Mutual Information Maximization for Optimal Spatial Modulation MIMO system

Haoyue Tang, Jintao Wang, Longzhuang He
Department of Electronic Engineering, Tsinghua University
Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084, P. R. China
Email: tanghaoyue13@mails.tsinghua.edu.cn wangjintao@tsinghua.edu.cn helongzhuang@126.com

***Abstract*—Combining Spatial Modulation (SM) with multiple-input multiple-output (MIMO) systems constitutes a promising technique for exploiting full spatial multiplexing. To further maximize its capacity, we propose to optimize the activation probability with an iternative optimization method. Moreover, we propose to incorporate the newly-proposed distribution matcher encoder for probability shaping. Numerical results are provided to substantiate the proposed system's improvement in terms of mutual information lower bound, and to validate the efficiency of the proposed encoding scheme.**

***Index Terms*—spatial modulation; spacial multiplexing; MIMO systems**

I. Introduction

Spatial Modulation (SM) is a promising technique stemmed from the conventional multiple-input multiple-output (MIMO) systems [1]. Conventional MIMO systems require a dedicated Radio Frequency (RF) chain for each of the transmit antennas, which severely reduces the energy efficiency due to the high power dissipation of the RF chains. Besides, conventional MIMOs also suffer from inter channel inference (ICI) and inter antenna synchronization (IAS). With one antenna activated at each time slot, SM-MIMO systems require only a single radio frequency chain and are capable of avoiding the ICI and IAS. The essence of SM is that the position of active antenna conveys information.

Conventionally, SM requires each antenna to be randomly activated according to a uniform distribution. Mutual information of SM-MIMO systems in conjunction with equal-probability activation has been analyzed by An *et al.* [2] . However, it is widely acknowledged that, aided with the transmitter's knowledge of the channel state information (CSI), a variety of transmitter pre-processing techniques can be applied to further benefit the performance of SM-MIMO systems [3]. The performance of SM-MIMO systems are improved with phase preprocessing. [4] [5]. Precoding for SM-MIMO systems is studied in [6] and minimum Euclidean distance is enlarged..

Recently, Liu *et al.* proposed an optimal spatial-domain design for capacity maximation [7]. Unlike the conventional SM-MIMO systems, transmitted antennas are activated with different probability. Besides, the authors proposed an iterative heuristic approach to find the suboptimal probability distribution for MISO systems. However, the proposed model and algorithm deal only with the multiple antenna systems with one receive antenna, and lacks a practical coding algorithm for non-identical activation distribution.

It is worth noting that, in practice, the input bit stream is always assumed to be uniformly distributed. Therefore, an encoding regime with *probability reshaping* is needed so that different activation probability can be assigned to different transmit antennas. In 2016, W. Wang and W. Zhang introduced Huffmann coding into the field of Spatial Modulation systems [8], which receives a better performance. However, Huffmann coding is a suboptimal coding scheme with huge redundancy and various block lengths, thus which could be impractical in a realistic communication scenario. Without coding, the gap between the real capacity and the optimal SM-MIMO capacity could not be dropped. Coded Modulation is to develop practical tools that provably allow to overcome the shaping gap. In 2016, P. Schulte and G. Boecher proposed a practical, invertible, fixed to fixed (f2f) length distribution matcher. [9] The matcher could satisfy the needs of SM-MIMO systems.

The main contributions of the paper are as follows:

1) A novel SM-MIMO construction is proposed along with probability shaping. Aided with mutual information analysis, the corresponding capacity lower bound is also proposed. Based on the proposed bound, an optimization algorithm is proposed to design the antenna activation probability in terms of mutual information lower bound maximization. A low complexity heuristic probability distribution algorithm that could achieve mutual information gain with full Tx CSI is also proposed in the paper.

2) To facilitate the practical application of the proposed system, the concept of constant composition distribution matching is exploited for probability shaping, which is shown to be effective via numerical simulations.

The rest of the paper is organized as follows. We introduce the adaptive activation SM-MIMO system model and the fading channel model in Section II. Section III discussed the probability distribution optimization algorithm. Section IV provides the heuristic antenna distribution allocation algorithm, both under full CSI and statistical CSI. Numerous simulation results are provided in Section IV. Section V summarizes the paper and briefly discuss future directions.

II. System Model

In our research, we consider an $(m \times n)$ SM-MIMO system, where m and n denote the numbers of transmit and receive

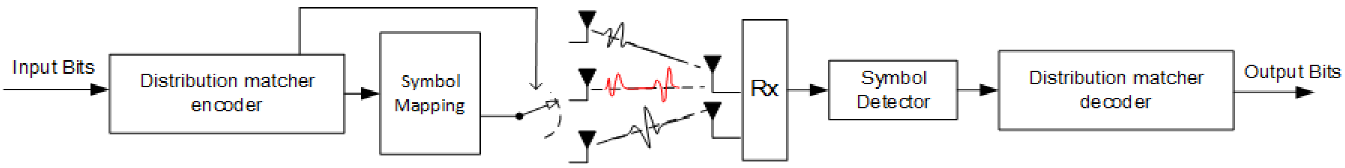

Fig. 1. Block graph of optimal SM-MIMO system

antennas. One antenna X_{ch} is randomly activated at each time slot and transmit an amplitude-phase modulation (APM) symbol X to the receive antenna, $\mathbf{y} \in \mathbb{C}^n$ is the received vector. The channel state information (CSI) $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_M]$ is known at the receiver side. On the receiver side, there exists an additive Gaussian white noise (AWGN) $\mathbf{z} \sim \mathbb{C}\mathcal{N}(0, \sigma_z \mathbf{I}_N)$. For information transmission, the probability shaping encoder first encodes the input bits into active antenna position with dirferent activation probabilities, the chosen antenna send the APM symbols through the channel, the receiver demodulates the received bit sequence and then demaps the symbol to bit stream.

Suppose the current activated antenna is m, the received signal could be expressed as

$$\mathbf{y} = \mathbf{h}_m X + \mathbf{z} \tag{1}$$

Different from the conventional SM-MIMO system, the activation probability of each antenna is various. Thus, the capacity of the proposed SM-MIMO system is

$$C = \max_{\mathbf{p}_{SM}} I(X, X_{ch}; \mathbf{Y}) \tag{2}$$

The vector $\mathbf{p}_{SM}$ denotes the activation probability distribution of the antenna. The mutual information for a given activation distribution is:

$$I(X, X_{ch}; \mathbf{Y}) = I(X; \mathbf{y}|X_m) + I(X_m; \mathbf{Y}) \tag{3}$$

Denote the first item as I_1 and the second item as I_2. I_1 represents the mutual information between the channel output and the input APM symbols. By applying the capacity of conventional MIMO systems, the APM domain mutual information could be formed as

$$I(X; \mathbf{Y}|X_{ch}) = \sum_{m=1}^{M} p(X_{ch} = m) \log(1 + \frac{\|\mathbf{h_m}\|_2^2}{\sigma_n^2}) \tag{4}$$

As the transmitted antenna, received symbols and the estimated antenna $X_{ch} - \mathbf{Y} - \hat{X}_{ch}$ form a Markov Chain, Data processing inequality could be applied here as a lower bound for the mutual information. By further applying the Chain Rule for mutual information:

$$I(X_{ch}; \mathbf{Y}) \geq I(X_{ch}; \hat{X}_{ch}) = H(\hat{X}_{ch}) - H(\hat{X}_{ch}|X_{ch}) \tag{5}$$

Suppose the probability distribution for active antennas is not fully known at the receiver side, the cross entropy of $H(\hat{X}_{ch}|X_{ch})$ could be estimated via a Maximum Likelyhood (ML) detector. In high SNR scenerios, the pairwise probability of judging the m^{th} antenna as k^{th} is upper bounded by

$$Pr(m \to k) \leq P(|\mathbf{h}_{k,0}^{\mathcal{H}} \mathbf{y}| \geq |\mathbf{h}_{m,0}^{\mathcal{H}} \mathbf{y}|) \tag{6}$$

The closed form the the pairwise probability is given [2]

$$Pr(m \to k) = \frac{1}{2} - \frac{1 - |\eta_{m,k}|^2 - \frac{\sigma_z^2}{\sigma_m^2}(1 - |\eta_{m,k}|^2)}{2\sqrt{[1 + |\eta_{m,k}|^2 + \frac{\sigma_z^2}{\sigma_m^2}(1 - |\eta_{m,k}|^2)]^2 - 4|\eta_{m,k}|^2}} \tag{7}$$

$\eta_{m,k}$ describes the relevance of the transmitting channel

$$\eta_{m,k} = \frac{\mathbf{h}_m^{\mathcal{H}} \mathbf{h}_k}{\|\mathbf{h}_m\|_2 \|\mathbf{h}_k\|_2} \tag{8}$$

The pairwise transition probability is fixed as long as the channel is invariant. In reliabele communication system scenerio, where SNR is large, the entropy of the distribution of the estimated active antenna probability could be computed via

$$p(\hat{X}_{ch} = k) = \sum_{m=1}^{M} p(X_{ch} = m) Pr(m \to k) \tag{9}$$

The pairwise transition matrix is given by

$$\mathbf{P}_{trans} = \begin{bmatrix} Pr(1 \to 1) & Pr(2 \to 1) & ... & Pr(M \to 1) \\ Pr(2 \to 1) & Pr(2 \to 2) & ... & Pr(M \to 2) \\ ... & & & ... \\ Pr(M \to 1) & Pr(M \to 2) & ... & Pr(M \to M) \end{bmatrix} \tag{10}$$

Denote the probability of activation antenna, the APM capacity of each channel and the conditional entropy as vectors

$$\mathbf{p}_{SM} = [p(X_{ch} = 1), p(X_{ch} = 2), ..., p(X_{ch} = m)]^T \tag{11}$$

$$\mathbf{C}_{APM} = \left[\log\left(1 + \frac{\|\mathbf{h}_1\|_2^2)}{\sigma_z^2}\right), ..., \log(1 + \frac{\|\mathbf{h}_M\|_2^2)}{\sigma_z^2}),\right] \tag{12}$$

$$\mathbf{H}_{pair} = \left[H(\hat{X}_{ch}|X_{ch} = 1), ..., H(\hat{X}_{ch}|X_{ch} = M)\right] \tag{13}$$

The channel capacity lower bound is given by

$$C = \max_{\mathbf{p}_{SM}} \mathbf{C}_{APM}^T \mathbf{p}_{SM} + H(\mathbf{P}_{trans}\mathbf{p}_{SM}) - \mathbf{H}_{pair}^T \mathbf{p} \tag{14}$$

III. Capacity Bounds and Optimizaton

We first prove the optimization problem we are facing is a concave maximization optimization problem. The mutual information is a linear function to the probability distribution. As the probability of the estimated antennas could be viewed as an affine mapping from the active antenna probability distribution and the entropy function is concave, thus $h(\hat{X}_{ch})$ is concave. The estimization transition probability is fixed for a typical channel, thus the conditional entropy proves to be a linear function. Thus, the capacity function is concave in the domain, and the global optimizer could be found and the optimial capacity could be reached with appropriate probability distribution.

With the constraints that the sum of probability distribution equals 1, the optimization problem is organized as follows:

$$C=\max_{\mathbf{p}_{SM}} \mathbf{C}_{APM}^T \mathbf{p}_{SM} + H(\mathbf{P}_{trans}\mathbf{p}_{SM}) - \mathbf{H}_{pair}^T \mathbf{p}_{SM}, \text{s.t } \mathbf{1}^T\mathbf{p}=1 \tag{15}$$

We start with conventional SM-MIMO system as an initial point, whose antenna activation distribution is uniform. As the initial point is within the domain, phase I optimization could be skipped. Using the Backtracking algorithm and the Newton Method with Equality constraints, the descent direction could be calculated via

$$\begin{bmatrix} \nabla_{\mathbf{p}}^2 C & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{p} \\ \Delta\nu \end{bmatrix} = \begin{bmatrix} -\nabla_{\mathbf{p}} C \\ 0 \end{bmatrix} \tag{16}$$

The Hessian matrix is computed via

$$\nabla_{\mathbf{p}}^2 = \mathbf{P}_{trans}\text{diag}(\mathbf{P}_{trans}\mathbf{p})^{-1}\mathbf{P}_{trans}^T \tag{17}$$

Combining the Newton Method with equality constraints and the backtracking line search algorithm [13], the optimal capacity converges after a few iterations.

Initial Feasible Point Selection $\mathbf{p}_{SM} = \frac{1}{M}\mathbf{1}$
while $\Delta C_{improve} > threshold$ **do**
  $t := 1$
  $[\Delta\mathbf{p}_{SM}^T, \Delta\nu^T] = \begin{bmatrix} \nabla^2 C(\mathbf{p}_{SM}) & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla C(\mathbf{p}_{SM}) \\ \mathbf{0} \end{bmatrix}$
  if $\mathbf{p}_{SM} + t\Delta\mathbf{p}_{SM} \notin domC$ *or* $C(\mathbf{p}+t\Delta\mathbf{p}_{SM}) \geq C(\mathbf{p}) + \alpha t \nabla C(\mathbf{p})^T \Delta\mathbf{p}$ **then**
    $t := \beta t$
  end
end

Algorithm 1: Newton Method for Convex Optimization with Linear Constraints

IV. Heuristic Activation Probability Allocation Algorithm

Suppose the channel state $\mathbf{H}$ is fully known at the transmitter. Using the Singular Eigenvalue Decomposition

$$\mathbf{H}^H\mathbb{H} = \mathbf{V}_{Tx}\mathbf{\Psi}_{Tx}\mathbf{V}_{Tx}^H \tag{18}$$

The power allocation scheme is the following:

$$\mathbf{p}_{SM} = \mathbf{v}_{Tx,max} \otimes \mathbf{v}_{Tx,max}^* \tag{19}$$

V. Numerical Simulations

A. Realistic lower bound with infinite block lengths

In this part we investigate encoding system with infinite lengths. With CCDM encoder, the desired output could be perfectly matched and the lower bound could reach its optimum. Experiments are carried out with random generalized Rayleigh flat fading Channels of 1000 experiments.

From the figure, we could see the lower bound has been improved with optimization. The improvement is especially significant in, when the number of receive antennas is less than the transmit antennas. This is often the case in the downlink channel, where the antenna array on the base station is large, but the equipment is limited.

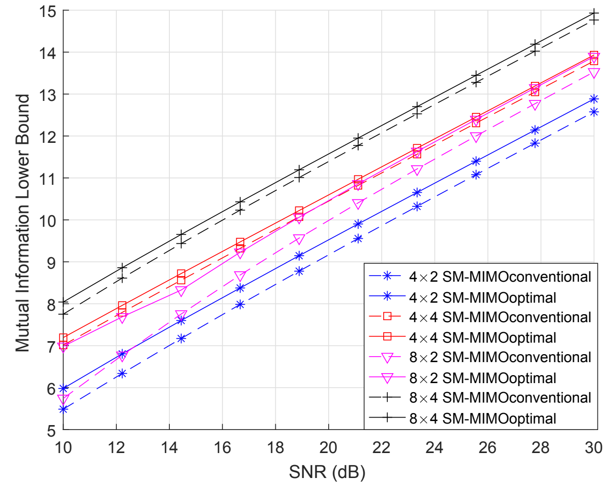

Fig. 2. Capacity Comparison of conventional SM-MIMO and optimal capacity SM-MIMO systems

B. Capacity lower bounds with increasing input block length

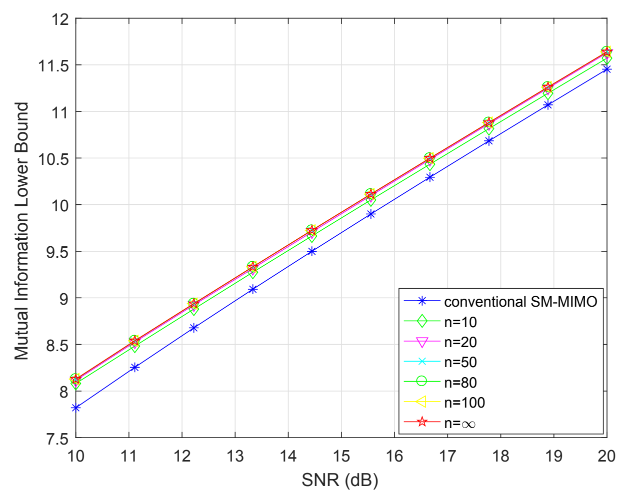

Fig. 3. Capacity Comparison of capacity lower bounds with different input block length for a 8×4 SM-MIMO system

The figure above describes the relationship between the block length and capacity, experiment is carried out with 8 transmit antennas, 4 receive antennas in Rayleigh flat fading channels through over 1000 realizations. In accordance with the theoretic analysis of the ccdm codes, it approaches the theoretic upper bound as the block length increases. The proposed encoder reaches a 10^{-4} bits gap under high SNR scenarios with block length $n = 100$, which proves its effectiveness and is of practical use.

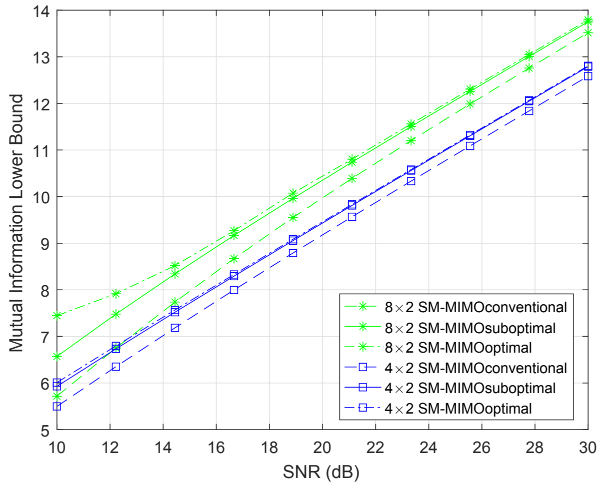

Fig. 4. Mutual information lower bounds for SM-MIMO systems with full CSI

C. with Full Tx CSI

Simulations are carried out in Rayleigh Fading Channels. In the figure, optimal SM-MIMO system with fast probability allocation algorithm overbeats the conventional MIMO systems. The gain increase with the increase of transmit antennas. The proposed algorithm receives better performance under low SNR scenario. From the figure, the gap between the proposed low complexity probability allocation scheme and the scheme via convex optimization is nearly 0.016 bits, which proves its effectiveness.

VI. Conclusion

In this paper, a novel SM-MIMO system is proposed to maximize the capacity of SM-MIMO system. The lower bound of the novel system has been derived and proved to be convex as a function of probability distribution. Furthermore, an optimal optimization algorithm based on actual channel state information can be adopted to maximize the capacity. Numerical simulation results proved its efficiency. Aided with a probability distribution matcher encoder ccdm, the novel system could be implemented into real communication system.

Future work will seek to engage in more theoretical analysis into the performance of SM-MIMO systems with statistical CSI. We will also try to design better precoding algorithm for the SM-MIMO to achieve optimal performance under statistical CSI. Moreover, theoretical studies on communicating with no CSI at both the transmitter side and the receiver side are also needed to get a thorough understanding of SM-MIMO systems.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 61471221).

References

[1] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation - a new low complexity spectral efficiency enhancing technique," in *2006 First International Conference on Communications and Networking in China*, Oct 2006, pp. 1–5.

[2] Z. An, J. Wang, J. Wang, S. Huang, and J. Song, "Mutual Information Analysis on Spatial Modulation Multiple Antenna System," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 826–843, Mar. 2015.

[3] C. Masouros, "Improving the diversity of spatial modulation in miso channels by phase alignment," *IEEE Communications Letters*, vol. 18, no. 5, pp. 729–732, May 2014.

[4] M. Maleki, H. R. Bahrami, S. Beygi, M. Kafashan, and N. H. Tran, "Space modulation with csi: Constellation design and performance evaluation," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 4, pp. 1623–1634, May 2013.

[5] L. He, J. Wang, C. Zhang, and J. Song, "Improving the performance of spatial modulation by phase-only pre-scaling," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 3210–3215.

[6] X. Guan, Y. Cai, and W. Yang, "On the mutual information and precoding for spatial modulation with finite alphabet," *IEEE Wireless Communications Letters*, vol. 2, no. 4, pp. 383–386, August 2013.

[7] C. Liu, M. Ma, Y. Yang, and B. Jiao, "Optimal Spatial-Domain Design for Spatial Modulation Capacity Maximization," *IEEE Communications Letters*, vol. 20, no. 6, Jun. 2016.

[8] W. Wang and W. Zhang, "Adaptive spatial modulation using huffman coding," in *2016 IEEE Global Communications Conference, GLOBECOM 2016, Washington, DC, USA, December 4-8, 2016*, 2016, pp. 1–6.

[9] P. Schulte and G. Boecherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, Jan 2016.

[10] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, Mar 2000.

[11] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic mimo radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, Aug 2002.

[12] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in mimo wireless systems under correlated fading," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 637–650, Mar 2002.

[13] B. Stephen and V. Lieven, *Convex Optimization*. Cambridge University Press, 2009.