Optimal Diversity–Multiplexing Tradeoff and Code Constructions of Some Constrained Asymmetric MIMO Systems

Hsiao-Feng (Francis) Lu, *Member, IEEE*, and Camilla Hollanti, *Member, IEEE*

***Abstract*—In multiple-input–multiple-output (MIMO) communications, the notion of asymmetric channel refers to the situation when the number of transmit antennas is strictly larger than the number of receive antennas. Such channels can often be found in MIMO downlink transmissions. While existing cyclic-division-algebra (CDA)-based codes can still be employed to achieve the optimal diversity–multiplexing tradeoff (DMT) at high signal-to-noise ratio (SNR) regime, such codes cannot be directly decoded using, for example, the pure sphere decoding method. Although other means of decoding methods such as minimum mean square error generalized decision feedback equalizer (MMSE-GDFE) with lattice search and regularized lattice decoding are available, an alternative approach is to constrain the number of active transmit antennas in each channel use to be no larger than the number of receive antennas. The resulting system is coined *constrained asymmetric MIMO system*. Two general types of asymmetrical channels are considered in this paper, namely, 1) when there are two receive antennas and the number of transmit antennas is arbitrary, and 2) when the number of transmit antennas is one larger than the number of receive antennas. Explicit optimal transmission schemes as well as the corresponding code constructions for such constrained asymmetric MIMO channels are presented, and are shown to achieve the same DMT performance as their unconstrained counterparts.**

***Index Terms*—Constrained asymmetric multiple-input-multiple-output (MIMO) channels, cyclic-division algebra, diversity–multiplexing tradeoff, transmit antenna selection schemes, space-time codes..**

I. Introduction

The use of multiple antennas in wireless communication has been proven to be able to linearly increase the channel capacity [1], improve the diversity gain, and provide better reliability [2]. In an $(n_t \times n_r)$ multiple-input–multiple-output (MIMO) communication channel consisting of n_t transmit and n_r receive antennas, most of the existing literature [3]–[13] has focused on the case of $n_t \leq n_r$, and the corresponding code designs have been extensively investigated. However, in MIMO downlink transmissions, it is often found that there can be more transmit antennas available at the base station than receive antennas at mobile user ends. That is, it corresponds to the case of $n_t > n_r$. Such MIMO channel is commonly referred to as an *asymmetric MIMO channel* [14] or as an *underdetermined system* [15].

If all the n_t transmit antennas are active during each channel use and if the channel state information (CSI) is known completely to the receiver but not to the transmitter, Telatar [1] showed that the ergodic channel capacity of such $(n_t \times n_r)$ MIMO channel approximates $\min\{n_t, n_r\} \log_2 \mathrm{SNR}$ at high signal-to-noise ratio (SNR) regime, regardless of the relation between n_t and n_r. Furthermore, assuming the transmitter communicates at rate $R = r \log_2 \mathrm{SNR}$ bits per channel use (bpcu), where r, $0 \leq r \leq \min\{n_t, n_r\}$, is termed *multiplexing gain*, Zheng and Tse [16] proved that given r, the smallest bit error probability that can be achieved by any coding schemes is given by $P_{\mathrm{e,min}}(\mathrm{SNR}) \doteq \mathrm{SNR}^{-d^*(r)}$, where by $\doteq$, we mean the exponential equality defined in [16, Def. 1]. The negative exponent $d^*(r)$ is termed *diversity gain*. For quasi-static Rayleigh fading channels where channel remains fixed for at least n_t uses, $d^*(r)$ is a piecewise linear function in r and is given by [9], [16]

$$d^*(r) = (n_t - r)(n_r - r) \tag{1}$$

for integral values of r. $d^*(r)$ represents an optimal tradeoff between the multiplexing gain r and the diversity gain. It is, therefore, termed the diversity–multiplexing tradeoff (DMT) [16].

Motivated by this remarkable result, a considerable amount of research activity has been devoted to constructing coding schemes to achieve the optimal tradeoff $d^*(r)$; see, for example, [8]–[10], [12], [17], and [18]. In particular, for any n_t, using a cyclic division algebra (CDA) with degree n_t^2 over its center $\mathbb{Q}(\imath)$, where $\imath = \sqrt{-1}$, Elia *et al.* [9] have provided an algebraic construction of $(n_t \times n_t)$ matrix codes meeting the optimal tradeoff $d^*(r)$.

While all the aforementioned coding schemes are DMT optimal, they do require all the n_t transmit antennas to be active during all channel uses. Such requirement could result in some difficulty in decoding if $n_r < n_t$, i.e., the case of asymmetric channels. To see this, note that the channel matrix H is of size $(n_r \times n_t)$ with $n_t > n_r$. H has no left multiplicative matrix inverse, hence it is impossible to use zero-forcing (ZF) decoder for decoding. Similarly, pure sphere decoding [3] of full lattices

Manuscript received May 19, 2008; revised September 12, 2009. Current version published April 21, 2010. The work of H.-F. Lu was supported in part by Taiwan National Science Council under Grants NSC 96-2628-E-009-172-MY3 and NSC 98-2221-E-009-045-MY3. The work of C. Hollanti was supported by the TUCS, the Finnish Academy of Science and Letters, and by the Academy of Finland under Grant 210280. The material in this paper was presented in part at the IEEE Global Communications Conference, New Orleans, LA, November 2008.

H.-F. Lu is with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: francis@mail.nctu.edu.tw).

C. Hollanti is with the Department of Mathematics, University of Turku, Turku FI-20014, Finland and also with the Department of Mathematics and Statistics, University of Tampere, Tampere FI-33014, Finland (e-mail: cajoho@utu.fi).

Communicated by B. S. Rajan, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2010.2043779

would also fail since the sphere decoder relies on the QR decomposition of channel matrix H, and H has linearly dependent column vectors. For linear minimum-mean square error (MMSE) detector, Ahmadreza and Aria [19] showed that at high spectral efficiencies the diversity gain achieved by linear MMSE receiver is at most $n_r - n_t + 1$. Hence, the performance of MMSE decoding might not be good when $n_t > n_r$. To improve the performance of MMSE-based receiver, Damen *et al.* [15] proposed a three-stage receiver that consists of a minimum mean square error generalized decision feedback equalizer (MMSE-GDFE) [8] front-end, a lattice reduction with a greedy ordering technique, and a lattice search stage. This receiver is significantly less complex than the original sphere decoder. Through simulations, it was found [15] that it can offer error performance within fractions of a decibel from the maximal-likelihood (ML) receiver. Early termination of the last search stage is possible for further complexity reduction, but it would result in a performance loss. Recently, Jaldén and Elia [20], [21] proposed a class of *regularized lattice decoders* that includes the MMSE-GDFE receiver as a special case, and showed these regularized lattice decoders are DMT optimal.

An alternative solution to the problem of decoding in asymmetric channels, if one still wishes to employ the simple sphere decoders, is to restrict the number of active transmit antennas in each channel use to be $\leq n_r$. With this additional constraint, the resulting system is termed *constrained asymmetric MIMO system* in this paper, and coding schemes satisfying this additional requirement are coined *constrained asymmetric space–time codes*. Similarly, codes without this constraint will be termed *unconstrained codes*. An additional advantage of constrained asymmetric MIMO system is a significant reduction on the number of radio frequency (RF) chains required at the transmitter end. Typically the multiple antennas (patch or dipole) needed in a MIMO system are cheap, and the additional digital signal processors are even cheaper. But, the RF chains, including low-noise amplifiers, up- and downconverters, and digital-to-analog and analog-to-digital converters, are expensive and do not follow Moore's law [22]. The constrained system can reduce the required RF chains at the transmitter from n_t to at most n_r, and therefore, offers a significant cost reduction.

In [14], Hollanti and Ranto considered different asymmetric coding methods. As a special case, they considered four transmit and two receive antennas, i.e., $n_t = 4$ and $n_r = 2$, and proposed a block-diagonal coding method for constructing constrained asymmetric space–time codes. They first partition the four transmit antennas into two groups, each consisting of two transmit antennas, and then perform a joint-encoding between these two groups by making use of a multiblock space–time code [23] with two blocks. As the entries of the channel matrix H are independent identically distributed (i.i.d.), given multiplexing gain r, following [16] and [23] it can be easily shown that the DMT achieved by this scheme is $d(r) = 2(2-r)(2-r)$. Therefore, the achieved DMT performance is far from being optimal for all $r \in (0,2)$, compared to that of unconstrained system which has DMT $d^*(r) = (4-r)(2-r)$ in this case.

In this paper, we will investigate the optimal DMT of some constrained asymmetric MIMO systems, and in particular, we will focus on the following two cases.

1) $n_t > n_r = 2$: it corresponds to the case when there are two receive antennas and when the number of transmit antennas is strictly larger than two.
2) $n_t = n_r + 1$: it is the case when the number of transmit antennas is one larger than the number of receive antennas.

These two cases cover many practical MIMO downlink scenarios, for example, the (3×2), (4×2), and (4×3) asymmetric MIMO channels that can be widely found in existing MIMO-based wireless communication standards [24]–[27].

As for the case of $n_t > n_r = 1$, we remark that by regarding the $(n_t \times 1)$ asymmetric MIMO channel as a single-input–single-output (SISO) channel with n_t independent fading blocks, it can be easily shown [16], [23] that the optimal DMT of this multiblock fading channel is $d^*(r) = n_t(1-r)$, for $0 \leq r \leq 1$, and is exactly the same as that of unconstrained channel. To achieve $d^*(r)$, one can employ a multiblock single-antenna vector code of length n_t as proposed in [23]. Each entry is transmitted by one of the transmit antennas in one channel use. Then, the DMT achieved by such scheme is exactly $d^*(r)$.

This paper is organized as follows. In Section II, we will present DMT optimal transmission schemes for constrained asymmetric MIMO systems with $n_t > n_r = 2$ or with $n_t = n_r + 1$, and show that these constrained schemes can achieve the same DMT $d^*(r)$ in (1) as the unconstrained. Thus, if the codes are properly designed, there will be no performance loss in the DMT sense even when we limit the number of active transmit antennas in each channel use. Our proposed scheme is reminiscent of the antenna selection method proposed by Molisch and Win [22] except that we do not require any CSI at the transmitter end.[1] For the ease of presentation, detailed proofs of the DMT optimality of the proposed transmission schemes will be relegated to Appendixes I and II. Having obtained the optimal transmission schemes, the corresponding DMT optimal codes will be given in Section III.

II. DMT Optimal Transmission Scheme for Some Constrained Asymmetric MIMO Systems

Given an $(n_t \times n_r)$ constrained asymmetric MIMO channel with $n_t > n_r$, in this section, we present several DMT optimal transmission schemes in which at most n_r transmit antennas are used for transmission at any instant. We focus on two cases: 1) when $n_t > n_r = 2$, and 2) when $n_t = n_r + 1$. For both cases, we present DMT optimal transmission schemes that can achieve the same optimal DMT $d^*(r)$ of the unconstrained asymmetric channels. To describe our schemes, we first define the following.

Definition 1: In an $(n_t \times n_r)$ constrained asymmetric MIMO channel, let $\mathcal{T} = \{T_1, \ldots, T_{n_t}\}$ be the set of indices of n_t transmit antennas. We say $\mathcal{S} := \{(\mathcal{T}_1, n_1), \ldots, (\mathcal{T}_s, n_s)\}$ is a transmit antenna selection (TAS) scheme if the antenna selection patterns $\mathcal{T}_i$ are distinct proper subsets of $\mathcal{T}$ and have size

[1] It is also worthwhile to note that if CSI is known to transmitter, then DMT stops being a fundamental performance measure in such systems.

$1 \leq |\mathcal{T}_i| \leq n_r < n_t$ for all i. Moreover, each antenna selection pattern $\mathcal{T}_i$ will be used for n_i transmissions and it is assumed that the MIMO channel remains fixed for T channel uses with $T \geq \sum_{i=1}^{s} n_i$. □

The block-diagonal coding method proposed in [14] for the (4×2) constrained asymmetric MIMO channel can be regarded as a TAS scheme with $\mathcal{S}_{\rm BD} = \{(\{T_1,T_2\},2),(\{T_3,T_4\},2)\}$. We have already seen in Section I that scheme $\mathcal{S}_{\rm BD}$ is not DMT optimal in the (4×2) constrained asymmetric MIMO channel. However, it should be noted that $\mathcal{S}_{\rm BD}$ can achieve almost the same ergodic capacity as the unconstrained system. In particular, for any TAS scheme $\mathcal{S} = \{(\mathcal{T}_1,n_1),\ldots,(\mathcal{T}_s,n_s)\}$ with $|\mathcal{T}_i| = n_r$, it is easy to see the ergodic channel capacity achieved by $\mathcal{S}$ equals $C(\mathrm{SNR}) = n_r \log_2 \mathrm{SNR} + O(1)$ and differs from the capacity of the unconstrained system only in the $O(1)$ part. For example, the ergodic capacity achieved by $\mathcal{S}_{\rm BD}$ is roughly only 1.3 bpcu less than that of the unconstrained system at high SNR regime.

We remark that the challenge of designing coding schemes for constrained asymmetric channels lies more on the design of good TAS scheme than the design of space–time codes. The best space–time code can still perform poorly if the TAS scheme has poor DMT performance since the DMT behaves as a performance upper bound for any space–time codes.

A. Proposed Optimal Transmission Scheme for $n_t > n_r = 2$

While scheme $\mathcal{S}_{\rm BD}$ fails to be DMT optimal, for any $n_t > n_r = 2$ below we provide another transmission scheme and we will show it achieves the same optimal DMT $d^*(r)$ as the unconstrained system. Clearly, the maximal value of multiplexing gain r is upper bounded by $\min\{n_t,n_r\} = 2$, hence $0 \leq r \leq 2$. The proposed scheme is the following.

Theorem 1: In an $(n_t \times 2)$ constrained asymmetric MIMO system with $n_t > 2$, let $\{T_1,\ldots,T_{n_t}\}$ be the set of indices of n_t transmit antennas. Given the desired multiplexing gain r, if $r \in [1,2]$, then the TAS scheme $\mathcal{S}_1 := \{(\{T_1,T_2\},2),(\{T_2,T_3\},2),\ldots,(\{T_{n_t-1},T_{n_t}\},2)\}$ achieves the optimal DMT $d^*(r)$ of (1). If $r \in [0,1)$, then the scheme $\mathcal{S}_2 := \{(\{T_1,T_2\},4),(\{T_2,T_3\},2),\ldots,(\{T_{n_t-2},T_{n_t-1}\},2), (\{T_{n_t-1},T_{n_t}\},4)\}$ achieves DMT $d^*(r)$. □

The only difference between $\mathcal{S}_1$ and $\mathcal{S}_2$ is that when $0 \leq r < 1$, the sets $\{T_1,T_2\}$ and $\{T_{n_t-1},T_{n_t}\}$ are used twice more than the other sets. Thus, compared to $\mathcal{S}_{\rm BD}$ or the unconstrained CDA-based (4×4) space–time codes, the proposed scheme requires two more channel uses for $r \geq 1$ and six more for $r < 1$. However, the price of using more transmissions is well paid off by having a much better DMT performance (compared to $\mathcal{S}_{\rm BD}$) or having lesser RF chains (compared to unconstrained CDA codes). Furthermore, the proposed constrained scheme can achieve the same DMT performance as the unconstrained codes.

The proof of Theorem 1 involves the outage performance analysis of the proposed transmission scheme and is relegated to Appendix I for the ease of reading. In particular, we will prove in Appendix I that the DMTs achieved by $\mathcal{S}_1$ and $\mathcal{S}_2$ are given by

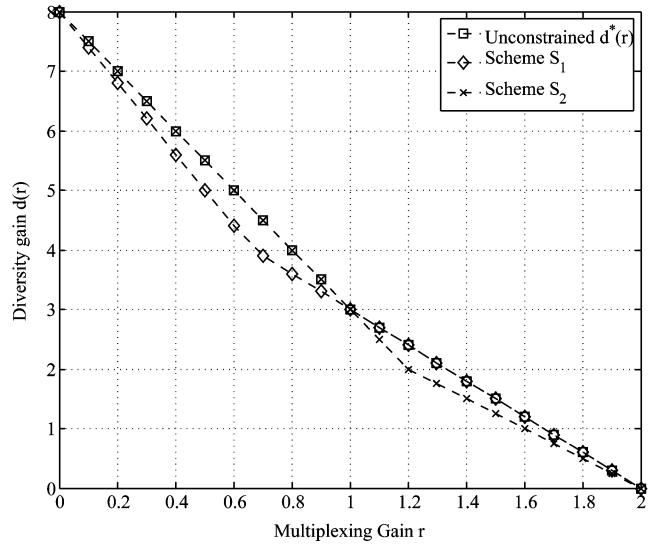

Fig. 1. DMT performances of schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ for the (4×2) constrained asymmetric MIMO system.

$$d(r) \geq \max\{(n_t-1)(2-r), 2n_t - 2(n_t-1)r\} \quad (2)$$

for $\mathcal{S}_1$, and

$$d(r) \geq \max\left\{2n_t - (n_t+1)r, \frac{1}{2}(n_t+1)(2-r)\right\} \quad (3)$$

for $\mathcal{S}_2$. In Fig. 1, we show the DMTs of schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ for the (4×2) constrained asymmetric MIMO system. It can be easily seen that each scheme is DMT optimal in its designated region.

B. Proposed Optimal Transmissions Schemes for $n_t = n_r + 1$

Now we move our attention to the case of $n_t = n_r + 1$. Let $\{T_1,T_2,\ldots,T_{n_r},T_{n_r+1}\}$ be the set of transmit antennas. The proposed scheme is a two-phase transmission. In the first phase, we use the set $\mathcal{T}_1 = \{T_1,\ldots,T_{n_r-1},T_{n_r}\}$ of transmit antennas for the first round transmission. During the second phase, we change the transmission set to $\mathcal{T}_2 = \{T_1,\ldots,T_{n_r-1},T_{n_r+1}\}$. This scheme can be applied to signal transmission in, e.g., (3×2), (4×3), or (5×4) constrained asymmetric MIMO communication systems. It turns out that the above scheme is DMT optimal and achieves the same DMT performance as the unconstrained ones. We have the following theorem.

Theorem 2: In an $(n_t \times n_r)$ constrained asymmetric MIMO system with $n_t = n_r+1$, let $\{T_1,\ldots,T_{n_r},T_{n_r+1}\}$ be the set of indices of n_t transmit antennas. Given the desired multiplexing gain r, the TAS scheme $\mathcal{P} = \{(\mathcal{T}_1,n_r),(\mathcal{T}_2,n_r)\}$ achieves the optimal DMT $d^*(r)$ of (1), where $\mathcal{T}_1$ and $\mathcal{T}_2$ are defined as the above.

Proof: The proof is relegated to Appendix II for the ease of reading. □

Unlike the case of $n_t > n_r = 2$ presented in Theorem 1 where we have to use two different TAS schemes to achieve the optimal DMT $d^*(r)$ for different values of multiplexing gain r, here for the case of $n_t = n_r + 1$ in Theorem 2, only one TAS scheme is sufficient for achieving $d^*(r)$ for all values of $r \in [0, n_r]$.

Compared to the unconstrained CDA-based space–time codes [9] where only n_t channel uses are required for code transmission, the TAS scheme proposed in Theorem 2 requires $2n_r$ channel uses. For example, in the (3×2) constrained asymmetric MIMO system, scheme $\mathcal{P}$ would require four channel uses, which is only one more than the unconstrained (3×3) CDA code.

Before concluding this section, we remark that the TAS schemes $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{P}$ proposed in Theorems 1 and 2 are *not* approximately universal in general. A simple way to see this is to note that any coding method derived from these TAS schemes would have a code matrix with zero entries; see, for example, the code given in Example 1. Then, following [28, Prop. 5], we see such codes (as well as the TAS schemes) cannot be approximately universal. In fact, it is straightforward to extend this argument to show that any constrained TAS schemes cannot be approximately universal at all in general. On the other hand, albeit not explicitly shown, it can be inferred from the proofs in Appendixes I and II that the TAS schemes $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{P}$ are still able to achieve the same optimal DMT performance as their unconstrained counterparts as long as the entries of the channel matrix are i.i.d. and have a joint probability distribution function that is invariant under unitary transformations. Hence, these schemes should be regarded as *partially approximately universal*.

III. DMT Optimal Codes for Constrained Asymmetric MIMO Systems

In Section II, we have identified two DMT optimal transmission schemes for the constrained asymmetric MIMO systems. These schemes use only n_r out of n_t transmit antennas during each transmission, and therefore can be decoded by simple decoding methods, such as ZF, sphere, and MMSE decoders. Further, the proposed schemes successfully transform the underdetermined system [15] into a determined system and significantly reduce the number of required RF chains at the transmitter end without any performance loss in the DMT sense. In this section, we aim at providing explicit code constructions that are specifically dedicated to such TAS schemes.

A. Proposed Code Constructions

To achieve the optimal DMT performance promised by Theorems 1 and 2, we simply employ the multiblock space–time codes proposed in [23] with certain necessary modifications. Specifically, given any TAS scheme $\mathcal{S} := \{(\mathcal{T}_1, n_1), \ldots, (\mathcal{T}_s, n_s)\}$ with s patterns and $n := |\mathcal{T}_1| = \cdots = |\mathcal{T}_s| \le n_r$, we assume without loss of generality that n divides n_i for all i. Otherwise, we could replace n_i by ℓn_i for some suitable ℓ and the resulting scheme will have the same DMT performance as the original scheme.

Given the multiplexing gain r, the space–time coding over the TAS scheme $\mathcal{S}$ calls for a number field $\mathbb{E}$ that is cyclic Galois over $\mathbb{F} = \mathbb{Q}(\imath)$ of degree mn, where $m := \frac{1}{n}\sum_{i=1}^{s} n_i$ is an integer by construction. Let σ be the generator of Galois group $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Methods for constructing such number fields can be found in [9] and [23]. Next, let $\mathbb{L}$ be the intermediate number field that is fixed by the Galois group $\langle \tau = \sigma^m \rangle$, i.e., we have $\mathrm{Gal}(\mathbb{E}/\mathbb{L}) = \langle \tau \rangle$. Since $\mathbb{E}$ is cyclic Galois over $\mathbb{L}$, with proper choice of $\gamma \in \mathbb{L}^*$, we can construct a CDA $\mathfrak{D} = (\mathbb{E}/\mathbb{L}, \tau, \gamma) = \{\sum_{i=0}^{n-1} z^i x_i : x_i \in \mathbb{E}\}$, where z is some indeterminate satisfying $z^n = \gamma$. We refer the interested readers to [6], [9], and [23] for details on constructing $\mathfrak{D}$.

Let $\psi : \mathfrak{D} \to \mathbb{E}^{n \times n}$ be the map of left-regular representation [9] of elements in $\mathfrak{D}$; then following [23], the multiblock space–time code for scheme $\mathcal{S}$ is

$$\mathcal{X} = \{\kappa(X, \ldots, \sigma^{m-1}(X)) : X = \psi(x), x \in \mathfrak{A}(\mathrm{SNR})\} \quad (4)$$

where

$$\mathfrak{A}(\mathrm{SNR}) = \left\{ \sum_{i=0}^{n-1} z^i \sum_{j=0}^{nm-1} a_{i,j} e_j : a_{i,j} \in \mathcal{A}(\mathrm{SNR}) \right\}.$$

$\{e_0, \ldots, e_{nm-1}\}$ is an integral basis for $\mathbb{E}/\mathbb{F}$, $\mathcal{A}(\mathrm{SNR})$ is a quadrature amplitude modulation (QAM) constellation set of size $\mathrm{SNR}^{r/n}$, and r is the desired multiplexing gain. $\kappa^2 = \mathrm{SNR}^{1-r/n}$ is a constant.

It was proven in [23, Th. 2] that the multiblock code $\mathcal{X}$ is actually approximately universal and can achieve the optimal DMT performance of the underlying MIMO channel. Thus, to adapt $\mathcal{X}$ to the TAS scheme $\mathcal{S}$, we propose the following transmission method. Note that each codeword of $\mathcal{X}$ consists of m blocks, and each block is a matrix of size $(n \times n)$. As n divides n_i for all i by construction, we can partition these m blocks into s groups, each of size $\frac{n_i}{n}$, $i = 1, \ldots, s$. Then, for each group i, the code matrices within that group will be transmitted via transmit antennas in set $\mathcal{T}_i$. In this manner, the code $\mathcal{X}$ will be transmitted exactly according to the TAS scheme $\mathcal{S}$, hence will achieve the DMT performance associated with $\mathcal{S}$ due to the approximately universal property of $\mathcal{X}$ shown in [23, Th. 2]. We have proven the following result.

Theorem 3: Given the TAS scheme $\mathcal{S}$, let $\mathcal{X}$ be the multiblock code constructed as the above. By transmitting $\mathcal{X}$ according to scheme $\mathcal{S}$, it achieves exactly the DMT performance of $\mathcal{S}$. ☐

In particular, following the steps outlined above one can easily construct multiblock space–time codes for the TAS schemes $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{P}$ proposed in Section II. Below we give an example construction for the (3×2) constrained asymmetric MIMO system.

Example 1: For the (3×2) constrained asymmetric MIMO system, one could follow either Theorem 1 which gives the scheme $\mathcal{S}_1 = \{(\{T_1, T_2\}, 2), (\{T_2, T_3\}, 2)\}$ for $r \in [1, 2]$ and $\mathcal{S}_2 = \{(\{T_1, T_2\}, 4), (\{T_2, T_3\}, 4)\}$ for $r \in [0, 1)$, or Theorem 2 that suggests $\mathcal{P} = \{(\{T_1, T_2\}, 2), (\{T_2, T_3\}, 2)\}$. Clearly, both approaches offer the same DMT performance $d^*(r) = (3 - r)(2 - r)$. To minimize decoding delay, we will choose $\mathcal{P}$. Such scheme calls for a (2×2) multiblock space–time code with $m = 2$ blocks, and it can be constructed from a number field that is cyclic Galois over $\mathbb{Q}(\imath)$ with degree $mn = 4$. So, we may choose $\mathbb{E} = \mathbb{Q}(\imath, \zeta_5)$ where ζ_5 is a complex, primitive fifth root of unity. In this setting, $\sigma(\zeta_5) = \zeta_5^2$, and it is easy to see $\mathbb{L} = \mathbb{Q}(\imath, \zeta_5 + \zeta_5^{-1})$. To construct the CDA, we can set $\gamma = \imath$, which can be shown to be a suitable non-norm element. Then, for any $x \in \mathfrak{A}(\mathrm{SNR})$, we get

$$X = \psi(x) = \begin{bmatrix} \underline{x}_1^t \\ \underline{x}_2^t \end{bmatrix}$$

where $\underline{x}_i$ are column vectors of length 2. The actual transmitted code matrix, when reformulating as a (3×4) matrix to correspond to the three transmit antennas and four channel uses, is

$$S = \kappa \begin{bmatrix} \underline{x}_1^t & \\ \underline{x}_2^t & \sigma\left(\underline{x}_1^t\right) \\ & \sigma\left(\underline{x}_2^t\right) \end{bmatrix}.$$

To decode using sphere decoders, note that $\mathcal{X}$ can be regarded as a 16-dimensional real lattice with full rank 16. On the other hand, multiplying the code matrix S by the (2×3) channel matrix H gives a received signal matrix of size (2×4) over complex field $\mathbb{C}$, hence can be written as a real vector of length 16. Now, as entries of H are i.i.d. and H is of full rank with probability one, decoding the 16 coordinates of the lattice code $\mathcal{X}$ using sphere decoder is straightforward. □

B. Simulation Results

In Fig. 2, we present simulation results of the example code given in Example 1 for the (3×2) asymmetric MIMO Rayleigh block fading channel. Two transmission rates, 4 and 6 bpcu, are simulated. Error performance of the proposed code is compared to that of the unconstrained (3×3) CDA-based space–time code [9]. The (3×3) CDA code can be regarded as a full-rank real lattice of 18 dimensions, however, when it is sent through the (3×2) channel, the received signal matrix, when written as a real vector, has dimension only 12. Hence, direct sphere decoding of the (3×3) code is not possible. While it still can be decoded by various MMSE–GDFE-based receivers [15], [20], [21], these receivers have different complexities and error performances. As pointed out in [15], the error performances of these receivers can differ by a few decibels. Therefore, here in Fig. 2, we resort to the ML decoding of the (3×3) CDA code, and the resulting performance can be seen as a benchmark for all MMSE–GDFE-based receivers. Another reason for using ML decoders is that we wish to focus on the difference in error performance between the constrained MIMO system and the unconstrained one, neither the performance nor the complexity of decoders. In Fig. 2, it is seen that the example code is roughly 1.7 dB better than the (3×3) CDA at 4 bpcu, but is about 0.7 dB worse at 6 bpcu. The reason for such disagreement is the following. At 4 bpcu, the proposed code uses 4-QAM as the underlying QAM set. On the other hand, we have to randomly puncture six out of 18 dimensions from the lattice of (3×3) CDA code such that it has the same rate of 4 bpcu. Probably due to puncturing, the (3×3) code does not perform well at this rate. For 6 bpcu, the (3×3) CDA uses all 18 dimensions available with binary phase-shift keying (BPSK) signaling, and the proposed code uses its all eight complex dimensions (or equivalently 16 real dimensions) with an amplitude-modulated phase-shift keying (AM-PSK) signaling of $\{\pm 1, \pm \imath, \pm 1 \pm \imath\}$. Thus, from Fig. 2, we conclude that the proposed TAS scheme $\mathcal{P}$ when transmitted using multiblock space–time code can achieve an error performance close to that of the unconstrained code, and in particular, achieves the same diversity gain and DMT.

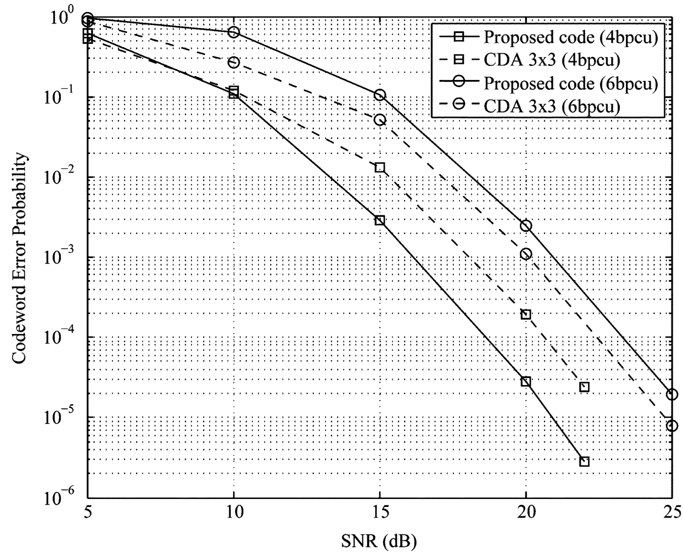

Fig. 2. Error performances of the proposed and the (3×3) CDA codes over (3×2) asymmetric MIMO channel.

Finally, we remark that both codes, proposed and unconstrained (3×3) CDA, can be further optimized by using techniques such as ideals and maximal orders [17] to improve their error performance.

Appendix I
Proof of Theorem 1

As $\mathcal{S}_1$ and $\mathcal{S}_2$ differ only in the number of channel uses for each $\mathcal{T}_i$, to simplify notations here we focus on the general scheme $\mathcal{S} := \{(\{T_1, T_2\}, n_1), \dots (\{T_{n_t-1}, T_{n_t}\}, n_{n_t-1})\}$, where the ith selection $\{T_i, T_{i+1}\}$ will be used n_i times during transmission. Moreover, for each $\{T_i, T_{i+1}\}$, let $\underline{x}_{i_j}$ be a length-n_r, zero-mean, complex Gaussian random vector with covariance matrix $K_i = \frac{\mathrm{SNR}}{2} I_{n_r}$. The subindex j, $j = 0, 1, \dots, n_i - 1$, represents the jth use of selection pattern $\{T_i, T_{i+1}\}$. Thus, given $\underline{x}_{i_j}$, the received signal vector is

$$\underline{y}_{i_j} = H_i \underline{x}_{i_j} + \underline{w}_{i_j} \tag{5}$$

where $H_i := [\underline{h}_i \ \underline{h}_{i+1}]$ and $\underline{h}_i$ is length-2 vector consisting of the fading coefficients between the ith transmit antenna T_i and the receive antennas. $\underline{w}_{i_j}$ is the corresponding additive complex Gaussian noise with $\mathbb{C}\mathcal{N}(0, 1)$ entries. Given the channel matrix H_i, the mutual information between the transmit and receive signal vectors is

$$\begin{aligned} I\left(\underline{x}_{i_j}; \underline{y}_{i_j} \middle| H_i\right) &= \log_2 \det\left(I_2 + \frac{\mathrm{SNR}}{2} H_i H_i^\dagger\right) \\ &= \log_2 \det(I_2 + \mathrm{SNR} H_i H_i^\dagger) + O(1) \end{aligned}$$

where dropping the constant factor of 2 does not affect the asymptotic mutual information at high SNR regime. Given the desired multiplexing gain r, the outage probability of $\mathcal{S}$ is

$$\Pr\left\{\sum_{i=1}^{n_t-1} n_i \log_2 \det(I_2 + \mathrm{SNR} H_i H_i^\dagger) \le Nr \log_2 \mathrm{SNR}\right\} \tag{6}$$

where $N := \sum_{j=1}^{n_t-1} n_j$ is the total number of channel uses. We begin by analyzing the sum of mutual information in (6). First, the summand associated with $\{T_1, T_2\}$, i.e., H_1, can be rewritten as

$$\begin{aligned}
&\log_2 \det(I_2 + \mathrm{SNR} H_1 H_1^\dagger)\\
&= \log_2 \det(I_2 + \mathrm{SNR}\, \underline{h}_1 \underline{h}_1^\dagger + \mathrm{SNR}\, \underline{h}_2 \underline{h}_2^\dagger)\\
&= \log_2 \det(I_2 + \mathrm{SNR}\, D_1)\\
&\quad + \log_2(1 + \mathrm{SNR}\, \underline{h}_2^\dagger U_1 (I_2 + \mathrm{SNR}\, D_1)^{-1} U_1^\dagger \underline{h}_2)\\
&= \log_2\left(1 + \mathrm{SNR}\|\underline{h}_1\|^2\right)\\
&\quad + \log_2(1 + \mathrm{SNR}\, \underline{g}_2^\dagger (I_2 + \mathrm{SNR} D_1)^{-1} \underline{g}_2)
\end{aligned}$$

where $U_1 D_1 U_1^\dagger$ is the eigendecomposition of the rank-1 matrix $\underline{h}_1 \underline{h}_1^\dagger$, $\|\underline{h}_1\|$ is the Frobenius norm of vector $\underline{h}_1$, and where $\underline{g}_2 := U_1^\dagger \underline{h}_2$. Since H_1 is a complex Gaussian matrix with i.i.d. $\mathbb{CN}(0,1)$ entries, $\underline{g}_2$ has the same joint probability density function (pdf) as $\underline{h}_2$. Furthermore, we rewrite the outage probability in the following form:

$$\begin{aligned}
&\Pr\{\text{channel outage}\}\\
&= \int \Pr\{\text{channel outage} \,|\, \underline{h}_1\} f(\underline{h}_1)\, d\underline{h}_1\\
&\overset{\text{(i)}}{=} \int \left[\int \Pr\{\text{channel outage} \,|\, \underline{h}_1, \underline{h}_2\} f(\underline{h}_2) d\underline{h}_2\right] f(\underline{h}_1) d\underline{h}_1\\
&\overset{\text{(ii)}}{=} \int \left[\int \Pr\{\text{channel outage} \,|\, \underline{h}_1, \underline{g}_2\} f(\underline{g}_2) d\underline{g}_2\right] f(\underline{h}_1) d\underline{h}_1
\end{aligned}$$

where $f(\underline{h}_i)$ is the joint pdf of complex Gaussian vector $\underline{h}_i$, and where (i) follows from that $\underline{h}_2$ is random and independent of $\underline{h}_1$. Equality (ii) is because U_1 is fixed and unitary, and because $\underline{h}_2$ and $\underline{g}_2 = U_1^\dagger \underline{h}_2$ have the same joint pdf. Hence, transforming $\underline{h}_2$ to $\underline{g}_2$ does not affect the value of outage probability (6). On the other hand, in evaluating $\Pr\{\text{channel outage} \,|\, \underline{h}_1, \underline{g}_2\}$, given the fading coefficients $\underline{h}_1$ and $\underline{g}_2$, the channel matrix associated with the second pattern $\{T_2, T_3\}$ changes to $H_2' = [\underline{g}_2 \quad \underline{h}_3]$, and following the same reasoning the second summand in (6) becomes

$$\begin{aligned}
I\left(\underline{x}_{2_j}; \underline{y}_{2_j} | H_2'\right) &= \log_2\left(1 + \mathrm{SNR}\|\underline{g}_2\|^2\right)\\
&\quad + \log_2\left(1 + \mathrm{SNR}\, \underline{g}_3^\dagger (I_2 + \mathrm{SNR}\, D_2)^{-1} \underline{g}_3\right)
\end{aligned}$$

where $U_2 D_2 U_2^\dagger$ is the eigendecomposition of the rank-1 matrix $\underline{g}_2 \underline{g}_2^\dagger$ and $\underline{g}_3 = U_2^\dagger \underline{h}_3$. Again, we could replace the random vector $\underline{h}_3$ by $\underline{g}_3$, and then proceed to investigate the conditional probability $\Pr\{\text{channel outage} \,|\, \underline{h}_1, \underline{g}_2, \underline{g}_3\}$. Literally, the above process is closely related to the chain rule of conditional probability. Instead of handing all the channel vectors $\{\underline{h}_1, \ldots, \underline{h}_{n_t}\}$ to the receiver to determine whether an outage occurs, in the above derivation, we first reveal $\underline{h}_1$ to the receiver. Given $\underline{h}_1$, hence U_1 is fixed, there is no difference in probability whether we tell the receiver the second fading vector is $\underline{g}_2$ or $\underline{h}_2$, and so on. The whole process does not change value of outage probability since $\underline{g}_i$ and $\underline{h}_i$ are of the same pdf. By repeating the above process, we see $\Pr\{\text{channel outage}\} = \int \cdots \int \mathbf{1}\{\text{channel outage} \,|\, \underline{h}_1, \underline{g}_2, \ldots, \underline{g}_{n_t}\} f(\underline{h}_1) f(\underline{g}_2) \cdots f(\underline{g}_{n_t}) d\underline{h}_1 d\underline{g}_2 \ldots d\underline{g}_{n_t}$, where $\mathbf{1}\{\cdot\}$ is the indicator function, i.e., telling whether an outage does occur. Thus, given the channel vectors $\underline{h}_1, \underline{g}_2, \ldots, \underline{g}_{n_t}$, we can rewrite the overall mutual information resulting from $\mathcal{S}$ as

$$\sum_{i=1}^{n_t-1} n_i \left[\log_2(1 + \mathrm{SNR}\|\underline{g}_i\|^2) + \log_2\left(1 + \frac{\mathrm{SNR}|g_{i+1,1}|^2}{1 + \mathrm{SNR}\|\underline{g}_i\|^2} + \mathrm{SNR}|g_{i+1,2}|^2\right)\right] \tag{7}$$

where we have set $\underline{g}_1 = \underline{h}_1$ for convenience. Set $|g_{i,j}|^2 = \mathrm{SNR}^{-\alpha_{i,j}}$; then, at high SNR, we can rewrite (7) as

$$\log_2 \mathrm{SNR} \cdot \sum_{i=1}^{n_t-1} n_i \left[\left(\underbrace{\max_j\{(1-\alpha_{i,j})^+\}}_{:=(1-\beta_i)^+}\right) + (\max\{(1-\alpha_{i+1,1} - (1-\beta_i)^+)^+, (1-\alpha_{i+1,2})^+\})\right]$$

where $(x)^+ := \max\{0, x\}$. Thus, following the Laplace method as in [16], the diversity gain achieved by scheme $\mathcal{S}$ is

$$d(r) = \inf_{\mathcal{A}(r)} \sum_{i=1}^{n_t} \sum_{j=1}^{2} \alpha_{i,j} \tag{8}$$

where the constraint set is

$$\mathcal{A}(r) = \left\{(\alpha_{1,1}, \ldots, \alpha_{n_t,2}) : \sum_{i=1}^{n_t-1} n_i [(1-\beta_i)^+ + \max\{(1-\alpha_{i+1,1} - (1-\beta_i)^+)^+, (1-\alpha_{i+1,2})^+\}] \le Nr, \alpha_{i,j} \ge 0\right\}.$$

While minimizing $d(r)$ over $\mathcal{A}(r)$ appears to be a nonlinear optimization problem, below we show how it can be converted to a linear programming problem. First note that each $\alpha_{i,j}$ has probability exponentially zero if $\alpha_{i,j} < 0$. Second, to minimize $d(r)$, it suffices to consider only $\alpha_{i,j} \le 1$ as $(1-\alpha_{i,j})^+ = 0$ for $\alpha_{i,j} \ge 1$ and setting $\alpha_{i,j} \le 1$ minimizes the cost $d(r)$. Thus, we have the following sets of linear constraints:

$$0 \le \alpha_{i,j} \le 1, \qquad \text{for all } i = 1, \ldots, n_t - 1, j = 1, 2 \tag{9}$$

Next, for $i = 1, 2, \ldots, n_t - 1$, we set

$$r_{i,1} := (1-\beta_i)^+ = \max\{(1-\alpha_{i,1})^+, (1-\alpha_{i,2})^+\}$$

and (9) implies $0 \le r_{i,1} \le 1$. As $(x)^+ \ge x$ by definition, we see $r_{i,1} \ge \max\{(1-\alpha_{i,1}), (1-\alpha_{i,2})\}$. Hence, it follows that

$$\alpha_{i,1} \ge 1 - r_{i,1} \quad (10)$$
$$\alpha_{i,2} \ge 1 - r_{i,1} \quad (11)$$
$$1 \ge r_{i,1} \ge 0. \quad (12)$$

Similarly, for $i = 1, 2, \ldots, n_t - 1$

$$r_{i,2} := \max\{(1-\alpha_{i+1,1} - (1-\beta_i)^+)^+, (1-\alpha_{i+1,2})^+\} \ge \max\{(1-\alpha_{i+1,1} - r_{i,1}), (1-\alpha_{i+1,2})\}$$

and this leads to the following linear constraints:

$$\alpha_{i+1,1} \ge 1 - r_{i,1} - r_{i,2} \quad (13)$$
$$\alpha_{i+1,2} \ge 1 - r_{i,2} \quad (14)$$
$$1 \ge r_{i,2} \ge 0. \quad (15)$$

To achieve the desired multiplexing gain r, the constraint on the $r_{i,j}$ is given by

$$\sum_{i=1}^{n_t-1} n_i (r_{i,1} + r_{i,2}) \le Nr. \quad (16)$$

Prior to minimizing $d(r)$ of (8) subject to constraints (9)–(16), we observe the following.

O1) $r_{1,1}$ is the only variable appearing in inequalities (10) and (11) for $\alpha_{1,1}$ and $\alpha_{1,2}$, and does not appear anywhere else. Adding these two inequalities suggests $\alpha_{1,1} + \alpha_{1,2} \ge 2 - 2r_{1,1}$, and we will set $r_{1,1}$ to its maximal possible value to minimize $d(r)$.

O2) There is a symmetry between the variables $r_{i,1}$ and $r_{i,2}$ for $i = 2, \ldots, n_t - 2$ in (10), (11), (13), and (14) for both schemes $\mathcal{S}_1$ and $\mathcal{S}_2$. So we may set these variables at the same value.

With the above observations in mind, we now begin to find the minimal possible $d(r)$.

1) For the scheme $\mathcal{S}_1$, i.e., $n_i = 2$ for all i, we have $N = 2(n_t - 1)$ and we will show

$$d(r) \ge (n_t - 1)(2 - r) \quad (17)$$

and

$$d(r) \ge 2n_t - 2(n_t - 1)r. \quad (18)$$

To show the above, from O1 and (16), the maximal possible value for $r_{1,1}$ is $\min\{1, \frac{Nr}{2}\}$. That is, for $r \le \frac{1}{n_t-1}$, we have $r_{1,1} = \frac{Nr}{2}$ and $r_{i,j} = 0$ for the remaining. Hence, $\alpha_{i,j} = 1$ for all $i = 2, 3, \ldots$ and $j = 1, 2$. Thus, the diversity gain in this interval is $\sum_{i,j} \alpha_{i,j} \ge 2 - 2r_{i,j} + 2(n_t - 1) = 2n_t - 2(n_t - 1)r$ and shows (18). When $r \ge \frac{1}{n_t-1}$, $r_{1,1} = 1$. By observation O2, we set $r_{2,1} = r_{2,2} = r_{3,1} = \cdots = r_{n_t-2,2} = Z$ for simplicity. As $n_i = 2$ for all i, again by symmetry, we can without loss of generality set $r_{n_t-1,1} = Z$. Then, we can rewrite (16) as

$$1 + r_{1,2} + (2n_t - 5)Z + r_{n_t-1,2} \le (n_t - 1)r \quad (19)$$

and from (10)–(16) we have $\alpha_{1,1} = \alpha_{1,2} = 0$, and $\alpha_{2,2} \ge 1 - Z$, $\alpha_{2,2} \ge 1 - r_{1,2}$, $\alpha_{n_t,1} \ge 1 - Z - r_{n_t,2}$, $\alpha_{n_t,2} \ge 1 - r_{n_t-1,2}$, and $\alpha_{i,j} \ge 1 - Z$ for all other values of i and j. Combining the inequality $\alpha_{2,2} \ge 1 - r_{1,2}$ with the ones for other $\alpha_{i,j}$'s, we get $\sum_{i,j} \alpha_{i,j} \ge 2 - Z - r_{1,2} + (n_t - 3)(2 - 2Z) + 2 - Z - 2r_{n_t-2,2} \ge 2n_t - 1 - Z - r_{n_t-1,2} - (n_t - 1)r$, where the last inequality follows from constraint (19). Similarly, following the same approach with inequality $\alpha_{2,2} \ge 1 - Z$, we end up with $\sum_{i,j} \alpha_{i,j} \ge 2 - 2Z + (n - 3)(2 - 2Z) + 2 - Z - 2r_{n_t-1,2} \ge 2n_t - 1 + r_{1,2} - 2Z - r_{n_t-1,2}$. As $r_{n_t-1,2}$ appears in negative in both inequalities, we will set $r_{n_t-1,2}$ to its maximal value. Together with observation O1, we set $r_{1,1} = r_{n_t-1,2} = \min\{1, \frac{Nr}{4}\}$. It is easy to verify that with the above, (18) remains to hold. For $\frac{2}{n_t-1} \le r \le 2$, such setting gives $r_{1,1} = r_{n_t-1,2} = 1$ and $\alpha_{n_t,j} = 0$ for $j = 1, 2$. Equation (19) is replaced by $2 + r_{1,2} + (2n_t - 5)Z \le (n_t - 1)r$. Combining $\alpha_{2,2} \ge 1 - r_{1,2}$ with the remaining inequalities gives $\sum_{i,j} \alpha_{i,j} \ge 2 - Z - r_{1,2} + (n_t - 3)(2 - 2Z) \ge 2(n_t - 1) - (n_t - 1)r$, and proves (17).

We remark that it is also possible to solve the linear programming (10)–(16) directly using any linear programming computer packages and verify the results (17) and (18). Finally, we see for $1 \le r \le 2$, the DMT achieved by $\mathcal{S}_1$ is given by

$$d(r) = (n_t - 1)(2 - r), \qquad \text{for } 1 \le r \le 2. \quad (20)$$

2) For scheme $\mathcal{S}_2$, i.e., the case when $n_1 = n_{n_t-1} = 4$ and the remaining $n_i = 2$, we have $N = 2n_t + 2$, and we will show

$$d(r) \ge 2n_t - (n_t + 1)r \quad (21)$$

and

$$2d(r) \ge (n_t + 1)(2 - r). \quad (22)$$

First, again from observation O1, for $0 \le r \le \frac{2}{n_t+1}$, we can set $r_{1,1} = \frac{Nr}{4}$ and obtain $\sum_{i,j} \alpha_{i,j} \ge 2 - 2r_{1,1} + 2(n_t - 1) = 2n_t - (n_t + 1)r$. This shows (21) in this interval of interest. For $\frac{2}{n_t+1} \le r \le 2$, we will set $r_{1,1} = 1$, and $r_{2,1} = \cdots = r_{n_t-2,2} = Z$ due to observation O2. Constraint (16) now changes to

$$2 + 2r_{1,2} + 2(n_t - 3)Z + 2r_{n_t-1,1} + 2r_{n_t-1,2} \le (n_t + 1)r. \quad (23)$$

Adding the inequalities of $\alpha_{2,2} \ge 1 - r_{1,2}$, $\alpha_{n_t-1,1} \ge 1 - r_{n_t-1,1}$, $\alpha_{n_t-1,2} \ge 1 - Z$, and the remaining ones gives $\sum_{i,j} \alpha_{i,j} \ge 2 - Z - r_{1,2} + (n_t - 4)(2 - 2Z) + 2 - Z - r_{n_t-1,1} + 2 - r_{n_t-1,1} - 2r_{n_t-1,2} = (2n_t - 2) - r_{1,2} - 2(n_t - 3)Z - 2r_{n_t-1,1} - 2r_{n_t-1,2} \ge 2n_t + r_{1,2} - (n_t + 1)r$, where the last inequality follows from (23). Hence, we will make $r_{1,2}$ as small as possible. From (23), we can set $r_{1,2} = 0$ whenever $r \le \frac{2n_t}{n_t+1}$ and $r_{1,2} = \frac{(n_t+1)r}{2} - n_t$ if $r \in [\frac{2n_t}{n_t+1}, 2]$. Thus, for $r \in [\frac{2}{n_t+1}, \frac{2n_t}{n_t+1}]$, we see $\sum_{i,j} \alpha_{i,j} \ge 2n_t - (n_t + 1)r$, and this gives (21). On the other hand, combining $\alpha_{1,2} \geq 1 - Z$, $\alpha_{n_t-1,1} \geq 1 - r_{n_t-1,1}$, $\alpha_{n_t-1,2} \geq 1 - Z$, with the others gives $\sum_{i,j} \alpha_{i,j} \geq 2-2Z+(n_t-4)(2-2Z)+2-Z-r_{n_t-1,1}+2-r_{n_t-1,1}-2r_{n_t-1,2} = (2n_t-2)-Z-2(n_t-3)Z-2r_{n_t-1,1}-2r_{n_t-1,2} \geq 2n_t+2r_{1,2}-(n_t+1)r-Z$. This suggests that we will maximize Z. Pushing Z to 1 requires $r \geq \frac{2(n_t-2)}{n_t+1}$ from (23). With $Z=1$, (23) gives $r_{1,2}+r_{n_t-1,1}+r_{n_t-1,2} \leq \frac{1}{2}[(n_t+1)r-2(n_t-2)]$. Further, setting $Z=1$ gives $\alpha_{2,1} \geq 0$, $\alpha_{2,2} \geq 1-r_{1,2}$, $\alpha_{n_t-1,1} \geq 1-r_{n_t-1,1}$, and $\alpha_{n_t-1,2} \geq 1-r_{n_t-1,1}$. Hence, $\sum_{i,j} \alpha_{i,j} \geq 5-r_{1,2}-3r_{n_t-1,1}-2r_{n_t-1,2} \geq (2n_t+1)+r_{1,2}-r_{n_t-1,1}-(n_t+1)r$, and one should bear in mind that $r_{1,2} = \frac{(n_t+1)r}{2} - n_t$ in this interval of interest. It then suggests that we should set $r_{n_t-1,1}$ to its maximal, i.e., $r_{n_t-1,1}=1$. Finally, with the above, we have $\sum_{i,j} \alpha_{i,j} \geq \frac{(n_t+1)(2-r)}{2}$ and this gives (22). In particular, it is easy to show that when $r \in [\frac{2}{n_t+1}, \frac{2n_t}{n_t+1}]$ we can set $r_{n_t-1,2}=1$ as its optimal value in this region. Finally, for the region of $0 \leq r \leq 1$, the DMT achieved by scheme $\mathcal{S}_2$ is given by

$$d(r) = 2n_t - (n_t+1)r, \qquad \text{for } 0 \leq r \leq 1. \tag{24}$$

The proof is now complete after noting that the DMTs (20) and (24) achieved, respectively, by schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ in the region of $r \in [1,2]$ and $r \in [0,1)$ match exactly the optimal DMT $d^*(r)$ given in (1).

Remark 1: To help the readers understand better the DMT analysis presented in this section, below we show how our method can be applied to analyze the DMT of conventional (2×2) MIMO system. The result will be a rediscovery of the DMT found by Zheng and Tse [16], but our approach is rather elementary in the sense that we do not require the knowledge of the complicated eigenvalue distribution of Wishart matrices as in [16]. To this end, substituting $n_t=2$ into linear constraints (9)–(16) gives $\alpha_{1,1} \geq 1-r_{1,1}$, $\alpha_{1,2} \geq 1-r_{1,1}$, $\alpha_{1,2} \geq 1-r_{1,1}-r_{1,2}$, $\alpha_{2,2} \geq 1-r_{1,2}$, $r_{1,1}+r_{1,2} \leq r$, $0 \leq \alpha_{i,j} \leq 1$, and $0 \leq r_{1,j} \leq 1$ with $0 \leq r \leq 2$. Adding all the linear constraints of $\alpha_{i,j}$ gives $\sum_{i,j} \alpha_{i,j} \geq 4-3r_{1,1}-2r_{1,2} \geq 4-2r-r_{1,1}$. It shows that we will maximize $r_{1,1}$ and set it as $r_{1,1}=\min\{1,r\}$. Thus, for $r \in [0,1]$, we see $r_{1,1}=r$ and $r_{1,2}=0$ and this gives $d(r) \geq 4-3r$. For $r \in [1,2]$, we then have $r_{1,1}=1$ and $r_{1,2}=r-1$. Hence, the minimal $d(r)$ is achieved at $\alpha_{1,1}=\alpha_{1,2}=\alpha_{2,1}=0$ and $\alpha_{2,2}=1-r_{1,2}=2-r$, and this gives $d(r)=2-r$. Overall, we show the DMT of (2×2) system equals $d^*(r)=\max\{4-3r, 2-r\}$. This agrees with the DMT shown by Zheng and Tse [16]. □

APPENDIX II
PROOF OF THEOREM 2

First note that in the proposed transmission scheme $\mathcal{P}$, the set of transmit antennas $\mathcal{T}_0 := \{T_1, \ldots, T_{n_r-1}\}$ is used for both transmissions, and in each channel use, we add a new transmit antenna to the set $\mathcal{T}_0$, i.e., $\mathcal{T}_i := \mathcal{T}_0 \cup \{T_{n_r-i+1}\}$ for $i=1,2$. With a similar argument as in Appendix I, the channel outage probability of scheme $\mathcal{P}$ is given by

$$\Pr\left\{\sum_{i=1}^{2} \log_2 \det(I_{n_r} + \mathrm{SNR} H_i H_i^\dagger) \leq 2r \log_2 \mathrm{SNR}\right\}$$

where $H_i = [H_0 \underline{h}_i]$, random matrix H_0 represents the fading coefficients between transmit antenna set $\mathcal{T}_0$ and receive antennas, and $\underline{h}_i$ is the channel vector associated with transmit antenna T_{n_r+i-1}, $i=1,2$.

Let $UD_0U^\dagger$ be the eigendecomposition of $H_0H_0^\dagger$ with nondecreasing eigenvalues $0=\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n_r-1}$, where $\lambda_1 > 0$ with probability 1. As $H_iH_i^\dagger = H_0H_0^\dagger + \underline{h}_i\underline{h}_i^\dagger$, following similar arguments as in Appendix I shows

$$\begin{aligned} &\log_2 \det(I_{n_r} + \mathrm{SNR} H_0 H_0^\dagger + \mathrm{SNR}\, \underline{h}_i \underline{h}_i^\dagger) \\ &\quad = \log_2 \det(I_{n_r} + \mathrm{SNR} D_0) \\ &\qquad + \log_2[1 + \mathrm{SNR}\underline{g}_i^\dagger (I_{n_r} + \mathrm{SNR}\, D_0)^{-1} \underline{g}_i] \\ &\quad = \left[\sum_{j=1}^{n_r-1} \log_2(1+\mathrm{SNR}\lambda_j)\right] \\ &\qquad + \log_2\left(1 + \sum_{j=1}^{n_r-1} \frac{\mathrm{SNR}|g_{i,j}|^2}{1+\mathrm{SNR}\lambda_j} + \mathrm{SNR}|g_{i,n_r}|^2\right) \end{aligned} \tag{25}$$

where $\underline{g}_i := U^\dagger \underline{h}_i = [g_{i,1} \cdots g_{i,n_r}]^t$. Clearly, since entries of H_i are i.i.d. $\mathbb{C}\mathcal{N}(0,1)$ random variables, $\underline{h}_i$ and $\underline{g}_i$ are the same joint pdfs. Set $\lambda_i := \mathrm{SNR}^{-\alpha_i}$ and $|g_{i,j}|^2 := \mathrm{SNR}^{-\beta_{i,j}}$; then we can rewrite the outage event as

$$\begin{aligned} \mathcal{A}(r) := \Bigg\{ &(\alpha_1, \ldots, \alpha_{n_r-1}, \beta_{1,1}, \ldots, \beta_{2,n_r}) : \left[2\sum_{j=1}^{n_r-1} (1-\alpha_j)^+\right] \\ &+ \sum_{i=1}^{2} \max_{1 \leq j \leq n_r-1} \left\{(1-\beta_{i,j}-(1-\alpha_j)^+)^+,\ (1-\beta_{i,n_r})^+\right\} \\ &\leq 2r, \beta_{i,j} \geq 0, \alpha_i \geq 0 \Bigg\}. \end{aligned} \tag{26}$$

Following [16] and [29], the joint pdf for $\alpha_1 \geq \cdots \geq \alpha_{n_r-1}$ and $\beta_{i,j}$'s is given by

$$f(\alpha_1, \ldots, \alpha_{n_r-1}, \beta_{1,1}, \ldots, \beta_{2,n_r}) \doteq \begin{cases} 0, & \text{if any } \alpha_i, \beta_{i,j} < 0 \\ \mathrm{SNR}^{-d}, & \text{otherwise} \end{cases}$$

where

$$d = \left[2\sum_{i=1}^{n_r-1} i \cdot \alpha_i\right] + \sum_{i=1}^{2}\sum_{j=1}^{n_r} \beta_{i,j}.$$

Note that there is a symmetry between the $\beta_{i,j}$ in the above and in (26). Hence, in finding the dominant case of $\beta_{i,j}$'s, we are free to set $\beta_{2,j} = \beta_{1,j}$ and the diversity gain $d(r)$ achieved by $\mathcal{P}$ equals

$$d(r) = 2 \inf_{\mathcal{A}'(r)} \left\{\left[\sum_{i=1}^{n_r-1} i \cdot \alpha_i\right] + \sum_{j=1}^{n_r} \beta_{1,j}\right\}$$

where $\mathcal{A}'(r) := \{(\alpha_1, \ldots, \alpha_{n_r-1}, \beta_{1,1}, \ldots, \beta_{1,n_r}) \in \mathcal{A}(r), \beta_{1,j} = \beta_{2,j}\}$. For $n_t = n_r + 1$, it can be shown that when $K - 1 \leq r \leq K$ for some positive integer K, the infimum is achieved at

$$\alpha_i = \begin{cases} 1, & \text{if } i = 1, \ldots, n_r - K - 1 \\ K - r, & \text{if } i = n_r - K \\ 0, & \text{if } i \geq n_r - K + 1 \end{cases}$$

$$\beta_{1,j} = \begin{cases} \alpha_j, & \text{if } 1 \leq j \leq n_r - 1 \\ 1, & \text{if } j = n_r \text{ and } K = 1, \ldots, n_r - 1 \\ n_r - r, & \text{if } j = n_r \text{ and } K = n_r \end{cases}$$

and the resulting diversity gain $d(r) = d^*(r)$. This completes the proof.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[3] B. Hassibi and M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[4] H. F. Lu and P. V. Kumar, "A unified construction of space-time codes with optimal rate-diversity tradeoff," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1709–1730, May 2005.

[5] H. El Gamal and A. R. Hammons, Jr., "On the design of algebraic space-time codes for MIMO block fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 151–163, Jan. 2003.

[6] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2596–2616, Oct. 2003.

[7] M. O. Damen, H. El Gamal, and N. C. Beaulieu, "Linear threaded algebraic space-time constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2372–2388, Oct. 2003.

[8] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 968–985, Jun. 2004.

[9] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit construction of space-time block codes achieving the diversity–multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869–3884, Sep. 2006.

[10] T. Kiran and B. S. Rajan, "STBC-schemes with non-vanishing determinant for certain number of transmit antennas," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2984–2992, Aug. 2005.

[11] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space time codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3885–3902, Sep. 2006.

[12] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden code: A 2×2 full-rate space-time code with non-vanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[13] K. R. Kumar and G. Caire, "Space-time codes from structured lattices," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 547–556, Feb. 2009.

[14] C. Hollanti and K. Ranto, "Asymmetric space-time block codes for MIMO systems," in *Proc. IEEE Inf. Theory Workshop*, Solstrand, Norway, Jul. 2007, pp. 101–105.

[15] M. O. Damen, H. E. Gamal, and G. Caire, "MMSE-GDFE lattice decoding for solving under-determined linear systems with integer unknowns," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, p. 538.

[16] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[17] R. Vehkalahti, C. Hollanti, J. Lahtonen, and K. Ranto, "On the densest MIMO lattices from cyclic division algebras," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3751–3780, Aug. 2009.

[18] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, Jul. 2006.

[19] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007.

[20] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," [Online]. Available: http://arxiv.org/abs/0905.4023

[21] J. Jalden and P. Elia, "LR-aided MMSE lattice decoding is DMT optimal for all approximately universal codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. 2009, pp. 1263–1267.

[22] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[23] H. F. Lu, "Explicit constructions of multi-block space-time codes that achieve the diversity–multiplexing tradeoff," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3790–3796, Aug. 2008.

[24] *IEEE P802.11 Wireless LANs: TGn Sync Proposal Technical Specification*, IEEE 802.11-04/0889r7.

[25] *Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems—Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE 802.16e D11, Sep. 13, 2005.

[26] *Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals*, ETSI Standard: EN 302 304 V1.1.1, Nov. 2004.

[27] *Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*, ETSI Standard: EN 300 744 V1.5.1, Nov. 2004.

[28] P. Elia and J. Jaldén, "Construction criteria and existence results for approximately universal linear space-time codes with reduced decoding complexity," in *Proc. Allerton Conf. Commun. Control Comput.*, Monticello, IL, Sep. 2008, pp. 1359–1364.

[29] A. Edelman, "Eigenvalues and conditional numbers of random matrices," Ph.D. dissertation, Dept. Math., Massachusetts Inst. Technol., Cambridge, MA, 1989.

Hsiao-Feng (Francis) Lu (S'98–M'04) received the B.S. degree from Tatung University, Taipei, Taiwan, in 1993, and the M.S.E.E. and Ph.D. degrees from the University of Southern California (USC), Los Angeles, in 1999 and 2003, respectively, all in electrical engineering.

He was a Postdoctoral Research Fellow at the University of Waterloo, ON, Canada, during 2003–2004. In February 2004, he joined the faculty of the Department of Communications Engineering, National Chung-Cheng University, Chiayi, Taiwan, where he was promoted to Associate Professor in August 2007. Since August 2008, he has been with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan. His research is in the area of space-time codes, MIMO systems, error correcting codes, wireless communication, optical fiber communication, and multiuser detection.

Dr. Lu is an Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is a recipient of several research awards, including the 2006 IEEE Information Society Taipei Chapter and IEEE Communications Society Taipei/Tainan Chapter Best Paper Award for Young Scholars, the 2007 Wu Da You Memorial award from Taiwan National Science Council, the 2007 IEEE Communication Society Asia Pacific Outstanding Young Researchers Award, and the 2008 Academia Sinica Research Award for Junior Research Investigators.

Camilla Hollanti (M'09) received the M.Sc. and Ph.D. degrees from the University of Turku, Turku, Finland, in 2003 and 2009, respectively, both in pure mathematics.

Since June 2004, she has been with the Department of Mathematics, University of Turku. In 2005, she visited the Department of Algebra at Charles' University, Prague, Czech Republic, for six months. In 2009–2012, she will be leading the project "Applications of Class Field Theory in Present and Future Multiuser Communications" at the University of Turku. She will join the Department of Mathematics, University of Tampere, Finland, as a Lecturer for the academic year 2009–2010. Her research is in the area of applications of algebraic number theory in wireless communications.

Dr. Hollanti is a recipient of several grants from various foundations, including the Finnish Cultural Foundation research grant in 2007 and the Finnish Academy of Science research grant in 2008. She has also won the prize for the best presentation at the 2007 European Women in Mathematics (EWM) conference, Cambridge, U.K.