Capacity Analysis and Structured Sparse Detection of Generalized Media-based Modulation

Bharath Shamasundar and A. Chockalingam
Department of ECE, Indian Institute of Science, Bangalore 560012

***Abstract*—Media-based modulation (MBM) is an attractive channel modulation scheme with rate, performance, and hardware advantages. In this paper, we are concerned with two important aspects of MBM. The first one is on the capacity of MBM and the other is on low-complexity detection of high-rate MBM signals using structured spare recovery techniques. We derive closed-form expression for the capacity of generalized MBM (GMBM). The main idea in the capacity analysis is to recognize that MBM uses two alphabets to convey information, namely, the source alphabet (e.g., QAM/PSK) and the channel alphabet (fade coefficients). This observation allows us to show that the capacity is achieved when the mutual information is maximized over the source-channel product alphabet. We then propose a greedy structured sparse recovery algorithm for the detection of GMBM signal vectors. The proposed algorithm is a two-stage algorithm in which support recovery is done in greedy manner first, followed by data detection in the non-zero positions. Simulation results show that the proposed algorithm achieves good performance at low complexity even when the system is highly underdetermined.**

***Index Terms*—Media-based modulation, channel alphabet, capacity, structured sparsity, signal detection.**

I. Introduction

Media-based modulation (MBM) is a channel modulation scheme in which a collection of channel fade coefficients/vectors is used as a channel modulation alphabet, in addition to a conventional modulation alphabet such as QAM/PSK [1]-[7]. The set of fade coefficients/vectors that form the channel modulation alphabet is created using radio frequency (RF) mirrors placed in the near field of the transmit antenna. These RF mirrors act as digitally controlled scatterers. Each RF mirror either reflects (ON state) or allows (OFF state) the RF signal radiated by the transmit antenna to pass through, depending on the digital input given to it. The digital control inputs, in turn, are decided based on information bits to be transmitted. Suppose there are m_{rf} RF mirrors at the transmitter, then there are $N_m \triangleq 2^{m_{rf}}$ different ON/OFF combinations, called 'mirror activation patterns' (MAP). Each of these MAPs creates a different near field geometry for the transmitted signal. In a rich scattering environment, even a small perturbation in the near field will be augmented by random reflections. This leads to different end-to-end fade realizations for different MAPs. If n_r is the number of receive antennas, then each of the N_m channel realizations is an $n_r \times 1$ vector of channel coefficients. The collection of these N_m fade vectors form the channel alphabet, denoted by $\mathcal{H}$. The transmitter activates one of the MAPs (or equivalently selects one of the N_m channels from $\mathcal{H}$) based on m_{rf} information bits. Also, a symbol from a source alphabet $\mathbb{A}$ (e.g., QAM/PSK) is transmitted by the transmit antenna, which conveys $M \triangleq \log_2 |\mathbb{A}|$ bits. Therefore, MBM uses two alphabets to convey information bits, namely, the source alphabet $\mathbb{A}$ with M symbols and the channel alphabet $\mathcal{H}$ with N_m symbols. The achieved rate in MBM is, therefore, given by $\eta_{\text{MBM}} = m_{rf} + \log_2 M$ bits per channel use (bpcu). MBM has been shown to achieve good bit error performance in both the point-to-point and multiuser settings [1]-[7]. Also, MBM has the advantage of RF hardware simplicity because it needs only a single RF chain and RF mirrors can be realized using simple varactors/PIN diodes.

Generalized MBM (GMBM) [4] is a scheme in which MBM is used along with generalized spatial modulation (GSM). In GSM [8],[9], there are n_t antennas and n_{rf} transmit RF chains. In each channel use, n_{rf} out of n_t antennas are activated, and which n_{rf} antennas are active conveys information bits, in addition to the bits conveyed by n_{rf} symbols from $\mathbb{A}$ sent on the n_{rf} activated antennas. The achieved rate in GSM is therefore given by $\eta_{\text{GSM}} = \lfloor \log_2 \binom{n_t}{n_{rf}} \rfloor + n_{rf} \log_2 M$ bpcu.

In a GMBM system, there are n_{tu} transmit units (TU) and n_{rf} transmit RF chains, and each TU consists of one transmit antenna and m_{rf} RF mirrors (see Fig. 1). In a given channel use, information bits are conveyed through 1) spatial indexing of TUs (i.e., by activating n_{rf} out of n_{tu} TUs), 2) m_{rf} RF mirrors in each active TU, and 3) n_{rf} symbols from a conventional modulation alphabet $\mathbb{A}$, resulting in the achieved rate of $\eta_{\text{GMBM}} = \lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor + n_{rf}(m_{rf} + \log_2 M)$ bpcu, where $M \triangleq \log_2 |\mathbb{A}|$. It has been shown that GMBM can achieve superior bit error performance compared to MBM and other index modulation schemes [4]. In terms of capacity, with only channel alphabet and no source alphabet (tone), [2] has shown that MBM achieves the capacity of n_r parallel AWGN channels asymptotically as $m_{rf} \to \infty$. This work does not consider capacity analysis in the non-asymptotic regime. Filling this gap is one of our new contributions in this paper. Specifically, we derive a closed-form expression for GMBM capacity, where the key idea is to recognize that MBM uses two alphabets to convey information, namely, the source alphabet and the channel alphabet, and that the capacity is achieved when the mutual information is maximized over the source-channel product alphabet.

Next, signal detection is another important issue in high-rate

This work was supported in part by the J. C. Bose National Fellowship, Department of Science and Technology, Government of India, and the Intel India Faculty Excellence Program.

GMBM systems. Algorithms based on message passing have been proposed for GMBM signal detection in the literature [12]. Sparse recovery techniques offer a promising approach to low-complexity signal detection [13]-[17]. Here, we take note of the fact GMBM signals are inherently sparse with an interesting structure, and exploit this structured sparsity for low-complexity detection of GMBM signals. Specifically, we propose a detection algorithm using greedy structured sparse recovery techniques. This forms the second contribution in this paper. We call the proposed algorithm as GMBM matching pursuit (GMBM-MP) algorithm. Simulation results show that the proposed GMBM-MP algorithm achieve good performance at low computational complexity even in highly underdetermined systems.

II. System model

In this section, we first present the GMBM transmitter structure and the transmission scheme. Then, we present the GMBM channel alphabet, signal set, and the input-output system model.

A. GMBM transmitter

Figure 1 shows the block diagram of a GMBM transmitter. An MBM transmit unit (MBM-TU) consists of a transmit antenna and m_{rf} RF mirrors placed near that transmit antenna. The GMBM transmitter consists of n_{tu} MBM-TUs and n_{rf} transmit RF chains. In a channel use, n_{rf} out of n_{tu} MBM-TUs are selected based on $\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor$ information bits. An $n_{rf} \times n_{tu}$ switch connects the n_{rf} RF chains to the transmit antennas of the n_{rf} selected MBM-TUs. The mirror activation pattern (MAP) for each active MBM-TU is selected based on m_{rf} information bits, and hence $n_{rf} m_{rf}$ information bits are conveyed by the MAPs of the n_{rf} active MBM-TUs. Further, on an active MBM-TU, a symbol from an M-ary source alphabet $\mathbb{A}$ (e.g., QAM/PSK) is transmitted, which conveys $\log_2 M$ bits per active MBM-TU. The achieved rate in GMBM is therefore given by

$$\eta_{\text{GMBM}} = \left\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \right\rfloor + n_{rf}(m_{rf} + \log_2 M) \quad \text{bpcu.} \tag{1}$$

B. GMBM channel alphabet

Before presenting the GMBM channel alphabet, we first present the MBM channel alphabet.

1) MBM channel alphabet: Let $\mathbf{h}_k^m$ denote the $n_r \times 1$ channel gain vector corresponding to mth MAP of the kth MBM-TU, where $\mathbf{h}_k^m = [h_{1,k}^m \; h_{2,k}^m \cdots h_{n_r,k}^m]^T$, $h_{i,k}^m$ is the channel gain from kth MBM-TU to ith receive antenna when mth MAP is used, $m = 1, \cdots, N_m$, $k = 1, \cdots, n_{tu}$, $i = 1, \cdots, n_r$, and $N_m \triangleq 2^{m_{rf}}$. The $h_{i,k}^m$s are assumed to be i.i.d and distributed $\mathcal{CN}(0,1)$. The MBM channel alphabet of the kth MBM-TU, denoted by $\mathcal{H}_k$, is the set of N_m channel gain vectors, $\mathcal{H}_k = \{\mathbf{h}_k^1, \mathbf{h}_k^2, \cdots, \mathbf{h}_k^{N_m}\}$.

2) GMBM channel alphabet: In GMBM, n_{rf} out of n_{tu} MBM-TUs are selected based on $\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor$ bits, which is equivalent to selecting a combination of n_{rf} channel alphabets $\mathcal{H}_{k_1}, \cdots, \mathcal{H}_{k_{n_{rf}}}$ out of the $2^{\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor}$ available combinations. For each active MBM-TU, a MAP is chosen based on m_{rf} bits, which is equivalent to selecting one of the channels $\mathbf{h}_{k_j}^{m_l}$ out of the N_m available channels in each $\mathcal{H}_{k_j}$, $j = 1, \cdots, n_{rf}$, $l = 1, \cdots, N_m$. Overall, the MBM-TU indexing and the MAP indexing together selects a combination of n_{rf} channels, $\mathbf{h}_{k_1}^{m_1}, \mathbf{h}_{k_2}^{m_2}, \cdots, \mathbf{h}_{k_{n_{rf}}}^{m_{n_{rf}}}$. The set of all such combinations form the GMBM channel alphabet, given by

$$\mathcal{H} = \Big\{ [\mathbf{h}_{k_1}^{m_1} \mathbf{h}_{k_2}^{m_2} \cdots \mathbf{h}_{k_{n_{rf}}}^{m_{n_{rf}}}] \Big| (k_1, \cdots, k_{n_{rf}}) \in \mathcal{B}_{n_{tu}}^{n_{rf}} \text{ and } m_i \in \{1, \cdots, N_m\}, i = 1, \cdots, n_{rf} \Big\}, \tag{2}$$

where $\mathcal{B}_{n_{tu}}^{n_{rf}}$ is the set of MBM-TU activation patterns (MTAP), where an MTAP is an n_{rf}-tuple, which contains the indices of the n_{rf} active MBM-TUs. The size of the GMBM channel channel alphabet is $|\mathcal{H}| = 2^{(\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor + n_{rf} m_{rf})}$, and we denote the symbols of this channel alphabet by $\mathcal{H}^l$, $l = 1, 2, \cdots, |\mathcal{H}|$. Therefore, in GMBM, one of the $|\mathcal{H}|$ symbols is chosen from the GMBM channel alphabet $\mathcal{H}$ based on the $\log_2 |\mathcal{H}| = (\lfloor \log_2 \binom{n_{tu}}{n_{rf}} \rfloor + n_{rf} m_{rf})$ bits, thereby conveying these information bits using the channel alphabet.

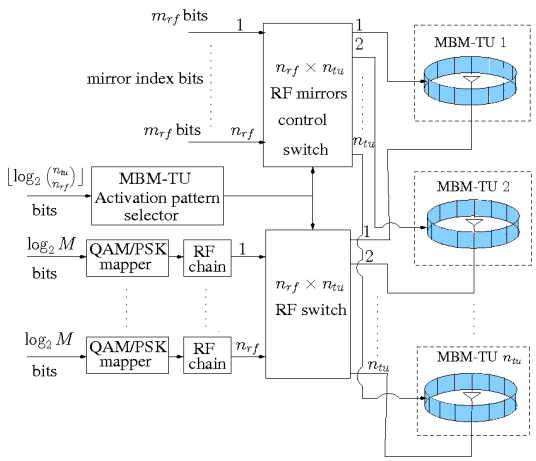


Fig. 1: GMBM system block diagram.

C. GMBM signal set

We first present the MBM signal set and then extend it to the GMBM signal set.

1) MBM signal set: Define $\mathbb{A}_0 \triangleq \mathbb{A} \cup 0$. The MBM signal set, denoted by $\mathbb{S}_{\text{MBM}}$, is the set of $N_m \times 1$-sized MBM signal vectors given by

$$\mathbb{S}_{\text{MBM}} = \{ \mathbf{s}_{m,q} \in \mathbb{A}_0^M : m = 1, \cdots, N_m, \; q = 1, \cdots, |\mathbb{A}| \}$$
$$\text{s.t} \;\; \mathbf{s}_{m,q} = [0, \cdots, 0, \underbrace{s_q}_{m\text{th coordinate}} 0, \cdots, 0]^T, s_q \in \mathbb{A}, \tag{3}$$

where m is the index of the MAP.

2) GMBM signal set: In GMBM, MBM signal vectors are transmitted on the n_{rf} active MBM-TUs and the remaining $(n_{tu} - n_{rf})$ MBM-TUs are silent, which is equivalent to zero vectors being transmitted. With this, the GMBM signal set is given by

$$\mathbb{S}_{\text{GMBM}}=\Bigg\{\mathbf{x}=[\mathbf{x}_1^T\mathbf{x}_2^T\cdots\mathbf{x}_{n_{tu}}^T]^T \text{ s.t } \mathbf{x}_i\in\mathbb{S}_{\text{MBM}}\cup\mathbf{0}, i=1,\cdots,n_t, \|\mathbf{x}\|_0=n_{rf}, \text{ and } (k_1,\cdots,k_{n_{rf}})\in\mathcal{B}_{n_{tu}}^{n_{rf}}\Bigg\}, \quad (4)$$

where $(k_1,\ldots,k_{n_{rf}})$ are the indices of the n_{rf} active MBM-TUs and $\mathcal{B}_{n_{tu}}^{n_{rf}}$ is the set of MTAPs as defined earlier.

D. GMBM received signal

Let n_r denote the number of receive antennas and $\mathcal{A}\triangleq\mathbb{A}^{n_{rf}}$ denote the GMBM source alphabet, which is the set of all combinations of n_{rf} symbols from $\mathbb{A}$. Then, the $n_r\times 1$ received GMBM signal vector is given by

$$\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{n}=\mathcal{H}^l\mathbf{s}+\mathbf{n}, \quad (5)$$

where $\mathbf{H}=[\mathcal{H}_1\cdots\mathcal{H}_{n_{tu}}]$ is $n_r\times n_{tu}N_m$ channel matrix, $\mathbf{x}\in\mathbb{S}_{\text{GMBM}}$ is the $n_{tu}N_m\times 1$ GMBM transmit vector, $\mathcal{H}^l\in\mathcal{H}$ is an $n_r\times n_{rf}$ GMBM channel symbol, $\mathbf{s}\in\mathcal{A}$ is the GMBM source symbol which forms the non-zero part of $\mathbf{x}$, and $\mathbf{n}\in\mathbb{C}^{n_r\times 1}$ is the additive white Gaussian noise vector. The entries of $\mathbf{n}$ are modeled as i.i.d complex Gaussian with zero mean and variance of σ^2.

Assuming perfect channel state information at the receiver (which is equivalent to the channel alphabet being known at the receiver), the maximum likelihood (ML) detection rule is to jointly detect the channel and source symbols as follows:

$$\left[\hat{\mathcal{H}}^l,\hat{\mathbf{s}}\right]=\underset{\mathcal{H}^l\in\mathcal{H},\mathbf{s}\in\mathcal{A}}{\operatorname{argmin}}\|\mathbf{y}-\mathcal{H}^l\mathbf{s}\|^2. \quad (6)$$

The ML detection has exponential complexity and hence ML detection becomes infeasible for GMBM systems with higher number of MBM-TUs and RF chains. We consider low-complexity detection of GMBM signals in Sec. IV. Before that, we carry out the capacity analysis of GMBM in the following section.

III. GMBM CAPACITY ANALYSIS

In this section, we derive an expression for the capacity of GMBM in closed-form. The main idea here is to recognize that, in GMBM, information is conveyed using channel symbols $\mathcal{H}^l\in\mathcal{H}$ and the source symbols $\mathbf{s}\in\mathcal{A}$. Therefore, the mutual information is the amount of information gained about both the channel and the source symbols on receiving $\mathbf{y}$, which is given by

$$I(\mathbf{y};\mathcal{H},\mathcal{A})=H(\mathbf{y})-H(\mathbf{y}|\mathcal{H},\mathcal{A}), \quad (7)$$

where

$$H(\mathbf{y})=-\mathbb{E}_{\mathbf{y}}(\log_2 p(\mathbf{y})) \quad (8)$$

is the entropy of $\mathbf{y}$, with $p(\mathbf{y})$ being the probability density function (pdf) of $\mathbf{y}$, and

$$H(\mathbf{y}|\mathcal{H},\mathcal{A})=H(\mathbf{n})=-\mathbf{E}_{\mathbf{n}}(\log_2 p(\mathbf{n})) \quad (9)$$

is the noise entropy, with $p(\mathbf{n})$ being the pdf of noise. The pdf of $\mathbf{y}$ can be simplified as

$$\begin{aligned}p(\mathbf{y})&=\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}p(\mathbf{y},\mathcal{H}^l,\mathbf{s})\\&=\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}p(\mathbf{y}|\mathcal{H}^l,\mathbf{s})p(\mathcal{H}^l,\mathbf{s}).\end{aligned} \quad (10)$$

Since the channel and the source symbols are independent, the pdf of $\mathbf{y}$ can be further simplified as

$$\begin{aligned}p(\mathbf{y})&=\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}p(\mathbf{y}|\mathcal{H}^l,\mathbf{s})p(\mathcal{H}^l)p(\mathbf{s})\\&=\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}\frac{1}{2^{\eta_{\text{GMBM}}}}\frac{1}{(\pi\sigma^2)^{n_r}}e^{-\frac{\|\mathbf{y}-\mathcal{H}^l\mathbf{s}\|^2}{\sigma^2}}.\end{aligned} \quad (11)$$

Using (11) in (8), the entropy of $\mathbf{y}$ can be simplified as

$$H(\mathbf{y})=-n_r\log_2(\pi\sigma^2)-\mathbb{E}_{\mathbf{y}}\left\{\log_2\left(\frac{1}{2^{\eta_{\text{GMBM}}}}\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}e^{-\frac{\|\mathbf{y}-\mathcal{H}^l\mathbf{s}\|^2}{\sigma^2}}\right)\right\}. \quad (12)$$

For the circularly symmetric Gaussian noise, the conditional entropy in (9) becomes [11]

$$H(\mathbf{y}|\mathcal{H},\mathcal{A})=H(\mathbf{n})=n_r\log_2(\pi\sigma^2 e). \quad (13)$$

Using (12) and (13) in (7), the mutual information is given by

$$I(\mathbf{y};\mathcal{H},\mathcal{A})=-n_r\log_2(e)-\mathbb{E}_{\mathbf{y}}\left\{\log_2\left(\frac{1}{2^{\eta_{\text{GMBM}}}}\sum_{\mathcal{H}^l\in\mathcal{H}}\sum_{\mathbf{s}\in\mathcal{A}}e^{-\frac{\|\mathbf{y}-\mathcal{H}^l\mathbf{s}\|^2}{\sigma^2}}\right)\right\}. \quad (14)$$

To find the capacity of GMBM, the mutual information has to be maximized jointly over the pdfs of the channel and source alphabets. Thus, the capacity of GMBM is given by

$$\begin{aligned}C&=\max_{p_{\mathcal{H}},p_{\mathcal{A}}}I(\mathbf{y};\mathcal{H},\mathcal{A})\\&=\max_{p_{\mathcal{H}},p_{\mathcal{A}}}H(\mathbf{y})-H(\mathbf{y}|\mathcal{H},\mathcal{A}),\end{aligned} \quad (15)$$

where $p_{\mathcal{H}}$ and $p_{\mathcal{A}}$ denote the pdfs of channel and source alphabets, respectively. It should be noted that there is no averaging on the channel in (15) since channel serves as an alphabet to convey information and is known at the receiver, like the source alphabet is known at the receiver.

From (13) it is clear that, the $H(\mathbf{y}|\mathcal{H},\mathcal{A})$ in (15) does not depend on $\mathcal{H}$ and $\mathcal{A}$, hence they need not be considered for maximization. Therefore, to maximize the mutual information $I(\mathbf{y};\mathcal{H},\mathcal{A})$ in (15), we need to maximize the entropy of $\mathbf{y}$, i.e., maximize $H(\mathbf{y})$. It is known that the distribution maximizing the entropy of $\mathbf{y}$ is Gaussian [11]. However, the key idea is to observe from (5) that, for $\mathbf{y}$ to have Gaussian distribution, the

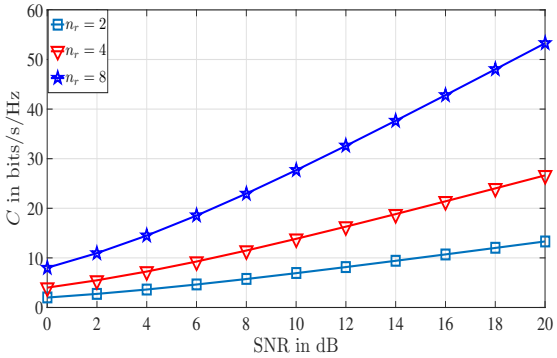


Fig. 2: Capacity of GMBM with $n_r = 2$, 4, and 8.

product $\mathcal{H}^l$s should be Gaussian distributed. That is, the entropy of $\mathbf{y}$ is maximized when the product of the source and channel symbols is Gaussian distributed. Combining this with the fact that there is no averaging over the channel in (15), the capacity of GMBM is achieved when the source-channel product alphabet $\mathcal{HA}$ is Gaussian. However, since the channel alphabet is fixed, a suitable signal shaping of the source symbols is required for the product alphabet to be Gaussian. It has been shown in [10] in the context of space modulation techniques that, when the channel is Rayleigh, the product alphabet $\mathcal{HA}$ is Gaussian distributed if the symbols of the source alphabet have constant amplitude and uniformly distributed (in $[-\pi, \pi]$) phases. With this, let the product distribution $\mathcal{HA} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{n_r})$, where $\mathbf{I}_{n_r}$ is the $n_r \times n_r$ identity matrix. Then, the maximum entropy of $\mathbf{y}$ is

$$\max_{p_{\mathcal{H}}, p_{\mathcal{A}}} H(\mathbf{y}) = n_r \log_2(\pi e(1+\sigma^2)). \tag{16}$$

Now, using (13), (16) in (15), the capacity of GMBM is obtained as

$$C = n_r \log_2\left(1 + \frac{1}{\sigma^2}\right) = n_r \log_2(1+\gamma), \tag{17}$$

where $\gamma \triangleq \frac{1}{\sigma^2}$ is the signal-to-noise ratio (SNR).

It can be seen from (17) that the capacity of GMBM depends only on the SNR and the number of receive antennas. A similar result has been recently shown in [10] in the context of spatial modulation. It is noted that the earlier literature on MBM [3],[5] has shown only the asymptotic ($m_{rf} \to \infty$) capacity of MBM, when only the channel alphabet is used for conveying information, to be equal to the expression in (17). However, our current analysis is more general, since it considers both channel and source alphabets. Further, the result in (17) is not asymptotic and is valid for any number of RF mirrors. This result has not been reported before for GMBM. Figure 2 shows the capacity of GMBM with $n_r = 2$, 4, and 8, where we can observe the capacity increase with increase in the number of receive antennas. This shows that GMBM in well suited for large-scale MIMO systems that employ large number of receive antennas.

IV. Structured sparsity exploiting detection

In this section, we develop a sparsity exploiting algorithm for GMBM signal detection. The MBM signal vectors in (3) are inherently sparse with only one non-zero element out of $N_m = 2^{m_{rf}}$ elements. Therefore, GMBM signal vectors which are constructed using n_{rf} MBM signal vectors and $(n_{tu} - n_{rf})$ zero-vectors is also sparse with n_{rf} non-zeros out of the $n_{tu}N_m$ elements. For example, for a GMBM system with $n_t = 16$, $n_{rf} = 8$, and $m_{rf} = 6$ ($N_m = 2^{m_{rf}} = 64$), the sparsity factor is $\frac{8}{16 \times 64} = \frac{1}{128}$. Also, there is a structure in the GMBM signal vectors. A GMBM signal vector consists of n_{rf} subvectors each of length N_m with only one non-zero entry (MBM vectors of the active MBM-TUs) and $(n_{tu} - n_{rf})$ zero subvectors each of length N_m (of the inactive MBM-TUs). This can be seen from the GMBM signal set in (4). Conventional sparse recovery algorithms like those in [13]-[15] can not be used for GMBM signal detection since these algorithms are generic and have no constraint on the support of the signal vectors to be reconstructed. Further, incorporating the structure in the sparse recovery process can lead to significant performance gains in the signal detection [16],[17]. Therefore, we propose a greedy *structured sparsity* exploiting detection algorithm for the detection of GMBM signal vectors. We call the proposed algorithm is as the GMBM matching pursuit (GMBM-MP) algorithm.

A. GMBM signal detection as structured sparse recovery

The signal detection of GMBM can be formulated as a structured sparse recovery problem as follows:

$$\begin{aligned} &\min \|\mathbf{x}\|_1 \text{ s.t} \\ &(i)\ \mathbf{y} = \mathbf{Hx} + \mathbf{n} \\ &(ii)\ \|\mathbf{x}_i\|_0 \in \{0,1\}, i = 1, \cdots, n_{tu} \\ &(iii)\ \|\mathbf{x}\|_0 = n_{rf}. \end{aligned} \tag{18}$$

Note that the formulation in (18) takes into account all the constraints discussed before. We take greedy sparse recovery approach to reconstruct the GMBM signal vector, where the first step is to recover the support of the vector and the next step is to recover the non-zero values corresponding to the detected support. The main challenge is in the support recovery stage, since the recovered support has to satisfy the constraints in (18). In the proposed algorithm we carefully refine and restrict the support in each step so that the constraints in (18) are satisfied.

B. GMBM matching pursuit (GMBM-MP)

The listing of the proposed GMBM-MP algorithm is given in **Algorithm 1**. The algorithm starts with an initial support $\mathcal{S}^0$, which contains the initial estimates of the non-zero positions of $\mathbf{x}$. To find $\mathcal{S}^0$, the algorithm first computes the inner products of the received vector $\mathbf{y}$ with the channel vectors corresponding to the different MAPs of all the MBM-TUs. Let $\mathbf{W}^0$ be the $N_m \times$

Algorithm 1 GMBM matching pursuit

1: Inputs: $\mathbf{y}, \mathbf{H}, n_{rf}, n_{tu}, m_{rf}$
2: Initialize:
3: $\mathcal{S}^0 = \{l_1^0, l_2^0, \cdots, l_{n_{rf}}^0\}$,
where $l_j^0 = k_j^0(N_m - 1) + m_j^0$, $j = 1, \cdots, n_{rf}$,
with (m_j^0, k_j^0) = pos. of jth max. entry in $\mathbf{W}^0$ s.t $\{k_j^0\}_{j=1}^{n_{rf}}$ are distinct, where $\mathbf{W}^0(m,k) = (\mathbf{h}_k^m)^H\mathbf{y}$, $m = 1, \cdots, N_m$, $k = 1, \cdots, n_{tu}$
4: $\mathbf{a}^0 = \left(\mathbf{H}_{\mathcal{S}^0}^H\mathbf{H}_{\mathcal{S}^0} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\mathcal{S}^0}^H\mathbf{y}$
5: $\mathbf{r}^0 = \mathbf{y} - \mathbf{H}_{\mathcal{S}^0}\mathbf{a}^0$
6: Iteration: In the ith iteration, do the following:
7: $\tilde{\mathcal{S}}^i = \mathcal{S}^{i-1} \cup \{l'\}$ s.t $l' = k'(N_m - 1) + m'$,
where (m', k') = pos. of max. entry of $\mathbf{W}^i$,
with $\mathbf{W}^i(m,k) = (\mathbf{h}_k^m)^H\mathbf{r}^{i-1}$
8: $\mathbf{z} = \left(\mathbf{H}_{\tilde{\mathcal{S}}^i}^H\mathbf{H}_{\tilde{\mathcal{S}}^i} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\tilde{\mathcal{S}}^i}^H\mathbf{y}$
9: $\mathcal{S}^i = \{l_1^i, \cdots, l_{n_{rf}}^i\}$, where $l_1^i, \cdots, l_{n_{rf}}^i$ are n_{rf} elements of $\tilde{\mathcal{S}}^i$ corresponding to distinct MBM-TUs and resulting in highest n_{rf} values of $\mathbf{z}$
10: $\mathbf{a}^i = \left(\mathbf{H}_{\mathcal{S}^i}^H\mathbf{H}_{\mathcal{S}^i} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\mathcal{S}^i}^H\mathbf{y}$
11: $\mathbf{r}^i = \mathbf{y} - \mathbf{H}_{\mathcal{S}^i}\mathbf{a}^i$
12: If $\|\mathbf{r}^i\| > \|\mathbf{r}^{i-1}\|$, let $\mathcal{S}^i = \mathcal{S}^{i-1}$ and quit the iteration
13: $\mathcal{S} = \mathcal{S}^l$ and $\mathbf{a} = \left(\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\mathcal{S}}^H\mathbf{y}$
14: $\hat{s}_k = \text{argmin}_{s\in\mathbb{A}}\|a_k - s\|^2$, $\forall k = 1, \cdots, n_{rf}$
15: Output: The estimated GMBM signal vector $\widehat{\mathbf{x}}$ satisfying $\widehat{\mathbf{x}}_{\{1,\cdots,n_{tu}N_m\}\setminus\mathcal{S}} = \mathbf{0}$ and $\hat{\mathbf{x}}_{\mathcal{S}} = \widehat{\mathbf{s}} = [\hat{s}_1 \cdots \hat{s}_{n_{rf}}]^T$.

n_{tu} matrix of inner products of $\mathbf{y}$ with the N_m channel vectors corresponding to different MAPs of the n_{tu} MBM-TUs, with kth column of $\mathbf{W}^0$ containing the inner products of $\mathbf{y}$ with different MAPs of kth MBM-TU. The (m,k)th entry of $\mathbf{W}^0$ is given by

$$\mathbf{W}^0(m,k) = (\mathbf{h}_k^m)^H\mathbf{y}, \tag{19}$$

where $\mathbf{h}_k^m$ denotes the channel gain vector from kth MBM-TU using mth MAP to n_r receive antennas. Now, the n_{rf} pairs $\{(m_1^0, k_1^0), \cdots, (m_{n_{rf}}^0, k_{n_{rf}}^0)\}$ are selected which are the n_{rf} (m,k) positions of $\mathbf{W}^0$ with highest inner product values such that $k_1^0, \cdots, k_{n_{rf}}^0$ are distinct. The $k_1^0, \cdots, k_{n_{rf}}^0$ are the initial estimates of the n_{rf} active MBM-TUs and $m_1^0, \cdots, m_{n_{rf}}^0$ are the initial estimates of the MAPs corresponding to the active MBM-TUs. The initial support $\mathcal{S}^0$ of $\mathbf{x}$ is then given by $l_j^0 = (k_j^0 - 1)N_m + m_j^0$, $j = 1, \cdots, n_{rf}$. This is shown in step 3 of the algorithm. The initial estimates of the non-zero values corresponding to the detected initial support is then obtained as

$$\mathbf{a}^0 = \left(\mathbf{H}_{\mathcal{S}^0}^H\mathbf{H}_{\mathcal{S}^0} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\mathcal{S}^0}^H\mathbf{y}, \tag{20}$$

where γ is the SNR, $\mathbf{H}_{\mathcal{S}^0}$ denotes the $n_r \times n_{rf}$ matrix obtained by restricting the columns of $\mathbf{H}$ to $\mathcal{S}^0$. Note that this is the MMSE detection done for the non-zero part of $\mathbf{x}$. This is shown in step 4 of the algorithm. The initial residue $\mathbf{r}^0$ is found in step 5.

The algorithm then refines the support in each iteration and finds the non-zeros corresponding to the refined support. In the ith iteration, the inner product of the residue of $(i-1)$th iteration, $\mathbf{r}^{i-1}$, with the channel vectors $\mathbf{h}_k^m$s are computed to obtain the matrix $\mathbf{W}^i$. The (m,k)th entry of $\mathbf{W}^i$ is given by

$$\mathbf{W}^i(m,k) = (\mathbf{h}_k^m)^H\mathbf{r}^{i-1}. \tag{21}$$

Let $l' = (k'-1)N_m + m'$, where (m', k') is the position of the highest inner product in $\mathbf{W}^i$. Now, an intermediate support set $\tilde{\mathcal{S}}^i$ is formed by appending l' to the support set of $(i-1)$th iteration, $\mathcal{S}^{i-1}$. This is shown in the step 7 of the algorithm. This support set is refined by first computing the non-zeros corresponding to the support $\tilde{\mathcal{S}}^i$ in step 8 and taking the n_{rf} elements of $\tilde{\mathcal{S}}^i$ corresponding to different MBM-TUs, that result in n_{rf} maximum non-zero values. This is shown in step 9 of the algorithm, where $\mathcal{S}^i$ denotes the refined support set of ith iteration. The non-zeros corresponding to the refined support are computed as

$$\mathbf{a}^i = \left(\mathbf{H}_{\mathcal{S}^i}^H\mathbf{H}_{\mathcal{S}^i} + \frac{1}{\gamma}\mathbf{I}\right)^{-1}\mathbf{H}_{\mathcal{S}^i}^H\mathbf{y}, \tag{22}$$

which is shown in step 10 of the algorithm. The residue in the ith iteration is computed in step 11 and this residue is compared to the residue of the previous iteration in step 12. If the current residue is less than the residue of previous iteration, then the algorithm proceeds to the next iteration; otherwise the iteration is stopped. Finally, the computed non-zeros are mapped to the valid symbols from $\mathbb{A}$ in step 14.

Complexity: The main steps contributing to the complexity of GMBM-MP are the computation of $\mathbf{W}^i$ and $\mathbf{a}^i$. The $\mathbf{W}^i$ can be efficiently computed as $\mathbf{H}^H\mathbf{y}$ and reshaping and has the complexity of $\mathcal{O}(n_r n_{tu} N_m)$. The computation of $\mathbf{a}^i$ has the complexity of $\mathcal{O}(n_{rf}^3)$. Therefore, the complexity of the GMBM-MP is $\mathcal{O}(n_r n_{tu} N_m + n_{rf}^3)$, which is only polynomial in the system parameters, unlike ML detection whose complexity is exponential in system parameters.

C. Results and discussions

In this section, we present the simulated GMBM BER performance and complexity results using the proposed GMBM-MP algorithm. We compare these results with those of GMBM using a multi-layered message passing (MLMP) algorithm in [12].

BER performance with GMBM-MP and MLMP detection: Figure 3 shows a BER performance comparison between GMBM-MP detection and MLMP detection. The following two GMBM system configurations are considered: i) system-1: $n_{tu} = 8$, $n_{rf} = 4$, $n_r = 48$, $m_{rf} = 5$, 4-QAM, and 34 bpcu, ii) system-2: $n_{tu} = 16$, $n_{rf} = 8$, $n_r = 96$, $m_{rf} = 6$, 4-QAM, and 77 bpcu. It is noted that both the systems are underdetermined with system-1 having 256 transmit and 48 receive dimensions, and system-2 having 1024 transmit and 96 receive dimensions. From Fig. 3, it can

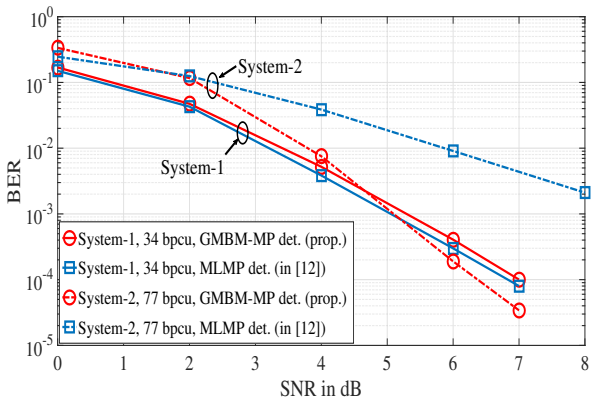


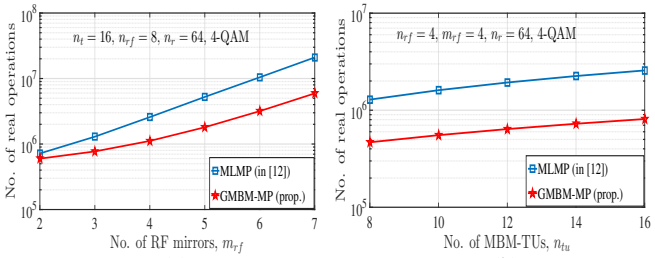
Fig. 3: BER performance of GMBM systems 1 and 2 with GMBM-MP and MLMP detection.

be seen that system-1 achieves almost similar BER performance with both GMBM-MP and MLMP detection. However, the BER results for system-2 show that the performance with GMBM-MP detection is much superior compared to the performance with MLMP detection. For example, in system-2, GMBM-MP detection achieves about 3.5 dB better performance compared to that with MLMP detection at a BER of 2×10^{-3}. Note that system-2 is more underdetermined than system-1. This illustrates the effectiveness of the GMBM-MP algorithm in exploiting the structured sparsity inherent in the GMBM signal vectors in highly underdetermined GMBM system settings.

Complexity of GMBM-MP and MLMP detection: The complexity of GMBM signal detection using GMBM-MP and MLMP as a function of number of RF mirrors and the number of MBM-TUs are shown in Figs. 4(a) and 4(b), respectively. In Fig. 4(a), GMBM with $n_t = 6$, $n_{rf} = 8$, $n_r = 64$, and 4-QAM is considered, and the complexity of GMBM-MP and MLMP, in number of real operations, is plotted as a function of m_{rf}. From this figure, it is observed that GMBM-MP has lesser complexity compared to MLMP. A similar complexity advantage in favor of GMBM-MP can be seen in Fig. 4(b), where the complexity is plotted as a function of the number of MBM-TUs. Thus, the BER and complexity results show that the proposed GMBM-MP detection achieves good BER performance even when the system is highly underdetermined and also that the GMBM-MP has lesser computational complexity compared to MLMP.

V. Conclusions

We investigated two important aspects of GMBM, namely, capacity and low-complexity signal detection. Recognizing that GMBM uses both source and channel alphabets to convey information, and that the capacity is achieved when the mutual information is maximized over the source-channel product alphabet, we derived closed-form capacity expression for GMBM. It was shown that the GMBM capacity is achieved when the source-channel product alphabet is Gaussian distributed, and that the



(a) (b)

Fig. 4: Complexity of GMBM-MP and MLMP as a function of (a) number of RF mirrors and (b) number of MBM-TUs.

capacity depends only on the number of receive antennas and SNR. Regarding the detection of high-rate MBM signal vectors, we proposed a structured sparse recovery algorithm. Simulation results showed that the proposed algorithm achieves good BER performance at low complexity even when the system is highly underdetermined.

References

[1] A. K. Khandani, "Media-based modulation: a new approach to wireless transmission," *Proc. IEEE ISIT'2013*, pp. 3050-3054, Jul. 2013.

[2] A. K. Khandani, "Media-based modulation: converting static Rayleigh to AWGN," *IEEE ISIT*, Jun. 2014.

[3] E. Seifi, M. Atamanesh, A. K. Khandani, "Media-based MIMO: outperforming known limits in wireless," *Proc. IEEE ICC'2016*, May 2016.

[4] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors," *IEEE Trans. Veh. Tech.*, vol. 66, no. 6, pp. 4967-4983, Jun. 2017.

[5] B. Shamasundar, S. Jacob, T. Lakshmi Narasimhan, A. Chockalingam, "Media-based modulation for the uplink in massive MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 67, no. 9, pp. 8169-8183, Sep. 2018.

[6] E. Basar and I. Altunbas, "Space-time channel modulation," *IEEE Trans. Veh. Tech.*, vol. 66, no. 8, pp. 7609-7614, Aug. 2017.

[7] I. Yildirim, E. Basar, and I. Altunbas, " Quadrature channel modulation," *IEEE Wireless Comm. Lett.*, vol. 6, no. 6, pp. 790-793, Dec. 2017.

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

[9] T. Lakshmi Narasimhan, P. Raviteja, and A. Chockalingam, "Generalized spatial modulation in large-Scale multiuser MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 7, pp. 3764-3779, Jul. 2015.

[10] A. Younis and R. Mesleh, "Information-theoretic treatment of space modulation MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 67, no. 8, pp. 6969-6969, Aug. 2018.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* 2nd ed., Wiley Publishers, Jul. 2006.

[12] K. M. Krishnan, S. Jacob, and A. Chockalingam, "Detection of generalized media-based modulation signals using multi-layered message passing," *Proc. IEEE VTC'2018-Spring*, Jun. 2018.

[13] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, vol. 12, pp. 4655-4666, Dec. 2007.

[14] D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis* 26.3 (2009): 301-321.

[15] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 5, pp. 2230-2249, May 2009.

[16] A. Bandi and C. R. Murthy, "Structured sparse recovery algorithms for data decoding in media based modulation," *Proc. IEEE ICC'2017*, Jul. 2017.

[17] Z. Gao, L. Dai, Z. Wang, S. Chen, and L. Hanzo, "Compressive-sensing-based multiuser detector for the large-scale SM-MIMO uplink," *IEEE Trans. Comm.,* vol. 63, no. 7, pp. 2565-2579, Oct. 2016.