An Uplink Frequency-Time Index Modulation Multiple Access for 6G Networks

Muhammad Sajid Sarwar[1], I Nyoman[1], Apraz Ramatryana[1], Gelar Budiman[1], and Soo Young Shin[1]

[1]Affiliation not available

January 03, 2025

An Uplink Frequency-Time Index Modulation Multiple Access for 6G Networks

Muhammad Sajid Sarwar, I Nyoman Apraz Ramatryana, Gelar Budiman, Soo Young Shin, *Senior Member, IEEE*

***Abstract*—This investigation introduces an uplink frequency-time (FT) index modulation multiple access (IMMA) (FT-IMMA) scheme designed to augment the spectral efficiency (SE) of conventional orthogonal frequency-division multiple access (OFDMA). The fundamental principle of this approach involves allocating FT resources within OFDMA to users through two-dimensional index modulation (IM). A subset of input bits from each user dictates the indices of subcarriers and time slots within the FT resources of OFDMA, while the remaining bits undergo a conventional M-ary modulation. FT-IMMA obviates the necessity for a prior scheduling mechanism, thereby reducing communication overhead. While this grant-free access approach may potentially introduce data collisions within the system, the collision probability of the proposed model is observed to be markedly lower compared to traditional IMMA systems. However, the two-dimensional IM approach increases the computational complexity of conventional maximum likelihood detection. To address this, the authors propose using orthogonal matching pursuit, a compressive sensing technique that effectively reduces the computational burden. Through analytical investigation and simulation results, FT-IMMA demonstrates superior performance compared to its counterparts, showcasing notable advantages in both bit error rate and SE.**

***Index Terms*—Compressive sensing, index modulation, multiple access, NOMA, OFDMA, spectral efficiency, unscheduled communications.**

I. Introduction

Spectrum and energy-efficient multiple access (MA) techniques are required to serve rapidly growing smart devices and enormous data traffic. Previous studies divide MA techniques into two major categories: orthogonal MA (OMA) and non-orthogonal MA (NOMA) schemes [1]. OMA includes time/frequency/code division multiple access (TDMA/FDMA/CDMA) techniques that allocate orthogonal resources to multiple users. The number of users served by OMA schemes is limited by the number of available orthogonal resources, which results in low spectral efficiency (SE). NOMA techniques can improve user connectivity by providing simultaneous access to spectrum resources through distinct power levels, constellation design, or code. The SE was achieved at the expense of user interference, which is more prevalent in the case of massive connectivity [2]. User pairing and optimal power allocation in the power domain of NOMA and constellation design in the code domain of NOMA constitute other challenges [1].

Recently, index modulation (IM) has been introduced to enhance system performance, which employs a fraction of resource entities, such as subcarriers, antennas, time slots, channel states, etc., to convey information [3]. The active index pattern of the resources, which was determined by input bits, was employed as supplementary information in addition to the conventional constellation symbols [4]–[7]. This technique has improved energy efficiency (EE) and bit error rate (BER); however, partial inactivation results in a loss in SE compared with conventional techniques [8], [9]. To improve SE, more than one transmit entity was utilized simultaneously to carry information [10]. Another approach that was used to improve SE involved varying the number of subcarrier activations, where a group of subcarriers could range from having only one active subcarrier to all subcarriers [11].

An uplink MA scheme inspired by the concept of IM was proposed in [12], [13], termed index modulation MA (IMMA). This technique is categorized as NOMA and operates in a time-slotted manner, where a particular slot is assigned to a user based on a portion of the input bit block of that user. Therefore, it contains two information-carrying units: the index of the active time slot and the constellation symbol. The IMMA principle was applied in the frequency domain and termed IM orthogonal frequency division MA (IM-OFDMA) [14], [15]. In IM-OFDMA, the subcarrier block of OFDMA was partitioned into chunks comprising a certain number of subcarriers, and each chunk could serve an arbitrary number of users. Users could modulate their data over single or multiple subcarriers in each chunk, and active indices were also considered a source of information. These techniques did not require prior scheduling and therefore improved EE [12], [14]. Data collisions between different users could occur at the common base station because no coordination existed among the operating users for resource selection [16], [17]. Suppose the intended users are assigned to the same resource owing to similar index bits. In that case, a collision can occur. Moreover, to achieve higher SE, more active subcarriers are required, which can even worsen the collision scenario. This will result in a waste of resource entities and spectral inefficiency. A summary of time/frequency-domain IMMA techniques proposed in the literature is given in Table I.

Compressive sensing (CS) has emerged as a powerful technique for recovering sparse signals from a limited set of linear measurements [18]. It has found widespread applications in areas such as channel estimation [19], interference cancellation [20], and signal detection [21]. In particular, a CS-based low-complexity decoder was introduced in [22]–[24] for

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2019R1A2C1089542).

M. S. Sarwar and S. Y. Shin are with the WENS Laboratory, Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea (e-mail: sajidsarwarch@hotmail.com, wdragon@kumoh.ac.kr).

I. N. A. Ramatryana is with the Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany (e-mail: ramatryana@gmail.com).

G. Budiman is with the School of Electrical Engineering, Telkom University, Bandung, Indonesia (email: gelarbudiman@telkomuniversity.ac.id)

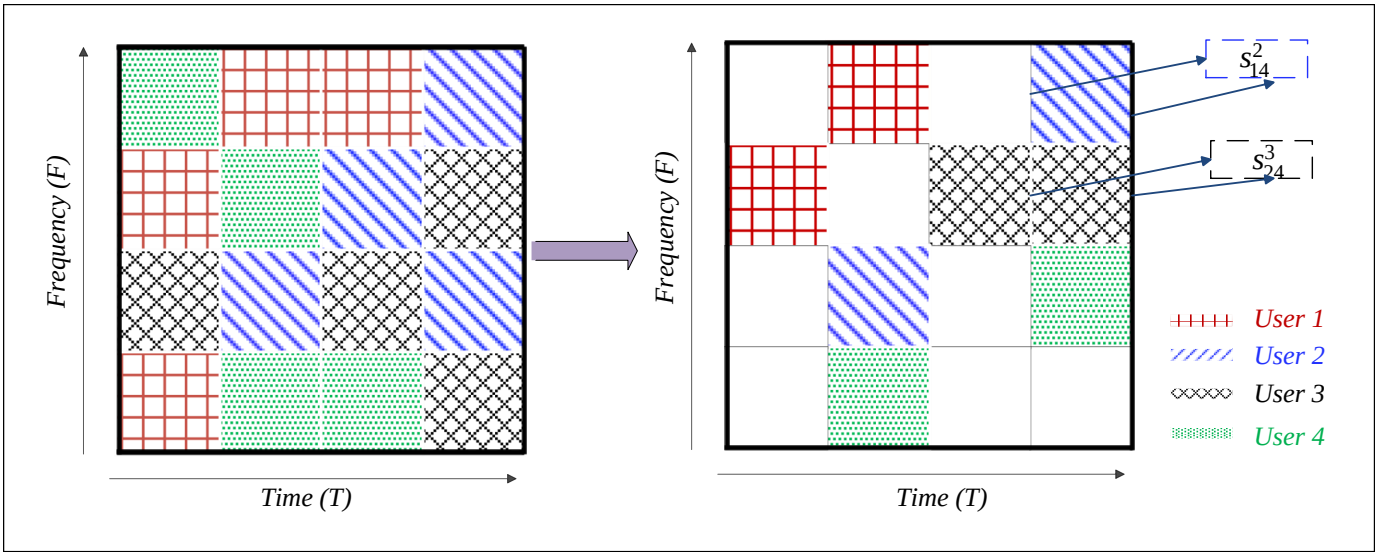


Fig. 1: Resource allocation (OFDMA (left), FT-IMMA (right)).

spatial modulation (SM) in massive multi-input, multi-output (MIMO) systems. CS was also utilized for uplink multiuser signal decoding in quadrature SM MIMO systems in [25], demonstrating its efficiency in multiuser detection (MUD) scenarios. Additionally, partially active subcarriers facilitated CS-based decoders to recover the sparse signal of OFDM-IM [26]–[28]. This concept has also been extended to space-frequency domain IM [29], space-time domain IM [30], and space-time-frequency domain IM [31].

Common algorithms used in the CS framework include basis pursuit (BP), least absolute shrinkage and selection operator (LASSO), and greedy algorithms like orthogonal matching pursuit (OMP). While BP and LASSO offer high accuracy in recovering sparse signals, OMP provides a faster implementation with near-optimal performance, making it particularly suitable for real-time applications in wireless communication systems [30]. The advantages of OMP in fast and efficient sparse signal recovery were investigated and highlighted in [32], positioning it as a strong candidate for use in communication systems requiring both low complexity and high performance. OMP was used to reduce the detection complexity of orthogonal chirp division multiplexing in a dual-function radar communication system [33].

Against this background, our study proposes frequency-time (FT)-based IMMA (FT-IMMA) to improve SE and decrease collision rates in both IMMA and IM-OFDMA. Unlike conventional techniques, the proposed technique chooses subcarriers over specific time slots, e.g., a FT resource of classical OFDMA, as an example is shown in Fig. 1. A resource grid of OFDMA is partitioned into chunks, and each chunk of FT resources is simultaneously accessible to an arbitrary number of users. Each user's input bits contain three portions. Two portions are utilized to select one or more FT resources, and the third is modulated by utilizing standard M-ary constellation symbols. Users randomly access FT resources, which can lead to collisions; however, the proposed system exhibits a significantly low collision probability to reach the same SE as its traditional IMMA counterparts. A relatively lower-order modulation is required because more information can be transmitted virtually via two-dimensional resource index selection. Additionally, the process does not require any prior scheduling of users, resulting in EE. The SE, EE, and BER for the proposed system are examined and compared with those of OFDMA, IMMA, and IM-OFDMA. Furthermore, it is observed that FT-IMMA achieves performance gains at the cost of computational complexity. Consequently, an OMP-based suboptimal MUD is suggested, which leverages the inherited sparsity of IM to minimize the number of calculations.

The rest of the article is arranged as follows: Section II discusses the system model for FT-IMMA. Sections III-A, III-B, III-C, and III-D provide the key performance indicators, including collision probability, EE, achievable data rate, and BER bound, respectively. The theoretical and simulation results of the proposed scheme are discussed in Section IV. Finally, the research work is concluded with a discussion of future challenges in Section V.

Notation: The following notations are consistently employed throughout this article: Boldface lowercase letters signify vectors, while boldface uppercase letters denote matrices. The symbol s^u_{ft} represents the data symbol from the u^{th} user, mapped onto the f^{th} subcarrier and t^{th} time slot. The matrix $\mathbf{S}^u_{ft}$ signifies the symbol matrix mapped across multiple active resources. The transpose operation is denoted by $(\cdot)^T$. On the receiving end, the estimated vectors are indicated by $(\hat{\cdot})$.

II. System Model

The block representation of the FT-IMMA transmitter is shown in Fig. 2. Consider a system with a total bandwidth of B divided into F_s orthogonal subcarriers across T_s independent time slots. In this case, F_sT_s time-frequency resources exist. F_s is further partitioned into $G_F = F_s/F$ chunks, where F is

the length of a chunk spread across $T = T_s/G_T$ time slots and G_T denotes the number of chunks of T_s. In Fig. 1, $T = 4$ and $F = 4$ make a two-dimensional grid of $FT = 16$ resources. The number of such resource grids is represented by $G = \frac{F_s T_s}{FT}$. Each FT grid can serve U users, and the system can accommodate GU users. For simplicity, the system model is explained only for a single resource grid, and the model can be readily expanded to include G grids. In a grid, each user can modulate data over single or multiple FT resources; for simplicity, it is assumed that all users activate the same number of FT resources, e.g., two resources are activated by each user in Fig. 1.

The incoming data of each user is divided into three parts: b_1, b_2, and b_3, which represent the frequency index, time index, and constellation bits, respectively. An example of a predefined index mapping table (IMT) is shown in Table II. Notably, b_1 can be paired with b_2 in four alternative methods in each cell. This provides an extra degree of freedom compared with IMMA and IM-OFDMA. b_1 is utilized to select f subcarriers from F subcarriers, whereas b_2 is utilized to select t time slots from T slots. It allows the user to select K 'FT resources' to map b_3 bits by utilizing the conventional M-ary constellation. The transmitted bit block of the u^{th} user is expressed as $b = b_1 + b_2 + b_3$ and is expanded as follows:

$$b = \left\lfloor \log_2 \binom{F}{f} \right\rfloor + \left\lfloor \log_2 \binom{T}{t} \right\rfloor + K \log_2(M), \quad \text{where } K = \begin{cases} f, \text{ if } f \geq t \\ t, \text{ if } t \geq f \end{cases} \tag{1}$$

where M is the modulation order adopted by the user. In this scenario, if $f = 2$ and $t = 1$, two different frequency resources

TABLE I: Summary of Time Domain and Frequency Domain IMMA Schemes

Scheme	IM Type	Features
Uplink IMMA [12], [34], [35]	One-dimensional	An IM concept was introduced for uplink multiple access scenarios to allow a substantial number of users to simultaneously occupy a limited number of time slots. Each user endeavors to access a single time slot by utilizing a portion of transmission data. This method can transmit $b = \lfloor \log_2(T) \rfloor + \log_2(M)$ bits per user.
Uplink IM-OFDMA [14]	One-dimensional	A generalized form of frequency domain IMMA has been introduced, enabling users to occupy multiple frequency channels concurrently to enhance SE. The number of bits transmitted per user in IM-OFDMA is calculated as $b = \left\lfloor \log_2 \binom{F}{f} \right\rfloor + \log_2(M)$.
Quadrature uplink IMMA [15]	One-dimensional	The first two portions of each user's bits were employed to select a time slot for the independent transmission of the in-phase and quadrature components of the constellation symbol, respectively. The third portion was used to transmit the M-ary symbol. The number of transmitted bits for a user is given by $b = \lfloor \log_2(T^2) \rfloor + \log_2(M)$.
Low-complexity uplink IMMA [13]	One-dimensional	A low-complexity detector based on the log-likelihood ratio was suggested, approaching optimal ML detection. Its SE aligns with that of classical IMMA.
CNN-Based uplink IM-OFDMA [36]	One-dimensional	A detector employing a convolutional neural network was applied to conventional IM-OFDMA utilizing transmitted signals and channel state information as input. This approach has reduced the computational complexity of the ML detector, albeit with a marginal decline in BER.
Precoded uplink IMMA [16]	One-dimensional	A three-step precoding process encompassing channel pre-equalization, power scaling, and phase adjustment, has been implemented to better distinguish the combined received signal from the users. The principal outcome of this research was BER improvement, whereas SE is the same as conventional IMMA.
An overview of uplink/downlink IMMA for 6G [17]	One-dimensional	This research has outlined the fundamental principles of IMMA across spatial, temporal, frequency, code, and power domains. In addition, it has investigated the potential applications of IMMA technology in vehicular networks, reconfigurable intelligent surface-aided networks, cooperative networks, and secure networks.
Uplink IM-OFDMA [37]–[40]	One-dimensional	These articles showcase that IM-OFDMA exhibits superior performance in terms of BER than traditional OFDMA in the presence of carrier frequency offset, transmitter/receiver in-phase, and quadrature imbalance.
Downlink IM-NOMA [41]	One-dimensional	In contrast to classical downlink NOMA, it has offered an SE gain for a two-user scenario. Each user allocates a portion of bits to randomly select a power level from a set of P power levels. For each user, it can transmit $b = \lfloor \log_2(P) \rfloor + \log_2(M)$ bits.
Downlink NOMA-SM [42]	One-dimensional	This study explored how NOMA with spatial modulation (NOMA-SM) can offer an SE gain in a multi-user scenario. Each user allocates a portion of its bits to randomly select a transmitting antenna, and the base station transmits a superimposed SM signal to all users. For each user, the achievable number of bits is given by $b = \lfloor \log_2(N_t) \rfloor + \log_2(M)$, where N_t denotes the total number of antennas at the transmitter.

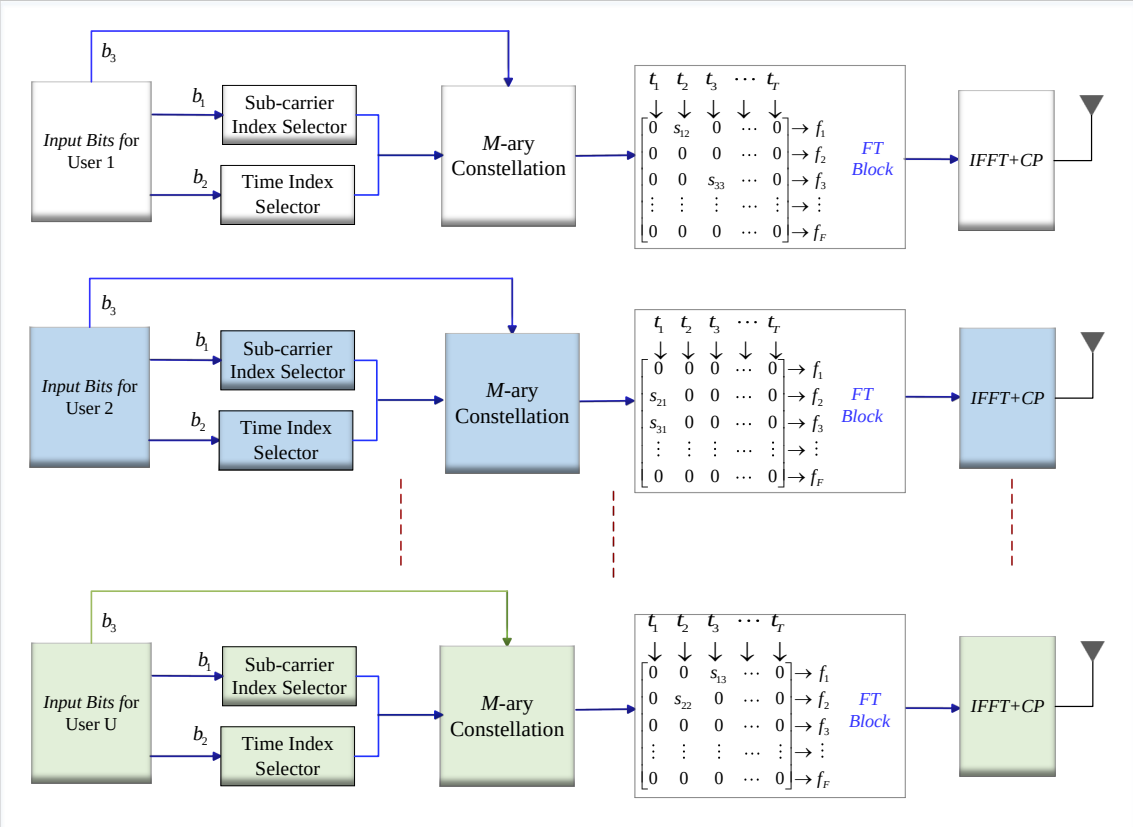


Fig. 2: Transmitter block diagram for FT-IMMA

will be active within the same time slot. If $f = 1$ and $t = 2$, the same frequency resource will be active across two different time slots. Conversely, if $f = 2$ and $t = 2$, two different frequency resources will be active, each in a different time slot.

It is possible to select K resources from the FT grid directly, rather than choosing frequency (f) and time (t) resources separately. While selecting K resources allows for the transmission of more index bits, it can also increase complexity due to the larger number of possible combinations of active resources. As an example, consider an OFDM grid with a total of 16 FT (4×4 grid) resources. If a user selects 3 active resources, 9 bits can be transmitted via index modulation. On the other hand, if the resources are selected separately using $\left\lfloor \log_2 \binom{F}{f} \right\rfloor + \left\lfloor \log_2 \binom{T}{t} \right\rfloor$, for $F = 4$ and $T = 4$, activating $f = 3$ frequency and $t = 3$ time resources can transmit 4 index bits in total, which drastically reduces the search space from 2^9 to 2^4. This smaller search space reduces the complexity of the detection process at the receiver. The detailed complexity analysis is provided in Section II-A. To maintain a balance between SE and computational complexity, the authors recommend using separate selections of frequency and time resources. Additionally, separate selection allows for unequal choices of f and t, providing a means to control both frequency and time selectivity while adjusting the size of the index bits. For instance, an IMT can be designed to activate different subcarriers in the same time slot, or different time slots for the same subcarrier, depending on the desired performance characteristics.

TABLE II: IM Example ($U = 4$, $F = 4$, $T = 4$, $K = 1$).

Users	b_1	Subcarrier Index	b_2	Time Index
U_1	00	f_1	00	t_1
U_2	01	f_2	01	t_2
U_3	10	f_3	10	t_3
U_4	11	f_4	11	t_4

In the proposed scenario, the total number of transmitted bits per grid is specified as $\beta = Ub$. Consider the u^{th} user with six bits: $b_1 = 00$, $b_2 = 01$, and $b_3 = 11$, which select the $f_1 \in F$ subcarrier at the $t_2 \in T$ time slot according to Table II. Thus an $F \times T$ transmission matrix will have a non-zero element indicating the QAM symbol $(1 - j)$ at the $\blacksquare_{12}$ position. In general, if K resources are selected, for instance, b_1 activates $f_1, f_2, \cdots, f_K$ subcarriers and b_2 selects $t_1, t_2, \cdots, t_K$ time slots. Then, the transmission matrix for the u^{th} user is expressed as follows:

$$\mathbf{X}_u = \begin{bmatrix} 0 & s^u_{12} & \cdots & 0 \\ 0 & 0 & 0 & s^u_{2T} \\ s^u_{f_t} & \vdots & \ddots & \vdots \\ s^u_{F1} & 0 & 0 & s^u_{FT} \end{bmatrix}, \tag{2}$$

where $\mathbf{X}_u$ has dimensions of $F \times T$ and constellation symbol s^u_{ft} is mapped over $f^{th} \in F$ row and $t^{th} \in T$ column. The overall transmission power of each user is normalized and distributed equally among the active FT resources. After this mapping,

each user's signal is processed through an inverse fast Fourier transform (IFFT), and a cyclic prefix (CP) of length C_p is appended. The matrix of each user is received by the base station (BS) via a frequency-selective Rayleigh fading channel, denoted by $\mathbf{H}_u = IFFT\{\mathbf{J}_u\}$ with dimensions $T \times F$ where $\mathbf{J}_u \sim \mathcal{CN}(0, 1/L)$ denotes the L-tap fading channel in the time domain. During the transmission of a data block, the channel is regarded as unaltered and $C_p > L$. The channel is assumed to be independent and identically distributed and follows the distribution $\mathcal{CN}(0, 1)$. Following the CP removal and fast Fourier transform operation at the BS, the received signal model is as follows:

$$\mathbf{Y} = \sum_{u=1}^{U} \mathbf{X}_u \odot \mathbf{H}_u + \mathbf{W}_u \,, \tag{3}$$

where $\mathbf{Y}$ represents the received matrix containing the information for all users and $\odot$ denotes the element-wise product. For each user, $\mathbf{W}_u \sim \mathcal{CN}\left(0, \sigma^2\right)$ is the additive white Gaussian noise with zero mean and variance σ^2. The signal received at an FT resource is expressed as y_{ft},

$$y^u_{ft} = \sum_{u=1}^{U} h^u_{ft} s^u_{ft} + w^u_{ft} \,, \tag{4}$$

where h^u_{ft} represents the channel gain at the ft resource between the u^{th} user and BS.

In grant-free MA scenarios, users transmit signals without predetermined or scheduled resource allocations. The objective is to effectively segregate the amalgamated data received at the BS and attribute each signal to its corresponding user. This facilitates successful communication in situations where users transmit without prior coordination or the presence of reference signals. The BS adapts the blind minimum mean square estimation (MMSE) to the observed received signal [43]. Given the MMSE estimate, it applies maximum likelihood detection under the assumption of perfect channel estimation. The goal is to jointly decode the index information and corresponding transmitted symbols. Mathematically, it is expressed as follows:

$$< \mathbf{f}, \mathbf{t}, \hat{\mathbf{X}}_u >= \underset{f,t,\mathbf{X}_i \in \mathcal{X}}{\arg\min} \left\| \mathbf{Y}_u - \mathbf{X}_i \odot \mathbf{H}_u \right\|^2, \forall i = 1, \cdots, 2^{\beta}, \tag{5}$$

where $\mathbf{f} = \{1, 2, \cdots, f\}$ and $\mathbf{t} = \{1, 2, \cdots, t\}$ denote active FT resource indices, $\| \cdot \|^2$ represents the Frobenius norm, and $\mathcal{X}$ contains all possible transmission matrices with $2^{\beta} = 2^{Ub}$.

A. Complexity

The complexity of the proposed scheme is higher than IM-OFDMA and IMMA because it decodes both time and frequency indices to demodulate constellation symbols over the active resources. The complexity is computed by calculating the number of real multiplications required in the detection process. An IMT is required at the receiver to evaluate all possible combinations of active resources and determine the maximum likelihood (ML) detection. This IMT contains all potential activation patterns corresponding to the time and frequency index bits, which are crucial for decoding and achieving the ML solution. The complexity order for the proposed scheme is given as follows:

$$\begin{aligned} Complexity\ Order &= \mathcal{O}\left(2^{index\ bits} \times M^{active\ resources}\right) \\ &= \mathcal{O}\left(2^{(b_1+b_2)} \times M^{K}\right) \end{aligned} \tag{6}$$

According to [12], [14], the SE of IM-OFDMA is given by $b = \left\lfloor \log_2 \binom{F}{f} \right\rfloor + f \times \log_2(M)$. Consequently, its complexity is relatively low compared to the proposed scheme, due to the smaller number of index bits involved. The complexity order for IM-OFDMA is given as $\mathcal{O}\left(2^{b_1} \times M^{f}\right)$. The complexity of IMMA is similarly defined, with the key distinction that time slots are used in place of frequency subcarriers.

In contrast, FT-IMMA introduces the additional complexity of detecting more index bits. While ML detection provides optimal performance, it is computationally expensive due to the exhaustive search over all possible index combinations. Therefore, a sub-optimal solution, such as the CS-MUD algorithm, is suggested for FT-IMMA. CS-MUD is based on OMP, a greedy algorithm widely used for sparse signal estimation.

B. Low Complexity Detection Using OMP

OMP is a representative method in the greedy algorithm family [44]. It is well suited to estimate sparse signals. Typically, the number of active users is much smaller than the number of all potential users in the system. Additionally, each user's transmission signal has zeros that IM has inherited; as a result, the MUD problem in the proposed scenario is naturally a sparse signal recovery problem.

The OMP algorithm takes as input the received signal, the channel matrix, and the sparsity level S. The sparsity level corresponds to the number of active users, and the BS can estimate it by analyzing the received pilot signals from the users who are transmitting in a given time frame.

In standard OMP, a measurement matrix is typically used to transform the original signal into a lower-dimensional representation. However, in a communication system, the channel matrix replaces the measurement matrix, as it models the transformation undergone by the sparse transmitted signal during communication. Although the channel matrix does not perform dimensionality reduction, it still provides the necessary information about how the transmitted signal propagates through the channel.

The OMP algorithm begins by initializing the residual signal to the received signal. The residual represents the difference between the received signal and the current estimate of the sparse signal. Initially, the algorithm starts with an empty support set (Γ_u), which will track the indices of the active resources (subcarriers and time slots) used by the active users.

OMP operates iteratively, where in each iteration, it identifies the column of the channel matrix $\mathbf{H}_u$ that is most correlated with the current residual $\mathbf{R}$. This step helps identify the indices of the non-zero elements (active resources). The identified indices are added to the support set Γ_u, and the residual is updated by subtracting the signal projection onto the selected columns of $\mathbf{H}_u$.

The algorithm continues iterating until either all active users are detected or the residual becomes sufficiently small,

Algorithm 1 OMP Detection for FT-IMMA

Input: Received signal $\mathbf{Y}$, Channel matrix $\mathbf{H}_u$, Sparsity level S
Output: Estimated sparse signal $\hat{\mathbf{X}}_u$
Initialization:
    Set residual $\mathbf{R} = \mathbf{Y}$
    Set support set $\Gamma_u = \emptyset$ ▷ Support set to track selected indices
    Set iteration counter $i = 1$
while $i \leq S$ **do**
    Find index of column in $\mathbf{H}_u$ most correlated with residual:

$$j = \arg\max_k \left| \mathbf{H}_u(:,k)^H \mathbf{R} \right|$$

    Update support set: $\Gamma_u = \Gamma_u \cup \{j\}$
    Solve least squares problem to estimate sparse signal:

$$\hat{\mathbf{X}}_{\Gamma_u} = \mathbf{H}^{\dagger}_{u,\Gamma_u} \mathbf{Y}$$

    Update residual:

$$\mathbf{R} = \mathbf{Y} - \mathbf{H}_{u,\Gamma_u} \hat{\mathbf{X}}_{\Gamma_u}$$

    Increment iteration counter: $i = i + 1$
end while
Return estimated sparse signal $\hat{\mathbf{X}}_u$

indicating that the estimated signal has converged. Finally, the reconstructed signal $\hat{\mathbf{X}}_u$ is computed by solving a least-squares problem, which reveals the active part of the FT grid used by the active users. This process is summarized mathematically in Algorithm 1.

OMP iteratively selects the best-matching columns from the sensing matrix to approximate the signal, significantly reducing the need to evaluate all possible combinations. The complexity of OMP in the FT-IMMA detection process can be expressed as:

$$\text{Complexity Order} = \mathcal{O}\left(S \times \mathcal{S}_{\text{sensing}} \times \text{Iterations}\right) \quad (7)$$

Here, S denotes the sparsity level, $\mathcal{S}_{\text{sensing}}$ specifies the sensing parameter (the number of measurements or rows in the sensing matrix, typically a fraction of the original signal length), and the number of iterations generally matches S. This complexity is often significantly lower than that of ML, especially when $\mathcal{S}_{\text{sensing}}$ is considerably reduced relative to the original signal dimensions. However, the accuracy of OMP depends on selecting an optimal $\mathcal{S}_{\text{sensing}}$ value and may suffer in very low SNR conditions compared to ML. In Algorithm 1, $\mathcal{S}_{\text{sensing}}$ is determined by the number of rows in $\mathbf{Y}$ and $\mathbf{H}_u$, establishing the level of dimensionality reduction. This parameter is pre-selected before detection to strike a balance between computational efficiency and detection accuracy.

III. Key Performance Indicators

A. Collision Probability

The probability that a user will occupy a certain resource block is $\rho = \frac{K}{F \times T}$ if $\begin{pmatrix} F \\ f \end{pmatrix}$ and $\begin{pmatrix} T \\ t \end{pmatrix}$ have values equal to a power of two when all possible index combinations will be used. In contrast, ρ is slightly impacted because some index combinations are dropped based on the IM notion. Eliminating certain combinations is an NP-hard index selection problem involving combinatorial optimization. However, in this article, the authors stick to ρ by assuming the number of index combinations equal to a power of two. The resource allocation depends on the incoming bit stream of users, multiple users may access the same resource simultaneously. The probability of u users out of U having identical bits and attempting to occupy the same resource is computed as follows (the number of users in the same FT resource follows a Binomial distribution)

$$P_u(u \geq 2) = \sum_{u=2}^{U} \begin{pmatrix} U \\ u \end{pmatrix} (\rho)^u (1-\rho)^{U-u}. \quad (8)$$

This expression is further simplified as follows:

$$P_u = 1 - \left(1 - \frac{K}{F \times T}\right)^U \left(1 + \frac{U \times K}{(F \times T) - K}\right). \quad (9)$$

The likelihood of collision depends on U, F, T, and K. For fixed values of F and T, increasing U or K will result in more collisions. Conversely, for fixed U and K, increasing F and T will improve the performance by lowering the number of collisions.

B. Energy Efficiency

The EE of the u^{th} user of FT-IMMA compared to IM-OFDMA and OFDMA can be calculated as follows:

$$\epsilon = \frac{K \log_2(M) - F \log_2(\hat{M})}{F \log_2(\hat{M})} = \left\{ \frac{K \log_2(M)}{F \log_2(\hat{M})} - 1 \right\} \quad (10)$$

where K is the number of active FT resources in FT-IMMA and F represents the total subcarriers in OFDMA. To transmit 12 bits, FT-IMMA requires $F = 4$, $T = 4$, $f = t = K = 2$, and $M = 16$. On the other hand, OFDMA needs $F = 4$ and $\hat{M} = 8$ to deliver the same amount of bits. In this case, the FT-IMMA can save 33.34% energy compared to OFDMA. Similarly, to compare IM-OFDMA with the proposed scenario, F in (10) can be replaced with f, i.e., the active number of subcarriers in IM-OFDMA. To achieve the same SE, it employs $f = 2$ and $\hat{M} = 32$ which requires 20% more energy compared to FT-IMMA. The higher EE of FT-IMMA is because of additional bits that are being transmitted via two-dimensional indices.

C. Achieveable Data Rate

This subsection dissects the intricacies of data rate components, denoted as C_1, C_2, and C_3, each contributing to the overall information conveyed through the activation of time and frequency indices, coupled with constellation symbols. The achievable data rate of FT-IMMA can be derived by maximizing the mutual information (MI) I over the probability

distribution functions (PDF) of indices and constellation symbols. Therefore, the MI for a user in the system is expressed as follows:

$$C = \underbrace{\mathrm{E}\left[\max_{p_s}\left\{I\left(\mathbf{S}^u_{ft};\mathbf{Y}^u_{ft} \mid \mathbf{f},\mathbf{t},\mathbf{H}^u_{ft}\right)\right\}\right]}_{C_1} + \underbrace{I\left(\mathbf{f};\mathbf{Y}^u_{ft} \mid \mathbf{H}^u_{ft}\right)}_{C_2} + \underbrace{I\left(\mathbf{t};\mathbf{Y}^u_{ft} \mid \mathbf{H}^u_{ft}\right)}_{C_3} \tag{11}$$

where $\mathbf{S}^u_{ft}$, $\mathbf{Y}^u_{ft}$, and $\mathbf{H}^u_{ft}$ represent the modulated signal over the active $\mathbf{f}$, $\mathbf{t}$ resources, the received signal, and the channel matrix of u^{th} user, respectively. As the data rate calculation for each user follows a similar method, for brevity, the superscript u and subscript ft will be omitted in the subsequent descriptions. Furthermore, p_s denotes the PDF of the constellation symbol s that maximizes C_1. It is assumed that this choice of p_s also maximizes both C_2 and C_3 which represent the portion of information transmitted through frequency and time index activations. C_1 that maximizes the MI between constellation symbols and received signal is expanded as below

$$\begin{aligned} C_1 =& \mathrm{E}\left[\max_{p_s}\left\{I\left(\mathbf{S};\mathbf{Y} \mid \mathbf{f},\mathbf{t},\mathbf{H}\right)\right\}\right] \\ =& \mathrm{E}\left[\max_{p_s}\left\{\mathcal{H}\left(\mathbf{Y}\right)\right\} - \mathcal{H}\left(\mathbf{Y} \mid \mathbf{S},\mathbf{H}\right)\right] \end{aligned} \tag{12}$$

where $\mathcal{H}$ indicates the entropy. By maximizing the entropies, C_1 is analytically solvable for Gaussian random variables with perfect channel state information (CSI). The first part of (12) is expanded as follows:

$$\mathcal{H}(\mathbf{Y}) = -\int_{-\infty}^{\infty} p_{\mathbf{Y}}(\mathbf{Y}) \log_2 p_{\mathbf{Y}}(\mathbf{Y})\, d\mathbf{Y} = \log_2(\pi e \sigma_y^2) \tag{13}$$

where the PDF of $\mathbf{Y}$, a complex Gaussian random variable, is expressed as $p_{\mathbf{Y}}(\mathbf{Y}) = \frac{1}{\sqrt{\pi\sigma_y^2}} e^{-(x-\mu)^2/\sigma_y^2}$ with mean μ, variance σ_y^2. It follows the distribution $\mathcal{CN}\left(\mu, \sigma_y^2\right)$ and $\sigma_y^2 = E\left\{\|\mathbf{Y}\|^2\right\}$. Similarly, the second part, $\mathcal{H}\left(\mathbf{Y} \mid \mathbf{S},\mathbf{H}\right)$ in (12), simplifies to $\mathcal{H}\left(\mathbf{W} \mid \mathbf{S}\right)$ because of the known CSI at the receiver, where $\mathbf{S}$ and $\mathbf{W}$ are independent. It can be solved as follows:

$$\mathcal{H}\left(\mathbf{Y} \mid \mathbf{S},\mathbf{H}\right) = \mathcal{H}\left(\mathbf{W}\right) = \log_2\left(\pi e \sigma_w^2\right) \tag{14}$$

where σ_w^2 represents noise power. Now, substituting (13) and (14) in (12) gives the following expression:

$$C_1 = \sum_{i=1}^{F} \log_2\left(1 + \rho\left|h_i\right|^2\right) \tag{15}$$

The conditional MI C_2 in (11) can be defined in terms of joint and marginal probabilities as given below

$$C_2 = \sum_{\mathbf{f}} \int_{\mathbf{Y}} p_{(\mathbf{f},\mathbf{Y}|\mathbf{H})}(\mathbf{f},\mathbf{Y} \mid \mathbf{H}) \log_2\left(\frac{p_{(\mathbf{f},\mathbf{Y}|\mathbf{H})}(\mathbf{f},\mathbf{Y} \mid \mathbf{H})}{p_{\mathbf{f}|\mathbf{H}}(\mathbf{f} \mid \mathbf{H}) p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y} \mid \mathbf{H})}\right) d\mathbf{Y} \tag{16}$$

where,

$$p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y} \mid \mathbf{f},\mathbf{H}) = \frac{p_{(\mathbf{f},\mathbf{Y}|H)}(\mathbf{f},\mathbf{Y} \mid H)}{p_{\mathbf{f}|H}(\mathbf{f} \mid H)} \tag{17}$$

By substituting (17) in (16), it takes the following form:

$$\begin{aligned} C_2 &= \sum_{\mathbf{f}} \int_{\mathbf{Y}} p_{\mathbf{f}}(\mathbf{f}) p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y} \mid \mathbf{f},\mathbf{H}) \log_2\left(\frac{p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y} \mid \mathbf{f},\mathbf{H})}{p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y} \mid \mathbf{H})}\right) d\mathbf{Y} \\ C_2 &= \mathrm{E}_{\mathbf{f}}\left\{\int_{\mathbf{Y}} p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y} \mid \mathbf{f},\mathbf{H}) \log_2\left(\frac{p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y} \mid \mathbf{f},\mathbf{H})}{p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y} \mid \mathbf{H})}\right) d\mathbf{Y}\right\} \end{aligned} \tag{18}$$

The expectation operator $\mathrm{E}_{\mathbf{f}}$ indicates the average over all possible values of $\mathbf{f}$ based on the probability distribution $p_{\mathbf{f}}(\mathbf{f})$.

$$\begin{aligned} C_2 = & \underbrace{\mathrm{E}_{\mathbf{f}}\left\{\int_{Y} p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y}|\mathbf{f},\mathbf{H}) \log_2\left(p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y}|\mathbf{f},\mathbf{H})\right)\, d\mathbf{Y}\right\}}_{\mathcal{H}(\mathbf{Y}|\mathbf{S},\mathbf{H})} \\ & -\mathrm{E}_{\mathbf{f}}\left\{\int_{\mathbf{Y}} p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y}|\mathbf{f},\mathbf{H}) \log_2\left(p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y}|\mathbf{H})\right)\, d\mathbf{Y}\right\} \end{aligned} \tag{19}$$

where

$$\begin{aligned} p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y}|\mathbf{f}=f_j,\mathbf{H}) = & \frac{1}{(\pi\sigma_w^2)^F} \times \prod_{i\in f_j} \frac{1}{1+\rho|H_i|^2 F/(F-f)} \\ & \times e^{-\frac{1}{\sigma_w^2}\frac{|Y_i|^2}{1+\rho|H_i|^2F/(F-f)}} \times \prod_{i\in\overline{f}_j} e^{-\frac{|Y_i|^2}{\sigma_w^2}} \end{aligned} \tag{20}$$

and

$$\begin{aligned} p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y} \mid \mathbf{H}) = & \sum_{j=1}^{C(F,f)} \Pr\left(\mathbf{f}=f_j\right) p_{\mathbf{Y}|\mathbf{f},\mathbf{H}}(\mathbf{Y}|\mathbf{f}=f_j,\mathbf{H}) \\ = & \frac{1}{C(F,f)(\pi\sigma_w^2)^F} \sum_{j=1}^{C(F,f)} \prod_{i\in f_j} \frac{1}{1+\rho|H_i|^2F/(F-f)} \\ & \times e^{-\frac{1}{\sigma_w^2}\frac{|Y_i|^2}{1+\rho|H_i|^2F/(F-f)}} \times \prod_{i\in\overline{f}_j} e^{-\frac{|Y_i|^2}{\sigma_w^2}}. \end{aligned} \tag{21}$$

Furthermore, $\mathcal{H}\left(\mathbf{Y} \mid \mathbf{S},\mathbf{H}\right) = \log_2\left(\pi e \sigma_w^2\right)$ from (14). The closed-form expression for (19) is not available and therefore can be written as below:

$$\begin{aligned} C_2 = & -\log_2\left(\pi e \sigma_w^2\right) - \mathrm{E}_{\mathbf{f}}\left\{\log_2\left(1+\rho\left|h_i\right|^2\right)\right\} \\ & - \mathrm{E}_{\mathbf{f}}\left\{\mathrm{E}_{\mathbf{Y}|\mathbf{f},\mathbf{H}}\left\{\log_2\left(p_{\mathbf{Y}|\mathbf{H}}(\mathbf{Y} \mid \mathbf{H})\right)\right\}\right\}. \end{aligned} \tag{22}$$

The solution for C_3 follows a similar approach as C_2 in (22). The maximization of (11) depends on the distribution of $\mathbf{f}$ and $\mathbf{t}$ indices as well as the signal constellation.

D. BER Bound

Suppose $\mathbf{X}_c$ is transmitted and incorrectly detected as $\mathbf{X}_d$. In that case, the conditional pairwise error probability (PEP) is expressed by utilizing the Q-function

$$\Pr\left(\mathbf{X}_c \rightarrow \mathbf{X}_d \mid \mathbf{H}\right) = Q\left(\sqrt{\frac{\eta}{2N_{0,F}}}\right), \tag{23}$$

where $\eta = \|\mathbf{H}\mathbf{\Lambda}\|_F^2$ and $\mathbf{\Lambda} = \mathbf{X}_c - \mathbf{X}_d$. Thus, $\eta = \sum_{n=1}^{N} \| \sum_{u=1}^{U} h_{nu}\Lambda_{un}\|^2$ where Λ_{un}, when N is assumed as total FT resources, and an element of $\mathbf{\Lambda}$ is mapped over the n^{th} resource block for the u^{th} user.

η is assumed to follow an Erlang distribution. This distribution was initially developed to assess the number of simultaneous phone calls [45]. This is suitable for the proposed scenario, in which multiple users may attempt to simultaneously access the same resource. Let R be a nonzero block and $R \leq N$ where R is an exponential random variable (RV) with a mean $\mu_n = \sum_{u=1}^{U} \|\Lambda_{un}\|^2$. Out of R, $S \leq R$ distinct means represented by $\mu_1, \mu_2, \cdots, \mu_S$ exist. The number of RVs with similar means μ_s, $1 \leq s \leq S$, is m_s, which implies that $m_1 + m_2 + \cdots + m_S = R$. The PDF of η is expressed as [45]

$$f_\eta(\eta) = \sum_{s=1}^{S} \lambda_s^{m_s} e^{-\eta\lambda_s} \sum_{w=1}^{m_s} A\frac{(-1)^{m_s-w}}{(w-1)!}\eta^{w-1}. \tag{24}$$

The number m_s reflects how many users have similar behavior in terms of accessing a particular resource, while $\lambda_s = 1/\mu_s$ denotes the arrival rate. $\lambda_s^{m_s}$ indicates the combined effect of users that show similar behavior. In uplink random access, multiple users can attempt to access the same resources. $e^{-\lambda_s\eta}$ represents the attempts that could not result in success due to collisions. w denotes the number of users that contribute to the interference power. The term A is expressed as follows:

$$A = \sum_{\substack{b_1+\cdots+b_S=m_s-w \\ b_s=0}} \prod_{\substack{q=1 \\ q\neq s}}^{S} \binom{m_q + b_q - 1}{b_q} \frac{\lambda_q^{m_q}}{\left(\lambda_q - \lambda_s\right)^{m_q+b_q}}. \tag{25}$$

The factor A represents the combinatorial ways in which users can interfere with each other while accessing the same resources. When $S = 1$, $A = 0$ if $m_s - w > 0$; otherwise, $A = 1$. Here $A = 0$ means that there are remaining users that cannot be distributed, indicating unresolved collisions. Whereas $A = 1$ accounts for a configuration with no unresolved collisions. This simplifies the PEP calculations because the unresolved collisions are excluded. Therefore, the PEP is obtained by integrating (23) over $f_\eta(\eta)$:

$$\mathrm{P}\left(\mathbf{X}_c \rightarrow \mathbf{X}_d\right) = \int_{-\infty}^{+\infty} \Pr\left(\mathbf{X}_c \rightarrow \mathbf{X}_d \mid \mathbf{H}\right) f_\eta(\eta) d\eta. \tag{26}$$

By utilizing Craig's formula $Q\left(\sqrt{\frac{\eta}{2N_{0,F}}}\right) = \frac{1}{\pi}\int_0^{\pi} \exp\left(-\frac{\eta}{4N_{0,F}\sin^2\theta}\right) \mathrm{d}\theta$ to circumvent the approximation of the Q-function, (26) can be represented as follows:

$$\mathrm{P}\left(\mathbf{X}_c \rightarrow \mathbf{X}_d\right) = \frac{1}{\pi}\sum_{s=1}^{S} \lambda_s^{m_s} \sum_{w=1}^{m_s} A\frac{(-1)^{m_s-w}}{(w-1)!} \times \int_0^{\pi/2}\int_0^{\infty} \exp\left(-\frac{\eta}{4N_{0,F}\sin^2\theta}\right) e^{-\eta\lambda_s}\eta^{w-1} \cdot d\eta d\theta. \tag{27}$$

According to [46], the inner integral over η is approximated as follows

$$\mathrm{P}\left(\mathbf{X}_c \rightarrow \mathbf{X}_d\right) = \frac{1}{\pi}\sum_{s=1}^{S} \lambda_s^{m_s} \sum_{w=1}^{m_s} A(-1)^{m_s-w}\left(\frac{1}{\lambda_s}\right)^w \times \int_0^{\frac{\pi}{2}} \left(\frac{\sin^2\theta}{\sin^2\theta + \frac{1}{4N_{0,F}\lambda_s}}\right)^w d\theta. \tag{28}$$

The integral over θ can be solved as below [47],

$$\mathrm{P}\left(\mathbf{X}_c \rightarrow \mathbf{X}_d\right) = \sum_{s=1}^{S} \lambda_s^{m_s} \sum_{w=1}^{m_s} A(-1)^{m_s-w}\left(\frac{\chi}{\lambda_s}\right)^w \sum_{i=0}^{w-1} \binom{w-1+i}{i} (1-\chi)^i, \tag{29}$$

where $\chi = \frac{1}{2}\left(1 - \frac{1}{\sqrt{(1+4N_{0,F}\lambda_s)}}\right)$. The average BER is expressed as follows:

$$\mathrm{BER} = \frac{1}{b2^b}\sum_{c=1}^{2^b}\sum_{d=1}^{2^b} \mathrm{P}\left(\mathbf{X}_c - \mathbf{X}_d\right)\gamma\left(\mathbf{X}_c, \mathbf{X}_d\right), \tag{30}$$

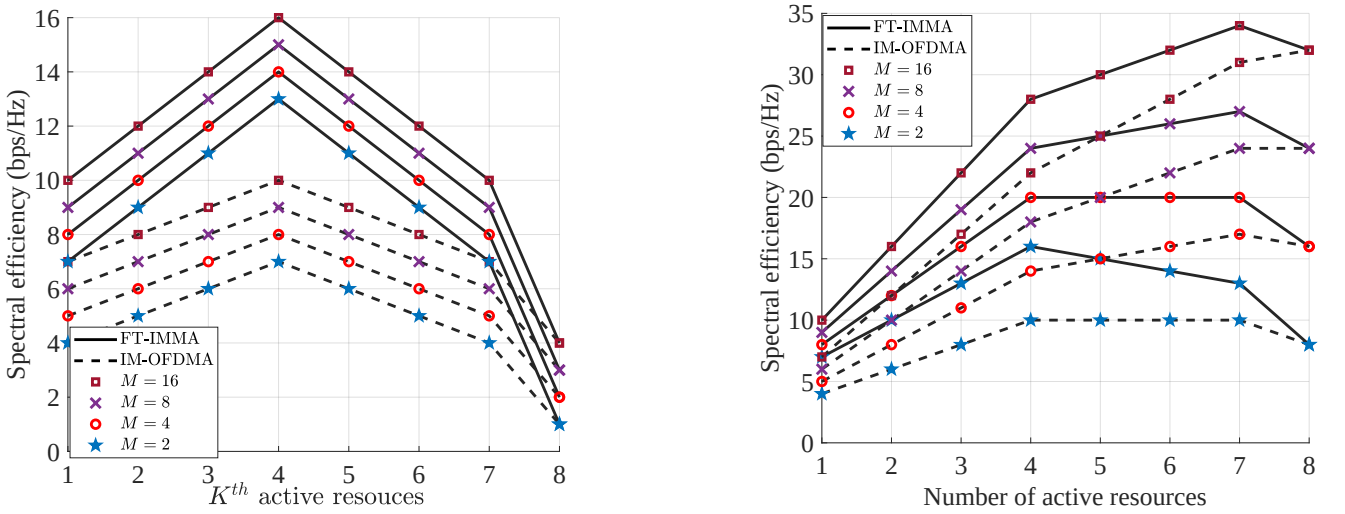
where $\gamma\left(\mathbf{X}_c, \mathbf{X}_d\right)$ denotes the hamming distance between $\mathbf{X}_c$ and $\mathbf{X}_d$.

IV. Results and Discussion

This section discusses the results of FT-IMMA in terms of SE, collision probability, and BER. The performance is evaluated against traditional benchmarks including OFDMA, IM-OFDMA, and IMMA. This analysis assumes that the transmitter and receiver are perfectly synced and channel state information is known at the receiver. Additionally, while describing each outcome, the pertinent simulation settings are provided.

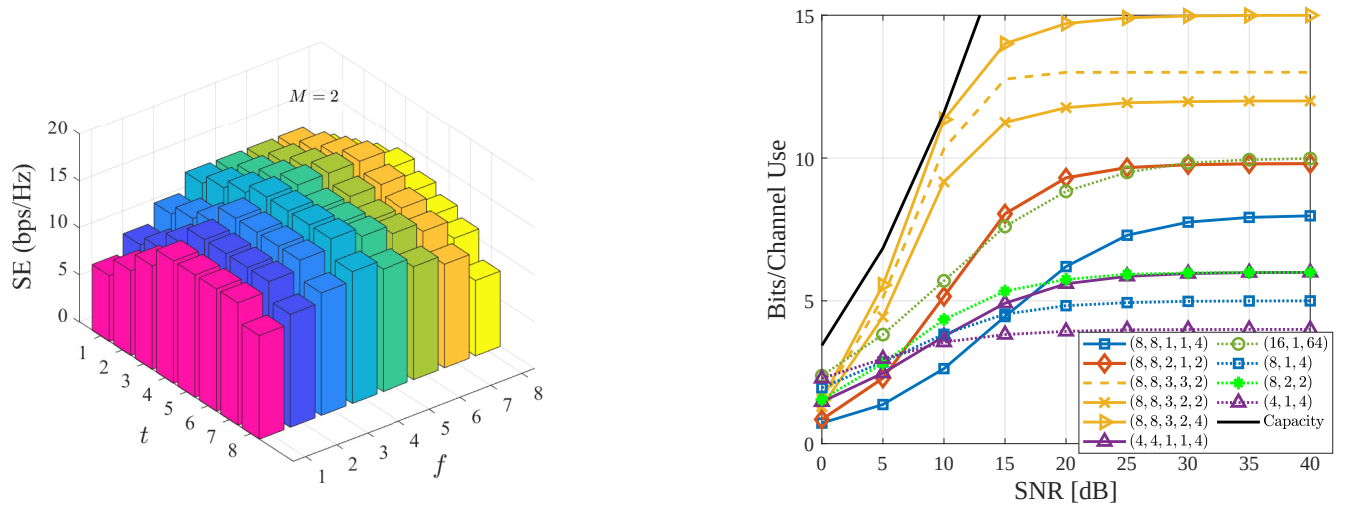
1) Spectral Efficiency: Fig. 3a presents a comparison of SE per user between IM-OFDMA and FT-IMMA under the condition where only a single resource (denoted as the K^{th} resource) is active for a user, as described in [14] for IM-OFDMA. This setup represents a specific case of FT-IMMA known as single-symbol FT-IMMA (SFT-IMMA). In SFT-IMMA, a single resource K^{th} is used to transmit the constellation symbol, while the position of the K^{th} resource determines the number of index bits. In the case of FT-IMMA, the maximum achievable number of index bits is given by $\left\lfloor \log_2\binom{F}{F/2} \right\rfloor + \left\lfloor \log_2\binom{T}{T/2} \right\rfloor$, which occurs when a user selects the $\frac{F}{2}$th frequency and $\frac{T}{2}$th time slot, thereby transmitting the maximum number of index bits. The SE grows initially, peaks in the middle, and subsequently declines. It is minimum when $\binom{8}{8} = 1$, which is the standard OFDMA scenario. An investigation of SE for various modulation orders revealed that, at $M = 2$, FT-IMMA could achieve higher SE compared with IM-OFDMA and OFDMA at $M = 16$.

The SE per user for multiple active resources carrying different constellation symbols is shown in Fig. 3b. A maximum of 8 resources can be assigned to a user, which represents the typical OFDMA scenario. FT-IMMA has a higher SE



(a) (b)

Fig. 3: (a) SE per user for multiple active resources ($f = t$), transmitting the same constellation symbol (a scenario of SFT-IMMA). (b) SE per user for multiple active resources ($f = t$), transmitting different constellation symbols.



(a) (b)

Fig. 4: (a) SE per user for $f \neq t$ and $M = 2$. (b) Mutual Information Comparison (The parameters listed as $\{F, T, f, t, M\}$ and $\{F, f, M\}$ represent FT-IMMA (1st column) and IM-OFDMA (2nd column), respectively).

compared with other techniques in each scenario due to implicitly embedded information via two-dimensional resource index selection. Notably, IMMA provides the same SE as IM-OFDMA; however, time slots are considered rather than frequency slots.

The MI performance of FT-IMMA compared to IM-OFDMA over a Rayleigh fading channel is depicted in Fig. 4b. The figure illustrates the maximum number of transmitted *bits per channel use* for various configurations, with parameters listed as F, T, f, t, M for FT-IMMA and F, f, M for IM-OFDMA. Additionally, the plot includes the Shannon capacity curve, representing the theoretical upper bound, achieved when the system operates with zero mean and unit variance. It incorporates the singular value decomposition method to accurately capture optimal capacity values. This sets the fundamental limit for MI at a given SNR.

For the proposed FT-IMMA scheme, it is evident that MI approaches the channel capacity as the number of active FT resources increases, or as a higher constellation size is employed. The figure shows that the gap between MI and capacity diminishes with more active resources. For instance, when three FT resources are active, the MI is significantly higher compared to the case where only one FT resource is active in the 8×8 FT grid.

The more active subcarriers or time slots that FT-IMMA utilizes, the higher the MI. This trend is demonstrated in comparisons such as $(8, 8, 3, 2, 2)$ (3 active subcarriers, 2 active time slots, BPSK) versus $(8, 8, 2, 1, 2)$ (2 active subcarriers, 1 active time slot, BPSK). In scenarios like $(8, 8, 2, 1, 2)$, the MI performance of FT-IMMA is comparable to that of IM-

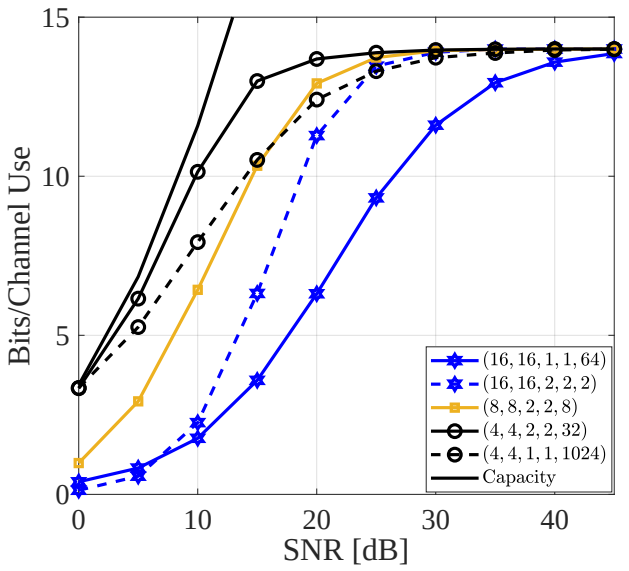


Fig. 5: Comparison of various cases to achieve the same mutual information.

OFDMA $(16, 1, 64)$, but the latter requires a greater number of total subcarriers and a higher modulation order to achieve similar MI levels.

The difference between configurations such as $(8, 8, 3, 3, 2)$ and $(8, 8, 3, 2, 2)$ highlights the additional information attached to the indices when more f or t resources are activated, despite using the same amount of overall active FT resources (as described in Section II, Equation (1)). This increased use of index resources directly translates into higher MI, showcasing the efficiency of FT-IMMA in leveraging both frequency and time domains for enhanced SE.

Furthermore, FT-IMMA can transmit 6, 8, and 15 bits with parameters $\{4, 4, 1, 1, 4\}$, $\{8, 8, 1, 1, 4\}$, and $\{8, 8, 3, 2, 4\}$, respectively. The difference between $\{8, 8, 3, 3, 2\}$ and $\{8, 8, 3, 2, 2\}$ is more information attached to indices despite the same utilized resources. The proposed technique with $\{4, 4, 1, 1, 4\}$, $\{8, 8, 1, 1, 4\}$ can transmit 33.3%, 37.5%, and 28.6% more bits than IM-OFDMA with parameters $\{4, 1, 4\}$, $\{8, 1, 4\}$, and $\{16, 1, 64\}$, respectively.

Fig. 5 shows a scenario where different configurations of FT-IMMA are used to achieve the same MI, assuming no collision occurs between active resources. The key observation is that increasing the number of active FT resources while maintaining the same overall MI can reduce the modulation order, which leads to performance improvements in low-SNR regions.

For example, in the case of a 16×16 FT grid, moving from a single active FT resource to two active FT resources allows the modulation size to decrease from $M = 64$ to BPSK. This reduction in modulation size mitigates the negative impact of noise, enhancing the performance in lower SNR scenarios. A similar trend is observed in the 4×4 FT grid, where the system with two active FT resources demonstrates better robustness in low SNR compared to the system with a higher modulation order but a single active FT resource.

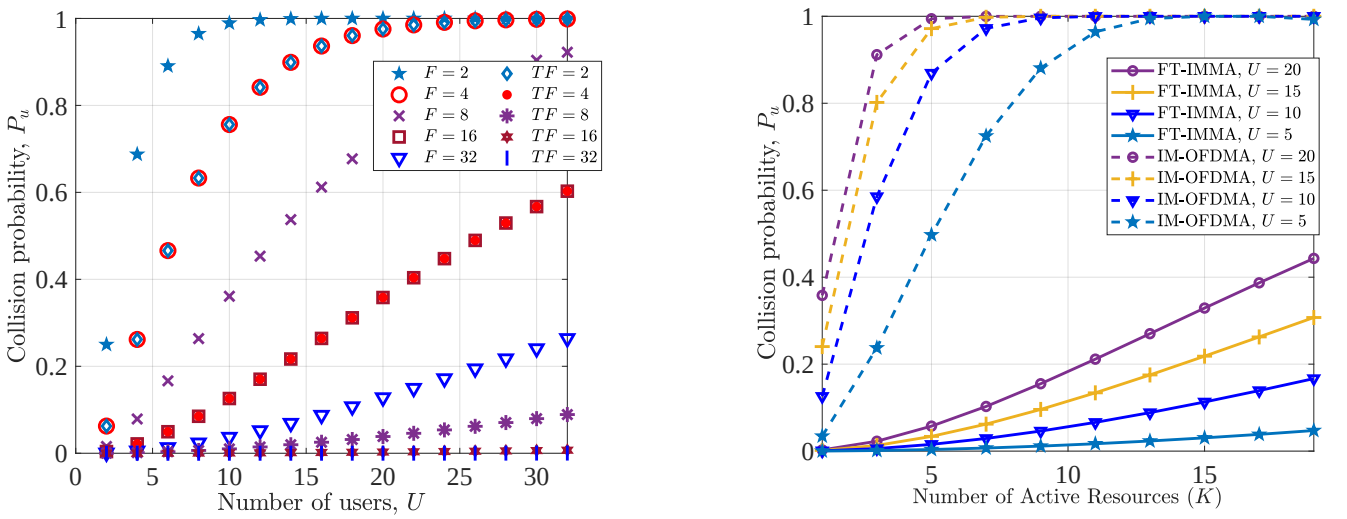
2) Collision Probability: A comparison of the probability of collision among simultaneous users of FT-IMMA and IM-OFDMA is shown in Fig. 6a. The collision rate increases with the number of users in the system. This decreases when the number of resources exceeds the number of users. This worsens in the opposite scenario, in which more users occupy fewer available resources. In the case of an OFDMA scenario in which 4 users can be served by 16 resources as shown in Fig. 1, each user is assigned four 4 resources, each of which can carry the symbol with $M = 4$. Each user can send 8 bits in this manner. On the other hand, each user can transmit 10 bits in IM-OFDMA with the same parameters; however, only 2 of the 16 subcarriers are active per user. For 4 users, 8 subcarriers are utilized, and the remaining 8 are deactivated. More active subcarriers are required to further improve SE, which exacerbates collisions. In contrast, only one active subcarrier can transmit 10 bits in the FT-IMMA. Compared with the case of IM-OFDMA, the probability of a collision decreases because 2D index selection offers an extra degree of freedom in FT-IMMA.

Fig. 6b illustrates the collision probability as a function of the number of active resources for varying user counts. As the number of active resources or users increases, the probability of collisions also rises. In FT-IMMA, this increase follows a linear trend, suggesting a more controlled growth in collisions. In contrast, IM-OFDMA exhibits an exponential rise in collision probability, quickly approaching the maximum collision threshold as the number of active resources increases. This demonstrates that FT-IMMA is more resilient to collisions in scenarios with a higher density of active users compared to IM-OFDMA.

3) Bit Error Rate Analysis: The BER performance of FT-IMMA is generally better than that of IM-OFDMA because FT-IMMA requires a smaller constellation size to achieve the same SE. Additionally, FT-IMMA outperforms classical OFDMA in terms of BER since it needs fewer active resources than OFDMA to reach the same SE. The BER results presented are averaged over multiple iterations, with users transmitting random bits.

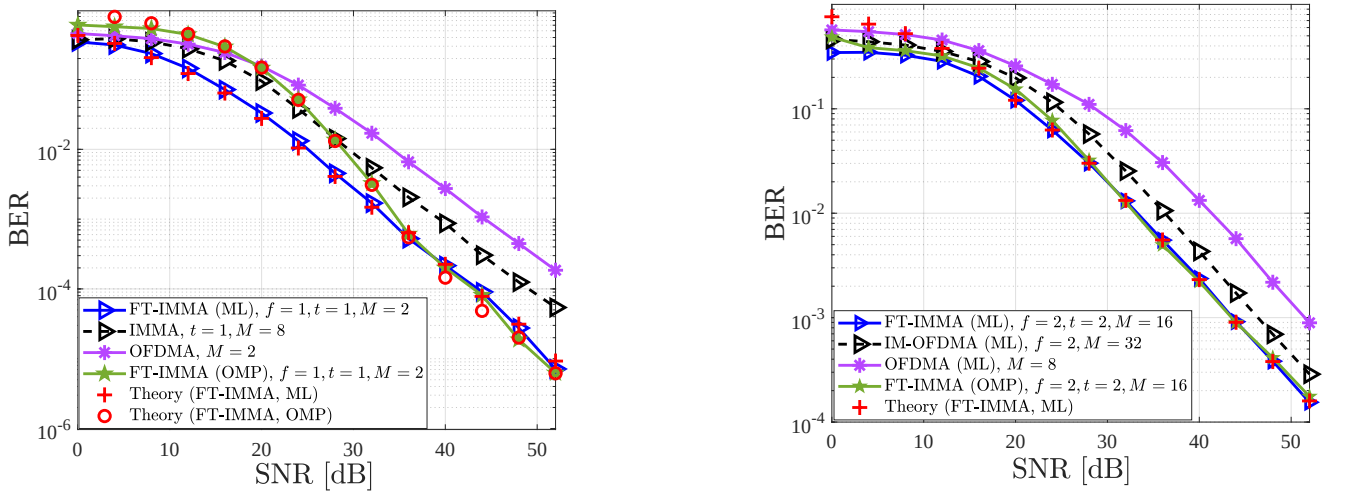
The BER of the proposed scheme is compared with its classical counterparts at the same SE by adjusting various parameters such as f, t, and M. FT-IMMA outperformed conventional IMMA and OFDMA in terms of the BER while transmitting 5 bits in each case with $F = 4$ and $T = 4$, as shown in Fig. 7a. A single active resource of the FT-IMMA requires $M = 2$ to achieve an SE similar to that of the IMMA with $M = 8$. Therefore, it gains approximately 6 dB to achieve a BER of $\sim 5 \times 10^{-05}$.

Due to the high computational complexity of ML detection, which exhaustively searches all possible combinations in the predefined IMT (see Section II-A), OMP is utilized to reduce the MUD complexity by leveraging the inherent sparsity of FT-IMMA. The performance of OMP relies on factors such as sparsity level, measurement matrix, and threshold settings, which collectively help to accurately reconstruct the transmitted signal with minimal information loss. As shown in Fig. 7a, ML outperforms OMP in low-SNR regions, where



(a) (b)

Fig. 6: (a) Collision probability of FT-IMMA and IM-OFDMA (The legends F and FT represent IM-OFDMA and FT-IMMA, respectively.). (b) Collision probability for varying numbers of users with increasing active resources from an $F \times T$ grid.

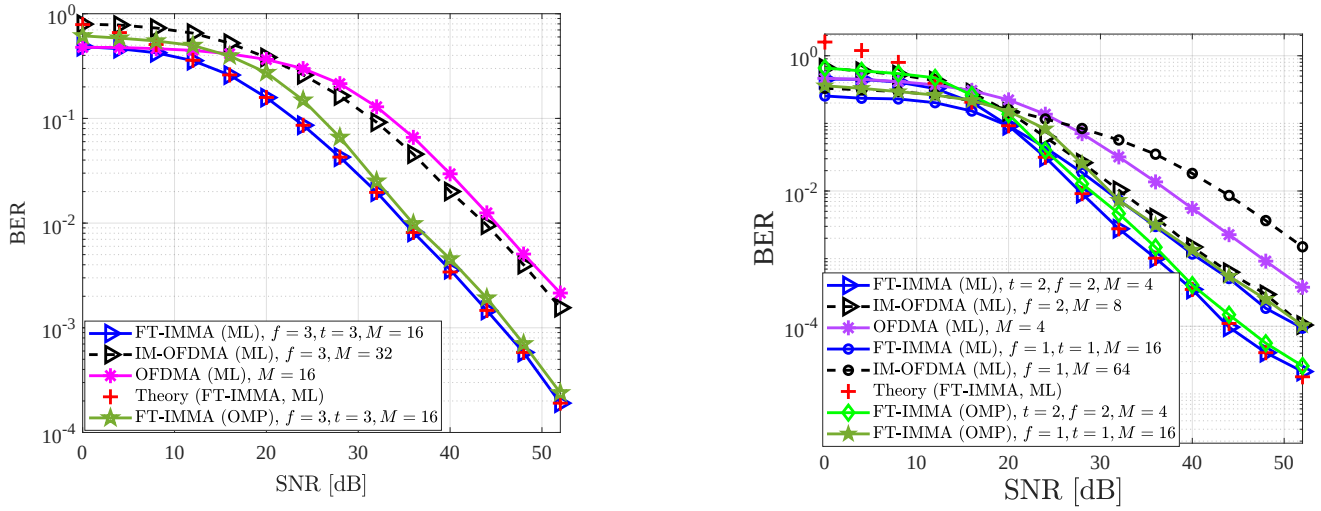


(a) (b)

Fig. 7: (a) BER performance at 1.25 bps/Hz. (b) BER performance at 3 bps/Hz.

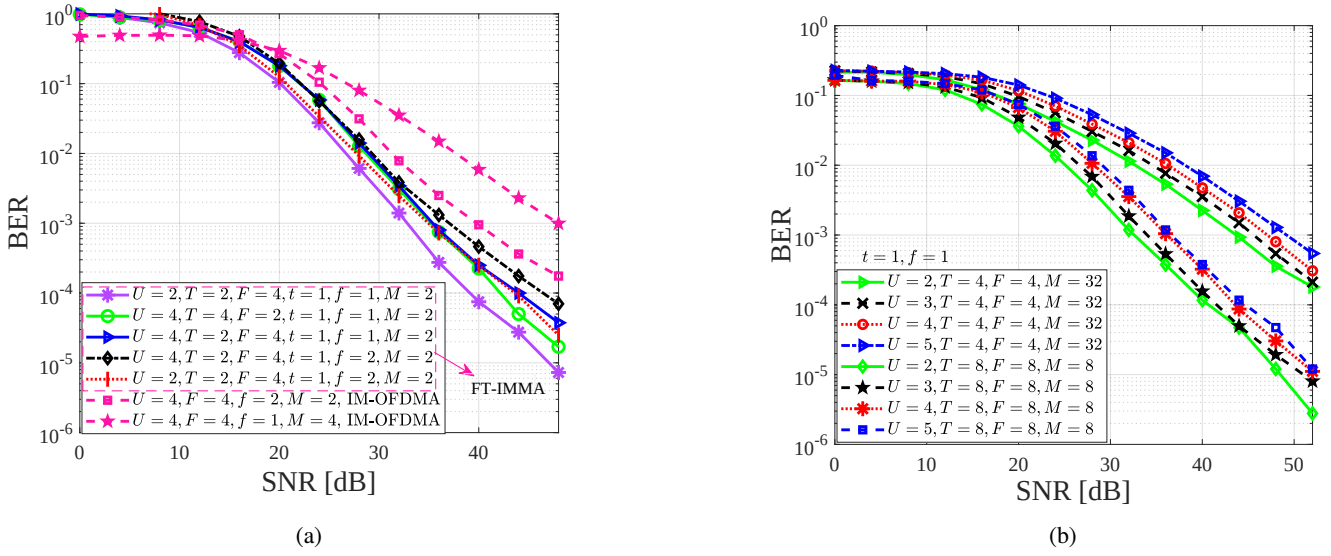
the higher noise levels make it challenging for OMP to detect the sparse signal accurately. However, as SNR increases, the performance of OMP approaches that of ML, as reduced channel noise allows OMP to recover the transmitted symbols more precisely with fewer measurements, thereby achieving lower BER. In this context, 70% ($S_{sensing} = 0.7$) of the original signal dimensions are used for data acquisition, with the channel matrix effectively acting as an analogous sensing matrix. The threshold is set to 0.1, helping to distinguish the actual transmitted signal from noise when decoding the received signal.

In Fig. 7b, the BER performance of FT-IMMA is compared with those of IM-OFDMA and OFDMA to transmit 12 bits in each scenario. FT-IMMA activates two FT resources and transmits symbols by utilizing $M = 16$ whereas IM-OFDMA requires $M = 32$ over the two active subcarriers to send the same number of bits. OFDMA employs $M = 8$ but requires 4 FT resources per user. This results in inter-carrier interference (ICI) enhancement and BER deterioration. In this scenario, FT-IMMA provides nearly 3 and 8 dB gains compared with IM-OFDMA and OFDMA, respectively, to attain a BER of $\sim 9 \times 10^{-04}$. Furthermore, the theoretical BER of the proposed technique follows the numerical results. To reduce the computational complexity of detection, OMP leverages a reduced search space by utilizing only 80% ($S_{sensing} = 0.8$) of the original signal dimensions for data recovery. As shown, the BER performance achieved with OMP closely approaches that of ML detection in medium to high SNR regions, demonstrating OMP's efficiency and accuracy in these scenarios. By using a relatively higher sensing parameter, which captures a greater portion of the original signal dimensions than in the case shown in Fig. 7a, the reconstruction accuracy improves,



(a) (b)

Fig. 8: (a) BER performance at 4 bps/Hz. (b) BER performance at 2 bps/Hz.



(a) (b)

Fig. 9: (a) BER performance for unequal T and F resources. (b) BER performance for different numbers of users in the FT-IMMA system transmitting 9 bits.

resulting in a lower BER.

BER comparison to transmitting 16 bits with 3 active resources per user in IM cases is shown in Fig. 8a, whereas the classical OFDMA employs 4 resources per user. The BER trend is similar to that of previously discussed scenarios. A comparison of the BER of FT-IMMA and IM-OFDMA with single and two active resources per user is shown in Fig. 8b. Cases with two active resources require lower modulation sizes to achieve an SE similar to scenarios with a single active resource. FT-IMMA outperformed its counterparts in each case. IM-OFDMA with $f = 1$ fails to outperform OFDMA because the predecessor requires $M = 64$, whereas the successor provides the same SE at $M = 4$.

As shown in Fig. 8a and 8b, activating more patterns allows for smaller constellation sizes to achieve comparable SE. However, this also increases the search space for ML detection, as it must evaluate all possible signal combinations, as discussed in Section II-A. To address this, OMP is applied to reduce the search space by using 75% ($\mathcal{S}_{sensing} = 0.75$) of the original signal dimensions, striking a balance between computational complexity and BER performance. Notably, OMP's BER closely aligns with ML detection in high SNR regions, benefiting from reduced data processing requirements at the receiver.

Fig. 9a shows the BER performance of FT-IMMA when $F \neq T$ and compares it with IM-OFDMA. Also, cases with different numbers of active users and $f \neq t$ are considered. It is observed in the scenario that BER slightly deteriorates when the number of active users or FT resources increases. Fig. 9b shows the impact on BER performance with varying numbers of active users. The system has an average 4 dB loss when the number of users increases from 2 to 5. Additionally, the

comparison with a higher number of available FT resources is depicted with the same number of users but at a lower modulation order to transmit the same amount of bits. In that case, BER improves due to the low probability of collision and lower constellation size. However, it could result in overall SE loss due to the increased number of inactive resources.

V. Conclusion

This article introduced an uplink FT-IMMA, leveraging a two-dimensional resource block within conventional OFDMA through the implementation of IM principles. The enhancement in EE and SE was attributed to the effective utilization of IM in two dimensions. Compared with OFDMA, fewer active FT resource blocks could send more bits. The proposed scheme was operated without a scheduling mechanism, leading to occasional data collisions. However, collision occurrences were observed to be lower than those associated with traditional IM-OFDMA and IMMA. Moreover, when compared to traditional methods, the proposed FT-IMMA has attained a comparable SE while employing a smaller modulation size, thereby resulting in an improvement in BER. However, FT-IMMA has higher detection complexity compared with IM-OFDMA and IMMA. Therefore, OMP-MUD has been proposed for fast and low-complexity implementation with near-optimal performance.

A. Future Challenges

This technique is particularly well-suited for scenarios requiring random access, such as multichannel ALOHA, where multiple users access subchannels in a stochastic manner. Future challenges include the design of low-complexity decoders and possible solutions for collision resolution. One approach to further reduce collisions is by utilizing the concept of composite multi-mode OFDM from [48], where different users transmit using varying energy coefficients and distinguishable constellations. Another approach to reducing collisions in future work will be to utilize the concept of uplink NOMA, where signals from multiple users will be received at different power levels. This will enable the recovery of collided users through successive interference cancellation, as users with distinct power levels can be separated and decoded efficiently.

To further enhance SE and reduce collisions, a more generalized scheme, such as quadrature FT-IMMA, could be considered in future work. In this approach, separate FT resources would be allocated to transmit the in-phase (I) and quadrature (Q) components of the signal, allowing more data to be transmitted via index bits. This would provide a more flexible and efficient use of resources, further improving system performance.

References

[1] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2294–2323, 2018.

[2] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, 2021.

[3] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 315–348, 2018.

[4] E. Başar, Ü. Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, 2013.

[5] Q. Li, M. Wen, B. Clerckx, S. Mumtaz, A. Al-Dulaimi, and R. Q. Hu, "Subcarrier index modulation for future wireless networks: Principles, applications, and challenges," *IEEE Wireless Commun.*, vol. 27, no. 3, pp. 118–125, 2020.

[6] T. Mao and Z. Wang, "Terahertz wireless communications with flexible index modulation aided pilot design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1651–1662, 2021.

[7] T. Mao, Z. Zhou, Z. Xiao, C. Han, and Z. Wang, "Index-modulation-aided terahertz communications with reconfigurable intelligent surface," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, 2024.

[8] S. A. Nambi and K. Giridhar, "Lower order modulation aided BER reduction in OFDM with index modulation," *IEEE Wireless Commun. Lett.*, vol. 22, no. 8, pp. 1596–1599, 2018.

[9] Y. Yang, S. Dang, M. Wen, and M. Guizani, "Millimeter wave MIMO-OFDM with index modulation: A pareto paradigm on spectral- energy efficiency trade-off," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6371–6386, 2021.

[10] S. Doğan Tusha, A. Tusha, E. Basar, and H. Arslan, "Multidimensional index modulation for 5G and beyond wireless networks," *Proc. IEEE*, vol. 109, no. 2, pp. 170–199, 2021.

[11] M. Wen, J. Li, S. Dang, Q. Li, S. Mumtaz, and H. Arslan, "Joint-mapping orthogonal frequency division multiplexing with subcarrier number modulation," *IEEE Trans. on Commun.*, vol. 69, no. 7, pp. 4306–4318, 2021.

[12] S. Althunibat, R. Mesleh, and T. F. Rahman, "A novel uplink multiple access technique based on index-modulation concept," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4848–4855, 2019.

[13] J. Li, Q. Li, S. Dang, M. Wen, X.-Q. Jiang, and Y. Peng, "Low-complexity detection for index modulation multiple access," *IEEE Wireless Commun.*, vol. 9, no. 7, pp. 943–947, 2020.

[14] S. Althunibat, R. Mesleh, and K. A. Qaraqe, "IM-OFDMA: A novel spectral efficient uplink multiple access based on index modulation," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 315–10 319, 2019.

[15] S. Althunibat, R. Mesleh, and K. Qaraqe, "Quadrature index modulation based multiple access scheme for 5G and beyond," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2257–2261, 2019.

[16] W. Belaoura, S. Althunibat, K. Qaraqe, and K. Ghanem, "Precoded index modulation based multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12 912–12 920, 2020.

[17] J. Li, S. Dang, M. Wen, Q. Li, Y. Chen, Y. Huang, and W. Shang, "Index modulation multiple access for 6G communications: Principles, applications, and challenges," *IEEE Netw.*, vol. 37, no. 1, pp. 52–60, 2023.

[18] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[19] Z. Han, H. Li, and W. Yin, *Compressive sensing for wireless networks*. Cambridge Univ. Press, 2013.

[20] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge univ. press, 2012.

[21] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[22] Z. Gao, L. Dai, C. Qi, C. Yuen, and Z. Wang, "Near-optimal signal detector based on structured compressive sensing for massive SM-MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1860–1865, 2017.

[23] M. Huang, K. Xu, D. Zhang, X. Xia, W. Xie, and J. Xu, "Improved OMP algorithm for cell-free massive MIMO networks with multi-user spatial index modulation," in *IEEE Intl. Conf. Computer and Commun.*, 2020, pp. 612–617.

[24] C. Zhang, H. Lu, and J. Liu, "IMRecoNet: Learn to detect in index modulation aided MIMO systems with complex valued neural networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 1791–1805, 2023.

[25] S. Adnan, Y. Fu, N. U. R. Junejo, Z. Chen, and H. Esmaiel, "Sparse detection with orthogonal matching pursuit in multiuser uplink quadrature spatial modulation MIMO system," *IET Commun.*, vol. 13, no. 20, pp. 3472–3478, 2019.

[26] H. Zhang, L.-L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.

[27] J. Choi, "Single-carrier index modulation and CS detection," in *IEEE Int. Conf. Commun.*, 2017, pp. 1–6.
[28] ——, "Single-carrier index modulation for iot uplink," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 6, pp. 1237–1248, 2019.
[29] T. Datta, H. S. Eshwaraiah, and A. Chockalingam, "Generalized space-and-frequency index modulation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 4911–4924, 2016.
[30] B. Shamasundar, S. Bhat, S. Jacob, and A. Chockalingam, "Multidimensional index modulation in wireless communications," *IEEE Access*, vol. 6, pp. 589–604, 2018.
[31] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed-sensing-aided space-time frequency index modulation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6259–6271, 2018.
[32] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
[33] S. Li, F. Wang, Y. Zhang, R. Li, S. Shi, Y. Li, X. Li, and D. B. d. Costa, "Orthogonal chirp division multiplexing assisted dual-function radar communication in IoT networks," *IEEE Internet Things J.*, vol. 11, no. 13, pp. 23 752–23 764, 2024.
[34] Y.-T. Lai, Y.-R. Ciou, and J.-M. Wu, "Index modulation multiple access," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun., Bologna, Italy*, 2018, pp. 1–5.
[35] A. Almohamad, S. Althunibat, M. Qaraqe, R. Mesleh, and M. Hasna, "Performance analysis of index modulation based multiple access under imperfect channel estimation," in *Proc. IEEE Int. Conf. Commun. Netw.*, 2020, pp. 1–5.
[36] O. Alaca, S. Althunibat, S. Yarkan, S. L. Miller, and K. A. Qaraqe, "CNN-based signal detector for IM-OFDMA," in *IEEE Glob. Commun. Conf.*, 2021, pp. 01–06.
[37] O. Alaca, S. Althunibat, S. L. Miller, and K. A. Qaraqe, "Analysis of receiver IQ-imbalance in IM-OFDMA uplink systems," *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 917–921, 2022.
[38] O. Alaca, S. Althunibat, S. Yarkan, and Miller, "Performance analysis of uplink IM-OFDMA systems in the presence of CFO and Rx-IQI," in *IEEE Int. Conf. Inf. Netw.*, 2023, pp. 168–173.
[39] O. Alaca, S. Althunibat, S. Yarkan, S. L. Miller, and K. A. Qaraqe, "Secrecy analysis of uplink IM-OFDMA systems in the presence of IQ imbalance," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14 411–14 425, 2023.
[40] O. Alaca, S. Althunibat, S. Yarkan, and Miller, "Analysis and compensation of Tx and Rx IQ imbalances in uplink IM-OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 12 956–12 969, 2023.
[41] X. Pei, Y. Chen, M. Wen, H. Yu, E. Panayirci, and H. V. Poor, "Next-generation multiple access based on NOMA with power level modulation," *IEEE JSAC*, vol. 40, no. 4, pp. 1072–1083, 2022.
[42] X. zhu, Z. Wang, and J. Cao, "NOMA-based spatial modulation," *IEEE Access*, vol. 5, pp. 3790–3800, 2017.
[43] Z. Yuan, Y. Hu, W. Li, and J. Dai, "Blind multi-user detection for autonomous grant-free high-overloading multiple-access without reference signal," in *IEEE 87th VTC Spring*, 2018, pp. 1–7.
[44] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6202–6216, 2012.
[45] H. Jasiulewicz and W. Kordecki, "Convolutions of Erlang and of Pascal distributions with applications to reliability," *Demonstratio Math.*, vol. 36, no. 1, pp. 231–238, 2003.
[46] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Cambridge, MA, USA: Academic, 2014.
[47] M.-S. Alouini and A. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, 1999.
[48] J. Li, S. Dang, Y. Huang, P. Chen, X. Qi, M. Wen, and H. Arslan, "Composite multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3748–3761, 2023.