Generalized Space-and-Frequency Index Modulation

Tanumay Datta, Harsha S. Eshwaraiah, and A. Chockalingam, *Senior Member, IEEE*

***Abstract*—Unlike in conventional modulation where information bits are conveyed only through symbols from modulation alphabets defined in the complex plane [e.g., quadrature amplitude modulation (QAM) and phase shift keying (PSK)], in index modulation (IM), additional information bits are conveyed through indexes of certain transmit entities that get involved in the transmission. Transmit antennas in multiantenna systems and subcarriers in multicarrier systems are examples of such transmit entities that can be used to convey additional information bits through indexing. In this paper, we introduce *generalized space and frequency IM*, where the indexes of active transmit antennas and subcarriers convey information bits. We first introduce IM in the spatial domain, which is referred to as generalized spatial IM (GSIM). For GSIM, where bits are indexed only in the spatial domain, we derive the expression for achievable rate and easy-to-compute upper and lower bounds on this rate. We show that the achievable rate in GSIM can be more than that in spatial multiplexing and analytically establish the condition under which this can happen. It is noted that GSIM achieves this higher rate using fewer transmit radio-frequency (RF) chains compared with spatial multiplexing. We also propose a Gibbs-sampling-based detection algorithm for GSIM and show that GSIM can achieve better bit error rate (BER) performance than spatial multiplexing. For generalized space–frequency IM (GSFIM), where bits are encoded through indexing in both active antennas and subcarriers, we derive the achievable rate expression. Numerical results show that GSFIM can achieve higher rates compared with conventional multiple-input-multiple-output orthogonal frequency division multiplexing (MIMO-OFDM). Moreover, BER results show the potential for GSFIM performing better than MIMO-OFDM.**

***Index Terms*—Achievable rate, detection, multiantenna systems, multicarrier systems, space–frequency index modulation, spatial index modulation, transmit radio-frequency (RF) chains.**

I. Introduction

MULTIANTENNA wireless systems have become very popular due to their high spectral efficiencies and improved performance compared with single-antenna systems [1]–[3]. Practical multiantenna systems are faced with the problem of maintaining multiple radio-frequency (RF) chains at the transmitter and receiver, as well as the associated RF hardware complexity, size, and cost [4]. Spatial modulation, which is a transmission scheme that uses multiple transmit antennas but only one transmit RF chain, can alleviate the need for multiple transmit RF chains [5]–[7]. In spatial modulation, at any given time, only one among the transmit antennas will be active, and the other antennas remain silent. The index of the active transmit antenna will also convey information bits, in addition to the information bits conveyed through the conventional modulation symbol (e.g., chosen from the quadrature amplitude modulation (QAM) and phase shift keying (PSK) alphabet) sent on the active antenna. An advantage of spatial modulation over conventional modulation is that, for a given spectral efficiency, conventional modulation requires a larger modulation alphabet size than spatial modulation, and this can lead to spatial modulation performing better than conventional modulation [8], [9].

In this paper, we take the view that spatial modulation is an instance of the general idea of "index modulation (IM)." Unlike in conventional modulation where information bits are conveyed only through symbols from modulation alphabets defined in the complex plane (e.g., QAM and PSK), in IM, additional information bits are conveyed through indexes of certain transmit entities that get involved in the transmission. Transmit antennas in multiantenna systems, subcarriers in multicarrier systems, and precoders are examples of such transmit entities that can be used to convey information bits through indexing. Indexing in spatial domain (e.g., spatial modulation and space shift keying, which is a special case of spatial modulation) is a widely studied and reported IM technique (see [7] and the references therein). Far fewer works have been reported in frequency and precoder IM techniques, e.g., subarrier IM (SIM) in [10]–[13] and precoder IM in [14]. The focus of this paper is twofold: 1) generalization of the idea of spatial modulation, which we refer to as generalized spatial IM (GSIM); and 2) generalization of the idea of IM to both spatial domain (multiple antennas) and frequency domain (subcarriers), which we refer to as generalized space–frequency IM (GSFIM).

In spatial modulation, the choice of the transmit antenna to activate in a channel use is made based on a group of m bits, where the number of transmit antennas is $n_t = 2^m$. On the chosen antenna, a symbol from an M-ary modulation alphabet $\mathbb{A}$ (e.g., M-QAM) is sent. The remaining $n_t - 1$ antennas remain silent. Therefore, the achieved rate in spatial modulation, in bits per channel use (bpcu), is $\log_2 n_t + \log_2 M$. The error performance of spatial modulation has been studied extensively, and it has been shown that spatial modulation can achieve performance gains compared with spatial multiplexing [15], [16]. Space shift keying is a special case of spatial modulation [17], where instead of sending an M-ary modulation symbol, a signal known to the receiver, e.g., $+1$, is sent on the chosen antenna. Therefore, the achieved rate in space shift keying is $\log_2 n_t$ bpcu. In spatial modulation and space shift keying, the number

Manuscript received October 13, 2014; accepted June 23, 2015. Date of publication September 25, 2015; date of current version July 14, 2016. This work was presented in part at the IEEE Wireless Communications and Networking Conference, Shanghai, China, April 2013. The review of this paper was coordinated by Dr. T. Jiang.

T. Datta is with Centrale Supelec, 91190 Gif sur Yvette, France (e-mail: tanumaydatta@gmail.com).

H. S. Eshwaraiah was with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India (e-mail: seharsha@gmail.com).

A. Chockalingam is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India (e-mail: achockal@ece.iisc.ernet.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2015.2451095

of transmit RF chains is restricted to one, and the number of transmit antennas is restricted to powers of two. The first contribution in this paper consists of generalization of spatial modulation, which removes these restrictions [18]–[21], an analysis of achievable rate, and proposal of a detection algorithm. In GSIM, the transmitter has n_t transmit antenna elements and $n_{\rm rf}$ transmit RF chains; $1 \leq n_{\rm rf} \leq n_t$, and $n_{\rm rf}$ out of n_t antennas are activated at a time, and therefore, $\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \rfloor$ additional bits are conveyed through antenna indexing. Spatial modulation and spatial multiplexing turn out to be special cases of GSIM for $n_{\rm rf} = 1$ and $n_{\rm rf} = n_t$, respectively. We derive the expression for the achievable rate in GSIM and easy-to-compute upper and lower bounds on this rate. We show that the achievable rate in GSIM can be more than that in spatial multiplexing and analytically establish the condition under which this can happen. It is noted that GSIM achieves this higher rate using fewer transmit RF chains compared with spatial multiplexing. We also propose a Gibbs-sampling-based detection algorithm for GSIM and show that GSIM can achieve better bit error rate (BER) performance than spatial multiplexing.

In the second contribution in this paper, we introduce GSFIM that uses both spatial and frequency domains to encode bits through indexing. GSFIM can be viewed as a generalization of the GSIM scheme by exploiting indexing in the frequency domain as well. Index modulation that exploits the frequency domain alone—referred to as subcarrier IM (SIM)—has been studied in [10]–[13]. These works have shown that OFDM with SIM achieves better performance than conventional OFDM, particularly at medium-to-high SNRs. These works have not exploited indexing in the spatial domain in multiple-input–multiple-output (MIMO) systems. Our contribution addresses, for the first time, indexing both in space and frequency in MIMO systems. In particular, we 1) propose a signaling architecture for combined space and frequency indexing, 2) study in detail its achieved rate in comparison with conventional MIMO-OFDM, and 3) show that better performance compared with that in conventional MIMO-OFDM can be achieved in the medium-to-high SNR regime. The proposed GSFIM system has N subcarriers, n_t transmit antennas, and $n_{\rm rf}$ transmit RF chains $1 \leq n_{\rm rf} \leq n_t$. In the spatial domain, $n_{\rm rf}$ out of n_t transmit antennas are chosen for activation based on $\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \rfloor$ bits. In the frequency domain, in a space–frequency block of size $n_{\rm rf} \times N$, information bits are encoded in multiple subblocks where each subblock is of size $n_{\rm rf} \times n_f$, and N/n_f is the number of subblocks. We characterize the achievable rate in GSFIM as a function of the system parameters. We show that GSFIM can offer better rates and fewer transmit RF chains compared with those in conventional MIMO-OFDM. It is also shown that GSFIM can achieve better BER performance than MIMO OFDM.

The remainder of this paper is organized as follows. In Section II, we present the GSIM system model, and a detailed analysis of achievable rate and rate bounds in GSIM. We quantify rate gains and savings in transmit RF chains in GSIM compared with spatial multiplexing. The proposed detection algorithm for GSIM and its BER performance are also presented. In Section III, we present the GSFIM system model, analysis of achievable rate in GSFIM, and BER performance of GSFIM. Conclusions and scope for future work are presented in Section IV.

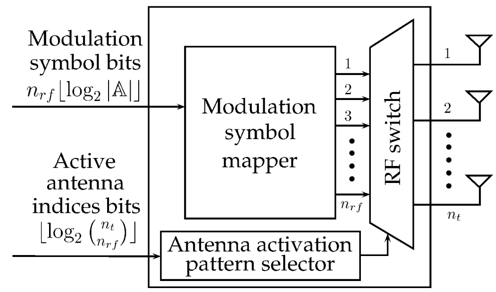


Fig. 1. GSIM transmitter.

II. Generalized Spatial Index Modulation

Here, we consider GSIM that encodes bits through indexing in the spatial domain. In GSIM, the transmitter has n_t transmit antennas and $n_{\rm rf}$ transmit RF chains, $1 \leq n_{\rm rf} \leq n_t$. In any given channel use, $n_{\rm rf}$ out of n_t antennas are activated. Information bits are conveyed through both conventional modulation symbols and the indexes of the active antennas. Spatial multiplexing becomes a special case of GSIM with $n_{\rm rf} = n_t$. We present an analysis of the achievable rates in GSIM, which shows that the maximum achievable rate in GSIM can be more than the rate in spatial multiplexing and that using fewer transmit RF chains.

A. System Model

A GSIM transmitter is shown in Fig. 1. It has n_t transmit antennas and $n_{\rm rf}$ transmit RF chains: $1 \leq n_{\rm rf} \leq n_t$. An $n_{\rm rf} \times n_t$ switch connects the RF chains to the transmit antennas. In a given channel use, $n_{\rm rf}$ out of n_t transmit antennas are chosen, and $n_{\rm rf}$ M-ary modulation symbols are sent on these chosen antennas. The remaining $n_t - n_{\rm rf}$ antennas remain silent (i.e., they can be viewed as transmitting the value zero). Therefore, if $\mathbb{A}$ denotes the M-ary modulation alphabet used on the active antennas, the effective alphabet becomes $\mathbb{A}_0 \triangleq \mathbb{A} \cup 0$.

Define an antenna activation pattern to be an n_t-length vector that indicates which antennas are active (denoted by a “1” in the corresponding antenna index) and which antennas are silent (denoted by a “0”). There are $L = \binom{n_t}{n_{\rm rf}}$ antenna activation patterns possible, and $K = \lfloor \log_2 \binom{n_t}{n_{\rm rf}} \rfloor$ bits are used to choose an activation pattern for a given channel use. Note that not all L activation patterns are needed, and any 2^K patterns out of them are adequate. Take any 2^K patterns out of L patterns and form a set called the “antenna activation pattern set” $\mathbb{S}$. Let us illustrate this using the following example. Let $n_t = 4$ and $n_{\rm rf} = 2$. Then, $L = \binom{4}{2} = 6, K = \lfloor \log_2 6 \rfloor = 2$, and $2^K = 4$. The six antenna activation patterns are given by

$$\left\{[1,1,0,0]^T, [1,0,1,0]^T, [0,1,0,1]^T, [0,0,1,1]^T \right.$$
$$\left. [0,1,1,0]^T, [1,0,0,1]^T \right\}.$$

Out of these six patterns, any $2^K = 4$ patterns can be taken to form the set $\mathbb{S}$. Accordingly, let us take the antenna activation pattern set as

$$\mathbb{S} = \left\{[1,1,0,0]^T, [1,0,1,0]^T, [0,1,0,1]^T, [0,0,1,1]^T \right\}.$$

TABLE I
DATA BITS TO GSIM SIGNAL MAPPING FOR $n_t = 4$, $n_{\rm rf} = 2$. $\mathbb{A}$: M-ARY MODULATION ALPHABET

Data bits $K = 2$	Antenna activity pattern	Antenna status			
		Ant.1	Ant.2	Ant.3	Ant.4
0 0	$[1,\ 1,\ 0,\ 0]^T$	$\in \mathbb{A}$	$\in \mathbb{A}$	OFF	OFF
0 1	$[1,\ 0,\ 1,\ 0]^T$	$\in \mathbb{A}$	OFF	$\in \mathbb{A}$	OFF
1 0	$[0,\ 1,\ 0,\ 1]^T$	OFF	$\in \mathbb{A}$	OFF	$\in \mathbb{A}$
1 1	$[0,\ 0,\ 1,\ 1]^T$	OFF	OFF	$\in \mathbb{A}$	$\in \mathbb{A}$

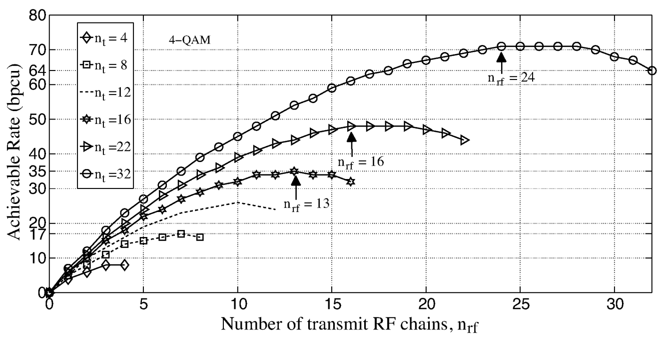


Fig. 2. Achievable rate in GSIM $R_{\rm gsim}$ as a function of $n_{\rm rf}$ for different values of n_t and 4-QAM.

Table I shows the mapping of data bits to GSIM signals for $n_t = 4, n_{\rm rf} = 2$ for the given activation pattern set. Suppose 4-QAM is used to send information on the active antennas. Let $\mathbf{x} \in \mathbb{A}_0^{n_t}$ denote the n_t-length transmit vector. Let 010011 denote the information bit sequence. GSIM translates these bits to the transmit vector $\mathbf{x}$ as follows: 1) The first two bits are used to choose the activity pattern; 2) the second two bits form a 4-QAM symbol; and 3) the third two bits form another 4-QAM symbol, so that, with Gray mapping, the transmit vector $\mathbf{x}$ becomes

$$\mathbf{x} = [1+\mathbf{j},\ 0,\ -1-\mathbf{j},\ 0]^T$$

where $\mathbf{j} = \sqrt{-1}$.

B. Achievable Rates in GSIM

The transmit vector in a given channel use in GSIM is formed using 1) antenna activation pattern selection bits and 2) M-ary modulation bits. The number of activation pattern selection bits is $\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \rfloor$. The number of M-ary modulation bits is $n_{\rm rf} \log_2 M$. By combining these two parts, the achievable rate in GSIM with n_t transmit antennas, $n_{\rm rf}$ transmit RF chains, and M-QAM can be written as

$$R_{\rm gsim} = \underbrace{\left\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \right\rfloor}_{\text{Antenna index bits}} + \underbrace{n_{\rm rf} \log_2 M}_{\text{modulation symbol bits}} \quad \text{bpcu}. \tag{1}$$

Let us examine the GSIM rate $R_{\rm gsim}$ in (1) in some detail. In particular, let us examine how $R_{\rm gsim}$ varies as a function of its variables. Fig. 2 shows the variation of $R_{\rm gsim}$ as a function of $n_{\rm rf}$ for different values of $n_t = 4, 8, 12, 16, 22, 32$ and 4-QAM. The value of $n_{\rm rf}$ in the x-axis is varied from 0 to n_t. As mentioned earlier, $n_{\rm rf} = n_t$ corresponds to spatial multiplexing. The $R_{\rm gsim}$ versus $n_{\rm rf}$ plot for a given n_t shows an interesting behavior, namely, for a given n_t, there is an optimum $n_{\rm rf}$ that maximizes the achievable rate $R_{\rm gsim}$. Let $R_{\rm gsim}^{\max}$ denote the maximum achievable rate, i.e.,

$$R_{\rm gsim}^{\max} = \max_{1 \le n_{\rm rf} \le n_t} R_{\rm gsim}. \tag{2}$$

In Fig. 2, it is interesting to see that $R_{\rm gsim}^{\max}$ does not necessarily occur at $n_{\rm rf} = n_t$ but at some $n_{\rm rf} < n_t$. $R_{\rm gsim}$ can exceed the spatial multiplexing rate of $n_t \log_2 M$ whenever the first term in (1) exceeds $(n_t - n_{\rm rf}) \log_2 M$. The following theorem formally establishes the condition under which the $R_{\rm gsim}^{\max}$ will be more than the spatial multiplexing rate of $n_t \log_2 M$.

Theorem 1: The maximum achievable rate in GSIM is strictly greater than the rate achieved in spatial multiplexing (i.e., $R_{\rm gsim}^{\max} > n_t \log_2 M$) if and only if $n_t \ge 2M$.

Proof: Consider the two terms on the right-hand side (RHS) of the rate expression (1). The first term (contribution due to antenna index bits) increases when $n_{\rm rf}$ is increased from 0 to $\lfloor n_t/2 \rfloor$ and then decreases, i.e., it peaks at $n_{\rm rf} = \lfloor n_t/2 \rfloor$. The second term (contribution due to modulation symbol bits), on the other hand, increases linearly with $n_{\rm rf}$. These two terms when added can cause a peak at some $n_{\rm rf}$ in the range $\lfloor n_t/2 \rfloor \le n_{\rm rf} \le n_t$. Observe that, as we reduce $n_{\rm rf}$ below n_t, we gain rate from the first term but lose rate in the second term. The rate loss in the second term is $\log_2 M$ bpcu per RF chain reduced. Therefore, we can rewrite (1) as

$$R_{\rm gsim} = n_t \log_2 M + \left\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \right\rfloor - (n_t - n_{\rm rf}) \log_2 M. \tag{3}$$

Case 1: $n_t \ge 2M$: If $n_t \ge 2M$, then $\lfloor \log_2 n_t \rfloor > \log_2 M$. By putting $n_{\rm rf} = n_t - 1$ in (3), we get

$$R_{\rm gsim} = n_t \log_2 M + \lfloor \log_2 n_t \rfloor - \log_2 M. \tag{4}$$

Therefore, in this case, the $R_{\rm gsim}$ in (4) is more than $n_t \log_2 M$, i.e., GSIM with $n_{\rm rf} = n_t - 1$ RF chains achieves more rate than spatial multiplexing. This implies $R_{\rm gsim}^{\max} > n_t \log_2 M$, i.e., the maximum rate available in GSIM is more than the spatial multiplexing rate. Conversely, if $n_t < 2M$, we show below that $R_{\rm gsim}^{\max}$ is not more than the spatial multiplexing rate.

Case 2: $n_t < 2M$: If $n_t < 2M$

$$\log_2 n_t < 1 + \log_2 M. \tag{5}$$

From the properties of binomial coefficients, we have

$$\begin{aligned} \binom{n_t}{n_{\rm rf}} &= \binom{n_t}{n_t - n_{\rm rf}} \\ &= \frac{n_t(n_t-1), \ldots, (n_{\rm rf}+1)}{1\,2, \ldots, (n_t - n_{\rm rf})} < \frac{n_t^{n_t - n_{\rm rf}}}{2^{n_t - n_{\rm rf} - 1}}. \end{aligned} \tag{6}$$

Hence

$$\begin{aligned} \left\lfloor \log_2 \binom{n_t}{n_{\rm rf}} \right\rfloor &\le \lfloor (n_t - n_{\rm rf}) \log_2 n_t - n_t + n_{\rm rf} + 1 \rfloor && (7) \\ &< \lfloor (n_t - n_{\rm rf})(1 + \log_2 M) - n_t + n_{\rm rf} + 1 \rfloor && (8) \\ &< (n_t - n_{\rm rf} + 1) \log_2 M && (9) \\ &\le (n_t - n_{\rm rf}) \log_2 M. && (10) \end{aligned}$$

TABLE II
PERCENTAGE SAVING IN TRANSMIT RF CHAINS AND PERCENTAGE INCREASE IN RATE IN GSIM COMPARED TO SPATIAL MULTIPLEXING FOR $N_t = 16, 32$, AND BPSK, 4-QAM, 8-QAM, 16-QAM

M-ary alphabet	Percentage saving in no. of Tx RF chains at $R = R_{\text{gsim}}^{max}$		Percentage saving in no. of Tx RF chains at $R = n_t \log_2 M$		Percentage increase in rate at $R = R_{\text{gsim}}^{max}$	
	$n_t = 16$	$n_t = 32$	$n_t = 16$	$n_t = 32$	$n_t = 16$	$n_t = 32$
BPSK	31.25	40.63	68.75	71.88	43.75	46.88
4-QAM	18.75	25	37.5	43.75	9.385	10.94
8-QAM	6.25	12.5	18.75	21.88	2.08	3.13
16-QAM	6.25	3.13	6.25	9.38	0	0.78

The inequality in (7) is obtained by taking logarithm in (6), and (8) is obtained from (7) and (5). Hence, using (3), we obtain $R_{\text{gsim}} \leq n_t \log_2 M$, for $1 \leq n_{\text{rf}} \leq n_t$, and thus, for $n_t < 2M$, $R_{\text{gsim}}^{\max} \leq n_t \log_2 M$. Combining the arguments in Cases 1 and 2, we get **Theorem 1**. ■

From Fig. 2, the following interesting observations can be made.

1) By choosing the optimum (n_t, n_{rf}) combination (i.e., using fewer RF chains than transmit antennas, $n_{\text{rf}} < n_t$), GSIM can achieve a higher rate than that of spatial multiplexing, where $n_{\text{rf}} = n_t$.
2) One can operate GSIM at the same rate as that of spatial multiplexing but with even fewer RF chains.

For example, for $n_t = 32$, the optimum n_{rf} that maximizes R_{gsim} is 24, and the corresponding maximum rate $R_{\text{gsim}}^{\max}$ is 71 bpcu. Compare this rate with $32 \log_2 4 = 64$ bpcu, which is the rate achieved in spatial multiplexing. This is an 11% gain in rate in GSIM compared with spatial multiplexing. Interestingly, this rate gain is achieved using a fewer number of RF chains; 24 RF chains in GSIM versus 32 RF chains in spatial multiplexing. This is a 25% saving in transmit RF chains in GSIM compared with spatial multiplexing. Further, if GSIM were to achieve the spatial multiplexing rate of 64 bpcu in this case, then it can achieve it with even fewer RF chains, i.e., using only 18 RF chains, which is a 43% saving in RF chains compared with spatial multiplexing. Table II gives the percentage gains in number of transmit RF chains at achieved rate $R = R_{\text{gsim}}^{\max}$ and $R = n_t \log_2 M$, and the percentage gains in rates achieved by GSIM compared with spatial multiplexing for $n_t = 16, 32$ with BPSK, 4-QAM, 8-QAM, and 16-QAM.

C. Bounds on Achievable Rates in GSIM

We now proceed to obtain bounds on the achievable rate in GSIM. From (1), we observe that

$$R_{\text{gsim}} \leq \log_2 \left(\frac{n_t!}{n_{\text{rf}}!(n_t - n_{\text{rf}})!} \right) + n_{\text{rf}} \log_2 M \quad (11)$$

$$R_{\text{gsim}} > \log_2 \left(\frac{n_t!}{n_{\text{rf}}!(n_t - n_{\text{rf}})!} \right) + n_{\text{rf}} \log_2 M - 1. \quad (12)$$

From the properties of the factorial operator [22], we have

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \leq n! \leq e\sqrt{n} \left(\frac{n}{e}\right)^n \qquad \forall\, n \in \mathbb{N}. \quad (13)$$

Let us define the function $f(n_t, n_{\text{rf}}, \log_2 M)$ as

$$f(n_t, n_{\text{rf}}, \log_2 M) \triangleq n_t \log_2 n_t - n_{\text{rf}} \log_2 n_{\text{rf}} - (n_t - n_{\text{rf}}) \log_2 (n_t - n_{\text{rf}}) + n_{\text{rf}} \log_2 M. \quad (14)$$

By substituting (13) in (11), using (14), and simplifying, we obtain

$$R_{\text{gsim}} \leq \log_2 \frac{e}{2\pi} + 0.5 \log_2 \frac{n_t}{n_{\text{rf}}(n_t - n_{\text{rf}})} + f(n_t, n_{\text{rf}}, \log_2 M). \quad (15)$$

In a similar way, using (13) in (12), we can write

$$R_{\text{gsim}} > \log_2 \frac{\sqrt{2\pi}}{e^2} + 0.5 \log_2 \frac{n_t}{n_{\text{rf}}(n_t - n_{\text{rf}})} + f(n_t, n_{\text{rf}}, \log_2 M) - 1. \quad (16)$$

Let us rewrite (15) and (16) in the following way:

$$R_{\text{gsim}} \leq f_1(n_t, n_{\text{rf}}) + f_2(n_t, n_{\text{rf}}) + c_1 \quad (17)$$

$$R_{\text{gsim}} > f_1(n_t, n_{\text{rf}}) + f_2(n_t, n_{\text{rf}}) + c_2 \quad (18)$$

where $f_1(n_t, n_{\text{rf}}) = 0.5 \log_2(n_t / n_{\text{rf}}(n_t - n_{\text{rf}}))$, $f_2(n_t, n_{\text{rf}}) = f(n_t, n_{\text{rf}}, \log_2 M)$, $c_1 = \log_2(e/2\pi)$, and $c_2 = \log_2(\sqrt{2\pi}/e^2) - 1$. For a fixed n_t, the maximum value of $f_1(n_t, n_{\text{rf}})$ in the range $1 \leq n_{\text{rf}} \leq n_t - 1$ is obtained at $n_{\text{rf}} = 1$ or $n_{\text{rf}} = n_t - 1$, and the maximum value is $0.5 \log_2(n_t / n_t - 1)$. Hence

$$\max\{f_1(n_t, n_{\text{rf}})\} = 0.5 \log_2 \frac{n_t}{n_t - 1}. \quad (19)$$

Moreover, the term $f_1(n_t, n_{\text{rf}})$ is minimized for $n_{\text{rf}} = \lfloor n_t/2 \rfloor$, and the minimum value is $0.5 \log_2(4/n_t) = 1 - 0.5 \log_2 n_t$ for even n_t, and is $0.5 \log_2(n_t/(n_t/2)^2 - 0.25) \geq 1 - 0.5 \log_2 n_t$ for odd n_t. Hence

$$\min\{f_1(n_t, n_{\text{rf}})\} \geq 1 - 0.5 \log_2 n_t. \quad (20)$$

Therefore, from (14), (15), (19), and (20), we obtain the upper bound on R_{gsim} as

$$R_{\text{gsim}} \leq f(n_t, n_{\text{rf}}, \log_2 M) + 0.5 \log_2 \frac{n_t}{n_t - 1} + \log_2 \frac{e}{2\pi}. \quad (21)$$

In a similar way, from (14) and (16), we obtain the lower bound on R_{gsim} as

$$R_{\text{gsim}} > f(n_t, n_{\text{rf}}, \log_2 M) - 0.5 \log_2 n_t + \log_2 \frac{\sqrt{2\pi}}{e^2}. \quad (22)$$

Since n_t, n_{rf} and M take finite positive integer values and because of the floor operation in the first term on the right-hand side of (1), we can rewrite the bounds in (21) and (22) as

$$R_{\mathrm{gsim}} \leq \left\lfloor f(n_t, n_{\mathrm{rf}}, \log_2 M) + 0.5 \log_2 \frac{n_t}{n_t - 1} + \log_2 \frac{e}{2\pi} \right\rfloor \quad (23)$$

$$R_{\mathrm{gsim}} \geq \left\lceil f(n_t, n_{\mathrm{rf}}, \log_2 M) - 0.5 \log_2 n_t + \log_2 \frac{\sqrt{2\pi}}{e^2} \right\rceil. \quad (24)$$

Note that the above bounds on R_{gsim} can be computed easily for any n_t, n_{rf}, without the need for the computation of factorials of large numbers in the actual rate expression in (1). Further, noting that the optimum n_{rf} that maximizes $f_2(n_t, n_{\mathrm{rf}})$ is given by

$$n_{\mathrm{rf}}^* = \frac{n_t M}{M + 1} \quad (25)$$

we obtain the upper and lower bounds on $R_{\mathrm{gsim}}^{\max}$, by substituting n_{rf}^* in (25) into (23) and (24), respectively, as

$$R_{\mathrm{gsim}}^{\max} \leq \left\lfloor n_t \log_2(M+1) + 0.5 \log_2 \frac{n_t}{n_t - 1} + \log_2 \frac{e}{2\pi} \right\rfloor \quad (26)$$

$$R_{\mathrm{gsim}}^{\max} \geq \left\lceil f\left(n_t, \left\lfloor n_t \frac{M}{M+1} \right\rceil, \log_2 M\right) - 0.5 \log_2 n_t + \log_2 \frac{\sqrt{2\pi}}{e^2} \right\rceil. \quad (27)$$

These bounds on $R_{\mathrm{gsim}}^{\max}$ can be calculated for any given n_t and M directly, without exhaustive computation of the rate for all possible values of n_{rf}. From (26) and (27), we observe that as $n_t \to \infty$, $R_{\mathrm{gsim}}^{\max}$ can be approximated by $n_t \log_2(M + 1)$. Note that a spatial multiplexing system that uses a zero-augmented alphabet $\mathbb{A}_0$ achieves the rate of $n_t \log_2(M + 1)$ if all the symbols in $\mathbb{A}_0$ are equiprobable.

In Fig. 3(a), we plot the upper and lower bounds of R_{gsim} computed using (23) and (24), respectively, along with exact R_{gsim}, for $n_t = 16$ and BPSK ($M = 2$). The number of RF chains n_{rf} is varied from 1 to 15. It can be observed that the upper and lower bounds are tight (within 2 bpcu of the actual rate). In Fig. 3(b), we plot the upper and lower bounds of $R_{\mathrm{gsim}}^{\max}$ obtained from (26) and (27), respectively, for different values of n_t and $M = 2, 4$ (i.e., BPSK, 4-QAM). The corresponding exact $R_{\mathrm{gsim}}^{\max}$ values are also plotted for comparison. It can be observed that the lower and upper bounds of $R_{\mathrm{gsim}}^{\max}$ are within 2 bpcu of the exact $R_{\mathrm{gsim}}^{\max}$.

D. GSIM Signal Detection

Here, we consider detection of GSIM signals. Let $\mathbf{H}$ denote the $n_r \times n_t$ channel matrix, where n_r is the number of receive antennas. Assume rich scattering environment where the entries of $\mathbf{H}$ are modeled as circularly symmetric complex Gaussian with zero mean and unit variance. Let $\mathbf{y}$ denote the $n_r \times 1$-sized received vector, which is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (28)$$

where $\mathbf{x}$ is the $n_t \times 1$-sized transmit vector, and $\mathbf{n}$ is the $n_r \times 1$-sized additive white Gaussian noise vector at the receiver,

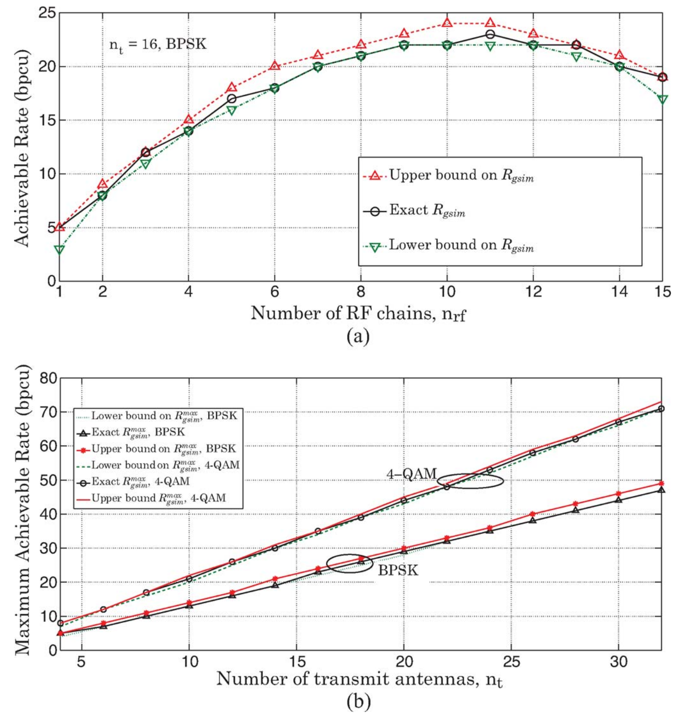


Fig. 3. (a) Bounds on R_{gsim} with BPSK for $n_t = 16$ and varying n_{rf}. (b) Bounds on $R_{\max}^{\mathrm{gsim}}$ with BPSK and 4-QAM for varying n_t.

whose ith element $n_i \sim \mathcal{CN}(0, \sigma^2) \, \forall i = 1, 2, \ldots, n_r$. Let $\mathbb{U}$ denote the set of all possible transmit vectors given by

$$\mathbb{U} = \left\{ \mathbf{x} | \mathbf{x} \in \mathbb{A}_0^{n_t \times 1}, \|\mathbf{x}\|_0 = n_{\mathrm{rf}}, \mathbf{t}^{\mathbf{x}} \in \mathbb{S} \right\} \quad (29)$$

where $\|\mathbf{x}\|_0$ denotes the zero norm of vector $\mathbf{x}$ (i.e., the number of nonzero entries in $\mathbf{x}$), and $\mathbf{t}^{\mathbf{x}}$ denotes the antenna activation pattern vector corresponding to $\mathbf{x}$, where $t_j^{\mathbf{x}} = 1$, iff $x_j \neq 0, \forall j = 1, 2, \ldots, n_t$. Note that $|\mathbb{U}| = 2^{R_{\mathrm{gsim}}}$. The activation pattern set $\mathbb{S}$ and the mapping between elements of $\mathbb{S}$ and antenna selection bits are known at both transmitter and receiver. Hence, from (28) and (29), the maximum-likelihood decision rule for GSIM signal detection is given by

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{U}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \quad (30)$$

For small values of n_t and n_{rf}, the set $\mathbb{U}$ may be fully enumerated, and ML detection as per (30) can be done. However, for medium and large values of n_t and n_{rf}, brute-force computation of $\widehat{\mathbf{x}}$ in (30) becomes computationally prohibitive. Here, we propose a low-complexity algorithm for detection of GSIM signals.

The proposed approach is based on Gibbs sampling, where a Markov chain is formed with all possible transmitted vectors as states. As the total number of nonzero entries in the solution vector has to be equal to n_{rf}, one cannot sample each coordinate individually as is done in the case of Gibbs-sampling-based detection in conventional MIMO systems [23]. To address this issue, we propose the following sampling approach: Sample two coordinates at a time jointly, keeping other $(n_t - 2)$ coordinates fixed that contain $(n_{\mathrm{rf}} - 1)$ nonzero entries.

1) Proposed Modified Gibbs Sampler: For any vector $\mathbf{x}^{(t)} \in \mathbb{A}_0^{n_t}, \|\mathbf{x}^{(t)}\|_0 = n_{\mathrm{rf}}$, where the t in the superscript of $\mathbf{x}^{(t)}$ refers to the iteration index in the algorithm. Let $i_1, i_2, \ldots, i_{n_{\mathrm{rf}}}$ denote the locations of nonzero entries and $j_1, j_2, \ldots, j_{n_t - n_{\mathrm{rf}}}$ denote the locations of zero entries in $\mathbf{x}^{(t)}$. We will sample $x_{i_l}^{(t)}$ and $x_{j_k}^{(t)}$ jointly, keeping other coordinates fixed, where $l = 1, 2, \ldots, n_{\mathrm{rf}}$ and $k = 1, 2, \ldots, (n_t - n_{\mathrm{rf}})$. As any possible transmitted vector can have only n_{rf} nonzero entries, the next possible state $\mathbf{x}^{(t+1)}$ can only be any one of the following $2|\mathbb{A}|$ candidate vectors denoted by $\{\mathbf{z}^w, w = 1, 2, \ldots, 2|\mathbb{A}|\}$, which can be partitioned into two sets. In the first set corresponding to $w = 1, 2, \ldots, |\mathbb{A}|$, we enlist the vectors that have the same activity pattern as $\mathbf{x}^{(t)}$. Hence, $z_{i_l}^w = \mathbb{A}^w$, $z_{j_k}^w = 0$, $z_q = x_q^{(t)}, q = 1, 2, \ldots, n_t$, $q \neq i_l, j_k \,\forall\, w = 1, 2, \ldots, |\mathbb{A}|$. For $w = |\mathbb{A}| + 1, |\mathbb{A}| + 2, \ldots, 2|\mathbb{A}|$, we enlist the vectors whose activity pattern differs from that of $\mathbf{x}^{(t)}$ in locations j_k and i_l. Hence, $z_{i_l}^w = \mathbb{A}^w, z_{j_k}^w = 0, z_q = x_q^{(t)}, q = 1, 2, \ldots, n_t, q \neq i_l, j_k \,\forall\, w = 1, 2, \ldots, |\mathbb{A}|$. For $w = |\mathbb{A}| + 1, |\mathbb{A}| + 2, \ldots, 2|\mathbb{A}|$, we enlist the vectors whose activity pattern differs from that of $\mathbf{x}^{(t)}$ in locations j_k and i_l. Hence, $z_{j_k}^w = \mathbb{A}^{(w - |\mathbb{A}|)}$, $z_{i_l}^w = 0, z_q = x_q^{(t)}, q = 1, 2, \ldots, n_t, q \neq i_l, j_k \,\forall\, w = |\mathbb{A}| + 1, |\mathbb{A}| + 2, \ldots, 2|\mathbb{A}|$.

To simplify the sampling process, we calculate the best vectors from the two sets corresponding to not swapping and swapping the zero and nonzero locations, and choose among these two vectors. Let $\mathbf{x}^{\mathrm{NS}}$ denote the best vector from the first set corresponding to no swap. We set $\mathbf{x}^{\mathrm{NS}} = \mathbf{x}^{(t)} + \lambda \mathbf{e}_{i_l}$ and minimize $\|\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{NS}}\|^2$ over λ. For this, we have

$$\begin{aligned}\|\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{NS}}\|^2 &= \left\|\mathbf{y} - \mathbf{H}\left(\mathbf{x}^{(t)} + \lambda \mathbf{e}_{i_l}\right)\right\|^2 \\ &= \mathbf{y}^H\mathbf{y} - 2\Re\left(\mathbf{y}^{\mathrm{MF}}\mathbf{x}^{(t)}\right) + \mathbf{x}^{(t)^H}\mathbf{R}\mathbf{x}^{(t)} \\ &\quad - 2\Re\left(\lambda \mathbf{y}^{\mathrm{MF}}\mathbf{e}_{i_l}\right) \\ &\quad + 2\Re\left(\lambda \mathbf{x}^{(t)^H}\mathbf{R}\mathbf{e}_{i_l}\right) + |\lambda|^2 R_{i_l, i_l} \end{aligned} \tag{31}$$

where $\mathbf{y}^{\mathrm{MF}} = \mathbf{y}^H\mathbf{H}$, and $\mathbf{R} = \mathbf{H}^H\mathbf{H}$. Differentiating (31) with respect to λ and equating it to zero, we get

$$\lambda_{\mathrm{opt}} = \frac{\left(y_{i_l}^{\mathrm{MF}} - \mathbf{x}^{(t)^H}\mathbf{r}_{i_l}\right)^H}{R_{i_l, i_l}} \tag{32}$$

where $\mathbf{r}_{i_l}$ is the i_lth column vector of $\mathbf{R}$. We obtain $\mathbf{x}^{\mathrm{NS}} = [\mathbf{x}^{(t)} + \lambda_{\mathrm{opt}}\mathbf{e}_{i_l}]_{\mathbb{A}}$, where $[\mathbf{x}]_{\mathbb{A}}$ denotes the element-wise quantization of $\mathbf{x}$ to its nearest point in $\mathbb{A}$. Similarly, we obtain $\mathbf{x}^S$, the best vector from the second set corresponding to swap. The next state $\mathbf{x}^{(t+1)}$ is chosen between $\mathbf{x}^S$ and $\mathbf{x}^{\mathrm{NS}}$ with probability p^S and p^{NS}, respectively, where $p^S = (1 - q)\widetilde{p}^S + q/2, p^{\mathrm{NS}} = 1 - p^S$, and

$$\widetilde{p}^S = \frac{\exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}^S\|^2 - \|\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{NS}}\|^2}{\sigma^2}\right)}{1 + \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}^S\|^2 - \|\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{NS}}\|^2}{\sigma^2}\right)}. \tag{33}$$

Here, q gives the probability of mixing between Gibbs sampling and sampling from uniform distribution. We use $q = 1/n_t$ because the simulation plots of BER as a function of q have shown that the best BER is achieved at around $q = 1/n_t$. After sampling, the best vector obtained so far is updated. The above sampling process is repeated for all l and k. The algorithm is stopped after it meets the stopping criterion or reaches the maximum number of allowable iterations and outputs the best vector in terms of ML cost obtained so far.

2) Stopping and Restart Criteria: The following stopping and restart criteria are employed in the algorithm. Let us denote the best vector so far as $\mathbf{z}$. The stopping criterion works as follows: Compute a metric $\Theta_s(\mathbf{z}) = \lceil \max(c_{\min}, c_1 \exp(\phi(\mathbf{z}))) \rceil$, where $\phi(\mathbf{z}) = (\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 - n_r\sigma^2)/\sqrt{n_r}\sigma^2$ is the normalized ML cost of $\mathbf{z}$. If $\mathbf{z}$ has not changed for $\Theta_s(\mathbf{z})$ iterations, then stop. This concludes one restart, and $\mathbf{z}$ is declared as the output of this restart. Now, check whether $\mathbf{t}^{\mathbf{x}}$ belongs to $\mathbb{S}$ or not to check its validity. Several such runs, each starting from a different initial vector, are carried out until the best valid output obtained so far is reliable in terms of ML cost. Let us denote the best vector among restart outputs as $\mathbf{s}$ and the number of restarts that has given $\mathbf{s}$ as output as r_s. We calculate another metric $\Theta_r(\mathbf{s}) = \lfloor \max(0, c_2\phi(\mathbf{s})) \rfloor + 1$ and compare r_s with this. If r_s is equal to $\Theta_r(\mathbf{s})$ or the maximum number of restarts is reached, we terminate the algorithm. The listing of the proposed algorithm is given in Algorithm 1.

Algorithm 1 Proposed Gibbs-sampling-based algorithm for GSIM detection

---

1: **input:** $\mathbf{y}$, $\mathbf{H}$, n_t, n_{rf}; MAX-ITR: max. no. of iterations; MAX-RST: max. no. of restarts;
2: Compute $\mathbf{y}^{\mathrm{MF}} = \mathbf{y}^H\mathbf{H}$ and $\mathbf{R} = \mathbf{H}^H\mathbf{H}$; initialize $r = 0$, $\kappa = 10^{10}$, $q = (1)/(n_t)$;
3: $\phi(\cdot)$: ML cost fn; $\Theta_s(\cdot)$: *stopping criterion* fn; $\Theta_r(\cdot)$: *restart criterion* fn;
4: **while** $r <$ MAX-RST **do**
5:   $\mathbf{x}^{(0)}$: initial vector $\in \mathbb{A}_0^{n_t \times 1}; \|\mathbf{x}^{(0)}\|_0 = n_{\mathrm{rf}}$; $\beta = \phi(\mathbf{x}^{(0)})$; $\mathbf{z} = \mathbf{x}^{(0)}$; $t = 0$;
6:   **while** $t <$ MAX-ITR **do**
7:     **for** $l = 1$ to n_{rf} **do**
8:       **for** $k = 1$ to $n_t - n_{\mathrm{rf}}$ **do**
9:         find i_l and j_k indexes;
10:         Compute λ_{opt} from (32); compute $\mathbf{x}^{\mathrm{NS}} = [\mathbf{x}^{(t)} + \lambda_{\mathrm{opt}}\mathbf{e}_{i_l}]_{\mathbb{A}}$; compute $\mathbf{x}^{\mathrm{NS}}$;
11:         Compute $\widetilde{p}^S$ from (33); compute $p^S = (1 - q)\widetilde{p}^S + (q)/(2)$, $p^{\mathrm{NS}} = 1 - p^S$;
12:         Choose $\mathbf{x}^{(t+1)}$ between $\mathbf{x}^S$ and $\mathbf{x}^{\mathrm{NS}}$ with probability p^S & p^{NS};
13:         $\gamma = \phi(\mathbf{x}^{(t+1)})$;
14:         **if** $(\gamma \leq \beta)$ **then**
15:           $\mathbf{z} = \mathbf{x}^{(t+1)}$; $\beta = \gamma$; calculate $\Theta_s(\mathbf{z})$;
16:         **end if**
17:         $t = t + 1$; $\beta_v^{(t)} = \beta$;
18:       **end for**
19:     **end for**
20:     **if** $\Theta_s(\mathbf{z}) < t$ **then**
21:       **if** $\beta_v^{(t)} == \beta_v^{(t - \Theta_s(\mathbf{z}))}$ **then**
22:         go to step 26

23: **end if**
24: **end if**
25: **end while**
26: $r = r + 1$;
27: **if** $\mathbf{t}^{\mathbf{z}} \in \mathbb{S}$ **then**
28: **if** $\beta < \kappa$ **then**
29: $\kappa = \beta$; $r_s = 1$; $\mathbf{s} = \mathbf{z}$; Compute $\Theta_r(\mathbf{s})$;
30: **end if**
31: **if** $\beta == \kappa$ **then**
32: $r_s = r_s + 1$;
33: **end if**
34: **if** $r_s == \Theta_r(\mathbf{s})$ **then**
35: go to step 39
36: **end if**
37: **end if**
38: **end while**
39: **output:** $\mathbf{s}$. $\mathbf{s}$: output solution vector

3) Complexity: The complexity of the proposed Gibbs-sampling-based detector can be separated into three parts: 1) computation of starting vectors; 2) computation of $\mathbf{y}^{\text{MF}}$ and $\mathbf{R}$; and 3) computations involved in the sampling and updating process. In our simulations, we use MMSE output as the starting vector for the first restart and random starting vectors for the subsequent restarts. The MMSE output needs the computation of $\left(\mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I}_{n_t}\right)^{-1}\mathbf{H}^H\mathbf{y}$, whose complexity is $\mathcal{O}(n_t^3)$. Note that this operation includes the computations of $\mathbf{y}^{\text{MF}}$ and $\mathbf{R}$. For the sampling and updating process, in each iteration, i.e., for each choice of l and k, the algorithm needs to compute $\mathbf{x}^{(t)^H}\mathbf{r}_{i_l}$ and $\mathbf{x}^{(t)^H}\mathbf{r}_{j_k}$, which requires $\mathcal{O}(n_{\text{rf}})$ computations. The rest of the computations are $\mathcal{O}(1)$. The number of iterations before the algorithm terminates is found to be $\mathcal{O}(n_{\text{rf}}(n_t - n_{\text{rf}}))$ by computer simulations. Thus, the total number of computations involved in 3) is $\mathcal{O}(n_{\text{rf}}^2(n_t - n_{\text{rf}}))$. Hence, the total complexity of the proposed algorithm for GSIM detection is $\mathcal{O}(n_t^3) + \mathcal{O}(n_{\text{rf}}^2(n_t - n_{\text{rf}}))$.

E. BER Performance Results

We now present the BER performance of GSIM. For systems with small n_t, we present brute-force ML detection performance. For systems with large n_t where brute-force ML detection is prohibitive, we present the performance using the proposed detection algorithm. We also compare the performance of GSIM with the performance of spatial multiplexing. For notation purposes, a GSIM system with n_t transmit antennas and n_{rf} transmit RF chains is referred to as "(n_t, n_{rf})-GSIM" system. Moreover, we use the term "(n_t, n_{rf})-SM" system to refer the spatial multiplexing system where $n_t = n_{\text{rf}}$. The following parameters are used in proposed detection algorithm: $c_{\min} = 10n_{\text{rf}}(n_t - n_{\text{rf}})$, $c_1 = 10n_{\text{rf}}(n_t - n_{\text{rf}})\log_2 M$, $\text{MAX-ITR} = 8n_t n_{\text{rf}}(n_t - n_{\text{rf}})\sqrt{M}$, $\text{MAX-RST} = 20$, $c_2 = 0.5(1 + \log_2 M)$. Let $n_{\text{rf}}^{\text{mid}}$ denote the minimum number of RF chains in GSIM that achieves the same rate as in spatial multiplexing for a given n_t and M. Let $n_{\text{rf}}^{\text{opt}}$ denote the number of RF chains that achieves $R_{\text{gsim}}^{\max}$ for a given n_t and M.

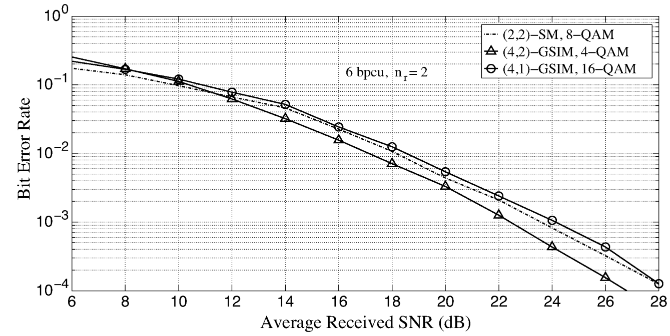


Fig. 4. BER comparison between (4, 2)-GSIM, (4, 1)-GSIM, and (2, 2)-SM systems with 6 bpcu, $n_r = 2$, and brute-force ML detection.

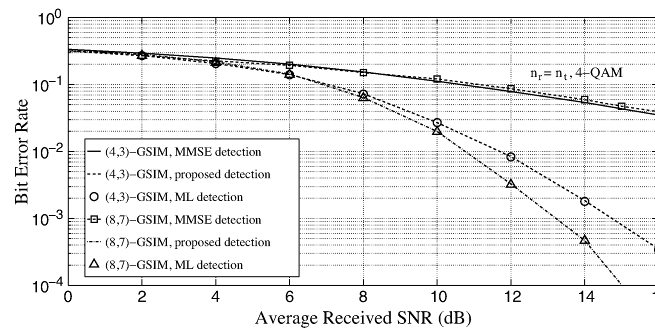


Fig. 5. BER comparison between MMSE detection, proposed detection, and brute-force ML detection in (4,3)-GSIM and (8,7)-GSIM systems with $n_r = n_t$ and 4-QAM.

In Fig. 4, we show the BER comparison between 1) (4, 2)-GSIM with 4-QAM, 2) (4,1)-GSIM with 16-QAM, and 3) (2, 2)-SM with 8-QAM, using $n_r = 2$. Note that, in all the three systems, the modulation alphabets have been chosen such that the rate is the same 6 bpcu. Since the systems are small, brute-force ML detection is used. It can be seen that (4, 2)-GSIM system performs better than (2, 2)-SM system. That is, for the same rate of 6 bpcu and $n_{\text{rf}} = 2$, GSIM achieves better performance than spatial multiplexing by about 1 dB at 0.01 uncoded BER. As we will see in Figs. 6 and 7, this improvement increases to about 1.5–2 dB for 24- and 48-bpcu systems. It is noted that GSIM needs extra transmit antennas than spatial multiplexing to achieve this improvement. However, the additional resources used in GSIM are not the transmit RF chains (which are expensive), but only the transmit antenna elements (which are not expensive). It can also be seen that even (4, 1)-GSIM performs close to within 0.5 dB of (2, 2)-SM performance in medium-to-high SNRs. This shows that GSIM can save transmit RF chains without losing much performance compared with spatial multiplexing.

Fig. 5 shows the BER performance of different detection schemes for GSIM. (4, 3)-GSIM and (8, 7)-GSIM with $n_r = n_t$ and 4-QAM are considered. Note that the choice of n_{rf} in both systems corresponds to $n_{\text{rf}}^{\text{opt}}$. Three detectors, namely, MMSE detector, proposed detector, and brute-force ML detector are considered. It can be seen that MMSE detector yields very poor performance, but the proposed detector yields a performance

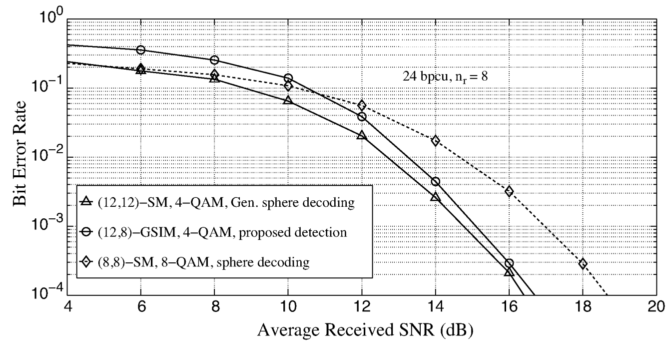


Fig. 6. BER comparison among three systems achieving 24 bpcu: 1) (8,8)-SM system with 8-QAM; 2) (12,8)-GSIM system with 4-QAM; and 3) (12,12)-SM system with 4-QAM, $n_r = 8$.

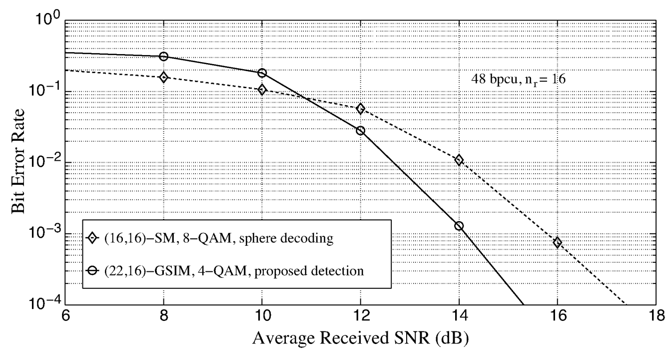


Fig. 7. BER comparison between GSIM and SM systems using same RF resources for $n_{\rm rf} = n_{rf}^{\rm opt}$, $n_r = n_{\rm rf}$ to achieve 48 bpcu.

that almost matches the ML detector performance. The proposed detector achieves this almost ML performance in just cubic complexity in n_t, whereas ML detection has exponential complexity in n_t.

In Fig. 6, we compare the performance of three systems, each achieving 24 bpcu: 1) (8, 8)-SM with 8-QAM and ML detection using sphere decoder (SD), 2) (12,8)-GSIM with 4-QAM and proposed detection, and 3) (12,12)-SM system with 4-QAM using generalized SD (GSD).[1] All the three systems use $n_r = 8$. Fig. 6 shows that the (12, 8)-GSIM with proposed detection outperforms (8, 8)-SM with SD employing same RF resources by about 2 dB in high SNR regime by using four extra transmit antennas. The performance of (12, 8)-GSIM with proposed detection is very close to that of (12,12)-SM system with GSD, which uses more RF resources to achieve the same rate. Moreover, the proposed detector has much lower complexity than GSD, which has exponential complexity in n_t.

Fig. 7 shows the BER comparison between GSIM and SM using same RF resources for $n_{\rm rf} = n_{\rm rf}^{\rm opt}$, and $n_r = n_{\rm rf}$ to achieve 48 bpcu. GSIM uses $n_t = 22$ and 4-QAM, whereas (16,16)-SM scheme uses 8-QAM modulation alphabet to match the rate. For GSIM, the proposed detection is used. For SM, sphere decoding is used. It can be seen that, (22,16)-GSIM scheme outperforms (16,16)-SM scheme using same RF resources by about 2 dB in the medium-to-high SNR regime by using six extra transmit antennas. Moreover, the proposed detection has much lower complexity than SD.

In Figs. 6 and 7, we also observe that, at low SNRs, the SM schemes have better BER performance compared with the corresponding GSIM schemes. This can be explained as follows. First, it can be observed that, to achieve the same rate, GSIM needs smaller constellation compared with SM. Hence, GSIM will have a larger minimum distance among the constellation points than that in SM. Second, unlike in SM where there are no antenna index bits, the following two types of error events are observed in GSIM: 1) the antenna activity pattern itself is decoded wrongly, and thus both the antenna index bits and modulation symbol bits are incorrectly decoded; and 2) the antenna activity pattern is decoded correctly, but the modulation symbol bits are wrongly decoded. At medium-to-high SNRs, the error event of the second type is more likely to occur; therefore, this type of error events dominates the resulting performance. Coupled with this, a larger minimum distance among constellation points in GSIM than that in the corresponding SM makes GSIM to outperform SM in medium-to-high SNRs. However, at low SNRs, the error event of the first type is more likely to occur, and this error event type dominates the resulting performance. Since there are no antenna index bits in SM, error events of the first type do not occur in SM, leading to better performance for SM in the low-SNR regime.

III. Generalized Space–Frequency Index Modulation

Here, we propose a GSFIM scheme that encodes bits through indexing in both spatial and frequency domains. GSFIM can be viewed as a generalization of the GSIM scheme presented earlier by exploiting indexing in the frequency domain as well. In the proposed GSFIM scheme, information bits are mapped through antenna indexing in the spatial domain, frequency indexing in the frequency domain, and M-ary modulation. After mapping, the signal is modulated using OFDM and is transmitted through the selected antennas. We obtain the rate equation for the proposed GSFIM system and study its achievable rate, rate variation as a function of the parameters involved, and the rate gain compared with conventional MIMO-OFDM.

A. System Model

The proposed GSFIM system uses n_t transmit antennas, $n_{\rm rf}$ transmit RF chains, $1 \le n_{\rm rf} \le n_t, N$ subcarriers, and n_r receive antennas. The channel between each transmit and receive antenna pair is assumed to be frequency-selective fading with L multipaths. The block diagrams of the GSFIM transmitter and receiver are shown in Fig. 8. At any given time, only $n_{\rm rf}$ transmit antennas are active, and the remaining $n_t - n_{\rm rf}$ antennas

[1]Since $n_r = 8$, the (12,12)-SM system is an underdetermined system. Therefore, we have used the GSD in [24], which achieves ML detection in such underdetermined systems. GSD for spatial modulation has been reported in [25].

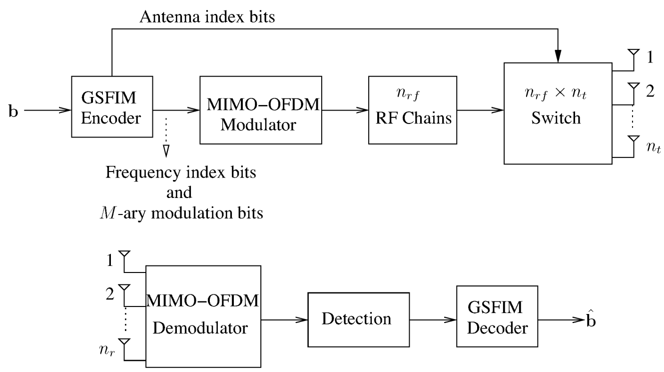


Fig. 8. Block diagram of GSFIM transmitter and receiver.

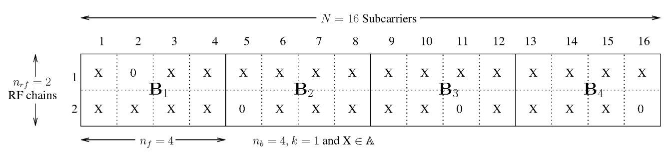


Fig. 9. Frequency indexing in GSFIM.

remain silent. The GSFIM encoder takes $\lfloor \log_2 \binom{n_t}{n_{\mathrm{rf}}} \rfloor$ bits and maps to n_{rf} out of n_t transmit antennas (antenna index bits). It also takes additional bits to index subcarriers (frequency index bits) and bits for M-ary modulation symbols on subcarriers. The frequency and antenna indexing mechanisms are detailed as follows.

1) Frequency Indexing: Consider a matrix $\mathbf{B}$ of size $n_{\mathrm{rf}} \times N$ whose entries belong to $\mathbb{A}_0$, where $\mathbb{A}_0 = \mathbb{A} \cup 0$ with $\mathbb{A}$ denoting an M-ary modulation alphabet. The frequency index bits and M-ary modulation bits are embedded in $\mathbf{B}$ as follows. The matrix $\mathbf{B}$ is divided into n_b submatrices $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_{n_b}$, each of size $n_{\mathrm{rf}} \times n_f$, where $n_f = (N)/(n_b)$ is the number subcarriers per submatrix (see Fig. 9). Let k, $1 \le k \le n_{\mathrm{rf}} n_f$, denote the number of nonzero elements in each submatrix, where each of the nonzero elements belong to $\mathbb{A}$. This k is a design parameter. Then, for each submatrix, there are $l_f = \binom{n_{\mathrm{rf}} n_f}{k}$ possible "frequency activation patterns." A frequency activation pattern for a given submatrix refers to a possible combination of zero and nonzero entries in that submatrix. Note that not all l_f activation patterns are needed for frequency indexing. Any 2^{k_f} patterns out of them, where $k_f = \lfloor \log_2 \binom{n_{\mathrm{rf}} n_f}{k} \rfloor$, are adequate. Take any 2^{k_f} patterns out of l_f patterns and form a set called the "frequency activation pattern set," denoted by $\mathbb{S}_f$. The frequency activation pattern for a given submatrix is then formed by choosing one among the patterns in the set $\mathbb{S}_f$ using k_f bits. These k_f bits are the frequency index bits for that submatrix. Therefore, there are a total of $n_b k_f$ frequency index bits in the entire matrix $\mathbf{B}$. In addition to these frequency index bits, $k n_b \log_2 M$ bits are carried as M-ary modulation bits in the nonzero entries of $\mathbf{B}$.

Example: Let us illustrate this using the following example. Let $n_{\mathrm{rf}} = 2$, $N = 16$, $n_b = 4$, and $k = 7$. Then, $n_f = (16/4) = 4$, $l_f = \binom{8}{7} = 8$, $k_f = \lfloor \log_2 8 \rfloor = 3$, and $2^{k_f} = 8$. In this example, $l_f = 2^{k_f} = 8$, i.e., all the eight possible patterns are in the frequency activation pattern set, given by

$$\mathbb{S}_f = \left\{ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \right.$$
$$\times \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$
$$\times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$
$$\left. \times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \right\}.$$

Suppose $\mathbb{A}$ is 4-QAM. Let [00101001111000110] denote the information bit sequence for submatrix $\mathbf{B}_1$. The GSFIM encoder translates these bits to the submatrix $\mathbf{B}_1$ as follows: The first three bits are used to choose the frequency activity pattern (i.e., 001 chooses the activation pattern $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ in the set $\mathbb{S}_f$ above), and the next 14 bits are mapped to seven 4-QAM symbols so that one 4-QAM symbol gets mapped to one active subcarrier. The submatrix $\mathbf{B}_1$ then becomes

$$\mathbf{B}_1 = \begin{bmatrix} -1 - \mathbf{j} & 0 & -1 + \mathbf{j} & 1 - \mathbf{j} \\ 1 - \mathbf{j} & -1 + \mathbf{j} & -1 - \mathbf{j} & 1 + \mathbf{j} \end{bmatrix}$$

where $\mathbf{j} = \sqrt{-1}$. Likewise, the submatrices $\mathbf{B}_i$, $i = 2, 3, 4$ are formed. The full matrix $\mathbf{B}$ of size $n_{\mathrm{rf}} \times N$ is then formed as

$$\mathbf{B} = [\mathbf{B}_1\ \mathbf{B}_2\ \mathbf{B}_3\ \mathbf{B}_4].$$

Each row of the matrix $\mathbf{B}$ is of dimension $1 \times N$. There are n_{rf} rows. Each N-length row vector in $\mathbf{B}$ is fed to the IFFT block in the OFDM modulator to generate an N-length OFDM symbol. A total of n_{rf} such OFDM symbols, i.e., one for each row in $\mathbf{B}$, are generated. These n_{rf} OFDM symbols are then transmitted through n_{rf} active transmit antennas in parallel. The choice of these n_{rf} active transmit antennas among the n_t available antennas is made through antenna indexing as described in the following.

2) Antenna Indexing: The selection of n_{rf} out of n_t antennas for transmission is made based on antenna index bits. The antenna index bits choose an "antenna activation pattern," which tells which n_{rf} antennas out of n_t antennas are used for transmission. There are $l_a = \binom{n_t}{n_{\mathrm{rf}}}$ antenna activation patterns possible, and $k_a = \lfloor \log_2 \binom{n_t}{n_{\mathrm{rf}}} \rfloor$ bits are used to choose one among them. These k_a bits are the antenna index bits. Note that not all l_a activation patterns are needed, and any 2^{k_a} patterns out of them are adequate. Take any 2^{k_a} patterns out of l_a patterns and form a set called the "antenna activation pattern set," which is denoted by $\mathbb{S}_a$.

Example: Let us illustrate this using the following example. Let $n_t = 3$, $n_{\mathrm{rf}} = 2$. Then, $l_a = \binom{3}{2} = 3$, $k_a = \lfloor \log_2 \binom{3}{2} \rfloor = \lfloor \log_2 3 \rfloor = 1$, and $2^{k_a} = 2$. The possible antenna activation patterns are given by $\{[1, 1, 0]^T, [1, 0, 1]^T, [0, 1, 1]^T\}$. The set $\mathbb{S}_a$ is formed by selecting any two patterns out of the above three patterns. For example, $\mathbb{S}_a$ can be

$$\mathbb{S}_a = \left\{ [1, 1, 0]^T, [1, 0, 1]^T \right\}.$$

An $n_{\mathrm{rf}} \times n_t$ switch connects the transmit RF chains to the transmit antennas. The chosen n_{rf} out of n_t transmit antennas transmit the MIMO-OFDM symbol constructed using the frequency index bits and M-ary modulation bits. The active transmit antennas can change from one MIMO-OFDM symbol to the other.

B. Achievable Rate, Rate Variation, and Rate Gain

In GSFIM, the information bits are encoded using: 1) frequency indexing over each submatrix $\mathbf{B}_i$, $i = 1, 2, \ldots, n_b$; 2) M-ary modulation symbols in each submatrix; and 3) antenna indexing. The number of frequency indexing bits per submatrix is $\lfloor \log_2 \binom{n_{\mathrm{rf}} n_f}{k} \rfloor$. The number of M-ary modulation bits in each submatrix is $k \log_2 M$. The number of antenna indexing bits is $\lfloor \log_2 \binom{n_t}{n_{\mathrm{rf}}} \rfloor$. Combining these three parts, the achievable rate in GSFIM with n_t transmit antennas, n_{rf} transmit RF chains, N subcarriers, n_b submatrices, and M-ary modulation is given by

$$R_{\mathrm{gsfim}} = \underbrace{\left(\frac{\left\lfloor \log_2 \binom{n_t}{n_{\mathrm{rf}}} \right\rfloor}{N + L - 1} \right)}_{R_{\mathrm{A}}} + \underbrace{\left(\frac{\left\lfloor \log_2 \binom{n_{\mathrm{rf}} n_f}{k} \right\rfloor n_b}{N + L - 1} \right)}_{R_{\mathrm{F}}} + \underbrace{\left(\frac{k n_b \log_2 M}{N + L - 1} \right)}_{R_{\mathrm{Q}}} \quad \text{bpcu}. \tag{34}$$

Note that, in a conventional MIMO-OFDM system, there is no contribution to the rate by antenna or frequency indexing, and the achieved rate is only through M-ary modulation symbols. Moreover, in MIMO-OFDM, M-ary modulation symbols are mounted on all N subcarriers on each of the n_{rf} active transmit antennas. Therefore, the achieved rate in MIMO-OFDM (with no antenna and frequency indexing) for the same parameters as in GSFIM is given by

$$R_{\mathrm{mimo-ofdm}} = \left(\frac{1}{N + L - 1} \right) n_{\mathrm{rf}} N \log_2 M \quad \text{bpcu}. \tag{35}$$

From (34) and (35), we can make the following observations.

- Conventional MIMO-OFDM becomes a special case of GSFIM for $n_{\mathrm{rf}} = n_t$, $n_f = N$ (i.e., $n_b = 1$).
- GSIM presented in Section II becomes a special case of GSFIM for $N = n_f = n_b = 1$, $k = n_{\mathrm{rf}}$.
- For $n_{\mathrm{rf}} < n_t$, $R_{\mathrm{A}} > 0$, which is the additional rate contributed by antenna indexing. In this case, R_{gsfim} in (34) can be more or less compared with $R_{\mathrm{mimo-ofdm}}$ depending on the choice of parameters. For example, the parameter k can take values in the range 1 to $n_{\mathrm{rf}} n_f$. An instance where R_{gsfim} is more than $R_{\mathrm{mimo-ofdm}}$ happens when $k = n_{\mathrm{rf}} n_f$; in which case, $R_{\mathrm{F}} = 0$, and $R_{\mathrm{Q}} = R_{\mathrm{mimo-ofdm}}$. Therefore, R_{A} is the excess rate (rate gain) in GSFIM compared with MIMO-OFDM. Likewise, an instance where R_{gsfim} is less than $R_{\mathrm{mimo-ofdm}}$ happens when $k = 1$, in which case, R_{gsfim} becomes $((n_b \log_2(n_{\mathrm{rf}} n_f M))/(N + L - 1))$, which is less than $R_{\mathrm{mimo-ofdm}}$ given by $((n_b n_{\mathrm{rf}} n_f \log_2(M))/(N + L - 1))$.
- The sum of rates R_{F} and R_{Q} in (34) as a function of k reaches its maximum for a value of k in the range$\lfloor n_{\mathrm{rf}} n_f / 2 \rfloor$ and $n_{\mathrm{rf}} n_f$ as does the total rate R_{gsfim}. The maximum R_{gsfim} will be more than or equal to $R_{\mathrm{mimo-ofdm}}$.

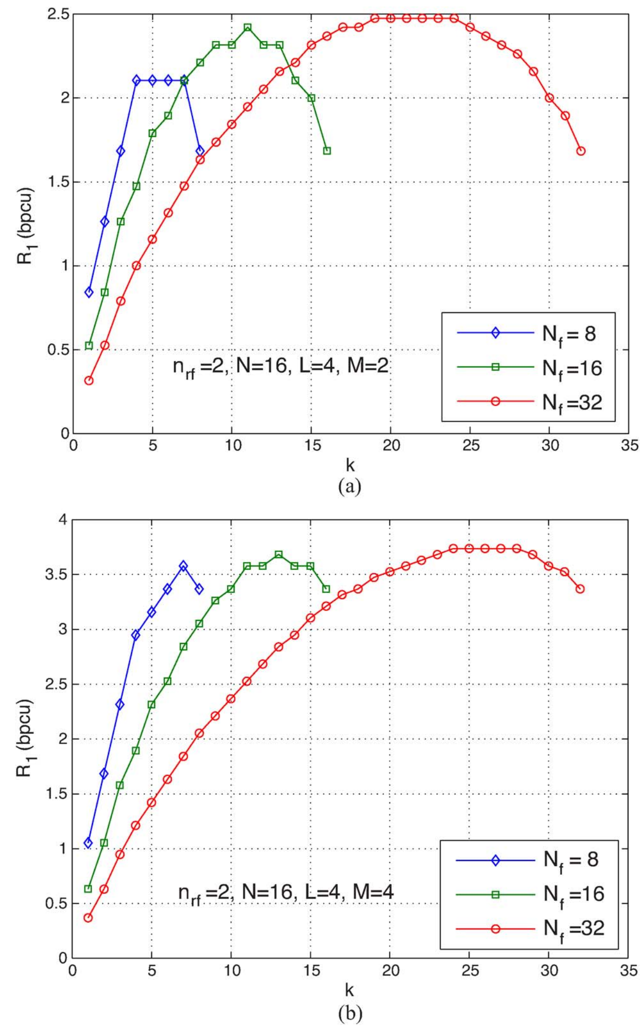


Fig. 10. Rate $R_1 = R_F + R_Q$ as a function of k for different values of $N_f = n_{\mathrm{rf}} n_f$. (a) $M = 2$. (b) $M = 4$.

We now illustrate the above observations through numerical results. Define $R_1 \triangleq R_{\mathrm{F}} + R_{\mathrm{Q}}$ and $N_f \triangleq n_{\mathrm{rf}} n_f$. In Fig. 10, we plot R_1 as a function of k, for different values of $N_f = 8, 16, 32$, $L = 4$, and $M = 2, 4$, We observe that R_1 reaches its maximum value for k between $\lfloor N_f/2 \rfloor$ and N_f. Moreover, the maximum R_1 increases as N_f increases because the R_{F} term in (34) increases with N_f.

In Fig. 11, we plot the maximum R_{gsfim} as a function of n_t for $n_{\mathrm{rf}} = 8$, $N = 32$, $L = 4$, and $n_f = 1, 2, 4, 8, 16, 32$. $R_{\mathrm{mimo-ofdm}}$ is also plotted for comparison. We observe that, for a given n_f, the maximum R_{gsfim} increases with n_t because of the increase in antenna index bits carried. For a given n_t

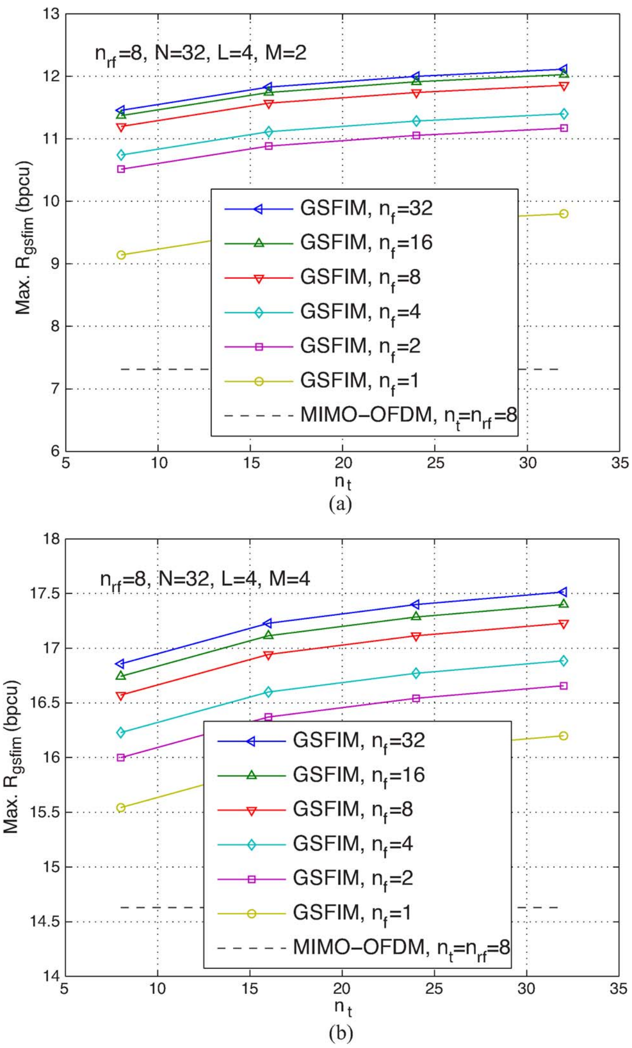


Fig. 11. Maximum R_{gsfim} as a function of n_t, for $n_{\text{rf}} = 8$ and different values of n_f. (a) $M = 2$. (b) $M = 4$.

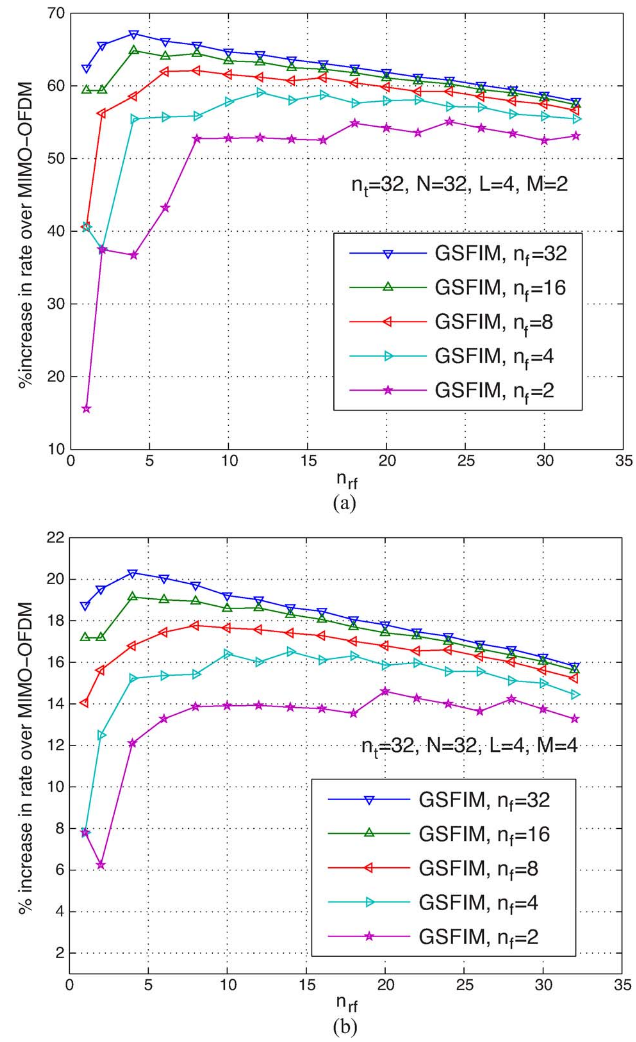


Fig. 12. Percentage rate gain in GSFIM compared with MIMO-OFDM as a function of n_{rf} and n_f. (a) $M = 2$. (b) $M = 4$.

and n_{rf}, the maximum R_{gsfim} increases with increase in n_f because of increase in N_f and the associated increase in R_{F}. From this figure, we can see that GSFIM can achieve a rate gain of up to 65% for $M = 2$ and up to 19% for $M = 4$, compared with MIMO-OFDM. In Fig. 12, we have plotted the percentage rate gain in GSFIM compared with MIMO-OFDM (i.e., difference between maximum R_{gsfim} and $R_{\text{mimo-ofdm}}$ in percentage), as a function of n_{rf} for $n_t = 32$, $N = 32$, $L = 4$, and $n_f = 2, 4, 8, 16, 32$. As can be observed in Fig. 12, GSFIM can achieve rate gains up to 65% for $M = 2$ and 20% for $M = 4$, compared with MIMO-OFDM.

In Fig. 13, we plot the maximum R_{gsfim} as a function of n_{rf} for a given $n_t = 32$, $N = 32$, $L = 4$, and $n_f = 1, 32$. We can observe that, for a given n_f, the rate increases with n_{rf} because of the increase in R_{F}. For a given n_{rf}, the maximum R_{gsfim} increases with the increase in n_f. In Fig. 14, we have plotted bar graphs showing the percentage savings in transmit RF chains in GSFIM compared with MIMO-OFDM $n_t = N = 32$, $L = 4$, and $n_f = 1, 4, 32$. It can be observed that this savings is high for small-sized modulation alphabets, e.g., the savings is up to 42% for $M = 2$ and 20% for $M = 4$.

In Fig. 15, we plot the maximum R_{gsfim} as a function of n_f for $n_t = 32$, $n_{\text{rf}} = 8$, $L = 4$, and $N = 32$. We can observe that the maximum R_{gsfim} increases for up to certain n_f, and thereafter, it saturates. This is because the maximum R_1 saturates to a value $((N_f n_b \log_2(M+1))/(N+L-1))$ for large N_f.

C. GSFIM Signal Detection and Performance

Here, we consider GSFIM signal detection and performance. Let $\mathbf{H}_n$ denote $n_r \times n_t$ channel matrix on subcarrier n. Let $\mathbf{H}_n^{\mathbf{a}}$ denote the $n_r \times n_{\text{rf}}$ channel matrix corresponding to the chosen n_{rf} antennas. The superscript $\mathbf{a}$ in $\mathbf{H}_n^{\mathbf{a}}$ refers to the antenna activation pattern that tells which n_{rf} antennas are chosen. Let us denote the $n_r \times 1$-sized received vector on subcarrier n as $\mathbf{y}_n$, which can be written as

$$\mathbf{y}_n = \mathbf{H}_n^{\mathbf{a}} \mathbf{z}_n + \mathbf{w}_n, \quad n = 1, 2, \ldots, N \tag{36}$$

where $\mathbf{z}_n$ is the $n_{\text{rf}} \times 1$-sized transmitted vector on subcarrier n, and $\mathbf{w}_n$ is the $n_r \times 1$-sized additive white Gaussian noise

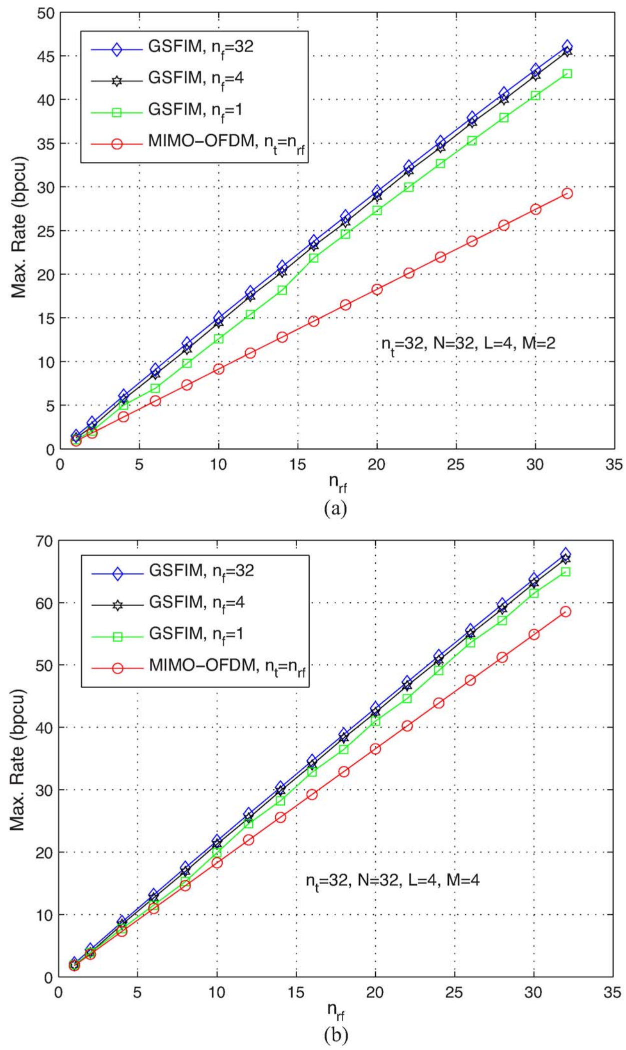


Fig. 13. Maximum $R_{\rm gsfim}$ as a function of $n_{\rm rf}$, for $n_t = N = 32$, and $n_f = 1, 4, 32$. (a) $M = 2$. (b) $M = 4$.

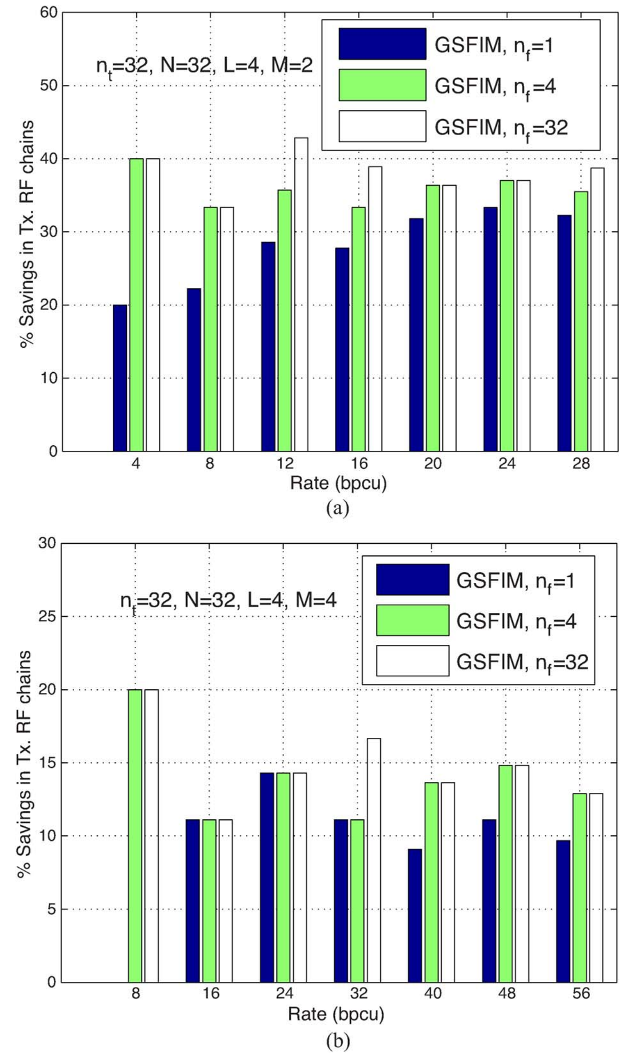


Fig. 14. Percentage savings in number of transmit RF chains in GSFIM compared with MIMO-OFDM, for $n_t = N = 32$, $n_f = 1, 4, 32$. (a) $M = 2$. (b) $M = 4$.

vector at the receiver $\mathbf{w}_n \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{n_r})$. Consider the system model in (36) for the ith submatrix, which is given by

$$\mathbf{y}_l = \mathbf{H}_l^{\mathbf{a}} \mathbf{z}_l + \mathbf{w}_l, \quad l = i_1, i_2, \ldots, i_j, \ldots, i_{n_f} \tag{37}$$

where $i_j = (i-1)n_f + j$. Write (37) as

$$\mathbf{y}^i = \mathbf{G}_i^{\mathbf{a}} \mathbf{z}^i + \mathbf{w}^i, \quad i = 1, 2, \ldots, n_b \tag{38}$$

where

$$\mathbf{y}^i = \begin{bmatrix} \mathbf{y}_{i_1} \\ \mathbf{y}_{i_2} \\ \vdots \\ \mathbf{y}_{i_{n_f}} \end{bmatrix} \quad \mathbf{z}^i = \begin{bmatrix} \mathbf{z}_{i_1} \\ \mathbf{z}_{i_2} \\ \vdots \\ \mathbf{z}_{i_{n_f}} \end{bmatrix}$$

$$\mathbf{G}_i^{\mathbf{a}} = \begin{bmatrix} \mathbf{H}_{i_1}^{\mathbf{a}} & & & 0 \\ & \mathbf{H}_{i_2}^{\mathbf{a}} & & \\ & & \ddots & \\ 0 & & & \mathbf{H}_{i_{n_f}}^{\mathbf{a}} \end{bmatrix}.$$

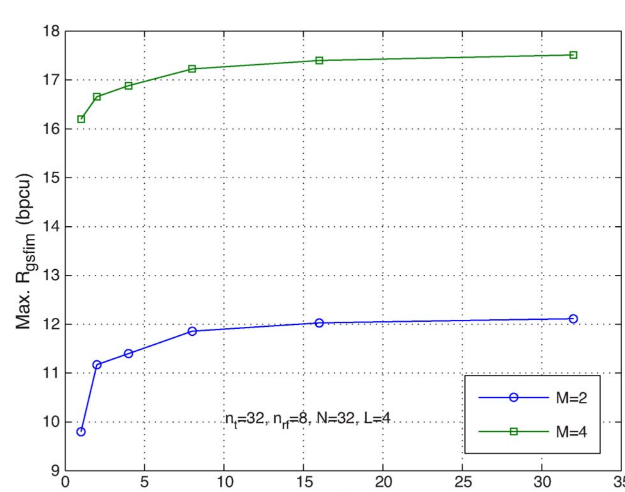


Fig. 15. Maximum $R_{\rm gsfim}$ as a function of n_f, for fixed n_t, $n_{\rm rf}$.

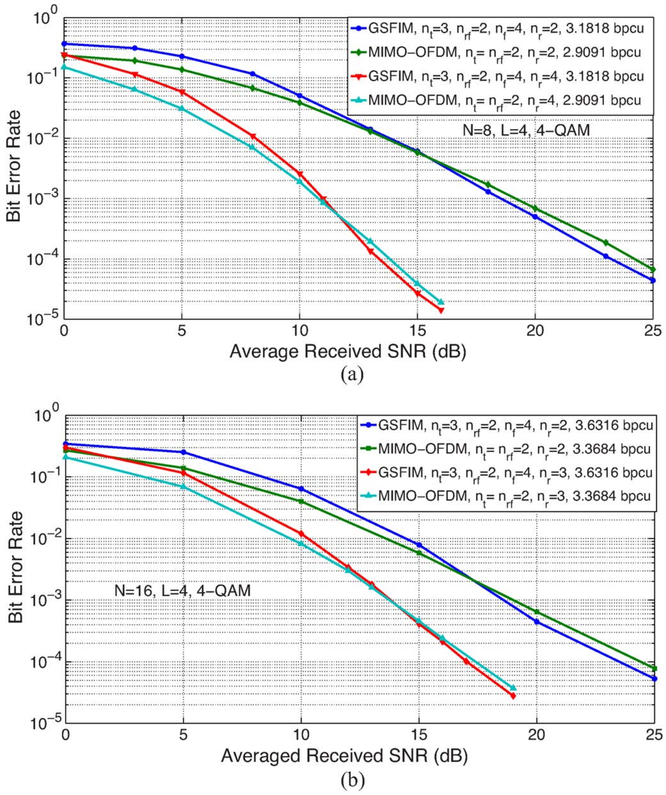


Fig. 16. BER performance of GSFIM and MIMO-OFDM under ML detection. (a) GSFIM with $n_t = 3$, $n_{\rm rf} = 2$, $N = 8$, $n_f = 4$, $n_r = 2, 4$, $L = 4$, 4-QAM, and 3.1818 bpcu, and MIMO-OFDM with $n_t = n_{\rm rf} = 2$, $N = 8$, $n_r = 2, 4$, $L = 4$, 4-QAM, and 2.9091 bpcu. (b) GSFIM with $n_t = 3$, $n_{\rm rf} = 2$, $N = 16$, $n_f = 4$, $n_r = 2, 3$, $L = 4$, 4-QAM, and 3.6316 bpcu, and MIMO-OFDM with $n_t = n_{\rm rf} = 16$, $N = 16$, $n_r = 2, 3$, $L = 4$, 4-QAM, and 3.3684 bpcu. (a) $N = 8$. (b) $N = 16$.

The ML metric for a given antenna activation pattern $\mathbf{a}$ and vectors $\mathbf{z}^i, i = 1, \ldots, n_b$ representing the frequency activation pattern and M-ary modulation bits is

$$d(\mathbf{a}, \mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n_b}) = \sum_{i=1}^{n_b} \|\mathbf{y}^i - \mathbf{G}_i^{\mathbf{a}} \mathbf{z}^i\|^2. \tag{39}$$

Let $\mathbb{U}$ denote the set of all possible N_f-length transmit vectors corresponding to a submatrix. Then, $\mathbb{U}$ is given by

$$\mathbb{U} = \left\{ \mathbf{x} | \mathbf{x} \in \mathbb{A}_0^{N_f \times 1}, \|\mathbf{x}\|_0 = k, \mathbf{t}^{\mathbf{x}} \in \mathbb{S}_f \right\} \tag{40}$$

where $\mathbf{t}^{\mathbf{x}}$ denotes the frequency activity pattern corresponding to $\mathbf{x}$, where $t_j^{\mathbf{x}} = 1, \text{iff} x_j \neq 0, \forall j = 1, 2, \ldots, N_f$. The antenna activation and frequency activation pattern sets ($\mathbb{S}_a$, $\mathbb{S}_f$), and the antenna and frequency index bit maps are known at both transmitter and receiver. Therefore, from (39) and (40), the ML decision rule for GSFIM signal detection is given by

$$(\hat{\mathbf{a}}, \hat{\mathbf{z}}^1, \hat{\mathbf{z}}^2, \ldots, \hat{\mathbf{z}}^{n_b}) = \mathop{\arg\min}_{\mathbf{a} \in \mathbb{S}_a, \mathbf{z}^i \in \mathbb{U}, \forall i} d(\mathbf{a}, \mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n_b}). \tag{41}$$

By inverse mapping, the antenna index bits are recovered from $\hat{\mathbf{a}}$, and the frequency index bits are recovered from $\hat{\mathbf{z}}^1, \hat{\mathbf{z}}^2, \ldots, \hat{\mathbf{z}}^{n_b}$.

In Fig. 16(a) and (b), we show the BER performance of GSFIM in comparison with MIMO-OFDM under ML detection. In Fig. 16(a), the GSFIM system has $n_t = 3$, $n_{\rm rf} = 2$, $N = 8$, $n_f = 4$, $n_r = 2, 4$, 4-QAM, and the achieved rate is $R_{\rm gsfim} = 3.1818$ bpcu. The MIMO-OFDM has $n_t = n_{\rm rf} = 2$, $N = 8$, $n_r = 2, 4$, 4-QAM, and the achieved rate is $R_{\rm mimo-ofdm} = 2.9091$ bpcu. In Fig. 16(b), the GSFIM system has $n_t = 3$, $n_{\rm rf} = 2$, $N = 16$, $n_f = 4$, $n_r = 2, 3$, $L = 4$, 4-QAM, and the achieved rate is $R_{\rm gsfim} = 3.6316$ bpcu. The MIMO-OFDM system has $n_t = n_{\rm rf} = 2$, $N = 16$, $n_r = 2, 3$, $L = 4$, 4-QAM, and the achieved rate is $R_{\rm mimo-ofdm} = 3.3684$ bpcu. It is seen that in Fig. 16(a) and (b), GSFIM has higher rates than MIMO-OFDM. In terms of error performance, while MIMO-OFDM performs better at low SNRs, GSFIM performs better at moderate to high SNRs. This performance crossover can be explained in the same way as explained in the case of GSIM earlier (see Section II-E and Figs. 6 and 7), i.e., at moderate-to-high SNRs, errors in index bits are less likely, and this makes GSFIM perform better; at low SNRs, index bits and hence the associated modulation bits are more likely to be in error, making MIMO-OFDM to perform better. Similar performance crossovers have been reported in the literature for single-antenna OFDM with/without subcarrier indexing (e.g., in [12]), where it has been shown that OFDM with subcarrier indexing outperforms classical OFDM without subcarrier indexing at moderate-to-high SNRs, whereas classical OFDM outperforms OFDM with subcarrier indexing at low SNRs. The plots in Fig. 16(a) and (b) essentially capture a similar phenomenon when there are index bits both in frequency and spatial domains.

IV. Conclusion

We introduced IM, where information bits are encoded in the indexes of the active antennas (spatial domain) and subcarriers (frequency domain), in addition to conveying information bits through conventional modulation symbols. For GSIM, where bits are indexed only in the spatial domain, we derived the expression for achievable rate and easy-to-compute upper and lower bounds on this rate. We showed that the achievable rate in GSIM can be more than that in spatial multiplexing and analytically established the condition under which this can happen. We also proposed a Gibbs-sampling-based detection algorithm for GSIM and showed that GSIM can achieve better BER performance than spatial multiplexing. GSIM achieved this better performance using fewer transmit RF chains compared with spatial multiplexing. For GSFIM, where bits are encoded in the indexes of both active antennas and subcarriers, we derived the achievable rate expression. Numerical results showed that GSFIM can achieve higher rates compared with conventional MIMO-OFDM. Moreover, BER results using ML detection showed the potential for GSFIM performing better than MIMO-OFDM at moderate-to-high SNRs. Low-complexity detection methods for GSFIM can be taken up for future extension to this work.

Acknowledgment

The authors would like to thank T. L. Narasimhan and B. Chakrapani for their valuable contributions to the discussions on IM techniques.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[2] H. Bölcskei, D. Gesbert, C. B. Papadias, and Alle-Jan van der Veen, Eds. *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[3] A. Chockalingam and B. Sundar Rajan, *Large MIMO Systems*. Cambridge, U.K.: Cambridge Univ. Press, Feb. 2014.

[4] A. Mohammadi and F. M. Ghannouchi, "Single RF front-end MIMO transceivers," *IEEE Commun. Mag.*, vol. 50, no. 12, pp. 104–109, Dec. 2011.

[5] R. Mesleh, H. Haas, S. Sinaovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[6] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 50, no. 12, pp. 182–191, Dec. 2011.

[7] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[8] N. Serafimovskil, S. Sinanovic, M. Di Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, 2012, p. 299.

[9] T. L. Narasimhan, P. Raviteja, and A. Chockalingam, "Large-scale multiuser SM-MIMO versus massive MIMO," in *Proc. ITA*, *San Diego, CA, USA*, Feb. 2014, pp. 1–9.

[10] R. Abu-alhiga and H. Haas, "Subcarrier index modulation OFDM," in *Proc. IEEE PIMRC*, Sep. 2009, pp. 177–181.

[11] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced Subcarrier Index Modulation (SIM) OFDM," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 728–732.

[12] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with indexing," in *Proc. IEEE GLOBECOM*, Dec. 2012, pp. 4741–4746.

[13] Y. Xiao *et al.*, "OFDM with interleaved subcarrier-index modulation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1447–1450, Aug. 2014.

[14] T. L. Narasimhan, Y. Naresh, T. Datta, and A. Chockalingam, "Pseudo-random phase precoded spatial modulation and precoder index modulation," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 3868–3873.

[15] M. Di Renzo andH. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.

[16] M. Di Renzo and H. Haas, "On transmit-diversity for spatial modulation MIMO: Impact of spatial-constellation diagram and shaping filters at the transmitter," *IEEE Trans. Veh. Technol.*, vol. 62, no. 6, pp. 2507–2531, Jul. 2013.

[17] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.

[18] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1498–1502.

[19] J. Fu, C. Hou, W. Xiang, L. Yan, and Y. Hou, "Generalised spatial modulation with multiple active transmit antennas," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 839–844.

[20] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605–1615, Apr. 2012.

[21] T. Datta and A. Chockalingam, "On generalized spatial modulation," in *Proc. IEEE WCNC*, Apr. 2013, pp. 2716–2721.

[22] B. Schmuland, "Factorials!", Jul. 26, 2013. [Online]. Available: www.stat.ualberta.ca/people/schmu/preprints/factorial.pdf

[23] T. Datta, N. A. Kumar, A. Chockalingam, and B. S. Rajan, "A novel Monte Carlo sampling based receiver for large-scale uplink multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3019–3038, Sep. 2013.

[24] P. Wang and T. Le-Ngoc, "A low-complexity generalized sphere decoding approach for underdetermined linear communication systems: Performance and complexity evaluation," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3376–3388, Nov. 2009.

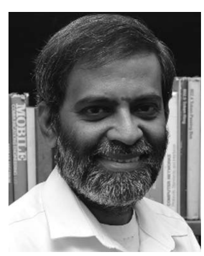
[25] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh, and H. Haas, "Generalized sphere decoding for spatial modulation," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2805–2815, Jul. 2013.

Tanumay Datta received the B.E. degree in electronics and telecommunication engineering from Jadavpur University, Kolkata, India, in 2008 and the Ph.D. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 2014.

From July 2013 to March 2015, he was with Broadcom Communication Technologies, Bangalore, as Scientist Staff II in the Systems Engineering Group. He is currently working toward post-doctoral research with Centrale Supelec, Gif sur Yvette, France. His research interests include low-complexity near-maximum-likelihood detection algorithms for large multiple-input–multiple-output and multiuser systems, spatial modulation, multimode fiber communications, and wireless relay networks.

Harsha S. Eshwaraiah received the B.E. degree in electronics and communication engineering from Jagadguru Sri Shivarathreeshwara Academy of Technical Education, Bangalore, India, in 2007. He received the M.Sc. (Engg.) degree in electrical communication engineering, Indian Institute of Science, Bangalore, India, in 2014.

From 2007 to 2009, he was a Lecturer with the Department of Telecommunications Engineering, Siddaganga Institute of Technology, Tumkur, India. His current research interests include detection algorithms for single-carrier (SC) communications, SC frequency-division multiple access, and spatial modulation.



A. Chockalingam (S'92–M'93–SM'98) was born in Rajapalayam, Tamil Nadu, India. He received the B.E. (Honors) degree in electronics and communication engineering from the PSG College of Technology, Coimbatore, India, in 1984; the M.Tech. degree in electronics and electrical communications engineering (with specialization in satellite communications) from the Indian Institute of Technology, Kharagpur, India, in 1985; and the Ph.D. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1993.

From 1986 to 1993, he worked with the Transmission R&D Division, Indian Telephone Industries Limited, Bangalore. From December 1993 to May 1996, he was a Postdoctoral Fellow and an Assistant Project Scientist with the Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA. From May 1996 to December 1998, he was with Qualcomm, Inc., San Diego, as a Staff Engineer/Manager in the Systems Engineering Group. Since December 1998, he has been with the faculty of the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, where he is currently a Professor, working in the area of wireless communications and networking.

Dr. Chockalingam served as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on Multiuser Detection for Advanced Communication Systems and Networks) and for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (Special Issue on Soft Detection on Wireless Transmission). He received the Swarnajayanti Fellowship from the Department of Science and Technology, Government of India. He is a Fellow of the Indian National Academy of Engineering; the National Academy of Sciences, India; the Indian National Science Academy; and the Indian Academy of Sciences.