

# Robust Reconfigurable Intelligent Surfaces via Invariant Risk and Causal Representations

Sumudu Samarakoon\*, Jihong Park†, and Mehdi Bennis\*

\*Centre for Wireless Communication, University of Oulu, Finland, email: {sumudu.samarakoon,mehdi.bennis}@oulu.fi

†School of Information Technology, Deakin University, Geelong, VIC 3220, Australia, email: jihong.park@deakin.edu.au

**Abstract**—In this paper, the problem of robust reconfigurable intelligent surface (RIS) system design under changes in data distributions is investigated. Using the notion of invariant risk minimization (IRM), an invariant causal representation across multiple environments is used such that the predictor is simultaneously optimal for each environment. A neural network-based solution is adopted to seek the predictor and its performance is validated via simulations against an empirical risk minimization-based design. Results show that leveraging invariance yields more robustness against unseen and out-of-distribution testing environments.

**Index Terms**—Invariant risk minimization, reconfigurable intelligent surfaces, causality.

## I. INTRODUCTION

Reconfigurable intelligent surfaces (RISs) have recently gained remarkable attention as a low-cost, hardware-efficient, and highly scalable technology capable of offering dynamic control of electromagnetic wave propagation [1]–[3]. Due to the design based on nearly passive multiple reflective elements, the channel acquisition and the dynamic configuration of RIS parameters are two main challenges in RIS-assisted wireless communication.

The overwhelming majority of the existing literature on communication aided by RISs relies on the availability of perfect/estimated channel state information (CSI) to develop machine learning (ML)-based RIS configuration [2]–[5]. Therein, frequent channel sampling via dedicated control signaling and pilot-based estimation techniques that introduce additional overhead and communication complexity to the system under channel dynamics have been used. To reduce such overheads imposed by the need of perfect CSI, the works of [6]–[8] have developed neural network (NN)-based RIS configuration solutions utilizing system properties such as locations and relative distances of transmitters, receivers, and RISs instead of relying on CSI measurements. All current designs mainly focus on exploiting statistical correlations within the observed data. However, these works neglect the data generation process and the underlying causal relationships between the environment and RIS configurations. Moreover, current approaches yield high inference accuracy for a given environment, mostly when the training and test data distribution are identical. As

a result, these approaches fail to generalize out-of distribution (OOD) data and other unseen environmental changes.

The main contribution of this article is to fill this void and develop a **novel robust learning framework for predicting RIS configurations** rooted in learning over representations that are invariant across different environments. Towards this goal, it is essential to discover the causal dependencies of system components (e.g., RIS configuration, angles of arrivals/departures at transmitters, receivers, and reflectors) that are invariant over different environments, i.e., *invariant representation of network parameters* in contrast to exploiting environment-dependent spurious correlations (e.g., scatterer distribution affecting CSI). In this work, we use a supervised learning approach to determine the optimal RIS configurations based on CSI and network-wide parameters pairs collected over different network settings (referred to as *environments*). The conventional approach is to seek predictors based on the observed CSI that minimizes the empirical loss; a process known as empirical risk minimization (ERM). In contrast, we formulate an invariant risk minimization (IRM) problem by casting the prediction of RIS phase configuration as a loss minimization problem leveraging an invariant representation of the data (CSI) that is optimal for all environments. Based on the IRM formulation, we train a NN-based robust RIS configuration predictor and show that the proposed predictor outperforms an ERM-based predictor in terms of OOD generalization and unseen test environments with up to 15% higher prediction accuracy.

The rest of the paper is organized as follows. Section II describes the system model and the conventional design of the optimal RIS phase configuration predictor using ERM. The IRM based RIS configuration design that utilizes learning over invariant representation is discussed in III. Section IV evaluates and compares the proposed solution with the baselines by means of extensive simulations in terms of generalization and OOD robustness. Finally, conclusions are drawn in Section V.

## II. SYSTEM MODEL & PROBLEM FORMULATION

As illustrated in Fig. 1, we consider a set of multiple downlink RIS communication environments  $\mathcal{E}$  defined from the Rx's perspective, such that an environment  $e \in \mathcal{E}$  consists of a single Tx-Rx pair without line-of-sight (LOS), an RIS, randomly located scatterers in the Tx's vicinity, and an area  $\mathcal{A}^e$  that the Rx occupies. The RIS enables non line-of-sight

This research was partly supported by Academy of Finland 6G Flagship (grant no.318927) and projects NEGEIN, SMARTER, EU-ICT IntelliIoT, EUCHISTERA LearningEdge, and Infotech-NOOR.

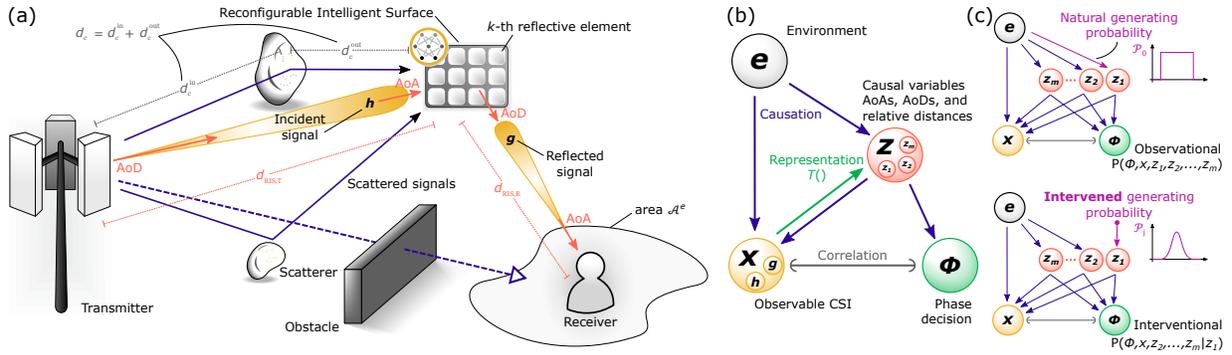


Fig. 1. System model illustrating RISs-aided Tx-Rx connectivity under blocked line-of-sight within an instance of environment  $e$  is given in (a). The causal relation among the observable features, decision, and the environment along the observable correlation is illustrated under (b). The impact of intervening the generating probability of  $z_1$  yielding an intervention outcome rather than of the observable one is illustrated in (c).

(NLOS) communication from Tx to Rx, via reflecting signals transmitted from Tx and diffracted signals from a set of scatterers ( $\mathcal{S}$ ). Therein, the goal is to maximize the downlink data rate by adjusting the phases of the RIS elements. The system model and problem formulation are elaborated next.

#### A. Channel model

Let  $\mathbf{h} = [h_k]_{k \in \mathcal{K}}$  and  $\mathbf{g} = [g_k]_{k \in \mathcal{K}}$  be the channel vectors of the incident (Tx-RIS) and reflected (RIS-Rx) signals defined over the reflective elements  $\mathcal{K}$  in the RIS. The channel model is based on the work in [9], in which, the link between Tx and RIS is composed of LOS channels as well as NLOS channels due to the presence of scatterers, while the RIS-Rx link has LOS connectivity due to their close proximity. Let  $d_o$ ,  $\phi_o$ , and  $\theta_o$  be the distance, azimuth angle, and elevation angle of an object  $o \in \{\text{Tx}, \text{Rx}\} \cup \mathcal{S}$  with respect to the RIS. With a uniformly distributed random phase  $\eta \sim \mathcal{U}[0, 2\pi]$  and  $v^2 = -1$ , under the assumption that the scatterers are only in the vicinity of the Tx, the RIS-Rx channel is modeled as follows:

$$\mathbf{g} = \sqrt{G(\theta_{\text{RIS,R}})L(d_{\text{RIS,R}})}e^{j\eta}\mathbf{a}(\phi_{\text{RIS,R}}, \theta_{\text{RIS,R}}), \quad (1)$$

where  $G(\cdot)$ ,  $L(\cdot)$ , and  $\mathbf{a}(\cdot)$  are the RIS element radiation pattern, distance-dependent path loss, and array response, respectively [9]. Similar to (1), the LOS component of the channel between RIS and Tx is modeled by,

$$\mathbf{h}_{\text{LOS}} = \sqrt{G(\theta_{\text{RIS,T}})L(d_{\text{RIS,T}})}e^{j\eta}\mathbf{a}(\phi_{\text{RIS,T}}, \theta_{\text{RIS,T}}), \quad (2)$$

as defined in the fifth generation (5G) channel model [9].

The NLOS links between the Tx and RIS are due to the presence of scatterers. Let  $d_s$ ,  $\phi_s$ , and  $\theta_s$  be the traveled-distance of the reflected signal from Tx to RIS at scatterer  $s$  and the azimuth and elevation angles of scatterer  $s$  with respect to the RIS, respectively. Then, the NLOS channel is modeled as follows;

$$\mathbf{h}_{\text{NLOS}} = \frac{1}{S} \sum_s \gamma_s \sqrt{G(\theta_s)L(d_s)}\mathbf{a}(\phi_s, \theta_s), \quad (3)$$

where  $\gamma_s \sim \mathcal{CN}(\mathbf{0}, \mathbf{1})$  is a scatterer-dependent random path gain. In this view, the channel between Tx and RIS becomes  $\mathbf{h} = \mathbf{h}_{\text{LOS}} + \mathbf{h}_{\text{NLOS}}$ .

#### B. Rate maximization

At the RIS, the phases of incident signals are altered to enhance the capacity at the Rx. Denoting the phase change decision of the RIS over its reflective elements by  $\psi = [\psi_k]_{k \in \mathcal{K}}$  with  $\text{abs}(\psi_k) = 1$ , the received signal  $y$  at the Rx is given by  $\mathbf{g}^\dagger \psi \mathbf{h} v + \tilde{n}$ , where  $v$  is the transmit signal with  $\mathbb{E}[v^2] = p$  and  $\tilde{n} \sim \mathcal{N}(0, N_0)$  is the noise. The data rate at the Rx is  $r(\psi, \mathbf{h}, \mathbf{g}) = W \log_2 \left( 1 + \frac{|\mathbf{g}^\dagger \psi \mathbf{h}|^2 p}{W N_0} \right)$  where  $W$  is the bandwidth and In this view, the data rate maximization at the Rx is cast as follows:

$$\max_{\psi \in \mathcal{C}} r(\psi, \mathbf{h}, \mathbf{g}) = W \log_2 \left( 1 + \frac{|\mathbf{g}^\dagger \psi \mathbf{h}|^2 p}{W N_0} \right), \quad (4)$$

where  $\mathcal{C}$  is the feasible set of RIS configurations.

Under the perfect knowledge of CSI over all links, the optimal phase decision  $\phi$  can be obtained through an exhaustive search. However, this poses two challenges: (i) the complexity of a heuristic search increases with the number of reflective elements and their configurations and (ii) assuming perfect CSI at RIS requires a huge number of channel measurements. The issue (i) can be addressed by the aid of ML-based regression. Therein, given a dataset  $\mathcal{D} = \{(\mathbf{x}, \phi)_n | n \in \{1, \dots, N\}\}^1$  consisting of observed CSI  $\mathbf{x} = (\mathbf{h}, \mathbf{g})$  and optimal configuration labels, a mapping function  $\hat{f}_w(\cdot)$  parameterized by  $\mathbf{w}$ , referred to as a *predictor*,  $\hat{\psi} = \hat{f}_w(\mathbf{x})$  is obtained by solving the ERM as follows,

$$\min_{\mathbf{w}} \frac{1}{N} \sum_n \ell(\phi_n, \hat{\psi}_n), \quad (5)$$

where  $\ell(\phi, \hat{\psi}) = \sum_k \ell_0 \sin^2 \left( \frac{\hat{\psi}_k - \phi_k}{2} \right)$  is the loss (or risk) function in terms of phase prediction with an arbitrary scaling coefficient  $\ell_0 > 0$ . Nevertheless, due to the issue in (ii), training over a larger dataset collected over different channel realizations and network configurations, as well as embedding CSI measurement capabilities at RISs for inference are impractical. In short the application of ERM for RISs poses algorithmic and practical difficulties in terms of lack of robustness and generalization across different environments. This calls for

<sup>1</sup>Note that the subscript  $n$  is neglected from all parameters for simplicity unless the notion of sample is significant.

developing a new learning framework for predicting phase configurations that are robust across multiple environments.

### III. IRM-BASED PHASE OPTIMIZATION

The main limitation of analyzing a limited set of observations is due to the resultant predictor  $\hat{f}_w$  that depends on spurious correlations among observables. Overfitting to spurious correlations based on the knowledge of the scattered signals while neglecting the underlying causal relations that are invariant across multiple environments prevents an ERM-based system design to operate as a robust predictor against changes in the environment. Therefore, it is essential to develop an invariant predictor based on causality rather than exploiting spurious correlations.

Towards the design of an invariant predictor across multiple environments, we first identify  $(\mathbf{x}, \phi)_n^e$  as the  $n$ -th input-output tuple (sample) of environment  $e$ . Hence, the entire dataset can be seen as a composite of data collected over different environments, i.e.,  $\mathcal{D} = \cup_e \mathcal{D}_e$ . Concretely, we seek a parameterized mapping function (predictor)  $\hat{\psi}_n^e = f_w(\mathbf{z}_n^e)$  that is robust across all environments  $\mathcal{E}$ , where  $\mathbf{z}_n^e = T(\mathbf{x}_n^e)$  is the vector of causal variables for the input  $\mathbf{x}_n^e$  obtained through a representation function  $T(\cdot)$ . In this setting, we derive an invariant mapping that is simultaneously optimal over all environments [10]. Formally, the design of the invariant predictor is cast as follows:

$$\min_{T(\cdot), \mathbf{w}} \sum_{e \in \mathcal{E}} \frac{1}{N_e} \sum_n \ell(\phi_n^e, f_w(\mathbf{z}_n^e)), \quad (6a)$$

$$\text{s. t. } \mathbf{w} \in \arg \min_{\mathbf{w}'} \frac{1}{N_e} \sum_n \ell(\phi_n^e, f_{\mathbf{w}'}(\mathbf{z}_n^e)) \quad \forall e \in \mathcal{E}, \quad (6b)$$

$$\mathbf{z}_n^e = T(\mathbf{x}_n^e) \quad \forall (\mathbf{x}, \phi)_n^e \in \mathcal{D}. \quad (6c)$$

It is worth highlighting that *without* (6b), the above problem is equivalent to *multi-task learning* denoting environments as tasks, and solved using metalearning or similar solutions [11]. However, with the simultaneous optimality introduced by (6b), we seek for a different solution as discussed next.

Due to the constraints defined over environments, (6) boils down to a bilevel optimization problem that requires lower-level optimizations per environment, which makes solving (6) challenging. Alternatively, following the IRM framework [10], we recast (6b) as a penalized loss as follows:

$$\min_{T(\cdot), \mathbf{w}} \sum_{e \in \mathcal{E}} \frac{1}{N_e} \sum_n \underbrace{\left( \ell(\phi_n^e, f_w(\mathbf{z}_n^e)) + \lambda \|\nabla_{\mathbf{w}} \ell(\phi_n^e, f_w(\mathbf{z}_n^e))\|_2 \right)}_F, \quad (7)$$

where  $\lambda > 0$  is a constant hyperparameter. By choosing a large  $\lambda$ , we enforce  $\|\nabla_{\mathbf{w}} \ell(\phi_n^e, f_w(\mathbf{z}_n^e))\|_2 \rightarrow 0$  and thus,  $\nabla_{\mathbf{w}} \ell(\phi_n^e, f_w(\mathbf{z}_n^e)) \approx 0$  constitutes the per-environment optimality defined under (6b).

Based on the channel models (1)-(3), it can be noticed that the composite CSI at the Rx depends on several features including distances, azimuth and elevation angles of Tx, Rx, and RIS. Hence, the invariant properties of angle of arrivals (AoAs), angle of departures (AoDs), and relative distances can be considered as an invariant representation such that

$\mathbf{z} = T(\mathbf{x})$  of CSI irrespective of the environment in which the channels are generated.

Due to the nature of the choice of  $f_w(\cdot)$ , determining the optimal  $\mathbf{w}$  could be derived analytically (e.g.,  $f_w(\cdot)$  as a linear regression) or obtained via gradient decent methods (e.g.,  $f_w(\cdot)$  as a NN). The objective of (7) can be modified as follows:

$$F = \ell_0 \sum_k \sin^2 \left( \frac{[f_w(\mathbf{z}_n^e)]_k - \phi_{k,n}^e}{2} \right) + \frac{\lambda \ell_0}{2} \sqrt{\sum_i \left( \sum_k \sin([f_w(\mathbf{z}_n^e)]_k - \phi_{k,n}^e) \frac{\partial [f_w(\mathbf{z}_n^e)]_k}{\partial w_i} \right)^2}, \quad (8)$$

by differentiating the loss function  $\ell(\phi_n^e, f_w(\mathbf{z}_n^e))$  with respect to  $\mathbf{w}$ . Here, (8) is used as the loss function of the regression task of  $f_w(\cdot)$ . Then, using the training data  $\mathcal{D}$  gathered over multiple environments, we learn a NN-based invariant phase predictor.

### IV. SIMULATION RESULTS

The RIS-aided communication is simulated using the ‘‘Sim-RIS’’ channel simulator [9]. Therein, an indoor environment with the Tx located at the coordinate<sup>2</sup> (0, 35, 1) using the 28 GHz carrier frequency. The RIS consists of  $K = 10 \times 10$  grid of reflective elements with half-wavelength inter-element distances located at (10, 30, 1) facing the  $y$ -axis direction. The Rx can be located at one of three environments, ENV#1, ENV#2, and ENV#3, such that the distances from RIS to the environment specific region  $\mathcal{A}^e$  centers are 2 m, 6 m, and 4 m, respectively. From each environment, 1000 channel realizations are generated, CSI of incident and reflected signals decoupled over real and imaginary components ( $\mathbf{x} \in \mathbb{R}^{200 \times 2}$ ) as well as AoDs, AoAs (including azimuth and elevation), and distances of Tx-RIS and RIS-Rx ( $\mathbf{z} \in \mathbb{R}^{10}$ ) are recorded, and the configuration  $\phi$  yielding the highest SNR is labeled as CLASS#1 if  $\phi = [0, 0, \dots, 0]$  and CLASS#2 if  $\phi = [0, \pi, \dots, 0, \pi]$ .

For training,  $N = 600$  samples are collected from ENV#1 and ENV#2 while ENV#3 is used as the testing dataset unless stated otherwise. The fraction of samples from ENV#1 within the training dataset is denoted by  $\alpha_E$ . Out of two configuration classes, we select  $\alpha_C : 1 - \alpha_C$  ratio of samples from CLASS#1 and #2 within ENV#1 while  $1 - \alpha_C : \alpha_C$  ratio of samples from CLASS#1 and #2 within ENV#2. Note that the testing dataset of ENV#3 has equal number of samples from each class. For both the baseline ERM as per (5) and the proposed method that solves (7) (referred to as IRM hereinafter), NNs with two fully-connected hidden layers of sizes  $\{16, 4\}$  within the multi layer perceptrons (MLP) architecture are trained as the phase predictors. For the purpose of the benchmark, the ground truth results obtained via *exhaustive search*, i.e., the best unique configurations  $\phi_n$  yielding the maximal rates, are referred to as BEST while the random phase decision making is indicated by RAND, hereinafter. For sake of comparison, three different metrics are used for method  $m \in \{\text{BEST}, \text{IRM}, \text{ERM}, \text{RAND}\}$ : (i) *accuracy* =  $\frac{1}{N} \sum_n \mathbb{I}(\hat{\psi}_n = \phi_n)$  measures the predictors’

<sup>2</sup>All distances are measured in meters within the Cartesian coordinates.

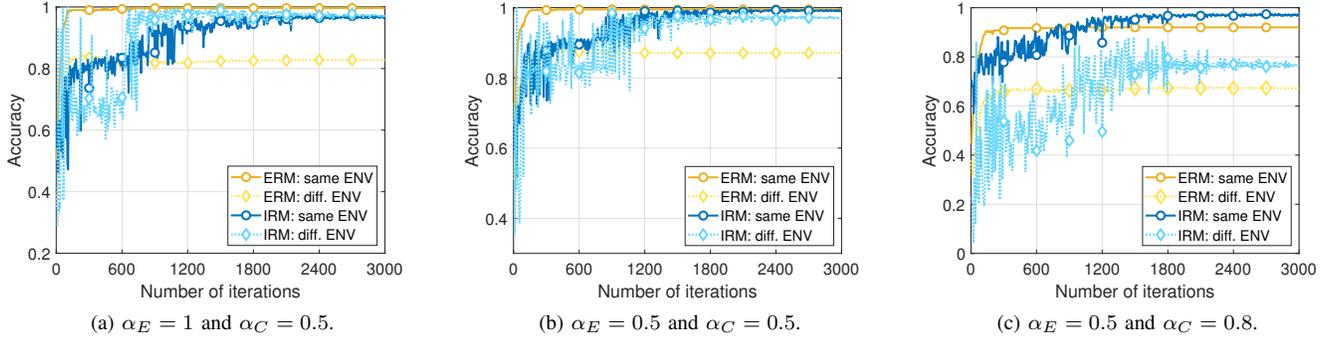


Fig. 2. Training convergence of the proposed IRM and baseline ERM methods is compared within the *same* environment (ENV#1 and/or ENV#2) as training data as well as over a *different* environment (ENV#3).

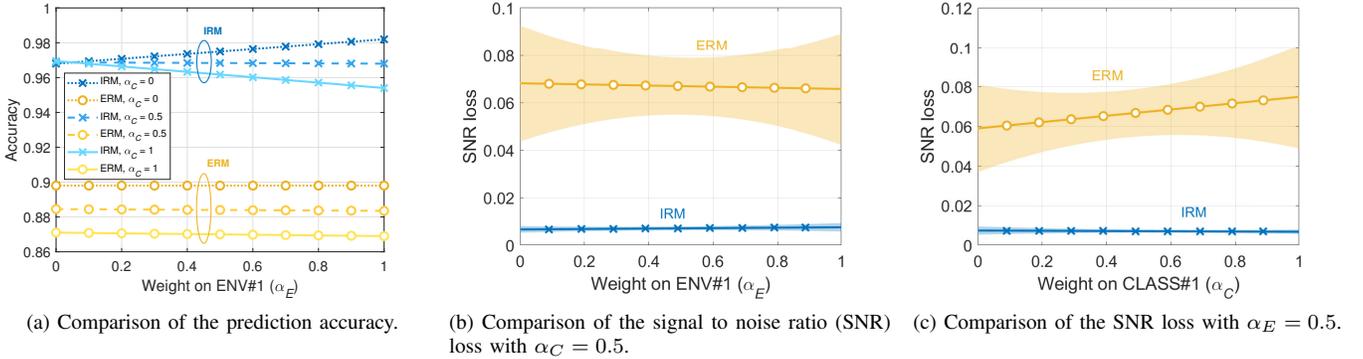


Fig. 3. Performance comparison of ERM and IRM methods for OOD testing datasets based on different choices of  $\alpha_E$  and  $\alpha_C$ . The shaded areas in right and middle figures correspond to the standard deviation of the SNR loss.

ability to infer the phase configuration as BEST where  $\mathbb{I}(\cdot)$  is the indicator function, (ii) *spectral efficiency* indicates the communication performance in terms of  $\log_2(1 + \text{SNR}_m)$ , and (iii) *SNR loss* evaluates the communication degradation with respect to the BEST case that is independent from the choices of  $p$ ,  $N_0$ , and  $W$ , i.e.,  $\text{SNR loss} = 1 - \text{SNR}_m / \text{SNR}_{\text{BEST}}$ .

**Training Convergence and Inference Accuracy.** Within the same environment, ERM predictor achieves higher test accuracy (high generalization) over IRM when equal number of samples from all classes are used for the training as shown in Figs. 2a and 2b. However, testing in a different environment (ENV#3) incurs a huge loss in accuracy under ERM highlights its lack of robustness. In contrast, IRM yields slightly lower accuracy (low generalization) within the same environment under balanced datasets, yet maintains high accuracy (high robustness) over ENV#3 yielding beyond 95% accuracy compared to 80-85% with ERM as shown in Figs. 2a and 2b. When the dataset is biased towards specific classes, accuracy degradation (loss of generalization) is observed in both methods as illustrated in Fig. 2c. However, IRM predictor achieves about 78% accuracy while ERM provides an accuracy of 67%.

**OOD Robustness.** In Fig. 3, both ERM and IRM methods are trained over a dataset with  $\alpha_E = 0.5$  and  $\alpha_C = 0.5$ . During testing, instead of using ENV#3, we used datasets ENV#1 and ENV#2 with different choices of  $\alpha_E$  and  $\alpha_C$ . From Fig. 3a,

it can be noted that changing  $\alpha_E$  and  $\alpha_C$  has a slight impact on IRM method for  $\alpha_C = \{0, 1\}$ . In contrast, the accuracy of ERM is consistent over  $\alpha_E$ , yet much lower compared to IRM due to averaging of training samples. Additionally, the SNR loss comparisons in Figs. 3b and 3c indicate that IRM not only results in about 90.2% lower loss, but also yields about 86.1% lower variance in errors compared to ERM. Compared to IRM, ERM exhibits higher error variance when the testing data deviates from the training data distribution. The error variance when similar distributions are used for training and testing is about  $1.2 \times 10^{-2}$  while it increases to about  $2.4 \times 10^{-2}$  when the distribution is biased towards a single environment, and to about  $2.5 \times 10^{-2}$  under the bias towards a single class. These comparisons further highlight the robustness of the proposed IRM design over the environment-unaware ERM method.

**Spectral Efficiency.** The achievable spectral efficiency in ENV#3 with BEST, IRM, ERM, and RAND methods for different training sample complexities are compared in Fig. 4. Here, balanced training datasets are used, i.e.,  $\alpha_E = 0.5$  and  $\alpha_C = 0.5$ . It can be noted that random choices result in 36.3% loss of spectral efficiency compared to the best case. With training over a small dataset consisting of 60 samples, ERM results in a loss of 29.5% whereas for IRM the loss is 22.7%. As the dataset size is increased to 300 samples, the improvements in ERM results in a loss of 13.4% whereas IRM yields a loss of 0.13%. Such close to the optimal performance

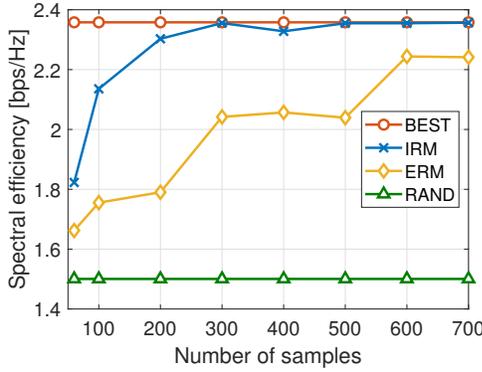


Fig. 4. The impact of training sample complexity on the achievable spectral efficiency for all methods.

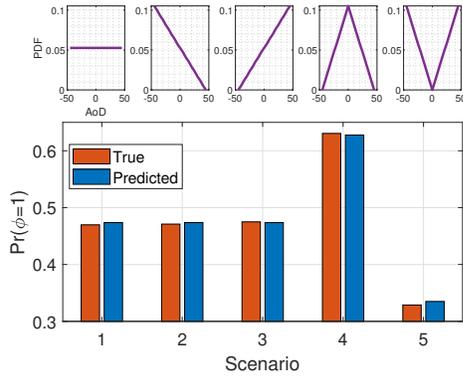


Fig. 5. Validation of IRM's capability of predicting over interventions on AoD at the RIS following the distributions shown at the top.

with the proposed IRM predictor using fewer training samples compared to ERM is because of the learning over an invariant causal representation across multiple environments. Even with the increment of training samples to 700, ERM results in 4.3% loss in terms of achievable spectral efficiency highlighting the lack of robustness using the ERM design.

**Impact of Causal Interventions.** Here, we define five different scenarios by intervening on the probabilistic generation process of the AoD at the RIS similar to Fig. 1c. In this view, we consider interventions that are constraining into 19 regular discrete steps spanning from  $-50^\circ$  to  $50^\circ$  following the distributions  $\{\mathcal{P}_j\}_{j \in \{1, \dots, 5\}}$  as illustrated at the top row of Fig. 5. Note that the IRM model is trained with a dataset consisting of AoD at the RIS (denoted by  $z_1$  hereinafter) following a uniform distribution over  $[-180^\circ, 180^\circ]$ , which we denote by  $\mathcal{P}_0$ . For each intervention  $j$ , we use *do-calculus* [12] to compute the probability of the configuration being CLASS#1, i.e.,  $\phi = c_1$ , by

$$\Pr(\hat{\psi} = c_1) = \sum_{z'} \mathcal{P}_0(\hat{\psi} = c_1 | \text{do}(z_1 = z')) \mathcal{P}_j(z'), \quad (9)$$

where  $\text{do}(z_1 = z')$  represents the intervention taken with probability  $\mathcal{P}_j(z')$  rather than collecting data for the events  $z_1 = z'$  with probability  $\mathcal{P}_j(z')$ . For the numerical calculations, we generate  $N' = N\mathcal{P}_j(z')$  samples and obtain

the RIS configurations using the IRM predictor, after which,  $\mathcal{P}_0(\hat{\psi} = c_1 | \text{do}(z_1 = z')) = \frac{1}{N'} \sum_n \mathbb{I}(\hat{\psi} = c_1)$  is computed. The predicted value is then compared with the simulated outcome of BEST method that uses a dataset generated based on  $\mathcal{P}_j$  and the results are presented in Fig. 5 under the corresponding scenario  $j$ . It can be noted that under the selected five scenarios, the predictions based on IRM coincide with the true probability of obtaining the CLASS#1 under interventions. The maximum deviation occurs with scenario 5 with a difference of  $6.3 \times 10^{-3}$  in the probability underscoring the IRM's potential inference over interventions without the need for data collection or retraining.

## V. CONCLUSIONS

This paper investigates a novel system design for robust RIS based on leveraging the underlying causal structure that is invariant over different environments. The problem is cast as an IRM problem as opposed to environment-unaware ERM approach. Neural network-based phase configuration predictors are trained in a supervised learning manner and evaluated over different training and testing environments. The results indicate that the proposed IRM-based predictor is robust over ERM design across different environments and generalizes out-of-distribution. Considering multiple transmitters, receivers, and RISs with several antennas are interesting future extensions.

## REFERENCES

- [1] J. He, H. Wymeersch, L. Kong, O. Silvén, and M. Juntti, "Large intelligent surface for positioning in millimeter wave MIMO systems," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. IEEE, 2020, pp. 1–5.
- [2] Ö. Özdogan and E. Björnson, "Deep learning-based phase reconfiguration for intelligent reflecting surfaces," *preprint arXiv:2009.13988*, 2020.
- [3] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," *arXiv preprint arXiv:1912.03619*, 2019.
- [4] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Unsupervised learning for passive beamforming," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1052–1056, 2020.
- [5] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in *Proc. of IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.
- [6] J. Park, S. Samarakoon, H. Shiri, M. K. Abdel-Aziz, T. Nishio, A. Elgabri, and M. Bennis, "Extreme urllc: Vision, challenges, and key enablers," *arXiv preprint arXiv:2001.09683*, 2020.
- [7] B. Sheen, J. Yang, X. Feng, and M. M. U. Chowdhury, "A deep learning based modeling of reconfigurable intelligent surface assisted wireless communications for phase shift configuration," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 262–272, 2021.
- [8] G. C. Alexandropoulos, S. Samarakoon, M. Bennis, and M. Debbah, "Phase configuration learning in wireless networks with multiple reconfigurable intelligent surfaces," in *Proc. of 2020 IEEE Globecom Workshops*, 2020, pp. 1–6.
- [9] E. Basar, I. Yildirim, and I. F. Akyildiz, "Indoor and outdoor physical channel modeling and efficient positioning for reconfigurable intelligent surfaces in mmWave bands," *arXiv preprint arXiv:2006.02240*, 2020.
- [10] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz, "Invariant risk minimization," *arXiv preprint arXiv:1907.02893*, 2019.
- [11] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International Conference on Machine Learning*. PMLR, 2017, pp. 1126–1135.
- [12] R. R. Tucci, "Introduction to judea pearl's do-calculus," *arXiv preprint arXiv:1305.5506*, 2013.