Active Sensing for Reciprocal MIMO Channels

Tao Jiang and Wei Yu
Electrical and Computer Engineering Department, University of Toronto, Canada
Emails: taoca.jiang@mail.utoronto.ca, weiyu@ece.utoronto.ca

***Abstract*—This paper tackles the problem of precoding and decoding matrices design for a time-division duplexing (TDD) massive MIMO system to support N_s independent data streams. The optimal design requires estimating the top-N_s singular vectors of the high-dimensional channel matrix, which typically involves significant pilot overhead if conventional channel estimation methods are used. Alternatively, some prior works seek to estimate the precoding and decoding matrices directly by exploiting channel reciprocity and the power iteration principle, but their performances suffer in the low SNR regime. To address this issue, this paper proposes a novel active sensing framework, where both transmitter and receiver send pilots alternatingly using their sensing beamformers that are actively designed as functions of previously received pilots. This is accomplished by a proposed active sensing unit, which first employs recurrent neural networks to summarize information from historical observations into their hidden state vectors then uses fully connected neural networks to obtain the sensing beamformers and precoding/decoding matrix. Simulations demonstrate that the proposed method outperforms existing approaches significantly and maintains superior performance even in low SNR regimes.**

***Index Terms*—Active sensing, channel estimation, deep learning, massive MIMO, power iteration method.**

I. Introduction

Massive multiple-input multiple-output (MIMO) technology has been a key enabler for high spectral efficiency communication in 5G and future wireless systems [1], particularly in mmWave bands, where a massive number of antennas can be packed into a small volume due to the short wavelength [2]. However, the optimal design of beamforming matrices at the transmitter and receiver requires high-dimensional channel state information (CSI), which involves significant pilot training overhead. In this paper, we investigate the design of precoding and decoding beamformers for a time-division duplexing (TDD) massive MIMO system. Inspired by recent developments in deep learning for reducing pilot training overhead [3], we propose an active sensing framework that directly learns the beamformers from received pilots without estimating the entire channel.

The conventional approach to massive MIMO system design involves first estimating the channel then optimizing the beamformers based on the estimated channel. To reduce the pilot training overhead, the knowledge of the underlying structure of the channel, e.g., sparsity in mmWave channels, can be utilized in the channel estimation step [4]. However, such an approach relies heavily on the geometry of the antenna arrays and the sparsity assumption on the channel model, which is not necessarily accurate. Moreover, in conventional approaches, the sensing beamformers for pilot transmission are often generated randomly, but sensing in random directions is not necessarily optimal, because typically only a low-dimensional subspace of the channel is of interest for designing the precoders/decoders.

Recently, active sensing has demonstrated superior performance for various applications including beam alignment [5], [6], beam tracking [7] and localization [8]. The main idea of active sensing is to gradually focus on the low-dimensional part of the channel by actively designing the sensing beamformers based on previously received pilots. For example, [5], [8] show that the active sensing approach can learn to probe broader beams at the beginning and gradually narrows down the searching range for angle-of-arrival estimation or localization. In this paper, we propose an active sensing framework to adaptively design the sensing beamformers to focus the sensing energy towards the top-N_s singular vector directions to learn the optimal precoding and decoding matrices.

The paper is most closely related to the works [6], [9], [10] about actively learning the singular vector pairs of the channel matrix. The paper [6] focuses on the problem of learning the top singular vector pair of the channel where the transmitter and receiver are equipped with only one RF chain. In a fully digital MIMO setup, [9] proposes to send pilots back and forth from both sides while the pilot sequences (or equivalently, the sensing beamformers) are designed based on the received pilots to mimic the power iteration methods. The power iteration method is also exploited in the design of the analog beamformers in [11]. However, the algorithm can converge to a highly suboptimal solution in the low signal-to-noise-ratio (SNR) regime, which is a typical scenario in mmWave initial alignment phase. The paper [10] proposes some techniques to address this problem, but it introduces extra feedback overhead between both sides.

In this paper, we show that the proposed active sensing framework can achieve superior performance in a wide range of SNRs without introducing extra feedback overhead. This is accomplished by a novel deep neural network architecture consisting of parallel gated recurrent units (GRUs) to abstract information from historical observations, and parallel fully connected deep neural networks (DNNs) to design the sensing beamformers and precoding/decoding matrix. Simulations show that the proposed framework works effectively even in the most challenging Rayleigh channel model scenario.

II. System Model

We consider a narrowband point-to-point MIMO system in which agent A with M_t fully digital antennas communicates with agent B with M_r fully digital antennas. The channel matrix from agent A to agent B is denoted as $\boldsymbol{G} \in \mathbb{C}^{M_\mathrm{r} \times M_\mathrm{t}}$. A block fading channel model is assumed. We assume the system operates in TDD mode and that the channel reciprocity holds, i.e., the uplink channel can be represented as $\boldsymbol{G}^\mathsf{H}$. The received signal $\boldsymbol{y} \in \mathbb{C}^{M_\mathrm{r}}$ at the agent B can be written as:

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{C}^{M_\mathrm{t}}$ is the signal transmitted from agent A and $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, \sigma_\mathrm{dl}^2 \boldsymbol{I})$ is the additive white Gaussian noise. Moreover, the signal $\boldsymbol{x}$ can be represented as $\boldsymbol{x} = \boldsymbol{W}_\mathrm{t}\boldsymbol{s}$, where $\boldsymbol{s} \in \mathbb{C}^{N_\mathrm{s}}$ denotes the data symbol vector with $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^\mathsf{H}] = \boldsymbol{I}$ and $\boldsymbol{W}_\mathrm{t} \in \mathbb{C}^{M_\mathrm{t} \times N_\mathrm{s}}$ is the precoding matrix at agent A. We choose the number of independent data streams to be less than the rank of the channel, i.e., $N_\mathrm{s} \leq \mathrm{rank}(\boldsymbol{G})$. We assume that the transmitted signal must satisfy a power constraint, i.e., $\mathbb{E}[\|\boldsymbol{x}\|_2^2] = \|\boldsymbol{W}_\mathrm{t}\|_F^2 \leq 1$.

After receiving the signal $\boldsymbol{y}$, agent B applies a decoding matrix $\boldsymbol{W}_\mathrm{t} \in \mathbb{C}^{M_\mathrm{r} \times N_\mathrm{s}}$ to obtain:

$$\hat{\boldsymbol{s}} = \boldsymbol{W}_\mathrm{r}^\mathsf{H}\boldsymbol{y} = \boldsymbol{W}_\mathrm{r}^\mathsf{H}\boldsymbol{G}\boldsymbol{W}_\mathrm{t}\boldsymbol{s} + \boldsymbol{W}_\mathrm{r}^\mathsf{H}\boldsymbol{n}. \tag{2}$$

The channel capacity of this system can be written as [12]:

$$R = \log_2 \det(\boldsymbol{I} + \boldsymbol{C}^{-1}\boldsymbol{W}_\mathrm{r}^\mathsf{H}\boldsymbol{G}\boldsymbol{W}_\mathrm{t}\boldsymbol{W}_\mathrm{t}^\mathsf{H}\boldsymbol{G}^\mathsf{H}\boldsymbol{W}_\mathrm{r}), \tag{3}$$

where $\boldsymbol{C} = \sigma_\mathrm{dl}^2 \boldsymbol{W}_\mathrm{r}^\mathsf{H}\boldsymbol{W}_\mathrm{r}$. This paper considers the capacity maximization problem via optimizing the precoding matrix $\boldsymbol{W}_\mathrm{t}$ and decoding matrix $\boldsymbol{W}_\mathrm{r}$.

If the channel $\boldsymbol{G}$ is known perfectly, the optimization of $\boldsymbol{W}_\mathrm{t}$ and $\boldsymbol{W}_\mathrm{r}$ has an analytic solution, namely $\boldsymbol{W}_\mathrm{t}$ and $\boldsymbol{W}_\mathrm{r}$ should match the singular vector pairs of the channel matrix $\boldsymbol{G}$ corresponding to the largest N_s singular values. Let the singular value decomposition (SVD) of $\boldsymbol{G}$ be

$$\boldsymbol{G} = [\boldsymbol{U}_1, \boldsymbol{U}_2] \begin{bmatrix} \boldsymbol{\Sigma}_1, & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_1^\mathsf{H} \\ \boldsymbol{V}_2^\mathsf{H} \end{bmatrix}, \tag{4}$$

where $\boldsymbol{U}_1 \in \mathbb{C}^{M_\mathrm{r} \times N_\mathrm{s}}$ and $\boldsymbol{V}_1 \in \mathbb{C}^{M_\mathrm{t} \times N_\mathrm{s}}$ have orthogonal and unit 2-norm columns, and $\boldsymbol{\Sigma}_1$ is a diagonal matrix with top-N_s singular values. To achieve the channel capacity, the optimal $\boldsymbol{W}_\mathrm{t}^*$ and $\boldsymbol{W}_\mathrm{r}^*$ are given by [12]

$$\boldsymbol{W}_\mathrm{t}^* = \boldsymbol{V}_1\boldsymbol{D}, \tag{5a}$$
$$\boldsymbol{W}_\mathrm{r}^* = \boldsymbol{U}_1, \tag{5b}$$

where $\boldsymbol{D}$ is a diagonal matrix with the diagonal terms given by water-filling power allocations scheme [12]. For simplicity, we assume a uniform power allocation scheme, i.e., $\boldsymbol{D} = \frac{1}{N_\mathrm{s}}\boldsymbol{I}$, which is near optimal in the high SNR regime.

This paper addresses the case where the channel matrix $\boldsymbol{G}$ is unknown. A conventional technique for designing $\boldsymbol{W}_\mathrm{t}$ and $\boldsymbol{W}_\mathrm{r}$ would have required first estimating the channel in a pilot stage, followed by SVD of the estimated channel. However, the channel estimation would require significant pilot overhead, especially for massive MIMO systems.

To reduce pilot training overhead, this paper aims to directly learn the precoding and decoding matrices from received pilots. This is based on the key observation from (5) that the optimal precoding and decoding matrices are only functions of the top-N_s singular vectors of $\boldsymbol{G}$ instead of the entire channel matrix. The proposed active sensing framework is based on a ping-pong pilot transmission scheme, where the two agents send pilot sequences alternatingly, and each agent actively designs the pilots based on previous observations so that both agents gradually discover the direction of the top-N_s singular vectors.

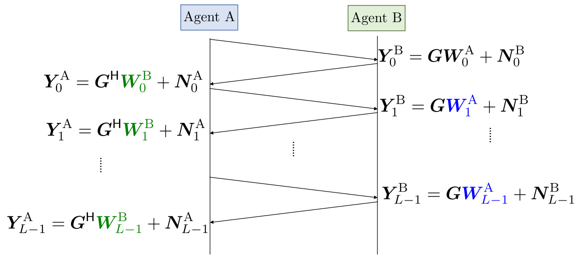


Fig. 1: Ping-pong pilot transmission protocol of L rounds.

III. Active Sensing via Ping-pong Pilots

The proposed ping-pong pilot transmission protocol is presented in Fig. 1, which consists of L rounds of ping-pong pilot transmission from both sides. In the ℓ-th round of ping-pong pilot transmission, agent A sends a sequence of pilots $\boldsymbol{W}_\ell^\mathrm{A} \in \mathbb{C}^{M_\mathrm{t} \times N_\mathrm{s}}$ of length N_s, and agent B observes $\boldsymbol{Y}_\ell^\mathrm{B}$ according to:

$$\boldsymbol{Y}_\ell^\mathrm{B} = \boldsymbol{G}\boldsymbol{W}_\ell^\mathrm{A} + \boldsymbol{N}_\ell^\mathrm{B}, \quad \ell = 0, \ldots, L-1, \tag{6}$$

where $\boldsymbol{N}_\ell^\mathrm{B}$ is the additive white Gaussian noise with each column independently distributed as $\mathcal{CN}(\boldsymbol{0}, \sigma_\mathrm{B}^2\boldsymbol{I})$. The pilot vector $\boldsymbol{w}_{\ell,i}^\mathrm{A}$, which is the i-th column of $\boldsymbol{W}_\ell^\mathrm{A}$, is subjected to the power constraint $\|\boldsymbol{w}_{\ell,i}^\mathrm{A}\|_2 \leq 1$. After receiving the pilots, agent B sends back a sequence of pilots $\boldsymbol{W}_\ell^\mathrm{B} \in \mathbb{C}^{M_\mathrm{r} \times N_\mathrm{s}}$. Due to channel reciprocity, agent A receives pilots $\boldsymbol{Y}_\ell^\mathrm{A}$ as follows:

$$\boldsymbol{Y}_\ell^\mathrm{A} = \boldsymbol{G}^\mathsf{H}\boldsymbol{W}_\ell^\mathrm{B} + \boldsymbol{N}_\ell^\mathrm{A}, \quad \ell = 0, \ldots, L-1, \tag{7}$$

where $\boldsymbol{N}_\ell^\mathrm{A}$ is the additive white Gaussian noise with each column independently distributed as $\mathcal{CN}(\boldsymbol{0}, \sigma_\mathrm{A}^2\boldsymbol{I})$. The i-th column of $\boldsymbol{W}_\ell^\mathrm{B}$, denoted by $\boldsymbol{w}_{\ell,i}^\mathrm{B}$, satisfies a power constraint $\|\boldsymbol{w}_{\ell,i}^\mathrm{B}\|^2 \leq 1$. The pilot vectors $\boldsymbol{w}_{\ell,i}^\mathrm{A}$ and $\boldsymbol{w}_{\ell,i}^\mathrm{B}$ can be thought of as sensing beamformers because they are used to sense the channel for estimating the top-N_s singular vectors of $\boldsymbol{G}$.

The conventional channel estimation based approach often generates the sensing beamformers (i.e., pilot sequence) randomly and estimates the entire channel matrix, which is not the most efficient way to probe the low-dimensional part of the channel matrix corresponding to the top-N_s singular vectors. In this paper, we propose to actively design the sensing beamformers based on the pilots received so far, so that the sensing beamformers learn to focus energy towards the directions of the top-N_s singular vectors.

Specifically, the sensing beamformers $\boldsymbol{W}_\ell^\mathrm{A}$ and $\boldsymbol{W}_\ell^\mathrm{B}$ are designed based on the past received pilots on each side as follows:

$$\boldsymbol{W}_\ell^\mathrm{A} = f_\ell^\mathrm{A}(\boldsymbol{Y}_0^\mathrm{A}, \ldots, \boldsymbol{Y}_\ell^\mathrm{A}), \quad \ell = 0, \ldots, L-2, \tag{8a}$$
$$\boldsymbol{W}_\ell^\mathrm{B} = f_\ell^\mathrm{B}(\boldsymbol{Y}_0^\mathrm{B}, \ldots, \boldsymbol{Y}_\ell^\mathrm{B}), \quad \ell = 0, \ldots, L-1, \tag{8b}$$

where $f_\ell^{\rm A}$ and $f_\ell^{\rm B}$ are the sensing strategies of the ℓ-th rounds of transmission at agent A and agent B, respectively. After L rounds of ping-pong pilot transmission, each agent utilizes all the received pilots on each side to design the data transmission precoding matrix and decoding matrix as follows:

$$\boldsymbol{W}_{\rm t} = g^{\rm A}(\boldsymbol{Y}_0^{\rm A}, \ldots, \boldsymbol{Y}_{L-1}^{\rm A}), \quad (9a)$$

$$\boldsymbol{W}_{\rm r} = g^{\rm B}(\boldsymbol{Y}_0^{\rm B}, \ldots, \boldsymbol{Y}_{L-1}^{\rm B}), \quad (9b)$$

where $g^{\rm A}$ and $g^{\rm B}$ are functions to map all the received pilots to the desired solutions.

The goal of this paper is to find the mappings $\{f_\ell^{\rm A}\}_{\ell=0}^{L-2}$, $\{f_\ell^{\rm B}\}_{\ell=0}^{L-1}$, $g^{\rm A}$ and $g^{\rm B}$ such that the final precoding and decoding matrices, $\boldsymbol{W}_{\rm t}$ and $\boldsymbol{W}_{\rm r}$, closely match $\boldsymbol{V}_1$ and $\boldsymbol{U}_1$, respectively. This is a highly nontrivial problem because it involves searching for solutions in a high-dimensional functional space. To address this problem, we propose a deep active sensing framework that parameterizes the functional mappings by a set of carefully designed deep neural networks.

IV. Conventional Power Iteration Method

To motivate the proposed deep learning approach, we review a method proposed in [9] for designing the sensing beamformers based on power iteration in the ping-pong pilot training stage. We use the example of estimating the top singular vector (i.e., $N_{\rm s} = 1$) in the noiseless scenario to illustrate the idea.

In the ℓ-th ping-pong round, each agent sends a pilot symbol and simply sets the next sensing vectors as the current received pilot vector, which means $\boldsymbol{w}_{\ell+1}^{\rm A} = \boldsymbol{y}_\ell^{\rm A}$ and $\boldsymbol{w}_\ell^{\rm B} = \boldsymbol{y}_\ell^{\rm B}$. Starting for a random vector $\boldsymbol{w}_0^{\rm A}$, we obtain the following equations after ℓ rounds of transmission:

$$\boldsymbol{y}_\ell^{\rm A} = (\boldsymbol{G}^{\sf H}\boldsymbol{G})^\ell \boldsymbol{w}_0^{\rm A} = \sum_i \sigma_i^{2\ell}\beta_i \boldsymbol{v}_i, \quad (10a)$$

$$\boldsymbol{y}_\ell^{\rm B} = (\boldsymbol{G}\boldsymbol{G}^{\sf H})^{\ell-1}\boldsymbol{G}\boldsymbol{w}_0^{\rm A} = \sum_i \sigma_i^{2\ell-1}\beta_i \boldsymbol{u}_i, \quad (10b)$$

where we use the fact that $\boldsymbol{w}_0^{\rm A}$ can be expressed as $\boldsymbol{w}_0^{\rm A} = \sum_{i=1}^{{\rm rank}(\boldsymbol{G})} \beta_i \boldsymbol{v}_i + \bar{\boldsymbol{v}}_{{\rm t},0}$, with $\bar{\boldsymbol{v}}_{{\rm t},0}$ being a vector in the null space of $\boldsymbol{G}$. As ℓ increases, the vectors $\boldsymbol{y}_\ell^{\rm A}$ and $\boldsymbol{y}_\ell^{\rm B}$ in (10) will be dominated by the top singular vectors $\boldsymbol{v}_1$ and $\boldsymbol{u}_1$, respectively, with a linear convergence rate [9]. Furthermore, the power normalization in each iteration $\boldsymbol{w}_\ell^{\rm B} = \boldsymbol{y}_\ell^{\rm B}/\|\boldsymbol{y}_\ell^{\rm B}\|_2$ and $\boldsymbol{w}_{\ell+1}^{\rm A} = \boldsymbol{y}_\ell^{\rm A}/\|\boldsymbol{y}_\ell^{\rm A}\|_2$ does not alter the direction of $\boldsymbol{y}_\ell^{\rm A}$ and $\boldsymbol{y}_\ell^{\rm B}$. The fast convergence rate implies a significant pilot overhead reduction as compared to estimating the entire channel matrix.

The power iteration method can be extended to the $N_{\rm s} > 1$ scenario. Specifically, agent B performs QR decomposition on the received pilots $\boldsymbol{Y}_\ell^{\rm B}$ in (6), i.e., $\boldsymbol{Y}_\ell^{\rm B} = \boldsymbol{Q}_\ell^{\rm B}\boldsymbol{R}_\ell^{\rm B}$, then sends back N_s pilots to agent A by setting $\boldsymbol{W}_\ell^{\rm B} = \boldsymbol{Q}_\ell^{\rm B}$. Similarly, agent A performs QR decomposition on the received pilots $\boldsymbol{Y}_\ell^{\rm A}$ in (7), i.e., $\boldsymbol{Y}_\ell^{\rm A} = \boldsymbol{Q}_\ell^{\rm A}\boldsymbol{R}_\ell^{\rm A}$, and sets $\boldsymbol{W}_{\ell+1}^{\rm A} = \boldsymbol{Q}^{\rm A}$ for the next transmission.

The primary impediment to the practical implementation of such a power iteration method is that the algorithm can converge to a highly suboptimal solution in the presence of noise. In this paper, we leverage a data driven approach to mimic the power iteration method while mitigating its limitation in noisy scenarios by incorporating a neural network based active sensing unit.

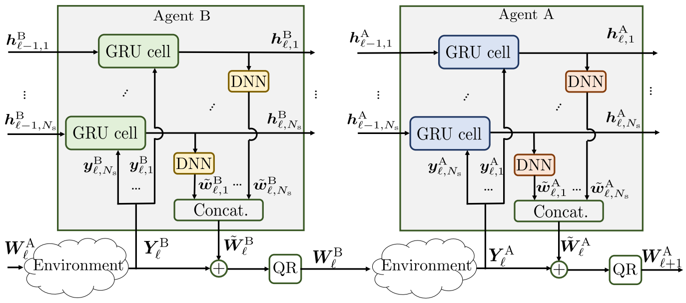


Fig. 2: Proposed active sensing unit in the ℓ-th pilot round.

Algorithm 1 Proposed active sensing framework.

Initial $\boldsymbol{W}_0^{\rm A} \in \mathbb{C}^{M_t \times N_{\rm s}}$, $\boldsymbol{h}_{-1,i}^{\rm B} \in \mathbb{C}^{N_{\rm h}^{\rm B}}$, $\boldsymbol{h}_{-1,i}^{\rm A} \in \mathbb{C}^{N_{\rm h}^{\rm A}}$.
for $\ell = 0, \ldots, L-1$ **do**
  A→ B: $\boldsymbol{Y}_\ell^{\rm B} = \boldsymbol{G}\boldsymbol{W}_\ell^{\rm A} + \boldsymbol{N}_\ell^{\rm B}$
  for $i = 1, \ldots, N_{\rm s}$ **do**
    $\boldsymbol{h}_{\ell,i}^{\rm B} = {\rm GRUCell}^{\rm B}(\boldsymbol{h}_{\ell-1,i}^{\rm B}, \boldsymbol{y}_{\ell,i}^{\rm B})$
    $\tilde{\boldsymbol{w}}_{\ell,i}^{\rm B} = {\rm DNN}^{\rm B}(\boldsymbol{h}_{\ell,i}^{\rm B})$
  end for
  $\tilde{\boldsymbol{W}}_\ell^{\rm B} = [\tilde{\boldsymbol{w}}_{\ell,1}^{\rm B}, \ldots, \tilde{\boldsymbol{w}}_{\ell,N_{\rm s}}^{\rm B}]$
  QR decomposition: $(\tilde{\boldsymbol{W}}_\ell^{\rm B} + \boldsymbol{Y}_\ell^{\rm B}) = \boldsymbol{Q}_\ell^{\rm B}\boldsymbol{R}_\ell^{\rm B}$
  Set $\boldsymbol{W}_\ell^{\rm B} = \boldsymbol{Q}_\ell^{\rm B}$.

  B→ A: $\boldsymbol{Y}_\ell^{\rm A} = \boldsymbol{G}^{\sf H}\boldsymbol{W}_\ell^{\rm B} + \boldsymbol{N}_\ell^{\rm A}$
  for $i = 1, \ldots, N_{\rm s}$ **do**
    $\boldsymbol{h}_{\ell,i}^{\rm A} = {\rm GRUCell}^{\rm A}(\boldsymbol{h}_{\ell-1,i}^{\rm A}, \boldsymbol{y}_{\ell,i}^{\rm A})$
    $\tilde{\boldsymbol{w}}_{\ell,i}^{\rm A} = {\rm DNN}^{\rm A}(\boldsymbol{h}_{\ell,i}^{\rm A})$
  end for
  $\tilde{\boldsymbol{W}}_\ell^{\rm A} = [\tilde{\boldsymbol{w}}_{\ell,1}^{\rm A}, \ldots, \tilde{\boldsymbol{w}}_{\ell,N_{\rm s}}^{\rm A}]$
  QR decomposition: $(\tilde{\boldsymbol{W}}_\ell^{\rm A} + \boldsymbol{Y}_\ell^{\rm A}) = \boldsymbol{Q}_\ell^{\rm A}\boldsymbol{R}_\ell^{\rm A}$
  Set $\boldsymbol{W}_{\ell+1}^{\rm A} = \boldsymbol{Q}_\ell^{\rm A}$
end for

V. Proposed Active Sensing Framework

We now describe the proposed active sensing framework for learning the sensing matrices in the ping-pong pilot transmission stage. The proposed framework is designed to be robust to channel noise and can achieve excellent performance even in low SNR scenarios. Moreover, the proposed framework includes the power iteration method as a special case.

The proposed active sensing framework parameterizes the sensing strategies $f_\ell^{\rm A}$ and $f_\ell^{\rm B}$ with an active sensing unit as shown in Fig. 2. Specifically, both agent A and agent B consist of $N_{\rm s}$ GRU cells to extract useful information from the received pilots and $N_{\rm s}$ DNNs to map the updated hidden state vectors to the next pilot sequence. More specifically, given the observations $\boldsymbol{Y}_\ell^{\rm B}$ in (6), agent B takes the i-th column of $\boldsymbol{Y}_\ell^{\rm B}$

(denoted by $\boldsymbol{y}^{\mathrm{B}}_{\ell,i}$) as input to the i-th GRU cell. The GRU cell updates its hidden state vector $\boldsymbol{h}^{\mathrm{B}}_{\ell,i} \in \mathbb{C}^{N_{\mathrm{h}}^{\mathrm{B}}}$ as follows:

$$\boldsymbol{h}^{\mathrm{B}}_{\ell,i} = \mathrm{GRUCell}^{\mathrm{B}}(\boldsymbol{h}^{\mathrm{B}}_{\ell-1,i}, \boldsymbol{y}^{\mathrm{B}}_{\ell,i}), \quad i = 1, \ldots, N_{\mathrm{s}}, \tag{11}$$

where $\mathrm{GRUCell}^{\mathrm{B}}$ follows the standard gated recurrent unit implementation as proposed in [13]. Here, the ability of GRU to summarize useful information from historical observations into a fixed dimensional vector enables the active sensing framework to scale up to any number of transmission rounds L since the same unit in Fig. 2 can be applied to different transmission rounds.

The hidden state vector $\boldsymbol{h}^{\mathrm{B}}_{\ell,i}$ is then mapped to a vector $\tilde{\boldsymbol{w}}^{\mathrm{B}}_{\ell,i} \in \mathbb{C}^{M_{\mathrm{t}}}$ using a fully connected neural network $\mathrm{DNN}^{\mathrm{B}}$:

$$\tilde{\boldsymbol{w}}^{\mathrm{B}}_{\ell,i} = \mathrm{DNN}^{\mathrm{B}}(\boldsymbol{h}^{\mathrm{B}}_{\ell,i}), \quad i = 1, \ldots, N_{\mathrm{s}}. \tag{12}$$

The vectors $\tilde{\boldsymbol{w}}^{\mathrm{B}}_{\ell,i}$'s are collected into a matrix $\tilde{\boldsymbol{W}}^{\mathrm{B}}_{\ell}$, given by

$$\tilde{\boldsymbol{W}}^{\mathrm{B}}_{\ell} = [\tilde{\boldsymbol{w}}^{\mathrm{B}}_{\ell,1}, \ldots, \tilde{\boldsymbol{w}}^{\mathrm{B}}_{\ell,N_{\mathrm{s}}}]. \tag{13}$$

After power normalization in each column, the matrix $\tilde{\boldsymbol{W}}^{\mathrm{B}}_{\ell}$ can already be used as the designed sensing beamformers in the ℓ-th transmission round at agent B.

To further improve the performance, we propose to incorporate the power iteration method into the neural network through the following steps:

$$(\tilde{\boldsymbol{W}}^{\mathrm{B}}_{\ell} + \boldsymbol{Y}^{\mathrm{B}}_{\ell}) = \boldsymbol{Q}^{\mathrm{B}}_{\ell}\boldsymbol{R}^{\mathrm{B}}_{\ell} \quad \text{(QR decomposition)}, \tag{14a}$$

$$\boldsymbol{W}^{\mathrm{B}}_{\ell} = \boldsymbol{Q}^{\mathrm{B}}_{\ell}. \tag{14b}$$

This ensures that our proposed active sensing framework performs at least as well as the conventional power method. If the neural network generates a matrix $\tilde{\boldsymbol{W}}^{\mathrm{B}}_{\ell} = \boldsymbol{0}$, the proposed active sensing framework reduces to the conventional power method proposed in [9].

The neural network architecture at agent A is the same as that of agent B, except for using different parameters. The overall algorithm in the ping-pong pilot transmission stage is listed in Algorithm 1. The initial matrix $\boldsymbol{W}^{\mathrm{A}}_{0} \in \mathbb{C}^{M_{\mathrm{t}} \times N_{\mathrm{s}}}$ is learned from channel statistics in the training stage and remains fixed in the testing stage. The initial hidden state vectors $\boldsymbol{h}^{\mathrm{B}}_{-1,i} \in \mathbb{C}^{N_{\mathrm{h}}^{\mathrm{B}}}$ and $\boldsymbol{h}^{\mathrm{A}}_{-1,i} \in \mathbb{C}^{N_{\mathrm{h}}^{\mathrm{A}}}$ are both set to the vector with all entries equal to one. To improve scalability, the GRU cells and DNNs on each side share the same parameters, respectively.

To design the final precoding/decoding matrices in the data transmission stage, both agents use another set of DNNs to map the latest hidden state vectors to the corresponding matrices, which are further processed with QR decompositions. After L rounds of pilot transmission, agent B generates the final decoding matrix $\boldsymbol{W}_{\mathrm{r}}$ as follows:

$$\tilde{\boldsymbol{w}}_{\mathrm{r},i} = \mathrm{DNN}^{\mathrm{r}}(\boldsymbol{h}^{\mathrm{B}}_{L,i}), \quad i = 1, \ldots, N_{\mathrm{s}}, \tag{15a}$$

$$\tilde{\boldsymbol{W}}_{\mathrm{r}} = [\tilde{\boldsymbol{w}}_{\mathrm{r},1}, \ldots, \tilde{\boldsymbol{w}}_{\mathrm{r},N_{\mathrm{s}}}], \tag{15b}$$

$$(\tilde{\boldsymbol{W}}_{\mathrm{r}} + \boldsymbol{Y}^{\mathrm{B}}_{L-1}) = \boldsymbol{Q}^{\mathrm{r}}\boldsymbol{R}^{\mathrm{r}}, \quad \text{(QR decomposition)}, \tag{15c}$$

$$\boldsymbol{W}_{\mathrm{r}} = \boldsymbol{Q}^{\mathrm{r}}. \tag{15d}$$

Similarly, agent A uses its hidden state vector $\boldsymbol{h}^{\mathrm{A}}_{L,i}$ and received pilots $\boldsymbol{Y}^{\mathrm{A}}_{L-1}$ to design its precoding matrix $\boldsymbol{W}_{\mathrm{t}}$ following the same procedure in (15), but replaces the $\mathrm{DNN}^{\mathrm{r}}$ and $\boldsymbol{Y}^{\mathrm{B}}_{L-1}$ with $\mathrm{DNN}^{\mathrm{t}}$ and $\boldsymbol{Y}^{\mathrm{A}}_{L-1}$, respectively.

The proposed active sensing framework is trained by concatenating L active sensing units sequentially to optimize the policy for the entire sensing trajectory during the training stage. The training objective is to enforce the neural network's outputs $\boldsymbol{W}_{\mathrm{r}}$ and $\boldsymbol{W}_{\mathrm{t}}$ to match the optimal solution $\boldsymbol{V}_1$ and $\boldsymbol{U}_1$. While the loss function $\mathbb{E}\big[\|\boldsymbol{W}_{\mathrm{r}} - \boldsymbol{V}_1\|_2^2 + \|\boldsymbol{W}_{\mathrm{t}} - \boldsymbol{U}_1\|_2^2\big]$ can achieve this goal, it requires performing SVD for each channel matrix in the training data to obtain the true labels $\boldsymbol{V}_1$ and $\boldsymbol{U}_1$. To avoid SVD calculations, we propose to use unsupervised training, where the loss function is set as $-\mathbb{E}\big[\det(\boldsymbol{W}_{\mathrm{r}}^{\mathsf{H}}\boldsymbol{G}\boldsymbol{W}_{\mathrm{t}})\big]$ since the function $\det(\boldsymbol{W}_{\mathrm{r}}^{\mathsf{H}}\boldsymbol{G}\boldsymbol{W}_{\mathrm{t}})$ is maximized when $\boldsymbol{W}_{\mathrm{r}} = \boldsymbol{V}_1$ and $\boldsymbol{W}_{\mathrm{t}} = \boldsymbol{U}_1$. Furthermore, to train a neural network that is generalizable to different ping-pong round L, we set the loss function as $-\mathbb{E}\Big[\sum_{\ell=0}^{L-1}\det(\boldsymbol{W}_{\mathrm{r},\ell}^{\mathsf{H}}\boldsymbol{G}\boldsymbol{W}_{\mathrm{t},\ell})\Big]$, where $\boldsymbol{W}_{\mathrm{r},\ell}$ and $\boldsymbol{W}_{\mathrm{t},\ell}$ are the data transmission precoding and decoding matrices generated by the neural network in round ℓ.

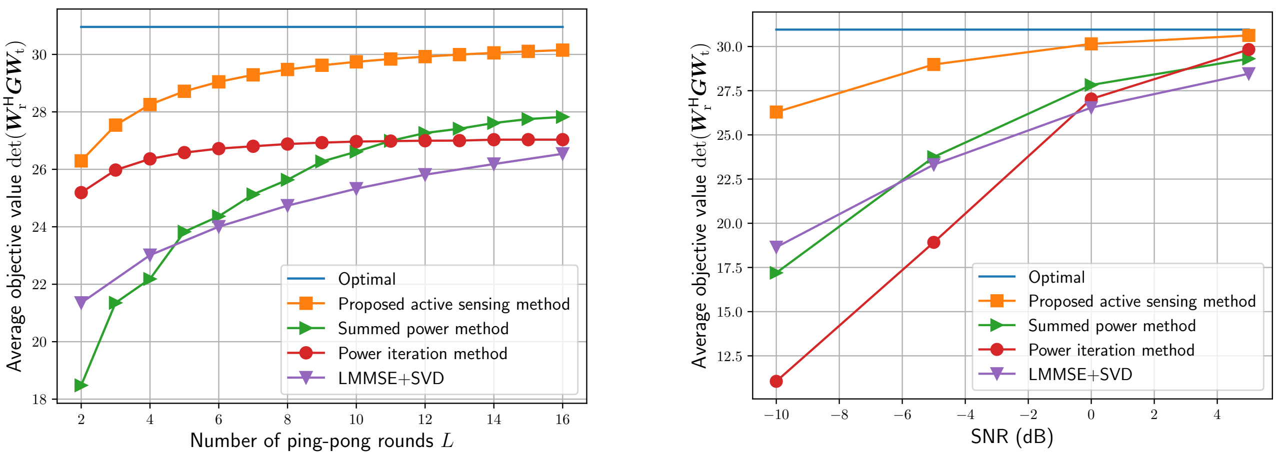
VI. Simulation Results

In this section, we evaluate the performance of the proposed active sensing framework for a MIMO system with $M_{\mathrm{t}} = 64$ transmit antennas and $M_{\mathrm{r}} = 64$ receive antennas. The number of independent data streams is set to $N_{\mathrm{s}} = 4$. The channel matrix follows the Rayleigh fading model, namely the entries of $\boldsymbol{G}$ follow i.i.d. Gaussian distribution $\mathcal{CN}(0,1)$. In practice, the model should be trained using the site-specific channel data to achieve the best performance. For simplicity, we assume the SNRs in both the uplink and the downlink are the same, i.e., $\mathrm{SNR} \triangleq 1/\sigma_{\mathrm{A}}^2 = 1/\sigma_{\mathrm{B}}^2$.

We set the dimensions of all the hidden states of the gated recurrent units (GRUs) to be $N_{\mathrm{h}}^{\mathrm{A}} = N_{\mathrm{h}}^{\mathrm{B}} = 512$. All the DNNs are two-layer fully connected neural networks with the dimension $[512, 1024, 2M_{\mathrm{t}}]$. The entire neural network is implemented on PyTorch [14]. We train the neural network offline using the Adam optimizer [15], with an initial learning rate progressively decreasing from 10^{-3} to 10^{-5}. We generate as much training data as needed, and stop training if the performance on the validation set does not improve after several epochs. The testing dataset consists of 1000 randomly generated channel matrices.

We compare the proposed active sensing approach with the following benchmarks. i) *LMMSE+SVD*: Each agent first estimates the channel matrix based on the received pilots using linear minimum mean square error (LMMSE) estimator then performs SVD on the estimated channel matrix to design precoding/decoding matrices. ii): *Power iteration method* [9]. iii) *Summed power method* [10]: Both agents calculate their next sensing beamformers based on the accumulated sum of previously received pilots to effectively average out noise.

In Fig. 3a, we present the average objective value against the number of ping-pong transmission rounds at SNR = 0dB. The neural network is trained at $L = 16$. It can be seen from the figure that the proposed active sensing method outperforms



(a) Objective value vs. number of ping-pong rounds for SNR $= 0$dB.

(b) Objective value vs. SNR for $L = 16$.

Fig. 3: Performance comparison of different methods for a MIMO system with $M_{\mathrm{t}} = M_{\mathrm{r}} = 64$, $N_{\mathrm{s}} = 4$.

other benchmarks significantly. Although the power iteration method converges quickly, it yields a highly suboptimal solution due to the impact of noise. The summed power method can eventually achieve better performance than the power iteration method, but it still performs much worse than the proposed method. The channel estimation-based method is not as competitive as the other approaches, even though the LMMSE estimator is optimal in terms of the mean squared error metric in this Rayleigh fading scenario. This implies that estimating the entire channel matrix is inefficient if we only need partial information on the channel matrix, i.e., the top-N_{s} singular vectors.

In Fig. 3b, we plot the average objective value against the SNR for a fixed number of ping-pong rounds $L = 16$. We observe that the proposed active sensing approach can still perform well even when the SNR drops to -10dB. However, the performance of the other benchmarks degrades significantly as the SNR decreases. This indicates that the proposed approach is much more robust to noise as compared to the other benchmarks.

VII. Conclusion

This paper proposes an active sensing framework to directly estimate the optimal precoding and decoding matrices for a TDD massive MIMO system in a ping-pong pilot training stage, where the sensing beamformers are actively designed based on historically received pilots on both sides. Compared to previous approaches, the proposed algorithm achieves significantly better performance and maintains superior performance in low SNR regimes. Simulations also show that the proposed algorithm works well even in the most challenging Rayleigh fading channel model.

References

[1] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays," *Digit. Signal Process.*, vol. 94, pp. 3–20, Nov. 2019.

[2] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[3] W. Yu, F. Sohrabi, and T. Jiang, "Role of deep learning in wireless communications," *IEEE BITS Inf. Theory Mag.*, vol. 2, no. 2, pp. 56–72, Nov. 2022.

[4] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[5] F. Sohrabi, T. Jiang, W. Cui, and W. Yu, "Active sensing for communications by learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1780–1794, June 2022.

[6] T. Jiang, F. Sohrabi, and W. Yu, "Active sensing for two-sided beam alignment using ping-pong pilots," in *Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, California, USA, Nov. 2022, pp. 913–918.

[7] H. Han, T. Jiang, and W. Yu, "Active beam tracking with reconfigurable intelligent surface," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, Rhodes, Greece, June 2023.

[8] Z. Zhang, T. Jiang, and W. Yu, "Active sensing for localization with reconfigurable intelligent surface," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, June 2023.

[9] T. Dahl, N. Christophersen, and D. Gesbert, "Blind MIMO eigenmode transmission based on the algebraic power method," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2424–2431, Sept. 2004.

[10] D. Ogbe, D. J. Love, and V. Raghavan, "Noisy beam alignment techniques for reciprocal MIMO channels," *IEEE Trans. Signal Process.*, vol. 65, no. 19, pp. 5092–5107, Oct. 2017.

[11] P. Xia, R. W. Heath, and N. Gonzalez-Prelcic, "Robust analog precoding designs for millimeter wave MIMO transceivers with frequency and time division duplexing," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4622–4634, Nov. 2016.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 2006.

[13] K. Cho, B. van Merrienboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. Conf. on Empirical Methods in Natural Lang. Process. (EMNLP)*, Doha, Qatar, Oct. 2014, pp. 1724–1734.

[14] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Pytorch: An imperative style, high-performance deep learning library," in *Proc. Conf. on Neural Info. Process. Systems (NeurIPS)*, Vancouver, Canada, Dec. 2019, pp. 8024–8035.

[15] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Int. Conf. on Learning Representations (ICLR)*, Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, May 2015, pp. 1–13.