

# Diversity of MMSE MIMO Receivers

Ahmed Hesham Mehana, *Student Member, IEEE*, and Aria Nosratinia, *Fellow, IEEE*

**Abstract**—In most multiple-input multiple-output (MIMO) systems, the family of waterfall error curves, calculated at different spectral efficiencies, are asymptotically parallel at high signal-to-noise ratio. In other words, most MIMO systems exhibit a single diversity value for all fixed rates. The MIMO minimum mean square error (MMSE) receiver does not follow this pattern and exhibits a varying diversity in its family of error curves. This paper analyzes this interesting behavior of the MMSE MIMO receiver and produces the MMSE MIMO diversity at all rates. The diversity of the quasi-static flat-fading MIMO channel consisting of any arbitrary number of transmit and receive antennas is fully characterized, showing that full spatial diversity is possible if and only if the rate is within a certain bound which is a function of the number of antennas. For other rates, the available diversity is fully characterized. At sufficiently low rates, the MMSE receiver has a diversity similar to the maximum likelihood receiver (maximal diversity), while at high rates, it performs similarly to the zero-forcing receiver (minimal diversity). Linear receivers are also studied in the context of the MIMO multiple-access channel. Then, the quasi-static frequency selective MIMO channel is analyzed under zero-padding and cyclic-prefix (CP) block transmissions and MMSE reception, and lower and upper bounds on diversity are derived. For the special case of SIMO under CP, it is shown that the aforementioned bounds are tight.

**Index Terms**—Diversity, linear receiver, minimum mean square error (MMSE), multiple-input multiple-output (MIMO).

## I. INTRODUCTION

LINEAR receivers are widely used for their low complexity compared to maximum likelihood (ML) receivers. In the context of multiple-input multiple-output (MIMO) systems, linear receivers such as the minimum mean square error (MMSE) receiver are adopted in some of the emerging standards, e.g., IEEE 802.11n and 802.16e. Therefore, the analysis of MMSE receivers is strongly motivated by both theoretical and practical considerations.

A significant amount of research has focused on linear receivers; however, their performance is not fully understood in the MIMO channel. For instance, the distribution of the output signal-to-interference-plus-noise ratio (SINR) of the linear MIMO receiver is still unknown except in asymptotic regimes [large number of antennas, and high/low signal-to-noise ratio (SNR)] [1]–[4]. The outage and diversity of MMSE receiver

Manuscript received January 30, 2011; accepted February 15, 2012. Date of publication August 03, 2012; date of current version October 16, 2012. This work was supported in part by the Texas Higher Education Coordinating Board under Grant 009741-0084-2007. This paper was presented in part at the 2010 and 2011 IEEE International Symposia on Information Theory.

The authors are with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083-0688 USA (e-mail: ahmed.mehana@student.utdallas.edu; aria@utdallas.edu).

Communicated by L. Zheng, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2012.2204954

have also been a subject of interest. It has been observed [5]–[7] that while the MMSE receiver can extract the full spatial diversity of the MIMO quasi-static channel at low rates, it does not enjoy this feature at high rates.

Fig. 1 shows the outage probabilities (for various spectral efficiencies  $R$  b/s/Hz) of MMSE and ML receivers respectively. Clearly, one of the main differences between the two characteristics is the slope of the error curves, i.e., the *diversity*. Fig. 1 shows that in a  $2 \times 2$  MIMO system, the ML receiver achieves diversity 4 at all rates. However, the MMSE receiver diversity varies with the operating spectral efficiency. From a system design perspective, obtaining the MMSE diversity is important in order to understand the broad tradeoffs involved in the determination of the operating point of the system and predicting its performance.

In this paper, we seek answers for the following questions: when can the MMSE receiver exploit the full diversity in MIMO channel? More generally, how does the diversity of the MMSE receiver vary with the system parameters such as spectral efficiency  $R$ , the number of antennas, and in case of intersymbol interference channel (ISI), the channel memory?

The well-known and powerful framework of diversity-multiplexing tradeoff (DMT) is not sufficient to answer the above questions, because the DMT framework cannot distinguish between different spectral efficiencies that correspond to the same multiplexing gain. In the MIMO MMSE receiver, rates that correspond to the same multiplexing gain can produce different diversities.

We approach the problem of MMSE reception in MIMO flat-fading channels through a rate-dependent approximation of the outage probability and then proceed with bounding the pairwise error probability (PEP) from both sides using the outage. This leads to a closed-form expression for the diversity-rate tradeoff which reveals the relationship between diversity, spectral efficiency, and number of transmit and receive antennas. The approximation of outage and PEP as functions of rate requires more delicate handling compared with the DMT analysis, as certain ratios and terms that simply vanish in the DMT analysis are in our case relevant and must be carefully handled.

We then analyze the *frequency-selective*, quasi-static MIMO channel. Specifically, we consider single-carrier (SC) MMSE equalization under zero-padding (ZP) and cyclic-prefix (CP) transmission. SC-MMSE provides an attractive alternative to orthogonal frequency-division multiplexing (OFDM) due to its low complexity and natural avoidance of the peak-to-average power ratio problem. The use of CP and ZP has been investigated in the literature, but the explicit tradeoff between the spectral efficiency and diversity of MIMO SC-MMSE under these two schemes has been unknown and is the subject of our work. We show that the diversity is a function of number of antennas, channel memory, and spectral efficiency, and obtain the explicit tradeoff in the special case of SIMO under CP transmission.

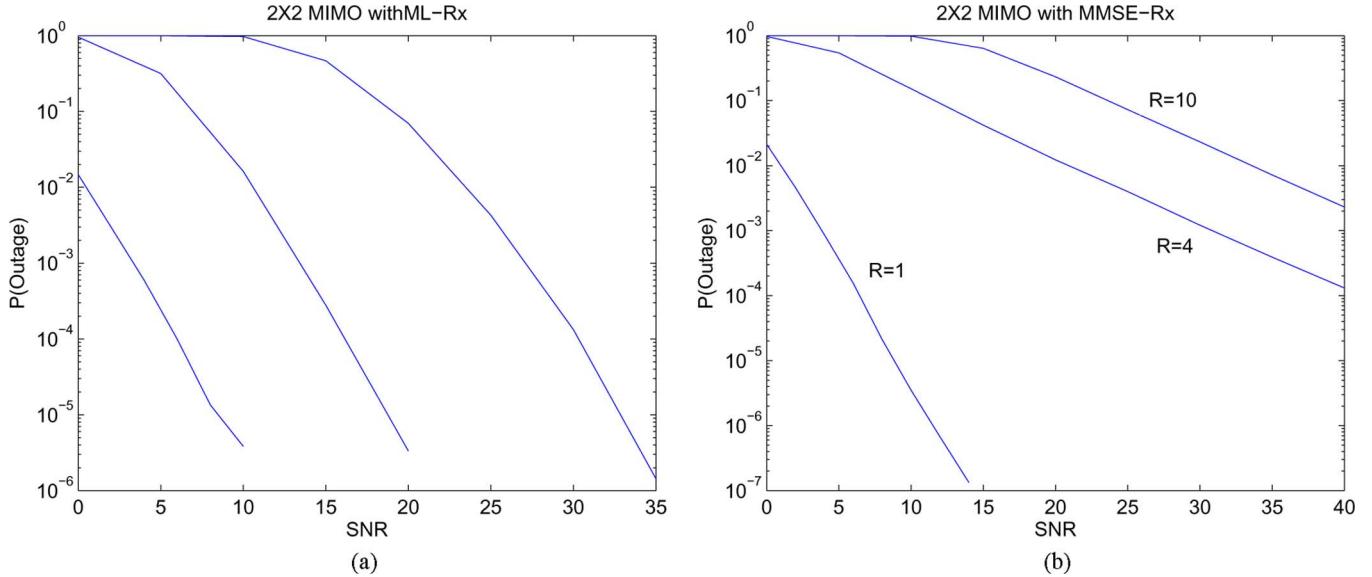


Fig. 1. Outage probability of ML receiver (left) and MMSE (right) with  $M = N = 2$  antennas and for rates  $R = 1, 4$ , and  $10$  b/s/Hz.

The results of this paper fully characterize the MIMO MMSE diversity in the fixed rate flat quasi-static regime. We analyze both the cases  $N \geq M$  and  $N < M$ , showing that in either case it is possible for the system to be limited to a diversity strictly less than  $MN$ . More specifically, the central result of the paper is as follows: with  $M$  transmit and  $N$  receive antennas (for any  $N$  and  $M$ ), the diversity is  $d = \lceil (M2^{-\frac{R}{M}} - (M-N)^+)^2 + |N-M| \lceil (M2^{-\frac{R}{M}} - (M-N)^+)^+ \rceil \rceil$ , where  $(\cdot)^+$  =  $\max(0, \cdot)$  and  $\lceil \cdot \rceil$  denotes rounding up to the next higher integer. Our results confirm and refine the earlier approximate results on the diversity of MMSE MIMO receivers that were obtained for very high and very low rates [5]–[7]. The MIMO multiple-access channel (MAC) channel is also studied.

Some of the related literature is as follows. The performance of MMSE receiver in terms of reliability goes back to [8] where outage analysis was performed for MMSE SIMO diversity combiner in a Rayleigh fading channel with multiple interferers. In the context of point-to-point MIMO systems, Gore *et al.* [9] compared the performance of MMSE D-BLAST with the ordered successive cancellation V-BLAST. They show that the former has better throughput at low and moderate SNR. Ongosanusi *et al.* [5] studied MMSE and zero-forcing (ZF) MIMO receivers and noticed their distinct outage performance at high SNR, specifically for large number of transmit antennas and low spectral efficiencies  $R$ , but provided no analysis.

Hedayat and Nosratinia [6] considered the outage probability as a function of fixed rates  $R$  under joint and separate spatial encoding, but for MMSE they obtained results only in the extremes of very high and very low rates. Kumar *et al.* [7] provided a DMT analysis for the system of [6] and observed that the DMT analysis does not predict the diversity of MMSE receivers at lower rates. We note that all existing analyses are limited to the case where the number of receive antennas ( $N$ ) is greater than or equal to the number of the transmit antennas ( $M$ ).

This paper is organized as follows. Section II describes the system model. Section III finds the exponential order of outage. Section IV bounds the codeword error probabilities using the

outage values and derives the final result. Section V extends the result to the MAC channel. Section VI calculates the diversity of MIMO MMSE reception in frequency-selective block-transmission systems. Section VII provides simulations that illuminate our results.

## II. LINEAR RECEIVERS

The input–output system model for flat-fading MIMO channel with  $M$  transmit and  $N$  receive antennas is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{H} \in \mathbb{C}^{N \times M}$  is the channel matrix whose entries are independent and identically distributed (i.i.d.) complex Gaussian,  $\mathbf{x} \in \mathbb{C}^{M \times 1}$  is the transmitted vector,  $\mathbf{n} \in \mathbb{C}^{N \times 1}$  is the Gaussian noise vector. The vectors  $\mathbf{x}$  and  $\mathbf{n}$  are assumed independent. We assume a quasi-static flat-fading channel and perfect channel state information (CSI) at the receiver (CSIR) and no CSI at the transmitter (CSIT), therefore transmit antennas operate with equal power.

We aim to characterize the diversity gain,  $d(R, M, N)$ , as a function of the spectral efficiency  $R$  (b/s/Hz) and the number of transmit and receive antennas. This requires a PEP analysis which is not directly tractable. Instead, we find the exponential order of outage probability and then demonstrate that outage and PEP exhibit identical exponential orders.

Following the notation of [10], we define the outage-type quantities

$$P_{\text{out}}(R, N, M) \triangleq \mathbb{P}(I(\mathbf{x}; \mathbf{y}) < R) \quad (2)$$

$$d_{\text{out}}(R, N, M) \triangleq -\lim_{\rho \rightarrow \infty} \frac{\log P_{\text{out}}(R, M, N)}{\log \rho} \quad (3)$$

where  $\rho$  is the per-stream SNR.

We say that the two functions  $f(\rho)$  and  $g(\rho)$  are *exponentially equal*, denoted by  $f(p) \doteq g(p)$  when

$$\lim_{\rho \rightarrow \infty} \frac{\log f(\rho)}{\log(\rho)} = \lim_{\rho \rightarrow \infty} \frac{\log g(\rho)}{\log(\rho)}$$

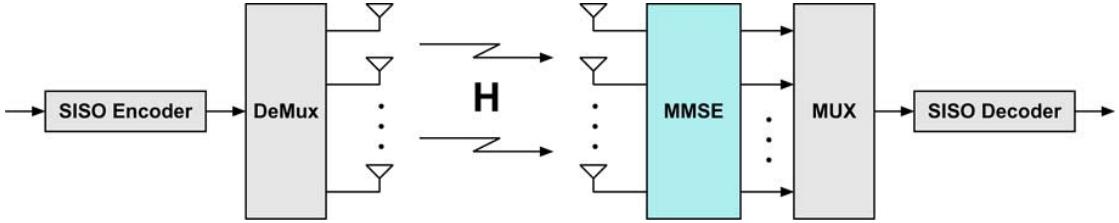


Fig. 2. MIMO system with linear MMSE receiver.

The ordering operators  $\lessdot$  and  $\gtrdot$  are also defined accordingly. If  $f(\rho) \doteq \rho^d$ , we say that  $d$  is the *exponential order* of  $f(p)$ .

#### A. MMSE Equalizer

The equalizer, denoted by  $\mathbf{W}$ , decouples the  $M$  transmitted data streams at the receiver (see Fig. 2). The MMSE equalizer is obtained by minimizing the mean square error (MSE) defined as  $\mathbb{E}[\|\mathbf{x} - \mathbf{W}^H \mathbf{y}\|^2]$ . It is usually assumed [6], [7] that the number of transmit antennas  $M$  is no more than that of receive antennas  $N$ . In the following, we start with  $N \geq M$  but later generalize it to  $N < M$  as well.

For  $N \geq M$ , using the orthogonality principle, the MMSE equalizer is given by [5], [11]

$$\begin{aligned} \mathbf{W} &= \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \rho^{-1} \mathbf{I})^{-1} \\ &= (\mathbf{H}^H \mathbf{H} + \rho^{-1} \mathbf{I})^{-1} \mathbf{H}^H. \end{aligned} \quad (4)$$

The corresponding SINR of the output stream  $k$  of the MMSE detector is

$$\gamma_k = \frac{1}{(\mathbf{I} + \rho \mathbf{H}^H \mathbf{H})_{kk}^{-1}} - 1, \quad 1 \leq k \leq M \quad (5)$$

where  $(\cdot)^H$  denotes matrix Hermitian,  $(\cdot)_{kk}^{-1}$  denotes the diagonal element  $k$  of the matrix inverse.

For the case  $N < M$ , it can be shown using a technique<sup>1</sup> very similar to [8, Appendix A] that the SINR expression (5) is again valid.

The square matrix  $\mathcal{W} = \mathbf{H}^H \mathbf{H}$  is random, nonnegative definite, and obeys the *Wishart Distribution* [12], [13]. In this study, the joint distribution of the eigenvalues of this equivalent channel matrix opens the door to the development of our analysis, as is also the case in many other MIMO results.

The equalizer output is

$$\mathbf{y} = \mathbf{W} \mathbf{H} \mathbf{x} + \mathbf{W} \mathbf{n}. \quad (6)$$

The signal streams of the transmit antennas may be either separately or jointly encoded. Separate encoding is simpler and has been fully analyzed [6], but we mention the central result for completeness.

**Theorem 1** [6], [7]: In a MIMO system consisting of  $M$  transmit and  $N$  receive antennas ( $N \geq M$ ), under separate spatial encoding, the MMSE receiver achieves the diversity

$$d_{\text{out}}(R, N, M) = N - M + 1 \quad (7)$$

under either uniform or nonuniform rate assignment.

<sup>1</sup>In [8], an MMSE diversity combiner is used at the receiver in the presence of one transmit antenna and  $M$  interferers.

Furthermore, it has been established [6], [7] that the ZF equalizer achieves diversity  $N - M + 1$  under both joint or separate spatial encoding.

According to Theorem 1, an MMSE receiver operating under separate spatial encoding (e.g., horizontal encoding V-BLAST) will have no more diversity gain than ZF receiver.

### III. OUTAGE ANALYSIS

We now consider the MMSE diversity where the data stream is first encoded then multiplexed into  $M$  sub-streams, each transmitted by one antenna. This approach is known to improve the performance compared with separate coding of the streams [14]. Outage occurs if the channel fails to support the target rate [12]. After channel equalization, the  $M$  substreams  $x_k$  are decoupled, and thus, the mutual information between the transmitted vector  $\mathbf{x}$  and the received vector  $\mathbf{y}$  given CSIR is [5]

$$I(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^M I(x_k, y_k). \quad (8)$$

Thus, from (2) and (8),  $P_{\text{out}}$  is given by

$$P_{\text{out}} = \mathbb{P} \left( \sum_{k=1}^M \log(1 + \gamma_k) < R \right). \quad (9)$$

Substituting MMSE SINR ( $\gamma_k$ ) from (5) in (9), we get

$$P_{\text{out}} = \mathbb{P} \left( \sum_{k=1}^M \log((\mathbf{I} + \rho \mathcal{W})_{kk}^{-1}) > -R \right). \quad (10)$$

The dependence on the diagonal elements of the random matrix  $(\mathbf{I} + \rho \mathcal{W})_{kk}^{-1}$  makes further analysis intractable. We instead proceed to provide lower and upper bounds on the outage probability. In Section IV, we will show that outage probability ( $P_{\text{out}}$ ) and PEP exhibit identical exponential error.

#### A. Outage Upper Bound

**Lemma 1:** For an MMSE MIMO system consisting of  $M$  transmit and  $N$  receive antennas, under quasi-static Rayleigh fading, we have  $P_{\text{out}}(R, M, N) \lessdot \rho^{-d_{\text{out}}(R, M, N)}$  where

$$\begin{aligned} d_{\text{out}}(R, M, N) &= \left[ (M 2^{-\frac{R}{M}} - (M - N)^+)^+ \right]^2 + \\ &\quad |N - M| \left[ (M 2^{-\frac{R}{M}} - (M - N)^+)^+ \right] \end{aligned} \quad (11)$$

where  $(\cdot)^+$  denotes the  $\max(0, \cdot)$ .

*Proof:* We begin by bounding the sum in (10) via Jensen's inequality

$$\begin{aligned} \sum_{k=1}^M \log (\mathbf{I} + \rho \mathcal{W})_{kk}^{-1} &\leq M \log \left( \sum_{k=1}^M \frac{1}{M} (\mathbf{I} + \rho \mathcal{W})_{kk}^{-1} \right) \\ &= M \log \left( \frac{1}{M} \text{tr}((\mathbf{I} + \rho \mathcal{W})^{-1}) \right) \\ &= M \log \left( \frac{1}{M} \sum_{k=1}^M \frac{1}{1 + \rho \lambda_k} \right) \quad (12) \end{aligned}$$

where (12) is true because trace is equal to the sum of eigenvalues.

Notice that for  $N < M$ , only  $N$  eigenvalues are nonzero. Hence, (12) can be written as

$$M \log \left( \frac{1}{M} \sum_{k=1}^L \frac{1}{1 + \rho \lambda_k} + (M - N)^+ \right) \quad (13)$$

where  $L = \min(M, N)$ .

Substituting (13) into (10), we have

$$P_{\text{out}} \leq \mathbb{P} \left( \sum_{k=1}^L \frac{1}{1 + \rho \lambda_k} \geq M 2^{-\frac{R}{M}} - (M - N)^+ \right). \quad (14)$$

Define

$$\alpha_k \triangleq -\frac{\log \lambda_k}{\log \rho}, \quad \text{for } k = 1, \dots, n \quad (15)$$

based on which we can write the exponential equality

$$\frac{1}{1 + \rho \lambda_k} \triangleq \begin{cases} \rho^{\alpha_k - 1}, & \alpha_k < 1 \\ 1, & \alpha_k \geq 1. \end{cases} \quad (16)$$

Define  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]$  and a new random variable

$$M(\boldsymbol{\alpha}) \triangleq \sum_{\alpha_k > 1} 1. \quad (17)$$

This definition is based on the observation that the term  $\frac{1}{1 + \rho \lambda_k}$  defined in (16) is either 0 or 1 at high SNR; therefore, to characterize  $\sum_k \frac{1}{1 + \rho \lambda_k}$  at high SNR, we count the ones. Thus

$$\sum_{k=1}^n \frac{1}{1 + \rho \lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \quad (18)$$

$$\doteq M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1} \quad (19)$$

$M(\boldsymbol{\alpha})$  inherits its randomness from  $\lambda_1, \dots, \lambda_n$ . The bound in (14) is evaluated by computing the probability of  $\{\boldsymbol{\alpha} \in \mathcal{A}\}$ , where  $\mathcal{A} = \{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1} > M 2^{-\frac{R}{M}} - (M - N)^+\}$  denotes the outage event based on the approximation in (14). In order to evaluate the probability of this event, we need the joint distribution of the eigenvalues, or equivalently the distribution of  $\boldsymbol{\alpha}$ . The distribution follows Wishart distribution and was initially discovered by [13]. The distribution of  $\boldsymbol{\alpha}$  can be easily evaluated as follows [15].

Let  $\mathbf{R}$  be an  $m \times n$  ( $m \geq n$ ) random matrix whose entries are  $\mathcal{CN}(0, 1)$ . The joint PDF of the ordered random variables  $\boldsymbol{\alpha}$  (defined in (15) for the eigenvalues of  $\mathbf{R}^H \mathbf{R}$ ) is given by

$$\begin{aligned} \mathbb{P}(\boldsymbol{\alpha}) &= K_{m,n}^{-1} (\log \rho)^n \prod_{i=1}^n \rho^{-(m-n+1)\alpha_i} \times \\ &\quad \prod_{i < j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \exp \left[ -\sum_{i=1}^n \rho^{-\alpha_i} \right] \quad (20) \end{aligned}$$

where  $K_{m,n}^{-1}$  is a normalizing factor.

Using the distribution of  $\boldsymbol{\alpha}$  for the defined matrix  $\mathbf{R}$ , the asymptotic outage bound is

$$\begin{aligned} P_{\text{out}} &\dot{\leq} \int_{\mathcal{A}} \mathbb{P}(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \\ &= K_{m,n}^{-1} (\log \rho)^n \int_{\mathcal{A}} \prod_{i=1}^n \rho^{-(m-n+1)\alpha_i} \prod_{i < j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \times \\ &\quad \exp \left[ -\sum_{i=1}^n \rho^{-\alpha_i} \right] d\boldsymbol{\alpha}. \quad (21) \end{aligned}$$

The simplification of the integral follows from [15]. The term outside the integral has no effect on the exponent. The term  $|\rho^{-\alpha_i} - \rho^{-\alpha_j}|$  is dominated by  $\rho^{-\alpha_i}$  at high SNR. We now divide the integration range into  $\mathcal{A}' = \mathcal{A} \cap \mathbb{R}_+^n$  and its complement. If  $\boldsymbol{\alpha} \notin \mathcal{A}'$ , the exponential term will dominate the other terms and will drive the integral to zero. If  $\boldsymbol{\alpha} \in \mathcal{A}'$ , the exponential term is approximately 1 at high SNR and will disappear. Therefore

$$\begin{aligned} P_{\text{out}} &\dot{\leq} \int_{\mathcal{A}'} \prod_{i=1}^n \rho^{-(m-n+1)\alpha_i} \prod_{i < j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 d\boldsymbol{\alpha} \\ &\doteq \int_{\mathcal{A}'} \prod_{i=1}^n \rho^{-(2i-1+m-n)\alpha_i} d\boldsymbol{\alpha} \quad (22) \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}' &= \{M(\boldsymbol{\alpha}) > M 2^{-\frac{R}{M}} - (M - N)^+\} \\ &= \{\alpha_1 > 1, \dots, \alpha_S > 1, \alpha_{S+1} > 0, \dots, \alpha_L > 0\} \quad (23) \end{aligned}$$

and  $S = \lceil (M 2^{-\frac{R}{M}} - (M - N)^+)^+ \rceil$ . The integration region  $\mathcal{A}'$  has boundaries that are parallel to nonnegative orthant  $\mathbb{R}_+^n$ ; therefore, the integration over multiple variables in (22) can be separated

$$P_{\text{out}} \dot{\leq} \prod_{i=1}^n \int_{\mathcal{A}'} \rho^{-(2i-1+m-n)\alpha_i} d\boldsymbol{\alpha} \quad (24)$$

$$= \rho^{-\sum_{i=1}^S (2i-1+m-n)} \quad (25)$$

$$= \rho^{-(S^2 + (m-n)S)}, \quad \text{for } m \geq n \quad (25)$$

$$= \rho^{-(S^2 + |m-n|S)}, \quad \text{for general } m, n \quad (26)$$

$$= \rho^{-d_{\text{out}}} \quad (26)$$

which establishes the proof of Lemma 1.  $\square$

### B. Outage Lower Bound

**Lemma 2:** For an MMSE MIMO system consisting of  $M$  transmit and  $N$  receive antennas (and  $L = \min\{M, N\}$ ), operating under quasi-static Rayleigh fading, we have  $P_{\text{out}}(R, M, N) \geq \rho^{-d_{\text{out}}(R, M, N)}$  where

$$d_{\text{out}}(R, M, N) = \left[ (M2^{-\frac{R}{M}} - (M-N)^+)^+ \right]^2 + |N-M| \left[ (M2^{-\frac{R}{M}} - (M-N)^+)^+ \right].$$

*Proof:* The lower bound is also based on Jensen's inequality. Recall

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\sum_{k=1}^M \log(1+\gamma_k) < R\right) \\ &= \mathbb{P}\left(\sum_{k=1}^M \log \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} < R\right) \\ &\geq \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^M \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} < R\right). \end{aligned} \quad (27)$$

Let the eigendecomposition of  $\mathbf{H}^H \mathbf{H}$  be given by  $\mathbf{H}^H \mathbf{H} = \mathbf{U}^H \Lambda \mathbf{U}$ , where  $\mathbf{U}$  is unitary and  $\Lambda$  is a diagonal matrix that has the eigenvalues of the Wishart matrix  $\mathcal{W}$  on its diagonal. Let the vector  $\mathbf{u}_k$  be the column  $k$  of the matrix  $\mathbf{U}$  and  $u_{\ell k}$  be the element  $\ell$  of this column; we have

$$\begin{aligned} (\mathbf{I} + \rho\mathcal{W})_{kk}^{-1} &= \mathbf{u}_k^H (\mathbf{I} + \rho\Lambda)^{-1} \mathbf{u}_k \\ &= \sum_{\ell=1}^M \frac{|u_{\ell k}|^2}{1 + \rho\lambda_\ell} \\ &\triangleq S_k. \end{aligned} \quad (28)$$

Let  $\bar{k} = \arg \min_k S_k$ . Using (28), we can bound the sum in (27)

$$\begin{aligned} \frac{1}{M} \sum_{k=1}^M \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} &= \frac{1}{M} \sum_{k=1}^M \frac{1}{S_k} \\ &\leq \frac{1}{\min_k S_k} \end{aligned} \quad (29)$$

$$= \frac{1}{S_{\bar{k}}}. \quad (30)$$

Thus, the outage bound in (27) can be further bounded using (29)

$$\begin{aligned} P_{\text{out}} &\geq \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^M \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} < R\right) \\ &\geq \mathbb{P}\left(M \log \frac{1}{S_{\bar{k}}} < R\right) \\ &= \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}}\right). \end{aligned} \quad (31)$$

We now bound (31) by conditioning on the event  $\mathcal{B} \triangleq \{|u_{\ell\bar{k}}|^2 \geq \frac{a}{M}\}$ , where  $a$  is a positive real number that

is slightly smaller than 1, i.e.,  $a = 1 - \epsilon$ , and  $\epsilon$  is a small positive number. We then have

$$\begin{aligned} \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}}\right) &\geq \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \mid \mathcal{B}\right) \mathbb{P}(\mathcal{B}) \\ &= \mathbb{P}\left(\sum_{\ell=1}^M \frac{|u_{\ell\bar{k}}|^2}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}} \mid \mathcal{B}\right) \mathbb{P}(\mathcal{B}) \\ &\geq \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^M \frac{a}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}}\right) \mathbb{P}(\mathcal{B}) \\ &\stackrel{a}{=} \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^M \frac{a}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}}\right) \\ &= \mathbb{P}\left(\sum_{\ell=1}^M \frac{1}{1 + \rho\lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}}\right) \\ &= \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^L \frac{1}{1 + \rho\lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}} - (M-N)^+\right) \end{aligned} \quad (32)$$

where (32) follows because  $\mathbb{P}(\mathcal{B})$  is finite and independent of  $\rho$ ; this can be proved similarly to [7, Appendix A]. To make the upcoming expressions compact, we introduce a new variable  $\kappa \triangleq \frac{M}{a} 2^{-\frac{R}{M}} - (M-N)^+$

$$\mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^L \frac{1}{1 + \rho\lambda_\ell} > \kappa\right). \quad (34)$$

Whenever  $M2^{-\frac{R}{M}}$  is noninteger, the constant  $a$  can be chosen such that  $\lceil (M2^{-\frac{R}{M}} - (M-N)^+)^+ \rceil = \lceil (\frac{M}{a} 2^{-\frac{R}{M}} - (M-N)^+)^+ \rceil$ . We note this is satisfied for all rates, with the exception of an isolated set of points. As long as  $M2^{-\frac{R}{M}} \notin \mathbb{N}$ , we have

$$\begin{aligned} P_{\text{out}} &\geq \mathbb{P}\left(\sum_{\ell=1}^L \frac{1}{1 + \rho\lambda_\ell} > \kappa\right) \\ &\stackrel{a}{=} \mathbb{P}\left(\sum_{k=1}^L \frac{1}{1 + \rho\lambda_\ell} > \lceil \kappa \rceil\right). \end{aligned} \quad (35)$$

The remaining steps follow similarly to the proof of Lemma 1. Thus,  $P_{\text{out}} \geq \rho^{-d_{\text{out}}}$  with  $d_{\text{out}}$  is given by Lemma 2.

On the set of isolated points  $M2^{-\frac{R}{M}} \in \mathbb{N}$ , the right-hand side of (35) obeys a slightly weaker upper bound by replacing  $\kappa$  with  $\kappa + 1$ . We can combine the cases where  $M2^{-\frac{R}{M}}$  is integer and noninteger to write the upper bound compactly as follows:

$$\begin{aligned} d_{\text{out}}(R, M, N) &\leq \left[ (M2^{-\frac{R}{M}} + 1 - (M-N)^+)^+ \right]^2 + \\ &\quad |N-M| \left[ (M2^{-\frac{R}{M}} + 1 - (M-N)^+)^+ \right]. \end{aligned}$$

Inspection shows that this bound is tight against the lower bound everywhere except its discontinuity points. In other words, the upper bound is left-continuous while the lower bound was right-continuous at the discontinuity points.  $\square$

#### IV. PEP ANALYSIS

Recalling that the diversity is roughly defined as the slope of PEP at high SNR, we now proceed to bound the PEP tightly from both sides using the outage results already obtained.

##### A. PEP Upper Bound

We start by a lower bound that is inspired by [15, Lemma 5] but requires a more careful treatment since we are analyzing rate, not the DMT (see Section I).

*Lemma 3:* For a quasi-static fading MIMO channel with MMSE receiver, we have  $d_{\text{out}}(R, M, N) \geq d(R, M, N)$ .

*Proof:* Denote  $E$  for an error event, and let  $x \in \mathcal{C}$  be the transmitted codeword from a codebook  $\mathcal{C}$  of size  $2^{Rl}$  where  $R$  and  $l$  are code rate and code length respectively. Define  $\mathbf{f} = \mathbf{W}\mathbf{H}\mathbf{x}$  that accounts for the combined effect of channel and equalizer. The transmit messages are assumed equiprobable so the entropy  $\mathcal{H} = \log |\mathcal{C}| = Rl$ . Applying the Fano inequality [16]

$$\mathbb{P}(E|\mathbf{f} = f) \geq \frac{Rl - I(\mathbf{x}; \mathbf{y}|\mathbf{f} = f)}{Rl} - \frac{\mathcal{H}(\mathbb{P}(E)|\mathbf{f} = f)}{Rl}. \quad (36)$$

By defining  $\mathcal{D}_\delta$  for any  $\delta > 0$  as  $\mathcal{D}_\delta \triangleq \{f : I(\mathbf{x}; \mathbf{y}|\mathbf{f} = f) < l(R - \delta)\}$ , and noting that  $\mathcal{H}(\mathbb{P}(E)|f \in \mathcal{D}_\delta) \leq \mathcal{H}(\mathbb{P}(E))$  from (36), we get

$$\begin{aligned} \mathbb{P}(E|\mathbf{f} \in \mathcal{D}_\delta) &\geq \frac{Rl - I(\mathbf{x}; \mathbf{y}|\mathbf{f} \in \mathcal{D}_\delta)}{Rl} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl} \\ &\geq \frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl}. \end{aligned} \quad (37)$$

Also by using the definition of  $P_{\text{out}}$ , we have

$$\mathbb{P}(\mathbf{f} \in \mathcal{D}_\delta) = \mathbb{P}(I(\mathbf{x}; \mathbf{y}) < l(R - \delta)) \doteq \rho^{-d_{\text{out}}(R - \delta, M, N)}. \quad (38)$$

For small enough values of  $\delta > 0$ , we have  $d_{\text{out}}(R, M, N) = d_{\text{out}}(R - \delta, M, N)$  since  $d_{\text{out}}(R, M, N)$  is left-continuous with respect to  $R$ . Hence, by invoking (37) and (38), the error probability is given by

$$\begin{aligned} \mathbb{P}_{\text{err}}(R, M, N) &= \mathbb{P}(E|f \in \mathcal{D}_\delta)\mathbb{P}(f \in \mathcal{D}_\delta) + \\ &\quad \mathbb{P}(E|f \notin \mathcal{D}_\delta)\mathbb{P}(f \notin \mathcal{D}_\delta) \\ &\geq \mathbb{P}(E|f \in \mathcal{D}_\delta)\mathbb{P}(f \in \mathcal{D}_\delta) \\ &\geq \left( \frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl} \right) \rho^{-d_{\text{out}}} \\ &\doteq \rho^{-d_{\text{out}}} \end{aligned} \quad (39)$$

where we have used  $\left( \frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl} \right) \doteq 1$ , which was derived in [10]. This establishes the proof of the PEP upper bound.  $\square$

##### B. PEP Lower Bound

We begin by writing the error probability in terms of error event  $E$  and outage event  $O$

$$\mathbb{P}_{\text{err}}(R, M, N) = \mathbb{P}(E|O) \cdot P_{\text{out}} + \mathbb{P}(E, \bar{O})$$

In Section III-A, we have shown that, based on the event  $\left\{ \sum_{k=1}^L \frac{1}{1+\rho\lambda_k} \geq M 2^{-\frac{R}{M}} - (M-N)^+ \right\}$ , the outage probability is upper bounded by  $P_{\text{out}} \leq \rho^{-d_{\text{out}}}$ . Hence, the error probability can be bounded as

$$\begin{aligned} \mathbb{P}_{\text{err}}(R, M, N) &\leq \mathbb{P}(E|O) \rho^{-d_{\text{out}}} + \mathbb{P}(E, \bar{O}) \\ &\leq \rho^{-d_{\text{out}}} + \mathbb{P}(E, \bar{O}) \end{aligned} \quad (40)$$

We intend to show that  $\rho^{-d_{\text{out}}} \geq \mathbb{P}(E, \bar{O})$ , and thus  $\mathbb{P}_{\text{err}}(R, M, N) \leq \rho^{-d_{\text{out}}}$  which produces the following lemma.

*Lemma 4:* For a quasi-static fading MIMO channel with MMSE receiver, we have  $d_{\text{out}}(R, M, N) \leq d(R, M, N)$ .

*Proof:* We begin by giving a sketch of the proof then we proceed with the details. The first part of the proof consists of developing a bound on PEP conditioned on  $H$ , namely  $P[s_k \rightarrow s_j | \mathbf{H} = H]$ . To do this, we obtain an upper bound of the variance of the SINR which is expressed in terms of the eigenvalues of the Wishart matrix  $\mathcal{W}$ , resulting in  $\mathbb{P}[E|\mathbf{H} = H] \leq 4 \exp\left(-\left(\sum_{k=1}^L \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}\right)$ . The PEP is used to derive a conditional union bound on error. We then divide the channel events into two sets based on the exponential order of the eigenvalues: the set where  $M(\boldsymbol{\alpha}) = 0$  and otherwise. We apply the Bayes theorem on the union bound using these two sets. The calculation of the terms of the Bayesian gives  $\mathbb{P}(E, \bar{O}) \leq \rho^{-MN} \leq \rho^{-d_{\text{out}}}$  as desired.

We now proceed in detail. We want to compute the probability that the transmitted symbol  $x(k) = s_l$  is erroneously detected as  $x(k) = s_j$ .

Recalling the equalizer output given by (6), define the noise-plus-interference signal

$$\tilde{\mathbf{n}} = \mathbf{y} - \sqrt{\rho}\mathbf{x} = \sqrt{\rho}(\mathbf{W}\mathbf{H} - \mathbf{I})\mathbf{x} + \mathbf{W}\mathbf{n}. \quad (41)$$

Using the eigendecomposition of  $H$  and noting that  $E(\mathbf{n}) = 0$  and  $E(\mathbf{n}\mathbf{n}^H) = I$ , we have

$$\mu_{\tilde{\mathbf{n}}} \triangleq E(\tilde{\mathbf{n}}) = \sqrt{\rho}(\mathbf{W}\mathbf{H} - \mathbf{I}) = -\rho^{\frac{1}{2}}(\mathcal{W} + \rho^{-1}\mathbf{I})^{-1}\mathbf{x} \quad (42)$$

$$\mathbf{R}_{\tilde{\mathbf{n}}} \triangleq E(\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H) = (\mathcal{W} + \rho^{-1}\mathbf{I})^{-1}. \quad (43)$$

Thus, the variance of the noise sample  $\tilde{n}(k)$  is given by

$$\begin{aligned} \sigma_{\tilde{n}}^2(k) &= \mathbf{R}_{\tilde{\mathbf{n}}}(k, k) - |\mu_{\tilde{\mathbf{n}}}(k)|^2 \\ &= (\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-1} - \rho^{-1}(\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-2} \end{aligned} \quad (44)$$

where  $|\mu_{\tilde{\mathbf{n}}}(k)|^2$  is the  $k$ th diagonal of the matrix  $E(\tilde{\mathbf{n}})E(\tilde{\mathbf{n}}^H)$  and  $k$  counts from 1 to  $M$ .

By defining  $e_{jl} \triangleq \frac{s_j - s_l}{|s_j - s_l|}$ , the probability of erroneous detection for channel realization is given by

$$\begin{aligned} &\mathbb{P}[s_l \rightarrow s_j | \mathbf{H} = H] \\ &= \mathbb{P}\left[\frac{\rho}{4}|s_j - s_l|^2 \leq |e_{jl}^*(y(k) - \sqrt{\rho}s_l)|^2 \middle| \mathbf{H} = H\right] \\ &\leq \mathbb{P}\left[\frac{\rho}{4}|s_j - s_l|^2 \leq |\tilde{n}_k|^2 \middle| \mathbf{H} = H\right] \end{aligned} \quad (45)$$

where the inequality holds since  $|e_{jl}^*(y(k) - \sqrt{\rho}s_l)| \leq |e_{jl}^*|(y(k) - \sqrt{\rho}s_l)| = |(y(k) - \sqrt{\rho}s_j)| = |\tilde{n}(k)|$ .

Denoting the real and imaginary parts of  $\tilde{n}(k)$  by  $\tilde{n}_r(k) \sim \mathcal{N}(\mu_r(k), \sigma_r^2(k))$  and  $\tilde{n}_i(k) \sim \mathcal{N}(\mu_i(k), \sigma_i^2(k))$ , respectively, we then have

$$\begin{aligned} & \left\{ \frac{\rho}{4}|s_j - s_l|^2 \leq |\tilde{n}(k)|^2 \right\} \\ & \subset \left\{ \frac{\rho}{16}|s_j - s_l|^2 \leq |\tilde{n}_r(k)|^2 \right\} \cup \left\{ \frac{\rho}{16}|s_j - s_l|^2 \leq |\tilde{n}_i(k)|^2 \right\}. \end{aligned} \quad (46)$$

Applying the property of the Gaussian tail function  $Q(x) \leq e^{(-x^2/2)}$  for the PEP, we obtain

$$\begin{aligned} & \mathbb{P}[s_k \rightarrow s_j | \mathbf{H} = H] \\ & \leq e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| - \mu_r(k))^2}{\sigma_r^2(k)} \right)} + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| + \mu_r(k))^2}{\sigma_r^2(k)} \right)} \\ & + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| - \mu_i(k))^2}{\sigma_i^2(k)} \right)} + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| + \mu_i(k))^2}{\sigma_i^2(k)} \right)} \\ & \leq e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| - \mu_r(k))^2}{\sigma_{\bar{n}}^2(k)} \right)} + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| + \mu_r(k))^2}{\sigma_{\bar{n}}^2(k)} \right)} \\ & + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| - \mu_i(k))^2}{\sigma_{\bar{n}}^2(k)} \right)} + e^{\left( -\frac{(\sqrt{\rho}|s_j - s_l| + \mu_i(k))^2}{\sigma_{\bar{n}}^2(k)} \right)} \end{aligned} \quad (47)$$

where the last step holds as  $\sigma_{\bar{n}}^2(k) = \sigma_r^2(k) + \sigma_i^2(k) \geq \sigma_r^2(k), \sigma_i^2(k)$ .

Now we proceed by showing that  $\mu_i(k) \leq \rho^{1/2}$ . Consider the eigendecomposition of

$$\begin{aligned} [\mathcal{W} + \rho^{-1}\mathbf{I}]^{-1} &= \mathbf{U}^H[\Lambda + \rho^{-1}\mathbf{I}]^{-1}\mathbf{U} \\ &= \mathbf{U}^H[\text{diag}\left\{\frac{1}{\lambda_k + \rho^{-1}}\right\}]\mathbf{U} \end{aligned} \quad (48)$$

where  $\mathbf{U}$  is unitary matrix, and  $\Lambda$  is the eigendecomposition of  $\mathcal{W}$ . Note that  $\lambda_k + \rho^{-1} \geq \rho^{-1}$  or  $\frac{1}{\lambda_k + \rho^{-1}} \leq \rho$ . Therefore, all elements of the matrix  $\pm \mathbf{U}^H[\Lambda + \rho^{-1}\mathbf{I}]^{-1}\mathbf{U}$ , being linear combination of  $\{\frac{1}{\lambda_k + \rho^{-1}}\}$ , cannot grow faster than  $O(\rho)$ , and thus the elements of  $\pm \rho^{1/2}[\mathcal{W} + \rho^{-1}\mathbf{I}]^{-1}$  cannot grow faster than  $O(\rho^{1/2})$ , i.e.,  $\pm \mu_{\bar{n}}(k) \leq \rho^{1/2}$  and therefore  $\rho^{1/2} \pm \mu_{\bar{n}}(k) \leq \rho^{1/2}$ . The same result holds for  $\mu_r(k)$  and  $\mu_i(k)$ .

As a result, for any  $s_j$  and  $s_l$ ,  $\frac{\sqrt{\rho}}{4}|s_j - s_l| \pm \mu_r(k) \leq \rho^{1/2} \pm \mu_r(k) \leq \rho^{1/2}$  and similarly  $\frac{\sqrt{\rho}}{4}|s_j - s_l| \pm \mu_i(k) \leq \rho^{1/2}$ . Thus, from (47), we have

$$\mathbb{P}[s_k \rightarrow s_j | \mathbf{H} = H] \leq 4e^{-\frac{\rho}{\sigma_{\bar{n}}^2(k)}}. \quad (49)$$

Now we bound the variance in (44) and apply it in (49)

$$\begin{aligned} \sigma_{\bar{n}}^2(k) &\leq \sum_{k=1}^L \left[ (\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-1} - \rho^{-1}(\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-2} \right] \\ &= \sum_{k=1}^L \left[ \frac{\rho}{1 + \rho\lambda_k} - \frac{\rho}{(1 + \rho\lambda_k)^2} \right] = \sum_{k=1}^L \frac{\rho^2\lambda_k}{(1 + \rho\lambda_k)^2}. \end{aligned} \quad (50)$$

Denoting the error event  $E$  and using (50), the probability of erroneous detection in (49) is bounded as

$$\mathbb{P}[E | \mathbf{H} = H] \leq 4e^{-\left(\sum_{k=1}^L \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}}. \quad (51)$$

Applying the union bound, we get

$$\mathbb{P}(E | \mathbf{H} = H) \leq 2^{RL} e^{-\left(\sum_{k=1}^L \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}}. \quad (52)$$

Based on (52), we can evaluate  $P(E, \bar{O})$  in (40) as follows. Recalling the exponential inequality

$$\sum_{k=1}^n \frac{1}{1 + \rho\lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \quad (53)$$

$$\doteq M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1} \doteq M(\boldsymbol{\alpha}). \quad (54)$$

Consider the two regions:  $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) = 0\}$  and  $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \geq 1\}$ . At high SNR, the event  $\bar{O}$  is equivalent to  $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \leq [M2^{-\frac{R}{M}} - (M - N)^+]\}$ .

In the first region  $\{M(\boldsymbol{\alpha}) = 0\}$ , at any rate  $R \geq 0$ , we have  $\{\boldsymbol{\alpha} : [M2^{-\frac{R}{M}} - (M - N)^+] \geq M(\boldsymbol{\alpha}) = 0$  so there is no outage.

In the second region  $\{M(\boldsymbol{\alpha}) \geq 1\}$ , the exponent order of the outage probability depends on the rate. We investigate these two regions separately.

In the region  $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) = 0\}$ , we have  $\max_k \alpha_k < 1$  since all  $\alpha'_k < 1$ . From (52) and (54), we conclude that

$$\begin{aligned} \mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) = 0) &\leq 2^{RL} e^{-\rho^{\left(\max_k \alpha_k - 1\right)^{-1}}} \\ &= 2^{RL} e^{-\rho^{\left(1 - \max_k \alpha_k\right)}}. \end{aligned} \quad (55)$$

Since exponential function dominates all polynomials and  $1 - \max_k \alpha_k > 0$ , we get

$$\lim_{\rho \rightarrow \infty} \frac{e^{-\rho^{\left(1 - \max_k \alpha_k\right)}}}{\rho^{-MN}} = 0$$

which in turn yields

$$\begin{aligned} \mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) = 0) &\leq 2^{RL} e^{-\rho^{\left(1 - \max_k \alpha_k\right)}} \\ &\leq \rho^{-MN}. \end{aligned} \quad (56)$$

We next show that the same result holds for the other region  $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \geq 1\}$ .

Following the same line of argument as we did for (56) but for  $M(\boldsymbol{\alpha}) \geq 1$ , we have

$$\begin{aligned} \mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) \geq 1) &\leq 2^{RL} e^{-\left(\sum_{k=1}^L \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}} \\ &\leq e^{2^{RL}} e^{-\left(\sum_k \frac{1}{1+\rho\lambda_k} - \sum_k \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}} \\ &= e^{2^{RL}} e^{-\left(\sum_k \frac{1}{1+\rho\lambda_k}\right)^{-1}} \times \\ &\quad \doteq 1 \text{ since } M(\boldsymbol{\alpha}) \geq 1 \\ &\quad e^{\left[-\frac{\sum_k \frac{1}{(\rho\lambda_k + 1)^2}}{\left(\sum_k \frac{1}{1+\rho\lambda_k}\right)\left(\sum_k \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)}\right]} \\ &\leq e^{2^{RL}} e^{\left[-\frac{L M(\boldsymbol{\alpha})}{L M(\boldsymbol{\alpha}) - \min_k |1 - \alpha_k|}\right]} \\ &\doteq e^{-\rho^{\min_k |1 - \alpha_k|}} \\ &\leq e^{\rho^{1 - \max_k \alpha_k}} \leq \rho^{-MN} \end{aligned} \quad (57)$$

$$\leq \rho^{-MN} \quad (58)$$

where (57) is direct application of (54) for  $M(\boldsymbol{\alpha}) \geq 1$ , and (58) follows from the fact that  $|1 - \alpha_k| \geq 1$ . Note that (58) is true for any code length  $l$ . Invoking the results of (56) and (58), we can now evaluate  $\mathbb{P}(E, \bar{O})$  as follows:

$$\begin{aligned} \mathbb{P}(E, \bar{O}) &= \int_{M(\boldsymbol{\alpha})=0} \mathbb{P}(E, \bar{O}|M(\boldsymbol{\alpha})=0)\mathbb{P}(\boldsymbol{\alpha})d\boldsymbol{\alpha} \\ &\quad + \int_{M(\boldsymbol{\alpha}) \geq 1} \mathbb{P}(E, \bar{O}|M(\boldsymbol{\alpha}) \geq 1)\mathbb{P}(\boldsymbol{\alpha})d\boldsymbol{\alpha} \end{aligned} \quad (59)$$

$$\leq \rho^{-MN} \int_{M(\boldsymbol{\alpha})=0} \mathbb{P}(\boldsymbol{\alpha})d\boldsymbol{\alpha} + \rho^{-MN} \int_{M(\boldsymbol{\alpha}) \geq 1} \mathbb{P}(\boldsymbol{\alpha})d\boldsymbol{\alpha} \quad (60)$$

$$= \rho^{-MN}. \quad (61)$$

Therefore,  $\mathbb{P}(E, \bar{O}) \leq \rho^{-MN}$  for all regions of  $\boldsymbol{\alpha}$ . Finally, (40) becomes

$$\begin{aligned} \mathbb{P}_{\text{err}}(R, M, N) &\leq \mathbb{P}(E|O)\rho^{-d_{\text{out}}} + \mathbb{P}(E, \bar{O}) \\ &\leq \rho^{-d_{\text{out}}} + \mathbb{P}(E, \bar{O}) \\ &\doteq \rho^{-d_{\text{out}}} + \rho^{-MN} \\ &\doteq \rho^{-d_{\text{out}}} \\ &= P_{\text{out}}(R, M, N) \end{aligned} \quad (62)$$

which establishes the lemma.  $\square$

From Lemma 3 and Lemma 4, we thus get the following.

**Theorem 2:** For MMSE MIMO receiver under quasi-static channel and joint spatial encoding, the PEP and the outage probability  $P_{\text{out}}$  are exponentially equal and the diversity gain is  $d(R, M, N) = d_{\text{out}}(R, M, N)$ , where  $d_{\text{out}}(R, M, N)$  is given in (11).

## V. MULTIPLE ACCESS CHANNEL (MAC)

We now extend the result to the MAC channel. Consider a MIMO MAC channel with  $K$  users,  $M$  transmit antennas per user, and  $N$  receive antennas (there is no condition on  $M$ ,  $N$ , and  $k$ ). Assume flat-fading MIMO channel, the system model is given by

$$\mathbf{y} = \sum_{i=1}^K \mathbf{H}_i \mathbf{x}_i + \mathbf{n} = \mathbf{H}_e \mathbf{X} + \mathbf{n} \quad (63)$$

where  $\mathbf{H}_i \in \mathbb{C}^{N \times M}$  is the user  $i$  channel matrix whose entries are i.i.d. complex Gaussian,  $\mathbf{H}_e = [\mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_K]$  is the overall equivalent channel matrix,  $\mathbf{x}_i \in \mathbb{C}^{M \times 1}$  is the transmitted vector of user  $i$ ,  $\mathbf{X} = [\mathbf{x}_1^T \mathbf{x}_2^T \cdots \mathbf{x}_K^T]^T$  is the overall transmitted vector, and  $\mathbf{n} \in \mathbb{C}^{N \times 1}$  is the Gaussian noise vector. The vectors  $\mathbf{X}$  and  $\mathbf{n}$  are assumed independent. We keep the same assumptions about the channel. That is we assume a quasi-static flat-fading channel and perfect CSIR and no CSIT. We have the following theorem

**Theorem 3:** In a MIMO MAC system with MMSE receiver consisting of  $K$  users,  $M$  transmit antennas per user, and  $N$  receive antennas, the lower and upper bounds on the per user diversity are, respectively, given by  $d_L^{\text{MAC}}(R)$  and  $d_U^{\text{MAC}}(R)$

$$\begin{aligned} d_L^{\text{MAC}}(R) &= \left[ (M2^{-R/M} - (M-N)^+)^+ \right]^2 + \\ &\quad |N-KM| \left[ (M2^{-R/M} - (M-N)^+)^+ \right] \end{aligned} \quad (64)$$

$$\begin{aligned} d_U^{\text{MAC}}(R) &= \left[ (KM2^{-R/KM} - (M-N)^+)^+ \right]^2 + \\ &\quad |N-KM| \left[ (KM2^{-R/KM} - (M-N)^+)^+ \right]. \end{aligned} \quad (65)$$

From (64), it is straightforward to verify the single-user case. The machinery of the proof is mostly similar to the single-user case. However, the outage upper and lower bounds are obtained in a different manner that is pointed out in the following analysis for  $N \geq M$ . The case  $N < M$  can be similarly obtained.

### A. MAC Outage Upper Bound

The user  $i$  outage probability can be written as

$$P_{\text{out}}^i = \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(1 + \gamma_k^i) < R\right) \quad (66)$$

where  $\gamma_k^i$  is the SINR of the stream  $k$  of user  $i$ . Specializing this to MMSE receiver, we get

$$P_{\text{out}}^i = \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right). \quad (67)$$

Using Jensen's Inequality, the outage probability can be bounded as

$$\begin{aligned} P_{\text{out}}^i &\leq \mathbb{P}\left(\log\left(\sum_{k=(i-1)M+1}^{iM} \frac{1}{M} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1}\right) > \frac{-R}{M}\right) \\ &\leq \mathbb{P}\left(\log\left(\sum_{k=1}^{KM} \frac{1}{M} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1}\right) > \frac{-R}{M}\right) \end{aligned} \quad (68)$$

$$= \mathbb{P}\left(\sum_{k=1}^{KM} \frac{1}{1 + \rho \lambda_k} > M 2^{-\frac{R}{M}}\right) \quad (69)$$

where (68) is true since the summation in the left-hand side of the inequality adds more positive terms (recall that  $(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)$  is a positive-definite matrix [12]). Following similar steps that were used to obtain (26), we can easily show that  $P_{\text{out}}^i \leq \rho^{-d_L^{\text{MAC}}}$ , where  $d_L^{\text{MAC}}$  is given by (64).

### B. MAC Outage Lower Bound

The outage probability can be lower bounded as follows:

$$\begin{aligned} P_{\text{out}}^i &= \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \\ &\geq \mathbb{P}\left(\sum_{k=1}^{KM} \log(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \quad (70) \end{aligned}$$

$$\geq \mathbb{P}\left(\sum_{k=1}^{KM} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > \frac{KM}{a} 2^{\frac{-R}{M}}\right) \quad (71)$$

where (70) is a trivial bound based on dedicating all  $KM$  antennas to *one* user, and (71) uses the same technique as in Section III-B, and  $a$  is a positive number slightly less than one. Following similar steps that were used to obtain (26), we can easily show that  $P_{\text{out}}^i \geq \rho^{-d_U^{\text{MAC}}}$ , where  $d_U^{\text{MAC}}$  is given by (65).

## VI. FREQUENCY-SELECTIVE CHANNEL

Broadband wireless systems usually operate in frequency-selective channels where, in addition to the spatial diversity obtained in MIMO broadband systems, frequency diversity can be achieved. Broadband systems usually employ OFDM or SC transmission [17]. Specifically, SC was shown to be attractive for broadband wireless channels due to its lower complexity, lower peak-to-average power ratio, and reduced sensitivity to carrier frequency errors compared to OFDM [17], [18].

In this section, we investigate the diversity achieved by SC-MMSE receivers for two block transmission schemes, namely CP and ZP schemes. The CP and ZP are commonly used for guard intervals in block quasi-static channels. Although CP was initially proposed for both single- and multicarrier systems, ZP was lately shown to be an attractive alternative for both systems [19], [20].

### A. System Model

We consider a general MIMO system in a rich scattering quasi-static environment. The equivalent baseband channel is given by multipath model with  $\nu$  paths referred to as the ISI channel in the sequel. The  $(\nu+1)$ -tap channel impulse response between the transmit antenna  $m$  and receive antenna  $n$  is denoted by the vector  $\mathbf{h}_{mn} = [h_{mn,0}, h_{mn,1}, \dots, h_{mn,\nu}]$ . We assume a block-fading model where  $\mathbf{h}_{mn}$  remains unchanged during a transmission block. Assuming  $M$  transmit and  $N$  receive antennas, the received vector  $\mathbf{y}_k$  at time instant  $k$  is given by [10], [21]

$$\mathbf{y}_k = \sum_{i=0}^{\nu} \mathbf{H}_i \mathbf{x}_{k-i} + \mathbf{n}_k \quad (72)$$

where  $\mathbf{H}_i$  is the  $M \times N$  channel matrix that has  $h_{mn,i}$  as its  $(m, n)$  element,  $\mathbf{x}_{k-i}$  is  $M \times 1$  transmitted vector at time index  $k-i$ ,  $\mathbf{y}_k$  is the  $N \times 1$  received vector and  $\mathbf{n}_k$  is the  $N \times 1$  Gaussian noise vector at time index  $k$ .

Consider a transmission of  $L_d + L_e$  spatial vectors each of size  $M \times 1$ , where  $L_d$  is an integer representing the number of transmissions over the quasi-static channel and  $L_e$  is the length of data extension to avoid interblock interference, in the form of either ZP or CP. The receiver discards the  $L_e$  vectors in the case of CP transmission [21]. Stacking the transmitted vector in an  $M(L_d + L_e) \times 1$  vector, we can write the stacked  $M(L_d + L_e) \times 1$  transmitted as follows:

$$\bar{\mathbf{x}}_k = [\mathbf{x}_{k(L_d+L_e)}^T, \dots, \mathbf{x}_{k(L_d+L_e)+L_d+L_e-1}^T]^T.$$

We can then rewrite (72) as

$$\bar{\mathbf{y}}_{cp} = \bar{\mathbf{H}} \bar{\mathbf{x}} + \bar{\mathbf{n}} \quad (73)$$

where  $\bar{\mathbf{y}}_{cp}$  is the  $NL_d \times 1$  received vector,  $\bar{\mathbf{x}}$  is the  $M(L_d + L_e) \times 1$  transmitted vector,  $\bar{\mathbf{n}}$  is the white Gaussian noise vector in  $\mathbb{C}^{NL_d \times 1}$  and  $\bar{\mathbf{H}}$  is the channel matrix given by

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & 0 & \cdots & 0 \\ 0 & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu \end{bmatrix}. \quad (74)$$

The linear data extension operation maps the data vector  $\hat{\mathbf{x}}$  to the transmitted vector  $\bar{\mathbf{x}}$  and is shown by

$$\bar{\mathbf{x}} = \mathbf{U}_{cp} \hat{\mathbf{x}} \quad (75)$$

where  $\mathbf{U}_{cp}$  is given by

$$\mathbf{U}_{cp} = \begin{bmatrix} \mathbf{I}_{ML_d} & \\ \mathbf{I}_{ML_e} & \mathbf{0}_{ML_e \times (L_d-L_e)M} \end{bmatrix} \quad (76)$$

The system model in (73) can now be written in terms of the unpadded data vector  $\hat{\mathbf{x}}$  and an equivalent channel matrix  $\mathbf{H}_e$  as follows:

$$\bar{\mathbf{y}}_{cp} = \mathbf{H}_e \hat{\mathbf{x}} + \bar{\mathbf{n}} \quad (77)$$

where in a CP system,  $\mathbf{H}_e = \bar{\mathbf{H}} \mathbf{U}_{cp}$  is a  $NL_d \times ML_d$  block circulant matrix constructed by block circulations of the matrix  $[\mathbf{H}_0, \mathbf{H}_1, \dots, \mathbf{H}_\nu, 0, \dots, 0]^T$ .

For the ZP transmission, we can rewrite (72) as

$$\bar{\mathbf{y}}_{zp} = \mathbf{H}_e \hat{\mathbf{x}} + \bar{\mathbf{n}} \quad (78)$$

where  $\bar{\mathbf{y}}_{zp}$  is the  $N(L_d+L_e) \times 1$  received vector,  $\bar{\mathbf{x}}$  is the  $ML_d \times 1$  transmitted vector,  $\bar{\mathbf{n}}$  is the white Gaussian noise vector in  $\mathbb{C}^{N(L_d+L_e) \times 1}$  and  $\bar{\mathbf{H}}$  is the channel matrix given by

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{H}_0 & 0 & \cdots & 0 \\ \vdots & \mathbf{H}_1 & \ddots & \vdots \\ \mathbf{H}_\nu & \vdots & \ddots & \mathbf{H}_0 \\ 0 & \mathbf{H}_\nu & \ddots & \vdots \\ \vdots & \vdots & \vdots & \mathbf{H}_\nu \end{bmatrix}. \quad (79)$$

Assuming perfect CSIR and that the channel remains unchanged during the transmission of  $L_d + L_e$  vectors, the MMSE equalizer  $\mathbf{W}$  is applied to decouple the received streams (after

removing the  $L_e$  extension vectors in case of CP transmission). The MMSE equalizer is given by

$$\mathbf{W} = (\rho^{-1}\mathbf{I} + \mathbf{H}_e^H \mathbf{H}_e)^{-1} \mathbf{H}_e^H \quad (80)$$

and the unbiased decision-point SINRs of the equalizers output for detecting the  $k$ th transmitted stream are

$$\gamma_k = \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} - 1, \quad k = 1, \dots, ML_d. \quad (81)$$

In Sections VII and VIII, we analyze the outage diversity for the ZP and CP systems. The PEP analysis follows in a direct manner as in the flat-fading case so we omit it.

### B. ZP MMSE Receiver

It is known that in a point-to-point *single-antenna* ISI channel, linear receivers can achieve full multipath diversity under ZP transmission [20], [22], [23]. In this section, we investigate the similar question for MIMO systems whose receivers use linear MMSE operations in both the spatial and temporal dimensions. We provide lower and upper bounds on diversity. The bounds are not always tight, but the diversity is fully characterized for SIMO systems.

We begin by analyzing the tradeoff between the spectral efficiency  $R$  and the diversity of MMSE receiver in the *single-antenna* ISI channel  $d_{\text{MMSE}}^{\text{ISI}}$  under ZP transmission. Tajer and Nosratinia [10] shows that  $d_{\text{MMSE}}^{\text{ISI}}$  varies with  $R$  under CP transmission and MMSE equalization, in particular, for a quasi-static single-antenna ISI channel with  $\nu + 1$  taps, the diversity of the SC-MMSE receiver under CP transmission is  $d_{\text{MMSE}}^{\text{CP}} = 1 + \min(\nu, \lfloor 2^{-R} L_d \rfloor)$ , where  $L_d$  is the transmission data block length. We show that the same is not true for ZP transmission.

*Lemma 5:* For a quasi-static single-antenna ISI channel with  $\nu + 1$  taps, the diversity of the SC-MMSE receiver under ZP transmission is  $d_{\text{MMSE}}^{\text{ZP}} = \nu + 1$  irrespective of  $R$ .

*Proof:* See Appendix A. ■

We proceed with lower and upper bounds on diversity for MIMO ISI channel.

1) *Diversity Upper Bound:* Applying the MMSE equalizer given by (80) to the received vector in (77), the effective mutual information between  $\hat{\mathbf{x}}$  and  $\mathbf{W}\bar{\mathbf{y}}$  is equal to the sum of mutual information of their components [5]

$$I(\hat{\mathbf{x}}, \mathbf{W}\bar{\mathbf{y}}) = \frac{1}{L_d} \sum_{k=1}^{ML_d} I(x_k, y_k).$$

Thus, the outage probability is given by

$$P_{\text{out}} = \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log(1 + \gamma_k) < R\right) \quad (82)$$

$$= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} < R\right) \quad (83)$$

$$\geq \mathbb{P}\left(M \log \frac{1}{ML_d} \sum_{k=1}^{ML_d} \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} < R\right) \quad (84)$$

where we have used Jensen's inequality as in Section III-B. Let the eigendecomposition of  $\mathbf{H}_e^H \mathbf{H}_e$  be given by  $\mathbf{H}_e^H \mathbf{H}_e = \mathbf{U}^H \Lambda \mathbf{U}$  where  $\mathbf{U}$  is unitary and  $\Lambda$  is a diagonal matrix that has the eigenvalues of the matrix  $\mathbf{H}_e^H \mathbf{H}_e$  on its diagonal. Let the eigenvalues of  $\mathbf{H}_e^H \mathbf{H}_e$  be given by  $\{\lambda_\ell\}$  with  $\lambda_1 \geq \lambda_2 \dots \geq \lambda_{ML_d}$ . Let the vector  $\mathbf{u}_k$  be the column  $k$  of the matrix  $\mathbf{U}$ ; we have

$$\begin{aligned} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk} &= \mathbf{u}_k^H (\mathbf{I} + \rho \Lambda)^{-1} \mathbf{u}_k \\ &= \sum_{\ell=1}^{ML_d} \frac{|u_{\ell k}|^2}{1 + \rho \lambda_\ell} \\ &\stackrel{\triangle}{=} S_k. \end{aligned}$$

Let  $\bar{k} = \arg \min_k S_k$ . We can bound the sum in (84)

$$\begin{aligned} \frac{1}{ML_d} \sum_{k=1}^{ML_d} \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} &= \frac{1}{ML_d} \sum_{k=1}^{ML_d} \frac{1}{S_k} \\ &\leq \frac{1}{\min_k S_k} \\ &= \frac{1}{S_{\bar{k}}}. \end{aligned} \quad (85)$$

Thus, the outage bound in (84) can be further bounded

$$\begin{aligned} P_{\text{out}} &\geq \mathbb{P}\left(M \log \frac{1}{ML_d} \sum_{k=1}^M \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} < R\right) \\ &\geq \mathbb{P}\left(M \log \frac{1}{S_{\bar{k}}} < R\right) \\ &= \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}}\right). \end{aligned} \quad (86)$$

We now bound (86) by conditioning on the event

$$\mathcal{B} \stackrel{\triangle}{=} \left\{ |u_{\ell \bar{k}}|^2 \geq \frac{a}{M}, \ell = ML_d - M + 1, \dots, ML_d \right\} \quad (87)$$

where  $a$  is a positive real number that is slightly smaller than one  $a = 1 - \epsilon_1$ , and  $\epsilon_1$  is a small positive number. We then have

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}}\right) \\ &\geq \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \mid \mathcal{B}\right) \mathbb{P}(\mathcal{B}) \\ &= \mathbb{P}\left(\sum_{\ell=1}^{ML_d} \frac{|u_{\ell \bar{k}}|^2}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}} \mid \mathcal{B}\right) \mathbb{P}(\mathcal{B}) \\ &\geq \mathbb{P}\left(\sum_{\ell=ML_d-M+1}^{ML_d} \frac{|u_{\ell \bar{k}}|^2}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}} \mid \mathcal{B}\right) \mathbb{P}(\mathcal{B}) \end{aligned} \quad (88)$$

$$\geq \mathbb{P}\left(\frac{1}{M} \sum_{\ell=ML_d-M+1}^{ML_d} \frac{a}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}}\right) \mathbb{P}(\mathcal{B}) \quad (89)$$

$$= \mathbb{P}\left(\sum_{\ell=ML_d-M+1}^{ML_d} \frac{1}{1 + \rho \lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}}\right) \quad (90)$$

where (88) follows by removing some of the elements of the sum corresponding to the largest eigenvalues. The steps used to obtain (89) are similar to the steps used in Section III-B.

Note that  $\mathbf{H}_e^H \mathbf{H}_e$  is not a Wishart matrix; hence, the analysis of Section II does not directly apply here. The block diagonal elements of  $\mathbf{H}_e^H \mathbf{H}_e$  are similar and are given by

$$\mathbf{D} = \sum_{i=0}^{\nu} \mathbf{H}_i^H \mathbf{H}_i. \quad (91)$$

The matrix  $\mathbf{H}_e^H \mathbf{H}_e$  is Toeplitz and Hermitian. Moreover, the matrix  $\mathbf{D}$  given by (91) is a Wishart matrix.<sup>2</sup>

Observe that the probability in (90) depends on the  $M$  smallest eigenvalues. We now bound these eigenvalues with the eigenvalues of the matrix  $\mathbf{D}$  via the Sturmian separation theorem [24, P.1077].

*Theorem 4 (Sturmian Separation Theorem):* Let  $\{\mathbf{A}_r, r = 1, 2, \dots\}$  be a sequence of symmetric  $r \times r$  matrices such that each  $\mathbf{A}_r$  is a submatrix of  $\mathbf{A}_{r+1}$ . Then, if  $\{\lambda_k(\mathbf{A}_r), k = 1, \dots, r\}$  denote the ordered eigenvalues of each matrix  $\mathbf{A}_r$  in descending order, we have

$$\lambda_{k+1}(\mathbf{A}_{i+1}) \leq \lambda_k(\mathbf{A}_i) \leq \lambda_k(\mathbf{A}_{i+1}).$$

For our purposes, we consider a special case of the Sturmian Theorem by constructing a set of matrices  $\mathbf{A}_M, \mathbf{A}_{M+1}, \dots, \mathbf{A}_{L_d M}$  starting by the largest one  $\mathbf{A}_{L_d M} \triangleq \mathbf{H}_e^H \mathbf{H}_e$  and making all other matrices  $\mathbf{A}_i$  to be (successively embedded)  $i \times i$  principal submatrices of  $\mathbf{H}_e^H \mathbf{H}_e$ , such that the smallest matrix is  $\mathbf{A}_M = \mathbf{D}_{L_d}$ . Then, we repeatedly apply the first inequality in the Sturmian to get

$$\begin{aligned} \lambda_{ML_d}(\mathbf{A}_{ML_d}) &\leq \lambda_{ML_d-1}(\mathbf{A}_{ML_d-1}) \leq \dots \leq \lambda_M(\mathbf{A}_M) \\ \lambda_{ML_d-1}(\mathbf{A}_{ML_d}) &\leq \lambda_{ML_d-2}(\mathbf{A}_{ML_d-1}) \leq \dots \leq \lambda_{M-1}(\mathbf{A}_M) \\ &\vdots &&\vdots \\ \lambda_{ML_d-M+1}(\mathbf{A}_{ML_d}) &\leq \lambda_{ML_d-M}(\mathbf{A}_{ML_d-1}) \leq \dots \leq \lambda_1(\mathbf{A}_M). \end{aligned}$$

This implies that the smallest  $M$  eigenvalues of  $\mathbf{H}_e^H \mathbf{H}_e$  are bounded above by the  $M$  eigenvalues of  $\mathbf{D}$ , respectively. Hence

$$P_{\text{out}} \geq \mathbb{P}\left(\sum_{k=1}^M \frac{1}{1 + \rho \lambda_k(\mathbf{D})} > \frac{M}{a} 2^{-\frac{R}{M}}\right). \quad (92)$$

$\mathbf{D}$  is a sum of  $(\nu + 1)$  central Wishart matrices each with  $N$  degrees of freedom and with identity covariance matrix, i.e.,  $\mathbf{D} \in \mathcal{W}((\nu + 1)N, I)$ . Therefore, the analysis of Section II applies here and we have the following lemma.

*Lemma 6:* In a MIMO quasi-static frequency-selective system (with channel memory  $\nu$ ) consisting of  $M$  transmit and

<sup>2</sup>Let  $\mathcal{W}(n, \Sigma)$  denote a Wishart distribution with degree of freedom  $n$  and covariance (also called scale) matrix  $\Sigma$ . Any of the diagonal block matrices  $\mathbf{D}_j$  given by (91) follows a Wishart distribution since if  $\mathbf{B}_1 \in \mathcal{W}(n_1, \Sigma)$  and  $\mathbf{B}_2 \in \mathcal{W}(n_2, \Sigma)$  then  $\mathbf{B}_1 + \mathbf{B}_2 \in \mathcal{W}(n_1 + n_2, \Sigma)$ .

$N$  receive antennas, the MMSE receiver diversity under joint spatial encoding and ZP transmission is upper bounded as

$$\begin{aligned} d^{\text{ZP}} &\leq \left[ (M 2^{-\frac{R}{M}} + 1 - (M - N)^+)^+ \right]^2 \\ &+ |(\nu + 1)N - M| \left[ (M 2^{-\frac{R}{M}} + 1 - (M - N)^+)^+ \right]. \end{aligned} \quad (93)$$

*2) Diversity Lower Bound:* We can upper bound the outage probability as follows:

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log(1 + \gamma_k) < R\right) \\ &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \\ &\leq \mathbb{P}\left(M \log \frac{1}{ML_d} \sum_{k=1}^{ML_d} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \quad (94) \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^{ML_d} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \\ &= \mathbb{P}\left(\sum_{k=1}^{ML_d} \frac{1}{1 + \rho \lambda_k(\mathbf{H}_e^H \mathbf{H}_e)} > M 2^{-\frac{R}{M}}\right) \\ &\leq \mathbb{P}\left(\sum_{k=1}^M \frac{1}{1 + \rho \lambda_k(\mathbf{H}_e^H \mathbf{H}_e)} + L_d M - M > M 2^{-\frac{R}{M}}\right) \quad (95) \end{aligned}$$

$$= \mathbb{P}\left(\sum_{k=1}^M \frac{1}{1 + \rho \lambda_k(\mathbf{H}_e^H \mathbf{H}_e)} > M 2^{-\frac{R}{M}} - (ML_d - M)\right) \quad (96)$$

where (94) follows from Jensen's inequality and (95) follows from setting the smallest  $L_d M - M$  eigenvalues to zero.

Now we repeatedly use the second inequality in the Sturmian theorem to get

$$\begin{aligned} \lambda_M(\mathbf{A}_M) &\leq \dots \leq \lambda_M(\mathbf{A}_{ML_d-1}) \leq \lambda_M(\mathbf{A}_{ML_d}) \\ \lambda_{M-1}(\mathbf{A}_M) &\leq \dots \leq \lambda_{M-1}(\mathbf{A}_{ML_d-1}) \leq \lambda_{M-1}(\mathbf{A}_{ML_d}) \\ &\vdots &&\vdots \\ \lambda_1(\mathbf{A}_M) &\leq \dots \leq \lambda_1(\mathbf{A}_{ML_d-1}) \leq \lambda_1(\mathbf{A}_{ML_d}) \end{aligned}$$

with  $\mathbf{A}_{ML_d} \triangleq \mathbf{H}_e^H \mathbf{H}_e$  and  $\mathbf{A}_M \triangleq \mathbf{D}$ , similar to the earlier case. Therefore, the largest  $M$  eigenvalues of  $\mathbf{H}_e^H \mathbf{H}_e$  are bounded below by the  $M$  eigenvalues of  $\mathbf{D}$ , respectively. Therefore

$$P_{\text{out}} \leq \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^M \frac{1}{1 + \rho \lambda_k(\mathbf{D})} > Q\right) \quad (97)$$

where  $Q = \max(0, M 2^{-\frac{R}{M}} - (ML_d - M))$ . Recall that  $\mathbf{D}$  is a Wishart matrix; therefore, the analysis of Section II follows and we obtain the following lemma.

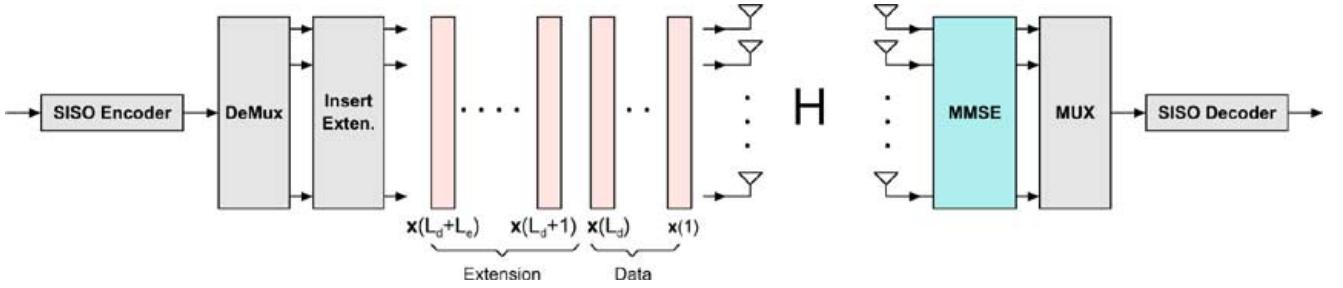


Fig. 3. SC block transmission in a frequency-selective channel. In the case of CP, the extension is removed at the receiver prior to equalization.

*Lemma 7:* In a MIMO quasi-static frequency-selective system (with channel memory  $\nu$ ) consisting of  $M$  transmit and  $N$  receive antennas, the MMSE receiver diversity is lower bounded as

$$d^{ZP} \geq [Q]^2 + |(\nu + 1)N - M| [Q] \quad (98)$$

under joint spatial encoding and ZP transmission.  $Q = \max(0, M2^{-\frac{R}{M}} - (ML_d - M))$ .

*Remark 1:* Notice that both lower and upper bounds differ only in the second term of  $Q$ , i.e.,  $(ML_d - M)$ . The diversity lower bound for  $L_d = 1$  is tight against the upper bound, but for  $L_d > 1$  the lower bound (98) is trivial.

### C. CP MMSE Receiver

For the single-antenna ISI channel under CP transmission, the explicit tradeoff between spectral efficiency and diversity was found [10] to be  $d_{\text{MMSE}}^{\text{CP}} = 1 + \min(\nu, \lfloor 2^{-R} L_d \rfloor)$ . In this section, we extend the analysis to the MIMO case. The system model is shown in Fig. 3. We start with the general  $M \times N$  MIMO system.

The system model is again given by (77) where  $\mathbf{H}_e = \bar{\mathbf{H}} \mathbf{U}_{cp}$  and  $\hat{\mathbf{x}}$  is generated by taking the IDFT of the information vector  $\mathbf{x}$  [25], i.e.,

$$\hat{\mathbf{x}} = \mathbf{Q}_{Tx}^H \mathbf{x} \quad (99)$$

where  $\mathbf{Q}_{Tx}$  is the augmented DFT matrix given by  $\mathbf{Q}_{Tx} = \mathbf{Q} \otimes \mathbf{I}_M$ , where  $\mathbf{I}_M$  is the identity matrix,  $\mathbf{Q}$  is the normalized DFT matrix, and  $\otimes$  is the Kroenecker product.

The  $NL_d \times ML_d$  block-circulant matrix  $\mathbf{H}_e$  has eigendecomposition  $\mathbf{H}_e = \mathbf{Q}_{Rx}^H \Lambda \mathbf{Q}_{Tx}$ , where  $\mathbf{Q}_{Rx} = \mathbf{Q} \otimes \mathbf{I}_N$ . Both  $\mathbf{Q}_{Tx}$  and  $\mathbf{Q}_{Rx}$  are unitary matrices. The block diagonal matrix  $\Lambda$  is given by

$$\Lambda = \begin{pmatrix} \mathbf{B}_1 & & & 0 \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ 0 & & & \mathbf{B}_{L_d} \end{pmatrix} \quad (100)$$

where the matrix  $\mathbf{B}_k$  is given by [26]

$$\mathbf{B}_k = \sum_{i=0}^{\nu} \mathbf{H}_i e^{-j \frac{2\pi i (k-1)}{L_d}}, \quad \text{for } k = 1, \dots, L_d \quad (101)$$

and  $\mathbf{H}_i$  is the instantaneous MIMO channel (cf., Section VI-A).

Analogous to the proof of [10], we first consider the case where the transmission data block length is equal to the number of channel taps, i.e.,  $L_d = \nu + 1$ . In this case, the entries of  $\mathbf{B}_k$ 's are i.i.d. normal complex Gaussian.

*1) Outage Upper Bound:* The outage probability of the MMSE receiver is given by

$$P_{\text{out}} = \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log\left(\frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1}}\right) < R\right) \quad (102)$$

$$\begin{aligned} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log((\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1}) > -R\right) \\ &\leq \mathbb{P}\left(M \log \sum_{k=1}^{ML_d} \frac{1}{ML_d} (\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)_{kk}^{-1} > -R\right) \end{aligned} \quad (103)$$

$$= \mathbb{P}\left(M \log \sum_{k=1}^{ML_d} \frac{1}{ML_d} (\mathbf{I} + \rho \Lambda^H \Lambda)_{kk}^{-1} > -R\right) \quad (104)$$

$$\begin{aligned} &= \mathbb{P}\left(\sum_{k=1}^{ML_d} (\mathbf{I} + \rho \Lambda^H \Lambda)_{kk}^{-1} > ML_d 2^{-\frac{R}{M}}\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{L_d} \text{tr}(\mathbf{I} + \rho \mathbf{B}_i^H \mathbf{B}_i)^{-1} > ML_d 2^{-\frac{R}{M}}\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^M \frac{1}{(1 + \rho \lambda_{k,i})} > ML_d 2^{-\frac{R}{M}}\right) \end{aligned} \quad (105)$$

where (103) follows from Jensen's inequality, (104) follows from the eigendecomposition of  $\mathbf{H}_e$ , and  $\lambda_{k,i}$  is  $k$ th eigenvalue of the  $i$ th Wishart matrix  $\mathbf{B}_i^H \mathbf{B}_i$ .

Recall from Section III that the eigenvalues of a Wishart matrix have the asymptotic property

$$\sum_{k=1}^M \frac{1}{1 + \rho \lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \quad (106)$$

based on which we established in Lemmas 1 and 2 the following

$$\mathbb{P}\left(\sum_{k=1}^M \frac{1}{1 + \rho \lambda_k} \geq s\right) \doteq \rho^{-(s^2 + |N-M|s)} \quad (107)$$

where  $\alpha_k$  is defined in (15) and  $s$ ,  $M$ , and  $N$  are arbitrary integers. Define

$$\theta_i \triangleq \sum_{\alpha_{k,i} > 1} 1$$

$\theta_i$  are i.i.d. discrete random variables with the following asymptotic distribution [cf., Section III, (22)–(26)]

$$\mathbb{P}(\theta_i = n_i) \doteq \rho^{-(n_i^2 + |N-M|n_i)}, \quad \text{for } n_i = 1, \dots, M. \quad (108)$$

Using (107), the outage probability in (105) can be evaluated as

$$\begin{aligned} P_{\text{out}} &\dot{\leqslant} \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^M \frac{1}{(1+\rho\lambda_{k,i})} > ML_d 2^{-\frac{R}{M}}\right) \\ &\doteq \mathbb{P}\left(\sum_{i=1}^{L_d} \theta_i \geq \Omega\right) \end{aligned} \quad (109)$$

where  $\Omega = \lceil ML_d 2^{-\frac{R}{M}} \rceil$ . Evaluating the probability in (109) in a combinatorial manner, we get

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{L_d} \theta_i \geq \Omega\right) &\doteq \mathbb{P}\left(\sum_{i=1}^{L_d} \theta_i = \Omega\right) \\ &\doteq \sum_{n_1, n_2, \dots, n_{L_d}} \rho^{-(n_1^2 + |N-M|n_1)} \dots \rho^{-(n_{L_d}^2 + |N-M|n_{L_d})} \\ &\doteq \max_{n_1, n_2, \dots, n_{L_d}} \rho^{-(n_1^2 + |N-M|n_1)} \dots \rho^{-(n_{L_d}^2 + |N-M|n_{L_d})} \end{aligned} \quad (110)$$

where  $n_i \in [0, M]$  for  $(i = 1, 2, \dots, L_d)$  is the value of the  $i$ th discrete random variable  $\theta_i$ , and (111) is true since the summation in (110) is dominated by the maximum element.

Let the set  $\{n_k^*, k = 1, \dots, L_d\}$  be the set of indices of the optimal solution of (111). The set  $\{n_k^*\}$  is obtained by solving the following optimization problem:

$$\begin{aligned} \min_{n_1, n_2, \dots, n_{L_d}} &\sum_{k=1}^{L_d} (n_k^2 + |N-M|n_k) \\ \text{subject to } &\sum_{k=1}^{L_d} n_k = \Omega \\ &0 \leq n_k \leq M \end{aligned}$$

or equivalently

$$\begin{aligned} \min_{n_1, n_2, \dots, n_{L_d}} &\sum_{k=1}^{L_d} n_k^2 \\ \text{subject to } &\sum_{k=1}^{L_d} n_k = \Omega \\ &n_k \geq 0. \end{aligned} \quad (112)$$

The problem in (112) is a quadratic integer-programming (QIP) problem (see, e.g., [27]). Integer programming problems are in general NP-hard. However, due to the simple structure of the objective function in (112), we can efficiently solve it, thus obtain a closed-form expression for  $\{n_k^*\}$  and hence (111).

*Lemma 8:* For the QIP given by (112), the optimum solution is given by

$$\begin{aligned} n_i^* &= u, \quad \text{for } 1 \leq i \leq t \\ n_j^* &= u+1, \quad \text{for } t+1 \leq j \leq L_d \end{aligned}$$

where  $u = \lfloor \frac{\Omega}{L_d} \rfloor$  and  $t = L_d(u+1) - \Omega$ .

*Proof:* See Appendix B ■

Using Lemma 8, we can now evaluate the outage upper bound given by (111) as

$$P_{\text{out}} \dot{\leqslant} \rho^{-d_{cp}} \quad (113)$$

where  $d_{cp} = \Omega(2u+1) - uL_d(u+1) + |N-M|\Omega$  and  $u = \lfloor \frac{\Omega}{L_d} \rfloor$ .

2) *Outage Lower Bound:* The bound is obtained using the same steps to obtain the lower bound in Section VI-B1. It can be shown that

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log((\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}) > -R\right) \\ &\geq \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^M \frac{1}{(1+\rho\lambda_{k,i})} > ML_d 2^{-\frac{R}{M}}\right). \end{aligned} \quad (114)$$

$$\geq \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^M \frac{1}{(1+\rho\lambda_{k,i})} > ML_d 2^{-\frac{R}{M}}\right). \quad (115)$$

The bound in (115) is the same as the upper bound in (105); thus, the bound is tight and the diversity is given by (113). The PEP analysis follows in a manner similar to Section IV.

Recall that so far we have considered data block length  $L_d = \nu + 1$ . It can be shown that the diversity for any  $L_d > \nu + 1$  is upper bounded by the computed diversity for the case  $L_d = \nu + 1$ . This bounding is derived from (104) via FFT arguments similar to those used in [10], which we omit for brevity. A tight diversity *lower* bound for data block lengths  $L_d > \nu + 1$  remains an open problem, except for the SIMO system, as discussed in Section VII.

### 3) Diversity of CP Transmission in the SIMO Channel:

*Theorem 5:* In a SIMO quasi-static frequency-selective channel with memory  $\nu$ ,  $N$  receive antennas, and data block length  $L_d$ , the MMSE receiver diversity is  $d_{MMSE}^{CP} = N \min(\nu + 1, \lfloor 2^{-R} L_d \rfloor + 1)$  under joint spatial encoding and CP transmission.

In order to prove Theorem 5, we first analyze the case of  $L_d = \nu + 1$  and then generalize the result for  $L_d > \nu + 1$ . The system model is given by (77) where the  $NL_d \times L_d$  equivalent channel matrix is given by

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_{\nu} & 0 & \cdots & 0 \\ 0 & \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_{\nu} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_{\nu} & 0 & \cdots & \mathbf{h}_0 \end{bmatrix} \quad (116)$$

where  $\mathbf{h}_i$  (for  $i = 0, 1, \dots, \nu$ ) is  $N \times 1$  SIMO channel. Note that the diagonal elements of  $(\mathbf{H}_e^H \mathbf{H}_e)$  are identical and equal to  $\sum_{i=0}^{\nu} \mathbf{h}_i^H \mathbf{h}_i$ . Thus, the MMSE SINR for each output information stream is

$$\gamma_k = \frac{1}{(\mathbf{I} + \mathbf{H}_e^H \mathbf{H}_e)_{kk}} - 1 = \frac{1}{\frac{1}{L_d} \text{tr}(\mathbf{I} + \mathbf{H}_e^H \mathbf{H}_e)_{kk}} - 1. \quad (117)$$

Evaluating the outage probability as in (102)

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{L_d} \log\left(\frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}}\right) < R\right) \\ &= \mathbb{P}\left(\log \frac{1}{L_d} \sum_{k=1}^{L_d} \frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}} < R\right) \quad (118) \\ &= \mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{(1+\rho\lambda_k)} > L_d 2^{-R}\right) \quad (119) \end{aligned}$$

where (118) follows from (117) and (119) follows similarly to (105).

In a manner similar to (105), we have  $\lambda_k = \mathbf{B}_k^H \mathbf{B}_k$  because now  $\mathbf{B}$  is simply a  $N \times 1$  vector. For the case  $L_d = \nu + 1$ , the eigenvalues  $\{\lambda_k\}$  are distributed according to Gamma distribution with shape parameter  $N$  and scale parameter 1, i.e.,  $\lambda_k \sim \Gamma(N, 1)$ . For  $L_d > \nu + 1$ , the Gaussian variables in  $\mathbf{B}_k$  are no longer independent and thus analyzing this case requires the unknown distribution  $\{\lambda_k\}$ . Instead, we indirectly show that the diversity of  $L_d = \nu + 1$  also holds for  $L_d > \nu + 1$ .

*Lemma 9:* In a SIMO quasi-static frequency-selective channel with memory  $\nu$ ,  $N$  receive antennas and data block length  $L_d = \nu + 1$ , the MMSE receiver diversity is  $d_{MMSE}^{CP} = N(\lfloor L_d 2^{-R} \rfloor + 1)$  under joint spatial encoding and CP transmission.

*Proof:* The outage probability can be written as

$$P_{\text{out}} = \mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{(1 + \rho \lambda_k)} > L_d 2^{-R}\right) \doteq \mathbb{P}(M(\alpha) > L_d 2^{-R}) \quad (120)$$

where we use  $M(\alpha) = \sum_{\alpha k > 1} 1$  from (106). We thus need to evaluate  $\mathbb{P}(\alpha > 1)$ . The probability density function of  $\lambda_k$  is

$$f_{\lambda_k}(x) = \frac{1}{\Gamma(N)} x^{N-1} e^{-x}. \quad (121)$$

The distribution of  $\alpha_k$  is thus given by

$$f_{\alpha_k}(x) = \frac{1}{\Gamma(N)} \rho^{-Nx} e^{\rho^{-x}} \ln \frac{1}{\rho}. \quad (122)$$

The cumulative distribution function of  $\alpha_k$  is

$$F_{\alpha_k}(x) = \int_{-\infty}^x f_{\alpha_k}(y) dy = \frac{1}{\Gamma(N)} \int_{\rho^{-x}}^{\infty} r^{N-1} e^{-r} dr \quad (123)$$

$$= \frac{1}{\Gamma(N)} \left( \int_0^{\infty} r^{N-1} e^{-r} dr - \int_0^{\rho^{-x}} r^{N-1} e^{-r} dr \right) \quad (124)$$

$$= e^{-\rho^{-x}} \sum_{k=0}^{N-1} \frac{\rho^{-xk}}{k!} \quad (125)$$

where we have made a change of variables  $r = \rho^{-x}$  in (123), and evaluate the integral according to [24, P.334 and P.336]. Thus, we have

$$P(\alpha_k > 1) = 1 - e^{-\rho} \sum_{k=0}^{N-1} \frac{\rho^{-k}}{k!} \doteq 1 - \left(1 - \frac{1}{N!} \rho^{-N}\right) \quad (126)$$

$$\doteq \rho^{-N} \quad (127)$$

where (126) follows from the Taylor expansion for (125).

From the independence of  $\{\lambda_k\}$ , and subsequently the independence of  $\{\alpha_k\}$ , we conclude that  $M(\alpha)$  in (120) is binomially distributed with parameter  $\rho^{-N}$ . Hence, similar to [10], we have

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{1 + \rho \lambda_k} > L_d 2^{-R}\right) &\doteq \mathbb{P}(M(\alpha) > L_d 2^{-R}) \\ &= \sum_{i=\lfloor L_d 2^{-R} \rfloor + 1}^{L_d} \mathbb{P}(M(\alpha) = i) \\ &\doteq \sum_{i=\lfloor L_d 2^{-R} \rfloor + 1}^{L_d} \binom{L_d}{i} \rho^{-Ni} \underbrace{(1 - \rho^{-N})^{n-i}}_{\doteq 1} \\ &\doteq \rho^{-N(\lfloor L_d 2^{-R} \rfloor + 1)} \end{aligned}$$

which concludes the proof for  $L_d = \nu + 1$  ■

For  $L_d > \nu + 1$ , we follow steps similar to [10].

*Lemma 10:* Consider two SIMO systems both operating under quasi-static frequency-selective channels with memory  $\nu$ . One system has data block length  $L_{d_1} > \nu + 1$  and the other  $L_{d_2} \geq L_{d_1}$ , we have the following property:

$$\mathbb{P}\left(\sum_{k=1}^{L_{d_1}} \frac{1}{(1 + \rho \lambda_k)} > m\right) \doteq \mathbb{P}\left(\sum_{k=1}^{L_{d_2}} \frac{1}{(1 + \rho \lambda_k)} > m\right)$$

for any  $m \in \mathbb{R}$ .

*Proof:* The proof has similarities with the SISO case developed in [10, Lemma 2], but is not a trivial extension (see Appendix C). ■

Using Lemma 10 and the results in [10, Th. 2], Theorem 5 is established.

## VII. SIMULATION RESULTS

Simulations generate Monte Carlo random channel realizations and calculate outage probability by checking the appropriate linear MIMO receiver mutual information for the quasi-static flat-fading model. Fig. 4 shows the case  $M = N = 3$ . According to Theorem 2,  $d_{\text{out}} = 1$  for  $R \geq 4.755$ ,  $d_{\text{out}} = 4$  for  $4.755 > R \geq 1.755$ , and  $d_{\text{out}} = 9$  for  $R < 1.7549$ . Fig. 4 shows the diversity step between  $R = 4.5$  and  $4.8$  b/s/Hz. The slope of diversity 9 is difficult to measure precisely with simulations, but it is approximately observed. Fig. 5 shows the outage probability for  $R = 1, 4$ , and  $10$  with the Jensen bound, with a diversity transition at  $R = 2$ . Fig. 6 shows the case of  $M = 2$ , and  $N = 3$  again with transition at  $R = 2$ . In Fig. 7, simulations results for  $N = 2$  and  $M = 3$  are given and compared with  $N = 3$  and  $M = 2$ . Theorem 2 gives the diversity for both systems. It is observed that when  $N > M$ , the break point of the slopes occurs before its counterparts in  $M > N$  case. Lower rates were difficult to simulate precisely.

## VIII. CONCLUSION

This paper settles the long standing problem of the diversity of the MMSE MIMO receivers under all fixed rates for any number of transmit ( $M$ ) and receive ( $N$ ) antennas, giving the

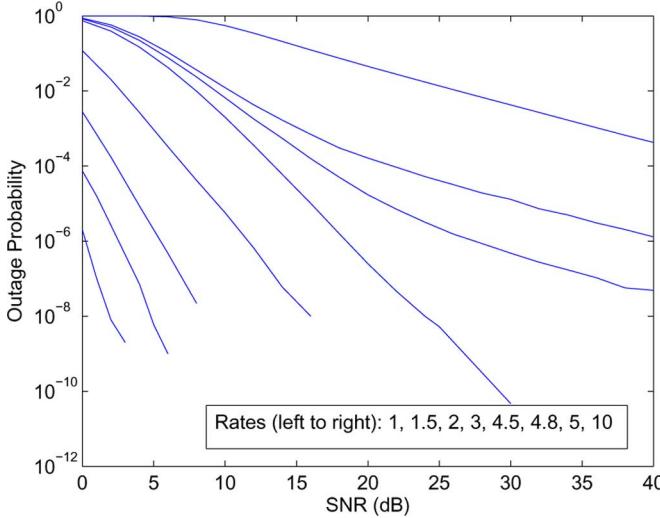


Fig. 4. Outage probability of MMSE Receiver,  $M = N = 3$  for  $R = 1, 1.5, 2, 3, 4.5, 4.8, 5, 10$  b/s/Hz.

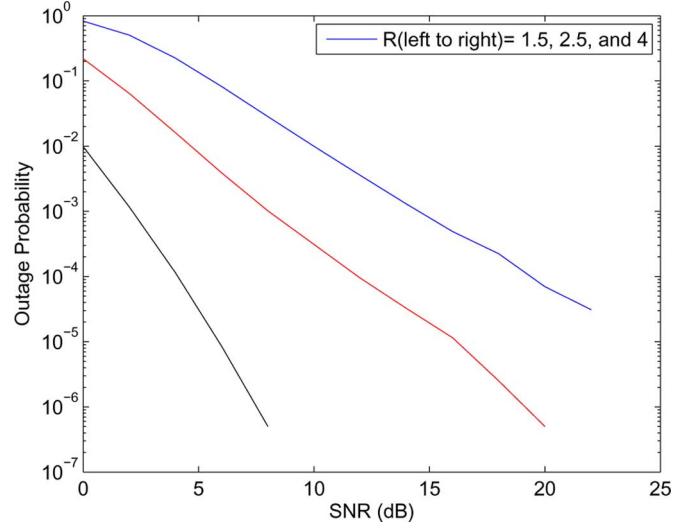


Fig. 6. Outage probability of MMSE Receiver,  $M = 2$ , and  $N = 3$  for  $R$  (left to right) = 1.5, 2.5, 4 b/s/Hz.

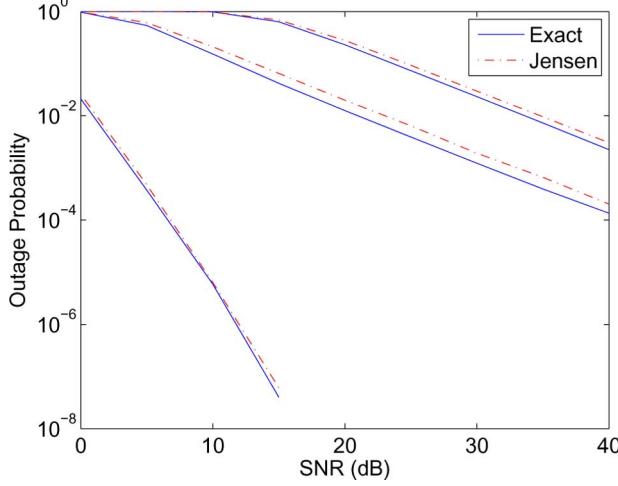


Fig. 5. Outage probability of MMSE Receiver,  $M = N = 2$  for  $R$  (left to right) = 1, 4, 10 b/s/Hz.

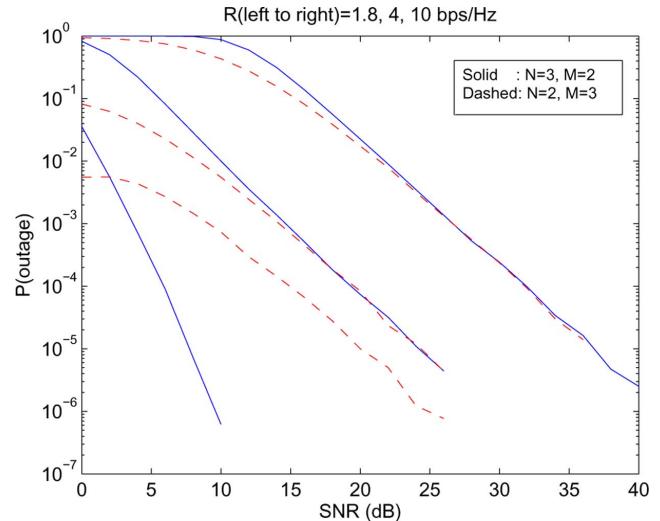


Fig. 7. Outage probability of MMSE Receiver for both cases  $N > M$  (solid) and  $M > N$  (dashed). The spectral efficiency  $R$  (left to right) = 1.8, 4, and 10 b/s/Hz.

result as  $d = \lceil M2^{-\frac{R}{M}} - (M - N)^+ \rceil^2 + |N - M| \lceil M2^{-\frac{R}{M}} - (M - N)^+ \rceil$ . The analysis confirms the earlier approximate results [6], [7] showing that the system diversity can be as high as  $MN$  for low spectral efficiency and as low as  $N - M + 1$  for high spectral efficiency. The result is extended to the MAC. We also analyze the case of frequency-selective MIMO channel under CP and ZP transmission, and obtain the explicit tradeoff between rate and diversity.

## APPENDIX

### A. Proof of Lemma 5

Consider a single-antenna ISI channel  $\mathbf{h} = [h_0, \dots, h_\nu]$ , where  $\nu$  is channel memory. The transmitter sends a block of  $L_d + \nu$  symbols (i.e., the extension  $L_e = \nu$ ), the last  $\nu$  symbols of which are zeros to remove the interblock interference. The system model is given by

$$\mathbf{y} = \mathbf{H}_e \mathbf{x} + \mathbf{n} \quad (128)$$

where  $\mathbf{x}$  is the transmitted length- $(L_d + \nu)$  vector. We consider the case where the padding length is equal to the memory of the channel. The results are also valid for  $L_e > \nu$  as a direct result of [10, Th. 2].

The outage probability of MMSE receiver under ZP transmission is given by [10]

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{L_d} \log\left(\frac{1}{(\mathbf{I} + \rho \mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}}\right) < R\right) \\ &\leq \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{L_d} \log\left(1 + \frac{\rho}{(\mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk}}\right) < R\right) \end{aligned} \quad (129)$$

$$\leq \mathbb{P}\left(\log \frac{1}{L_d} \sum_{k=1}^{L_d} \frac{1}{\rho} (\mathbf{H}_e^H \mathbf{H}_e)^{-1}_{kk} > -R\right) \quad (130)$$

$$= \mathbb{P}\left(\frac{L_d 2^{-R}}{\text{tr}(\mathbf{H}_e^H \mathbf{H}_e)^{-1}} < \rho^{-1}\right) \quad (131)$$

where (129) represents the outage probability of ZF equalizer which upper bounds that of the MMSE. The bound in (130) follows from Jensen's inequality.

We want to show that  $\text{tr}(\mathbf{H}_e^H \mathbf{H}_e)^{-1}$  in (131) is proportional to  $\|\mathbf{h}\|^{-2}$ . Thus, it is straightforward to obtain full diversity at any  $R$  since [15]

$$\mathbb{P}(c\|\mathbf{h}\|^2 < \rho^{-\alpha}) \doteq \rho^{-L\alpha} \quad (132)$$

where  $c$  is a constant that is independent of  $\mathbf{h}$ .

To show that this is indeed the case, we use the result of Tepedelenlioglu [22], [28] which provides a family of linear ZF equalizers that is capable of achieving full multipath diversity in zero-padded systems under certain constraints. We paraphrase the result for convenience.

*Lemma 11* [22], [28]: Under zero-padded transmission, there exists a family of left-inverses of  $\mathbf{H}_e$ , denoted by  $\mathbf{G}$ , such that  $\|\mathbf{G}\|^{-2} \geq C\|\mathbf{h}\|^2$  for some constant  $C$  independent of the channel vector  $\mathbf{h}$ . Moreover, we have  $\|\mathbf{W}_{ZF}\| \leq \|\mathbf{G}\|$ , for any  $\mathbf{G}$  satisfying  $\mathbf{G}\mathbf{H}_e = \mathbf{I}$ , and  $\mathbf{W}_{ZF}$  is given by

$$\mathbf{W}_{ZF} = (\mathbf{H}_e^H \mathbf{H}_e)^{-1} \mathbf{H}_e^H. \quad (133)$$

Applying the ZF equalizer  $\mathbf{W}_{ZF}$  on the channel output given by (128), we get the equalized signal  $\tilde{\mathbf{y}} = \mathbf{x} + \mathbf{z}$ , where  $\mathbf{z} = \mathbf{W}_{ZF}\mathbf{n}$ . The filtered noise power  $P_z$  can be evaluated as

$$\begin{aligned} P_z &= \mathbb{E}\text{tr}[\mathbf{z}\mathbf{z}^H] \\ &= \text{tr}[\mathbb{E}((\mathbf{H}_e^H \mathbf{H}_e)^{-1} \mathbf{H}_e^H \mathbf{n} \mathbf{n}^H \mathbf{H}_e (\mathbf{H}_e^H \mathbf{H}_e)^{-1})] \\ &= \text{tr}[(\mathbf{H}_e^H \mathbf{H}_e)^{-1}] \end{aligned} \quad (134)$$

where we assume the noise is uncorrelated and has variance equal to 1.

Using the properties of the Frobenius norm,  $P_z$  can be bounded as

$$\begin{aligned} P_z &= \mathbb{E}(\|\mathbf{W}_{zf}\mathbf{n}\|^2) \\ &\leq \mathbb{E}(\|\mathbf{W}_{zf}\|^2 \|\mathbf{n}\|^2) = L_d \|\mathbf{W}_{zf}\|^2. \end{aligned} \quad (135)$$

Using (134), (135), and Lemma 11, the trace in (131) can be bounded by

$$\text{tr}[(\mathbf{H}_e^H \mathbf{H}_e)^{-1}] \leq L_d \|\mathbf{W}_{zf}\|^2 \leq \frac{L_d}{C\|\mathbf{h}\|^2}. \quad (136)$$

Thus, from (131), we have

$$\begin{aligned} P_{out} &\leq \mathbb{P}(C_2\|\mathbf{h}\|^2 < \rho^{-1}) \\ &\doteq \rho^{-(\nu+1)}. \end{aligned} \quad (137)$$

where  $C_2 = C 2^{-R}$  is a constant independent of  $\mathbf{h}$  and  $\rho$ .

Note that the constraints and construction methods in [22] and [28] for the ZF equalizers to achieve full multipath diversity in ZP systems do not apply in CP systems. That is, Lemma 11 is not true for CP transmission. This is because the equivalent channel

in CP systems does not have the same properties that were used in [22] and [28].

### B. Proof of Lemma 8 (QIP Problem)

Consider the following QIP problem:

$$\begin{aligned} \min_{n_1, n_2, \dots, n_\ell} \quad & \sum_{k=1}^{\ell} n_k^2 \\ \text{subject to} \quad & \sum_{k=1}^{\ell} n_k = \Omega \\ & n_k \geq 0. \end{aligned} \quad (138)$$

where  $\Omega$  and  $\ell$  are integers.

Consider a candidate solution vector  $[n_1, \dots, n_k, \dots, n_\ell]$ . We partition the variables in this vector according to their values into  $\Omega + 1$  sets  $\mathcal{N}_j = \{n_k : n_k = j\}$  for  $0 \leq j \leq \Omega$ ; clearly some of these sets may be empty. Denote the membership of each set  $S_j = |\mathcal{N}_j|$ . Furthermore, let  $\Omega = m\ell + K$  where  $m$  is the divisor and  $K$  is the remainder of the division of  $\Omega$  by  $\ell$ . From the constraint in (138), we have

$$\sum_{k=1}^{\ell} n_k = \sum_{j=0}^{\Omega} j S_j = m\ell + \sum_{j=0}^{\Omega} (j - m) S_j = m\ell + K. \quad (139)$$

Evaluating the objective function

$$\begin{aligned} \sum_{k=1}^{\ell} n_k^2 &= \sum_{j=0}^{\Omega} (m + j - m)^2 S_j \\ &= \ell m^2 + 2m \sum_{j=0}^{\Omega} (j - m) S_j + \sum_{j=0}^{\Omega} (j - m)^2 S_j \\ &= \ell m^2 + 2mK + \sum_{j=0}^{\Omega} (j - m)^2 S_j \end{aligned} \quad (140)$$

$$\begin{aligned} &\geq \ell m^2 + 2mK + \sum_{j=0}^{\Omega} (j - m) S_j \\ &= \ell m^2 + 2mK + K \end{aligned} \quad (141) \quad (142)$$

where (140) and (142) use  $\sum_{j=0}^{\Omega} (j - m) S_j = K$ , which follows from (139).

We now propose that one may achieve optimality when all variables take values either  $m$  or  $m + 1$ . In that case

$$\begin{aligned} \sum_k n_k &= mS_m + (m + 1)(\ell - S_m) = m\ell + (\ell - S_m) \\ \sum_k n_k^2 &= m^2 S_m + (m + 1)^2 (\ell - S_m) = \ell m^2 + 2mK + K. \end{aligned}$$

where we substituted the value of  $\ell - S_m$  from the first equation into the second equation above. This shows that the variables taking values  $m$  or  $m + 1$  achieves the lower bound in (142). At optimality,  $S_m = (m + 1)\ell - \Omega$ .

### C. Proof of Lemma 10

We begin by showing that for any integer multiplier of  $L_{d_1} = \nu + 1$  denoted by  $L_{d_2} = TL_{d_1}$  ( $T \in \mathbb{N}$ ) and any real-valued  $m \in (0, L_{d_1})$ , we have

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_1}} \frac{1}{(1+\rho\lambda_q)} > m\right) \doteq \mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) \quad (143)$$

Note that for SIMO-CP system,  $\lambda_q = \mathbf{b}_q^H \mathbf{b}_q$ , where  $\mathbf{b}_q$  is the  $N \times 1$  vector given by

$$\mathbf{b}_q^{(i)} = \sum_{n=0}^{\nu} \mathbf{h}_n e^{-j \frac{2\pi(n-1)}{L_{d_i}}} \quad \text{for } q = 1, \dots, L_{d_i} \quad (144)$$

where  $\mathbf{h}_n$  is the channel gain as a function of the tap index  $n$ , and the superscript  $i = 1, 2$  is used to distinguish the variables in two systems with data block lengths  $L_{d_1}$  and  $L_{d_2}$ .

Recall that we can take a  $L_{d_1}$ -point signal and apply a  $L_{d_2}$ -point DFT on it (after ZP), which will result in a resampling in the Fourier domain at  $L_{d_2}$  points. Following [10], we can write the explicit relationship between entries of  $\mathbf{b}^{(1)}$  and  $\mathbf{b}^{(2)}$  as

$$b_{q,l}^{(1)} = \sum_{i=1}^{L_{d_1}} b_{i,l}^{(2)} \psi_i, \quad q = 1, 2, \dots, L_{d_2} \text{ and } l = 1, 2, \dots, N \quad (145)$$

where

$$\psi_i = \frac{1}{L_{d_1}} \frac{1 - e^{-j \frac{(q-1)2\pi L_{d_1}}{L_{d_2}}}}{1 - e^{-j \left( \frac{2\pi(q-1)}{L_{d_2}} - \frac{2\pi(i-1)}{L_{d_1}} \right)}}.$$

Define  $\alpha_{q,l}^{(i)} = \frac{\log |b_{q,l}^{(i)}|^2}{\log \rho}$ . Note that  $b_{T(q-1),l}^{(1)} = b_{q,l}^{(2)}$  and  $\alpha_{T(q-1),l}^{(2)} = \alpha_{q,l}^{(1)}$  for  $q = 1, 2, \dots, L_{d_1}$  since  $L_{d_2} = TL_{d_1}$ . From (145), we have

$$|b_{q,l}^{(1)}|^2 = \sum_{i=1}^{L_{d_1}} |\psi_i|^2 |b_{i,l}^{(2)}|^2 + \underbrace{\sum_{i=1}^{L_{d_1}} \sum_{s=1}^{L_{d_1}} \psi_i \psi_s b_{i,l}^{(2)} b_{s,l}^{*(2)}}_{\triangleq \eta}. \quad (146)$$

We now analyze each part of the sum in (146). For the set of indices  $\mathcal{A} \triangleq \{i : i = T(k-1) + 1, k = 1, \dots, L_{d_1}\}$ , the coefficients  $\{\psi_i\}$  are nonzero constants, then  $|\psi_i|^2 |b_{i,l}^{(2)}|^2 \doteq |b_{i,l}^{(2)}|^2 \forall l$ . Noting that  $\eta$  must be real-valued, and defining  $\alpha_\eta \triangleq -\frac{\log |\eta|}{\log \rho}$ , (146) can be written as

$$\begin{aligned} \rho^{-\alpha_{q,l}^{(2)}} &\doteq \sum_{i=1}^{L_{d_1}} \rho^{-\alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta} \\ &\doteq \rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta}. \end{aligned} \quad (147)$$

Note that if  $\eta < 0$  the second term in (147) should be smaller than the first term since otherwise the right-hand side of (147)

will be negative while the left-hand side is positive. Thus, for  $\eta < 0$ , we have  $\alpha_\eta \geq \min_i \alpha_{i,l}^{(1)}$ . Also, for  $a \geq 0$  we have  $\rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta} \geq \rho^{-\min_i \alpha_{i,l}^{(1)}}$ . Thus, we always have  $\rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta} \geq \rho^{-\min_i \alpha_{i,l}^{(1)}}$ , leading to the following lemma.

**Lemma 12:** For  $\alpha_{q,l}^{(1)}$  and  $\alpha_{q,l}^{(2)}$  defined above, we have  $\rho^{-\alpha_{q,l}^{(2)}} \geq \rho^{-\min_i \alpha_{i,l}^{(1)}} \Rightarrow \alpha_{q,l}^{(2)} \leq \min_i \alpha_{i,l}^{(1)}$  for  $q \in \mathcal{A}$ .

We now partition the DFT points into two sets  $\mathcal{A} = \{T(i-1) + 1, i = 1, \dots, L_{d_1}\}$  and  $\mathcal{B} = \{1, \dots, L_{d_1}\} \setminus \{T(i-1) + 1, i = 1, \dots, L_{d_1}\}$ . We now define the event

$$\mathcal{D} \triangleq \{\min_i \alpha_{i,1}^{(1)} < 1, \min_i \alpha_{i,2}^{(1)} < 1, \dots, \min_i \alpha_{i,N}^{(1)} < 1\}$$

and proceed to evaluate the probability

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) = \mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{1 + \rho \sum_{l=1}^N |b_{q,l}^{(1)}|^2} > m\right) \quad (148)$$

$$\begin{aligned} &= \mathbb{P}\left(\sum_{q \in \mathcal{A}} \frac{1}{1 + \rho \sum_{l=1}^N |b_{q,l}^{(1)}|^2} + \sum_{q \in \mathcal{B}} \frac{1}{1 + \rho \sum_{l=1}^N |b_{q,l}^{(1)}|^2} > m\right) \\ &\doteq \mathbb{P}(S_1 + S_2 > m) \end{aligned} \quad (149)$$

where (148) follows since  $\lambda_q = \mathbf{b}_q^H \mathbf{b}_q$  and  $S_1$  and  $S_2$  are given by

$$\begin{aligned} S_1 &\triangleq \sum_{q=1}^{L_{d_1}} \frac{1}{1 + \sum_{l=1}^N \rho^{1-\alpha_{q,l}^{(1)}}} \\ S_2 &\triangleq \sum_{q \in \mathcal{B}} \frac{1}{1 + \sum_{l=1}^N \rho^{1-\alpha_{q,l}^{(2)}}}. \end{aligned}$$

We now evaluate (149)

$$\begin{aligned} \mathbb{P}(S_1 + S_2 > m) &= \mathbb{P}(S_1 + S_2 > m \mid \mathcal{D}) \times \mathbb{P}(\mathcal{D}) + \\ &\quad \mathbb{P}(S_1 + S_2 > m \mid \bar{\mathcal{D}}) \times \mathbb{P}(\bar{\mathcal{D}}). \end{aligned} \quad (150)$$

Note that subject to the event  $\mathcal{D}$ , we have

$$S_2 = \sum_{q \in \mathcal{B}} \frac{1}{1 + \sum_{l=1}^N \rho^{1-\alpha_{q,l}^{(2)}}} \doteq 0.$$

Therefore, this term can be asymptotically ignored. Also subject to  $\bar{\mathcal{D}}$ , we have

$$S_1 = \sum_{q=1}^{L_{d_1}} \frac{1}{1 + \sum_{l=1}^N \rho^{1-\alpha_{q,l}^{(1)}}} \doteq L_{d_1}$$

and since with probability one,  $L_{d_1} \geq m$ , the other (nonnegative) term can be asymptotically ignored. Thus, both the terms involving the set  $\mathcal{B}$  can be altogether ignored and we have

$$\begin{aligned} \mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) &\doteq \mathbb{P}(S_1 > m|\mathcal{D})\mathbb{P}(\mathcal{D}) + \\ &\quad \mathbb{P}(S_1 > m|\bar{\mathcal{D}})\mathbb{P}(\bar{\mathcal{D}}) \\ &\doteq \mathbb{P}\left(\sum_{q=1}^{L_{d_1}} \frac{1}{(1+\rho\lambda_q)} > m\right). \end{aligned}$$

We have thus established (143) when  $L_{d_1}|L_{d_2}$ . We must now show that the same result holds for any  $T'$  when  $L_{d_1} \nmid T'$ . To do so, let  $L_{d_2} = T'L_{d_1}$ ; then, we have

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) \doteq \mathbb{P}\left(\sum_{q=1}^{T'} \frac{1}{(1+\rho\lambda_q)} > m\right). \quad (151)$$

Using (143) when  $L_{d_1}|L_{d_2}$  and (151) when  $T'|L_{d_2}$  together establishes (143) for any two positive integers.

## REFERENCES

- [1] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 271–286, Jan. 2006.
- [2] A. L. Moustakas, K. R. Kumar, and G. Caire, "Performance of MMSE MIMO receivers: A large n analysis for correlated channels," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2009, pp. 1–5.
- [3] E. A. Jorswieck and H. Boche, "Information theory outage probability in multiple antenna system," *Eur. Trans. Telecommun.*, vol. 18, no. 3, pp. 217–233, Jun. 2006.
- [4] Y. Jiang, M. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An in-depth study of the high SNR regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2008–2026, Apr. 2011.
- [5] E. N. Onggosanusi, A. G. Dabak, T. Schmidl, and T. Muharemovic, "Capacity analysis of frequency-selective MIMO channels with sub-optimal detectors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2002, vol. 3, pp. 2369–2372.
- [6] 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.
- [7] K. R. Kumar, G. Caire, and A. L. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.
- [8] H. Gao, P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 666–672, May 1998.
- [9] D. Gore, A. Gorokhov, and A. Paulraj, "Joint MMSE versus v-BLAST and antenna selection," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2002, vol. 1, pp. 505–509.
- [10] A. Tajer and A. Nosratinia, "Diversity order in ISI channels with single-carrier frequency-domain equalizer," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1022–1032, Mar. 2010.
- [11] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [12] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.
- [13] A. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.
- [14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [15] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [17] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304–306, Jul. 2001.
- [18] H. Sari, G. Karan, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.
- [19] B. Muquet, Z. Wang, G. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.
- [20] Z. Wang, X. Ma, and G. Giannakis, "Optimality of single-carrier zero-padded block transmissions," in *Proc. IEEE Wireless Commun. Network. Conf.*, Mar. 2002, vol. 2, pp. 660–664.
- [21] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.
- [22] C. Tepedelenlioglu, "Low complexity linear equalizers with maximum multipath diversity for zero-padded transmissions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2003, vol. 4, pp. 636–639.
- [23] L. Grokop and D. Tse, "Diversity-multiplexing tradeoff in ISI channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 109–135, Jan. 2009.
- [24] I. S. Gradshteyn and I. M. Ryzbk, *Tables of Integrals, Series, and Products*, 6th ed. New York: Academic, 2000.
- [25] A. Stamoulis, S. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2451–2464, Oct. 2002.
- [26] A. Kaveh and H. Rahami, "Block circulant matrices and applications in free vibration analysis of cyclically repetitive structures," *Acta Mechanica*, vol. 217, pp. 1–12, 2010.
- [27] S. Bradely, A. Hax, and T. Magnanti, *Applied Mathematical Programming*. Reading, MA: Addison-Wesley, 1977, ch. 6.
- [28] C. Tepedelenlioglu, "Maximum multipath diversity with linear equalization in precoded OFDM systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 232–235, Jan. 2004.
- [29] A. H. Mehana and A. Nosratinia, "Diversity of MMSE MIMO receivers," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2163–2167.

**Ahmed Hesham Mehana** (S'10) received his B.S. and M.S. degrees from Cairo University in 2004 and 2007, respectively, both in Electrical Engineering. He is currently pursuing the Ph.D. degree in electrical engineering at the University of Texas at Dallas. He was an intern at Research in Motion Co. Ltd during 2010 and research assistant in Texas A&M at Qatar in 2008. His current interests include MIMO precoding, linear receivers, and interference management.

**Aria Nosratinia** (S'87–M'97–SM'04–F'10) is Jonsson Distinguished Professor of Engineering at the University of Texas at Dallas. He received his Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 1996. He has held visiting appointments at Princeton University, Rice University, and UCLA. His interests lie in the broad area of information theory and signal processing, with applications in wireless communications. He was the secretary for the IEEE Information Theory Society in 2010–2011 and was the treasurer for ISIT 2010 in Austin, Texas. He has served as editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE Wireless Communications (Magazine). He has been the recipient of the National Science Foundation career award, and is a fellow of IEEE.