Capacity Analysis for Parallel and Sequential MIMO Equalizers

Xinying Zhang and Sun-Yuan Kung, *Fellow, IEEE*

Abstract—**Employing multiple antennas at both the transmitter and receiver ends offers a promising channel capacity. Unfortunately, equalizers that can deliver better theoretical capacity performance usually incur higher implementation cost. To facilitate the trade-off design in practice, this paper explores and compares the asymptotic capacity performance of different multiple-input–multiple-output (MIMO) equalizers for both deterministic and stochastic (in particular, Rayleigh fading) channel models, in terms of two measurements: "capacity gap" and "capacity ratio." Based on linear algebra and matrix operations, the capacity results are given in analytical form as functions of the coupling terms in the channel transfer function and the signal-to-noise ratio (SNR). The closed-form solutions enable our theoretical work serve as a reference for practical system designers. Our theoretical finding concludes that the interplay among the individual equalizers associated with each input stream plays a much more important role than the detailed filter selection, which is further verified by the Monte Carlo simulations. Although we focus on flat-fading channels in this paper, the result is naturally extendible to the inter-symbol-interference (ISI) MIMO systems.**

Index Terms—**Capacity, DFE, equalizer, MIMO, Rayleigh.**

I. Introduction

MULTIPLE-input–multiple-output (MIMO) techniques are attracting extensive attentions from both the researchers and system designers in recent days and are broadly recognized as having the potential to greatly increase the (Shannon) capacity of the modern wireless systems [1]–[5]. Communication scheme designs for MIMO system have demonstrated a promising performance improvement [6]–[8]. Practically, equalizers and coder/decoders that serve to suppress the interference from other antennas and channel distortion inevitably suffer a certain capacity loss. Various MIMO diversity-combining equalizers have been investigated in the past [9]–[11], [13]. It has been well known that the simple linear equalizers sacrifice a lot of performance due to the possible severe noise enhancement. On the other hand, minimum mean square error (MMSE) equalizers, combined with a nonlinear successive interference cancellation procedure, can achieve the full channel capacity [15], [16]. Modified versions of the finite impulse response (FIR) decision-feedback equalizer (DFE) have been designed and reported [17], [18]. However, the difficulty of the MMSE filter lies in its requirement of accurate signal-to-noise ratio (SNR) estimate and a big implementation complexity. Such additional implementation cost will become much more serious when the ISI effect can no longer be neglected with much higher data rates in next-generation wireless communications. Therefore, any practical system designers should be keenly aware of the vital importance of reaching an optimal tradeoff between the implementation complexity and the achievable capacity. To facilitate such design tradeoff, it will be critical to have available theoretical analysis of how much performance is lost when employing suboptimal equalizer designs and, whenever possible, in simple closed-form formulations.

To this end, this paper studies the asymptotic capacity performance of different MIMO equalizers with distinct filter or structure design. Mathematical analysis is provided first for deterministic channel models, and thereafter, the capacity performance under Rayleigh fading channels are investigated based on the deterministic results. The major findings are summarized in the five theorems, three comparisons, and one conjecture throughout the paper. The beauty inside our theoretical work is that most of the capacity results are given in simple closed-form equations as functions of the channels parameters and SNR, which makes it directly applicable to practical system designs. The paper is outlined as follows. Section II describes the most common equalizers used nowadays and the design considerations, where the design criterion and constraint for each equalizer and the corresponding mathematical derivations are provided. In Section III, the achievable capacity performance of these equalizers under deterministic channels are addressed and compared. The asymptotic performance when the SNR goes to zero or infinity is given in analytical form. Finally, the performance under stochastic Rayleigh fading channels is investigated in Section IV. Simulation results are provided throughout the paper to confirm and support the theoretical findings.

The following notations are used throughout the paper: $*$ for conjugate transpose, $+$ for the Moore-Penrose pseudo-inverse of matrices, $\mathbf{I}_n$ for the $n \times n$ identity matrix, $\mathbf{0}_{m\times n}$ for the $m \times n$ zero matrix, $\mathcal{E}[\cdot]$ for expectation, $\det[\cdot]$ for determinant, $\mathrm{tr}[\cdot]$ for matrix trace, and $\doteq$ for the asymptotic equivalence. All the matrices are written in boldface capital letters and vectors are written in lowercase letters with upper arrows.

We consider a general MIMO communication system adopting t transmit and r receive antenna elements, as depicted in Fig. 1:

$$\vec{x}(k) = \mathbf{H}\vec{s}(k) + \vec{n}(k) \tag{1}$$

Manuscript received December 15, 2002; revised June 3, 2003. This work was supported in part by a grant from Mitsubishi Electric Research Laboratories. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dhananjay A. Gore.

The authors are with the Electrical Engineering Department, Princeton University, Princeton, NJ 08544 USA (e-mail: xinying@ee.princeton.edu; kung@ee.princeton.edu).

Digital Object Identifier 10.1109/TSP.2003.818151

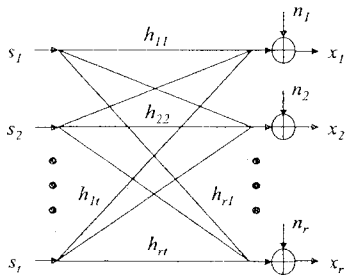


Fig. 1. Mathematical MIMO channel model.

where $\vec{x}(k) \in \mathcal{C}^r$, $\vec{s}(k) \in \mathcal{C}^t$ are sample stacks of the complex-valued receiver data and source sequences, and $\mathbf{H}$ is the $r \times t$ channel transfer function. In this paper, we assume that $r \geq t$ and the channel is generic, i.e., $\mathbf{H}$ has full column rank. The channel realization is assumed to be tracked at the receiver end. The total transmission power is constrained to P and is shared by the t transmitting antennas with the distribution factors $\{\phi_1, \ldots, \phi_t\}$: $\mathcal{E}[s_i(k)s_i^*(k)] = \phi_i P$, where $\sum_{i=1}^t \phi_i = 1$. The thermal noises $\vec{n}(k) \in \mathcal{C}^t$ are both spatially and temporally white i.i.d. Gaussian random processes with independent real and imaginary parts and variance $N\mathbf{I}_r$. Assuming independent inputs, the channel capacity of the MIMO system in (1) is well known as

$$\begin{aligned} C &= \log_2 \det\left(\mathbf{I}_r + \frac{P}{N}\mathbf{H}\mathbf{\Phi}\mathbf{H}^*\right) \\ &= \log_2 \det(\mathbf{I}_t + \rho\mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2}) \\ &= \sum_{i=1}^{t} \log_2(1 + \rho|\lambda_i|^2) \end{aligned} \tag{2}$$

where $\mathbf{\Phi} = \text{diag}\{\phi_i\}_{i=1}^t$, $\rho = P/N$ is the nominal SNR, and $\{|\lambda_i|^2\}_{i=1}^t$ are the singular values of $\mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2}$.

Here, we use the term "channel capacity" with a minor abuse. The standard Shannon capacity characterizes the maximal information rate supported in the channel with error-free transmission when the channel state information (CSI) is available at the transmitter and optimal power allocation is adopted (e.g., water filling; cf. [12]). In practice, the power allocation $\mathbf{\Phi}$ may be different when the CSI is unavailable or when the optimal allocation is not adopted due to the specific QoS and complexity consideration in the system. The power allocation is beyond the scope of this paper. Our channel capacity in (2) describes the maximal supportable rate for any given power distribution $\mathbf{\Phi}$. Our capacity results in Section III hold for any valid $\mathbf{\Phi}$, where the optimal case (with water filling) and the suboptimal case (for example, even distribution) could be viewed as special cases.

In practice, the capacity is sacrificed not only from the suboptimal equalizers but also from other imperfect devices in the transceiver, such as the codec. In this paper, we only address the capacity loss caused by different equalizers. The underlying assumption is that the information rate allocation over different antennas is known to the transmitter and is moderated by link layer such that the codec will be able to support a perfect error-correction. In other words, we have made the assumption that there is no error propagation, which is reasonable as far as the capacity (as opposed to BER) analysis is concerned. (Recall that there always exists a code that will permit the error-free transmission across the channel as long as the signal is transmitted at a rate less than the channel capacity.) Even for small SNR, the capacity results in the paper could be approached as long as proper codings with sufficient error-correction ability have been adopted.

II. Design Considerations for MIMO Equalizers

While MIMO channels offer an opportunity for enormous capacity by providing multiple copies of the same information stream, the interference from different users present an additional difficulty for the correct retrieval of each stream. Strategically, the design and overall performance of a MIMO equalizer depend on the following two factors.

1) **Filter Selection:**
   What kind of individual filter is applied for the recovery of each input? There are several possible choices for the filter selection.
2) **Structure Selection:**
   The structural design places emphasis on the interplay of the t individual filters between each other: one for each input.

A. Filter Design

Different selection of filter designs can lead to very distinctive performance in capacity, error resilience, BER, and computational complexity. It is therefore an important factor in the design tradeoff. An intuitive way to equalize is to apply linear filter arrays on the receiver samples and get the input estimate via the weighted summation of observations. The (linear) estimate filter design is conducted with respect to different criteria, among which, the two popular candidates are zero-forcing (ZF) and MMSE filters.

1) ZF Constraint and ZF Equalizer: For the recovery of the ith input stream $s_i(k)$, the ZF constraint on the corresponding diversity-combiner, which is denoted by the row vector $\vec{g}_{\text{ZF},i}$, is

$$\vec{g}_{\text{ZF},i}\mathbf{H} = \vec{e}_i \tag{3}$$

where $\vec{e}_i$ is a unit (row) vector with all elements zero, except 1 at position i. Specifically, the application of $\vec{g}_{\text{ZF},i}$ in (3) yields a virtual single-input–single-output (SISO) channel

$$\vec{g}_{\text{ZF},i}\vec{x}(k) = s_i(k) + \vec{g}_{\text{ZF},i}\vec{n}(k). \tag{4}$$

The associated capacity after equalization is then

$$C_{\text{ZF},i} = \log_2\left(1 + \frac{\phi_i\rho}{\|\vec{g}_{\text{ZF},i}\|^2}\right). \tag{5}$$

The optimal ZF equalizers satisfying the constraint in (3) and maximizing the capacity in (5) are given by the pseudo-inverse of the channel transfer function:

$$\vec{g}_{\text{ZF},i} = \vec{e}_i\mathbf{H}^+. \tag{6}$$

2) MMSE Criterion and MMSE Equalizer: The individual MMSE filters are designed to minimize the MSE. For the retrieval of input i, the MMSE criterion on the equalizer is

$$\begin{aligned} \vec{g}_{M,i} &= \arg\min_{\vec{g}} \mathcal{E}[|\vec{g}\vec{x}(k) - s_i(k)|^2] \\ &= \arg\min_{\vec{g}} \mathcal{E}[|(\vec{g}\mathbf{H} - \vec{e}_i)\vec{s}(k) + \vec{g}\vec{n}(k)|^2]. \end{aligned} \tag{7}$$

Denote ith column in $\mathbf{H}$ as $\vec{h}_i$. The solution to the equation above can be obtained via the application of orthogonality principle in estimation theory:

$$\begin{aligned}\vec{g}_{M,i} =& \mathcal{E}[s_i(k)\vec{x}^*(k)](\mathcal{E}[\vec{x}(k)\vec{x}^*(k)])^{-1} \\ =& \rho\phi_i\vec{h}_i^*(\mathbf{I}_r + \rho\mathbf{H}\Phi\mathbf{H}^*)^{-1}. \end{aligned} \tag{8}$$

The MMSE filter delivers the following signal-to-noise-and-interference ratio (SNIR) after equalization:

$$\begin{aligned}\text{SNIR}_i =& \frac{\mathcal{E}[|\vec{g}_{M,i}\vec{h}_i s_i(k)|^2]}{\mathcal{E}[|\vec{g}_{M,i}\mathbf{H}\vec{s}(k) + \vec{g}_{M,i}\vec{n}(k) - \vec{g}_{M,i}\vec{h}_i s_i(k)|^2]} \\ =& \frac{\rho\phi_i|\vec{g}_{M,i}\vec{h}_i|^2}{\vec{g}_{M,i}(\mathbf{I}_r + \rho\mathbf{H}\Phi\mathbf{H}^* - \rho\phi_i\vec{h}_i\vec{h}_i^*)\vec{g}_{M,i}^*} \\ =& \frac{\rho\phi_i\vec{h}_i^*(\mathbf{I}_r + \rho\mathbf{H}\Phi\mathbf{H}^*)^{-1}\vec{h}_i}{1 - \rho\phi_i\vec{h}_i^*(\mathbf{I}_r + \rho\mathbf{H}\Phi\mathbf{H}^*)^{-1}\vec{h}_i}. \end{aligned} \tag{9}$$

The SNIR derivation above suggests the following achievable capacity in each equalized subchannel:

$$\begin{aligned}C_{\text{MMSE},i} =& \log_2(1 + \text{SNIR}_i) \\ =& -\log_2[1 - \rho\phi_i\vec{h}_i^*(\mathbf{I}_r + \rho\mathbf{H}\Phi\mathbf{H}^*)^{-1}\vec{h}_i^*] \\ =& -\log_2[\vec{e}_i(\mathbf{I}_r + \rho\Phi^{1/2}\mathbf{H}^*\mathbf{H}\Phi^{1/2})^{-1}\vec{e}_i^*]. \end{aligned} \tag{10}$$

The equality in the last equation can be verified via the Sherman-Morrison-Woodbury identity: $\mathbf{A}(\mathbf{I} + \mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^* = \mathbf{I} - (\mathbf{I} + \mathbf{A}\mathbf{A}^*)^{-1}$.

B. Structure Design

Another important aspect of equalizer design, which will be momentarily shown to play a more dominant role than filter choice, is the structure selection. Two kinds of commonly used structures are parallel design and sequential design.

1) Parallel Equalizer: The design of parallel equalizers is simple both conceptually and in implementation. The recovery of each input falls on the responsibility of one filter only. In the parallel structure, a total of t independently processing filters (one-for-one input) will be adopted. Simultaneous application of these separately designed filters leads to the parallel structure depicted in Fig. 2.

2) Sequential Equalizer: Like its parallel counterpart, a sequentially structured equalizer also comprises t filters. Unlike the parallel equalizer, the t filters are no longer independently processed. The coupling component is via a process often called "decision feedback." The detection of one input stream now consists of 1) a feedforward session with linear space-time combiners and 2) a feedback session with nonlinear devices. Each feedback part is in turn composed of a decision device and an interference canceller. The decision device is usually a decoder with error correction ability. The contributions of the detected symbols are promptly eliminated from the observation so that the decision feedback section will produce interference-reduced data samples ready for further extraction of other streams.

The sequential structure has also been well known as DFE or vertical layering. In BLAST (V-BLAST) [13], [14], the sequential equalizer with ZF has been proposed to realize very high data rates over the rich-scattering wireless channel. Similarly, the capacity performance of sequential MMSE equalizer has been addressed in [15].

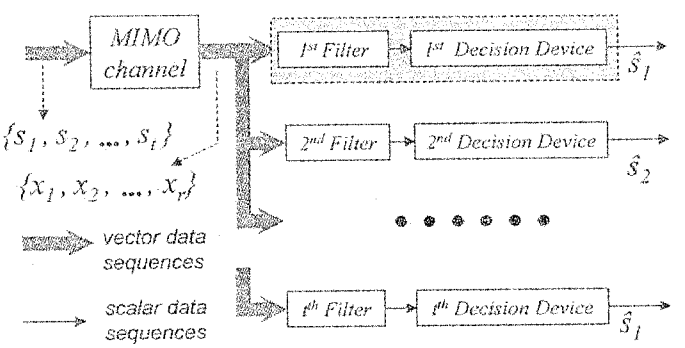


Fig. 2. Block diagram of a parallel equalizer. It can be decoupled into t individual equalizers: one for each input.

III. Capacity Analysis of MIMO Equalizers Under Deterministic Channels

This section studies the achievable capacities via different MIMO equalizers under deterministic channel models.

A. Parallel Equalizers

1) Parallel ZF (PZF) Equalizers: The application of PZF equalizers on the receiver converts the original MIMO system into t independent and interference-free channels that can be separately decoded. The total information rate supported by the PZF equalization is therefore the summation of the subchannel capacities $C_{PZF,i}$ over all the indices $i = 1, \ldots, t$.

To facilitate the analysis, we introduce the following Cholesky factorization:

$$\Phi^{1/2}\mathbf{H}^*\mathbf{H}\Phi^{1/2} = \mathbf{R}^*\mathbf{R} \tag{11}$$

where $\mathbf{R} = [r_{ij}]$ is a $t \times t$ nonsingular upper-triangular matrix. It turns out that the quantitative analysis hinges on both $\mathbf{R}$ and the inverse matrix $\mathbf{R}^{-1}$. For convenience, the elements of $\mathbf{R}^{-1}$ are denoted as $\hat{r}_{ij}$, i.e., $\mathbf{R}^{-1} = [\hat{r}_{ij}]$. In the context of capacity analysis, of particular interests are the sets of correlation factors defined below.

Definition (Correlation Factors in $\mathbf{R}$ and $\mathbf{R}^{-1}$): For the upper-triangular matrix $\mathbf{R}$ defined in (11) and its matrix inverse $\mathbf{R}^{-1}$, we define three sets of so called "correlation factors" as follows:

$$\begin{aligned}\alpha_i =& \frac{\|\mathbf{R}\vec{e}_i^*\|^2}{|r_{ii}|^2} - 1 = \frac{\sum_{j=1}^{i-1}|r_{ji}|^2}{|r_{ii}|^2} \geq 0 \\ \beta_i =& |r_{ii}|^2\|\vec{e}_i\mathbf{R}^{-1}\|^2 - 1 = \frac{\sum_{j=i+1}^{t}|\hat{r}_{ij}|^2}{|\hat{r}_{ii}|^2} \geq 0 \\ \gamma_i =& \frac{\|\vec{e}_i\mathbf{R}^{-1}\mathbf{R}^{-*}\|^2}{\|\vec{e}_i\mathbf{R}^{-1}\|^4} - 1 \geq 0. \end{aligned} \tag{12}$$

These parameters represent the degree of correlation between the channel vectors regarding different inputs. For example, β_i and γ_i are, respectively, the 2-norm ratio of the total off-diagonal terms to the diagonal terms in the ith row of matrix $\mathbf{R}^{-1}$ and $\mathbf{R}^{-1}\mathbf{R}^{-*}$. In fact, $\{\alpha_i\}$ and $\{\beta_i\}$ depend only on the channel transfer function $\mathbf{H}$, whereas $\{\gamma_i\}$ depend on both $\mathbf{H}$ and $\mathbf{\Phi}$. If the channel vector for different transmitters are mutually orthogonal, all the correlation factors will vanish.

Our evaluation on the capacity performance is based on the following two measurements.

Definition 2 (Capacity Gap and Ratio): Given a MIMO channel and a corresponding equalizer, the capacity gap is defined as the absolute difference between the capacity achieved by this equalizer (say, C_e) and that given by the original channel (say, C). The capacity ratio is then defined as C_e/C. These two terms are both functions of the SNR ρ. ■

The standard Landau symbols $O(\cdot)$ and $o(\cdot)$ are adopted in the rest of the paper, i.e., $f(x) = O(x)$ means $|f(x)| < Ax$ for some (finite) constant A and all values of x, and $f(x) = o(x)$ means that $\lim_{x\to\infty} f(x)/x = 0$.

Theorem 1 (Capacity Gap and Ratio of PZF Equalizers): The asymptotic capacity gap and ratio of the optimal PZF equalizer in (6) are as follows.

1) In the high SNR regime

$$C - C_{\mathrm{PZF}} = \sum_{i=1}^{t} \log_2(1+\beta_i) + O[\rho^{-2}]$$

$$\frac{C_{\mathrm{PZF}}}{C} = 1 - \frac{\sum_{i=1}^{t} \ln(1+\beta_i)}{t \ln \rho} + o[(\rho \ln \rho)^{-1}]. \quad (13)$$

2) In the low SNR regime

$$C - C_{\mathrm{PZF}} = \frac{\rho}{\ln 2} \sum_{i=1}^{t} |r_{ii}|^2 \left(1 + \alpha_i - \frac{1}{1+\beta_i}\right) + O[\rho^2]$$

$$\frac{C_{\mathrm{PZF}}}{C} = \frac{\sum_{i=1}^{t} \frac{|r_{ii}|^2}{(1+\beta_i)}}{\sum_{i=1}^{t} |r_{ii}|^2 (1+\alpha_i)} + O[\rho]. \quad (14)$$

Proof: By (5) and (6), the capacity achieved by the optimal PZF equalizer is

$$\begin{aligned} C_{\mathrm{PZF}} &= \sum_{i=1}^{t} \log_2 \left(1 + \frac{\phi_i \rho}{\|\vec{g}_{PZF,i}\|^2}\right) \\ &= \sum_{i=1}^{t} \log_2 \left[1 + \frac{\rho}{\vec{e}_i (\mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2})^{-1} \vec{e}_i^*}\right] \\ &= \sum_{i=1}^{t} \log_2 \left[1 + \frac{\rho}{\vec{e}_i (\mathbf{R}^*\mathbf{R})^{-1} \vec{e}_i^*}\right] \\ &= \sum_{i=1}^{t} \log_2 \left(1 + \frac{\rho |r_{ii}|^2}{1+\beta_i}\right). \end{aligned} \quad (15)$$

Comparing (15) with the original MIMO capacity in (2), we have

$$C - C_{\mathrm{PZF}} = \sum_{i=1}^{t} \log_2(1+\beta_i) + \sum_{i=1}^{t} \log_2 \frac{|\lambda_i|^2}{|r_{ii}|^2} + \sum_{i=1}^{t} \log_2 \frac{1 + \frac{1}{|\lambda_i|^2 \rho}}{1 + \frac{1+\beta_i}{\rho |r_{ii}|^2}} \quad (16)$$

As $\{|\lambda_i|^2\}_{i=1}^{t}$ are the singular values for the Hermitian matrix $\mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2} = \mathbf{R}^*\mathbf{R}$, we have the equalities

$$\begin{aligned} \prod_{i=1}^{t} |\lambda_i|^2 &= \det[\mathbf{R}^*\mathbf{R}] = \prod_{i=1}^{t} |r_{ii}|^2 \\ \sum_{i=1}^{t} \frac{1}{|\lambda_i|^2} &= \mathrm{tr}[(\mathbf{R}^*\mathbf{R})^{-1}] = \sum_{i=1}^{t} \frac{1+\beta_i}{|r_{ii}|^2}. \end{aligned} \quad (17)$$

Based on the two equalities , when $\rho \to \infty$, the capacity gap in (16) is simplified as

$$C - C_{\mathrm{PZF}} = \sum_{i=1}^{t} \log_2(1+\beta_i) + O[\rho^{-2}] \quad (18)$$

by using $\log_2(1+x) = x/\ln 2 + O(x^2)$ for small x. The other three equations in the theorem can be verified in a similar way, and the proof is omitted here. ■

2) Parallel MMSE (PMMSE) Equalizers: Qualitatively, it has been well known that MMSE can outperform its ZF counterpart. In this section, we will investigate the quantitative aspect of this improvement.

Theorem 2 (Improvement of PMMSE Equalizers): The capacity performance improvement of PMMSE over PZF equalizers is shown to be the following.

1) In the high SNR regime

$$C_{\mathrm{PMMSE}} - C_{\mathrm{PZF}} = \frac{1}{\rho \ln 2} \sum_{i=1}^{t} \frac{\gamma_i(1+\beta_i)}{|r_{ii}|^2} + O[\rho^{-2}] \quad (19)$$

$$\frac{C_{\mathrm{PMMSE}}}{C} - \frac{C_{\mathrm{PZF}}}{C} = \frac{1}{t\rho \ln \rho} \sum_{i=1}^{t} \frac{\gamma_i(1+\beta_i)}{|r_{ii}|^2} + o[(\rho \ln \rho)^{-1}]. \quad (20)$$

2) In the low SNR regime

$$C_{\mathrm{PMMSE}} - C_{\mathrm{PZF}} = \frac{\rho}{\ln 2} \sum_{i=1}^{t} |r_{ii}|^2 \left(1 + \alpha_i - \frac{1}{1+\beta_i}\right) + O[\rho^2]$$

$$\frac{C_{\mathrm{PMMSE}}}{C} = 1 + O[\rho]. \quad (21)$$

Proof: The proof of (19) is given below, whereas the others are omitted. Comparing (15) and (10), the gap is

$$C_{\mathrm{PMMSE},i} - C_{\mathrm{PZF},i} = \log_2 \frac{1}{\vec{e}_i (\mathbf{I}_r + \rho \mathbf{R}^*\mathbf{R})^{-1} \vec{e}_i^* \left(1 + \frac{\rho}{\vec{e}_i (\mathbf{R}^*\mathbf{R})^{-1} \vec{e}_i^*}\right)}. \quad (22)$$

Note that for large ρ

$$\begin{aligned}\vec{e}_i(\mathbf{I}_r+\rho\mathbf{R}^*\mathbf{R})^{-1}\vec{e}_i^* &= \vec{e}_i(\rho\mathbf{R}^*\mathbf{R})^{-1}[\mathbf{I}_r+(\rho\mathbf{R}^*\mathbf{R})^{-1}]^{-1}\vec{e}_i^* \\ &= \vec{e}_i[\rho^{-1}(\mathbf{R}^*\mathbf{R})^{-1}-\rho^{-2}(\mathbf{R}^*\mathbf{R})^{-2}]\vec{e}_i^* \\ &\quad + O[\rho^{-3}] \\ \Longrightarrow C_{\mathrm{PMMSE},i} - C_{\mathrm{PZF},i} &= \rho^{-1}\left[\frac{\vec{e}_i(\mathbf{R}^*\mathbf{R})^{-2}\vec{e}_i^*}{\vec{e}_i(\mathbf{R}^*\mathbf{R})^{-1}\vec{e}_i^*}-\vec{e}_i(\mathbf{R}^*\mathbf{R})^{-1}\vec{e}_i^*\right] \\ &\quad + O[\rho^{-2}] \\ &= \rho^{-1}\frac{\gamma_i(1+\beta_i)}{|r_{ii}|^2}+O[\rho^{-2}]. \end{aligned} \tag{23}$$

■

Theorems 1 and 2 demonstrate a constant capacity degradation of the amount $\sum_{i=1}^{t}\log_2(1+\beta_i)$ for both the linear MMSE and ZF schemes. (Thus, this gap is a function of $\mathbf{H}$ only.) The channel capacity can be asymptotically achieved only when *all* the channel correlation factors $\beta_i = 0$, which is rarely the case in practice. The argument above inspires possible alteration for improving capacity performance in bad channel scenario (i.e., highly aligned channel vectors with large β_i) via inserting nonlinear operations in the equalizer to artificially reduce the correlation factors. The subsequent analysis will further demonstrate that the degradation caused by such correlations can be eliminated via successive deletions of the last column vector in $\mathbf{H}$, which is equivalent to a nonlinear sequential structure in the equalizer.

B. Sequential Equalizers

1) Sequential ZF (SZF) Equalizers: In BLAST design [13], Foschini proved that the successive ZF can asymptotically approach the capacity lower bound for Rayleigh fading MIMOs when $r = t$. In this section, we give a comprehensive capacity analysis of SZF equalizers in the entire SNR range for any deterministic channel realization and antenna settings.

For notational simplicity, we assume that the input streams are sequentially retrieved in the order of $t, t-1, \ldots, 1$. In the SZF equalizer, the interferences associated with the already-detected inputs are successively nulled from the observation data before the filters for the other inputs are applied. In the sense of capacity, a perfect codec could be assumed to provide an error-free estimate on the information stream as long as the bit rate is within the theoretical limit. Therefore, no error propagation could be assumed in the calculation of capacity. Under this assumption, the interference-reduced receiver data, at the input of ith ZF filter, is

$$\vec{x}^{(i)}(k) = \mathbf{H}_i\vec{s}^{(i)}(k)+\vec{n}(k) \tag{24}$$

where $\mathbf{H}_i$ is the first i columns of $\mathbf{H}$, denoting the virtual channel after the inputs $i+1, i+2, \ldots, t$ have been detected and eliminated. The vector $\vec{s}^{(i)}(k)$ is the first i rows of $\vec{s}(k)$. The ZF constraint for the sequential equalizers is now

$$\vec{g}_{\mathrm{SZF},i}\mathbf{H}_i = \vec{e}_i, \quad 1 \le i \le t. \tag{25}$$

In addition, we denote $\mathbf{R}_i(\Phi_i)$ as the left-upper $i \times i$ minors of $\mathbf{R}(\Phi)$. Exploiting the upper-triangular structure of $\mathbf{R}$, we get

$$\Phi_i^{1/2}\mathbf{H}_i^*\mathbf{H}_i\Phi_i^{1/2} = \mathbf{R}_i^*\mathbf{R}_i. \tag{26}$$

Theorem 3 (Capacity Gap and Ratio for SZF Equalizers): The capacity achieved by the optimal SZF equalizer is

$$C_{\mathrm{SZF}} = \sum_{i=1}^{t}\log_2(1+\rho|r_{ii}|^2). \tag{27}$$

The corresponding capacity gap and ratio are as follows.

1) In the high SNR region

$$\begin{aligned}\frac{C_{\mathrm{SZF}}}{C} &= 1-\frac{1}{t\rho\ln\rho}\sum_{i=1}^{t}\frac{\beta_i}{|r_{ii}|^2}+o[(\rho\ln\rho)^{-1}] \\ C-C_{\mathrm{SZF}} &= \frac{1}{\rho\ln 2}\sum_{i=1}^{t}\frac{\beta_i}{|r_{ii}|^2}+O[\rho^{-2}]. \end{aligned} \tag{28}$$

2) In the low SNR region

$$\begin{aligned}C-C_{\mathrm{SZF}} &= \frac{\rho}{\ln 2}\sum_{i=1}^{t}\alpha_i|r_{ii}|^2+O[\rho^2] \\ \frac{C_{\mathrm{SZF}}}{C} &= \frac{\sum_{i=1}^{t}|r_{ii}|^2}{\sum_{i=1}^{t}|r_{ii}|^2(1+\alpha_i)}+O[\rho]. \end{aligned} \tag{29}$$

Proof: Just like the parallel case, the optimal SZF equalizers $\vec{g}_{\mathrm{SZF,i}}$ satisfying the constraint (25) are

$$\vec{g}_{\mathrm{SZF,i}} = \vec{e}_i\mathbf{H}_i^+ \tag{30}$$

with the corresponding capacity

$$\begin{aligned}C_{\mathrm{SZF}} &= \sum_{i=1}^{t}C_{\mathrm{SZF,i}} = \sum_{i=1}^{t}\log_2\left[1+\frac{\rho\phi_i}{\|\vec{g}_{\mathrm{SZF,i}}\|^2}\right] \\ &= \sum_{i=1}^{t}\log_2\left[1+\frac{\rho}{\vec{e}_i\left(\Phi_i^{1/2}\mathbf{H}_i^*\mathbf{H}_i\Phi_i^{1/2}\right)^{-1}\vec{e}_i^*}\right] \\ &= \sum_{i=1}^{t}\log_2\left[1+\frac{\rho}{\vec{e}_i\mathbf{R}_i^{-1}\mathbf{R}_i^{-*}\vec{e}_i^*}\right] \\ &= \sum_{i=1}^{t}\log_2(1+\rho|r_{ii}|^2). \end{aligned} \tag{31}$$

The capacity gap is

$$\begin{aligned}C-C_{\mathrm{SZF}} &= \sum_{i=1}^{t}\log_2\frac{1+\rho|\lambda_i|^2}{1+\rho|r_{ii}|^2} \\ &= \sum_{i=1}^{t}\log_2\frac{1+\frac{1}{\rho|\lambda_i|^2}}{1+\frac{1}{\rho|r_{ii}|^2}}+\sum_{i=1}^{t}\log_2\frac{|\lambda_i|^2}{|r_{ii}|^2}. \end{aligned} \tag{32}$$

In the large SNR region, by applying (17), we obtain the first-order expansion

$$C - C_{\mathrm{SZF}} = \frac{\rho^{-1}}{\ln 2} \sum_{i=1}^{t} \frac{\beta_i}{|r_{ii}|^2} + O[\rho^{-2}]. \tag{33}$$

The other derivations basically follow the same argument. ∎

2) Sequential MMSE (SMMSE) Equalizers: It is well known that SMMSE equalizers can achieve the full channel capacity [15], [16], provided an accurate SNR estimate. In [15], a derivation was given to prove $C_{\mathrm{SMMSE}} = C$ at any SNR point with any detection ordering based on a capacity region vertex analysis. The same statement can be made naturally following our notations and evaluations.

C. Detection Order and Error Propagation in Sequential Equalizers

The detection order is of vital importance for sequential equalizers. In the layered structure, the signal stream detected later benefits more from the interference cancellation step as more estimated signals have been subtracted from the observation data, which leads to a larger SNR gain of the current information stream upon recovery. The optimal ordering for V-BLAST has been addressed in [14]. It is, however, easy to verify that the optimal SZF equalizer asymptotically achieves the original channel capacity when the SNR ρ approaches either infinity or zero: $C_{\mathrm{SZF}} \doteq C$, regardless of the detection order.

To show this, recall that the capacity-achieving claim is obvious with reference to (28) and (29). When $\rho \to \infty$, (31) leads to

$$C_{\mathrm{SZF}} \doteq C \doteq \sum_{i=1}^{t} \log_2(\rho |r_{ii}|^2) = \log_2 \rho^t \det[\mathbf{R}^*\mathbf{R}]. \tag{34}$$

Assume that the t input streams are to be retrieved in a different order. This can be accomplished by rearranging the columns of $\mathbf{H}$ by a permutation matrix $\mathbf{P}$, i.e., $\mathbf{H}' = \mathbf{H}\mathbf{P}$. Accordingly, $\mathbf{\Phi}' = \mathbf{\Phi}\mathbf{P}$. Denote $\mathbf{R}'$ as the Cholesky factorization matrices of $(\mathbf{\Phi}')^{1/2}(\mathbf{H}')^*\mathbf{H}'(\mathbf{\Phi}')^{1/2}$. Based on (34), the proof is completed by noting that $\mathbf{P}\mathbf{P}^* = \mathbf{I}_t$, and hence

$$\det[(\mathbf{R}')^*\mathbf{R}'] = \det[\mathbf{R}^*\mathbf{R}]. \tag{35}$$

The argument above claims the detection-order-invariance property of SZF with infinite SNR. However, in the practical range with finite SNR, the capacity performance depends on the specific detection order. In the small SNR region, the performance variation with respect to different detection orders might be severe. With t input streams, there are altogether $t!$ different arrangements of the ordering, each associated with its own performance. The optimal ordering may be selected according to different criteria. One possible option is based on the maximization of achievable capacity. An example is given in this section to demonstrate the selection on detection ordering.

For pure capacity analysis, there is no need to consider the estimate error as perfect codec can be assumed to support the information rate. However, error propagation inevitably occurs in a practical system and bit-error-rate (BER) is a more reasonable criterion for evaluation of a communication system. To best prevent the potentially catastrophic error propagation, it is heuristically optimal to first estimate and decode the j^*th input stream, whose individual equalizer (in parallel design) delivers the highest post-processing SNR on recovery. The detection order in the subsequent layers can then be determined in the same manner.

The optimal detection orders with capacity and BER criterion could be very different. The real system should be designed according to the specific QoS requirement. As an example, we consider the MIMO channel with transfer function

$$\mathbf{H} = \begin{bmatrix} 1 & a\sin\theta \\ 0 & a\cos\theta \end{bmatrix}. \tag{36}$$

Any nontrivial 2×2 MIMO channel could be characterized by the model above (if $\mathbf{H}$ is not upper-diagonal, replace it with the Cholesky factorization on $\mathbf{H}^*\mathbf{H}$, and this does not change the capacity results below). The parameter a characterizes the ratio of fading on the two transmitting streams; when $a \geq 1$, the second stream gets magnified more by the channel. The capacity performance of SZF under such a channel can be exactly evaluated via our former derivations. With equal power allocation and the detection order of (2, 1), the result is given as

$$C_{\mathrm{SZF}} = \log_2 \left(\frac{a^2\cos^2\theta}{4}\rho^2 + \frac{1 + a^2\cos^2\theta}{2}\rho + 1 \right). \tag{37}$$

Note that the correlation factor is $\beta_1 = \tan^2\theta$, which is a monotonically increasing function of θ within $[0, \pi/2]$. When $\theta = 0$, i.e., the channel is an interference-free Gaussian channel, it is exactly the same as the channel capacity.

To demonstrate the role of detection order on the capacity performance of SZF equalizers, we calculate the achievable capacity with a different ordering (1, 2), which is given below:

$$C'_{\mathrm{SZF}} = \log_2 \left(\frac{a^2\cos^2\theta}{4}\rho^2 + \frac{a^2 + \cos^2\theta}{2}\rho + 1 \right). \tag{38}$$

Comparing (37) and (38), it is obvious that when $a = 1$, the two detection orderings have exactly the same performance. However, when $a^2 \geq 1$, we have $C'_{\mathrm{SZF}} \geq C_{\mathrm{SZF}}$. Consequently, from the perspective of maximizing the capacity, assuming no error propagation, the signals should be detected in the order of (1, 2). Intuitively, when $a^2 \geq 1$, the power of the second stream is magnified by the channel while the first stream remains unchanged. Therefore, the capacity gets boosted by a larger amount if the second stream is detected after the cancellation of the first one. On the other hand, from the perspective of BER and error propagation, it might be argued that the ordering should actually be switched. The reason is that the second stream delivers a better SNR upon recovery (a SNR gain of $a^2\cos^2\theta$ versus $\cos^2\theta$ for the first stream). Therefore, the reversed order (2, 1) will minimize the detection error and its propagation in the sequential structure.

D. Comparisons of Equalizers

Theorems 1–3 provide a quantitative insight concerning the capacity performance of various MIMO equalizers with choices

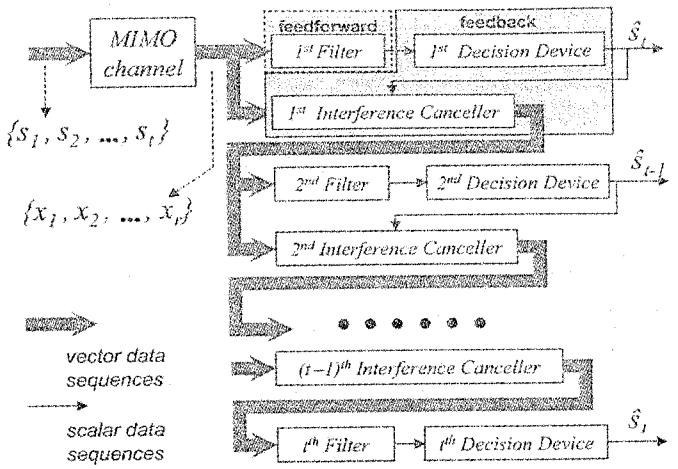


Fig. 3. Block diagram of a sequential equalizer.

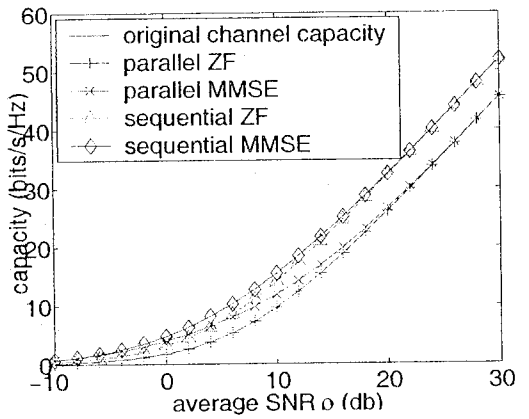


Fig. 4. Capacity performance comparison for different equalizers.

in both filter and structure designs. Comparison of the three theorems yields an in-depth and more unified understanding on the tradeoffs in different MIMO equalizers.

Comparison 1 (Minor Advantage of MMSE Over ZF Filter Design): The analysis in Sections III-A and B demonstrates that the capacity gain of MMSE over ZF is insignificant in modest and high SNR regions. Quantitatively speaking, their difference in terms of absolute capacity degradation (i.e., capacity gap) is on the order of ρ^{-1} for high SNR and ρ for low SNR. In the very rare situation with extremely low SNR (around 0), the PMMSE does deliver much superior capacity ratio (approximately 100%). However, as the theoretical channel capacity is very low at that range, the advantage of MMSE over ZF is, in general, negligible. ■

Fig. 3 shows a block diagram of a sequential equalizer, whereas Fig. 4 illustrates the simulation result on a randomly generated six-input–six-output MIMO channel with even power distribution. The Frobenius norm of the transfer function is normalized to $t \times r = 36$, i.e., $\|H\|_F^2 = 36$. Therefore, the nominal SNR ρ represents the average SNR at each receiver. The capacity performance of different equalizers is displayed. For clarity, the capacity gaps and ratios are magnified in Fig. 5. The shapes of these curves again agree with Theorems 1–3. Note also that in the small SNR region ($\rho < 0$ db), MMSE delivers an impressive performance. However, the situation becomes very different when we consider a more practical SNR region. For $\rho > 10$ dB, the difference between MMSE receiver and its ZF counterpart can be ignored in terms of both capacity gap and capacity ratio.

Comparison 2 (Major Advantage of Sequential Over Parallel Design): The structural selection in the equalizer design proves to be of much greater importance than the filter choice in the modest and high SNR regime. In contrast to the minor performance difference with filter selection, the sequential structure delivers an (asymptotic) capacity improvement of $\sum_{i=1}^{t} \log_2(1+\beta_i)$, bridging the capacity loss of its parallel counterpart, which could be especially large when the channels regarding different transmitters are highly correlated. ■

For the very same channel, the asymptotically constant capacity degradations of PZF or PMMSE are illustrated by the two lower curves in Fig. 4. The sequential equalizers deliver a capacity increment of about 6 bits/s/Hz over its parallel counterpart. In terms of capacity ratio, the sequential equalizers approaches 100% asymptotically at a much faster speed (in the order of $\rho \ln \rho$) than parallel ones (in the order of $\ln \rho$). For example, at $\rho = 20$ dB, the difference between them is as big as 20%. When the SNR $\rho > 0$ dB, the SZF significantly outperforms the PMMSE equalizer. In fact, when the SNR reaches 10 dB or higher, the SZF is very close to achieving full capacity.

Comparison 3 (Tradeoff Between Complexity and Performance): The most complex operation in these equalizer designs are those involving matrix inversions (or Cholesky decompositions). PZF, PMMSE, and SZF equalizers all require only one matrix inversion for the t users, whereas SMMSE requires t such operations. Moreover, the MMSE receivers require an accurate estimate on the SNR for satisfactory performance. Altogether, it appears that SZF equalizers represent an attractive solution in terms of the optimal tradeoff between capacity performance and complexity overhead. ■

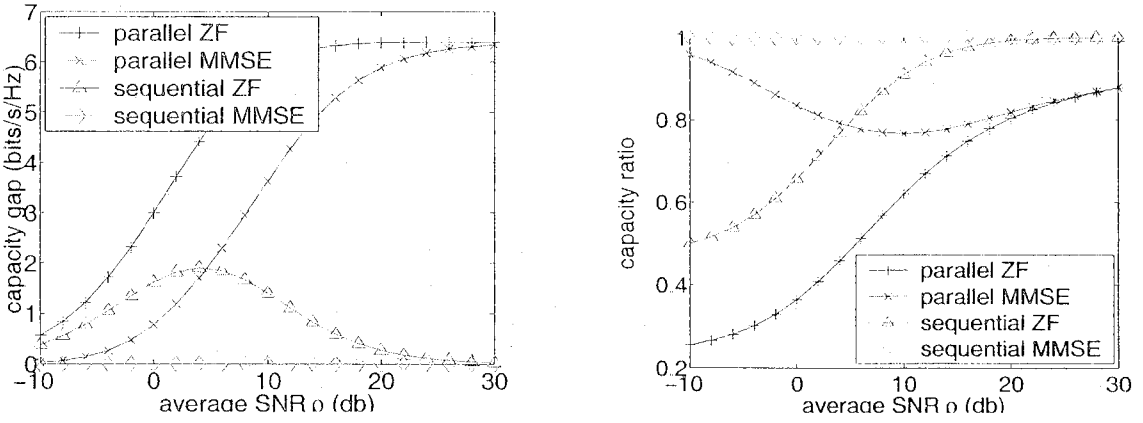
E. Extension to Frequency-Selective Fading Channels

Although our analysis in this section is focused on flat-fading MIMO channels, the results can be directly extended to the frequency-selective situation. For ISI channels, the capacity is treated in frequency domain, where the entire frequency band is divided into infinitesimal strips, each of which could be treated as flat-faded [12]. In addition, the channel capacity in (2) is replaced by the integration of supportable information rates over all the sub-bands

$$C = \frac{1}{2\pi} \int_0^{2\pi} \sum_{i=1}^{t} \log_2(1+\rho|\lambda_i(\theta)|^2) d\theta \qquad (39)$$

where $\{\lambda_i(\theta)\}_{i=1}^{t}$ are the singular values of $\mathbf{H}(\theta)$, which is the transfer function in the frequency domain. Similarly, the capacity performance of the different equalizers in the ISI case are obtainable from the flat-fading results summarized in Theorems 1–3, with exactly the same derivations integrated over the frequency band, and all the terms that characterize the MIMO channel (such as the correlation factors) are now defined in frequency domain.

A major difficulty in the analysis for ISI channel is the FIR constraint. The matrix inverse (or pseudo-inverse) in frequency band, which is required in the ZF and MMSE filter derivations



(a) (b)

Fig. 5. Comparison of (a) capacity gap and (b) capacity ratio for different equalizers. Asymptotically, only SZF and SMMSE approach a zero capacity gap as ρ increases

in (6) and (8), in general leads to an IIR solution in the temporal domain. Therefore, the results above describe the capacity performance of different equalizers with IIR filters, which provides an upper bound on the performance of any FIR implementation. When only finite taps are allowed, the calculation is much more complex. The analysis is based on the theorem of generalized Bezout Identity (for polynomial matrices) and other mathematical results, like the asymptotic eigenvalue distribution for block Toeplitz matrices. In addition, the number of taps allowed in the filter design and the system delay play an important role on the capacity performance [20], [21]. For more details on the capacity analysis for the ISI MIMO channels, see [20].

IV. Stochastic Analysis: Gaussian Channels With Rayleigh Fading

An important category of stochastic flat fading channels is the stationary Rayleigh fading channel, where all the entries in the channel transfer function $\mathbf{H}$ are assumed to be i.i.d. complex, circularly symmetric Gaussian random variables with independent real and imaginary parts, each with covariance $1/2$: $h_{ij} \sim \mathcal{N}_C(0,1)$. It has served as one of the most popular and practical models of the communication environment when there are a large number of superimposed non-line-of-sight (NLOS) propagation paths [19]. As SZF equalizers seem to be an attractive option for complexity-performance tradeoff, in this section, we focus on the analytical investigation of SZF equalizers under Rayleigh fading channels, whereas the performance of other equalizers are given via Monte Carlo simulations for comparison.

For flat Rayleigh fading channels, it is well known [4] that the optimal transmitter power allocation to maximize the average channel capacity is the equal power distribution, i.e., $\mathbf{\Phi} = (1/t)\mathbf{I}_t$, which is assumed by default in our discussion here. The method adopted for the deterministic channels is still applicable to the stochastic case. Recall that the analysis in the previous sections hinges on the Cholesky matrix $\mathbf{R}$, which is defined as $\mathbf{R}^*\mathbf{R} = \mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2}$ and is related to the QR factorization $\mathbf{H} = \mathbf{Q}\bar{\mathbf{R}}$ by the following equation:

$$\mathbf{R}^*\mathbf{R} = \mathbf{\Phi}^{1/2}\mathbf{H}^*\mathbf{H}\mathbf{\Phi}^{1/2} = \mathbf{\Phi}^{1/2}\bar{\mathbf{R}}^*\mathbf{Q}^*\mathbf{Q}\bar{\mathbf{R}}\mathbf{\Phi}^{1/2} = \frac{1}{t}\bar{\mathbf{R}}^*\bar{\mathbf{R}}. \tag{40}$$

Therefore, $\mathbf{R}$ is equal to $\frac{1}{\sqrt{t}}\bar{\mathbf{R}}$ (modular a diagonal unitary matrix). The distribution of matrix $\mathbf{R}$ ($\bar{\mathbf{R}}$) is given by the Bartlett decomposition in the following lemma [22], [23].

Lemma 1 (Distribution of $\bar{\mathbf{R}}$): Given the $r \times t$ $(r \geq t)$ matrix $\mathbf{H}$ with i.i.d. Rayleigh fading entries, let $\mathbf{H} = \mathbf{Q}\bar{\mathbf{R}}$ be its standard QR factorization, where $\bar{\mathbf{R}}$ is an upper-diagonal matrix. Then, we have the following.

1) All the nonzero elements in $\bar{\mathbf{R}}$ are independent random variables.
2) All the off-diagonal terms are i.i.d. complex Gaussian distributed: $\bar{r}_{ij} \sim \mathcal{N}_C(0,1)$ $(i < j)$.
3) The 2-norm of all the diagonal elements are Γ distributed with decreasing freedom
$$|\bar{r}_{ii}|^2 \sim \Gamma(r+1-i,1) \quad 1 \leq i \leq t.$$
(The pdf of $X \sim \Gamma(\alpha,\beta)$ is $p_X(x) = (1/\beta^\alpha\Gamma(\alpha))$ $x^{\alpha-1}e^{-x/\beta}$, $x > 0$. The Γ function is $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$. For a positive integer α, $\Gamma(\alpha) = (\alpha-1)!$.)

For the distribution of $\bar{\mathbf{R}}^{-1}$, we have the following result.

Lemma 2 (Stochastic Property of $\bar{\mathbf{R}}^{-1}$): With the same assumptions and notations as in Lemma 1 and denoting the inverse of $\bar{\mathbf{R}}$ as $\bar{\mathbf{R}}^{-1} = \begin{bmatrix} \hat{\bar{r}}_{11} & \hat{\bar{r}}_{12} & \cdots & \hat{\bar{r}}_{1t} \\ 0 & \hat{\bar{r}}_{22} & \cdots & \hat{\bar{r}}_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\bar{r}}_{tt} \end{bmatrix}$, we claim that

$$\mathcal{E}[|\hat{\bar{r}}_{ij}|^2] = \begin{cases} \frac{1}{r-i}, & \text{if } i = j \\ \frac{1}{(r-j)(r-j+1)}, & \text{otherwise.} \end{cases} \tag{41}$$

Proof: First, we consider $\hat{\bar{r}}_{11}$

$$\begin{aligned} \mathcal{E}[|\hat{\bar{r}}_{11}|^2] &= \mathcal{E}\left[\frac{1}{|\bar{r}_{11}|^2}\right] \\ &= \int_0^\infty \frac{1}{\Gamma(r)}x^{r-2}e^{-x}dx \\ &= \frac{1}{r-1}. \end{aligned} \tag{42}$$

Thus, (41) holds for $i = j = 1$. Next, we assume that the statement in (41) holds for any $j \leq k-1$, and then, we consider the case $j = k$. Denote the left-upper $(k-1)\times(k-1)$ minor of

$\bar{\mathbf{R}}$ as $\bar{\mathbf{R}}_1$. As $\bar{\mathbf{R}}$ is upper triangular, $\bar{\mathbf{R}}_1^{-1}$ is also the left-upper $(k-1)\times(k-1)$ minor of $\bar{\mathbf{R}}^{-1}$. It is then obvious that

$$\begin{bmatrix}\hat{\bar{r}}_{1k} & \hat{\bar{r}}_{2k} & \cdots & \hat{\bar{r}}_{k-1,k}\end{bmatrix}^T = -\frac{1}{\bar{r}_{kk}}\bar{\mathbf{R}}_1^{-1}\begin{bmatrix}\bar{r}_{1k} & \bar{r}_{2k} & \cdots & \bar{r}_{k-1,k}\end{bmatrix}^T. \quad (43)$$

Due to the upper-triangular structure of $\bar{\mathbf{R}}$, all the elements in $\bar{\mathbf{R}}_1^{-1}$ are independent of $\{\bar{r}_{ik}\}_{i=1}^k$ and satisfy (41). Therefore

$$\mathcal{E}\left[\left|\hat{\bar{r}}_{ik}\right|^2\right] = \mathcal{E}\left[\left|\sum_{j=i}^{k-1}\frac{\hat{\bar{r}}_{ij}\bar{r}_{jk}}{\bar{r}_{kk}}\right|^2\right] = \frac{\mathcal{E}\left[\sum_{j=i}^{k-1}|\hat{\bar{r}}_{ij}|^2\right]}{r-k} = \frac{1}{(r-k)(r-k+1)}. \quad (44)$$

Thus, it is verified by induction. ∎

The pdf of $\bar{\mathbf{R}}$ investigated in Lemmas 1 and 2 then helps to describe the stochastic characteristics of $\mathbf{R}$, and capacity formulas in Rayleigh fading channels in turn, which are summarized in Theorems 4 and 5.

Theorem 4 (SZF Performance Under Rayleigh Fading at Low SNR): For flat Rayleigh fading channels, in the low SNR regime, the mean value of capacity ratio and capacity gap for SZF equalizers are

$$\mathcal{E}\left[\frac{C_{\mathrm{SZF}}}{C}\right] = \frac{2r-t+1}{2r} + O[\rho] \quad (45)$$

$$\mathcal{E}[C - C_{\mathrm{SZF}}] = \frac{\rho(t-1)}{2\ln 2} + o[\rho^2]. \quad (46)$$

Proof: Equations (45) and (46) can be derived from (29) and the stochastic result in Lemma 1. Recalling (29), we have

$$\begin{aligned}\mathcal{E}[C - C_{\mathrm{SZF}}] &= \frac{\rho}{\ln 2}\mathcal{E}\left(tr[\mathbf{R}^*\mathbf{R}] - \sum_{i=1}^{t}|r_{ii}|^2\right) + O[\rho^2] \\ &= \frac{\rho}{\ln 2}\mathcal{E}\left[\frac{1}{t}\sum_{i=1}^{t-1}\sum_{j=i+1}^{t}|\bar{r}_{ij}|^2\right] + O[\rho^2] \\ &= \frac{\rho(t-1)}{2\ln 2} + O[\rho^2] \end{aligned} \quad (47)$$

and the mean value of capacity ratio at low SNR is

$$\begin{aligned}\mathcal{E}\left[\frac{C_{\mathrm{SZF}}}{C}\right] &= \mathcal{E}\left[\frac{\sum_{i=1}^{t}|r_{ii}|^2}{\sum_{i=1}^{t}|r_{ii}|^2 + \sum_{i=1}^{t-1}\sum_{j=i+1}^{t}|r_{ij}|^2}\right] + O[\rho] \\ &= \mathcal{E}\left[\frac{\sum_{i=1}^{t}|\bar{r}_{ii}|^2}{\sum_{i=1}^{t}|\bar{r}_{ii}|^2 + \sum_{i=1}^{t}\sum_{j=i+1}^{t-1}|\bar{r}_{ij}|^2}\right] + O[\rho]. \end{aligned} \quad (48)$$

Note that $\sum_{i=1}^{t}|\bar{r}_{ii}|^2 \sim \Gamma(t(2r-t+1)/2, 1)$ are independent of the random variables $\sum_{i=1}^{t-1}\sum_{j=i+1}^{t}|\bar{r}_{ij}|^2 \sim \Gamma(t(t-1)/2, 1)$. We denote $a \equiv t(2r-t+1)/2$ and $b \equiv t(t-1)/2$, and the formula in (48) can then be evaluated as

$$\begin{aligned}\mathcal{E}\left[\frac{C_{\mathrm{SZF}}}{C}\right] &= \int_0^\infty\int_0^\infty \frac{x}{x+y}\frac{1}{\Gamma(a)\Gamma(b)}x^{a-1}y^{b-1} \times e^{-x-y}dxdy \\ &= \int_0^\infty\int_0^1 \frac{1}{\Gamma(a)\Gamma(b)}x^{a+b-1}z^{-b}(1-z)^{b-1} \times e^{-x/z}dzdx \quad \left(z \equiv \frac{x}{x+y}\right) \\ &= \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}z^a(1-z)^{b-1}dz \\ &= \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+b+1)\Gamma(a)} \\ &= \frac{a}{a+b} = \frac{2r-t+1}{2r}. \end{aligned} \quad (49)$$

Thus, it is proved. ∎

Theorem 5 (SZF Performance Under Rayleigh Fading at High SNR): For flat Rayleigh fading channels, in the high SNR regime, the mean value of capacity ratio and capacity gap of SZF equalizers are the following.

1) When $r > t$

$$\mathcal{E}\left[\frac{C_{\mathrm{SZF}}}{C}\right] = 1 - \frac{1}{\rho\ln\rho}\sum_{i=1}^{t}\left(\frac{1}{r-t} - \frac{1}{r-i}\right) + o[(\rho\ln\rho)^{-1}] \quad (50)$$

$$\mathcal{E}[C - C_{\mathrm{SZF}}] = \frac{t}{\rho\ln 2}\sum_{i=1}^{t}\left(\frac{1}{r-t} - \frac{1}{r-i}\right) + O[\rho^{-2}]. \quad (51)$$

2) When $r = t$

$$\begin{aligned}\mathcal{E}\left[\frac{C_{\mathrm{SZF}}}{C}\right] &= 1 - \frac{1}{\rho\ln\rho}\sum_{i=1}^{t-1}\left(1 - \frac{1}{i}\right) + o[(\rho\ln\rho)^{-1}] \\ \mathcal{E}[C - C_{\mathrm{SZF}}] &= \frac{t}{\rho\ln 2}\sum_{i=1}^{t-1}\left(1 - \frac{1}{i}\right) + O[\rho^{-2}]. \end{aligned} \quad (52)$$

Proof: Equation (28) shows that the evaluation of mean capacity ratio and gap is determined by deriving the expectation of $\sum_{i=1}^{t}\frac{\beta_i}{|r_{ii}|^2}$, which is equivalent to the 2-norm summation of all the off-diagonal terms in matrix $\mathbf{R}^{-1}$. The statement for $r > t$ in the theorem can be easily verified by plugging the result of Lemma 2 into (28).

The high SNR analysis in (50) and (51) meets serious numerical problems with symmetric antenna setting $r = t$, which suggests an infinite term $\sum_{i=1}^{t}[1/(r-t) - 1/(r-i)]$. The reason is that our previous analysis for deterministic channels are only reasonable when assuming that the transfer function has full rank. However, under Rayleigh fading channels with $r = t$, from Lemma 1, the last element in $\bar{\mathbf{R}}$ has the distribution $\bar{r}_{tt} \sim \Gamma(1,1)$, i.e., $p_{\bar{r}_{tt}}(x) = e^{-x}$, which has nonzero probability at

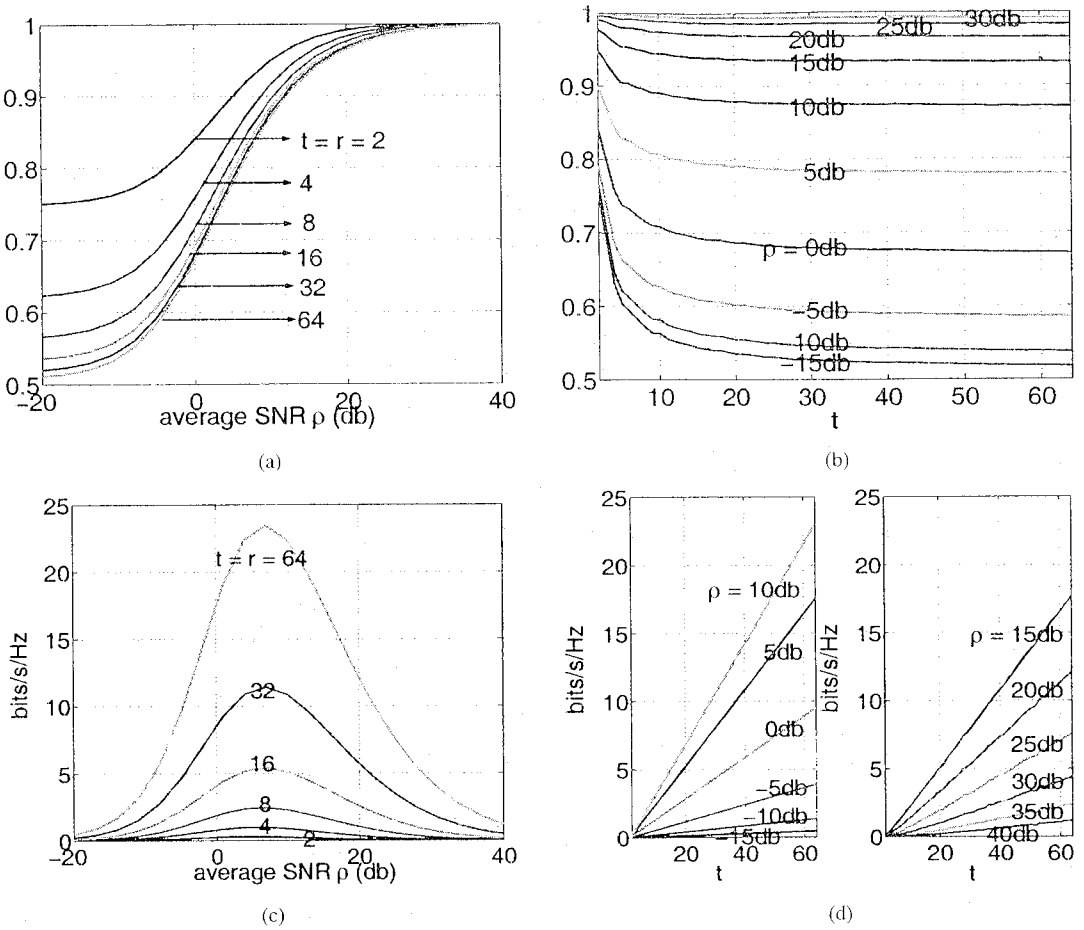


Fig. 6. Mean value of capacity ratio and capacity gap achieved by SZF under Rayleigh fading channel and symmetric antenna setting $r = t$ versus ρ and t.

$x = 0$. The possible singularity in Rayleigh fading channels invalidates our high SNR analysis by inserting undesired infinite terms in the equations mentioned above. To fix this problem, note that the singularity only happens at the last column, and therefore, the high SNR analysis in the $r = t$ case can be proceeded by ignoring the last column of $\mathbf{R}^{-1}$. This leads to the result in the theorem.

It is worth noting that with extra spatial diversities, as long as $r > t$, the singularity mentioned above diminishes. This can be easily checked by noting that all the diagonal elements in $\mathbf{R}$ now have zero probability density at point $x = 0$. ∎

Conjecture 1 (Monotonicity of Mean Capacity Ratio): For Rayleigh fading channels, the mean value of capacity ratio achieved by SZF is a monotonically increasing function of the SNR ρ. This has been confirmed by Monte Carlo simulations for some typical antenna setting values. In fact, the conjecture can be upgraded to be a theorem for the case $t = 2$. ∎

When $t = 2$, the claim can be analytically established by making use of the following two inequalities:

$$\sum_{i=1}^{2} |r_{ii}|^2 \le \sum_{i=1}^{2} |\lambda_i|^2, \quad \sum_{i=1}^{2} |r_{ii}|^{-2} \le \sum_{i=1}^{2} |\lambda_i|^{-2}$$

and the fact that $\prod_{i=1}^{2} |r_{ii}| = \prod_{i=1}^{2} |\lambda_i|$. In addition, we also observe that the monotonicity of capacity ratio almost always holds for the generic channels (i.e., when $\min(|\lambda_i|^2)/\max(|\lambda_i|^2)$ is above certain QoS-safeguard threshold). Under Rayleigh fading assumption, the distribution of the elements in $\mathbf{R}$ has been established in closed form. This suggests that the channel should be statistically guarded from being singular, and the condition number will, in general, fall into a respectable region.

Theorems 4 and 5 and Conjecture 1 together predict the shapes of average capacity gap and ratio achieved by SZF equalizers in Rayleigh fading channels. The monotonicity property is clearly depicted throughout the capacity ratio plots in the paper, especially those in Figs. 6(a) and 7(a), which will be explained shortly. Theorem 4 serves as a simple prediction of the lower bound on the average capacity ratio in the entire SNR range. In contrast, Theorem 5 provides the stochastic behavior at high SNR. Specifically, the range of average capacity ratio runs roughly from

$$\eta = \frac{2r - t + 1}{2r}$$

(for $\rho \to 0$) to 1.0 (for $\rho \to \infty$). Note that the range is dependent only on the antenna diversities t and r. The theorems also suggest a bell-shape curve of the average capacity gap versus ρ, as will be evidenced by the simulation results in Figs. 6(c) and 7(c).

Example ($r = t$ Analysis): For the symmetric case with equal number of transmitters and receivers $r = t$, by Theorems 4 and 5, the average capacity ratio and gap are asymptotically as follows.

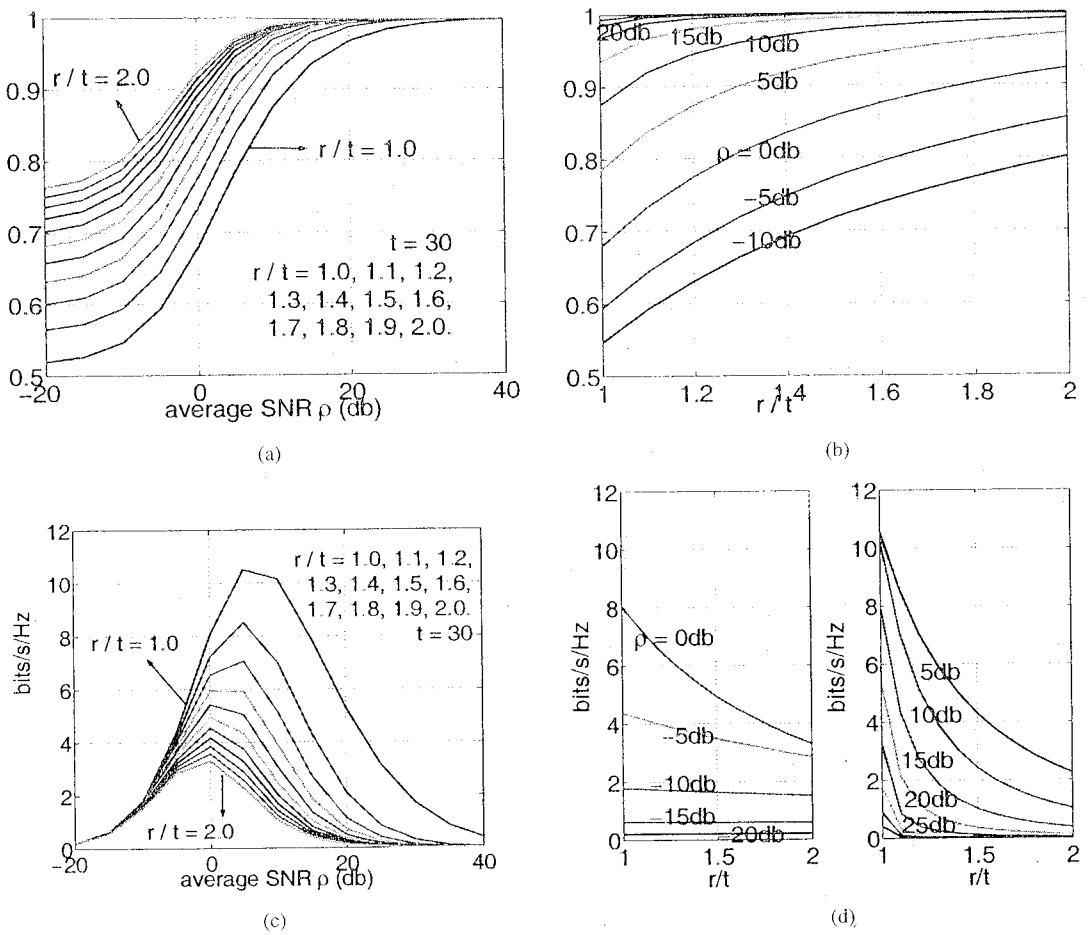


Fig. 7. Impact of extra dimension of spatial diversity on the performance of SZF equalizers. Mean values of capacity ratio and capacity gap achieved by SZF under Rayleigh fading channel with respect to ρ and r/t.

1) $\rho \to 0$: $\mathcal{E}[C_{\mathrm{SZF}}/C] \to 1/2 + 1/(2t)$ is a monotonically decreasing function of t, which provides a lower bound on the average capacity ratio in the entire SNR range.
2) $\rho \to \infty$: $\mathcal{E}[C_{\mathrm{SZF}}/C] \to 1 - 1/(\rho \ln \rho) \sum_{i=1}^{t} \left(1 - \frac{1}{i}\right)$.

   When $t \to \infty$, $\sum_{i=1}^{t}(1-1/i) \to t - \ln t$, and therefore, $\mathcal{E}[C_{\mathrm{SZF}}/C] \to 1 - (t - \ln t)/\rho \ln \rho$. It is also a decreasing function of t.
3) $\rho \to 0$: $\mathcal{E}[C - C_{\mathrm{SZF}}] \to \rho(t-1)/(2 \ln 2)$ is a linearly increasing function of ρ and t.
4) $\rho \to \infty$: $\mathcal{E}[C - C_{\mathrm{SZF}}] \to t(t - \ln t)/(\rho \ln 2)$ (as $t \to \infty$) is an increasing function of t.

■

The analysis is further verified by Monte Carlo simulations, displayed in Fig. 6, where 1000 tests are conducted for each plot. The channels are drawn from $\mathcal{C}^{t \times t}$ Rayleigh fading assemblies. Fig. 6(a) and (c) show the expectation of capacity ratio and gap versus ρ for different numbers of transmitters $r = t$. With $r = t$ increasing from 2 to 64, the corresponding average capacity ratio curve goes from upper to lower ones, as expected, and the (absolute) capacity gap also becomes larger. The average capacity ratio values at $\rho = -20$ dB in Fig. 6(a) perfectly match our lower bound $1/2 + 1/4$. The same simulation results are displayed in another angle in Fig. 6(b) and (d), where each curve represents the capacity performance versus the number of antennas at one SNR point. As suggested by the analysis above, at low SNRs, the average capacity ratio is almost inversely proportional to t, whereas in the high SNR regime, it looks like a slowly decreasing linear function. It is interesting to see that the average capacity gap is a linear function of t almost in the entire SNR range, as demonstrated by Fig. 6(d). As $\rho < 10$ dB [in the left subplot of (d)], the curve goes up with increasing ρ, whereas the case is reversed when $\rho > 10$ dB.

Example 2 ($r > t$ Analysis): Nonsquare MIMO channels with $r > t$ should outperform the square MIMO due to the expanded receiver diversity. We evaluate the asymptotic result when $r/t = k > 1$ is fixed and $t \to \infty$.

1) $\rho \to 0$: $\mathcal{E}[C_{\mathrm{SZF}}/C] \to 1 - (1/2k)$ is a monotonically increasing function of k.
2) $\rho \to \infty$: $\mathcal{E}[C_{\mathrm{SZF}}/C] \to 1 - [1/(k-1) + \ln(1 - 1/k)]/(\rho \ln \rho)$ is also an increasing function of k.
3) $\rho \to 0$: $\mathcal{E}[C - C_{\mathrm{SZF}}] \to \rho t/2 \ln 2$ is irrelevant for k.
4) $\rho \to \infty$: $\mathcal{E}[C - C_{\mathrm{SZF}}] \to [1/k - 1 + \ln(1 - 1/k)]t/\rho \ln 2$ is a decreasing function of k.

■

Example 2 is confirmed by the Monte Carlo results shown in Fig. 7. Here, we choose a fixed large number of transmitters $t = 30$ and vary the number of receivers r from 30 to 60. The performance is displayed in the same manner as in Fig. 6, except that (c) and (d) are plotted versus r/t instead of t.

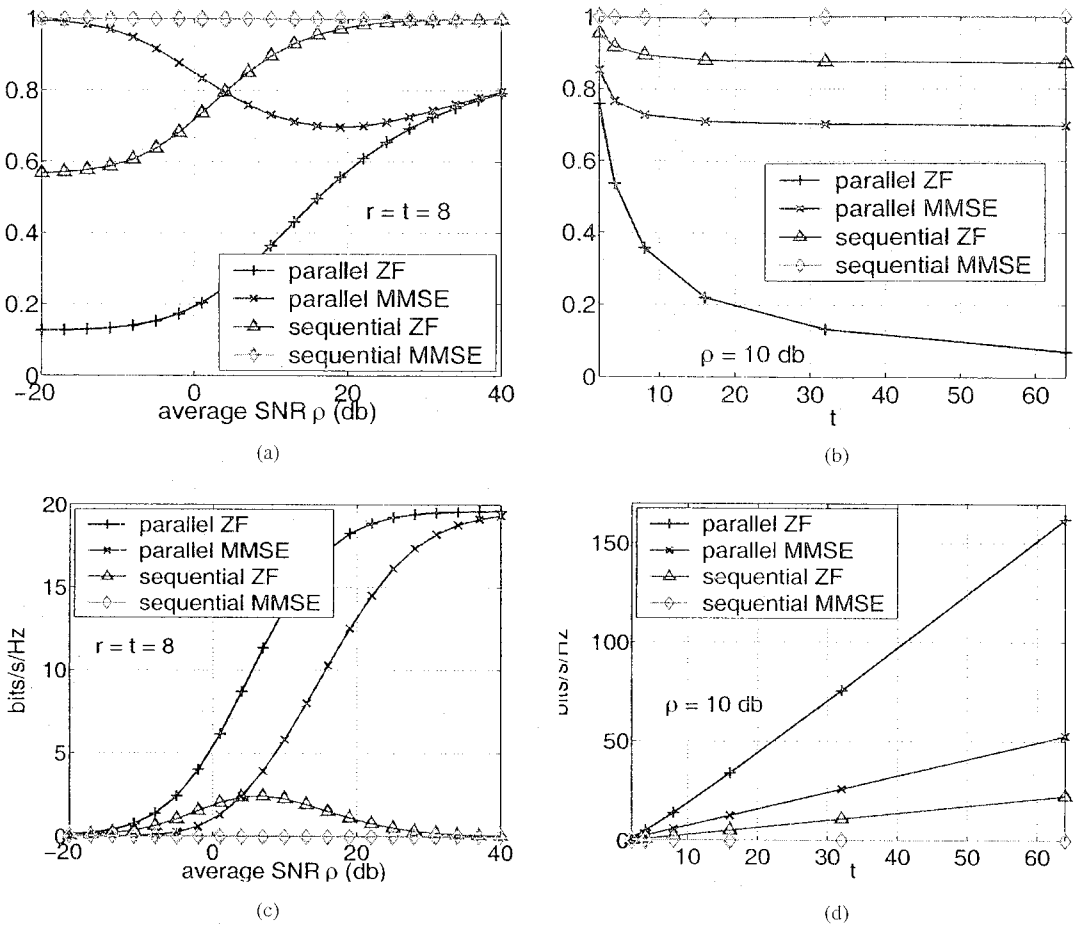


Fig. 8. Comparison of the average capacity ratio/gap achieved by PZF, PMMSE, SZF, and SMMSE equalizers under Rayleigh fading channel with respect to ρ and t ($r = t$).

Up to now, the average capacity performance for SZF equalizers has been addressed with both analytical formulas and simulation plots. Next we give some simulation results on comparison of PZF, PMMSE, SZF, and SMMSE equalizers under Rayleigh fading channels. The average capacity ratio/gap of the four equalizers with the antenna setting $r = t = 8$ are plotted w.r.t. ρ in Fig. 8(a)–(c), respectively. At $\rho \geq 20$ dB, the SZF almost coincides the curve of SMMSE, which achieves the full channel capacity. Fig. 8(b) and (d) display the performance at $\rho = 10$ dB w.r.t. the number of antennas t ($r = t$). The PZF equalizer shows to be very sensitive to t, which degrades dramatically when the number of antennas increases. The PMMSE equalizer delivers about twice the capacity gap of that by SZF at large t. The SZF equalizer in general delivers an average capacity ratio of about 90% at $\rho = 10$ dB.

To highlight the theoretical finding that the structural selection plays a much more important role than the filter selection, extensive simulation based on 10 000 tests for PZF, PMMSE, and SZF equalizers are conducted. Here, the MIMO channels $\mathbf{H}$ are drawn from the $\mathcal{C}^{8\times 6}$ Rayleigh fading assembly. Fig. 9(a) and (b) display the empirical pdfs of capacity ratio achieved by the three equalizers at $\rho = 10$ dB and $\rho = 20$ dB, respectively. At $\rho = 10$ dB, the average capacity ratios of PZF, PMMSE, and SZF are about 74.96, 83.34, and 97.07%, respectively. For $\rho = 20$ dB, the three values are correspondingly 86.46, 87.27, and 99.75%. In both cases, SZF significantly outperforms the others. It is worth noting that the capacity ratio of SZF, as a random variable, has a much smaller variance than the other choices. This means SZF not only delivers a superior average performance over PZF and PMMSE but also has a more reliable guarantee on the capacity. The performance difference between PZF and PMMSE goes down with increasing ρ. At $\rho = 20$ dB, their curves almost coincide, and the capacity ratio of SZF is densely located around 0.99 in the shape of an impulse function.

Our theoretical finding also predicts that an increasing receiver diversity will further enhance the performance of SZF equalizers. This is also confirmed by an extensive simulation study. Fig. 10(a) and (b) display the histograms of capacity ratio (based on 10 000 tests) for SZF equalizers at SNR $\rho = 20$ dB. It clearly suggests that while SZF for $\mathcal{C}^{6\times 6}$ and $\mathcal{C}^{8\times 6}$ MIMO channels both display a convincing capacity achieving property, the latter is far better than the former.

V. Conclusion

This paper is motivated by the practical necessity for a MIMO system designer to consider an optimal tradeoff between the implementation cost and the achievable capacity. Our study provides simple (mostly in closed-form) theoretical analysis of the achievable capacities among various equalizers. Both deterministic and stochastic channels were investigated. The asymptotic

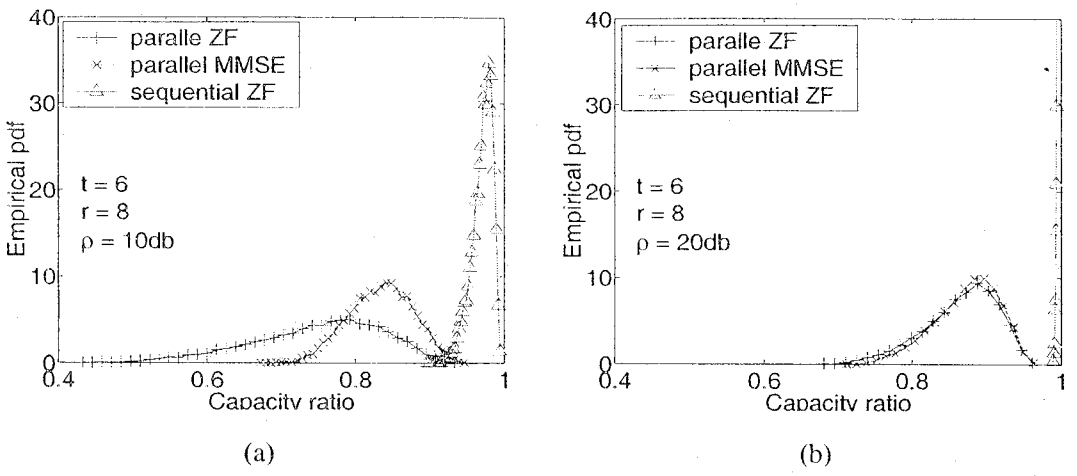


Fig. 9. Empirical probability density functions of capacity ratio (based on 10 000 tests) for PZF, PMMSE, and SZF equalizers at (a) $\rho = 10$ dB and (b) $\rho = 20$ dB for $\mathcal{C}^{8\times 6}$ MIMO channels.

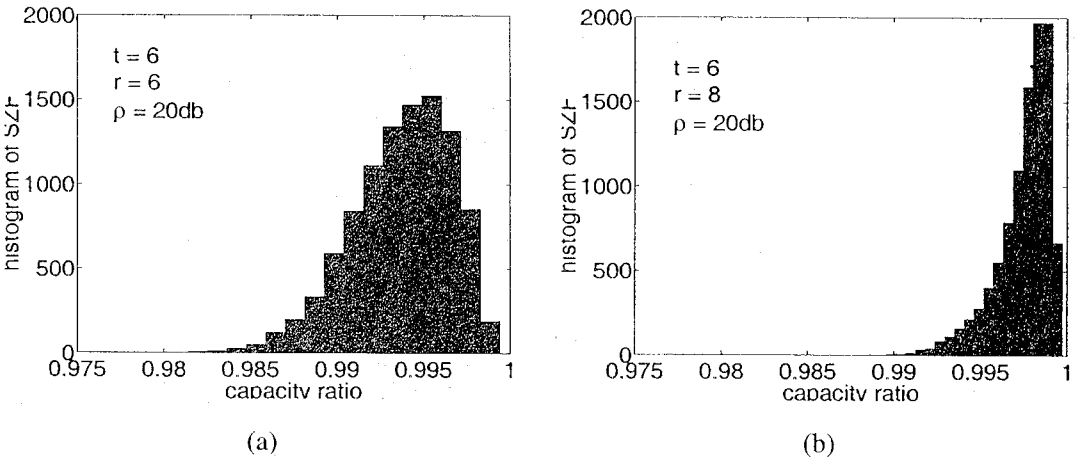


Fig. 10. Histograms of capacity ratio (based on 10 000 tests) for SZF equalizers at SNR $\rho = 20$ dB under (a) $\mathcal{C}^{6\times 6}$ Rayleigh fading channels and (b) $\mathcal{C}^{8\times 6}$ Rayleigh fading channels.

performance (when the SNR approaches either zero or infinity) is derived in closed form. Our theoretical finding concludes that the structural selection plays a much more important role than the filter selection. In short, the sequential detection structure can basically approach the Shannon capacity, regardless of whether MMSE or ZF filter is adopted. Considering the fact that ZF filters incur much simpler implementation than MMSE, the SZF equalizer represents a very promising and feasible design tradeoff between performance and cost. It is worth noting that the tradeoff will become more important with the presence of ISI since the implementation cost will become more expensive. Fortunately, the analysis here can be naturally extended to deal with the ISI MIMO channels, which will be reported on in a future publication.

Acknowledgment

The authors are very grateful to the editor and the anonymous reviewers for their invaluable comments and constructive suggestions.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in fading environments when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.

[2] J. H. Winters, "On the capacity of radio communications systems with diversity in rayleigh fading environments," *IEEE J. Select. Areas Commun.*, vol. JSAC-5, pp. 871–878, June 1987.

[3] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[4] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, Nov.–Dec. 1999.

[5] L. Zheng and D. N. C. Tse, "Diversity and freedom: A Fundamental tradeoff in multiple antenna channels," in *Proc. IEEE Int. Symp. Inform. Theory*, 2002, p. 476.

[6] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block code from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[7] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

[8] R. W. Heath and A. J. Paulraj, "Switching between multiplexing and diversity based on constellation distance," in *Proc. Allerton Conf. Commun. Contr. Comput.*, Oct. 2000.

[9] J. Salz, "Digital transmission over cross-coupled linear channels," *AT&T Tech. J.*, vol. 64, no. 6, pp. 1147–1159, 1985.

[10] A. Duel-Hallen, "Equalizers for multiple input/output channels and PAM systems with cyclostationary input sequences," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 630–639, Apr. 1992.
[11] B. A. Bjerke and J. G. Proakis, "Equalization and decoding for multiple-input-multiple-output wireless channels," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 3, pp. 249–266, Mar. 2002.
[12] G. G. Robert, *Information Theory and Reliable Communications*. New York: Wiley, 1968.
[13] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multiple antennas," *Bell Labs. Tech. J.*, vol. 2, pp. 41–59, 1996.
[14] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron.*, Pisa, Italy, Sept. 1998, pp. 295–300.
[15] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Conf. Rec. 31st Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 1997, pp. 1405–1409.
[16] A. J. Viterbi, "Spread spectrum multiple access with binary psk modulation can approach shannon capacity for the aggregate Gausian channel," unpublished, private communication, 1986.
[17] N. Al-Dhahir, A. F. Naguib, and A. R. Calderbank, "Finite-length MIMO decision feedback equalizers for space-time block-coded signals over multipath-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 1176–1182, July 2001.
[18] A. Lozano and C. Papadias, "Layered space-time receivers for frequency-selective wireless channels," *IEEE Trans. Commun.*, vol. 50, pp. 65–73, Jan. 2002.
[19] H. V. Poor and G. W. Wornell, *Wireless Communications*. Upper Saddle River, NJ: Prentice-Hall, 1998.
[20] S. Y. Kung, Y. Wu, and X. Zhang, "Bezout space-time precoding and equalization for MIMO channels," *IEEE Trans. Signal Processing*, vol. 50, pp. 2499–2514, Oct. 2002.
[21] X. Zhang and S. Y. Kung, "Capacity bound analysis for MIMO equalizers in frequency-selective fading channels," *IEEE Trans. Signal Processing*, submitted for publication.
[22] M. S. Bartlett, "On the theory of statistical regression," *Proc. R. Soc. Edinb.*, no. 53, pp. 260–283, 1933.
[23] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1989.

Xinying Zhang received the B.S. degree from Electronics Engineering Department, Tsinghua Univerisity, Beijing, China, in 1998. She is currently pursuing the Ph.D. degree from the Department of Electrical Engineering, Princeton University, Princeton, NJ.

Her research interests lie in the areas of signal processing and communications, including channel equalization, space-time coded systems, multicarrier communication, and MIMO systems.

Sun-Yuan Kung (F'88) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

Since 1987, he has been a Professor of electrical engineering at Princeton University, Princeton, NJ. In 1974, he was an Associate Engineer with Amdahl Corporation, Sunnyvale, CA. From 1977 to 1987, he was a Professor of electrical engineering–systems, University of Southern California, Los Angeles. In 1984, he was a Visiting Professor with the Stanford University and the Delft University of Technology, Delft, The Netherlands. In 1994, he was a Toshiba Chair Professor at Waseda University, Waseda, Japan, and a Honorary Professor with the Central China University of Science and Technology, Wuhan, China. In 2001, he was a Distinguished Chair of Multimedia Signal Processing, Hong Kong Polytechnic University. His research interests include VLSI array processors, image/video/multimedia signal processing, neural networks for biometric and bioinformatic signal processing, and wireless digital communications. He has authored more than 300 technical publications, including three books: *VLSI Array Processors* (Englewood Cliffs, NJ: Prentice-Hall, 1988) (with Russian and Chinese translations); *Digital Neural Networks* (Englewood Cliffs, NJ: Prentice-Hall, 1993), and *Principal Component Neural Networks* (New York: Wiley, 1996). He has edited numerous reference books, including *VLSI and Modern Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, 1985) (with Russian translation), *VLSI Signal Processing, Vol III* (Piscataway, NJ: IEEE Press), *Neural Networks for Signal Processing, Vols. I, II, and III* (Piscataway, NJ: IEEE Press), *Multimedia Signal Processing, Vol. I* (Piscataway, NJ: IEEE Press), *Systolic Arrays* (Menlo Park, CA: IEEE Comput. Soc. Press), and *Application-Specific Array Processors* (Menlo Park, CA: IEEE Comput. Soc. Press). He has recently co-edited a book entitled *Multimedia Image and Video Processing* (Boca Raton, FL: CRC, 2001).

Dr. Kung has served as an Editor-In-Chief of the *Journal of VLSI Signal Processing Systems* since 1990. He was appointed as the first Associate Editor in the VLSI Area in 1984 and the first Associate Editor in the Neural Network Area in 1991 of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He served as a member of IEEE Signal Processing Society Administration Committee from 1989 to 1991. He was a founding member of IEEE-SPS Technical Committees on VLSI Signal Processing, Neural Networks, and Multimedia Signal Processing. He chaired various international conferences, including IEEE Workshops on VLSI Signal Processing, Neural Networks, and Signal Processing, Multimedia Signal Processing, the International Conference on Application Specific Array Processors, and the International Computer Symposium. He was the recipient of 1992 IEEE Signal Processing Society Technical Achievement Award for his contributions on "parallel processing and neural network algorithms for signal processing." He was appointed an IEEE-SP Distinguished Lecturer in 1994. He received 1996 IEEE Signal Processing Society's Best Paper Award for his publication on principal component neural networks. He was a recipient of the IEEE Third Millennium Medal in 2000.