GFlowNet-Based Antenna Selection for ISAC Systems under the Presence of Eavesdroppers

Spilios Evmorfos
Electrical and Computer Engineering
Rutgers, the State University of New Jersey
New Brunswick, USA
spilios.evmorfos@rutgers.edu

Athina P. Petropulu
Electrical and Computer Engineering
Rutgers, the State University of New Jersey
New Brunswick, USA
athinap@soe.rutgers.edu

2024 58th Asilomar Conference on Signals, Systems, and Computers | 979-8-3503-5405-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/IEEECONF60004.2024.10942636

***Abstract*—This work addresses Physical Layer Security (PHY) in Integrated Sensing and Communication (ISAC) systems through antenna selection. We consider a MIMO transmitter communicating with a legitimate multi-antenna receiver while simultaneously sensing a target, with a multi-antenna eavesdropper intercepting the communication. Previous studies show that increasing the number of active antennas improves the ergodic secrecy rate up to a certain point, after which performance deteriorates. The instantaneous secrecy rate, for fixed channel conditions, depends not only on the number of active antennas but also on their positions. This work focuses on optimizing both the number and positions of active antennas to maximize the instantaneous secrecy rate, while ensuring that the Cramer Rao Bound (CRB) of the sensing functionality remains below a specified threshold. We propose an unsupervised approach based on Generative Flow Networks (GFlowNets), modeling antenna selection as a deterministic Markov Decision Process (MDP) where terminal states correspond to active antenna subsets of all sizes and configurations. GFlowNets parametrize the action-sampling distribution, ensuring the probability of reaching a final antenna subset configuration is proportional to its combination of the resulting secrecy rate and the CRB.**

***Index Terms*—ISAC, PHY, antenna selection, GFlowNets**

I. Introduction

The 6th generation of wireless networks promises transformative applications such as smart cities, connected vehicles, and smart manufacturing [1]. These applications require both reliable communication and advanced sensing perception of the environment. Since mmWave sensing and communication share frequency bands and hardware, this has led to the Integrated Sensing and Communication (ISAC) [2] paradigm, where both functions are jointly optimized on a single platform.

Security is a key concern in ISAC systems, especially when sensitive information is exchanged. Wireless security can be addressed from two perspectives: cryptography at higher protocol layers and Physical Layer Security (PHY) [3] at lower layers, which exploits physical properties of the wireless channel through signal processing techniques. This work focuses on enhancing PHY in ISAC while maintaining its sensing performance.

Work supported by ARO under Grants W911NF2110071 and W911NF2320103 and by NSF under Grant ECCS-2335876

Although the PHY aspect of ISAC is relatively new, some studies have addressed it from various angles. For example, [4] proposes an approach for ISAC systems with multiple eavesdroppers, using Capon approximate maximum likelihood estimation to locate the eavesdroppers and optimize a combination of the secrecy rate and Cramer-Rao Bound (CRB). In [5], the authors combine successive convex approximations and Block Coordinate Descent (BCA) to design the beamforming matrix in order to optimize both the secrecy rate and CRB of an ISAC system.

While previous research on antenna selection for ISAC systems exists, it has not focused on PHY. For instance, [6] uses deep reinforcement learning in an automotive ISAC scenario to select active transmit antennas and design a beamforming matrix that focus the maximum amount of energy on both the communication receiver and the tracked target. [7] explores antenna selection for ISAC systems with a single communication receiver and multiple radar targets, combining convex optimization and dynamic programming to balance communication rate and radar tracking accuracy.

Our work addresses antenna selection for ISAC systems from a PHY perspective, a topic which, to the best of our knowledge has not been previously explored. In [8], the authors show that in a MIMO system, the ergodic secrecy rate increases as more antennas are activated, up to a point where adding additional antennas degrades secrecy performance. We observe a similar phenomenon in a related ISAC system (a MIMO transmitter, multi-antenna receiver, eavesdropper, and radar target). Additionally, for instantaneous secrecy rate (fixed channel realizations), we find that performance depends on both the number of active antennas and their specific positions.

The current work proposes an unsupervised method for selecting both the number and positions of active antennas to maximize secrecy rate while maintaining the CRB of the radar component below a threshold. Antenna selection is framed as a deterministic Markov Decision Process (MDP). The root state of the MDP is represented by a vector, where each element is a placeholder that signifies an undefined antenna position. This vector has the same size as the transmit antenna array. At the start of the process, all antenna positions are undefined, indicating that it is unknown whether each antenna will be active or inactive. As actions are taken, one by one, these

placeholder values are replaced by binary values—either active (1) or inactive (0)—corresponding to specific antenna positions. Terminal states are reached when all antenna positions are defined, i.e., when the vector contains no more placeholder values. Rewards are assigned only to terminal states based on a combination of the secrecy rate and the CRB of the radar component. Intermediate states, where some antenna positions remain undefined, are not rewarded. The Generative Flow Network (GFlowNet) [9] is employed to parametrize an action-sampling policy, ensuring that the probability of reaching a terminal state (i.e., a specific configuration of active/inactive antennas) is proportional to its reward.

Previous works on GFlowNets for antenna selection, such as [10] and [11], have focused on selecting a fixed apriori defined number of antennas to optimize beampatterns or a combination of communication rate and CRB. In contrast, our approach optimizes both the number and configuration of active antennas, which requires a different MDP formulation. Unlike prior work where terminal states correspond to subarrays of a predefined size, our MDP allows terminal states to represent subarrays of all possible sizes.

Notation: We denote the matrices and vectors by bold uppercase and bold lowercase letters, respectively. The operators $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{H}$ denote transposition and conjugate transposition respectively. Caligraphic letters will be used to denote sets. The ℓ_p-norm of $\boldsymbol{x} \in \mathbb{R}^n$ is $\|\boldsymbol{x}\|_p \triangleq \left(\sum_{i=1}^{n} |x(i)|^p\right)^{1/p}$, for all $\mathbb{N} \ni p \geq 1$. The expectation of a random vector $\mathbf{x}$ is denoted as $\mathbb{E}(\mathbf{x})$. Continuous sets are denoted by $[\cdot]$ while discrete sets are denoted by $\{\cdot\}$. The logarithm of the determinant of a matrix $\mathbf{A}$ is denoted as $\log|\mathbf{A}|$.

II. System Model

We consider an ISAC system such as the one depicted in Fig. 1. It consists of one source with N_s antenna elements and communicates with a destination that is equipped with N_d antenna elements. At the same time the system tracks a target and the communication is being overheard by an eavesdropper that is equipped with N_e antennas. The source employs a precoding matrix $\mathbf{F} \in \mathbb{C}^{N_s \times N_d}$. The channel from the source to the legitimate receiver is denoted as $\mathbf{H}_d \in \mathbb{C}^{N_d \times N_s}$. The transmitter performs hybrid precoding and the transmitted signal is:

$$\mathbf{V} = \mathbf{F}\mathbf{X} \tag{1}$$

where $\mathbf{X}$ consists of N_s unit power streams of length L, which means $\frac{1}{L}\mathbb{E}\{\mathbf{X}\mathbf{X}^H\} = \mathbf{I}_{N_s}$.

Assuming that only a subset of the N_s antennas on the transmitter to be active, the selection vector is denoted as $\mathbf{x} \in \{0,1\}^{N_s}$. The corresponding selection matrix is $\mathbf{S}(\mathbf{x}) = \mathrm{diag}(\mathbf{x})$.

The secrecy rate of the sparse array is:

$$\mathcal{R}(\mathbf{x},\mathbf{F}) = \log_2|\mathbf{I}_{N_d} + \rho_d \mathbf{H_d}\mathbf{S}(\mathbf{x})\mathbf{F}\mathbf{F}^H\mathbf{S}(\mathbf{x}\mathbf{H}_d^H| - \log_2|\mathbf{I}_{N_e} + \rho_e \mathbf{H_e}\mathbf{S}(\mathbf{x})\mathbf{F}\mathbf{F}^H\mathbf{S}(\mathbf{x})\mathbf{H}_e^H|, \tag{2}$$

where ρ_d denotes the average SNR at each of the receiver's antennas while the ρ_e denotes the SNR at each of the eavesdropper's antennas. The corresponding CRB of the sparse array is :

$$\mathrm{CRB}(\mathbf{x},\mathbf{F}) = \frac{1}{L}\mathrm{trace}\left((\mathbf{S}(\mathbf{x})\mathbf{F}\mathbf{F}^H\mathbf{S}(\mathbf{x})^H)^{-1}\right) \tag{3}$$

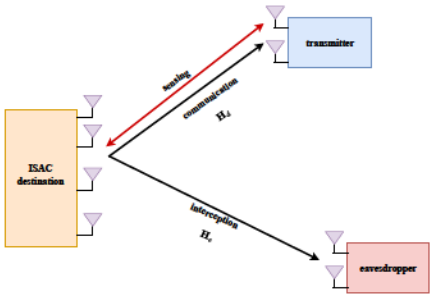

Fig. 1. The ISAC system that consists of a MIMO source, legitimate receiver that is also being tracked by the radar component and a multi-antenna eavesdropper that overhears the communication.

The problem formulation is to find the sparse selection vector $\mathbf{x} \in \{0,1\}^{N_s}$ and the precoding vector $\mathbf{F} \in \mathbb{C}^{N_s \times N_d}$ such that the secrecy rate of the system is maximized subject to the CRB being less than a predefined threshold.

$$\begin{aligned} &\text{maximize}_{\mathbf{x},\mathbf{F}} && \mathcal{R}(\mathbf{x},\mathbf{F}) \\ &\text{subject to} && \mathrm{CRB}(\mathbf{x},\mathbf{F}) \leq \mathrm{CRB}_{low} \end{aligned} \tag{4}$$

III. GFlowNets for Antenna Selection

A. Overview of the GFlowNet Framework

This section presents an overview of the GFlowNet framework [12]. Consider a deterministic MDP, where $\mathcal{S}$ is the set of all states, and $\mathcal{X} \subset \mathcal{S}$ is the set of terminal states. Let $\mathcal{A}$ denote the set of discrete actions, and $\mathcal{A}(\mathbf{s})$ represent the set of permissible actions at a given state $\mathbf{s}$. The MDP can be modeled as a Directed Acyclic Graph (DAG), with a unique root node $\mathbf{s}_0$ and terminal states serving as the leaf nodes.

In this setting, each terminal state is associated with a positive reward, while intermediate states yield zero reward ($\mathcal{R}(\mathbf{s}) = 0, \forall \mathbf{s} \notin \mathcal{X}$). The DAG is non-injective, meaning that different sequences of actions starting from the root may lead to the same state. The main objective is to learn an action-selection policy such that the probability of reaching a terminal state is proportional to its reward.

GFlowNet frames this MDP as a flow network, where flow originates at the root and sinks at each terminal state. If action $\mathbf{a}$ is taken in state $\mathbf{s}$, the resulting state is denoted by $T(\mathbf{s},\mathbf{a}) = \mathbf{s}'$. Given the deterministic nature of the MDP, $T(\mathbf{s},\mathbf{a})$ is uniquely determined for each pair. The flow along the edge $(\mathbf{s},\mathbf{a})$ is represented by $F(\mathbf{s},\mathbf{a})$, and the total flow passing through state $\mathbf{s}$ is denoted as $F(\mathbf{s})$.

To maintain flow balance, the inflow at each state must equal the outflow. The inflow to a state $\mathbf{s}'$ is:

$$F(\mathbf{s}') = \sum_{\mathbf{s},\mathbf{a}:T(\mathbf{s},\mathbf{a})=\mathbf{s}'} F(\mathbf{s},\mathbf{a}). \tag{5}$$

Conversely, the outflow from a state $\mathbf{s}'$ is given by:

$$F(\mathbf{s}') = \sum_{\mathbf{a}' \in \mathcal{A}(\mathbf{s}')} F(\mathbf{s}', \mathbf{a}'). \tag{6}$$

For any non-terminal state $\mathbf{s}$, the total flow is the sum of its outgoing flows and its reward, which is:

$$F(\mathbf{s}) = R(\mathbf{s}) + \sum_{\mathbf{a} \in \mathcal{A}(\mathbf{s})} F(\mathbf{s}, \mathbf{a}). \tag{7}$$

For a terminal state $\mathbf{s}_f$, the flow is simply the reward:

$$F(\mathbf{s}_f) = R(\mathbf{s}_f) > 0. \tag{8}$$

Therefore, the flow-matching condition for each state $\mathbf{s}'$ can be expressed as:

$$\sum_{\mathbf{s},\mathbf{a}:T(\mathbf{s},\mathbf{a})=\mathbf{s}'} F(\mathbf{s}, \mathbf{a}) = R(\mathbf{s}') + \sum_{\mathbf{a}' \in \mathcal{A}(\mathbf{s}')} F(\mathbf{s}', \mathbf{a}'). \tag{9}$$

Given that the state flows $F(\mathbf{s})$ and edge flows $F(\mathbf{s}, \mathbf{a})$ satisfy this equation, the action selection policy at each state (starting from the root) can be derived as:

$$\pi(\mathbf{a}|\mathbf{s}) = \frac{F(\mathbf{s}, \mathbf{a})}{F(\mathbf{s})}. \tag{10}$$

According to [13], two key results follow:

1) The flow at the root (also called the partition function of the MDP) is the sum of all terminal state rewards:

$$F(\mathbf{s}_0) = \sum_{\mathbf{s}_f \in \mathcal{X}} R(\mathbf{s}_f).$$

2) The probability of reaching a terminal state $\mathbf{s}_f$ is proportional to its reward, relative to the partition function:

$$\pi(\mathbf{s}_f) = \frac{R(\mathbf{s}_f)}{\sum_{\mathbf{s}'_f \in \mathcal{X}} R(\mathbf{s}'_f)} = \frac{R(\mathbf{s}_f)}{F(\mathbf{s}_0)}.$$

The GFlowNet paradigm parameterizes the flow F using a function approximator, denoted as $F_{\mathbf{w}}$, chosen from a sufficiently expressive class, such as a deep neural network. Trajectories of the MDP are sampled, and for each state $\mathbf{s}'$, the flow-matching objective is minimized via gradient descent on the parameters $\mathbf{w}$:

$$L_{\mathbf{w}}(\mathbf{s}') = \sum_{\mathbf{s},\mathbf{a}:T(\mathbf{s},\mathbf{a})=\mathbf{s}'} F_{\mathbf{w}}(\mathbf{s}, \mathbf{a}) - R(\mathbf{s}') - \sum_{\mathbf{a}' \in \mathcal{A}(\mathbf{s}')} F_{\mathbf{w}}(\mathbf{s}', \mathbf{a}'). \tag{11}$$

B. GFlowNets for Antenna Selection

To apply the GFlowNet paradigm to solve problem 4, an MDP must be designed to model the antenna selection process. In previous works on antenna/sensor selection using GFlowNet [10], [11], the number of active sensors was predefined (selecting k out of m antennas), with the goal of choosing the optimal subset from the $\binom{m}{k}$ possible combinations. These MDPs began with a zero vector, and each action activated an antenna (switching a 0 to 1), leading to leaf states representing all possible subsets of k active antennas.

For problem 4, the selection vector $\mathbf{x}$ can have any l_0 norm, meaning the number of active antennas can vary from 0 to N_s. Therefore, the MDP is structured as follows: the root state ($\mathbf{s}_0$) is a vector of size N_s, where each element is initialized as $\emptyset$, indicating that the status of each antenna (active or inactive) is undefined. The action space consists of assigning each $\emptyset$ element to either 0 (inactive) or 1 (active). Each root-to-leaf trajectory consists of N_s actions, and the leaf states represent all possible binary vectors of size N_s, covering all antenna configurations ($\binom{m}{k}$ for $k \in \{0, m\}$). Finally, to define the reward function for the MDP, we begin by noting that all intermediate states receive a reward of 0. The objective is to maximize the secrecy rate while ensuring the CRB remains below a specified threshold. Since the reward inherently depends on both the selection vector $\mathbf{x}$ and the precoding matrix $\mathbf{F}$, we first need to determine the optimal precoding matrix $\mathbf{F}^*$. This is achieved by maximizing the following expression with respect to $\mathbf{F}$:

$$\mathbf{F}^* = \arg\max_{\mathbf{F}} \left[\mathcal{R}(\mathbf{x}, \mathbf{F}) - c \left(\mathrm{CRB}(\mathbf{x}, \mathbf{F}) - \mathrm{CRB}_{low} \right) \right], \tag{12}$$

where c is a scaling factor to balance the two terms. This is achieved by a fixed number of gradient ascent steps. Since the optimization matrix $\mathbf{F}$ is complex, we employ the Wirtinger conjugate gradient definition [14].

With $\mathbf{F}^*$ determined, the reward for a terminal state $\mathbf{s}_f$, corresponding to the selection vector $\mathbf{x}$, is then given by:

$$R(\mathbf{s}_f = \mathbf{x}, \mathbf{F}^*) = \mathcal{R}(\mathbf{x}, \mathbf{F}^*) - c \left(\mathrm{CRB}(\mathbf{x}, \mathbf{F}^*) - \mathrm{CRB}_{low} \right). \tag{13}$$

A visual representation of the MDP is provided in Fig. 2.

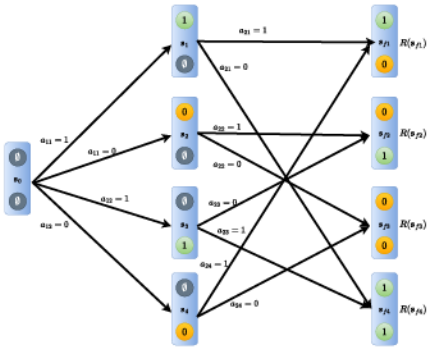

Fig. 2. An illustration of the MDP structure. It corresponds to an MDP whose transmit antenna array is of size 2.

The MDP flow is parameterized by a function approximator, $F_{\mathbf{w}}$, implemented as a deep neural network with parameters $\mathbf{w}$. Trajectories from the MDP are sampled, and the parameters are updated via gradient descent to minimize the loss in equation (11). The training algorithm is provided below (**Algorithm 1**) and is denoted as **GFLOW-TAS-PHY**.

IV. Experiments

We conducted a series of experiments to validate the effectiveness of the proposed framework. Specifically, we

Algorithm 1 GFLOW-TAS-PHY

Initialize $F_{\mathbf{w}}$, $\zeta \in [0,1]$ for exploration, learning rate η
for `all root-to-leaf trajectories` **do**
  $\mathbf{s} = \mathbf{s}_0 = [\emptyset]^{N_s}$
  for $N_s - 1$ transitions **do**
    Sample $z \sim \mathcal{U}(0,1)$ (Uniform distribution)
    If $z < \zeta$ choose $\mathbf{a} \in \mathcal{A}(\mathbf{s})$ randomly
    If $z >= \zeta$ choose $\mathbf{a} = \arg\max_{\mathbf{a}'} F_{\mathbf{w}}(\mathbf{s}, \mathbf{a}')$
    Apply action $\mathbf{a}$, compute sub. state $\mathbf{s}'$
    $\mathbf{w}' \rightarrow \mathbf{w} - \eta \nabla_{\mathbf{w}} L_{\mathbf{w}}(\mathbf{s}')$ eq. (11) $R(\mathbf{s}') = 0$
    $\mathbf{s} = \mathbf{s}'$
  end for
  Sample $z \sim \mathcal{U}(0,1)$
  If $z > \zeta$ choose $\mathbf{a} \in \mathcal{A}(\mathbf{s})$ randomly
  If $z \leq \zeta$ choose $\mathbf{a} = \arg\max_{\mathbf{a}'} F_{\mathbf{w}}(\mathbf{s}, \mathbf{a}')$
  Apply action $\mathbf{a}$, compute terminal state $\mathbf{x}$
  Compute $\mathbf{F}^* = \arg\max_{\mathbf{F}} R(\mathbf{x}, \mathbf{F})$ (fixed num of grad steps)
  $\mathbf{w}' \rightarrow \mathbf{w} - \eta \nabla_{\mathbf{w}} L_{\mathbf{w}}(\mathbf{x})$ $R(\mathbf{x}) = R(\mathbf{x}, \mathbf{F}^*)$ eq. (13)
end for

utilized a MIMO transmit array with $N_s = 50$ antennas, while both the eavesdropper and destination arrays consisted of $N_e = N_d = 5$ antennas. The communication channels from the source to the destination ($\mathbf{H}_d$) and from the source to the eavesdropper ($\mathbf{H}_e$) were assumed to be known and sampled from a complex normal distribution. The CRB threshold was set to $\text{CRB}_{low} = 2.5$. The parameters ρ_d and ρ_e are chosen to be 1 and 0.1, respectively. The optimal precoding matrix $\mathbf{F}^*$ was computed using 10 Wirtinger conjugate gradient ascent steps for each root-to-leaf trajectory.

The flow was parameterized using a deep neural network as proposed in [15], consisting of three layers: the first layer is a learnable Fourier kernel, initialized from a zero-mean Gaussian with variance $1e$-3, while the second and third layers are linear. A Rectified Linear Unit (ReLU) activation function was applied between the second and third layers. Each layer consisted of 150 neurons. The flow parameters were updated using the Adam optimizer [16], with a learning rate of $3e$-4. Training involved 15,000 root-to-leaf trajectories, with each trajectory corresponding to $N_s = 50$ gradient steps.

Upon completion of training, the optimal antenna subset can be sampled from the trained flow network by starting at the root and, at each state, selecting the action that maximizes the estimated outgoing flow, continuing this process until reaching the final state.

Fig. 3 illustrates the performance of the best antenna subset selected by **GFLOW-TAS-PHY**. Each point on the "Average Performance" line is calculated as follows: For a given value k on the x-axis (representing the number of active antennas), we sample 10 distinct subsets of k active antennas (wherever possible; for example, when $k = 50$, there is only one possible subset). For each subset, we compute the optimal precoding matrix $\mathbf{F}^*$ and calculate the corresponding secrecy rate. The final point is the average secrecy rate across the 10 sampled subsets.

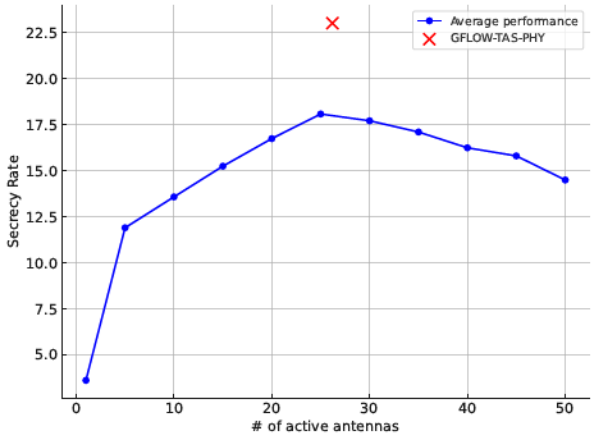

Fig. 3. Plot of the secrecy rate achieved by **GFLOW-TAS-PHY**. Each point on the "Average Performance" line (x-axis value is k) represents the average secrecy rate over 10 subsets with k active antennas and $N_s - k$ inactive antennas, computed using the corresponding $\mathbf{F}^*$ matrix. The final values are averaged over 5 seeds, each representing a new channel realization where **GFLOW-TAS-PHY** is trained from scratch.

Several observations can be made from Fig. 3. First, the behavior noted in [8] for PHY in MIMO systems also holds for ISAC systems. Specifically, the secrecy rate initially increases as more antennas are activated, but beyond a certain point, adding more antennas results in a decline in secrecy rate. Secondly, since the experiments are conducted on fixed channel realizations, the focus is on optimizing the instantaneous secrecy rate rather than the ergodic secrecy rate. As a result, the specific positions of the selected antennas become crucial. The proposed GFlowNet paradigm not only identifies the optimal number of active antennas but also selects the optimal antenna positions, leading to an average improvement of approximately 30% in the secrecy rate.

Fig. 4 shows the CRB performance of the antenna subsets selected by **GFLOW-TAS-PHY**. The points on the "Average Performance" line are computed in the same way as in Fig. 3, where each point represents the average CRB derived from multiple randomly sampled subsets of active antennas. For each subset size k, several random subsets are generated, and their CRB values are averaged.

Unlike the behavior observed with the secrecy rate, where adding more active antennas beyond a certain point can degrade performance, the CRB consistently improves as more active antennas are added. However, the subsets selected by **GFLOW-TAS-PHY** consistently achieve better CRB performance compared to the average performance of randomly selected subsets with the same number of active antennas, demonstrating the effectiveness of the method in optimizing both the number and positioning of antennas.

To illustrate the reduction in complexity, **GFLOW-TAS-PHY** searches a solution space consisting of over 10^{15} possible combinations, represented by $\sum_{k=1}^{50} \binom{50}{k}$. The algorithm runs for 15,000 root-to-leaf trajectories, allowing it to explore

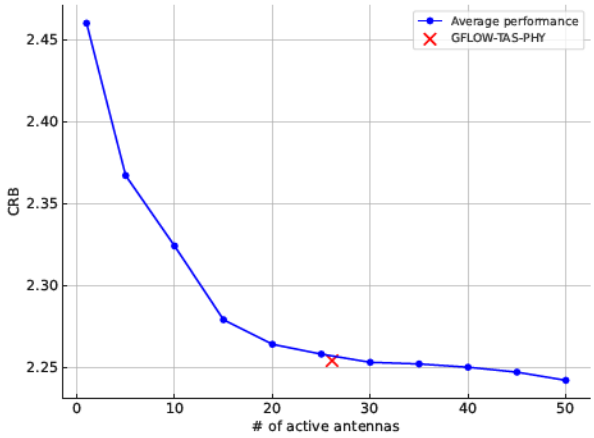

Fig. 4. Plot of the CRB achieved by **GFLOW-TAS-PHY**. Each point on the "Average Performance" line represents the average CRB over 10 subsets with k active antennas and $N_s - k$ inactive antennas, computed using the corresponding $\mathbf{F}^*$ matrix. The final values are averaged over 5 seeds, each representing a new channel realization where **GFLOW-TAS-PHY** is trained from scratch.

only a tiny fraction of the solution space. By parametrizing the flow, we leverage the generalization capabilities of the function approximator, enabling the GFlowNet to extrapolate and infer reward estimations for subsets not encountered during training.

V. Conclusions

This paper examines the impact of antenna selection on the PHY performance of ISAC systems. The scenario involves a MIMO source communicating with a multi-antenna destination while tracking a target, with an eavesdropper attempting to intercept the communication. The system's instantaneous secrecy rate increases as more antennas are activated, but beyond a certain point, additional antennas degrade secrecy performance. Conversely, the sensing performance, measured by the CRB, improves consistently with more active antennas. The goal is to optimize the number and configuration of active antennas to maximize the secrecy rate while ensuring the CRB remains below a set threshold. We propose an unsupervised approach based on the GFlowNet framework, modeling antenna selection as a MDP where actions determine whether antennas are active or inactive, leading to terminal states representing different antenna configurations. The reward for each terminal state combines the secrecy rate and the difference between the CRB and its threshold. GFlowNet learns an action-sampling policy that makes the probability of reaching a terminal state proportional to its reward. The method consistently selects antenna subsets that achieve a 30% improvement in secrecy rate over random subsets, while also slightly improving average CRB performance for subsets of the same size. This is achieved after being trained on a minute fraction of the possible antenna configurations (15×10^3 of the 10^{15} possible ones).

References

[1] H. Pennanen, T. Hänninen, O. Tervo, A. Tölli, and M. Latva-aho, "6g: The intelligent network of everything–a comprehensive vision, survey, and tutorial," *arXiv preprint arXiv:2407.09398*, 2024.

[2] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4264–4279, 2018.

[3] I. Ara and B. Kelley, "Physical layer security for 6g: Toward achieving intelligent native security at layer-1," *IEEE Access*, vol. 12, pp. 82 800–82 824, 2024.

[4] N. Su, F. Liu, and C. Masouros, "Sensing-assisted eavesdropper estimation: An isac breakthrough in physical layer security," *IEEE Transactions on Wireless Communications*, vol. 23, no. 4, pp. 3162–3174, 2024.

[5] H. Jia, X. Li, and L. Ma, "Physical layer security optimization with cramér–rao bound metric in isac systems under sensing-specific imperfect csi model," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 5, pp. 6980–6992, 2024.

[6] L. Xu, S. Sun, Y. D. Zhang, and A. Petropulu, "Joint antenna selection and beamforming in integrated automotive radar sensing-communications with quantized double phase shifters," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[7] M. Palaiologos, M. H. Castañeda García, T. Laas, R. A. Stirling-Gallacher, and G. Caire, "Joint antenna selection and covariance matrix optimization for isac systems," in *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2024, pp. 2071–2076.

[8] S. Asaad, A. Bereyhi, A. M. Rabiei, R. R. Müller, and R. F. Schaefer, "Optimal transmit antenna selection for massive mimo wiretap channels," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 817–828, 2018.

[9] E. Bengio, M. Jain, M. Korablyov, D. Precup, and Y. Bengio, "Flow network based generative models for non-iterative diverse candidate generation," in *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., vol. 34. Curran Associates, Inc., 2021, pp. 27 381–27 394.

[10] S. Evmorfos, Z. Xu, and A. Petropulu, "Gflownets for sensor selection," in *2023 IEEE 33rd International Workshop on Machine Learning for Signal Processing (MLSP)*, 2023, pp. 1–6.

[11] ——, "Sensor selection via gflownets: A deep generative modeling framework to navigate combinatorial complexity," *arXiv preprint arXiv:2407.19736*, 2024.

[12] Y. Bengio, S. Lahlou, T. Deleu, E. J. Hu, M. Tiwari, and E. Bengio, "Gflownet foundations," *Journal of Machine Learning Research*, vol. 24, no. 210, pp. 1–55, 2023.

[13] E. Bengio, M. Jain, M. Korablyov, D. Precup, and Y. Bengio, "Flow network based generative models for non-iterative diverse candidate generation," *Advances in Neural Information Processing Systems*, vol. 34, pp. 27 381–27 394, 2021.

[14] Z. Wei, W. Chen, C.-W. Qiu, and X. Chen, "Conjugate gradient method for phase retrieval based on the wirtinger derivative." *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, vol. 34, no. 5, pp. 708–712, 2017.

[15] S. Evmorfos, K. I. Diamantaras, and A. P. Petropulu, "Reinforcement learning for motion policies in mobile relaying networks," *IEEE Transactions on Signal Processing*, vol. 70, pp. 850–861, 2022.

[16] D. P. Kingma, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.