# Reliable Projection Based Unsupervised Learning for Efficient Constrained Beamforming Optimization

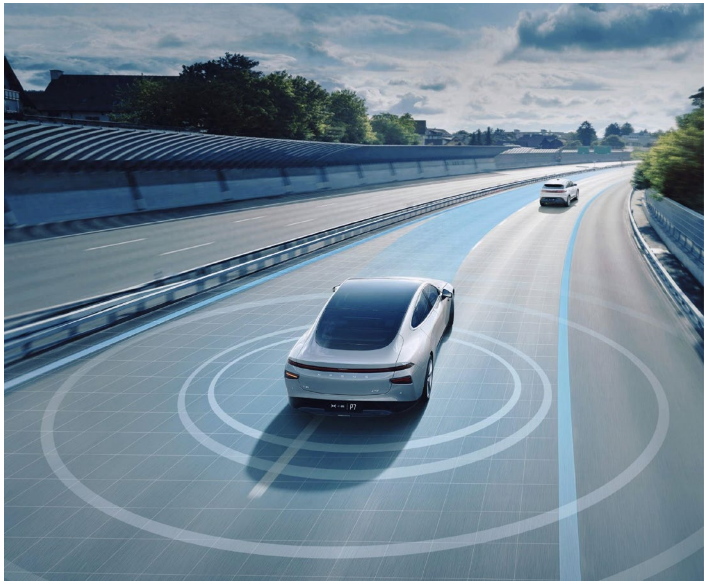
Xiucheng Wang

School of Telecommunications Engineering
Xidian University, Xi'an, China

2024/04/22

# Real-Time Constrained Optimization is Popular

**Safety Constrained IoV**

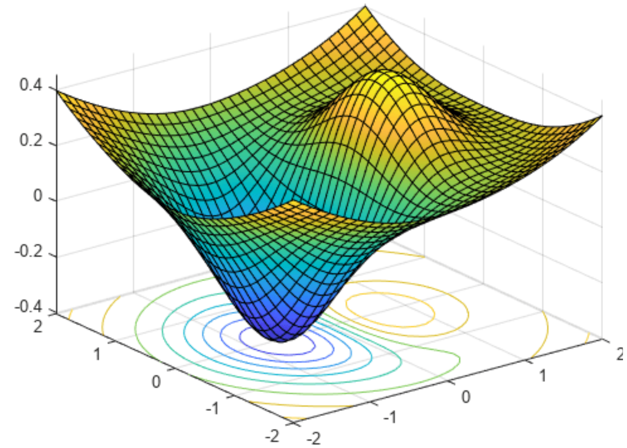


**Input**: vehicle state, sensor data
**Output**: control information
**Constraint**: safety guarantee

**User Demand Constrained Network Service**

**Input**: network state, user demand
**Output**: resource allocation
**Constraint**: network resource, user demand

# Conventional Iterative Methods **Struggle**




$$\min_{x\in\mathbb{R}^n} f(\boldsymbol{x};\boldsymbol{\phi})$$
$$s.t.\ \boldsymbol{x}\in\mathcal{B}_{\phi}$$

$\mathcal{B}$ is the feasible region

$\boldsymbol{\phi}$ is the factor matrix to define the function

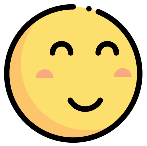
- Give $\boldsymbol{\phi}$, apply iterative strategies to pursue $\boldsymbol{x}^*$
  - Iterative Process
    - $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-g(\boldsymbol{x}_k;\boldsymbol{\phi})$
  - Example
    - PGD
    - Newton method with KKT conditions
  - Drawback
    - **High run-time complexity**
    - **Poor performance for Non-Convex Optimization**

# Conventional Iterative Methods **Struggle**

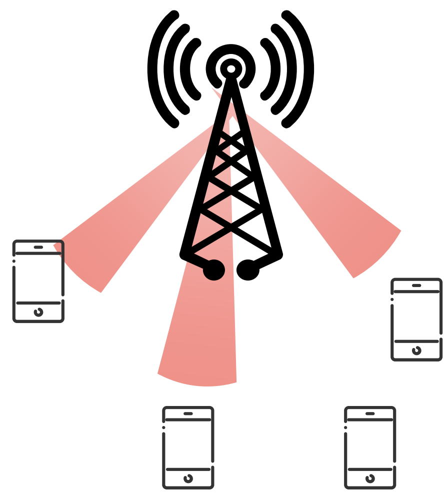
$$\min_{x \in \mathbb{R}^n} f(\boldsymbol{x}; \boldsymbol{\phi})$$
$$s.t.\ \boldsymbol{x} \in \mathcal{B}_{\phi}$$

$\mathcal{B}$ is the feasible region

$\boldsymbol{\phi}$ is the factor matrix to define the function

- Convex Objection
  w/ Convex Constraint 
- Convex Objection
  w/ Non-Convex Constraint 
- Non-Convex Objection
  w/ Convex Constraint 
- Non-Convex Objection
  w/ Non-Convex Constraint 

# Constrained Optimization in Wireless Network is **Inevitable**



MU-MISO Beamforming

$$\min_{\boldsymbol{w}} ||\boldsymbol{w}||^2$$
$$s.t.\ \frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} \geq \gamma_i$$

$\boldsymbol{h}_i$ is the channel state

$\sigma_i$ is the noise power

$\gamma_i$ is the SNR constraint

How hard of the above problem?

$$\min_{x}\ \boldsymbol{x}^H \boldsymbol{A}_0 \boldsymbol{x}$$
$$s.t.\ \boldsymbol{x}^H \boldsymbol{A}_i \boldsymbol{x} \leq c_i\ \forall i \in \{1, \cdots, M\}$$

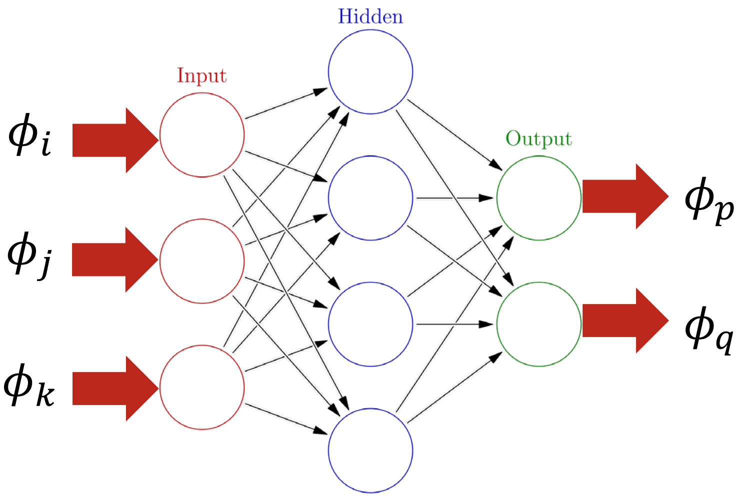
When $M \geq 4$ the above problem is NP-Hard

# NN-based Methods **Promising**

$$\min_{x \in \mathbb{R}^n} f(\boldsymbol{x}; \boldsymbol{\phi})$$
$$s.t.\ \boldsymbol{x} \in \mathcal{B}_{\boldsymbol{\phi}}$$

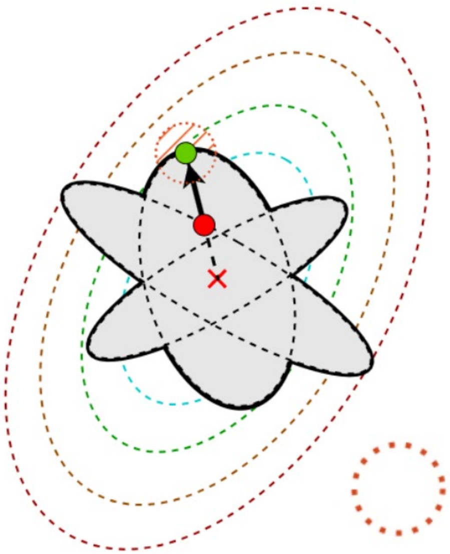
$\mathcal{B}$ is the feasible region

$\boldsymbol{\phi}$ is the factor matrix to define the function




- Give $\boldsymbol{\phi}$, obtain $\boldsymbol{x}^*$ by forward inference
  - Iterative Process
    - $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - f(\boldsymbol{x}_x; \boldsymbol{\phi})$
  - Example
    - PCNet
    - MPGNN
  - Drawback
    - **High training cost**
    - **Low feasible solution rate**
    - **Poor performance for non-convex** optimization

# Ensuring NN Solution Feasibility is **Challenging**




- Only **NN with infinitely wide** can map solution into feasible region without error: **universal approximation theorem**
- Backward propagation based naïve **NN cannot both effective and robustness**[1]

| Method | Solution Feasibility Guarantee | Performance Guarantee | Low Run-Time Complexity |
|---|---|---|---|
| Penalty | ✗ | ✗ | ✓ |
| Sampling | ✓ | ✗ | ✓ |
| Naïve Projection | ✓ | ✗ | ✗ |

[1] Zhang J J, Meng D. Quantum-Inspired Analysis of Neural Network Vulnerabilities: The Role of Conjugate Variables in System Attacks[J] *National Science Review*, 2024; nwae14.

# Motivation of Reliable Projection

**Lemma 1**. Given any non-zero vector $\boldsymbol{w}$ there exists a scalar $t$ such that the vector scaled by $t$, expressed as $t\boldsymbol{w}$, adheres to all constraints satisfied in MU-MISO beamforming problem.

*Proof*: Consider the function $f(t) = \frac{||\boldsymbol{h}_i t\boldsymbol{w}||^2}{\sigma^2} = t^2 \frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2}$.

Since $\frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} > 0$, the if $t \to +\infty$, the $f(t) \to +\infty$.

Moreover, $\gamma_i$ is bounded.

- A naïve method: If the solution output by NN is infeasible, we can project such a solution by scaling a large enough $t$.

# Optimization Analyzing of MU-MISO

**Theorem 1**. The optimal solution to MU-MISO problem must be on the boundary of the feasible region

*Proof*: The Lagrangian dual function is

$$L(\boldsymbol{w}, \boldsymbol{\lambda}) = ||\boldsymbol{w}||^2 - \sum_{i=1}^{M} \lambda_i \left( \gamma_i - \frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} \right)$$

Recall the KKT condition

$$\nabla_{\mathbf{w}} L = 2\boldsymbol{w} - \sum_{i=1}^{M} \lambda_i \left( \frac{2\boldsymbol{h}_i \boldsymbol{h}_i^H}{\sigma_i^2} \right) \boldsymbol{w} = 0, \text{ (1)}$$

$$\lambda_i \left( \gamma_i - \frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} \right) = 0, \text{ (2)}$$

$$\frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} \geq \gamma_i, \text{ (3)}$$

If all $\lambda_i = 0$, then the equation (1) is feasible only when $\boldsymbol{w} = 0$.

However, is $\boldsymbol{w} = 0$ the equation (3) cannot hold.

Therefor, there is at least one $\lambda_i = 0$.

# Motivation of Reliable Projection

Can we **project** the infeasible solution output by NN **into** the feasible region?

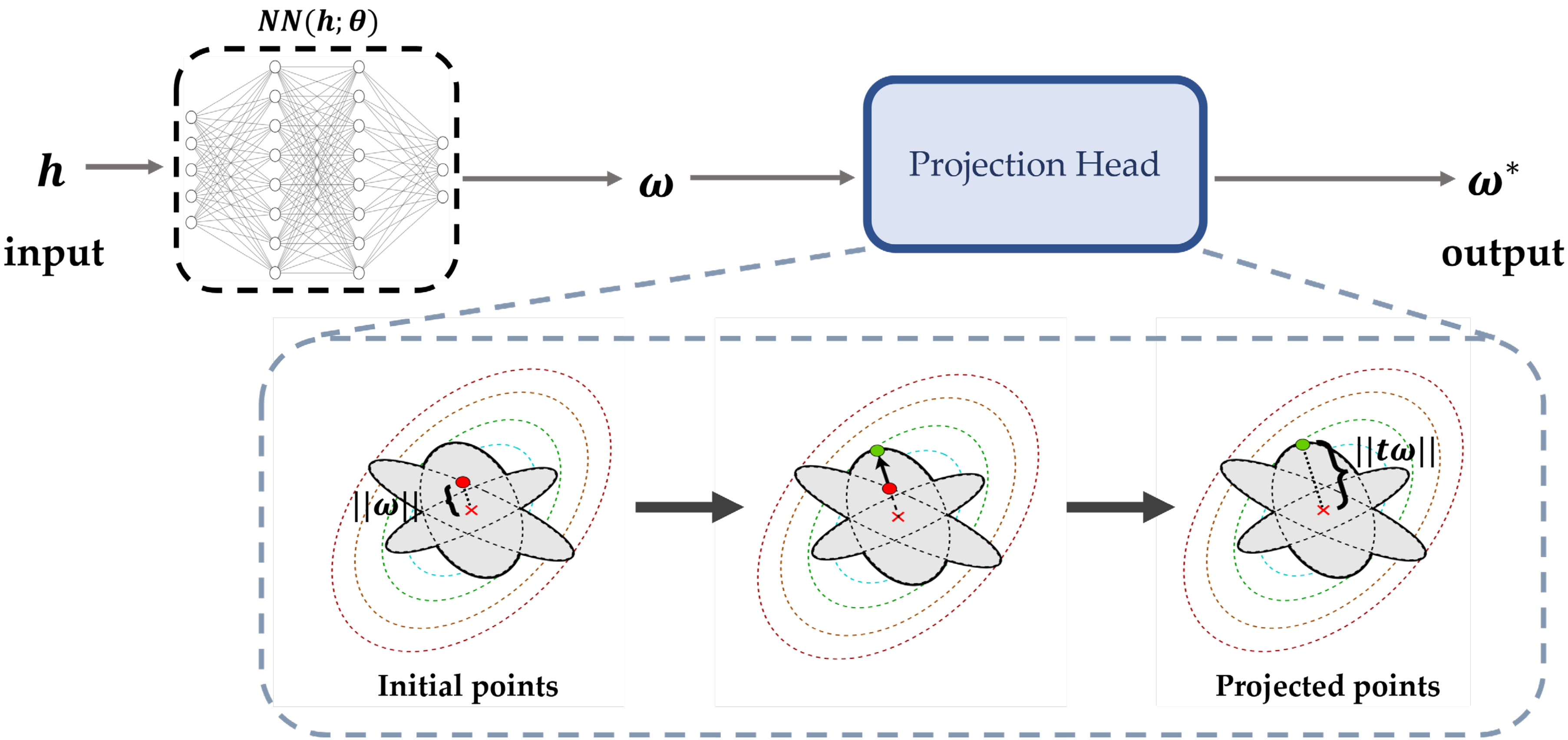
# Reliable Projection Function

$$t = \begin{cases} 1, \text{if all constrats are satisfied} \\ \max\limits_{i \in \{1,\cdots,M\}} \sqrt{\dfrac{\sigma_i \gamma_i}{||\boldsymbol{h}_i \boldsymbol{w}||^2}}, \text{otherwise} \end{cases}, \quad (4)$$

**Corollary 1**. Given any non-zero vectors $h$ and $w$ the projection function (4) ensures that all constraints are satisfied upon scaling $\boldsymbol{w}$ by $t$ to yield $t\boldsymbol{w}$.

The proof is omitted, see the paper for details.

# Reliable Projection-Based L2O Method



# Reliable Projection-Based L2O Method

## Can we further improve the projection?

- Recall that the optimal solution of the problem is on the boundary of the feasible region.
- Can the feasible solution output be NN project onto the boundary of the feasible region.

# Reliable Projection-Based L2O Method

The answer is quite simple.

$$t = \max_{i \in \{1,\cdots,M\}} \sqrt{\frac{\sigma_i \gamma_i}{||\boldsymbol{h}_i \boldsymbol{w}||^2}}, \ (5)$$

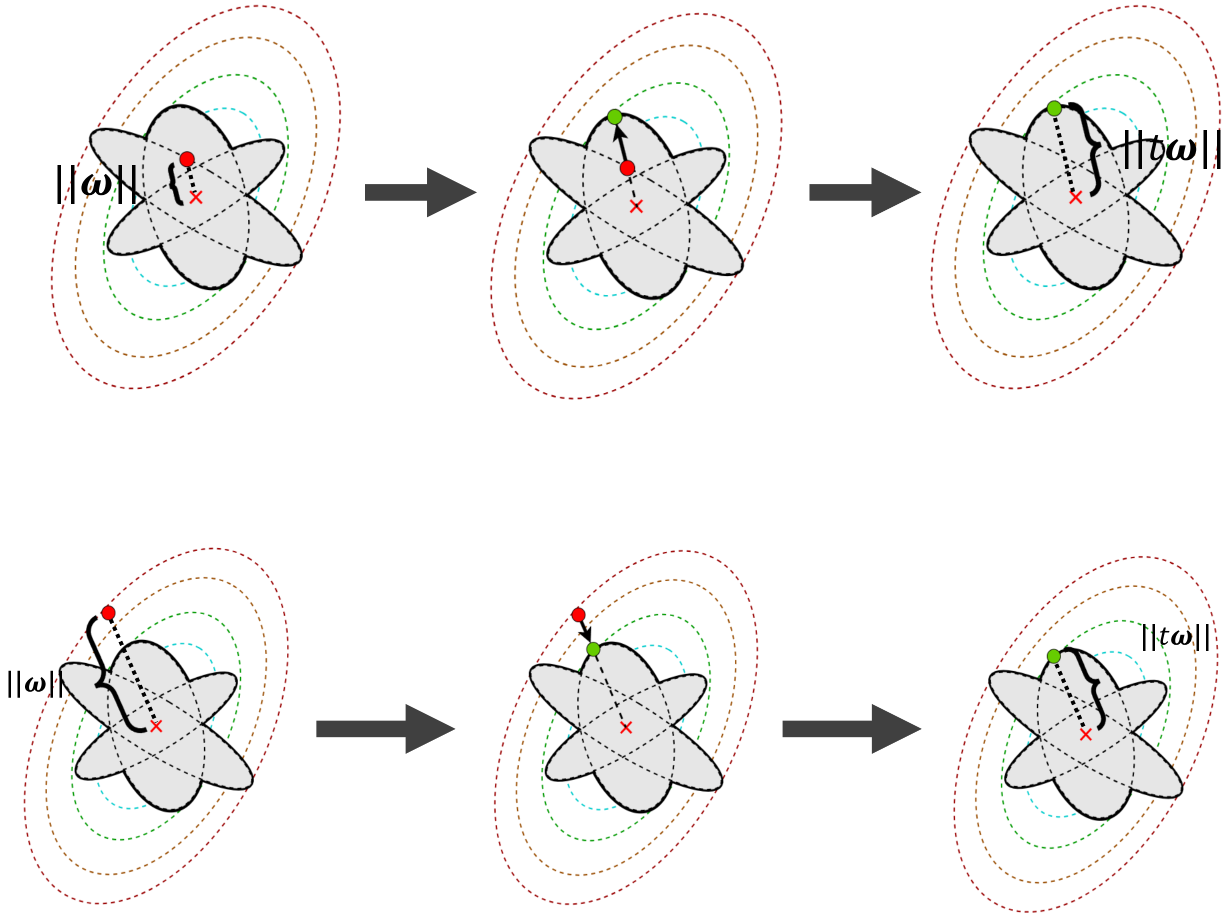
**Corollary 2**. Given any non-zero vectors $\boldsymbol{h}$ and $\boldsymbol{w}$ the projection function (5) not only projects any infeasible solution into the feasible region but also enhances the performance of an initial solution that is already feasible.

The proof is omitted, see the paper for details.

# Reliable Projection-Based L2O Method

**Infeasible Initial Solution**

**Feasible Initial Solution**




**Initial point**

**Projected point**

**Coordinate origin**

**Feasible region boundary**

**Constraint boundary**

**Infeasible region**

# Property Analysis

$$\min_{w} ||\boldsymbol{w}||^2$$
$$s.t. \frac{||\boldsymbol{h}_i \boldsymbol{w}||^2}{\sigma^2} \geq \gamma_i$$

$\Leftrightarrow$

$$\min_{w} ||t\boldsymbol{w}||^2$$
$$s.t. t = \max_{i \in \{1,\cdots,M\}} \sqrt{\frac{\sigma_i \gamma_i}{||\boldsymbol{h}_i \boldsymbol{w}||^2}}$$

- Constrained optimization for $w$
- Extremely large optimal solution search space

- Unconstrained optimization for $w$.
- Limited optimal solution search space

The proof is omitted, see the paper for details.

# Property Analysis

**Theorem 2**. The inferencing complexity of the proposed method is $\mathcal{O}(MN)$, where $N$ is the number of antennas.

The proof is omitted, see the paper for details.

**Corollary 3**. The proposed projection can reduce the inferencing complexity of NN -- Qualitative analyze by the universal approximation theorem.

**Corollary 4**. The proposed projection can reduce the training complexity of NN -- Qualitative analyze by using theorem in [2]

The analysis is omitted, see the paper for details.

[2] Xiao L. On the convergence rates of policy gradient methods[J]. Journal of Machine Learning Research, 2022, 23(282): 1-36.

# Training the L2O is also Challenging

- Supervised Learning
  - Cannot obtain enough optimal solution as label in acceptable latency.
  - The performance of the NN is limited by the label-finding method.
- Reinforcement Learning
  - Extremely long training latency.
  - Unreliable performance.

# Unsupervised Training Method

The parameters $\theta$ of the NN can be updated as follows.

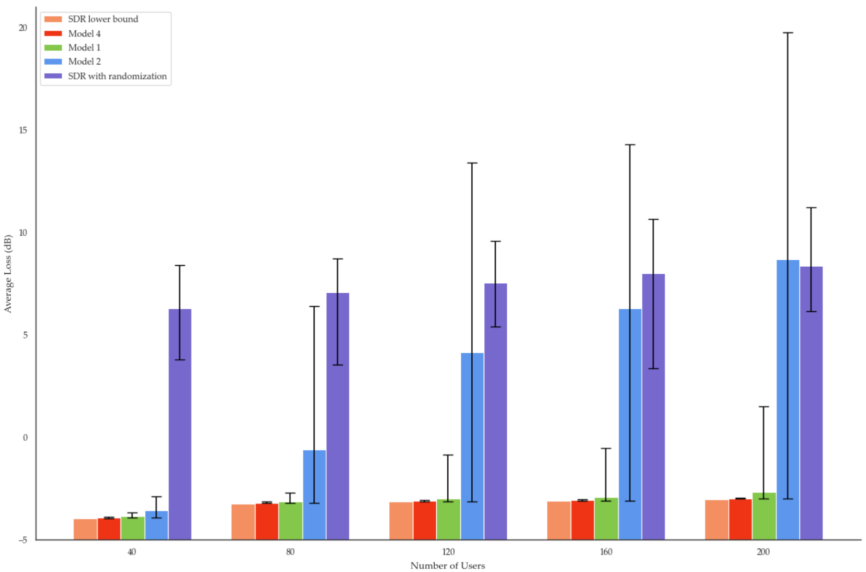
$$\frac{\partial \|t\boldsymbol{w}\|^2}{\partial \boldsymbol{\theta}} = 2\frac{\partial t\boldsymbol{w}}{\partial \boldsymbol{w}}\bigg|_{\boldsymbol{w}=NN(h_1,\cdots,h_M;\boldsymbol{\theta})} \frac{\partial \boldsymbol{w}}{\partial \boldsymbol{\theta}}.$$

$$\frac{\partial t\boldsymbol{w}}{\partial \boldsymbol{w}} = -\frac{\gamma_k \sigma_k^2 \boldsymbol{H}_k \boldsymbol{w}}{(\boldsymbol{w}^H \boldsymbol{H}_k \boldsymbol{w})^{3/2}} + \max_{i\in\{1,\ldots,M\}} \sqrt{\frac{\gamma_i \sigma_i^2}{\|\boldsymbol{h}_i^H \boldsymbol{w}\|^2}},$$

$$k = \arg\max_{i\in\{1,\ldots,M\}} \sqrt{\frac{\gamma_i \sigma_i^2}{\|\boldsymbol{h}_i^H \boldsymbol{w}\|^2}}$$

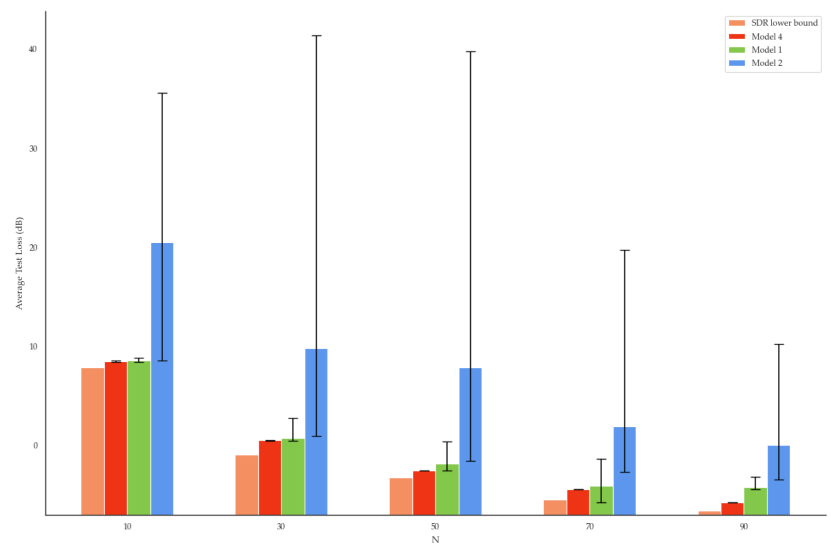
The proof is omitted, see the paper for details.

# Simulation Analysis




- As the number of users increases, the performance of proposed method is always similar to the lower bound.

- As the number of antennas increases, the performance of proposed method is always better than other methods.



# Conclusion

- Instead of focusing only on the probability that the constraint is satisfied as in the previous work, to our best knowledge, we first propose a L2O method where all constraints can be satisfied by reliable projection.
- An unsupervised learning method is used to reduce the training cost of L2O
- Future work will focus on how to combine the L2O and iteration based method to solve the CO both effectively and efficiently.

# Thanks for Listening

# Author Contributions

- Conceptualization: Xiucheng Wang.
- Validation: Xiucheng Wang, Qi Qiu.
- Writing original draft preparation: Xiucheng Wang, Qi Qiu.
- Theorem proof: Xiucheng Wang, Xuan Zhao.