Branch-and-Bound Search and Machine Learning-Based Transmit Antenna Selection in MIMOME Channels

DONGMIN SEO \(^1\) AND HYUKMIN SON \(^2\), (Member, IEEE)

\(^1\)Multidimensional Communication and Signal Processing Laboratory, Department of Electronic Engineering, Wonkwang University, Iksan, Jeonbuk 54538, Republic of Korea
\(^2\)Department of Electronic Engineering, Gachon University, Sujeong-gu, Seongnam-si, Gyeonggi-do 13120, Republic of Korea

Corresponding author: Hyukmin Son (hson102@gachon.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korea Government (MSIT) under Grant 2021R1F1A1051075, and in part by the Gachon University Research Fund of 2022 under Grant GCU-202205720001.

ABSTRACT
The studies on the conventional transmit antenna selection (TAS) in multiple-input, multiple-output, multiple-antenna eavesdropper (MIMOME) channel have been focused on determining the optimal number of transmit antennas or selecting the optimal transmit antenna indices. To maximize the secrecy capacity, we need to develop a novel TAS scheme to simultaneously determine the optimal number of transmit antennas and select the corresponding optimal transmit antenna indices. In this paper, we propose the iterative branch-and-bound search-based TAS (IB-TAS) scheme to determine the optimal transmit antenna set in MIMOME channel. To reduce the computational complexity caused by TAS, we also propose the machine learning-based TAS (ML-TAS) schemes utilizing neural network (NN), support vector machine (SVM), and naive-Bayes (NB). Through the simulation and numerical results, it is demonstrated that the IB-TAS scheme achieves the optimal secrecy capacity. In addition, through comparative analysis of the proposed ML-TAS schemes, it was shown that the NN-based TAS scheme minimizes the computational complexity while minimizing the loss of secrecy capacity compared to other ML-TAS schemes.

INDEX TERMS
Machine learning, MIMOME channel, physical layer security, transmit antenna selection.

I. INTRODUCTION
Wireless communication has become indispensable in personal and military fields \[1\]. Security in wireless communication is essential to prevent eavesdropping of transmitted data from unauthorized users. In recent years, there has been increasing attention to simple and efficient security techniques such as physical layer security (PLS) instead of traditional encryption techniques, which have inherent difficulties and vulnerabilities related to complicated key encryption and management procedures \[2\].

A pioneering study on PLS is investigated on a wiretap channel \[3\] with a single antenna for each of a legitimate transmitter (Alice), a legitimate receiver (Bob), and an Eavesdropper (Eve). It has been extended to multi-antenna environments such as multiple-input, single-output, multiple-antenna eavesdropper (MISOME) and multiple-input, multiple-output, multiple-antenna eavesdropper (MIMOME) channels \[4\], \[5\]. In MISOME and MIMOME channels, various PLS techniques have been proposed, such as transmit-beamforming \[6\], \[7\], artificial noise \[8\], \[9\], and antenna selection \[10\], \[11\], \[12\], \[13\], \[14\] based on channel state information (CSI) of Bob and Eve. In particular, the transmit antenna selection (TAS) based PLS schemes effectively maintain communication security at the cost of low feedback overhead.

The initial works on TAS in MIMOME channel focused on single antenna selection that maximizes the instantaneous signal-to-noise ratio (SNR) at Bob \[10\], \[11\], \[12\] when Bob’s CSI is only available at Alice. These studies additionally considered receive combining schemes such as maximum ratio combining, selection combining, and generalized...
selection combining. The authors in [13] and [14] investigate the multiple transmit antenna selection schemes to maximize the achievable secrecy capacity depending on whether Eve’s perfect CSI is available at Alice or not. In [13], the transmit antenna is selected, which has the largest channel gain one-by-one until the selected transmit antenna set maximizes the secrecy capacity. The authors in [14] propose the branch-and-bound search-based TAS (B-TAS) scheme to select the optimal transmit antenna indices when the number of transmit antennas to choose is fixed. The B-TAS scheme shows nearly optimal performance and reduces the computational complexity compared to the exhaustive search-based TAS scheme. Therefore, we need to develop a novel scheme determining the optimal transmit antenna set, i.e. the optimal number of transmit antennas and the corresponding antenna indices, based on Eve’s CSI available at Alice.

On the other hand, as the number of transmit antennas increases, the TAS schemes cause high computational complexity. Thus, it is necessary to reduce the complexity while maintaining the secrecy capacity for practicality. Machine learning is one of the candidate techniques to minimize computational complexity effectively. The machine learning model is implemented based on offline training data set. As a result, the trained model operates online, reducing the computational complexity compared with the iterative algorithms [15].

From this point of view, machine learning-based antenna selection schemes have been proposed in multiple-input and multiple-output (MIMO) channel [16], [17], [18], [19]. The authors in [16] propose the receive antenna selection scheme based on the k-nearest neighbor (k-NN) and the support vector machine (SVM). These schemes minimize the bit error ratio (BER) and reduce feedback overhead and computational complexity. In [17], the authors propose the k-NN, the SVM, and the neural network (NN) based TAS schemes to minimize the BER. The NN-based scheme shows the best performance because it utilizes a robust structure with parametric forms. Since the output layer neurons of the NN-based scheme proposed in [17] are mapped to possible combinations of transmit antennas, the number of output layer neurons increases exponentially as the number of transmit antennas increases. It causes high computational complexity.

The authors in [18] and [19] propose the NN-based TAS schemes to maximize minimum SNR and capacity. The output layer neurons are mapped to the transmit antenna indices in these schemes. Thus, if the output value of the neuron in the output layer is greater than or equal to 0.5, the corresponding transmit antenna is selected. Otherwise, the transmit antennas are not selected. In this structure, the number of output layer neurons linearly increases as the number of transmit antennas increases. Thus, the NN-based schemes proposed in [18] and [19] efficiently reduce the computational complexity compared to the structure that the output layer neurons are mapped to possible combinations of transmit antennas.

The authors in [20] and [21] utilize machine learning and reinforcement learning algorithms to select a single transmit antenna in MIMOME channel. In [20], the authors use SVM and the naive-Bayes (NB) to reduce the feedback overhead and computational complexity while maximizing secrecy performances depending on whether Eve’s CSI is available at Alice or not. In [21], the authors use the deep Q network, i.e. one of the reinforcement learning algorithms, to maximize the secrecy outage probability and minimize the BER when the perfect CSI of Eve is not available at Alice. However, these methods only focus on a single transmit antenna. To the best of our knowledge, the machine learning-based TAS (ML-TAS) schemes have not been investigated to determine the optimal transmit antenna set. Therefore, it is necessary to develop the novel ML-TAS schemes in MIMOME channel.

In this paper, we propose the iterative branch-and-bound search-based TAS (IB-TAS) schemes to find the optimal number of transmit antennas and the corresponding optimal transmit antenna indices based on Eve’s CSI available at Alice, i.e. perfect and partial CSI cases. We also design supervised learning-based machine learning models to reduce the computational complexity of TAS. The main contributions of this paper are following:

1) When the perfect CSI of Eve is available at Alice, we propose the IB-TAS scheme to maximize the achievable secrecy capacity. The IB-TAS simultaneously determines the optimal number of transmit antennas and the corresponding optimal transmit antenna indices.

2) When Eve’s partial CSI, i.e. channel statistics, is available at Alice, we derive the lower bound on the expected capacity between Alice and Eve. The upper bound on the estimated secrecy capacity is then derived based on this lower bound. As a result, we propose the IB-TAS scheme to maximize the estimated secrecy capacity.

3) We propose the ML-TAS schemes utilizing NN, SVM, and NB. These ML-TAS schemes are trained in a supervised learning manner based on the proposed IB-TAS schemes. The simulation and numerical results demonstrate that the NN-based TAS scheme minimizes the computational complexity at the cost of the minimal secrecy capacity loss, compared to the other ML-TAS schemes.

For notations, uppercase and lowercase boldface represent matrices and vectors, respectively. $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$, and $\det(\mathbf{A})$ denote the transpose, the conjugate transpose, the inverse (or pseudo-inverse), and the determinant of a matrix $\mathbf{A}$, respectively. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. $\mathbf{a}[n]$ denotes the $n^{th}$ element of a vector $\mathbf{a}$ and $|\mathcal{I}|$ denotes the cardinality of a set $\mathcal{I}$. Further, $\setminus$ denotes the set difference operator and $\mathbb{E}[\cdot]$ denotes the expectation operator.

II. SYSTEM MODEL

We consider MIMOME channel in which Alice, Bob, and a passive Eve are equipped with $N_A$, $N_B$, and $N_E$ antennas, respectively. A large number of antennas at Alice are
considered, i.e. \( N_A > N_B \) and \( N_A > N_E \). The channel matrix from Alice to Bob is denoted as \( \mathbf{H} \in \mathbb{C}^{N_A \times N_A} \), which consists of channel vectors, i.e. \( \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_A} \). The channel matrix from Alice to Eve is denoted as \( \mathbf{G} \in \mathbb{C}^{N_E \times N_A} \), which consists of channel vectors, i.e. \( \mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{N_A} \). It is assumed that the channel matrices of \( \mathbf{H} \) and \( \mathbf{G} \) are fast Rayleigh fading channels, which are statistically independent of each other in a rich scattering environment [22]. The entries of \( \mathbf{H} \) and \( \mathbf{G} \) are independent and identically distributed (i.i.d.) complex-valued Gaussian random variables with zero-mean and unit-variance.

The channel matrices \( \mathbf{H} \) and \( \mathbf{G} \) are defined as

\[
\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_A}) \quad \text{and} \quad \mathbf{G} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_E})
\]

where the channel matrices \( \mathbf{H} \) and \( \mathbf{G} \) are zero-mean and unit-variance complex AWGN, i.e. \( \mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_A}) \) and \( \mathbf{G} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_E}) \).

In (1) and (2), the transmit power of each rank is given as \( \rho_B \) and \( \rho_E \) because the rank of \( \mathbf{H} \) is \( N_B \), i.e. \( \text{rank}(\mathbf{H}) = \min(N_A, N_B) = N_B \). Based on (1) and (2), the achievable secrecy capacity, \( C_S \), is defined as

\[
C_S = [C_B - C_E]^+,
\]

where \( [x]^+ = \max(x, 0) \), \( C_B = \log_2 \det(\mathbf{I}_{N_B} + \frac{\rho_B}{N_B} \mathbf{H} \mathbf{H}^H) \), and \( C_E = \log_2 \det(\mathbf{I}_{N_E} + \frac{\rho_E}{N_E} \mathbf{G} \mathbf{G}^H) \).

Depending on the availability of Eve’s CSI at Alice, we consider two scenarios. Scenario A is that the perfect CSI of \( \mathbf{G} \) is available at Alice, and Scenario B is that the partial CSI, i.e. channel statistics of \( \mathbf{G} \), is available at Alice.

### III. ITERATIVE BRANCH-AND-BOUND SEARCH-BASED TAS

The conventional B-TAS focuses on selecting optimal \( k \) transmit antennas when the number of transmit antennas to select, i.e. \( k \), is given and fixed in the system [14]. Thus, the conventional B-TAS determines the optimal transmit antenna set, \( S_k \), i.e. the optimal \( k \) transmit antenna indices, to maximize the secrecy capacity. Thus, there is an opportunity to improve the secrecy capacity by determining the optimal transmit antenna set, i.e. the optimal number of transmit antennas and the corresponding antenna indices. Meanwhile, most existing studies focus on whether the perfect CSI of Eve is available at Alice or not. However, from a practical point of view, the partial CSI of Eve, such as the channel statistics, should be considered. By borrowing the branch-and-bound search concept in the conventional B-TAS, we propose the IB-TAS scheme to find the optimal number of transmit antennas and the corresponding optimal transmit antenna indices based on Eve’s CSI available at Alice, i.e. perfect and partial CSI cases.

### A. SCENARIO A: PERFECT CSI OF G IS AVAILABLE AT ALICE

The IB-TAS finds total \( N_A \) transmit antenna sets, i.e. \( S_1, S_2, \ldots, S_{N_A} \), and calculates the corresponding achievable secrecy capacities. It then selects the optimal transmit antenna set, \( S_{opt} \), with the largest secrecy capacity.

Initially, \( S_k \) is set to an empty set and the antenna index set, \( J \), includes all transmit antenna indices. Let \( \mathbf{H}_n \in \mathbb{C}^{N_A \times n} \) and \( \mathbf{G}_n \in \mathbb{C}^{N_E \times n} \) denote the submatrices of \( \mathbf{H} \) and \( \mathbf{G} \) for the selected \( n \) (i.e. \( n < k \)) transmit antennas. Based on \( \mathbf{H}_n \) and \( \mathbf{G}_n \), the secrecy capacity for the selected \( n \) transmit antennas, \( C_{S,n} \), is defined as

\[
C_{S,n} = [C_{B,n} - C_{E,n}]^+ \]

\[
= \log_2 \det(\mathbf{I}_{N_B} + \frac{\rho_B}{N_B} \mathbf{H}_n \mathbf{H}_n^H) - \log_2 \det(\mathbf{I}_{N_E} + \frac{\rho_E}{N_E} \mathbf{G}_n \mathbf{G}_n^H),
\]

where \( m \) is the rank of \( \mathbf{H}_n \), i.e. \( m = \text{rank}(\mathbf{H}_n) = \min(k, N_B) \), and \( \mathbf{H}_k \) is the selected channel matrix corresponding to \( S_k \). Since the IB-TAS selects \( k \) transmit antennas for determining \( S_k \), the allocated power for each rank is given as \( \frac{\rho_B}{m} \) and \( \frac{\rho_E}{m} \) in (4).

Suppose that \( J_{n+1} \) is the \((n + 1)\)th selected transmit antenna index, and \( \mathbf{h}_{n+1} \) and \( \mathbf{g}_{n+1} \) are the \((n + 1)\)th column vectors of \( \mathbf{H} \) and \( \mathbf{G} \), respectively. Then, \( S_k \) and \( J \) are updated to \( S_k = S_k \cup \{J_{n+1}\} \) and \( J = J \setminus \{J_{n+1}\} \), respectively. After \( n + 1 \) transmit antennas are selected, the updated channel matrices, i.e. \( \mathbf{H}_{n+1} \in \mathbb{C}^{N_A \times (n+1)} \) and \( \mathbf{G}_{n+1} \in \mathbb{C}^{N_E \times (n+1)} \), are then given as \( \mathbf{H}_{n+1} = [\mathbf{H}_n, \mathbf{h}_{n+1}] \) and \( \mathbf{G}_{n+1} = [\mathbf{G}_n, \mathbf{g}_{n+1}] \), respectively. Based on the updated channel matrix \( \mathbf{H}_{n+1} \), \( C_{B,n+1} \) is derived as

\[
C_{B,n+1} = \log_2 \det(\mathbf{I}_{N_B} + \frac{\rho_B}{N_B} \mathbf{H}_{n+1} \mathbf{H}_{n+1}^H)
\]

\[
= \log_2 \det(\mathbf{I}_{N_B} + \frac{\rho_B}{N_B} \mathbf{H}_n \mathbf{H}_n^H + \frac{\rho_B}{N_B} \mathbf{h}_{n+1} \mathbf{h}_{n+1}^H)
\]

\[
= \log_2 \left( 1 + \frac{\rho_B}{N_B} \mathbf{h}_{n+1}^H \mathbf{T}_{B,n} \mathbf{h}_{n+1} \right)
\]

where \( \mathbf{T}_{B,n} \) is the channel statistics of \( \mathbf{H} \) from Alice to Bob. From this sense, scenario B can be regarded as the practical scenario.
\[ C_{E,n+1} = \log_2 (1 + \frac{\rho_B v_{B,n+1}}{1 + \rho_B v_{B,n+1}}) \]  
where \( \rho_B = \frac{\rho_B}{m} \) and \( T_{B,n} = (I_{N_B} + \rho_B H_n H_n^H)^{-1} \). The equation of (a) in (5) holds for the Sylvester’s determinant identity, i.e., \( \det(X + AB) = \det(X) \det(I + BX^{-1}A) \) [23]. Based on the updated channel matrix \( G_{n+1} \), \( C_{E,n+1} \) is derived as

\[
C_{E,n+1} = \log_2 \det \left( I_{N_E} + \rho_E G_{n+1} G_{n+1}^H \right) = C_{E,n} + \log_2 \left( \frac{1 + \rho_E v_{E,n+1} T_{E,n} g_{n+1}}{1 + \rho_E v_{E,n+1}} \right),
\]

where \( \rho_E = \frac{\rho_E}{m} \) and \( T_{E,n} = (I_{N_E} + \rho_B H_n G_n H_n^H)^{-1} \). Based on (6) and (8), \( C_{S,n+1} \) is derived as

\[
C_{S,n+1} = [C_{B,n+1} - C_{E,n+1}]^2 = C_{S,n} + \log_2 \left( \frac{1 + \rho_B v_{B,n+1}}{1 + \rho_B v_{B,n+1}} \right),
\]

Thus, the most recently selected transmit antenna index only affects the term of \( \Delta_{n+1} \).

Suppose that \( J_{n+2} \) is the \((n+2)^{th}\) selected transmit antenna index, and \( h_{n+2} \) and \( g_{n+2} \) are the \( J_{n+2} \) column vectors of \( H \) and \( G \), respectively. Then, \( S_k \) and \( J \) are updated to \( S_k = S_k \cup \{J_{n+2}\} \) and \( J = J \setminus \{J_{n+2}\} \), respectively. The updated channel matrices are thus \( H_{n+2} = [H_{n+1}, h_{n+2}] \) and \( G_{n+2} = [G_{n+1}, g_{n+2}] \), respectively. Similar to (5) and (6), \( C_{B,n+2} \) corresponding to \( h_{n+2} \) is derived as

\[
C_{B,n+2} = C_{B,n+1} + \log_2 \left( 1 + \rho_B v_{B,n+2} T_{B,n+1} h_{n+2} \right)
\]

\[
= C_{B,n+1} + \log_2 \left( \frac{1 + \rho_B v_{B,n+2}}{1 + \rho_B v_{B,n+2}} \right),
\]

Similar to (7) and (8), \( C_{E,n+2} \) corresponding to \( g_{n+2} \) is derived as

\[
C_{E,n+2} = C_{E,n+1} + \log_2 \left( \frac{1 + \rho_E v_{E,n+2}}{1 + \rho_E v_{E,n+2}} \right).
\]

Using the Sherman-Morrison formula, i.e., \((A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1} \) [23], the matrices \( T_{B,n+1} \) in (10) and \( T_{E,n+1} \) in (12) are given as

\[
T_{B,n+1} = (I_{N_B} + \rho_B H_n H_n^H + \rho_B h_{n+1} h_{n+1}^H)^{-1}
\]

\[
= T_{B,n} - \frac{T_{B,n} h_{n+1} h_{n+1}^H T_{B,n}}{\rho_B v_{B,n+1}}
\]

\[
= T_{B,n} - T_{B,n+1} h_{n+1}^H
\]

\[
T_{E,n+1} = (I_{N_E} + \rho_B g_{n+1} G_n G_n^H + \rho_E g_{n+1} g_{n+1}^H)^{-1}
\]

\[
= T_{E,n} - \frac{T_{E,n} g_{n+1} g_{n+1}^H T_{E,n}}{\rho_E v_{E,n+1}}
\]

\[
= T_{E,n} - T_{E,n+1} g_{n+1}^H
\]

\[
\text{FIGURE 1. An example of search tree when } N_A = 6 \text{ and } k = 2.
\]

where \( t_{B,n+1} = \frac{T_{B,n} h_{n+1}}{\sqrt{\rho_B v_{B,n+1}}} \) and \( t_{E,n+1} = \frac{T_{E,n} g_{n+1}}{\sqrt{\rho_E v_{E,n+1}}} \). To reduce the computational complexity for calculating \( v_{B,n+2} \) in (11) and \( v_{E,n+2} \) in (13), both terms can be rewritten as

\[
v_{B,n+2} = h_{n+2}^H \left( T_{B,n} - T_{B,n+1} h_{n+1} h_{n+1}^H \right) h_{n+2} + \hat{v}_{B,n+1} - \left| h_{n+2}^H \right|^2
\]

\[
v_{E,n+2} = g_{n+2}^H \left( T_{E,n} - T_{E,n+1} g_{n+1}^H \right) g_{n+2} + \hat{v}_{E,n+1} - \left| g_{n+2}^H \right|^2.
\]

In (16) and (17), \( \hat{v}_{B,n+1} \) is given as \( h_{n+2}^H T_{B,n} h_{n+2} \) and \( \hat{v}_{E,n+1} \) is given as \( g_{n+2}^H T_{E,n} g_{n+2} \). Based on (11) and (13), \( C_{S,n+2} \) is derived as

\[
C_{S,n+2} = [C_{B,n+2} - C_{E,n+2}]^2
\]

\[
= C_{S,n+1} + \log_2 \left( \frac{1 + \rho_B v_{B,n+2}}{1 + \rho_B v_{B,n+2}} \right).
\]

In the conventional B-TAS, the branch-and-bound search algorithm in [24] is efficient for a monotonic objective function. By borrowing the branch-and-bound search concept in the conventional B-TAS, the IB-TAS uses the new objective function, \( \tilde{C}_{S,n} \), instead of \( C_{S,n} \) in (4). \( \tilde{C}_{S,n} \) is defined as

\[
\tilde{C}_{S,n} = C_{S,n} - \sum_{a=0}^{n-1} Z_a,
\]

where \( Z_a = \log_2 (1 + \rho_B c_a) - \log_2 (1 + \rho_E \eta_a), \) \( \xi_a = \max_{j \in I_a} h_j^H h_j, \) \( \eta_a = \min_{j \in I_a} g_j^H (I_{N_E} + \rho_E G a G^H)^{-1} g_j, \) \( Z_0 = 0, \) and the index set, \( I_a, \) consists of all the candidate transmit antenna indices in the step of selecting the \( a^{th} \) optimal transmit antenna. Based on (9), (18), and (19), the recursive form of the monotonic objective function is derived as

\[
\tilde{C}_{S,n+1} = \tilde{C}_{S,n} + \Delta_{n+1} - Z_n.
\]

Based on (20), the optimal \( S_k \) is determined by the IB-TAS.

Fig. 1 shows an example of search tree for the IB-TAS when \( k = 2 \) and \( N_A = 6 \). This example shows the search procedure to determine the optimal \( S_2 \). In Fig. 1, the number
on each node, i.e. node index, denotes the transmit antenna index. In level 1, each \( C_{S,1} \) corresponding to each element of \( I_1 \) is first calculated by (20), where \( I_1 \) is the index set for level 1, i.e. \( I_1 = \{1, 2, 3, 4, 5\} \). Let us assume that the node 2 in level 1 has the largest \( C_{S,1} \) among the other nodes in level 1. If \( C_{S,1} \) of the node 2 in level 1 is smaller than the global lower bound, \( B \), all child nodes corresponding to the node 2 in level 1 are pruned. In general, the initial value of \( B \) is very small, e.g. \(-\infty\). If \( C_{S,1} \) of the node 2 in level 1 is larger than \( B \), each \( C_{S,2} \) corresponding to each child node of the node 2 in level 1 is calculated by (20). Let us assume that \( C_{S,1} \) of the node 2 in level 1 is larger than \( B \) and the node 4 in level 2 among child nodes of the node 2 in level 1 has the largest \( C_{S,2} \). If its \( C_{S,2} \) is larger than \( B \), \( B \) is updated to the largest \( C_{S,2} \) and the optimal transmit antenna set, \( S_2 \), is updated to \( \{2, 4\} \). Then, it returns to the top level, i.e. level 1, and repeats the search procedure. If its \( C_{S,2} \) is smaller than \( B \), it returns to the top level, i.e. level 1, and repeats the search procedure. This is repeated until all the nodes in the search tree are pruned or visited for calculating the monotonic objective function, i.e. (20). Then, the optimal \( S_2 \) can be determined.

The achievable secrecy capacity corresponding to \( S_k \) can be calculated as

\[
C_{S_k} = \log_2 \det \left( I_{N_k} + \frac{\rho_b}{m} H_k H_k^H \right) - \log_2 \det \left( I_{N_k} + \frac{\rho_E}{m} G_k G_k^H \right),
\]

(21)

where \( H_k \) and \( G_k \) are the channel matrices corresponding to \( S_k \). The IB-TAS determines total transmit antenna sets, i.e. \( S_1, S_2, \ldots, S_{N_A} \), and calculates the corresponding achievable secrecy capacities based on (22). The optimal transmit antenna set with the largest secrecy capacity, \( S_{opt} \), is then determined among total transmit antenna sets, i.e.

\[
S_{opt} = \arg \max_{S_k} C_{S_k}.
\]

(22)

Therefore, the IB-TAS can determine the optimal number of transmit antennas and the corresponding optimal transmit antenna indices to maximize the achievable secrecy capacity.

**B. SCENARIO B: CHANNEL STATISTICS OF G ARE AVAILABLE AT ALICE**

When the channel statistics of \( G \) are only available at Alice, we derive the lower bound on the expected capacity between Alice and Eve, i.e. \( \mathbb{E}[C_E] \). The upper bound on the estimated secrecy capacity is then derived based on this lower bound. As a result, we propose the IB-TAS scheme to maximize the estimated secrecy capacity.

Similar to the operation in scenario A, the IB-TAS determines total \( N_A \) transmit antenna sets, i.e. \( S_1, S_2, \ldots, S_{N_A} \), and calculates the corresponding estimated secrecy capacities.

Initially, \( S_k \) is set to an empty set and the antenna index set, \( \mathcal{J} \), includes all transmit antenna indices. Let \( H_k \) and \( G_k \) denote the submatrices of \( H \) and \( G \) for selected \( n \) transmit antennas. Based on \( H_k \) and \( G_k \), the estimated secrecy capacity for the selected \( n \) transmit antennas, \( C_{S,n} \), is defined as

\[
C_{S,n} = \left[ C_{B,n} - \mathbb{E}[C_{E,n}] \right]^+ + \log_2 \det \left( I_{N_k} + \frac{\rho_b}{m} H_k H_k^H \right) - \mathbb{E} \left[ \log_2 \det \left( I_{N_k} + \frac{\rho_E}{m} G_k G_k^H \right) \right],
\]

(23)

where \( m \) is the rank of \( H_k \), i.e. \( m = \text{rank}(H_k) = \min(k, N_B) \), and \( H_k \) is the selected channel matrix corresponding to \( S_k \). Since the IB-TAS selects \( k \) transmit antennas for determining \( S_k \), the allocated power for each rank is given as \( \frac{\rho_b}{m} \) and \( \frac{\rho_E}{m} \) in (23).

From (24), the expected \( C_{E,n} \), \( \mathbb{E}[C_{E,n}] \), is derived as

\[
\mathbb{E}[C_{E,n}] = \mathbb{E} \left[ \log_2 \det \left( I_{N_k} + \frac{\rho_E}{m} W_n \right) \right] = u_n \mathbb{E} \left[ \log_2 \det \left( I_{N_k} + \frac{\rho_E}{m} W_n \right)^{n/2} \right] \geq u_n \mathbb{E} \left[ \log_2 \left( 1 + \frac{\rho_E}{m} \text{det} (W_n) \right)^{n/2} \right] \geq u_n \log_2 \left( 1 + \frac{\rho_E}{m} \text{det} (W_n) \right)
\]

(24)

(25)

(26)

(27)

where \( \rho_E = \frac{\rho_E}{m} \) and \( u_n = \min(n, N_E) \). In addition, \( W_n = \begin{cases} G_k^H G_n, & \text{if } n < N_E \\ \mathcal{N}(0, \sigma), & \text{if } n \geq N_E \end{cases} \) is a central complex Wishart matrix, i.e. \( W_n \sim \mathcal{W}(n, \sigma \mathcal{I}) \), where \( \sigma = \max(n, N_E) \). The inequalities of (a) and (c) in (24) and (b) in (26) are due to the Minkowski determinant theorem [25] and the Jensen’s inequality [27], respectively. In addition, the equality of (b) in (25) holds for the Jacobi’s formula [26]. Since the term of \( \log_2 (1 + \exp(\cdot)) \) in (25) is a convex function, it is given as \( \mathbb{E}(f(X)) \geq f(\mathbb{E}(X)) \). The term \( \alpha_n \) in (27) is given as

\[
\alpha_n = \frac{1}{u_n} \sum_{r=0}^{u_n-1} \psi \left( \frac{n-r}{\gamma} \right)
\]

(28)

(29)

where \( \psi(\cdot) \) is the digamma function and \( \gamma \) is the Euler-Mascheroni constant [28, Th. 2.11]. Based on (23) and (27), the upper bound of \( C_{S,n} \) is given as

\[
C_{S,n} \leq C_{B,n} - u_n \log_2 \left( 1 + \frac{\rho_E}{m} \exp(\alpha_n) \right).
\]

(30)

Suppose that \( J_{n+1} \) is the \((n+1)th\) selected transmit antenna index, and \( H_{n+1} \) and \( g_{n+1} \) are \((n+1)th\) column vectors of \( H \) and \( G \), respectively. Then, \( S_k \) and \( \mathcal{J} \) are updated to \( S_k = S_k \cup \{J_{n+1}\} \) and \( \mathcal{J} = \mathcal{J} \setminus \{J_{n+1}\} \), respectively. After \( n+1 \) transmit antennas are selected, the selected channel matrix, i.e. \( G_{n+1} \), is given as \( [G_n, g_{n+1}] \). Based on the updated channel matrix
$\mathbf{H}_{n+1}$, $\mathbf{C}_{B,n+1}$ is given as (6). Based on the updated channel matrix $\mathbf{G}_{n+1}$, the expected $C_{E,n+1}$, $\mathbb{E} \left[ C_{E,n+1} \right]$, is derived as

$$
\mathbb{E} \left[ C_{E,n+1} \right] \\
\geq u_{n+1} \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \frac{1}{u_{n+1}} \mathbb{E} \left[ \ln \det \left( \mathbf{W}_{n+1} \right) \right] \right) \right) \tag{31}
$$

$$
= u_{n+1} \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_{n+1} \right) \right) \\
+ u_n \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right) - u_n \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right)
$$

$$
= u_n \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right) \\
+ \log_2 \left( \frac{\left( 1 + \tilde{\rho}_E \exp \left( \alpha_{n+1} \right) \right)^{u_{n+1}}}{\left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right)^{u_n}} \right) \tag{32}
$$

$$
\text{where } u_{n+1} = \min \left( n + 1, N_E \right) \text{ and } \\
\mathbf{W}_{n+1} = \begin{cases} \\
\mathbf{G}_{n+1}^H \mathbf{G}_{n+1}, \text{ if } n + 1 < N_E, \\
\mathbf{G}_{n+1} \mathbf{G}^H_{n+1}, \text{ if } n + 1 \geq N_E,
\end{cases}
$$

i.e. $\mathbf{W}_{n+1} \sim \mathcal{W}_{n+1} \left( \mathbf{v}_{n+1}, \mathbf{I}_{u_{n+1}} \right)$ in which $\mathbf{v}_{n+1} = \max \left( n + 1, N_E \right)$. The lower bound of (31) is derived using the Minkowski determinant theorem, the Jacobi’s formula, and the Jensen’s inequality. Similar to (28) and (29), the term $\alpha_{n+1}$ in (32) is given as

$$
\alpha_{n+1} = \frac{1}{u_{n+1}} \sum_{r=0}^{u_{n+1}-1} \psi \left( v_{n+1} - r \right) \\
= \frac{1}{u_{n+1}} \sum_{r=0}^{u_{n+1}-1} \sum_{i=1}^{v_{n+1}-r-1} \frac{1}{i} - \gamma \tag{34}
$$

Based on (6) and (33), the estimated secrecy capacity for the selected $n + 1$ transmit antennas, $C_{S,n+1}$, is derived as

$$
C_{S,n+1} = \left[ C_{B,n+1} - \mathbb{E} \left[ C_{E,n+1} \right] \right]^+ \\
\leq C_{B,n} - u_n \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right) \\
+ \log_2 \left( \frac{1 + \tilde{\rho}_B \psi_{B,n+1}}{1 + \tilde{\rho}_E \exp \left( \alpha_{n+1} \right)} \right) \tag{35}
$$

$$
\Rightarrow \Delta_{n+1} \tag{36}
$$

Thus, the most recently selected transmit antenna index only affects the term of $\Delta_{n+1}$.

Suppose that $J_{n+2}$ is the $(n+2)^{th}$ selected transmit antenna index, and $\mathbf{h}_{n+2}$ and $\mathbf{g}_{n+2}$ are the $J_{n+2}$ column vectors of $\mathbf{H}$ and $\mathbf{G}$, respectively. Then, $\mathbf{S}_k$ and $\mathcal{J}$ are updated to $\mathbf{S}_k = \mathbf{S}_k \cup \{ J_{n+2} \}$ and $\mathcal{J} = \mathcal{J} \setminus \{ J_{n+2} \}$, respectively. The updated channel matrices are thus $\mathbf{H}_{n+2} = [\mathbf{H}_{n+1}, \mathbf{h}_{n+2}]$ and $\mathbf{G}_{n+2} = [\mathbf{G}_{n+1}, \mathbf{g}_{n+2}]$, respectively. $C_{B,n+2}$ corresponding to $\mathbf{H}_{n+2}$ is given as (11). To reduce the computational complexity for calculating $\psi_{B,n+2}$ in (11), the matrix $\mathbf{T}_{B,n+1}$ and the term of $\psi_{B,n+2}$ can be updated using (14) and (16), respectively.

Similar to (30), $\mathbb{E} \left[ C_{E,n+2} \right]$ corresponding to $\mathbf{G}_{n+2}$ is derived as

$$
\mathbb{E} \left[ C_{E,n+2} \right] \geq u_{n+2} \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_{n+2} \right) \right) + \log_2 \left( \frac{1 + \tilde{\rho}_E \exp \left( \alpha_{n+2} \right)}{1 + \tilde{\rho}_E \exp \left( \alpha_{n+1} \right) \exp \left( \alpha_{n+2} \right)} \right) \tag{36}
$$

$$
\text{where } u_{n+2} = \min \left( n + 2, N_E \right) \text{ and } \nu_{n+2} = \max \left( n + 2, N_E \right). \text{ In addition, the term } \alpha_{n+2} \text{ in (36) is given as }
$$

$$
\alpha_{n+2} = \frac{1}{u_{n+2}} \sum_{r=0}^{u_{n+2}-1} \sum_{i=1}^{\nu_{n+2}-r-1} \frac{1}{i} - \gamma \tag{37}
$$

Based on (11) and (36), the estimated secrecy capacity for the selected $n + 2$ transmit antennas, $C_{S,n+2}$, is derived as

$$
C_{S,n+2} = \left[ C_{B,n+2} - \mathbb{E} \left[ C_{E,n+2} \right] \right]^+ \\
\leq C_{B,n+1} - u_n \log_2 \left( 1 + \tilde{\rho}_E \exp \left( \alpha_n \right) \right) \\
+ \log_2 \left( \frac{1 + \tilde{\rho}_B \psi_{B,n+2}}{1 + \tilde{\rho}_E \exp \left( \alpha_{n+1} \right) \exp \left( \alpha_{n+2} \right)} \right) \tag{38}
$$

$$
\Rightarrow \Delta_{n+2} \tag{39}
$$

Similar to scenario A, the monotonic objective function, $\hat{C}_{S,n}$, instead of $C_{S,n}$ is used and it is defined as

$$
\hat{C}_{S,n} = C_{S,n} - \sum_{a=0}^{n-1} Z_a \tag{39}
$$

$$
\text{where } Z_a = \log_2 \left( \frac{1 + \tilde{\rho}_B \psi_{a}}{1 + \tilde{\rho}_E \exp \left( \alpha_a \right)} \right), \quad \zeta_a = \max_{i \in T_a} \mathbf{h}_i^H \mathbf{h}_j, \quad Z_0 = 0, \quad \alpha_0 = \psi \left( N_E \right) = \frac{n+1}{2} - \gamma, \text{ and the index set, } T_a, \text{ consists of all of the candidate transmit antenna indices in the step of selecting the } a^{th} \text{ optimal transmit antenna. Based on (35), (38), and (39), the recursive form of the monotonic objective function is derived as }
$$

$$
\hat{C}_{S,n+1} = \hat{C}_{S,n} + \Delta_{n+1} - Z_n \tag{40}
$$

Based on (40), the optimal $S_k$ is determined by the IB-TAS. $S_k$ can be determined using the IB-TAS algorithm in section III-A to which the monotonic objective function, i.e. (40), is applied. Then, the estimated secrecy capacity corresponding to $S_k$ can be calculated as

$$
C_{S_k} = \log_2 \det \left( \mathbf{I}_{N_E} + \tilde{\rho}_E \mathbf{G}_k \mathbf{H}_k^H \right) - \mathbb{E} \left[ \log_2 \det \left( \mathbf{I}_{N_E} + \tilde{\rho}_E \mathbf{G}_k \mathbf{H}_k^H \right) \right] \tag{41}
$$

$$
\text{where } \mathbf{H}_k \text{ and } \mathbf{G}_k \text{ are the channel matrices corresponding to } S_k. \text{ After total transmit antenna sets, i.e. } S_1, S_2, \ldots, S_{N_A}, \text{ and the corresponding estimated secrecy capacities are calculated, the optimal transmit antenna set with the largest secrecy capacity, } S_{opt}, \text{ is then determined using (22).}$$
C. ALGORITHM DESCRIPTION

Fig. 2 represents the block diagram for the IB-TAS algorithm. After initializing the parameters and variables for the given $k$, the IB-TAS constructs the search tree for selecting optimal $k$ transmit antennas. It then determines the optimal transmit antenna set, $S_k$, based on branch-and-bound search (BAB), which includes $k$ optimal transmit antenna indices. This process from the first to the third blocks in Fig. 2 is repeated until $N_A$ optimal transmit antenna sets corresponding to the given $k$, i.e., $S_1, S_2, \ldots, S_{N_A}$, are determined. Finally, in the fourth block of Fig. 2, $S_{\text{opt}}$ among $N_A$ optimal transmit antenna sets is determined, which maximizes the secrecy capacity.

Algorithm 1 summarizes the pseudo-codes for the IB-TAS of scenarios A and B. In Algorithm 1, the sub-algorithms for the cases of scenarios A and B are Algorithms 2, 4, and 6, and Algorithms 3, 5, and 6, respectively. Algorithms 2 and 3 construct the search tree for the given $k$. Algorithms 4 and 5 update the parameters for further exploration of the visited node in the current search level. Algorithm 6 prunes the child nodes below the visited node in the current search level. Note that Algorithms 4, 5, and 6 are used to determine $S_k$ through BAB. In Algorithms 1-6, $I_n$ is the index set for the nodes in level $a$, $I_{a,b}$ is the index set for the child nodes of node $b$ in level $a$, and $s$ is the selected transmit antenna index vector.

In Algorithm 1, the process from step 1 to step 35 represents determining $S_k$ and calculating $C_{S_k}$ for $1 \leq k \leq N_A$. This process corresponds from the first block to the third block in Fig. 2. Step 2 initializes the parameters and variables for the given $k$, as shown in the first block of Fig. 2. Since the search procedure starts from the root node in level 0, the currently visited antenna index, $J$, is set to 0, i.e., $J = 0$.

$S_k$ is determined through the search procedure from steps 6 to 32 when $1 \leq k \leq N_A - 1$. When $k$ is $N_A$, all transmit antennas are selected in steps 3 to 4. The search tree for the given $k$ is constructed through Algorithm 2 or Algorithm 3 in step 6. Algorithms 2 and 3 correspond to the second block of Fig. 2. In Algorithms 2 and 3, $Z_a$ is calculated as given in (19) and (39), respectively. Based on the calculated $Z_a$, the monotonic objective function $c_j$ corresponding to $I_{a,j}$ is calculated in the following steps of Algorithm 1.

The process from step 7 to step 31 determines $S_k$, which corresponds to the third block in Fig. 2. When searching the nodes of level $k - 1$, i.e., $n = k - 1$, in steps 8 to 16, each $c_j$ corresponding to $I_{n,j}$ is calculated based on 20 and 40 for scenarios A and B, respectively. If $c_{\text{max}}$ in step 10 is larger than $B, s, B$, and $S_k$ are updated by step 12. If $c_{\text{max}}$ is smaller than $B, s, B$, and $S_k$ are not updated, and it returns to step 7. After those procedures, the index $J$ is eliminated from $I_{n-1,a(n-1)}$. The search procedures from steps 8 to 16 are repeatedly performed until all nodes are visited.

When searching the nodes in the other levels, i.e., $n < k - 1$, in steps 18 to 30, each $c_j$ corresponding to $I_{n,j}$ is given as in step 9. In these steps, $c_j$ is the largest value among $c_j$ in the current search level $n$, and $\hat{n}$ is the level to which the child nodes of node $J$ in level $n$ belong. Note that $\hat{n}$ is the lower level than $n$. If $c_j$ is larger than $B$, $s$ and $\hat{C}_S$ are updated through Algorithms 4 and 5 in step 24. Then, the child nodes of node $J$ in level $n$ are visited to calculate $\Delta_j$ corresponding to $I_{n,j}$ and $n$ is updated to $n = \hat{n}$. If $c_j$ is smaller than $B$, the child nodes of node $J$ in level $n$ are pruned by step 27 through Algorithm 6, and it returns to step 7. If $I_{0,0}$ becomes the empty set in step 7, the optimal $S_k$ is determined to $s$, which was updated in step 12. In steps 33 and 34, the achievable secrecy capacity corresponding to the optimal $S_k$ is calculated using 21 in scenario A and 41 in scenario B. After determining the optimal transmit antenna sets for all values of $k$, the optimal transmit antenna set with the largest secrecy capacity, $S_{\text{opt}}$, is determined by step 36, which corresponds to the fourth block in Fig. 2.

D. COMPUTATIONAL COMPLEXITY

We use the big O notation to measure the computational complexity of the IB-TAS schemes proposed in Algorithm 1. Since all nodes, except for the pruned node, in the search tree are visited for calculating the monotonic objective function, the iterative calculation of the objective function causes high computational complexity. When calculating the objective
Algorithm 1 IB-TAS

Input: $H = [h_1, h_2, \ldots, h_N] \in \mathbb{C}^{N_h \times N_A}$, $G = [g_1, g_2, \ldots, g_N] \in \mathbb{C}^{N_e \times N_N}$, $N_A$, $N_B$, $N_E$, $\rho_B$, $\rho_E$

Output: $S_{opt}$

1: for $k = 1 : N_A$ do
2: $T_B = I_{N_B}$, $T_E = I_{N_E}$, $B = -\infty$, $\hat{C}_S = 0$, $n = 0$, $J = 0$, $Z_0 = 0$, $s = \emptyset$, $J = \{1, 2, \ldots, N_A\}$, $K = \{1, 2, \ldots, k\}$, $m = \min(k, N_B)$, $\rho_B = \frac{m}{m}$, $\rho_E = \frac{m}{m}$
3: if $k = N_A$ then
4: $S_k = J$
5: else
6: Perform Algorithm 2 for scenario A
7: Perform Algorithm 3 for scenario B
8: if $\mathcal{I}_{0,0} \neq \emptyset$ then
9: if $n = k - 1$ then
10: $c_j = \hat{C}_S + \Delta_j - Z_a$, $\forall j \in \mathcal{I}_{n,j}$
11: $c_{max} = \max_{j \in \mathcal{I}_{n,j}} c_j$
12: if $c_{max} > B$ then
13: $s[k] = \arg \max_{j \in \mathcal{I}_{n,j}} c_j$, $B = c_{max}$, $S_k = s$
14: end if
15: $\mathcal{I}_{n-1,n-1} = \mathcal{I}_{n-1,n-1} \setminus \{J\}$
16: $n = J + 1$, $s = \emptyset$, $\mathcal{I}_B = I_{N_B}$, $\mathcal{I}_E = I_{N_E}$, $\nu_{B,j} = h_j^H h_j$, $\nu_{E,j} = g_j^H g_j$, $\forall j \in J$
17: go to step 7
18: else
19: $c_j = \hat{C}_S + \Delta_j - Z_a$, $\forall j \in \mathcal{I}_{n,j}$
20: sort $c_j$, $\forall j \in \mathcal{I}_{n,j}$ in a descending order to get an ordered index vector $d \in \mathbb{R}^{1 \times |\mathcal{I}_{n,j}|}$
21: $J = d[1]$, $\hat{n} = n + 1$
22: $\mathcal{T}_B = \mathcal{T}_B$, $\mathcal{T}_E = \mathcal{T}_E$
23: $\hat{\nu}_{B,j} = \nu_{B,j}$, $\hat{\nu}_{E,j} = \nu_{E,j}$, $\forall j \in J$
24: if $c_j > B$ then
25: Perform Algorithm 4 for scenario A
26: Perform Algorithm 5 for scenario B
27: go to step 8
28: else
29: Perform Algorithm 6
30: go to step 7
31: end if
32: end if
33: $H_k = H[:; S_k[1 : k]]$, $G_k = G[:; S_k[1 : k]]$
34: calculate $C_{S_k}$ using (21)
35: end for
36: determine $S_{opt}$ using (22)

Algorithm 2 Construct the Search Tree for Scenario a

1: $\mathcal{I}_a = \{a, a + 1, \ldots, N_A + a - k\}$, $\forall a \in K$
2: $\mathcal{I}_{0,0} = \mathcal{I}_1$, $\mathcal{I}_{a,b} = \{b + 1, b + 2, \ldots, N_A + a - k + 1\}$, $\forall a \in K \setminus \{k\}$, $\forall b \in \mathcal{I}_a$
3: $\nu_{B,j} = h_j^H h_j$, $\nu_{E,j} = g_j^H g_j$, $\forall j \in J$
4: $\Delta_j = \log_2 \left( \frac{1 + \rho_B \nu_{B,j}}{1 + \rho_B \nu_{E,j}} \right)$, $\forall j \in J$
5: $\xi_a = \max_{j \in \mathcal{I}_a} \nu_{B,j}$, $\eta_a = \min_{j \in \mathcal{I}_a} \nu_{E,j}$, $\forall a \in K$
6: $Z_a = \log_2 \left( \frac{1 + \rho_B \nu_{B,a}}{1 + \rho_B \nu_{E,a}} \right)$, $\forall a \in K \setminus \{k\}$

Algorithm 3 Construct the Search Tree for Scenario B

1: $\mathcal{I}_a = \{a, a + 1, \ldots, N_A + a - k\}$, $\forall a \in K$
2: $\mathcal{I}_{0,0} = \mathcal{I}_1$, $\mathcal{I}_{a,b} = \{b + 1, b + 2, \ldots, N_A + a - k + 1\}$, $\forall a \in K \setminus \{k\}$, $\forall b \in \mathcal{I}_a$
3: $\nu_{B,j} = h_j^H h_j$, $\nu_{E,j} = g_j^H g_j$, $\forall j \in J$
4: $\alpha_0 = 0$, $\alpha_a = \frac{1}{\nu_a} \sum_{b=0}^{a-1} \sum_{t=1}^{b-1} \frac{1}{t} - y$, $\forall a \in K$
5: $\Delta_j = \log_2 \left( \frac{1 + \rho_B \nu_{B,j}}{1 + \rho_B \nu_{E,j}} \right)$, $\forall j \in J$
6: $\xi_a = \max_{j \in \mathcal{I}_a} \nu_{B,j}$, $\eta_a = \min_{j \in \mathcal{I}_a} \nu_{E,j}$, $\forall a \in K$
7: $Z_a = \log_2 \left( \frac{1 + \rho_B \nu_{B,a}}{1 + \rho_B \nu_{E,a}} \right)$, $\forall a \in K \setminus \{k\}$

Algorithm 4 Update the Parameters for Further Exploration of the Visited Node for Scenario a

1: $s[\hat{n}] = J$, $\hat{C}_S = c_j$
2: $t_B = \frac{T_{B,j}}{\sqrt{\rho_B + \nu_{B,j}}} t_E = \frac{T_{E,j}}{\sqrt{\rho_E + \nu_{E,j}}}$
3: $\mathcal{T}_B = \mathcal{T}_B - t_B h_j^H h_j$, $\mathcal{T}_E = \mathcal{T}_E - t_E g_j^H g_j$
4: $\nu_{B,j} = \nu_{B,j} - h_j^H h_j$, $\nu_{E,j} = \nu_{E,j} - g_j^H g_j$, $\forall j \in \mathcal{I}_{\hat{n},j}$
5: $\Delta_j = \log_2 \left( \frac{1 + \rho_B \nu_{B,j}}{1 + \rho_B \nu_{E,j}} \right)$, $\forall j \in \mathcal{I}_{\hat{n},j}$
6: $n = \hat{n}$

and O $(N_A N_B N_T)$, respectively. Note that $N_T$ is the total number of visited nodes to find the optimal transmit antenna set, which increases as the number of transmit antennas increases.

IV. MACHINE LEARNING-BASED TAS SCHEMES FOR PLS

Although the IB-TAS maximizes the secrecy capacity, the iterative search process causes high computational complexity. Machine learning is one of the candidate techniques to minimize computational complexity effectively, and the TAS can be interpreted as a multi-class classification problem. In this section, we propose the machine learning-based TAS (ML-TAS) schemes to solve the multi-class classification problem using neural network (NN), support vector machine (SVM), and naive-Bayes (NB).

A. TRAINING DATA SET GENERATION

We generate a training data set, which contains the $M$ training channel samples for scenarios A and B. The $m$th training
TABLE 1. Example of mapping table between class label (l) and indicated $S_{opt}$ ($N_A = 3$).

| class label (l) | indicated $S_{opt}$ | Binary representation of the indicated $S_{opt}$ | class label (l) | indicated $S_{opt}$ | Binary representation of the indicated $S_{opt}$ |
|----------------|---------------------|-----------------------------------------------|----------------|---------------------|-----------------------------------------------|
| l = 1          | $S_{opt} = \{1\}$   | $y = [1, 0, 0]$                              | l = 5          | $S_{opt} = \{1, 3\}$ | $y = [1, 0, 1]$                              |
| l = 2          | $S_{opt} = \{2\}$   | $y = [0, 1, 0]$                              | l = 6          | $S_{opt} = \{2, 3\}$ | $y = [0, 1, 1]$                              |
| l = 3          | $S_{opt} = \{3\}$   | $y = [0, 0, 1]$                              | l = 7          | $S_{opt} = \{1, 2, 3\}$ | $y = [1, 1, 1]$                              |
| l = 4          | $S_{opt} = \{1, 2\}$| $y = [1, 1, 0]$                              |               |                     |                                               |

Algorithm 5 Update the Parameters for Further Exploration of the Visited Node for Scenario B

1: $s[\hat{n}] = J, \hat{C}_{n} = c_{J}$
2: $t_B = \frac{T_B}{\sqrt{\bar{\rho} + \bar{\nu}_B}}$
3: $T_B = T_B - t_B T_B^H$
4: $v_{B,j} = \bar{v}_B, - h_j^H t_B^2, \forall j \in I_{\hat{n}, J}$
5: $\Delta_j = \log_2 \left( \frac{1 + \bar{\rho} \bar{\nu}_B}{(1 + \bar{\rho} \bar{\nu}_B)^{s_1} + \bar{\nu}_B^{s_2}} \right), \forall j \in I_{\hat{n}, J}$
6: $n = \hat{n}$

Algorithm 6 Prune the Child Nodes Below the Visited Node

1: if $n = 1$ then
2: $I_{0,0} = I_{0,0} \setminus \{J\}$
3: else
4: $I_{n-1,s[n-1]} = I_{n-1,s[n-1]} \setminus \{J\}$
5: end if
6: $n = J = 0, s = 0_k, T_B = I_{N_B}, T_E = I_{N_E}, v_{B,j} = h_j^H h_j, \forall j \in J$ for scenario A

$n = J = 0, s = 0_k, T_B = I_{N_B}, v_{B,j} = h_j^H h_j, \forall j \in J$ for scenario B

channel sample is denoted as $(H_m, G_m)$ for scenario A and $H_m$ for scenario B, where $1 \leq m \leq M$. Each training channel sample is converted to a feature vector, and then each feature vector is classified into one of $M$ classes, where $M$ is the number of possible transmit combinations in TAS, i.e., $M = \sum_{m=1}^{M} (N_A)^m = 2^{N_A} - 1$.

1) FEATURE VECTOR GENERATION

Because machine learning models take the real-valued feature vectors as input, the complex-valued training channel samples should be converted to the real-valued feature vector. The $m^{th}$ feature vector, $f_m \in \mathbb{R}^{1 \times N}$, is generated for each scenario. For scenario A, $f_m$ is given as

$$f_m = [h_{1,1}^m, \cdots, h_{N_B,N_A}^m, |g_{1,1}^m|, \cdots, |g_{N_E,N_A}^m|],$$

and for scenario B, $f_m$ is given as

$$f_m = [h_{1,1}^m, \cdots, h_{N_A}^m, |g_{1,1}^m|, \cdots, |g_{N_E,N_A}^m|].$$

where $h_{p,q}^m$ and $g_{p,q}^m$ are the $(p, q)^{th}$ elements of $H_m$ and $G_m$, respectively. The total number of features, $N$, are then given as $N_A (N_B + N_E)$ and $N_A N_B$ for scenarios A and B, respectively. These feature vectors should be normalized to avoid the local minima when training the machine learning models. Thus, the $n^{th}$ element of the $m^{th}$ normalized feature vector, $\tilde{f}_m$, is given as

$$\tilde{f}_m[n] = \frac{f_m[n] - \mu(f_m)}{\sigma(f_m)}.$$

2) LABELING OF THE FEATURE VECTORS

The optimal transmit antenna set, $S_{opt}$ for a given $m^{th}$ training channel sample is determined by the IB-TAS. Thus, we can determine the class label for each feature vector. Table 1 shows an example of a mapping table between class label and indicated $S_{opt}$ when $N_A = 3$. If $S_{opt}$ is $\{2, 3\}$, the class label is 6. Therefore, each feature vector can be mapped to one of the class labels based on $S_{opt}$ determined by the IB-TAS.

In particular, we utilize the binary representation of $S_{opt}$ for the supervised learning of the NN model, as shown in Table 1. The binary representation of $S_{opt}$ corresponding to the $m^{th}$ feature vector, $y_m \in \mathbb{R}^{N_A \times 1}$, is defined as

$$y_m[n] = \begin{cases} 1, & \text{if } n^{th} \text{ transmit antenna is selected,} \\ 0, & \text{otherwise,} \end{cases}$$

where $1 \leq n \leq N_A$.

B. NN-BASED TAS

Inspiring by a biologic neural network, the NN model has been widely utilized as a machine learning model [29]. The NN model consists of the input layer, one or more hidden layers, and the output layer. Each neuron in the current layer is connected to neurons in the previous and subsequent layers. The proposed NN model consists of the input layer, one hidden layer, and the output layer. The number of neurons in each layer is $N$ in the input layer, $N_h$ in the hidden layer, and $N_A$ in the output layer. When the number of neurons in the hidden layer is small or large, the NN causes underfitting or overfitting. Generally, $N_h$ is set to approximately $\frac{2}{3} N$ [30]. The output of hidden layer, $\tilde{h}_m \in \mathbb{R}^{N_h \times 1}$, is calculated as

$$\tilde{h}_m = \phi_h \left( W_h \tilde{f}_m + b_h \right),$$

where $\phi_h(\cdot)$ is the rectified linear unit (ReLU) function, $W_h \in \mathbb{R}^{N_h \times N}$ is the weight matrix associated with $\tilde{f}_m$, and $b_h \in \mathbb{R}^{N_h \times 1}$ is the bias vector in the hidden layer. Based on (46), the output of output layer, $\tilde{y}_m \in \mathbb{R}^{N_A \times 1}$, is calculated as

$$\tilde{y}_m = \phi_y \left( W_y \tilde{h}_m + b_y \right).$$
where $\varphi(\cdot)$ is the sigmoid function, $W_v (\in \mathbb{R}^{N_v \times N_y})$ is the weight matrix associated with $\bar{b}_m$, and $b_y (\in \mathbb{R}^{N_v \times 1})$ is the bias vector in the output layer.

The errors between $y_m$ and $\tilde{y}_m$ can be measured by the binary cross entropy loss function based on (45) and (47) [31]. The loss function, $L(y_m, \tilde{y}_m)$, is defined as

$$L(y_m, \tilde{y}_m) = -\frac{1}{N_A} \sum_{n=1}^{N_A} [y_m[n] \log(\tilde{y}_m[n])$$

$$+ (1 - y_m[n]) \log(1 - \tilde{y}_m[n])].$$

(48)

The proposed NN model is trained to minimize $L(y_m, \tilde{y}_m)$ through forward and backward propagation processes based on $M$ normalized feature vectors [15]. In the backpropagation process, the parameters of each layer, i.e., $W_h$, $W_v$, $b_h$, and $b_v$, are updated using the adaptive moment estimation algorithm [32], [33], [34].

After completing the training process, the proposed NN-based TAS can select the optimal transmit antenna set for a given channel sample. The given channel sample should be transformed into the normalized feature vector as an input to the NN-based TAS. Since the activation function of the output layer is the sigmoid function, the output value of each neuron in the output layer is determined to be between 0 and 1, i.e., $0 \leq \tilde{y}[n] \leq 1$ and $1 \leq n \leq N_A$. When $\tilde{y}[n]$ is greater than 0.5, it indicates that the $n^{th}$ transmit antenna is selected to maximize the secrecy capacity. Conversely, if $\tilde{y}[n]$ is less than 0.5, the $n^{th}$ transmit antenna is not selected.

### C. SVM-BASED TAS

SVM is used for the classification, which constructs a hyperplane or a set of hyperplanes to separate classes. In SVM, the hyperplane is designed to have the largest distance to the nearest training data point of any class.

The proposed SVM model applies the one-vs-rest strategy to classify the $M$ classes using $M$ hyperplanes [14], [20]. The objective function for constructing the hyperplane, which is used to separate between class $l$ and the others, is given as

$$f(w_l, b_l)$$

$$= \min_{w_l} \frac{1}{2} \|w_l\|^2 + C_l \sum_{m=1}^{M} \left[ \tilde{y}_m^{2l} \cos^2 \left( w_l^T \varphi(\tilde{f}_m) + b_l \right)$$

$$+ (1 - \tilde{y}_m) \cos^2 \left( w_l^T \varphi(\tilde{f}_m) + b_l \right) \right].$$

(49)

where $w_l$ and $b_l$ are learning parameters, $C_l$ is a non-negative scalar, $\cos z = \max((-1)^z z + 1.0)$, and

$$\tilde{y}_m^{2l} = \begin{cases} 1, & \text{if the class label corresponding to } \tilde{f}_m \text{ is } l, \\ 0, & \text{if the class label corresponding to } \tilde{f}_m \text{ is not } l. \end{cases}$$

In (49), $\varphi(\cdot)$ is a Gaussian radial-based kernel function and the $n^{th}$ element of $\varphi(\tilde{f}_m)$ is given as exp $\left(-\frac{\|\tilde{f}_m - \tilde{f}_l\|^2}{2\sigma^2}\right)$ with variance, $\sigma$. To minimize $f(w_l, b_l)$, $w_l$ and $b_l$ are updated iteratively using gradient descent method. That is,

$$w_l = w_l - \eta_0 \frac{\partial}{\partial w_l} f(w_l, b_l)$$

and

$$b_l = b_l - \eta_0 \frac{\partial}{\partial b_l} f(w_l, b_l),$$

(51)

where $\eta_0$ is the learning rate.

After completing the training process, the proposed SVM-based TAS can select the optimal transmit antenna set, i.e. the optimal class label $l^*$ among $M$ classes, for any given channel sample with the normalized feature vector. The optimal class label $l^*$ is given as

$$l^* = \arg \max_{l \in \{1, 2, \ldots, M\}} \left( w_l^T \varphi(\tilde{f}) + b_l \right).$$

(52)

### D. NB-BASED TAS

NB is a simple probabilistic classification applying Bayes’ theorem with the assumption of independence between the elements of the feature vector [35]. The NB-based TAS classifies $M$ normalized feature vectors by their class labels in the training process. Based on the normalized feature vectors whose class label belongs to $l$, we calculate mean, $\mu_{l,n}$, and variance, $\sigma_{l,n}$, of the $n^{th}$ element of them.

The posterior probability that any channel sample with the normalized feature vector, $\tilde{f}(\in \mathbb{R}^{1 \times N})$, belongs to class $l$ is given as

$$P(l|\tilde{f}) = \frac{P(\tilde{f}|l)P(l)}{P(\tilde{f})} = \prod_{n=1}^{N} \frac{P(\tilde{f}[n]|l)P(l)}{P(\tilde{f})},$$

(53)

where $P(\tilde{f}|l)$ is the probability of the occurrence of $\tilde{f}$ given the class label $l$, $P(l)$ is the prior probability of the class label $l$, and $P(\tilde{f})$ is the probability of the occurrence of $\tilde{f}$.

Thus, selecting the optimal transmit antenna set in the NB-based TAS is equivalent to choosing the class label with the maximal posterior probability for a given $\tilde{f}$ [20]. Since $P(\tilde{f})$ is independent of the class label and $P(l)$ is constant for all classes, the optimal class label $l^*$ is determined by

$$l^* = \arg \max_{l \in \{1, 2, \ldots, M\}} \prod_{n=1}^{N} \frac{P(\tilde{f}[n]|l)}{P(\tilde{f})}$$

$$= \arg \max_{l \in \{1, 2, \ldots, M\}} \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi \sigma_{l,n}^2}} \exp \left(-\frac{(\tilde{f}[n] - \mu_{l,n})^2}{2\sigma_{l,n}^2}\right).$$

(54)

### E. COMPUTATIONAL COMPLEXITY ANALYSIS

Since the machine learning models are implemented based on offline training using the data set, the computational complexity is generally measured only for online processing. In this subsection, we evaluate the computational complexity of ML-TAS schemes, which is required for the prediction, i.e. TAS.

In the NN-based TAS, the computational complexity from the input layer to the hidden layer is associated with (46) and it is given as $O(NN_h)$. Moreover, the computational complexity from the hidden layer to the output layer is associated
with (47) and it is given as \(O(N_BN_A)\). Therefore, the total computational complexity of the proposed NN-based TAS scheme is given as \(O(NN_BN_A)\).

When selecting the class label \(\ell^*\) for a given channel sample in the SVM-based TAS, the calculation of \(\varphi(\tilde{f})\) causes high computational complexity. Furthermore, it needs to be calculated \(M\) times. Thus, its computational complexity is dominant compared to the others. If the number of features, \(N\), is larger than \(M\), the total computational complexity of the SVM-based TAS is given as \(O(N^2)\) [20]. As the number of transmit antennas increases, \(M\) will exponentially increase. When \(M\) is larger than \(N\), the total computational complexity of the SVM-based TAS becomes \(O(MN)\).

When selecting the class label \(\ell^*\) for a given channel sample in the NB-based TAS, the calculation of \(\prod_{n=1}^{N} P(\tilde{f}[n] | \ell^*)\) causes high computational complexity. Furthermore, it needs to be calculated \(M\) times. Since its computational complexity is dominant compared to the others, the total computational complexity of the NB-based TAS is given as \(O(MN)\).

**V. SIMULATION RESULTS**

In this section, we evaluate the secrecy capacity and the computational complexity of the proposed IB-TAS and the ML-TAS schemes for scenarios A and B. The IB-TAS simultaneously determines the optimal number of transmit antennas and the corresponding transmit antenna indices in MIMOME channel. To the best of our knowledge, there have been no studies on TAS to simultaneously determine the optimal number of transmit antennas and the corresponding antenna indices. To evaluate the performance of the IB-TAS, the exhaustive search-based TAS (ES-TAS) that guarantees optimal performance is used as a reference scheme.

Fig. 3 shows the secrecy capacity comparison between the ES-TAS and the IB-TAS. The ES-TAS uses (21) and (41) to determine the optimal transmit antenna set for scenarios A and B. Both schemes show nearly identical secrecy performance in each scenario. It can be interpreted that the IB-TAS guarantees optimal performance. As shown in Fig. 3, the secrecy capacity shown in scenario A is higher than that shown in scenario B due to the accurate CSI of Eve available at Alice.

However, the performance loss between scenarios A and B is relatively small. Table 2 shows the number of transmit antennas selected by the ES-TAS and the IB-TAS in scenarios A and B. As shown in Table 2, the ES-TAS and the IB-TAS in scenario A select 0.3~1.2 more transmit antennas than those in scenario B.

Table 3 shows the accuracy of the selected transmit antenna indices determined by the IB-TAS compared to the ES-TAS in each scenario. For example, 100% means that the transmit antenna set determined by the IB-TAS includes the optimal transmit antenna set determined by the ES-TAS. Table 3 thus shows that the accuracy of transmit antenna indices selected by the IB-TAS are nearly identical to the optimal transmit antenna set determined by the ES-TAS in each scenario. As shown in Table 3, the accuracy of transmit antenna indices selected by the IB-TAS is nearly identical to the optimal transmit antenna set determined by the ES-TAS in each scenario. Based on Tables 2 and 3, it is observed that the number of selected transmit antennas causes the performance loss between scenarios A and B rather than the accuracy of selected transmit antenna indices. However, its difference is only in the range of 0.3 to 1.2. Therefore, the performance loss between scenarios A and B is relatively small.

**A. SECRECY CAPACITY**

Fig. 4 shows the secrecy capacity of the IB-TAS and the ML-TAS when SNR at Bob is fixed as 10 dB and SNR at Eve is 0 to 20 dB. In Fig. 4, the solid and dotted lines show the

**TABLE 2.** Number of selected transmit antennas versus \(\rho_E\) \((N_A = 32, N_B = 4, N_E = 4, and \rho_B = 10\ dB)\).

| Schemes   | \(\rho_E\) | 0 dB | 5 dB | 10 dB | 15 dB | 20 dB |
|-----------|------------|------|------|-------|-------|-------|
| ES-TAS in scenario A | 13.9 | 9.2  | 6.4  | 5.7   | 2.2   |
| IB-TAS in scenario A | 13.6 | 8.9  | 6.1  | 5.4   | 2.0   |
| ES-TAS in scenario B | 12.8 | 8.0  | 5.3  | 4.5   | 1.8   |
| IB-TAS in scenario B | 12.4 | 7.8  | 5.1  | 4.2   | 1.7   |

**TABLE 3.** Accuracy of selected transmit antenna indices between the ES-TAS and the IB-TAS versus \(\rho_E\) \((N_A = 32, N_B = 4, N_E = 4, and \rho_B = 10\ dB)\).

| Schemes   | \(\rho_E\) | 0 dB | 5 dB | 10 dB | 15 dB | 20 dB |
|-----------|------------|------|------|-------|-------|-------|
| IB-TAS in scenario A | 94.7% | 95.4% | 94.1% | 93.4% | 94.8% |
| IB-TAS in scenario B | 94.9% | 94.7% | 94.6% | 94.9% | 95.1% |
TABLE 4. Classification accuracy in scenario A.

| Schemes       | $P_B$ | 0 dB | 5 dB  | 10 dB | 15 dB | 20 dB |
|---------------|-------|------|-------|-------|-------|-------|
| NN-based TAS  |       | 93.6%| 94.0% | 93.7% | 94.0% | 94.0% |
| SVM-based TAS |       | 89.8%| 82.5% | 90.6% | 90.2% | 91.6% |
| NB-based TAS  |       | 88.1%| 79.1% | 87.9% | 87.8% | 87.8% |

TABLE 5. Classification accuracy in scenario B.

| Schemes       | $P_B$ | 0 dB | 5 dB  | 10 dB | 15 dB | 20 dB |
|---------------|-------|------|-------|-------|-------|-------|
| NN-based TAS  |       | 93.6%| 94.2% | 93.2% | 92.9% | 93.8% |
| SVM-based TAS |       | 88.7%| 82.4% | 89.4% | 91.0% | 92.7% |
| NB-based TAS  |       | 87.4%| 78.4% | 86.9% | 84.4% | 89.4% |

FIGURE 4. Secrecy capacity versus $P_E$ ($N_A = 32, N_B = N_E = 4, \text{ and } P_B = 10 \text{ dB}$).

The IB-TAS shows superior achievable secrecy capacity compared to the ML-TAS schemes because the IB-TAS achieves near-optimal performance as shown in Fig. 4. Table 4 shows the classification accuracy when comparing the optimal transmit antenna set determined by each ML-TAS scheme with the optimal transmit antenna set determined by the IB-TAS in scenario A. As shown in Table 4, the classification accuracy gradually decreases in the order of the NN-based TAS, the SVM-based TAS, and the NB-based TAS. This is because NB is a probabilistic method based on conditional probability, whereas SVM is a feature vector-based deterministic classification method for maximizing the margin between different classes. In addition, unlike SVM and NB, NN solves nonlinear problems through its parametric forms. In addition, it is more fault-tolerant and can handle incomplete data and noise. Therefore, as shown in Fig. 4, the NN-based TAS shows better secrecy capacity than other ML-TAS schemes.

Table 5 shows the classification accuracy when comparing the optimal transmit antenna set determined by each ML-TAS scheme with the optimal transmit antenna set determined by the IB-TAS in scenario B. The performance trend of the proposed TAS schemes is identical to those in scenario A. As mentioned in Fig. 3, the secrecy capacities of the proposed TAS schemes shown in scenario A are higher than those shown in scenario B due to the accurate CSI of Eve available at Alice.

Fig. 5 shows the secrecy capacity of the IB-TAS and the ML-TAS when SNR at Eve is fixed as 10 dB and SNR at Bob is 0 to 20 dB. In Fig. 5, the solid and dotted lines show the secrecy capacities of TAS schemes in scenario A, and the no-line markers show the secrecy capacities of TAS schemes in scenario B. The secrecy capacity is maximized when SNR at Bob is much greater than SNR at Eve, i.e. the point when SNR at Bob = 20 dB in Fig. 5. This is because

As the number of epochs increases, the performances of the ML-TAS increase and eventually converge. The classification accuracy of the NN-based TAS and the SVM-based TAS increases as the number of epochs increases because the parameters of each method gradually have optimal values. On the other hand, since the NB-based TAS determines the optimal transmit antenna set through the posterior probability for the training channels, the classification accuracy is constant even when the number of epochs increases.
FIGURE 5. Secrecy capacity versus $\rho_B$ ($N_A = 32, N_B = N_E = 4$, and $\rho_E = 10$ dB).

FIGURE 6. Secrecy capacity versus $\rho_E$ in scenario A ($N_A = 32(64), N_B = 4(8), N_E = 4(8)$, and $\rho_B = 10$ dB).

the increased capacity between Alice and Bob leads to an increase in the secrecy capacity. Conversely, if SNR at Eve is much greater than SNR at Bob, the capacity between Alice and Eve, i.e. the eavesdropping ability of Eve, increases. It then leads to a decrease the secrecy capacity.

Fig. 6 shows the secrecy capacity according to the number of antennas for Alice, Bob, and Eve, in scenario A. The number of selected transmit antennas for each case is summarized in Table 6. The increase in the number of antennas at Alice provides the opportunity to increase the capacity of Alice-Bob channel or to decrease the capacity of Alice-Eve channel due to the increase in the antenna selection diversity.

| Table 6. Number of selected transmit antennas versus $\rho_E$ in scenario A ($N_A = 32(64), N_B = 4(8), N_E = 4(8)$, and $\rho_B = 10$ dB). |
|---|
| Schemes | $\rho_E$ | 0 dB | 5 dB | 10 dB | 15 dB | 20 dB |
| IB-TAS | $\rho_E = 5$ dB | 13.9 | 8.6 | 5.3 | 4.3 | 2.1 |
| IB-TAS | $\rho_E = 10$ dB | 16.1 | 11.2 | 11.2 | 6.2 | 3.2 |
| IB-TAS | $\rho_E = 15$ dB | 31.2 | 30.4 | 30.4 | 29.4 | 29.2 |
| IB-TAS | $\rho_E = 20$ dB | 5.3 | 3.9 | 3.9 | 1 | 1 |

As a result, as shown in Table 6, the number of selected transmit antennas increases.

The increase in the number of antennas at Bob increases the rank of Alice-Bob channel. It then leads to the increase in the capacity of Alice-Bob channel so that the secrecy capacity is also increased. In addition, the capacity of Alice-Bob channel increases more than the Alice-Eve capacity as the number of selected transmit antennas increases because the number of antennas at Eve is less than that at Bob. Thus, as shown in Table 6, the number of selected transmit antennas has increased significantly as the number of antennas at Bob increases.

The increase in the number of antennas at Eve increases the rank of Alice-Eve channel. It then leads to the increase in the capacity of Alice-Eve channel so that the secrecy capacity is decreased. In addition, the Alice-Eve capacity increases more than the Alice-Bob capacity as the number of selected transmit antennas increases because the number of antennas at Bob is less than that at Eve. Thus, as shown in Table 6, the number of selected transmit antennas has decreased significantly as the number of antennas at Eve increases.

B. COMPUTATIONAL COMPLEXITY ANALYSIS

Table 7 summarizes the computational complexity of the IB-TAS and the ML-TAS schemes when Alice selects the optimal transmit antenna set. In the IB-TAS, $N_T$ is the total number of visited nodes to find the optimal transmit antenna set, and it can be determined by Algorithm 1. In the ML-TAS,

As a result, as shown in Table 6, the number of selected transmit antennas increases.

The increase in the number of antennas at Bob increases the rank of Alice-Bob channel. It then leads to the increase in the capacity of Alice-Bob channel so that the secrecy capacity is also increased. In addition, the capacity of Alice-Bob channel increases more than the Alice-Eve capacity as the number of selected transmit antennas increases because the number of antennas at Eve is less than that at Bob. Thus, as shown in Table 6, the number of selected transmit antennas has increased significantly as the number of antennas at Bob increases.

The increase in the number of antennas at Eve increases the rank of Alice-Eve channel. It then leads to the increase in the capacity of Alice-Eve channel so that the secrecy capacity is decreased. In addition, the Alice-Eve capacity increases more than the Alice-Bob capacity as the number of selected transmit antennas increases because the number of antennas at Bob is less than that at Eve. Thus, as shown in Table 6, the number of selected transmit antennas has decreased significantly as the number of antennas at Eve increases.

B. COMPUTATIONAL COMPLEXITY ANALYSIS

Table 7 summarizes the computational complexity of the IB-TAS and the ML-TAS schemes when Alice selects the optimal transmit antenna set. In the IB-TAS, $N_T$ is the total number of visited nodes to find the optimal transmit antenna set, and it can be determined by Algorithm 1. In the ML-TAS,
The number of transmit antennas causes the exponential increase in the IB-TAS and two ML-TAS schemes (i.e. SVM-based TAS and NB-based TAS), respectively. Since the increase in the number of transmit antennas causes the exponential increase in $N_T$ and $\mathcal{M}$, the asymptotic computational complexity of the IB-TAS, the SVM-based TAS, and the NB-based TAS increases exponentially, as shown in Fig. 7.

In the NN-based TAS, the values of both $N$ and $N_h$ increase linearly as the number of transmit antennas increases. Therefore, the computational complexity of the NN-based TAS is significantly reduced compared to other schemes, as shown in Fig. 7. The computational complexity of the IB-TAS and the ML-TAS schemes in scenario B also shows the same trend as in scenario A for the same reason as above.

VI. CONCLUSION

In this paper, we propose the IB-TAS schemes to maximize secrecy capacity based on Eve’s available CSI at Alice, i.e. perfect and partial cases, in MIMOME channel. The IB-TAS can determine the optimal number of transmit antennas and select the corresponding optimal transmit antenna indices, simultaneously. However, most TAS schemes generally increase the computational complexity as the number of transmit antennas. To reduce the computational complexity caused by TAS, we also propose the ML-TAS schemes such as the NN-based TAS, the SVM-based TAS, and the NB-based TAS. Simulation results show that the IB-TAS guarantees the near-optimal secrecy capacity, and the NN-based TAS minimizes the computational complexity at the cost of minimal secrecy capacity loss. Through the secrecy capacity and computational complexity comparisons among the proposed schemes, we provide an insight into the potential benefits of integrating machine learning with TAS.

REFERENCES

[1] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. Di Renzo, “Safeguarding 5G wireless communication networks using physical layer security,” IEEE Commun. Mag., vol. 53, no. 4, pp. 20–27, Apr. 2015.
[2] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, “Physical layer security in wireless networks: A tutorial,” IEEE Wireless Commun., vol. 18, no. 2, pp. 66–74, Apr. 2011.
[3] A. D. Wyner, “The wire-tap channel,” Bell Syst. Tech. J., vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
[4] A. Khisti and G. W. Wornell, “Secure transmission with multiple antennas I: The MISOME wiretap channel,” IEEE Trans. Inf. Theory, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.
[5] A. Khisti and G. Wornell, “Secure transmission with multiple antennas—Part II: The MIMO wiretap channel,” IEEE Trans. Inf. Theory, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
[6] A. Mukherjee and A. L. Swindlehurst, “Robust beamforming for security in MIMO wiretap channels with imperfect CSI,” IEEE Trans. Signal Process., vol. 59, no. 1, pp. 351–361, Jan. 2011.
[7] T. Gucluoglu and T. M. Duman, “Performance analysis of transmit and receive antenna selection over flat fading channels,” IEEE Trans. Wireless Commun., vol. 7, no. 8, pp. 3056–3065, Aug. 2008.
[8] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” IEEE Trans. Wireless Commun., vol. 7, no. 6, pp. 2180–2189, Jun. 2008.
[9] S. Yin, S. Im, J. Kim, and J. Han, “On the secrecy rate and optimal power allocation for artificial noise assisted MIMO-MIMO channels,” IEEE Trans. Veh. Technol., vol. 67, no. 4, pp. 3098–3113, Apr. 2018.
[10] N. Yang, H. A. Suraweera, I. B. Collings, and C. Yuen, “Physical layer security of TAS/MRC with antenna correlation,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 1, pp. 254–259, Jan. 2013.
[11] N. Yang, P. L. Yeoh, M. Elkashlan, R. Schober, and I. B. Collings, “Transmit antenna selection for security enhancement in MIMO wiretap channels,” IEEE Trans. Commun., vol. 61, no. 1, pp. 144–154, Jan. 2013.
[12] N. Yang, P. L. Yeoh, M. Elkashlan, R. Schober, and J. Yuan, “MIMO wiretap channels: Secure transmission using transmit antenna selection and receive generalized selection combining,” IEEE Commun. Lett., vol. 17, no. 9, pp. 1574–1577, Sep. 2013.
[13] S. Assaad, A. Berenzy, A. M. Rahbar, R. R. Müller, and R. F. Schaefar, “Optimal transmit antenna selection for massive MIMO wiretap channels,” IEEE J. Sel. Areas Commun., vol. 36, no. 4, pp. 817–828, Apr. 2018.
[14] C. Ouyang, Z. Ou, L. Zhang, and H. Yang, “Optimal transmit antenna selection algorithm in massive MIMOME channels,” in Proc. IEEE Wireless Commun. Networking Conf. (WCNC), Apr. 2019, pp. 1–6.
[15] J. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, Cambridge, MA, USA: MIT Press, 2016.
[16] J. Joung, “Machine learning-based antenna selection in wireless communications,” IEEE Commun. Lett., vol. 20, no. 11, pp. 2241–2244, Nov. 2016.
[17] P. Yang, Y. Xiao, M. Xiao, Y. L. Guan, S. Li, and W. Xiang, “Adaptive spatial modulation MIMO based on machine learning,” IEEE J. Sel. Areas Commun., vol. 37, no. 9, pp. 2117–2131, Sep. 2019.
[18] M. S. Ibrahim, A. S. Zamzam, X. Fu, and N. D. Sidiropoulos, “Learning-based antenna selection for multicasting,” in Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Kalamata, Greece, Jun. 2018, pp. 1–5.
[19] W. Yu, T. Wang, and S. Wang, “Multi-label learning based antenna selection in massive MIMO systems,” IEEE Trans. Veh. Technol., vol. 70, no. 7, pp. 7255–7260, Jul. 2021.
[20] D. He, C. Liu, T. Q. S. Quek, and H. Wang, “Transmit antenna selection in MIMO wiretap channels: A machine learning approach,” IEEE Wireless Commun. Lett., vol. 7, no. 4, pp. 634–637, Aug. 2018.
[21] Y. Hu, L. Li, J. Yin, H. Zhang, W. Liang, A. Gao, and Z. Han, “Optimal transmit antenna selection strategy for MIMO wiretap channel based on deep reinforcement learning,” in Proc. IEEE/CIC Int. Conf. Commun. China (ICCC), Aug. 2018, pp. 803–807, Apr. 2019.
[22] T. S. Rappaport, Wireless Communications: Principles and Practices, Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[23] R. A. Horn, Matrix Analysis. New York, NY, USA: Cambridge Univ. Press, 1986.
[24] P. M. Narendra and K. Fukunaga, “A branch and bound algorithm for feature subset selection,” IEEE Trans. Comput., vol. C-26, no. 9, pp. 917–922, Sep. 1977.
[25] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*. Boston, MA, USA: Allyn and Bacon, 1964.

[26] B. Hall, *Lie Groups, Lie Algebras, Representations: An Elementary Introduction*. Cham, Switzerland: Springer, 2015.

[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Norwell, MA, USA: Now Publishers, 2004.

[29] J. Heaton, *Artificial Intelligence for Humans*, vol. 3. North Charleston, SC, USA: CreateSpace Independent Publ. Platform, Oct. 2015.

[30] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: A survey,” *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.

[31] S. Mannor, D. Peleg, and R. Rubinstein, “The cross entropy method for classification,” in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 561–568.

[32] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[33] D. P. Kingma and J. Ba, “ADAM: A method for stochastic optimization,” in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.

[34] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[35] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yü, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, “Top 10 algorithms in data mining,” *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, Jan. 2008.

**DONGMIN SEO** received the B.S. and M.S. degrees in electronic engineering from Wonkwang University, Iksan, South Korea, in 2018 and 2021, respectively. He is currently an Aspiring Researcher in beyond 5G and 6G communication systems. His current research interests include MIMO, physical layer security, and transmit antenna selection in wireless communication fields.

**HYUKMIN SON** (Member, IEEE) was born in South Korea, in 1981. He received the B.S., M.S., and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, South Korea, in 2004, 2006, and 2011, respectively. From March 2011 to November 2012, he was working with the Advanced Communication Technology Research Laboratory, LG Electronics. He was responsible for the advanced receiver design and 3GPP LTE (-A) standard, in 2011. For a year in 2012, he was researching for 5G wireless communication. From December 2012 to November 2013, he held a postdoctoral position with the Communications and Signal Processing Group, Department of Electrical and Electronic Engineering, Imperial College London. He was also with the Multi-Dimensional Insight Laboratory, Yonsei University, as a Postdoctoral Research Fellow, from December 2013 to February 2014. From March 2014 to February 2017, he was a Senior Research Engineer with the Advanced Communication Research Team, Samsung Electronics, where he was responsible for 3GPP 5G NR standard. From March 2017 to February 2022, he was an Associate Professor with the Department of Electronic Engineering, Wonkwang University. He is currently an Associate Professor with the Department of Electronic Engineering, Gachon University, Seongnam, South Korea. His current research interests include MIMO, cooperative communication, physical layer security, V2X, NTN, and beyond 5G and 6G communication systems.

***