



**HAL**  
open science

## Time-varying space-only code: A new paradigm for coded MIMO communication

Dieter Duyck, Sheng Yang, Fambirai Takawira, Joseph Boutros, Marc Moeneclaey

► **To cite this version:**

Dieter Duyck, Sheng Yang, Fambirai Takawira, Joseph Boutros, Marc Moeneclaey. Time-varying space-only code: A new paradigm for coded MIMO communication. 2012 5th International Symposium on Communications, Control and Signal Processing (ISCCSP), May 2012, Roma, France. 10.1109/ISCCSP.2012.6217779 . hal-01798790

**HAL Id: hal-01798790**

**<https://hal.science/hal-01798790>**

Submitted on 30 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Time-Varying Space-Only Code: A New Paradigm for Coded MIMO Communication

Dieter Duyck, Sheng Yang, Fambirai Takawira, Joseph J. Boutros, and Marc Moeneclaey

**Abstract**—Multiple antennas are used to increase reliability and bit rate for a given bandwidth. For a fixed transmission rate, discrete input alphabets and without channel state information at the transmitter, optimal space-time codes (STCs) achieving both gains (full rate and full diversity) are well known. However, the complexity of maximum likelihood decoding increases exponentially with the number of space and time dimensions of the STC. Previous work reducing the complexity of decoding STCs has focused on the decoding algorithm, because the dimensions of the STC cannot be reduced without losing rate or diversity for *uncoded* communication. However, for *coded* communication (assuming the presence of an outer code), the dimensions of the STC may be reduced. We propose a new full-rate and full-diversity suboptimal *time-varying* space-only code, adding a new dimension to the work on complexity reduction. Full diversity is proved in terms of the outage probability.

## I. INTRODUCTION

Channels with multiple antenna at in- and output (MIMO) have more degrees of freedom which can be used to increase spectral efficiency (through spatial multiplexing) and reliability (through diversity). When the spectral efficiency is taken constant, both diversity and multiplexing gains can be fully achieved simultaneously, through full-rate full-diversity space-time codes (STCs). STCs for uncoded communication, i.e., without outer code, have been designed by minimizing the pairwise error probability under maximum likelihood (ML) decoding (e.g. [1], [8], [12], [15]). For coded communication, STCs have been designed assuming a genie (e.g. [7]). An STC taking into account both criteria was proposed in [4].

In the worst case, without sphere decoding [18], the complexity of exhaustive ML decoding of an  $u \times v$  STC (represented in matrix form) at the receiver of a MIMO channel scales as  $O(M^{uv})$ , where  $M$  is the modulation size. The optimal STCs, achieving full transmit diversity by coding over  $n$  space dimensions and  $n$  time dimensions, are of dimension  $n \times n$ , where  $n$  is the number of transmit antennas. The decoding complexity may be too high, e.g.  $O(16^{16})$  when 16-QAM and  $n = 4$  transmit antennas are used. Orthogonal STCs [15] have much lower complexity, but in general at the expense of a significantly reduced bit rate. Significant effort has been put in reducing the decoding complexity of non-orthogonal STCs (e.g. see [10], [14] and references therein).

Dieter Duyck and Marc Moeneclaey are with Ghent University, St-Pietersnieuwstraat 41, B-9000 Gent, Belgium, {dduyck,mm}@telin.ugent.be.

Sheng Yang is with Supélec, 3 rue Joliot-Curie, 91192 Gif sur Yvette, France, sheng.yang@supelec.fr

Fambirai Takawira is with the University of Kwazulu Natal, King George V Avenue, Durban, South Africa, ftakaw@ukzn.ac.za

Joseph J. Boutros is with Texas A&M University at Qatar, PO Box 23874 Doha, Qatar, boutros@tamu.edu

For example, in [10], [14], the complexity of decoding squared  $2 \times 2$  STCs for the  $2 \times 2$  MIMO channel was said to be reduced to  $O(M^{2.5})$ . However, the achieved reductions are not valid for all STCs and all MIMO channels.

We add a new dimension to the quest for low-complexity STCs, by reducing the *dimensions* of the STC itself, without compromising on rate or diversity. In the context of coded communication, i.e., assuming an error-correcting code with coding rate  $R_c < 1$ , we prove that the outage probability<sup>1</sup> of a  $2 \times 2$  MIMO channel can have full transmit diversity using  $2 \times 1$  (i.e. space-only) STCs<sup>2</sup>, limiting its dimension by a factor of two with respect to optimal STCs. This idea is inspired by [19], where the DMT has been analysed for Gaussian alphabets in a multihop scenario.

*Notation:* we write scalars, vectors and matrices as  $x$ ,  $\mathbf{x}$  and  $X$ . The Landau symbols  $f(n) = O(g(n))$  and  $f(n) = \Omega(g(n))$  mean  $f(n) \leq kg(n)$  and  $f(n) \geq kg(n)$  for some positive  $k$ . The equation sign  $f(\gamma) \doteq g(\gamma)$ , introduced in [16], means that  $\lim_{\gamma \rightarrow \infty} \frac{\log f(\gamma)}{\log \gamma} = \lim_{\gamma \rightarrow \infty} \frac{\log g(\gamma)}{\log \gamma}$ . Similar meanings hold for  $\leq$  and  $\geq$ .

## II. SYSTEM MODEL

We consider a point-to-point MIMO channel  $H = [h_{i,j}] \in \mathbb{C}^{r \times n}$  with  $n$  transmit antennas and  $r$  receive antennas, where  $h_{i,j}, h_{i,j} \sim \mathcal{CN}(0, 1)$  is the complex path gain from transmit antenna  $j$  to receive antenna  $i$ . We assume that all path gains are independent. The channel state information is perfectly known at the receiver side, but unknown at the transmitter side. The channel is assumed to vary slowly, so that it remains constant during the transmission of at least one codeword. At the  $t$ -th channel use, the output of the matched filter at the destination is

$$\boldsymbol{\mu}_t = \sqrt{\gamma} H \boldsymbol{\zeta}_t + \boldsymbol{\nu}_t, \quad t = 1, \dots, \infty, \quad (1)$$

where  $\boldsymbol{\mu}_t$  and  $\boldsymbol{\nu}_t$  are in  $\mathbb{C}^{r \times 1}$ ,  $\boldsymbol{\zeta}_t \in \mathbb{C}^{n \times 1}$  and  $\gamma$  is the average signal-to-noise ratio per symbol (SNR) at each transmit antenna<sup>3</sup>. The additive white Gaussian noise vector  $\boldsymbol{\nu}_t$  has i.i.d. entries,  $\nu_{b,t} \sim \mathcal{CN}(0, 1)$ . The components  $\zeta_{b,t}$  of the transmit vector  $\boldsymbol{\zeta}_t$  satisfy  $\mathbb{E}[|\zeta_{b,t}|^2] = 1, \forall b, t$ . The overall vector belongs to  $\Omega_{\boldsymbol{\zeta}}$ . The overall received signal-to-noise ratio per information bit is denoted as  $\frac{E_b}{N_0}$ , so that  $\frac{E_s}{N_0} = \gamma = \frac{RE_b/N_0}{nr}$ ,

<sup>1</sup>The outage probability is a fundamental and achievable lower bound on the average word error rate (WER) of coded systems [2], [13].

<sup>2</sup>The dimension of the  $2 \times 1$  code suggests that only spatial coding is performed. However, by using a *time-varying* code, also coding in time is performed, without increasing the complexity.

<sup>3</sup>In some papers, the average signal-to-noise ratio at each receive antenna is considered, which is  $n\gamma$ .

where  $R = \log_2(|\Omega_\zeta|)R_c$  is the spectral efficiency and  $R_c$  is the coding rate. The mutual information  $I(\zeta_t; \mu_t|H)$  depends on  $\Omega_\zeta$  and the channel realization  $H$ .

The channel realization  $H$  can be decomposed by a singular value decomposition (SVD) as  $H = U\Sigma V^\dagger$ , where  $U \in \mathbb{C}^{r \times r}$  and  $V \in \mathbb{C}^{n \times n}$  are unitary matrices, uniformly distributed with respect to the Haar measure [9],  $V^\dagger$  is the Hermitian transpose of  $V$ , and  $\Sigma \in \mathbb{R}^{r \times n}$  is a diagonal matrix with the non-negative singular values  $\sigma_b, b = 1, \dots, \min(r, n)$  of  $H$  on its diagonal. Because  $H$  is known at the receiver, the following transformation can be performed,

$$\mathbf{y}_t = U^\dagger \mu_t = \sqrt{\gamma} \Sigma V^\dagger \zeta_t + \mathbf{w}_t, \quad t = 1, \dots, \infty \quad (2)$$

where  $\mathbf{w}_t$  follows the same distribution as  $\nu_t$ . Denote  $\mathbf{x}_t = V^\dagger \zeta_t$ , where  $\mathbf{x}_t \in \Omega_\mathbf{x}$ . Because  $U^\dagger \mu_t$  is an invertible transformation, no information is lost, i.e.,

$$I(\mathbf{x}_t; \mathbf{y}_t | \Sigma, V) = I(\zeta_t; \mu_t | H). \quad (3)$$

The outage probability is expressed as

$$P_{\text{out}} = \Pr(\mathbb{E}_t [I(\mathbf{x}_t; \mathbf{y}_t | \Sigma, V)] < R), \quad (4)$$

where  $\mathbb{E}_t[\cdot]$  is the temporal mean. The SNR-exponent of the outage probability, known as the diversity order, is

$$d_{\text{out}} = \lim_{\gamma \rightarrow \infty} -\frac{\log P_{\text{out}}}{\log \gamma}. \quad (5)$$

Traditionally,  $n \times n$  STCs (containing  $n^2$  elements) are obtained as follows. Through a linear precoder in  $\mathbb{C}^{n^2 \times n^2}$ , a symbol vector of length  $n^2$  is transformed into another vector of length  $n^2$ . Folding this vector into  $n$  column vectors  $\zeta_t$ ,  $t = 1, \dots, n$ , an  $n \times n$  STC is obtained. Here, we propose a new unitary linear precoder  $P_t \in \mathbb{C}^{n \times n}, t = 1, \dots, \infty$ , so that  $\zeta_t = P_t \mathbf{z}_t$ , where  $z_{b,t} \in \Omega_z$  (and  $\mathbf{z}_t \in \Omega_z$ )<sup>4</sup>,  $\mathbb{E}[|z_{b,t}|^2] = 1, \forall b, t$ . Note that  $P_t$  is time-variant, so that also  $\Omega_\zeta$  changes in time. The overall linear precoder over  $N$  channel uses is therefore block diagonal with the matrices  $P_t$  on its diagonal. We denote  $m = n \log_2 |\Omega_z|$ , which is the maximal number of bits that can be conveyed per channel use. For simplicity, we assume  $n = r = 2$  in the remainder of the paper. We denote our new precoder type by the EMI code, where EMI (*Ergodic Mutual Information*) refers to the temporal mean of the mutual information.

### III. ACHIEVING FULL RATE AND FULL DIVERSITY WITH TIME-VARYING SPACE-ONLY PRECODERS

We first consider a toy example, where we remove  $V$  in (2), so that  $\mathbf{x}_t = \zeta_t$ , and consider a fixed precoder  $P$ , so that the time-index  $t$  can be dropped. Hence, the toy channel comprises 2 parallel standard flat fading channels and the precoder can be used to achieve signal-space diversity (SSD) [3], [5].

Let us recall a well known result [3], [5] on SSD. In a point-to-point flat parallel fading channel with a fixed precoder  $P$ , full diversity is achieved for any coding rate  $R_c < 1$ ,

<sup>4</sup>The constellation  $\Omega_z$  is considered to be the Cartesian product of  $\Omega_z$ :  $\Omega_z = (\Omega_z)^n$ , which means that all the components of  $\mathbf{z}_t$  have identical marginal distributions. In the case that  $n = 2$ , it is easy to prove that under a unitary transformation, this property is maintained, so that also  $\Omega_\zeta = (\Omega_\zeta)^n$ .

if and only if  $s_b \neq 0, \forall b$ ,  $\mathbf{x}' \neq \mathbf{x}$ , where  $s_b = x_b - x'_b$ . Hence, precoding parallel fading channels does not achieve full diversity when *bad* precoders are used.

*Definition 1:* We define *bad* precoders  $P_{\text{bad}}$  as the set of precoders that transform  $\Omega_z$  in  $\Omega_\mathbf{x}$  so that  $\exists b \in \{1, 2\}, \mathbf{x}, \mathbf{x}' \neq \mathbf{x}$ , satisfying  $x_b = x'_b$ .

Back to Eq. (2), we see that  $V$  is a *random* unitary precoder and denote  $V_t = V^\dagger P_t$ , so that  $\mathbf{x}_t = V_t \mathbf{z}_t$ . At first, consider a unitary but *fixed* precoder  $P_t = P$ . Hence,  $V_t$  has the same distribution as  $V$ . The probability of having *bad* precoders,  $V_t \in P_{\text{bad}}$ , is zero. However, it is well known that this scheme only achieves the receive diversity. Indeed,  $p(V_t) = p(V)$ , so that the precoder  $P$  does not change the outage probability with respect to the case without precoder. We explain the loss of transmit diversity by defining *corrupt* precoders.

*Definition 2:* We define *corrupt* precoders  $P_c$  as the set of precoders that transform  $\Omega_z$  in  $\Omega_\mathbf{x}$  so that  $\exists b, \mathbf{x}, \mathbf{x}' \neq \mathbf{x}$ , satisfying  $|s_b| \leq \gamma^{-0.5}$ .

*Lemma 1:* In a point-to-point  $2 \times 2$  MIMO flat fading channel with a fixed  $2 \times 2$  precoder  $P_t = P$ , full transmit diversity is not achieved (i.e.,  $d_{\text{out}} < 4$ ) for any coding rate  $R_c < 1$  due to corrupt precoders  $V_t = V^\dagger P = V'$ .

*Proof:* We give the proof in App. A, mainly to illustrate the proof techniques used in the following proposition. ■ Interestingly,  $\lim_{\gamma \rightarrow \infty} -\frac{\log(p(P_c))}{\log \gamma} = 1$ , or the occurrence probability of corrupt precoders is non-negligible with respect to the error rate as it decreases significantly slower than  $\frac{1}{\gamma^2}$ .

As a solution to this problem, we propose to marginalize the effect of corrupt precoders by averaging the mutual information over all precoders during the transmission of one codeword. Therefore, we let  $P_t$  be uniformly distributed in the Stiefel manifold. Since  $V$  is unitary and independent of  $P_t$  and  $P_t$  is invariant to unitary transformation,  $V_t = V^\dagger P_t$  is also uniformly distributed in the Stiefel manifold [9].

*Proposition 1:* In a point-to-point  $2 \times 2$  MIMO flat fading channel using a coding rate  $R_c < 1$  and using  $2 \times 2$  precoders  $P_t$ , randomly generated for each channel use, following a uniform distribution with respect to the Haar measure, full diversity is achieved.

*Proof:* See Appendix B. ■

When  $P_t$  is changed at each channel use, we denote the code simply by *EMI code*. When  $P_t$  is changed only a finite number of times  $N$  during the transmission of a codeword, we denote the code by *EMI- $N$  code*. As expected, the performance of the EMI- $N$  code should converge to the performance of the EMI-code for increasing  $N$ .

### IV. NUMERICAL RESULTS

We determined the outage probability by means of Monte-Carlo simulations. We assume a  $2 \times 2$  MIMO channel, with  $\Omega_z = 4$ -QAM, and assume a coding rate  $R_c = 0.9$  so that the spectral efficiency is  $R = 0.9 * 2 * 2 = 3.6$  bpcu. We compare the outage probability for different STCs (the Aladdin-Pythagoras code and our proposed EMI code) with the outage probability corresponding to a Gaussian input (Fig. 1). We also show that the performance of the EMI-10 code is close to that of the EMI-code. With the current system

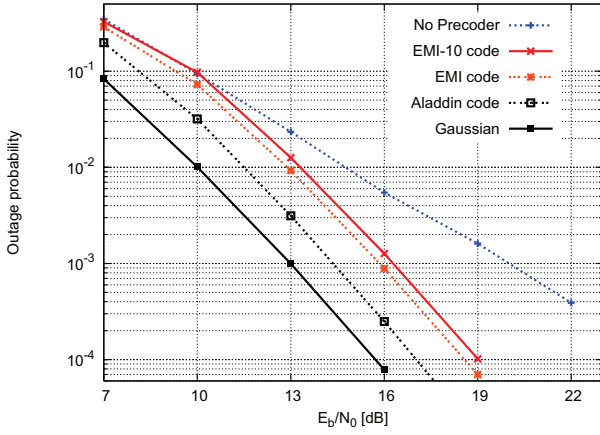


Fig. 1. We compare the outage probability of several STCs for  $2 \times 2$  MIMO. We use  $\Omega_z = 4$ -QAM and a coding rate  $R_c = 0.9$  so that  $R = 3.6$  bpcu.

parameters, the performance loss of the EMI-code with respect to Aladdin-Pythagoras code is approximately 1 dB. However, the complexity is much lower, which becomes considerable when using larger modulation sizes and more antennas.

## V. CONCLUSION

Through a lossless transformation at the receiver, the MIMO channel is converted into a parallel channel including a random precoder  $V$  that depends on the channel realization. The new channel has an identical mutual information. Using this new channel model, we study the fundamental reason why full diversity is not achieved without space-time coding. A set of corrupt precoders, with a non-negligible probability of occurrence, causes the loss in transmit diversity order. Therefore, we propose a new full-diversity full-rate space-time code which we refer to as the EMI-code. The latter averages the mutual information of the MIMO channel over all random precoders  $V$ , thereby marginalizing corrupt precoders. We prove that the new precoder achieves full diversity if an error-correcting code with coding rate  $R_c < 1$  is used. The EMI-code is said to be full rate because  $n$  new symbols are transmitted each channel use, where  $n$  is the number of transmit antennas. The main asset of this new coding scheme is that its complexity of decoding is significantly lower, as its dimension is  $n$  instead of  $n^2$ , at the expense of a small loss in performance. In a future work, we shall extend this work to general  $n \times r$  MIMO channels.

## APPENDIX

### A. Proof of Lemma 1

The outage probability is

$$P_{\text{out}} = \int_{P_c} p(V') P_{\text{out}|V'} dV' + \int_{\bar{P}_c} p(V') P_{\text{out}|V'} dV' \quad (6)$$

$$\geq P_{\text{out}|V'=V_l} \int_{P_c} p(V') dV' = P_{\text{out}|V'=V_l} \Pr(P_c) \quad (7)$$

where  $\bar{P}_c$  is the complement of  $P_c$  and  $V_l = \operatorname{argmin}_{V' \in P_c} P_{\text{out}|V'}$ .

First, we prove that the probability of occurrence of corrupt precoders,  $\Pr(V' \in P_c)$ , behaves as  $\frac{1}{\gamma}$ . Consider  $\mathbf{v} =$

$V' \frac{\mathbf{z}_t - \mathbf{z}'_t}{\|\mathbf{z}_t - \mathbf{z}'_t\|}$  with  $\mathbf{z}'_t \neq \mathbf{z}_t$ .  $V'$  being uniformly distributed in the Stiefel manifold, it is readily shown that  $\mathbf{v}$  is uniformly distributed in  $\mathbf{v}^\dagger \mathbf{v} = 1$  with respect to the Haar measure [9]. It implies that  $|v_b|^2$  follows a uniform distribution in  $[0, 1]$ ,  $\forall b = 1, 2$ . Because  $|s_b|^2 = \|\mathbf{z}_t - \mathbf{z}'_t\|^2 |v_b|^2$ , and  $\|\mathbf{z}_t - \mathbf{z}'_t\|^2$  is independent of  $\gamma$ , it is easy to show that  $\Pr(V' \in P_c) = \Pr(|s_b|^2 \leq \frac{1}{\gamma}) \doteq \frac{1}{\gamma}$ , so that  $P_{\text{out}} \geq \frac{1}{\gamma} P_{\text{out}|V'=V_l}$ .

Now, let us study  $P_{\text{out}|V'=V_l}$ . The mutual information between discrete input and output of a parallel channel is well known [5], [6], [17] and recalled in Eq. (9), where  $d^2(\mathbf{c}, \mathbf{d}) = \sum_{b=1}^2 |c_b - d_b|^2$ . As in [6], [16], we define the ordered normalized fading gains  $\alpha_b = -\frac{\log \sigma_b^2}{\log \gamma}$ ,  $\alpha_1 \geq \alpha_2$ , where the joint distribution  $p(\boldsymbol{\alpha})$  is given by [16]

$$p(\boldsymbol{\alpha}) = (\log \gamma)^2 e^{-\gamma^{-\alpha_1} - \gamma^{-\alpha_2}} \gamma^{-\alpha_1 - \alpha_2} (\gamma^{-\alpha_2} - \gamma^{-\alpha_1})^2. \quad (8)$$

The mutual information can be expressed as in Eq. (10), where

$$f(\alpha_b, s_b, w_b) = e^{-\gamma^{1-\alpha_b} |s_b|^2 - 2\sqrt{\gamma^{1-\alpha_b}} \mathcal{R}\{w_b s^*(b)\}},$$

$s(b) = (x_b - x'_b)$  and  $\mathcal{R}\{\cdot\}$  takes the real part. The normalized fading gains  $\boldsymbol{\alpha} \in \mathbb{R}^2$ , but because we only want to find the SNR-exponent, we only consider the region that is dominant,  $\boldsymbol{\alpha} \in \mathbb{R}_+^2$  (see [6], [16]). In order to determine the diversity order (Eq. (5)), we can apply the dominated convergence theorem [6], so that

$$\begin{aligned} \lim_{\gamma \rightarrow \infty} \mathbb{E}_{\mathbf{w}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_{\mathbf{x}}} \prod_{b=1}^2 f(\alpha_b, s_b, w_b) \right) \right] \\ = \mathbb{E}_{\mathbf{w}} \left[ \lim_{\gamma \rightarrow \infty} \log_2 \left( \sum_{\mathbf{x}' \in \Omega_{\mathbf{x}}} \prod_{b=1}^2 f(\alpha_b, s_b, w_b) \right) \right] \quad (11) \end{aligned}$$

For corrupt precoders<sup>5</sup>,  $\forall k \in \{1, 2\}$ ,  $\exists \mathbf{x} = \mathbf{x}_c, \mathbf{x}' \neq \mathbf{x}_c$ , satisfying  $|s_k| \leq \gamma^{-0.5}$ . In that case, we observe that for large  $\gamma$ ,  $\lim_{\gamma \rightarrow \infty} f(\alpha_k, s_k, w_k) = 1$  if  $\alpha_k > 0$ . Consider for example  $k = 2$ . For large  $\gamma$ , the mutual information is

$$I(\mathbf{x}, \mathbf{y}|\Sigma) = m - 2^{-m} \sum_{\mathbf{x} \in \Omega_{\mathbf{x}}} \mathbb{E}_{\mathbf{w}} [\log_2 (1 + g(\boldsymbol{\alpha}, \mathbf{x}))],$$

$$\text{where } g(\boldsymbol{\alpha}, \mathbf{x}_c) = \begin{cases} O(e^{-\gamma^{1-\alpha_1}}), & \text{if } \alpha_1 < 1 \\ \Omega(e^{-(\frac{1}{\gamma})^{\alpha_1-1}}), & \alpha_1 > 1 \end{cases}.$$

In the event that  $\boldsymbol{\alpha} \in \mathcal{B}_\epsilon$ , the mutual information is  $m - \Omega(e^{-\gamma^{-\epsilon}})$ , where  $\mathcal{B}_\epsilon = \{\boldsymbol{\alpha} : \alpha_1 \geq \alpha_2 > 0 : \alpha_1 \geq 1 + \epsilon\}$ . Hence, there exists a coding rate so that the spectral efficiency  $mR_c$  is larger then or equal to the mutual information. For this coding rate, the outage probability is

$$P_{\text{out}|V'=V_l} \geq \int_{\mathcal{B}_\epsilon} p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}. \quad (12)$$

Noting that  $P_{\text{out}} \geq \frac{1}{\gamma} P_{\text{out}|V'=V_l}$  and following the same lines as in [16] (see also [5], [11]) to obtain the SNR-exponent from

<sup>5</sup>The definition of corrupt precoders involves only that  $|s_k^{(1)}| \leq \gamma^{-0.5}$  for one particular combination of  $\mathbf{x}^{(1)} = (x_{1,c}, x_{2,c}), \mathbf{x}'^{(1)} = (x'_1, x'_2) \neq \mathbf{x}$  and  $k \in \{1, 2\}$ . Without loss of generality, assume  $k = 1$ . Because the components of  $\mathbf{x}$  have identical marginal distributions, the points  $\mathbf{x}^{(2)} = (x_{2,c}, x_{1,c})$  and  $\mathbf{x}'^{(2)} = (x'_2, x'_1)$  are also in  $\Omega_{\mathbf{x}}$ , so that  $|s_2^{(2)}| \leq \gamma^{-0.5}$ , where  $\mathbf{s}^{(2)} = \mathbf{x}^{(2)} - \mathbf{x}'^{(2)}$ .



$$I(\mathbf{x}, \mathbf{y} | \Sigma, V') = m - 2^{-m} \sum_{\mathbf{x} \in \Omega_{\mathbf{x}}} \mathbb{E}_{\mathbf{y} | \mathbf{x}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_{\mathbf{x}}} \exp \left[ (d^2(\mathbf{y}, \sqrt{\gamma} \Sigma \mathbf{x}) - d^2(\mathbf{y}, \sqrt{\gamma} \Sigma \mathbf{x}')) \right] \right) \right] \quad (9)$$

$$I(\mathbf{x}, \mathbf{y} | \Sigma, V') = m - 2^{-m} \sum_{\mathbf{x} \in \Omega_{\mathbf{x}}} \mathbb{E}_{\mathbf{w}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_{\mathbf{x}}} \prod_{b=1}^2 f(\alpha_b, s_b, w_b) \right) \right], \quad (10)$$

(12), we obtain that

$$d_{\text{out}} \leq d_{\text{out}}(\epsilon) = 1 + \inf_{\alpha \in \mathcal{B}_\epsilon} \sum_{b=1}^2 (2b-1) \alpha_b. \quad (13)$$

It is clear that the infimum is  $1 + \epsilon$ , which is achieved when  $\alpha_2 = 0$  and  $\alpha_1 = 1 + \epsilon$ . This holds for each  $\epsilon > 0$ , and the bound in Eq. (13) can be made tight taking the infimum  $\inf_{\epsilon} d_{\text{out}}(\epsilon)$  (see e.g. [5], [11]), so that we obtain that  $d_{\text{out}} \leq 2$ .

### B. Proof of Prop. 1

From Eq. (4), the outage probability can be upper bounded

$$P_{\text{out}} = \Pr(\mathbb{E}_t [I(V_t)] \leq R) \leq \Pr(I_{\text{in}} \leq R), \quad (14)$$

where  $I_{\text{in}} \leq \mathbb{E}_t [I(V_t)]$ , where  $I(V_t)$  denotes  $I(\mathbf{x}_t; \mathbf{y}_t | \Sigma, V_t)$ . Similarly to  $P_c$ , we define a larger set  $S_c$ , which is the set of precoders that transform  $\Omega_{\mathbf{z}}$  in  $\Omega_{\mathbf{x}}$  so that  $\exists b, \mathbf{x}, \mathbf{x}' \neq \mathbf{x}$ , satisfying  $|s_b| \leq (\log \gamma)^{-p}$ , for any  $p > 0$  (note that  $\Pr(S_c) \rightarrow 0$  for large  $\gamma$ ). We can now write  $\mathbb{E}_t [I(V_t)]$  as

$$\mathbb{E}_t [I(V_t)] = \int_{S_c} p(V_t) I(V_t) dV_t + \int_{\bar{S}_c} p(V_t) I(V_t) dV_t. \quad (15)$$

A lower bound  $I_{\text{in}}$  is

$$I_{\text{in}} = I(V_l) \int_{V_t \in \bar{S}_c} p(V_t) dV_t = I(V_l)(1 - \Pr(S_c)), \quad (16)$$

where  $V_l = \text{argmin}_{V_t \in \bar{S}_c} I(V_t)$  (worst case). By definition of  $\bar{S}_c$ ,  $|s_b| > (\log \gamma)^{-p}$ , so that for large  $\gamma$ ,

$$f(\alpha_b, s_b, w_b) = \begin{cases} \Omega \left( e^{-\left(\frac{1}{\gamma}\right)^{\alpha_b - 1}} \right), & \alpha_b > 1 \\ O \left( e^{-\gamma^{1 - \alpha_b}} \right), & \alpha_b < 1. \end{cases} \quad (17)$$

Hence, if  $\sum_{b=1}^2 \mathbf{1}\{\alpha_b < 1\} \geq 1$ , then  $I(V_l) \rightarrow m$  for large  $\gamma$ . More specifically, consider

$$\mathcal{A}_\epsilon = \left\{ \alpha : \sum_{b=1}^2 \mathbf{1}\{\alpha_b < 1 - \epsilon\} \geq 1 \right\}, \epsilon > 0$$

so that  $I(V_l | \alpha \in \mathcal{A}_\epsilon) = m - O(e^{-\gamma^\epsilon})$ . Note that for large  $\gamma$  and  $R_c = 1 - \epsilon_2$ ,  $\epsilon_2 > 0$ ,

$$\lim_{\gamma \rightarrow \infty} \Pr(I(V_l | \alpha \in \mathcal{A}_\epsilon)(1 - \Pr(S_c)) \leq R_c m) = 0.$$

Hence, for large  $\gamma$  and any  $\epsilon, \epsilon_2 > 0$ , the outage probability is upper bounded, i.e.,  $P_{\text{out}} \leq \Pr(\bar{\mathcal{A}}_\epsilon)$ , where  $\bar{\mathcal{A}}_\epsilon = \left\{ \alpha : \sum_{b=1}^2 \mathbf{1}\{\alpha_b \geq 1 - \epsilon\} = 2 \right\}$ . Following the same lines as in App. A,

$$d_{\text{out}} \geq \sup_{\epsilon} \inf_{\alpha \in \bar{\mathcal{A}}_\epsilon} \sum_{b=1}^2 (2b-1) \alpha_b \quad (18)$$

By letting  $\epsilon_2 \rightarrow 0^+$ , we obtain that  $d_{\text{out}} = 4$  for any  $R_c < 1$ .

## REFERENCES

- [1] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: a 2x2 full rate space-time code with non-vanishing determinants," *IEEE Trans. on Inf. Theory*, vol. 51, no. 4, pp. 1432-1436, Apr. 2005.
- [2] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. on Inf. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.
- [3] J.J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. on Inf. Theory*, vol. 44, no. 4, pp. 1453-1467, July 1998.
- [4] J.J. Boutros and H. Randriambololona, "The Aladdin-Pythagoras space-time code," in *Proc. IEEE Intern. Symp. on Inf. Theory (ISIT)*, pp. 2823-2827, July 2009.
- [5] A. Guillén i Fàbregas and G. Caire, "Multidimensional coded modulation in block-fading channels," *IEEE Trans. on Inf. Theory*, vol. 54, no. 5, pp. 2367-2372, 2008.
- [6] A. Guillén i Fàbregas and G. Caire, "Coded modulation in the block-fading channel: coding theorems and code construction," *IEEE Trans. on Inf. Theory*, vol. 52, no. 1, pp. 91-114, Jan. 2006.
- [7] N. Gresset, L. Brunel, and J.J. Boutros, "Space-time coding techniques with bit-interleaved coded modulations for MIMO block-fading channels," *IEEE Tr. on Inf. Theory*, vol. 54, no. 5, pp. 2156-2178, May 2008.
- [8] E.R. Larsson and P. Stoica, "Space-Time Block Coding for Wireless Communications," *Cambridge University Press*, 2003.
- [9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.
- [10] L.P. Natarajan, P.K. Srinath and B.S. Rajan, "Generalized Distributive Law for ML Decoding of STBCs," in *Proc. Information Theory Workshop (ITW)*, Paraty, Brasil, Oct. 2011.
- [11] K.D. Nguyen, A. Guillén i Fàbregas, L.K. Rasmussen, "A Tight Lower Bound to the Outage Probability of Discrete-Input Block-Fading Channels," *IEEE Tr. on Inf. Theory*, vol. 53, no. 11, pp. 4314-4322, Nov. 2007.
- [12] C. Oestges and B. Clerckx, "MIMO Wireless Communications: from realworld propagation to space-time code design," *Academic Press, Elsevier*, 2007.
- [13] L.H. Ozarow, S. Shamai and A.D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. on Vehicular Technology*, vol. 43, no. 2, pp. 359-379, May 1994.
- [14] K.P. Srinath and B.S. Rajan, "Low ML-Decoding Complexity, Large Coding Gain, Full-Rate, Full-Diversity STBCs for  $2 \times 2$  and  $4 \times 2$  MIMO Systems," *IEEE J. of Sel. Topics in Sign. Proc.*, vol. 3, no. 6, pp. 916-927, Dec. 2009.
- [15] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. on Inf. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.
- [16] L. Zheng and D.N.C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. on Inf. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.
- [17] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. on Inf. Theory*, vol. IT-28, no. 1, pp. 55-67, Jan. 1982.
- [18] E. Viterbo, J.J. Boutros, "A Universal Lattice Code Decoder for Fading Channels," *IEEE Trans. on Inf. Theory*, vol. 45, no. 5, pp. 1639-1642, 1999.
- [19] S. Yang, J.-C. Belfiore, "Distributed rotation recovers spatial diversity," *IEEE Intern. Symp. on Inf. Theory (ISIT)*, pp. 2158-2162, June 2010.