Spatial Modulation vs. Single-Antenna Transmission: When is Indexing Helpful?

Bharath Shamasundar* and Aria Nosratinia†
* Samsung R&D Institute Bangalore, India
† The University of Texas at Dallas, Richardson, TX, USA
Email: *bharath.s1@samsung.com, †aria@utdallas.edu

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206660

***Abstract*—In spatial modulation, even though only one antenna is active in each transmission interval, the receiver needs channel estimates with respect to all transmit antennas at all times. Thus, spatial modulation is more sensitive to the cost of the estimation of channel state information (CSI) compared with conventional single-antenna transmission. Despite its importance, an accurate, joint characterization of the impact of CSI cost and CSI imperfections on the capacity of spatial modulation has been unavailable thus far. As a result, the marginal cost/benefit of each additional antenna, and hence the optimal antenna alphabet, has been unclear. This work calculates the spectral efficiency of spatial modulation subject to training through a tight characterization of the dependence of the achievable rate on the power and degrees of freedom dedicated to pilots. Our results reliably characterize the cases when conventional single-antenna transmission is superior to spatial modulation or vice versa.**

I. Introduction

Spatial modulation selects one out of M available transmit antennas per channel use, and transmits a modulation symbol from the selected antenna [1], [2], [3]. The selection index and modulation symbol both carry information [4], [5]. The motivation for spatial modulation is often reducing hardware complexity, but in the process the achievable rate may be reduced. A large part of the literature on spatial modulation assumes free CSI at the receiver [4], [5], [1], including spectral efficiency calculations with free and perfect CSI [6], [7], [8], [9]. Practical systems, however, obtain CSI via pilots and estimation, introducing two important features: first, acquiring CSI incurs a *cost* in power and transmission time (degrees of freedom). The cost of CSI can be a predominant design issue in some spatial modulation scenarios. Second, the CSI obtained by pilots and estimation is imperfect and contains some noise. In several studies of the *bit-error performance* of spatial modulation [10], [11], [12], [13], [14], [15], imperfect CSI has featured prominently but the *cost* of pilots has been disregarded.

In the analysis of *spectral efficiency,* the effect of these two features (the cost of CSI and its imperfection) cannot be meaningfully separated from one another, requiring a more careful analysis. Rajashekar *et al.* [7] derived a lower bound on the training-based capacity of spatial modulation which used a loose approximation for the rate of the index, and disregarded the influence of the rate of the index in data/pilot power optimization. Both these issues are resolved in the present work. Further, some important design and operational questions have remained unanswered in prior works on spatial modulation. Among them: it has been unclear when spatial modulation is superior to SIMO and vice versa, and how many antennas must optimally participate in spatial modulation (antenna alphabets). The manner of dependence of these issues on the channel dynamics and SNR is of significant interest, but has remained open thus far.

This work presents an accurate characterization of the spectral efficiency of spatial modulation via the derivation of a tight lower bound that accounts for training overhead and training error. In contrast to [7], the present work develops and utilizes an exact expression for the rate carried by the index. For example, at 6 bits/s/Hz, our lower bound is 2.5 dB tighter than the best available bound [7] for 4×2 spatial modulation. Our results also address broader operational questions that have so far remained unanswered. For instance: when should we use spatial modulation, and when should we simply use fixed single-antenna signaling? In the former, the RF chain will switch among antennas, allowing a component rate to be transmitted via the antenna index, but it also requires pilots for each antenna. In the latter, the RF chain is always connected to the same antenna, reducing the training requirement, but also giving up on the rate that could be emitted via the antenna index. It has been speculated that SIMO is preferable to spatial modulation in highly dynamic channels and at low SNR; our analysis for the first time corroborates this intuition with rigor, and produces the SNR-coherence length boundary that characterizes which approach should be chosen.

Our results also determine the optimal antenna alphabet for spatial modulation, i.e., the number of antennas that must participate in spatial modulation to yield maximal spectral efficiency, subject to training and pilots.[1] It is not always optimal to utilize all the available antennas; sometimes it is better to leave some of them completely unused and save the pilots that would be required for them. We show that the optimal antenna alphabet size grows with SNR and channel coherence time, and demonstrate the exact manner of this dependence.

This work was supported in part by the NSF grants 1956213 and 2148211.

[1] This can be considered a generalization of the question of SIMO vs. spatial modulation.

II. System Model

Consider a multi-antenna system with M transmit and N receive antennas. The $M \times N$ spatial modulation activates one transmit antenna per channel use and transmits a symbol. The system model is characterized by

$$\mathbf{y} = \sqrt{\rho}\left(\sum_{i=1}^{M} \mathbf{g}_i v_i\right) z + \mathbf{w}, \tag{1}$$

where ρ denotes the signal-to-noise ratio, $\mathbf{g}_i$ is the channel gain vector from transmit antenna i to N receive antennas, v_i is a binary variable that is zero if antenna i is inactive and one if antenna i is activated, and z denotes the modulation symbol transmitted from the active antenna. The noise $\mathbf{w}$ and the channel gain $\mathbf{g}_i$ are $N \times 1$ vectors whose components are i.i.d. and obey $\mathcal{CN}(0,1)$. The system model in (1) can be written compactly as follows

$$\mathbf{y} = \sqrt{\rho}\mathbf{G}\mathbf{v}z + \mathbf{w}, \tag{2}$$

where $\mathbf{G} = [\mathbf{g}_1 \ \mathbf{g}_2 \ \cdots \ \mathbf{g}_M]$ and $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_M]^T$ such that $\mathbf{v}$ belong to the canonical basis $\{\mathbf{e}_i\}$. Transmission is carried out in frames, with each frame divided into a training phase followed by a data transmission phase. In the training phase, pilot signals are transmitted, and the channel is estimated at the receiver. In the data transmission phase, spatial modulation vectors are transmitted, which are decoded at the receiver based on the received signal and the knowledge of the estimated channel.

III. Achievable Rate

Let $\widehat{\mathbf{G}}$ denote the MMSE estimate of $\mathbf{G}$ and $\tilde{\mathbf{G}}$ the estimation error, which are uncorrelated, zero-mean complex Gaussian matrices. The variance of the entries of $\tilde{\mathbf{G}}$ is denoted σ_e^2, thus entries of $\widehat{\mathbf{G}}$ have variance $1-\sigma_e^2$. The system model is equivalent to

$$\begin{aligned}\mathbf{y} &= \sqrt{\rho}\widehat{\mathbf{G}}\mathbf{v}z + \sqrt{\rho}\tilde{\mathbf{G}}\mathbf{v}z + \mathbf{w}\\ &= \sqrt{\rho}\widehat{\mathbf{G}}\mathbf{v}z + \tilde{\mathbf{w}},\end{aligned} \tag{3}$$

where $\tilde{\mathbf{w}} \triangleq \sqrt{\rho}\tilde{\mathbf{G}}\mathbf{v}z + \mathbf{w}$ is the effective noise at the receiver. Let $\widehat{\mathbf{g}}_i$ denote the column i of $\widehat{\mathbf{G}}$.

Proposition 1. *The capacity of spatial modulation with estimated CSIR satisfies the following lower bound:*

$$\begin{aligned}C \geq & \mathbb{E}_{\widehat{\mathbf{g}}_1}\left[\log_2\left(1+\frac{\|\widehat{\mathbf{g}}_1\|^2\rho}{1+\rho\sigma_e^2}\right)\right] + \log_2 M - \\ & \mathbb{E}_{\mathbf{y},\widehat{\mathbf{G}}}\left[\sum_{i=1}^{M}\frac{f_i(\mathbf{y},\widehat{\mathbf{G}})}{\sum_{j=1}^{M} f_j(\mathbf{y},\widehat{\mathbf{G}})}\log_2\frac{\sum_{j=1}^{M} f_j(\mathbf{y},\widehat{\mathbf{G}})}{f_i(\mathbf{y},\widehat{\mathbf{G}})}\right],\end{aligned} \tag{4}$$

where $z \sim \mathcal{CN}(0,1)$ and

$$f_i(\mathbf{y},\widehat{\mathbf{G}}) \triangleq \mathbb{E}_z\left[\frac{1}{(\pi(1+|z|^2\rho\sigma_e^2))^N}\exp\left(\frac{-\|\mathbf{y}-\sqrt{\rho}\widehat{\mathbf{g}}_i z\|^2}{1+|z|^2\rho\sigma_e^2}\right)\right]. \tag{5}$$

Proof. We aim to characterize a channel whose input is excited with a pilot sequence, antenna index, and modulation symbol, and whose output is the channel observations due to the pilot and data. Since channel estimates are a deterministic function of channel observations during pilot transmissions, by data processing inequality:

$$\begin{aligned}C &\geq \max_{p(\mathbf{v},z):\mathbb{E}(|z|^2)\leq 1} I(\mathbf{v},z;\mathbf{y},\widehat{\mathbf{G}})\\ &\overset{(a)}{=} \max_{p(\mathbf{v},z):\mathbb{E}(|z|^2)\leq 1}[I(\mathbf{v},z;\mathbf{y}|\widehat{\mathbf{G}}) + I(\mathbf{v},z;\widehat{\mathbf{G}})]\\ &\overset{(b)}{\geq} I(\mathbf{v},z;\mathbf{y}|\widehat{\mathbf{G}}) + I(\mathbf{v},z;\widehat{\mathbf{G}})\\ &\overset{(c)}{=} I(\mathbf{v},z;\mathbf{y}|\widehat{\mathbf{G}}),\end{aligned} \tag{6}$$

where (a) follows from chain rule, (b) replaces the optimal input distribution with an arbitrary distribution satisfying $\mathbb{E}(|z|^2) \leq 1$, and (c) follows from the independence of transmitted data from channel gains, and therefore of its estimate. We now calculate $I(\mathbf{v},z;\mathbf{y}|\widehat{\mathbf{G}})$ when $\mathbf{v}$ is uniformly distributed, and z obeys $\mathcal{CN}(0,1)$ independent of $\mathbf{v}$.

$$I(\mathbf{v},z;\mathbf{y}|\widehat{\mathbf{G}}) = I(z;\mathbf{y}|\mathbf{v},\widehat{\mathbf{G}}) + I(\mathbf{v};\mathbf{y}|\widehat{\mathbf{G}}). \tag{7}$$

The first term is a SIMO mutual information subject to estimated CSIR:

$$\begin{aligned}I(z;\mathbf{y}|\mathbf{v},\widehat{\mathbf{G}}) &\overset{(a)}{=} I(z;\mathbf{y}|\mathbf{v}=\mathbf{e}_1,\widehat{\mathbf{G}})\\ &\overset{(b)}{=} I(z;\mathbf{y}|\mathbf{v}=\mathbf{e}_1,\widehat{\mathbf{g}}_1)\\ &\overset{(c)}{\geq} \mathbb{E}_{\widehat{\mathbf{g}}_1}\left[\log_2\left(1+\frac{\|\widehat{\mathbf{g}}_1\|^2\rho}{1+\rho\sigma_e^2}\right)\right],\end{aligned} \tag{8}$$

where (a) follows from the statistical symmetry between transmit antennas, (b) from the independence of channels from different transmit antennas, and (c) follows from the worst-case noise property [16] while utilizing the expectation form for conditional mutual information[2]. We now evaluate $I(\mathbf{v};\mathbf{y}|\widehat{\mathbf{G}})$ as follows:

$$\begin{aligned}I(\mathbf{v};\mathbf{y}|\widehat{\mathbf{G}}) &= h(\mathbf{v}|\widehat{\mathbf{G}}) - h(\mathbf{v}|\widehat{\mathbf{G}},\mathbf{y})\\ &= \log_2 M - h(\mathbf{v}|\widehat{\mathbf{G}},\mathbf{y}),\end{aligned} \tag{9}$$

where we used independence of the uniformly distributed antenna index from channel gains. We now tend to the second term in (9)

$$h(\mathbf{v}|\widehat{\mathbf{G}},\mathbf{y}) = \mathbb{E}_{\mathbf{y},\widehat{\mathbf{G}}}\Big[\sum_{i=1}^{M} p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}},\mathbf{y})\log_2\frac{1}{p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}},\mathbf{y})}\Big], \tag{10}$$

where

$$\begin{aligned}p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}},\mathbf{y}) &= \frac{p(\mathbf{y}|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}})}{\sum_{j=1}^{M} p(\mathbf{y}|\mathbf{v}=\mathbf{e}_j,\widehat{\mathbf{G}})p(\mathbf{v}=\mathbf{e}_j|\widehat{\mathbf{G}})}\\ &= \frac{p(\mathbf{y}|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})}{\sum_{j=1}^{M} p(\mathbf{y}|\mathbf{v}=\mathbf{e}_j,\widehat{\mathbf{G}})},\end{aligned}$$

[2]The worst-case noise property applies here since, given the index, the channel between z and $\mathbf{y}$ is similar to [16], i.e., a Rayleigh fading SIMO channel with imperfect channel state information at the receiver.

using once again the uniform distribution $p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}}) = \frac{1}{M}$.

$$\begin{aligned} p(\mathbf{y}|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}}) &= \int_z p(\mathbf{y},z|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})dz \\ &= \int_z p(\mathbf{y}|z,\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})p(z|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})dz \\ &= \mathbb{E}_z\big[p(\mathbf{y}|z,\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})\big], \end{aligned}$$

where the last equality follows since $p(z|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}}) = p(z)$ due to the independence of z from $\mathbf{v}$ and $\widehat{\mathbf{G}}$. Expressing $p(\mathbf{y}|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})$ as an expectation of $p(\mathbf{y}|z,\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})$ enables further calculations since the distribution $p(\mathbf{y}|z,\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}})$ is known from the model in (3),

$$(\mathbf{y}|z,\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}}) \sim \mathcal{CN}(\sqrt{\rho}\widehat{\mathbf{g}}_i z, (1+|z|^2\rho\sigma_e^2)\mathbf{I}),$$

and hence

$$\begin{aligned} &p(\mathbf{y}|\mathbf{v}=\mathbf{e}_i,\widehat{\mathbf{G}}) = \\ &\quad \mathbb{E}_z\left[\frac{1}{(\pi(1+|z|^2\rho\sigma_e^2))^N}\exp\left(\frac{-\|\mathbf{y}-\sqrt{\rho}\widehat{\mathbf{g}}_i z\|^2}{1+|z|^2\rho\sigma_e^2}\right)\right] \\ &\quad \triangleq f_i(\mathbf{y},\widehat{\mathbf{G}}). \end{aligned}$$

Therefore,

$$p(\mathbf{v}=\mathbf{e}_i|\widehat{\mathbf{G}},\mathbf{y}) = \frac{f_i(\mathbf{y},\widehat{\mathbf{G}})}{\sum_{j=1}^M f_j(\mathbf{y},\widehat{\mathbf{G}})}. \tag{11}$$

Using (11) in (10), combining the resulting expression with (9), and substituting (9) and (8) in (7) proves the proposition. □

Let T be the coherence interval of the channel in number of channel uses. Let T_τ and T_d denote the training interval and data transmission interval, respectively. Also, let ρ_τ and ρ_d denote the training and data SNR, respectively. Then, by conservation of time and energy

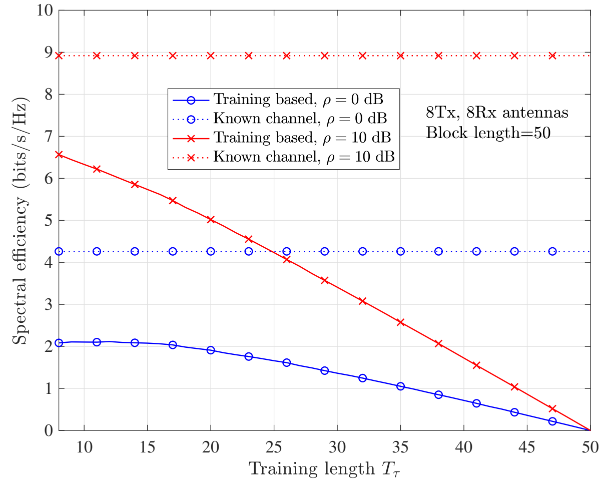
$$\begin{aligned} T &= T_\tau + T_d, \\ \rho T &= \rho_\tau T_\tau + \rho_d T_d. \end{aligned} \tag{12}$$

We now express the spectral efficiency of spatial modulation after accounting for the training time and energy in Proposition 1. With an abuse of notation, we redefine f_i from Eq. (5) by replacing ρ with ρ_d:

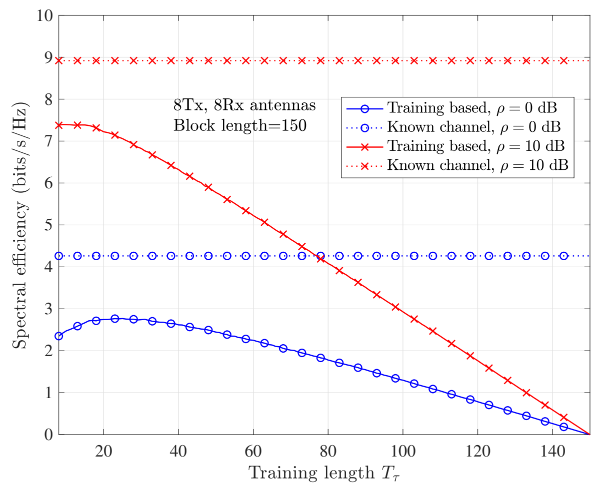
$$\begin{aligned} &f_i(\mathbf{y},\widehat{\mathbf{G}}) = \\ &\quad \mathbb{E}_z\left[\frac{1}{(\pi(1+|z|^2\rho_d\sigma_e^2))^N}\exp\left(\frac{-\|\mathbf{y}-\sqrt{\rho_d}\widehat{\mathbf{g}}_i z\|^2}{1+|z|^2\rho_d\sigma_e^2}\right)\right]. \end{aligned} \tag{13}$$

Then,

$$\begin{aligned} C_\tau \geq &\frac{T-T_\tau}{T}\left[\mathbb{E}_{\widehat{\mathbf{g}}_1}\left[\log_2\left(1+\frac{\|\widehat{\mathbf{g}}_1\|^2\rho_d}{1+\rho_d\sigma_e^2}\right)\right] + \log_2 M - \right. \\ &\left.\mathbb{E}_{\mathbf{y},\widehat{\mathbf{G}}}\left[\sum_{i=1}^M \frac{f_i(\mathbf{y},\widehat{\mathbf{G}})}{\sum_{j=1}^M f_j(\mathbf{y},\widehat{\mathbf{G}})}\log_2\frac{\sum_{j=1}^M f_j(\mathbf{y},\widehat{\mathbf{G}})}{f_i(\mathbf{y},\widehat{\mathbf{G}})}\right]\right], \end{aligned} \tag{14}$$



(a) Block length $T = 50$



(b) Block length $T = 150$

Fig. 1: Spectral efficiency of 8×8 spatial modulation as a function of training length T_τ when training and data powers are equal ($\rho_\tau = \rho_d$).

where [17]

$$\sigma_e^2 = \frac{1}{MN}\mathrm{Tr}\left\{\left(\frac{1}{N}\mathbf{I}_M + \frac{\rho_\tau}{N}\mathbf{X}_\tau\mathbf{X}_\tau^H\right)^{-1}\right\}, \tag{15}$$

where $\mathbf{X}_\tau$ is the $M \times T_\tau$ pilot sequence.

IV. Results and Discussions

Figure 1 shows the spectral efficiency of 8×8 spatial modulation (Eq. (14)) as a function of training duration T_τ for coherence intervals (also referred to as block lengths) $T = 50$ and 150, when the data and training powers are equal. The spectral efficiency with perfectly known CSIR is also shown in the figure. At low SNR (0 dB in the figure), the optimal training requires transmitting more pilots than transmit antennas, and the optimal training duration increases with block length T. For example, the optimal training at 0 dB

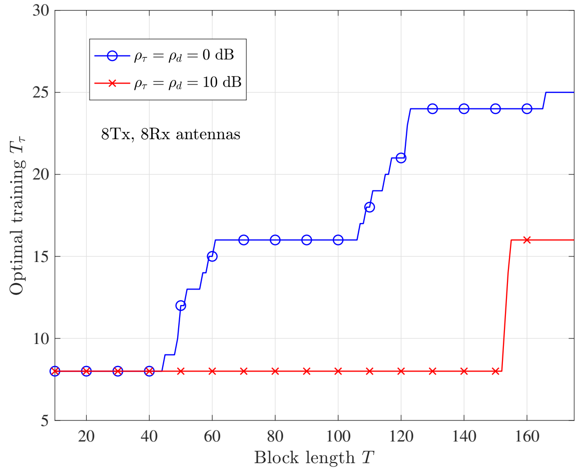


Fig. 2: Optimal training length as a function of block length with equal training and data power.

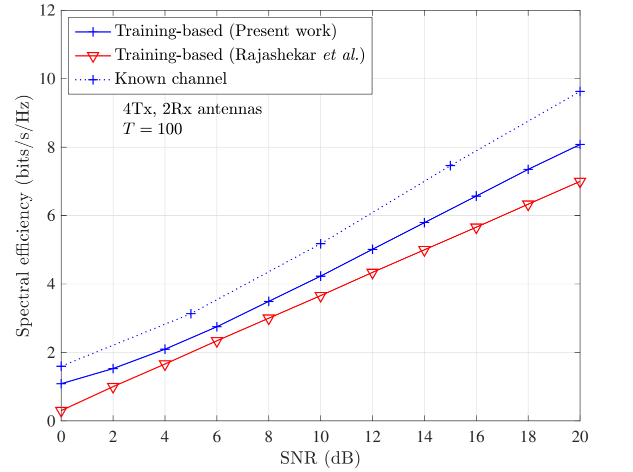


Fig. 4: Comparing Eq. (14) with Rajashekar *et al.* bound [7].

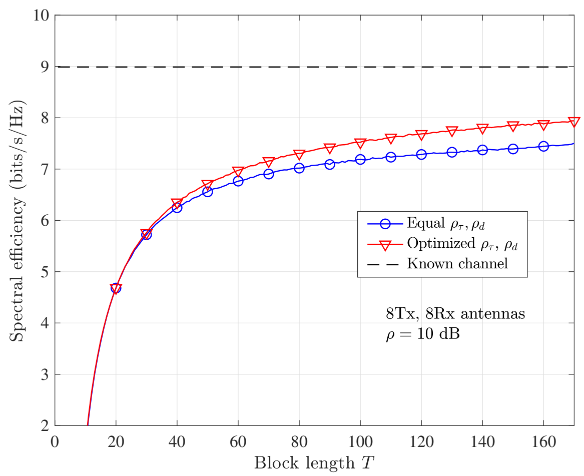


Fig. 3: Spectral efficiency of spatial modulation as a function of block length.

and $T = 150$ requires $T_\tau = 24$ pilots, which is three times the number of transmit antennas. Whereas, at high SNR (10 dB in the figure), optimally, as many training symbols are employed as transmit antennas even at $T = 150$.

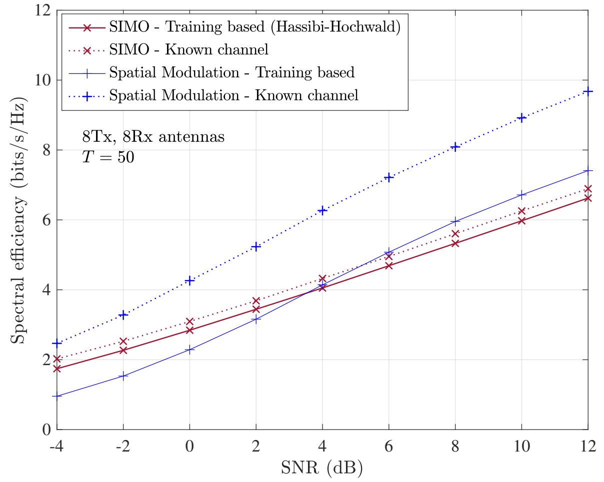
To further illustrate these observations, Fig. 2 shows the optimal training length as a function of block length at 0 dB and 10 dB SNR. At 0 dB the optimal training length increases rapidly with block length. Whereas, at 10 dB, the optimal training length is equal to the number of transmit antennas ($T_\tau = M = 8$) upto block length 150, beyond which it increases to $T_\tau = 2M = 16$.

Figure 3 shows the spectral efficiency of 8×8 spatial modulation as a function of block length under two cases: $i)$ equal ρ_τ, ρ_d (optimized over training time) and $ii)$ optimized ρ_τ, ρ_d. The spectral efficiency with known CSIR is also shown for reference. Allowing the training and data powers to vary improves the spectral efficiency compared with equal training and data powers. Also, the gain with optimized ρ_τ, ρ_d compared with equal ρ_τ, ρ_d increases with block length.

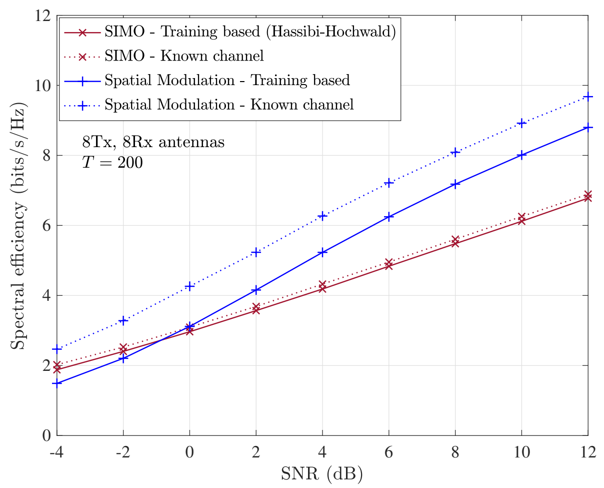
In Fig. 4, we compare the spectral efficiency bound in Eq. (14) with Rajashekar *et al.* [7]. Both the results are optimized with respect to training and data powers. The bound in Eq. (14) is tighter than Rajashekar *et al.* bound across all SNR values. For example, at 6 bits/s/Hz, our bound is tighter by 2.5 dB.

Figure 5 shows the spectral efficiency of spatial modulation (with optimal ρ_τ, ρ_d) as a function of SNR, for block lengths $T = 50$ and 200. The figure also shows the spectral efficiency of spatial modulation with CSIR, SIMO with training [18], and SIMO with CSIR. Up to a certain critical SNR, which depends on the block length, the training-based spectral efficiency of spatial modulation is less than SIMO (e.g., 4 dB for $T = 50$). Figure 6 shows this critical SNR as a function of block length, demonstrating that spatial modulation has smaller spectral efficiency than SIMO in highly dynamic channels. This is a natural consequence of the higher training overhead of spatial modulation (M pilots) compared with SIMO (one pilot). Spatial modulation has the same training overhead as MIMO, but lacks its spatial multiplexing gain. Fig. 6 shows that spatial modulation achieves higher spectral efficiency compared with SIMO in slowly varying channels.

As indicated by the above results, spatial modulation is inferior to SIMO at low-SNR and in highly dynamic channels. The spectral efficiency of spatial modulation can be further improved by choosing the optimal antenna alphabet. For a given coherence interval and operating SNR, the optimal antenna alphabet achieves the best tradeoff between the training overhead and the achievable rate. When $M^* = 1$, the optimal transmission scheme reduces to SIMO. Figure 7 shows the optimal antenna alphabet for spatial modulation as a function of SNR and channel dynamics. While SIMO is optimal in low-SNR and low-block length regime, the optimal antenna alphabet size increases with SNR and block length. Therefore, if M antennas are available at the transmitter, the transmitter-



(a) Block length $T = 50$



(b) Block length $T = 200$

Fig. 5: Spectral efficiency of 8×8 spatial modulation as function of SNR.

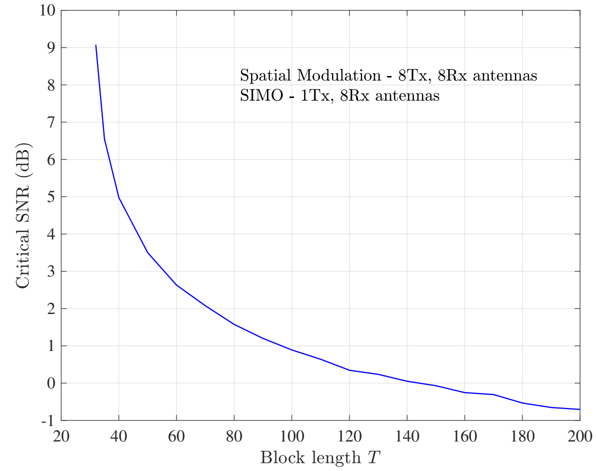


Fig. 6: Critical SNR for SIMO vs. spatial modulation. Above critical SNR, spatial modulation achieves a higher rate, and below critical SNR, SIMO achieves a higher rate.

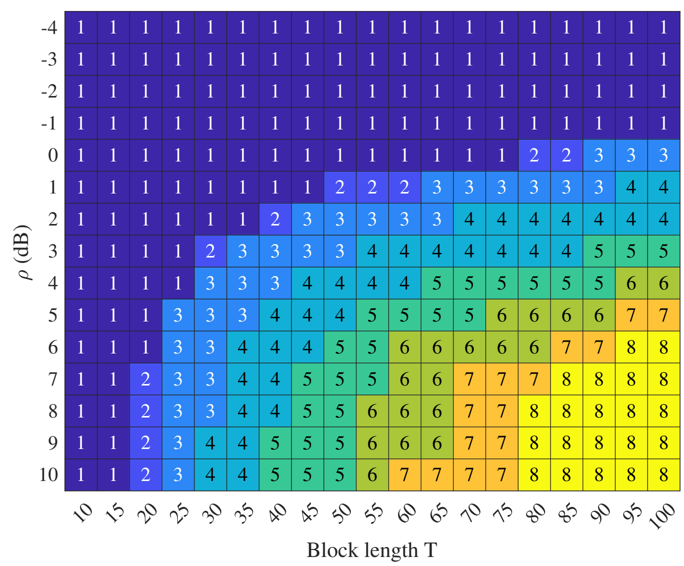


Fig. 7: Optimal antenna alphabet with one RF chain.

receiver pair can agree on employing $M^*(1 \leq M^* \leq M)$ antennas for spatial modulation based on SNR and channel dynamics.

V. Conclusions

We studied the spectral efficiency of spatial modulation with channel training accounting for the training overhead and error. Our results showed that in highly dynamic channels and at low-SNR, spatial modulation has lesser spectral efficiency than SIMO. Our results also provided the optimal antenna alphabet for spatial modulation subject to pilots and training. Extending the present work to generalized spatial modulation and to multi-user index modulation systems will be considered in the journal version of this work.

References

[1] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2228–2241, 2008.

[2] 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, 2014.

[3] M. Wen, B. Zheng, K. J. Kim, M. Di Renzo, T. A. Tsiftsis, K. Chen, and N. Al-Dhahir, "A survey on spatial modulation in emerging wireless systems: Research progresses and applications," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 9, pp. 1949–1972, 2019.

[4] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 168–175, 2016.

[5] E. Basar, M. Wen, R. Mesleh, M. Di Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16 693–16 746, 2017.

[6] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Communications Letters*, vol. 12, no. 4, pp. 225–227, 2008.

[7] R. Rajashekar, K. Hari, and L. Hanzo, "Reduced-complexity ML detection and capacity-optimized training for spatial modulation systems," *IEEE Transactions on Communications*, vol. 62, no. 1, pp. 112–125, 2014.

[8] B. Shamasundar and A. Nosratinia, "Spectral efficiency of multi-antenna index modulation," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 3291–3295.

[9] ——, "On the capacity of index modulation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 9114–9126, 2022.

[10] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Performance of spatial modulation in the presence of channel estimation errors," *IEEE Communications Letters*, vol. 16, no. 2, pp. 176–179, 2012.

[11] X. Wu, H. Claussen, M. Di Renzo, and H. Haas, "Channel estimation for spatial modulation," *IEEE Transactions on Communications*, vol. 62, no. 12, pp. 4362–4372, 2014.

[12] X. Wu, M. Di Renzo, and H. Haas, "Effect of pilot ratio on channel estimation for spatial modulation," in *2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2013, pp. 144–148.

[13] S. Sugiura and L. Hanzo, "Effects of channel estimation on spatial modulation," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 805–808, 2012.

[14] M. M. U. Faiz, S. Al-Ghadhban, and A. Zerguine, "Recursive least-squares adaptive channel estimation for spatial modulation systems," in *2009 IEEE 9th Malaysia International Conference on Communications (MICC)*, 2009, pp. 785–788.

[15] X. Wu, M. Di Renzo, and H. Haas, "Optimal power allocation for channel estimation in spatial modulation," in *2014 IEEE International Conference on Communications (ICC)*, 2014, pp. 5481–5485.

[16] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, 2006.

[17] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.

[18] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.