AI-empowered Fluid Antenna Systems: Opportunities, Challenges, and Future Directions

Chao Wang, Zan Li, Kai-Kit Wong, Ross Murch, Chan-Byoung Chae, and Shi Jin

***Abstract*—Fluid antennas embrace all forms of flexible-position antennas, both movable and non-movable. The concept of fluid antenna system (FAS) hence introduces a new dimension to enhance multiple-input multiple-output (MIMO) antenna systems, which is essential for achieving more ambitious goals in wireless communications. FAS fundamentally changes the way MIMO systems are optimized. In addition to optimizing precoding and decoding matrices, a flexible-position MIMO system, referred to as MIMO-FAS, needs to optimize the positions (i.e., ports) of the antennas to achieve the best performance. Unfortunately, due to the near-continuous nature of antenna position adjustment as well as the resulting high dimensionality, optimizing MIMO-FAS is NP-hard, complicated by the coupling between the optimization variables. Given the rapid advances in artificial intelligence (AI), it is fitting to harness its capabilities to alleviate the challenges of MIMO-FAS. This article explores a vision of how AI empowers FAS to overcome these hurdles using a learning-based approach, enabling FAS to excel. Furthermore, we use emerging integrated sensing and communication (ISAC) scenarios as a case study to illustrate the potential of enhancing FAS with AI capability.**

***Index Terms*—Artificial intelligence, flexible-position MIMO, fluid antenna system, learning-aided optimization.**

I. Introduction

Multiple-input multiple-output (MIMO) has been carrying the physical layer of wireless communications since its third generation (3G) in 2000. Its ability to create bandwidth from space has positioned it as an indispensable mobile communication technology. In the fourth generation (4G), MIMO came in the form of multiuser MIMO to accommodate multiple users on the same channel while MIMO has evolved into the massive version in the fifth generation (5G). A 5G base station (BS) has 64 antennas to support a maximum of 12 user equipments (UEs) delivering super-directivity on the same time-frequency resource unit. Looking ahead, recent trend suggests that an even greater number of antennas be deployed at the BS.

However, the same increase in the number of antennas at the UE has not been seen. Three decades have passed but the standard only requires just 4 receive antennas at the UE for the core 5G bands. The main reason is due to the small available space at the mobile devices. Electrically small metamaterial-based antennas have emerged to overcome this limitation [1] but it is not just the antenna size and mutual coupling between elements but also the radio frequency (RF) components that come with it. Traditionally, it is strongly believed that antennas should have at least a half-wavelength separation among them to justify the cost. The rationale is that antennas should ensure to receive statistically independent signals to be worthy.

Recent advances in reconfigurable antennas have led to the concept of fluid antennas. Fluid antenna represents *all forms of movable and non-movable flexible-position antennas* [2].[1] This provides a new way to obtain spatial diversity with less number of RF chains in a relatively small space [3]. Of relevance is the non-movable flexible-position antenna realizable by pixel based antennas that can influence the physical layer of wireless communications without noticeable delay [4].

Fluid antenna systems (FASs) were first introduced by Wong *et al.* in [5] where the receiver used a flexible-position antenna to maximize the signal-to-noise ratio (SNR). Subsequent work in [6] provided deeper analysis on the impact of the size and resolution of FAS on the performance while [7] made a critical contribution on refining the spatial correlation model for the analysis of FAS. Additionally, flexible-position MIMO system, referred to as MIMO-FAS, was studied recently in [8], which showed that there is enormous diversity gain over traditional fixed-position MIMO systems, see TABLE I. The substantial diversity should translate into saving of energy consumption over fixed antenna systems though a deeper analysis is needed when taking into account of the energy consumption due to the practical implementation of FAS.

The ability to flexibly change antenna position also presents a whole new technique for multiuser communications. In [9], it was hypothesized that if FAS could switch to the position with the minimum sum-interference plus noise signal on a per symbol basis, a UE could deal with hundreds of interferers utilizing a small-sized FAS with one RF chain. More recently, [10] considered a more practical setting where each UE's FAS only switches its position if the channel changes, and when it does, the FAS maximizes the average signal-to-interference plus noise ratio (SINR). The slow fluid antenna multiple access

The work of C. Wang is supported in part by the National Natural Science Foundation of China under Grants 62371357 and 61801518, in part by the International Postdoctoral Exchange Fellowship Program under Grant PC2021060, and in part by the China Postdoctoral Science Foundation under Grant 2020M683428. The work of Z. Li is supported by the National Science Fund for Distinguished Young Scholars of China under Grant 61825104. The work is supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/W026813/1. The work of C. B. Chae is supported by the Institute of Information and Communication Technology Promotion (IITP) grant funded by the Ministry of Science and ICT (MSIT), Korea (2021-0-02208, 2022R1A5A1027646). The work of R. Murch is supported by the Hong Kong Research Grants Council Area of Excellence grant AoE/E-601/22-R.

C. Wang and Z. Li are both with the Integrated Service Networks Lab, Xidian University, Xi'an 710071, China.

K. K. Wong is with the Department of Electronic and Electrical Engineering, University College London, Torrington Place, United Kingdom. He is also affiliated with Yonsei Frontier Laboratory and School of Integrated Technology, Yonsei University, Seoul, 03722, Korea.

R. Murch is with the Department of Electronic and Computer Engineering and Institute for Advanced Study (IAS), The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, China.

C. B. Chae is with the School of Integrated Technology, Yonsei University, Seoul, 03722, Korea.

S. Jin is is with the Frontiers Science Center for Mobile Information Communication and Security, Southeast University, Nanjing, China.

[1]In a broader sense, fluid antenna even includes flexibility in size and shape for other reconfigurabilities such as operating frequency and pattern.

TABLE I
DIVERSITY ORDER COMPARISON BETWEEN MIMO AND MIMO-FAS UNDER RICH SCATTERING ENVIRONMENTS.

Size (in λ^2)[‡]	MIMO	MIMO-FAS
0.5×0.5	$4 \times 4 = 16$	$13 \times 13 = 169$
1×1	$9 \times 9 = 81$	$23 \times 23 = 529$
1.5×1.5	$16 \times 16 = 256$	$34 \times 34 = 1156$
2×2	$25 \times 25 = 625$	$48 \times 48 = 2304$
2.5×2.5	$36 \times 36 = 1296$	$60 \times 60 = 3600$
3×3	$49 \times 49 = 2401$	$73 \times 73 = 5329$

[‡]The size corresponds to the size of a two-dimensional (2D) FAS surface at each end where λ denotes the wavelength.

(FAMA) scheme in [10] could handle several UEs on the same channel without the need of power control and precoding.

It is fair to say that FAS research is still in their infancy. A majority of results thus far largely considered the operation of FAS under ideal conditions, concerning less on how channel state information (CSI) is obtained, and how the optimization of position change (a.k.a. port selection) is done. For instance, FAS entails a large number of ports while the available RF chains are limited, making CSI prediction for each port challenging. Traditional optimization-based prediction approaches would struggle to address channel prediction problems in this scale. Furthermore, integrating MIMO with FAS comes with additional challenges, especially in scenarios where complete CSI is not available. This integration further complicates the joint optimization problem of non-convex precoding and NP-hard port selection. It is evident that traditional optimization methods are ill-equipped to deal with such high-dimensional mixed optimization problems. There are also desires to synergize with other emerging technologies such as reconfigurable intelligent surfaces (RIS), integrated sensing and communications (ISAC) and more. Even more complex optimization is therefore anticipated if FAS is employed.

With the explosive growth of artificial intelligence (AI), it is hopeful that the signal processing of FAS can become feasible under practical conditions. FAS exploits a series of spatially correlated signals while AI is known for its uncanny ability to recognize hidden correlations. This makes a great marriage between the two. The purpose of this article is to discuss some key areas in FAS in which learning-based approaches can be very effective. We will start with a brief review on the types of FAS, highlighting the anticipated capability of single and multiuser MIMO-FAS in Section II. Section III then outlines the design challenges in FAS before presenting how ideas in AI, such as deep learning, can tackle the challenges in Section IV. A case study on ISAC will be discussed in Section V where we show how AI-assisted FAS can deliver great performance. Finally, Section VI gives a list of future research topics and we have our concluding remarks in Section VII.

II. PRELIMINARIES AND OPPORTUNITIES

A. Types of Fluid Antenna

For decades, antennas were made of blocks of metal materials and they have been very efficient. But the emergence of cognitive radio and related applications has created the demand for highly reconfigurable antennas that motivate some to study using soft materials as antennas for additional reconfigurability. Recently, it is even proposed that an antenna can be made of many reconfigurable RF pixels [4]. These advances in antennas enable various kinds of flexibility, one of which is *position flexibility* that is believed to be game changing.

The paradigm of fluid antenna embraces all forms of movable and non-movable flexible-position antennas. Apparently, liquid-based antennas are movable while pixel-based antennas are not. A detailed coverage of the types of fluid antennas can be found in [2, Secs. 2.1 & 3.5]. It is worth pointing out that movable antennas may have limited use because to impact the physical layer, it will require a position change in the order of 10 cm in milliseconds, which equates to having an acceleration of 2×10^5 ms^{-2}.[2] Ignoring whether such unreal acceleration is practically possible, such antenna would be too dangerous to be allowed on a handset.[3] For this reason, the only feasible option for FAS is using pixel antennas (see Fig. 1), or packing many sub-wavelength metamaterial-based small antennas[4] in space. Fig. 1 shows the channel hardening phenomenon when using FAS as a receiver with different sizes. The channels for FAS are not typical channels as are traditional fixed-position MIMO systems. In FAS, the channels at different positions can be strongly correlated to each other. Under rich scattering, the spatial correlation follows the well-known Jake's or Clarke's model. For more information on the model, see [3].

To some people, Fig. 1 might look like placing a mini-RIS onto a handset. It is worth realizing that the working principles of RIS and FAS are fundamentally different. RIS operates on passive elements while FAS is based on active elements. FAS is hence closer to MIMO than RIS. As explained earlier, FAS brings position flexibility to liberate MIMO and makes MIMO much better, see e.g., TABLE I. Another point is that in FAS, each pixel is much smaller than a normal antenna which may be formed by joining several pixels together. Architecturally, RIS and FAS are thus not similar. Their distinct characteristics makes them suitable to different setups. While RIS is more of an intelligent repeater situated between BS and UE, FAS can form part of the BS, UE and RIS if flexible-position elements are deployed. Therefore, RIS and FAS are complementary, not mutually exclusive. Also, one possible application of FAS is to simplify the processing of RIS [2, Sec. 3.1].

B. Opportunities arising from FAS

1) Precoding: It goes without saying that MIMO is absolutely essential for mobile communications. Space-time codes and beamforming were the main techniques when MIMO was first introduced. As low-mobility, indoor UEs have begun to dominate and mobile networks move to small cells, precoding becomes more realistic as accurate CSI is possible at the BS. Nowadays precoding is a core technology for 5G and has been considered by many to synergize with other technologies.

[2]This is same as the acceleration of a bullet leaving the muzzle of a rifle with a speed of 600 ms^{-1}, with a 0.9 m long barrel.

[3]That said, movable antennas can be of great interest in situations where the position change responds only to statistical channel knowledge.

[4]A compact massive antenna array can also be viewed as a form of FAS.

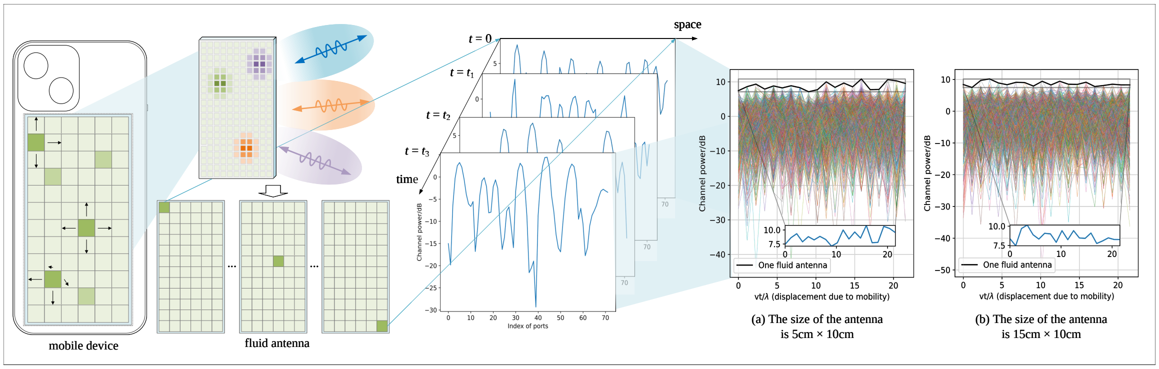

Fig. 1. FAS in the form of pixels or densely-packed small antennas, with results on the right illustrating its channel hardening effects at 39 GHz and 50×20 ports. Increasing the size from 5 cm $\times$ 10 cm to 15 cm $\times$ 10 cm decreases the variance of channel power from 1.18 to 0.47.

Upscaling the number of BS antennas is an obvious way to give mobile networks more power to serve more UEs, deliver higher rates and more. Nonetheless, this normally comes with an increased number of RF chains that can be very costly. FAS hence offers an interesting alternative that comes with additional spatial diversity but not necessarily the RF chains. Given the same number of RF chains, MIMO-FAS delivers greater diversity and capacity than fixed-position MIMO [8]. As a consequence, replacing fixed-position MIMO by MIMO-FAS in communication systems, will add essential degree-of-freedom (DoF) to enhance wireless system performance.

In Fig. 2, a vision that illustrates how FAS can be adopted to improve wireless system performance under different application scenarios, is presented. For example, we see that FAS can be used at the BS to liberate MIMO for improved precoding to UEs and/or RISs. Additionally, it is also possible to equip RIS with FAS technology on an unmanned aerial vehicle (UAV) to strengthen the channel from the BS to UEs when the UAV serves as a mobile RIS or relay. Intriguingly, FAS in RIS can play a key role in providing physical layer security. Note that traditional RIS uses precoding to construct a secrecy channel for reducing information leakage to eavesdroppers but its effect goes down when the UE and the eavesdropper locate along the same spatial direction. FAS in this case can give the needed dimension to differentiate the UE and eavesdropper's channels. Clearly, FAS can also be useful in vehicular communications when it is deployed at vehicles to provide reliable connections to mobile infrastructures. ISAC will benefit from the adoption of FAS as well because MIMO-FAS can have more DoF to balance between sensing and communication performance. In Section V, we will use ISAC as a case study to help understand the great potential of FAS over conventional MIMO.

2) Fading/Interference-immune Receiver: The fading phenomenon is widely regarded as the curse for wireless communications. It leads to an unpredictable signal fluctuation at the receiver end, causing the signal to disappear in deep fades. In fixed MIMO, the principle has been to mix the signals received at multiple fixed locations in a clever way to produce a more stable, stronger signal. FAS on the other hand provides a novel alternative, one that collects a large number of correlated signal samples in space and then selects the best signal in a manner of selection diversity.[5] It is well known that FAS has a decent channel hardening effect if the position resolution is good and the size is reasonable [6], as also depicted in Fig. 1. Unlike MIMO, FAS works even with only one RF chain. Evidently, if FAS is added to MIMO, the power is even greater [8].

Besides, FAS provides a whole new way of handling interference. The traditional understanding is that interference is either avoided or eliminated. Zero-forcing precoding utilizing fixed MIMO is an example of avoiding interference at designated receivers while successive interference cancellation lets interference exist but attempts to subtract (or eliminate) it from the receivers after the interference is estimated. Both methods work really well under ideal conditions. However, precoding requires accurate CSI and enough DoF at the transmitter side whereas interference cancellers will fail if there are too many interferers. By contrast, FAMA by adopting FAS looks at the interference mitigation problem in a very different way. FAMA is a receiver-based approach that needs neither precoding nor CSI at the transmitter side, and does not perform interference cancellation. Also, one RF chain is enough to operate.

The rationale behind FAMA is that with fading, signals in the spatial domain go up and down naturally, and it is possible to find a spatial window in which the aggregate interference suffers from a deep fade but the desired signal prevails. FAS gives UE the ability to access such opportunity to prevent from interference. FAMA was first proposed in [9] when the FAS was assumed to track and switch to the position in which the ratio of the instantaneous energy of the desired signal to that of the sum-interference plus noise signal was maximized on a per symbol basis. Recently in [10], a more practical scheme, referred to as slow FAMA, that limited the position change to only once during the channel coherence time, was proposed. Both approaches can deal with interference without precoding nor interference cancellation, with fast FAMA [9] capable of supporting a large number of UEs on the same channel while slow FAMA [10] can still handle several UEs.

This is shown by the results in Fig. 3 where an interference channel is considered and each transmitter uses a fixed antenna but each receiver is equipped with FAS. As we can see, for both slow and fast FAMA, the larger the size of FAS or the

[5]Different from conventional selection diversity, in FAS, the number of signal samples involved is massive and they are generally correlated.

Fig. 2. A vision showing pervasive use of FAS in wireless networks for different applications.

more the number of flexible positions (i.e., ports), the higher the network sum-rate. Additionally, the results show that slow FAMA can support 5 or 6 UEs on the same channel but will find it difficult to cope if there are too many UEs. The situation for fast FAMA is, however, very different. The results indicate that for fast FAMA, the sum-rate continues to increase even with hundreds of UEs, suggesting that fast FAMA is able to deal with a massive number of interferers.

Presumably, there will be many situations in wireless networks where FAMA can play a role in managing interference. FAMA can help remove some burden from the BS in terms of overheads for CSI acquisition. Furthermore, FAMA would be particularly useful in mitigating out-of-cell interference where accurate CSI is just too difficult to obtain.

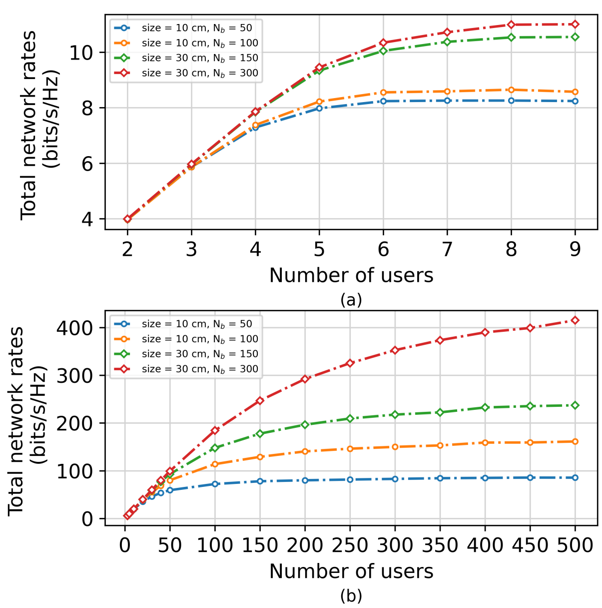

Fig. 3. Data rates of (a) slow FAMA [10] and (b) fast FAMA [9] against the number of UEs with different sizes and resolutions of FAS at each UE. One-dimensional FAS with N_b flexible positions is assumed at each UE. Also, the data rate is computed assuming binary symmetric channels under uncoded quadrature phase shift keying (QPSK) transmissions. The operating frequency is 39 GHz and the channel has a Rice factor of 7 with two scattered paths.

III. Design Challenges

The great potential of FAS inevitably comes with considerable challenges. Here, we briefly cover three main ones.

A. Estimation of Nearly Continuous CSI

Position flexibility in FAS not only gives a new dimension for performance enhancement but also the need to acquire the CSI in this additional dimension. A high-resolution FAS would mean that a nearly continuous CSI function in space needs to be acquired so that FAS can be intelligently optimized. In this case, if conventional estimation methods are to be used, it will incur an unbearable system overhead. Imagine if FAS needs to switch amongst a massive number of positions to estimate the CSI, then it will take a very long time and by the time the process is completed, the CSI may have already changed. As a result, a more feasible approach would be to estimate the CSI for only a small number of FAS positions and then adopt extrapolation techniques to predict the rest [11]. To this end, deep learning-based extrapolation stands out for its data-driven nature and model-free characteristics. Unlike traditional approaches, it can intelligently learn the channel correlations in real-world scenarios, and mitigate the generalization issues between theoretical models and practical environments.

B. Antenna Position Optimization

Antenna position optimization, a.k.a. port selection, aims to find the best position out of all the accessible positions of FAS provided the CSI is known. This is also the basic operation of FAS. Normally, if CSI is available and the performance metric is easily accessible, port selection will be a straightforward task. Unfortunately, given the high dimensionality of FAS, this is non-trivial. It is also much more challenging if the metric is not directly computable. For example, in fast FAMA, each UE should calculate the ratio of the instantaneous energy of the desired signal to that of the sum-interference and noise signal at all the ports, based on the received signals at all the positions. Estimating these ratios itself is a challenge, not to mention the difficulty of obtaining the received signals at all the positions if we only allow to observe received signals from a few positions. On the other hand, if FAS is equipped with multiple RF chains, then several ports can be activated and port selection will involve choosing and activating a number of best ports simultaneously. For this reason, the powerful fitting capability of deep learning should hold promise to directly fit the optimal port positions. By adopting a data-driven approach, it bypasses the need to solve extremely challenging signal decomposition and combination optimization problems.

C. Joint Antenna Position and Beamforming Optimization

For MIMO-FAS, the great benefits come from jointly optimizing the antenna positions and the beamforming matrices. With FAS at both ends, the beamforming design and multi-port selection are strongly coupled that makes the joint optimization problem extremely challenging to solve. To alleviate this, approximate optimization techniques are often used to simplify the objective function and attempt to convexify the problem. What approximations would be appropriate and how efficient they are to obtain valid solutions in MIMO-FAS, are not well understood. The challenges are also more severe if multiuser scenarios and/or other technologies are considered. Due to the strong coupling of variables and the complexity of the solution space, it is difficult for traditional optimization methods to obtain an efficient solution. Fortunately, deep reinforcement learning (DRL) can uncover solutions from extensive datasets generated through interactions with environments, bypassing the intractability of nonconvex optimization problems.

IV. Learning-aided Solutions

This article sees the great potential of using deep learning to provide tractable approaches to handle the challenges in the design of FAS-aided wireless communications. Here, we shed light on how deep learning can be used to optimize FAS.

A. Deep Learning for Channel Extrapolation

The extreme high-dimensionality in the CSI of FAS causes concern if FAS can indeed be useful. Without the fine resolution in space, FAS would lose its ability to obtain rich diversity and mitigate interference. The spatial correlation amongst the CSI in the spatial domain, fortunately, gives a lifeline to FAS. Channel extrapolation provides a tractable solution, which can predict the full CSI based on a partial one [11]. To this end, compressed sensing and deep learning are widely used, where compressed sensing based approaches exploit channel sparsity to recover the CSI under the approximation error constraint. Its implementation, however, relies on the sparse structure of the channel, which appears only if there are only a small number of propagation paths. By contrast, deep learning is data-driven, which can learn the nonlinear mapping from partial CSI to full CSI, ideal for CSI prediction. For FAS, the strong correlation between the CSI at different positions will allow extrapolation to accurately derive the full CSI using deep learning.

To provide an initial evaluation of this idea, we adopt the masked autoencoder (MAE) in [12] for channel extrapolation in FAS based on partially observed ports. The block diagram of the MAE estimator is illustrated in Fig. 4, which includes two key blocks: encoder and decoder, with the “encoder” block responsible for learning the latent representation of partial CSI and the “decoder” block able to learn the mapping from the latent representation to the full CSI. Specifically, we employ the vision transformer-based encoder including linear projection, position encoding, and transformer modules, where the position encoding is used to characterize the spatial relationship of the CSI at different ports and the transformer module uses the multi-head self-attention mechanism to capture the fine-grained interrelationships among the spatial signatures. In Fig. 5, the normalized minimum square error (NMSE) results of the predicted CSI versus the number of observed ports are shown for the FAS with 72 ports. The results demonstrate that observing less than one-third of the ports would be sufficient to obtain decent performance though a larger-sized FAS should need more port observation due to weaker correlation.

B. Learning-induced Multi-port Selection

Multi-port selection in FAS is a very challenging combinatorial optimization problem [2]. First, the huge search space makes exhaustive search inapplicable. Also, existing methods tend to have poor scalability and may fall into a local optimal solution whose performance is not assured. Though supervised learning can find a meaningful relationship between the CSI and the optimal ports, they cannot be applied to variable sized FASs, which will limit its practical significance.

To tackle this, pointer network has emerged as a new neural network architecture that can provide a learning-induced approach with great generalization capability. Pointer network has been confirmed to be effective in handling combinational optimization problems, e.g., the travelling salesman problem [13]. Therefore, we can employ a pointer network to optimize

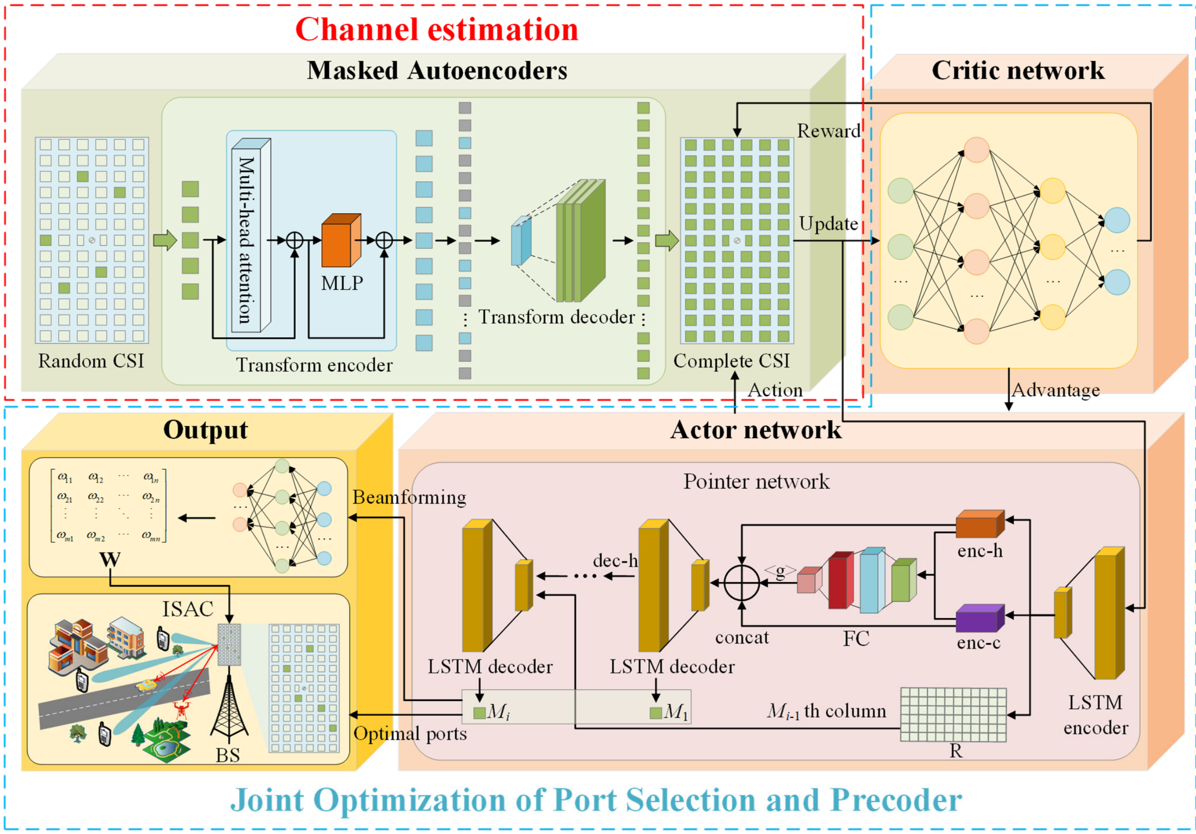

Fig. 4. The deep learning framework for channel estimation and optimization for FAS in the downlink ISAC channels.

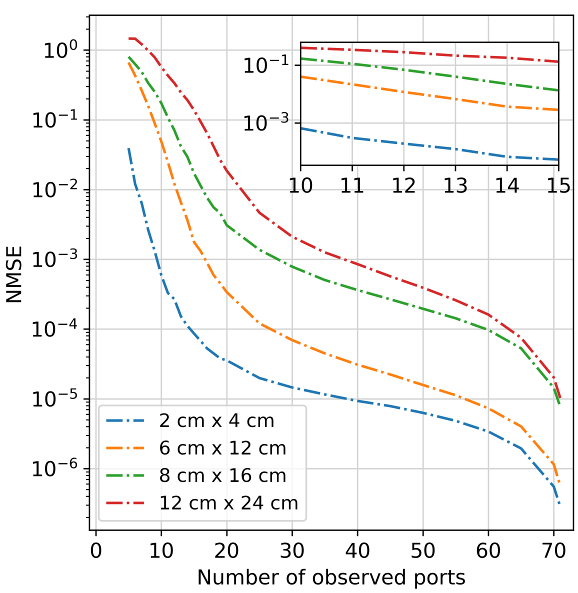

Fig. 5. CSI extrapolation performance of the MAE-based CSI estimator for a rectangular FAS with 72 available ports, at 3.4 GHz.

the port selection for FAS of variable sizes [13]. As depicted in Fig. 4, the pointer network adopts the content based input attention model including encoding and decoding modules. The long short-term memory (LSTM) encoder first transforms the CSI into an embedding matrix $\mathbf{R}$, the hidden state $\mathrm{enc\text{-}h}$, and the cell state $\mathrm{enc\text{-}c}$. Subsequently, a fully-connected network adopts $\mathrm{enc\text{-}h}$ and $\mathrm{enc\text{-}c}$ as the input to learn the training parameters <g>. Then the first LSTM decoder concatenates <g>, $\mathrm{enc\text{-}h}$, and $\mathrm{enc\text{-}c}$ to form the input and uses an attention-based pointing mechanism to learn the optimal port of the first fluid antenna, i.e., M_1 [14]. Afterwards, the i-th decoder adopts the M_{i-1}-th column of $\mathbf{R}$ and the latent memory state of the $(i-1)$-th decoder $\mathrm{dec\text{-}h}$ as the input to select the port for the succeeding fluid antenna, and so on. This technique will be considered and tested in the ISAC case study later.

C. DRL for Multiuser Multi-port Selection and Precoding

For multiuser MIMO-FAS in the downlink,[6] the port selection at the BS and UEs, the precoding matrix at the BS and the beamforming matrices at all the UEs need to be optimized jointly. Supervised learning could be effective to acquire the associations that exist between the input features (e.g., CSI) and their corresponding output labels (i.e., the joint solution). Unfortunately, such labeled training data do not exist. For this reason, reinforcement learning is a more suitable tool that can bypass the need of labeled training data.

Of great relevance is the work of [14] that proposed a neural combinatorial optimization approach utilizing reinforcement learning, which can be adopted to build an end-to-end learning framework to learn the stochastic policy for assigning a high

[6]Downlink is considered here as an example. Similar challenges arise in uplink and interference channels.

probability to the activated ports and precoder that maximize the sum-rate. In particular, we can use the pointer network to first learn the activation probability of different ports, which can be used later to generate the port selection with sampling. With the activated ports, we can employ the unsupervised learning approach to design the neural precoding network to build the relationship between the input CSI and the optimized precoder. Finally, it is possible to apply the advantage actor and critic (A2C) based DRL approach to establish an end-to-end learning framework including the two neural networks, see Fig. 4. We consider the CSI of all current ports as states in reinforcement learning, with the sum rate as the reward. The action is defined based on a specific task. For instance, in the port selection task, the action corresponds to the selected port. In the joint design task of port selection and precoding, the action may also include the precoding matrix.

Note that the critic network uses a deep neural network to approximate the excepted reward obtained by the policy while the actor network involves the pointer network and the neural precoding network to design the optimal policy. Moreover, the advantage function is adopted as the loss function which is the difference between the action-value function and the baseline function that is approximated by the critic network. Different from supervised learning, the A2C based DRL approach does not need labeled training data, but uses interactive feedback from the environment to train the neural networks.

It is important to note that the instability of reinforcement learning during the training phase significantly impacts the convergence of the model [15]. Combining ensemble learning with reinforcement learning is expected to mitigate learning variance, thereby significantly improving training stability and expediting model convergence. Specifically, the critic quantitatively scores the actor's strategies, known as Q-values, with current research revealing significant variance in the predicted Q-values. It is expected that employing ensemble learning to merge evaluations from various critic model architectures can address the issue of excessive Q-values variance, expediting model convergence. Moreover, the ϵ-greedy strategy that encourages actor exploration by randomly selecting actions, can help prevent the policy from converging prematurely. Despite this, its inherent randomness hampers exploration efficiency, potentially slowing down model convergence. Manual tuning of ϵ at different stages is however required to achieve model convergence. Leveraging the Bagging method in ensemble learning to train multiple actors is expected to realize adaptive exploration, a fine-tuning-free training process and robust strategy generation through data and model diversity.

V. Case Study: FAS-aided ISAC

To understand the effectiveness of the learning-aided FAS described above, we consider a downlink multiuser multiple-input single-output (MISO)-FAS for ISAC. There is a FAS-aided BS, multiple UEs and sensing targets each with a single fixed antenna. The objective is to maximize the network sum-rate subject to a radar sensing power constraint by optimizing the multi-port selection and the precoding matrix at the BS.

Fig. 4 illustrates a typical deep learning-aided joint optimization scheme for port selection and precoding. It comprises a MAE model for channel estimation, an A2C based DRL for port selection, and a learning-induced neural network for precoding design. Firstly, the MAE extrapolates the CSI of other ports from the observed port CSI. Then DRL is adopted to perform multi-port selection using the completed CSI. Finally, the precoding network utilizes the selection as input to optimize precoding. The A2C based DRL approach adopts the feedback from multiple downlink UEs to jointly design the pointer and neural precoding networks via gradient back-propagation. However, the sensing constraint poses challenges to implementation. Existing works usually utilize a penalized deep learning approach which introduces a penalty parameter to involve the constraints into the loss function. Nevertheless, there are no tractable approaches to design the penalty parameter, and inappropriate parameters can deteriorate the performance seriously. In order to devise an efficient design, we propose a primal-dual based unsupervised deep learning network to handle the constrained precoder design problem. In particular, we introduce the Lagrange multiplier to involve the constraint into the loss function and obtain a Lagrange dual function. Then, using the alternating gradient ascent approach, the Lagrange multiplier and the neural precoding network parameters are optimized iteratively to minimize the Lagrange dual function. Consequently, the neural precoding network can be trained to maximize the downlink sum-rate subject to the sensing power constraints, which is built in with the capability of mastering the constrained optimization problem.

Fig. 6 shows the sum-rate results of the multiuser MISO-FAS using the proposed learning-based approach against the percentage of unobserved ports for different physical sizes of FAS. The FAS is optimized for ISAC with a radar sensing

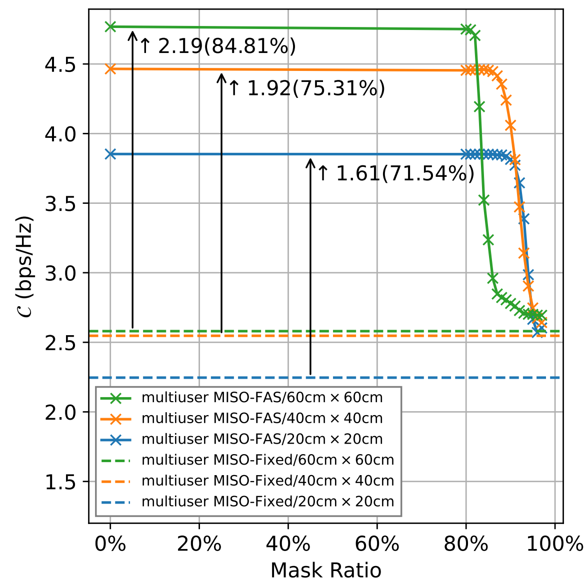

Fig. 6. Network sum rate of a downlink 3-user MISO-FAS using the learning structure in Fig. 4 against the percentage of unobserved ports with one target having the radar sensing threshold of -3 dB. The operating frequency is 800 MHz. The power budget at the BS is 15 dBm while the noise power at each UE is -80 dBm. The FAS at the BS has 10×10 flexible positions to select 5 activated ports. The UEs are located at different distances from the BS uniformly distributed within the range of 40 m to 60 m.

power constraint. The mask ratio indicates the proportion of ports that are not observed and require channel extrapolation. It is evident that fewer observed ports will result in increased prediction channel extrapolation errors, leading to a reduction in the sum-rate performance. Fortunately, owing to the spatial correlation of the channel, observing just 20% of the ports is adequate to achieve performance comparable to observing all ports. The larger the size of the FAS, the greater the number of observed ports required due to weaker spatial correlation. Nevertheless, the rate performance of MISO-FAS is better with larger size for spatial diversity. The performance of multiuser MISO with fixed antennas is also included in the figure for comparison. As shown, the results illustrate significant sum-rate gains of the proposed multiuser MISO-FAS over the fixed-position MISO counterpart, confirming the great ability of the learning method to take advantages of FAS.

VI. Future Directions

There are many other situations that can benefit greatly from FAS but this article is unable to cover them due to insufficient space. Some future research directions are listed below:

- A natural and important extension will be to study the joint optimization of a multiuser MIMO-FAS where both BS and UEs are equipped with flexible-position antennas. The CSI estimation process will be an order of magnitude harder to do, and so will be the overall optimization.
- A much larger FAS surface is possible in UAV or RIS and will be another interesting area with great potential to explore. The FAS-aided UAV or RIS will be able to achieve much more, but the challenges will need to be addressed. Similarly, it will be intriguing to learn how beneficial it is to incorporate FAS into other emerging technologies such as non-orthogonal multiple access (NOMA).
- It will also be interesting to see if the high-dimensionality of FAS would be useful in coding design. A differential coding scheme using FAS can be an effective approach to benefit from the diversity while reducing the burden for CSI estimation in fast fading scenarios. Also, the high-dimensionality in the FAS channel can be a useful feature to design secrecy codes for information security.
- The fact that the required speed to change the antenna's position may be unrealistic, suggests that compact pixel-based switchable antennas present a more realistic future for flexible-position antennas. However, in this case, the correlation among different pixel configurations will need to be properly examined, in light of mutual coupling.
- Last but not least, it is important to investigate how each FAS-aided UE could estimate the instantaneous energy ratio to obtain the optimal position on a per-symbol basis for fast FAMA under practical conditions.

VII. Conclusions

With the new paradigm of FAS providing additional DoF to enhance wireless communications performance using flexible-position antennas, this article discussed the opportunities and challenges. In particular, we highlighted the unique characteristics of FAS that promises superb diversity and interference mitigation capability. But concerns in acquiring the nearly continuous CSI and the complexity for finding the best position(s) and precoding are valid. This article has presented a machine learning framework that injects the FAS-assisted terminal (BS or UE) the ability to estimate the CSI with partial observations and optimize port selection and precoding jointly. MAE, DRL and pointer network are the main tools that make possible the overall system design. We then used ISAC as an application scenario to showcase the great power of multiuser MISO-FAS. The results revealed considerable gains in the sum-rate over fixed MISO with the same sensing performance.

References

[1] Y. Dong and T. Itoh, "Metamaterial-based antennas," *Proc. IEEE*, vol. 100, no. 7, pp. 2271–2285, Jul. 2012.

[2] K.-K. Wong, K.-F. Tong, Y. Shen, Y. Chen, and Y. Zhang, "Bruce Lee-inspired fluid antenna system: Six research topics and the potentials for 6G," *Front. Commun. and Netw.*, no. 3, pp. 1–31, Mar. 2022.

[3] K.-K. Wong, W. K. New, X. Hao, K.-F. Tong and C.-B. Chae, "Fluid antenna system—Part I: Preliminaries," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 1919–1923, Aug. 2023.

[4] L. Jing, M. Li and R. Murch, "Compact pattern reconfigurable pixel antenna with diagonal pixel connections," *IEEE Trans. Antennas & Propag.*, vol. 70, no. 10, pp. 8951–8961, Oct. 2022.

[5] K. K. Wong, A. Shojaeifard, K.-F. Tong, and Y. Zhang, "Performance limits of fluid antenna systems," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2469–2472, Nov. 2020.

[6] K.-K. Wong, A. Shojaeifard, K.-F. Tong and Y. Zhang, "Fluid antenna systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1950–1962, Mar. 2021.

[7] M. Khammassi, A. Kammoun, and M.-S. Alouini, "A new analytical approximation of the fluid antenna system channel," *IEEE Trans. Wireless Commun.*, vol. 22, no. 22, pp. 8843–8858, Dec. 2023.

[8] W. K. New, K. K. Wong, X. Hao, K. F. Tong, and C. B. Chae, "An information-theoretic characterization of MIMO-FAS: Optimization, diversity-multiplexing tradeoff and q-outage capacity," *IEEE Trans. Wireless Commun.*, DOI:10.1109/TWC.2023.3327063, 2023.

[9] K.-K. Wong and K.-F. Tong, "Fluid antenna multiple access," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4801–4815, Jul. 2022.

[10] K.-K. Wong, D. Morales-Jimenez, K.-F. Tong and C.-B. Chae, "Slow fluid antenna multiple access," *IEEE Trans. Commun.*, vol. 71, no. 5, pp. 2831–2846, May 2023.

[11] S. Zhang *et al.*, "Deep learning based channel extrapolation for large-scale antenna systems: Opportunities, challenges and solutions," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 160–167, Dec. 2021.

[12] K. He *et al.*, "Masked autoencoders are scalable vision learners," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR)*, New Orleans, USA, 2022, pp. 16 000–16 009.

[13] O. Vinyals, M. Fortunato, and N. Jaitly, "Pointer networks," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 28. Red Hook, NY, USA: Curran, 2015.

[14] ——, "Neural combinatorial optimization with reinforcement learning," in *Proc. Int. Conf. Learn. Representations*, Toulon, France: OpenReview.net, 2017, pp. 1–5.

[15] P. V. R. Ferreira *et al.*, "Multiobjective reinforcement learning for cognitive satellite communications using deep neural network ensembles," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 5, pp. 1030–1041, May 2018.

Biographies

Chao Wang is an Associate Professor with the School of Telecommunications Engineering, Xidian University, China. He is an Editor of IEEE Wireless Communications Letters and IEEE Open Journal of the Communications Society.

Zan Li is Professor at the School of Telecommunications Engineering, Xidian University, China. She is Fellow of the IET, the China Institute of Electronics (CIE), and the China Institute of Communications (CIC). She serves as an

Associate Editor for the IEEE Transactions on Cognitive Communications and Networking and China Communications.

Kai-Kit Wong is Chair in Wireless Communications at the Department of Electronic and Electrical Engineering, University College London, UK. His current research centers around 6G communications. He is Fellow of IEEE and IET.

Ross Murch is Chair Professor in the Department of Electronic and Computer Engineering and an Institute for Advanced Study (IAS) Senior Fellow at the Hong Kong University of Science and Technology. He is a Fellow of IEEE, IET and HKIE. He has also served as Editor and Chair of several international journals.

Chan-Byoung Chae is an Underwood Distinguished Professor with the School of Integrated Technology, Yonsei University, Korea. He served as an Editor-in-Chief of the the IEEE Trans. on Molecular, Biological, and Multi-scale Communications from 2019 to 2022. He has also served as an Editor of several international journals. He is an IEEE Distinguished Lecturer (ComSoc) and an IEEE Fellow.

Shi Jin is Professor with the faculty of National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. He serves as an Area Editor for the IEEE Transactions on Communications and IET Electronics Letters. He has been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory.