# Active Sensing for Localization with Reconfigurable Intelligent Surface

Zhongze Zhang, Tao Jiang, and Wei Yu

Department of Electrical and Computer Engineering, University of Toronto

{ufo.zhang,taoca.jiang}@mail.utoronto.ca, weiyu@ece.utoronto.ca

*Abstract*—This paper addresses an uplink localization problem in which the base station (BS) aims to locate a remote user with the aid of reconfigurable intelligent surface (RIS). This paper proposes a strategy in which the user transmits pilots over multiple time frames, and the BS adaptively adjusts the RIS reflection coefficients based on the observations already received so far in order to produce an accurate estimate of the user location at the end. This is a challenging active sensing problem for which finding an optimal solution involves a search through a complicated functional space whose dimension increases with the number of measurements. In this paper, we show that the long short-term memory (LSTM) network can be used to exploit the latent temporal correlation between measurements to automatically construct scalable information vectors (called hidden state) based on the measurements. Subsequently, the state vector can be mapped to the RIS configuration for the next time frame in a codebook-free fashion via a deep neural network (DNN). After all the measurements have been received, a final DNN can be used to map the LSTM cell state to the estimated user equipment (UE) position. Numerical result shows that the proposed active RIS design results in lower localization error as compared to existing active and nonactive methods. The proposed solution produces interpretable results and is generalizable to early stopping in the sequence of sensing stages.

## I. INTRODUCTION

Reconfigurable intelligent surface (RIS) is a planar surface consisting of a large number of passive elements, with each element capable of altering the phase of the incident electromagnetic wave with very low power consumption [1]. The technology is envisioned as one of the enabling technologies for 6G localization due to its ability to: i) establish a reliable reflected link when the direct path between the transceivers is weak or blocked, and ii) provide new anchor points without needing expensive RF chains [2], [3]. In the area of RIS-aided localization, fruitful research progress has been made in the direction of theoretical bounds and algorithms [4]–[11].

This paper considers an uplink localization problem in the presence of RIS, where a single user repeatedly transmits pilot symbols and the base station (BS) receives the pilots through the reflection at the RIS and determines the location of the user based on the received pilots. Specifically, this paper investigates a scenario in which the BS can control the reflection coefficients at the RIS and tackles the problem of designing the RIS configuration in an *active* manner to minimize localization error. By active sensing, we mean that the RIS reflection coefficients are sequentially designed as a function of previous measurements. As a result, the RIS can be used to focus the beam progressively to locate the user

over time as more measurements become available. This is however a challenging problem as a solution needs to be able to optimize over the complicated functional landscape from the sequence of observations to the RIS reflection coefficients, while being scalable in the number of measurements.

In the context of localization algorithms in RIS-assisted network, the design of RIS configuration to enhance localization accuracy is of great interest. The authors in [5] devise three classes of RIS profiles and evaluated the impact of each class on the 3D position error bound of the user equipment (UE). Two of the three classes are heuristically designed to produce narrow/broad beams, and the remaining class consists of random RIS profiles. In [6], [7], the indoor localization accuracy is improved by designing a set of RIS profiles that enlarges the differences between the received signal strength (RSS) values of adjacent locations. The set of RIS profiles is designed via local and global search methods. The authors in [8] consider joint localization and communication optimization problem and design the appropriate RIS profiles to improve localization accuracy and transmission throughput. In [9], the authors consider a fingerprinting localization problem enabled by an RIS. In particular, the authors employ machine learning to identify a subset of RIS profiles from a codebook to generate a more diverse set of fingerprints to improve fingerprint matching accuracy. All of the above works have demonstrated the importance of designing/selecting an appropriate set of RIS configurations and the associated performance gain. However, *active* design of RIS configuration, which could further improve the performance, has not been considered.

In this paper, we focus on the active design of RIS reflection coefficients based on the sequential historical measurements of the environment to improve the localization accuracy. Designing such an adaptive sequence of RIS patterns is a challenging problem. To make the problem tractable, the authors in [10] design a hierarchical codebook for the RIS that enables adaptive bisection search over the angular space in a 2D localization setting. However, a hierarchical codebook-based method is not necessarily the best approach. For example, in the context of angle of arrival (AoA) estimation problem, the performance of the hierarchical codebook-based method is severely restricted by the quality of the codebook and is far from optimal [12]. In another work on active sensing with RIS [11], the authors use gradient descend method to optimize RIS configuration with a goal of minimizing the Cramér-Rao lower bound (CRLB) in each step. However, the method is based on

first estimating the fading coefficients and AoAs instead of the location, which results in significant pilot overhead.

This paper proposes a learning-based localization solution, where a sequence of RIS configurations is adaptively designed based on the historical measurements in a codebook-free fashion. In particular, we use the long short-term memory (LSTM) based network, for its ability to capture the temporal relationship between different measurements over a long period [13]. We use a chain of LSTM cells, corresponding to the sequence of measurements. At each measurement time frame, an LSTM cell accepts new measurement of the environments, i.e., received pilots, and automatically uses the newly received measurement (along with historical measurements) to update a hidden state vector of fixed dimension. Subsequently, a deep neural network (DNN) is used to extract a correct representation of the knowledge of location so far from the hidden state to design the RIS profile for the next time frame. After multiple measurements, the cell state of the final LSTM cell is passed through another DNN to obtain the estimated UE position. Numerical result shows that the proposed active RIS profile design achieves lower localization error as compared to the existing nonactive and active RIS profile design.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

We consider a localization problem in an RIS-assisted system with a single-antenna BS, a single-antenna UE, and a planar RIS. The BS and RIS are placed as in Fig. 1 to localize the potential users in the area. We assume that the position of the BS and the RIS are known. By adopting the Cartesian coordinate system, let $\boldsymbol{p}^{\mathrm{BS}} = [x^{\mathrm{BS}}, y^{\mathrm{BS}}, z^{\mathrm{BS}}]^\top$ and $\boldsymbol{p}^{\mathrm{RIS}} = [x^{\mathrm{RIS}}, y^{\mathrm{RIS}}, z^{\mathrm{RIS}}]^\top$ denote the position of the BS and the RIS respectively. The unknown UE position is denoted as $\boldsymbol{p} = [x, y, z]^\top$.

The reflection coefficients of the RIS are controlled by an RIS controller which receives controlling signals from the BS. Let $N$ be the number of reflection coefficients at the RIS. Then the RIS reflection coefficient is denoted as

$$\boldsymbol{\theta} = [e^{j\delta_1}, e^{j\delta_2}, \cdots, e^{j\delta_N}]^\top \in \mathbb{C}^N, \tag{1}$$

with $\delta_n \in [0, 2\pi)$ as the phase shift of the $n$-th element.

We adopt a *block-fading* model in which the channels are assumed to be constant across multiple time frames within a coherence period, then change independently in subsequent coherence periods. As shown in Fig. 1, $h_{\mathrm{d}} \in \mathbb{C}$ denotes the direct channel from the BS to the UE, $\boldsymbol{h}_{\mathrm{r}} \in \mathbb{C}^N$ denotes the reflection channel from the RIS to the UE, and $\boldsymbol{g}_{\mathrm{r}}^\top \in \mathbb{C}^N$ denotes the channel from the BS to the RIS. We further assume that the reflection channel $\boldsymbol{h}_{\mathrm{r}}, \boldsymbol{g}_{\mathrm{r}}$ and the direct channel $h_{\mathrm{d}}$ follow Rician fading model:

$$\boldsymbol{h}_{\mathrm{r}} = \kappa \left( \sqrt{\epsilon/(1+\epsilon)}\tilde{\boldsymbol{h}}_{\mathrm{r}}^{\mathrm{LOS}} + \sqrt{1/(1+\epsilon)}\tilde{\boldsymbol{h}}_{\mathrm{r}}^{\mathrm{NLOS}} \right), \tag{2a}$$

$$\boldsymbol{g}_{\mathrm{r}} = \xi \left( \sqrt{\epsilon/(1+\epsilon)}\tilde{\boldsymbol{g}}_{\mathrm{r}}^{\mathrm{LOS}} + \sqrt{1/(1+\epsilon)}\tilde{\boldsymbol{g}}_{\mathrm{r}}^{\mathrm{NLOS}} \right), \tag{2b}$$

$$h_{\mathrm{d}} = \rho \left( \sqrt{\epsilon/(1+\epsilon)}\tilde{h}_{\mathrm{d}}^{\mathrm{LOS}} + \sqrt{1/(1+\epsilon)}\tilde{h}_{\mathrm{d}}^{\mathrm{NLOS}} \right), \tag{2c}$$
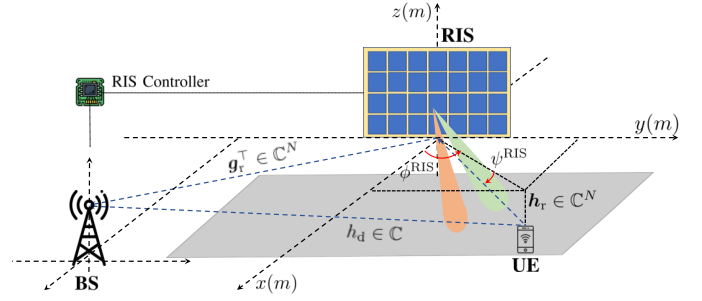


Fig. 1: RIS-assisted network.

where $\rho$ denotes the pathloss between the BS and the UE, and $\kappa$ and $\xi$ denote the path losses between the RIS and the UE/BS. Here, $\tilde{h}_{\mathrm{d}}^{\mathrm{NLOS}}$, $\tilde{\boldsymbol{h}}_{\mathrm{r}}^{\mathrm{NLOS}}$ and $\tilde{\boldsymbol{g}}_{\mathrm{r}}^{\mathrm{NLOS}}$ denote the non-line-of-sight components and their entries are generated independently according to $\mathcal{CN}(0,1)$. The line-of-sight component of the reflection channel contains angular information about the location of UE. For example, $\tilde{\boldsymbol{h}}_{\mathrm{r}}^{\mathrm{LOS}}$ is a function of UE location and RIS location. Let $\phi^{\mathrm{RIS}}$ and $\psi^{\mathrm{RIS}}$ denote the azimuth and elevation angle of arrival from the UE to the RIS. The line-of-sight component of $\boldsymbol{h}_{\mathrm{r}}$ can be written as

$$\tilde{\boldsymbol{h}}_{\mathrm{r}}^{\mathrm{LOS}} = \boldsymbol{a}^{\mathrm{RIS}}(\phi^{\mathrm{RIS}}, \psi^{\mathrm{RIS}}), \tag{3}$$

where the steering vector of the $n$-th element of the RIS can be expressed as [14]

$$[\boldsymbol{a}^{\mathrm{RIS}}(\phi^{\mathrm{RIS}}, \psi^{\mathrm{RIS}})]_n$$
$$= e^{j\frac{2\pi d_{\mathrm{R}}}{\lambda_c} \{v_1(n,N_c)\sin(\phi^{\mathrm{RIS}})\cos(\psi^{\mathrm{RIS}})+v_2(n,N_c)\sin(\psi^{\mathrm{RIS}})\}}, \tag{4}$$

where $d_{\mathrm{R}}$ is the distance between two reflective elements of the RIS, $\lambda_c$ is the carrier wavelength, and $v_1(n, N_c) = \mathrm{mod}(n - 1, N_c)$ and $v_2(n, N_c) = \left\lfloor \frac{n-1}{N_c} \right\rfloor$. Here, $N_c$ is the number of columns of the RIS.

Recall that $\boldsymbol{p}^{\mathrm{RIS}} = [x^{\mathrm{RIS}}, y^{\mathrm{RIS}}, z^{\mathrm{RIS}}]^\top$ denote the position of the RIS and $\boldsymbol{p} = [x, y, z]^\top$ denote the position of the UE, we have the following [14]:

$$\sin(\phi^{\mathrm{RIS}})\cos(\psi^{\mathrm{RIS}}) = (y - y^{\mathrm{RIS}})/d, \tag{5a}$$

$$\sin(\psi^{\mathrm{RIS}}) = (z - z^{\mathrm{RIS}})/d, \tag{5b}$$

where $d$ is the distance between the RIS and the UE.

Let $\eta^{\mathrm{RIS}}$ and $\vartheta^{\mathrm{RIS}}$ denote the azimuth and elevation angle of departures (AoD) from the RIS to the BS. Since the BS only has a single antenna, the line-of-sight component of $\boldsymbol{g}_{\mathrm{r}}$ is given by

$$\tilde{\boldsymbol{g}}_{\mathrm{r}}^{\mathrm{LOS}} = \boldsymbol{a}^{\mathrm{RIS}}(\eta^{\mathrm{RIS}}, \vartheta^{\mathrm{RIS}})^{\mathrm{H}}, \tag{6}$$

where the line-of-sight component of $h_{\mathrm{d}}$ is given by

$$\tilde{h}_{\mathrm{d}}^{\mathrm{LOS}} = 1. \tag{7}$$

### B. Signal Transmission

When there is a localization request, the UE sends a sequence of $T$ uplink pilot symbols to the BS over $T$ time frames. Let $x_t \in \mathbb{C}$ be the pilot symbol to be transmitted from

the UE to the BS at the $t$-th time frame. The BS receives a combination of the signal from the direct path and the signal reflected off the RIS, so the received pilots at the BS can be expressed as

$$y_t(\boldsymbol{\theta}_t) = \sqrt{P_u}(h_{\mathrm{d}} + \boldsymbol{v}_{\mathrm{r}}^\top \boldsymbol{\theta}_t)x_t + n_t, \quad t = 0, \cdots, T-1, \quad (8)$$

where $P_u$ is the uplink transmission power, $\boldsymbol{v}_{\mathrm{r}} = \mathrm{diag}(\boldsymbol{h}_{\mathrm{r}})\boldsymbol{g}_{\mathrm{r}}^\top \in \mathbb{C}^N$ is the cascade channel between the BS and the UE through the reflection at the RIS, and $n_t \sim \mathcal{CN}(0, \sigma_u^2)$ is the uplink additive white Gaussian noise. Here, the received pilot is a function of the RIS configuration at the $t$-th time frame $\boldsymbol{\theta}_t$, where the reflection coefficients can be configured randomly or designed according to some criteria to enable a better localization outcome.

### C. Problem Formulation

The goal of the localization problem is to estimate the unknown UE position $\boldsymbol{p}$ based on the $T$ observations $[y^{(t)}]_{t=0}^{T-1}$, and the known BS and RIS positions. The design of RIS configuration is of critical interest here. It has been shown in [6], [7], [9] that by strategically selecting a fixed set of RIS configurations, a more favourable RSS distribution can be obtained, which allows improvement in localization accuracy. However, most existing works are based on the *fixed* sensing framework, which passively collects all the observations of the environment according to a fixed set of RIS configurations. In essence, the RIS probes the search area using fixed beams along multiple random directions. Here, we instead propose an *active* sensing framework to gradually narrow down the searching area using more directional beams.

Specifically, we consider the following active localization setup. In the $t$-th time frame, the BS designs the next RIS configuration $\boldsymbol{\theta}_{t+1}$ based on the existing observations, which is used to make the next measurement $y_{t+1}$ in the $(t+1)$-th time frame. Thus, the design of RIS configuration is a function of historical measurements:

$$\boldsymbol{\theta}_{t+1} = \mathcal{G}_t(\{y_\tau\}_{\tau=0}^t), \quad t = 0, \cdots, T-1. \quad (9)$$

As no prior observation exists when $t < 0$, the first set of RIS configurations $\boldsymbol{\theta}_0$ is produced via function $\mathcal{G}_{-1}(\emptyset)$. The function accepts an empty set as input and always produces the same initialization of RIS configuration.

The estimated UE position $\hat{\boldsymbol{p}}$ can be written as a function of all $T$ historical observations.

$$\hat{\boldsymbol{p}} = \mathcal{F}(\{y_t\}_{t=0}^{T-1}). \quad (10)$$

The location estimation problem can now be formulated as:

$$\min_{\{\mathcal{G}_t(\cdot)\}_{t=0}^{T-1}, \mathcal{F}(\cdot)} \mathbb{E}\left[\|\hat{\boldsymbol{p}} - \boldsymbol{p}\|_2^2\right] \quad (11a)$$

$$\text{subject to} \quad |[\boldsymbol{\theta}_t]_n| = 1, \ \forall n, t, \quad (11b)$$

$$\boldsymbol{\theta}_{t+1} = \mathcal{G}_t(\{y_\tau\}_{\tau=0}^t), \ t = 0, \cdots, T-1, \quad (11c)$$

$$\hat{\boldsymbol{p}} = \mathcal{F}(\{y_t\}_{t=0}^{T-1}). \quad (11d)$$

Solving the optimization problem (11) is challenging, as the problem amounts to optimizing the function expression
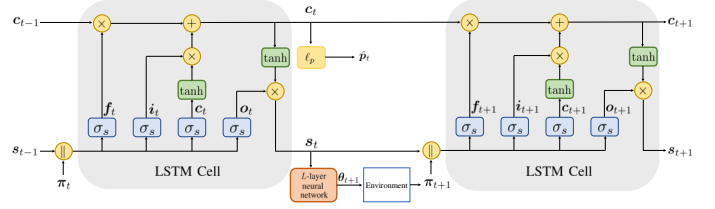


Fig. 2: Proposed active localization framework.

$\{\mathcal{G}_t(\cdot)\}_{t=0}^{T-1}$ in (9) and $\mathcal{F}(\cdot)$ in (10). To make the problem more tractable, a common approach is to select the RIS configuration adaptively from a predefined set of RIS patterns, namely the *codebook*, based on heuristics. For example, [10] considers a 2D RIS-assisted localization problem, where the sequence of RIS patterns is heuristically selected from hierarchical codebooks, based on feedback from the UE. However, a codebook-based approach is not ideal as the freedom of designing RIS configuration is restricted.

In this work, we propose to employ a neural network to parameterize the function $\mathcal{F}(\cdot)$ and $\mathcal{G}_t(\cdot)$ to adaptively design a sequence of RIS configurations, without restricting the design space of RIS coefficients to a pre-defined codebook.

## III. PROPOSED ACTIVE LOCALIZATION FRAMEWORK

In this section, we introduce the deep learning approach to solving the active localization problem discussed in (11). Here in problem (11), the next RIS configuration $\boldsymbol{\theta}_{t+1}$ is designed based on historical measurements up to $t$, i.e., $\{y_\tau\}_{\tau=0}^t$. The main design challenge is that as the dimension of historical observations increases linearly with $t$, it is difficult to use the entire history of measurements to design the next RIS configuration for large $t$. Thus, the key design question for the neural network is how to extract useful information from the historical measurements and how to map it to an information vector of fixed dimension.

In this paper, we develop an LSTM network capable of automatically constructing information vector of fixed dimension from existing measurements and extracting temporal features and long-term dependencies from a sequence of temporal input. The specific LSTM network developed here is similar to the one in [15] in which the information vector is the hidden state, but with modifications to the neural network architecture tailored to the localization problem in an RIS-assisted network. We also propose a new loss function to make the LSTM network generalizable to the number of time frames.

### A. Neural Network Architecture

The overall neural network architecture is shown in Fig. 2. At the $t$-th time frame, an LSTM cell takes new feature as input to update the hidden state vector $\boldsymbol{s}_t$ and the cell state vector $\boldsymbol{c}_t$. The new feature, denoted as $\boldsymbol{\pi}_t$, can be the RSS $|y_t|^2$, or the real and imaginary component of received pilots $[\mathcal{R}(y_t), \mathcal{I}(y_t)]$, depending on the constraints of the detecting

hardware at the receiver. We use $\circ$ to denote the element-wise product. The updating rules of the state vectors are as follows:

$$\boldsymbol{c}_t = \boldsymbol{f}_t \circ \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \circ \tanh\left(\boldsymbol{u}_c \boldsymbol{\pi}_t + \boldsymbol{w}_c \boldsymbol{s}_{t-1}\right), \qquad (12a)$$

$$\boldsymbol{s}_t = \boldsymbol{o}_t \circ \tanh(\boldsymbol{c}_t), \qquad (12b)$$

where $\boldsymbol{u}_c$ and $\boldsymbol{w}_c$ are linear layers with a number of fully connected units. Here, $\boldsymbol{f}_t$, $\boldsymbol{i}_t$, and $\boldsymbol{o}_t$ are the activation vectors of the forget gate, input gate and output gate within the LSTM cell respectively. The element-wise updating rules of the different gates are as follows:

$$\boldsymbol{f}_t = \sigma_s\left(\boldsymbol{u}_f \boldsymbol{\pi}_t + \boldsymbol{w}_f \boldsymbol{s}_{t-1}\right), \qquad (13a)$$

$$\boldsymbol{i}_t = \sigma_s\left(\boldsymbol{u}_i \boldsymbol{\pi}_t + \boldsymbol{w}_i \boldsymbol{s}_{t-1}\right), \qquad (13b)$$

$$\boldsymbol{o}_t = \sigma_s\left(\boldsymbol{u}_o \boldsymbol{\pi}_t + \boldsymbol{w}_o \boldsymbol{s}_{t-1}\right), \qquad (13c)$$

where $\sigma_s(x) = 1/(1 + e^{-x})$ is the sigmoid function, and $\boldsymbol{u}_f$, $\boldsymbol{u}_i$, $\boldsymbol{u}_o$, $\boldsymbol{w}_f$, $\boldsymbol{w}_s$ and $\boldsymbol{w}_o$ are linear layers. As per convention, the initial value of the cell state $\boldsymbol{c}_0$ is obtained by setting $\boldsymbol{s}_{-1} = \boldsymbol{c}_{-1} = \boldsymbol{0}$, and $y_0 = 1$.

We design the RIS configuration for time frame $t+1$ based on the hidden state vector $\boldsymbol{s}_t$. The hidden state vector is used as the input to a fully connected neural network of $L$ layers to design the RIS configuration for the next time frame,

$$\bar{\boldsymbol{\theta}}_{t+1} = \beta_L(\boldsymbol{A}_L \beta_{L-1}(\cdots \beta_1(\boldsymbol{A}_1 \boldsymbol{s}_t + \boldsymbol{b}_1)\cdots) + \boldsymbol{b}_L), \quad (14)$$

where $\beta_l$, $l \in \{1, \cdots, L\}$ is the activation function of the $l$-th layer, i.e., $\text{relu}(x) = \max(0, x)$, $\{\boldsymbol{A}_l\}_{l=1}^L$ and $\{\boldsymbol{b}_l\}_{l=1}^L$ are sets of trainable weights and biases. Here, $\bar{\boldsymbol{\theta}}_{t+1}$ contains the real and imaginary components of the RIS reflection coefficients

$$\bar{\boldsymbol{\theta}}_{t+1} = [\mathcal{R}(\boldsymbol{\theta}_{t+1})^\top, \mathcal{I}(\boldsymbol{\theta}_{t+1})^\top]^\top. \qquad (15)$$

The dimensions of $\boldsymbol{A}_L$ and $\boldsymbol{b}_L$ are designed to ensure the output is of correct dimension, such that $\bar{\boldsymbol{\theta}}_{t+1} \in \mathbb{R}^{2N}$. To enforce unit modulus constraint on each element of the RIS, an element-wise normalization is performed

$$[\boldsymbol{\theta}_{t+1}]_n = [\mathcal{R}(\boldsymbol{\theta}_{t+1})]_n / \sqrt{[\mathcal{R}(\boldsymbol{\theta}_{t+1})]_n^2 + [\mathcal{I}(\boldsymbol{\theta}_{t+1})]_n^2}$$
$$+ j[\mathcal{I}(\boldsymbol{\theta}_{t+1})]_n / \sqrt{[\mathcal{R}(\boldsymbol{\theta}_{t+1})]_n^2 + [\mathcal{I}(\boldsymbol{\theta}_{t+1})]_n^2}. \quad (16)$$

While the hidden state vector $\boldsymbol{s}_t$ is used to design the RIS configuration for the next time frame, the cell state vector $\boldsymbol{c}_T$ is used to obtain the estimated UE position at the $T$-th time frame. After $T$ time frames, the final estimated UE position $\hat{\boldsymbol{p}}_T$ is obtained through a fully connected neural network, based on the final cell state $\boldsymbol{c}_T$:

$$\hat{\boldsymbol{p}}_T = \boldsymbol{\ell}_p \boldsymbol{c}_T, \qquad (17)$$

where $\boldsymbol{\ell}_p$ denotes a neural network with linear layers. Here, we find that a linear network already works well as the final DNN.

### B. Loss Functions

To train the LSTM network, we employ Adam optimizer [16] to minimize the average mean squared error between the estimated position $\hat{\boldsymbol{p}}_T$ and the true position as follows

$$\mathbb{E}\left[\|\hat{\boldsymbol{p}}_T - \boldsymbol{p}\|_2^2\right]. \qquad (18)$$

TABLE I: Parameters of the LSTM Network.

| Label | | | Dimension | | |
|---|---|---|---|---|---|
| $\boldsymbol{u}_c, \boldsymbol{u}_f, \boldsymbol{u}_i, \boldsymbol{u}_o$ $\boldsymbol{w}_c, \boldsymbol{w}_f, \boldsymbol{w}_i, \boldsymbol{w}_o$ | | | 512 | | |
| $\boldsymbol{A}_1$ | $\{\boldsymbol{A}_l\}_{l=2}^{L-1}$ | $\boldsymbol{A}_L$ | $512 \times 1024$ | $1024 \times 1024$ | $1024 \times 2N$ |
| $\{\boldsymbol{b}_l\}_{l=1}^{L-1}$ | | $\boldsymbol{b}_L$ | 1024 | | $2N$ |
| $\boldsymbol{\ell}_p$ | | | $512 \times 3$ | | |

Here, the choice of loss function in (18) encourages the LSTM network to design a series of $T$ RIS configurations to minimize localization error at the final time frame. We note that this loss function only accounts for the estimation error at the final stage. This is a good choice since it gives the neural network freedom to design the sensing strategy across the entire $T$ measurement stages.

In some cases, we may need to have earlier stopping in the sequence of sensing stages, e.g., due to some latency constraints. This requires the neural network to output the best estimates of the location before the final stage, but this is not considered in the loss function (18). For this scenario, we propose an alternative loss function as follows:

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\|\hat{\boldsymbol{p}}_t - \boldsymbol{p}\|_2^2\right], \qquad (19)$$

where $\hat{\boldsymbol{p}}_t = \boldsymbol{\ell}_p \boldsymbol{c}_t$. By minimizing the average localization error across $t$, we can encourage the LSTM network to design RIS reflection coefficients to reduce estimation error at each time frame. The new loss function in (19) helps the learning framework to achieve good generalizability with respect to the number of stages.

### IV. NUMERICAL RESULTS

#### A. Simulation Environment

In simulations, the BS is located at $\boldsymbol{p}^{\text{BS}} = (40m, -40m, -10m)$ and an $8 \times 8$ RIS is located at $\boldsymbol{p}^{\text{RIS}} = (0m, 0m, 0m)$. The unknown user locations $\boldsymbol{p}$ are uniformly generated within a rectangular area on the $x$-$y$ plane $(20 \pm 15m, 0 \pm 35m, -20m)$. In subsequent simulation, the Rician factor $\epsilon$ is set to 10, and $2\pi d_R/\lambda_c = 1$ without loss of generality. The path-loss models of the direct and reflected paths are $32.6 + 36.7\log(d_1)$ and $30 + 22\log(d_2)$, respectively, where $d_1$ and $d_2$ denote the corresponding link distance.

#### B. Baseline Schemes

The proposed LSTM network for adaptive RIS-assisted localization is implemented using parameters in Table I. We set $L = 4$ and develop the model using Tensorflow [17]. The model samples a total of 2,048,000 training data in 2000 epochs. We compare the localization performance of the proposed algorithm against the following baseline schemes.

*Fingerprint based localization with random RIS configurations* [9]: In this scheme, the sequence of $T$ uplink RIS configurations is non-adaptive and random. Every $1m \times 1m$ block within the $30m \times 70m$ area is associated with a
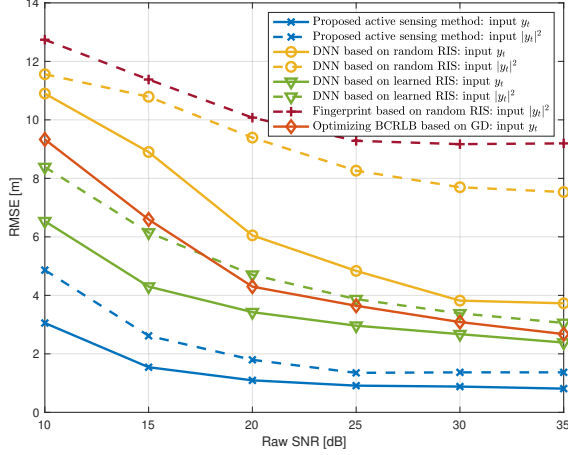
Fig. 3: Localization accuracy vs. raw SNR, $N = 64$, $T = 6$.



Fig. 4: Localization accuracy vs. pilot length, $N = 64$, SNR= 20dB.

vector of RSS as its fingerprint, i.e., $\left[|y_1|^2, \cdots, |y_T|^2\right]$. The fingerprints are stored in an offline database. A weighted $k$-nearest neighbour (wKNN) classifier is used to match online and offline measurements. Here, we set $k = 5$.

*DNN with random or learned RIS configurations*: The sequence of $T$ uplink RIS configurations is non-adaptive. Here, we can have two different possibilities: i) the RIS configurations are randomly chosen, or ii) the sequence of RIS configurations is learned from training data, but are not adaptive as a function of previous measurements. A fully connected neural network is used to map the received pilot symbols over $T$ time frames $\{\mathcal{R}(y_t), \mathcal{I}(y_t)\}_{t=0}^{T-1}$ or RSS values over $T$ time frames to an estimated UE position. The dimension of the neural network is $[200, 200, 200, 3]$.

*Optimizing CRLB using gradient descent (GD)* [11]: Here, we test the idea of designing an active sensing strategy based on minimizing CRLB in every time step. This is an adaptive scheme based on estimation theory rather than machine learning. In each time frame, based on received measurements, we use maximum a posterior (MAP) estimator to estimate the UE location, which is used to update the CRLB of the location error. Here, the CRLB is calculated by replacing the unknown true location of UE with the estimated one. Subsequently, the RIS reflection coefficients are designed using the GD algorithm to minimize the approximated CRLB for the next pilot transmission.

### C. Simulation Results

We first examine the localization performance versus raw SNR, i.e., $P_u = 10^{\text{SNR}/10}$, when $T = 6$. In Fig. 3, whether the input is the pilot symbols or RSS value, the proposed adaptive RIS design is seen to have the highest localization accuracy across different SNRs as compared to other benchmarks with non-adaptive RIS design. This implies that the proposed algorithm is effectively utilizing the current and historical measurements to design a more suitable RIS configuration for future time frames to minimize localization error.
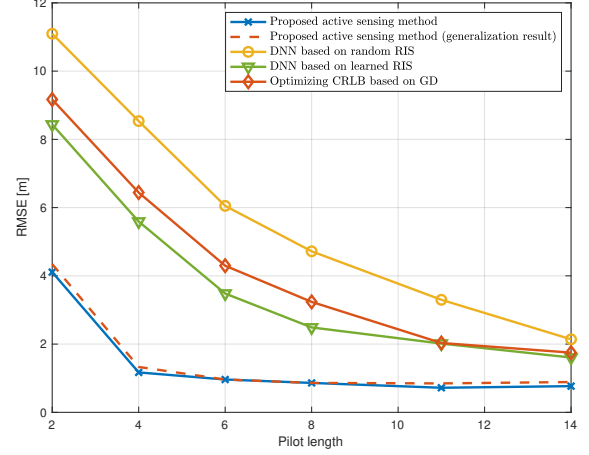
Here, we point out that CRLB-minimization based adaptive RIS design does not serve as an optimal RIS design in reducing the location MSE for the following reasons: i) an active sensing solution to (11) should design a series of sensing configuration jointly across $t$ to minimize the location MSE, i.e., $\{\mathcal{G}_t(\cdot)\}_{t=0}^{T-1}$ as in (11). However, the CRLB-based approach greedily treats the active sensing problem by designing individual $\mathcal{G}_t(\cdot)$ at each time frame, ii) the CRLB can be a loose lower bound of the MSE, especially when SNR is low and/or the number of observations is limited [18], [19] iii) using the estimated UE location in computing the CRLB can introduce errors, iv) finally, the optimization of CRLB is a nonconvex optimization problem; it is difficult to find its true optimal solution.

We point out that the fingerprinting based approach along with wKNN classifier is seen to experience poor performance in this simulation setup. This is due to the randomness in the non-line-of sight Rician channel model which adversely influences the fingerprint matching accuracy.

We next examine the localization performance with varying numbers of time frames. From Fig. 4, we observe that the proposed algorithm demonstrates a robust performance over other benchmarks. Moreover, the LSTM based neural network with the proposed training algorithm generalizes well to early stopping in the sequence of time frames. The model is trained to adaptively design RIS configurations to minimize loss function (19) when $T = 14$. We test the trained model when the number of time frames is fewer than $T$ without retraining, and find the localization accuracy to be as good as if the LSTM is trained for each specific $T$.

### D. Interpretation

We visually interpret the RIS design obtained from the proposed LSTM based network. Here, we test the neural network for one single user and use the RSS distribution (or radio map) as a means to illustrate the beamforming pattern produced by the adaptively designed RIS reflection
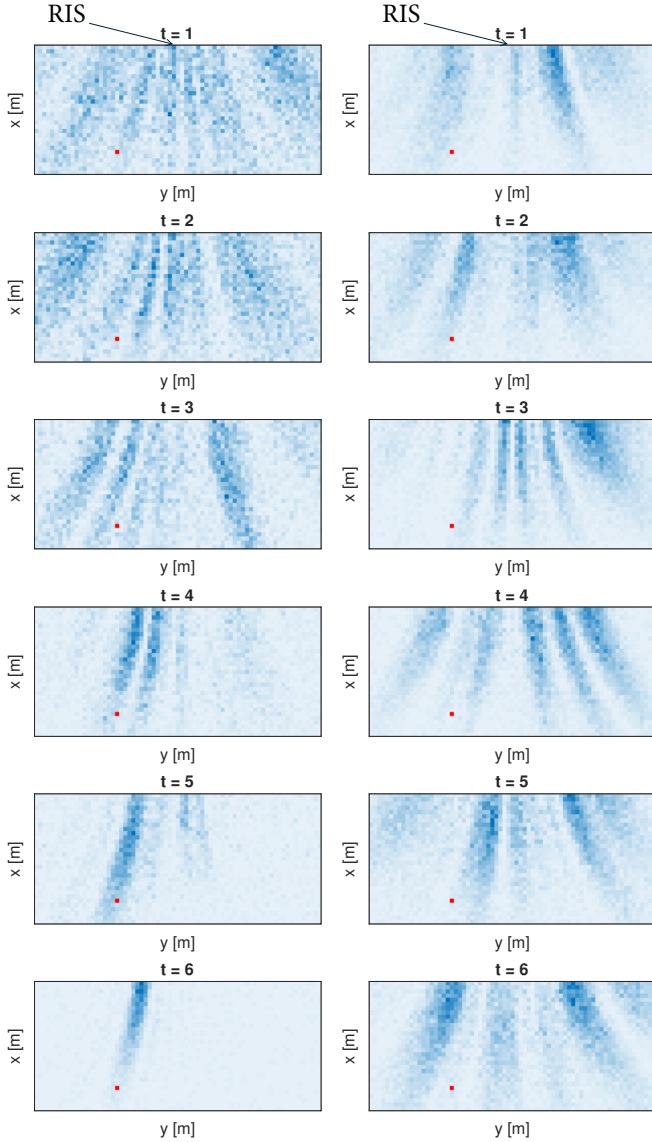
(a) The proposed active sensing method.   (b) Learned RIS (non-active).

Fig. 5: RSS distribution from time frame 1 to 6, $T = 6$, SNR = 25dB.

coefficients. To do so, at each time frame, we record the designed RIS configuration and plot the RSS obtained at every $1m \times 1m$ block in the area of interest across the $x - y$ plane as shown in Fig. 5. The red dot denotes the true UE position. From Fig. 5(a), it is clear that the RIS is probing broader beams at the first several time frames to search for the user, then gradually focusing the beam towards the UE as $t$ increases. This implies that the proposed neural network is indeed designing meaningful RIS configurations based on the measurements.

## V. CONCLUSION

This paper shows that active sensing can significantly improve the performance of an uplink RIS-assisted localization task in which the BS adaptively designs the sequence of RIS reflection coefficients based on the received pilots from the user so far to enhance the localization accuracy. By employing an LSTM based neural network, the proposed solution can effectively utilize historical measurements to design RIS configurations for subsequent measurements in a codeboook-free fashion for the purpose of minimizing localization error. Numerical results indicate a lower localization error as compared to existing active and nonactive benchmarks. The proposed solution demonstrates interpretable results and is generalizable to early stopping in the sequence of sensing stages.

## REFERENCES

[1] E. Basar, M. Di Renzo, J. de Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Aug. 2019.

[2] A. Bourdoux *et al.*, "6G white paper on localization and sensing," 2020. [Online]. Available: https://arxiv.org/abs/2006.01779

[3] C. Pan *et al.*, "An overview of signal processing techniques for RIS/IRS-aided wireless systems," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1–35, Aug. 2022.

[4] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, and H. Wymeersch, "SISO RIS-enabled joint 3D downlink localization and synchronization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[5] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, "Near-field localization with a reconfigurable intelligent surface acting as lens," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[6] H. Zhang, H. Zhang, B. Di, K. Bian, Z. Han, and L. Song, "Towards ubiquitous positioning by leveraging reconfigurable intelligent surface," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 284–288, Sept. 2021.

[7] ——, "Metalocalization: Reconfigurable intelligent surface aided multi-user wireless indoor localization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7743–7757, Jun. 2021.

[8] R. Wang, Z. Xing, and E. Liu, "Joint location and communication study for intelligent reflecting surface aided wireless communication system," 2021. [Online]. Available: https://arxiv.org/abs/2103.01063

[9] C. L. Nguyen, O. Georgiou, G. Gradoni, and M. Di Renzo, "Wireless fingerprinting localization in smart environments using reconfigurable intelligent surfaces," *IEEE Access*, vol. PP, pp. 1–1, Sept. 2021.

[10] J. He, H. Wymeersch, T. Sanguanpuak, O. Silven, and M. Juntti, "Adaptive beamforming design for mmWave RIS-aided joint localization and communication," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–6.

[11] Y. Liu, E. Liu, R. Wang, and Y. Geng, "Reconfigurable intelligent surface aided wireless localization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[12] F. Sohrabi, Z. Chen, and W. Yu, "Deep active learning approach to adaptive beamforming for mmWave initial alignment," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing (ICASSP)*, Jun. 2021, pp. 4940–4944.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, p. 1735–1780, Nov. 1997.

[14] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Select. Areas Commun.*, vol. 39, no. 7, pp. 1931 – 1945, Jul. 2021.

[15] F. Sohrabi, T. Jiang, W. Cui, and W. Yu, "Active sensing for communications by learning," *IEEE J. Select. Areas Commun.*, pp. 1–1, Mar. 2022.

[16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2015. [Online]. Available: http://arxiv.org/abs/1412.6980

[17] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *Proc. USENIX Conf. Operating Syst. Des. and Implementation (OSDI)*, 2016, p. 265–283.

[18] C. Prévost, E. Chaumette, K. Usevich, D. Brie, and P. Comon, "On cramér-rao lower bounds with random equality constraints," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing (ICASSP)*, May 2020, pp. 5355–5359.

[19] Y. Noam and H. Messer, "Notes on the tightness of the hybrid cramÉr–rao lower bound," *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2074–2084, Feb. 2009.