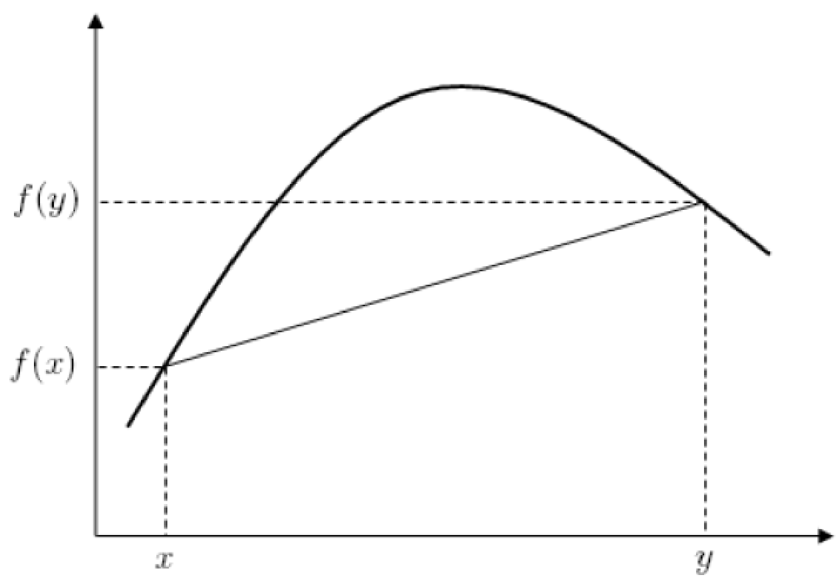
Figure: Illustration of a Convex Set [Moura 14]

Concave Function

- A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is called *concave* if for all $x, y \in \mathbb{R}^K$ and for all $\lambda \in [0, 1]$, we have

$$f[\lambda x + (1-\lambda)y] \geq \lambda f(x) + (1-\lambda)f(y).$$

- **Question:** A linear function is convex or concave function?



Convex Optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X \end{array}$$

- Here f is continuously differentiable, X is a convex set
- Convex set X means we allow the following types of constraints
  1. $g(x) \leq 0$ where $g(x)$ is a convex function
  2. $h(x) = 0$ where $h(x)$ is an affine function: $Cx + d = 0$

Convex means local minimum = global minimum

Algorithms to Find Optimum

- Now we have settled the question of when global min = local min
- We are then interested in finding such "global min"
- **Easiest way**: Simply solve $\nabla f(\mathbf{x}) = 0$!
- Suppose $\nabla f(\mathbf{x}) = 0$ is not easy, don't know how to solve
- Then what?

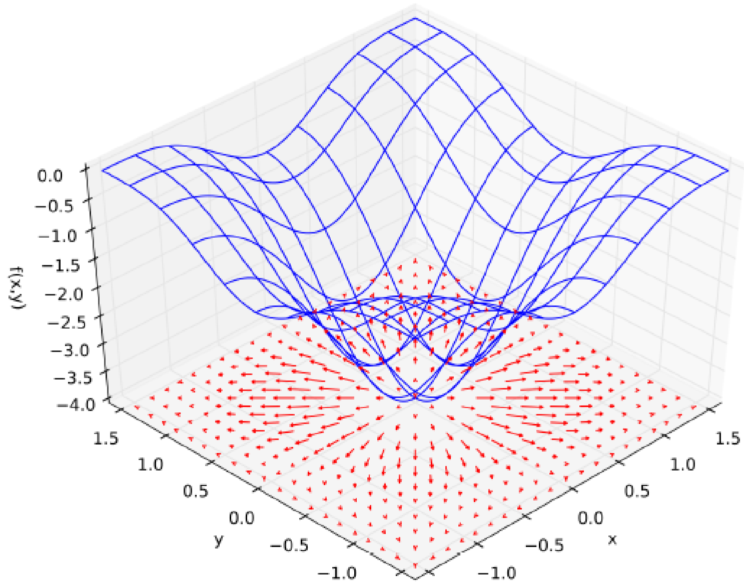


Figure: The gradients of a function (Wikipedia: Gradient)

Gradient Descent Method

- If $\nabla f(\mathbf{x}) = 0$, then $\mathbf{x}$ is a candidate solution (satisfying first-order sufficient condition); Done
- If $\nabla f(\mathbf{x}) \neq 0$, there is an interval $(0, \delta)$ of stepsizes such that

$$f(\mathbf{x} - \alpha \nabla f(\mathbf{x})) < f(\mathbf{x}), \ \forall\, \alpha \in (0, \delta).$$

- Show this using Mean Value Theorem [on board]?
- More generally, if a given direction $\mathbf{d}$ that is with obtuse angle with $\nabla f(\mathbf{x})$

$$\langle \nabla f(\mathbf{x}), \mathbf{d} \rangle < 0$$

there is an interval $(0, \delta)$ of stepsizes such that

$$f(\mathbf{x} + \alpha \mathbf{d}) < f(\mathbf{x}), \ \forall\, \alpha \in (0, \delta).$$

Iterative Descent Method

$$\mathbf{x}^{r+1} = \mathbf{x}^r + \alpha_r \mathbf{d}^r, \ r = 0, 1, \cdots$$

where, if $\nabla f(\mathbf{x}^r) \neq 0$, the direction $\mathbf{d}^r$ satisfies $\nabla f(\mathbf{x}^r)\mathbf{d}^r < 0$, and α^r is a positive stepsize

- General Case: Gradient descent methods

$$\mathbf{x}^{r+1} = \mathbf{x}^r - \alpha_r \mathbf{D}^r \nabla f(\mathbf{x}^r), \ r = 0, 1, \cdots$$

where $\mathbf{D}^r$ is a positive definite matrix

- Special case I: Steepest descent

$$\mathbf{x}^{r+1} = \mathbf{x}^r - \alpha_r \nabla f(\mathbf{x}^r), \ r = 0, 1, \cdots$$

- Special case II: Newton's method

$$\mathbf{x}^{r+1} = \mathbf{x}^r - \alpha_r \left(\nabla^2 f(\mathbf{x}^r)\right)^{-1} \nabla f(\mathbf{x}^r), \ r = 0, 1, \cdots$$

Shortcoming of Gradient Method

However, in practice steepest descent may have slow convergence

Figure: The Steepest Descent (ERASIP: DSPHumour)

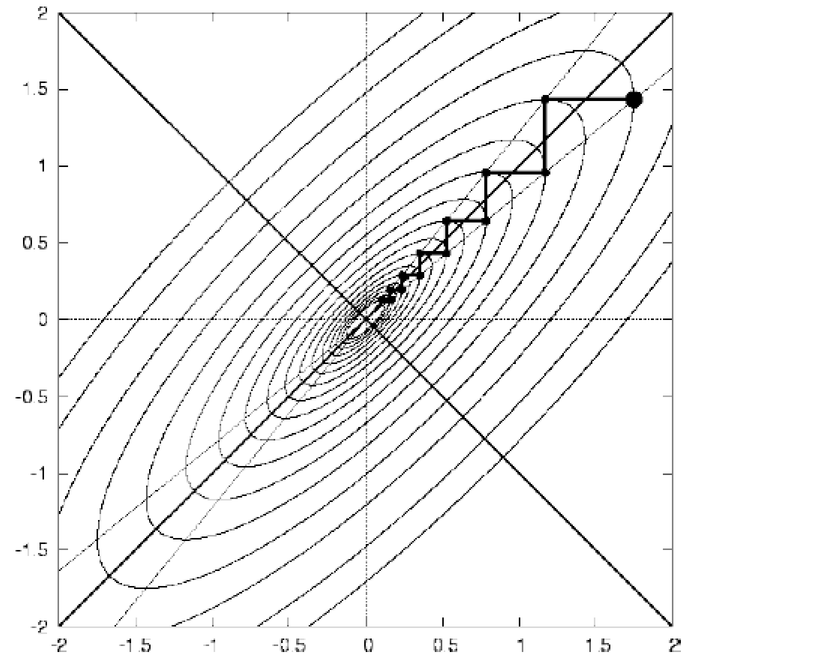


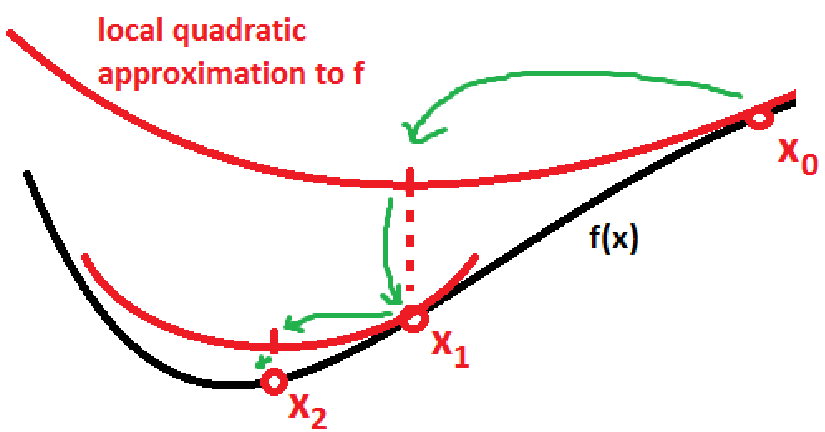
Figure: The Steepest Descent in Practice (Komarix.org)

Newton's Method

- Newton's method: generally fast convergence
  - Basically it treats the objective (locally) as a quadratic problem around $\mathbf{x}^r$

$$f(\mathbf{x}) \approx f(\mathbf{x}^r) + \langle \nabla f(\mathbf{x}^r), \mathbf{x} - \mathbf{x}^r \rangle + \frac{1}{2}(\mathbf{x} - \mathbf{x}^r)^T \nabla^2 f(\mathbf{x}^r)(\mathbf{x} - \mathbf{x}^r)$$

  - Question: how many iterations does it take for Newton method to minimize a quadratic function f?
  - Caution: very difficult to make it numerically stable, needs more information than the steepest descent method



Lagrangian Multiplier

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h_i(x) = 0, \quad i = 1, \cdots, m \\ & g_j(x) \leq 0, \quad j = 1, \cdots, n \end{array}$$

- **Reminder**: The problem is called convex problem if
  1. $f(x)$ is a convex function
  2. $h_i(x)$ is an affine function, i.e., $h_i(x) = Ax + b$
  3. $g_j(x)$ is a convex function
- The Lagrangian can be formed using the Lagrangian multipliers $\lambda_i \geq 0$ and $\nu_i \in \mathbb{R}$

$$L(x, \lambda, \nu) = f(x) + \sum_{j=1}^{n} \lambda_j g_j(x) + \sum_{i=1}^{m} \nu_i h_i(x)$$

Duality

- The Lagrangian dual function

$$L^*(\lambda, \nu) = \inf_{x \in X} L(x, \lambda, \nu) = \inf_{x \in X} f(x) + \sum_{j=1}^{n} \lambda_j g_j(x) + \sum_{i=1}^{m} \nu_i h_i(x)$$

- The Dual Problem

$$\max_{\lambda, \nu} \quad L^*(\lambda, \nu), \quad \text{s.t. } \lambda \geq 0$$

- λ_i and ν_i's can be viewed as "prices" for violating the constraints
- Let f^* be the optimal value of $f(x)$
- The Lagrangian dual L^* is
  1. A concave function: even when the original problem is not convex
  2. A lower bound: for $\lambda \geq 0$, $L^*(\lambda, \nu) \leq f^*$

Duality

- Let d^* be the optimal objective of the dual
- Weak duality: $d^* \leq f^*$
  1. Always true
  2. Non-trivial lower bound for hard problems
  3. Useful in approximation algorithms
- Strong duality: $d^* = f^*$
  1. Does not hold in general
  2. If holds, sufficient to solve the dual
  3. How to check if it holds?
- Constraint qualification
  1. Normally true for convex problems
  2. True if the problem is convex; And it is strictly feasible, i.e. there exists a $x \in X$ such that
  $$h_i(x) = 0, \quad g_j(x) < 0$$
  3. The above condition is known as the Slater's condition

KKT Condition: When to Stop?

$$
\begin{aligned}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & h_i(x) = 0, \quad i = 1, \cdots, m \qquad (1) \\
& g_j(x) \le 0, \quad j = 1, \cdots, n \qquad (2)
\end{aligned}
$$

Any optimal and dual pairs $\tilde{x}$ and $(\tilde{\lambda}, \tilde{\nu})$ must satisfy

Albert Tucker

$$\nabla f(\tilde{x}) + \sum_{j=1}^{n} \tilde{\lambda}_j \nabla g_j(\tilde{x}) + \sum_{i=1}^{m} \tilde{\nu}_i \nabla h_i(\tilde{x}) = 0_{K\times 1}$$

$g_j(\tilde{x}) \le 0, \forall j = 1, ..., n,$ (primal feasibility)

$h_i(\tilde{x}) = 0, \forall i = 1, ..., m,$ (primal feasibility)

$\tilde{\lambda}_j \ge 0, \forall j = 1, ..., n,$ (dual feasibility)

$g_j(\tilde{x}) \times \tilde{\lambda}_j = 0, \forall j$ (complementarity).

Harold Kuhn

Barrier Function

reformulation of (1) via indicator function:

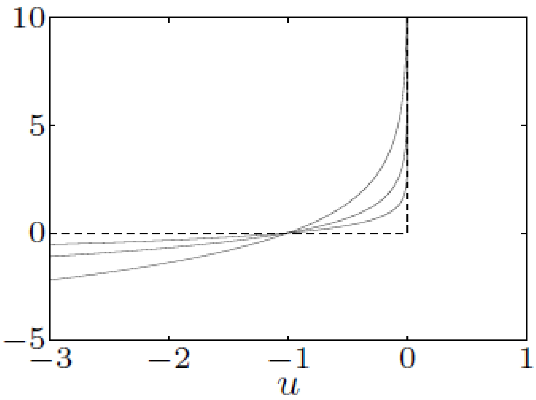
$$\begin{array}{ll}\text{minimize} & f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \\ \text{subject to} & Ax = b\end{array}$$

where $I_-(u) = 0$ if $u \leq 0$, $I_-(u) = \infty$ otherwise (indicator function of $\mathbf{R}_-$)

approximation via logarithmic barrier

$$\begin{array}{ll}\text{minimize} & f_0(x) - (1/t)\sum_{i=1}^m \log(-f_i(x)) \\ \text{subject to} & Ax = b\end{array}$$

- an equality constrained problem
- for $t > 0$, $-(1/t)\log(-u)$ is a smooth approximation of I_-
- approximation improves as $t \to \infty$

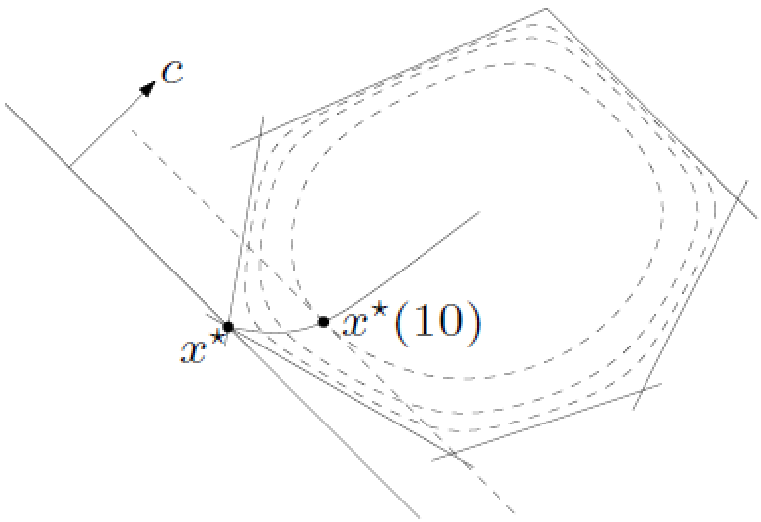


Interior Point Method

given strictly feasible x, $t := t^{(0)} > 0$, $\mu > 1$, tolerance $\epsilon > 0$.

repeat

1. *Centering step.* Compute $x^\star(t)$ by minimizing $tf_0 + \phi$, subject to $Ax = b$.
2. *Update.* $x := x^\star(t)$.
3. *Stopping criterion.* **quit** if $m/t < \epsilon$.
4. *Increase* t. $t := \mu t$.



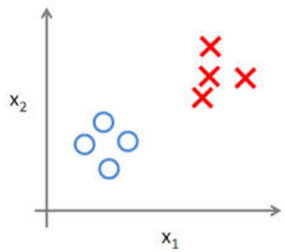
Learning Methods

- Classical Machine Learning
- Deep Learning
- Reinforcement Learning

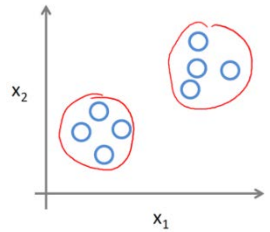
Classical Machine Learning

"Computers: learn without being explicitly programmed"

- Types:
  - Supervised Learning:
    - Example inputs (features) and their desired outputs (labels)
    - Goal: learn a general rule that maps inputs to outputs
    - SVM, neural networks, etc.

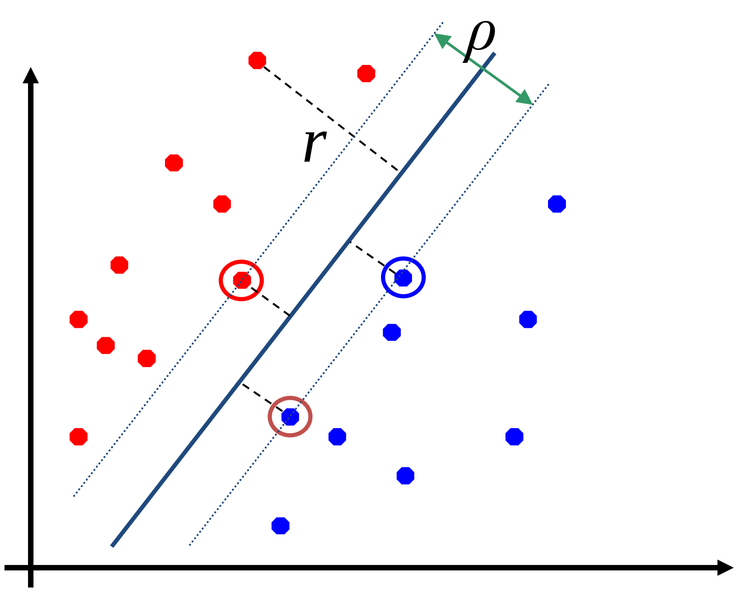


  - Unsupervised Learning:
    - No labels
    - Find structure in its input
    - Goal: discover hidden patterns in data
    - Clustering, K-means, etc.



Supervised Learning: SVM

- Distance from sample x_i to the separator: r
- Support vectors: samples closest to the hyperplane
- Margin ρ: the distance between support vectors
- Objective: maximize the margin ρ



$$r = \frac{y_i(w^T x_i + b)}{|w|} = \frac{1}{|w|}$$

$$\rho = \frac{2}{|w|}$$

$$y = -1 : w^T x_i + b \leq -\frac{\rho}{2}$$

$$y = 1 : w^T x_i + b \geq \frac{\rho}{2}$$

Supervised Learning: Applications

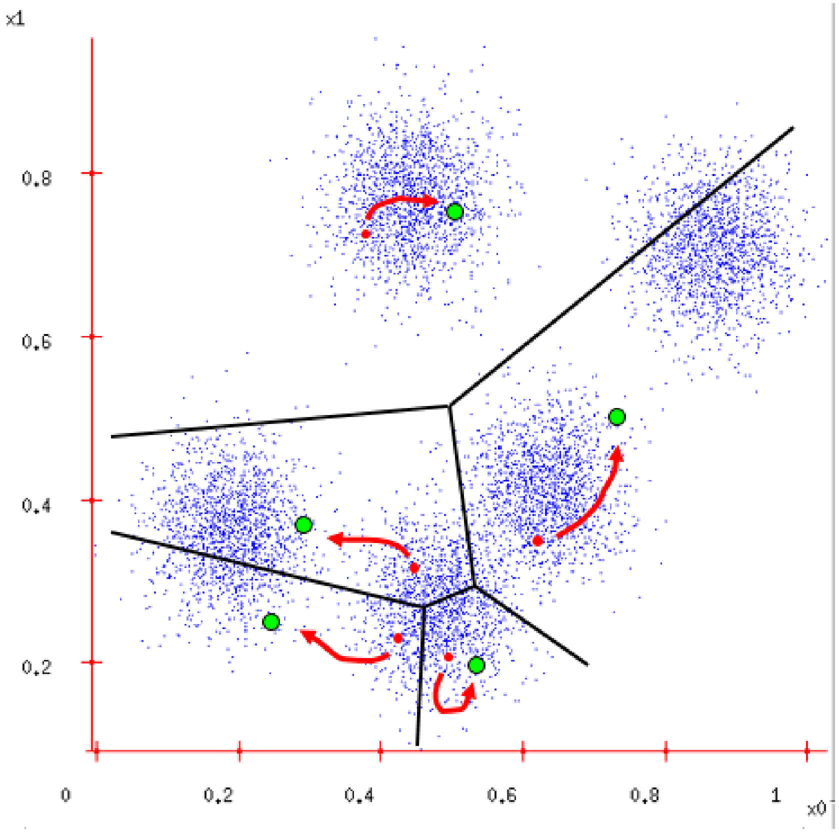
- The best performers for a number of classification tasks ranging from text to genomic data.

- Complex data types beyond feature vectors (e.g. graphs, sequences, relational data) by designing kernel functions for such data.

- Extend to a number of tasks such as regression, principal component analysis, etc.

Unsupervised Learning: K-Means

- Ask user how many clusters they'd like. (e.g. *k*=5)
- Randomly guess *k* cluster center locations
- Each data point: find out which center it's closest to
- Each center: find the centroid of the points it owns
- Change center
- Repeat until terminated

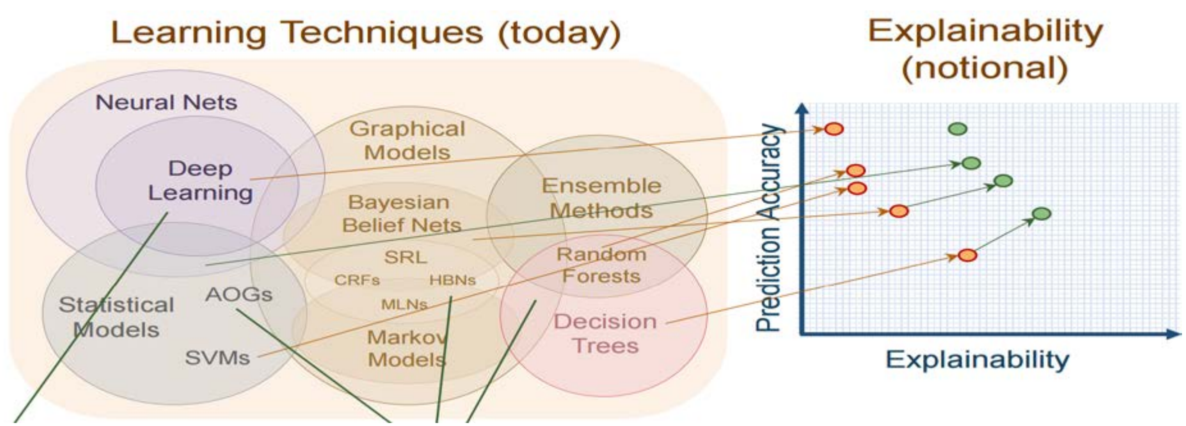


Unsupervised Learning: Applications

- Data mining
- Acoustic data in speech understanding to convert waveforms into one of k categories (known as Vector Quantization or Image Segmentation)
- Also used for choosing color palettes on old fashioned graphical display devices and Image Quantization

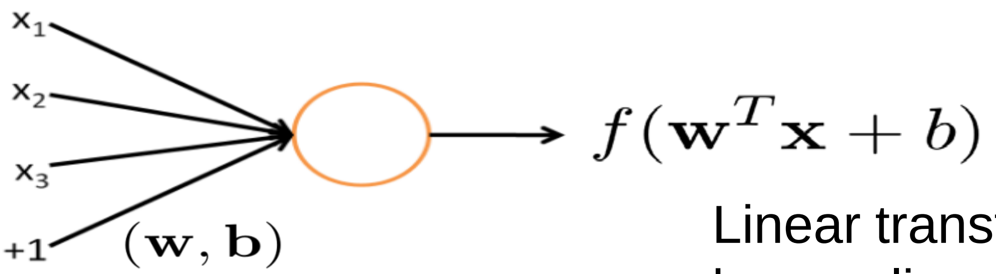
Deep Learning: Motivations

- Classic Methods
  - Do not have a lot of data, or
  - Training data have categorical features
  - A more explainable model
  - A high run-time speed
- Deep Learning
  - A lot of training data of the same or related domain
  - Improve Domain Adaptation
  - Appropriate scaling and normalization have to be done
  - Much slower

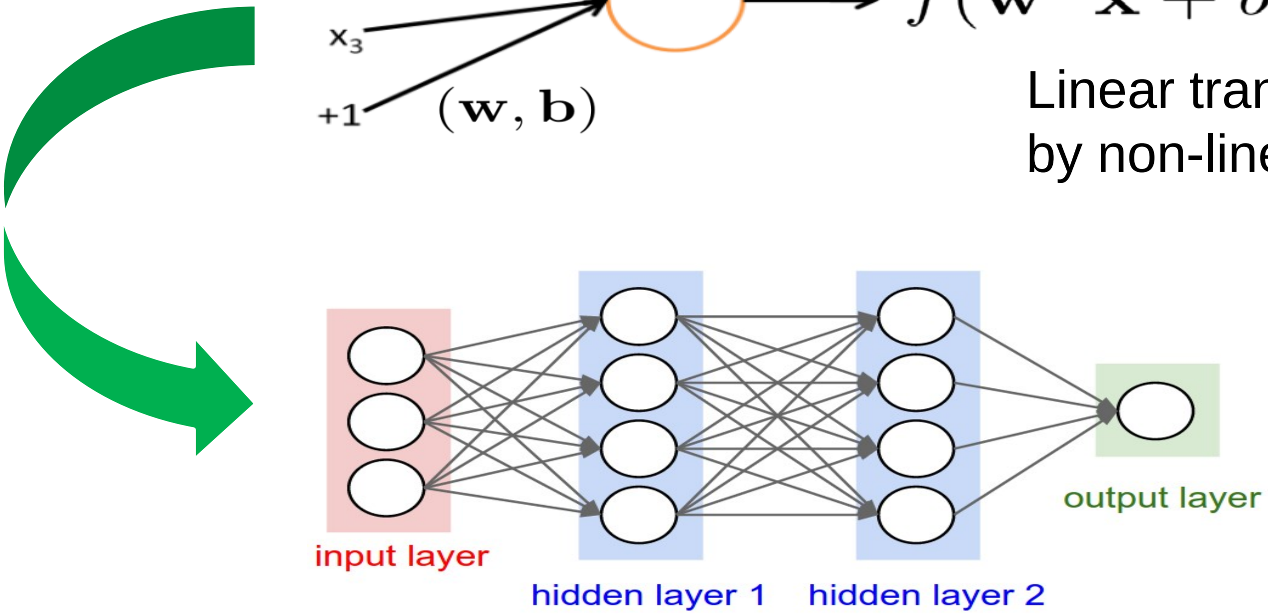


Deep Learning: Basic Idea

- Add Hidden Layers in Neural Networks



Linear transformation followed by non-linear rectification



- ☐ More parameters
- ☐ More non-linear parts

Typical Deep Neural Networks

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Deep Belief Networks

Convolutional Neural Networks (CNNs)

INPUT IMAGE

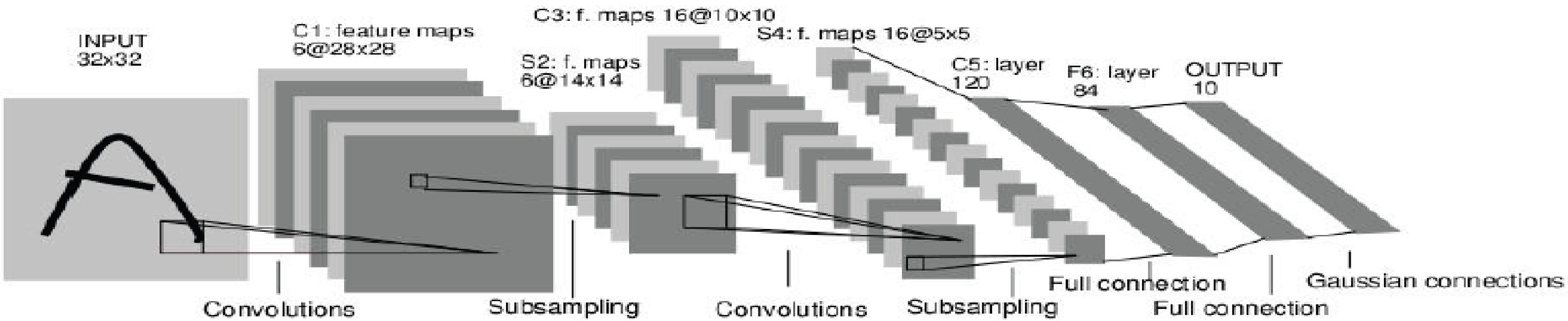
CONVOLUTION (LEARNED)

NON-LINEAR RECTIFIER

POOLING

FEATURE MAPS

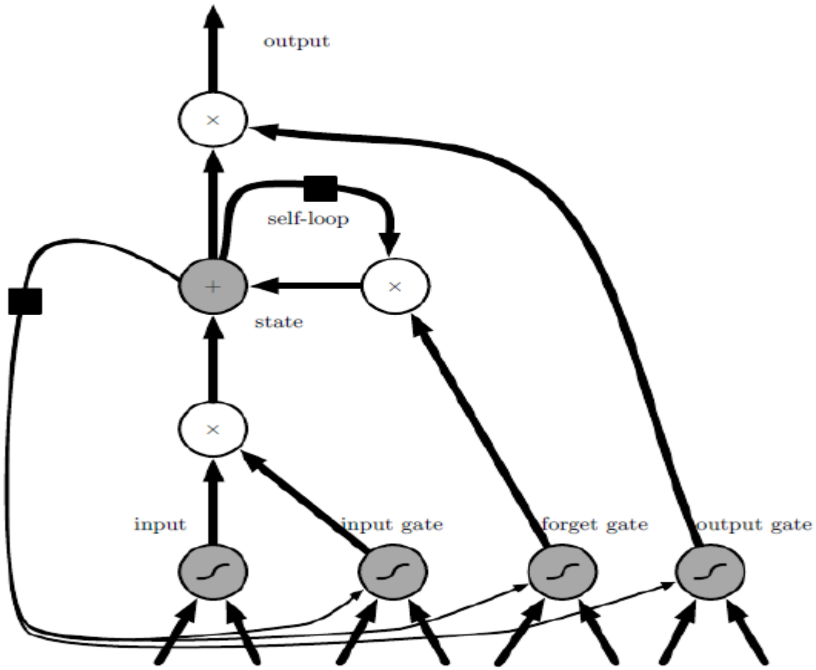
Work well in image processing!



[4] LeCun, Yann. "LeNet-5, convolutional neural networks". Retrieved Nov. 2013.

Recurrent Neural Networks (RNNs)

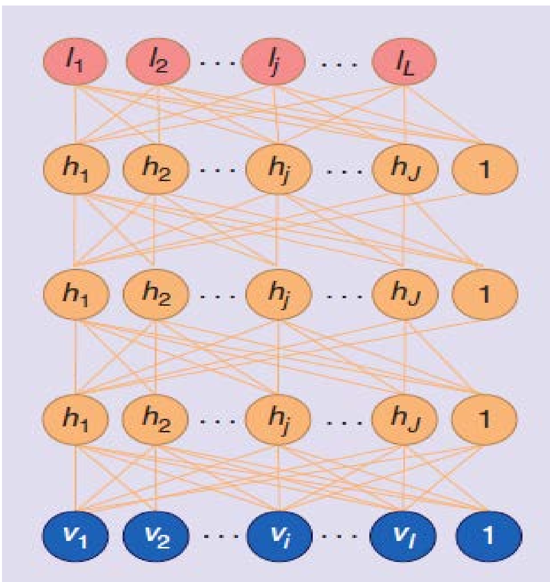
- Produce an output at each time step and have recurrent connections between hidden units
  - Long Short-Term Memory

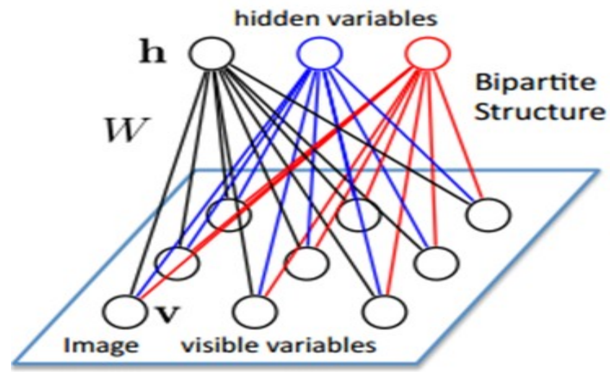


- Unconstrained handwriting recognition (Graves et al., 2009),
- Speech recognition (Graves et al., 2013; Graves and Jaitly, 2014)
- Handwriting generation (Graves, 2013),
- Machine translation (Sutskever et al., 2014)
- Image captioning (Kiros et al., 2014;Vinyals et al., 2014; Xu et al., 2015)
- Parsing (Vinyals et al., 2014a).

Deep Belief Networks

- Each link associates with a probability
- Parametric





The energy of the joint configuration:

$$E(\mathbf{v},\mathbf{h};\theta) = -\sum_{ij} W_{ij} v_i h_j - \sum_i b_i v_i - \sum_j a_j h_j$$

$\theta = \{W, a, b\}$ model parameters.

- Applied in clustering

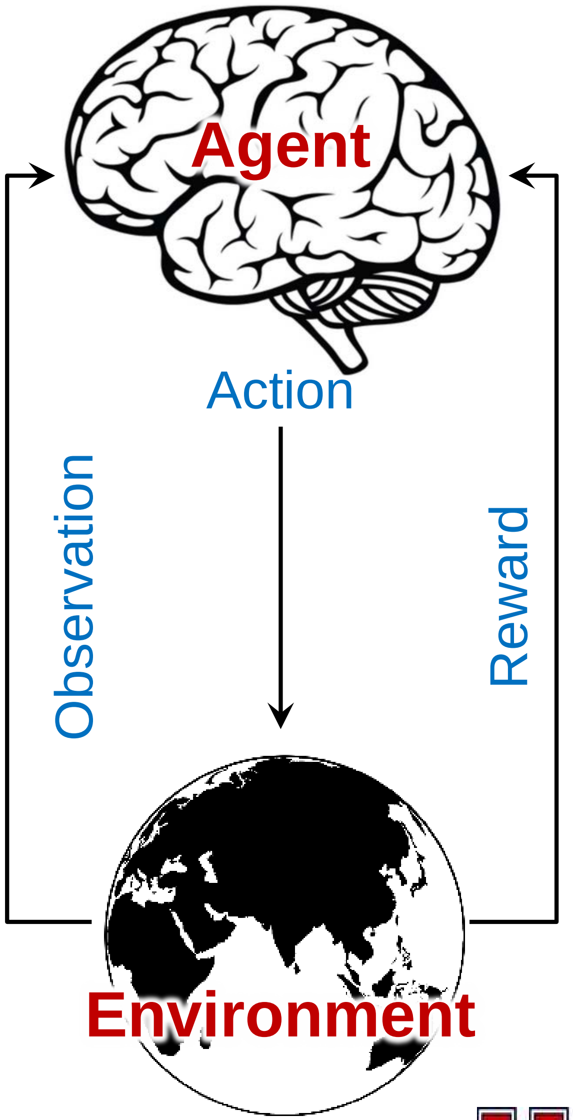
[5] G. E. Hinton, S. Osindero, and Y. W. Teh, "A Fast Learning Algorithm for Deep Belief Nets", *Neural Computation*. Vol. 18, No. 7, pp.1527–1554, 2006.

Comparison

<table>
<tr><th></th><th>Similarities</th><th>Differences</th></tr>
<tr><td>Convolutional Neural Networks</td><td rowspan="3">1. Multiple Layers
2. Use Back-propagation Algorithm for training
3. Can be combined together to create more powerful networks</td><td>1. More suitable for data with grid structures
2. Much fewer parameters
3. Very efficient training with GPUs</td></tr>
<tr><td>Recurrent Networks</td><td>1. Having memory of past (suitable for tasks like speech recognition)
2. Not able to take big input such as images or videos</td></tr>
<tr><td>Deep Belief Networks</td><td>1. Generative model (can generate realistic looking data after initializing at random variable)
2. Used much less due to inefficiency</td></tr>
</table>

Reinforcement Learning

- **Agent** —— An intelligent individual
- **Environment** —— Changes with the agent's action, then provides reward
- The agent observes the environment, takes action, and gets reward iteratively
- **Target of the agent** —— To maximize the total reward in the long run
- In case of full observation:
  - System states can be modelled as **Markov Decision Process (MDP)**



Markov Decision Process

- A Markov decision process further includes action in the Markov reward process, written as a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$
- $\mathcal{S}$ is a **finite set of the states** that have Markov property:

$$\mathbb{P}[S_{t+1} \mid S_t] = \mathbb{P}[S_{t+1} \mid S_1, ..., S_t]$$

  - Any finite set of discrete states $\xrightarrow[\text{transformed}]{\text{can be}}$ Markov states

- $\mathcal{A}$ is a **finite set of the actions**, from which the agent can choose to perform at each current state
- $\mathcal{P}$ is the **state transition probability matrix**, defined as the probability of state transition based on a given action:

$$\mathcal{P}_{ss'}^a = \mathbb{P}[S_{t+1} = s' \mid S_t = s, A_t = a]$$

Markov Decision Process

- A Markov decision process further includes action in the Markov reward process, written as a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$
- $\mathcal{R}$ is a **reward function**, showing the average reward of the next step when the current state is S_t, given as

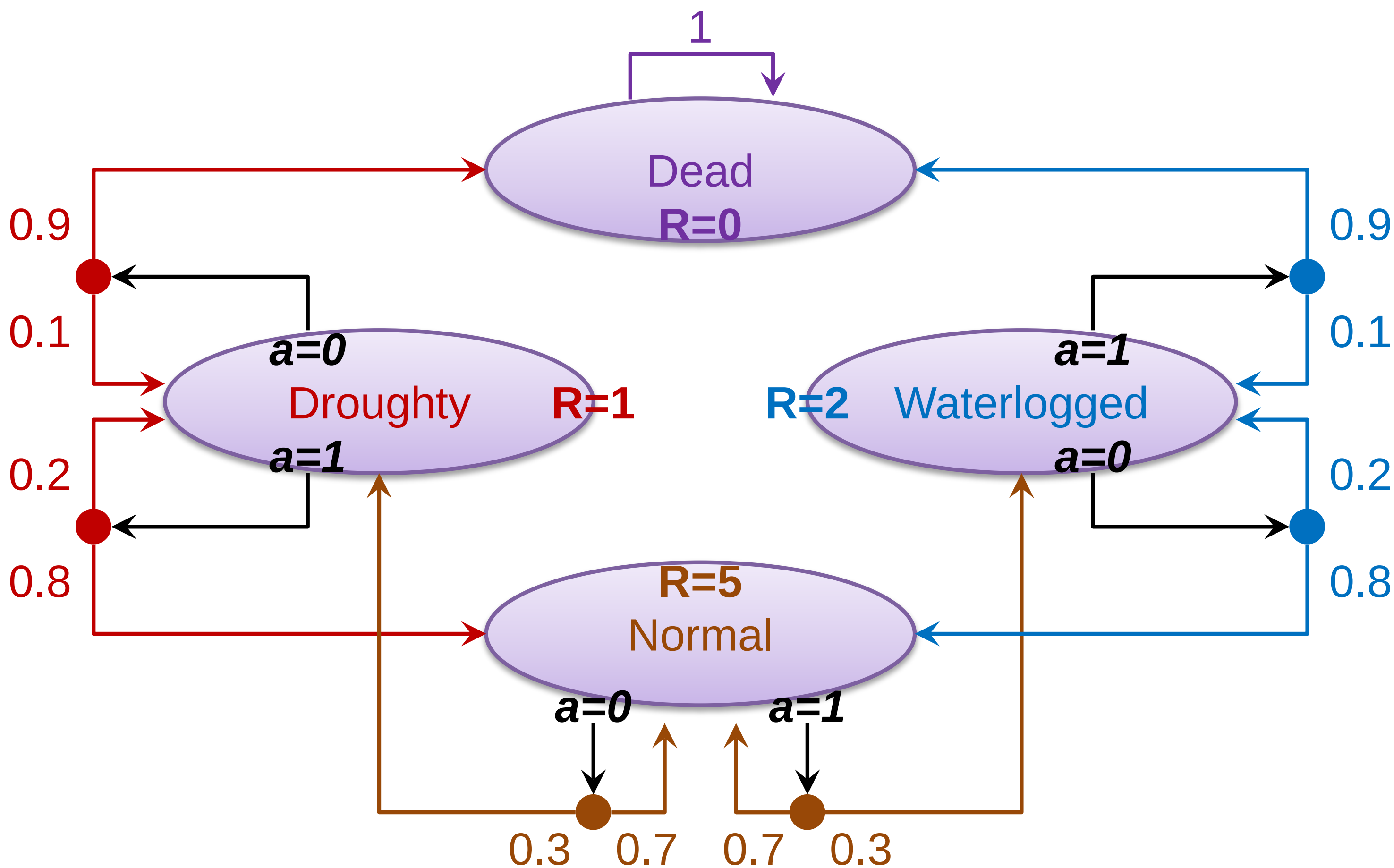
$$\mathcal{R}_s = \mathbb{E}\left[R_{t+1} \mid S_t = s\right]$$

- γ is a **discount factor**, which is used to weaken the reward of future, given by $\gamma \in [0, 1]$, avoiding infinite returns in cyclic state transitions
- Denote accumulative future reward as:

$$G_t = R_{t+1} + \gamma R_{t+2} + \ldots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

Markov Decision Process: Example

- Cultivating the flower by deciding whether to water it (***a=1*** for watering and ***a=0*** for non-watering)



Policy and Value Function

- **Policy** is the agent's behavior
- It is a ***map*** from ***state*** to ***action***
- Deterministic: $a = \pi(s)$ Stochastic: $\pi(a|s) = \mathbb{P}[A_t = a | S_t = s]$
- **Value function** is a prediction of future reward
- Used to evaluate the goodness/badness of states
- Defined as the average accumulative future reward from the current state based on the given policy π:
  - **State value function**: $v_\pi(s) = \mathbb{E}_\pi \left[G_t \mid S_t = s \right]$
  - **Action value function**: $q_\pi(s,a) = \mathbb{E}_\pi \left[G_t \mid S_t = s, A_t = a \right]$

Optimal Policy

- What is the optimal policy?
  - A policy that leads to the highest value for any state:

$$\pi \geq \pi' \text{ if } v_{\pi}(s) \geq v_{\pi'}(s), \forall s$$

- **Theorem**: For any Markov decision process, there exists an optimal policy π_* that is better than / equal to any other policy:

$$\pi_* \geq \pi, \forall \pi$$

- There is always a **deterministic optimal policy** for any MDP, given as

$$\pi_*(a|s) = \begin{cases} 1 & \text{if } a = \operatorname*{argmax}_{a \in \mathcal{A}} q_*(s, a) \\ 0 & otherwise \end{cases}$$

- Non-linear, no closed solution, many iterative methods

Q-Learning

Off-policy learning

- Experience from behavior $\mu(a|s)$
- E.g., learning from existing chess movement records

- A slightly different update function

$$Q(S,A) \leftarrow Q(S,A) + \alpha \left(R + \gamma \max_{a'} Q(S',a') - Q(S,A) \right)$$

- Use the best successor action instead of the action form the behavior to update the current Q(S,A)

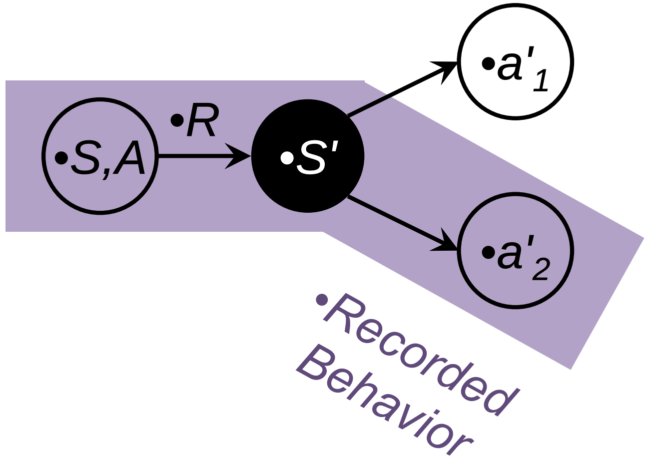
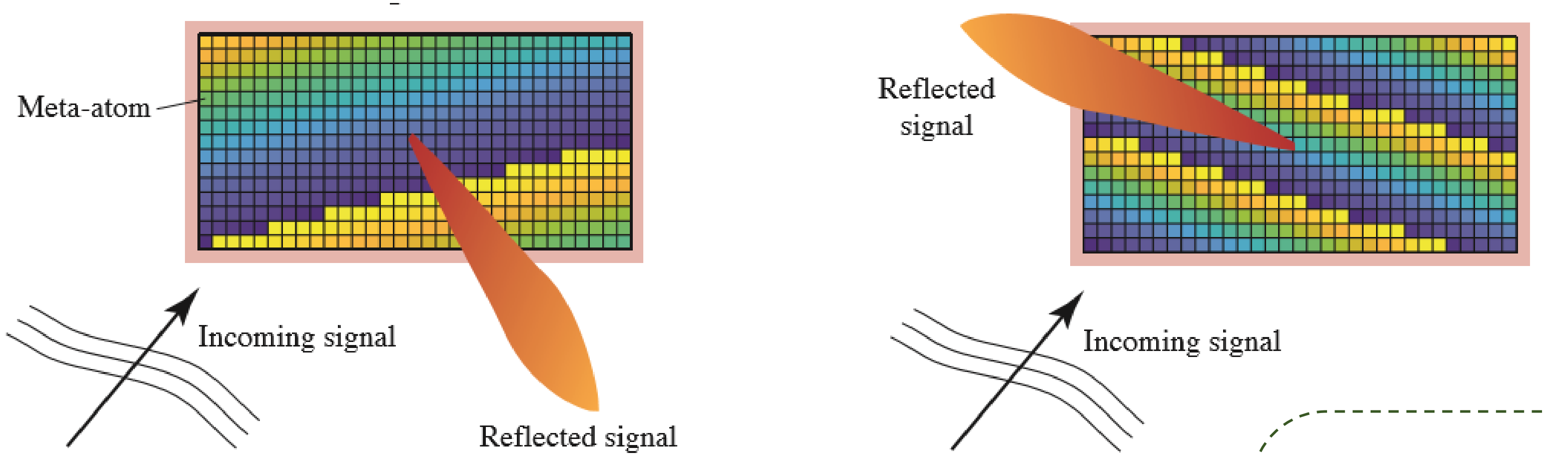


Table of Contents

RIS-aided Cellular Communications

Programmable propagation: inherent analog beamforming

- Different selections of the phase-shifts lead to beamforming from the RIS in different directions.

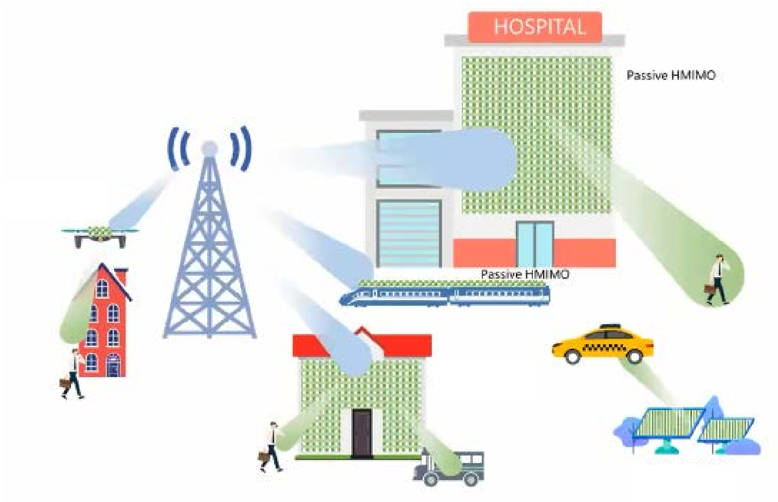


Propagation Environment = Channel Propagation h × Reflection Propagation → Frequency Response q × Reflection Co-efficient φ

Application Scenarios

Macro Cell Extension

- To extend the coverage
- To mitigate the inter-cell interference



Distributed Deployment

- Replace distributed antennas at lower cost and power (active)
- Distributed relays in ultra-dense networks (passive)
- Active RIS for energy harvesting

Figure source: China Mobile

Goals and Challenges

Goals

- Higher energy efficiency
- Higher capacity / Lower interference
- Larger coverage

Challenges

- How to design the number of phase shifts?
- How to deploy the RIS (orientation and location)?
- How does the size of RIS influence the performance?
- How to design the RIS configuration (phase shift)?
- How to coordinate multi-user access?

Case Study I: Limited Phase Shifts Effect

Reconfigurable Intelligent Surfaces assisted Communications with Limited Phase Shifts: How Many Phase Shifts Are Enough?

[6] H. Zhang, et al, "Reconfigurable Intelligent Surfaces assisted Communications with Limited Phase Shifts: How Many Phase Shifts Are Enough?" IEEE Transactions on Vehicle Technology, vol. 69, no. 4, pp. 4498-4502, Apr. 2020.

Motivations and Contributions

Problems

- Initial works focus on phase shifts optimization, the performance limits of RIS assisted communications
- Most works assume continuous phase shifts, which are hard to be implemented
- It is worthwhile to study the impact of the limited phase shifts on the achievable data rate

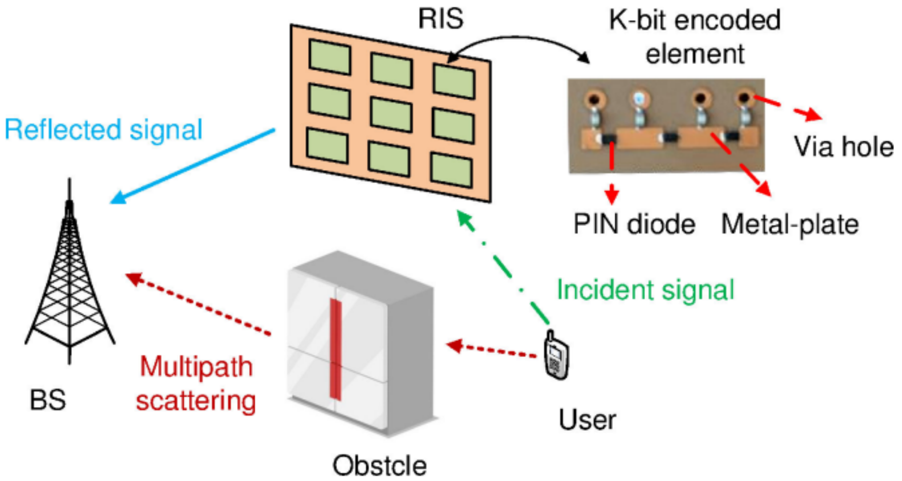
Contributions

- We provide an analysis on the achievable data rate with continuous phase shifts of the RIS, to evaluate the performance limits of the RIS assisted communications.
- We discuss how the limited phase shifts influence the data rate based on the derived achievable data rate.

System Model

System Description

- Uplink network
- No LoS between the BS and the user
- RIS to reflect the signal
  - Size: $M \times N$
  - *K* bits quantized



Reflection Response

$$\Gamma_{m,n} = \Gamma e^{-j\theta_{m,n}},$$

Phase shift for each element

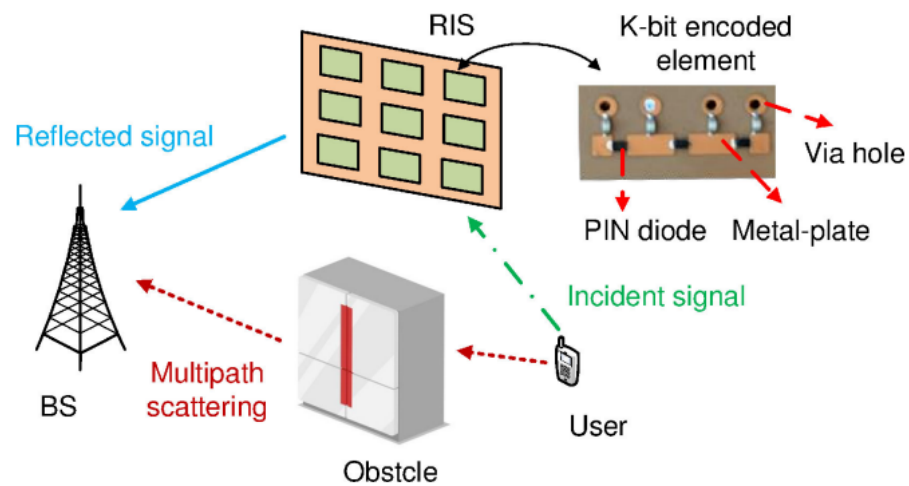
Constant reflection amplitude

System Model

Channel Model

- Rician model

$$\tilde{h}_{m,n} = \sqrt{\frac{\kappa}{\kappa+1}}\underbrace{h_{m,n}}_{\text{LoS}} + \sqrt{\frac{1}{\kappa+1}}\underbrace{\hat{h}_{m,n}}_{\text{NLoS}},$$



- LoS component

$$h_{m,n} = \sqrt{G}\underbrace{\left[\sqrt{D_{m,n}^{-\alpha}}e^{-j\frac{2\pi}{\lambda}D_{m,n}}\right]}_{\text{User-RIS}} \cdot \underbrace{\left[\sqrt{d_{m,n}^{-\alpha}}e^{-j\frac{2\pi}{\lambda}d_{m,n}}\right]}_{\text{RIS-BS}}$$

- NLoS component

$$\hat{h}_{m,n} = \underbrace{PL(D_{m,n})PL(d_{m,n})}_{\text{Pathloss}}\underbrace{g_{m,n}}_{\text{Fading}},$$

Achievable Data Rate Analysis

- Received Signals

Transmit power

$$y = \sum_{m,n} \Gamma_{m,n} \tilde{h}_{m,n} \underline{\sqrt{P}} s + w,$$

- Received Signal-to-Noise Ratio (SNR)

$$\gamma = \frac{P}{\sigma^2} \left(\sum_{m,n} \Gamma_{m,n} \tilde{h}_{m,n} \sum_{m',n'} \Gamma^*_{m',n'} \tilde{h}^*_{m,n} \right)$$

- The data rate can be expressed by

LoS path loss

Channel phase

NLoS path loss

$$\mathbb{E}\left[\log_2(1+\gamma)\right] \approx \log_2 \left(1 + \frac{\eta_{LoS}}{\kappa+1} MN + \frac{\kappa \eta_{NLoS}}{\kappa+1} \sum_{m,m',n,n'} e^{-j[\phi_{m,n} - \phi_{m',n'} + \theta_{m,n} - \theta_{m',n'}]} \right)$$

- The achievable data rate and the phase shift requirements

Arbitrary constant

$$R = \log_2 \left(1 + \frac{\eta_{NLoS}}{\kappa+1} \boxed{MN} + \frac{\kappa \eta_{LoS}}{\kappa+1} \boxed{M^2 N^2} \right) \quad \theta^*_{m,n} + \phi_{m,n} = \underline{C},$$

Rayleigh

Pure LoS

Analysis on Number of Phase Shifts

- Phase shifts errors

$$\delta_{m,n} = \underline{\theta^*_{m,n}} - \hat{\theta}_{m,n}.$$

Continuous Discrete

- With *K* bit quantized, the errors can be limited by

$$-\frac{2\pi}{2^{K+1}} \leq \delta_{m,n} < \frac{2\pi}{2^{K+1}}$$

- SNR expectation with limited phase shifts

$$\mathbb{E}[\hat{\gamma}] \geq \frac{\eta_{NLoS}}{\kappa+1} MN + \frac{\kappa\eta_{LoS}}{\kappa+1} M^2 N^2 \cos^2\left(\frac{2\pi}{2^{K+1}}\right)$$

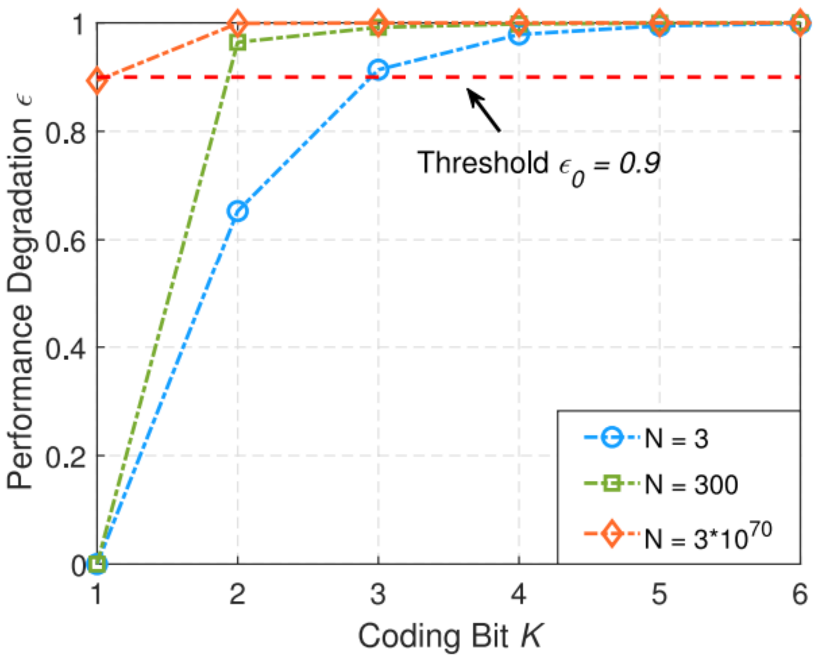
- To guarantee the performance

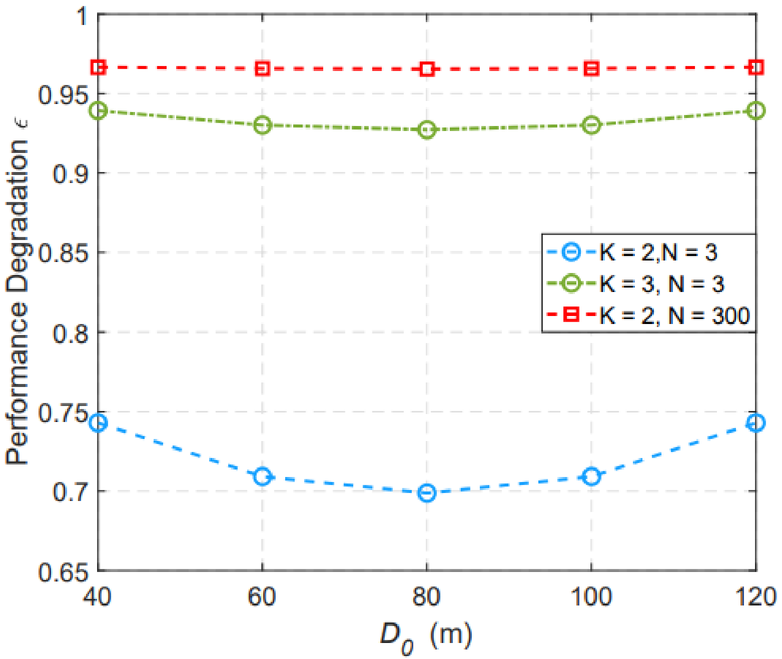
$$\epsilon = \log_2(1 + \mathbb{E}[\hat{\gamma}]) / \log_2(1 + \mathbb{E}[\gamma]) \geq \epsilon_0.$$ Threshold

- The required phase shifts

$$K_{req} = \log_2 \pi - \log_2 \arccos \sqrt{\frac{\kappa+1}{\kappa\eta_{LoS} M^2 N^2} \left(\left(1 + \frac{\eta_{NLoS}}{\kappa+1} MN + \frac{\kappa\eta_{LoS}}{\kappa+1} M^2 N^2 \right)^{\epsilon_0} - 1 - \frac{\eta_{NLoS}}{\kappa+1} MN \right)},$$

Simulation Results





- Required quantized bits decrease as the number of RIS elements grows, and 1 bit is enough when the RIS size goes to infinity
- We can easily observe that the data rate degradation will decrease first and then increase as the distance between the RIS and the BS increases given RIS size and quantized bit

Brief Summary

RIS assisted uplink cellular network

- Derive the achievable data rate
- Obtain the requirements on the quantized bits to ensure the data rate degradation is lower than a threshold

Remarks

- Pure LoS: asymptotic SNR of the squared number of RIS elements
- Rayleigh: asymptotic SNR of the number of RIS elements
- Required number of phase shifts decreases as the RIS size grows given data degradation threshold
- A number of phase shifts are necessary when RIS size is small, while 2 phase shifts are enough when the size goes infinity

Case Study II: RIS aided Coverage Extension

Reconfigurable Intelligent Surface (RIS) Assisted Wireless Coverage Extension: RIS Orientation and Location Optimization

[7] S. Zeng, et al, "Reconfigurable Intelligent Surface (RIS) Assisted Wireless Coverage Extension: RIS Orientation and Location Optimization," IEEE Communications Letter, under revision.

Motivations and Contributions

Problem:

- RIS is capable of shaping the propagation environment into desired form
- Existing works only utilized the RIS for coverage extension given the location of the RIS
- How to deploy the RIS to maximize the cell coverage has not been studied

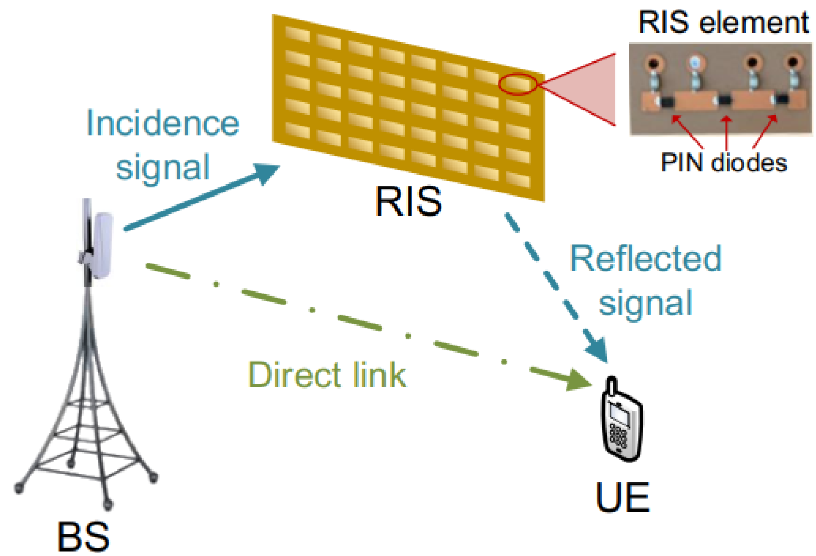
Contributions:

- We derive the cell coverage for an RIS assisted downlink cellular network with one BS and one UE.
- We maximize the cell coverage by optimizing the RIS orientation and the horizontal distance between the RIS and the BS.

System Model

Scenario Description

- Downlink scenario
- One BS, one UE, one RIS
- RIS:
  - Size: $M \times N$
  - Continuous phase shifts



Channel Model

- Direct link + Multiple RIS-based reflected channels
- Only consider pathloss to evaluate the average performance

Reflected links

$$h = \frac{1}{\sqrt{\xi}} \left(\sum_{m,n} \underline{\Gamma_{m,n}} \boxed{h_{m,n}} + \boxed{h_D} \right)$$

Reflected coefficients

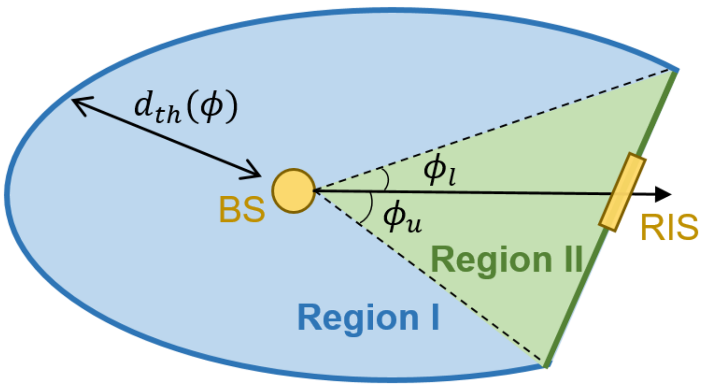
Direct links

Problem Formulation

Cell Coverage

- The cell coverage is defined as an area where the received SNR at the UE is larger than a certain threshold

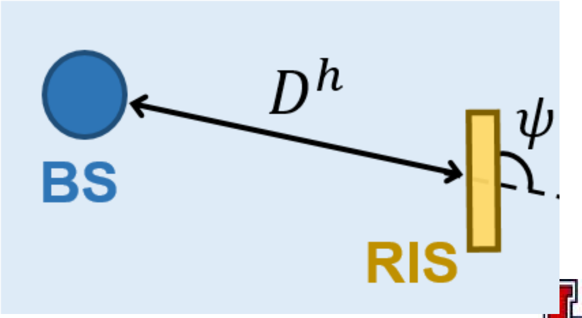
$$S = \underbrace{\int_{\phi_l}^{\phi_u} \frac{1}{2} d_{th}^2(\phi) d\phi}_{\text{Region I}} + \underbrace{\frac{1}{2}\sin(\phi_l - \phi_u) l(\phi_l) l(\phi_u)}_{\text{Region II}}.$$



- $d_{th}(\phi)$: coverage in a direction
- ϕ_l, ϕ_u: the boundary of the two regions

Optimization Problem

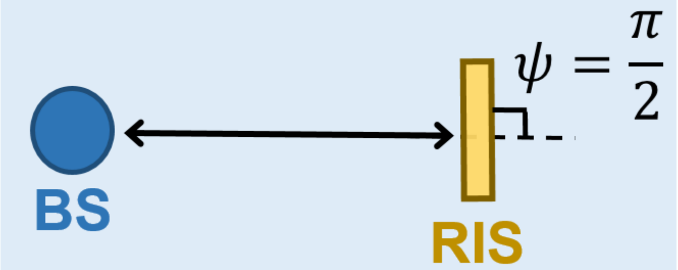
$$\max_{\underbrace{D^h}_{\text{Horizontal distance}},\ \underbrace{\psi}_{\text{RIS orientation}}} S$$



Solution

RIS Orientation Optimization

- Optimal RIS orientation: $\psi = \frac{\pi}{2}$
- For any horizontal distance

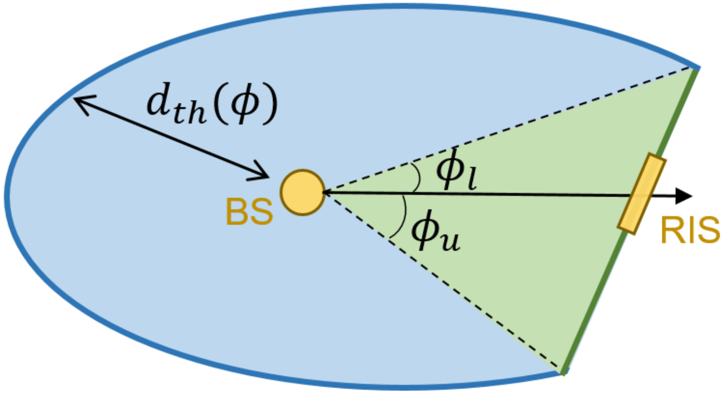


Horizontal Distance Optimization

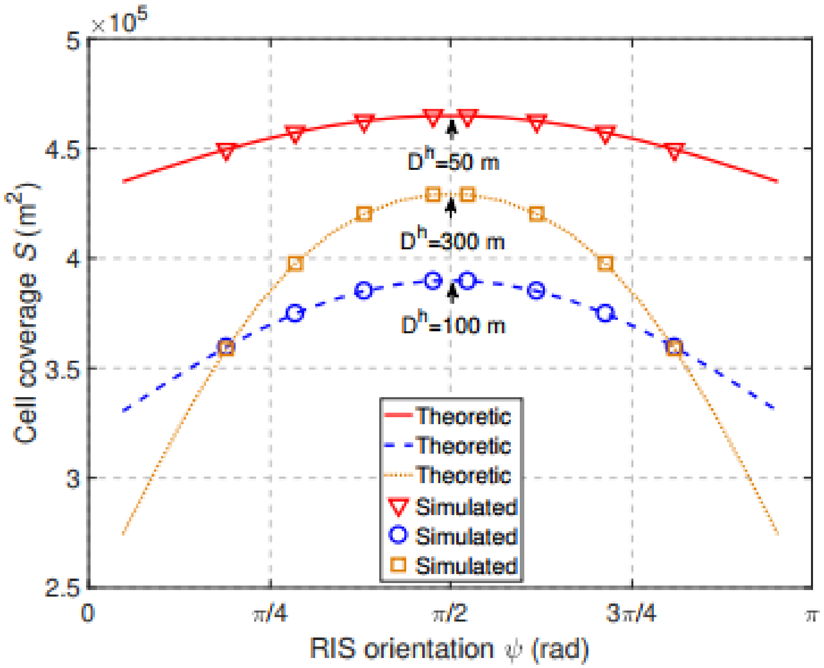
- $\phi_l, \phi_u, d_{th}(\phi)$: cannot be written in closed form
- Discretization

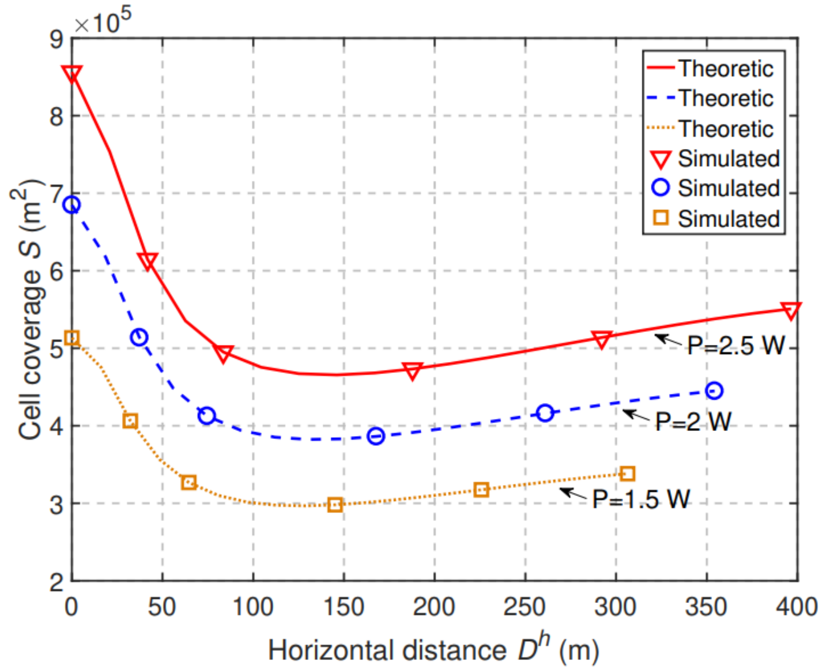
$$\int_{\phi_l}^{\phi_u} \frac{1}{2} d_{th}^2(\phi) d\phi \approx \sum_{i=0}^{K-1} \frac{1}{2} d_{th}^2(\phi_l + i\Delta)\Delta,$$

- Solved by interior point method



Simulation Results





- Theoretic and simulated results match well and it justifies that the optimal RIS orientation is $\frac{\pi}{2}$
- Cell coverage can be improved by placing the RIS close to the BS or the UE

Brief Summary

RIS assisted Downlink Cellular network

- Derive the cell coverage
- Formulate cell coverage extension problem
- Propose an algorithm to obtain the RIS placement solution

Remarks

- The RIS should be deployed vertical to the direction from the BS to the RIS.
- We should place the RIS close to either the BS or the UE.

Case Study III: RIS aided MIMO Communications

Hybrid Beamforming for Reconfigurable Intelligent Surface based Multi-user Communications: Achievable Rates with Limited Discrete Phase Shifts

[8] B. Di, et al, "Hybrid Beamforming for Reconfigurable Intelligent Surface based Multi-user Communications: Achievable Rates with Limited Discrete Phase Shifts," IEEE Journal of Selected Areas in Communications, to appear.

Motivation

Problems

- RIS configuration:
  - Multi-user case: Inter-user interference
  - Limited discrete electromagnetic responses
- How to design the size of RIS and perform beamforming?

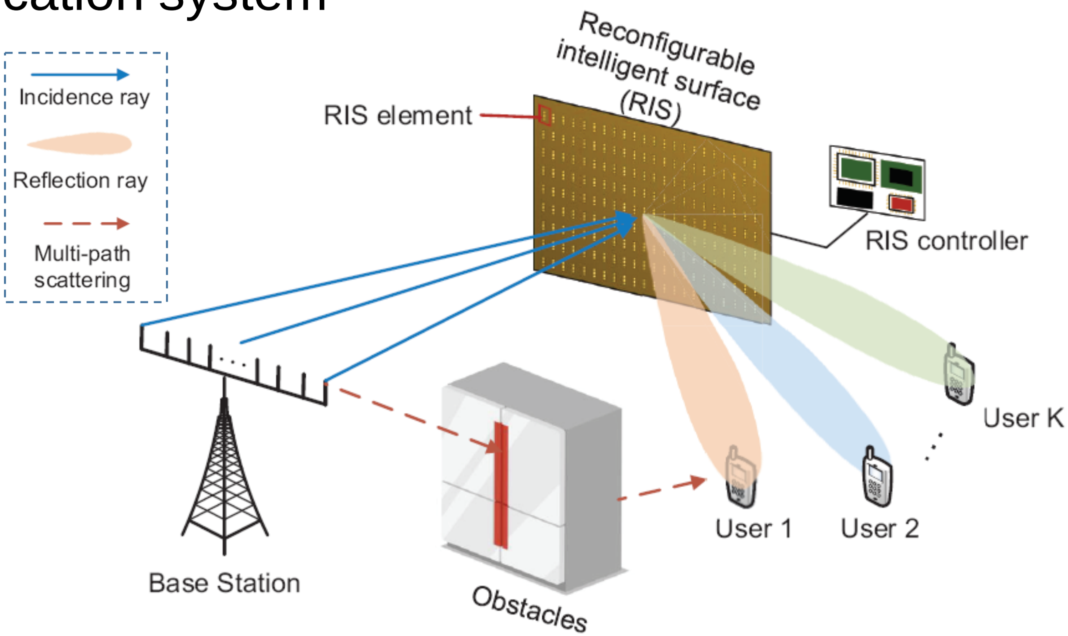
Challenges

- The channel propagation and the RIS configuration are coupled.
- Discrete electromagnetic responses renders the sum rate maximization to be a mixed integer programming problem.

System Model

System Model

- Downlink multi-user communication system
- N_t -antennas BS
- K single-antenna users
- $N_R \times N_R$ RIS elements
  - Phase shifts
  - Amplitudes
  - b-bit re-programmable



Channel Model

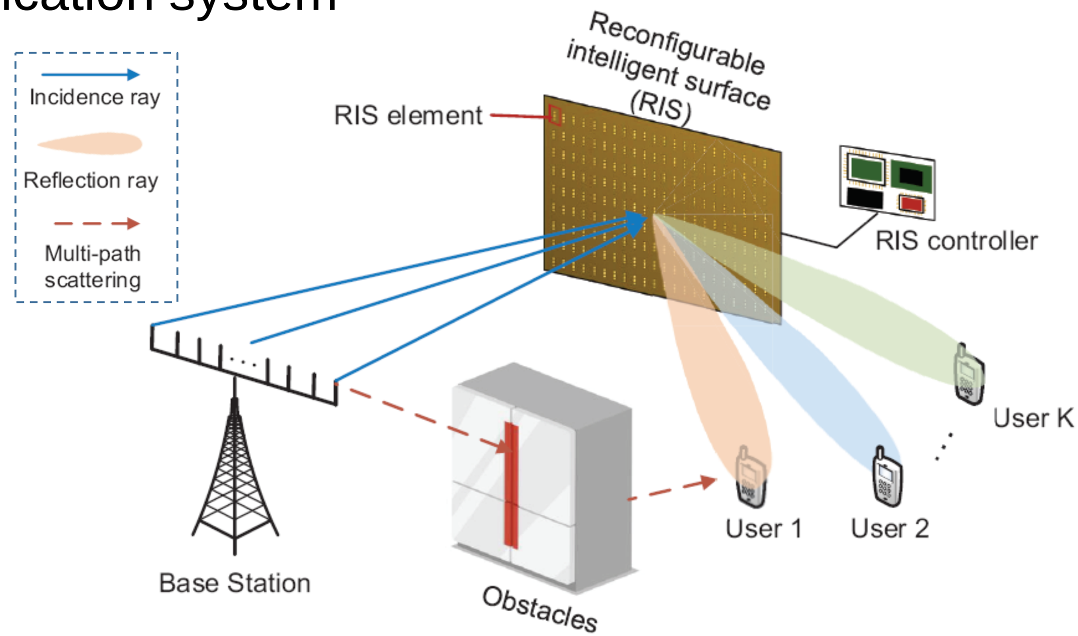
- Two-hop ray
  - The incident waves go through two paths sequentially.
- BS-user k: Rician model

Reflection Propagation = Frequency Response q × Reflection Co-efficient φ

System Model

System Model

- Downlink multi-user communication system
- N_t-antennas BS
- K single-antenna users
- $N_R \times N_R$ RIS elements
  - Phase shifts
  - Amplitudes
  - b-bit re-programmable

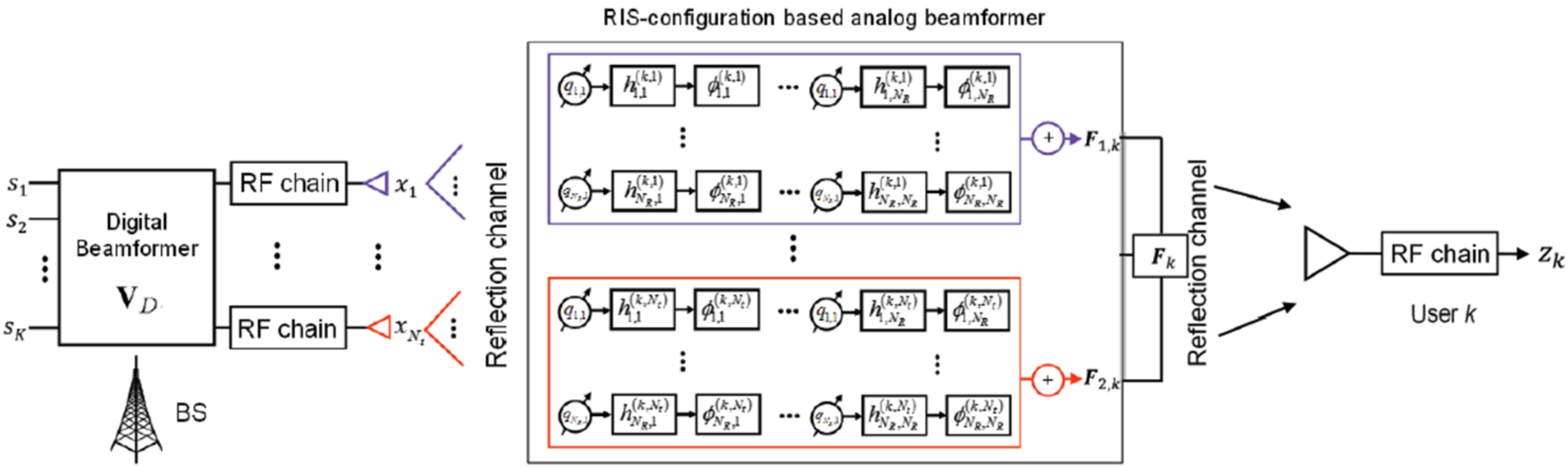


Channel Model

- Two-hop ray
  - The incident waves go through two paths sequentially.
- BS-user k: Rician model

		NLoS		LoS
Channel Propagation h	=	Multi-path Scattering	+	BS-RIS-user k

Hybrid Beamforming



RIS: Analog beamforming

- Different selections of the phase-shifts lead to beamforming from the RIS in different directions.

BS: Digital beamforming

- RIS elements have no digital processing capability.

Problem Formulation

Beamformers

- Digital beamformer $\boldsymbol{V}_D$

  $(\boldsymbol{V}_D)_k$ is a vector with respect to user *k*

- Coupled channel propagation and phase shifts

$$f_{k,n} = Tr\left(\phi^{(k)}\boldsymbol{Q}^T\underline{\boldsymbol{H}^{(k,n)}}\right)$$

Channel co-efficient

q_{l_1,l_2} is the phase shift of the (l_1, l_2)th RIS element (in matrix $\boldsymbol{Q}$)

Available Rate

$$R_k = \log_2\left(1 + \frac{\left|\mathbf{F}_k^H\mathbf{V}_{D,k}\right|^2}{\underline{\sum_{k'\neq k}\left|\mathbf{F}_k^H\mathbf{V}_{D,k'}\right|^2} + \sigma^2}\right)$$

Inter-user interference

Problem Formulation

Available Rate

$$R_k = \log_2\left(1 + \frac{\left|\mathbf{F}_k^H \mathbf{V}_{D,k}\right|^2}{\sum_{k' \neq k} \left|\mathbf{F}_k^H \mathbf{V}_{D,k'}\right|^2 + \sigma^2}\right)$$

Sum Rate Maximization Problem

$$\begin{aligned} \underset{\mathbf{V}_D, \{q_{l_1,l_2}\}}{\text{maximize}} \quad & \sum_{1 \leq k \leq K} R_k \\ \text{subject to} \quad & \operatorname{Tr}\left(\mathbf{V}_D^H \mathbf{V}_D\right) \leq P_T, \end{aligned}$$

Phase shift $q_{l_1,l_2} = \frac{j + e^{j\theta_{l_1,l_2}}}{2}$,

Discrete $\theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}}$

decouple

Digital Beamforming

$$\begin{aligned} \underset{\mathbf{V}_D}{\text{maximize}} \quad & \sum_{1 \leq k \leq K} R_k, \\ \text{subject to} \quad & \operatorname{Tr}\left(\mathbf{V}_D^H \mathbf{V}_D\right) \leq P_T. \end{aligned}$$

$$\begin{aligned} \underset{\{\theta_{l_1,l_2}\}}{\text{maximize}} \quad & \sum_{1 \leq k \leq K} R_k, \\ \text{subject to} \quad & \theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}} \end{aligned}$$

Analog Beamforming

Algorithm Design

Digital Beamforming Algorithm

- Digital beamformer: ZF beamforming based

$$\boldsymbol{V}_D = \boldsymbol{F}^H(\boldsymbol{F}\boldsymbol{F}^H)^{-1}\boldsymbol{P}^{\frac{1}{2}}$$

$$\left|\boldsymbol{F}_k^H(\boldsymbol{V}_D)_k\right| = \sqrt{p_k}$$

$$\left|\boldsymbol{F}_k^H(\boldsymbol{V}_D)_{k'}\right| = 0, \forall k' \neq k$$

$$\max_{p_k \geq 0} \sum_{1 \leq k \leq K} R_k = log_2\left(1 + \frac{p_k}{\sigma^2}\right)$$

Water-filling

$\boldsymbol{P}$

$\boldsymbol{V}_D$

Algorithm Design

Analog Beamforming Algorithm

1. To maximize the sum rate, the power should be maximized

   Analog beamforming is only related to the following constraint

$$\mathrm{Tr}\left(\boldsymbol{P}^{\frac{1}{2}}\tilde{\boldsymbol{V}}_D^H\tilde{\boldsymbol{V}}_D\boldsymbol{P}^{\frac{1}{2}}\right) \le P_T$$

$$\boldsymbol{V}_D = \boldsymbol{F}^H(\boldsymbol{F}\boldsymbol{F}^H)^{-1}\boldsymbol{P}^{\frac{1}{2}} = \tilde{\boldsymbol{V}}_D\boldsymbol{P}^{\frac{1}{2}}$$

$$\min_{\{q\}} \mathrm{Tr}(\underline{\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{F}\cdot(\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{F})^H})^{-1}$$

$$\text{s.t. } q \in \mathit{Limited, discreted\ set}$$

symmetric, positive
semi-definite matrix

2. A semi-define programming (SDP) problem:

   Schur complement

3. A mix-integer SDP with linear constraints

Branch-and-bound

Optimal solution

Check the feasibility

NO

Gradient cuts

YES

End

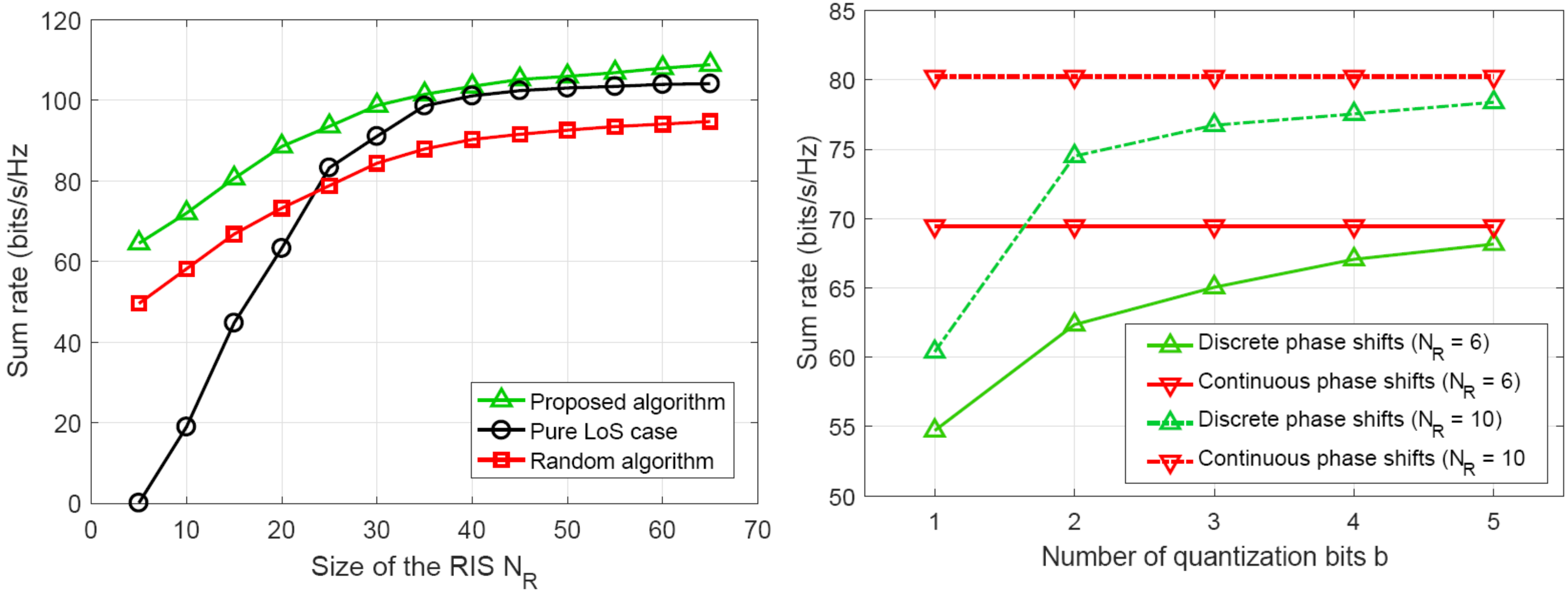
Theoretical Analysis

Required Number of Transmit Antennas

For the RIS-based hybrid beamforming scheme with a large enough number of RIS elements, i.e., $N_R^2 \geq KN_t$, to achieve any fully digital beamforming scheme, the number of transmit antennas at the BS should not be smaller than the number of single-antenna users, i.e., $N_t \geq K$.

- The size of the RIS should be larger than the product of the number of users and the size of the antenna array at the BS
- The number of minimum RF chains required by the RIS-based HBF is reduced by half compared to the traditional scheme

Simulation Results



- The sum rate grows rapidly with a small size of RIS and gradually flattens as the size of RIS continues to increase.
- The performance of our proposed algorithm is much better than that of the pure LoS case when the size of RIS is small.
- As the number of quantization bits increases, the sum rate obtained by our proposed algorithm approaches that in the continuous case.

Brief Summary

Hybrid Beamforming

- BS: digital beamforming
- RIS: analog beamforming
- Sum rate maximization

Three Remarks

- Sum rate increases rapidly with the number of quantization bits when the number of bits is small, and gradually approaches to the continuous case
- Sum rate increases with the size of RIS and converges to a stable value
- Required antennas at the BS is only half of that in traditional hybrid beamforming schemes

Case Study IV: RIS aided D2D Communications

Reconfigurable Intelligent Surface Assisted Device-to-Device Communications

[9] Yali Chen et al, “Reconfigurable Intelligent Surface Assisted Device-to-Device Communications”, IEEE Transactions on Wireless Communications, under revision. Arxiv: https://arxiv.org/abs/2007.00859.

Motivation

Problems

- Device-to-Device (D2D) communications: users in proximity can communicate directly
- Interference due to spectrum sharing

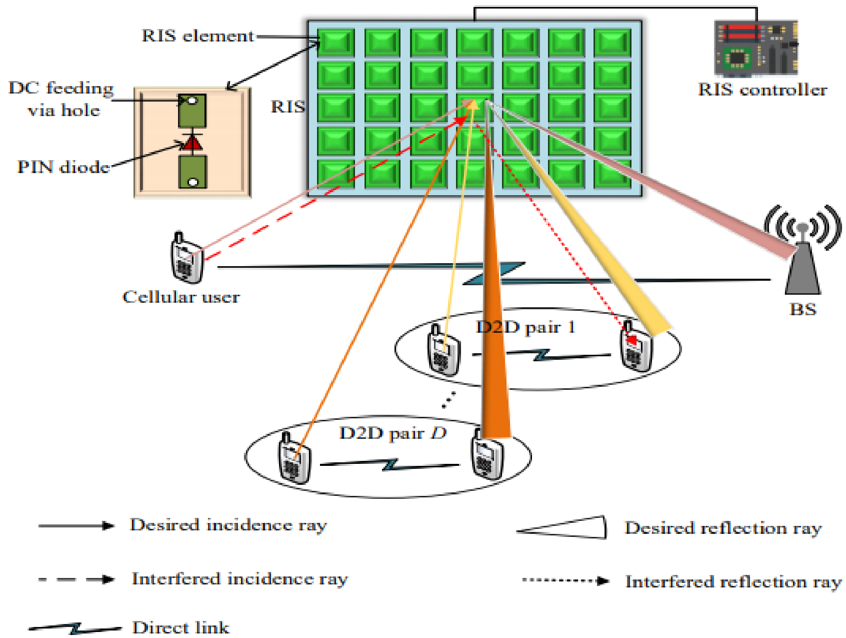
Contributions

- Propose to utilize RIS to alleviate the interference in D2D communications due to its capability of beamforming
- Formulate a system sum-rate maximization problem by optimizing the transmit power of the transmitters and the phase shifts of the RIS
- Design an iterative algorithm to solve this problem

System Model

Scenario Description

- Uplink RIS aided heterogeneous network
- One cellular user
- Multiple D2D users
- Share the same spectrum



Interference Analysis

- D2D receiver: interference from the cellular and other D2D links
- BS: interference from D2D links
- Data rate:

$$R_{r_i} = \log_2\left(1 + \frac{|\overset{\text{LoS}}{\boxed{\mathbf{H}^L_{r_i,t_i}}} + \overset{\text{Reflected}}{\boxed{\mathbf{F}_{r_i,t_i}}}|^2 p_i}{\underbrace{\sum_{j\in\mathbf{L}, j\neq i} |\mathbf{H}^L_{r_i,t_j} + \mathbf{F}_{r_i,t_j}|^2 p_j}_{\text{Interference}} + \sigma^2}\right)$$

Problem Formulation

Sum Rate Maximization Problem

$$\max_{\mathbf{P},\boldsymbol{\Theta}} \sum_{i=1}^{D+1} R_{r_i}$$

$$s.t.$$

$$(a)\ \Gamma_{r_i} \geq \gamma_{min},\ \forall i = 1,2,...,D+1,$$

$$(b)\ 0 \leq p_i \leq P_{max},\ \forall i = 1,2,...,D+1,$$

$$(c)\ q_{l_z,l_y} = e^{j\theta_{l_z,l_y}},\ \theta_{l_z,l_y} = \frac{2m_{l_z,l_y}\pi}{2^e-1},$$

$$m_{l_z,l_y} = \{0,1,...,2^e-1\}, 1 \leq l_z, l_y \leq N,$$

Sum rate

QoS constraint

Power constraint

Phase shift constraint

problem ***decomposition***

Maximize data rate
s.t. QoS constraint
Power constraint

power allocation

Maximize data rate
s.t. QoS constraint
Phase shift constraint

phase shift optimization

Algorithm Design

Power Allocation Subproblem

- Due to the interference, the data rate is not convex
- However, it can be written by the difference two concave functions.

Reformulation of data rate

$$\sum_{i=1}^{D+1} f_i(\mathbf{P}) \triangleq g_i(\mathbf{P}) - \varphi_i(\mathbf{P})$$

First order Tylor expansion

$$g_i(\mathbf{P}) = g_i(\mathbf{P}^{(n)}) + \sum_{k=1}^{D+1} \frac{\partial g_i(\mathbf{P})}{\partial p_k}\bigg|_{\mathbf{P}=\mathbf{P}^{(n)}} (p_k - p_k^{(n)}).$$

Convex Optimization

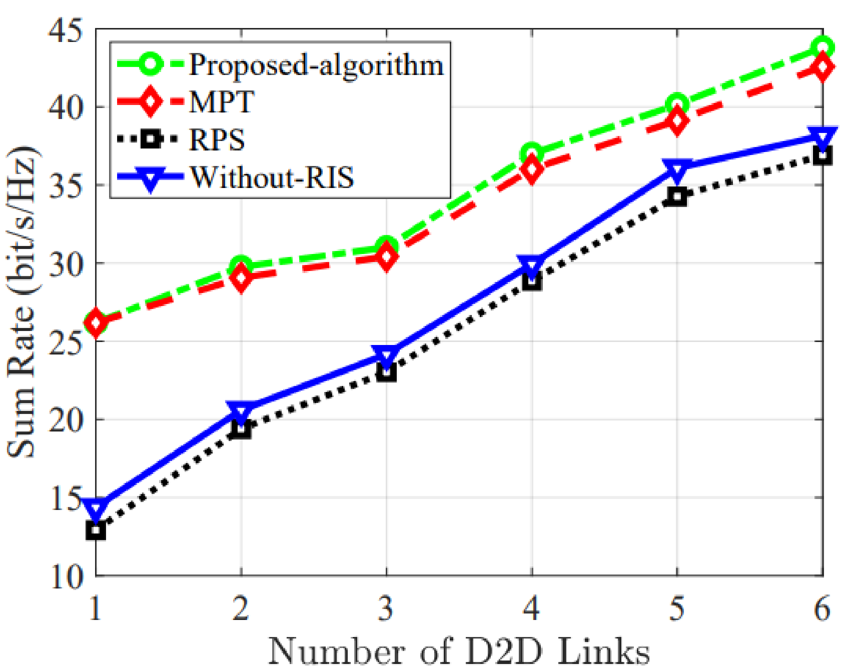
Power

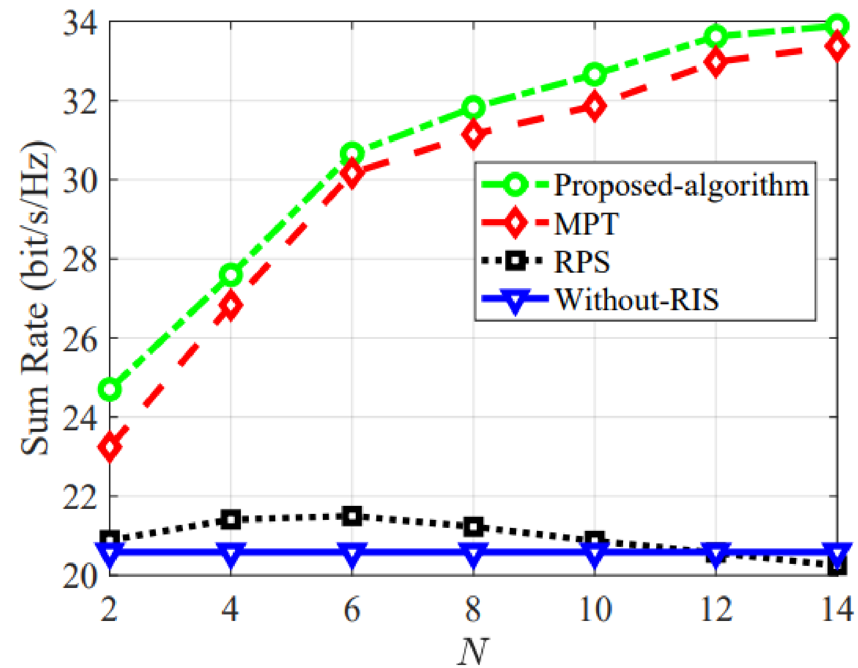
$\mathbf{P}$

Phase Shift Optimization Subproblem

- Discrete phase shifts: fixed others and select the optimal one

Simulation Results





- The utilization of RIS can effectively increase the system sum rate
- Proposed phase shift optimization algorithm can achieve a better sum rate than that obtained by the random one
- The system sum rate will increase as the size of the RIS grows

Brief Summary

Uplink D2D Communications

- Propose to utilize RIS to alleviate the interference
- Formulate the sum rate maximization problem
- Design a joint power allocation and phase shift optimization algorithm

Remarks

- Utilization of RIS can improve the system sum rate
- Proper phase shifts design will be important in the RIS assisted D2D communications
- Sum rate will increase as the RIS size grows

Case Study V: RIS assisted MAC Access

Reconfigurable Intelligent Surface Assisted MAC for 6G: Protocol Design, Analysis, and Optimization

[10] Xuelin Cao et al, "Reconfigurable Intelligent Surface Assisted MAC for 6G: Protocol Design, Analysis, and Optimization", IEEE Internet-of-Things Journal, under revision.

Motivation

Problems

- Most of existing papers focus on the physical layer, the multi-user communications on the medium access control (MAC) layer have not been effectively addressed yet
- How to connect RIS with the traditional communications on MAC layer?

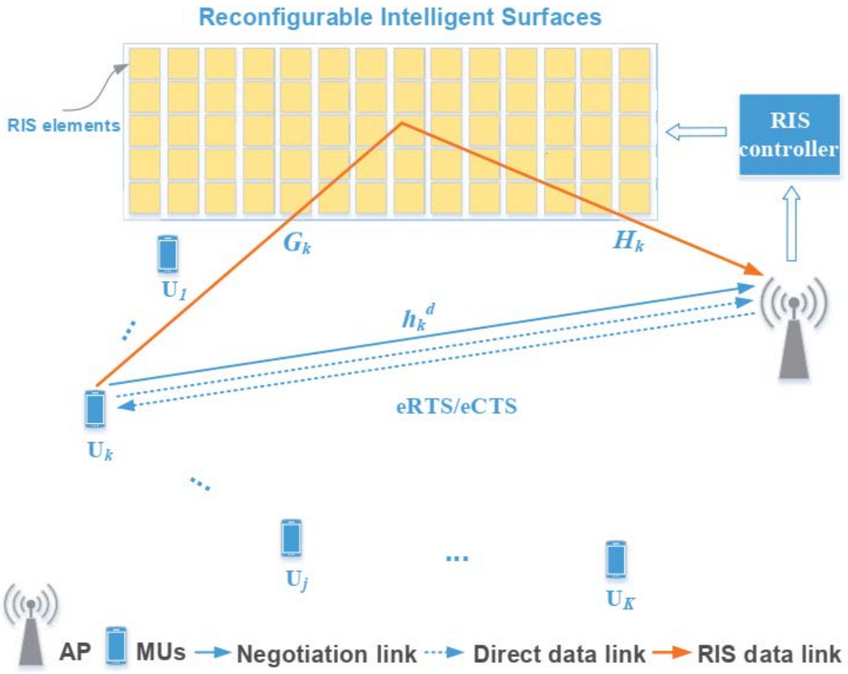
Contributions

- Propose an RIS-assisted MAC protocol with reservation to avoid the collision at the RIS side
- Solve a system sum-rate maximization problem by optimizing the transmit power, transmission bandwidth and reflection coefficients at the RIS

System Model

Scenario Description

- Multi-user uplink system
- K users communicates with an AP by an RIS over one channel
- AP and users: a single antenna
- RIS controller can communicate with AP to adjust the coefficients



Received Signals

- The received signal at AP from user *k* by the direct data link and RIS data link, y_k, can be denoted as

$$y_k = \underbrace{h_k^d s_k}_{\text{Direct data link}} + \underbrace{H_k \Theta_k G_k s_k}_{\text{RIS data link}} + w_k$$

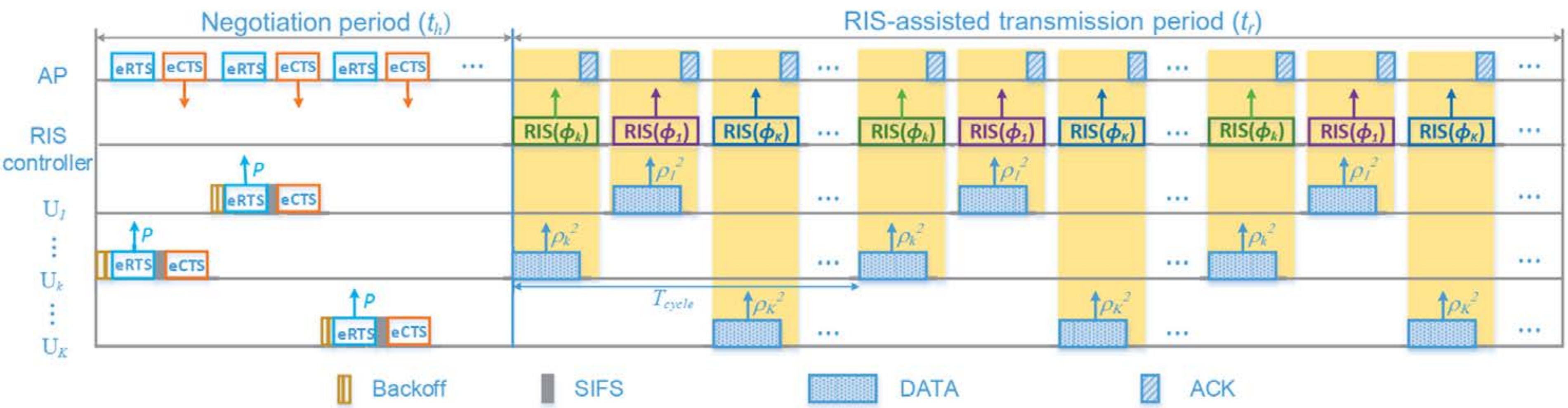
RIS-assisted MAC Protocol

Phase 1: Negotiation Phase

- Reservation between users and AP can be achieved by sending eRTS/eCTS control packets, and the RIS controller gets these from AP.

Phase 2: RIS-Assisted Transmission Phase

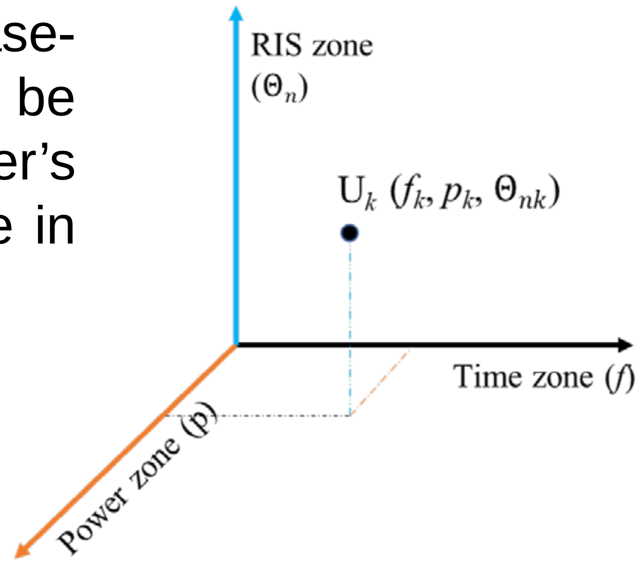
- In the reserved slots, the reflection coefficients updating of RIS is operated by the RIS controller to support the user's transmission



RIS-assisted MAC Protocol

Reserved Resources

- Dimension 1 (D1): time zone (f). The slots that are in the RIS-assisted transmission phase can be reserved, which is decided by the transmission frequency (f) and the initial transmission time;
- Dimension 2 (D2): power zone (p). The transmit power can be used at the requested user while starting the RIS-assisted data transmission;
- Dimension 3 (D3): RIS zone (Θ_n). The phase-shift of each RIS element, which can be controlled by the AP to align with the user's direct data transmission at the reserved time in the RIS-assisted transmission phase



Problem Formulation

Sum Rate Maximization Problem

- Objective: sum rate

$$\Im_{Total}^{SCMU} = \frac{t_p t_h \zeta_s}{T t_s} B \sum_{k=1}^{K} f_k \log_2\left(1 + \mathrm{SNR}_k^{SCMU}\right)$$

$$\begin{aligned}
\mathbf{P1}: \quad & \max_{\rho_1^2,\cdots,\rho_K^2;\Theta_1,\cdots,\Theta_K;f_1\cdots,f_K} \quad \Im_{Total}^{SCMU}, \\
\text{s.t.} \quad & \text{C1:} \quad 0 \le \rho_k^2 \le P - P_{RIS}, \quad \forall k, \\
& \text{C2:} \quad |\phi_k^n| = 1, \quad \forall k, n, \\
& \text{C3:} \quad \theta_k^n \in \Omega, \quad \forall k, n, \\
& \text{C4:} \quad t_h + t_r = T, \quad \forall k \\
& \text{C5:} \quad \frac{t_r}{t_h} \le \frac{t_p}{t_s} f_{max} \zeta_s, \quad \forall k \\
& \text{C6:} \quad 1 \le f_k \le f_{max}, \quad \forall k,
\end{aligned}$$

- C1 → Transmit power constraint
- C2, C3 → RIS coefficients constraints
- C4, C5 → Negotiation and transmission periods constraints
- C6 → Bandwidth constraint

Algorithm Design

Problem Decomposition

- Fix bandwidth $f_1, \cdots, f_K$

$$\mathbf{P2.1}: \max_{\rho_k^2;\Theta_k} \mathrm{SNR}_k^{SCMU} \quad \text{s.t.} \quad \text{C1 - C3.}$$

$$\mathbf{P2.1.1}: \max_{\rho_k^2;\theta_k^1,\cdots,\theta_k^N} \left|\left(h_k + H_k\Theta_k G_k\right)\rho_k\right|^2 \quad \text{s.t.} \quad \text{C1 - C3.}$$

  - Closed-form suboptimal solution

$$\theta_k^{n^*} = \arg(h_k\rho_k^*) - \arg(h_{k,RIS}^n) - \arg(g_{k,RIS}^n, \rho_k^*),$$

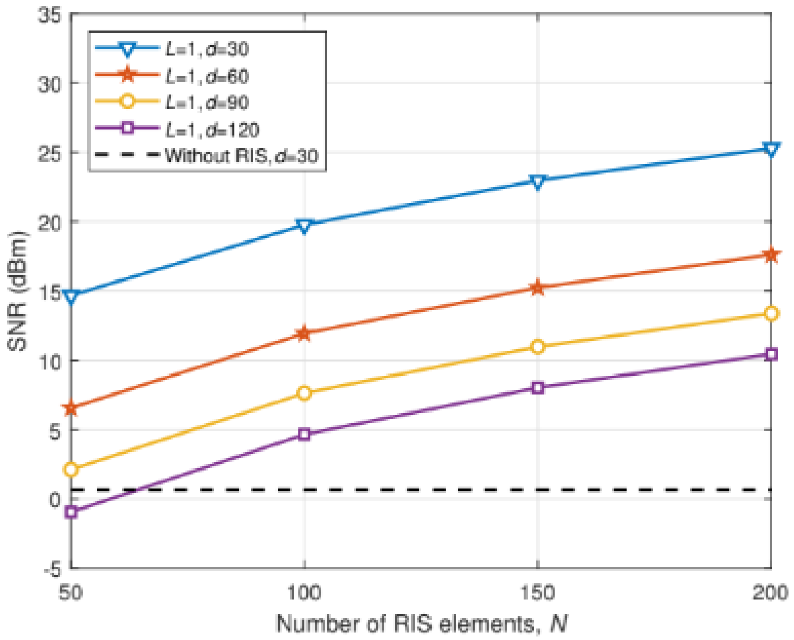
$$\rho_k^* = \sqrt{P - P_{RIS}} \frac{h_k + H_k\Theta_k^* G_k}{\|h_k + H_k\Theta_k^* G_k\|},$$

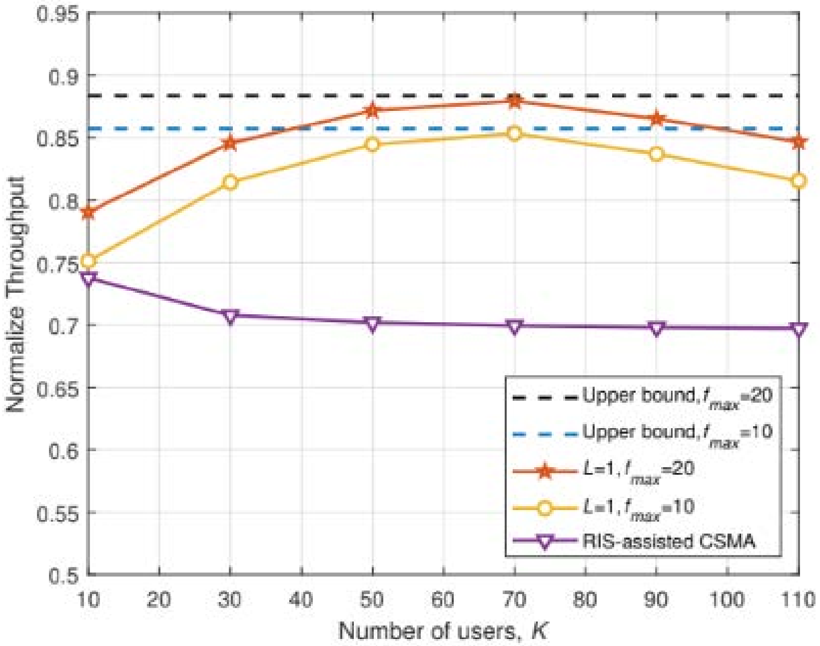
- Fixed the optimal RIS configuration $\Theta_1, \cdots, \Theta_K$ and power $\rho_1^2, \cdots, \rho_1^K$

$$\mathbf{P2.2}: \max_{f_1,\cdots,f_K} \sum_{k=1}^{K} f_k \quad \text{s.t.} \quad \text{C4 - C6.}$$

Linear program

Simulation Results





- The SNR increases as the number of RIS elements grows
- The normalized throughput increases first and then decreases as the number of users increases

Brief Summary

RIS-assisted MAC Access

- Consider multi-user single-channel uplink scenario
- Propose an RIS-assisted MAC protocol
- Formulate the sum rate maximization problem
- Design a joint bandwidth allocation, power allocation, and phase shift optimization algorithm

Remarks

- Throughput will increases as the number of RIS elements grows
- There exists an optimal number of accessed users
- The same method can be extended to multi-channel case

Potential Directions

RIS-based Multi-cell / Multi-user Coordination

- RIS-based cognitive radio
- RIS-based NOMA
- RIS-based cell-free MIMO

Channel Estimation and Modelling

- Semi-passive RIS channel estimation
- Passive RIS channel estimation

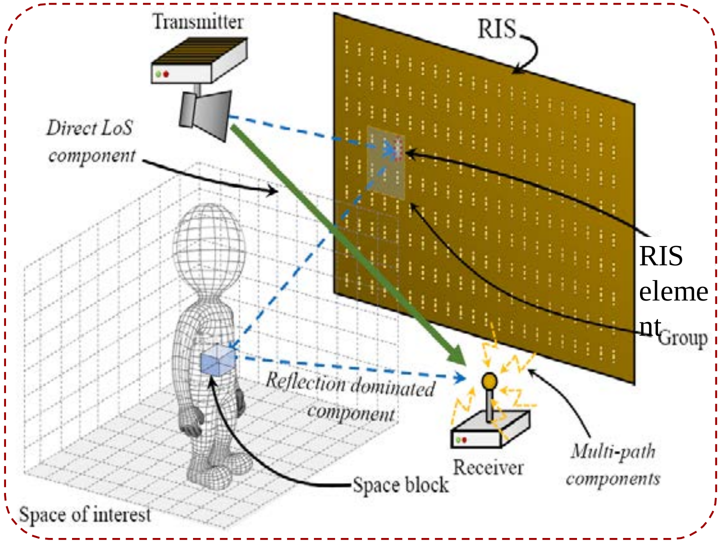
Other issues

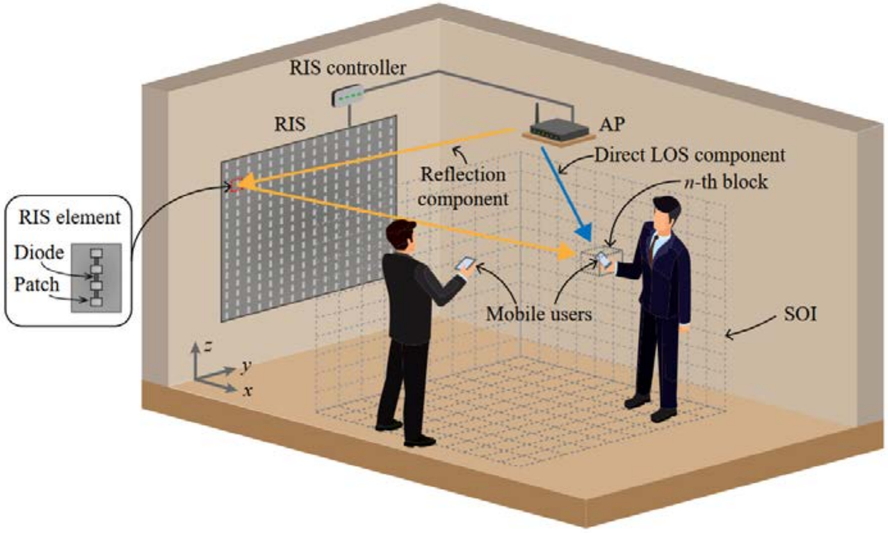
- Joint coding and transmission
- Physical-layer security
- Energy efficiency

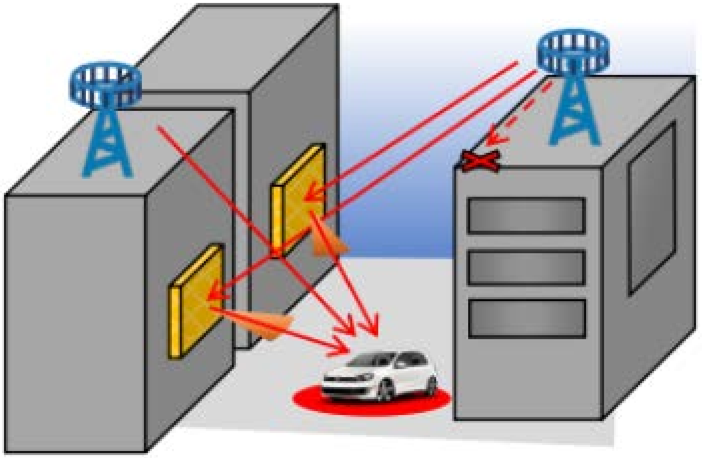
Table of Contents

RIS aided Internet of Things

- Due to the capability of beamforming, RIS is able to improve spatial resolution
- RIS can be utilized to enhance the sensing and localization accuracy







Localization under LOS Blockage

RIS-aided Sensing

RF Sensing

- Use RF signals to recognize objects or monitoring parameters
- Benefits: contactless, leveraging existing communication infrastructure

Objective

- Minimize sensing errors

Challenges

- Recognize object/monitor parameter through the received RF signal
- Customize propagation environment
- Multiple object recognition

RIS-aided Localization

Criteria

- Measurement: time of arrival (TOA); phase of arrival (POA); received signal strength (RSS)
- Reference system: geographical or relative

Objective

- Maximize localization accuracy

Challenges

- Enhance the differences for two adjacent positions
  - Proper phase shifts design for RIS
- Real-time implementation
  - Low-complexity recover algorithm for observation

Case Study VI: RIS aided RF Sensing

Reconfigurable Intelligent Surfaces based RF Sensing: Design, Optimization, and Implementation

[11] J. Hu, et al, "Reconfigurable Intelligent Surfaces based RF Sensing: Design, Optimization, and Implementation," IEEE Journal of Selected Areas in Communications, to appear.

Motivation

RF sensing: Receivers recognize the influence of the sensing targets on the wireless signal propagation

- No contact with sensing targets
- Limited by complicated wireless environments

RIS-based **RF sensing system**

- Human posture recognition:
    - Recognize different human postures automatically

Challenges

- RIS configuration design
    - Large number of RIS elements
    - Different states in each element
- RIS configuration and decision function are coupled.

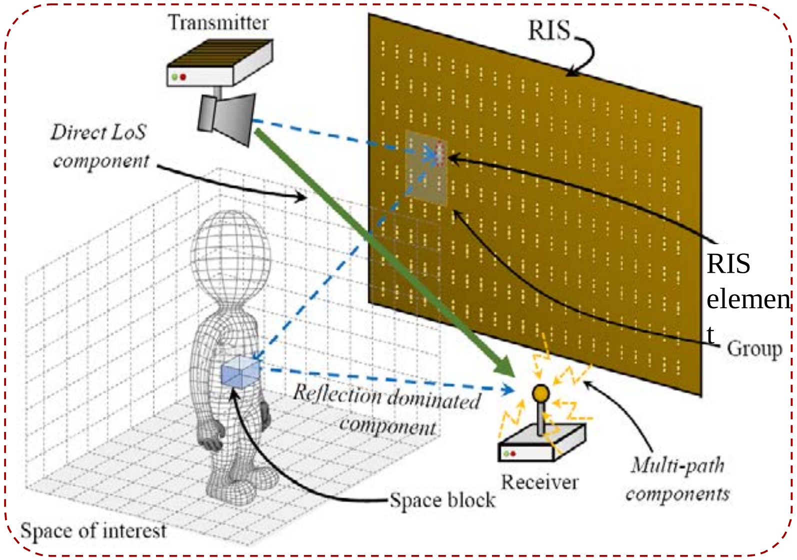
Model Description

System Structure

- Transmitter: A directional antenna which is pointed towards the RIS
- Receiver: An omni-directional vertical antenna below the RIS
- Human： Space reflection vector carries the information of postures.
- RIS: RIS elements in the same group are in the same state.

Channel Model

- Multi-path component:
  - Environment scattering
- LoS component:
  - Transmitter → Receiver
- Reflection dominated components
  - Transmitter → RIS → Human → Receiver



Periodic Configuring Protocol

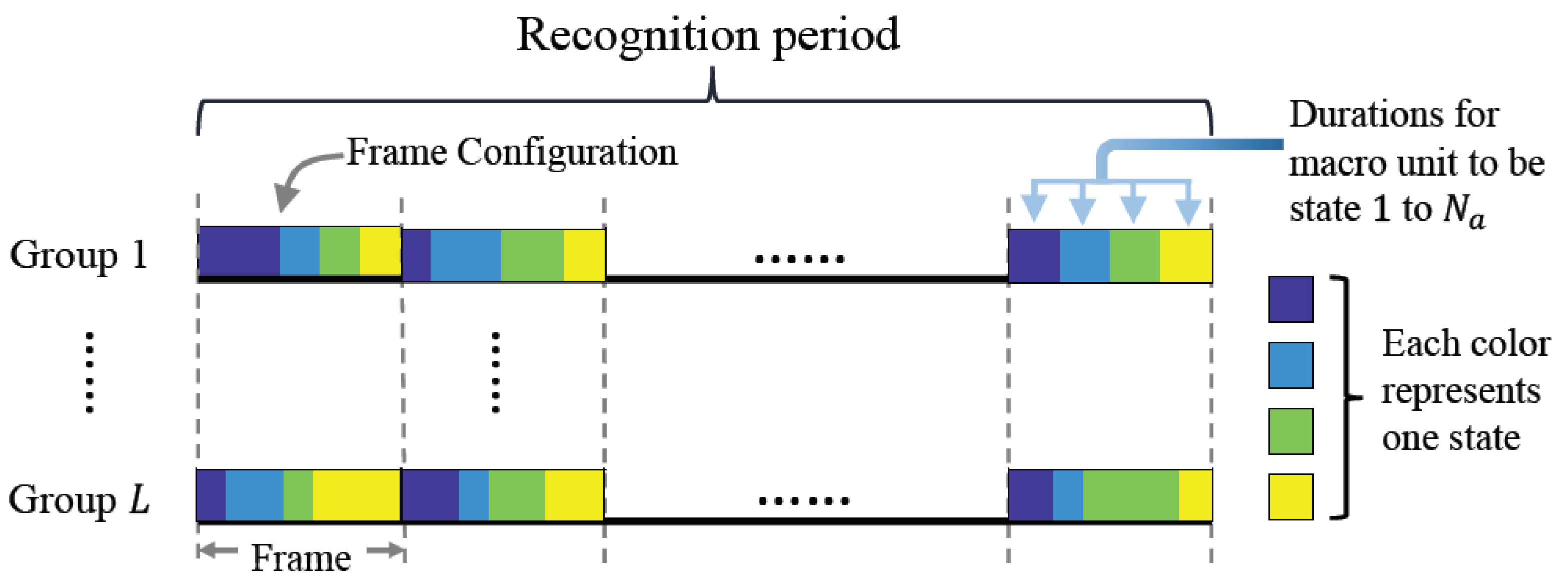
Recognition Period:

- Contains K *frames*, during which the human posture is fixed
- Received signals during a recognition period are used for recognition

Frame Configuration:

Different *states* correspond to different *phase shifts*

- Each group of RIS elements sequentially changes from State 1 to N_a.
- Constituted by the durations that each group stays in the N_a states



Problem Formulation

Decision Function: The receiver use the decision function to generate the probabilities for deciding on different human postures.

Optimization Problem: Minimize the false recognition cost

$$\text{(P1)} \quad \min_{\boldsymbol{T},\boldsymbol{\mathcal{L}}} C_{FR}(\boldsymbol{T},\boldsymbol{\mathcal{L}}) = \sum_{i,i'} \mathrm{Pr}(\mathrm{pos}_i) \cdot cost(i,i') \cdot \mathbb{E}_{\boldsymbol{y}}[\mathrm{Pr}(\boldsymbol{y}|\mathrm{pos}_i) \cdot \mathcal{L}_{i'}(\boldsymbol{y})]$$

Optimization Variables

$\boldsymbol{T}$: Frame configurations in a recognition period

$\boldsymbol{\mathcal{L}}$: Decision function

$\mathrm{Pr}(\mathrm{pos}_i)$: Probability of Posture i to appear

$cost(i,i')$: Cost for recognizing Posture i as i'

$\boldsymbol{y}$: Measured signals in a period

$\mathcal{L}_{i'}(\boldsymbol{y})$: Probability for deciding on Posture i' given $\boldsymbol{y}$

Problem Decomposition:

- Decomposing (P1) into the frame configuration optimization and the decision function optimization.

$C_{FR}(\boldsymbol{T},\boldsymbol{\mathcal{L}})$ → $C_{FR}(\boldsymbol{T})$, $C_{FR}(\boldsymbol{\mathcal{L}})$

Coupled optimization objective in (P1)

$C_{FR}(\boldsymbol{T})$: *RIS frame configuration optimization objective*

$C_{FR}(\boldsymbol{\mathcal{L}})$: *Decision function optimization objective*

Algorithm Design

Frame Configuration Optimization:

- **Objective:**
  - Based on compressive sensing technique, minimizing mutual coherence of measurement matrix improves sensing accuracy.
  - Thus, instead of recognition cost, the average mutual coherence of the *measurement matrix* for the space blocks is minimized.
- **Methods:** Frame Configuration Alternating Optimization (FCAO)
  - Alternatively minimizing the K frame configurations.
  - Adopt *augmented Lagrangian method* to optimize configurations

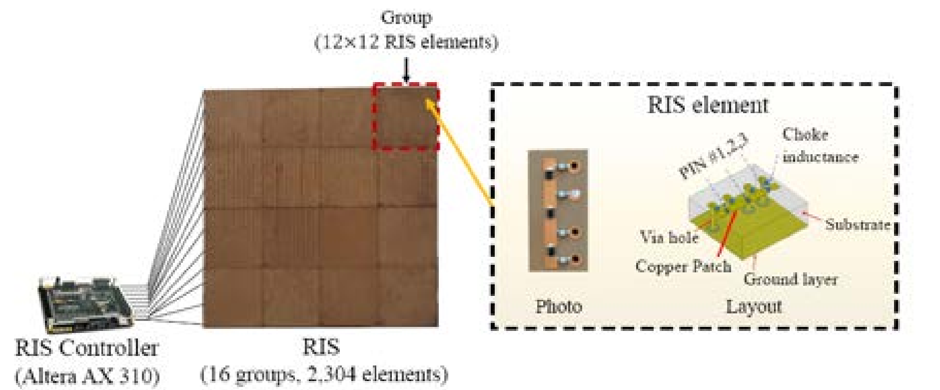
Decision Function Optimization:

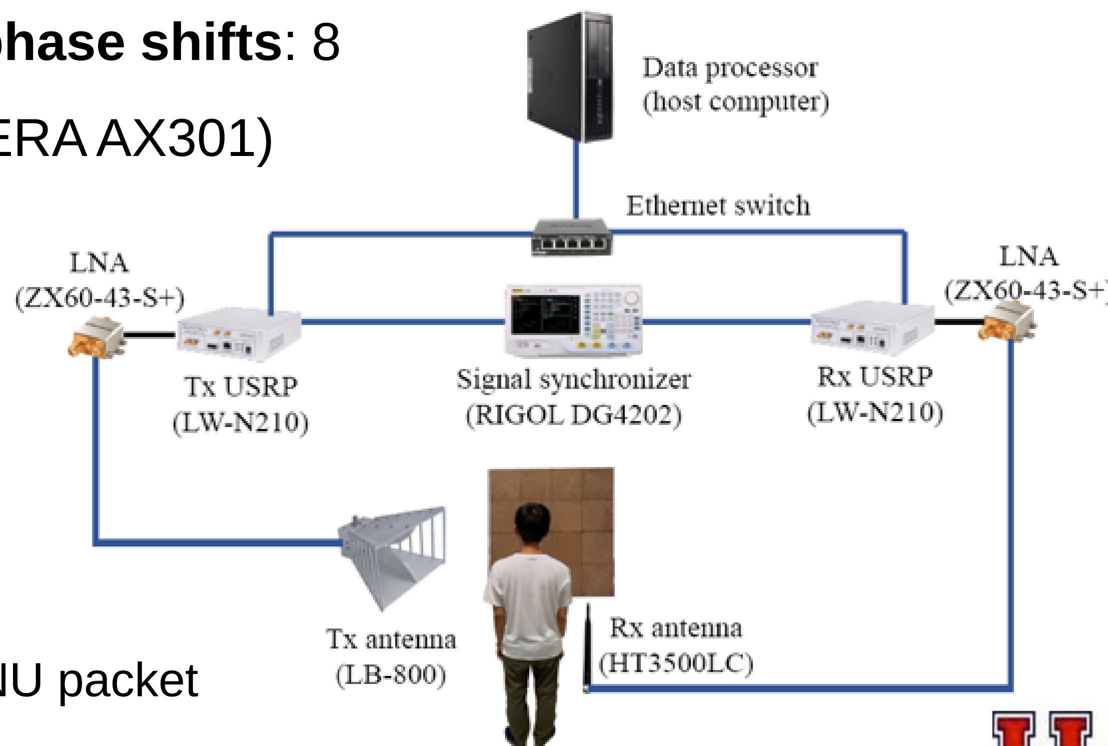
- **Objective:** Recognition cost given the optimized frame configurations
- **Methods:** Supervised learning algorithm.

Implementation

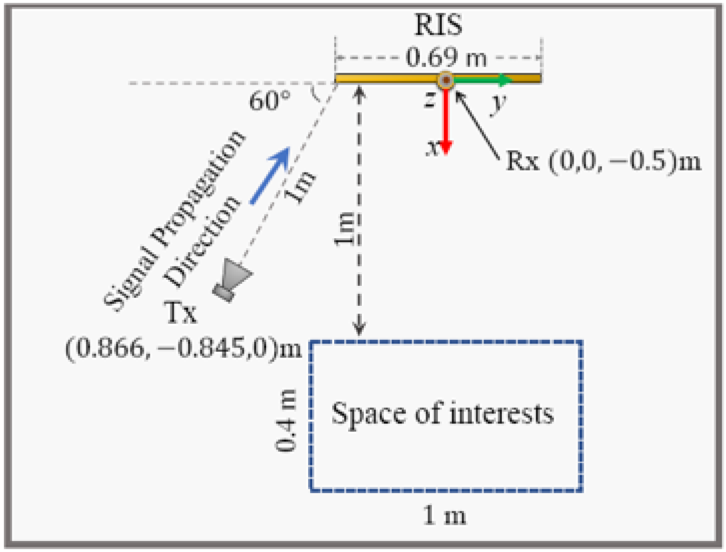
- **Size:** 69 × 69 × 0.52 cm^3
- **Dielectric substrate:** Rogers 3010

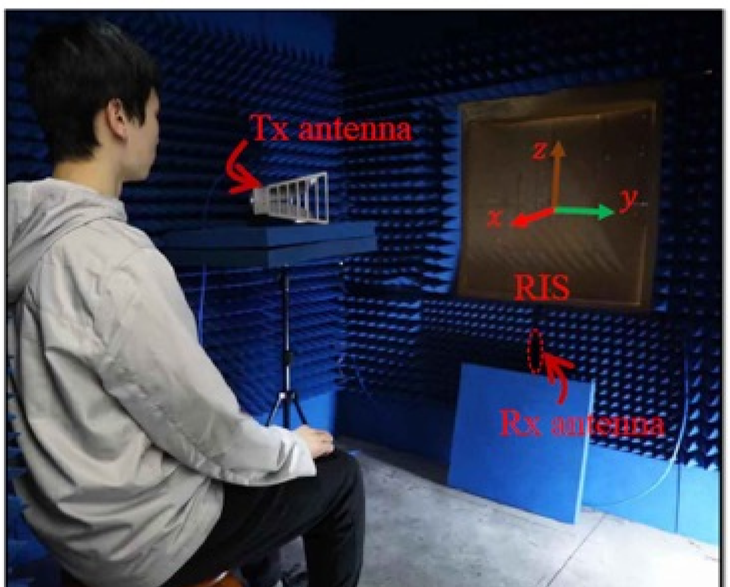
  (dielectric constant:10.2)
- **PIN diode:** BAR 65-02L
- **Total number of possible phase shifts**: 8
- **RIS controller:** FPGA (ALTERA AX301)
- **Transceivers:** USRPs
- **Others:**
  - Low-noise amplifiers
  - Signal synchronizer
  - Ethernet switch
  - Personal computer with GNU packet





Experimental Settings





Standing

Sitting

Bending

Lying down

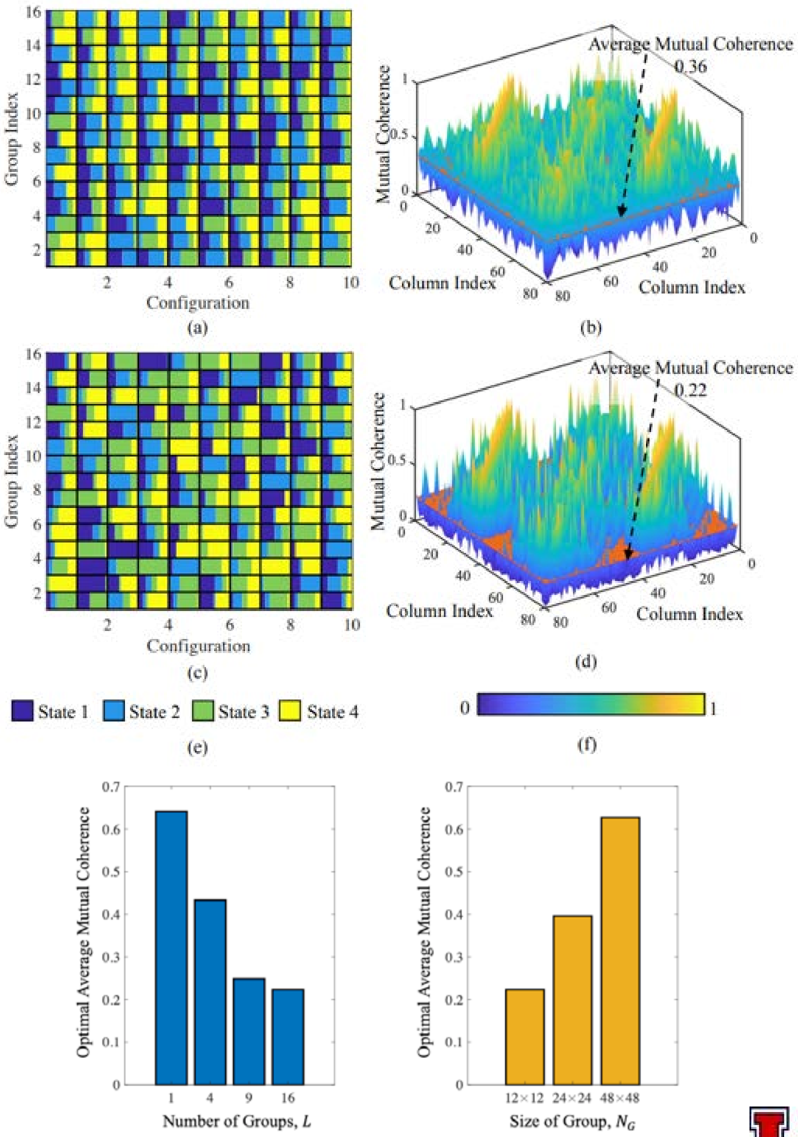
Experimental Results

Effectiveness:

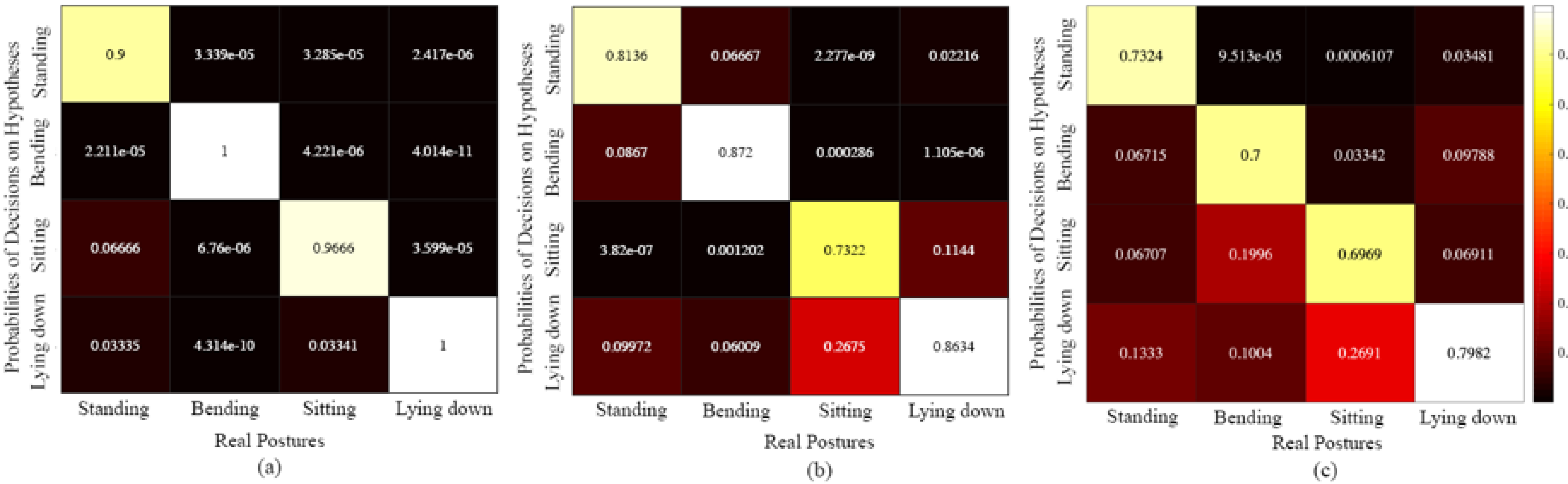
- The average mutual coherence of the measurement matrix is reduced, which can improve sensing accuracy.

Insights:

- The optimized average mutual coherence decreases with the number of groups that the RIS contains, i.e., the size of the RIS.
- The optimal average mutual coherence increases with the size of groups.



Experimental Results



- Compared with traditional RF sensing systems, RIS increases the posture recognition accuracy with 23.5%.
- Compared with the system with random frame configurations, the system with optimized frame configurations achieves 14.6% higher recognition accuracy.

Brief Summary

RIS-based posture recognition

- Propose a periodic configuring protocol
- Formulate the false recognition cost minimization problem
- Propose the FCAO algorithm to obtain the configuration matrix
- Use supervised learning to learn the optimal decision function

Remarks:

- Configuration matrix with lower average mutual coherence has higher recognition accuracy and a lower false recognition cost
- Optimized configuration can achieve 14.6% and 23.5% higher recognition accuracy, respectively.

Case Study VII: RIS aided Localization

Towards Ubiquitous Positioning by Leveraging Reconfigurable Intelligent Surface

[12] H. Zhang, et al, "Towards Ubiquitous Positioning by Leveraging Reconfigurable Intelligent Surface," IEEE Communications Letter, under revision.

Motivation

RSS based positioning: Users' locations are obtained by comparing the measured RSS and the stored RSS distribution in the indoor environment.

- Easy to implement
- Limited by unfavorable RSS distributions

RIS aided RSS based multi-user positioning:

- Users receive the signals from the AP and the RIS.
- RIS adjusts the RSS distribution by changing its configuration.

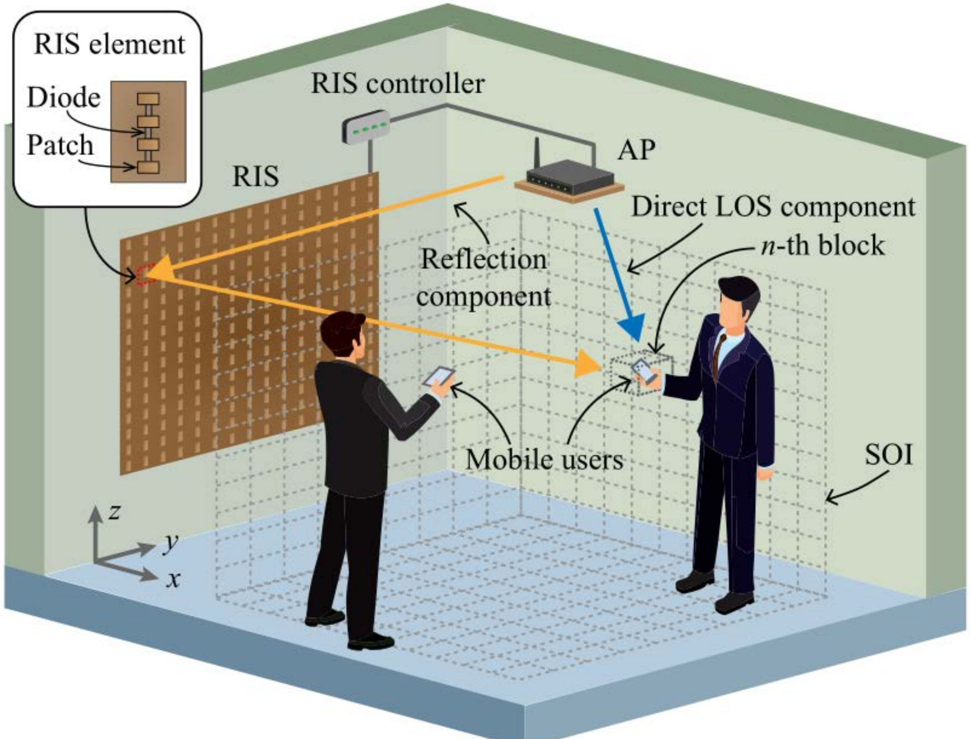
Challenge

- RIS configuration design
  - Large number of RIS configurations.
  - Complicated relation between the RIS configuration and the RSS distribution.

System Model

Positioning Scenario

- AP: sends signals to the RIS and mobile users.
- RIS: reflects the signals from the AP to the users.
- Users: measure the RSS for positioning.
- Space of Interest (SOI): is discretized into *N* blocks to represent users' positions.



RIS Model

- *M* elements.
- Each element has *C* states with different reflection coefficients.

$$r_m(c_m) = \underbrace{r(c_m)}_{\text{Amplitude}}\, e^{\underbrace{-jc_m\Delta\theta}_{\text{Phase shift}}}$$

- Configuration $\boldsymbol{c}$: the vector of all the elements' states

System Model

RSS Model

- Direct LOS channel h_{lo}: AP → User at the n-th block
- Reflection channel $h_{m,n}(c_m)$: AP → element m → User at the n-th block

$$h_{m,n}(c_m) = \frac{\lambda}{4\pi} \cdot \frac{\sqrt{g_m^t g_{m,n}^r}\, r_m(c_m) e^{-j2\pi(l_m^r + l_{m,n}^r)/\lambda}}{l_m^r l_{m,n}^r}.$$

Wavelength of the RF signal

Power gains of AP and user antennas

Distance between AP and the m-th element

Distance between the m-th element and the user

- RSS at the n-th block under configuration $\boldsymbol{c}$

$$s_n(\boldsymbol{c}) = s^t + 20 \log_{10} \left| h_{\mathrm{lo}} + \sum_{m \in \mathcal{M}} h_{m,n}(c_m) \right| + \xi.$$

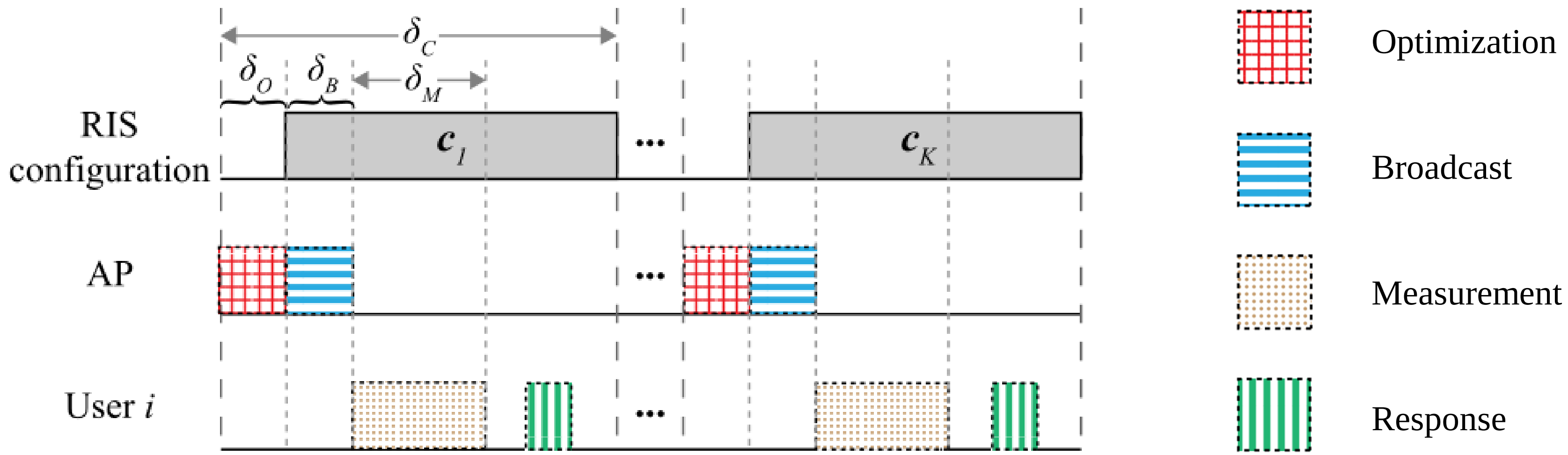
Transmission power of AP

Log-normal shadowing component

Positioning Protocol

The positioning process has *K* cycles, and each cycle contains four steps:

- **Optimization:** AP selects the optimal configuration $\boldsymbol{c}_k$ for this cycle.
- **Broadcast:** AP broadcasts $\boldsymbol{c}_k$ to users and the RIS.
- **Measurement:** AP sends single-tone signal with frequency f_c, and users record the RSS under configuration $\boldsymbol{c}_k$.
- **Response:** Users send the RSS information to the AP.



Problem Formulation

Objective: Minimize the average positioning loss (weighted probabilities of false positioning) in every cycle.

$$l(\boldsymbol{c}^k) = \sum_{i \in I} \sum_{\substack{n,n' \in \mathcal{N} \\ n \neq n'}} p_{i,n}^k \gamma_{n,n'}^k \int_{\mathcal{R}_{i,n'}^k} \mathbb{P}\left(s_i^k \middle| \boldsymbol{c}^k, n\right) \cdot ds_i^k$$

- $p_{i,n}^k$: prior probability that user i is at the n-th block in the k-th cycle.
- $\gamma_{n,n'}^k$: loss parameter when the positioning result is the n'-th block while the user is at the n-th block.
- $\mathbb{P}\left(s_i^k \middle| \boldsymbol{c}^k, n\right)$: probability that user i receives s_i^k under $\boldsymbol{c}^k$ at the n-th block.
- $\mathcal{R}_{i,n'}^k$: decision region for block n'.
  - Obtained using the maximum likelihood estimation method [2].
  - If $s_i^k \in \mathcal{R}_{i,n'}^k$, we estimate that user i's position is n' in the k-th cycle.

[2] M. A. Youssef, et al, "WLAN location determination via clustering and probability distributions," in Proc. IEEE PerCom, Fort Worth, TX, Mar. 2003.

Algorithm & Analysis

Configuration Optimization (CO) Algorithm

- Initialization phase
  - Find Z^l different local minimum configurations, denoted by $\mathcal{C}$.
  - Alternating optimization method.
- Global search phase
  - Iteratively infer other local minimum configurations using $\mathcal{C}$.
  - Steepest descent method.

Complexity

$$O(Z^u\, I^2\, M^2\, N^2)$$

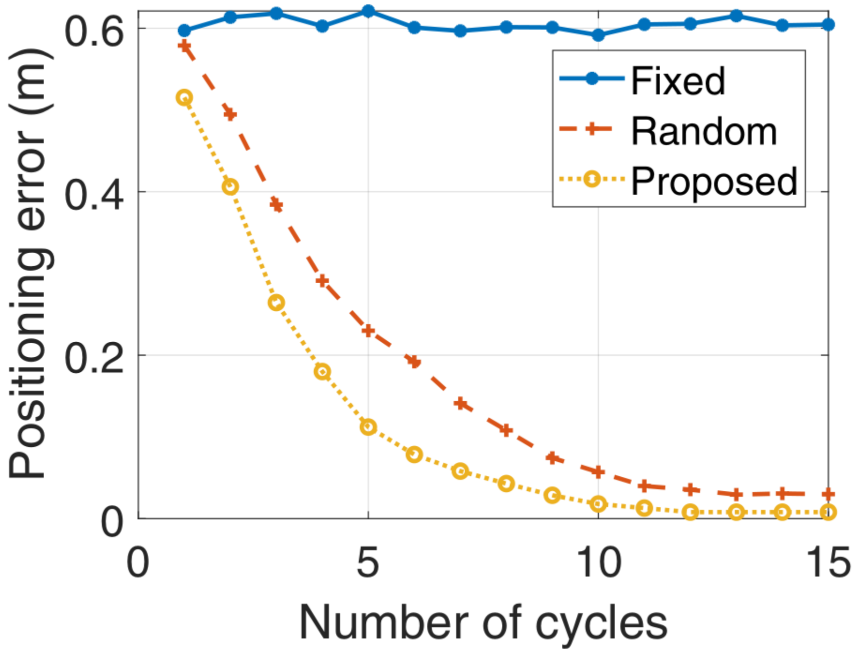
Number of local minimum configurations found using the CO algorithm

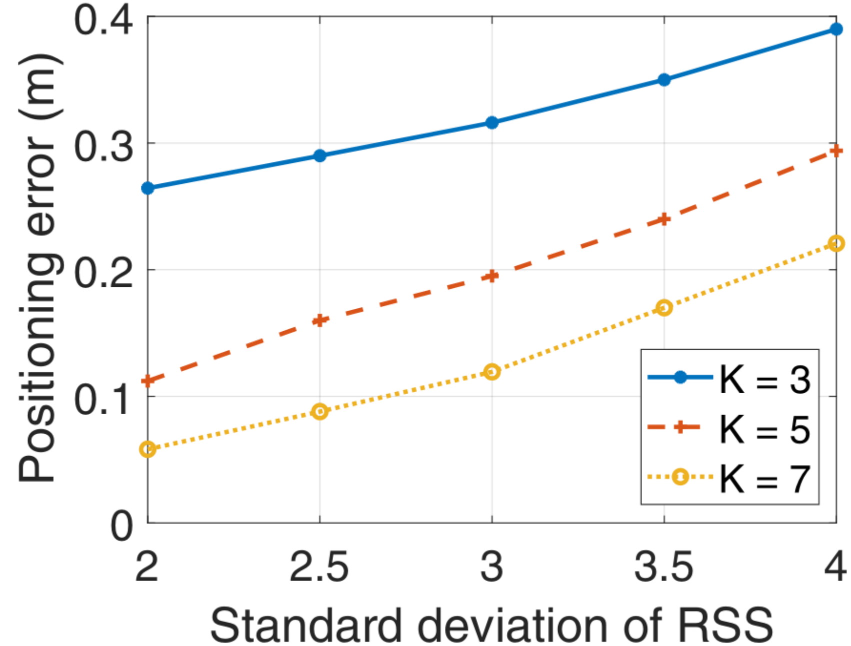
Number of mobile users

Number of RIS elements

Number of blocks in the SOI

Simulation Results





- The positioning error obtained by the proposed scheme is much lower and has a faster convergence speed than that of the random configuration scheme.
- The positioning error increases when the standard deviation increases and number of cycles *K* decreases.

Brief Summary

RIS-based positioning

- Propose a positioning protocol
- Formulate the average positioning minimization problem
- Propose the CO algorithm to obtain the optimal state of the RIS

Remarks:

- Positioning error obtained by the proposed scheme is much lower than the benchmarks
- Positioning error increases when the standard deviation increases and the number of cycles decreases

Potential Directions

Convergence with communications

- Coexistence: cognitive radio manner
- Cooperation: exchange information to support each other
- Co-design: efficient waveform design for different purposes

Signal processing

- Mobility and Doppler resolution
- Angular resolution and non-uniform illumination

Other issues

- Context-awareness
- Security and privacy
- Energy efficiency

Conclusions

- RIS is a promising solution for 6G providing an intelligent paradigm to shape the environments
  - Improve spectrum efficiency and network capacity
  - Extend the coverage and serve cell-edge users
  - Integrate imaging, sensing, and wireless communications
- We explore different aspects related to RIS-aided communications, sensing, and positioning
  - Limited phase shift effect
  - RIS orientation and placement for coverage extension
  - Hybrid beamforming and enhanced transmission
  - RF sensing for posture recognition
  - Ubiquitous positioning

Publications (1)

RIS aided Cellular Communications

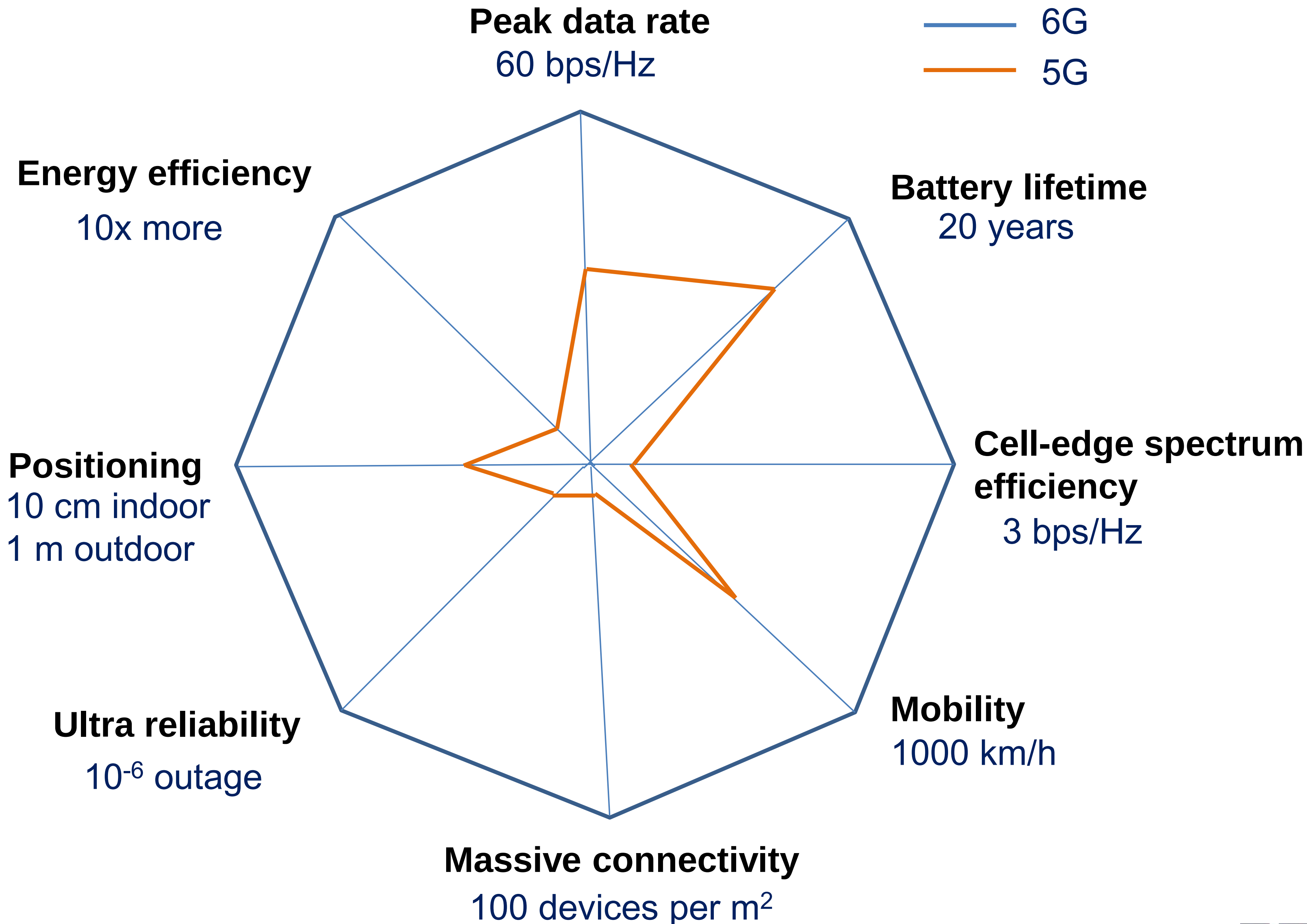
1. B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, “Hybrid Beamforming for Reconfigurable Intelligent Surface based Multi-user Communications: Achievable Rates with Limited Discrete Phase Shifts”, IEEE J. Sel. Areas Commun., to be published.
2. H. Zhang, B. Di, L. Song, and Z. Han, “Reconfigurable Intelligent Surfaces assisted Communications with Limited Phase Shifts: How Many Phase Shifts Are Enough?” IEEE Trans. Veh. Technol., vol. 69, no. 4, pp. 4498-4502, Apr. 2020.
3. B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, “Practical Hybrid Beamforming with Limited-Resolution Phase Shifters for Reconfigurable Intelligent Surface based Multi-user Communications”, IEEE Trans. Veh. Technol., vol. 69, no. 4, pp. 4565-4570, Apr. 2020.
4. M. A. Elmossallamy, H. Zhang, L. Song, K. Seddik, Z. Han, and G. Y. Li, “Reconfigurable Intelligent Surfaces for Wireless Communications: Principles, Challenges, and Opportunities,” IEEE Trans. Cognitive Commun. Netw., to be published.
5. D. Ma, L. Li, H. Ren, D. Wang, X. Li and Z. Han, “Distributed Rate Optimization for Intelligent Reflecting Surface with Federated Learning,” IEEE International Conference on Communications (ICC), Dublin, Ireland, Jun. 2020.
6. S. Zeng, H. Zhang, B. Di, Z. Han, and L. Song, “Reconfigurable Intelligent Surface (RIS) Assisted Wireless Coverage Extension: RIS Orientation and Location Optimization,” IEEE Commun. Lett., under revision.
7. Y. Chen, B. Ai, H. Zhang, Y. Niu, L. Song, Z. Han, and H. V. Poor, “Reconfigurable Intelligent Surface Assisted Device-to-Device Communications,” IEEE Trans. Wireless Commun., under revision. Arxiv: https://arxiv.org/abs/2007.00859.
8. X. Cao, B. Yang, H. Zhang, C. Yuen, and Z. Han, “Reconfigurable Intelligent Surfaces Assisted MAC for 6G: Protocol Design, Analysis and Optimization,” IEEE Internet Things J., under revision.

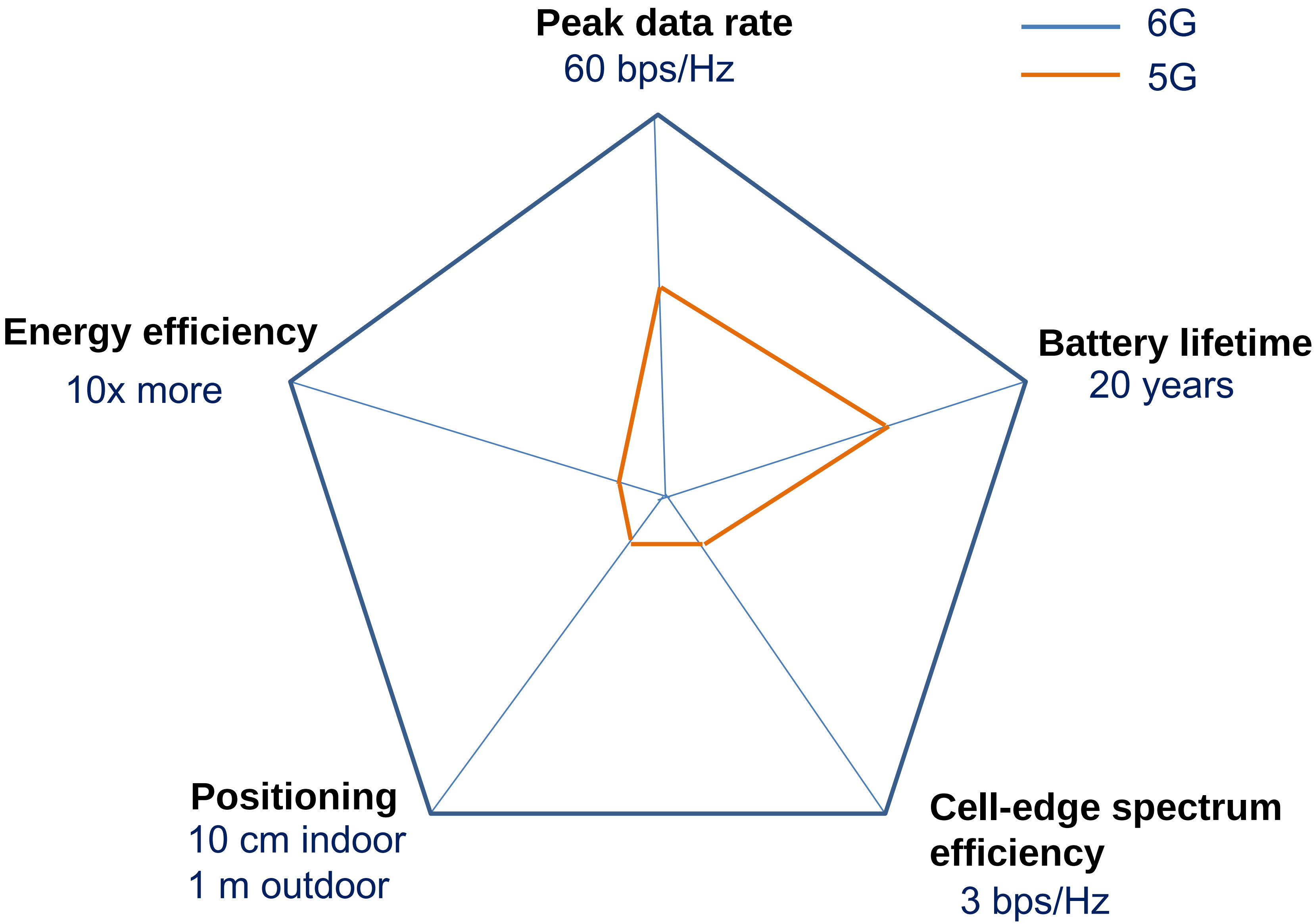
Publications (2)

RIS aided Internet of Things

1. J. Hu, H. Zhang, B. Di, L. Li, L. Song, Y. Li, Z. Han, and H. V. Poor, “Reconfigurable Intelligent Surfaces based RF Sensing: Design, Optimization, and Implementation,” IEEE J. Sel. Areas Commun., to be published.
2. J.Hu, H. Zhang, K. Bian, M. D. Renzo, Z. Han, and L. Song, “MetaSensing: Intelligent Metasurface Assisted RF 3D Sensing by Deep Reinforcement Learning,” ,” IEEE J. Sel. Areas Commun., submitted.
3. H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, “Towards Ubiquitous Positioning by Leveraging Reconfigurable Intelligent Surface,” IEEE Commun. Lett., under revision.
4. H. Zhang, J. Hu, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, “MetaRadar: Indoor Localization by Reconfigurable Metamaterials,” IEEE Trans. Mobile Comput., submitted.

Thanks for your attending


Peak data rate
60 bps/Hz
6G
5G
Battery lifetime
20 years
Cell-edge spectrum efficiency
3 bps/Hz
Mobility
1000 km/h
Massive connectivity
100 devices per m2
Ultra reliability
10-6 outage
Positioning
10 cm indoor
1 m outdoor
Energy efficiency
10x more


Peak data rate
60 bps/Hz
6G
5G
Energy efficiency
10x more
Battery lifetime
20 years
Positioning
10 cm indoor
1 m outdoor
Cell-edge spectrum efficiency
3 bps/Hz