Improving the DMT Performance for MIMO Communication With Linear Receivers

Ti-Wen Tang, *Student Member, IEEE*, Min-Kun Chen, and Hsiao-Feng Lu, *Senior Member, IEEE*

***Abstract*—Multiple-input–multiple-output (MIMO) linear receivers are often of more practical interest than maximum-likelihood (ML) receivers due to their low decoding complexity but at the cost of worse diversity gain performance. Such a statement on performance loss is due to the assumption of using an independent identically distributed complex Gaussian vector as channel input. By removing this assumption, we find that the diversity performance of MIMO linear receivers can be significantly improved. In an extreme case, it can be the same as that of ML receivers. Specifically, in this paper, we investigate the diversity–multiplexing tradeoff (DMT) performance of MIMO linear receivers with colored and possibly degenerate Gaussian channel inputs. By varying the rank of the covariance matrix of the channel input vector and by allowing temporal coding across multiple channel uses, we show that the MIMO linear receiver can achieve a much better DMT performance than the currently known one. Explicit optimal code constructions are provided, along with simulation results, to justify the above findings. For the case of (2 × 2) and (3 × 3) MIMO linear receivers, simulation results show that the newly proposed codes provide significant gains of 10 and 12.08 dB in E_b/N_0 at bit error rate 10^{-4} compared to the conventional schemes, respectively.**

***Index Terms*—Diversity–multiplexing tradeoff (DMT), explicit optimal code constructions, linear receivers, precoding matrix, quasi-static multiple-input–multiple-output (MIMO) channel.**

I. Introduction

The wireless multiple-input–multiple-output (MIMO) communication system equipped with n_t transmit and n_r receive antennas can simultaneously provide a much higher rate and a much better error performance than the conventional single-input–single-output (SISO) system [1], [2]. Such performance advantages are measured by the notions of (spatial) multiplexing gain and diversity gain, respectively. There is also a fundamental tradeoff between these two quantities, termed diversity-multiplexing gain tradeoff (DMT) [3], which can be achieved by any MIMO system.

Consider a quasi-static $(n_t \times n_r)$ MIMO Rayleigh fading channel with a targeting transmission rate $R = r \log \text{SNR}$ (in bits per channel use or bpcu), where r is termed the multiplexing gain, and the logarithm is taken to the base 2. Assuming that the number of channel uses $\geq n_t + n_r - 1$, Zheng and Tse [3] showed that the optimal diversity gain $d^*(r)$ that can be achieved by any space–time coding scheme is a piecewise linear function[1] connecting the points $(r, (n_t - r)(n_r - r))$ for $r = 0, 1, \ldots, K$, where $K := \min\{n_t, n_r\}$ is commonly known as the channel degree of freedom. Moreover, given n_t and r, there are space–time codes constructed explicitly from some cyclic division algebra (CDA) [4] that can achieve the optimal diversity gain $d^*(r)$ in n_t channel uses at the high signal-to-noise ratio (SNR) regime provided that maximum-likelihood (ML) or bounded-distance decoding is employed. While the computational complexity of ML decoding is extremely high, the CDA-based codes can be represented in a linear-dispersion form [5] and hence could be decoded by the use of sphere decoding [6], [7]. The sphere decoder is known to be ML equivalent but has a relatively lower complexity. In [8], El Gamal *et al.* introduced the minimum mean-square error generalized decision-feedback equalizer (MMSE-GDFE) to decode the class of LAttice Space Time codes, which is DMT optimal. The MMSE-GDFE receiver also falls in the class of sphere decoding [8].

Sphere decoding of linear-dispersion space–time codes in general has a complexity in the form of $O(|\mathcal{A}|^e)$, where $\mathcal{A} \cup \mathbb{Z}$ is the base alphabet, and e is some exponent depending upon the dimension of the code lattice. The value e is often quite large and can be much larger if the MIMO system is underdetermined, i.e., when $n_r < n_t$. For the overdetermined MIMO system, i.e., $n_t \leq n_r$, it is known [9], [10] that the expected number of visited nodes in the sphere decoder is lower bounded by $(|\mathcal{A}|^{\eta n_t} - 1)/(\sqrt{|\mathcal{A}|} - 1)$, where $\eta = (1/2)(1 + (4(|\mathcal{A}| - 1)/3\sigma^2)\text{SNR})^{-1}$, and $\sigma^2 = \mathbb{E}_{a \in \mathcal{A}} |a|^2$. Thus, in case of large-size base alphabet, i.e., $|\mathcal{A}| \gg 1$, there is a great concern whether the sphere decoder could meet the requirement on decoding latency and low power consumption in practical mobile and vehicular MIMO communications. Hence, it is very often of much practical interest to replace the sphere decoder with the linear detector, whose complexity is roughly $O(n_t^{2.3})$ [11] for matrix inversion and is independent of $|\mathcal{A}|$, even knowing that the latter could lead to considerable performance degradation beforehand [12]–[15]. Implementations of MIMO communication using linear receivers have been extensively investigated for mobile and vehicular communications, see, for example, [12]–[22].

Manuscript received February 7, 2012; revised August 17, 2012 and November 1, 2012; accepted November 9, 2012. Date of publication November 29, 2012; date of current version March 13, 2013. This paper was presented in part at the IEEE International Symposium on Information Theory, St. Petersburg, Russia, August 2011. The review of this paper was coordinated by Prof. A. M. Tonello.

T.-W. Tang and H.-F. Lu are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: dennistang.cm98g@nctu.edu.tw; francis@mail.nctu.edu.tw).

M.-K. Chen was with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan. He is now with the Taiwan Semiconductor Manufacturing Company Limited, Tainan 74144, Taiwan (e-mail: mkchenj@tsmc.com).

Digital Object Identifier 10.1109/TVT.2012.2230348

[1]For simplicity, in this paper we will write $d^*(r) = (n_t - r)(n_r - r)$ directly.

Assuming that the receiver has full channel state information, the DMT performance of MIMO linear receivers such as zero-forcing (ZF) or minimum mean-square error (MMSE) receiver has been studied in several works. In [16], it is shown that if the code has finite size (hence r approaches 0 as SNR $\to \infty$) and is transmitted in single channel use, the maximal diversity gain is given by $n_r - n_t + 1$, which is significantly smaller than the optimal diversity value $d^*(r = 0) = n_t n_r$. Kumar *et al.* [23, Th. 1] extended the DMT analysis to the case when the code has size SNR^r, growing as the SNR increases, and showed that ZF and MMSE receivers have the same DMT performance given by $(n_r - n_t + 1)(1 - (r/n_t))^+$ for multiplexing gain r, where $(x)^+ := \max\{0, x\}$. For the sake of completeness, we reproduce this result in the following theorem.

Theorem 1 [23, Th. 1]: The DMT of the n_t-transmit n_r-receive independent identically distributed (i.i.d.) Rayleigh MIMO channel with $n_r \geq n_t$, constrained to use Gaussian codes under either MMSE or ZF linear receivers, is given by

$$d^*_{lin}(r) = (n_r - n_t + 1)\left(1 - \frac{r}{n_t}\right)^+$$

for both cases of coding across antennas or pure spatial multiplexing.

It should be noted that the above result is based on the assumption of an i.i.d. complex Gaussian channel input in single channel use. To compensate for the performance loss due to the use of linear receivers, several modified MMSE receivers have been proposed [17], [18], [24]–[35]. In particular, [17], [18], and [24]–[29] proposed performing the Lenstra–Lenstra–Lováz (LLL) lattice basis reduction algorithm [36] on the channel matrix prior to linear equalization. It is shown [37] that with the additional LLL reduction, the maximal diversity gain can be increased to n_r for a finite code rate R, corresponding to the case of $r \downarrow 0$, in single channel use, where $r \downarrow 0$ means r approaches 0 from the right, i.e., $r \to 0^+$. A recent result given by Jaldén and Elia [26] shows that with LLL lattice basis reduction in preprocessing, the MMSE receivers can achieve the optimal ML performance in the DMT sense.

A. Main Contributions of This Paper

The contributions of this paper are the following. First, we remark that although the i.i.d. complex Gaussian channel input is sufficient for achieving the optimal DMT in general MIMO channel under ML decoding, as was established by Zheng and Tse [3], such channel input is in general not optimal for linear receivers. Second, based on [23], we extend the DMT analysis to the case when multiple channel uses and temporal coding are allowed. As most of the existing space–time codes can be represented in linear-dispersion forms, it makes so much sense to study the DMT performance for MIMO linear receivers in multiple channel uses by taking into account the possibility of temporal coding, even when the channel remains fixed throughout. Motivated by the above, in this paper, we investigate the DMT performance of MIMO linear receivers with general Gaussian codebooks. By which we mean that colored (and possibly degenerate) complex Gaussian random matrices will be used as channel input. We show that with such relaxation, a much better DMT performance is indeed achievable by linear receivers. Specifically, we report the following contributions in this paper.

1) Our main result is reported in Theorem 5, where we show that by varying the rank of the covariance matrix of the colored, degenerate, Gaussian input and by allowing temporal coding at the transmitter, the linear receiver indeed can achieve the maximal possible diversity gain $n_t n_r$ as r approaches 0. This is much larger than all previously known results.
2) Similar improvements of DMT performance are also obtained for $r > 0$ through our approach. We report in Theorem 4 that the optimal DMT for MMSE linear receiver is lower bounded by

$$d^n_{\text{MMSE}}(r) \geq \max_{\substack{n,m \in \mathbb{Z}^+: \\ 1 \leq nm \leq n_t}} (nn_r - m + 1)\left(1 - \frac{nr}{m}\right)^+$$

while the existing best result [23, Th. 1] had $(n_r - n_t + 1)(1 - (r/n_t))^+$ due to an assumption of a full rank i.i.d. Gaussian input.
3) We provide a systematic construction of temporal codes in Section IV that can achieve the above lower bound on the DMT for linear receivers. For the (2×2) and (3×3) MIMO linear receivers, the simulation results show that the newly proposed codes provide significant gains of 10 and 12.08 dB in bit SNR E_b/N_0 at bit error rate 10^{-4} compared to the conventional schemes, respectively.

Our findings are well justified by the simulations of outage probabilities of the newly proposed schemes as well as the bit error rates of the proposed codes.

Notation: Underlined lowercase letter $\underline{x}$ represents a vector, and uppercase letter A denotes a matrix of certain size. The (i, j)th entry of matrix A is denoted by $A_{i,j}$. $A^\dagger$ (respectively $A^\top$) denotes the Hermitian transpose (respectively transpose) of matrix A, and $\|A\|$ denotes its Frobenius norm. I_n is the $(n \times n)$ identity matrix, and $\mathbf{0}_n$ denotes the all-zero $(n \times n)$ matrix. Matrix inequalities such as $\succeq$, $\preceq$, $\succ$, and $\prec$ represent the partial orderings of positive semidefinite matrices [38, Sec. 7.7]. For example, we say $A \succeq B$ if $A - B$ is a nonnegative definite matrix. $\underline{x} \sim \mathbb{CN}(\underline{m}, K_x)$ stands for a circularly symmetric complex Gaussian random vector $\underline{x}$ with mean $\underline{m}$ and covariance matrix K_x. $f(x)$ and $g(x)$ are said to be exponentially equal, denoted by $f(x) \doteq g(x)$, if $\lim_{x\to\infty}(\log f(x)/\log x) = \lim_{x\to\infty}(\log g(x)/\log x)$, provided that the limits exist. Exponential inequalities, indicated by $\dot{\leq}$, $\dot{\geq}$, $\dot{<}$, and $\dot{>}$, are defined similarly. $\mathcal{U}(n)$ denotes the multiplicative group of $(n \times n)$ unitary matrices over $\mathbb{C}$. $d^{(n)}_{\text{MMSE}}(r)$ is the DMT of MIMO linear receivers with multiplexing gain r in n channel uses. The unit "bpsu" is for bits per subchannel use.

II. Improved Diversity-Multiplexing Tradeoff for Linear Receivers

In this section, we investigate the DMT performance of MIMO linear receivers using colored and possibly degenerate complex Gaussian vector as channel input. For brevity, we will restrict ourselves to the study of MMSE receivers. Our techniques can be directly applied to the study of ZF receivers, and all the results will hold the same. Although there are modified MMSE receivers with improved performance such as those reported in [30]–[34], in this paper, we will focus only on the elementary one.

For simplicity, in Section II-A, we will first conduct the DMT analysis of the MIMO linear receiver subject to single channel use. It will be shown that even in this case, by varying the rank of the covariance matrix of the channel input vector, the resulting maximal diversity gain can already be increased to the value n_r compared to the original value of $(n_r - n_t + 1)$ given in [16], [23]. In Section III, we will extend the results to the case of multiple channel uses with temporal coding allowed.

A. DMT of Linear Receiver Under Single Channel Use

Consider an $(n_t \times n_r)$ MIMO channel with the following channel input–output relation:

$$\underline{y} = \sqrt{\text{SNR}} H \underline{x} + \underline{w} \tag{1}$$

where $\underline{x} \sim \mathbb{CN}(\underline{0}, K_x)$ is the transmitted code vector subject to a power constraint of $\text{Tr}(K_x) \leq 1$. H is the $(n_r \times n_t)$ channel matrix with i.i.d. $\mathbb{CN}(0,1)$ entries and is assumed to be known to the receiver, but not to the transmitter. $\underline{w} \sim \mathbb{CN}(\underline{0}, I_{n_r})$ is the additive complex Gaussian noise. We assume $n_t \leq n_r$ throughout.

From the spectral theorem for Hermitian matrices [38, Th. 2.5.6], let $K_x = U_x \Lambda_x U_x^\dagger$ be the eigen-decomposition of K_x, where $U_x \in \mathcal{U}(n_t)$ is an $(n_t \times n_t)$ unitary matrix, and Λ_x is a diagonal matrix consisting of eigenvalues of K_x. Then, we can rewrite (1) as

$$\underline{y} = \sqrt{\text{SNR}} \tilde{H} \underline{\tilde{x}} + \underline{w}$$

where $\tilde{H} = H U_x$ and $\underline{\tilde{x}} = U_x^\dagger \underline{x}$ such that $K_{\underline{\tilde{x}}} = \mathbb{E}\underline{\tilde{x}}\underline{\tilde{x}}^\dagger = \Lambda_x$. Let $m := \text{rank}(\Lambda_x) \leq n_t$. We can assume without loss of generality that the first m diagonal entries of Λ_x are all nonzero. The transmitted vector $\tilde{x}$ is statistically equivalent to

$$\underline{\tilde{x}} = \begin{bmatrix} \underline{s} \\ \underline{0}_{(n_t - m)\times 1} \end{bmatrix}$$

for some $\underline{s} \sim \mathbb{CN}(\underline{0}, \Lambda)$, where Λ is the $(m \times m)$ diagonal submatrix of Λ_x, consisting of the nonzero eigenvalues of K_x. Finally, the channel model (1) simplifies to

$$\underline{y} = \sqrt{\text{SNR}} H_s \underline{s} + \underline{w} \tag{2}$$

where H_s is the $(n_r \times m)$ equivalent channel matrix formed by the leftmost m columns of $\tilde{H}$. Observing that $\tilde{H} = H U_x$ and H share the same statistical distribution, and that H_s is an $(n_r \times m)$ channel matrix with i.i.d. $\mathbb{CN}(0,1)$ entries, (2) can be regarded as an $(m \times n_r)$ MIMO channel with channel matrix H_s and channel input vector $\underline{s} \sim \mathbb{CN}(\underline{0}, \Lambda)$ subject to the constraints of $\Lambda \succ \mathbf{0}_m$ and $\text{Tr}(\Lambda) \leq 1$. It then follows from [23, Th. 1] that we have the following theorem.

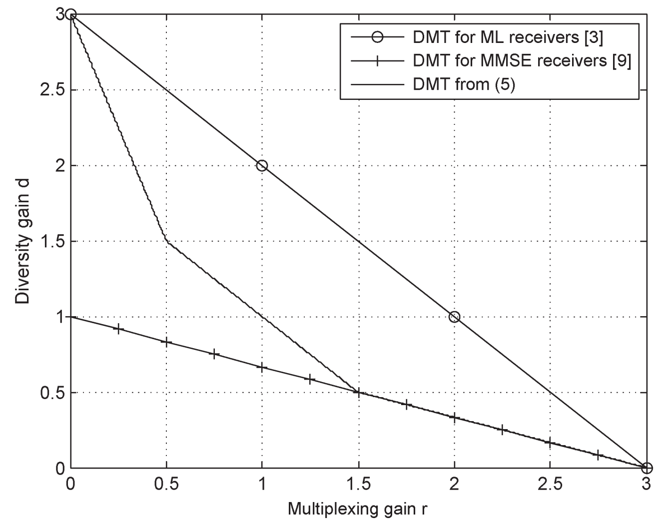

Fig. 1. DMTs for ML receivers [3, Sec. IV], MMSE receivers [23], and that from (5) for the (3×3) MIMO in single channel use.

Theorem 2: Given $\Lambda = (1/m) I_m$, the DMT of $(n_t \times n_r)$ i.i.d. MIMO Rayleigh channel with $m \leq n_t \leq n_r$ and multiplexing gain r, using either linear MMSE or ZF receivers in single channel use, is given by

$$d_{\text{MMSE}}^{(1)}(r)|_{\Lambda = \frac{1}{m} I_m} = (n_r - m + 1)\left(1 - \frac{r}{m}\right)^+. \tag{3}$$

■

It is clear from the foregoing theorem that even in the case of single channel use, the maximal diversity gain of either MMSE or ZF linear receiver can be significantly improved to n_r after setting $m = 1$ and $r = 0$. The value is much larger than the previously known $(n_r - n_t + 1)$ in [16] and [23].

While Theorem 2 focuses only on the case $\Lambda = (1/m) I_m$, below we give a much stronger result that the same DMT (3) holds for all covariance matrices K_x satisfying $K_x \succeq \mathbf{0}_{n_t}$, $\text{Tr}(K_x) \leq 1$, and $\text{rank}(K_x) = m$.

Theorem 3: The DMT of the $(n_t \times n_r)$ i.i.d. MIMO Rayleigh channel with $\text{rank}(K_x) = m \leq n_t \leq n_r$ and multiplexing gain r, using either linear MMSE or ZF receivers in single channel use, is given by

$$d_{\text{MMSE}}^{(1)}(r)|_{rank(K_x) = m} = (n_r - m + 1)\left(1 - \frac{r}{m}\right)^+. \tag{4}$$

■

By varying the values of m, in Fig. 1 we show the best DMT performance offered by Theorem 3 for the (3×3) MIMO channel in single channel use. That is, for each r, we find

$$d_{\text{MMSE}}^{(1)}(r) = \max_{\substack{m \in \mathbb{Z}: \\ 1 \leq m \leq n_t}} (n_r - m + 1)\left(1 - \frac{r}{m}\right)^+. \tag{5}$$

It can be clearly seen that the colored degenerate Gaussian input can achieve a much larger DMT than that in [23], which uses an i.i.d. Gaussian input. We shall also emphasize that the DMT for the ML receiver in the case of single channel use is $n_r(1 - (r/n_t))$ in Fig. 1. See [3, Sec. IV-D] for details.

B. Proof of Theorem 3

To prove the theorem, we note that the linear MMSE equalizer for (2) is given by

$$W_{\text{MMSE}} := \sqrt{\text{SNR}}\Lambda H_s^\dagger \left(\text{SNR} \cdot H_s \Lambda H_s^\dagger + I_{n_r}\right)^{-1} \tag{6}$$

$$= \sqrt{\text{SNR}} \left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right)^{-1} \Lambda H_s^\dagger \tag{7}$$

where the second equality is immediate from the following identity based on the associative law for matrix[2]

$$\left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right) \Lambda H_s^\dagger = \Lambda H_s^\dagger \left(\text{SNR} \cdot H_s \Lambda H_s^\dagger + I_{n_r}\right).$$

The corresponding equalized output is then given by

$$\underline{\hat{y}} := W_{\text{MMSE}}\, \underline{y} = G\underline{s} + \underline{n} \tag{8}$$

where the equivalent channel matrix is

$$\begin{aligned} G &:= \sqrt{\text{SNR}} \cdot W_{\text{MMSE}} H_s \\ &= \text{SNR} \cdot \left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right)^{-1} \Lambda H_s^\dagger H_s \\ &= \left(\text{SNR}\Lambda H_s^\dagger H_s + I_m\right)^{-1} \left[\left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right) - I_m\right] \\ &= I_m - \left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right)^{-1}. \end{aligned} \tag{9}$$

After obtaining $\underline{\hat{y}}$, the MMSE receiver makes an independent decision for each entry of $\underline{s}$. Hence, the channel model (8) can be interpreted as m parallel fading subchannels. Each subchannel is given by

$$\hat{y}_\ell = \underline{g}_\ell^\top \underline{s} + n_\ell = G_{\ell,\ell} s_\ell + z_\ell, \quad \ell = 1, 2, \ldots, m \tag{10}$$

where $\underline{g}_\ell^\top$ is the ℓth row vector of G. z_ℓ is the lump noise in the ℓth subchannel, i.e., it is the sum of the Gaussian noise n_ℓ, which is correlated with other noise n_i, and the interference from other transmitted antennas. Specifically, it is given by

$$z_\ell = \sum_{i=1, i\neq \ell}^{m} G_{\ell,i} s_i + n_\ell.$$

Although the lump noises z_ℓ are correlated, the MMSE detector simply ignores this fact and makes an independent decision for each s_ℓ, thereby yielding a low complexity detection. ■

For convenience, let

$$P := \left(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m\right)^{-1}.$$

Then, the channel input–output mutual information for the nonergodic ℓth subchannel is given by

$$I(s_\ell; \hat{y}_\ell) = \log(1 + \text{SINR}_\ell) = -\log P_{\ell,\ell} \tag{11}$$

where SINR_ℓ denotes the signal-to-interference-plus-noise power ratio (SINR) of the ℓth subchannel. The proof of the second equality in (11) is relegated to the Appendix for ease of reading. A further justification of $0 < P_{\ell,\ell} \leq 1$ will also be provided therein.

We also emphasize that the overall spatial multiplexing gain is kept constant for having a fair comparison between the previous and proposed schemes. Specifically, a system with $\text{rank}(K_x) = m$ has m subchannels, as shown in (10). The total sum rate of the m subchannels is set at $R = r \log \text{SNR}$ bpcu, or equivalently at spatial multiplexing gain r. It follows that each subchannel transmits on the average at a rate of (R/m) bpsu, where we have assumed an equal-rate split. This is the best rate-split scheme since the transmitter has no channel knowledge.

Now, given Λ with rank m and transmission rate R, the overall channel outage probability is

$$P_{\text{out,MMSE}}(r|\Lambda) = \Pr\left\{\bigcup_{\ell=1}^{m} \left\{H_s : I(s_\ell; \hat{y}_\ell) \leq \frac{R}{m}\right\}\right\}.$$

Note that the foregoing outage probability is conditioned on the explicit choice of Λ. It can also be easily seen that at high SNR regime the above outage probability is exponentially dominated by the worst outage probability associated with each subchannel, i.e.,

$$P_{\text{out,MMSE}}(r|\Lambda) \doteq \max_{1\leq \ell \leq m} \Pr\left\{H_s : I\left(s_\ell; \hat{y}_\ell | H_s\right) \leq \frac{R}{m}\right\}.$$

Thus, given constraint of $\text{rank}(K_x) = m$ and $R = r \log \text{SNR}$, the smallest outage probability with respect to the choice of Λ in our case, when minimized over all possible Λs, is given by

$$P_{\text{out,MMSE}}\left(r|\text{rank}(K_x) = m\right) \tag{12}$$

$$\doteq \inf_{\substack{\Lambda \succ \mathbf{0}_m \\ Tr(\Lambda)\leq 1}} \max_{1\leq \ell\leq m} \Pr\left\{H_s : I(s_\ell; \hat{y}_\ell | H_s) \leq \frac{R}{m}\right\}. \tag{13}$$

With the above, the proof of Theorem 3 calls for upper and lower bounds of (13). Clearly, setting $\Lambda = (1/m)I_m$ yields a valid upper bound on the outage probability, which in turn gives a lower bound on the diversity gain. It is already given in Theorem 2, and we only need to develop a lower bound for (13).

To bound $P_{\text{out,MMSE}}(r\,|\text{rank}(K_x) = m)$ from below, we let $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ for convenience. It should be noted that $\lambda_1, \ldots, \lambda_m$ are constants for power allocation. Specifically, λ_i denotes the ratio of the power used by the ith subchannel to the total power. It is a predetermined, fixed, unitless constant and is independent of the SNR.

As the matrix P can be factored into

$$P = \left(\text{SNR} \cdot H_s^\dagger H_s + \Lambda^{-1}\right)^{-1} \Lambda^{-1}$$

we will focus on the first term in the foregoing product. From the obvious inequality $\mathbf{0}_m \prec \Lambda \prec I_m$, we have

$$\left(\text{SNR} \cdot H_s^\dagger H_s + \Lambda^{-1}\right) \prec \left(\text{SNR} \cdot H_s^\dagger H_s + \lambda_{\min}^{-1} I_m\right)$$

[2] An alternative derivation of (7), (9), (11) for the case of $\Lambda = (1/m)I_m$ can be found in [19], [39]. The approach presented here is more general and follows from a much simpler and more elementary technique.

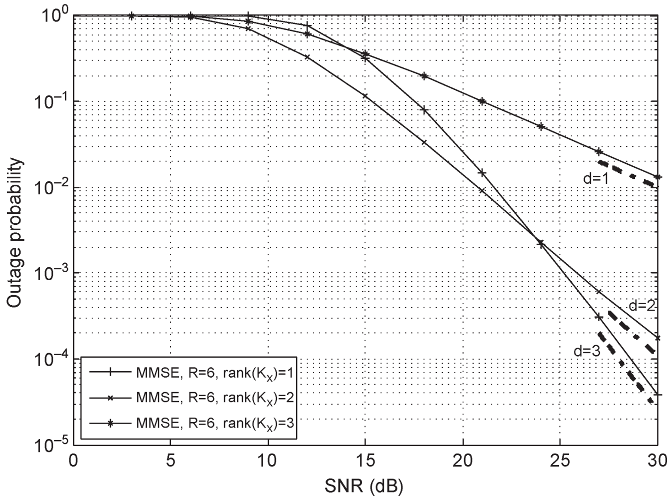

Fig. 2. Outage probabilities for MMSE receivers with various values of $\text{rank}(K_x)$ for (3 × 3) MIMO in single channel use at a constant sum rate $R = 6$ bpcu. The dash lines represent theoretical high SNR slopes for diversity gains.

where $\lambda_{\min} := \min_{1\le\ell\le m} \lambda_\ell > 0$. It follows that

$$\begin{aligned}&\left(\text{SNR}\cdot H_s^\dagger H_s + \Lambda^{-1}\right)^{-1}\\ &\quad\succ \left(\text{SNR}\cdot H_s^\dagger H_s + \lambda_{\min}^{-1} I_m\right)^{-1}\\ &\left[\left(\text{SNR}\cdot H_s^\dagger H_s + \Lambda^{-1}\right)^{-1}\right]_{\ell,\ell}\\ &\quad> \left[\left(\text{SNR}\cdot H_s^\dagger H_s + \lambda_{\min}^{-1} I_m\right)^{-1}\right]_{\ell,\ell}\end{aligned}$$

for $\ell = 1, \ldots, m$ since both matrices are positive definite. It is then clear that

$$\begin{aligned}I(s_\ell; \hat{y}_\ell | H_s) =& -\log P_{\ell,\ell}\\ &< -\log\left[\left(\lambda_{\min}\text{SNR}\cdot H_s^\dagger H_s + I_m\right)^{-1}\right]_{\ell,\ell}\\ &\quad+ \log\lambda_\ell - \log\lambda_{\min}.\end{aligned}$$

Since λ_ℓ is independent of SNR and $\lambda_\ell \doteq \text{SNR}^0$ for all ℓ, following the Laplace principle as used in [3] and [40], we can neglect the last two terms in the above and

$$\begin{aligned}&P_{\text{out,MMSE}}\left(r|\text{rank}(K_x) = m\right)\\ &\quad\dot{>} \max_{1\le\ell\le m} \Pr\left\{-\log\left[\left(\text{SNR}\cdot H_s^\dagger H_s + I_m\right)^{-1}\right]_{\ell,\ell} \le \frac{R}{m}\right\}\\ &\quad\doteq \text{SNR}^{-(n_r - m+1)(1-\frac{r}{m})^+}\end{aligned} \tag{14}$$

where the last exponential equality can be obtained following arguments similar to [23, Th. 1], and details are omitted for brevity. Finally, combining the upper bound (3) (obtained by setting $\Lambda = (1/m)I_m$) and lower bound (14) proves Theorem 3.

In Fig. 2, we provide computer simulations of the outage probabilities of the (3 × 3) MIMO channel in single channel use at a constant rate of $R = 6$ bpcu for various values of m. We emphasize that we have assumed throughout this paper that the overall spatial multiplexing gain is kept constant for having a fair comparison between the previous and proposed schemes. Thus, for example, in the case of $\text{rank}(K_x) = 1$, there is only one subchannel, and hence, it transmits at 6 bpsu. The case of $\text{rank}(K_x) = 3$, on the other hand, has three subchannels, and each transmits at 2 bpsu. In other words, a system with a lower rank of K_x, in keeping the total sum rate R a constant, each subchannel would have to transmit at a higher rate. Therefore, it does not necessarily imply that systems with smaller m would yield a better outage performance, as seen from Fig. 2.

The dash lines in Fig. 2 represent theoretical high SNR slopes for diversity gains based on (4) in Theorem 3. The simulated diversity gains match perfectly with the result given in Theorem 3 at high SNR regime. We note that while the scheme of $\text{rank}(K_x) = 1$ yields the largest diversity gain of 3 at the high SNR regime, it does not necessarily mean that the scheme has the smallest outage probability at low and moderate SNR values. Moreover, we find that for SNR value below 24 dB, the scheme of $\text{rank}(K_x) = 2$ actually delivers the best possible outage performance in single channel use.

III. Diversity -Multiplexing Tradeoff for Linear Receiver With Temporal Coding Allowed

So far, we have characterized the exact DMT performance of MIMO linear receivers in single channel use. In this section, we will extend the work to the case of multiple channel uses, when temporal coding is allowed and the channel remains fixed throughout. Almost all of the existing space–time codes include temporal coding. One such scheme, termed linear dispersion space–time code [5], [41], can be easily incorporated into the context of MIMO linear receivers. To elaborate, in a linear dispersion code, each code matrix is associated with a point in some lattice. It is obtained from an integer combination of some basis matrices, which together form a generator matrix of the lattice. In other words, each transmitted code matrix is obtained by a fixed lattice generator matrix multiplied by a coordinate vector [see (17)]. At the receiving end, the code matrix will be left-multiplied by the channel matrix. Thus, the linear channel equalization can be taken upon the product of the channel matrix and the fixed lattice-generator matrix. Now, we give the following theorem to provide a general lower bound on the DMT of the linear MMSE or ZF receiver.

Theorem 4: For an $(n_t \times n_r)$ i.i.d. quasi-static MIMO Rayleigh fading channel with $n_t \le n_r$ that is fixed for at least n channel uses, given multiplexing gain r, the DMT of the linear MMSE or ZF receiver is lower bounded by

$$d_{\text{MMSE}}^{(n)} \ge \max_{\substack{m,n\in\mathbb{Z}^+:\\ 1\le nm\le n_t}} (nn_r - m + 1)\left(1 - \frac{nr}{m}\right)^+ \tag{15}$$

where m is the rank of Gaussian input covariance matrix.

Proof: Assume that temporal coding is applied in n uses of the $(n_t \times n_r)$ quasi-static MIMO channel within which the channel matrix remains fixed. The channel input–output relation is given by

$$\underline{y}_i = \sqrt{\text{SNR}} H \underline{x}_i + \underline{w}_i, \quad i = 1, \ldots, n$$

where $\underline{x}_i$ satisfies the constraint $\mathbb{E}\|\underline{x}_i\|^2 \le 1$. By vertically stacking the receive signal vectors, the n-shot channel is equivalent to

$$\underline{y} = \begin{bmatrix}\underline{y}_1\\ \vdots\\ \underline{y}_n\end{bmatrix} = \sqrt{\text{SNR}} \underbrace{\begin{bmatrix}H & & \\ & \ddots & \\ & & H\end{bmatrix}}_{:=\tilde{H}} \underline{x} + \underline{w} \tag{16}$$

where $\underline{x} = [\underline{x}_1^\top \ \cdots \ \underline{x}_n^\top]^\top$ and $\underline{w} = [\underline{w}_1^\top \ \cdots \ \underline{w}_n^\top]^\top$. Let $\underline{x} \sim \mathbb{CN}(\underline{0}, K_x)$ with $\text{rank}(K_x) = m$ and power constraint $\text{Tr}(K_x) \leq n$. Decompose K_x as $K_x = U_x \Lambda U_x^\dagger$, where U_x is an $(nn_t \times m)$ matrix $1 \leq m \leq nn_t$ satisfying $U_x^\dagger U_x = I_m$ but $U_x U_x^\dagger \neq I_{nn_t}$. Λ is a $(m \times m)$ diagonal matrix with nonzero diagonal entries. Then, the stacking channel (16) can be rewritten as

$$\underline{y} = \sqrt{\text{SNR}}\tilde{H}U_x\underline{s} + \underline{w} = \sqrt{\text{SNR}}H_s\underline{s} + \underline{w} \tag{17}$$

where $\underline{s} \sim \mathbb{CN}(\underline{0}, \Lambda)$ and $H_s = \tilde{H}U_x$. We remark that unlike the case of single channel use, the entries of H_s are not of i.i.d. $\mathbb{CN}(0, 1)$. This is because $\tilde{H}$ is a block diagonal matrix consisting of n repetitions of the same matrix H. ∎

Given transmission rate $R = r \log \text{SNR}$ (bits per channel use) with $\text{rank}(K_x) = m$ in n channel uses, the overall channel outage probability is

$$P_{\text{out,MMSE}}(r|\text{rank}(K_x) = m, n\text{channel uses}, U_x)$$
$$\doteq \inf_{\Lambda \succ \mathbf{0}_m, Tr(\Lambda) \leq n} \max_{1 \leq \ell \leq m} \Pr\left\{H_s : I(s_\ell; \hat{y}_\ell | H_s) \leq \frac{nR}{m}\right\}.$$

The remaining discussion is similar to that in the previous section. The upper bound on the smallest outage probability can be obtained by setting $\Lambda = (n/m)I_m$ with an extra condition of $U_x^\dagger U_x = I_m$ that should be taken into account. Specifically, it is given by

$$\begin{aligned} &P_{\text{out,MMSE}}(r|\text{rank}(K_x) \\ &\quad = m, n\text{channel uses}) \\ &\quad \dot{\leq} \inf_{U_x \in \mathcal{U}} \max_{1 \leq \ell \leq m} \\ &\qquad \times \Pr\left\{\left[\left(\frac{n}{m}\text{SNR}H_s^\dagger H_s + I_m\right)^{-1}\right]_{\ell,\ell} \geq \text{SNR}^{-\frac{nr}{m}}\right\} \\ &\quad \leq \inf_{U_x \in \mathcal{U}} \Pr\left\{\text{Tr}\left[\left(\frac{n}{m}\text{SNR}H_s^\dagger H_s + I_m\right)^{-1}\right] \geq \text{SNR}^{-\frac{nr}{m}}\right\} \end{aligned} \tag{18}$$

where $\mathcal{U} = \{U_x \in \mathbb{C}^{nn_t \times m} : U_x^\dagger U_x = I_m\}$. The second inequality follows from the fact that the diagonal entries of $((n/m)\text{SNR}\, H_s^\dagger H_s + I_m)^{-1}$ are all positive. It is seen from (18) that the upper bound actually depends on the choice of "precoding matrix" U_x, which would affect the outage performance of MIMO linear receivers.

For simplicity, we consider the following specific type of U_x

$$\mathcal{U}_x := \left\{U_x = \begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix} : U_i \in \mathbb{C}^{n_t \times m}, U_i^\dagger U_j = \frac{\delta_{i,j}}{n} I_m\right\} \tag{19}$$

where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. The column vectors of each submatrix U_i are of length n_t and are mutually orthogonal to each other due to the second constraint of $U_i^\dagger U_j = (\delta_{i,j}/n)I_m$. It follows that the set $\mathcal{U}_x$ is nonempty if and only if $nm \leq n_t$. Moreover, we remark that there is a one–one correspondence between elements in the set $\mathcal{U}_x$ and those in the unitary group $\mathcal{U}(n_t)$. Now, given any $U_x \in \mathcal{U}_x$, H_s can be expressed as

$$H_s = \tilde{H}U_x = \begin{bmatrix} HU_1 \\ \vdots \\ HU_n \end{bmatrix}.$$

To find the statistical distribution of entries of H_s, note that any two rows of H_s are in the form of $\underline{h}_i^\top U_k$ and $\underline{h}_j^\top U_l$ for some i, j, k, l, where $\underline{h}_i^\top$ and $\underline{h}_j^\top$ are the ith and jth rows of H, respectively. The correlation matrix formed by the rows is $\mathbb{E}\{(\underline{h}_j^\top U_l)^\dagger \underline{h}_i^\top U_k\} = \delta_{i,j} U_l^\dagger U_k = (1/n)\delta_{i,j}\delta_{l,k} I_m$, where the last equality follows from (19). Thus, the entries of H_s are i.i.d. $\mathbb{CN}(0, (1/n))$.

Thus, with the specific type $\mathcal{U}_x$, the channel model (17) can be seen as an $(m \times nn_r)$ i.i.d. MIMO Rayleigh fading channel in single channel use. Note that $\mathcal{U}_x \cup \mathcal{U}$. By directly applying the results from the previous section to the present case, the smallest outage probability of MIMO linear receivers in n channel uses and with $\text{rank}(K_x) = m$ can be bounded from the above by

$$\begin{aligned} &P_{\text{out,MMSE}}\,(r|\text{rank}(K_x) = m, n \text{ channel uses}) \\ &\dot{\leq} \inf_{U_x \in \mathcal{U}_x} \Pr\left\{\text{Tr}\left[\left(\frac{n}{m}\text{SNR}\, H_s^\dagger H_s + I_m\right)^{-1}\right] \geq \text{SNR}^{-\frac{nr}{m}}\right\} \\ &\doteq \text{SNR}^{-(nn_r - m + 1)(1 - \frac{nr}{m})^+}. \end{aligned}$$

We remark that by setting $n = n_t$ and $m = \text{rank}(K_x) = 1$, Theorem 4 provides the following series of inequality on the DMT:

$$d^*(r) \geq d^*_{\text{MMSE}}(r) \geq d^{(n_t)}_{\text{MMSE}}(r) \geq n_t n_r (1 - n_t r)^+$$

where $d^*_{\text{MMSE}}(r)$ denotes the best possible DMT that can be achieved by MIMO MMSE receivers. Thus, for multiplexing gain $r \downarrow 0$, the maximal diversity gain $n_t n_r$ can indeed be achieved by the use of MIMO linear receivers without any further help from preprocessing, such as the LLL lattice reduction.

Before concluding this section, we further discuss the design of precoding matrix U_x. We offer the following theorem for the case of $mn \leq n_t$.

Theorem 5: For an $(n_t \times n_r)$ i.i.d. quasi-static Rayleigh fading MIMO channel with $n_t \leq n_r$ that is fixed for at least n channel uses, let

$$U_x = \begin{bmatrix} \underline{u}_{11} & \cdots & \underline{u}_{1m} \\ \vdots & \ddots & \vdots \\ \underline{u}_{n1} & \cdots & \underline{u}_{nm} \end{bmatrix}$$

be the $(nn_t \times m)$ precoding matrix defined as before, i.e., we have $K_x = U_x \Lambda U_x^\dagger$, $\text{rank}(K_x) = m$ and $\text{Tr}(K_x) \leq n$. If the set of vectors $\{\underline{u}_{11}, \ldots, \underline{u}_{nm}\}$ is linearly independent over $\mathbb{C}$, then

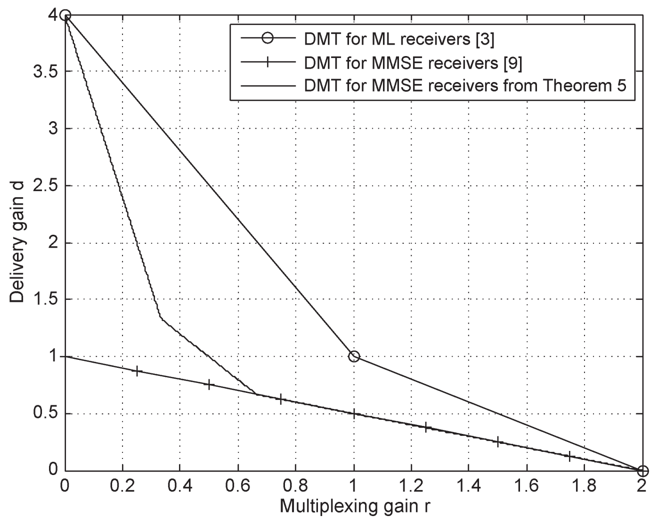

Fig. 3. DMTs for ML receivers [3], MMSE receivers [23], and temporal coded MMSE receivers for (2 × 2) MIMO in multiple channel uses.

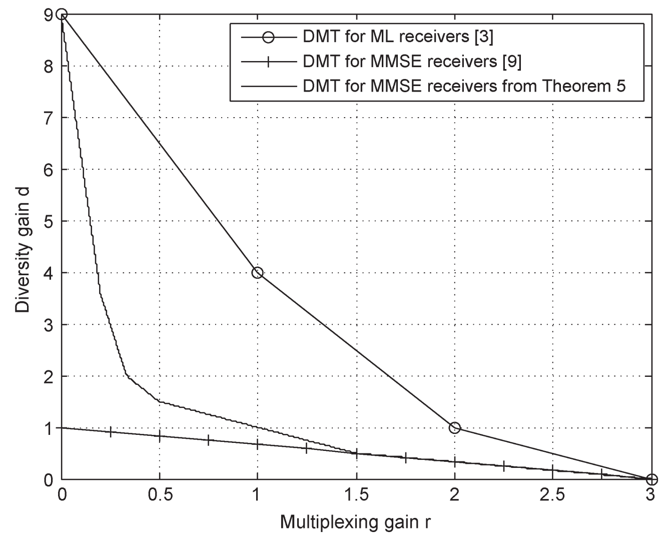

Fig. 4. DMTs for ML receivers [3], MMSE receivers [23], and temporal coded MMSE receivers from Theorem 5 for (3 × 3) MIMO in multiple channel uses.

given multiplexing gain r, the DMT resulting from the use of linear MMSE or ZF receiver is exactly

$$d^n_{\text{MMSE}}|_{U_x} = (nn_r - m + 1)\left(1 - \frac{nr}{m}\right)^+.$$

■

We remark that the condition of Theorem 5 is much weaker than that in (19) and is therefore giving more insight to the design of precoders. Furthermore, as the vectors $\underline{u}_{ij}$ are of length n_t, a necessary and sufficient condition for Theorem 5 to hold is $mn \le n_t$.

Proof: Although the entries of the equivalent channel matrix $H_s = \tilde{H}U_x$ are correlated, with the set of vectors $\{\underline{u}_{11}, \ldots, \underline{u}_{nm}\}$ being linearly independent over $\mathbb{C}$, the $(nn_r m \times nn_r m)$ covariance matrix of the entries of H_s must have full rank $nn_r m$. It then follows from the proof of [42, Th. 3] that the joint ordered eigenvalue distribution of $H_s^\dagger H_s$ is exponentially equal to that of $\hat{H}^\dagger \hat{H}$, where $\hat{H}$ is an $(nn_r \times m)$ random matrix with i.i.d. $\mathbb{C}\mathcal{N}(0,1)$ entries. Having seen the above, following similar arguments as in Section II-A, it can be shown that the outage probability for such precoding matrix U_x is upper and lower bounded by

$$\begin{aligned} &P_{\text{out,MMSE}}(r|\,\text{rank}\ m,\ n\ \text{channel uses}, U_x) \\ &\quad \le \Pr\left\{\text{Tr}\left[\left(\frac{\text{SNR}}{m}H_s^\dagger H_s + I_m\right)^{-1}\right] \ge \text{SNR}^{-\frac{nr}{m}}\right\} \\ &\quad \doteq \text{SNR}^{-(nn_r-m+1)(1-\frac{nr}{m})^+} \\ &P_{\text{out,MMSE}}(r|\,\text{rank}\ m,\ n\ \text{channel uses}, U_x) \\ &\quad \dot{\ge} \max_{1\le \ell \le m} \Pr\left\{-\log\left[\left(\text{SNR}\cdot H_s^\dagger H_s + I_m\right)^{-1}\right]_{\ell,\ell} \le \frac{nR}{m}\right\} \\ &\quad \doteq \text{SNR}^{-(nn_r-m+1)(1-\frac{nr}{m})^+}. \end{aligned}$$

Combining the upper and lower bounds proves Theorem 5. ■

We provide in Figs. 3 and 4 the DMT performances of (2 × 2) and (3 × 3) MIMO linear receivers based on results obtained in Theorem 5, respectively. It is seen that the DMT performance of the MMSE receiver can be dramatically improved by using degenerate Gaussian channel input. We further remark that in Fig. 6 our scheme has the same DMT performance as [23] when $1 \le r \le 3$. This is because that in such a region the DMT is dominated by the case of single channel use. Similar observations can also be made in Fig. 3.

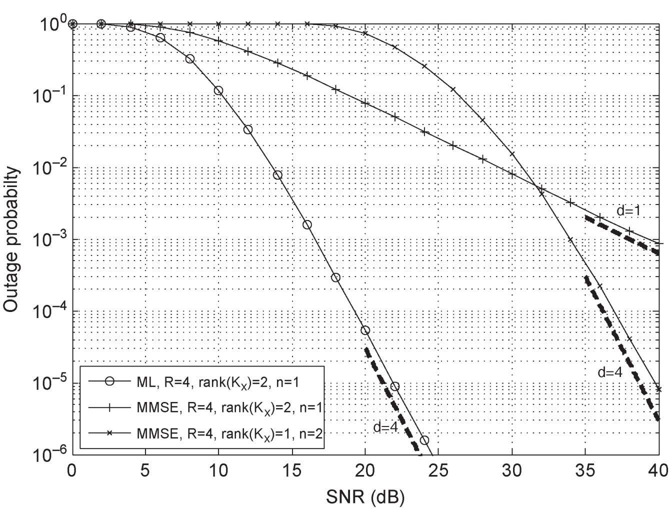

Fig. 5. Outage probabilities of ML receivers [3], MMSE receivers [23], and temporal coded MMSE receivers for (2 × 2) MIMO in multiple channel uses at a constant sum rate $R = 4$ bpcu. The dash lines represent theoretical high SNR slopes for diversity gains.

In Figs. 5 and 6, simulations of outage probabilities of various ranks of the covariance matrices and various multiple channel uses at a constant sum rate of $R = 4$ bpcu are provided. In particular, we see from Fig. 6 that the outage probabilities of ZF and MMSE receiver with i.i.d. complex Gaussian channel input are very close. The simulation results also justify our claims that using degenerate Gaussian random matrices as channel inputs could yield a much higher diversity gain. Furthermore, we observe that in all temporal coding schemes with

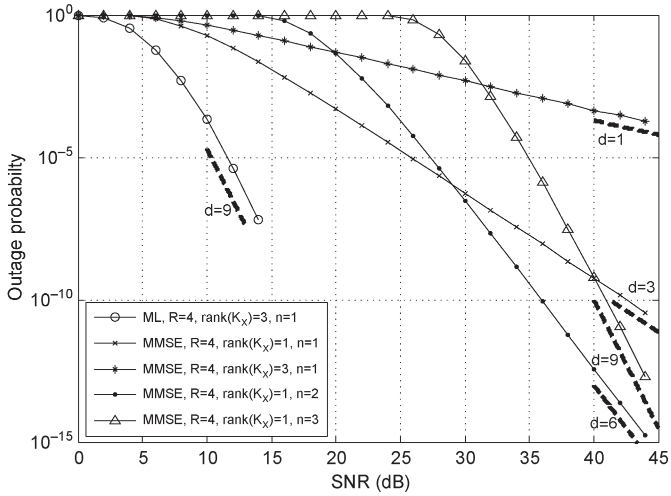

Fig. 6. Outage probabilities of ML receivers [3], MMSE receivers [23], and temporal coded MMSE receivers for (3 × 3) MIMO in multiple channel uses at a constant sum rate $R = 4$ bpcu. The dash lines represent theoretical high SNR slopes for diversity gains.

$\text{rank}(K_x) = 1$, the diversity gains increase with the number of channel uses.

We remark also that all the outage probabilities associated with rank-1 schemes follow directly from the χ^2 distribution and can therefore be explicitly evaluated. For example, the outage probability for $n=1$ and $\text{rank}(K_x)=1$, for the (3×3) MIMO, can be obtained in close form by

$$\begin{aligned} &Pr\left\{\log\left(1+\text{SNR}\|\underline{h}_1\|^2\right) \le R\right\} \\ &= \Pr\left\{\|\underline{h}_1\|^2 \le \frac{e^R-1}{\text{SNR}}\right\} \\ &= F_6\left(\frac{e^R-1}{\text{SNR}}\right) \end{aligned}$$

where $\underline{h}_1$ is a (3×1) random channel vector, and $F_6(\cdot)$ is the cumulative distribution function of the central χ^2 random variable with six° of freedom. We have also numerically simulated these outage probabilities, and they match perfectly with the above theoretical values, shown in dash lines in Fig. 6.

Finally, we point out that it is not necessarily true that schemes having larger diversity gain would have lower outage probabilities as well. It can be seen from the curves in Figs. 5 and 6 that such a statement is true only at the high SNR regime. It should be noted that the diversity gain is obtained by evaluating the outage/error performance at high SNR regime. Therefore, for low and moderate SNR values, it is often possible that schemes with inferior diversity gain could have smaller outage/error probability. System designers are well advised to consult the results in Figs. 5 and 6 and find what the best scheme is in the SNR region of their interest before designing codes for the use of MIMO linear receivers.

IV. Optimal Code Construction and Performance Simulations

In Section II, we have seen that in the case of single channel use, assuming that the covariance matrix of the channel input has rank m, the output of MIMO MMSE and similarly ZF equalizer can be regarded a set of m parallel SISO fading channels, each given by

$$\hat{y}_\ell = G_{\ell,\ell} s_\ell + z_\ell, \quad \ell = 1, 2, \ldots, m$$

where the equivalent channel coefficient $G_{\ell,\ell} = 1 - [(\text{SNR} \cdot \Lambda H_s^\dagger H_s + I_m)^{-1}]_{\ell,\ell}$. We then showed that the DMT performance achieved by the rank-m scheme is given by

$$d^{(1)}_{\text{MMSE}}|_{rank(K_x)=m} = (n_r - m + 1)\left(1 - \frac{r}{m}\right)^+ \quad (20)$$

and is independent of the choice of matrix K_x as well as Λ once $\text{rank}(K_x) = m$. As long as Gaussian random codebook is of concern, one can simply randomly choose s_ℓ following an i.i.d. $\mathbb{CN}(0, (1/m))$ distribution. Then, the resulting random code achieves the desired DMT performance given in (20). However, random Gaussian codebooks are hardly useful in practical systems. For designing a deterministic code, it turns out that one could replace the i.i.d. random variables s_ℓ by scaled quadrature amplitude modulation (QAM) constellation symbols, and the resulting deterministic code is still DMT optimal, in the sense of having a MIMO linear receiver. We have the following theorem for constructing DMT optimal codes for MIMO linear receivers.

Theorem 6: For the single-channel-use case, given the desired rank m and multiplexing gain r, the following code achieves the optimal DMT of a MIMO linear receiver:

$$\mathcal{X} := \left\{ \underline{x} = \kappa \cdot U_x \begin{bmatrix} \underline{s} \\ \underline{0}_{n_t - m} \end{bmatrix} : \begin{matrix} s_\ell \in \mathbb{Z}[\imath], \\ |s_\ell|^2 \dot{\le} \text{SNR}^{\frac{r}{m}} \\ \ell = 1, \ldots, m \end{matrix} \right\}$$

for any U_x taken from $\mathcal{U}(n_t)$, the unitary group of n_t. κ is some constant such that $\mathbb{E}_\mathcal{X}\|\underline{x}\|^2 = 1$ and $\kappa \doteq \text{SNR}^{-(r/2m)}$ for large SNR. $\mathbb{Z}[\imath]$ with $\imath = \sqrt{-1}$ denotes the usual ring of Gaussian integers, and $\mathbb{Z}[\imath] = \{a + b\imath | a, b \in \mathbb{Z}\}$.

Proof: First, it is straightforward to see that $|\mathcal{X}| \doteq \text{SNR}^r$ and $\mathbb{E}_\mathcal{X}\|\underline{x}\|^2 \dot{\le} 1$ from construction. Hence, $\mathcal{X}$ has the desired rate and meets the required power constraint. To show the DMT optimality, it suffices to note that each entry s_ℓ in the foregoing code is a QAM symbol and is therefore an approximately universal code [4], [43] for a SISO channel with any type of channel statistics. In other words, simply by the property of approximately universal, the QAM symbol s_ℓ forms a DMT optimal code for the ℓth SISO subchannel with channel coefficient $G_{\ell,\ell}$, which is a random variable of certain distribution. ■

A similar result holds for the case of multiple channel uses as well.

Corollary 7: Given the desired rank m, multiplexing gain r, and number of channel uses n, let the precoding matrix

$$U_x = \begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix}$$

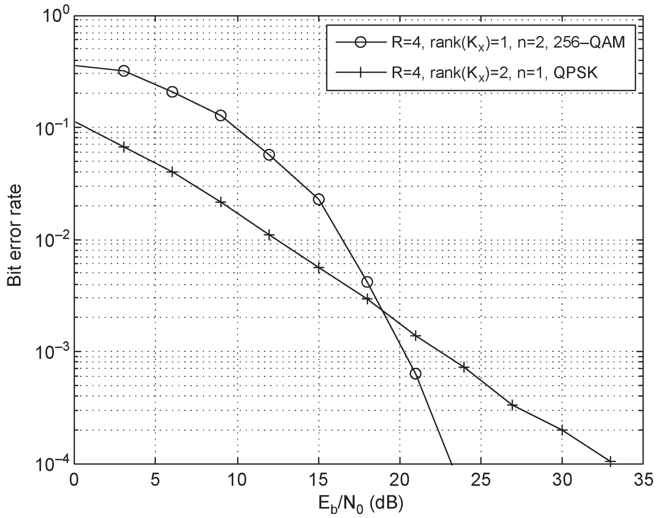

Fig. 7. Bit error rate for codes constructed from Theorem 6 and Corollary 7 for the (2 × 2) MIMO system at $R = 4$ bpcu.

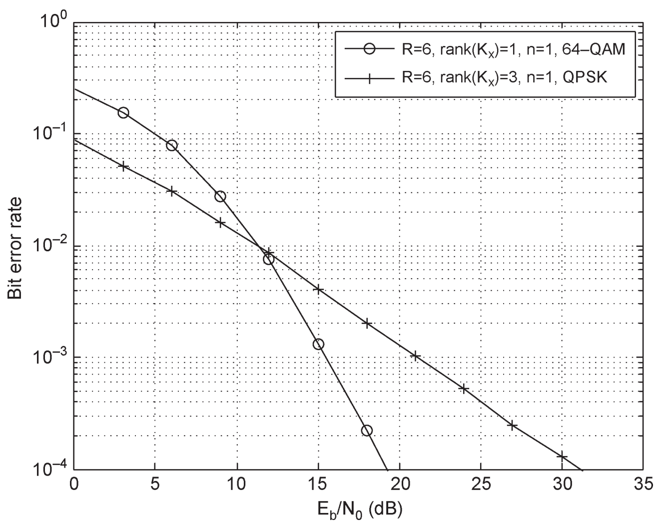

Fig. 8. Bit error rate for codes constructed from Theorem 6 for the (3 × 3) MIMO system at $R = 6$ bpcu in single channel use.

be any $(nn_t \times m)$ matrix satisfying the linearly independent criterion given in Theorem 5, where each submatrix U_i is of size $(n_t \times m)$. Then, the following code

$$\mathcal{X} = \left\{ X = [\underline{x}_1, \ldots, \underline{x}_n] : \begin{array}{c} \underline{x}_i = \kappa \cdot U_i \underline{s}, \\ s_\ell \in \mathbb{Z}[\imath], \\ |s_\ell|^2 \dot{\leq} \mathrm{SNR}^{\frac{nr}{m}} \\ \ell = 1, \ldots, m \end{array} \right\}$$

achieves DMT of $(nn_r - m + 1)(1 - (nr/m))^+$. κ is some constant such that $\mathbb{E}_{\mathcal{X}} \|X\|^2 = n$ and $\kappa \doteq \mathrm{SNR}^{-(nr/2m)}$ for large SNR. ■

Based on the optimal construction of codes given in Theorem 6 and Corollary 7, we provide performance simulations of the proposed codes in Figs. 7 and 8 for the (2 × 2) and (3 × 3) MIMO systems, respectively. The error performances are plotted in bit error rate. The diversity gain achieved by each scheme coincides perfectly with the theoretical result. We remark that in Fig. 7 (and similarly in Fig. 8), to achieve a constant sum rate of $R = 4$ bpcu, the two simulated coding schemes take on a different base alphabet. For example, in the case of $\mathrm{rank}(K_x) = 2$ and $n = 1$, there are two subchannels, and the transmission is done in one channel use. It means that each subchannel transmits at 2 bpcu; hence, a quadrature phase-shift keying (QPSK) constellation is sufficient from Theorem 6. For the other case of $\mathrm{rank}(K_x) = 1$ and $n = 2$, there is only one subchannel but the transmission requires two channel uses to complete. Therefore, this subchannel must transmit at $2 \times R = 8$ bpsu, resulting in a 256-QAM constellation from Corollary 7.

Furthermore, the results in Figs. 7 and 8 echo exactly what we have seen in the outage probability simulations, that is, schemes having larger diversity gain do not necessarily have lower bit error rate. For example, in Fig. 7, we see that to reach 4 bpcu in (2 × 2) MIMO system using linear receivers, if the E_b/N_0 value is below 19.16 dB, a better scheme would be to send two independent QPSK symbols in two transmit antennas in each channel use. On the other hand, if the E_b/N_0 value is above 19.16 dB, the scheme of sending one 256-QAM symbol, but rotated in four different ways for two transmit antennas in two channel uses, would yield a much lower bit error rate. Finally, one of the most important observations from the simulation results is the following. Compared to the common use of MIMO linear receivers, where system designers are used to make use of full multiplexing gains, i.e. set $\mathrm{rank}(K_x) = nn_t$, the simulation results show the newly proposed degenerate schemes can provide a significant performance gain. For example, in Fig. 7, the rank-1 scheme beats the conventional full multiplexing scheme by approximately 10 dB in SNR at bit error rate 10^{-4} for the (2 × 2) MIMO linear receiver. A similar result, but with a much larger gain of 12.08 dB in bit error rate, is reported in Fig. 8 for the (3 × 3) MIMO linear receiver in single channel use again at bit error rate equal to 10^{-4}.

V. Conclusion and Future Work

In this paper, we proposed using colored and possibly degenerate Gaussian random vectors as channel input to the MIMO linear receiver. The DMT resulting from such input is much larger than the currently known one which was based on an assumption of i.i.d. Gaussian input. By taking into account the possibility of having temporal coding at the transmitter, we show that the DMT of the MIMO linear receiver can be further improved so that in the extreme case the simple MIMO linear receiver achieves the same maximal diversity gain of $n_t n_r$ as the ML receiver but at a significantly lower complexity. We also observed that in terms of outage probability, it is not necessarily true that a rank-m scheme with larger diversity gain would have lower outage probability. System designers should first determine the value m of the rank-m scheme that yields the best outage performance in the SNR region of interest. Given any m, optimal constructions of codes for MIMO linear receivers are also provided. For the case of (2 × 2) and (3 × 3) MIMO linear receivers, we reported by simulation that the newly proposed rank-1 codes provide significant gains of 10 and 12.08 dB in E_b/N_0 at bit error rate 10^{-4} compared to the conventional schemes, respectively.

In our main result, Theorem 5, we have seen that the design of U_x for different combinations of n and m affects the range of DMT. In this paper, we only managed to prove the DMT performance for our specific U_x, thereby obtaining a lower bound, yet significantly larger than the current best, on the optimal DMT for linear receivers. In general, it is possible that our bound is not tight, and there might be better designs of the precoding matrices U_x that can lead to larger DMT. At the moment, we do not know what the optimal design is. It will require more substantial research in this direction in the future. Second, it should be noted that Theorem 5 only states for a fixed combination of n and m, the precoding matrices U_x satisfying the prescribed condition given therein would have the same diversity gain performance. It does not necessarily mean that all such U_x would yield the same error performance, as each of them can provide a different coding gain. We also remark that this resulting performance is still rather far away from the performance of ML detection. Therefore, the problem of how to design precoding matrices that have better coding gains remains open. Furthermore, in view of the decision feedback type of detectors that has the favorable property of somewhat closing the performance gap between linear and ML detectors with limited complexity, the proposed degenerate Gaussian input may also be beneficial in this case. However, it still requires further efforts to investigate the resulting DMT performance.

APPENDIX

To determine the SINR of the ℓth subchannel in (11), we recall that

$$\underline{\hat{y}} := W_{\mathrm{MMSE}}\left(\sqrt{\mathrm{SNR}}\cdot H_s\underline{s}+\underline{w}\right) = G\underline{s}+\underline{n}$$

and that

$$\begin{aligned} G &:= \mathrm{SNR}\cdot\left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s + I_m\right)^{-1}\Lambda H_s^\dagger H_s \\ &= I_m - \left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1} = I_m - P \end{aligned}$$

from (7) and (9), where

$$\begin{aligned} W_{\mathrm{MMSE}} &:= \sqrt{\mathrm{SNR}}\left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1}\Lambda H_s^\dagger \\ P &:= \left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1}. \end{aligned}$$

The covariance matrix of $\underline{\hat{y}}$ is given by

$$\begin{aligned} K_{\underline{\hat{y}}} &= \mathbb{E}\underline{\hat{y}}\underline{\hat{y}}^\dagger = G\Lambda G^\dagger + W_{\mathrm{MMSE}}W_{\mathrm{MMSE}}^\dagger \\ W_{\mathrm{MMSE}}W_{\mathrm{MMSE}}^\dagger &= \mathrm{SNR}\cdot\left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1}\Lambda H_s^\dagger H_s \\ &\quad\times\Lambda\left[\left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1}\right]^\dagger \\ &= G\Lambda\left[\left(\mathrm{SNR}\cdot\Lambda H_s^\dagger H_s+I_m\right)^{-1}\right]^\dagger \\ &= G\Lambda\left(I_m - G^\dagger\right) = G\Lambda - G\Lambda G^\dagger. \end{aligned}$$

Combining the above shows $K_{\underline{\hat{y}}} = G\Lambda$. Furthermore, we remark that as $K_{\underline{\hat{y}}}$ is Hermitian symmetric and Λ is a diagonal positive definite matrix, the former equality implies that $0 \le G_{\ell,\ell} \in \mathbb{R}$ for all $\ell = 1,2,\ldots,m$.

Next, we let $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_m)$, where $\lambda_\ell = \mathbb{E}|s_\ell|^2$. Then, the SINR of the ℓth subchannel is given by

$$\begin{aligned} \mathrm{SINR}_\ell &= \frac{G_{\ell,\ell}^2\mathbb{E}|s_\ell|^2}{\mathbb{E}|\hat{y}_\ell|^2 - G_{\ell,\ell}^2\mathbb{E}|s_\ell|^2} \\ &= \frac{G_{\ell,\ell}^2\lambda_\ell}{G_{\ell,\ell}\lambda_\ell - G_{\ell,\ell}^2\lambda_\ell} = \frac{1}{1-G_{\ell,\ell}} - 1 = \frac{1}{P_{\ell,\ell}} - 1. \end{aligned}$$

This proves the second equality in (11).

Furthermore, we remark that (11) shows $0 < P_{\ell,\ell} \le 1$ for all ℓ due to the nonnegativity of mutual information. The same conclusion can be obtained from an algebraic reasoning. Below we provide an alternative proof.

Theorem 8: Let $P = (\mathrm{SNR}\,\Lambda H_s^\dagger H_s + I_m)^{-1}$, where H_s is of size $(n_r \times m)$ with $m \le n_r$ and has rank m; then, $P_{\ell,\ell} \in (0,1]$ for all ℓ.

Proof: Recall that $G = \sqrt{\mathrm{SNR}}W_{\mathrm{MMSE}}H_s = I_m - P$. Hence, it suffices to show $G_{\ell,\ell} \in [0,1)$. We also recall from (7) that W_{MMSE} can be represented alternatively as $\sqrt{\mathrm{SNR}}\Lambda H_s^\dagger(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r})^{-1}$. Hence, we can rewrite G as

$$\begin{aligned} G &= \sqrt{\mathrm{SNR}}W_{\mathrm{MMSE}}H_s \\ &= \mathrm{SNR}\Lambda H_s^\dagger\left(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r}\right)^{-1}H_s. \end{aligned}$$

Since $H_s^\dagger(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r})^{-1}H_s$ is positive definite, we must have $G_{\ell,\ell} \ge 0$, and the boundary point 0 is achieved if and only if $\mathrm{SNR} = 0$. It suffices to consider the case of $\mathrm{SNR} > 0$, and we only need to show for every $\ell = 1,\ldots,m$

$$\left[H_s^\dagger\left(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r}\right)^{-1}H_s\right]_{\ell,\ell} < \frac{1}{\mathrm{SNR}\Lambda_{\ell,\ell}}.$$

To establish the above, let $H_s = U\Sigma V^\dagger$ be the singular-value decomposition of H_s. Then, we can rewrite the matrix at the left-hand side as

$$\begin{aligned} &H_s^\dagger\left(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r}\right)^{-1}H_s \\ &\quad= V\Sigma^\top U^\dagger\left(\mathrm{SNR}\cdot U\Sigma V^\dagger\Lambda V\Sigma^\top U^\dagger + I_{n_r}\right)^{-1}U\Sigma V^\dagger \\ &\quad= V\Sigma^\top\left(\mathrm{SNR}\cdot\Sigma V^\dagger\Lambda V\Sigma^\top + I_{n_r}\right)^{-1}\Sigma V^\dagger. \end{aligned} \tag{21}$$

Note that since $m \le n_r$, the matrix Σ must take the following form:

$$\Sigma = \begin{bmatrix} \tilde{\Sigma}_{m\times m} \\ \mathbf{0}_{(n_r-m)\times m} \end{bmatrix}$$

where $\tilde{\Sigma}_{m\times m} \succ \mathbf{0}$ is an invertible diagonal matrix. Then, after a few algebraic manipulations, we can simplify (21) to

$$\begin{aligned} & V\Sigma^\top \left(\mathrm{SNR}\cdot\Sigma V^\dagger \Lambda V \Sigma^\top + I_{n_r}\right)^{-1}\Sigma V^\dagger \\ &= V\tilde{\Sigma}\left(\mathrm{SNR}\cdot\tilde{\Sigma}V^\dagger\Lambda V\tilde{\Sigma}^\top + I_m\right)^{-1}\tilde{\Sigma}V^\dagger \\ &\overset{(a)}{=} \left(\mathrm{SNR}\cdot\Lambda + \left(H_s^\dagger H_s\right)^{-1}\right)^{-1} \\ &\overset{(b)}{\prec} (\mathrm{SNR}\cdot\Lambda)^{-1} \end{aligned}$$

where (a) is because V is unitary and $\tilde{\Sigma}$ is invertible, and (b) follows from

$$\left(\mathrm{SNR}\cdot\Lambda + \left(H_s^\dagger H_s\right)^{-1}\right) \succ (\mathrm{SNR}\cdot\Lambda).$$

In summary, we have shown that

$$\left[H_s^\dagger\left(\mathrm{SNR}\cdot H_s\Lambda H_s^\dagger + I_{n_r}\right)^{-1}H_s\right] \prec (\mathrm{SNR}\cdot\Lambda)^{-1}.$$

This completes the proof. ■

REFERENCES

[1] G. J. Foschini, "Layered space–time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Lab. Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

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

[3] 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.

[4] 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.

[5] 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.

[6] M. O. Damen, A. Chkeif, and J. C. Belfiore, "Sphere decoding of space–time codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Aug. 2000, p. 362.

[7] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, 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] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[10] J. W. Choi, B. Shim, A. C. Singer, and N. I. Cho, "Low-complexity decoding via reduced dimension maximum-likelihood search," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1780–1793, Mar. 2010.

[11] D. Coppersmith and S. Winograd, "Matrix multiplication via arithmetic progressions," *J. Symbol. Comput.*, vol. 9, no. 3, pp. 251–280, Mar. 1990.

[12] L. Hong and A. G. Armada, "Bit error rate performance of MIMO MMSE receivers in correlated Rayleigh flat-fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 313–317, Jan. 2011.

[13] M. Kiessling and J. Speidel, "Analytical performance of MIMO MMSE receivers in correlated Rayleigh fading environments," in *Proc. IEEE 58th Veh. Technol. Conf.*, 2003, vol. 3, pp. 1738–1742.

[14] P. L. Kafle, A. B. Sesay, and J. McRory, "Capacity of MIMO OFDM systems in spatially correlated indoor fading channels," in *Proc. IEEE 60th Veh. Technol. Conf.*, 2004, vol. 1, pp. 129–133.

[15] A. L. Moustakas, K. R. Kumar, and G. Caire, "Performance of MMSE MIMO receivers: A large n analysis for correlated channels," in *Proc. IEEE 69th Veh. Technol. Conf.*, Apr. 2009, pp. 1–5.

[16] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 309–318, Feb. 2008.

[17] D. Wubben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "MMSE extension of V-BLAST based on sorted QR decomposition," in *Proc. IEEE 58th Veh. Technol. Conf.*, 2003, vol. 1, pp. 508–512.

[18] D. Wubben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction," in *Proc. IEEE Int. Conf. Commun.*, 2004, vol. 2, pp. 798–802.

[19] E. K. Onggosanusi, A. G. Dabak, T. Schmidl, and T. Muharemovic, "Capacity analysis of frequency-selective MIMO channels with sub-optimal detectors," in *Proc. IEEE ICASSP*, 2002, pp. III-2369–III-2372.

[20] H. A. A. Saleh and W. Hamouda, "BER performance of MIMO-SM with zero-forcing in spatially correlated Ricean fading," in *Proc. IEEE 69th Veh. Technol. Conf.*, Apr. 2009, pp. 1–5.

[21] J. Akhtar, "Diversity and spatial-multiplexing tradeoff under linear decoding," in *Proc. IEEE 59th Veh. Technol. Conf.*, 2004, vol. 2, pp. 589–593.

[22] K. K. Wong and A. Paulraj, "On the decoding order of MIMO maximum-likelihood sphere decoder: Linear and non-linear receivers," in *Proc. IEEE 59th Veh. Technol. Conf.*, 2004, vol. 2, pp. 698–702.

[23] K. Kumar, G. Caire, and A. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.

[24] C. Windpassinger and R. F. H. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proc. ITW*, Paris, France, Mar. 2003, pp. 345–348.

[25] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE GLOBECOM*, Nov. 2002, pp. 424–428.

[26] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4765–4780, Oct. 2010.

[27] A. Sezgin, E. A. Jorswieck, and E. Costa, "Lattice-reduction aided detection: Spatial multiplexing versus quasi-orthogonal STBC," in *Proc. IEEE 60th Veh. Technol. Conf.*, 2004, vol. 4, pp. 2389–2393.

[28] Y. Yang and J. Kim, "Fixed-complexity LLL-based signal detection for MIMO systems," *IEEE Trans. Veh. Technol.*, Early Access Articles.

[29] I. Berenguer and X. Wang, "MIMO antenna selection with lattice-reduction-aided linear receivers," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1289–1302, Sep. 2004.

[30] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 3, pp. 630–639, Apr. 1992.

[31] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 421–434, Feb.–Apr. 1995.

[32] G. Raleigh and J. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[33] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[34] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1334–1347, Sep. 1994.

[35] A. D. Murugan, H. El Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: Rediscovering the sequential decoder," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 933–953, Mar. 2006.

[36] A. Lenstra, H. Lenstra, and Lováz, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, no. 4, pp. 515–534, Dec. 1982.

[37] M. Taherzadeh, A. Mobasher, and A. Khandani, "LLL lattice-basis reduction achieves the maximum diversity in MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory*, 2005, pp. 1300–1304.

[38] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1988.

[39] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[40] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*. Boston, MA: Jones & Bartlett, 1992.

[41] D. Deng, M. Zhao, and J. Zhu, "Transmit antenna selection for linear dispersion codes based on linear receiver," in *Proc. IEEE 63rd Veh. Technol. Conf.*, 2006, vol. 6, pp. 2927–2931.

[42] K. R. Kumar and G. Caire, "Outage analysis and code design for correlated MIMO fading channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 24–29, 2007, pp. 2276–2280.

[43] 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.

Ti-Wen Tang (S'10) was born in Tainan, Taiwan. He received the B.S. degree in optoelectronics and communication engineering from the National Kaohsiung Normal University, Kaohsiung, Taiwan, in 2007 and the M.S. degree in communication engineering from the National Chung Cheng University, Chiayi, Taiwan, in 2009. He is currently working toward the Ph.D. degree in communications engineering under the supervision of Prof. H.-F. Lu with the National Chiao Tung University, Hsinchu, Taiwan.

His research interests include wireless communication systems, multiuser detection, and error correcting codes.

Min-Kun Chen received the B.S. degree in communication engineering from the National Chung Cheng University, Chayi, Taiwan, in 2009 and the M.S. degree from the National Chiao Tung University, Hsinchu, Taiwan, in 2011.

He is currently with the Taiwan Semiconductor Manufacturing Company, Tainan, Taiwan. His research is in the area of wireless communication, multiple-input–multiple-output systems, diversity, and multiplexing gain tradeoff.

Hsiao-Feng (Francis) Lu (S'98–M'04–SM'12) received the B.S. degree in electrical engineering from Tatung University, Taipei, Taiwan, in 1994 and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1999 and 2003, respectively.

He was a Postdoctoral Research Fellow with the University of Waterloo, Waterloo, ON, Canada, during 2003–2004. In February 2004, he joined 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, where he is currently a Full Professor. His research is in the area of space–time codes, multiple-input–multiple-output 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 has received several research awards, including the 2006 IEEE Information Society Taipei Chapter and the IEEE Communications Society Taipei/Tainan Chapter Best Paper Award for Young Scholars, the 2007 Wu Da You Memorial Award from the 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.